

Л.Шварц
АНАЛИЗ
ТОМ 1

Имя Лорана Шварца — одного из крупнейших математиков современности — хорошо известно советским специалистам.

Его двухтомный курс существенно отличается от всех имеющихся книг по анализу. Изложение характеризуется глубоким взаимопроникновением методов классического и функционального анализа, современной алгебры и топологии. Следует отметить также блестящий стиль курса, умение автора выделить основное, объяснить значение тех или иных идей.

Первый том включает теорию множеств, топологию, дифференциальное и интегральное исчисление.

Книга Л. Шварца, несомненно, заинтересует преподавателей математики, научных работников в области математики, физики и механики, а также инженеров и будет весьма полезна студентам университетов, педагогических институтов и высших технических учебных заведений с углубленным изучением математики.

Содержание

Предисловие редактора перевода	5
Предисловие	7
Глава I. Теория множеств	9
1. Множества. Элементарные операции	9
Части множества	9
Отношение включения. Дополнение	10
Объединение. Пересечение	10
Произведение множеств	11
2. Отображения. Функции	11
Примеры отображений	12
Инъекции. Сюръекции. Биекции	13
Образ и прообраз подмножества	13
Множество отображений. Семейства. Последовательности	15
Композиция отображений	15
Замена переменных и замена функций	16
3. Отношения эквивалентности. Фактормножество	17
Классы эквивалентности. Разбиения	18
Фактормножество	19
Факторгруппа по инвариантной подгруппе	19

Факторпространство векторного пространства по векторному подпространству	20
4. Отношения порядка	21
Примеры отношений порядка	22
Мажорируемые части. Мажоранты. Максимум. Точная верхняя грань	23
Возрастающие функции	25
Пополненная прямая	26
5. Мощности. Счетные множества	27
Мощности. Кардинальные числа	27
Счетные множества	31
Мощность континуума	33
Трансцендентные числа	33
Континуум-гипотеза	35
6. Некоторые основные понятия логики	36
Глава II. Топология	40
1. Метрические пространства. Элементарные примеры	40
Сфера. Шары	41
Нормированные векторные пространства	41
2. Открытые и замкнутые части. Окрестности. Внутренность. Граница.	44
Замыкание. Плотные подмножества	44
Открытые части	44
Замкнутые части	45
Окрестности	46
Внутренность	47
Внешность	48
Граница	48
Замыкание	49
Плотные подмножества	50
Подпространства. Индуцированная метрика	50
3. Непрерывные функции. Гомеоморфизмы	52
Гомеоморфизмы	54
4. Метрические пространства и топологические пространства	55
Топология пополненной прямой \bar{R}	60
5. Последовательности. Пределы. Сходимости	60
6. Топологическое произведение	63

Сходящиеся последовательности в произведении	64
Непрерывные функции многих переменных	65
Топологические группы. Топологические векторные пространства	66
Раздельная непрерывность функции двух переменных	67
7. Компактные пространства. Элементарные свойства	67
Локально компактные пространства	74
Точка сгущения последовательности	75
Верхний и нижний пределы вещественной последовательности	79
8. Свойства непрерывных функций на компактных пространствах	79
Равномерная непрерывность	86
9. Связные пространства	88
Линейно связные пространства	90
10. Дополнение по общей топологии связных пространств	92
Некоторые применения понятия связности. Критерии негомеоморфности	97
Существование и непрерывность обратной функции для строго монотонной непрерывной функции	98
Применение: метрики, определяющие топологию в \bar{R}	99
11. Полные метрические пространства	100
Продолжение равномерно непрерывных отображений	104
Частные свойства конечномерных топологических векторных пространств	106
12. Теорема о неподвижной точке	107
13. Элементарная теория нормированных векторных пространств и пространств Банаха	110
Ядро и образ непрерывного линейного отображения	113
Произведения нормированных векторных пространств	118
Билинейные непрерывные отображения произведения нормированных векторных пространств в нормированное векторное пространство	121
Мультилинейные непрерывные отображения	126
Алгебры. Нормированные алгебры	127
14. Ряды в нормированных векторных пространствах	128
Перестановка членов ряда	130
Суммирование по блокам безусловно сходящегося ряда	134
Действие линейного непрерывного отображения на ряд	136
Произведение двух числовых рядов. Применение билинейного	137

непрерывного отображения к двум рядам	
Обратимые отображения в банаховых пространствах	139
Критерий условной сходимости	142
15. Наиболее употребительные примеры функциональных пространств.	146
Сходимость простая и равномерная	146
Функциональные пространства	146
Простая сходимость последовательности функций	149
Равномерная сходимость последовательности функций	150
Другие применения выражения "равномерная сходимость"	152
Пространства, порожденные структурами пространств E и F	154
Непрерывность локально равномерного предела последовательности непрерывных функций	155
Некоторые контрпримеры	157
Ряды функций со значениями в нормированном векторном пространстве	159
16. Бесконечные произведения вещественных или комплексных чисел и функций	162
Бесконечные произведения и логарифмические ряды	164
Бесконечные произведения вещественных или комплексных функций	167
Применение к функции ζ Римана	168
Глава III. Дифференциальное исчисление	174
1. Аффинные пространства	174
Аффинные многообразия	176
Линейные отображения. Аффинные отображения	178
Аффинные нормированные пространства	179
Выпуклые множества в аффинных пространствах	183
Евклидовы векторные и евклидовы аффинные пространства	184
Эрмитовы векторные и эрмитовы аффинные пространства	187
Изоморфизм (или полуизоморфизм) конечномерного евклидова (или эрмитова) пространства и его сопряженного пространства	189
Ортонормированные базисы	190
Обобщенные евклидовы или эрмитовы пространства	192
2. Вещественные функции вещественной переменной. Непрерывность справа и слева	195
Разрывы первого рода. Правильные функции	196
Производная вещественной функции вещественной переменной	198
Монотонные функции	202

Дифференцируемые функции и теоремы о промежуточных значениях	204
Выпуклые функции	204
3. Производная отображения одного аффинного пространства в другое.	208
Производный вектор функции скалярной переменной	
Общий случай. Частная производная вдоль вектора	209
Матрица Якоби. Якобиан	211
Недостатки понятия производной вдоль вектора	212
Полная производная или производное отображение	213
Понятие дифференциала	217
Геометрическая интерпретация производного отображения: дифференцируемое многообразие и линейное касательное многообразие	218
Градиент вещественной функции в евклидовом пространстве	221
Случай, когда F является произведением аффинных пространств	223
Случай, когда E является произведением аффинных пространств. Частные производные отображения	224
Производная билинейного непрерывного отображения	225
Дифференцируемые функции. Непрерывно дифференцируемые функции	227
Примеры непрерывно дифференцируемых функций	228
Пространства дифференцируемых функций	229
4. Теорема о сложной функции	230
Примеры вычисления обычных производных	235
5. Формула конечных приращений	247
Полная дифференцируемость и частная дифференцируемость	253
6. Производные высших порядков	257
Последовательные производные	261
Случай произведения пространств. Полная и частная дифференцируемости	266
Пространства n раз дифференцируемых функций	267
Производная произведения (формула Лейбница)	268
7. Формула Тейлора. Максимум и минимум	272
Применение формулы Тейлора для вычисления производных	276
Формула Тейлора относительно некоторой системы координат	279
Применение к изучению максимумов и минимумов. Определения	285
Необходимые условия экстремума	285
Нахождение необходимых и достаточных условий экстремума функции	287

Частный случай вещественной функции f двух вещественных переменных x, y	290
Применение формулы Тейлора к изучению расположения гиперповерхности по отношению к касательной гиперплоскости	292
8. Теорема о неявной функции. Постановка задачи	293
Существование неявной функции	294
Дифференцируемость неявной функции	298
Дифференцируемость функции $u \rightarrow u^{-1}$ на $\mathcal{L}(\vec{F}, \vec{G})$	300
Частный случай, когда $E=F=G=K$ -скалярное поле	306
Случай, когда E, F, G конечномерны	308
Обратная функция как неявная функция	309
Вычисление производных высших порядков неявной функции	314
Техника замены переменных и замены функций	318
9. Дифференцируемые многообразия	319
Определение многообразия при помощи его параметрического представления	321
Определение многообразия с помощью неявных уравнений	331
Вещественные и комплексные многообразия	333
Абстрактные многообразия	334
Векторное пространство, касательное в точке к многообразию аффинного пространства E размерности N	338
Векторное пространство, касательное к абстрактному многообразию в точке	343
Теорема о постоянном ранге	345
Зависимые и независимые функции	350
Особые, или параметрические, многообразия	352
10. Условные максимумы и минимумы	353
Практический способ вычисления условного максимума или минимума	356
Применение теории условных максимумов. Неравенства Гёльдера и Минковского	358
11. Вариационное исчисление	369
Постановка задачи	369
Дифференцируемость J	372
Необходимые условия экстремума	378
Лемма Хаара	379

Простые случаи интегрируемости уравнений Эйлера	383
Уравнение геодезических на поверхности	389
Относительный экстремум	393
Замена переменных	394
Приложение к задаче о геодезических	396
Переменные концы. Условие трансверсальности	400
Применение к геодезическим кривым	405
Канонические уравнения Гамильтона	406
Применения к механике	409
Вариационное исчисление для кратных интегралов	410
Глава IV. Интегральное исчисление	416
1. Интеграл Римана на прямой	416
Ступенчатые функции	418
Верхний интеграл Римана от ограниченной функции $f > 0$ с компактным носителем	421
Интегрируемые функции со значениями в пространстве Банаха	423
Интеграл от интегрируемой функции	425
Примеры интегрируемых по Риману функций	433
Вычисление интеграла функции с помощью сумм Коши - Римана	435
Среднее значение функции на интервале	438
2. Меры Радона на локально компактном пространстве	439
Мера Радона на компактном пространстве	439
Примеры мер Радона	440
Меры на локально компактном пространстве	446
Примеры мер Радона	448
Применения к механике и физике	450
Векторные меры	450
Разложение единицы	452
Носитель меры Радона	465
Продолжение меры на непрерывные функции φ с некомпактным носителем	473
Принцип кусочной склейки мер	475
Комплексные и вещественные меры	476
Вещественные положительные меры	478
Решетки	481

3. Продолжение положительной меры. Теория Лебега	489
Внешние меры открытых множеств	490
Внутренняя мера компакта	492
Измеримые множества. Мера множеств	494
Множества нулевой меры	506
Свойства, выполняющиеся почти всюду	508
μ -измеримые функции со значениями в метризуемом сепарабельном пространстве	510
μ -этажные функции	513
Борелевские функции	517
Интеграл от векторной этажной функции	520
Верхний интеграл от вещественной неотрицательной функции	520
Интегрируемость функций с векторными значениями	524
Интеграл Лебега от функции с векторными значениями	524
Интегрируемость и интегралы от функций, определенных почти всюду	533
4. Теорема Лебега о сходимости. Пространство L^1	534
Примеры применений теоремы Лебега	540
Характеристика интегрируемых функций. Интегрируемость и измеримость	551
Теория интегрирования, основанная на свойствах непрерывных и полунепрерывных снизу функций	555
Пространства $\mathcal{L}^p(X, \mu; \vec{F})$	561
Пространства $\mathcal{L}^p(X, \mu; \vec{F})$. Теорема Фишера-Рисса	572
Пространства $\mathcal{L}^\infty(\vec{F})$ и $L^\infty(\vec{F})$	574
Продолжение мер, не обладающих свойством неотрицательности	578
5. Умножение меры на функцию	590
Произведение векторной меры на непрерывную скалярную функцию	590
Элементарные свойства	591
Случай когда μ - вещественная мера ≥ 0	591
Мера с базой μ . Мера с базой ≥ 0	594
Применение к продолжению меры с векторными значениями	610
Применение к интегрируемости функции по нескольким мерам	612
Сопряженность пространств L^p и $L^{p'}$	614
6. Образ меры при отображении	617

Случай, когда H является гомеоморфизмом X на Y	626
Обобщение теоремы 59 на случай, когда μ не ≥ 0	627
Различные примеры образов мер	628
7. Широкая сходимость мер Радона	630
Сходимость по норме. Локальная сходимость по норме	630
Широкая сходимость	632
Функции, ти-интегрируемые по Риману	634
Широкая сходимость и равномерная сходимость	641
Компактные подмножества пространства $C_k(X)$	645
Широкая сходимость последовательности мер к мере Дирака	646
Узкая сходимость последовательности мер конечной нормы	652
Сходимость широкая и сходимость узкая	654
8. Тензорное произведение мер. Кратные интегралы	657
Постановка задачи	657
Существование и единственность тензорного произведения	658
Примеры тензорных произведений	662
Элементарные свойства	663
Носитель меры $\mu \otimes \nu$	663
Случай, когда интегрируемая функция является произведением функции от x и функции от y	673
Окончание доказательства прямого утверждения теоремы	675
Обобщение на произвольные кратные интегралы	677
Широкая сходимость тензорных произведений	679
9. Частные свойства мер Радона на вещественной прямой R	682
Введение символа $\int_a^b d\vec{\mu}$	682
Неопределенные интегралы	683
Функции с ограниченной вариацией на прямой	686
Функция, удовлетворяющая условию Липшица на ограниченном интервале R_1 прямой R , имеет ограниченную вариацию	687
Функции ограниченной вариации и неопределенные интегралы	694
Длина пути в метрическом пространстве	701
Неопределенный интеграл и первообразная	706
Последовательные первообразные непрерывной функции на прямой	711

Формула интегрирования по частям	716
Замена переменных при вычислении простых интегралов	720
Несобственные интегралы на прямой	724
Примеры применения критерия Абеля	732
Главное значение в смысле Коши	736
10. Кратные интегралы на R^n. Длины, площади и объемы в конечномерном аффинном евклидовом пространстве. Замена переменных в кратных интегралах на R^n	742
Измерение объемов в аффинных евклидовых конечномерных пространствах	754
Измерение длин в аффинном евклидовом пространстве	756
Измерение n -мерных площадей в линейном многообразии размерности n аффинного евклидова конечномерного пространства	756
n -мерная площадь n -мерного параметрического многообразия	760
Вычисление объемов с помощью поверхностных интегралов	770
11. Функции, представимые рядами или интегралами	777
Функции, представимые рядами	777
Непрерывность суммы ряда	778
Интегрируемость суммы ряда относительно некоторой меры ≥ 0	778
Дифференцируемость суммы ряда	779
Дифференцируемость бесконечного произведения	787
Функции, представимые интегралами	792
Непрерывность функции, представимой интегралом	792
Интегрируемость функции, представимой интегралом	793
Дифференцируемость функции, представимой интегралом	794
Случай несобственных сходящихся интегралов	799
Применение к делимости дифференцируемых функций	804
Предметный указатель	811

Предметный указатель

Абсолютно мажорируемая векторная мера 488	Абстрактное многообразие размерности n класса C^m 334
- сходящееся бесконечное произведение 165	Аксиома отделимости Хаусдорфа 44
- сходящийся несобственный интеграл Римана 725	Алгебра 127
-- ряд 129, 160	σ -алгебра частей множества 505
	Антисимметричность 21

- Апроксимирующая последовательность 424, 524, 563
- Ассоциативность тензорного произведения 677
- Атлас многообразия 326
- Атомическая мера 441
- Аффинная контингенция 219
- система координат 176
- Аффинное многообразие 176
- нормированное пространство 179
- отображение 178
- пространство 174
- - касательное в точке к многообразию 338
- Аффинно линейная функция 151
- Банахова алгебра 127
- Банахово пространство 117
- Барицентр 184
- Безусловно сходящийся ряд 131
- Бесконечно дифференцируемое многообразие 319
- Бесконечное произведение 162
- Биективное отображение 13
- Биекция 13
- Билинейное отображение 121
- Борелевская σ -алгебра 505
- Борелевское отображение 517
- подмножество 505
- Брахистохрона 369, 384
- Вектор касательный в точке к множеству 218
- Векторная контингенция 219
- мера 450
- Векторное подпространство, касательное к многообразию в точке 220
- поле 222
- пространство, касательное к абстрактному многообразию в точке 338
- - - многообразию в точке 338
- Верхний интеграл 520
- - Римана 421, 635, 653
- предел последовательности 79
- Верхняя огибающая 25
- Вещественная мера Радона 476
- Вещественное многообразие 333
- Взвешенное геометрическое среднее 360
- Внешность множества 48
- Внешняя мера 490, 494
- Внутренность множества 47
- Внутренняя мера 492
- Возрастающее отображение 25
- Вполне нулевая мера 586
- отрицательная мера 586
- положительная мера 586
- упорядоченное множество 23
- Выпуклая функция 204
- Выпуклое множество (часть) 43
- - в аффинном пространстве 184
- Галилеева система координат 195
- Гармоническая функция 415
- Геодезическая кривая 389, 398
- Гиперплоскость 177
- Гиперповерхность 327
- Главное значение в смысле Коши 736
- Гомеоморфизм 54
- Гомеоморфные пространства 55
- Градиент функции 221
- Граница 48
- Двойная последовательность 61
- Двойной интеграл 657
- Делимость дифференцируемых функций 804
- Дизъюнктные меры 588
- C^m -дiffeоморфизм 336
- Дифференциал 213, 217
- Дифференцируемая гиперповерхность 292
- функция 227
- Дифференцируемое многообразие 218, 319
- Дифференцируемость бесконечного произведения 787
- неявной функции 298
- суммы ряда 779

- функции, представимой интегралом 794
- Длина пути 701
- Достаточные условия экстремума 287
- Дуга 90
- Евклидово аффинное пространство 185
 - векторное пространство 185
 - скалярное произведение 184
- Единичный шар 42
- Единственность тензорного произведения мер 658
- Естественная топология 64
- Зависимые функции 350
- Задача Дирихле 415
 - Плато 414
- Заключение 36
- Закон инерции 194
- Замена переменных 16, 394, 720, 742
- Замкнутая кривая 353
 - часть 45
- Замкнутый интервал 26, 46
 - параллелепипед 46
 - шар 41, 46
- Замыкание 49
- Измельчение разбиения 418
- Измеримое множество 500, 501
 - отображение 510, 537
- Изолированная точка 47
- Изотропные векторы 192
- Изотропный конус 192
- Индуктированная метрика 50
 - топология 59
- Интеграл Лебега от функции с векторными значениями 524
 - от векторной этажной функции по мере μ 520
 - - функции, определенной почти всюду 533
 - - - по мере μ 443
 - - Римана 426
 - - на прямой 416
 - - ступенчатой функции с компактным носителем 419
- сходящийся в смысле главного значения Коши 736
- Френеля 735
- Интегрируемая функция с векторными значениями 524
 - μ -интегрируемая функция с векторными значениями 524
- Интегрируемая функция со значениями в пространстве Банаха 423
- Интегрируемость в смысле Лебега 725
 - суммы ряда 778
 - уравнений Эйлера 383
 - функции по нескольким мерам 612
 - - представимой интегралом 793
- Инъективное отображение 13
- Инъекция 13
- Истинное параметрическое представление 321
- Исходное множество отображения 11
- Каноническая инъекция 12
 - мера прямой R 448
 - сюръекция 19
- Кардинальное число 28
- Карта 326
 - множества 326
 - покрывающая точку 326
- Касательная к пути 709
- Касательный вектор к многообразию в точке 344
- Квантор 37
- Класс эквивалентности 18
- Колебание функции в точке 196, 638
 - - на окрестности точки 638
- Компактная часть 68
- Компактное топологическое пространство 69
- Комплексная мера Радона 476
- Комплексное многообразие 333
- Комплексно сопряженная мера Радона 477
- Композиция отображений 15
- Конечное множество отображения 11

- Континуум-гипотеза 35
Кратный интеграл 677
- - на R^n 742
Кривая 327
- Жордана 753
- класса C^m 353
Криволинейная абсцисса 704
Криволинейный интеграл 705
Критерий Абеля 732
- равномерной сходимости 161
- условной сходимости 142
Круг 41
Лемма Хаара 379
Линейное многообразие, касательное
к многообразию в точке 220
- L-непрерывное отображение 579
- отображение 110
Линейно связное пространство 92
Локальная карта класса C^m 326
- равномерная сходимость
бесконечного произведения 168
- сходимость последовательности
мер Радона по норме 631
Локально компактное пространство
74
- линейно связное пространство 95
- равномерно сходящаяся
последовательность функций
153
- связное пространство 94
Локальный гомеоморфизм 312
Мажоранта 23
Мажорируемая часть 23
Максимум 23, 25
Матрица Якоби 211
Мера Дирака 441
- - в точке 440
- допускающая борелевское
продолжение 583
- Лебега 448
- множества 501
- плотности р по отношению к
вещественной мере 594
- Радона 440
- - - на вещественной прямой 682
- - - локально компактном
пространстве 447
- - - равная нулю 465
- с базой 594
- μ -сингулярная 589
- сосредоточения на части
пространства 589
 n -мерная площадь n -мерного
параметрического
многообразия 760, 763
Метризуемое топологическое
пространство 59
Метрическое пространство 40
- свойство 58
Минимум 24
Миноранта 23
Минорируемая часть 23
Многообразие, дифференцируемое m
раз 319
- принадлежащее классу C^m 320
- размерности n , принадлежащее
классу C^m 320
- m раз непрерывно
дифференцируемое 319
Множество 9
- линейных отображений 15
- непрерывных отображений 15
- нулевой меры 506
- отображений 15
Множитель Лагранжа 353
Монотонная функция 202
Мощность 27
- континуума 33
Натуральная параметризация 704
Независимые функции 350
Необходимые условия экстремума
285, 287
Неограниченно удаляющаяся
последовательность точек 619
Неопределенный интеграл 694, 706
- - от меры Радона 683

- - - функции по мере 684
- Неподвижная точка отображения 107
- Непрерывная функция многих переменных 65
 - - с компактным носителем 463
- Непрерывно дифференцируемая функция 227
- дифференцируемое многообразие 319
- Непрерывное в бесконечности отображение 618
 - отображение 52
 - слева отображение 195
 - справа отображение 195
- Непрерывность суммы ряда 778
 - функции, представимой интегралом 792
- Неравенство выпуклости 562
 - Гёльдера 364, 567
 - Коши - Буняковского - Шварца 43
 - Коши - Шварца 366
 - Минковского 361, 562
 - счетной выпуклости 546, 562, 565
- Несобственный интеграл 724
 - путь 703
- Нечетность 736
- Неявная функция 293
- Нижний предел последовательности 79
- Нижняя огибающая 25
- Норма 41
 - билинейного отображения 124
 - линейного отображения 113
- Нормальная система уравнений 331
- Нормальное пространство 462
- Нормально сходящийся ряд 129, 160
- Норма меры 449
 - порядка p 359, 561
- Нормированная алгебра 127
- Нормированное векторное пространство 42
- Носитель 417
 - меры 663
 - - Радона 466
- Обобщенное евклидово пространство 192
- эрмитово пространство 192
- Образ карты 326
 - - меры при отображении 617, 618, 621
 - множества при отображении 13
 - отображения 113
 - параметрического многообразия 352
- Обратное отображение 139
- Общая, или аффинная, система координат 176
 - теорема Лебега 539
- Объединение 10
- Ограниченнная часть 23
- Окрестность точки 46
- Окружность 41
- Ортонормированный базис 190
- Основная теорема измерения объемов 754
- Особое многообразие размерности n класса C^m 352
- Отделимое пространство 58
- Открытое множество 44
 - отображение 311
- Открытый интервал 26
 - параллелепипед 45
 - прямоугольник 64
 - шар 41, 45
- Относительно компактная часть 68
- Относительный максимум 285
 - минимум 285
 - экстремум 393
- Отношение порядка 21
 - - эквивалентности 17
- Отображение 11
 - μ -измеримое 510
 - класса C^k 336
 - непрерывное в точке 52
 - μ -собственное 623
- Отрезок 43
- Отрицательная мера Радона 478

- часть элемента 481
- Паракомпактное пространство 453
- Параллельный перенос меры Радона 632
- Параметрическое многообразие размерности n класса C^m 352
- Параметрические многообразия, C^m - эквивалентные 353
- Первообразная 706
- Пересечение множества частей 10
- Перестановка 13
 - членов ряда 130
 - Σ и \int для рядов с положительными членами 547
- Плотная часть 50
- Плотность по отношению к мере Лебега 449
- Площадь n -мерная 756
- Поверхность 327
- Подмножество (или часть) 9
- Подпокрытие 67
- Покрытие 67
- Полная вариация последовательности 142
 - - функции 686
 - дифференцируемость 253, 266
 - масса 474
 - производная 213
 - решетка 482
- Полное метрическое пространство 101
- Полный дифференциал 213
- Положительная мера Радона 478
 - часть элемента 481
- Полулинейная форма 187
- Полулинейное отображение 187
- Полунорма 564
- Полуоткрытый интервал 26
- Полусходящийся несобственный интеграл Римана 725
- Полуторалинейная форма 187
- Полуупорядоченное векторное пространство 480
- Пополненная прямая 26
- Последовательность 15
 - имеющая ограниченную вариацию 142
 - Коши 101
 - мер Радона, широко сходящаяся к предельной мере 632
 - L-сходящаяся 578
- Постоянное отображение 13
- Правило нахождения относительного максимума или минимума 289
- Правильная функция 197
- Предел последовательности 60
 - - функций 777
- Преобразование 13
- Принцип кусочной склейки мер 475
- Продолжение меры 473
 - - с векторными значениями 610
 - положительной меры 489
 - равномерно непрерывных отображений 104
- Проекция 12
- Произведение меры на функцию 590, 592
 - множеств 11
 - нормированных векторных пространств 118
 - топологий 64
 - числовых рядов 137
- Производная 199
 - билинейного непрерывного отображения 225
 - матрица 211
 - мера 698
 - отношения 238
 - отображения 198
 - - слева 198
 - - справа 198
 - порядка m 199
 - произведения 236, 268
- Производное отображение 213
- Производный вектор 215
- Прообраз множества при отображении 13

- функции при замене переменных 16
Простая сходимость
последовательности функций 149
Пространство n раз
дифференцируемых функций 267
Псевдоизмеримое множество 500, 501
- подмножество 496
Пустой носитель меры Радона 466
Путь 90
Равномерно дифференцируемая функция 253
- непрерывное отображение 86
- стягивающееся множество 748
- сходящаяся последовательность 150, 152
- сходящийся ряд 160
Разбиение 418
Раздельно непрерывная функция 67
Разложение единицы 452
- - подчиненное покрытию 452
Разложимая функция 463, 654
Рассеянная мера 443
Расстояние 40, 82
Расходящееся бесконечное произведение 163
Регулярное пространство 59
Решетка 481
Рефлексивность 17, 21
Свойство Больцано - Вейерштрасса 76
- выполняющееся почти всюду 508
- - μ -почти всюду 508
- Гейне - Бореля 68
- Лузина 536
Связанные точки 93
Связная компонента 93
Связное пространство 88
Сдвиг 150
Сепарабельное пространство, 50, 510
Сжатие 107

Симметричное билинейное отображение 259
- n -линейное отображение 262
Симметричность 17
Скачок функции в точке 196
Собственное отображение 618
Совершенное множество 507
Соотношение Шаля 175, 683
Сопряженное пространство 115
Спрямляемый несобственный путь 703
Среднее значение функции на интервале 438
- порядка p 359
Стационарная точка 287
Строгий относительный максимум 285
-- минимум 285
Строго возрастающее отображение 25
- убывающее отображение 25
Ступенчатая функция 418
Сумма безусловно сходящегося ряда 131
- Коши - Римана 436
- ряда 128
Суммирование по блокам безусловно сходящегося ряда 134
Существенная верхняя грань 574
Существенно ограниченная функция 574
Существование тензорного произведения мер 661
Сфера 41
Сходимость мер Радона по норме 631
Сходящаяся двойная последовательность 61
- последовательность 60
- - мер Радона 631
Сходящееся бесконечное произведение 162
Сходящийся несобственный интеграл 724
- ряд 128

- Счетное множество 27, 31
Сюръективное отображение 13
Сюръекция 13
Трансфинитное число 28
Трансцендентное число 33
Тензорное произведение мер 658
Теорема Абеля 143, 732
- Банаха-Штейнгауза 641
- Бернштейна 27
- Бэра 641
- - Вейерштрасса о полиномиальной аппроксимации 680
- Даламбера 84
- Данфорда - Петтиса 594
- Егорова 534
- Лебега - Никодима 594
- - о сходимости 534
- об интегрировании по частям 728
- - ортогональной проекции площадей гиперплоскостей 759
- о замене переменной 735
- - неподвижной точке 108
- - неявной функции 305
- - перестановке интеграла и линейного непрерывного отображения 430
- - переходе через таможню 92
- - плотности 659
- - постоянном ранге 345
- - применении билинейного непрерывного отображения 567
- - - линейного непрерывного отображения 566
- - сложной функции 230
- - среднем 433
- Рисса 106
- Ролля 200
- Фату 542
- Фишера - Рисса 573
- Фубини - Лебега 664
- - Фату 670
Тождественное отображение 12
Топологическая группа 66
Топологическое векторное пространство 66
- множество 320
- подпространство 59
- произведение 63
- пространство 58
- свойство 58
Топология пополненной прямой 60
Точка прикосновения 49
- разрыва первого рода 196
- сгущения 75
Точная верхняя грань 24
- нижняя грань 24
Транзитивность 17, 21
- образов мер 624
Убывающее отображение 25
Узкая сходимость последовательности мер конечной нормы 652
- - - Радона 654
Узко сходящаяся последовательность мер 653
Универсальный атлас 326
Уравнение колебания струны 239
- Эйлера 378
Уравнения Гамильтона 408
Условие 36
- Гёльдера 87
- Липшица 87
Условия трансверсальности 400
Условно сходящееся бесконечное произведение 165
- - сходящийся ряд 129
Условный максимум 353
- минимум 353
Факторгруппа 19
Фактормножество 19
Факторпространство 20
Формула интегрирования по частям 716
- конечных приращений 200, 247
- Лейбница 268
- Тейлора 201, 272, 276, 279, 292, 714,
715

- Фундаментальная последовательность 100
- система окрестностей 47
- Функции, μ -эквивалентные в смысле Лебега 572
- Функциональное пространство 146
- Функция 11
- Гамильтона 407
 - длины 702
 - имеющая интегрируемую р-ю степень 563
 - локально ограниченную вариацию 694
 - интегрируемая по Риману 423
 - μ -интегрируемая 610
 - по Риману 634
 - Лагранжа 407
 - полуунпрерывная сверху 85
 - снизу 85
 - представимая интегралом 777, 792
 - рядом 777
 - принадлежащая классу C^m 199
 - m раз непрерывно дифференцируемая 199, 266
 - Римана 168
 - с ограниченной вариацией 686
 - μ существенно ограниченная 574
 - удовлетворяющая условию Гельдера 87
 - - - Липшица 87
 - характеристическая 30
- Хевисайда 684
- Частная дифференцируемость 253, 266
- производная вдоль вектора 209
 - - отображения 224
 - сумма с индексом n 128
 - теорема Лебега 538
- Часть (подмножество) множества 9
- Четность 736
- Шар 41
- Широкая сходимость последовательности мер к мере Дирака 646
- - - Радона 630, 632, 654
 - тензорных произведений 679
- Эквивалентные метрики 56
- нормы 56
 - пути 702
 - функции 572
- Экстремаль 378
- Экстремум 285
- Эрмитова положительно определенная матрица 192
- Эрмитово аффинное пространство 187
- векторное пространство 187
 - скалярное произведение 187
- Этажная функция 513
- μ -этажная функция 513
- Ядро отображения 113
- Якобиан 212

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Мы предлагаем вниманию читателей двухтомный курс анализа, написанный известным французским математиком Лораном Шварцем. Задачи, которыеставил перед собой автор при написании курса, достаточно ярко описаны в его предисловии. Формально для чтения книги предварительных знаний не требуется, однако, по существу она доступна лишь тем, кто знаком с обычным курсом математического анализа и имеет определенную математическую культуру.

В курсе Л. Шварца излагается ряд математических дисциплин, которые по установившейся в нашей системе преподавания традиции изучаются раздельно. Мы считаем, что объединение всех этих дисциплин в единый курс вполне отвечает духу современной математической науки. Заметим, что на самом деле — это возврат на новом уровне к старым традициям: достаточно вспомнить классический курс математического анализа Э. Гурса или курс математики В. И. Смирнова.

В первой главе курса излагается элементарная теория множеств. Вторая глава содержит основы теоретико-множественной топологии и элементы функционального анализа, которые в дальнейшем широко используются. В третьей главе развивается аппарат многомерного дифференциального исчисления. Здесь же вводятся понятия, связанные с дифференцируемыми многообразиями. Последний параграф этой главы посвящен основам вариационного исчисления.

Четвертая глава содержит развернутое изложение теории интеграла и меры. После краткого изложения теории интеграла Римана вводится понятие меры Радона как непрерывного функционала на пространстве финитных непрерывных функций, которое лежит в основе всей теории интеграла. В главе, кроме общей теории, имеются и приложения интегрального исчисления.

Сравнительно небольшая пятая глава посвящена теории дифференциальных уравнений. В ней содержатся общие теоремы и теория линейных уравнений. (В русском издании в целях более равномерного распределения материала она перенесена из первого тома во второй.)

Большое место в курсе занимает изложение внешнего дифференциального исчисления (глава шестая). Это весьма отрадно, так как важный аппарат дифференциальных форм незаслуженно медленно пробивает себе дорогу в нашу учебную литературу. Здесь же освещены связи между теорией дифференциальных форм и алгебраической топологией.

В последней, седьмой главе изучаются функции комплексных переменных. Преимущественно исследуются функции одной комплексной переменной, правда, со значениями в банаховом пространстве, в связи с чем затрагивается ряд вопросов функционального анализа. Заметим, что в этом отношении седьмая глава не составляет исключения: всюду, где только возможно, автор рассматривает векторнозначные функции.

Лоран Шварц является превосходным лектором, и академический стиль его книги часто прерывается живой речью. Мы старались, чтобы эта эмоциональная речь автора не потерялась при переводе. Мы надеемся, что его курс будет весьма полезным пособием для студентов математических и физических факультетов университетов, для аспирантов и преподавателей вузов, а также для студентов технических вузов, желающих более глубоко изучить математику.

C. Г. Крейн

ПРЕДИСЛОВИЕ

Как часто говорят, для физиков и инженеров «нет математики без слез». Физикам и современным инженерам нужны необъятные математические познания в самых разнообразных областях. Абсолютно исключено, чтобы эти «прикладники» знали все нужные им результаты со всеми доказательствами на уровне строгости, принятом в математике. И вот сложилась такая ситуация. Имеются краткие курсы, содержащие сравнительно немного результатов, но доказанных весьма обстоятельно,— эти книги удовлетворяют математиков, но не физиков. Имеются также краткие курсы, богатые результатами, но с доказательствами, либо едва намеченными, либо вовсе отсутствующими,— эти книги не удовлетворяют читателей, воспитанных в декартовом духе.

Мы избрали третью возможность. Курс, который мы предлагаем, велик (пожалуй даже слишком велик), он содержит много теорем, и притом с достаточно полными доказательствами. Это, собственно говоря, не учебник, а книга, содержащая фактический, документальный материал.

Лекции, основанные на этой книге, должны сводиться лишь к кратким резюме. Студентам же для обязательного изучения следует каждый раз выбирать строго определенную часть книги, которая содержит много формулировок, но мало доказательств.

Учащиеся должны научиться схватывать новые идеи и структуры, понимать суть теорем и применять их там, где нужно, хотя бы с книгой в руках. Это не так легко, как кажется,— тот, кто никогда не думал над теоремой, не сможет и сразу применить ее, даже если у него есть книга. Подробно изучать нужно лишь те доказательства, которые являются наиболее типичными и поучительными.

Кроме того, учащиеся могут (и мы активно это рекомендуем) выбрать для факультативного изучения какой-либо из разделов курса в соответствии с собственными вкусами или recommendationами преподавателей. При этом можно удовлетворить различные вкусы и учесть разные уровни подготовки. Если студенты одного потока изберут не совсем одинаковые разделы, это можно только приветствовать.

Лоран Шварц

ГЛАВА I

Теория множеств

§ 1. МНОЖЕСТВА. ЭЛЕМЕНТАРНЫЕ ОПЕРАЦИИ

Множеством называется совокупность некоторых объектов. Примеры: Множество учащихся одного выпуска,

множество точек плоскости,

множество невырожденных поверхностей 2-го порядка в трехмерном пространстве,

множество \mathbb{N} неотрицательных целых чисел,

множество \mathbb{Z} произвольных целых чисел,

множество \mathbb{Q} рациональных чисел,

множество \mathbb{R} вещественных чисел,

множество \mathbb{C} комплексных чисел.

Части множества

Запись $x \in E$ означает: « x является элементом множества E ¹⁾». Если множество A состоит из элементов, принадлежащих некоторому другому множеству E , то оно называется *частью* или *подмножеством* этого множества. Множество эллипсоидов, например, является частью множества поверхностей второго порядка.

Частьями множества E являются: оно само; его пустая часть, обозначаемая через \emptyset ; часть, сводящаяся к одному элементу a , обозначаемая через $\{a\}$; часть, образованная тремя элементами a, b, c , обозначаемая через $\{a; b; c\}$. Не следует смешивать a как элемент E и $\{a\}$ как часть E , состоящую из одного элемента a . Для того чтобы обозначить множество элементов, обладающих некоторым свойством P , часто пользуются такой записью: $\{x; x \text{ обладает свойством } P\}$. Например, $\{x; x \in \mathbb{R}, x \geq 0\}$ означает множество вещественных чисел ≥ 0 , $\{\arg \operatorname{tg} x; x \in \mathbb{R}, e \leq x \leq 2e\}$ означает множество значений $\arg \operatorname{tg} x$, когда x принимает все вещественные значения между e и $2e$ включительно.

Через $\mathfrak{P}(E)$ обозначают *множество всех частей множества E* . Это некоторое новое множество, которое можно образовать,

¹⁾ Множество иногда называется также пространством, а его элементы часто называются точками.

исходя из множества E . Можно затем рассмотреть $\mathfrak{P}(\mathfrak{P}(E))$ и т. д.

Если E содержит n элементов, то $\mathfrak{P}(E)$ содержит 2^n элементов¹⁾.

Отношение включения. Дополнение

Пусть X и Y — две части множества E . Говорят, что X содержится в Y , или что Y содержит X , если каждый элемент из X является элементом из Y ; при этом пишут $X \subset Y$ или $Y \supset X$. Так, например, $\emptyset \subset X \subset E$, $\emptyset \subset \emptyset$, $E \subset E$. Если $X \subset Y$ и $Y \subset Z$, то $X \subset Z$.

Множество элементов E , не принадлежащих некоторой его части A , называется дополнением к A в E и обозначается через $C_E A$ или просто CA , когда это не может привести к путанице. Очевидно, $CCA = A$, и если $A \subset B$, то $CA \supset CB$. Если A и B — две части множества E и $A \supset B$, то через $A - B$ обозначают иногда множество элементов из A , не принадлежащих B .

Объединение. Пересечение

Объединением некоторого множества частей из E называют часть E , состоящую из элементов, принадлежащих по крайней мере одной из этих частей. Объединение частей A , B , C записывается в виде $A \cup B \cup C$.

Объединение семейства частей A_i , занумерованных с помощью некоторого множества индексов I , обозначается через $\bigcup_{i \in I} A_i$.

Пересечением множества частей E называется часть E , состоящая из элементов, принадлежащих одновременно всем этим частям. Для пересечения применяются аналогичные обозначения: $A \cap B \cap C$, $\bigcap_{i \in I} A_i$.

¹⁾ Через C_n^p , или лучше $\binom{n}{p}$, обозначается число частей, содержащих по p элементов некоторого множества, состоящего из n элементов. Если теперь подсчитать все части E (не пропуская ни \emptyset , ни самого E), то получим:

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n-1} + \binom{n}{n} = 2^n.$$

Действительно, если $E_n = \{1, 2, \dots, n\}$, то $\mathfrak{P}(E_n)$ для $n \geq 1$ можно считать состоящим из двух различных множеств: множества частей, содержащих элемент n , и множества частей, не содержащих его. Количество элементов в каждом из них такое же, как и в множестве $\mathfrak{P}(E_{n-1})$. Если $\mathfrak{P}(E_n)$ содержит A_n элементов, то $A_n = 2A_{n-1}$, и поскольку $A_0 = 1$ (E_0 — пустое множество, имеющее ровно одну часть \emptyset), то $A_n = 2^n$.

Две части называются непересекающимися, если их пересечение пусто.

Дополнение пересечения некоторого семейства частей совпадает с объединением их дополнений, и дополнение объединения семейства частей есть пересечение их дополнений:

$$\mathbf{C}(A \cap B) = \mathbf{C}A \cup \mathbf{C}B \quad \text{и} \quad \mathbf{C}(A \cup B) = \mathbf{C}A \cap \mathbf{C}B. \quad (I, 1; 1)$$

Преобразование, которое каждой части из E ставит в соответствие ее дополнение, является, следовательно, соответствием, переводящим \subset в \supset , \supset в \subset , \cup в \cap , \cap в \cup .

Произведение множеств

Z Произведением $E \times F$ двух множеств E и F называется множество всевозможных упорядоченных пар (x, y) , образованных из элементов x множества E и элементов y множества F . При этом пары (x, y) и (y, x) с $x \neq y$ считаются различными. Это особенно важно иметь в виду, когда множества E и F совпадают.

Точно так же можно определить произведение нескольких множеств или произвольного семейства множеств. Множества $(E \times F) \times G$, $E \times (F \times G)$ и $E \times F \times G$ отождествляются, так что произведение ассоциативно.

Произведения $E \times E$, $E \times E \times E$ и т. д. обозначаются также через E^2 , E^3 и т. д. Отсюда возникло обозначение \mathbb{R}^n для произведения n множеств, идентичных множеству \mathbb{R} всех вещественных чисел. Точка из \mathbb{R}^n является, следовательно, упорядоченной системой произвольных n вещественных чисел (x_1, x_2, \dots, x_n) .

§ 2. ОТОБРАЖЕНИЯ. ФУНКЦИИ

Пусть заданы два множества E и F . Отображением E в F , или функцией, определенной на E со значениями в F , называется соответствие f , которое каждому элементу x из E относит некоторый элемент из F , обозначаемый через $f(x)$. Обозначение $E \xrightarrow{f} F$ будет означать, что f является отображением E в F . Множество E называется исходным множеством, а множество F конечным множеством отображения. Следует различать отображение f и элемент $f(x)$, соответствующий x при этом отображении. Однако по чисто практическим соображениям при изложении это различие иногда не принимается во внимание. Так, например, неправильно (но удобно) говорить «функция $\sin x$ », хотя следовало бы читать «функция \sin », поскольку $\sin x$ является значением этой функции в точке x . Эту неточность обычно исправляют, говоря, что «задана

функция $x \rightarrow \sin x$, или «функция f , определенная равенством $f(x) = \sin x$ ».

Отображение множества в себя называют также оператором.

Примеры отображений

1°) Функция f , определенная формулой $f(x) = 1/x$, является отображением множества отличных от нуля элементов вещественной прямой \mathbb{R} в \mathbb{R} . Будет неточным говорить, что это — функция вещественной переменной, поскольку здесь переменная не может принимать все вещественные значения. Однако вполне возможно рассматривать отображение \mathbb{R} в \mathbb{R} , определенное соотношениями $g(x) = 1/x$ для $x \neq 0$ и $g(0) = 2$. Эта функция g отлична от функции f , так как имеет иную область определения. Можно также рассматривать отображение h , определенное как $h(x) = 1/x$ для $x \neq 0$ и $h(0) = +\infty$, множества $E = \mathbb{R}$ в множество F , образованное из \mathbb{R} и некоторого дополнительного элемента, обозначаемого $+\infty$.

2°) Если E — множество вещественных функций $\varphi(x)$, определенных и интегрируемых на вещественном интервале $[a, b]$, то интеграл $\int_a^b \varphi(x) dx$ является отображением E на вещественную прямую \mathbb{R} .

3°) Если E является множеством кривых конечной длины на евклидовой плоскости, то можно определить отображение E на полупрямую \mathbb{R}_+ (множество элементов ≥ 0 из \mathbb{R}), которое каждой кривой ставит в соответствие ее длину.

Отображение f множества E в E , определенное равенством $f(x) = x$, называется *тождественным*.

Если E составляет часть F , то отображение E в F , определенное равенством $f(x) = x$, называется *канонической инъекцией* E в F .

Если $E \times F$ — произведение двух множеств, то отображение $E \times F$ в E , ставящее в соответствие каждой паре $(x, y) \in E \times F$ элемент $x \in E$, называется *проекцией на* E . Точно так же определяется *проекция на* F .

Пусть E, F, G — три множества. Отображение f множества $E \times F$ в множество G каждой паре $(x, y) = z \in E \times F$ ставит в соответствие некоторый элемент из G , обозначаемый через $f(z) = f((x, y))$. Обычно двойную скобку заменяют на простую $f(x, y)$ и говорят, что f является функцией двух переменных. Например, если E является множеством вещественных функций вещественной переменной, интегрируемых на каждом конечном

интервале, то интеграл $\int_a^b \varphi(x) dx$ определяет некоторое ото-

бражение $E \times \mathbb{R} \times \mathbb{R}$ в \mathbb{R} , поскольку это отображение является функцией трех переменных: $\varphi \in E$, $a \in \mathbb{R}$ и $b \in \mathbb{R}$.

Рассмотрим теперь отображение E в $F \times G$. Оно имеет вид $x \rightarrow (f(x), g(x))$, где f (соответственно g) является отображением E в F (соответственно в G). Задание такого отображения эквивалентно заданию системы двух функций. В общем случае задание отображения некоторого произведения $E_1 \times E_2 \times \dots \times E_n$ в произведение $F_1 \times F_2 \times \dots \times F_m$ равносильно заданию системы m функций от n переменных.

Отображение f множества E в множество F называется *постоянным*, если $f(x)$ для всех $x \in E$ является одним и тем же элементом из F .

Инъекции, сюръекции, биекции

Говорят, что отображение f множества E в множество F *инъективно*, или является *инъекцией*, если два различных элемента из E имеют образами при отображении f два различных элемента F . Каноническая инъекция некоторого подмножества в само множество является инъективным отображением.

Отображение f называют *сюръективным*, или *сюръекцией*, если каждый элемент из F является образом при отображении f по крайней мере одного элемента из E .

Отображение f называют *биективным*, или *биекцией*, если каждый элемент F является образом при отображении f некоторого, и притом единственного, элемента из E . Отображение биективно тогда и только тогда, когда оно одновременно инъективно и сюръективно. Биекция множества на себя называется также *перестановкой*, или *преобразованием*.

Пусть f — некоторая биекция, и пусть $y \in F$. Обозначим через $f^{-1}(y)$ единственный элемент x из E , такой, что $f(x) = y$. Тем самым мы определили некоторое отображение f^{-1} множества F в E . Это — снова биекция. Ее называют обратным отображением, или обратной биекцией к f . Часто ее также называют просто обратной функцией.

Образ и прообраз подмножества

Пусть f — отображение E в F и A — подмножество E . Обозначим через $f(A)$ подмножество F , образованное из всех элементов $f(x)$, $x \in A$. Очевидно, $f(\emptyset) = \emptyset$. Исходя из отображения f , мы тем самым определим некоторое отображение $\bar{A} \rightarrow f(A)$ множества $\wp(E)$ в множество $\wp(F)$. Это отображение сохраняет символы \subset , \supset , \cup в том смысле, что

$$\begin{aligned} \text{если } A \subset B, \text{ то } f(A) \subset f(B), \\ f(A \cup B) = f(A) \cup f(B). \end{aligned} \tag{I, 2; 1}$$

Однако символы \cap , \subset при этом отображении не сохраняются, так как имеет место лишь включение

$$f(A \cap B) \subset f(A) \cap f(B)^1). \quad (\text{I}, 2; 2)$$

Часть $f(A)$ называется образом части A при отображении f .

Пусть теперь B является некоторой частью множества F . Будем обозначать через $f^{-1}(B)$ часть E , образованную из всех таких элементов x , что $f(x) \in B^2$. Очевидно, $f^{-1}(\emptyset) = \emptyset^3$. Мы связали с отображением f отображение $B \rightarrow f^{-1}(B)$ множества $\mathfrak{P}(F)$ в множество $\mathfrak{P}(E)$.

Это отображение сохраняет 5 символов \subset , \supset , \cup , \cap , \mathbf{C} в том смысле, что:

$$\text{если } A \subset B, \text{ то } f^{-1}(A) \subset f^{-1}(B),$$

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B), \quad f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B), \quad (\text{I}, 2; 3)$$

$$f^{-1}(\mathbf{C} A) = \mathbf{C} f^{-1}(A).$$

Мы видим, следовательно, что так определенное отображение f^{-1} проще описанного выше отображения f . Часть $f^{-1}(B)$ называется прообразом множества B при отображении f . Уместно заметить, что это определение вовсе не предполагает биективности f . Всё всяком случае, если $y \in F$, то можно говорить о $f^{-1}(\{y\})$, но это — некоторая часть E , а не элемент E . Эта часть может содержать более одного элемента, если f не является инъективным, и может оказаться пустой, если f не сюръективно⁴). Если f биективно, то $f^{-1}(\{y\}) = \{f^{-1}(y)\}$.

Символ $f^{-1}(B)$, кроме того, может иметь два возможных одинаковых значения: это прообраз B при отображении f или образ множества B при биективном обратном отображении f^{-1} . Естественно, если отображение f биективно, то образ также сохраняет все 5 символов: \subset , \supset , \cup , \cap , \mathbf{C} . Кроме того, имеем:

$$\begin{aligned} f^{-1}(f(A)) &\supseteq A \quad \text{для } A \in \mathfrak{P}(E), \\ f(f^{-1}(B)) &\subset B \quad \text{для } B \in \mathfrak{P}(F). \end{aligned} \quad (\text{I}, 2; 3_2)$$

¹⁾ Пусть f — постоянное отображение, т. е. $f(x) = b$ при любом $x \in E$. Пусть A и B — непересекающиеся части E . Так как $A \cap B = \emptyset$ то $f(A \cap B) = \emptyset$ что не совпадает с множеством $f(A) \cap f(B) = \{b\}$.

²⁾ Согласно кратким обозначениям, введенным на стр. 9,

$$f^{-1}(B) = \{x; f(x) \in B\}, \quad f(A) = \{f(x); x \in A\}.$$

³⁾ Может оказаться, что $f^{-1}(B) = \emptyset$ при $B \neq \emptyset$. Если, например, f является отображением $x \rightarrow x^2$ множества \mathbb{R} в \mathbb{R} , то $f^{-1}(\{-1\}) = \emptyset$.

⁴⁾ См. примечание 3.

Множество отображений. Семейства. Последовательности

Если E и F — два множества, то можно говорить о некотором новом множестве — множестве отображений E в F . Если E имеет только 2 элемента a_1 и a_2 , то множество отображений E в F может быть биективно отображенено на квадрат F^2 , ибо каждое отображение E в F полностью определяется парой $(x, y) \in F^2$ образов x, y элементов a_1, a_2 из E . Если E состоит из n элементов a_1, a_2, \dots, a_n , то множество отображений E в F можно биективно отобразить на F^n , ибо каждое такое отображение эквивалентно заданию системы $(x_1, x_2, \dots, x_n) \in F^n$ образов элементов a_1, a_2, \dots, a_n при этом отображении. Поэтому *множество отображений E в F принято обозначать через F^E* .

Естественно в дальнейшем рассматривать различные подмножества множества F^E : множество *непрерывных отображений E в F* , если E и F — метрические пространства; множество *линейных отображений E в F* , если E и F — векторные пространства и т. д. *Семейство $(x_i)_{i \in I}$ элементов из E , снабженных индексами из множества I ,* можно рассматривать как отображение I в E . Множество семейств элементов из E , снабженных индексами из I , является не чем иным, как множеством E^I отображений I в E .

В частности, *последовательность элементов из E* можно определить как семейство элементов из E , снабженных индексами из множества \mathbb{N} целых чисел ≥ 0 , или же как отображение \mathbb{N} в E . Множество всех последовательностей элементов из E является множеством $E^{\mathbb{N}}$. Можно также говорить о последовательностях с индексами из множества \mathbb{N}_1 целых чисел ≥ 1 , или о *конечной последовательности*, снабженной индексами из конечного множества целых чисел $1, 2, \dots, n$. Всегда будет важным уточнять смысл слова «последовательность», когда множество индексов не совпадает с \mathbb{N} .

Композиция отображений

Пусть E, F, G — три множества, и пусть f — некоторое отображение E в F , а g — отображение F в G . Композицией $g \circ f$ называется отображение E в G , определенное формулой $g \circ f(x) = g(f(x))$. Заметим, что запись $g \circ f$ производится в порядке, обратном тому, в котором производятся операции:

$$E \xrightarrow{f} F \xrightarrow{g} G. \quad (\text{I}, 2; 4)$$

Таким образом, в математике принято правило, согласно которому в композиции операций $g \circ f$ надо начинать с операцией f , расположенной справа.

Если A — часть E , то $g \circ f(A) = g(f(A))$. Если B — часть G , то $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$. Композиция отображений ассоциативна, т. е. если f , g , h — отображения E в F , F в G и G в H соответственно, то $(h \circ g) \circ f = h \circ (g \circ f)$, что проще записывается в виде $h \circ g \circ f$.

Если f^{-1} является биекцией, обратной к биекции f множества E на F , то $f^{-1} \circ f = I_E$, где I_E — тождественное отображение E , и $f \circ f^{-1} = I_F$, где I_F — тождественное отображение множества F . Обратно, если отображение f множества E в F и отображение g множества F в E таковы, что $g \circ f = I_E$ и $f \circ g = I_F$, то f является биекцией, а g — его обратной биекцией.

Если f — биекция E на F , а g — биекция F на G , то $g \circ f$ является биекцией E на G , а ее обратной биекцией будет $f^{-1} \circ g^{-1}$.

Пусть f — отображение E в F и A — некоторая часть E . Отображение f_A множества A в F , определяемое формулой $f_A(x) = f(x)$ для $x \in A$, называется *сужением* f на A и часто обозначается через $f|A$. Говорят также, что f является *продолжением* на E отображения f_A множества A в множество F . Чаще всего в этом случае продолжают писать f вместо f_A .

Точно так же, если f — некоторое отображение E в F и если $f(E) \subset B$, то f определяет некоторое отображение f_B множества E в множество B , задаваемое формулой $f_B(x) = f(x)$. Практически вместо f_B всегда пишут f .

Замена переменных и замена функций

Пусть f — функция, определенная на E , со значениями в F . Если u является отображением некоторого множества E_1 в множество E , то можно построить новую функцию $f_1 = f \circ u$, определенную на E_1 , со значениями в F . Говорят, что тем самым совершена замена *переменной* u или замена исходного множества $E_1 \xrightarrow{u} E$ и что f_1 является *прообразом* f при этой замене *переменных*. Производя в выражении $f(x)$ подстановку $x = u(x_1)$, получают выражение $f_1(x_1)$. Часто, если не указывается сама замена переменной u , функция f_1 обозначается через u^*f или даже через f^* . Таким образом,

$$u^*f(x_1) = f(u(x_1)) \quad \text{или} \quad f^*(x_1) = f(u(x_1)). \quad (\text{I}, 2; 5)$$

При беглом изложении, допуская вольность речи, иногда отождествляют f_1 и f , говоря, что это та же самая функция (*sic!*), представимая через переменную x_1 вместо переменной x , но такое упрощение изложения опасно, так как может привести к серьезным противоречиям.

Если теперь v является некоторым отображением F в множество F_2 , то можно определить новую функцию $f_2 = v \circ f$, определенную на E , со значениями в F_2 . Говорят, что произве-

дена замена v функции или замена множества значений $F: F \xrightarrow{v} F_2$ и что f_2 является образом f при этой замене.

Можно произвести одновременно и замену функции, рассматривая $f_3 = v \circ f \circ u$. Здесь f_3 является образом f при замене переменной u и замене функции v .

§ 3. ОТНОШЕНИЯ ЭКВИВАЛЕНТНОСТИ. ФАКТОРМНОЖЕСТВО

Говорят, что на множестве E определено некоторое бинарное отношение, если задано множество всех пар (x, y) элементов из E , удовлетворяющих этому отношению. Так, например, отношения $x = y$, $x \leq y$, $x > y$, $x^2 = y^3$, $(x^2 + y^2 = 1, x \neq 0)$ являются бинарными отношениями между вещественными числами. Очевидно, бинарное отношение на E есть не что иное, как подмножество R произведения $E \times E$.

В этом и следующем параграфах мы будем изучать бинарные отношения, имеющие особые значения.

Бинарное отношение R называется *отношением эквивалентности*, если оно обладает следующими тремя свойствами:

- рефлексивность*: $(x, x) \in R$ при любом $x \in E$,
- симметричность*: из $(x, y) \in R$ следует $(y, x) \in R$, (I, 3; 1)
- транзитивность*: если $(x, y) \in R$ и $(y, z) \in R$, то $(x, z) \in R$.

Вместо того чтобы писать $(x, y) \in R$ часто пишут $x \sim_R y$ или $x \equiv y \pmod{R}$ (читается: « x конгруэнтно y по модулю R ») или, проще, $x \sim y$ и $x \equiv y$, если нет необходимости каждый раз указывать, что речь идет об одном и том же отношении R .

Примеры. Отношения, приведенные в начале параграфа, кроме первого, отношениями эквивалентности не являются. Легко доказать, что все перечисленные ниже бинарные отношения являются отношениями эквивалентности.

1°) Будем исходить из множества \mathbb{Z} целых чисел произвольного знака. В качестве множества E возьмем подмножество множества $\mathbb{Z} \times \mathbb{Z}$ пар (p, q) , у которых $q \neq 0$.

Отношение эквивалентности определим формулой

$$(p, q) \equiv (p', q'), \text{ если } pq' - p'q = 0. \quad (\text{I, 3; 2})$$

2°) Возьмем в качестве E множество вещественных неотрицательных функций вещественной переменной, т. е. множество отображений \mathbb{R} в $\mathbb{R}_+ = \{x; x \in \mathbb{R}, x \geq 0\}$. Отношение эквивалентности выберем таким, в котором $f \equiv g$ тогда и только тогда, когда f и g «эквивалентны» при $x \rightarrow +\infty$, т. е. если, каково бы ни было $\varepsilon > 0$, существует x_0 , такое, что для всех $x \geq x_0$ имеют место неравенства $|f(x) - g(x)| \leq \varepsilon$.

3°) Определим на множестве целых чисел \mathbb{Z} отношение эквивалентности так, чтобы $p \equiv q$ тогда и только тогда, когда

$p - q$ делится на заранее заданное число m . В теории чисел такое отношение записывается в виде $p \equiv q \pmod{m}$.

4°) Пусть E — множество прямых на плоскости. Определим отношение эквивалентности для прямых D и D' , считая $D \equiv D'$, когда эти прямые параллельны или совпадают.

5°) Пусть E — множество всех векторов \vec{AB} на плоскости. Определим в E отношение эквивалентности $\vec{A'B'} \sim \vec{AB}$, если векторы $\vec{A'B'}$ и \vec{AB} параллельны, одинаково направлены и имеют одинаковую длину.

6°) В том же самом множестве E установим отношение эквивалентности, согласно которому $\vec{A'B'} \sim \vec{AB}$, если $\vec{A'B'}$ и \vec{AB} лежат на одной прямой, одинаково направлены и имеют одинаковую длину.

Классы эквивалентности. Разбиения

Пусть для элементов множества E определено некоторое отношение эквивалентности R . Часть E , образованную из всех эквивалентных между собой элементов, будем называть *классом эквивалентности*.

Теорема. *Два класса эквивалентности всегда либо совпадают, либо не пересекаются.*

Доказательство. Пусть A и B — два класса эквивалентности. Предположим, что они пересекаются, и x — некоторая точка пересечения. Класс A образован из элементов, эквивалентных одному из своих элементов a , но, так как x и a эквивалентны, то из свойства с) следует, что A образован также из элементов, эквивалентных x . Однако тогда по той же причине B обладает тем же свойством, и, значит, A и B совпадают.

Из проведенного доказательства, кроме того, вытекает, что класс эквивалентности является множеством всех элементов, эквивалентных произвольному элементу из этого класса. Таким образом, классы эквивалентности определяют некоторое *разбиение* множества E , т. е. *некоторое множество непустых попарно не пересекающихся частей* E , объединение которых равно E . Если известны эти части, то известно и отношение эквивалентности, ибо x и y эквивалентны тогда и только тогда, когда они принадлежат одному классу эквивалентности. Обратно, всякое разбиение множества E : $E = \bigcup_{i \in I} A_i$, где A_i не пусты и попарно не пересекаются, определяет некоторое отношение эквивалентности, а именно: $x \equiv y$, если существует такой индекс $i \in I$, что $x \in A_i$, $y \in A_i$. В этом случае множества A_i являются классами эквива-

лентности в этом соотношении. В частности, если f — некоторая сюръекция множества E на множество F , то отношение $f(x) = f(y)$ является отношением эквивалентности на E и множества $f^{-1}(\{z\})$, $z \in F$, являются классами эквивалентности.

Фактормножество

Фактормножеством множества E по отношению эквивалентности R называется множество, обозначаемое через E/R , каждый элемент которого является одним из классов эквивалентности. Если через x обозначить класс эквивалентности элемента x , то x является элементом фактормножества и $x \in x$.

Отображение, ставящее в соответствие каждому элементу x из E класс эквивалентности x , называется *канонической сюръекцией* E на E/R . Действительно, отображение сюръективно. Обратно, мы видели выше, что всякая сюръекция f множества E на F определяет некоторое отношение эквивалентности R . Отображение, ставящее в соответствие каждому $z \in F$ класс эквивалентности $f^{-1}(\{z\})$, является биекцией множества F на E/R , что позволяет рассматривать F как некоторую «модель» фактормножества E/R .

Можно дать простую интерпретацию фактормножества на примерах отношений эквивалентности, приведенных ранее на стр. 17—18.

В первом примере фактормножество является множеством \mathbb{Q} рациональных чисел, поскольку рациональное число обычно определяется как семейство пар (p, q) , $q \neq 0$, где при $pq' - p'q = 0$ две пары (p, q) и (p', q') определяют одно и то же рациональное число.

В третьем примере фактормножествами являются множества целых чисел, сравнимых по модулю m .

В четвертом — фактормножество есть множество направлений прямых на плоскости.

В пятом — фактормножеством является множество свободных векторов на плоскости. При этом свободным вектором называется как раз класс всех эквивалентных векторов.

В шестом примере фактормножеством является множество «скользящих векторов» плоскости, применяемых в механике.

Факторгруппа по инвариантной подгруппе

Пусть G — некоторая группа, закон композиции которой записываем мультипликативно: $(x, y) \rightarrow xy$. Заданная композиция является отображением $G \times G$ в G .

Пусть Γ — некоторая инвариантная подгруппа¹⁾. Будем рассматривать следующее бинарное отношение: « $x \equiv y$, если существует элемент $a \in \Gamma$, такой, что $y = ax$, или, что то же самое, если $yx^{-1} \in \Gamma$ ». Можно проверить, что такое отношение есть отношение эквивалентности. Классами эквивалентности являются «правые классы смежности» в G по подгруппе Γ . Поскольку Γ инвариантна, то отношение эквивалентности « $x \equiv y$, если существует элемент $a \in \Gamma$, такой, что $y = xa$, или если $x^{-1}y \in \Gamma$ », идентично предыдущему и правые классы смежности являются также левыми классами смежности. Пусть α, β — два из этих классов. Какими бы ни были $x \in \alpha, y \in \beta$, произведение xy всегда принадлежит одному и тому же классу. Будем обозначать его через $\alpha\beta$. Закон композиции $(\alpha, \beta) \rightarrow \alpha\beta$ превращает фактормножество в новую группу, обозначаемую через G/Γ и называемую факторгруппой группы G по инвариантной подгруппе Γ .

Факторпространство векторного пространства по векторному подпространству

Пусть E — векторное пространство над полем K , и пусть F — некоторое векторное подпространство. Рассмотрим бинарное отношение на E «разность $x - y$ принадлежит F ». Это — некоторое отношение эквивалентности. Впрочем, относительно сложения E является абелевой группой, а F — ее подгруппой. В абелевой группе всякая подгруппа инвариантна. Следовательно, рассматриваемое отношение является частным случаем предыдущего. Пусть E/F — факторгруппа. Эта факторгруппа абелева, и закон композиций в ней мы записываем аддитивно. Пусть $\lambda \in K$ — некоторый скаляр и α — некоторый класс экви-

валентности. Если x — произвольный элемент из α ($x = \alpha$), то λx всегда принадлежит одному и тому же классу эквивалентности, обозначаемому через $\lambda\alpha$. Закон сложения на E/F и закон умножения на скаляры $(\lambda, \alpha) \rightarrow \lambda\alpha$ превращают E/F в векторное пространство, называемое факторпространством векторного пространства E по векторному подпространству F . Согласно определению векторных операций в E/F , для $x \in E, y \in E, \lambda \in K$ имеем $(x + y)^\circ = x^\circ + y^\circ, (\lambda x)^\circ = \lambda^\circ x^\circ$. Другими словами, каноническая сюръекция $x \rightarrow x^\circ$ пространства E на E/F является линейным отображением.

Каноническая сюръекция не биективна (кроме случая, когда $F = \{0\}$). Однако, если G является векторным подпростран-

¹⁾ Подгруппа Γ называется инвариантной, если из $a \in G, x \in \Gamma$ следует, что $a\alpha a^{-1} \in \Gamma$.

ством E , дополнительным¹⁾ к F , то сужение на G канонического отображения является биекцией G на E/F . В самом деле:

а) Это сужение инъективно, ибо если x и y — два элемента G , таких, что $\dot{x} = \dot{y}$, то $x - y \in E$. Так как G — векторное подпространство, то $x - y \in G$. Поскольку F и G дополнительны друг другу, то их пересечение сводится к нулевому элементу, а, значит, $x = y$.

б) Рассматриваемое сужение сюръективно, ибо если α является некоторым классом эквивалентности и если x есть некоторый элемент этого класса, то x записывается в виде $x = x' + x''$, $x' \in F$, $x'' \in G$ (значит, $x - x' \in G$), а, следовательно, $x'' = x - x' = \alpha$, где $x'' \in G$, что доказывает сюръективность рассматриваемого отображения.

Итак, отображение $x \rightarrow \dot{x}$ подпространства G на E/F является линейной биекцией, т. е. биекцией, сохраняющей векторную структуру. Оно позволяет считать G моделью векторного факторпространства E/F . В частности, G имеет ту же размерность, что и E/F ²⁾.

§ 4. ОТНОШЕНИЯ ПОРЯДКА

Бинарное отношение R на множестве E называется *отношением порядка*, если оно обладает следующими свойствами:

- а) *рефлексивность*: $(x, x) \in R$;
- б) *транзитивность*: если $(x, y) \in R$ и $(y, z) \in R$, то $(x, z) \in R$;
(I, 4; 1)
- с) *антисимметричность*: если $(x, y) \in R$ и $(y, x) \in R$, то $x = y$.

Вместо того чтобы писать $(x, y) \in R$, пишут также $x \leqslant_R y$, или $x \leqslant y$, если отношение порядка было указано заранее и нет необходимости его повторять.

В этом случае $y \geqslant x$ означает $x \leqslant y$. Отношения (I, 4; 1) теперь можно записать в виде:

- а) $x \leqslant x$;
- б) если $x \leqslant y$ и $y \leqslant z$, то $x \leqslant z$;
- с) если $x \leqslant y$ и $y \leqslant x$, то $x = y$.

1) Каждый элемент E представляется единственным образом как сумма некоторого элемента из F и некоторого элемента из G .

2) Факторпространство E/F существует всегда, поскольку мы его явно определили. Однако всегда ли существует дополнение G к F в E ? Существование такого дополнения очевидно, если E конечномерно. То же справедливо и во всех других случаях. Этот факт мы примем без доказательства.

Примеры отношений порядка

1°) В множестве \mathbb{N} неотрицательных целых чисел, в множестве \mathbb{Z} всех целых чисел, в множестве \mathbb{Q} рациональных чисел и в множестве \mathbb{R} вещественных чисел отношение $((x, y) \in R)$ тогда и только тогда, когда $x \leqslant y$, является отношением порядка.

Уместно заметить, что отношение $((x, y) \in R)$ тогда и только тогда, когда $x \geqslant y$, является также отношением порядка, называемым противоположным предыдущему.

Символ \leqslant был использован по аналогии с символом \leqslant . По этой же аналогии можно рассматривать отношения:

$$x \leqslant y, \quad x < y, \quad y \geqslant x \quad \text{и} \quad y > x,$$

читаемые соответственно как « x предшествует y », « x строго предшествует y », « y следует за x », « y строго следует за x », а символ $x < y$ означает « $x \leqslant y$ и $x \neq y$ ».

Заметим, что мы порываем здесь с принятыми определениями, говоря « x предшествует y » вместо обычного «предшествует или совпадает» и «строго предшествует» вместо «предшествует». Причина этого изменения, полностью оправдываемая в дальнейшем, заключается в том, что чаще всего используется отношение \leqslant и его следует называть короче. Следует всякий раз, когда это возможно, вместо символа $<$ применять символ \leqslant . Когда пишется строгое неравенство с $<$, то это должно являться предупреждением для читателя в том, что здесь требуется осторожность, что неравенство \leqslant не подходит. Например, непрерывность вещественной функции f вещественной переменной в точке a следует записывать следующим образом:

«каково бы ни было $\varepsilon > 0$, существует $\eta > 0$, такое, что из $|x - a| \leqslant \eta$ следует $|f(x) - f(a)| \leqslant \varepsilon$ ».

Мы употребили знак более широкого неравенства там, где это было возможным, и применили строгое неравенство > 0 только там, где это было абсолютно необходимо.

Использованная аналогия между произвольным отношением порядка \leqslant и его частным случаем \leqslant в множестве вещественных чисел может создать некоторые затруднения, поскольку можно писать $x \leqslant y$ и говорить « x предшествует y » даже в том случае, когда отношение порядка определено неравенством $x \geqslant y$.

2°) В множестве слов существует отношение порядка, называемое алфавитным (если договориться отождествлять омонимы).

3°) В множестве $E = \mathfrak{P}(F)$ частей множества F существует естественное отношение порядка: $X \leqslant Y$, если $X \subset Y$.

4°) В множестве $E = \mathbb{R}^F$ функций, определенных на множестве F , с вещественными значениями также существует естественное отношение порядка: $f \leq g$, если при любом $x \in F$ выполняется неравенство $f(x) \leq g(x)$. Заметим, что здесь отношение $f < g$ означает, что, каково бы ни было x , $f(x) \leq g(x)$ и что, по меньшей мере для одного x , $f(x) < g(x)$. Рассматриваемое отношение вовсе не означает, что при любом x имеет место неравенство $f(x) < g(x)$.

5°) В множестве \mathbb{N}_1 целых чисел ≥ 1 существует отношение порядка, применяемое в арифметике: $a \leq b$, если a является делителем b .

6°) В произвольном множестве E отношение « $x \leq y$, если $x = y$ » является отношением порядка. Говорят, что это — хаотический порядок на E .

Говорят, что отношение порядка является полным или что множество E вполне упорядочено, если для любых двух элементов x, y из E необходимо имеем: или $x < y$, или $x = y$, или $x > y$.

Это справедливо для естественных порядков в множествах $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ в алфавитном порядке слов, но это не так для отношений порядка, рассмотренных в примерах 3°)—5°).

Z В каждом из таких случаев, если x и y не удовлетворяют никакому из трех указанных соотношений, говорят, что они *не сравнимы*. В примере 3°) две непустые не пересекающиеся части F не сравнимы; в 4°) — не сравнимы функции 0 и x ; в 5°) — не сравнимы целые числа 2 и 3; в хаотическом порядке примера 6°) два произвольных не совпадающих элемента не сравнимы.

Мажорируемые части, мажоранты, максимум, точная верхняя грань

Говорят, что некоторая часть упорядоченного множества E , *мажорируема*, если существует по крайней мере один элемент из E , следующий за всеми элементами этой части. Каждый такой элемент называется *мажорантой* этой части.

Аналогичное определение дается для *минорируемой* части и *миноранты*.

Часть, одновременно мажорируемая и минорируемая, называется *ограниченной*.

Говорят, что некоторая часть E имеет *максимум*, если существует мажоранта этой части, *ей принадлежащая*.

Часть не обязана иметь максимум, однако, если она его имеет, то этот максимум *единствен*. В самом деле, если a и b — два максимума одной и той же части, то одновременно имеем: $a \leq b$ и $b \leq a$, откуда $a = b$. Максимум некоторой

части A , если он существует, обозначается через $\max_{x \in A} x$ или $\max(A)$.

Аналогичное определение дается для *минимума*, обозначаемого через \min .

Говорят, что некоторая часть A множества E имеет *точную верхнюю грань*, если множество ее мажорант имеет минимум, и этот минимум называется *точной верхней гранью рассматриваемой части*.

Точная верхняя грань является, следовательно, наименьшей из мажорант. Любой элемент, мажорирующий часть A , мажорирует также и ее точную верхнюю грань.

Часть не обязана всегда иметь точную верхнюю грань, но если она ее имеет, то эта грань единственна. Точная верхняя грань некоторой части A , если она существует, обозначается через $\sup_{x \in A} x$ или $\sup A$.

Если точная верхняя грань принадлежит A , то она является максимумом, и обратно. Аналогичное определение дается для *точной нижней грани*, обозначаемой через \inf .

Теорема 2. *Всякая непустая мажорируемая часть вещественной прямой \mathbb{R} , снабженной отношением естественного порядка, имеет точную верхнюю грань, и всякая непустая минорируемая ее часть¹⁾ имеет точную нижнюю грань. Кроме того, точная верхняя грань b мажорируемой части A характеризуется соотношениями*

- a) при любом $x \in A$, $x \leq b$;
- b) каково бы ни было $b_1 < b$, существует по крайней мере одно число $x \in A$, такое, что $b_1 \leq x \leq b$.

Доказательство этой теоремы общеизвестно.

Естественно, что не мажорируемая часть не имеет точной верхней грани. Прямая \mathbb{R} ни мажорируема, ни минорируема.

В множестве E частей множества F при отношении порядка $X \subset Y$ (пример 3°) каждая часть имеет точные верхнюю и нижнюю грани. В самом деле, любая часть из E является некоторым множеством частей F . Точной верхней гранью будет объединение этих частей, а точной нижней гранью — их пересечение.

В множестве E вещественнозначных функций, определенных на множестве F (пример 4°), всякая мажорируемая часть имеет точную верхнюю грань и всякая минорируемая часть имеет точ-

¹⁾ Условие «непустая» существенно. Часть \emptyset мажорируема, так как любая точка \mathbb{R} является мажорантой для \emptyset . Множество мажорант совпадает с \mathbb{R} , не имеющим минимума. Хотя \emptyset и ограничена, она не имеет ни точной верхней, ни точной нижней граней.

ную нижнюю грань. Действительно, если A — эта часть, то ее точной верхней гранью будет функция f_0 , задаваемая формулой:

$$f_0(x) = \sup_{f \in A} f(x) \quad \text{для всех } x \in F. \quad (\text{I}, 4; 2)$$

Заметим, что имеется два *совершенно разных*, но полезных понятия. С одной стороны, точная верхняя грань части A в упорядоченном множестве \mathbb{R}^F . Это вещественная функция, называемая *верхней огибающей* функций $f \in A$. С другой стороны, точная верхняя грань в \mathbb{R} множества всех значений этих функций, $\sup_{\substack{x \in F \\ f \in A}} f(x)$, которая является, если она существует, некоторым

вещественным *числом* — точной верхней гранью функций $f \in A$. Если A состоит из одного элемента f , то говорят о точной верхней грани f , $\sup_{x \in F} f(x)$, — вещественном числе, являющемся точной верхней гранью множества значений функции f . Если она достигается для некоторого значения x , то ее называют *максимумом*.

Аналогичное замечание можно сделать для *нижней огибающей* и *точной нижней грани*. Эти часто употребляемые понятия в силу своей близости таят в себе некоторую опасность.

В множестве \mathbb{N}_1 целых чисел $\geqslant 1$ с отношением делимости 5°) всякая конечная часть имеет точную верхнюю грань — общее наименьшее кратное и точную нижнюю грань — общий наибольший делитель чисел, входящих в эту часть.

Возрастающие функции

Пусть E и F — два упорядоченных множества. Отображение f множества E в F называется *возрастающим*, если

$$\text{из } x \leqslant y \text{ следует } f(x) \leqslant f(y). \quad (\text{I}, 4; 3)$$

Если, сверх того,

$$\text{из } x < y \text{ следует } f(x) < f(y), \quad (\text{I}, 4; 4)$$

то отображение называется *строго возрастающим*.

Соответствующим образом измененное определение дается для *убывающего* и *строго убывающего отображений*.

Если E вполне упорядочено, то всякая одновременно возрастающая и убывающая функция на E постоянна. В самом деле, пусть x и y — произвольные элементы E . Для них либо $x \leqslant y$, либо $x \geqslant y$. Так как f — одновременно возрастающая и

убывающая функция, то $f(x) \leq f(y)$ и $f(x) \geq f(y)$, а, значит, $f(x) = f(y)$, т. е. f постоянна¹⁾.

Замкнутым интервалом $[a, b]$, $a \leq b$, упорядоченного множества E называется множество элементов x из E , удовлетворяющих неравенствам $a \leq x \leq b$. *Открытым интервалом* $]a, b[$ называется множество элементов x , удовлетворяющих отношениям $a < x < b$. Через $[a, b[$ обозначается множество элементов x , удовлетворяющих неравенствам $a \leq x < b$, а через $]a, b]$ — множество элементов x , для которых $a < x \leq b$. Последние множества называются также *полуоткрытыми интервалами*. Во всех случаях a является началом, а b — концом интервала. Иногда нам придется рассматривать интервалы с началом a и концом b без уточнения того, являются ли они открытыми, полуоткрытыми или замкнутыми. Мы их будем обозначать через $[a, b]$.

Множество таких элементов x , что $a \leq x$, называется правым замкнутым сечением $[a, \rightarrow)$, а множество таких x , что $a < x$, называется правым открытым сечением $]a, \rightarrow)$.

Аналогичное определение дается для левых сечений $(\leftarrow, a]$ и $(\leftarrow, a[$.

В случае вещественной прямой \mathbb{R} с естественным упорядочением такие сечения называются полупрямыми. Эти сечения, так же как и все множество E в целом, принято считать интервалами.

Пополненная прямая

Пополненной прямой $\bar{\mathbb{R}}$ называется множество, образованное, с одной стороны, из элементов вещественной прямой \mathbb{R} и, с другой стороны, — из двух элементов, обозначаемых $-\infty$ и $+\infty$. На $\bar{\mathbb{R}}$ устанавливается отношение порядка, обозначаемое через \leq и определяемое следующим образом: $a \leq b$, если a и b конечны и если они удовлетворяют неравенству \leq в \mathbb{R} , или же если $a = -\infty$ или $b = +\infty$. Множество $\bar{\mathbb{R}}$ вполне упорядочено и имеет, кроме того, минимум $-\infty$ и максимум $+\infty$. Любая часть $\bar{\mathbb{R}}$ ограничена, и теорема 2 справедлива в $\bar{\mathbb{R}}$ без предположения мажорируемости части A , поскольку эта мажорируемость всегда имеет место²⁾. Интервал из $\bar{\mathbb{R}}$ вида $[a, +\infty[$ при $a \in \mathbb{R}$ содержится в $\bar{\mathbb{R}}$, поэтому его часто называют полуупря-

1) Мы замечали, что отношение $x = y$ является таким отношением порядка в E , при котором два различных элемента не сравнимы. Если F — некоторое упорядоченное множество, то любое отображение E в F является одновременно возрастающим и убывающим, не являясь необходимо постоянным. Таким образом, предположение о полной упорядоченности E не излишне.

2) Точная верхняя грань пустой части \emptyset есть минимум $\bar{\mathbb{R}}$ (см. примечание на стр. 24), т. е. $-\infty$, точной нижней границей является $+\infty$. Для непустой части A всегда $\sup(A) \geq \inf(A)$. Это оказывается неверным для пустого множества.

мой $[a, \rightarrow)$ в \mathbb{R} . Точно так же, если A — некоторая не мажорируемая часть \mathbb{R} , то часто пишут $\sup(A) = +\infty$, это приводит к тому, что A лучше рассматривать как некоторую часть $\bar{\mathbb{R}}$.

§ 5. МОЩНОСТИ. СЧЕТНЫЕ МНОЖЕСТВА

Рассмотрим отображение \mathbb{N} в \mathbb{N} , ставящее в соответствие каждому целому числу удвоенное число. Это отображение взаимно однозначно и позволяет таким образом в некотором смысле считать, что существует столько же четных чисел, сколько существует целых. Отсюда следует также, что в случае бесконечных множеств может существовать биективное отображение некоторого множества на его часть, отличную от самого множества.

Может показаться слишком смелым, несмотря на сказанное выше, пытаться сравнивать между собой бесконечные множества. Тем не менее, благодаря понятию биективного отображения такое сравнение возможно.

Теорема 3 (Бернштейн). Пусть E и F — два произвольных множества. Тогда

1°) либо существует инъекция E в F , либо существует инъекция F в E (оба обстоятельства не исключают друг друга);

2°) если существуют одновременно инъекция E в F и инъекция F в E , то существует также биекция E на F .

Мы примем эту теорему без доказательства.

Следствие. Для заданных множеств E и F имеется только три возможности:

а) Существует инъекция E в F и не существует инъекции F в E . В этом случае говорят, что F имеет мощность, строго большую мощности E , или что E имеет мощность, строго меньшую F .

б) Существует инъекция F в E и не существует инъекции E в F . Тогда E имеет мощность, строго большую, чем F , или F по мощности строго меньше, чем E .

с) Существует биекция E на F . В этом случае говорят, что E и F имеют одинаковую мощность, или равномощны¹⁾.

Мощности. Кардинальные числа

Отношение « E равномощно F » является отношением эквивалентности между множествами²⁾. Класс эквивалентности,

¹⁾ Доказательство теоремы Бернштейна сложно, и мы его не приводим. Следствие вытекает непосредственно из этой теоремы. Если E равномощно с F или мощность его строго меньше, чем F , то говорят, что E имеет мощность, меньшую, чем мощность F .

²⁾ Мы намеренно не останавливаемся на логических трудностях, связанных с понятием множества всех множеств.

т. е. класс всех множеств, равномощных данному множеству, называется *мощностью*, или *кардиальным числом*. Конечные кардиальные числа являются классами эквивалентности конечных множеств. Эти числа по определению являются целыми натуральными числами 0, 1, 2 (Заметим, что мы приняли как первичное понятие целые натуральные числа, но их строгое математическое определение полно сложностей. В частности, не легко а priori определить *конечные множества*. Часто по определению считают множество конечным, если оно не равномощно никакой из его частей, отличных от самого множества, а затем доказывают, что кардиальные конечные числа обладают свойствами целых натуральных чисел.) Бесконечное кардиальное число, т. е. мощность бесконечного множества, называется *трансфинитным кардиальным числом*, или *трансфинитным числом*.

В классе кардиальных чисел существует отношение порядка: «если α является кардиальным числом некоторой части множества мощности β , то $\alpha \leqslant \beta$ ». Согласно теореме 3, это отношение обладает антисимметрией, а, значит, действительно является некоторым отношением порядка. Из этой же теоремы следует, что это отношение порядка является полным, т. е. любые два кардиальных числа сравнимы.

Заметим, что если существует сюръекция f множества E на множество F , то мощность F меньше мощности E . В самом деле, прообраз каждой точки F не пуст, и если в каждом из этих прообразов мы выберем по одному элементу, то получим некоторую часть E , равномощную $F^1)$. Например, фактормножество множества E по некоторому отношению эквивалентности всегда имеет меньшую мощность, чем E .

На множествах кардиальных чисел можно определить операции сложения, умножения, возведения в степень точно так же, как и в множестве натуральных чисел.

1°) Пусть α и β — два кардиальных числа, а E и F — множества мощности соответственно α и β . Через $\alpha + \beta$ обозначается мощность «суммы» E и F , т. е. всякого множества, допускающего *разбиение*, образованное из двух множеств, равномощных E и F соответственно.

2°) Через $\alpha\beta$ обозначается мощность произведения $E \times F$. Это — кардиальное число объединения α непересекающихся частей, каждая из которых имеет мощность β .

3°) Через α^β обозначается мощность множества E^F всех отображений F в E .

1) Нетрудно выбрать по одному элементу в каждом из конечного числа множеств. Производить подобный выбор в случае бесконечного числа множеств не легко. После больших споров в начале нашего века было признано, что возможность такого выбора может быть введена как аксиома теории множеств — *аксиома выбора*, или *аксиома Цермело*.

Теорема 4. Предыдущие операции, определенные на множестве кардинальных чисел, обладают следующими свойствами: ассоциативность и коммутативность сложения, ассоциативность и коммутативность умножения, дистрибутивность умножения по отношению к сложению, и, сверх того, имеют место соотношения

$$(\alpha^\beta)(\alpha^\gamma) = \alpha^{\beta+\gamma}, \quad \alpha^{\gamma\beta} = (\alpha\beta)^\gamma, \quad (\alpha^\beta)^\gamma = \alpha^{\beta\gamma}. \quad (I, 5; 1)$$

Доказательство. Не очевидны лишь равенства (I, 5; 1). Пусть E, F, G — множества мощностей α, β, γ соответственно. Тогда мощности множеств E^F и E^G равны α^β и α^γ . Для того чтобы определить отображение множества $F + G$ в E , достаточно определить сужение этого отображения на F и на G , т. е. произвольное отображение F в E и произвольное отображение G в E . Общий элемент E^{F+G} получается, таким образом, как пара, состоящая из произвольного элемента E^F и произвольного элемента E^G . Каждая такая пара является произвольным элементом $E^F \times E^G$, а следовательно, E^{F+G} равномощно $E^F \times E^G$ и $\alpha^{\beta+\gamma} = \alpha^\beta\alpha^\gamma$.

На стр. 13 мы видели, что всякое отображение G в $E \times F$, т. е. любой элемент $(E \times F)^G$, является парой, состоящей из некоторого отображения G в E и некоторого отображения G в F , т. е. элемента E^G и элемента F^G . Такая пара является произвольным элементом $E^G \times F^G$, а, следовательно, $(E \times F)^G$ равномощно множеству $E^G \times F^G$ и $(\alpha\beta)^\gamma = \alpha^\gamma\beta^\gamma$.

Пусть теперь f — некоторое отображение $F \times G$ в E . Отображение $x \rightarrow f(x, y)$ множества F в E при фиксированном y из G называется частным отображением f_y , так что по определению $f_y(x) = f(x, y)$. Таким образом, f из $E^{F \times G}$ определяет некоторое отображение $y \rightarrow f_y$ множества G в множество E^F отображений F в E , т. е. элемент $(E^F)^G$. Обратно, если $y \rightarrow g_y$ является отображением G в E^F , то оно порождается с помощью предыдущего процесса из отображения g множества $F \times G$ в E , определяемого по формуле $g(x, y) \rightarrow g_y(x)$. Тем самым нами установлена биекция $E^{F \times G}$ на $(E^F)^G$, откуда $\alpha^{\beta\gamma} = (\alpha^\beta)^\gamma$.

Эти формулы могут убедить читателя в том, что кардинальные числа, даже трансфинитные, обладают всеми свойствами обычных натуральных целых чисел, но это не так, как показывает приводимое ниже удивительное свойство, которое мы примем без доказательства.

Теорема 5. Если α и β — два кардинальных числа $\neq 0$ и если по крайней мере одно из них трансфинитно, то сумма $\alpha + \beta$ и произведение $\alpha\beta$ равны наибольшему из них¹⁾.

¹⁾ Доказательство весьма замысловато.

Отсюда, в частности, вытекает, что для кардинальных бесконечных чисел определить вычитание невозможно, ибо среди чисел, прибавление которых к α дает α , фигурирует не только 0, но и любое конечное число и даже само α .

Теорема 6. *Каково бы ни было множество E , множество его частей имеет мощность, строго большую мощности E .*

Эта теорема показывает, что последовательность кардинальных бесконечных чисел не ограничена.

В самом деле, предположим, что существует сюръекция f множества E на множество частей $\mathfrak{P}(E)$. Тогда, для $x \in E$, $f(x)$ является элементом $\mathfrak{P}(E)$, т. е. некоторой частью E . Обозначим через A часть E , образованную из таких $x \in E$, что $x \notin f(x)$. Так как $A \subset \mathfrak{P}(E)$, то в E существует по крайней мере один элемент y , такой, что $f(y) = A$. Если $y \in f(y) = A$, то, по определению множества A , $y \notin A$, что невозможно. Если $y \notin f(y) = A$, то $y \in A$. В обоих случаях мы приходим к противоречию.

Поскольку, однако, существует инъекция E в $\mathfrak{P}(E)$, а именно $x \rightarrow \{x\}$, то E имеет мощность, меньшую мощности $\mathfrak{P}(E)$, а значит, строго меньшую мощности $\mathfrak{P}(E)$.

Замечание 1. Если E бесконечно, то множество $\mathfrak{P}_f(E)$ конечных частей E равномощно множеству E . Это легко доказывается с помощью приводимой ниже теоремы 7. Отображение $(x_1, x_2, \dots, x_n) \rightarrow \{x_1, x_2, \dots, x_n\}$, ставящее в соответствие каждому элементу (x_1, x_2, \dots, x_n) из E^n ($n \geq 1$) часть E , образованную из этих элементов (не обязательно различных), является некоторой сюръекцией E^n на множество $\mathfrak{P}_n(E)$ непустых частей E , образованных не более чем из n элементов. Но тогда $\text{card } \mathfrak{P}_n(E) \leq \text{card } E^n = \text{card } E$ (теорема 5), и поскольку $\text{card } \mathfrak{P}_n(E) \geq \text{card } E$, то $\text{card } \mathfrak{P}_n(E) = \text{card } E$. Пусть теперь $f_n: x \rightarrow f_n(x)$ — некоторая биекция E на $\mathfrak{P}_n(E)$. Положим $f_0(x) = \emptyset$ для всех $x \in E$. Тогда $(n, x) \rightarrow f_n(x)$ будет некоторой сюръекцией $\mathbb{N} \times E$ на $\mathfrak{P}_f(E)$, а значит, $\text{card } \mathfrak{P}_f(E) \leq \text{card } (\mathbb{N} \times E) = v \text{ card } E = \text{card } E$ (в силу теоремы 7, неравенства $v \leq \text{card } E$ и теоремы 5). Поскольку обратное неравенство очевидно, окончательно получаем: $\text{card } \mathfrak{P}_f(E) = \text{card } E$.

Замечание 2. Характеристической функцией некоторой части A множества E называется функция φ_A , определенная на E , со значениями в двухэлементном множестве $\{0, 1\}$, такая, что

$$\varphi_A(x) = 1, \text{ если } x \in A, \text{ и } \varphi_A(x) = 0, \text{ если } x \notin A. \quad (I, 5; 2)$$

Задание этой функции однозначно определяет часть A . Впрочем, всякая функция на E , принимающая значения 0 и 1, единственным образом определяет некоторую часть E . Существует, таким образом, некоторая биекция множества $\mathfrak{P}(E)$ частей E

на множество отображений E в множество $\{0, 1\}$. Отсюда следует, что кардинальным числом множества $\mathfrak{P}(E)$ является $\text{card } E$.

Предыдущую теорему теперь можно сформулировать следующим образом:

Каково бы ни было кардинальное число α , $2^\alpha > \alpha^1$.

Мы сейчас изучим две наиболее важные трансфинитные мощности: мощность счетного множества и мощность континуума.

Счетные множества

Мощностью счетного множества v называется мощность множества \mathbb{N} целых натуральных чисел. Всякое множество, равномощное \mathbb{N} , называется счетным²⁾.

Теорема 7. v является наименьшим трансфинитным кардинальным числом.

Это означает лишь, что всякое бесконечное множество E содержит по меньшей мере одну счетную часть. Предположим, что для некоторого бесконечного множества E соотношение $\text{card } E > v$ не имеет места. Согласно теореме 3, это означает, что существует биекция E на некоторую бесконечную часть P множества \mathbb{N} . Отображение $n \rightarrow x_n$, где x_n есть $(n+1)$ -й (в порядке возрастания) элемент P , определяет биекцию \mathbb{N} на P . Окончательно получаем, что $\text{card } E = v$.

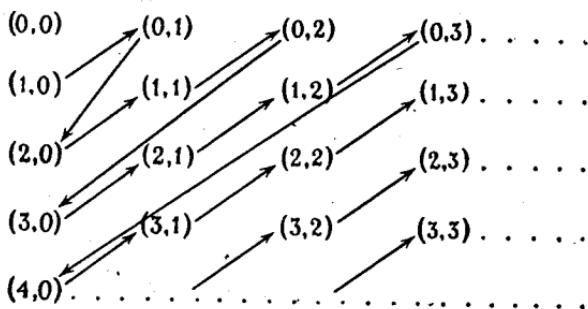
Теорема 8. Для любого конечного числа $m \geqslant 1$ выполняются равенства $mv = v$ и $v^m = v$.

Результат непосредственно вытекает из теоремы 5, согласно которой $mv = v$ при $m \leqslant v$ и $v^2 = vv = v$, а следовательно, по индукции $v^m = v$. Однако мы приняли теорему 5 без доказательства, в то время как равенство $vv = v$, из которого вытекает теорема 8, доказывается элементарно.

¹⁾ В частности, для любого целого $n \geqslant 0$ имеем $2^n > n$.

²⁾ Сказать, что некоторое множество счетно, означает сказать, что оно допускает по крайней мере одну биекцию на \mathbb{N} . Однако это не значит, что такая биекция задана. Это означает также, что возможно расположить элементы данного множества в некоторую последовательность x_0, x_1, x_2, \dots , хотя сама эта последовательность не задана. Часто кардинальное число множества обозначают через \aleph^0 (\aleph — «алеф» — первая буква древнееврейского алфавита). Во многих работах счетными называют множества, кардинальные числа которых $\leqslant v$. Множество E конечно или счетно ($\text{card } E \leqslant v$) тогда и только тогда, когда существует инъекция E в \mathbb{N} , или, если положить $E \neq \emptyset$, тогда и только тогда, когда существует сюръекция \mathbb{N} на E .

Покажем, что $\mathbb{N} \times \mathbb{N}$ равномощно \mathbb{N} . Если элементы $\mathbb{N} \times \mathbb{N}$ перенумеровать в порядке, определяемом последовательными линиями, параллельными диагонали квадратной таблицы:



то получится последовательность $(0, 0), (1, 0), (0, 1), (2, 0), (1, 1), (0, 2), \dots$, определяющая биекцию \mathbb{N} на $\mathbb{N} \times \mathbb{N}$.

Следствия. 1°) Объединение конечного или счетного множества конечных или счетных частей множества E конечно или счетно.

В самом деле, пусть I — некоторая часть \mathbb{N} и $A_i, i \in I$, — некоторые части E . Можно предположить, что ни одна из них не пуста, так как пустые множества ничего не изменяют в объединении. Пусть f_i — некоторая сюръекция \mathbb{N} на A_i . Тогда $(i, n) \rightarrow f_i(n)$ будет являться сюръекцией $I \times \mathbb{N}$ на $\bigcup_{i \in I} A_i$.

Поскольку $I \times \mathbb{N}$ счетно (или пусто, если I пусто), то $\bigcup_{i \in I} A_i$ конечно или счетно.

2°) Множество \mathbb{Z} всех целых чисел счетно (как объединение двух счетных множеств). Множество \mathbb{Q} рациональных чисел счетно.

В самом деле, отображение, ставящее каждой паре (p, q) , $q \neq 0$, множества $\mathbb{Z} \times \mathbb{Z}$ в соответствие рациональное число p/q , является сюръекцией подмножества $\mathbb{Z} \times \mathbb{Z}$ на \mathbb{Q} . Значит, \mathbb{Q} не более чем счетно, но так как оно содержит \mathbb{N} , то \mathbb{Q} счетно.

3°) Множество алгебраических вещественных чисел счетно.

Алгебраическим вещественным числом называется вещественный корень некоторого не равного тождественно нулю полинома с целыми коэффициентами.

Поскольку не равный тождественно нулю полином степени $\leq m$ с целыми коэффициентами имеет $m+1$ коэффициентов, которые являются произвольными целыми числами и не все равны нулю, то мощность множества этих полиномов равна $v^{m+1} = v$.

Каждый такой полином имеет не более m алгебраических вещественных корней, а следовательно, множество алгебраических чисел, корнем полиномов степени $\leq m$, можно рассматривать как объединение счетного множества конечных множеств. Такое множество не более чем счетно, а так как оно бесконечно, то оно счетно. Поскольку m принимает все значения, то множество всех алгебраических чисел получается как объединение счетного множества счетных множеств, а значит, является счетным множеством.

4°) *Множество всех точек \mathbb{R}^n с рациональными или алгебраическими координатами счетно, ибо его кардинальное число равно $v^n = v$.*

Мощность континуума

Теорема 9. Множество всех вещественных чисел несчетно.

Доказательство. Покажем, что даже множество E вещественных чисел, удовлетворяющих неравенствам $0 \leq x < 1$, несчетно. Допустим, что E счетно и существует некоторая биекция \mathbb{N}_1 на E , т. е. элементы E могут быть записаны в виде попарно различных элементов некоторой последовательности x_1, x_2, x_3, \dots

Кроме того, рассмотрим вещественное число ξ , определяемое следующим образом: перед запятой мы поставим 0, затем в качестве j -го десятичного знака выберем произвольное целое число между 1 и 8, отличное от j -го десятичного знака числа x_j . Таким путем мы образуем бесконечную дробь, определяющую некоторое число ξ . Поскольку n -й десятичный знак числа ξ отличен от n -го десятичного знака числа x_n и все десятичные знаки числа ξ отличны от 0 и 9, то $\xi \neq x_n$. (Применение десятичных знаков 0 и 9 может создать затруднения, поскольку число, в десятичном разложении которого, начиная с некоторого места, стоят только нули, допускает иное десятичное разложение, содержащее с некоторого места одни девятки, например $0,102000 \dots = 0,101999 \dots$) Итак, мы пришли к противоречию: множество $\{x_1, x_2, x_3, \dots\}$ не содержит числа, расположенного на отрезке $[0, 1]$.

Трансцендентные числа

Вещественные числа, не являющиеся алгебраическими, называются *трансцендентными*. Поскольку множество алгебраических чисел счетно, а множество вещественных несчетно, то существуют трансцендентные числа и даже «большинство» вещественных чисел трансцендентны. Однако, несмотря на это, не так легко указать явно трансцендентные числа. Можно

доказать (но это вовсе не очевидно!), например, что числа e и π трансцендентны.

Замечание. Часть множества $E = [0, 1]$, образованная из чисел, в десятичном разложении которых имеются лишь цифры 1, 2, 3, 4, 5, 6, 7, 8, является множеством, равномощным множеству $\{1, 2, 3, 4, 5, 6, 7, 8\}^{\mathbb{N}}$, т. е. имеет мощность $8^{\mathbb{N}}$.

Что же касается самого множества E , то его мощность $\leqslant 10^{\mathbb{N}}$ (мы пишем \leqslant , а не $=$ из-за указанного выше двоякого десятичного представления чисел). Поэтому $8^{\mathbb{N}} \leqslant \text{card } E \leqslant 10^{\mathbb{N}}$, откуда $2^{\mathbb{N}} \leqslant \text{card } E \leqslant 16^{\mathbb{N}} = (2^4)^{\mathbb{N}} = 2^{4\mathbb{N}} = 2^{\mathbb{N}}$, а значит, $\text{card } E = 2^{\mathbb{N}}$. Только что доказанная теорема является лишь частным случаем теоремы 6. Впрочем, приведенное доказательство является лишь частным случаем доказательства теоремы 6 (небольшое усложнение возникло в силу двойственного десятичного представления чисел).

Будем обозначать через γ мощность множества $E = [0, 1]$ и называть его *мощностью континуума*. Мощность континуума — это мощность множества \mathbb{R} вещественных чисел (ибо $x \rightarrow -\log x/(1-x)$ является биекцией $[0, 1]$ на \mathbb{R}).

Теорема 10. Имеют место равенства $t\gamma = \gamma\gamma = \gamma\gamma = \gamma^m = \gamma^n = \gamma$, где $m \geqslant 1$ — произвольное целое конечное число.

Доказательство. Все эти кардинальные числа $\leqslant \gamma^{\mathbb{N}}$ и $\geqslant \gamma$, поэтому достаточно доказать, что $\gamma^{\mathbb{N}} = \gamma$. Легко видеть, что $\gamma^{\mathbb{N}} = (2^{\mathbb{N}})^{\mathbb{N}} = 2^{\mathbb{N}^2} = 2^{\mathbb{N}} = \gamma$.

Следствие¹⁾. 1°) *Множество комплексных чисел имеет мощность континуума* (поскольку оно равномощно \mathbb{R}^2).

2°) *Любое векторное пространство конечного числа измерений n над полем вещественных или комплексных чисел имеет мощность континуума.*

Действительно, зафиксировав базис, легко определить биекцию такого пространства на \mathbb{R}^n , имеющего, согласно равенству $\gamma^n = \gamma$, мощность континуума. Отсюда, в частности, вытекает парадоксальное следствие о существовании биекции \mathbb{R} на плоскость \mathbb{R}^2 , поскольку множества \mathbb{R} и \mathbb{R}^2 равномощны. Когда сравнивают семейства точек, кривых, поверхностей, зависящих от некоторого числа параметров, то часто говорят, что одно из этих семейств больше другого, поскольку его элементы зависят от трех вещественных параметров, в то время как элементы другого зависят лишь от двух параметров. Такое рассуждение

¹⁾ Эти следствия, точно так же как и многие другие элементарные свойства счетных множеств и континуума, могут быть доказаны учащимися младших курсов непосредственно, т. е. без применения принятых нами без доказательства больших теорем 3, 4, 5.

не является строгим, поскольку \mathbb{R}^2 и \mathbb{R}^3 равномощны. Во второй главе, посвященной топологии, мы познакомимся с другими типами сравнений, но ни в коем случае не следует сравнивать бесконечности в описанном смысле.

3°) Множество всех последовательностей вещественных чисел и последовательностей комплексных чисел имеет мощность континуума, ибо их кардинальное число равно $\gamma^\nu = \gamma$.

4°) Множество E непрерывных вещественных функций вещественной переменной имеет мощность континуума. В самом деле, каждой такой функции можно поставить в соответствие последовательность вещественных чисел — значений функции в точках с рациональными абсциссами. Можно считать также, что эти точки находятся во взаимно однозначном соответствии с \mathbb{N} , поскольку \mathbb{Q} счетно. Конечно, последовательность этих значений не произвольна: если бы значения функции в рациональных точках выбирались произвольно, то невозможно было бы ее продолжить непрерывно на всю вещественную прямую. Однако последовательность вещественных чисел, соответствующая непрерывной функции, ее полностью определяет. Следовательно, можно установить биекцию множества непрерывных функций на часть множества последовательностей вещественных чисел. Значит, это множество E имеет мощность, не большую мощности континуума. А так как отображение, которое каждой непрерывной функции ставит в соответствие ее значение в одной точке, например, в начале координат, является сюръекцией E на \mathbb{R} , то E имеет мощность, не меньшую мощности континуума. Отсюда следует, что E имеет мощность, в точности равную мощности континуума.

5°) Множество всех вещественных функций вещественной переменной или даже множество всех функций, принимающих только значения 0 и 1, имеет мощность, строго большую мощности континуума, ибо их мощности равны соответственно γ^ν и 2^ν , а $\gamma^\nu = (2^\nu)^\nu = 2^{\nu\nu} = 2^\nu > \gamma$. Из 4°) и 5°) вытекает, что «большинство» функций имеет не менее одной точки разрыва.

Континуум-гипотеза

Согласно этой гипотезе, 2^ν является кардинальным числом, непосредственно следующим за ν . В общем случае *обобщенная континуум-гипотеза* заключается в предположении, что при любом кардинальном числе α кардинальное число 2^α непосредственно следует за α .

Недавно доказано (П. Коэн, 1963), что континуум-гипотеза *неразрешима* — ее невозможно ни доказать, ни опровергнуть, можно лишь принять ее или противоположное ей утверждение как аксиому.

§ 6. НЕКОТОРЫЕ ОСНОВНЫЕ ПОНЯТИЯ ЛОГИКИ

Всякая теорема заключается, вообще говоря, в том, что здается некоторое свойство A , называемое *условием*, из которого выводится свойство B , называемое *заключением*.

Коротко выражение « A влечет B » записывается формулой $A \Rightarrow B$. Обратная теорема, которая не всегда справедлива, записывается тогда в виде $B \Rightarrow A$.

Если данная теорема и её обратная обе справедливы, то свойства A и B эквивалентны, и такую теорему можно записать в виде $A \Leftrightarrow B$, что также выражается в форме: «Для того чтобы $A \dots$, необходимо и достаточно, чтобы $B \dots$ ».

Утверждение, противоположное некоторому свойству A , записывается «не A », или $\neg A$. Если x не принадлежит множеству E , то пишут $x \notin E$. Если X и Y — части E и если X не содержитя в Y , то пишут $X \not\subset Y$ (что вовсе не означает, что X содержит Y).

Во всех случаях «не не A » = A , и справедливо утверждение: «или A , или не A » (принцип исключенного третьего).

Теорема 11. Предложение $A \Rightarrow B$ справедливо тогда и только тогда, когда верно предложение $\neg B \Rightarrow \neg A$.

Доказательство. 1°) Предположим, что $A \Rightarrow B$ верно и в то же время имеет место $\neg B$. Тогда свойство A невозможно без того, чтобы не было справедливым B , что противоречит свойству $\neg B$. Мы имеем, следовательно, свойство $\neg A$, что доказывает справедливость предложения $\neg B \Rightarrow \neg A$.

2°) Если справедливо предложение $\neg B \Rightarrow \neg A$, то, согласно 1°), справедливо $\neg \neg A \Rightarrow \neg \neg B$. Так как $\neg \neg A = A$ и $\neg \neg B = B$, то справедливо предложение $A \Rightarrow B$.

Пример. Эквивалентными являются следующие предложения: «Любая вещественная функция, непрерывная на конечном вещественном интервале $[a, b]$, ограничена» и «Любая вещественная функция, определенная на $[a, b]$ и не ограниченная на нем, разрывна по крайней мере в одной точке».

Когда некоторый объект обладает свойством A или свойством B , то пишут, что он удовлетворяет $A \vee B$ или еще « A или B ». В математике связка *или* никогда не относится к событиям, обязательно исключающим одно другое. « A или B » вовсе не исключает одновременного выполнения A и B . Например, если A означает $x \leq 0$, а B означает $x \geq 0$, то имеет место A или B . Эти два свойства не исключают друг друга, поскольку возможно, что $x=0$.

Если имеют место одновременно свойства A и B , то это записывается в виде $A \wedge B$ или же « A и B ».

В математических теоремах часто используются выражения: «Для всех ...» и «Существует ... такое, что ...». Их обозначают соответственно через \forall и \exists и называют *кванторами*.

Предложения «Для всех ...» и «Существует ...» часто сопровождаются некоторыми ограничениями. Эти ограничения обычно записываются в круглых скобках.

Предположим, что для некоторой вещественной функции f вещественной переменной мы хотим сформулировать свойство быть непрерывной в каждой точке. Непрерывность функции в точке a означает, что, каково бы ни было $\varepsilon > 0$, существует $\eta > 0$, такое, что неравенство $|x - a| \leq \eta$ влечет за собой неравенство $|f(x) - f(a)| \leq \varepsilon$. Иначе говоря, каково бы ни было $\varepsilon > 0$, существует $\eta > 0$, такое, что для любого $x \in \mathbb{R}$, такого, что $|x - a| \leq \eta$, имеет место свойство P : $|f(x) - f(a)| \leq \varepsilon$.

Для того чтобы теперь выразить, что функция *непрерывна в любой точке*, мы должны записать:

(Для всех $a \in \mathbb{R}$) (для всех $\varepsilon > 0$) (существует $\eta > 0$, такое, что) (для всех $x \in \mathbb{R}$, таких, что $|x - a| \leq \eta$) имеем P : $|f(x) - f(a)| \leq \varepsilon$.

Эта фраза коротко записывается в виде:

$$(\forall a \in \mathbb{R}) (\forall \varepsilon > 0) (\exists \eta > 0) (\forall x \in \mathbb{R}, |x - a| \leq \eta) : |f(x) - f(a)| \leq \varepsilon. \quad (I, 6; 1)$$

Теорема 12. Отрицание свойства, содержащего некоторое число кванторов \forall, \exists и свойство P , получается заменой каждого квантора \forall на \exists и \exists на \forall и свойства P на его отрицание $\neg P$.

Так, например, свойство вещественной функции вещественной переменной не быть всюду непрерывной, т. е. иметь разрыв хотя бы в одной точке, может быть записано в одну строку

$$(\exists a \in \mathbb{R}) (\exists \varepsilon > 0) (\forall \eta > 0) (\exists x \in \mathbb{R}, |x - a| \leq \eta) : |f(x) - f(a)| > \varepsilon. \quad (I, 6; 2)$$

Этот метод отрицания свойства применяется так же автоматически, как и правило знаков в умножении или при раскрытии скобок.

Доказательство. Предположим, что имеется один квантор, например \forall . Наше свойство имеет тогда вид:

$$(\forall x, \text{ удовлетворяющий } S) : P.$$

Его отрицание очевидно: существует некоторое x , удовлетворяющее S , но не обладающее свойством P , т. е.

$$(\exists x, \text{ удовлетворяющий } S) : \neg P.$$

Теорема для данного случая доказана. Аналогичное доказательство имеет место для случая, когда имеется только квантор \exists . Для завершения доказательства теперь достаточно применить индукцию по произвольному числу кванторов.

Предположим, что теорема доказана в том случае, когда имеется $n - 1$ квантор, и докажем ее для случая, когда их имеется n . Пусть рассматриваемое свойство записано, например, в виде

$$(\forall x, \text{ удовлетворяющее } S) : Q,$$

где Q — свойство, содержащее $n - 1$ кванторов. Тогда, в силу доказанного выше, его отрицанием будет свойство:

$$(\exists x, \text{ удовлетворяющий } S) : \neg Q.$$

Однако $\neg Q$ можно получить, используя теорему, поскольку Q содержит только $n - 1$ квантор. Итак, теорема для этого случая верна. Тот же самый результат получится, если первым квантором является \exists , а следовательно, теорема имеет место в общем случае.

Замечание Каждый раз, когда квантор \exists предшествует некоторому числу других кванторов, следующая за ним буква может оказаться функцией всех букв, фигурирующих в предыдущих кванторах.

Например, в свойстве функции быть всюду непрерывной η зависит от a и от ε . Может случиться, что η можно выбирать в зависимости только от ε , а не от a . В этом случае говорят, что функция равномерно непрерывна. Свойство функции быть равномерно непрерывной более сильное, чем свойство быть непрерывной. Его можно записать следующим образом:

$$(\forall \varepsilon > 0) (\exists \eta > 0) (\forall a \in \mathbb{R}) (\forall x \in \mathbb{R}, |x - a| \leq \eta) : |f(x) - f(a)| \leq \varepsilon. \quad (I, 6; 3)$$

Очевидно, перестановка кванторов существенно изменяет заданное свойство. Непосредственно видно, что самым сильным является такое свойство, в котором символ \exists поставлен как можно раньше. Производя, например, еще одну перестановку, получим:

$$(\exists \eta > 0) (\forall a \in \mathbb{R}) (\forall \varepsilon > 0) (\forall x \in \mathbb{R}, |x - a| \leq \eta) : |f(x) - f(a)| \leq \varepsilon. \quad (I, 6; 4)$$

Фиксируем η , о котором говорится в самом начале, что оно существует. Для любого x из интервала $[a - \eta, a + \eta]$ имеет место неравенство $|f(x) - f(a)| \leq \varepsilon$, $\forall \varepsilon > 0$, что означает $|f(x) - f(a)| = 0$. Но тогда $f(x) = f(a)$ для любого a и любого x из интервала $[a - \eta, a + \eta]$. Выбирая в качестве a все числа

вида $p/10^b$, где целое p изменяется от $-\infty$ до $+\infty$ и выбрано так, что $p/10^b \leq \eta, \dots$, получаем: предложение (I, 6; 4) означает, что f постоянна. Получаем еще большее усиление свойства непрерывности!

Н. В. Мы рассматривали теорию множеств и логику с «наивной» точки зрения. Слова: множество, равный, каков бы ни был, существует, влечет и т. д. . . употреблялись в том смысле, какой они имеют в разговорном языке. Совершенно ясно, что при этом не проводилось никакого серьезного математического обоснования смысла этих понятий. Символы $\forall, \exists, \in, =, \Rightarrow$ должны были быть только знаками, подчиняющимися некоторым «правилам игры», как конь и ладья в шахматах являются фигурами с определенными правилами передвижений, не совпадающими с теми, какими обладают настоящие конь или ладья. Точно так же евклидова плоскость — это не поверхность воды, а сфера — это не апельсин. Математическая логика (содержащая теорию множеств) является ветвью современной математики. Заметим, что еще не доказана ее непротиворечивость. Логическая теория называется *противоречивой*, если существует такое предложение P , для которого можно одновременно доказать как P , так и $\neg P$. В этом случае будет справедливым любое предложение Q (а также и $\neg Q$). В самом деле, соотношения $P \Rightarrow \Rightarrow (P \text{ или } Q)$ и $((P \text{ или } Q) \text{ и } \neg P) \Rightarrow Q$ дают $(P \text{ и } \neg P) \Rightarrow Q$. Однако если современная теория множеств противоречива, то это еще не означает, что мы находимся в безнадежном положении. Надо лишь уменьшить число аксиом теории, что, без сомнения, не изменит заметно математики в целом!

ГЛАВА II

Топология

§ 1. МЕТРИЧЕСКИЕ ПРОСТРАНСТВА. ЭЛЕМЕНТАРНЫЕ ПРИМЕРЫ

Множество E называется *метрическим пространством*, если для его элементов определено понятие *расстояния*, т. е. определено некоторое отображение множества $E \times E$ в полуправую $R_+ = \{x, x \in \mathbb{R}, x \geq 0\}$, которое каждой паре (x, y) из $E \times E$ ставит в соответствие некоторое число $d(x, y) \geq 0$, называемое *расстоянием* между x и y .

Расстояние должно обладать следующими тремя свойствами:

- | | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| 1) симметрия: $d(x, y) = d(y, x);$
2) положительность $d(x, y) > 0$, если $x \neq y$,
и $d(x, x) = 0;$
3) неравенство треугольника:
$d(x, z) \leq d(x, y) + d(y, z)$ (каждая сторона треугольника не превосходит суммы двух других сторон). | } (II, 1; 1) |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|

Из неравенства треугольника как следствие вытекает, что каждая сторона треугольника не меньше разности двух других сторон:

$$d(x, z) \geq |d(x, y) - d(y, z)|. \quad (\text{II, 1; 2})$$

Из этого же неравенства вытекает также, что для произвольных n точек множества имеет место неравенство

$$d(x_1, x_n) \leq d(x_1, x_2) + d(x_2, x_3) + \dots + d(x_{n-1}, x_n). \quad (\text{II, 1; 3})$$

Приведем сразу же ряд важных примеров метрических пространств:

1) Вещественная прямая \mathbb{R} и комплексная плоскость \mathbb{C} , с расстоянием $d(x, y) = |x - y|$. Эту метрику называют естественной метрикой в \mathbb{R} и \mathbb{C} . Всюду, где специально не оговорено противное, в пространствах \mathbb{R} и \mathbb{C} метрика считается естественной.

2) Вещественная прямая \mathbb{R} , для точек которой расстояние определено формулой $d(x, y) = |F(x) - F(y)|$, где F — произвольная вещественная строго монотонная функция вещественной переменной.

3) Евклидово вещественное пространство \mathbb{R}^n или унитарное комплексное пространство \mathbb{C}^n n измерений, расстояние

между точками (x_1, x_2, \dots, x_n) и (y_1, y_2, \dots, y_n) в котором определено формулою $\left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{1/2}$.

Несколько позже мы строго докажем, что здесь действительно определено метрическое пространство (стр. 43). Введенную метрику называют естественной метрикой \mathbb{R}^n и \mathbb{C}^n . Если не оговорено противное, то метрика в пространствах \mathbb{R}^n и \mathbb{C}^n всегда будет считаться естественной.

4°) В произвольном множестве можно ввести дискретную метрику, положив $d(x, y) = 1$, если $x \neq y$, и $d(x, x) = 0$.

Сфера. Шары

Сферой с центром a конечного радиуса $R > 0$ называется множество точек x из E , таких, что $d(a, x) = R$. Из определения вовсе не следует, что такая сфера не пуста или что две сферы с различными центрами не совпадают (пример: в дискретной метрике все сферы радиуса 2 совпадают и пусты).

Множество точек x из E , таких, что $d(a, x) < R$ (соответственно $\leq R$), обозначается через $B_0(a, R)$ (соответственно $B(a, R)$) и называется *открытым* (соответственно *замкнутым*) шаром с центром в a и конечным радиусом $R > 0$ ¹⁾.

Когда говорят о шаре без каких-либо пояснений, то имеют в виду замкнутый шар. В дискретной метрике замкнутый шар радиуса < 1 сводится к его центру. Шар (замкнутый) радиуса ≥ 1 совпадает со всем пространством. (Для $R = 0$ открытый шар пуст, сфера же и замкнутый шар сводятся к их центрам. Мы будем всегда предполагать, даже если это специально не оговорено, что радиусы сфер или шаров конечны и > 0 .)

Часть метрического пространства называется *ограниченной*, если существует шар (как всегда, конечного радиуса), в котором она содержится. Множество \mathbb{R} и часть \mathbb{N} множества \mathbb{R} не ограничены. Однако если в E введена дискретная метрика, то E ограничено.

Нормированные векторные пространства

Пусть E — векторное пространство над полем \mathbb{K} вещественных или комплексных чисел. *Нормой* в векторном пространстве E называется любая функция, обозначаемая через $\vec{x} \rightarrow \|\vec{x}\|^2$,

¹⁾ Мы делаем строгое различие между понятиями сферы и шара. В случае плоскости \mathbb{R}^2 говорят *окружность* вместо сферы и *круг* вместо шара.

²⁾ Договоримся всегда ставить стрелку над элементами векторного пространства. При этом мы будем строго различать число 0 и нуль $\vec{0}$ векторного пространства.

обладающая следующими свойствами:

- 1) положительность: $\|\vec{x}\| > 0$ для $\vec{x} \neq \vec{0}$ и $\|\vec{0}\| = 0$;
- 2) преобразование при гомотетии: $\|\lambda\vec{x}\| = |\lambda| \|\vec{x}\|, \lambda \in K$;
- 3) условие выпуклости: $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.

Из 2) и 3) легко выводится обобщенное условие выпуклости

$$\|\vec{x}_1 + \vec{x}_2 + \dots + \vec{x}_n\| \leq \|\vec{x}_1\| + \|\vec{x}_2\| + \dots + \|\vec{x}_n\|, \quad (\text{II}, 1; 5)$$

а также неравенство

$$\|\vec{x} - \vec{y}\| \geq \|\vec{x}\| - \|\vec{y}\|. \quad (\text{II}, 1; 6)$$

Если в векторном пространстве E введена норма, обладающая указанными выше свойствами, то его называют *нормированным векторным пространством*.

Символ $\|\cdot\|$, естественно, можно употреблять лишь тогда, когда речь идет о некоторой раз и навсегда определенной норме. Если в одной и той же задаче для одного и того же векторного пространства вводится несколько различных норм, то их следует обозначать различными символами.

Пусть E — нормированное векторное пространство. В этом пространстве можно определить расстояние между элементами, удовлетворяющее всем аксиомам расстояния, по формуле: $d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$. Таким образом, любое нормированное векторное пространство является автоматически метрическим пространством. Расстояние в нем обладает дополнительными свойствами, совместимыми с его векторной структурой: расстояние инвариантно относительно сдвига, иначе говоря, $d(\vec{x} - \vec{a}, \vec{y} - \vec{a}) = d(\vec{x}, \vec{y})$, а гомотетия с коэффициентом гомотетии λ приводит к умножению расстояния на $|\lambda|$, т. е. $d(\lambda\vec{x}, \lambda\vec{y}) = |\lambda| d(\vec{x}, \vec{y})$. Обратно, легко видеть, что любое расстояние в векторном пространстве, обладающее двумя этими свойствами, определяет в нем некоторую норму, а именно: $\|\vec{x}\| = d(\vec{0}, \vec{x})$.

Под *открытым шаром радиуса R* в векторном пространстве, если не указан его центр, мы будем понимать шар радиуса R с центром в нуле векторного пространства. Аналогично определяется *замкнутый шар*. Они обозначаются соответственно через $B_0(R)$ и $B(R)$.

В частности, *единичный открытый* (соответственно замкнутый) шар есть открытый (замкнутый) шар с центром в $\vec{0}$ радиуса 1. *Единичный шар* без каких-либо уточнений означает просто единичный замкнутый шар.

Отрезком с концами a и b в векторном пространстве над полем вещественных или комплексных чисел называется множество всех точек вида $\vec{ta} + (1-t)\vec{b}$, $0 \leq t \leq 1$. Его обозначают через $[a, b]$ точно так же, как и замкнутый интервал в упорядоченном множестве. Говорят, что некоторая часть множества E выпукла, если вместе с любыми двумя различными точками этого множества она содержит и весь отрезок, соединяющий эти точки. Из условия выпуклости следует, что всякий шар нормированного векторного пространства является выпуклым множеством. В самом деле, если $\|\vec{x}\| \leq R$ и $\|\vec{y}\| \leq R$, то

$$\begin{aligned} \|\vec{tx} + (1-t)\vec{y}\| &\leq \|\vec{tx}\| + \|(1-t)\vec{y}\| = \\ &= |t| \|\vec{x}\| + |1-t| \|\vec{y}\| \leq [t + (1-t)]R = R. \end{aligned} \quad (\text{II}, 1; 7)$$

Приведем примеры нормированных векторных пространств:

1°) Функция $x \rightarrow |x|$, определенная над полем \mathbb{R} или \mathbb{C} , является нормой. Она определяет упоминавшуюся ранее (1°, стр. 40) естественную метрику.

2°) Функции

$$(x_1, x_2, \dots, x_n) \rightarrow \max_{i=1, \dots, n} |x_i|, \quad \text{или} \quad \sum_{i=1}^n |x_i|, \quad \text{или} \quad \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$$

в векторных пространствах \mathbb{R}^n или \mathbb{C}^n определяют некоторые нормы. Для первых двух норм это очевидно. Проверим выполнение аксиом нормы в третьем случае. Пусть (x_1, x_2, \dots, x_n) , (y_1, y_2, \dots, y_n) — две системы n комплексных чисел. Для доказательства условия выпуклости

$$\left(\sum |x_i + y_i|^2 \right)^{1/2} \leq \left(\sum |x_i|^2 \right)^{1/2} + \left(\sum |y_i|^2 \right)^{1/2} \quad (\text{II}, 1; 8)$$

заметим, что левая часть этого неравенства не превосходит $\left(\sum (|x_i| + |y_i|)^2 \right)^{1/2}$, а значит, достаточно доказать, что

$$\sum (|x_i| + |y_i|)^2 \leq \sum |x_i|^2 + \sum |y_i|^2 + 2 \left(\sum |x_i|^2 \right)^{1/2} \left(\sum |y_i|^2 \right)^{1/2} \quad (\text{II}, 1; 9)$$

или

$$\sum |x_i||y_i| \leq \left(\sum |x_i|^2 \right)^{1/2} \left(\sum |y_i|^2 \right)^{1/2}. \quad (\text{II}, 1; 10)$$

Это — известное неравенство Коши — Буняковского — Шварца¹⁾.

Норма $x = (x_1, \dots, x_n) \rightarrow \sqrt{\sum_{i=1}^n |x_i|^2}$ называется естественной нормой в \mathbb{R}^n или \mathbb{C}^n .

Позже мы познакомимся с другими важными примерами векторных нормированных бесконечномерных пространств.

¹⁾ См., например, книгу И. М. Гельфанд «Лекции по линейной алгебре», «Наука», 1966, стр. 35. — Прим. ред.

**§ 2. ОТКРЫТЫЕ И ЗАМКНУТЫЕ ЧАСТИ. ОКРЕСТНОСТИ.
ВНУТРЕННОСТЬ. ГРАНИЦА. ЗАМЫКАНИЕ.
ПЛОТНЫЕ ПОДМНОЖЕСТВА**

Открытые части

Пусть E — метрическое пространство. Часть A пространства E называется *открытой*, если вместе с каждой своей точкой она содержит некоторый открытый шар (радиуса > 0) с центром в этой точке¹⁾.

Открытые части пространства E обладают, очевидно, следующими свойствами:

- a) само пространство E и его пустая часть \emptyset открыты²⁾;
 - b) пересечение конечного числа открытых множеств открыто;
 - c) объединение конечного или бесконечного числа открытых множеств открыто.
- } (II,2;1)

Докажем, например, свойство b).

Пусть $\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_n$ — открытые подмножества E и x — произвольный элемент их пересечения. Для каждого i существует такое число $R_i > 0$, что шар $B(x, R_i)$ полностью содержится в \mathcal{O}_i . Если теперь положить $R = \min_{i=1, 2, \dots, n} (R_i)$, то шар $B(x, R)$ будет содержаться в рассматриваемом пересечении, а, значит, пересечение открыто.

Имеет место интересное четвертое свойство, называемое *аксиомой отделимости Хаусдорфа*.

d) Каковы бы ни были различные точки a и b пространства E , существуют две открытые непересекающиеся части E , содержащие соответственно a и b .

В самом деле, если взять открытые шары с центрами в точках a и b радиуса $d/3$, где d — расстояние $d(a, b)$, то эти шары

¹⁾ Короче: A является открытой частью в E , если $(\forall x \in A)(\exists \rho > 0)(\forall y \in E, d(x, y) < \rho) : y \in A$, или еще короче: $(\forall x \in A)(\exists \rho > 0) : B_0(x, \rho) \subset A$. Естественно, ρ зависит от x . В определении можно «открытый шар» заменить на «замкнутый шар», поскольку замкнутый шар радиуса ρ содержит открытый шар радиуса ρ , а открытый шар радиуса ρ содержит замкнутый шар радиуса $\rho/2$.

²⁾ Множество A открыто, если для любого $x \in A$ существует шар с центром в x , содержащийся в A . Если $A = \emptyset$, то найти точку $x \in A$ невозможно, а потому указанное свойство выполняется и \emptyset открыто. Приведем другой пример того же самого логического рассуждения. Будем говорить, что человек обладает свойством (P), если его рост больше роста его детей. Тогда люди, не имеющие детей, будут, очевидно, обладать свойством (P). Можно, впрочем, исходить из отрицания рассматриваемого свойства. Некоторая часть A не является открытой, если найдется хотя бы один элемент $x \in A$, для которого не существует шара с центром в x , содержащегося в A . Множество \emptyset не может не быть открытым, так как элементов $x \in \emptyset$ не существует. Пустое множество, таким образом, открыто.

не будут иметь общей точки, так как в противном случае для такой точки c из неравенства треугольника $d(a, b) \leq d(a, c) + d(c, b)$ получилось бы неравенство $d(a, b) \leq 2d/3$, которое неверно.

Примеры. *Открытые шары* являются открытыми, и это оправдывает их название. В частности, при введении естественной метрики на прямой \mathbb{R} открытые интервалы являются открытыми множествами. Легко проверяется от противного, что замкнутые или полуоткрытые интервалы открытыми множествами не являются.

Для того чтобы некоторая часть E была открытой, необходимо, согласно определению, чтобы она была объединением открытых шаров. Однако это условие и достаточно, поскольку каждый открытый шар открыт, а объединение открытых частей открыто (свойство с)). Приведенное условие является, следовательно, необходимым и достаточным для того, чтобы некоторое множество было открытым.

В пространстве \mathbb{R}^n , снабженном метрикой, которая определяется одной из указанных на стр. 43 норм, *открытый параллелепипед*, т. е. множество

$$\{x = (x_1, x_2, \dots, x_n); a_i < x_i < b_i, i = 1, 2, \dots, n\},$$

является открытой частью.

В любом метрическом пространстве множество точек x , удовлетворяющих неравенству $d(a, x) > R$, открыто.

В дискретной метрике (пример 4°), стр. 41) все части E открыты.

Замкнутые части

Замкнутой называется любая часть пространства E , дополнение к которой в E открыто.

Переходом к дополнениям в свойствах а), б), с) открытых частей непосредственно получаются эквивалентные свойства замкнутых частей:

- | | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| а') E и \emptyset замкнуты;
б') объединение конечного числа замкнутых частей замкнуто;
с') пересечение конечного или бесконечного числа замкнутых частей замкнуто ¹⁾ . | } (II, 2; 2) |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|

Переход от открытых множеств к замкнутым в аксиоме Хаусдорфа интересного результата не дает.

¹⁾ Следствие из теоремы 15 дает наилучший практический критерий для определения замкнутости множества.

Примеры. Всякая часть, сводящаяся к одной или конечному числу точек, замкнута. Замкнутый шар является замкнутой частью. В частности, каждый замкнутый интервал вещественной прямой \mathbb{R} с естественной метрикой замкнут, в то время как полуинтервал или открытый интервал не замкнуты. Множество $\{x : d(a, x) \geq R\}$ замкнуто. Любая сфера замкнута.

В пространстве \mathbb{R}^n с одной из введенных метрик, указанных на стр. 43, замкнутый параллелепипед, определенный неравенствами $\{x = (x_1, x_2, \dots, x_n), a_i \leq x_i \leq b_i, i = 1, 2, \dots, n\}$, является замкнутой частью.

В дискретной метрике все части E замкнуты.

Замечания. 1°) Рассматривая объединение и пересечение семейств открытых или замкнутых частей, очень важно тщательно различать случаи, когда рассматриваемое семейство конечно или бесконечно.

Например, на вещественной прямой \mathbb{R} с естественной метрикой часть, сводящаяся к точке, замкнута, но не открыта, и в то же время она является пересечением счетного множества открытых интервалов. Точно так же открытый интервал не замкнут, и в то же время он является объединением счетного множества вложенных в него замкнутых интервалов.

2°) Существуют части E , не являющиеся ни открытыми, ни замкнутыми; например, полуоткрытые интервалы в \mathbb{R} .

3°) Как показывает пример дискретной метрики, кроме \emptyset и E , могут существовать и другие множества, являющиеся одновременно и открытыми, и замкнутыми¹⁾.

Окрестности

Окрестностью точки a множества E называется любая часть E , содержащая по крайней мере одно открытое множество, содержащее точку a (или же содержащая открытый или замкнутый шар с центром в точке a).

Кроме свойства включать в себя точку a , окрестности точки a обладают следующими свойствами:

а'') всякая часть, содержащая окрестность a , является окрестностью a ;

б'') пересечение конечного числа окрестностей точки a является окрестностью a ;

с'') аксиома отделимости Хаусдорфа: каковы бы ни были различные точки a и b множества E , существуют не пересекающиеся окрестности точек a и b .

(II, 2; 3)

¹⁾ Сравните с изложенным в § 9.

Следует заметить, что слово «окрестность» может ввести в заблуждение: создается впечатление, что окрестность точки a является множеством, весьма близким к a , и что чем меньше множество, тем вероятнее, что оно является окрестностью точки a . В действительности свойство а) окрестностей показывает, что, напротив, чем часть A множества E больше, тем скорее она может быть окрестностью точки a . Все пространство E является окрестностью точки a , и, кроме исключительных случаев, таких, как случай дискретной метрики, часть, сходящаяся к самой точке a , не является окрестностью точки a .

Говорят, что некоторая точка E изолирована, если часть E , сходящаяся к этой точке, является ее окрестностью, т. е. является открытой частью. В дискретной метрике каждая точка изолирована. В множестве \mathbb{R} изолированных точек нет.

Пусть A — часть E . Окрестностью части A называется всякая часть E , которая содержит открытую часть, содержащую A . Если задано некоторое $\varepsilon > 0$, то окрестностью A порядка ε называется объединение всех шаров радиуса ε с центрами в точках из A . Это множество является окрестностью, так как оно содержит объединение всех открытых шаров радиуса ε с центрами в точках из A , которое открыто и содержит A .

Теорема 1. Для того чтобы некоторая часть A пространства E была открытой, необходимо и достаточно, чтобы она являлась окрестностью каждой своей точки.

Доказательство. Пусть A является окрестностью каждой своей точки. Тогда вместе с каждой своей точкой множество A должно содержать некоторую открытую часть E , содержащую эту точку, т. е. A должно содержать некоторый шар с центром в рассматриваемой точке. Это означает, что A является открытой частью. Обратно, если часть A открыта и $a \in A$, то A всегда содержит некоторое открытое множество, содержащее a , а именно: саму часть A . Это означает, что A является окрестностью точки a .

Пусть задано семейство $(\mathcal{U}_i)_{i \in I}$ окрестностей точки a в E . Говорят, что это семейство является фундаментальной системой окрестностей точки a , если любая окрестность точки a содержит одну из окрестностей \mathcal{U}_i . Например, в метрическом пространстве E открытые шары (или замкнутые шары) с центром в a или шары с центром в a и рациональными радиусами образуют фундаментальную систему окрестностей точки a .

Внутренность

Пусть A — некоторая часть метрического пространства. Внутренностью A называется объединение всех открытых частей

из E , содержащихся в A . Внутренность множества A обозначается через $\text{int } A$. Множество $\text{int } A$ есть открытое множество, содержащееся в A (свойство с), стр. 44) являющееся, очевидно, наибольшим открытым множеством, содержащимся в A . Конечно, $\text{int } A$ может быть пусто (см. пример на стр. 50 при $E = \mathbb{R}$, $A = \mathbb{Q}$).

Теорема 2. *Внутренность $\text{int } A$ части A является множеством точек из E , каждая из которых является центром по крайней мере одного шара, содержащегося в A , или, иначе, это множество точек E , у которых имеются окрестности, целиком лежащие в A , или, иначе, это множество точек E , для которых A является окрестностью.*

Доказательство. В самом деле, пусть a является центром некоторого открытого шара, содержащегося в A . Этот шар является открытым множеством, содержащимся в A , а, значит, он содержитя в $\text{int } A$ и тем более $a \in \text{int } A$.

Обратно, если $a \in \text{int } A$, то в силу того, что $\text{int } A$ открыто, найдется некоторый шар с центром в a , содержащийся в $\text{int } A$ и тем более в A .

Внешность

Внешностью A называется внутренность его дополнения. Это — наибольшее открытое подмножество E , не пересекающееся с A . Внешность A является множеством всех точек E , каждая из которых есть центр по крайней мере одного шара радиуса >0 , не пересекающегося с A . Внутренность и внешность A , очевидно, не пересекаются.

Граница

Множество точек E , не принадлежащих ни внутренности, ни внешности A , называется *границей A* и обозначается через ∂A .

Поскольку дополнение к границе, являясь объединением двух открытых частей, открыто, граница A замкнута.

Из свойств двух этих частей непосредственно вытекает

Теорема 3. *Для того чтобы точка a принадлежала границе ∂A множества A , необходимо и достаточно, чтобы любая окрестность a содержала бы одновременно как точки A , так и точки его дополнения.*

Доказательство. Прежде всего, если любая окрестность \mathcal{U} точки a пересекается одновременно с частями A и CA , то она не содержитя ни в одной из них. Но тогда a не лежит ни во внутренности, ни во внешности A и, следовательно, при-

надлежит границе A . Обратно, пусть a принадлежит границе \mathcal{V} — некоторая окрестность a . Окрестность \mathcal{V} пересекается с A , так как в противном случае она лежала бы в CA и точка a принадлежала бы внешности. Точно так же видно, что окрестность \mathcal{V} пересекается с CA , ибо в противном случае она лежала бы в A , а значит, a принадлежала бы внутренности A .

Внутренность, внешность и граница части A попарно не пересекаются, а их объединение совпадает с E .

Замыкание

Замыканием \bar{A} части A называется пересечение всех замкнутых частей E , содержащих A . Оно замкнуто (свойство c'), стр. 45). Это — наименьшая замкнутая часть, содержащая A . Точка, принадлежащая замыканию, называется *точкой прикосновения части A* .

Теорема 4. Замыкание множества A является объединением его внутренности и границы, иначе говоря, это дополнение к внешности A . Можно сказать также, что замыкание есть множество всех точек E , каждая окрестность которых пересекается с A ¹⁾.

Доказательство. \bar{A} является наименьшей замкнутой частью, содержащей A . Это утверждение эквивалентно тому, что CA является наибольшей открытой частью, содержащейся в CA , т. е. CA является внешностью A , а это как раз то, что требовалось доказать.

Пусть a — некоторая точка прикосновения части A . Тогда любая окрестность \mathcal{V} точки a пересекается с A . В самом деле, в противном случае при $\mathcal{V} \subset CA$ нашлось бы открытое множество O , содержащее a и содержащееся в \mathcal{V} , которое входит в CA . Но тогда O , а вместе с ним и a принадлежали бы объединению всех открытых частей, содержащихся в CA , которое является внешностью A , что противоречит предположению $a \in \bar{A}$.

Обратно, если точка a такова, что каждая из ее окрестностей пересекается с A , то a не может лежать во внешности A , ибо в противном случае эта внешность была бы некоторой окрестностью a , не пересекающейся с A . Отсюда следует, что a принадлежит замыканию \bar{A} .

Мы доказали тем самым, что a является точкой прикосновения части A тогда и только тогда, когда любая ее окрестность пересекается с A .

Приведем несколько примеров. Пусть A — некоторый открытый или замкнутый шар с центром a и радиусом $R > 0$.

¹⁾ В теореме 15 будет указана другая существенная характеристика замыкания.

в нормированном векторном пространстве. Его внутренностью является соответствующий открытый шар, его замыканием — соответствующий замкнутый шар, его внешностью — множество $\{x; d(a, x) > R\}$, а его границей — сфера с центром в a радиуса R . (Заметим, что, вообще говоря, это не всегда так. Если в E введена дискретная метрика и A — замкнутый шар радиуса 1, то он совпадает с E и, следовательно, со своей внутренностью, в то время как его граница и внешность пусты. Открытый шар радиуса 1 в этом случае вырождается в центр.)

Пусть A — множество рациональных чисел вещественной прямой \mathbb{R} . Тогда $\bar{A} = \emptyset$, $\dot{A} = \bar{A} = \mathbb{R}$.

Для того чтобы некоторая часть A пространства E была открытой, необходимо и достаточно, чтобы она совпадала со своей внутренностью.

Для того чтобы она была замкнутой, необходимо и достаточно, чтобы она совпадала со своим замыканием.

Плотные подмножества

Часть A метрического пространства E называется *плотной* в E , если каждая точка E является для нее точкой прикосновения, т. е. если ее замыкание совпадает с E . Это означает также, что любая открытая часть пространства E пересекается с A . Метрическое пространство E называется *сепарабельным*, если оно конечно или если оно содержит в себе некоторую счетную всюду плотную часть.

Примеры. В пространстве \mathbb{R} с естественной метрикой множество A рациональных чисел и множество B иррациональных чисел плотны. Поскольку множество рациональных чисел счетно, то \mathbb{R} сепарабельно.

Подпространства. Индуцированная метрика

Пусть F — некоторая часть метрического пространства E . Сужение функции расстояния, определенной в $E \times E$, на $F \times F$ превращает множество F в новое метрическое пространство. Его называют *метрическим подпространством* E , а определенную в нем метрику — «*индукцированной*» метрикой.

Пусть A — подмножество F . В этом случае следует четко различать свойство A быть открытым или замкнутым в F от аналогичного свойства в E . Так, например, множество F само по себе одновременно открыто и замкнуто, но если рассматривать F как подмножество E , то это утверждение, вообще говоря, не верно.

Теорема 5. Для того чтобы некоторая часть A множества F была открытой (соответственно замкнутой) в метрическом пространстве F , необходимо и достаточно, чтобы она была пересечением F и некоторой открытой (замкнутой) части метрического пространства E .

Для того чтобы некоторая часть A в метрическом пространстве F была окрестностью точки $a \in F$, необходимо и достаточно, чтобы она была пересечением F и некоторой окрестности a в E .

Доказательство. Обозначим через B шары пространства E и через β — шары пространства F .

Пусть A — часть F , являющаяся пересечением F и некоторой открытой части A_1 из E . Если $a \in F$ лежит в A , а, значит, и в A_1 , то существует некоторый шар $B_0(a, R)$, принадлежащий A_1 . Но тогда $\beta_0(a, R) = B_0(a, R) \cap F$ содержитя в A . Это означает, что A открыто в F .

Обратно, пусть часть A открыта в F . Тогда A является (стр. 45) объединением некоторого семейства $\beta(a_i, R_i)$, $i \in I$, открытых шаров, причем $a_i \in A$ и $R_i > 0$. Объединение шаров $B_0(a_i, R_i)$ является открытым множеством A_1 из E и $A = A_1 \cap F$. Таким образом, свойство, относящееся к открытым множествам, доказано.

Пусть теперь A_1 — некоторая замкнутая часть E . Пусть C_1 — ее дополнение в E . Части A_1 и C_1 разбивают F на два взаимно дополнительных подмножества A и C множества F . Так как C открыто, то $A = A_1 \cap F$ замкнуто.

Обратно, пусть A — некоторая замкнутая часть F . Пусть C — ее дополнение в F . Так как C открыто, то в E существует такая открытая часть C_1 , что $C_1 \cap F = C$. Пусть A_1 — дополнение к C_1 в E . Множество A_1 замкнуто. Поскольку A_1 и C_1 взаимно дополнительны в E , то их пересечения с F будут взаимно дополнительными в F . Так как $C_1 \cap F = C$, то $A = A_1 \cap F$. Из замкнутости A_1 теперь следует замкнутость A . Теорема доказана и в этом случае.

Пусть, наконец, \mathcal{V}_1 — некоторая окрестность точки $a \in F$ в E . Она содержит некоторую открытую часть A_1 , содержащую a . Но тогда $\mathcal{V} = \mathcal{V}_1 \cap F$ содержит открытую часть $A = A_1 \cap F$ пространства F , содержащую a . Эта часть является окрестностью a в F . Обратно, пусть \mathcal{V} — некоторая окрестность a в F . Она содержит некоторую открытую часть A из F , содержащую a . В этом случае существует открытая часть A_1 из E , такая, что $A = A_1 \cap F$. Поскольку множество $\mathcal{V}_1 = A_1 \cup \mathcal{V}$ содержит открытую часть A_1 , содержащую a , оно является некоторой окрестностью a в E , причем $\mathcal{V} = \mathcal{V}_1 \cap F$.

Из доказанного вытекают два следствия:

1°) Открытые, замкнутые части и, следовательно, окрестности метрического пространства F полностью определяются соответствующими множествами в E ; для их определения не нужно знать функцию расстояния. В важности этого замечания мы убедимся позже (см. § 4).

2°) Какова бы ни была часть F из E , если некоторая часть A из F открыта (соответственно замкнута) в метрическом пространстве E , то она открыта (замкнута) в метрическом подпространстве F ; если A является окрестностью точки $a \in F$ в E , то она заведомо является окрестностью точки a в F .

Обратное утверждение, вообще говоря, не верно, как показывает рассмотренный ранее случай $A = F$. Однако имеет место

Теорема 6. а) Если F — открытая часть E , то любая часть A из F , открытая в метрическом пространстве F , также открыта в метрическом пространстве E .

б) Если F — замкнутая часть E , то любая часть A из F , замкнутая в метрическом пространстве F , также замкнута в метрическом пространстве E .

с) Если F является окрестностью точки a в E , то любая часть A из F , являющаяся окрестностью точки a в F , есть окрестность точки a в E .

Доказательство. а) Пусть часть F открыта в E . Если A — открытая часть F , то, согласно теореме 5, существует часть A_1 , открытая в E и такая, что $A = A_1 \cap F$. Поскольку A_1 и F открыты в E , то часть $A = A_1 \cap F$ также открыта в E . Отсюда сразу же вытекает утверждение с).

б) Доказательство аналогично. Достаточно всюду заменить открытые части на замкнутые.

§ 3. НЕПРЕРЫВНЫЕ ФУНКЦИИ. ГОМЕОМОРФИЗМЫ

Пусть f — отображение метрического пространства E в метрическое пространство F . Говорят, что отображение f непрерывно в точке $a \in E$, если, каково бы ни было $\varepsilon > 0$, существует такое $\eta > 0$, что неравенство $d(a, x) \leq \eta$ влечет за собой неравенство $d(f(a), f(x)) \leq \varepsilon$.

Можно также сказать иначе: каким бы ни был шар с центром в $f(a)$, существует шар с центром в a , образ которого при отображении f содержится в предыдущем шаре.

Можно также сказать, что, какова бы ни была окрестность \mathcal{U} точки $f(a)$, существует такая окрестность \mathcal{U} точки a , что $f(\mathcal{U}) \subset \mathcal{U}$.

Можно также сказать, что прообраз при отображении f любой окрестности точки $f(a)$ является окрестностью точки a .

Отображение E в F называется непрерывным, если оно непрерывно в каждой точке a пространства E .

Первое определение непрерывности существенно использует метрику. Напротив, два последних определения используют только открытые множества и окрестности, а не саму метрику. Важность этого замечания будет ясна в дальнейшем (см. § 4).

Пример. Функция $1/x$, т. е. отображение $x \rightarrow 1/x$ метрического пространства E , дополнительного к началу координат вещественной прямой, в метрическое пространство $F = \mathbb{R}$ (вещественная прямая) всюду непрерывно.

Теорема 7. Для того чтобы отображение f метрического пространства E в метрическое пространство F было непрерывным, необходимо и достаточно, чтобы прообраз при отображении f любой открытой части F был открытой частью в E .

Доказательство. Докажем сначала, что это условие необходимо. Предположим, что отображение f непрерывно. Пусть B — открытая часть F и $A = f^{-1}(B)$. Пусть $a \in A$. Так как отображение f непрерывно в точке a и B является окрестностью $f(a)$, то часть A будет окрестностью точки a . Итак, A является окрестностью каждой своей точки, а это значит, что A открыта (теорема 1).

Докажем теперь его достаточность. Предположим, что прообраз при отображении f любой открытой части F открыт в E . Пусть a — произвольная точка E и \mathcal{U} — окрестность точки $f(a)$ в F . Тогда \mathcal{U} содержит некоторую открытую часть B , содержащую точку $b = f(a)$. Прообраз $f^{-1}(\mathcal{U})$ содержит открытую часть $f^{-1}(B)$, содержащую a . Значит, $f^{-1}(\mathcal{U})$ является окрестностью точки a , что доказывает непрерывность отображения f в точке a .

Теорема 8. Для того чтобы отображение f метрического пространства E в метрическое пространство F было непрерывным, необходимо и достаточно, чтобы прообраз при отображении f каждой замкнутой части F был замкнут в E .

Доказательство повторяет доказательство предыдущей теоремы. Надо лишь везде заменить открытые части из E и F на их замкнутые дополнения и применить формулу (I, 2; 3).

Замечание. Если в двух предыдущих теоремах заменить прообразы на прямые образы, то можно прийти к неверным результатам. Рассмотрим, например, постоянное отображение E в F . Такое отображение, очевидно, непрерывно. Однако, образ любой открытой части E , в том числе и самого множества E , сводится к единственной точке F , а подмножества, вырождающиеся в точку, вообще говоря, не являются открытыми. Если рассмотреть функцию $1/x$ из разобранного выше примера, то образ всего множества E , т. е. некоторого замкнутого множества при этом отображении, является в множестве $F = \mathbb{R}$

дополнением к началу координат, не являющимся замкнутым в \mathbb{R} .

Теорема 9. *Норма в нормированном векторном пространстве E , рассматриваемая как отображение E в прямую \mathbb{R} с естественной метрикой, является непрерывной функцией.*

Доказательство. Из (II, 1; б) вытекает, что для заданного $\varepsilon > 0$ достаточно выбрать $\eta = \varepsilon$, чтобы из неравенства $\|x - a\| \leq \eta$ вытекало неравенство $|\|x\| - \|a\|| \leq \varepsilon$.

Теорема 10. *Композиция двух непрерывных отображений непрерывна.*

Доказательство. Пусть E, F, G — три метрических пространства и $h = g \circ f$ — композиция отображения f пространства E в F и отображения g пространства F в G . Будем считать, кроме того, что функция f непрерывна в точке $a \in E$, а функция g непрерывна в точке $b = f(a)$ пространства F .

Пусть $c = g(b) = h(a)$ и \mathcal{W} — окрестность точки c в G . Так как функция g непрерывна в точке b , то прообраз $\mathcal{V} = g^{-1}(\mathcal{W})$ является окрестностью b в F . Поскольку f непрерывна в a , то прообраз $\mathcal{U} = f^{-1}(\mathcal{V})$ является окрестностью a в E . Из того, что $\mathcal{U} = f^{-1}(g^{-1}(\mathcal{W}))$ совпадает с прообразом $h^{-1}(\mathcal{W})$, следует непрерывность h в точке a^1 . Отсюда, очевидно, вытекает, что если f и g всюду непрерывны, то h тоже всюду непрерывна. Этот факт можно было бы получить непосредственно, если применить теорему 7 или теорему 8.

Гомеоморфизмы

Гомеоморфизмом метрического пространства E на метрическое пространство F называется любая биекция E на F , непрерывная вместе со своей обратной биекцией.

Теорема 11. *Для того чтобы биективное и непрерывное отображение f пространства E в F было гомеоморфизмом, необходимо и достаточно, чтобы образ при отображении f каждой открытой части E был открыт в F ; необходимо и достаточно также, чтобы образ при отображении f каждой замкнутой части E был замкнутым в F .*

В самом деле, эти образы есть не что иное, как прообразы относительно обратной биекции $g = f^{-1}$, и предыдущие условия совпадают с условиями непрерывности f^{-1} , указанными в теореме 7 и теореме 8.

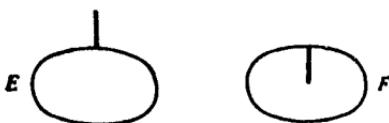
¹⁾ В качестве упражнения проведите другое доказательство, пользуясь первым (метрическим) определением непрерывности с помощью $\forall \varepsilon, \exists \eta, \dots$.

З а м е ч а н и е. Не надо думать, что любая непрерывная биекция является заведомо гомеоморфизмом. Так, например, если E является прямой \mathbb{R} с дискретной метрикой, а F — прямой \mathbb{R} с естественной метрикой, то тождественное отображение E в F непрерывно и биективно, но не является гомеоморфизмом.

Говорят, что два метрических пространства E и F *гомеоморфны*, если существует по крайней мере один гомеоморфизм E на F . В этом случае оба пространства обладают одинаковыми топологическими свойствами, т. е. свойствами, относящимися к открытым, замкнутым множествам и окрестностям.

П р и м е р. Внутренность круга и внутренность треугольника в евклидовой плоскости являются гомеоморфными метрическими пространствами. Полуплоскость $y > 0$, область $y > x^2$, расположенная над параболой $y = x^2$, область $y < x^2$, лежащая под этой параболой на плоскости \mathbb{R}^2 , гомеоморфны¹⁾.

Два метрических пространства, определенных кривыми, изображенными на следующем ниже рисунке, гомеоморфны:



Р и с. 1.

З Однако внимание! Это вовсе не означает, что существует гомеоморфизм первой плоскости на вторую, переводящий первое подпространство на второе!

Вещественная прямая с естественной метрикой и вещественная прямая с дискретной метрикой не гомеоморфны потому, что в последнем случае все подмножества открыты, а в первом случае это не так²⁾.

§ 4. МЕТРИЧЕСКИЕ ПРОСТРАНСТВА И ТОПОЛОГИЧЕСКИЕ ПРОСТРАНСТВА

Используя лишь понятие расстояния, мы смогли определить в метрическом пространстве E понятия открытого и замкнутого множеств, окрестности, замыкания, внутренней и внешней части множества, его границы, плотного множества, непрерывного

¹⁾ В качестве упражнения укажите в каждом случае гомеоморфизм соответствующих гомеоморфных метрических пространств.

²⁾ Мы показали выше, что в данном случае тождественное преобразование не является гомеоморфизмом между двумя рассматриваемыми пространствами. Здесь же мы показали, что между ними не существует никакого гомеоморфизма, — это более сильное утверждение.

отображения. Все эти понятия могут быть определены, исходя из понятия открытого множества.

Может случиться, что две различные метрики в одном и том же пространстве E определяют одну и ту же систему открытых множеств. Тогда они определяют одинаковые системы замкнутых подмножеств, окрестностей каждой точки и т. д.

Пусть, например, в метрическом пространстве E расстояние определяется функцией d и функцией $2d$, т. е. такой функцией, при которой расстояние между двумя элементами x и y считается равным $2d(x, y)$. Эти функции определяют одни и те же открытые множества.

Две метрики в пространстве E называются эквивалентными, если они определяют одну и ту же систему открытых множеств. Говорят также, что они определяют *одну и ту же топологию* в E . Можно также сказать и иначе: *тождественное отображение множества E , снабженного 1-й метрикой, на множество E , снабженное 2-й метрикой, является гомеоморфизмом*.

В векторных пространствах две нормы называются эквивалентными, если соответствующие им метрики эквивалентны.

Теорема 12. *Две нормы $\|\cdot\|_1$ и $\|\cdot\|_2$ на некотором векторном пространстве эквивалентны тогда и только тогда, когда существуют такие постоянные k' , $k'' > 0$, что для всех \vec{x} из E имеют место следующие неравенства:*

$$\|\vec{x}\|_2 \leq k' \|\vec{x}\|_1 \quad \text{и} \quad \|\vec{x}\|_1 \leq k'' \|\vec{x}\|_2. \quad (\text{II}, 4; 1)$$

Доказательство. Обозначим через $B_1(R)$ (соответственно через $B_2(R)$) замкнутый шар с центром в нуле пространства и радиусом R в смысле первой нормы (соответственно второй нормы). Предположим, что обе нормы эквивалентны, т. е. определяют одни и те же открытые множества.

Шар $B_1(1)$ содержит некоторое открытое в смысле 1-й метрики подмножество E , содержащее нуль пространства. Но тогда $B_1(1)$ содержит также некоторое множество, открытое в смысле 2-й метрики, содержащее этот нуль. Существует, следовательно, число $1/k''$, такое, что $B_1(1) \supset B_2(1/k'')$. С помощью гомотетии с множителем $k''R$ отсюда получается включение $B_1(k''R) \supset B_2(R)$, означающее, что соотношение $\|\vec{x}\|_2 \leq R$ влечет за собой неравенство $\|\vec{x}\|_1 \leq k''R$. Так как первое неравенство остается верным при $R = \|\vec{x}\|_2$, то второе неравенство дает при этом $\|\vec{x}\|_1 \leq k'' \|\vec{x}\|_2$.

Меняя ролями нормы $\|\cdot\|_1$ и $\|\cdot\|_2$, мы можем доказать необходимость условия теоремы. Покажем теперь, что оно и до-

статочно. Если это условие выполнено, то из неравенства $\|\vec{x}\|_2 \leqslant R/k''$ будет вытекать неравенство $\|\vec{x}\|_1 \leqslant R$, а значит, $B_2(R/k'') \subset B_1(R)$. Отсюда следует, что любой шар в первой метрике заведомо содержит некоторый шар во второй метрике, и обратно. С помощью параллельного переноса легко убедиться, что все только что доказанное для шаров с центром в нуле остается верным для произвольных шаров¹⁾.

Поскольку в топологии каждое открытое множество является таким множеством, которое вместе с любой точкой содержит некоторый шар с центром в этой точке, то из свойств, установленных нами для шаров, вытекает, что открытые множества в обеих метриках совпадают.

Замечание. Пусть d_1 и d_2 — расстояния, определяемые двумя эквивалентными нормами. Тогда найдется такая постоянная k , что при любых x и y имеют место неравенства

$$d_1(x, y) \leqslant k d_2(x, y) \quad \text{и} \quad d_2(x, y) \leqslant k d_1(x, y).$$

Это обстоятельство является характеристическим для эквивалентных метрик, определенных эквивалентными нормами на векторных пространствах.

Пусть теперь d — расстояние, определенное некоторой метрикой на множестве E . Легко проверяется, что $d' = \inf(d, 1)$, определяемое через $d'(x, y) = \min(d(x, y), 1)$, также является некоторым расстоянием (проверьте свойства (II, 1; 1)). Метрика, определяемая с помощью d' , очевидно, эквивалентна исходной метрике (шары радиусов $\leqslant 1$ одни и те же). Далее, если d изменяется от 0 до $+\infty$, то d' изменяется от 0 до 1. Это означает, что неравенства типа (II, 4; 2) для расстояний d и d' не имеют места. Если E — векторное пространство и d — расстояние, определенное в нем некоторой нормой, то расстояние d' таким уже не будет. Этот пример показывает, что *при любой заданной метрике в E всегда можно найти эквивалентную метрику, относительно которой E будет ограниченным*.

Следствие. Все три нормы, определенные в пространстве \mathbb{R}^n в начале этой главы, эквивалентны.

¹⁾ Доказательство необходимости. Если бы не существовало такого числа k'' , то для любого целого числа $n \geqslant 0$ нашлась бы точка $\vec{x}_n \neq \vec{0}$, такая, что $\|\vec{x}_n\|_1 \geqslant n \|\vec{x}_n\|_2$. Используя при необходимости гомотетию, можно всегда предполагать, что $\|\vec{x}_n\|_1 = 1$. При этом $\|\vec{x}_n\|_2 \leqslant 1/n$. Таким образом, последовательность \vec{x}_n сходится к нулю по 2-й норме и не сходится к нулю по 1-й норме, а это означает, что нормы не эквивалентны. (Это доказательство использует сходящиеся последовательности, которые будут рассмотрены позже.)

Это утверждение вытекает из неравенств:

$$\frac{1}{n} \sum_{i=1}^n |x_i| \leq \max_{i=1, 2, \dots, n} |x_i| \leq \sum_{i=1}^n |x_i|,$$

$$\frac{1}{\sqrt{n}} \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \leq \max_{i=1, 2, \dots, n} |x_i| \leq \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}. \quad (\text{II}, 4; 2)$$

Имеет место следующее более общее утверждение, которое мы примем без доказательства¹⁾:

Теорема 13. В конечномерном векторном пространстве над вещественным или комплексным полем любые две нормы эквивалентны. Существует единственная система открытых множеств, замкнутых множеств и т. д. для любых введенных в нем норм.

Здесь уместно заметить, что, как мы увидим далее на примерах, это свойство не переносится на бесконечномерные векторные пространства.

Итак, мы пришли к тому, что для метрических пространств определяются два рода свойств: *метрические свойства*, зависящие явно от самой метрики, такие, как расстояние между двумя точками, свойство сторон треугольника или фигуры, образованной из нескольких точек, сферы и т. д., и *топологические свойства*, не зависящие явно от самой метрики, а зависящие только от множества открытых, замкнутых²⁾ и т. д. частей. Становится ясным, что топологию можно ввести, не используя в качестве посредника метрику.

Топологическим пространством E называется множество E , в котором выделено семейство частей, называемых *открытыми* в этой топологии.

Это — совершенно произвольное семейство частей, которые должны обладать лишь свойствами а), б), с), д), указанными на стр. 44³⁾.

¹⁾ Доказательство можно найти на стр. 72 (см. теорему 23).

²⁾ Ограниченность части A из E (см. определение на стр. 41) является метрическим, а не топологическим свойством: множество \mathbb{N}_1 целых чисел ≥ 1 не ограничено в метрике $d(p, q) = |p - q|$, но оно станет ограниченным, если метрику в нем определить по формуле $d(p, q) = \left| \frac{1}{p} - \frac{1}{q} \right|$, в то время как эти две метрики эквивалентны. Однако, в *нормированных векторных пространствах* две эквивалентные нормы приводят, согласно теореме 12, к одним и тем же ограниченным частям (см. замечание на стр. 57).

³⁾ Вообще говоря, на открытые множества в топологии налагаются лишь требования аксиом а), б), с). Если, кроме того, они подчиняются требованиям аксиомы д), то топологическое пространство называется *отделимым*. Неотделимые пространства в анализе используются очень редко, и нам они не по-

Мы видим, что метрическое пространство является частным случаем топологического пространства. Существуют топологические пространства, которые нельзя определить, исходя из какой-либо метрики.

Говорят, что топологическое пространство *метризуемо*, если существует метрика, порождающая его топологию.

Подавляющее большинство топологических пространств, с которыми мы встретимся в будущем, метризуемы.

Заметим, что почти все рассмотренные нами определения и теоремы относительно метрических пространств служили для введения в них топологии, а не самой метрики. Легко проверить, что они имеют место для произвольных топологических пространств. Однако теоремы 2, 9, 12 и 13 справедливы только для метрических пространств, ибо они существенно используют метрику даже в своих формулировках.

Каждая точка всегда имеет фундаментальную систему открытых окрестностей (поскольку в любой окрестности точки a содержится открытое множество, содержащее a , которое является открытой окрестностью этой точки). Однако такой системы замкнутых окрестностей может и не быть. Топологическое пространство называется *регулярным*, если каждая его точка обладает фундаментальной системой замкнутых окрестностей. Метризуемое пространство регулярно. С другой стороны, пересечение замкнутых окрестностей точки a сводится к точке a . Следовательно, пересечение всех окрестностей точки a тем более сводится к a . В самом деле, если $b \neq a$, то, согласно аксиоме Хаусдорфа, существуют открытые непересекающиеся множества \mathcal{A} и \mathcal{B} , содержащие точки a и b . Из того, что \mathcal{A} является замкнутой окрестностью точки a и $\mathcal{A} \subset \mathcal{C}\mathcal{B}$, следует, что $\mathcal{A} \subset \mathcal{C}\mathcal{B}$, а, значит, \mathcal{A} не содержит b . Точка b , таким образом, принадлежит не всем замкнутым окрестностям точки a .

Теорему 5 можно использовать для определения индуцированной топологии. Пусть E — некоторое топологическое пространство, а F — часть E . В множестве F можно ввести топологию, принимая за открытые множества в F пересечение с F открытых множеств в смысле топологии E . В этом случае говорят, что F является *топологическим подпространством* E или же что его *топология индуцирована топологией* E .

В дальнейшем всякий раз, когда это будет возможно, мы будем рассматривать теоремы в топологических пространствах.

требуются. Поэтому мы всегда будем предполагать, что *топологическое пространство отдельно, а его открытые множества удовлетворяют аксиоме d*).

В неотделимом пространстве часть, вырождающаяся в точку, не является необходимо замкнутой, и сходящаяся последовательность может иметь несколько различных пределов (см. теорему 14)!

Однако мы не постесняемся проводить доказательства только для случая метрических пространств, если это может привести к упрощению рассуждений.

Топология дополненной прямой $\bar{\mathbb{R}}$ (см. стр. 26 гл. I)

В $\bar{\mathbb{R}}$ можно ввести топологию, определяя открытые множества следующим образом. Часть U множества $\bar{\mathbb{R}}$ считается открытой при выполнении следующих условий:

- если U содержит точку x из \mathbb{R} , то она содержит также по крайней мере один открытый интервал, содержащий x ;
- если U содержит точку $-\infty$, то она содержит также по крайней мере один интервал вида $[-\infty, A]$;
- если U содержит точку $+\infty$, то она содержит также по крайней мере один интервал вида $[A, +\infty]$.

Можно доказать, что определенные таким образом открытые множества удовлетворяют всем аксиомам, необходимым для того, чтобы $\bar{\mathbb{R}}$ было топологическим пространством. Легко видеть, что в этой топологии $\bar{\mathbb{R}}$ метризуемо, а несколько позже мы приведем бесконечное множество эквивалентных метрик, порождающих эту топологию (теорема 38, стр. 99), но ни одна из них естественным образом не выделяется среди других. \mathbb{R} является подпространством $\bar{\mathbb{R}}$, и в этом подмножестве топология, порожденная топологией $\bar{\mathbb{R}}$, совпадает с топологией, определяемой естественной метрикой.

§ 5. ПОСЛЕДОВАТЕЛЬНОСТИ. ПРЕДЕЛЫ. СХОДИМОСТИ

Существенным свойством топологических пространств является возможность рассматривать в этих пространствах сходящиеся последовательности.

1°) Пусть $x_0, x_1, x_2, \dots, x_n, \dots$ — некоторая последовательность точек метрического пространства E . Говорят, что эта последовательность *сходится к точке l из E* , или что она имеет *пределом точку l* ¹⁾, если последовательность вещественных чисел $d(l, x_0), d(l, x_1), \dots, d(l, x_n), \dots$ сходится к нулю.

Это означает следующее: каково бы ни было $\varepsilon > 0$, существует такое целое число n_0 , что из неравенства $n \geq n_0$ следует неравенство $d(l, x_n) \leq \varepsilon$. Можно также сказать так: какова бы ни была окрестность \mathcal{U} точки l , существует такое целое число n_0 , что для $n \geq n_0$ все x_n принадлежат \mathcal{U} . Можно также ска-

¹⁾ Подразумевается: «при n , стремящемся к $+\infty$ ».

Если x_n — вещественные числа, то можно считать, что $l \in [-\infty, +\infty]$. Последовательность при этом рассматривается в $\bar{\mathbb{R}}$.

зать иначе: какова бы ни была окрестность \mathcal{U} точки l , все x_n принадлежат \mathcal{U} , за исключением конечного числа значений n^1).

Последние два определения пригодны и в том случае, когда E — топологическое пространство. Однако если даже E — метрическое пространство, то сходимость последовательности в нем является свойством топологическим, а не метрическим.

Существуют и другие определения пределов, не относящиеся к последовательностям.

2°) Рассмотрим двойную последовательность $x_{m, n}$, где m и n — целые неотрицательные числа, т. е. отображение $\mathbb{N} \times \mathbb{N}$ в E .

Говорят, что эта двойная последовательность сходится к элементу l из E при m и n , стремящихся к бесконечности, если, каково бы ни было $\varepsilon > 0$, существуют такие целые числа m_0, n_0 , что из неравенства $m \geq m_0, n \geq n_0$ вытекает неравенство $d(x_{m, n}, l) \leq \varepsilon^2$.

3°) Говорят, однако, что $x_{m, n}$ сходится к l , когда m или n стремятся к бесконечности, если, каково бы ни было $\varepsilon > 0$, существуют целые числа m_0 и n_0 , такие, что из неравенства $m \geq m_0$ или $n \geq n_0$ вытекает неравенство $d(x_{m, n}, l) \leq \varepsilon^3$.

4°) Если f является отображением вещественной прямой \mathbb{R} в E , то выражение « $f(x)$ стремится к l , когда x стремится к a справа», означает, что, каково бы ни было $\varepsilon > 0$, существует такое $\eta > 0$, что из $(|x - a| \leq \eta, x > a)$ следует $d(f(x), l) \leq \varepsilon$.

5°) В тех же самых условиях выражение « $f(x)$ стремится к l при x , стремящемся к $+\infty$ », означает, что, каково бы ни было $\varepsilon > 0$, существует такое вещественное число A , что из $x \geq A$ вытекает $d(f(x), l) \leq \varepsilon$.

Все эти пределы легко определяются, если E является топологическим и не обязательно метрическим пространством. Они входят в рамки следующего более общего определения.

Пусть X — топологическое пространство, A — некоторая часть X , a — точка X , принадлежащая замыканию \bar{A} , и f — отображение A в E . Говорят, что $f(x)$ сходится к l при x , стремящемся к a по значениям из A ⁴), если, какова бы ни была

1) Это определение показывает, кроме того, что сходимость последовательности не зависит от порядка ее членов. Изменить порядок членов последовательности $x_0, x_1, x_2, \dots, x_n, \dots$ — это значит заменить ее последовательностью $x_{p_0}, x_{p_1}, \dots, x_{p_n}, \dots$, где соотношение $n \rightarrow p_n$ является биекцией \mathbb{N} на себя. Если исходная последовательность сходится к l , то это будет верно и для преобразованной последовательности.

2) Короче: $(\forall \varepsilon > 0)(\exists m_0 \in \mathbb{N})(\exists n_0 \in \mathbb{N})(\forall m \geq m_0, \forall n \geq n_0): d(x_{m, n}, l) \leq \varepsilon$.

3) Короче: $(\forall \varepsilon > 0)(\exists m_0 \in \mathbb{N})(\exists n_0 \in \mathbb{N})(\forall (m, n) \in \mathbb{N} \times \mathbb{N}, m \geq m_0 \text{ или } n \geq n_0): d(x_{m, n}, l) \leq \varepsilon$.

4) Очень часто $a \in \bar{A}$, но $a \notin A$. Именно это имеют в виду, подчеркивая, что «когда x стремится к a , то предполагается, что $x \neq a$ ». Здесь мы не хотим ограничивать себя этим условием. Можно считать, что $a \in A$ или $a \notin A$.

окрестность \mathcal{U} точки l в E , существует окрестность \mathcal{V} точки a в X , такая, что $f(\mathcal{U} \cap A) \subset \mathcal{V}$. Так, например, в случае последовательностей (случай 1°): $X = \bar{\mathbb{R}}$, $A = \mathbb{N}$, $a = +\infty$. В случае 4°): $X = \mathbb{R}$, $A = \{x \in \mathbb{R}, x > a\}$. В случае 5°): $X = \bar{\mathbb{R}}$, $A = \mathbb{R}$, $a = +\infty$.

Сказать, что некоторое отображение f топологического пространства E в топологическое пространство F непрерывно в точке a из E , — то же самое, что сказать: $f(x)$ сходится к $f(a)$, когда x стремится к a .

Теорема 14. Если последовательность имеет предел, то этот предел обязательно единственен.

Доказательство. Предположим, что некоторая последовательность x_0, x_1, x_2, \dots элементов из E имеет два различных предела a и b из E . Согласно аксиоме отделимости Хаусдорфа, существует окрестность \mathcal{U} точки a и окрестность \mathcal{V} точки b , не имеющие общих точек. Тогда, с одной стороны, должно существовать такое целое число m_0 , что $x_n \in \mathcal{U}$ при $n \geq m_0$, а с другой — такое целое число n_0 , что $x_n \in \mathcal{V}$ при $n \geq n_0$. Для $n \geq \max(m_0, n_0)$ получим $x_n \in \mathcal{U} \cap \mathcal{V}$, а это невозможно, так как пересечение \mathcal{U} и \mathcal{V} пусто.

Теорема 15. Для того чтобы точка a метризуемого пространства E принадлежала замыканию части A из E , необходимо и достаточно, чтобы существовала последовательность элементов из A , сходящаяся к a ¹⁾.

Доказательство. Условие, очевидно, достаточно, так как если такая последовательность существует, то любая окрестность \mathcal{V} точки a содержит хотя бы одну точку этой последовательности, т. е. хотя бы одну точку A , а это означает, что a принадлежит замыканию A (см. теорему 4).

Обратно, пусть a принадлежит замыканию A . Выберем какую-либо метрику, определяющую топологию E . Тогда шар с центром в точке a радиуса $1/n$ содержит по крайней мере одну точку x_n , принадлежащую A . Образованная таким образом последовательность x_0, x_1, x_2, \dots принадлежит A и сходится к a .

Итак, замыкание \bar{A} множества A является множеством пределов последовательностей из A , сходящихся в E .

Следствие. Для того чтобы некоторая часть топологического метризуемого пространства E была замкнутой, необходимо и достаточно, чтобы она содержала пределы всех своих последовательностей, сходящихся в E .

¹⁾ Если E — топологическое, но не метризуемое пространство, то указанное условие будет достаточным, но не будет необходимым.

Теорема 16. Для того чтобы отображение f метризуемого пространства E в метризуемое пространство F было непрерывным в точке $a \in F$, необходимо и достаточно, чтобы образ при отображении f любой сходящейся к a последовательности точек из E был последовательностью точек из F , сходящейся к $f(a)$ ¹⁾.

1°) Условие необходимо. В самом деле, пусть отображение f непрерывно в точке a и x_0, x_1, x_2, \dots — сходящаяся к a последовательность из E . Тогда, какова бы ни была окрестность \mathcal{U} точки $f(a)$ в F , ее прообраз при отображении f является некоторой окрестностью \mathcal{U} точки a в E . Но тогда для всех целых n , кроме конечного числа, x_n лежат в \mathcal{U} и, следовательно, $f(x_n)$ находятся в \mathcal{U} , что доказывает сходимость последовательности $f(x_n)$ к $f(a)$.

2°) Условие достаточно. Предположим, что указанное в теореме условие выполнено и что в E выбрана некоторая метрика, определяющая его топологию. Если бы отображение f не было непрерывным в a , то нашлось бы такое число $\varepsilon > 0$, что при любом выборе $\eta > 0$ существовала бы точка x , такая, что $d(x, a) \leq \eta$ и в то же время $d(f(x), f(a)) > \varepsilon$.

В частности, при таком выборе ε для каждого целого n найдется хотя бы одна точка x_n , такая, что $d(x_n, a) \leq 1/n$ и $d(f(x_n), f(a)) > \varepsilon$.

Последовательность x_n , очевидно, сходится при этом в E к a , в то время как последовательность $f(x_n)$ в F к $f(a)$ не сходится, что противоречит предположению.

§ 6. ТОПОЛОГИЧЕСКОЕ ПРОИЗВЕДЕНИЕ

Обозначим через d_1 и d_2 соответственно расстояния, введенные в двух метрических пространствах E_1 и E_2 . На множестве произведения $E_1 \times E_2$ с помощью формулы

$$\delta((x_1, x_2), (y_1, y_2)) = \max(d_1(x_1, y_1), d_2(x_2, y_2)) \quad (\text{II}, 6; 1)$$

можно определить метрику, которую мы будем записывать в виде $\delta = \max(d_1, d_2)$. Точно так же по формуле $(d_1 + d_2)(x_1, x_2), (y_1, y_2)) = d_1(x_1, y_1) + d_2(x_2, y_2)$ можно ввести метрику, записываемую в виде $d_1 + d_2$, или же метрику $\sqrt{d_1^2 + d_2^2}$, определяемую выражением $\sqrt{d_1^2 + d_2^2}((x_1, x_2), (y_1, y_2)) = \sqrt{(d_1(x_1, y_1))^2 + (d_2(x_2, y_2))^2}$. Очевидно, ни одна из этих метрик не выделяется среди других. Впрочем, легко видеть, что эти метрики эквивалентны, т. е. они определяют в $E_1 \times E_2$ одну и ту же топологию. Это рассуждение подчеркивает равнозначенность метрик: ни одна из них существенно не выделяется среди

1) Если E и F — топологические, но не метризуемые пространства, то условие остается необходимым, но не достаточным.

других метрик на произведении. С другой стороны, легко определить *естественную топологию* на произведении двух топологических пространств:

Рассмотрим в пространстве E_1 открытую часть A_1 , а в пространстве E_2 открытую часть A_2 и определим произведение $A_1 \times A_2$, т. е. множество пар (x, y) из $E_1 \times E_2$, таких, что $x \in A_1$ и $y \in A_2$ ¹⁾.

Говорят, что некоторая часть $E_1 \times E_2$ открыта в топологии произведения, если вместе с любой точкой произведения она содержит хотя бы одно произведение открытых множеств $A_1 \times A_2$, содержащее эту точку.

Легко проверяется, что выделенные таким образом на $E_1 \times E_2$ системы множеств удовлетворяют всем аксиомам, которым должны удовлетворять открытые множества некоторой топологии. При этом произведение открытых частей открыто, но имеются и другие открытые множества. Определенная таким образом топология на $E_1 \times E_2$ называется *произведением топологий*, заданных на E_1 и E_2 . Точно так же определяется топология произведения нескольких пространств. (Например, если $E_1 = E_2 = \dots = E_n = \mathbb{R}$, то их произведение является известным топологическим пространством \mathbb{R}^n , определенным, например, с помощью естественной метрики.) Из самого определения этой топологии вытекает, что *канонические проекции* $E_1 \times E_2$ на E_1 и E_2 непрерывны. В самом деле, если f является проекцией $E_1 \times E_2$ на E_1 и если A_1 есть некоторая открытая часть E_1 , то ее прообразом будет множество $f^{-1}(A_1) = A_1 \times E_2$, которое открыто в $E_1 \times E_2$ и даже есть произведение открытых множеств.

Сходящиеся последовательности в произведении

Теорема 17. Для того чтобы последовательность $(x_0, y_0), (x_1, y_1), (x_2, y_2), \dots$ ²⁾ элементов из $E \times F$ сходилась к точке (a, b) относительно произведения топологий, необходимо и достаточно, чтобы последовательность x_n сходилась к a в E , а последовательность y_n сходилась к b в F .

Предположим сначала, что последовательность (x_n, y_n) сходится к (a, b) . Из теоремы 16, в силу непрерывности проекций $(x, y) \rightarrow x$ и $(x, y) \rightarrow y$, вытекает, что последовательность x_n сходится к a , а последовательность y_n сходится к b .

Обратно, предположим, что последовательность x_n сходится к a в E , а последовательность y_n сходится к b в F . Пусть \mathcal{U} — окрестность (a, b) в $E \times F$. Она содержит некоторое открытое

¹⁾ $A_1 \times A_2$ иногда называют *открытым прямоугольником* по аналогии с *прямоугольником* в \mathbb{R}^2 , являющимся произведением интервалов из \mathbb{R} .

²⁾ Мы обозначили здесь через E и F топологические пространства, чтобы сохранить нумерацию 1, 2, ... элементов последовательности.

множество, содержащее (a, b) , а, значит, согласно определению открытых множеств в $E \times F$, содержит некоторое открытое произведение $A \times B$, $a \in A$, $b \in B$. Согласно предположению о сходимости последовательностей x_n и y_n в этом случае существуют такие целые числа n_1 и n_2 , что $x_n \in A$ для $n \geq n_1$ и $y_n \in B$ для $n \geq n_2$. Но тогда при $n \geq \max(n_1, n_2)$ пара $(x_n, y_n) \in A \times B \subset \mathcal{Y}$, что означает сходимость последовательности (x_n, y_n) к (a, b) .

Пусть G — топологическое пространство. Отображение h множества E в множество $F \times G$ задается парой отображений f и g множества E в F и G соответственно, таких, что $h(x) = (f(x), g(x))$ (гл. I, стр. 13). Для того чтобы это отображение было непрерывным, необходимо и достаточно, чтобы каждое отображение f и g в отдельности было непрерывным. Доказательство очевидно.

Непрерывные функции многих переменных

В математике часто встречается понятие непрерывной функции двух переменных.

Пусть f — функция двух переменных: $(x, y) \rightarrow f(x, y)$. Это — отображение произведения пространств $E \times F$ в пространство G . Предположим, что все три пространства являются топологическими. Выясним, когда можно утверждать, что функция f непрерывна в точке (a, b) ?

В случае метрических пространств непрерывность означает, что, каково бы ни было $\varepsilon > 0$, существует такое $\eta > 0$, что из неравенств $d(x, a) \leq \eta$ и $d(y, b) \leq \eta$ следует неравенство $d(f(x, y), f(a, b)) \leq \varepsilon$ ¹⁾.

Однако это означает лишь, что *отображение f пространства $E \times F$ в G непрерывно в точке (a, b) в смысле произведения топологий на $E \times F$* . Именно так мы и будем определять непрерывность функции двух переменных в общем случае. Задание непрерывного отображения произведения $E_1 \times E_2 \times \dots \times E_n$ в произведение $F_1 \times F_2 \times \dots \times F_m$ равносильно заданию системы m непрерывных функций n переменных.

Теорема 17₂. *Если E — метрическое пространство, то функция расстояния d , отображающая $E \times E$ в вещественную прямую \mathbb{R} , непрерывна.*

В самом деле, каково бы ни было $\varepsilon > 0$, из неравенств $d(x, a) \leq \varepsilon/2$ и $d(y, b) \leq \varepsilon/2$ вытекает неравенство $|d(x, y) - d(a, b)| \leq d(x, a) + d(y, b) \leq \varepsilon$, означающее, что функция d непрерывна.

¹⁾ Расстояния в E , F , G для простоты мы обозначим тем же символом d .

Топологические группы.

Топологические векторные пространства

Топологической группой называется множество G , имеющее, с одной стороны, групповую структуру и наделенное, с другой стороны, такой топологией, что основные отображения, определяющие групповую структуру, т. е. отображение $(x, y) \rightarrow xy$ множества $G \times G$ в G и отображение $x \rightarrow x^{-1}$ множества G в G , являются непрерывными.

Топологическим векторным пространством называется множество E , имеющее, с одной стороны, структуру векторного пространства над полем вещественных или комплексных чисел и обладающее, с другой стороны, такой топологией, относительно которой сложение $(\vec{x}, \vec{y}) \rightarrow \vec{x} + \vec{y}$ является непрерывным отображением $E \times E$ в E^1), а умножение на скаляры $(\lambda, \vec{x}) \rightarrow \vec{\lambda x}$ является непрерывным отображением $K \times E$ в E , где K — поле скаляров с естественной топологией. Легко видеть, что *нормированное векторное пространство* является *топологическим векторным пространством*. В самом деле,

$$\begin{aligned} \|(\vec{x} + \vec{y}) - (\vec{a} + \vec{b})\| &\leqslant \|\vec{x} - \vec{a}\| + \|\vec{y} - \vec{b}\|, \\ \|\lambda \vec{x} - \vec{aa}\| &\leqslant |\lambda - a| \|\vec{x}\| + |a| \|\vec{x} - \vec{a}\|. \end{aligned} \quad (\text{II}, 6; 2)$$

Отсюда вытекает, что для $\|\vec{x} - \vec{a}\| \leqslant \varepsilon/2$, $\|\vec{y} - \vec{b}\| \leqslant \varepsilon/2$ справедливо неравенство $\|(\vec{x} + \vec{y}) - (\vec{a} + \vec{b})\| \leqslant \varepsilon$, выражающее непрерывность сложения. Далее, если последовательность (λ_n, \vec{x}_n) сходится к (α, \vec{a}) в $K \times E$, т. е. если последовательность \vec{x}_n сходится к \vec{a} и если последовательность λ_n сходится к α (теорема 17), то $\|\vec{x}_n - \vec{a}\|$, а вместе с ним и $|\alpha| \|\vec{x}_n - \vec{a}\|$, сходится к нулю. С другой стороны, $|\lambda_n - \alpha|$ стремится к нулю и $\|\vec{x}_n\|$ ограничена, а, значит, $|\lambda_n - \alpha| \|\vec{x}_n\|$ сходится к нулю. Из второго неравенства вытекает, что $\lambda_n \vec{x}_n$ сходится к $\alpha \vec{a}$. Согласно теореме 16, это доказывает непрерывность отображения $(\lambda, \vec{x}) \rightarrow \lambda \vec{x}$ множества $K \times E$ в E .

Теорема 13 обобщается следующим образом: в конечномерном векторном пространстве существует только одна топология топологического векторного пространства.

¹⁾ Непрерывность сложения означает, что предел суммы двух векторов является суммой пределов этих векторов.

Раздельная непрерывность функции двух переменных

Пусть f — некоторое отображение $E \times F$ в G и (a, b) — точка $E \times F$. Зафиксировав $x = a$, мы получим отображение F в G : $y \rightarrow f(a, y)$, определяемое с помощью f и a . Это отображение часто обозначают через f_a , так что $f_a(y) = f(a, y)$. Иногда его обозначают через $f(a, \cdot)$, опуская переменную y . Может случиться, что это отображение непрерывно в точке $y = b$. В этом случае говорят, что отображение f *раздельно непрерывно по y* в точке b для фиксированного $x = a$. Такое же определение дается для отображения f_b : $x \rightarrow f(x, b)$, обозначаемого также через $f(\cdot, b)$, и для понятия *раздельной непрерывности* в точке a по x при фиксированном $y = b$.

В обоих случаях речь идет о том, что сужение отображения f на подпространство $\{a\} \times F$ или подпространство $E \times \{b\}$ непрерывно в точке (a, b) . Говорят, что f *раздельно непрерывно* в точке (a, b) , если оно раздельно непрерывно по обеим переменным в указанном выше смысле. Отображение называется *раздельно непрерывным* на $E \times F$, если оно раздельно непрерывно в каждой точке (a, b) произведения $E \times F$.

Непрерывная функция является, очевидно, раздельно непрерывной, однако обратное, вообще говоря, не верно. Рассмотрим, например, функцию двух вещественных переменных, определенную формулой:

$$f(x, y) = \frac{xy}{x^2 + y^2} \text{ для } (x, y) \neq (0, 0) \text{ и } f(0, 0) = 0. \quad (\text{II}, 6; 3)$$

Эта функция, очевидно, непрерывна всюду на дополнении к началу координат в \mathbb{R}^2 . Она непрерывна на оси $x'x$ и на оси $y'y$, где она $\equiv 0$. Следовательно, она раздельно непрерывна в начале координат, а значит, и во всей плоскости. Однако, эта функция не является непрерывной по совокупности обеих переменных в начале координат, ибо на прямой $y = mx$ вне начала координат она принимает значение $m/(1 + m^2) \neq 0$ и не стремится к нулю, когда (x, y) стремится к началу координат.

§ 7. КОМПАКТНЫЕ ПРОСТРАНСТВА. ЭЛЕМЕНТАРНЫЕ СВОЙСТВА

Пусть E — некоторое топологическое пространство. Множество частей E , таких, что каждая точка E принадлежит *по крайней мере одной* из этих частей, называется *покрытием* множества E . *Подпокрытием* покрытия называется покрытие, образованное из частей первого покрытия. Покрытие называется *конечным*, если оно состоит из конечного числа частей E . Покрытие называется *открытым*, если все части этого покрытия открыты в E . Так, например, множество интервалов $[n - 1, n + 1]$ при n , пробегающем \mathbb{Z} , образует *открытое покрытие* \mathbb{R} . В этом

случае у этого покрытия других подпокрытий не существует, так как достаточно удалить какой-либо интервал $]n - 1, n + 1[$, как соответствующая точка не будет покрыта¹⁾.

Определение. Свойство Гейне — Бореля — Лебега. Топологическое пространство называется *компактным*, если из любого его открытого покрытия E можно выделить хотя бы одно конечное подпокрытие. Это означает, что для любого покрытия пространства E множеством открытых частей существует конечное множество этих частей, покрывающее E .

Заметим, что до настоящего момента мы изучали топологические свойства лишь некоторых частей E относительно самого множества E , например свойство части A из E быть открытой, замкнутой и т. д. Свойство же компактности является свойством самого топологического пространства. Однако можно говорить о компактности части A топологического пространства E , если в смысле топологии, индуцированной в A топологией из E , часть A является компактным пространством. При этом само E не обязано быть компактным! Всякая часть E , замыкание которой компактно, называется *относительно компактной* частью E .

Легко видеть, что *объединение конечного числа компактных частей E компактно*.

В самом деле, пусть A_1, A_2, \dots, A_n — компактные части E и \mathcal{R} — некоторое открытое покрытие объединения $A = A_1 \cup A_2 \cup \dots \cup A_n$. Так как пересечение открытого множества из \mathcal{R} с A_1 является открытым множеством в A_1 , то \mathcal{R} определяет открытое покрытие множества A_1 . Поскольку часть A_1 компактна, то ее можно покрыть конечным числом открытых множеств из \mathcal{R} . То же самое верно и для A_2, \dots, A_n . Значит, выбирая одновременно эти n конечных систем открытых множеств из \mathcal{R} , можно покрыть все A конечным числом открытых множеств из \mathcal{R} ; а это означает, что A компактно. В силу тех же соображений *объединение конечного числа относительно компактных множеств относительно компактно*.

Примеры. Пространство, содержащее только конечное число точек, компактно. вещественная прямая \mathbb{R} , векторное пространство \mathbb{R}^n , векторное нормированное конечномерное или бесконечномерное пространство не компактны. В самом деле, рассмотрим множество открытых шаров с центром в нуле и радиусом > 0 . Они образуют, очевидно, покрытие пространства.

¹⁾ Множество интервалов $]n - \frac{3}{4}, n + \frac{3}{4}[$ также является открытым покрытием \mathbb{R} . Оно образовано из открытых частей, меньших, чем предыдущие части. Однако это — не подпокрытие рассматриваемого покрытия, так как оно является множеством частей, существенно отличных от первых.

Далее, любое конечное число этих шаров содержится в одном шаре конечного радиуса и, следовательно, не покрывает всего пространства. Более общо, любое *неограниченное подмножество метрического пространства*, т. е. не содержащееся хотя бы в одном шаре (конечного радиуса), *не компактно*.

Теорема 18. *Замкнутый ограниченный интервал $[a, b]$ вещественной прямой является компактным пространством.*

Доказательство. Пусть \mathcal{X} — открытое покрытие интервала $[a, b]$, и пусть c — середина $[a, b]$. Предположим, что из \mathcal{X} невозможно выбрать конечного числа частей, покрывающих полностью $[a, b]$. Тогда это невозможно сделать по крайней мере для одного из подинтервалов $[a, c]$ или $[c, b]$, например для $[a, c]$. Обозначим этот интервал через $[a_1, b_1]$. Разделив его на два, мы можем найти подинтервал $[a_2, b_2]$, в два раза меньший интервала $[a_1, b_1]$ и обладающий тем же свойством. Таким образом можно построить бесконечную последовательность $[a_0, b_0] = [a, b], [a_1, b_1], [a_2, b_2], \dots$ подинтервалов $[a, b]$, обладающих общим свойством: ни один из них не может быть покрыт конечным числом частей, принадлежащих \mathcal{X} .

Возрастающая ограниченная сверху последовательность a_n сходится к некоторому пределу α . В свою очередь убывающая ограниченная снизу последовательность b_n сходится к некоторому пределу β , а так как длина отрезка $[a_n, b_n]$ равна $(b - a)/2^n$, то необходимо $\alpha = \beta$. Согласно определению предела последовательности, любой открытый интервал, содержащий $\alpha = \beta$, содержит все a_n и b_n с достаточно большими номерами, а значит, содержит и все интервалы $[a_n, b_n]$ с этими номерами.

Далее, в покрытии \mathcal{X} существует открытое множество \mathcal{Y} , содержащее точку $\alpha = \beta$. Поскольку \mathcal{Y} открыто, то существует открытый интервал $]a', b'[$, содержащийся в \mathcal{Y} и содержащий эту точку. Для достаточно больших n интервал $[a_n, b_n]$ будет находиться в $]a', b'[$, а значит, и в \mathcal{Y} . Мы пришли к противоречию: по построению интервал $[a_n, b_n]$ не может быть покрыт конечным числом частей из \mathcal{X} , и в то же время он покрывается одной из них, а именно \mathcal{Y} . Это противоречие доказывает компактность $[a, b]$.

Более общо, в пространстве \mathbb{R}^m замкнутый ограниченный параллелепипед, т. е. множество точек (x_1, x_2, \dots, x_m) , определенных системой неравенств $a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, \dots, a_m \leq x_m \leq b_m$, где a_1, a_2, \dots, a_m и b_1, b_2, \dots, b_m — конечные числа, является *компактным пространством*. Этот факт доказывается тем же методом деления с той лишь разницей, что вместо деления на две части на каждом шаге операции приходится

разбивать параллелепипед на 2^m частей, деля на два каждую из m координат.

Теорема 19. *Полная прямая $\bar{\mathbb{R}}$, снабженная естественной топологией, является компактным пространством.*

Доказательство. Пусть \mathcal{A} — открытое покрытие $\bar{\mathbb{R}}$. Точка $-\infty$ принадлежит хотя бы одному открытому множеству, скажем \mathcal{U}_- , из \mathcal{A} , и точка $+\infty$ принадлежит хотя бы одному открытому множеству, скажем \mathcal{U}_+ , из \mathcal{A} . Отсюда, в частности, вытекает, что дополнение к объединению этих двух открытых множеств содержится в достаточно большом замкнутом ограниченном интервале $[A, B]$. Поскольку в этом случае существует конечное число открытых множеств из \mathcal{A} , покрывающих $[A, B]$ (теорема 18), то это конечное число открытых множеств, дополненное открытыми множествами \mathcal{U}_- и \mathcal{U}_+ , покрывает все $\bar{\mathbb{R}}$, откуда и следует компактность $\bar{\mathbb{R}}$.

Замечания. 1°) Из определения компактности вытекает следующее свойство:

Пусть компактное пространство E является объединением возрастающей последовательности открытых множеств $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n, \dots$. Тогда при некотором n открытое множество \mathcal{U}_n совпадает с E .

2°) Если два пространства гомеоморфны и одно из них компактно, то второе также компактно. Из предыдущих примеров следует, что \mathbb{R} и $\bar{\mathbb{R}}$ не гомеоморфны (не существует гомеоморфизма одного множества на другое). Замкнутый ограниченный интервал $[a, b]$ из \mathbb{R} не гомеоморден \mathbb{R} (в то время как открытый интервал гомеоморден \mathbb{R} : $x \rightarrow \operatorname{tg} x$ является, например, гомеоморфизмом $[\pi/2, \pi/2]$ на \mathbb{R}). Критерии, позволяющие установить негомеоморфность двух пространств, слишком сложны; доказать же, что два пространства гомеоморфны, обычно проще всего путем установления соответствующего гомеоморфизма.

Теорема 20. *Для того чтобы топологическое пространство E было компактным, необходимо и достаточно, чтобы из любого множества замкнутых частей E , пересечение которых пусто, можно было выбрать конечное множество частей с пустым пересечением.*

Доказательство. Для доказательства достаточно в формулировке теоремы заменить части их дополнениями. Объединение при этом перейдет в пересечение, а открытые части перейдут в замкнутые.

Следствие 1. *Если E — компактное пространство, а F_0, F_1, F_2, \dots — убывающая последовательность замкнутых подмно-*

жеств E , пересечение которых пусто, то существует такое n , что F_n пусто.

Заметим, что указанное свойство, очевидно, не верно для вещественной прямой \mathbb{R} , и это подтверждает отмеченный ранее факт некомпактности \mathbb{R} . В самом деле, если мы рассмотрим убывающую последовательность замкнутых интервалов $[n, +\infty[$, то их пересечение пусто, в то время как ни один из них не пуст.

Замечание. Если воспользоваться теоремой 11 гл. I, то можно заметить, что предыдущее следствие эквивалентно следующему:

Следствие 2. *Если E — компактное пространство, а $F_0, F_1, F_2 \dots$ — убывающая последовательность замкнутых подмножеств E , ни одно из которых не пусто, то их пересечение также не пусто.*

Теорема 21. *Пусть E — топологическое пространство, а F — компактная часть E . Тогда заведомо F является замкнутой частью E ¹⁾.*

Доказательство. Для простоты при доказательстве мы ограничимся тем случаем, когда E является метрическим пространством. Пусть a — точка приоснования F . Нам надо доказать, что $a \in F$. Если мы через F_n обозначим пересечение F и замкнутого шара с центром в a радиуса $1/n$, n — целое $\geqslant 1$, то мы увидим, что F_n образуют убывающую последовательность замкнутых частей F (теорема 5). Ни одно из F_n не пусто, ибо a является точкой приоснования F (теорема 4), а так как F , по предположению, компактно, то их пересечение не пусто. Однако пересечение рассматриваемых шаров сводится к точке a . Следовательно, a — единственная точка пересечения всех F_n . Отсюда вытекает, что a принадлежит всем F_n , а значит, и множеству F , что доказывает замкнутость F в E .

Очевидно, что обратное утверждение к этой теореме, вообще говоря, не верно. Произвольное замкнутое множество топологического пространства не обязательно компактно, так как в противном случае пространство, являясь замкнутым, всегда было бы компактным. Однако имеет место следующее обратное утверждение:

Теорема 22. *Всякая замкнутая часть компактного пространства является компактным пространством.*

Доказательство. Пусть E — компактное пространство, F — замкнутая часть E . Пусть $(F_i)_{i \in I}$ — некоторое семейство

¹⁾) Внимание! Эта теорема сравнивает внутреннее свойство F — свойство быть компактным (в индуцированной топологии) — со свойством множества F относительно E быть замкнутым в E .

частей F , замкнутых в F , пересечение которых пусто. Поскольку F_i замкнуты в F и пространство F , по предположению, замкнуто в E , то F_i замкнуты в E (теорема 6, б)). Так как E , по предположению, компактно, то существует конечное число частей F_i , пересечение которых пусто, а это, согласно теореме 20, обеспечивает компактность F .

Из теорем 21 и 22 следует, что в компактном пространстве компактные подмножества совпадают с замкнутыми подмножествами. Так, например, для того чтобы некоторое подмножество пополненной прямой $\bar{\mathbb{R}}$ было компактным, необходимо и достаточно, чтобы оно было замкнутым в $\bar{\mathbb{R}}$.

Теорема 22₂. *Всякое компактное пространство регулярно: любая его точка обладает фундаментальной системой компактных окрестностей.*

Доказательство. Пусть \mathcal{U} — некоторая окрестность точки a и \mathcal{U}' — ее внутренность. Тогда $K = \mathbb{C} \setminus \mathcal{U}'$ замкнуто и, в силу теоремы 22, компактно. Пусть \mathcal{W} — замкнутая (и, следовательно, компактная) окрестность точки a . Тогда $\mathcal{W} \cap K$ является замкнутым подмножеством K . При всевозможных \mathcal{W} эти замкнутые подмножества из K имеют пустое пересечение, поскольку пересечение всех \mathcal{W} сводится к точке a (стр. 59), которая не принадлежит K . Так как K компактно, то из множества замкнутых окрестностей \mathcal{W} можно выбрать конечное число множеств $\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_n$, таких, что пересечение всех $\mathcal{W}_i \cap K$ пусто. Множество $\mathcal{W}_0 = \mathcal{W}_1 \cap \mathcal{W}_2 \cap \dots \cap \mathcal{W}_n$ является такой компактной окрестностью точки a , что $\mathcal{W}_0 \cap K$ пусто; при этом $\mathcal{W}_0 \subset \mathcal{W}$. Таким образом, любая окрестность \mathcal{U} точки a содержит компактную окрестность \mathcal{W}_0 , что и требовалось доказать.

Более общо, тем же самым методом доказывается, что в компактном пространстве каждая компактная часть имеет фундаментальную систему компактных окрестностей.

Теорема 23. *Для того чтобы часть конечномерного векторного нормированного пространства E была компактной, необходимо и достаточно, чтобы она была замкнутой и ограниченной.*

Доказательство. 1°) Указанное условие необходимо. Как мы видели вначале, компактная часть всегда ограничена, и в силу теоремы 21, она заведомо замкнута.

2°) Условие достаточно. Если даже полем скаляров является \mathbb{C} , мы можем рассматривать n -мерное векторное пространство над \mathbb{C} как векторное пространство над \mathbb{R} размерности $2n$. Поэтому можно предполагать, что полем скаляров является \mathbb{R} .

а) Предположим сначала, что пространством E является \mathbb{R}^n и что норма в нем определена функцией $(x_1, x_2, \dots, x_n) \rightarrow \max_{i=1, 2, \dots, n} |x_i|$.

Любая ограниченная часть содержится в некотором шаре, являющемся в нашем случае замкнутым ограниченным параллелепипедом, т. е. содержитя в некотором компакте. Поскольку рассматриваемая часть предполагается замкнутой, она является замкнутой частью компактного множества, и потому для доказательства достаточно будет применить теорему 22. (Более общо, если в метрическом пространстве любой замкнутый шар компактен, то любая его ограниченная замкнутая часть будет также компактной.)

б) Пусть теперь E является произвольным нормированным n -мерным векторным пространством над \mathbb{R} .

Докажем сначала утверждение теоремы 13. Выберем в E какой-либо базис $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. Тогда каждая точка \vec{x} может быть выражена через свои координаты x_1, x_2, \dots, x_n . Обозначим через $\|\vec{x}\|$ заданную норму в E и положим $|\vec{x}| = \max_{i=1, 2, \dots, n} |x_i|$.

Тогда $\|\vec{x}\| = \left\| \sum_i x_i \vec{e}_i \right\| \leq \sum_i |x_i| \|\vec{e}_i\| \leq \sum_i \|\vec{e}_i\| \max_{i=1, 2, \dots, n} |x_i| = \sum_i \|\vec{e}_i\| |\vec{x}|$.

Если положить $k = \sum_{i=1}^n \|\vec{e}_i\|$, то получим неравенство

$$\|\vec{x}\| \leq k |\vec{x}|. \quad (\text{II}, 7; 1)$$

Нам нужно получить противоположное неравенство. Из неравенства $(\text{II}, 7; 1)$ следует, что тождественное преобразование E с нормой $|\vec{x}_i|$ в E с нормой $\|\vec{x}\|$ непрерывно. (В самом деле, каково бы ни было $\varepsilon > 0$, полагая $\eta = \varepsilon/k$ для всех \vec{x} , таких, что $|\vec{x} - \vec{a}| \leq \eta$, получим $\|\vec{x} - \vec{a}\| \leq k |\vec{x} - \vec{a}| \leq k\eta = \varepsilon$ — это — определение непрерывности преобразования в точке \vec{a} .) Но тогда прообраз любого замкнутого множества замкнут (теорема 8). Другими словами, любая часть F пространства E , замкнутая в метрике $\|\cdot\|$, замкнута и в метрике $|\cdot|$. В частности, шар $\|\vec{x}\| \leq R$, будучи замкнутым в метрике $\|\cdot\|$, замкнут и в метрике $|\cdot|$. Обозначим через Q сферу (полый куб) $|\vec{x}| = 1$. Это — некоторая замкнутая и ограниченная в E часть по метрике $|\cdot|$ и, следовательно, компактная часть в силу того, что мы видели в начале доказательства. Пусть F_R — пересечение Q с

шаром $\|\vec{x}\| \leq R$, замкнутое по метрике $||$. Это — некоторая замкнутая часть компакта Q . Так как пересечение всех шаров $\|\vec{x}\| \leq R$ содержит лишь одну точку — начало координат, не принадлежащую Q , то пересечение всех рассмотренных замкнутых подмножеств пусто. Согласно теореме 20, из них можно выбрать конечное число с пустым пересечением. Другими словами, существует такое число $\rho > 0$, что шар $\|\vec{x}\| \leq \rho$ с множеством Q не пересекается. Но тогда из $\|\vec{x}\| \leq \rho$ следует $|\vec{x}| < 1$ (в противном случае нашлась бы точка \vec{x}_0 , удовлетворяющая неравенствам $\|\vec{x}_0\| \leq \rho$ и $|\vec{x}_0| \geq 1$, и тогда для $\lambda = 1/|\vec{x}_0|$ мы получили бы $\|\lambda\vec{x}_0\| \leq \lambda\rho \leq \rho$ и $|\lambda\vec{x}_0| = 1$, что противоречит изложенному выше). С помощью гомотетии с отношением μ отсюда получаем, что неравенство $\|\vec{x}\| \leq \mu\rho$ влечет за собой неравенство $|\vec{x}| < \mu$ и тем более $\leq \mu$. Первое неравенство справедливо при $\mu = \|\vec{x}\|/\rho$, но тогда второе принимает вид

$$|\vec{x}| \leq \frac{1}{\rho} \|\vec{x}\|, \quad (\text{II}, 7; 2)$$

что вместе с (II, 7; 1) доказывает эквивалентность норм $||$ и $\|\cdot\|$, т. е. теорему 13.

Теорема 23 теперь доказана во всех случаях, ибо понятия «замкнутый, ограниченный, компактный» одни и те же для двух эквивалентных норм (см. примечание²) на стр. 58), а то, что доказано в п. а) для специальной нормы в E , определенной с помощью базиса, сохраняется и для данной нормы, эквивалентной ей по теореме 13.

Локально компактные пространства

Говорят, что топологическое пространство E локально компактно, если каждая его точка имеет по крайней мере одну компактную окрестность. Так, например, любое конечномерное нормированное векторное пространство локально компактно, поскольку локально компактен замкнутый шар такого пространства.

Можно доказать, что бесконечномерное нормированное векторное пространство локально компактным быть не может (теорема 45₂). Замкнутый шар в таком пространстве не компактен. Для того чтобы некоторая часть пространства была компактной, необходимо, но не достаточно, чтобы она была замкнутой и ограниченной.

Если два топологических пространства гомеоморфны и одно из них локально компактно, то таким же будет и второе. Это

дает нам новый критерий, позволяющий выяснить, когда два пространства не являются гомеоморфными. Например, конечно-мерное и бесконечномерное векторные нормированные пространства гомеоморфными быть не могут.

Любое локально компактное пространство регулярно, и каждая точка в нем обладает фундаментальной системой компактных окрестностей. В самом деле, если \mathcal{U} является компактной окрестностью точки a , то, согласно теореме 22₂, точка a имеет в \mathcal{U} фундаментальную систему компактных окрестностей. Остается заметить, что каждая окрестность точки a в \mathcal{U} является также окрестностью точки a во всем пространстве (теорема 6, с)). Более общо, *каждый компакт локально компактного пространства обладает фундаментальной системой компактных окрестностей*.

Точка сгущения последовательности¹⁾

Пусть x_0, x_1, x_2, \dots — некоторая последовательность элементов топологического пространства E . Говорят, что точка a является *точкой сгущения* этой последовательности, если для любой окрестности \mathcal{U} точки a найдется бесконечное множество значений n , таких, что $x_n \in \mathcal{U}$. Если последовательность сходится к a , то a для нее является точкой сгущения.

Теорема 24. *Пусть E — метризуемое пространство и a — точка сгущения последовательности x_0, x_1, x_2, \dots элементов E . В этом, и только в этом случае, из этой последовательности можно выбрать подпоследовательность, сходящуюся к a .*

Доказательство. Подпоследовательностью последовательности x_0, x_1, x_2, \dots называется последовательность вида $x_{p_0}, x_{p_1}, x_{p_2}, \dots$, где $p \rightarrow p_n$ является строго возрастающим отображением \mathbb{N} в \mathbb{N} ²⁾. Исходя из определения, теперь легко проверить, что если существует подпоследовательность, сходящаяся к a , то точка a является точкой сгущения для исходной последовательности (и это так даже в том случае, когда E — не-метризуемое топологическое пространство). Докажем теперь обратное, предполагая, что в E выбрана некоторая метрика, определяющая ее топологию. Пусть a — точка сгущения исходной последовательности. Тогда для каждого шара $B(a, 1/n)$ существует бесконечное множество значений p , таких, что

¹⁾ Ее часто называют также *точкой приоснования* последовательности. Однако это может привести к ошибочному отождествлению с понятием точки приоснования множества. По этой причине мы используем различные слова.

²⁾ Полагая $p_n = n$, получаем, что последовательность является подпоследовательностью самой себя.

$x_p \in B(a, 1/n)$. Возьмем сначала такое целое число p_1 , что $x_{p_1} \in B(a, 1)$. Выберем затем такое целое $p_2 > p_1$, что $x_{p_2} \in B(a, 1/2)$, затем такое целое $p_3 > p_2$, что $x_{p_3} \in B(a, 1/3)$ и т. д. Шаг за шагом таким методом мы образуем подпоследовательность $n \rightarrow x_{p_n}$ исходной последовательности, которая, очевидно, сходится к a .

Теорема 25 (свойство Больцано—Вейерштрасса). Для того чтобы метризуемое пространство E было компактным, необходимо и достаточно, чтобы любая последовательность элементов из E имела по крайней мере одну точку сгущения¹⁾.

Доказательство. 1°) Пусть E компактно и x_0, x_1, x_2, \dots — некоторая последовательность его элементов. Обозначим через A_n множество $\{x_n, x_{n+1}, \dots\}$, а через \bar{A}_n — его замыкание. Множества \bar{A}_n образуют убывающую последовательность замкнутых множеств. Так как ни одно из этих множеств не пусто, то их пересечение не пусто. Пусть a — точка этого пересечения. Сказать, что a принадлежит \bar{A}_n или что a является точкой приоснования A_n , означает сказать, что каждая окрестность точки a содержит хотя бы одну точку из A_n , а так как это верно при любом n , то мы получаем, что a является точкой сгущения исходной подпоследовательности.

2°) Доказательство обратного утверждения не просто. Мы его получим, предварительно рассмотрев две леммы.

Лемма 1. Пусть E — метрическое пространство, каждая последовательность которого имеет по крайней мере одну точку сгущения. Пусть \mathcal{X} — открытое покрытие E . Тогда существует такое число $\varepsilon > 0$, что любой шар с произвольным центром радиуса $\leq \varepsilon$ будет полностью содержаться хотя бы в одном из открытых множеств покрытия.

Предположим, что это не верно. Тогда для каждого целого n можно будет найти в E такую точку a_n , что шар с центром в a_n радиуса $1/n$ не лежит целиком ни в одном из открытых множеств покрытия. Образуем таким образом бесконечную последовательность a_1, a_2, \dots элементов E . Эта последовательность имеет по крайней мере одну точку сгущения a . Поскольку \mathcal{X} — покрытие, то существует открытое множество \mathcal{V} из \mathcal{X} , содержащее a . Само это множество содержит некоторый шар с центром a радиуса α . Однако существует бесконечное множество значений n и, значит, хотя бы одно такое значение, что одновременно выполняются неравенства $1/n \leq \alpha/2$ и $d(a_n, a) \leq \alpha/2$. Теперь видно, что шар с центром a_n радиуса $1/n \leq \alpha/2$ целиком лежит в шаре $B(a, \alpha)$ и тем более в открытом множестве \mathcal{V} покрытия, что про-

¹⁾ Если E — топологическое, но не метризуемое пространство, то условие останется необходимым, но не будет достаточным.

тиворечит сделанному предположению относительно последовательности a_n . Тем самым лемма доказана.

Лемма 2. Пусть E — метрическое пространство, каждая последовательность которого имеет по крайней мере одну точку сгущения. Тогда, каково бы ни было $\varepsilon > 0$, множество E может быть полностью покрыто конечным числом шаров радиуса ε .

В самом деле, пусть a_0 — любая точка E . Если $B_0(a_0, \varepsilon) = E$, то лемма доказана. Если же это не так, то существует хотя бы одна точка a_1 , не принадлежащая шару $B_0(a_0, \varepsilon)$. Если $B_0(a_0, \varepsilon) \cup B_0(a_1, \varepsilon) = E$, то лемма доказана. Если же это не так, то можно найти точку a_2 , не принадлежащую $B_0(a_0, \varepsilon) \cup B_0(a_1, \varepsilon)$. Продолжая рассуждать таким образом далее, мы сможем образовать последовательность $B_0(a_0, \varepsilon), B_0(a_1, \varepsilon), B_0(a_2, \varepsilon), \dots$ шаров радиуса ε . Если такое построение мы не сможем закончить, то придет к бесконечной последовательности точек a_0, a_1, a_2, \dots , взаимные расстояния между которыми $\geq \varepsilon$. Но это невозможно, так как эта последовательность, как любая бесконечная последовательность, должна иметь точку сгущения a . Последнее означает, что существует бесконечное множество значений n , а значит, хотя бы два различных значения p и q , таких, что $d(a, a_p) \leq \varepsilon/3$ и $d(a, a_q) \leq \varepsilon/3$. Полученное противоречие показывает, что исходные построения закончатся при некотором целом n , а это означает, что E может быть покрыто $n + 1$ открытыми шарами радиуса ε .

С помощью этих двух лемм без труда доказывается вторая часть теоремы 25. Выберем некоторую метрику, определяющую топологию в E . Для доказательства компактности E рассмотрим некоторое открытое покрытие \mathcal{R} . Согласно лемме 1, существует число $\varepsilon > 0$, такое, что каждый шар радиуса $\leq \varepsilon$ полностью лежит по крайней мере в одном из открытых множеств покрытия \mathcal{R} . Согласно лемме 2, множество E может быть полностью покрыто конечным числом шаров $B_0, B_1, B_2, \dots, B_n$ радиуса ε . Поскольку каждый шар B_i целиком содержится в некотором открытом множестве \mathcal{V}_i покрытия \mathcal{R} , мы получаем, что для покрытия E достаточно иметь лишь конечное число открытых множеств $\mathcal{V}_0, \mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_n$ покрытия \mathcal{R} .

Замечание 1. Леммы 1 и 2 определяют свойства компактных метрических пространств¹⁾.

Замечание 2. Теорема Больцано — Вейерштрасса, очевидно, не верна для прямой \mathbb{R} : последовательность целых полу-

¹⁾ Слово «компактность» означает некоторую уплотненность и хорошо отражает суть дела. Лемма 2 отмечает, что если даже шары имеют маленький радиус, все равно E может быть покрыто конечным числом этих шаров. Множество E , таким образом, весьма уплотнено.

жительных чисел не имеет точки сгущения. Напротив, в \bar{R} , в силу ее компактности, такая последовательность сходится к $+\infty$.

Теорема 26. *Топологическое произведение двух компактных пространств компактно.*

Доказательство. Ограничимся доказательством для случая двух метрических компактных пространств E и F . Пусть $(x_0, y_0), (x_1, y_1), (x_2, y_2), \dots$ — некоторая последовательность элементов из $E \times F$. Согласно теореме 25, последовательность x_0, x_1, x_2, \dots имеет по крайней мере одну точку сгущения a в E . Согласно теореме 24, из этой последовательности можно извлечь подпоследовательность $x_{p_0}, x_{p_1}, x_{p_2}, \dots$, сходящуюся к a . Подпоследовательность $y_{r_0}, y_{r_1}, y_{r_2}, \dots$ в свою очередь имеет хотя бы одну точку сгущения b в F , и в силу той же теоремы 24 из нее можно извлечь подпоследовательность $y_{p_{q_0}}, y_{p_{q_1}}, y_{p_{q_2}}, \dots$, сходящуюся к b . Окончательно, подпоследовательность $(x_{p_{q_0}}, y_{p_{q_0}}), (x_{p_{q_1}}, y_{p_{q_1}}), (x_{p_{q_2}}, y_{p_{q_2}}), \dots$ исходной последовательности сходится к (a, b) . Отсюда вытекает, что исходная произвольная последовательность имеет точку сгущения, что полностью доказывает компактность произведения.

Теорема 27. *Пусть E — компактное пространство. Для того чтобы последовательность элементов из E сходилась к точке a , необходимо и достаточно, чтобы точка a для нее была единственной точкой сгущения.*

Доказательство. Условие это очевидным образом необходимо. Если последовательность сходится к a , то точка a для нее является точкой сгущения. Другой точки сгущения b быть не может. В самом деле, существуют окрестность \mathcal{U} точки a и не пересекающаяся с ней окрестность \mathcal{V} точки b , такие, что для всех n , кроме конечного числа их, $x_n \in \mathcal{U}$ и в то же время $x_n \in \mathcal{V}$ для бесконечного множества значений n , а это бессмысленно. Докажем теперь, что указанное условие достаточно. Пусть x_0, x_1, x_2, \dots — последовательность элементов E , для которой точка a является единственной точкой сгущения. Если бы эта последовательность не была сходящейся, то существовали бы по крайней мере одно открытое множество \mathcal{V} , содержащее a , и подпоследовательность x_{p_n} данной последовательности, такие, что все x_{p_n} находились бы в дополнении к \mathcal{V} . Поскольку это дополнение замкнуто, то оно, согласно теореме 22, компактно. Подпоследовательность x_{p_n} должна тогда иметь в $C\mathcal{V}$ хотя бы одну точку сгущения, которая также является точкой сгущения для исходной последовательности, а это

противоречит тому, что данная последовательность имеет лишь одну точку сгущения.

З а м е ч а н и е. Приведенное в теореме утверждение, очевидно, не верно для случая вещественной прямой \mathbb{R} . Например, последовательность $1, 1, 2, 1/3, 3, \dots, n, 1/n, \dots$ имеет единственную точку сгущения 0 ; однако, эта последовательность не сходится.

Верхний и нижний пределы вещественной последовательности

Пусть x_0, x_1, x_2, \dots — последовательность элементов $\bar{\mathbb{R}}$. Поскольку $\bar{\mathbb{R}}$ компактно, то множество ее точек сгущения не пусто. Это — некоторая часть F множества $\bar{\mathbb{R}}$. Множество F как пересечение замкнутых множеств \bar{A}_n , где $A_n = \{x_n, x_{n+1}, \dots\}$, всегда замкнуто (см. стр. 71). Следовательно, F является компактным множеством в $\bar{\mathbb{R}}$. В нем имеется максимальный элемент L и минимальный элемент l ; L называется верхним пределом последовательности, а l — ее нижним пределом. Они обозначаются через $L = \limsup_{n \rightarrow \infty} x_n$ и $l = \liminf_{n \rightarrow \infty} x_n$ соответственно. Верхний предел характеризуется следующим свойством: *какие бы ни были L_1 и L_2 , такие, что $L_1 < L < L_2$, все x_n , кроме конечного их числа, $\leq L_2$, и имеется бесконечное множество значений n , для которых $x_n \geq L_1$.*

Не надо смешивать верхний предел и точную верхнюю грань.

З Например, 1 является точной верхней гранью последовательности $1, 1/2, 1/3, \dots$. Но эта последовательность сходится к нулю, который будет ее единственной точкой сгущения и, следовательно, ее верхним пределом.

Для того чтобы некоторая последовательность из $\bar{\mathbb{R}}$ сходилась, необходимо и достаточно, чтобы ее верхний предел совпадал с ее нижним пределом (ибо это условие выражает единственность точки сгущения).

§ 8. СВОЙСТВА НЕПРЕРЫВНЫХ ФУНКЦИЙ НА КОМПАКТНЫХ ПРОСТРАНСТВАХ

Теорема 28. *Образ компакта при непрерывном отображении является компактом.*

Доказательство. Пусть f — непрерывное отображение компактного топологического пространства E в топологическое пространство F . Конечно, мы не хотим доказать, что F компактно. Надо доказать, что образ $f(E)$ пространства E при отображении f , как подпространство пространства F с индуцированной на нем топологией, компактен.

Пусть \mathcal{R} — открытое покрытие $f(E)$. Прообразы $f^{-1}(\mathcal{V}_i)$ открытых множеств \mathcal{V}_i , составляющих \mathcal{R} , образуют открытое покрытие E . В самом деле, если x — некоторая точка E , то ее образ $f(x)$ принадлежит хотя бы одному из открытых множеств покрытия \mathcal{R} , например \mathcal{V}_i и, следовательно, x принадлежит элементу $f^{-1}(\mathcal{V}_i)$. Множества $f^{-1}(\mathcal{V}_i)$ открыты в силу непрерывности f .

Поскольку E предполагалось компактным, то для его покрытия достаточно конечного числа открытых множеств $f^{-1}(\mathcal{V}_i)$, например $f^{-1}(\mathcal{V}_1), f^{-1}(\mathcal{V}_2), \dots, f^{-1}(\mathcal{V}_n)$. Но это значит, что открытые множества $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_n$ образуют покрытие образа $f(E)$. В самом деле, для $y \in f(E)$ прообраз $f^{-1}(\{y\})$ не пуст. Пусть x — некоторый элемент этого прообраза. Хотя бы одно из открытых множеств $f^{-1}(\mathcal{V}_1), f^{-1}(\mathcal{V}_2), \dots, f^{-1}(\mathcal{V}_n)$, например $f^{-1}(\mathcal{V}_k)$, содержит x , и, следовательно, \mathcal{V}_k содержит $y = f(x)$, что доказывает сказанное выше. В итоге получаем, что $f(E)$ удовлетворяет определению компактных пространств: из любого открытого покрытия можно выделить конечное подпокрытие.

Следствие. Всякая непрерывная биекция компактного пространства E на топологическое пространство F является гомеоморфизмом.

В самом деле, согласно теореме 22, образ каждой замкнутой части E является образом компактной части и, следовательно, по доказанному выше, компактен, а значит, согласно теореме 21, замкнут. Согласно теореме 11, любое биективное и непрерывное отображение, переводящее замкнутые множества в замкнутые, является гомеоморфизмом.

Замечание. Не следует думать, что прообраз компакта при непрерывном отображении является компактом. Пусть, например, f — постоянное отображение некомпактного пространства E в некоторое пространство F . Образ всего множества E является подмножеством, сводящимся к точке b из F , т. е. к некоторому компакту, тогда как прообраз $\{b\}$ при отображении f не компактен, поскольку прообразом является всё E . Заметим также, что если f — непрерывное отображение E в F , то прообразы открытых (или замкнутых) частей открыты (или замкнуты), а образы компактных частей компактны. Здесь имеется существенное различие между образами и прообразами множеств при отображении. Объединяя же два рода утверждений, можно получить следующий результат: если E компактно и если f — непрерывное отображение E в F , то образ любой замкнутой части E при этом отображении замкнут в F . В самом деле, поскольку E компактно, то его замкнутая часть также компактна, а, следовательно, ее образ компактен в F , а значит, и замкнут в F .

Теорема 29. *Непрерывное отображение непустого компактного пространства в $\bar{\mathbb{R}}$ достигает своего максимума и минимума.*

Доказательство. Напомним, что точной верхней гранью функции, определенной на множестве E , со значениями в $\bar{\mathbb{R}}$ называется точная верхняя грань ее значений. Эта грань называется максимумом функции, если она достигается для некоторого значения переменной. Пусть f — непрерывное отображение компактного пространства E в пополненную прямую $\bar{\mathbb{R}}$. Согласно предыдущей теореме, образ $f(E)$ является непустым компактом в $\bar{\mathbb{R}}$. Согласно теореме 23, это непустое замкнутое множество в $\bar{\mathbb{R}}$. Непустое множество в $\bar{\mathbb{R}}$, согласно теореме 2 гл. I (измененной так, как указано на стр. 26), имеет точную верхнюю грань, и, кроме того, согласно той же теореме, эта грань является точкой прикосновения рассматриваемого множества. Поскольку оно замкнуто, точная верхняя грань принадлежит этому множеству и, следовательно, является максимумом функции. Аналогичное доказательство имеет место и для минимума функции.

Теорема 29 имеет важное значение. Поэтому полезно в качестве упражнения дать различные ее доказательства. Доказательство, основанное на свойстве Гейне — Бореля — Лебега, будет дано на стр. 86. Приведем еще одно доказательство. Пусть M — точная верхняя грань функции f ; M_0, M_1, M_2, \dots — возрастающая последовательность чисел $\leq M$, стремящихся к M . Для каждого n множество F_n : $\{x; x \in E, M_n \leq f(x) \leq M\}$ не пусто (теорема 2 гл. I) и замкнуто как прообраз замкнутого интервала $[M_n, M]$ из $\bar{\mathbb{R}}$ при преобразовании f (теорема 8). Последовательность F_n является убывающей последовательностью непустых замкнутых подмножеств компакта E . Согласно следствию из теоремы 20, их пересечение не пусто. Это некоторое множество точек x из E , в которых $f(x) = M$. Таким образом, в множестве E существует хотя бы одна точка, в которой f достигает своего максимума M .

Можно рассуждать иначе. Сохраняя предыдущие обозначения, заметим, что при любом n существует хотя бы одна точка x_n из E , такая, что $M_n \leq f(x_n) \leq M$. Последовательность точек x_n имеет хотя бы одну точку сгущения a (теорема Больцано — Вейерштрасса), а следовательно, из нее можно извлечь подпоследовательность x_{p_0}, x_{p_1}, \dots , сходящуюся к a (теорема 24). Так как f непрерывна в a , то последовательность $f(x_{p_n})$ сходится к $f(a)$ (теорема 16). Поскольку $M_{p_n} \leq f(x_{p_n}) \leq M$, то $f(x_{p_n})$ сходится к M , а значит, $f(a) = M$, где M — максимум функции f .

Следствие 1. Вещественная (т. е. со значениями в \mathbb{R}) функция f , непрерывная на компакте E , ограничена. Если во всех точках x из E функция $f(x) > 0$, то существует такое число $\delta > 0$, что $f(x) \geq \delta$ для всех x из E .

В самом деле, можно считать, что значения f принадлежат $\bar{\mathbb{R}}$. Ее точная верхняя грань M , точнее максимум, достигается в некоторой точке a из E : $M = f(a) < +\infty$. Таким образом, f ограничена сверху. Точно так же она ограничена снизу. Если f повсюду > 0 , то ее минимум δ достигается в некоторой точке b и, следовательно, $\delta > 0$. Итак, $f(x) \geq f(b) = \delta$ при любом x из E .

Следствие 2. Пусть f — вещественная непрерывная функция, определенная на метрическом пространстве E , и пусть K — компакт в E . Тогда существует такая окрестность компакта K , на которой f ограничена. Если $f(x) > 0$ для всех x из K , то существуют такое число $\delta > 0$ и такая окрестность K , на которой $f(x) \geq \delta$.

В самом деле, согласно следствию 1, $f(K)$ является ограниченным подмножеством \mathbb{R} . Поэтому существует такое число M , что $f(K) \subset]-M, +M[$. Тогда $K \subset f^{-1}(]-M, +M[)$. Так как f непрерывна и интервал $]-M, +M[$ открыт, то множество $f^{-1}(]-M, +M[)$ открыто. Это множество содержит K , и f на нем ограничена по модулю числом M .

Пусть теперь $f(x) > 0$ для всех x из K . Из следствия 1 вытекает, что существует такое число $\delta' > 0$, что $f(x) \geq \delta'$ для всех $x \in K$. Положим $\delta = \delta'/2$. Тогда $f(K) \subset]\delta, +\infty[$, а, значит, $K \subset f^{-1}(]\delta, +\infty[)$, где $f^{-1}(]\delta, +\infty[)$ — открытое множество, содержащее K , на котором $f(x) \geq \delta$.

Замечание. Эти результаты, естественно, не верны для разрывных функций. Разрывная функция может быть всюду конечной и тем не менее не быть ограниченной; такой будет, например, функция $1/x$, определенная на дополнении к началу координат в компактном интервале $[0, 1]$ и принимающая значение 0 в начале координат. Это вещественная функция, определенная в компактном пространстве, но имеющая точку разрыва. Она повсюду конечна, но не является ограниченной. Впрочем, указанные результаты будут не верны и для непрерывной функции, но определенной на некомпактном пространстве. Например, на вещественной прямой функция x непрерывна, но не ограничена, функция $\operatorname{arc tg} x$ непрерывна и ограничена, но не достигает своей точной верхней грани, т. е. не имеет максимума. Конечно, точка максимума может быть не единственной, как показывает пример постоянной функции.

Приложение. Пусть E — метрическое пространство, F — замкнутая часть E , a — некоторая точка. Расстоянием $d(a, F)$

от точки a до множества F называется точная нижняя грань расстояний от a до точек множества F . Поскольку все эти расстояния ≥ 0 , то расстояние от a до F само ≥ 0 .

Если $a \in F$, то, очевидно, $d(a, F) = 0$. Если же $a \notin F$, то $d(a, F) > 0$. В самом деле, если бы это было не так, то любой шар с центром в a содержал хотя бы одну точку F , а, следовательно, точка a была бы точкой прикосновения F и, в силу замкнутости F , принадлежала бы этому множеству, что противоречит исходному предположению.

Покажем, что $a \rightarrow d(a, F)$ является непрерывным отображением E в \mathbb{R} , т. е. расстояние от некоторой точки до замкнутого множества меняется непрерывно с изменением точки.

Действительно, пусть a и a' — две точки. Согласно свойству точной нижней грани (теорема 2 гл. I), каково бы ни было $\delta > 0$, существует точка x из F , такая, что $d(a, x) \leq d(a, F) + \delta$. Так как $d(a', x) \leq d(a, x) + d(a, a') \leq d(a, F) + d(a, a') + \delta$, то $d(a', F) \leq d(a', x) \leq d(a, F) + d(a, a') + \delta$. Поскольку δ произвольно, $d(a', F) \leq d(a, F) + d(a, a')$. Меняя ролями a и a' , точно так же получаем, что $d(a, F) \leq d(a', F) + d(a, a')$, откуда окончательно вытекает неравенство

$$|d(a', F) - d(a, F)| \leq d(a, a'). \quad (\text{II}, 8; 1)$$

Из $d(a, a') \leq \varepsilon$ следует $|d(a', F) - d(a, F)| \leq \varepsilon$, что доказывает непрерывность отображения $a \rightarrow d(a, F)$.

Можно поставить вопрос: будет ли точная нижняя грань $d = d(a, F) = \inf_{x \in F} d(a, x)$ при заданной точке a минимумом; другими словами, существует ли точка $c \in F$, такая, что $d(a, c) = d$? Покажем, что это действительно так, если все замкнутые шары в F компактны (согласно теореме 23 этот случай имеет место тогда, когда E — конечномерное нормированное векторное пространство).

Рассмотрим шар B радиуса $d + 1$ с центром в a . По предположению этот шар замкнут и компактен. Его пересечение с замкнутым множеством F , являясь замкнутой частью компактного шара B (теорема 5), компактно (теорема 22). Отсюда следует, что непрерывная на компакте $B \cap F$ функция $x \rightarrow d(a, x)$ достигает своего минимума. Точка c , в которой достигается этот минимум, является решением поставленного вопроса. В самом деле, $d(a, c) \leq d(a, x)$ для любой точки $x \in B \cap F$. Кроме того, $d(a, c) \leq d(a, x)$ для любой другой точки x из F , поскольку все другие точки F находятся на расстоянии $\geq d + 1$ от точки a , и $d(a, c) \leq d + 1$.

Из неравенства $d(a, c) \leq d(a, x)$, имеющего место для всех $x \in F$, следует, что $d(a, c)$ действительно является минимумом функции $d(a, x)$ при $x \in F$ (и $d(a, c) = d$).

Пусть теперь E — некоторое метрическое пространство, а F_1 и F_2 — замкнутые части E . Расстоянием $d(F_1, F_2)$ между множествами F_1 и F_2 назовем точную нижнюю грань расстояний $d(x_1, x_2)$ при $x_1 \in F_1$, $x_2 \in F_2$. Если даже $F_1 \cap F_2 = \emptyset$, может случиться, что $d(F_1, F_2) = 0$, как видно из примера гиперболы и ее асимптоты в евклидовой плоскости. Этот пример также показывает, что точная нижняя грань не является минимумом.

Покажем, что если F_1 компактно и если $F_1 \cap F_2 = \emptyset$, то $d = d(F_1, F_2) > 0$. Кроме того, если замкнутые шары из E компактны, то эта точная нижняя грань является также и минимумом.

В самом деле, $d = \inf_{x_1 \in F_1} d(x_1, F_2)$. Мы знаем, что функция $x_1 \rightarrow d(x_1, F_2)$ непрерывна. Так как на компакте F_1 она всегда > 0 , то она имеет минимум, т. е. существует точка $c_1 \in F_1$, такая, что $d = d(c_1, F_2) > 0$. Предположим теперь, что все замкнутые шары из E компактны. Мы видели ранее, что в этом случае существует точка $c_2 \in F_2$, такая, что $d(c_1, c_2) = d(c_1, F_2)$, а, значит, $d(c_1, c_2) = d = d(F_1, F_2)$ действительно является минимумом.

Часто оказывается полезным такое утверждение: если Ω — открытое множество метрического пространства E и K — компакт, содержащийся в Ω , то $d = d(K, \partial\Omega) > 0$. Этот факт вытекает из того, что K является компактом, не имеющим общих точек с замкнутым множеством $\partial\Omega$.

Теорема 30 (Даламбера). Каждый полином комплексной переменной степени m с комплексными коэффициентами имеет m комплексных корней.

Доказательство. Когда мы говорим m корней, то предполагаем, естественно, что каждый корень считается столько раз, какова его кратность.

Для доказательства этой теоремы достаточно показать, что каждый полином при $m \geq 1$ имеет по крайней мере один корень a . Действительно, если мы затем разделим данный полином на $z - a$, то сведем его к полиному степени $m + 1$, к которому снова можно применить то же самое утверждение. Тогда теорема будет доказана индукцией по степени полинома. Итак, предположим, что полином

$$P(z) = a_0 + a_1 z + \dots + a_m z^m \quad (\text{II}, 8; 2)$$

не имеет корней, и попытаемся прийти к противоречию. Известно, что $|P(z)|$ стремится к $+\infty$ при $|z|$, стремящемся к $+\infty$. Но тогда существует число R , такое, что вне окружности радиуса R с центром в 0 комплексной плоскости выполняется неравенство $|P(z)| \geq |P(0)|$. Обозначим через μ минимум > 0 (следствие из теоремы 29) непрерывной функции $|P|$

на компакте $|z| \leq R$, а через z_0 такую точку, в которой $|P(z_0)| = \mu$. Поскольку $|P(z)| \geq |P(0)| \geq \mu$ для $|z| \geq R$, то неравенство $|P(z)| \geq \mu$ справедливо для всех z . Значит, μ является минимумом модуля полинома на всей комплексной плоскости.

Рассмотрим разложение Тейлора полинома P в точке z_0 :

$$P(z) = P(z_0) + c_k(z - z_0)^k + c_{k+1}(z - z_0)^{k+1} + \dots + c_m(z - z_0)^m. \quad (\text{II}, 8; 3)$$

Заметим, что первый член этого разложения отличен от нуля: $|P(z_0)| = \mu$. Далее, существует настолько малое число ρ , что на окружности Γ с центром в точке z_0 радиуса ρ имеет место неравенство

$$|c_{k+1}(z - z_0)^{k+1} + \dots + c_m(z - z_0)^m| < |c_k(z - z_0)^k| = |c_k|\rho^k. \quad (\text{II}, 8; 4)$$

Можно считать ρ столь малым, что $|c_k|\rho^k < \mu$.

Если z пробегает окружность Γ , точка $c_k(z - z_0)^k$ пробегает окружность с центром в начале радиуса $|c_k|\rho^k$ и, значит, $P(z_0) + c_k(z - z_0)^k$ пробегает всю окружность с центром в $P(z_0)$ радиуса $|c_k|\rho^k$. Следовательно, существует точка $z_1 \in \Gamma$, такая, что $P(z_0) + c_k(z_1 - z_0)^k$ находится на отрезке $[0, P(z_0)]$ комплексной плоскости. Имеем:

$$|P(z_0) + c_k(z_1 - z_0)^k| = \mu - |c_k|\rho^k. \quad (\text{II}, 8; 5)$$

В результате мы получаем оценку

$$\begin{aligned} |P(z_1)| &\leq |P(z_0) + c_k(z_1 - z_0)^k| + \\ &+ |c_{k+1}(z_1 - z_0)^{k+1} + \dots + c_m(z_1 - z_0)^m| < \\ &< (\mu - |c_k|\rho^k) + |c_k|\rho^k = \mu. \end{aligned} \quad (\text{II}, 8; 6)$$

Полученное неравенство $|P(z_1)| < \mu$ противоречит тому, что μ является минимумом модуля полинома.

Обобщение теоремы 29. Пусть f — функция на топологическом пространстве E со значениями в $\bar{\mathbb{R}}$. Говорят, что эта функция *полунепрерывна сверху в точке* $a \in E$, если, каково бы ни было $b_1 > f(a)$, в пространстве E существует такая окрестность \mathcal{V}' точки a , что из $x \in \mathcal{V}'$ следует $f(x) \leq b_1$. Говорят, что она *полунепрерывна снизу*, если, каково бы ни было $b_2 < f(a)$, в пространстве E существует такая окрестность \mathcal{V}'' точки a , что из $x \in \mathcal{V}''$ следует $f(x) \geq b_2$.

Вещественная функция непрерывна тогда и только тогда, когда она одновременно полунепрерывна сверху и снизу. Теорему 29 теперь можно обобщить следующим образом:

Теорема 30₂. *Всякая функция, полуунепрерывная сверху (снизу) на компактном пространстве, достигает на нем максимума (минимума).*

Доказательство. Пусть M — точная верхняя грань в $\bar{\mathbb{R}}$ функции f , которая, по предположению, полуунепрерывна сверху на компакте E . Предположим, что она не является максимумом. Тогда $f(x) < M$ для любой точки x из E . В силу полуунепрерывности сверху функции f на E , для любого $0 < \varepsilon_x < M - f(x)$ можно найти такую открытую окрестность \mathcal{U}_x точки x , в которой функция мажорируется числом $M_x = f(x) + \varepsilon_x < M$. При всевозможных значениях x получаем систему открытых окрестностей \mathcal{U}_x , покрывающих E . Поскольку пространство E компактно, из этого покрытия можно выбрать конечное покрытие. Другими словами, существует конечное множество точек x_1, x_2, \dots, x_n из E , такое, что E будет объединением множеств $\mathcal{U}_{x_1}, \mathcal{U}_{x_2}, \dots, \mathcal{U}_{x_n}$. Поскольку в каждом множестве \mathcal{U}_{x_i} функция f мажорируется числом $M_{x_i} < M$, то во всем пространстве E она мажорируется максимальным из чисел $M_{x_1}, \dots, M_{x_2}, \dots, M_{x_n}$, которое само $< M$. Однако это противоречит тому, что M является точной верхней гранью f . Мы видим, что M действительно является максимумом функции f .

Равномерная непрерывность

Свойства, которые мы сейчас рассмотрим, относятся только к метрическим пространствам и не могут быть распространены на общие топологические пространства. Впрочем, даже две эквивалентные метрики могут привести к различным результатам¹⁾.

Определение. Пусть E и F — два метрических пространства. Говорят, что отображение f пространства E в пространство F *равномерно непрерывно*, если, каково бы ни было $\varepsilon > 0$, существует $\eta > 0$, такое, что из неравенства $d(x', x'') \leq \eta$ следует $d(f(x'), f(x'')) \leq \varepsilon$.

Короче,

$$(\forall \varepsilon > 0) (\exists \eta > 0) (\forall x' \in E, \forall x'' \in E, d(x', x'') \leq \eta) :$$

$$d(f(x'), f(x'')) \leq \varepsilon. \quad (II, 8; 7)$$

Каждая равномерно непрерывная функция, очевидно, непрерыв-

¹⁾ Другими словами, отображение метрического пространства E в метрическое пространство F может быть равномерно непрерывным и перестать быть таковым, когда метрика заменяется другой эквивалентной метрикой. Тем не менее, если эти метрики определены нормами в векторных пространствах, то в силу теоремы 12 равномерная непрерывность сохранится при замене норм на эквивалентные.

на, но обратное не верно. Если функция f всюду непрерывна, то для любой точки $a \in E$ при любом $\varepsilon > 0$ существует $\eta > 0$, такое, что $d(x, a) \leq \eta$ влечет за собой $d(f(x), f(a)) \leq \varepsilon$; однако так определенное число η зависит одновременно и от a , и от ε . Когда же говорят, что функция равномерно непрерывна, то это означает, что η можно выбрать в зависимости только от ε (см. по этому поводу гл. I, стр. 38).

На вещественной прямой функция x равномерно непрерывна, однако функция x^2 этим свойством не обладает. В самом деле, $|x+h|^2 - x^2 = |2hx + h^2| \geq 2hx$ при $h \geq 0$ и $x \geq 0$. Для того чтобы при заданном ε найти h так, чтобы $|2hx + h^2| \leq \varepsilon$, необходимо взять $|h| \leq \varepsilon/2|x|$. Выбрать при этом условии h независимо от x невозможно.

Говорят, что отображение f метрического пространства E в метрическое пространство F удовлетворяет *условию Гельдера* или *условию Липшица порядка α* , $0 < \alpha \leq 1$, если существует такая постоянная k , что при любых x' и x'' из E

$$d(f(x'), f(x'')) \leq k [d(x', x'')]^\alpha. \quad (\text{II}, 8; 8)$$

При $\alpha = 1$ говорят просто, что f удовлетворяет *условию Липшица* (не уточняя, что ее порядок равен 1).

Функция, удовлетворяющая условию Гельдера порядка α и тем более условию Липшица, равномерно непрерывна. Обратное утверждение не верно, как показывает пример функции, равной $1/\ln x$ на интервале $[0, 1/2]$ и нулю в точке $x=0$. Она непрерывна, а, следовательно, как будет показано в теореме 31, равномерно непрерывна. Однако $|f(x) - f(0)| = 1/|\ln x|$ не может быть мажорируема выражением вида $k|x|^\alpha$.

Теорема 31. *Всякое непрерывное отображение компактного метрического пространства E в метрическое пространство F равномерно непрерывно.*

Доказательство. Предположим, что отображение f не равномерно непрерывно, и покажем, что это приведет к противоречию. То, что f не является равномерно непрерывным, означает, что

$$(\exists \varepsilon > 0)(\forall \eta > 0)(\exists x' \in E, \exists x'' \in E, d(x', x'') \leq \eta) : d(f(x'), f(x'')) > \varepsilon.$$

Пусть ε — число, стоящее в начале этого соотношения. Тогда, каково бы ни было целое число $n \geq 1$, можно найти две точки x'_n, x''_n , такие, что $d(x'_n, x''_n) \leq 1/n$ и $d(f(x'_n), f(x''_n)) > \varepsilon$.

¹⁾ Случай $\alpha > 1$ не представляет интереса. Например, для $E=F=\mathbb{R}$ при $\alpha > 1$ из $\left| \frac{f(x+h) - f(x)}{h} \right| \leq k|h|^{\alpha-1}$ следует, что функция f всюду имеет производную, равную нулю, а это означает, что эта функция постоянна.

Поскольку E компактно, из последовательности x'_n можно выделить подпоследовательность $n \rightarrow x'_{p_n}$, сходящуюся к некоторой точке x (теоремы 24 и 25). Из соотношения $d(x'_n, x''_n) \leq 1/n$ следует, что подпоследовательность x''_{p_n} также сходится к x . Поскольку отображение f непрерывно, то последовательности $f(x'_{p_n})$ и $f(x''_{p_n})$ заведомо сходятся к $f(x)$ (в силу теоремы 16). Тогда для достаточно больших n имеем $d(f(x_{p_n}), f(x)) \leq \varepsilon/2$, $d(f(x''_{p_n}), f(x)) \leq \varepsilon/2$, а, следовательно, $d(f(x'_{p_n}), f(x''_{p_n})) \leq \varepsilon$, что противоречит условию построения точек x'_{p_n} и x''_{p_n} .

Как мы только что видели на примере функции на вещественной прямой \mathbb{R} , функция может быть непрерывной на метрическом некомпактном пространстве и не быть равномерно непрерывной на нем.

§ 9. СВЯЗНЫЕ ПРОСТРАНСТВА

Интуитивно ясно, что некоторые топологические пространства можно рассматривать как нечто целое, как, например, сферу, шар в пространстве \mathbb{R}^n , в то время как другие пространства состоят из нескольких различных «кусков». Таким будет, например, пространство, являющееся объединением двух сфер, не имеющих общих точек, или дополнение к некоторой сфере в \mathbb{R}^n . Уточним это интуитивное понятие.

Определение. Топологическое пространство называется *связным*, если его невозможно разбить на две непересекающиеся открытые части, или иначе: если его невозможно разбить на две непересекающиеся замкнутые части, или иначе: если в E не существует одновременно открытых и замкнутых подмножеств, кроме самого пространства E и пустого множества \emptyset .

Эквивалентность этих трех определений следует из определения замкнутых множеств как дополнений к открытым. Как и компактность, связность является внутренним свойством самого топологического пространства; однако, говорят, что F является связной частью E , если F , рассматриваемое как пространство с индуцированной топологией, связно.

Теорема 32. Для того чтобы часть E пополненной вещественной прямой $\bar{\mathbb{R}}$ была связным топологическим пространством, необходимо и достаточно, чтобы она была открытым, полуоткрытым или замкнутым интервалом.

Доказательство. Пусть E — связная часть $\bar{\mathbb{R}}$, а x и y — две любые различные точки E . Покажем, что весь замкну-

тый интервал $[x, y]$ содержится в E . Если бы это было не так, то на этом интервале можно было бы найти хотя бы одну точку z , не принадлежащую E . Тогда открытые в $\bar{\mathbb{R}}$ множества $[-\infty, z[$ и $]z, +\infty]$, пересекаясь с E , делили бы E на два открытых непересекающихся (поскольку $z \notin E$) множества, а, значит, E не было бы связным. Если мы теперь обозначим через a (соответственно через b) точную верхнюю (соответственно нижнюю) грань E , то, как мы только что видели, E будет совпадать с одним из четырех интервалов $[a, b]$, $[a, b[$, $]a, b]$, $]a, b[$.

Обратно, пусть E — некоторый интервал в $\bar{\mathbb{R}}$. Пусть A — некоторое непустое подмножество E , одновременно открытое и замкнутое в E . Для доказательства связности E достаточно показать, что $A = E$. Пусть c — некоторый элемент A . Рассмотрим множество H всех $x \in E$, таких, что $[c, x] \subset A$, и пусть γ — точная верхняя грань H в $\bar{\mathbb{R}}$. Для каждого γ' , такого, что $c \leq \gamma' < \gamma$, существует элемент H , мажорирующий γ' . Следовательно, $\gamma' \in A$ и $[c, \gamma'] \subset A$. Поскольку A замкнуто, отсюда вытекает, что $\gamma \in A$ и $[c, \gamma] \subset A$, кроме того случая, когда $\gamma = b$ и $b \notin E$. Если бы элемент γ был $< b$, то в силу того, что A открыто, существовал бы элемент $\gamma'' > \gamma$, такой, что $[\gamma, \gamma''] \subset A$, а, следовательно, $[c, \gamma''] \subset A$ и $\gamma'' \in H$, а это невозможно, так как γ является точной верхней гранью H . Поэтому $\gamma = b$, $[c, b] \subset A$ и, кроме того, $b \in A$, если $b \in E$. Проводя слева от точки c рассуждения, аналогичные только что проведенным справа от этой точки, мы получаем точно так же, что $]a, c] \subset A$ и $a \in A$, если $a \in E$. Окончательно получаем, что $A = E$ и, следовательно, E связно.

Следствие. Множество \mathbb{Q} рациональных чисел не связно.

Теорема 33. Образ связного топологического пространства при непрерывном отображении является связным.

Доказательство. Пусть f — непрерывное отображение связного пространства E в F . Речь, естественно, идет не о том, чтобы доказать связность F ; мы докажем, что образ $f(E)$ пространства E в F является связным. В самом деле, если бы это было не так, то $f(E)$ можно было бы разбить на две открытые непересекающиеся части A и B . Тогда их прообразы образуют разбиение E (согласно (I, 2; 3)), причем их пересечение пусто. Если $x \in E$, то $f(x)$ принадлежит либо A , либо B , т. е. x принадлежит либо $f^{-1}(A)$, либо $f^{-1}(B)$, а, значит, E является объединением $f^{-1}(A)$ и $f^{-1}(B)$. Ни одно из этих множеств не пусто. В самом деле, A , например, не пусто и содержится в $f(E)$, а значит, его прообраз не пуст и открыт (согласно теореме 7), что противоречит предположению о связности E .

Следствие. Если f — непрерывная функция, определенная на связном пространстве E , со значениями в $\bar{\mathbb{R}}$, то множество ее значений является открытым, полуоткрытым или замкнутым интервалом в $\bar{\mathbb{R}}$.

Доказательство. В самом деле, это множество должно быть связным подмножеством $\bar{\mathbb{R}}$ и остается лишь применить предыдущую теорему.

Это свойство часто формулируют для частного случая, когда E само является открытым, полуоткрытым или замкнутым интервалом $\bar{\mathbb{R}}$. Наконец, часто просто говорят, что вещественная непрерывная функция, заданная на связном пространстве, вместе с любыми двумя значениями принимает и все промежуточные. Это так называемое *свойство промежуточных значений*. Уместно заметить, что оно не является характеристическим свойством непрерывных функций. В связных пространствах существуют разрывные функции, обладающие тем же самым свойством. Например, функция, заданная формулами $f(x) = \sin 1/x$ для $x \neq 0$ и $f(0) = 0$ и определенная на всей вещественной прямой, разрывна в точке $x = 0$ и все же обладает тем свойством, что множество ее значений образует замкнутый интервал $[-1, 1]$; в любом интервале $[a, b]$ она вместе с двумя значениями принимает и все промежуточные.

Однако можно привести следующую обратную теорему:

Теорема 34. Если топологическое пространство E таково, что любая определенная на нем вещественная непрерывная функция вместе с любыми своими двумя значениями принимает также и все промежуточные значения, то E связно.

Доказательство. Пусть E не связно. Тогда его можно разбить на две открытые части A и B , не имеющие общих точек. Вещественная функция, равная 0 на A и 1 на B , будет непрерывной, поскольку прообраз любого открытого в $\bar{\mathbb{R}}$ множества является одним из четырех открытых множеств \emptyset, A, B, E . Эта функция принимает значения 0 и 1, не принимая никаких промежуточных значений, что противоречит условию теоремы. Следовательно, E связно.

Линейно связные пространства

Полезно иметь некоторый критерий, позволяющий определять связность пространства. В связи с этим введем понятие линейно связного пространства.

Дугой, или путем, соединяющим точку a и точку b топологического пространства E , называется такое непрерывное отображение f интервала $[\alpha, \beta]$ вещественной прямой $\bar{\mathbb{R}}$ в простран-

ство E , что $f(\alpha) = a$ и $f(\beta) = b$. Точки a и b называют при этом началом и концом пути.

З Не следует смешивать отображение, входящее в определение пути, с образом интервала $[\alpha, \beta]$, который при этом отображении называется образом пути.

Если, например, отображение постоянно, то этот образ сводится к точке. Можно, впрочем, сказать в этом случае, что путь сводится к точке, но путь — это нечто иное, чем множество точек, полученных при отображении.

Точно так же лемнискату Бернулли можно «обойти» в двух различных направлениях, тогда как рассматриваемая как множество она остается одной и той же в обоих случаях. Два способа обхода соответствуют двум различным путям, т. е. двум различным отображениям интервала из \mathbb{R} на плоскость:

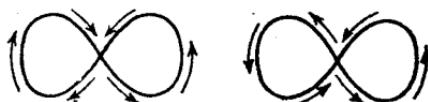


Рис. 2.

Говорят, что путь f проходит через точку c из E , если образ $f([\alpha, \beta])$ содержит c . Говорят также, что путь пересекает часть A множества E , если $f([\alpha, \beta]) \cap A \neq \emptyset$.

Непосредственно видно, что если две точки a и b могут быть соединены некоторым путем и в свою очередь точки b и c могут быть тоже соединены некоторым путем, то точки a и c также могут быть соединены путем.

Теорема 35. Если любые две точки топологического пространства E могут быть соединены некоторым путем, то E связано.

Доказательство. Предположим, что это не так. Тогда E можно разбить на два непересекающихся открытых множества A и B . Пусть a и b — элементы множеств A и B соответственно. По предположению, точки a и b можно соединить некоторым путем. Пусть K — образ этого пути. Тогда $K \cap A$ и $K \cap B$ будут открытыми дополнительными подмножествами K (теорема 5), и ни одно из них не пусто, поскольку a и b принадлежат соответственно первому и второму множествам. Мы получим разбиение K на два открытых непересекающихся множества, что противоречит теореме 33, в которой утверждается, что образ связного пространства $[\alpha, \beta]$ при непрерывном отображении связан. Противоречие доказывает связность множества E .

Обратное утверждение к этой теореме, вообще говоря, не верно. Если, например, мы рассмотрим в евклидовой плоскости \mathbb{R}^2 множество E , являющееся объединением кривой, опре-

деляемой функцией $y = \sin 1/x$, $x \neq 0$, и вертикального отрезка $x=0$, $|y| \leqslant 1$, то легко проверить, что E связно, и тем не менее точки $(0, 0)$ и $(a, \sin 1/a)$ соединить путем невозможно.

Говорят, что пространство *линейно связно*, если две его произвольные точки могут быть соединены некоторым путем. Это свойство более сильное, чем обычная связность.

Пример. Любое нормированное векторное пространство линейно связно. В самом деле, любые две его точки могут быть соединены отрезком с началом и концом в рассматриваемых точках.

Точно так же проверяется, что шар и сфера в нормированных векторных пространствах линейно связны.

Теорема 36 (о переходе через таможню). Пусть A — часть топологического пространства E . Любой путь, соединяющий внутреннюю точку A с точкой ее внешней части, пересекает границу A .

Доказательство. Поскольку образ K , согласно теореме 32, является связным множеством, то эта теорема является частным случаем следующего утверждения:

Любое связное подмножество $B \subset E$, имеющее общие точки с внутренней и внешней частями A , *заведомо пересекает его границу*. Это свойство очевидно. В самом деле, если бы это было не так, то B содержалось бы в объединении внутренней и внешней частей A , пересечение B с обеими этими частями определяло бы в этом случае разбиение B на два непересекающихся открытых множества, что невозможно.

§ 10. ДОПОЛНЕНИЕ ПО ОБЩЕЙ ТОПОЛОГИИ СВЯЗНЫХ ПРОСТРАНСТВ

Теорема 36₂. Объединение A любого семейства (A_i) , $i \in I$, связных частей топологического пространства E , имеющих попарно непустые пересечения, связно.

Доказательство. Допустим, что это не так. Тогда A можно рассматривать как объединение двух непересекающихся и открытых относительно A частей B' и B'' . Для каждого подмножества A_i множества $A_i \cap B'$ и $A_i \cap B''$ не пересекаются, являются открытыми относительно A_i и в объединении дают A_i . Поскольку A_i по условию связны, одна из этих двух частей необходимо пуста, а другая совпадает с A_i . Таким образом, A_i целиком содержится в B' или в B'' . Поскольку пересечение любых двух частей A_i и A_j не пусто, то обе они заведомо содержатся либо в B' , либо в B'' , так что, окончательно, все части A_i содержатся в B' или же все лежат в B'' , а значит, $A \subset B'$ или

$A \subset B''$. Это противоречит предположению, согласно которому A является объединением непустых непересекающихся частей B' и B'' . Полученное противоречие доказывает связность множества A .

Теорема 36₃. Замыкание \bar{A} связной части A в топологическом пространстве E связано.

Доказательство. Пусть это не так, т. е. пусть \bar{A} является объединением двух непустых непересекающихся частей B' и B'' , каждая из которых замкнута относительно \bar{A} . Поскольку \bar{A} замкнуто в E , то они замкнуты относительно E (теорема 6). Мы получаем, что A является объединением двух непересекающихся частей $A \cap B'$ и $A \cap B''$, замкнутых относительно A . Поскольку часть A предполагалась связной, то одно из этих пересечений должно быть пустым, так что мы имеем, например, $A \cap B' = \emptyset$ и $A = A \cap B'' : A \subset B''$. Отсюда, в силу замкнутости B в E , получаем $\bar{A} \subset B''$, а это противоречит предположению о том, что \bar{A} является объединением непересекающихся непустых частей B' и B'' . Полученное противоречие доказывает связность замыкания \bar{A} .

Определение. Две точки топологического пространства E называются *связанными*, если существует связное подмножество E , содержащее эти точки. Две точки, которые можно соединить некоторым путем, связаны.

Теорема 36₄. Отношение « x и y связаны в E » является отношением эквивалентности в E .

Доказательство. Это отношение, очевидно, рефлексивно и симметрично; поэтому достаточно доказать его транзитивность. Итак, пусть x и y связаны. Если, кроме того, связаны точки y и z , то существует, с одной стороны, связное подмножество, содержащее x и y , и, с другой стороны, связное подмножество, содержащее y и z . Так как оба подмножества связаны и их пересечение не пусто, то по теореме 36₂ их объединение связано. Так как точки x и z содержатся в общем связном подмножестве, то они связаны.

Определение. Класс эквивалентности E по отношению эквивалентности « x и y связаны в E » называется *связной компонентой* E . Пространство E является объединением связных попарно не пересекающихся компонент. *Связной компонентой* точки x из E называется содержащая ее связная компонента E .

Теорема 36₄. Связная компонента точки x из E совпадает с множеством точек E , связанных с x , и является наибольшей связной частью E , содержащей x . Связные компоненты E замкнуты.

Доказательство. 1°) По определению классов эквивалентности, связная компонента E_x точки x в E является множеством точек E , связанных с x . Каждая связная часть E , содержащая x , полностью лежит в E_x , поскольку все ее точки связаны с x . Обратно, если точка $y \in E_x$, то она связана с x , а, следовательно, содержится хотя бы в одной связной части E , содержащей x . Таким образом, E_x является объединением всех связных частей E , содержащих x . Так как все эти связные части содержат x , то попарное пересечение их не пусто. Из теоремы 36₂ вытекает, что это объединение E_x связно. Это — связная часть E , содержащая x , а также любую связную часть, содержащую x . Таким образом, E_x является наибольшей связной частью E , содержащей x .

2°) Согласно теореме 36₃, замыкание \bar{E}_x множества E_x в E связно. Поскольку \bar{E}_x является наибольшей связной частью E , содержащей x , то $\bar{E}_x = E_x$ и, следовательно, E_x замкнута.

Определение. Пространство E называется *локально связным*, если, какова бы ни была точка a из E и окрестность \mathcal{U} точки a в E , существует связная окрестность \mathcal{U}' точки a , содержащаяся в \mathcal{U} . Как показывает сам выбранный нами термин, свойство некоторого пространства быть локально связным является *локальным свойством*, в то время как обычная связность носит глобальный характер. Эти два свойства совершенно не связаны друг с другом¹⁾.

1°) Рассмотрим, например, в плоскости \mathbb{R}^2 множество E , образованное двумя прямыми, определяемыми уравнениями $y=0$ и $y=1$. Это множество, как подмножество \mathbb{R}^2 , *не связно*. Однако оно локально связно, так как любая окрестность \mathcal{U} точки a из E содержит некоторый горизонтальный интервал с центром в точке a , являющийся связной частью. Этот пример показывает, что пространство может быть локально связным, не будучи связным.

2°) Обозначим через E множество точек плоскости \mathbb{R}^2 , состоящее из всех прямых, параллельных оси Ox и имеющих *рациональные* ординаты, и *всей* оси Oy . Это топологическое пространство не является локально связным: если рассматривается произвольная точка E с координатами (a, b) , где $a \neq 0$, и некоторый шар с центром в этой точке радиуса $< |a|$, то эта окрестность не содержит никакой связной окрестности. Однако множество E связно и даже линейно связно. Любые две его точки можно соединить путем, составленным из трех отрезков: первый и третий отрезки параллельны оси Ox , а второй отрезок

¹⁾ Не следует проводить аналогий с понятием компактности: любое компактное пространство является и локально компактным (обратное не верно), тогда как связное пространство не обязано быть локально связным.

расположен на оси Oy . Поэтому рассмотренное пространство связно, но не локально связно.

Теорема 36₆. *Если топологическое пространство E локально связно, то каждая его связная компонента одновременно открыта и замкнута в E .*

Доказательство. В самом деле, пусть E_x — связная компонента x в E и y — некоторая точка E_x . Так как y обладает некоторой связной окрестностью, то все точки z этой окрестности связаны с y , а, следовательно, и с точкой x . Отсюда следует, что E_x содержит всю эту окрестность. Итак, если E_x содержит точку y , то она содержит и некоторую ее окрестность. Значит, E_x открыто. Согласно теореме 36₅, E_x одновременно и замкнуто.

Замечания. 1°) Обратное утверждение, вообще говоря, не верно. Так в приведенном выше примере, где E связно, но не локально связно, имеется лишь одна связная компонента E . Эта компонента одновременно открыта и замкнута, и тем не менее E не локально связно.

2°) *Если E является открытым множеством нормированного векторного пространства E_0 , то E локально связно.* В самом деле, каждая окрестность точки a из E заведомо содержит некоторый шар с центром a в E_0 , который, как известно, является связным, и даже линейно связным. В этом случае применима сформулированная теорема. Связные компоненты E заведомо одновременно открыты и замкнуты в E . Они тогда также открыты и в E_0 .

3°) Отметим еще один замечательный случай, когда все связные компоненты одновременно открыты и замкнуты: это случай, когда E имеет только *конечное* число связных компонент. Каждая из этих компонент, будучи уже замкнутой, является дополнением к объединению других компонент. Далее, так как конечное объединение замкнутых подмножеств само замкнуто, то рассматриваемые компоненты открыты.

Определение. Топологическое пространство называется *локально линейно связным*, если, каковы бы ни были точка a и ее окрестность \mathcal{U} , существует линейно связная окрестность \mathcal{U}' точки a , содержащаяся в \mathcal{U} . Поскольку линейно связное пространство связно, локально линейно связное пространство является локально связным. Открытое множество нормированного векторного пространства локально линейно связно.

Теорема 36₇. *Пусть E — локально линейно связное топологическое пространство. Если оно связно, то оно линейно связно.*

Если оно не связно, то каждая его связная компонента открыта, замкнута и линейно связна.

Доказательство. Пусть x — некоторая точка E . Обозначим через F_x множество точек E , которые могут быть соединены с точкой x некоторым путем. Если y — такая точка, то существует окрестность y , все точки z которой могут быть соединены с y некоторыми путями. Поскольку y может быть соединена с x некоторым путем, то z также может быть соединена с x некоторым путем. Другими словами, если F_x содержит точку y , то оно содержит целую ее окрестность. Это означает, что F_x открыто.

Докажем теперь, что множество F_x замкнуто. Для этого достаточно показать, что его дополнение открыто. Предположим, что y принадлежит этому дополнению, т. е. не принадлежит F_x . Существует окрестность \mathcal{U} точки y , все точки которой могут быть соединены с y некоторым путем. Тогда никакая точка z из \mathcal{U} не принадлежит F_x . Действительно, если бы можно было соединить путем точки x и z , то можно было бы соединить точки x и y , что противоречит нашему предположению. Итак, если дополнение к F_x содержит точку y , то оно содержит и некоторую окрестность этой точки. Это означает, что дополнение к F_x открыто, а, значит, само F_x замкнуто.

Пусть E связно. Множество F_x одновременно открыто, замкнуто и не пусто, поскольку оно содержит x . Это означает, что $F_x = E$, и, значит, E линейно связно.

В любом случае, если даже E не связно, для связной компоненты E_x точки x в E мы получаем следующий результат. Поскольку F_x , будучи линейно связным, связно, то $F_x \subset E_x$. Так как E_x связно, а F_x одновременно открыто и замкнуто в E , а, значит, и в E_x , то $F_x = E_x$. В итоге получаем, что F_x открыто, замкнуто и линейно связно.

Замечания. 1°) Две предыдущие теоремы будут применяться, в частности, когда E является некоторым многообразием (см. далее гл. III, § 9). В самом деле, каждая окрестность точки содержит тогда некоторую окрестность, гомеоморфную сфере, и, следовательно, локально связна.

2°) Легко доказываются и другие теоремы этого рода.

Предположим, например, что E является открытым множеством нормированного векторного пространства. Тогда оно локально связно по ломанным в том смысле, что каждая окрестность \mathcal{U} точки a содержит другую окрестность \mathcal{U}' (например, шар с центром в a), две произвольные точки которой могут быть соединены ломаной линией (т. е. путем, последовательно образованным из конечного числа отрезков прямой; для шара достаточно одного отрезка). То же самое доказательство, что

и выше¹⁾, показывает, что если E связно, то оно локально связно по ломанным, и что, даже если E не связно, каждая его связная компонента является связной по ломанным.

Если E является открытым множеством конечномерного нормированного векторного пространства с введенной в нем системой координат, то точно так же доказывается, что любая связная компонента E является связной «по ломанным линиям с отрезками, параллельными осям координат»; другими словами, любые две точки связной компоненты E могут быть соединены путем, последовательно образованным из конечного числа отрезков, параллельных осям координат.

Некоторые применения понятия связности. Критерии негомеоморфности

Так же как и понятие компактности, рассмотренное выше понятие связности позволяет нам установить в некоторых случаях, являются ли два данных топологических пространства гомеоморфными. В самом деле, если два пространства гомеоморфны и одно из них связно, то связным также должно быть и второе. Рассмотрим, например, вещественную прямую \mathbb{R} и плоскость \mathbb{R}^2 . Мы знаем, что эти множества имеют одинаковую мощность, значит, существует некоторая биекция одного из них на другое. Легко видеть также, что существует непрерывное сюръективное отображение плоскости на прямую. Это — проекция $(x, y) \rightarrow x$ плоскости \mathbb{R}^2 на \mathbb{R} . Довольно сложным путем Пеано доказал, что существует непрерывное сюръективное отображение прямой \mathbb{R} на плоскость \mathbb{R}^2 . Но эти отображения, как это ни кажется парадоксальным, не могут быть биективными: всегда существуют точки плоскости, имеющие несколько прообразов! Довольно просто можно доказать, что гомеоморфизма прямой на плоскость не существует. В самом деле, согласно теореме 32, прямая обладает следующим свойством: если из нее удаляют точку, то оставшееся дополнение, рассматриваемое как топологическое пространство, не связно. Напротив, плоскость этим свойством не обладает: дополнение к точке безусловно связно и даже линейно связно²⁾. Отсюда непосредственно вытекает, что прямая и плоскость не могут быть гомеоморфными. Точно так же, несмотря на то что интервал $[a, b]$ прямой \mathbb{R} и окружность плоскости \mathbb{R}^2 компактны, они не гомеоморфны. В самом деле, если из интервала $[a, b]$ выкинуть его середину, то дополнение этой точки не будет связным, в то

¹⁾ То есть нужно рассмотреть множество точек F_x , которые можно соединить с x ломаной линией.

²⁾ Две произвольные точки могут быть соединены путем, образованным из двух прямолинейных отрезков плоскости.

время как удаляя любую точку окружности, мы всегда получаем связное дополнение.

Можно доказать (но это не просто!), что два нормированных векторных пространства разной размерности не могут быть гомеоморфными. Сфера в \mathbb{R}^m и сфера в \mathbb{R}^n при $m \neq n$ не гомеоморфны.

Существование и непрерывность обратной функции для строго монотонной непрерывной функции

Теорема 37. Пусть f — некоторое непрерывное строго возрастающее отображение интервала $[a, b]$ (открытого, полуоткрытого или замкнутого) из $\bar{\mathbb{R}}$ в $\bar{\mathbb{R}}$. Тогда образ $f([a, b])$ является интервалом $[\alpha, \beta]$ того же вида, а f есть гомеоморфизм $[a, b]$ на $[\alpha, \beta]$. Продолжение \hat{f} отображения f , определенное с помощью равенств $\hat{f}(a) = \alpha$, $\hat{f}(b) = \beta$, является гомеоморфизмом $[a, b]$ на $[\alpha, \beta]$.

То, что образ является интервалом $[\alpha, \beta]$, вытекает из следствия теоремы 33. Инъективность f вытекает из того, что это отображение строго возрастающее. Следовательно, f является непрерывной биекцией $[a, b]$ на $[\alpha, \beta]$. Поскольку f сохраняет отношение порядка и $a = \inf([a, b])$, $\alpha = \inf([\alpha, \beta])$, то α принадлежит образу тогда и только тогда, когда интервал $[a, b]$ содержит a . То же самое справедливо для точек b и β . Поэтому интервал $[\alpha, \beta]$ имеет тот же вид (открытый, полуоткрытый или замкнутый), что и интервал $[a, b]$. Образ любого интервала $[c, d]$ из $[a, b]$, открытого в $[a, b]$, является интервалом того же вида, а, следовательно, является открытым множеством в $[\alpha, \beta]$. Образ любого открытого множества из $[a, b]$ при отображении f является открытым множеством из $[\alpha, \beta]$, а, следовательно, f является гомеоморфизмом $[a, b]$ на $[\alpha, \beta]$ (теорема 11). Обратное отображение, будучи, очевидно, строго возрастающим, является также биективным и непрерывным отображением $[\alpha, \beta]$ на $[a, b]$.

Когда x стремится к a строго справа, то $f(x)$ стремится к α (см. стр. 61). В самом деле, каково бы ни было $\alpha' > \alpha$, прообразом интервала $]\alpha, \alpha'[$ является интервал $]f_{(a)}^{-1}, f_{(a')}^{-1}[$, являющийся пересечением некоторой окрестности точки a в $\bar{\mathbb{R}}$ с интервалом $[a, b]$. Следовательно, если положить $\hat{f}(a) = \alpha$ и, точно так же $\hat{f}(b) = \beta$, то \hat{f} будет продолжением f , являющимся строго возрастающим непрерывным отображением, а, значит, гомеоморфизмом $[a, b]$ на $[\alpha, \beta]$. Если $a = -\infty$, $b = +\infty$, а α и β конечны, то \hat{f} определит некоторый гомеоморфизм $\bar{\mathbb{R}}$ на замкнутый ограниченный интервал в $\bar{\mathbb{R}}$.

З а м е ч а н и е. Предположение о строгой монотонности f естественно: непрерывное и инъективное отображение f интервала $[a, b]$ из $\bar{\mathbb{R}}$ в $\bar{\mathbb{R}}$ заведомо строго монотонно. В самом деле, предположим, что это не так. Тогда в $[a, b]$ будут существовать точки x_1, x_2, x_3 , $x_1 < x_2 < x_3$, такие, что $f(x_1) < f(x_2)$ и $f(x_2) > f(x_3)$, или такие, что $f(x_1) > f(x_2)$ и $f(x_2) < f(x_3)$. Рассмотрим, например, первый случай. Пусть y — общая точка интервалов $[f(x_1), f(x_2)]$ и $[f(x_3), f(x_2)]$. Тогда, согласно теореме о промежуточных значениях (следствие из теоремы 33), существуют точка $x' \in [x_1, x_2]$ и точка $x'' \in [x_2, x_3]$, в которых f принимает значение y . Это означает, что f не является инъекцией, а это противоречит исходному условию.

Применение: метрики, определяющие топологию в $\bar{\mathbb{R}}$

Теорема 38. *Пространство $\bar{\mathbb{R}}$ метризуемо. Если f — строго возрастающее непрерывное и ограниченное отображение \mathbb{R} в \mathbb{R} , то метрика*

$$d(x, y) = |f(x) - f(y)|, \quad x \in \bar{\mathbb{R}}, y \in \bar{\mathbb{R}}, \quad (\text{II}, 10; 1)$$

где \hat{f} — продолжение отображения f до отображения $\bar{\mathbb{R}}$ в $\bar{\mathbb{R}}$, определяемое равенствами $\hat{f}(-\infty) = \inf_{x \in \mathbb{R}} f(x)$ и $\hat{f}(+\infty) = \sup_{x \in \mathbb{R}} f(x)$, задает топологию в пространстве $\bar{\mathbb{R}}$.

Согласно теореме 37, f является гомеоморфизмом \mathbb{R} на интервал $[\alpha, \beta]$ из $\bar{\mathbb{R}}$. Так как f ограничено, то $\alpha > -\infty$, $\beta < +\infty$. Кроме того, $\alpha = \inf_{x \in \mathbb{R}} f(x)$, $\beta = \sup_{x \in \mathbb{R}} f(x)$ и продолжение \hat{f} отображения f , такое, что $\hat{f}(-\infty) = \alpha$ и $\hat{f}(+\infty) = \beta$, является гомеоморфизмом $\bar{\mathbb{R}}$ на $[\alpha, \beta] \subset \mathbb{R}$.

Поскольку топология в $[\alpha, \beta] \subset \mathbb{R}$ определяется естественной метрикой прямой \mathbb{R} , то топология в $\bar{\mathbb{R}}$ определяется метрикой, перенесенной от естественной метрики в $[\alpha, \beta]$ отображением \hat{f}^{-1} , т. е. такой метрикой, в которой расстояние между $x \in \bar{\mathbb{R}}$ и $y \in \bar{\mathbb{R}}$ является естественным расстоянием между $\hat{f}(x)$ и $\hat{f}(y)$, т. е. $|\hat{f}(x) - \hat{f}(y)|$, или (II, 10; 1), что и требовалось доказать.

Например, можно взять

$$f(x) = \arctg x, \quad \text{где } \hat{f}(-\infty) = -\frac{\pi}{2}, \quad \hat{f}(+\infty) = \frac{\pi}{2}, \quad (\text{II}, 10; 2)$$

что, по существу, сводится к метрике $d(x, y) = \left| \int_x^y d\xi / (1 + \xi^2) \right|$.

Более общо, если h является непрерывной функцией >0 , интегрируемой на \mathbb{R} , то можно положить

$$d(x, y) = \left| \int_x^y h(\xi) d\xi \right|. \quad (\text{II}, 10; 3)$$

Здесь можно принять

$$f(x) = \operatorname{th} x \quad \text{и} \quad f(-\infty) = -1, \quad f(+\infty) = +1. \quad (\text{II}, 10; 4)$$

Можно положить

$$f(x) = \frac{x}{1+|x|} \quad \text{и} \quad f(-\infty) = -1, \quad f(+\infty) = 1 \quad (\text{II}, 10; 5)$$

и т. д. Все эти весьма различные метрики эквивалентны в \mathbb{R} (а следовательно, в \mathbb{R}), поскольку они определяют одну и ту же топологию.

§ 11. ПОЛНЫЕ МЕТРИЧЕСКИЕ ПРОСТРАНСТВА

Пусть в метрическом пространстве E задана последовательность точек x_0, x_1, x_2, \dots , сходящаяся к x . Тогда эта последовательность обладает следующим свойством¹⁾:

$$(\forall \varepsilon > 0) (\exists p \in \mathbb{N}) (\forall m \in \mathbb{N}, \forall n \in \mathbb{N}, m \geq p, n \geq p) : d(x_m, x_n) \leq \varepsilon \quad (\text{II}, 11; 1)$$

или же

$$d(x_n, x_m) \text{ стремится к } 0, \text{ когда } m \text{ и } n \text{ стремятся к } +\infty. \quad (\text{II}, 11; 2)$$

Последовательность, обладающую этим свойством, называют *последовательностью Коши* (или *фундаментальной последовательностью*).

Из общего курса анализа известно, что если x_0, x_1, x_2, \dots — последовательность Коши в поле вещественных или комплексных чисел, то она сходится.

Это свойство является основным, поскольку оно позволяет выяснить, является ли последовательность сходящейся, даже если мы не знаем предела, к которому она может сходиться; общее определение сходимости не дает такой возможности. Именно это свойство приводит к построению различных критериев сходимости рядов и позволяет решить, сходится ли ряд, составленный из вещественных или комплексных чисел, даже если мы ничего не знаем о его сумме, что очень важно в анализе. Однако это свойство, справедливое в поле вещественных и комплексных чисел, не имеет места в произвольных метрических

¹⁾ Его часто называют *критерием Коши*.

пространствах. Изучению этого вопроса мы и посвятим этот параграф.

Рассмотрим прежде всего некоторые общие свойства последовательностей Коши.

Теорема 39. Любая подпоследовательность последовательности Коши также является последовательностью Коши. Каждая последовательность Коши ограничена.

Доказательство. Первое свойство очевидно. Докажем второе. Ясно, что существует целое p , такое, что из $m \geq p$, $n \geq p$ следует $d(x_m, x_n) \leq 1$. Но тогда все x_n с $n \geq p$ содержатся в шаре $B(x_p, 1)$ и, следовательно, вся последовательность находится в шаре $B(x_p, R)$, где

$$R = \max [d(x_p, x_0), d(x_p, x_1), \dots, d(x_p, x_{p-1}), 1].$$

Теорема 40. Если последовательность Коши x_0, x_1, x_2, \dots в метрическом пространстве E имеет точку сгущения a , то эта последовательность сходится к a .

Доказательство. При заданном $\varepsilon > 0$ существует такое p , что для $m \geq p$ и $n \geq p$ имеет место неравенство $d(x_m, x_n) \leq \varepsilon/2$. Существует, кроме того, бесконечное множество, a , значит, по крайней мере одно такое m , что $d(x_m, a) \leq \varepsilon/2$ для $m \geq p$. Отсюда получаем, что $d(x_n, a) \leq d(x_n, x_m) + d(x_m, a) \leq \varepsilon$ при $n \geq p$.

Следствие 1. Всякая последовательность Коши либо не имеет точек сгущения, либо имеет только одну такую точку, и тогда она сходится.

Следствие 2. Если последовательность Коши такова, что одна из ее подпоследовательностей сходится, то и сама последовательность сходится.

В самом деле, из предположения следует, что данная последовательность имеет точку сгущения.

Определение. Метрическое пространство E называется *полным*, если любая последовательность Коши в нем сходится.

Мы уже видели, сколь важным является понятие полноты. Необходимо поэтому найти критерии, позволяющие выяснить, является ли данное метрическое пространство полным.

Заметим здесь же, что понятие полного пространства не имеет смысла в случае неметрических топологических пространств. Впрочем, легко видеть, что для двух эквивалентных метрических пространств последовательности Коши не обязательно должны быть одинаковыми и что одно из таких пространств может быть полным, а второе — нет. *Понятие полного*

пространства является метрическим, а не топологическим понятием.

Рассмотрим в качестве примера вещественную прямую \mathbb{R} . Если на ней ввести естественную метрику, то, как мы увидим позже, она будет полной (это известно и из общего курса математического анализа). Последовательность \mathbb{N} целых чисел в этой метрике не является последовательностью Коши. Напротив, если мы рассмотрим на дополненной прямой $\bar{\mathbb{R}}$ одну из метрик, определенных на стр. 100, то она индуцирует на прямой \mathbb{R} некоторую метрику, эквивалентную естественной. Однако, в этой метрике \mathbb{R} не полно. В самом деле, последовательность целых чисел является последовательностью Коши на $\bar{\mathbb{R}}$, поскольку она сходится к $+\infty$, а, значит, она является последовательностью Коши на \mathbb{R} , где она вовсе не сходится.

Теорема 41. Любое метрическое пространство E , все замкнутые шары которого компактны, полно. В частности, полно любое метрическое компактное пространство и любое конечномерное нормированное векторное пространство.

Доказательство. Согласно теореме 29, каждая последовательность Коши x_0, x_1, x_2, \dots ограничена, а, следовательно, содержится в некотором замкнутом шаре B , т. е. в некотором компакте. По теореме Больцано — Вейерштрасса она имеет хотя бы одну точку сгущения. Согласно теореме 40, последовательность сходится, а, следовательно, E полно.

Замечания. 1°) Из условий теоремы следует, что E локально компактно. Однако в локально компактных метрических пространствах замкнутые шары не обязательно компактны¹⁾. Рассмотрим, например, пространство $\bar{\mathbb{R}}$, снабженное одной из метрик, указанных на стр. 100. Подпространство \mathbb{R} в этой метрике локально компактно, однако, как было замечено выше, оно не полно. Легко проверить, что все замкнутые шары в нем не компактны. В самом деле, в силу компактности дополненной прямой $\bar{\mathbb{R}}$ она ограничена. Следовательно, для достаточно большого r шар радиуса r в $\bar{\mathbb{R}}$ совпадает со всей прямой $\bar{\mathbb{R}}$. Тогда шар радиуса r в $\bar{\mathbb{R}}$ совпадает с некомпактным множеством \mathbb{R} .

2°) Бесконечномерные нормированные векторные пространства могут быть как полными, так и неполными. Позже мы приведем примеры тех и других пространств. Наиболее интересными, естественно, будут полные пространства.

¹⁾ Локальная компактность метрического пространства E является топологическим свойством. Свойство компактности замкнутых шаров является метрическим свойством.

3°) Существуют не локально компактные метрические пространства, которые являются полными. Именно об этом говорится в замечании 2°).

4°) Точно так же как и свойство компактности, свойство полноты метрического пространства является внутренним свойством самого пространства. Однако если F является подмножеством некоторого метрического пространства E , то мы будем говорить, что F полно, если оно полно как метрическое пространство с индуцированной метрикой.

Теорема 42. Полное подмножество F метрического пространства E замкнуто в E .

Доказательство. Пусть $a \in E$ является точкой прикосновения F . Согласно теореме 15, существует последовательность точек x_0, x_1, x_2, \dots из F , сходящаяся к a . Это — последовательность Коши в E , а, следовательно, и в F . Поскольку F полно, последовательность имеет предел a' в F , а значит, и в E . Последнее означает, что $a = a' \in F$. Так как $a \in F$, то F замкнуто.

Следствие 1. Пусть E — метрическое пространство. Полное подмножество F , не совпадающее с E , не обязательно полно.

В самом деле, замыканием F в E является $E \neq F$. Следовательно, F не замкнуто. Отсюда следует, например, что поле \mathbb{Q} рациональных чисел с естественной метрикой не полно. Легко, впрочем, образовать несходящуюся последовательность Коши в \mathbb{Q} . Достаточно взять последовательность из \mathbb{Q} , сходящуюся в \mathbb{R} к иррациональному числу. Необходимость расширения поля рациональных чисел и введения поля вещественных чисел диктуется двумя важными причинами: каждое ограниченное непустое множество должно иметь точную верхнюю грань и каждая последовательность Коши должна быть сходящейся.

Следствие 2. Конечномерное подпространство F нормированного векторного пространства E замкнуто в E .

В самом деле, согласно изложенному перед теоремой 41, подпространство F полно. Можно доказать, что этот результат сохраняется, если E — всего лишь векторное топологическое пространство.

Естественно, обратное утверждение к предыдущей теореме не верно. Например, $F = E$ всегда замкнуто в E , а тем не менее не обязательно полно. Однако имеет место

Теорема 43. Любое замкнутое подмножество F полного метрического пространства E полно.

Доказательство. Пусть x_0, x_1, x_2, \dots — последовательность Коши из F . Она является также последовательностью Коши в E и, так как E полно, сходится к некоторой точке

$a \in E$. Поскольку все x_n лежат в F , то a является точкой сгущения для F (теорема 15) и, в силу замкнутости F , принадлежит ему. Итак, любая последовательность Коши в F сходится в F , т. е. F полно.

Из теорем 42 и 43 следует, что если E полно, то его полные и замкнутые подмножества совпадают. Отметим сходство между этими свойствами и соответствующими свойствами компактных пространств.

Теорема 44. Пусть E_1 и E_2 — два полных метрических пространства. Тогда произведение $E_1 \times E_2$ будет также полным в смысле какой-либо из метрик, указанных на стр. 63.

Доказательство. Обозначим через d_1 и d_2 расстояния в E_1 и E_2 , и пусть δ — расстояние в произведении E_1 на E_2 , обладающее следующими свойствами:

1°) δ определяет в $E_1 \times E_2$ топологию произведения топологий в E_1 и E_2 , определенных расстояниями d_1 и d_2 ;

2°) существует такое число k , что для каждой пары (x_1, x_2) , (y_1, y_2) элементов $E_1 \times E_2$ имеют место неравенства:

$$k\delta((x_1, x_2), (y_1, y_2)) \geq d_1(x_1, y_1) \text{ и } \geq d_2(x_2, y_2). \quad (\text{II}, 11;3)$$

Докажем, что в так определенной метрике пространство $E_1 \times E_2$ полно. В самом деле, пусть $((x_1)_n, (x_2)_n)$ — некоторая последовательность Коши этого пространства. Из условия 2°) следует, что $(x_1)_n$ является последовательностью Коши в E_1 , а последовательность $(x_2)_n$ — последовательностью Коши в E_2 . Поскольку пространства E_1 и E_2 полны, эти последовательности сходятся соответственно к элементам $a_1 \in E_1$ и $a_2 \in E_2$. Но тогда в топологии произведения, т. е. в рассматриваемой метрике δ , последовательность $((x_1)_n, (x_2)_n)$ сходится к (a_1, a_2) в $E_1 \times E_2$, что означает полноту этого пространства.

Продолжение равномерно непрерывных отображений

Теорема 45. Пусть E и F — некоторые метрические пространства, E_1 — подпространство, плотное в E , а f_1 — отображение E_1 в F . Предположим также, что f_1 равномерно непрерывно на E_1 , а F полно.

Тогда существует, и притом единственное, непрерывное отображение f пространства E в F , являющееся продолжением f_1 . Это отображение, кроме того, равномерно непрерывно.

Доказательство. 1°) Если даже f_1 только непрерывно, а F не полно, то не может существовать более одного отображения, обладающего указанными свойствами. В самом деле, пусть имеется некоторое такое отображение f . Пусть x — точ-

ка E . Так как E_1 плотно в E , то, согласно теореме 15, существует последовательность x_0, x_1, x_2, \dots в E_1 , сходящаяся к x . В силу теоремы 16, из непрерывности f вытекает, что последовательность $f(x_n) = f_1(x_n)$ сходится к $f(x)$ в F . Следовательно, $f(x)$, являясь пределом $f_1(x_n)$, полностью определена, что доказывает ее единственность.

2°) Для доказательства существования необходимо предполагать, кроме того, что f равномерно непрерывна и F полно. Пусть теперь снова x — некоторая точка E и x_0, x_1, x_2, \dots — последовательность в E_1 , сходящаяся к x . Тогда это последовательность Коши в E . Отсюда немедленно вытекает, что последовательность $f_1(x_n)$ является последовательностью Коши в F . В самом деле, для любого заданного $\varepsilon > 0$, в силу равномерной непрерывности, существует такое число $\eta > 0$, что из $x' \in E_1, x'' \in E_1, d(x', x'') \leq \eta$ следует $d(f_1(x'), f_1(x'')) \leq \varepsilon$. Поскольку x_0, x_1, x_2, \dots является последовательностью Коши, существует такое целое p , что из $m \geq p, n \geq p$ следует $d(x_m, x_n) \leq \eta$, а, следовательно, и $d(f(x_m), f(x_n)) \leq \varepsilon$, что полностью доказывает наше утверждение.

Поскольку F по предположению полно, последовательность $f_1(x_n)$ имеет предел в F . Обозначим его через $f(x)$. Надо доказать прежде всего, что $f(x)$ полностью определяется заданием x , т. е. не зависит от выбора последовательности x_n . Рассмотрим с этой целью какие-либо две последовательности x'_n и x''_n , сходящиеся к одному и тому же x . «Смешанная» последовательность $x'_0, x''_0, x'_1, x''_1, x'_2, x''_2, \dots$ является некоторой последовательностью из E_1 , сходящейся к x . Но тогда последовательность $f(x'_0), f(x''_0), f(x'_1), f(x''_1), \dots, f(x'_n), f(x''_n), \dots$ должна сходиться к некоторому элементу F и это полностью доказывает тот факт, что обе последовательности $f(x'_n)$ и $f(x''_n)$ имеют один и тот же предел в F .

Показав это, мы определили некоторое отображение f пространства E в пространство F . Это отображение является тривиальным продолжением f_1 , ибо для любого $x \in E_1$ можно взять последовательность x, x, x, \dots , сходящуюся к x . При этом образ $f(x)$ является пределом последовательности $f_1(x), f_1(x), f_1(x), \dots$, т. е. совпадает с $f_1(x)$. Остается, следовательно, доказать равномерную непрерывность отображения f . Пусть $\varepsilon > 0$ — заданное число, а $\eta > 0$ — соответствующее ему число, входящее в определение равномерной непрерывности f_1 . Пусть x и y — произвольные точки E , такие, что $d(x, y) \leq \eta/2$. Докажем, что тогда $d(f(x), f(y)) \leq \varepsilon$, что обеспечит равномерную непрерывность f . Пусть x_n и y_n — последовательности из E_1 , сходящиеся соответственно к x и y . В силу непрерывности функции расстояния (теорема 17₂), существует такое целое число p , что для $n \geq p$

имеем $d(x_n, y_n) \leq d(x, y) + \eta/2$. Отсюда вытекает, что для $n \geq p$ выполняется неравенство $d(x_n, y_n) \leq \eta$ и, следовательно, $d(f_1(x_n), f_1(y_n)) \leq \varepsilon$. Поскольку в F функция расстояния непрерывна, а последовательности $f_1(x_n)$ и $f_1(y_n)$ сходятся соответственно к $f(x)$ и $f(y)$, то $d(f_1(x_n), f_1(y_n))$ сходится к $d(f(x), f(y))$. Следовательно, $d(f(x), f(y)) \leq \varepsilon$. Доказательство теоремы закончено.

Замечания. 1°) Предположение относительно полноты пространства F существенно. Если, например, мы возьмем $F = E_1$, а в качестве f_1 возьмем тождественное отображение E_1 в E_1 , то оно не может быть продолжено до непрерывного отображения f пространства E в E_1 .

В самом деле, пусть x — некоторая точка E , не принадлежащая E_1 . Пусть x_n — последовательность из E_1 , сходящаяся к x . Тогда, если бы существовало непрерывное продолжение f , последовательность $f_1(x_n) = x_n$ сходилась бы к $f(x)$ в E_1 , а, значит, и в E , и мы имели бы $f(x) = x$, что невозможно, так как $x \notin E_1$.

2°) Недостаточно предполагать, что f_1 только непрерывно. Рассмотрим, например, пространство $E = \bar{\mathbb{R}}$ и снабдим его одной из метрик, указанных в теореме 38; положим $E_1 = \mathbb{R}$. В качестве F выберем прямую \mathbb{R} с естественной метрикой. Тогда F полно. Легко видеть, что тождественное отображение f_1 пространства E_1 в F непрерывно, но не равномерно непрерывно. Его, очевидно, нельзя продолжить до непрерывного отображения f пространства $E = \bar{\mathbb{R}}$ в $F = \mathbb{R}$. В самом деле, точка $x = +\infty$ из $E = \bar{\mathbb{R}}$ является пределом в E последовательности целых чисел \mathbb{N} . Однако образ этой последовательности, т. е. само \mathbb{N} , не сходится в $F = \mathbb{R}$, тогда как он должен был бы сходиться к $f(+\infty)$, если бы отображение f существовало.

Частные свойства конечномерных топологических векторных пространств

Мы уже отмечали некоторые такие свойства: теорема 13, теорема 23, следствие 2 из теоремы 41. Остается доказать свойство, указанное на стр. 74:

Теорема 45₂ (Фредерик Рисс). Для того чтобы топологическое векторное пространство было локально компактным, необходимо и достаточно, чтобы оно было конечномерным.

Доказательство. Для простоты мы будем считать, что E является нормированным векторным пространством. Незначительное видоизменение доказательства даст общий результат. Мы уже знаем, что из конечномерности E вытекает его локальная компактность. Нам же надо доказать обратное утвержде-

ние. Предположим, что в E имеется некоторая компактная окрестность \mathcal{V} точки 0. В дальнейшем мы будем пользоваться следующими обозначениями. Если A — некоторая часть E , то λA , где λ — скаляр, является множеством точек λx , где $x \in A$. Если A и B — две части E , то через $A + B$ обозначается множество точек $x + y$, $x \in A$, $y \in B$. (Внимание! $2A$ содержится в $A + A$, но, вообще говоря, с ним не совпадает.) Векторное подпространство M пространства E характеризуется следующими свойствами: $M + M = M$ и $\lambda M \subset M$ для любого λ . Далее, $2\mathcal{V}$ является окрестностью 0 и $a + \mathcal{V}$ является открытой окрестностью точки a . Когда a пробегает множество $2\mathcal{V}$, множества $a + \mathcal{V}$ образуют открытое покрытие множества $2\mathcal{V}$. В силу предположения о компактности, существует конечное число точек a_1, a_2, \dots, a_n , таких, что $a_i + \mathcal{V}$ покрывают $2\mathcal{V}$. Пусть M — векторное подпространство, порожденное векторами a_i . Оно конечномерно, а $M + \mathcal{V}$ покрывает $2\mathcal{V}$.

Имеем: $M + \mathcal{V} = M + M + \mathcal{V} \supset M + 2\mathcal{V}$. Умножая на 2, получаем: $M + \mathcal{V} \supset M + 2\mathcal{V} = 2M + 2\mathcal{V} = 2(M + \mathcal{V}) \supset 2(M + 2\mathcal{V}) = = 2M + 4\mathcal{V} = M + 4\mathcal{V}$. Для любого n имеют место включения $M + \mathcal{V} \supset M + 2^n\mathcal{V} \supset 2^n\mathcal{V}$. Так как объединение всех $2^n\mathcal{V}$ дает все пространство, то $M + \mathcal{V}$ совпадает со всем пространством E . Отсюда можно получить, что M также совпадает с E , а, значит, E конечномерно. Если бы это было не так, то существовала бы точка $a \notin M$. Так как M замкнуто (следствие 2 теоремы 41), то тогда существовал бы шар с центром в a , не пересекающийся с M . Этот шар можно представить в виде $a + B$, где B — шар с центром в нуле пространства. Тогда $M + B$ не будет содержать a . Для произвольного $\rho > 0$ множество $M + \rho B = \rho M + \rho B$ не содержит ra , и мы получаем $ra \notin M + \rho B$. Однако окрестность \mathcal{V} компактна, а, значит, ограничена. При достаточно большом ρ она содержится в ρB , и тогда $ra \notin M + \mathcal{V}$, что противоречит ранее полученному результату $M + \mathcal{V} = E$.

§ 12. ТЕОРЕМА О НЕПОДВИЖНОЙ ТОЧКЕ

Определение. Отображение f метрического пространства E в E называется *сжатием*, если существует такая положительная постоянная $k < 1$, что для любой пары элементов x и y из E имеет место неравенство

$$d(f(x), f(y)) \leq k d(x, y). \quad (\text{II}, 12; 1)$$

Отсюда, очевидно, вытекает, что отображение f удовлетворяет условию Липшица и, следовательно, равномерно непрерывно.

Точка a называется *неподвижной точкой* отображения f , если $f(a) = a$.

Теорема 46. Каждое сжатие полного метрического пространства в себя имеет неподвижную точку, и притом единственную.

Доказательство. Единственность неподвижной точки очевидна, даже если E не полно. В самом деле, если a и b — две неподвижные точки, то

$$d(a, b) \leq kd(a, b) < d(a, b), \text{ если только } d(a, b) \neq 0. \quad (\text{II, 12; 2})$$

Поэтому $d(a, b) = 0$, т. е. a и b совпадают.

Докажем теперь существование неподвижной точки. Мы применим для этого метод, называемый методом последовательных приближений.

Пусть x_0 — произвольная точка E . Положим

$$x_1 = f(x_0), \quad x_2 = f(x_1), \quad \dots, \quad x_n = f(x_{n-1}), \quad \dots \quad (\text{II, 12; 3})$$

Мы образовали некоторую бесконечную последовательность x_0, x_1, x_2, \dots элементов из E . Докажем, что она является последовательностью Коши. Поскольку f — сжатие, мы можем записать последовательность неравенств:

$$\begin{aligned} d(x_2, x_1) &\leq kd(x_1, x_0), \\ d(x_3, x_2) &\leq kd(x_2, x_1) \leq k^2d(x_1, x_0), \\ &\dots \\ d(x_{n+1}, x_n) &\leq k^n d(x_1, x_0), \end{aligned} \quad (\text{II, 12; 4})$$

а, следовательно,

$$\begin{aligned} d(x_{n+p}, x_n) &\leq d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \dots \\ &\dots + d(x_{n+1}, x_n) \leq (k^{p-1} + k^{p-2} + \dots + k + 1) k^n d(x_1, x_0) \leq \\ &\leq \frac{k^n}{1-k} d(x_1, x_0). \end{aligned} \quad (\text{II, 12; 5})$$

Отсюда получаем, что $d(x_{n+p}, x_n)$ стремится к 0, когда p стремится к $+\infty$. Таким образом, последовательность x_n является последовательностью Коши. Так как E по условию полно, то эта последовательность сходится к некоторому пределу a . В силу непрерывности f последовательность $x_{n+1} = f(x_n)$ сходится к $f(a)$, а так как x_{n+1} также сходится к a , то $f(a) = a$ является неподвижной точкой отображения f . Описанная процедура не только позволяет установить существование неподвижной точки, но и является практическим методом ее вычисления. Неравенство

$$d(x_n, a) = \lim_{p \rightarrow \infty} d(x_n, x_{n+p}) \leq \frac{k^n}{1-k} d(x_1, x_0) \quad (\text{II, 12; 6})$$

дает оценку быстроты сходимости последовательности x_n к a .

Замечания. 1°) Предположение $k < 1$ необходимо, ибо условие $k \leq 1$ не является достаточным ни для существования, ни для единственности неподвижной точки.

Тождественное отображение метрического пространства в себя удовлетворяет неравенству (II, 12; 1) при $k = 1$; все точки пространства являются для него неподвижными. Сдвиг $x \rightarrow x + 1$ на прямой также удовлетворяет неравенству (II, 12; 1) с $k = 1$, но он не имеет ни одной неподвижной точки.

2°) Пусть f — отображение E в E , не обязательно являющееся сжатием. Если одна из его итераций f_p является сжатием, то отображение f имеет неподвижную точку, и притом единственную.

Напомним, что итерации отображения f множества E в себя определяются формулами:

$$f_2 = f \circ f, f_3 = f \circ f \circ f, \dots, f_p = f \circ f_{p-1} = f_{p-1} \circ f. \quad (\text{II, 12; 7})$$

Пусть a — неподвижная точка f . Она является неподвижной точкой для любой из итераций f , среди которых имеется итерация f_p , являющаяся сжатием, а, следовательно, имеющая только одну неподвижную точку. Отсюда следует, что само f может иметь только одну неподвижную точку. Для доказательства существования будем исходить из того, что сжатие f_p имеет единственную неподвижную точку a . Элемент $f_{p+1}(a)$ может быть записан двумя способами: либо $f(f_p(a)) = f(a)$, либо $f_p(f(a))$. Равенство $f_p(f(a)) = f_{p+1}(a) = f(a)$ означает, что $f(a)$ является новой неподвижной точкой итерации f_p . Поскольку f_p , будучи сжатием может иметь только одну неподвижную точку, то $f(a) = a$, т. е. a является неподвижной точкой отображения f . Таким образом, f имеет единственную неподвижную точку a . Мало того, ее можно получить теми же последовательными приближениями, так как последовательность x_0, x_1, x_2, \dots , определенная формулами (II, 12; 3), сходится к a . В самом деле, каждая из последовательностей

$$x_0, x_p, x_{2p}, \dots, x_{np}, \dots;$$

$$x_1, x_{p+1}, x_{2p+1}, \dots, x_{np+1}, \dots; \quad x_{p-1}, x_{2p-1}, \dots, x_{np+p-1}, \dots$$

сходится к a , поскольку она определяет последовательные приближения для итерации f_p . Последнее означает, что вся последовательность x_0, x_1, x_2, \dots сходится к a . (Отметим следующее: предположение о том, что отображение f_p является сжатием, не влечет за собой непрерывности f .)

Предположим теперь, что сжатие f достаточно хорошоим образом зависит от некоторого параметра λ . Тогда для каждого значения λ оно имеет, и притом единственную, неподвижную точку a_λ . Отыщем условия, при которых неподвижная точка a_λ непрерывно зависит от параметра λ .

Теорема 46₂. Пусть E — полное метрическое пространство, Λ — топологическое пространство, а f — отображение $E \times \Lambda$ в E . Предположим, что для каждой фиксированной точки x из E частное отображение $\lambda \mapsto f(x, \lambda)$ пространства Λ в пространство E непрерывно и что для каждого $\lambda \in \Lambda$ отображение $f_\lambda: x \mapsto f(x, \lambda)$ является сжатием E в E , соответствующим числу $k < 1$ (формула (II, 12; 1)), не зависящему от λ . Тогда единственная неподвижная точка a_λ отображения f_λ непрерывно зависит от параметра λ , иначе говоря, отображение $\lambda \mapsto a_\lambda$ пространства Λ в E непрерывно.

Доказательство. Обозначим через λ_0 какую-либо точку Λ . Тогда

$$\begin{aligned} d(a_\lambda, a_{\lambda_0}) &= d(f_\lambda(a_\lambda), f_{\lambda_0}(a_{\lambda_0})) \leq d(f_\lambda(a_\lambda), f_\lambda(a_{\lambda_0})) + d(f_\lambda(a_{\lambda_0}), f_{\lambda_0}(a_{\lambda_0})) \leq \\ &\leq kd(a_\lambda, a_{\lambda_0}) + d(f(a_{\lambda_0}, \lambda), f(a_{\lambda_0}, \lambda_0)), \quad (\text{II, 12; 8}) \end{aligned}$$

откуда

$$(1 - k)d(a_\lambda, a_{\lambda_0}) \leq d(f(a_{\lambda_0}, \lambda), f(a_{\lambda_0}, \lambda_0)).$$

Зададим $\varepsilon > 0$. В силу непрерывности f по λ при фиксированном $x = a_{\lambda_0}$ существует в Λ такая окрестность \mathcal{U} точки λ_0 , что соотношение $\lambda \in \mathcal{U}$ влечет за собой $d(f(a_{\lambda_0}, \lambda), f(a_{\lambda_0}, \lambda_0)) \leq \varepsilon(1 - k)$. Тогда из соотношения $\lambda \in \mathcal{U}$ следует также неравенство $d(a_\lambda, a_{\lambda_0}) \leq \varepsilon$, доказывающее непрерывность рассматриваемого отображения в точке λ_0 пространства Λ .

§ 13. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ НОРМИРОВАННЫХ ВЕКТОРНЫХ ПРОСТРАНСТВ И ПРОСТРАНСТВ БАНАХА

Пусть E и F — два векторных нормированных пространства над полем \mathbb{K} , которое мы всегда будем предполагать полем вещественных или полем комплексных чисел¹⁾. Пусть u — некоторое отображение E в F . Напомним, что u называется линейным, если

$$\begin{aligned} u(\vec{x} + \vec{y}) &= u(\vec{x}) + u(\vec{y}), \quad u(\lambda \vec{x}) = \lambda u(\vec{x}), \\ \lambda \in \mathbb{K}, \quad \vec{x} &\in E, \quad \vec{y} \in E. \end{aligned}$$

¹⁾ Любое векторное пространство над полем комплексных чисел является также векторным пространством над полем вещественных чисел. Поэтому, если заданы два векторных пространства E и F , одно над полем вещественных, другое над полем комплексных чисел, то оба их можно рассматривать как векторное пространство над полем вещественных чисел.

Принято одним и тем же символом $\vec{0}$ обозначать нейтральные элементы (вообще говоря, различные) пространства E и F . Очевидно, $u(\vec{0}) = \vec{0}$. Одним и тем же символом $\| \cdot \|$ принято часто обозначать нормы в E и F .

Если E конечномерно, то линейное отображение E в F заведомо непрерывно и даже удовлетворяет условию Липшица, а значит, равномерно непрерывно. В самом деле, поскольку все нормы в конечномерном пространстве эквивалентны (теорема 13), то для доказательства непрерывности, а также и равномерной непрерывности¹⁾ достаточно ввести в E некоторый базис $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ и определить норму любого элемента \vec{x} через его координаты в этом базисе x_1, x_2, \dots, x_n по формуле $\|\vec{x}\| = \max |x_i|$. При этом

$$\begin{aligned} \|u(\vec{x}') - u(\vec{x}'')\| &= \|u(\vec{x}' - \vec{x}'')\| = \\ &= \left\| u\left(\sum_{i=1}^n (x'_i - x''_i) \vec{e}_i\right) \right\| = \left\| \sum_{i=1}^n (x'_i - x''_i) u(\vec{e}_i) \right\| \leqslant \\ &\leqslant \left(\sum_{i=1}^n \|u(\vec{e}_i)\| \right) \max_{i=1, 2, \dots, n} |x'_i - x''_i| \leqslant k \|\vec{x}' - \vec{x}''\|, \end{aligned}$$

что полностью доказывает наше утверждение.

Если же, напротив, мы будем считать E бесконечномерным, то высказанное утверждение будет неверным. Как это ни кажется странным, существуют линейные разрывные отображения.

Приведем пример. Возьмем в качестве E векторное пространство полиномов с вещественными коэффициентами над полем вещественных чисел²⁾. Норму в E введем по формуле

$$\|P\| = \max_{0 \leqslant x \leqslant 1} |P(x)|. \quad (\text{II}, 13; 1)$$

В том, что это действительно норма, легко убедиться, ибо все необходимые для нормы неравенства проверяются без труда. (Существование максимума вытекает из непрерывности P и теоремы 29.) Очевидно, $\|\lambda P\| = |\lambda| \|P\|$ для любого вещественного λ , а для любых полиномов P и Q имеет место неравенство $\|P + Q\| \leqslant \|P\| + \|Q\|$. Если $P \not\equiv 0$, то $\|P\| > 0$; если же $\|P\| =$

¹⁾ Непрерывность является топологическим, а не метрическим свойством, чего нельзя сказать относительно равномерной непрерывности; однако теорема 12 утверждает, что в определении равномерной непрерывности некоторого отображения нормированного векторного пространства в другое пространство нормы можно заменить эквивалентными нормами.

²⁾ Сложение и умножение на вещественное число является обычным сложением полиномов и обычным умножением их на число. При этом E будет векторным пространством, нулевым элементом которого является полином $\equiv 0$. Векторное пространство полиномов степени $\leqslant m$ имеет размерность $m+1$. Один из его базисов составляют полиномы $1, x, x^2, \dots, x^m$. Векторное пространство всевозможных полиномов поэтому бесконечномерно.

$= 0$, то полином P , равный нулю на $[0, 1]$, тождественно равен нулю.)

Выберем в качестве u вещественную функцию над E , которая каждому полиному P ставит в соответствие его значение в точке $x = 3$: $u(P) = P(3)$. Это, очевидно, некоторая линейная форма над E . Покажем, что она разрывна. Для этого достаточно рассмотреть последовательность полиномов, определенных формулами:

$$P_n(x) = \left(\frac{x}{2}\right)^n. \quad (\text{II, 13; 2})$$

Очевидно, $\|P_n(x)\| = 2^{-n}$, и, следовательно, эта последовательность сходится к 0 в E . Тем не менее $P_n(3) = (3/2)^n$, и потому последовательность значений $u(P_n)$ стремится к $+\infty$, а это означает, что отображение u разрывно.

Теорема 47. Всякое линейное отображение нормированного векторного пространства E в нормированное векторное пространство F , непрерывное в нуле пространства, непрерывно всюду. Это отображение, кроме того, удовлетворяет условию Липшица, а, следовательно, равномерно непрерывно. Для того чтобы отображение обладало этим свойством, необходимо и достаточно, чтобы существовало такое постоянное число $k \geq 0$, что

$$\|u(\vec{x})\| \leq k \|\vec{x}\| \text{ для всех } \vec{x} \text{ из } E. \quad (\text{II, 13; 3})$$

Доказательство. Докажем прежде всего, что если отображение u непрерывно в нуле пространства E , то существует число k , при котором имеет место неравенство (II, 13; 3). Непрерывность u в нуле влечет за собой существование такого числа $\eta > 0$, при котором из $\|\vec{x}\| \leq \eta$ следует $\|u(\vec{x})\| \leq 1$. В силу линейности u , с помощью гомотетии получаем, что из $\|\vec{x}\| \leq \lambda \eta$ следует $\|u(\vec{x})\| \leq \lambda$. Так как первое из этих неравенств всегда выполняется, если взять $\lambda = \|\vec{x}\|/\eta$, то будет всегда выполняться и второе неравенство $\|u(\vec{x})\| \leq \|\vec{x}\|/\eta$, и мы получим искомое неравенство с $k = 1/\eta$. Обратно, предположим, что существует некоторое число k , при котором имеет место неравенство (II, 13; 3). Тогда отображение u не только будет непрерывным в нуле пространства E , но будет также непрерывным всюду и даже удовлетворяющим условию Липшица, ибо, в силу линейности u , имеет место неравенство

$$\|u(\vec{x}') - u(\vec{x}'')\| = \|u(\vec{x}' - \vec{x}'')\| \leq k \|\vec{x}' - \vec{x}''\|. \quad (\text{II, 13; 4})$$

Замечание. Это теорема включает в себя, как частный случай, теорему 12. В самом деле, утверждение «две нормы в векторном пространстве E эквивалентны» равносильно утверждению «тождественное отображение пространства E , снабженного одной из двух норм, на пространство E , снабженное другой нормой, является непрерывным». Но тогда из (II, 13; 3) следует (II, 4; 1).

Ядро и образ непрерывного линейного отображения

Пусть E и F — векторные пространства над полем K и u — линейное отображение E в F .

Ядром отображения u называется множество таких \vec{x} из E , что $u(\vec{x}) = \vec{0}$. Это множество является также прообразом $u^{-1}\{\vec{0}\}$. Ядро является векторным подпространством в E , ибо для любых \vec{x} и \vec{y} , принадлежащих ядру, $u(\vec{x}) = \vec{0}$, $u(\vec{y}) = \vec{0}$, а так как $u(\vec{x} + \vec{y}) = u(\vec{x}) + u(\vec{y}) = \vec{0}$, то $\vec{x} + \vec{y}$ также принадлежит ядру; $u(\lambda\vec{x}) = \lambda u(\vec{x}) = \vec{0}$ для любого скаляра λ , т. е. $\lambda\vec{x}$ также принадлежит ядру.

Образом отображения u называется образ $u(E)$ пространства E при отображении u . Это также некоторое векторное подпространство в F . Размерность этого подпространства называется **рангом отображения u** . Напомним, что если E и F конечномерны и если в рассматриваемых пространствах выбраны базисы, то u выражается с помощью некоторой матрицы. Ранг отображения u является наибольшим порядком отличных от нуля миноров этой матрицы. Напомним, что ранг является также разностью размерности E и размерности ядра отображения u .

Если пространства E и F нормированы, а u непрерывно, то ядро является замкнутым векторным подпространством, ибо оно является прообразом подмножества F , сводящегося к точке, которое замкнуто (теорема 8). Образ же не обязательно замкнут¹⁾.

Определение. Точная нижняя грань таких чисел k , при которых выполняется неравенство (II, 13; 3), называется **нормой линейного отображения u** ²⁾. Эта норма обозначается через $\|u\|$.

¹⁾ Образ замкнутого подмножества при непрерывном отображении во все не обязан быть замкнутым!

²⁾ Норму можно определить и тогда, когда линейное отображение u разрывно. В этом случае следует положить $\|u\|=+\infty$. Заметим, что эта точная нижняя грань является минимумом.

и определяется следующими соотношениями:

$$\|u\| = \sup_{\substack{\vec{x} \neq \vec{0} \\ \vec{x} \in E}} \frac{\|u(\vec{x})\|}{\|\vec{x}\|} = \sup_{\substack{\|\vec{x}\| \leq 1 \\ \vec{x} \neq \vec{0}}} \|u(\vec{x})\| = \sup_{\|\vec{x}\|=1} \|u(\vec{x})\|. \quad (\text{II, 13; 5})$$

Заметим, в частности, что для любого \vec{x}

$$\|u(\vec{x})\| \leq \|u\| \|\vec{x}\|. \quad (\text{II, 13; 6})$$

Теорема 48. *Множество $\mathcal{L}(E; F)$ линейных непрерывных отображений векторного нормированного пространства E в векторное нормированное пространство F имеет структуру нормированного векторного пространства с нормой, определенной формулой (II, 13; 5). Кроме того, если G — третье нормированное векторное пространство, $u \in \mathcal{L}(E; F)$, а $v \in \mathcal{L}(F; G)$, то $v \circ u \in \mathcal{L}(E; G)$ и имеет место неравенство*

$$\|v \circ u\| \leq \|v\| \|u\|. \quad (\text{II, 13; 7})$$

Доказательство. Докажем сначала, что $\mathcal{L}(E; F)$ имеет структуру векторного пространства над полем \mathbb{K} . Пусть u_1 и u_2 — два линейных непрерывных отображения E в F . Определим отображение $u_1 + u_2$ формулой:

$$(u_1 + u_2)(\vec{x}) = u_1(\vec{x}) + u_2(\vec{x}). \quad (\text{II, 13; 8})$$

Легко проверить, что так определенное отображение является новым линейным отображением E в F . С другой стороны, неравенство (полученное из (II, 13; 6)):

$$\|(u_1 + u_2)(\vec{x})\| \leq \|u_1(\vec{x})\| + \|u_2(\vec{x})\| \leq (\|u_1\| + \|u_2\|) \|\vec{x}\| \quad (\text{II, 13; 9})$$

показывает, что оно непрерывно, т. е. $u_1 + u_2 \in \mathcal{L}(E; F)$, и, кроме того, согласно (II, 13; 5), имеет место неравенство

$$\|u_1 + u_2\| \leq \|u_1\| + \|u_2\|. \quad (\text{II, 13; 10})$$

Мы видим, что в $\mathcal{L}(E; F)$ определено сложение элементов, превращающее его в абелеву группу. Нулем этой группы является нулевое отображение 0 , т. е. такое отображение, которое каждому элементу из E ставит в соответствие нулевой вектор $\vec{0}$ пространства F . Пусть теперь λ — некоторый скаляр. Для любого $u \in \mathcal{L}(E; F)$ определим отображение λu по формуле

$$(\lambda u)(\vec{x}) = \lambda(u(\vec{x})). \quad (\text{II, 13; 11})$$

Легко проверяется, что это снова некоторое линейное отображение. Кроме того, это линейное отображение непрерывно, т. е. $\lambda u \in \mathcal{L}(E; F)$. В самом деле,

$$\|(\lambda u)(\vec{x})\| = |\lambda| \|u(\vec{x})\| \leq |\lambda| \|u\| \|\vec{x}\|. \quad (\text{II}, 13; 12)$$

Кроме того,

$$\|\lambda u\| = \sup_{\|\vec{x}\|=1} \|(\lambda u)(\vec{x})\| = |\lambda| \sup_{\|\vec{x}\|=1} \|u(\vec{x})\|; \quad (\text{II}, 13; 13)$$

другими словами,

$$\|\lambda u\| = |\lambda| \|u\|. \quad (\text{II}, 13; 14)$$

Таким путем в множестве $\mathcal{L}(E; F)$ мы определили умножение на скаляры. Легко видеть, что это умножение обладает по отношению к сложению элементов из $\mathcal{L}(E; F)$ всеми свойствами, необходимыми для того, чтобы можно было считать $\mathcal{L}(E; F)$ векторным пространством над полем \mathbb{K} . Кроме того, из (II, 13; 10) и (II, 13; 14) следует, что отображение $u \rightarrow \|u\|$, действительно, определяет некоторую норму в $\mathcal{L}(E; F)$ (тот факт, что из $\|u\| = 0$ следует $u = 0$, очевидным образом вытекает из (II, 13; 6)). Таким образом, $\mathcal{L}(E; F)$ является нормированным векторным пространством. Далее, если $u \in \mathcal{L}(E; F)$ и $v \in \mathcal{L}(F; G)$, то, очевидно, $v \circ u \in \mathcal{L}(E; G)^1$. С другой стороны, из неравенства $\|(v \circ u)(\vec{x})\| \leq \|v\| \|u(\vec{x})\| \leq \|v\| \|u\| \|\vec{x}\|$ следует, что $\|v \circ u\| \leq \|v\| \|u\|$ — неравенство (II, 13; 7), указанное в формулировке теоремы.

В частности, когда F является полем скаляров \mathbb{K} , то линейное отображение E в \mathbb{K} называется линейной формой на E . Пространство линейных непрерывных форм обозначается через E' и называется сопряженным к E . Оно является векторным нормированным пространством²⁾.

Обозначения. Если $u \in \mathcal{L}(E; F)$ и $\vec{x} \in E$, то через $u \circ \vec{x}$ (или даже $u\vec{x}$) удобно обозначать образ $u(\vec{x})$ вектора \vec{x} при

¹⁾ Композиция двух линейных отображений линейна, а композиция двух непрерывных отображений непрерывна (теорема 10).

²⁾ В курсах алгебры и геометрии сопряженным к векторному пространству E обычно называют векторное пространство линейных форм на E . Это определение совпадает с нашим, если E конечномерно, так как в этом случае при любой норме на E каждая линейная форма непрерывна (стр. 111). Если же пространство E бесконечномерно и нормировано, то существуют разрывные линейные формы, но они не представляют интереса, и термин «сопряженный» предпочтительно оставить за пространством непрерывных линейных форм.

отображении u . Если, кроме того, имеется отображение $v \in \mathcal{L}(F; G)$, то через $v \cdot u$ или vu часто обозначают композицию $v \circ u$. Тогда соотношение $(v \circ u)(\vec{x}) = v(u(\vec{x}))$ может быть записано так: $(v \circ u) \circ \vec{x}$, или $v \circ (\vec{u} \circ \vec{x})$, или $v \circ u \circ \vec{x}$, или $v u \vec{x}$. Если $F = E$ и $u \in \mathcal{L}(E; F)$, то $u \circ u$ обозначается через u^2 , $u \circ u \circ u$ обозначается через u^3 и т. д.

Теорема 49. Пусть E и F — нормированные векторные пространства, E_1 — плотное векторное подпространство E и u_1 — линейное непрерывное отображение E_1 в F . Если F полно, то существует, и притом единственное, непрерывное продолжение u отображения u_1 в E . Это отображение E в F линейно, и норма u в $\mathcal{L}(E; F)$ равна норме u_1 в $\mathcal{L}(E_1; F)$.

Доказательство. Поскольку отображение u_1 линейно и непрерывно, оно равномерно непрерывно (теорема 47). Следовательно, согласно теореме 45, существует, и притом единственное, непрерывное отображение E в F , являющееся продолжением u_1 . Докажем сначала, что u линейно. Пусть \vec{x} и \vec{y} — элементы E , а λ — некоторый скаляр. Пусть $\vec{x}_0, \vec{x}_1, \vec{x}_2, \dots$ и $\vec{y}_0, \vec{y}_1, \vec{y}_2, \dots$ — последовательности элементов из E_1 , сходящиеся соответственно к \vec{x} и \vec{y} . В силу того, что E является топологическим векторным пространством (стр. 66), последовательность $\vec{x}_n + \vec{y}_n$ сходится к $\vec{x} + \vec{y}$, а последовательность $\lambda \vec{x}_n$ сходится к $\lambda \vec{x}$.

Имеем

$$\begin{aligned} u(\vec{x} + \vec{y}) &= \lim_{n \rightarrow \infty} u_1(\vec{x}_n + \vec{y}_n)^1 = \lim_{n \rightarrow \infty} [u_1(\vec{x}_n) + u_1(\vec{y}_n)]^2 = \\ &= \lim_{n \rightarrow \infty} u_1(\vec{x}_n) + \lim_{n \rightarrow \infty} u_1(\vec{y}_n)^3 = u(\vec{x}) + u(\vec{y})^4; \\ u(\lambda \vec{x}) &= \lim_{n \rightarrow \infty} u_1(\lambda \vec{x}_n) = \lim_{n \rightarrow \infty} \lambda u_1(\vec{x}_n) = \lambda \lim_{n \rightarrow \infty} u_1(\vec{x}_n) = \lambda u(\vec{x}), \end{aligned} \quad (\text{II, 13; 15})$$

что говорит о линейности отображения u .

Покажем теперь, что норма u в $\mathcal{L}(E; F)$ равна норме u_1 в $\mathcal{L}(E_1; F)$. Из определения нормы (формула (II, 13; 5)) очевидным образом следует, что $\|u_1\| \leq \|u\|$. Впрочем, если $x \in E$ яв-

¹⁾ Поскольку $\vec{x}_n + \vec{y}_n$ сходится к $\vec{x} + \vec{y}$ и u является функцией, непрерывно продолжающей u_1 .

²⁾ Потому, что u_1 линейно.

³⁾ Так как F является топологическим векторным пространством.

⁴⁾ Поскольку u непрерывно и является продолжением u_1 .

ляется пределом последовательности $\vec{x}_n \in E_1$, то $u(\vec{x})$ является пределом $u_1(\vec{x}_n)$. Из неравенства $\|u_1(\vec{x}_n)\| \leq \|u_1\| \|\vec{x}_n\|$ и непрерывности нормы (теорема 9), переходя к пределу, находим, что $\|u(\vec{x})\| \leq \|u_1\| \|\vec{x}\|$. Это означает, что $\|u\| \leq \|u_1\|$, а следовательно $\|u\| = \|u_1\|$.

Следствие. В условиях теоремы, т. е. если E_1 плотно в E и F полно, векторные нормированные пространства $\mathcal{L}(E; F)$ и $\mathcal{L}(E_1; F)$ можно отождествлять.

В самом деле, поставим в соответствие каждому линейному непрерывному отображению u пространства E в пространство F его сужение u_1 на E_1 . Теорема утверждает, что так определенное соответствие $u \rightarrow u_1$ пространств $\mathcal{L}(E; F)$ и $\mathcal{L}(E_1; F)$ является биекцией, сохраняющей структуру векторного пространства и нормы. Это означает, что два векторных нормированных пространства $\mathcal{L}(E; F)$ и $\mathcal{L}(E_1; F)$ могут быть отождествлены.

Определение. Векторное нормированное полное пространство над полем вещественных и комплексных чисел называется *банаховым пространством*.

Из изложенного ранее на стр. 102 (теорема 41) вытекает, что всякое конечномерное нормированное векторное пространство является банаховым.

Теорема 50. Если E и F — нормированные векторные пространства и F банахово пространство, то пространство $\mathcal{L}(E; F)$ также банахово. В частности, сопряженное пространство E' к пространству E банахово.

Доказательство. Достаточно доказать первое утверждение. Поскольку поле скаляров \mathbb{K} является банаховым, то второе утверждение немедленно вытекает из первого ($E' = \mathcal{L}(E; \mathbb{K})$). Итак, пусть u_0, u_1, u_2, \dots — некоторая последовательность Коши в $\mathcal{L}(E; F)$. Это означает, что при m и n , стремящихся к $+\infty$, норма $\|u_m - u_n\|$ стремится к нулю.

Неравенство $\|u_m(\vec{x}) - u_n(\vec{x})\| \leq \|u_m - u_n\| \|\vec{x}\|$, справедливое для каждого фиксированного \vec{x} , доказывает, что последовательность $u_n(\vec{x})$ является последовательностью Коши в F . Поскольку F по предположению полно, то эта последовательность сходится к некоторому элементу F , который мы будем обозначать через $u(\vec{x})$. Тем самым мы определили отображение $u: \vec{x} \rightarrow u(\vec{x})$ в F . Покажем, прежде всего, что отображение u линейно. Пусть \vec{x} и \vec{y} — элементы E и λ — некоторый скаляр.

Тогда

$$\begin{aligned} u(\vec{x} + \vec{y}) &= \lim_{n \rightarrow \infty} u_n(\vec{x} + \vec{y})^1 = \lim_{n \rightarrow \infty} (u_n(\vec{x}) + u_n(\vec{y}))^2 = \\ &= \lim_{n \rightarrow \infty} u_n(\vec{x}) + \lim_{n \rightarrow \infty} u_n(\vec{y})^3 = u(\vec{x}) + u(\vec{y}), \\ u(\lambda \vec{x}) &= \lim_{n \rightarrow \infty} u_n(\lambda \vec{x})^1 = \lim_{n \rightarrow \infty} \lambda u_n(\vec{x})^2 = \\ &= \lambda \lim_{n \rightarrow \infty} u_n(\vec{x})^3 = \lambda u(\vec{x}), \end{aligned} \tag{II, 13; 16}$$

что означает линейность отображения u .

Докажем теперь, что отображение u непрерывно. Согласно критерию Коши, существует такое целое число p , что из неравенств $m \geq p$ и $n \geq p$ следует неравенство $\|u_m - u_n\| \leq \varepsilon$. Отсюда получаем, что для $m \geq p$, $n \geq p$ $\|u_m\| \leq \|u_n\| + \varepsilon$. Кроме того, выполняются также неравенства

$$\|u_m(\vec{x}) - u_n(\vec{x})\| \leq \varepsilon \|\vec{x}\|, \tag{II, 13; 17}$$

$$\|u_m(\vec{x})\| \leq (\|u_n\| + \varepsilon) \|\vec{x}\|. \tag{II, 13; 18}$$

Фиксируя в (II, 13; 18) $n \geq p$ и вектор \vec{x} и устремляя m к $+\infty$, получаем:

$$\|u(\vec{x})\| \leq (\|u_n\| + \varepsilon) \|\vec{x}\|, \tag{II, 13; 19}$$

откуда следует непрерывность отображения u . Таким образом, u является элементом пространства $\mathcal{L}(E; F)$.

Если теперь при фиксированном \vec{x} и $n \geq p$ в соотношении (II, 13; 17) устремить m к $+\infty$, то получим

$$\|u(\vec{x}) - u_n(\vec{x})\| \leq \varepsilon \|\vec{x}\|, \tag{II, 13; 20}$$

а, следовательно,

$$\|u - u_n\| \leq \varepsilon. \tag{II, 13; 21}$$

Полученное неравенство показывает, что u_n стремится к u в нормированном пространстве $\mathcal{L}(E; F)$ при n , стремящемся к $+\infty$, а это означает, что $\mathcal{L}(E; F)$ полно.

Произведения нормированных векторных пространств

Пусть E_1 и E_2 — два векторных пространства над полем K . Известно, что их произведение $E_1 \times E_2$ может быть снабжено

¹⁾ По определению u .

²⁾ В силу линейности u_n .

³⁾ Так как F является топологическим векторным пространством.

структурой векторного пространства. В самом деле, если (\vec{x}_1, \vec{x}_2) и (\vec{y}_1, \vec{y}_2) — два элемента этого произведения, то можно положить

$$(\vec{x}_1, \vec{x}_2) + (\vec{y}_1, \vec{y}_2) = (\vec{x}_1 + \vec{y}_1, \vec{x}_2 + \vec{y}_2). \quad (\text{II, 13; 22})$$

Если λ — некоторый скаляр, то можно считать

$$\lambda(\vec{x}_1, \vec{x}_2) = (\lambda \vec{x}_1, \lambda \vec{x}_2). \quad (\text{II, 13; 23})$$

Введенные операции на $E_1 \times E_2$ превращают его в векторное пространство над полем \mathbb{K} .

Векторное пространство \mathbb{R}^n (соответственно \mathbb{C}^n) является не чем иным, как векторным пространством произведения n пространств, тождественных \mathbb{R} (соответственно \mathbb{C}), рассматриваемых как векторные пространства над самим \mathbb{R} (или \mathbb{C}).

В этом векторном пространстве существует два особых подпространства: пространство $A_1 = E_1 \times \{\vec{0}\}$ элементов вида $(\vec{x}_1, \vec{0})$ и пространство $A_2 = \{\vec{0}\} \times E_2$ элементов вида $(\vec{0}, \vec{x}_2)$. Эти пространства являются дополнительными в том смысле, что каждый элемент (\vec{x}_1, \vec{x}_2) произведения единственным образом записывается в виде суммы одного элемента из первого пространства и одного элемента из второго, так что $(\vec{x}_1, \vec{x}_2) = (\vec{x}_1, \vec{0}) + (\vec{0}, \vec{x}_2)$.

Когда два векторных подпространства некоторого векторного пространства являются дополнительными, исходное пространство может быть представлено в виде прямой суммы этих подпространств, и эта прямая сумма обозначается знаком \oplus . Поэтому мы можем записать

$$E_1 \times E_2 = E_1 \times \{\vec{0}\} \oplus \{\vec{0}\} \times E_2. \quad (\text{II, 13; 24})$$

Если E_1 и E_2 — два нормированных векторных пространства с нормами $\|\cdot\|_1$ и $\|\cdot\|_2$, то на произведении $E_1 \times E_2$ не существует нормы, которая была бы предпочтительнее других. Можно взять, например, какую-либо одну из следующих эквивалентных норм:

$$\|(\vec{x}_1, \vec{x}_2)\| = \max(\|\vec{x}_1\|_1, \|\vec{x}_2\|_2), \text{ или } \|\vec{x}_1\|_1 + \|\vec{x}_2\|_2, \text{ или } \sqrt{\|\vec{x}_1\|_1^2 + \|\vec{x}_2\|_2^2}. \quad (\text{II, 13; 25})$$

Для каждой из этих норм топология пространства $E_1 \times E_2$ является топологией произведения. Дополнительные подпространства $E_1 \times \{\vec{0}\}$ и $\{\vec{0}\} \times E_2$ замкнуты в E . Если не будет оговорено противное, пространство $E_1 \times E_2$ всегда будет снабжаться

одной из этих норм и один и тот же символ $\| \cdot \|$ будет обозначать нормы в E_1 , E_2 и $E_1 \times E_2$.

Теорема 51. Пусть E_1 и E_2 — два нормированных векторных пространства. Всякое линейное непрерывное отображение u из $E_1 \times E_2$ в векторное нормированное пространство F выражается, и притом единственным образом, в виде

$$u(\vec{x}_1, \vec{x}_2) = u_1(\vec{x}_1) + u_2(\vec{x}_2), \quad (\text{II}, 13; 26)$$

где u_1 (соответственно u_2) является линейным непрерывным отображением E_1 (соответственно E_2) в F . Обратно, если u_1 (соответственно u_2) является линейным непрерывным отображением E_1 (соответственно E_2) в F , то отображение u , определенное формулой (II, 13; 26), является линейным непрерывным отображением $E_1 \times E_2$ в F .

Доказательство. При заданном u определим отображение u_1 следующим образом: значением u_1 на элементе \vec{x}_1 из E_1 будет значение u на элементе $(\vec{x}_1, \vec{0})$ из $E_1 \times E_2$: $u_1(\vec{x}_1) = u(\vec{x}_1, \vec{0})$. Таким же образом определим отображение u_2 из E_2 в F . Отображения u_1 и u_2 , очевидно, являются линейными. Они непрерывны, если на $E_1 \times E_2$ с помощью формулы (II, 13; 25) ввести одну из эквивалентных норм, т. е. если на $E_1 \times E_2$ ввести топологию произведения.

В самом деле, в этом случае имеем неравенство

$$\|u_1(\vec{x}_1)\| = \|u(\vec{x}_1, \vec{0})\| \leq \|u\| \|(\vec{x}_1, \vec{0})\| = \|u\| \|\vec{x}_1\|, \quad (\text{II}, 13; 27)$$

которое доказывает наше утверждение и показывает в то же время, что имеют место неравенства

$$\|u_1\| \leq \|u\|, \quad \|u_2\| \leq \|u\|. \quad (\text{II}, 13; 28)$$

Непосредственно видно, что

$$u(\vec{x}_1, \vec{x}_2) = u(\vec{x}_1, \vec{0}) + u(\vec{0}, \vec{x}_2)^1 = u_1(\vec{x}_1) + u_2(\vec{x}_2), \quad (\text{II}, 13; 29)$$

а это доказывает, что u представлено в указанном выше виде.

Непосредственно видно, что других отображений u_1 и u_2 , обладающих тем же свойством, не существует, потому что, если бы v_1 и v_2 были бы такими линейными отображениями, что

$$u(\vec{x}_1, \vec{x}_2) = v_1(\vec{x}_1) + v_2(\vec{x}_2), \quad \text{то мы имели бы } u(\vec{x}_1, \vec{0}) = v_1(\vec{x}_1),$$

а следовательно, $v_1 = u_1$ и точно так же $v_2 = u_2$.

Обратно, если u_1 (соответственно u_2) является линейным непрерывным отображением E_1 (соответственно E_2) в F , то ото-

¹⁾ В силу линейности u , а также потому, что $(\vec{x}_1, \vec{x}_2) = (\vec{x}_1, \vec{0}) + (\vec{0}, \vec{x}_2)$.

бражение u , определенное формулой (II, 13; 26), очевидно, линейно и, кроме того, непрерывно, ибо

$$\begin{aligned} \|u(\vec{x}_1, \vec{x}_2)\| &\leq \|u_1(\vec{x}_1)\| + \|u_2(\vec{x}_2)\| \leq \|u_1\| \|\vec{x}_1\| + \|u_2\| \|\vec{x}_2\| \leq \\ &\leq (\|u_1\| + \|u_2\|) \max(\|\vec{x}_1\|, \|\vec{x}_2\|) \leq (\|u_1\| + \|u_2\|) \|(\vec{x}_1, \vec{x}_2)\|. \quad (\text{II, 13; 30}) \end{aligned}$$

Отсюда сразу же получаем, что оно удовлетворяет неравенству

$$\|u\| \leq \|u_1\| + \|u_2\|. \quad (\text{II, 13; 31})$$

Мы предоставляем читателю определить структуру нормированного векторного пространства произведения n векторных нормированных пространств E_1, E_2, \dots, E_n и обобщить теорему 51 на линейные непрерывные отображения $E_1 \times E_2 \times \dots \times E_n$ в F .

Билинейные непрерывные отображения произведения нормированных векторных пространств в нормированное векторное пространство

Пусть E, F, G — векторные пространства над полем K . Мы изучили линейные отображения векторного пространства $E \times F$ в пространство G . Теперь рассмотрим совершенно новое понятие *билинейного отображения* $E \times F$ в G .

Отображение u пространства $E \times F$ в пространство G называется *билинейным*, если при фиксировании одной из переменных оно линейно относительно другой переменной. Если мы зафиксируем \vec{x} , то u определит некоторое частное отображение $\vec{y} \rightarrow u(\vec{x}, \vec{y})$ пространства F в G . Это отображение обозначается через $u_{\vec{x}}$, так что

$$u_{\vec{x}}(\vec{y}) = u(\vec{x}, \vec{y}). \quad (\text{II, 13; 32})$$

Удобно также обозначать это отображение через $u(\vec{x}, \vec{y})$, опуская переменную \vec{y} ¹⁾.

Если u — билинейное отображение, то $u_{\vec{x}}$ должно быть линейным отображением F в G ; другими словами, должны выпол-

¹⁾ Если E и F различны, обозначение $u_{\vec{x}}$ для данного элемента \vec{x} из E не вызывает недоразумений. Но если $E=F$ и u является билинейным отображением $E \times E$ в G , то обозначение $u_{\vec{x}}$ для элемента \vec{x} из E не имеет точного смысла, ибо непонятно, о какой из частных функций идет речь — о первой или второй. В этом случае следует писать $u(\vec{x}, \vec{x})$ или $u(\vec{x}, \vec{x})$.

няться равенства:

$$\begin{aligned} u(\vec{x}, \vec{y}_1 + \vec{y}_2) &= u(\vec{x}, \vec{y}_1) + u(\vec{x}, \vec{y}_2), \\ u(\vec{x}, \lambda \vec{y}) &= \lambda u(\vec{x}, \vec{y}). \end{aligned} \quad (\text{II, 13; 33})$$

Аналогично, частное отображение u_y или $u(\cdot, \vec{y}): \vec{x} \rightarrow u(\vec{x}, \vec{y})$ должно быть линейным отображением E в G ; другими словами,

$$\begin{aligned} u(\vec{x}_1 + \vec{x}_2, \vec{y}) &= u(\vec{x}_1, \vec{y}) + u(\vec{x}_2, \vec{y}), \\ u(\lambda \vec{x}, \vec{y}) &= \lambda u(\vec{x}, \vec{y}). \end{aligned} \quad (\text{II, 13; 34})$$

Имея в виду соотношения, относящиеся к сложению, говорят, что u дистрибутивно по отношению к сложению. Из двух предыдущих формул легко получить соотношения:

$$\begin{aligned} u(\vec{x}_1 + \vec{x}_2, \vec{y}_1 + \vec{y}_2) &= u(\vec{x}_1, \vec{y}_1) + u(\vec{x}_1, \vec{y}_2) + u(\vec{x}_2, \vec{y}_1) + u(\vec{x}_2, \vec{y}_2)^1, \\ u(\lambda \vec{x}, \mu \vec{y}) &= \lambda \mu u(\vec{x}, \vec{y}). \end{aligned} \quad (\text{II, 13; 35})$$

Например, скалярное произведение векторов является билинейным отображением $\mathbb{R}^3 \times \mathbb{R}^3$ в \mathbb{R} ; векторное произведение является билинейным отображением $\mathbb{R}^3 \times \mathbb{R}^3$ в \mathbb{R}^3 . Прототипом билинейных отображений являются обычные произведения вещественных или комплексных чисел. Это — билинейные отображения $\mathbb{R} \times \mathbb{R}$ в \mathbb{R} или $\mathbb{C} \times \mathbb{C}$ в \mathbb{C} . Впрочем, любое билинейное отображение $\mathbb{R} \times \mathbb{R}$ в \mathbb{R} имеет вид $(x, y) \rightarrow cxy$, где c — некоторая вещественная постоянная, в то время как линейное отображение $\mathbb{R} \times \mathbb{R}$ в \mathbb{R} имеет вид $(x, y) \rightarrow ax + by$, где a и b — вещественные постоянные.

Теорема 52. *Всякое билинейное отображение произведения нормированных векторных пространств $E \times F$ в нормированное векторное пространство G , непрерывное в нуле пространства $E \times F$, непрерывно всюду (но не равномерно непрерывно). Для того чтобы оно было таким, необходимо и достаточно, чтобы существовала такая постоянная $k \geq 0$, что для всех \vec{x} из E и всех \vec{y} из F имело место неравенство*

$$\|u(\vec{x}, \vec{y})\| \leq k \|\vec{x}\| \|\vec{y}\|^2. \quad (\text{II, 13; 36})$$

¹⁾ Если бы u было линейным на $E \times F$, то в правой части равенства мы бы имели либо $u(\vec{x}_1, \vec{y}_1) + u(\vec{x}_2, \vec{y}_2)$, либо $u(\vec{x}_1, \vec{y}_2) + u(\vec{x}_2, \vec{y}_1)$.

²⁾ Для линейного отображения правая часть имеет, например, такой вид: $k(\|\vec{x}\| + \|\vec{y}\|)$.

Доказательство. Пусть E, F, G — векторные нормированные пространства, и пусть u — билинейное отображение $E \times F$ в G , непрерывное в нуле. Тогда, согласно определению топологического произведения, должно существовать такое число $\eta > 0$, что из неравенств $\|\vec{x}\| \leq \eta, \|\vec{y}\| \leq \eta$ следует неравенство $\|u(\vec{x}, \vec{y})\| \leq 1$. Если мы теперь произведем гомотетию с отношениями λ и $\mu \geq 0$, то из второй формулы (II, 13; 35) получим, что из $\|\vec{x}\| \leq \lambda\eta, \|\vec{y}\| \leq \mu\eta$ следует $\|u(\vec{x}, \vec{y})\| \leq \lambda\mu$.

Так как первые неравенства будут выполнены при $\lambda = \|\vec{x}\|/\eta, \mu = \|\vec{y}\|/\eta$, то мы всегда будем иметь неравенство $\|u(\vec{x}, \vec{y})\| \leq (1/\eta^2) \|\vec{x}\| \|\vec{y}\|$, являющееся формулой (II, 13; 36) с $k = 1/\eta^2$. Обратно, пусть u является билинейным отображением, удовлетворяющим неравенству вида (II, 13; 36). Докажем, что тогда оно непрерывно всюду. Пусть задана точка (\vec{a}, \vec{b}) пространства $E \times F$ и число $\varepsilon > 0$. Имеет место формула

$$u(\vec{x}, \vec{y}) - u(\vec{a}, \vec{b}) = u(\vec{x} - \vec{a}, \vec{y}) + u(\vec{a}, \vec{y} - \vec{b}), \quad (\text{II, 13; 37})$$

из которой вытекает оценка

$$\|u(\vec{x}, \vec{y}) - u(\vec{a}, \vec{b})\| \leq k \|\vec{x} - \vec{a}\| \|\vec{y}\| + k \|\vec{a}\| \|\vec{y} - \vec{b}\|. \quad (\text{II, 13; 38})$$

Положим $\eta_2 = \varepsilon / (2k\|\vec{a}\|)$. Из неравенства $\|\vec{y} - \vec{b}\| \leq \eta_2$ следует, очевидно, неравенство $k\|\vec{a}\| \|\vec{y} - \vec{b}\| \leq \varepsilon/2$. Выберем затем $\eta_1 = \varepsilon / [2k(\|\vec{b}\| + \eta_2)]$. Очевидно, для $\|\vec{y} - \vec{b}\| \leq \eta_2$ имеет место оценка $\|\vec{y}\| \leq \|\vec{b}\| + \eta_2$, откуда получаем, что из $\|\vec{x} - \vec{a}\| \leq \eta_1$ вытекает неравенство $k\|\vec{x} - \vec{a}\| \|\vec{y}\| \leq \varepsilon/2$.

Полученные оценки показывают, что из $\|\vec{x} - \vec{a}\| \leq \eta_1, \|\vec{y} - \vec{b}\| \leq \eta_2$ следует неравенство $\|u(\vec{x}, \vec{y}) - u(\vec{a}, \vec{b})\| \leq \varepsilon$, а это является определением непрерывности отображения u в точке (\vec{a}, \vec{b}) .

Заметим, напротив, что если билинейное отображение u не является тождественно нулевым, то оно не может быть и равномерно непрерывным. В самом деле, предположим, что существует хотя бы одна пара (\vec{a}, \vec{b}) , для которой $u(\vec{a}, \vec{b}) \neq 0$.

Рассмотрим теперь в пространстве $E \times F$ последовательность точек $\vec{X}_n = (\vec{na}, \vec{nb})$ и последовательность точек $\vec{Y}_n = ((n+1/n)\vec{a}, (n+1/n)\vec{b})$. Из неравенства $\|\vec{X}_n - \vec{Y}_n\| \leq (\|\vec{a}\| + \|\vec{b}\|)/n$ следует,

что эта величина стремится к нулю при n , стремящемся к $+\infty$. Тем не менее норма разности $u(\vec{X}_n) - u(\vec{Y}_n) = (n^2 - (n + 1/n)^2) \times \times u(\vec{a}, \vec{b}) = -(2 + 1/n^2) u(\vec{a}, \vec{b})$ при n , стремящемся к $+\infty$, стремится к $2\|u(\vec{a}, \vec{b})\| \neq 0$, а это означает, что функция u равномерно непрерывной быть не может¹⁾.

Определение. Точная нижняя грань таких чисел k , при которых имеет место неравенство (II, 13; 36), называется *нормой билинейного отображения* и обозначается через $\|u\|$. Таким образом, по определению,

$$\|u\| = \sup_{\substack{\vec{x} \neq \vec{0}, \vec{y} \neq \vec{0}}} \frac{\|u(\vec{x}, \vec{y})\|}{\|\vec{x}\| \|\vec{y}\|} = \sup_{\|\vec{x}\| \leq 1, \|\vec{y}\| \leq 1} \|u(\vec{x}, \vec{y})\|. \quad (\text{II, 13; 39})$$

Заметим, что всегда

$$\|u(\vec{x}, \vec{y})\| \leq \|u\| \|\vec{x}\| \|\vec{y}\|. \quad (\text{II, 13; 40})$$

Теорема 53. *Множество $\mathcal{L}_2(E, F; G)$ билинейных отображений произведения нормированных векторных пространств $E \times F$ в нормированное векторное пространство G будет иметь структуру нормированного векторного пространства, если в нем определить норму по формуле (II, 13; 39). Если, кроме того, G является банаховым пространством, то $\mathcal{L}_2(E, F; G)$ также будет банаховым пространством. В частности, выбирая в качестве G поле скаляров, получаем, что векторное пространство билинейных непрерывных форм на произведении нормированных векторных пространств является банаховым пространством.*

Доказательство аналогично доказательствам теорем 48 и 50.

Рассмотрим теперь непрерывное билинейное отображение u пространства $E \times F$ в пространство G . Мы видим, что для фиксированного \vec{x} оно определяет линейное отображение $u_{\vec{x}}$ из F в G . Это отображение, очевидно, непрерывно и его норма $\leq \|u\| \|\vec{x}\|$ в силу (II, 13; 40), а значит, $u_{\vec{x}}$ принадлежит пространству $\mathcal{L}(F; G)$. Тем самым мы нашли отображение, которое каждому элементу \vec{x} из E ставит в соответствие некоторый элемент $u_{\vec{x}}$ из $\mathcal{L}(F; G)$; другими словами, мы определили отображение \vec{U} пространства E в пространство $\mathcal{L}(F; G)$. Оба пространства E и $\mathcal{L}(F; G)$ являются нормированными векторными

¹⁾ Впрочем, известно, что произведение $(x, y) \rightarrow xy$ не является равномерно непрерывным отображением $\mathbb{R} \times \mathbb{R}$ в \mathbb{R} и что функция $x \rightarrow x^2$ не является равномерно непрерывной функцией из \mathbb{R} в \mathbb{R} .

пространствами. Докажем, что U является линейным непрерывным отображением E в $\mathcal{L}(F; G)$ и что норма $\|U\|$ этого линейного отображения есть не что иное, как норма $\|u\|$ рассматриваемого билинейного отображения.

Докажем прежде всего линейность отображения U ; для этого, естественно, не требуется использовать топологию. Нам надо показать, что

$$u_{\vec{x}_1 + \vec{x}_2} = u_{\vec{x}_1} + u_{\vec{x}_2} \quad \text{и} \quad u_{\lambda \vec{x}} = \lambda u_{\vec{x}}. \quad (\text{II, 13; 41})$$

Все выписанные элементы являются элементами пространства $\mathcal{L}(F; G)$, т. е. линейными отображениями F в G . Согласно определению суммы двух линейных отображений или произведения линейного отображения на скаляр, эти равенства означают, что для каждого элемента \vec{y} из F имеют место равенства:

$$u(\vec{x}_1 + \vec{x}_2, \vec{y}) = u(\vec{x}_1, \vec{y}) + u(\vec{x}_2, \vec{y}) \quad \text{и} \quad (\lambda \vec{x}, \vec{y}) = \lambda u(\vec{x}, \vec{y}). \quad (\text{II, 13; 42})$$

Но эти равенства справедливы, ибо они означают лишь то, что частное отображение $u_{\vec{y}}$ пространства F в пространство G для каждого фиксированного \vec{y} линейно.

Докажем теперь непрерывность отображения U . Норма линейного отображения $U(\vec{x}) = u_{\vec{x}}$ пространства F в G не превосходит величины $\|u\| \|\vec{x}\|$. Из этого неравенства следует, что U является линейным непрерывным отображением E в $\mathcal{L}(F; G)$ и что $\|U\| \leq \|u\|$. Так как $\|U(\vec{x})\| \leq \|U\| \|\vec{x}\|$, то $\|u(\vec{x}, \vec{y})\| = \|U(\vec{x})\| \|\vec{y}\| \leq \|U\| \|\vec{x}\| \|\vec{y}\|$, откуда $\|u\| \leq \|U\|$, а, значит, $\|u\| = \|U\|$.

Обратно, будем исходить из линейного непрерывного отображения U из E в $\mathcal{L}(F; G)$. Для каждого \vec{x} из E , $U(\vec{x})$ является элементом $\mathcal{L}(F; G)$, т. е. линейным непрерывным отображением F в G , и, следовательно, для каждого элемента \vec{y} из F $U(\vec{x}) \cdot \vec{y}$ является элементом G . Если мы положим $u(\vec{x}, \vec{y}) = U(\vec{x}) \cdot \vec{y}$, то получим, что $u: (x, y) \rightarrow u(x, y)$ есть отображение $E \times F$ в G . Без труда проверяется, что оно билинейно, а из неравенств $\|u(x, y)\| \leq \|U(x)\| \|y\| \leq \|U\| \|x\| \|y\|$, $\|u\| \leq \|U\|$ следует его непрерывность.

Частное отображение u_x , соответствующее u , для каждого фиксированного \vec{x} является не чем иным, как $U(\vec{x})$, а линейное

непрерывное отображение E в $\mathcal{L}(F; G)$, соответствующее u , есть не что иное, как исходное отображение U . Установленное соответствие между отображениями u и U является биекцией пространства $\mathcal{L}_2(E, F; G)$ билинейных непрерывных отображений из $E \times F$ в G на пространство $\mathcal{L}(E; \mathcal{L}(F; G))$ линейных непрерывных отображений E в $\mathcal{L}(F; G)$, и эта биекция сохраняет векторные структуры и нормы. Существует, естественно, аналогичная биекция $\mathcal{L}_2(E, F; G)$ на $\mathcal{L}(F; \mathcal{L}(E, G))$.

Теорема 54. Пусть E и F — нормированные векторные пространства. Тогда отображение $(u, x) \rightarrow u \cdot x$ пространства $\mathcal{L}(E; F) \times E$ в пространство F является билинейным непрерывным и имеет норму, равную 1 (за исключением того случая, когда E или F сводится к нулю).

Если E, F, G — три нормированных векторных пространства, то отображение $(u, v) \rightarrow v \circ u$ пространства $\mathcal{L}(E; F) \times \mathcal{L}(F; G)$ в пространство $\mathcal{L}(E; G)$ билинейно, непрерывно и имеет норму, равную 1 (за исключением того случая, когда E, F или G сводятся к нулю).

Доказательство. Билинейность рассматриваемого отображения очевидным образом вытекает из определения суммы двух линейных отображений и умножения линейного отображения на скаляр¹⁾. Непрерывность этих билинейных отображений и тот факт, что их норма ≤ 1 , сразу вытекает из оценок $\|u \cdot x\| \leq \|u\| \|x\|$, $\|v \circ u\| \leq \|v\| \|u\|$ (формулы (II, 13; 6) и (II, 13; 7)). Далее, при заданном u и любом заданном $\varepsilon > 0$, согласно определению нормы u , существует по крайней мере один элемент $x \neq \vec{0}$ из E , такой, что $\|u \cdot x\| \geq (1 - \varepsilon) \|u\| \|x\|$, откуда следует, что норма первого билинейного отображения $\geq 1 - \varepsilon$. Так как ε может быть выбрано сколь угодно малым, то эта норма ≥ 1 , и, следовательно, она равна 1.

Тот факт, что норма второго билинейного отображения также равна 1, мы примем без доказательства.

Мультилинейные непрерывные отображения

Пусть E_1, E_2, \dots, E_n, F — нормированные векторные пространства над одним и тем же полем \mathbb{K} (полем вещественных или комплексных чисел). Отображение u пространства

¹⁾ Билинейность этих отображений позволила ввести мультипликативные обозначения ux , vu и u^n на стр. 115.

$E_1 \times E_2 \times \dots \times E_n$ в пространство F называется *мультилинейным* (или n -линейным), если при фиксировании $n-1$ переменных в произвольных $n-1$ векторных пространствах это отображение является линейным относительно n -й переменной.

Например, обычное произведение $(x_1, x_2, \dots, x_n) \rightarrow x_1 x_2 \dots x_n$ является n -линейным отображением \mathbb{K}^n в \mathbb{K} . В пространстве \mathbb{R}^3 смешанное произведение трех векторов определяет трилинейное отображение $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ в \mathbb{R} .

Пространство n -линейных непрерывных отображений из $E_1 \times E_2 \times \dots \times E_n$ в F можно наделить структурой векторного нормированного пространства. Его тогда обозначают через $\mathcal{L}_n(E_1, E_2, \dots, E_n; F)$. Если F является банаевым пространством, то и рассматриваемое пространство также будет банаевым.

Изложенное перед теоремой 54 может быть обобщено следующим образом:

Теорема 54₂. Между пространствами

$$\mathcal{L}_n(E_1, E_2, \dots, E_n; F), \mathcal{L}_1(E_1; \mathcal{L}_{n-1}(E_2, \dots, E_n; F)),$$

$$\mathcal{L}_2(E_1, E_2; \mathcal{L}_{n-2}(E_3, \dots, E_n; F)), \dots, \mathcal{L}_{n-1}(E_1, E_2, \dots, E_{n-1}; \mathcal{L}_1(E_n; F))$$

и пространствами, полученными перестановкой индексов, существуют биекции, сохраняющие векторную структуру и норму. Элемент U в $\mathcal{L}_p(E_1, E_2, \dots, E_p; \mathcal{L}_{n-p}(E_{p+1}, \dots, E_n; F))$, соответствующий элементу $u \in \mathcal{L}_n(E_1, E_2, \dots, E_n; F)$, задается формулой

$$U(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) \cdot (\vec{x}_{p+1}, \dots, \vec{x}_n) = u(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p, \vec{x}_{p+1}, \dots, \vec{x}_n). \quad (\text{II, 13; 43})$$

Алгебры. Нормированные алгебры

Алгеброй \mathcal{A} над полем \mathbb{K} называется векторное пространство над полем \mathbb{K} , в котором определено билинейное отображение $\mathcal{A} \times \mathcal{A}$ в \mathcal{A} , называемое умножением и обозначаемое через $(\vec{x}, \vec{y}) \rightarrow \vec{xy}$. Это умножение ассоциативно: $(\vec{xy})z = \vec{x}(yz)$ и имеет единицу $\vec{I} \neq \vec{0}$ ($\vec{x}\vec{I} = \vec{Ix}$ для любого \vec{x}). Поле \mathbb{K} само является алгеброй. Нормированной алгеброй называется алгебра, представляющая собой такое нормированное векторное пространство, в котором

$$\|\vec{I}\| = 1, \quad \|\vec{xy}\| \leq \|\vec{x}\| \|\vec{y}\|. \quad (\text{II, 13; 44})$$

Полная нормированная алгебра называется *банаевой*. Пусть E — нормированное векторное пространство. Простран-

ство $\mathcal{L}(E; E)$ будет нормированной алгеброй, если в нем определить умножение элементов через их композицию: $(u, v) \rightarrow u \circ v$. Эта алгебра банахова, если пространство E банахово.

§ 14. РЯДЫ В НОРМИРОВАННЫХ ВЕКТОРНЫХ ПРОСТРАНСТВАХ

Пусть задана некоторая последовательность $\vec{u}_0, \vec{u}_1, \vec{u}_2, \dots$ элементов нормированного векторного пространства E . Сумма

$$\vec{S}_n = \vec{u}_0 + \vec{u}_1 + \dots + \vec{u}_n = \sum_{0 \leq m \leq n} \vec{u}_m \quad (\text{II}, 14; 1)$$

называется *частной суммой с индексом n* ряда $\sum_{k=0}^{\infty} \vec{u}_k$. Частная

сумма \vec{S}_n является элементом пространства E . Говорят, что ряд $\sum_{k=0}^{\infty} \vec{u}_k$ *сходится и имеет сумму \vec{S}* , если последовательность

\vec{S}_n сходится и имеет своим пределом \vec{S} ¹⁾. Понятие ряда сводится, таким образом, к понятию последовательности. Впрочем, верно и обратное: последовательность в нормированном векторном пространстве можно выразить через ряд. Последовательность $\vec{x}_0, \vec{x}_1, \vec{x}_2, \dots$ сходится и имеет пределом число \vec{a}

тогда и только тогда, когда сходится ряд $\vec{x}_0 + (\vec{x}_1 - \vec{x}_0) + (\vec{x}_2 - \vec{x}_1) + \dots + (\vec{x}_n - \vec{x}_{n-1}) + \dots$ и имеет сумму, равную \vec{a} .

Если пространство E конечномерно и если в нем выбран некоторый базис, то, вспоминая результаты, полученные нами ранее для последовательностей, можно утверждать, что ряд $\sum_n \vec{u}_n$ является сходящимся и имеет своей суммой \vec{S} тогда и только тогда, когда каждый ряд из координат $\sum_n (u_n)_i$ сходится к не-

которой сумме S_i , где $(u_n)_i$ есть i -я координата элемента \vec{u}_n , а S_i есть i -я координата \vec{S} .

Преимущество банаховых пространств заключается в том, что в них, как и в поле комплексных чисел, можно судить о сходимости ряда, не зная заранее о существовании суммы \vec{S} .

Для того чтобы ряд в банаховом пространстве E был сходящимся, необходимо и достаточно, чтобы он удовлетворял кри-

¹⁾ Сходимость или расходимость ряда сохраняется, если одну норму заменяют другой, ей эквивалентной, ибо сходимость является топологическим свойством.

терию Коши: при m и n , стремящихся к $+\infty$, $\|\vec{S}_n - \vec{S}_m\|$ стремится к 0, или же

$$(\forall \varepsilon > 0) (\exists n \in \mathbb{N}) (\forall m \geq n) (\forall k \in \mathbb{N}): \|\vec{u}_m + \vec{u}_{m+1} + \dots + \vec{u}_{m+k}\| \leq \varepsilon. \quad (\text{II}, 14; 2)$$

Отсюда получается основная теорема, дающая наиболее важный критерий сходимости векторных рядов:

Теорема 55. Пусть $\sum_{n=0}^{\infty} \vec{u}_n$ — ряд из элементов банаухова пространства E . Если ряд, составленный из норм, $\sum \|\vec{u}_n\|$ сходится, то и сам ряд также сходится и, кроме того,

$$\|\sum \vec{u}_n\| \leq \sum \|\vec{u}_n\|. \quad (\text{II}, 14; 3)$$

Доказательство. Проверим выполнение критерия Коши. Так как ряд, составленный из норм, по условию сходится, то для любого $\varepsilon > 0$ можно найти такое n , что для всех $m \geq n$ и любого k $\|\vec{u}_m\| + \|\vec{u}_{m+1}\| + \dots + \|\vec{u}_{m+k}\| \leq \varepsilon$. Отсюда тем более для $m \geq n$ и любого k имеем, что $\|\vec{u}_m + \vec{u}_{m+1} + \dots + \vec{u}_{m+k}\| \leq \varepsilon$, т. е. что заданный в E ряд удовлетворяет условию Коши, а так как пространство E предполагалось полным, то рассматриваемый ряд сходится. Кроме того, для конечного числа членов имеет место неравенство $\|\vec{S}_n\| \leq \sum_{0 \leq m \leq n} \|\vec{u}_m\|$, откуда, переходя к пределу при n , стремящемся к $+\infty$, и учитывая, что сходимость \vec{S}_n к \vec{S} влечет за собой сходимость $\|\vec{S}_n\|$ к $\|\vec{S}\|$ (теорема 9), получаем искомое неравенство (II, 14; 3).

Определение. Говорят, что ряд элементов банаухова пространства E *нормально сходится*, или *абсолютно сходится*¹⁾, если сходится ряд из норм, являющийся числовым рядом с положительными членами. Ряд называется *условно сходящимся*, если он сходится, не являясь нормально сходящимся.

Полнота пространства E является весьма существенным фактом. Впрочем, можно доказать следующее обратное утверждение:

Теорема 56. Пусть E — нормированное векторное пространство. Если в этом пространстве сходится любой ряд $\sum \vec{u}_n$, $\vec{u}_n \in E$, у которого сходится ряд из норм $\sum \|\vec{u}_n\|$, то пространство E полно.

1) В силу теоремы 12, нормальная сходимость сохраняется, когда норма E заменяется на другую, ей эквивалентную.

Доказательство. Нам надо доказать, что любая последовательность Коши из E является сходящейся.

Пусть $\vec{u}_0, \vec{u}_1, \vec{u}_2, \dots$ — некоторая последовательность Коши. Тогда, каким бы ни было целое число $k \geq 0$, всегда можно найти такое целое число p_k , что из $m \geq p_k, n \geq p_k$ следует неравенство $\|\vec{u}_m - \vec{u}_n\| \leq 1/2^k$.

Будем выбирать одно за другим целые числа p_k таким образом, чтобы последовательность p_k была строго возрастающей, и рассмотрим ряд $\vec{u}_{p_0} + (\vec{u}_{p_1} - \vec{u}_{p_0}) + (\vec{u}_{p_2} - \vec{u}_{p_1}) + \dots$. Ряд, составленный из его норм, мажорируется числовым рядом $\|\vec{u}_{p_0}\| + 1 + 1/2 + 1/4 + \dots$, сходящимся в силу известных свойств числовых рядов с положительными членами. Но тогда, в силу предположения относительно E , рассматриваемый ряд сходится, что означает сходимость подпоследовательности \vec{u}_{p_n} . Рассматриваемая последовательность Коши \vec{u}_n имеет, таким образом, сходящуюся подпоследовательность. Согласно следствию 2 из теоремы 40, она должна сходиться, а, значит, E полно.

Перестановка членов ряда

Пусть $\vec{u}_0 + \vec{u}_1 + \vec{u}_2 + \dots$ — ряд, составленный из векторов E . Изменить порядок членов последовательности означает рассмотреть биекцию $n \rightarrow p_n$ из \mathbb{N} на \mathbb{N} и заменить данный ряд новым $\vec{u}_{p_0} + \vec{u}_{p_1} + \vec{u}_{p_2} + \dots$.

Теорема 57. Если ряд $\vec{u}_0 + \vec{u}_1 + \vec{u}_2 + \dots$ в нормированном векторном пространстве E сходится и вместе с ним сходится ряд, составленный из норм $\|\vec{u}_0\| + \|\vec{u}_1\| + \|\vec{u}_2\| + \dots$, то перестановка его членов не отражается на сходимости ряда и, кроме того, не меняет его суммы.

Доказательство. Пусть \tilde{S} — сумма данного ряда. При любом заданном $\varepsilon > 0$ существует m , такое, что $\sum_{n>m} \|\vec{u}_n\| \leq \varepsilon/2$.

Далее, существует такое целое число m' , что множество целых чисел $\{p_0, p_1, p_2, \dots, p_{m'}\}$ содержит множество $\{0, 1, 2, \dots, m\}$. Тогда для $n' \geq m'$ частная сумма $\vec{u}_{p_0} + \vec{u}_{p_1} + \dots + \vec{u}_{p_{n'}}$ преобразованного ряда равна частной сумме $\vec{u}_0 + \vec{u}_1 + \dots + \vec{u}_m$, дополненной на некоторую конечную сумму членов с индексами $> m$.

Норма суммы этих дополнительных членов может быть промажорирована суммой $\sum_{m < n \leq p} \|\vec{u}_n\|$, где p — наибольшее из целых чисел p_0, p_1, \dots, p_n ; тем более она не превосходит $\sum_{n > m} \|\vec{u}_n\|$, а, значит, и $\varepsilon/2$. Таким образом, для $n' \geq m'$ имеем:

$$\|(\vec{u}_{p_0} + \vec{u}_{p_1} + \dots + \vec{u}_{p_{n'}}) - (\vec{u}_0 + \vec{u}_1 + \dots + \vec{u}_m)\| \leq \frac{\varepsilon}{2}. \quad (\text{II, 14; 4})$$

Поскольку

$$\|(\vec{u}_0 + \vec{u}_1 + \dots + \vec{u}_m) - \vec{S}\| = \left\| \sum_{n > m} \vec{u}_n \right\| \leq \sum_{n > m} \|\vec{u}_n\| \leq \frac{\varepsilon}{2}, \quad (\text{II, 14; 5})$$

то окончательно для $n' \geq m'$ получаем неравенство

$$\|(\vec{u}_{p_0} + \vec{u}_{p_1} + \dots + \vec{u}_{p_{n'}}) - \vec{S}\| \leq \varepsilon, \quad (\text{II, 14; 6})$$

доказывающее, что преобразованный ряд сходится и имеет сумму \vec{S} .

Следствие. Если ряд с вещественными положительными членами сходится, то его сходимость и сумма сохраняются при любой перестановке членов.

Определение. Пусть I — счетное множество индексов и $(\vec{u}_i)_{i \in I}$ — семейство векторов некоторого нормированного векторного пространства E . Говорят, что ряд $\sum_{i \in I} \vec{u}_i$ безусловно сходится и имеет сумму \vec{S} , если, какова бы ни была биекция $n \rightarrow p_n$ множества \mathbb{N} на множество I , обычный ряд $\sum_{n=0}^{\infty} \vec{u}_{p_n}$ сходится и имеет суммой вектор \vec{S} , не зависящий от выбора биекции.

Пусть u_i — вещественные числа ≥ 0 . Обозначим через S точную верхнюю грань в $\bar{\mathbb{R}}$ (конечную или равную $+\infty$) сумм $S_J = \sum_{i \in J} u_i$, соответствующих всем конечным подмножествам J множества индексов I . Если S конечна, то ряд безусловно сходится и имеет сумму S . В самом деле, пусть $n \rightarrow p_n$ — некоторая биекция \mathbb{N} на I . Тогда $u_{p_0} + u_{p_1} + \dots + u_{p_n} \leq S$. С другой стороны, при заданном $\varepsilon > 0$ существует конечное подмножество J из I , такое, что $S_J \geq S - \varepsilon$. Если m является наименьшим из целых чисел, при которых множество $\{p_0, p_1, \dots, p_m\}$ содержит множество J , то для $n \geq m$ $u_{p_0} + u_{p_1} + \dots + u_{p_n} \geq S_J \geq S - \varepsilon$.

$\geq S_J \geq S - \varepsilon$. Следовательно, ряд $u_{p_0} + u_{p_1} + \dots + u_{p_n} + \dots$ сходится и имеет сумму S . Если $S = +\infty$, то ряд $\sum_{n=0}^{\infty} u_{p_n}$ расходится для любой биекции $n \rightarrow p_n$ множества \mathbb{N} на множество I , ибо если бы он был сходящимся для некоторой частной биекции и имел суммой число σ , то тогда для любого J имели бы место неравенства $S_J \leq \sigma$ и $S \leq \sigma$, что противоречит нашему предположению. В этом случае можно сказать, что ряд безусловно расходится, и писать $\sum_{i \in I} u_i = +\infty$.

Для рядов с положительными членами нет необходимости говорить о безусловной сходимости. Она будет такой автоматически, и поэтому имеет смысл говорить лишь об обычной сходимости.

Если J является бесконечным подмножеством I , то всегда имеют место следующие соотношения: $S_J = \sum_{i \in J} u_i \leq S = \sum_{i \in I} u_i$.

Теорема 57 показывает, что если $\sum_{i \in I} \vec{u}_i$ является рядом из векторов банахова пространства E и если ряд из норм его членов $\sum_{i \in I} \| \vec{u}_i \|$ сходится, то сам ряд является безусловно сходящимся. К этому результату можно добавить следующее:

1°) Если речь идет о вещественных числовых рядах, то имеется обратное утверждение к предыдущему результату:

Числовой вещественный ряд может быть безусловно сходящимся только в том случае, когда он абсолютно сходится, т. е. когда сходится ряд $\sum_{i \in I} |u_i|$.

В самом деле, рассмотрим ряд из вещественных чисел. Выберем некоторую биекцию \mathbb{N} на I , или, что по существу то же самое, предположим, что $I = \mathbb{N}$. Члены ряда будем выбирать следующим образом.

Выпишем сначала из данного ряда подряд все положительные члены до тех пор, пока их сумма не превзойдет 1. Затем возьмем первый отрицательный член данного ряда. После этого из оставшихся положительных членов выберем подряд столько, чтобы получить сумму, большую 2. Затем снова добавим следующее отрицательное число из данного ряда (если оно есть) и будем продолжать добавлять положительные числа, пока сумма не превзойдет 3 и т. д.

Если подпоследовательность, составленная из положительных членов ряда, расходится, то проводимое нами построение можно продолжать неограниченно. Получаемый такой перестановкой ряд расходится, так как при любом $n \geq 0$ существуют его

частные суммы, превосходящие n . Поэтому, если ряд, составленный из положительных членов, расходится, то данный ряд безусловно сходиться не может. Точно такие же рассуждения можно провести и в том случае, когда расходится подпоследовательность, составленная из отрицательных членов ряда. Таким образом, данный ряд может безусловно сходиться только тогда, когда сходятся подпоследовательности положительных и отрицательных членов ряда, т. е. когда данный ряд абсолютно сходится.

2°) Такое же обратное утверждение имеет место, если E — нормированное конечномерное векторное пространство. Поскольку векторное пространство размерности k над полем \mathbb{C} является пространством размерности $2k$ над полем \mathbb{R} , то можно считать, что рассматривается конечномерное векторное пространство над полем \mathbb{R} . Поскольку, кроме того, результат не зависит от выбранной нормы (в силу эквивалентности всех норм (теорема 13)), то можно считать, что в пространстве выбран базис, и отождествить E с пространством \mathbb{R}^m , и в этом случае в качестве нормы взять $\|(x_1, x_2, \dots, x_m)\| = \sum_{1 \leq i \leq m} |x_i|$ на \mathbb{R}^m .

Если теперь ряд безусловно сходится, то непосредственно видно, что каждая из компонент ряда представляет собой безусловно сходящийся ряд вещественных чисел и, следовательно, согласно результату, полученному для рядов с вещественными членами, каждая из компонент ряда должна быть абсолютно сходящейся. Окончательно, по определению выбранной нормы, получаем, что данный ряд абсолютно сходится.

3°) Напротив, с помощью весьма тонких рассуждений можно доказать, что в любом бесконечномерном нормированном векторном пространстве E можно найти безусловно сходящийся ряд, не являющийся абсолютно сходящимся¹⁾.

Если ряд безусловно сходится, то его сходимость и сумма не зависят от порядка членов. Естественно предположить, что, как и в случае рядов с положительными членами, можно дать определение безусловно сходящегося ряда и его суммы, не выбирая определенную биекцию \mathbb{N} на I .

В самом деле, можно доказать следующее:

1) Пусть $\sum_{n=0}^{\infty} \vec{u}_n$ — ряд из векторов некоторого банахова пространства.

Тогда для него возможны следующие случаи:

- a) ряд сходится абсолютно и, следовательно, безусловно;
- b) ряд сходится безусловно, но не абсолютно (только тогда, когда E бесконечномерно);
- c) ряд сходится, но не безусловно;
- d) ряд расходится.

Теорема 58. Ряд $\sum_{i \in I} \vec{u}_i$ безусловно сходится и имеет сумму \vec{S} тогда и только тогда, когда для любого $\varepsilon > 0$ существует такое конечное подмножество J индексов из множества индексов I , что для любого конечного подмножества K множества I , содержащего J , имеет место неравенство

$$\|\vec{S}_K - \vec{S}\| = \left\| \sum_{i \in K} \vec{u}_i - \vec{S} \right\| \leq \varepsilon. \quad (\text{II}, 14; 7)$$

Читатель может доказать эту теорему в качестве упражнения.

Суммирование по блокам безусловно сходящегося ряда

Теорема 59. Предположим, что счетное множество индексов I является объединением семейства непустых и непересекающихся подмножеств I_α , $\alpha \in A$:

$$I = \bigcup_{\alpha \in A} I_\alpha, \quad I_\alpha \cap I_\beta = \emptyset \quad \text{для } \alpha \neq \beta^1).$$

Тогда, если ряд $\sum_{i \in I} \vec{u}_i$ из векторов банахова пространства E безусловно сходится к \vec{S} , то каждый из частных рядов $\sum_{i \in I_\alpha} \vec{u}_i$ безусловно сходится; если их суммы мы обозначим через \vec{S}_α , то ряд $\sum_{\alpha \in A} \vec{S}_\alpha$ будет безусловно сходиться и иметь сумму \vec{S} . Другими словами, имеет место формула суммирования по блокам:

$$\sum_{i \in I} \vec{u}_i = \sum_{\alpha \in A} \left(\sum_{i \in I_\alpha} \vec{u}_i \right). \quad (\text{II}, 14; 8)$$

Доказательство. Ограничимся доказательством теоремы для случая, когда данный ряд абсолютно сходится. Положим $M = \sum_{i \in I} \|\vec{u}_i\|$. Тогда, очевидно, $\sum_{i \in I_\alpha} \|\vec{u}_i\| \leq M$; это означает, что каждый из рядов $\sum_{i \in I_\alpha} \vec{u}_i$ является безусловно сходящимся (теоремы 55 и 57). Следовательно, если рассматривать некоторое конечное множество B элементов из множества индексов A , то мы получим неравенство $\sum_{\alpha \in B} \|\vec{S}_\alpha\| \leq \sum_{i \in \bigcup_{\alpha \in B} I_\alpha} \|\vec{u}_i\| \leq M$,

¹⁾ Множество A конечно или счетно, и каждое множество I_α конечно или счетно.

которое показывает, что ряд $\sum_{a \in A} \vec{S}_a$ является также безусловно сходящимся. Остается доказать формулу $\sum_{a \in A} \vec{S}_a = \vec{S}$.

Прежде всего при заданном $\varepsilon > 0$ можно найти такое конечное подмножество J индексов из I , при котором имеет место неравенство

$$\sum_{i \in C_J} \| \vec{u}_i \| \leq \frac{\varepsilon}{2}, \quad a, \text{ следовательно, } \| \vec{S} - \vec{S}_J \| \leq \frac{\varepsilon}{2}.$$

Пусть теперь $n \rightarrow p_n$ — некоторая биекция множества \mathbb{N} на множество A в случае, когда A счетно, или же подмножества $\{0, 1, \dots, m\}$ множества \mathbb{N} на A в случае, если A конечно. В обоих случаях существует такое целое m , что объединение множеств $I_{p_0}, I_{p_1}, \dots, I_{p_m}$ содержит J . Тогда для $n \geq m$ частная сумма $\vec{S}_{p_0} + \vec{S}_{p_1} + \dots + \vec{S}_{p_n}$ является суммой \vec{S}_J и некоторого безусловно сходящегося ряда, образованного из тех членов \vec{u}_i , индексы i которых принадлежат дополнению J .

При этом имеем неравенство

$$\| (\vec{S}_{p_0} + \vec{S}_{p_1} + \dots + \vec{S}_{p_n}) - \vec{S}_J \| \leq \frac{\varepsilon}{2}, \quad (II, 14; 9)$$

из которого вытекает неравенство

$$\| (\vec{S}_{p_0} + \vec{S}_{p_1} + \dots + \vec{S}_{p_n}) - \vec{S} \| \leq \varepsilon, \quad (II, 14; 10)$$

доказывающее исходную формулу $\sum_{a \in A} \vec{S}_a = \vec{S}$.

Замечания. 1°) Если $I = \mathbb{N}$ и если заданный ряд $\sum_{n=0}^{\infty} \vec{u}_n$ сходится, но не безусловно, то такое утверждение будет неверным. Например, если u_n — вещественные числа и если ряд $\sum_{n=0}^{\infty} u_n$ является условно сходящимся, то частный ряд, составленный из членов ≥ 0 , и частный ряд, составленный из членов < 0 , расходятся!

2°) Рассмотрим обратное утверждение к этой теореме. Пусть $(I_a)_{a \in A}$ — некоторое разбиение множества индексов I на непересекающиеся подмножества. Предположим, что для каждого a из A ряд $\sum_{i \in I_a} \vec{u}_i$ безусловно сходится и имеет сумму \vec{S}_a ; с другой стороны, предположим, что ряд $\sum_{a \in A} \vec{S}_a$ является

безусловно сходящимся и имеет сумму \vec{S} . Можем ли мы тогда утверждать, что исходный ряд $\sum_{i \in I} \vec{u}_i$ также безусловно сходится и имеет сумму \vec{S} ?

Если он безусловно сходится, то из теоремы следует, что его сумма заведомо равна \vec{S} , но *ничто не говорит о том, что этот ряд должен быть безусловно сходящимся*. В самом деле, достаточно рассмотреть пример, в котором I равно множеству \mathbb{Z} всех целых чисел, A является множеством \mathbb{N} всех целых чисел ≥ 0 и в котором для каждого α , т. е. для каждого целого $n \geq 0$, I_α совпадает с множеством $\{+n, -n\}$ ¹⁾. Тогда если мы рассмотрим ряд, в котором $u_i = i$, то увидим, что формула $\sum_{a \in A} \left(\sum_{i \in I_a} u_i \right) = \sum_{n \in \mathbb{N}} (n - n)$ дает безусловно сходящийся ряд с суммой, равной 0, в то время как ряд $\sum_{i \in \mathbb{Z}} i$ не является безусловно сходящимся ($\sum_{i \in \mathbb{Z}} |i| = +\infty$).

3°) Напротив, если все u_i — вещественные положительные числа и если $\sum_{a \in A} \left(\sum_{i \in I_a} u_i \right)$ дает в результате S , то ряд $\sum_{i \in I} u_i$, конечно, сходится (потому что все его частные суммы, составленные из конечного числа членов, ограничены) и, следовательно, имеет ту же сумму S .

Если в случае рядов с положительными членами договориться через $+\infty$ обозначать сумму расходящегося ряда, то без всяких дополнительных предположений имеет место равенство (II, 14; 8), обе части которого либо являются конечными числами, либо равны $+\infty$.

4°) Отсюда, наконец, следует, что если E является пространством Банаха и если сходится ряд $\sum_{a \in A} \left(\sum_{i \in I_a} \|\vec{u}_i\| \right)$, то можно утверждать, что ряд $\sum_{i \in I} \|\vec{u}_i\|$ сходится, и, следовательно, ряд $\sum_{i \in I} \vec{u}_i$ является безусловно сходящимся, а, значит, применима теорема 59 и имеет место формула (II, 14; 8).

Действие линейного непрерывного отображения на ряд

Теорема 60. Пусть E и F — два нормированных векторных пространства и L — линейное непрерывное отображение E

1) Сводящимся к одному элементу, если $n = 0$.

в F . Если $\sum_{n=0}^{\infty} \vec{u}_n$ является сходящимся рядом в E , то ряд

$\sum_{n=0}^{\infty} L \cdot \vec{u}_n$ сходится в F ; при этом

$$\sum_{n=0}^{\infty} (L \cdot \vec{u}_n) = L \cdot \left(\sum_{n=0}^{\infty} \vec{u}_n \right). \quad (\text{II, 14; 11})$$

Такое же утверждение имеет место для безусловной и абсолютной сходимости в случае, когда E и F являются пространствами Банаха. Кроме того, справедливо неравенство

$$\sum_{n=0}^{\infty} \|L \cdot \vec{u}_n\| \leq \|L\| \cdot \sum_{n=0}^{\infty} \|\vec{u}_n\|. \quad (\text{II, 14; 12})$$

Доказательство. В силу линейности L ,

$$L \cdot \left(\sum_{n \leq m} \vec{u}_n \right) = \sum_{n \leq m} (L \cdot \vec{u}_n). \quad (\text{II, 14; 13})$$

При m , стремящемся к $+\infty$, сумма $\sum_{n \leq m} \vec{u}_n$ сходится к сумме \vec{S} ; следовательно, в силу непрерывности L , левая часть последнего равенства сходится к $L \cdot \vec{S}$. К тому же пределу стремится правая часть, а это означает, что ряд $\sum_{n=0}^{\infty} L \cdot \vec{u}_n$ сходится и имеет сумму $L \cdot \vec{S}$, откуда мы получаем (II, 14; 11).

Из (II, 13; 6) следует, что

$$\sum_{n \leq m} \|L \cdot \vec{u}_n\| \leq \|L\| \sum_{n \leq m} \|\vec{u}_n\|, \quad (\text{II, 14; 14})$$

откуда, переходя к пределу при m , стремящемся к $+\infty$, мы получаем (II, 14; 12). (Написанные выражения могут быть конечными или равными $+\infty$.) Таким образом, если $\sum_{n=0}^{\infty} \vec{u}_n$ абсолютно сходится, то то же самое будет верно и для $\sum_{n=0}^{\infty} L \cdot \vec{u}_n$.

Произведение двух числовых рядов. Применение билинейного непрерывного отображения к двум рядам

Напомним, что самым простым билинейным непрерывным отображением является произведение — непрерывное билинейное отображение $\mathbb{R} \times \mathbb{R}$ в \mathbb{R} .

В общем курсе математического анализа показывается, что для представления произведения двух рядов в виде ряда необходимо предполагать, что ряды абсолютно сходятся. По этой причине здесь также не будет теорем, относящихся к просто сходящимся рядам.

Теорема 61. Пусть B — билинейное непрерывное отображение произведения $E \times F$ банаховых пространств в банахово пространство G . Пусть $\sum_{i \in I} \vec{u}_i$ и $\sum_{j \in J} \vec{v}_j$ — два абсолютно сходящихся ряда с элементами из E и F . Тогда ряд $\sum_{(i, j) \in I \times J} B(\vec{u}_i, \vec{v}_j)$ абсолютно сходится в G и имеет место формула

$$\sum_{(i, j) \in I \times J} B(\vec{u}_i, \vec{v}_j) = B\left(\sum_{i \in I} \vec{u}_i, \sum_{j \in J} \vec{v}_j\right). \quad (\text{II}, 14; 15)$$

Доказательство. Обозначим через \vec{U} и \vec{V} соответственно суммы $\sum_{i \in I} \vec{u}_i$ и $\sum_{j \in J} \vec{v}_j$. Пусть K — некоторое конечное подмножество $I \times J$. Существуют такие конечные подмножества L из I и M из J , что $L \times M$ содержит K и при этом имеют место оценки:

$$\begin{aligned} & \sum_{(i, j) \in K} \|B(\vec{u}_i, \vec{v}_j)\| \leq \sum_{(i, j) \in L \times M} \|B\| \|\vec{u}_i\| \|\vec{v}_j\| = \\ & = \|B\| \left(\sum_{i \in L} \|\vec{u}_i\| \right) \left(\sum_{j \in M} \|\vec{v}_j\| \right) \leq \|B\| \left(\sum_{i \in I} \|\vec{u}_i\| \right) \left(\sum_{j \in J} \|\vec{v}_j\| \right). \quad (\text{II}, 14; 16) \end{aligned}$$

Эти неравенства показывают, что все частные суммы конечного числа элементов ряда $\sum_{(i, j) \in I \times J} \|B(\vec{u}_i, \vec{v}_j)\|$ ограничены и что, следовательно, ряд $\sum_{(i, j) \in I \times J} B(\vec{u}_i, \vec{v}_j)$ абсолютно сходится. Мы можем применить к нему теорему 59 о суммировании по блокам. Произведение $I \times J$ допускает разбиение на непересекающиеся части — объединение множеств $\{i\} \times J$, когда i пробегает I . При этом

$$\sum_{(i, j) \in I \times J} B(\vec{u}_i, \vec{v}_j) = \sum_{i \in I} \left(\sum_{j \in J} B(\vec{u}_i, \vec{v}_j) \right).$$

Рассмотрим сначала сумму $\sum_{j \in J} B(\vec{u}_i, \vec{v}_j)$ для фиксированного значения i . При фиксированном \vec{u}_i отображение $\vec{v} \rightarrow \rightarrow B(\vec{u}_i, \vec{v})$ является линейным и непрерывным отображением F в G . Поэтому к абсолютно сходящемуся в F ряду $\sum_{j \in J} \vec{v}_j$ и

к этому линейному непрерывному отображению можно применить теорему 60 и записать формулу

$$\sum_{j \in J} B(\vec{u}_i, \vec{v}_j) = B\left(\vec{u}_i, \sum_{j \in J} \vec{v}_j\right) = B(\vec{u}_i, \vec{V}). \quad (\text{II}, 14; 17)$$

Но точно так же мы можем затем рассмотреть линейное и непрерывное отображение $\vec{u} \rightarrow B(\vec{u}, \vec{V})$ пространства E в G и применить к этому отображению и ряду $\sum_{i \in I} \vec{u}_i$ теорему 60. На этот раз мы получим следующее:

$$\sum_{i \in I} B(\vec{u}_i, \vec{V}) = B\left(\sum_{i \in I} \vec{u}_i, \vec{V}\right) = B(\vec{U}, \vec{V}), \quad (\text{II}, 14; 18)$$

и, следовательно, $\sum_{(i, j) \in I \times J} B(\vec{u}_i, \vec{v}_j) = B(\vec{U}, \vec{V})$, что совпадает с (II, 14; 15).

В случае, когда I и J являются множеством \mathbb{N} целых чисел ≥ 0 , удобно положить

$$\vec{w}_n = B(\vec{u}_n, \vec{v}_0) + B(\vec{u}_{n-1}, \vec{v}_1) + \dots + B(\vec{u}_0, \vec{v}_n) \quad (\text{II}, 14; 19)$$

и рассматривать ряд $\sum_{(i, j) \in \mathbb{N} \times \mathbb{N}} B(\vec{u}_i, \vec{v}_j)$ как ряд $\sum_{n=0}^{\infty} \vec{w}_n$.

Замечание. Из доказательства вытекает следующее: если заранее известно, что ряд $\sum_{i, j} B(\vec{u}_i, \vec{v}_j)$ безусловно сходится, то нет необходимости предполагать ряды абсолютно сходящимися.

Обратимые отображения в банаховых пространствах

Определение. Пусть u — линейное непрерывное отображение нормированного векторного пространства E в нормированное векторное пространство F . Говорят, что отображение u *обратимо*, если оно является биекцией и если обратная биекция u^{-1} (очевидно, линейная) также непрерывна. Если $E = F$, то это означает, что элемент u в алгебре $\mathcal{L}(E; E)$ имеет обратный. Учитывая соотношение « $u \circ u^{-1} = \text{ тождественное отображение}$ » и формулу (II, 13; 7), между нормами этих двух обратных биекций можно установить соотношение:

$$1 = \| \text{ тождественное отображение } \| \leq \| u \| \| u^{-1} \|, \text{ или } \| u^{-1} \|^{-1} \leq \| u \|. \quad (\text{II}, 14; 20)$$

Свойства рядов в банаховых пространствах позволяют доказать, что отображение, достаточно близкое к обратному отображению, само обратимо. Точнее, имеет место такое утверждение:

Теорема 62. Пусть u — линейное непрерывное обратимое отображение банахова пространства E в банахово пространство F , и пусть v — линейное непрерывное отображение E в F , удовлетворяющее неравенству

$$\|v\| < \|u^{-1}\|^{-1}. \quad (\text{II}, 14; 21)$$

Тогда линейное непрерывное отображение $u + v$ пространства E в пространство F также обратимо, и при этом

$$\|(u + v)^{-1}\| \leq \frac{1}{\|u^{-1}\|^{-1} - \|v\|}.$$

Доказательство. Рассмотрим сначала случай $F = E$ и $u = I$, где I — тождественное отображение E в E . Тогда v будет линейным непрерывным отображением E в E , удовлетворяющим неравенству $\|v\| < 1$. Знак $<$ не может быть заменен знаком \leq , ибо для $v = -I$ отображение $I - I = 0$ не обратимо.

Произведем сначала чисто формальные выкладки. Мы будем вычислять $(I + v)^{-1}$ как вычисляют $(1 + z)^{-1} = 1/(1 + z)$, записывая сумму в виде ряда геометрической прогрессии $1 - z + z^2 - \dots$. Напишем

$$(I + v)^{-1} = I - v + v^2 - v^3 + \dots + (-1)^n v^n + \dots \quad (\text{II}, 14; 22)$$

В этой формуле отображение v^n является композицией $v \circ v \circ \dots \circ v$, составленной из n отображений, совпадающих с v . Теперь обоснуем написанную формулу.

Ряд, стоящий в правой части, нормально сходится в пространстве Банаха $\mathcal{L}(E; E)$ (теорема 50), ибо $\|v^2\| \leq \|v \circ v\| \leq \|v\|^2$, $\|v^3\| \leq \|v\|^3$, ... и т. д. и по условию $\|v\| < 1$. Согласно теореме 55, этот ряд сходится и представляет собой элемент w из $\mathcal{L}(E; E)$.

Согласно теореме 54, функция $u \rightarrow uv$ является линейным непрерывным отображением пространства $\mathcal{L}(E; E)$ в пространство $\mathcal{L}(E; E)$. Теперь можно к этому отображению и ряду, стоящему в правой части (II, 14; 22), применить теорему 60, что дает:

$$\begin{aligned} w(I + v) &= w + w \cdot v = \\ &= (I - v + v^2 - v^3 + \dots) + (v - v^2 + v^3 - \dots) = I. \end{aligned} \quad (\text{II}, 14; 23)$$

Такое же рассуждение показывает, что $(I + v)w = I$. Тогда, в силу изложенного на стр. 16 гл. I, отображение $I + v$ будет являться биекцией, а отображение w — его обратной биекцией. Поскольку $w \in \mathcal{L}(E; E)$, то $I + v$ обратимо, и для рассматриваемого случая теорема доказана. Кроме того, имеем оценку:

$$\|w\| = \|(I + v)^{-1}\| \leq 1 + \|v\| + \|v\|^2 + \dots = \frac{1}{1 - \|v\|}. \quad (\text{II}, 14; 23_2)$$

Заметим, что обратимость $I + v$ вытекает также из теоремы о неподвижной точке. Пусть \vec{y} — некоторый элемент E . Проверим, существует ли такая точка \vec{x} из E , что $(I + v)\vec{x} = \vec{y}$, или $\vec{x} + v \cdot \vec{x} = \vec{y}$, или $\vec{x} = -v \cdot \vec{x} + \vec{y}$? Рассмотрим отображение $f: \vec{x} \rightarrow -v \cdot \vec{x} + \vec{y}$ пространства E в E . Это отображение — сжатие, ибо $\| -v \cdot (\vec{x}_1 - \vec{x}_2) \| \leq \| v \| \|\vec{x}_1 - \vec{x}_2\|$ и $\| v \| < 1$. В силу полноты E , существует, и притом единственный, элемент \vec{x} , такой, что $\vec{x} = f(\vec{x})$ или $(I + v)\vec{x} = \vec{y}$ (теорема 46). Следовательно, отображение $I + v$ является биекцией. Поскольку эта биекция линейна, ее обратная биекция также линейна. Кроме того, теорема 46₂ показывает, что решение \vec{x} непрерывно зависит от \vec{y} ; следовательно, отображение $(I + v)^{-1}$ непрерывно, а отображение $I + v$ обратимо.

Заметим, что выражение для \vec{x} может быть получено из элемента \vec{y} по методу последовательных приближений. Исходя из $\vec{x}_0 = \vec{0}$, получаем $\vec{x}_1 = \vec{y}$, $\vec{x}_2 = -v \cdot \vec{x}_1 + \vec{y} = -v \cdot \vec{y} + \vec{y}$, $\vec{x}_3 = -v \cdot \vec{x}_2 + \vec{y} = \vec{y} - v \cdot \vec{y} + v^2 \cdot \vec{y}$, ... и в пределе находим $\vec{x} = \vec{y} - v \cdot \vec{y} + v^2 \cdot \vec{y} - v^3 \cdot \vec{y} + \dots$. Это новый способ записи равенства $(I - v)^{-1} = I - v + v^2 - v^3 + \dots$.

Рассмотрим теперь общий случай произвольных пространств E и F , обратимого отображения u и $\| u \| < \| u^{-1} \|^{-1}$. Поскольку отображение u обратимо, можно записать, что

$$u + v = u(I + u^{-1}v). \quad (\text{II}, 14; 24)$$

Здесь $u^{-1}v$ и $I + u^{-1}v$ — линейные непрерывные отображения пространства E в E . В силу предположения, сделанного относительно v , $\| u^{-1}v \| \leq \| u^{-1} \| \| v \| < 1$. Теперь, в силу доказанного частного случая, $I + u^{-1}v$ обратимо и его обратное отображение может быть записано в виде

$$(I + u^{-1}v)^{-1} = I - u^{-1}v + u^{-1}vu^{-1}v - u^{-1}vu^{-1}vu^{-1}v + \dots \quad (\text{II}, 14; 25)$$

Но тогда, согласно (II, 14; 22), $u + v$ можно представить в виде композиции двух обратимых отображений, а, значит, обратимо само это отображение и его обратное отображение является композицией обратных отображений, взятых в обратном порядке:

$$(u + v)^{-1} = (I + u^{-1}v)^{-1}u^{-1} = u^{-1} - u^{-1}vu^{-1} + u^{-1}vu^{-1}vu^{-1} - u^{-1}vu^{-1}vu^{-1}vu^{-1} + \dots \quad (\text{II}, 14; 26)$$

Кроме того, имеет место оценка:

$$\begin{aligned} \|(u+v)^{-1}\| &\leqslant \|(I+u^{-1}v)^{-1}\| \|u^{-1}\| \leqslant \\ &\leqslant \frac{\|u^{-1}\|}{1-\|u^{-1}v\|} \leqslant \frac{\|u^{-1}\|}{1-\|u^{-1}\|\|v\|} = \frac{1}{\|u^{-1}\|^{-1}-\|v\|}. \quad (\text{II, 14; 27}) \end{aligned}$$

Замечание 1. Пусть \mathcal{U} — множество обратимых элементов пространства $\mathcal{L}(E; F)$ и \mathcal{U}^{-1} — множество обратимых элементов пространства $\mathcal{L}(F; E)$. Если $u_0 \in \mathcal{U}$, то любой элемент пространства $\mathcal{L}(E; F)$, принадлежащий открытому шару с центром u_0 радиуса $\|u_0^{-1}\|^{-1}$, лежит в \mathcal{U} . Следовательно, множество \mathcal{U} является открытым подмножеством пространства $\mathcal{L}(E; F)$, а если поменять ролями E и F , то мы получим, что таким же будет множество \mathcal{U}^{-1} . Предыдущие неравенства позволяют легко доказать, что $u \rightarrow u^{-1}$ является гомеоморфизмом \mathcal{U} на \mathcal{U}^{-1} . Это можно проверить в качестве упражнения. С этим фактом мы встретимся в теореме 27 гл. III.

Замечание 2. Пусть \mathcal{A} — банахова алгебра. То же самое рассуждение показывает, что если элемент $\vec{x} \in \mathcal{A}$ обратим, т. е. имеет в алгебре обратный, а \vec{y} является таким элементом \mathcal{A} , что $\|\vec{y}\| < \|\vec{x}^{-1}\|^{-1}$, то элемент $\vec{x} + \vec{y}$ обратим и обратный к нему элемент выражается формулой

$$(\vec{x} + \vec{y})^{-1} = \vec{x}^{-1} - \vec{x}^{-1}\vec{y}\vec{x}^{-1} + \vec{x}^{-1}\vec{y}\vec{x}^{-1}\vec{y}\vec{x}^{-1} - \dots \quad (\text{II, 14; 27}_2)$$

Если \mathcal{U} является множеством обратимых элементов алгебры \mathcal{A} , то \mathcal{U} является открытым подмножеством \mathcal{A} , а отображение $\vec{x} \rightarrow \vec{x}^{-1}$ является гомеоморфизмом \mathcal{U} на себя.

Критерий условной сходимости

Говорят, что последовательность $\vec{u}_0, \vec{u}_1, \vec{u}_2, \dots, \vec{u}_n, \dots$ элементов банахова пространства E имеет *ограниченную вариацию*, если ряд

$$\|\vec{u}_1 - \vec{u}_0\| + \|\vec{u}_2 - \vec{u}_1\| + \|\vec{u}_3 - \vec{u}_2\| + \dots \quad (\text{II, 14; 28})$$

сходится. Сумма (II, 14; 28) называется *полной вариацией* последовательности. Ограниченностю вариации последовательности приводит, естественно, к ее сходимости. В самом деле, сходимость последовательности равносильна сходимости ряда $\vec{u}_0 + (\vec{u}_1 - \vec{u}_0) + (\vec{u}_2 - \vec{u}_1) + \dots$, а сделанное выше предположение равносильно абсолютной сходимости этого ряда. Таким образом,

последовательность с ограниченной вариацией является сходящейся, так как абсолютно сходящийся ряд сходится. Если последовательность вещественных чисел u_n монотонна и ограничена, то она имеет ограниченную вариацию. В самом деле, например, если последовательность возрастающая и ограниченная, то

$$|u_1 - u_0| + |u_2 - u_1| + \dots = (u_1 - u_0) + (u_2 - u_1) + \dots = \\ = -u_0 + \lim_{n \rightarrow \infty} u_n < +\infty. \quad (\text{II}, 14; 29)$$

Впрочем, Жордан доказал, что и, обратно, если некоторая последовательность вещественных чисел имеет ограниченную вариацию, то она может быть записана в виде $u_n = a_n - b_n$, где a_0, a_1, a_2, \dots и b_0, b_1, b_2, \dots — две возрастающие и ограниченные последовательности¹⁾.

Дадим теперь наиболее важный критерий условной сходимости рядов, к которому практически можно свести все наиболее употребительные критерии (в особенности теоремы о знакопеременных рядах).

Теорема 63 (Абель). Пусть E, F, G — три пространства Банаха. Пусть \vec{u}_n — некоторая последовательность векторов из E с ограниченной вариацией, стремящихся к $\vec{0}$ при $n \rightarrow \infty$, и \vec{v}_n — последовательность векторов из F с ограниченными частными суммами, т. е. такая, что нормы величин

$$\vec{\sigma}_{m,n} = \vec{v}_m + \vec{v}_{m+1} + \dots + \vec{v}_n, \quad n \geq m, \quad (\text{II}, 14; 30)$$

ограничены. Тогда, если B является билинейным непрерывным отображением пространства $E \times F$ в пространство G , то ряд с общим членом $\vec{w}_n = B(\vec{u}_n, \vec{v}_n)$ сходится.

Кроме того, если положить $U_m = \|\vec{u}_{m+1} - \vec{u}_m\| + \|\vec{u}_{m+2} - \vec{u}_{m+1}\| + \dots$ и $V_m = \sup_{n \geq m} \|\vec{\sigma}_{m,n}\|$, то сумма \vec{S} и остаток $\vec{R}_m = \vec{w}_{m+1} + \vec{w}_{m+2} + \dots$ могут быть оценены следующим образом:

$$\|\vec{S}\| \leq \|B\| U_0 V_0 \quad \text{и} \quad \|\vec{R}_m\| \leq \|B\| U_{m+1} V_{m+1}. \quad (\text{II}, 14; 31)$$

¹⁾ Это почти очевидно. Достаточно положить $a_0 = (u_0)^+$, $a_1 = (u_0)^+ + (u_1 - u_0)^+$, $a_2 = (u_0)^+ + (u_1 - u_0)^+ + (u_2 - u_1)^+$, ..., где $x^+ = x$ для $x \geq 0$ и $x^+ = 0$ для $x < 0$. Затем $b_0 = (u_0)^-$, $b_1 = (u_0)^- + (u_1 - u_0)^-$, $b_2 = (u_0)^- + (u_1 - u_0)^- + (u_2 - u_1)^-$, ..., где $x^- = 0$ для $x > 0$ и $x^- = |x|$ для $x \leq 0$. Ясно, что $x = x^+ - x^-$ и $|x| = x^+ + x^-$.

²⁾ Эти величины стремятся к нулю при m , стремящемся к $+\infty$, потому что V_{m+1} остается ограниченным, а U_{m+1} стремится к нулю как остаток сходящегося ряда с положительными членами.

В большинстве приложений $E = F = G = \mathbb{C}$ — поле комплексных чисел, а B — произведение. Величины σ_m, n тогда заранее ограничены, если ряд $\sum_{n=0}^{\infty} \vec{v}_n$ сходится.

Доказательство. Имеем:

$$\begin{aligned}\vec{S}_n &= B(\vec{u}_0, \vec{v}_0) + B(\vec{u}_1, \vec{v}_1) + \dots + B(\vec{u}_n, \vec{v}_n) = \\ &= B(\vec{u}_0, \sigma_{0,0}) + B(\vec{u}_1, \sigma_{0,1} - \sigma_{0,0}) + \dots + B(\vec{u}_n, \sigma_{0,n} - \sigma_{0,n-1}) = \\ &= B(\vec{u}_0, \sigma_{0,0}) + [B(\vec{u}_1, \sigma_{0,1}) - B(\vec{u}_1, \sigma_{0,0})] + \dots \\ &\quad \dots + [B(\vec{u}_n, \sigma_{0,n}) - B(\vec{u}_n, \sigma_{0,n-1})] = \\ &= B(\vec{u}_0 - \vec{u}_1, \sigma_{0,0}) + B(\vec{u}_1 - \vec{u}_2, \sigma_{0,1}) + \dots \\ &\quad \dots + B(\vec{u}_{n-1} - \vec{u}_n, \sigma_{0,n-1}) + B(\vec{u}_n, \sigma_{0,n}).\end{aligned}\quad (\text{II}, 14; 32)$$

Отличный от других член $B(\vec{u}_n, \sigma_{0,n})$ мажорируется по норме величиной $\|B\| \| \vec{u}_n \| V_0$, а так как при n , стремящемся к $+\infty$, \vec{u}_n стремится к $\vec{0}$, то эта величина стремится к 0.

Остается доказать, что сумма

$$B(\vec{u}_0 - \vec{u}_1, \sigma_{0,0}) + B(\vec{u}_1 - \vec{u}_2, \sigma_{0,1}) + \dots + B(\vec{u}_{n-1} - \vec{u}_n, \sigma_{0,n-1})\quad (\text{II}, 14; 33)$$

имеет предел при n , стремящемся к $+\infty$, т. е. что ряд с общим членом

$$\vec{z}_n = B(\vec{u}_n - \vec{u}_{n+1}, \sigma_{0,n})\quad (\text{II}, 14; 34)$$

сходится. Ряд же из норм этих членов сходится, поскольку

$$\|\vec{z}_n\| \leq \|B\| \|\vec{u}_n - \vec{u}_{n+1}\| V_0\quad (\text{II}, 14; 35)$$

и последовательность \vec{u}_n по предположению имеет ограниченную вариацию. Так как пространство G по предположению полно, то в силу теоремы 55 сходимость данного ряда доказана. Доказав сходимость, мы можем, исходя из неравенства

$$\|\vec{S}_n\| \leq \|B\| \|\vec{u}_n\| V_0 + \|B\| U_0 V_0,\quad (\text{II}, 14; 36)$$

переходом к пределу при n , стремящемся к бесконечности, получить оценку для суммы $\vec{S} = \vec{w}_0 + \vec{w}_1 + \vec{w}_2 + \dots$

$$\|\vec{S}\| \leq \|B\| U_0 V_0.\quad (\text{II}, 14; 37)$$

Те же самые рассуждения, начатые с члена \vec{w}_{m+1} с учетом того, что $\vec{R}_m = \vec{w}_{m+1} + \vec{w}_{m+2} + \dots$, дадут оценку остатка (II, 14; 31).

Примеры. 1°) *Теорема о знакопеременных рядах.* Рассмотрим вещественный ряд $\sum_{n=0}^{\infty} (-1)^n u_n$, в котором последовательность чисел $u_0, u_1, u_2, \dots \geq 0$ является убывающей и сходящейся к 0 при n , стремящемся к $+\infty$. В этом случае применима доказанная теорема, если положить $v_n = (-1)^n$, $U_m = u_m$, $V_m = 1$. Она же дает известное неравенство

$$|R_m| \leq u_{m+1}. \quad (\text{II, 14; 38})$$

Известно также, что в этом случае остаток имеет знак первого отброшенного члена.

2°) *Случай тригонометрических рядов.* Рассмотрим ряд $\sum_{n=0}^{\infty} u_n e^{ni\theta}$, в котором предполагается, что последовательность комплексных чисел u_0, u_1, u_2, \dots имеет ограниченную вариацию и сходится к 0 при n , стремящемся к $+\infty$, и что θ является вещественным числом¹⁾. Примем $v_n = e^{ni\theta}$. По формуле суммы членов геометрической прогрессии при $\theta \neq 2k\pi$ получаем равенство

$$e^{mi\theta} + e^{(m+1)i\theta} + \dots + e^{ni\theta} = \frac{e^{(n+1)i\theta} - e^{mi\theta}}{e^{i\theta} - 1} \quad (\text{II, 14; 39})$$

и, следовательно, оценку

$$|\sigma_{m,n}| \leq \frac{2}{|e^{i\theta} - 1|}. \quad (\text{II, 14; 40})$$

Напротив, если $\theta = 2k\pi$, то мы получаем величину $e^{mi\theta} + \dots + e^{ni\theta} = n - m + 1$, которая не является ограниченной. Отсюда вытекает, что для $\theta \neq 2k\pi$ ряд $\sum_{n=0}^{\infty} u_n e^{ni\theta}$ сходится и имеют место оценки

$$|S| \leq U_0 \frac{2}{|e^{i\theta} - 1|}, \quad |R_m| \leq U_{m+1} \frac{2}{|e^{i\theta} - 1|}. \quad (\text{II, 14; 41})$$

Аналогичный результат, конечно, имеет место и для рядов $\sum_{n=0}^{\infty} u_n \cos n\theta$ и $\sum_{n=0}^{\infty} u_n \sin n\theta$. Последний же ряд сходится и при $\theta = 2k\pi$, поскольку все его члены при этом обращаются в нуль.

1) Если u_n вещественны, убывают и стремятся к 0 и если $\theta = \pi$, то мы получаем знакопеременный ряд.

**§ 15. НАИБОЛЕЕ УПОТРЕБИТЕЛЬНЫЕ ПРИМЕРЫ
ФУНКЦИОНАЛЬНЫХ ПРОСТРАНСТВ.
СХОДИМОСТЬ ПРОСТАЯ И РАВНОМЕРНАЯ**

Функциональные пространства

Под функциональным пространством понимают пространство, элементами которого являются функции, т. е. отображения одного множества в другое.

Пусть E и F — два множества. В гл. I (стр. 15) через F^E мы обозначали множество всевозможных отображений E в F . Если F обладает некоторой структурой (векторного пространства, топологического пространства и т. д.), то, вообще говоря, и в F^E можно ввести аналогичную структуру.

1°) Предположим, что F является векторным пространством над полем \mathbb{K} . Тогда, очевидно, в F^E можно также ввести структуру векторного пространства над полем \mathbb{K} . В самом деле, если \vec{f} и \vec{g} — два отображения E в F и λ — некоторый скаляр, то можно определить сумму $\vec{f} + \vec{g}$ и произведение $\lambda\vec{f}$ как новое отображение из E в F по формулам:

$$(\vec{f} + \vec{g})(x) = \vec{f}(x) + \vec{g}(x) \quad \text{для любого } x \in E, \quad (\text{II}, 15; 1)$$

$$(\lambda\vec{f})(x) = \lambda\vec{f}(x) \quad \text{для любого } x \in E. \quad (\text{II}, 15; 2)$$

Таким образом, на F^E мы определили закон сложения и закон умножения на скаляр из \mathbb{K} . Без труда проверяется, что эти законы удовлетворяют всем аксиомам, превращающим F^E в векторное пространство над полем \mathbb{K} ¹⁾.

Например, если F является полем скаляров \mathbb{K} , а последнее является полем вещественных или комплексных чисел, то множество \mathbb{K}^E вещественных или комплексных функций, определенных на E , является векторным пространством над полем вещественных или комплексных чисел.

2°) Предположим теперь, что F является метрическим пространством. Если f и g — два отображения E в F , то расстоянием между этими функциями будем называть величину, обозначаемую через $d(f, g)$ и определяемую по формуле:

$$d(f, g) = \sup_{x \in E} d(f(x), g(x)). \quad (\text{II}, 15; 3)$$

Это расстояние не обязательно конечно; точную верхнюю грань следует брать в пополненной прямой $\bar{\mathbb{R}}$. Отсюда вытекает,

¹⁾ Если E состоит из двух элементов, то F^E может быть отождествлено с $F^2 = F \times F$. Структура полученного при этом векторного пространства совпадает со структурой пространства произведения, рассмотренного на стр. 119 (формулы (II, 13; 22) и (II, 13; 23)).

что в множестве F^E невозможно непосредственно ввести естественную структуру метрического пространства. Поэтому мы ограничимся рассмотрением подпространства $(F^E)_b$ пространства F^E , состоящего из *ограниченных отображений* E в F . Говорят, что отображение f множества E в F ограничено, если область значений $f(E)$ — образ множества E — является ограниченной частью F . В этом случае расстояние между любыми двумя ограниченными отображениями f и g из E в F заведомо конечно. (В самом деле, пусть, например, $f(E)$ содержится в шаре с центром a радиуса α , а $g(E)$ лежит в шаре с центром b радиуса β и тем более в шаре с центром a радиуса $\beta + d(a, b)$ множества F . Очевидно, для всех x из E имеет место неравенство $d(f(x), g(x)) \leq \alpha + \beta + d(a, b)$ и, следовательно, $d(f, g)$ конечно.)

Проверим теперь, что отображение $d: (f, g) \rightarrow d(f, g)$ действительно определяет расстояние в пространстве $(F^E)_b$ ¹⁾. Для этого необходимо проверить все три свойства определения (II, 1; 1). Симметрия очевидна. Положительность также очевидна, ибо, с одной стороны, $d(f, g) \geq 0$, а, с другой стороны, для различных f и g найдется хотя бы один элемент x из E , такой, что $d(f(x), g(x)) > 0$, и тогда тем более $d(f, g) > 0$.

Остается проверить неравенство треугольника.

Пусть f, g, h — три отображения E в F . Для любого x из E имеем:

$$d(f(x), h(x)) \leq d(f(x), g(x)) + d(g(x), h(x)) \leq d(f, g) + d(g, h). \quad (\text{II, 15; 4})$$

Поскольку это верно для любого x , то получаем искомое неравенство:

$$d(f, h) \leq d(f, g) + d(g, h). \quad (\text{II, 15; 5})$$

Замечание. Только что введенные понятия являются *метрическими*, а не топологическими. Легко доказывается, что если в F заменить данную метрику на эквивалентную, то полностью изменяется в первую очередь пространство $(F^E)_b$, ибо ограниченность функции зависит от метрики, а не от топологии (см. стр. 57), и, кроме того, если даже это пространство не изменится, то новая метрика в $(F^E)_b$ не будет эквивалентна первой.

¹⁾ Если E состоит лишь из двух элементов, то F^E может быть отождествлено с $F^2 = F \times F$. Только что установленное расстояние в $(F^E)_b$ (которое в нашем случае совпадает с F^E) является одним из расстояний, которое мы в свое время выбирали для произведения на стр. 119:

$$d((x_1, x_2), (y_1, y_2)) = \sup [d(x_1, y_1), d(x_2, y_2)].$$

3°) Предположим теперь, что F имеет структуру нормированного векторного пространства. Тогда в пространстве $(F^E)_b$ также можно ввести структуру нормированного векторного пространства, положив

$$\|f\| = \sup_{x \in E} \|\vec{f}(x)\|^l. \quad (\text{II}, 15; 6)$$

Методы, аналогичные предыдущему, показывают, что мы определили некоторую норму. Кроме того, метрика, определяемая этой нормой, является метрикой, которую мы определили в п. 2°): $d(f, g) = \|f - g\|$.

Если, например, F является полем вещественных или комплексных чисел, то пространство вещественных или комплексных ограниченных функций, определенных на некотором множестве E , является нормированным векторным пространством²⁾.

Векторные пространства, полученные в п. 1°) и 3°), практически всегда бесконечномерны. В самом деле, предположим, что E является некоторым множеством из n элементов, которые для удобства мы обозначим через $1, 2, \dots, n$. Тогда множество F^E в частном случае, когда $F = \mathbb{R}$, является пространством-произведением \mathbb{R}^n , а норма, введенная в п. 3°), имеет вид

$$\|(x_1, x_2, \dots, x_n)\| = \sup(|x_1|, |x_2|, \dots, |x_n|). \quad (\text{II}, 15; 7)$$

Каждый раз, когда E содержит бесконечное множество элементов (а во всех практически важных случаях в качестве E берется вещественная прямая или некоторый интервал этой прямой), рассмотренное выше пространство F^E бесконечномерно.

Теорема 64. Если метрическое пространство F полно, то метрическое пространство $(F^E)_b$ ограниченных отображений E в F также полно³⁾.

Доказательство. Пусть f_0, f_1, f_2, \dots — последовательность Коши из $(F^E)_b$. Из определения расстояния следует, что $d(f_m(x), f_n(x)) \leq d(f_m, f_n)$, а, значит, для любой точки x из E

1) Через $\| \cdot \|$ мы обозначаем норму элемента F и через $\| \cdot \|$ норму ограниченного отображения E в F для того, чтобы избежать смешения нормированных векторных пространств F и $(F^E)_b$. Таким образом, $\|\vec{f}\|$ означает функцию $x \rightarrow \|\vec{f}(x)\|$, которая ≥ 0 , в то время, как $\|\vec{f}\|$ означает точную верхнюю грань этой функции, т. е. число ≥ 0 .

2) В поле скаляров \mathbb{K} норма совпадает с модулем, обозначаем через $| \cdot |$. Поэтому, если $f \in (\mathbb{K}^E)_b$, то $|f|$ является функцией $x \rightarrow |f(x)|$, которая ≥ 0 . Здесь будет уместным через $\|f\|$ обозначать точную верхнюю грань этой функции, т. е. норму f в $(\mathbb{K}^E)_b$.

3) В силу примечания на стр. 147, эта теорема содержит как частный случай теорему 44.

последовательность точек $f_n(x)$ является последовательностью Коши в F .

Поскольку пространство F по предположению полно, то эта последовательность сходится к некоторой точке F , которую мы обозначим через $f(x)$. Тем самым мы определили отображение f множества E в пространство F .

Докажем прежде всего, что это отображение ограничено. Зададим $\varepsilon > 0$. Тогда существует целое число p , такое, что при $m \geq p, n \geq p$ имеет место неравенство $d(f_m, f_n) \leq \varepsilon$, а, следовательно, $d(f_m(x), f_n(x)) \leq \varepsilon$ для всех x из E . Зафиксировав x из E , перейдем в последнем неравенстве к пределу при m , стремящемся к $+\infty$, и, учитывая непрерывность функции расстояния в F , получим неравенство $d(f(x), f_n(x)) \leq \varepsilon$ для $n \geq p$. Так как f_p ограничено, то множество $f_p(E)$ содержится в некотором шаре с центром a_p радиуса R_p . Из неравенства

$$d(a_p, f(x)) \leq d(a_p, f_p(x)) + d(f_p(x), f(x))$$

следует, что множество $f(E)$ содержитя в шаре с центром a_p радиуса $R_p + \varepsilon$; тем самым ограниченность f полностью доказана. Отображение f , таким образом, также является элементом $(F^E)_b$. Остается убедиться, что f_n сходится к f при n , стремящемся к $+\infty$. Действительно, при заданном ε и выбранном p для всех $x \in E$ и $n \geq p$ имеем неравенство $d(f(x), f_n(x)) \leq \varepsilon$, из которого следует неравенство $d(f, f_n) \leq \varepsilon$, означающее сходимость f_n к f . Окончательно получаем, что $(F^E)_b$ является полным метрическим пространством.

Следствие. Если F является пространством Банаха, то нормированное векторное пространство $(F^E)_b$ также является пространством Банаха. В частности, пространство $(K^E)_b$ ограниченных функций, определенных на множестве E , с вещественными или комплексными значениями является пространством Банаха.

Было бы полезным использовать полученные результаты при исследовании сходимости последовательностей функций, чтобы иметь возможность сказать, что некоторая последовательность функций f_n сходится к предельной функции f , если эти функции как точки f_n некоторого топологического пространства сходятся в этом пространстве к точке f .

Простая сходимость последовательности функций

Говорят, что последовательность функций f_n , т. е. последовательность отображений множества E в метрическое пространство F , просто сходится при n , стремящемся к $+\infty$, к предель-

ной функции f , если для любого x из E последовательность точек $f_n(x)$ из F сходится при n , стремящемся к $+\infty$, к точке $f(x)$ пространства F .

С помощью логических знаков можно это определение записать так:

$$(\forall x \in E) (\forall \varepsilon > 0) (\exists m \in \mathbb{N}) (\forall n \geq m): d(f_n(x), f(x)) \leq \varepsilon.$$

(II, 15; 8)

Отметим, что, как мы видели ранее (в гл. I, стр. 38), выбор m зависит в действительности одновременно от ε и от x . Возникает вопрос: можно ли в пространстве F^E отображений E в F ввести топологическую структуру таким образом, чтобы элементы этого топологического пространства f_n сходились к элементу f тогда и только тогда, когда функции f_n просто сходятся к функции f в указанном выше смысле? Это в действительности возможно, но не просто. Получаемое при этом топологическое пространство не метризуемо, и мы его здесь рассматривать не будем.

Равномерная сходимость последовательности функций

Говорят, что последовательность функций f_n сходится равномерно к функции f при n , стремящемся к $+\infty$, если целое число m , указанное в (II, 15; 8), может быть выбрано независимо от x , т. е. если оно является лишь функцией ε ; другими словами, если

$$(\forall \varepsilon > 0) (\exists m \in \mathbb{N}) (\forall x \in E) (\forall n \geq m): d(f_n(x), f(x)) \leq \varepsilon.$$

(II, 15; 9)

Это же можно записать короче:

$$(\forall \varepsilon > 0) (\exists m \in \mathbb{N}) (\forall n \geq m): d(f_n, f) \leq \varepsilon. \quad (\text{II, 15; 10})$$

Последнее означает, что расстояние между f_n и f стремится к нулю при n , стремящемся к бесконечности.

Очевидно, равномерная сходимость влечет за собой простую сходимость, но, как мы сейчас увидим, обратное не верно. Равномерная сходимость является гораздо более сильным свойством, чем обычная сходимость.

Пример 1. Рассмотрим вещественную функцию вещественной переменной, определенную формулой

$$g(x) = \frac{1}{1+x^2}. \quad (\text{II, 15; 11})$$

Назовем *сдвигом* этой функции на h функцию, полученную перемещением графика данной функции параллельно оси x на расстояние h , т. е. отображение $(x, y) \rightarrow (x + h, y)$. Другими словами, значениями новой функции $\tau_h g$ в точке x являются значения старой функции $g(x)$ в точке $x - h$:

$$(\tau_h g)(x) = g(x - h), \quad (\text{II}, 15; 12)$$

так что

$$(\tau_h g)(x) = \frac{1}{1 + (x - h)^2}. \quad (\text{II}, 15; 13)$$

Рассмотрим теперь последовательность сдвигов $\tau_n g$, $n \in \mathbb{N}$. Непосредственно видно, что эта последовательность функций

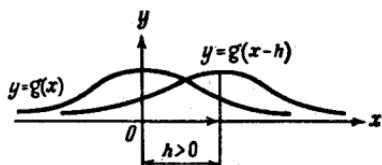


Рис. 3.

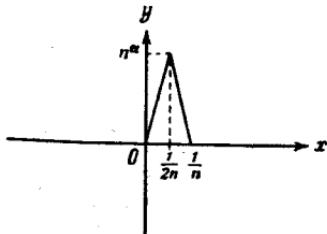


Рис. 4.

сходится к тождественно нулевой функции при n , стремящемся к $+\infty$. Действительно, для фиксированного x

$$\lim_{n \rightarrow \infty} \frac{1}{1 + (x - n)^2} = 0. \quad (\text{II}, 15; 14)$$

Впрочем, это означает лишь, что последовательность значений функции g в точках $x - n$ стремится к 0 при n , стремящемся к $+\infty$, что очевидно.

Однако, последовательность функций $\tau_n g$ при n , стремящемся к $+\infty$, равномерно сходиться к 0 не может, ибо расстояние от $\tau_n g$ до 0, будучи всегда равным 1, не зависит от n .

Пример 2. Рассмотрим вещественную функцию f_n ($n \geq 1$) вещественной переменной, определенную следующим образом: она равна 0 для $x \leq 0$ и $x \geq 1/n$, она равна n^α , $\alpha > 0$, для $x = 1/(2n)$, а в каждом из интервалов $[0, 1/(2n)]$, $[1/(2n), 1/n]$ она аффинно линейна¹⁾. График этой функции изображен на рис. 4.

1) Аффинно линейной, или аффинной, мы называем функцию $y = ax + b$. Слово «линейная» в соответствии с общим определением линейного отображения векторных пространств мы сохраняем за функцией $y = ax$.

Как это ни кажется странным, но последовательность f_n сходится к функции, тождественно равной 0, при n , стремящемся к $+\infty$. В самом деле, при любом $x > 0$ для достаточно большого n выполняется неравенство $1/n < x$, а, следовательно, $f_n(x) = 0$. Для всех $x \leq 0$, $f_n(x) = 0$, откуда и следует наше утверждение. Однако расстояние от f_n до 0 равно n^α ; эта величина стремится к бесконечности, а, значит, f_n не сходится равномерно к 0 при n , стремящемся к $+\infty$. Мы видим, что понятие простой сходимости на самом деле является не столь уж естественным, как это кажется с первого взгляда. В самом деле, тот факт, что две рассмотренные выше последовательности функций стремятся к 0 при n , стремящемся к $+\infty$, выглядит довольно странным. Для последовательности функций понятие равномерной сходимости является более естественным, чем понятие простой сходимости.

Из сказанного выше о связи между равномерной сходимостью и расстоянием между функциями видно, что топологическим пространством, приспособленным к равномерной сходимости, является метрическое пространство $(F^E)_b$. Сказать, что некоторая последовательность ограниченных отображений f_n из E в F сходится равномерно к ограниченному отображению f из E в F , означает сказать, что последовательность точек f_n метрического пространства $(F^E)_b$ сходится к точке f этого метрического пространства.

Другие применения выражения «равномерная сходимость»

Рассмотрим последовательность элементов $x_n(\lambda)$ метрического пространства F , зависящих от параметра λ , пробегающего множество Λ . Говорят, что эта последовательность элементов сходится к некоторому пределу $x(\lambda)$ из F (зависящему, очевидно, также от параметра λ) *равномерно по λ* , пробегающему множество Λ , если

$$(\forall \varepsilon > 0) (\exists m \in \mathbb{N}) (\forall \lambda \in \Lambda) (\forall n \geq m): d(x_n(\lambda), x(\lambda)) \leq \varepsilon.$$

Это, по существу, равносильно утверждению, что последовательность функций $\lambda \rightarrow x_n(\lambda)$, определенных на Λ со значениями в F , сходится равномерно к функции $\lambda \rightarrow x(\lambda)$. Это понятие сводится, следовательно, к предыдущему, но с психологической точки зрения мы находимся в разных ситуациях, когда рассматривается последовательность функций или когда мы имеем дело с последовательностью точек, зависящих от параметра λ .

Рассмотрим теперь последовательность отображений f_n вещественной прямой \mathbb{R} в метрическое пространство F . Что означает выражение: последовательность f_n сходится при n , стремящемся к $+\infty$, к предельной функции f *равномерно на каждом*

ограниченном интервале \mathbb{R} ? Очевидно, это значит, что каким бы ни был ограниченный интервал $[a, b]$ прямой \mathbb{R} , последовательность сужений f_n в этот интервал сходится равномерно к сужению f ; другими словами, это означает, что

$$(\forall a \in \mathbb{R}) (\forall b \in \mathbb{R}, b \geq a) (\forall \varepsilon > 0) (\exists m \in \mathbb{N})$$

$$(\forall x \in [a, b]) (\forall n \geq m): d(f_n(x), f(x)) \leq \varepsilon. \quad (\text{II, 15; 15})$$

Входящее в это определение число m не зависит от x , но является функцией, с одной стороны, числа ε , а с другой, — интервала $[a, b]$. Здесь можно, естественно, заменить прямую \mathbb{R} и ограниченные интервалы на пространство \mathbb{R}^n и ограниченные подмножества \mathbb{R}^n . Более общо, если заданы множество E , метрическое пространство F и семейство частей $(A_i)_{i \in I}$ множества E , то можно говорить о последовательности отображений f_n множества E в F , которая при n , стремящемся к $+\infty$, сходится к отображению f равномерно на каждой части A_i данного семейства. Если мы вернемся к примеру функций τ_{ng} из формулы (II, 15; 13), то увидим, что последовательность функций τ_{ng} сходится к 0 равномерно на каждом ограниченном интервале \mathbb{R} и даже равномерно на всей полуправой $]—\infty, b[$. В самом деле, при $n \geq b$ для всех $x \leq b$ имеем: $1/[1 + (x - n^2)] \leq 1/[1 + (b - n)^2]$, а эта величина стремится к 0 при n , стремящемся к $+\infty$. Если мы рассмотрим теперь второй пример (стр. 151), то увидим, что последовательность функций f_n сходится к функции, тождественно равной нулю, равномерно на дополнении к любому интервалу $[-\delta, +\delta]$, $\delta > 0$, с центром в начале, но она не сходится равномерно ни на каком интервале $]0, \delta[$.

Наконец, если E является топологическим пространством, то говорят, что последовательность функций f_n сходится к f локально равномерно на E , если каждая точка a из E имеет окрестность \mathcal{U}_a , на которой f_n сходится равномерно к f . Это можно записать следующим образом:

$$(\forall a \in E) (\exists \mathcal{U}, \text{окрестность } a) (\forall \varepsilon > 0) (\exists m \in \mathbb{N})$$

$$(\forall x \in \mathcal{U}) (\forall n \geq m): d(f_n(x), f(x)) \leq \varepsilon. \quad (\text{II, 15; 15}_2)$$

Если E локально компактно, то локально равномерная сходимость эквивалентна равномерной сходимости на каждом компакте E . В самом деле, если функции f_n сходятся к функции f равномерно на каждом компакте, то, поскольку каждая точка a из E имеет компактную окрестность, на которой f_n сходятся равномерно, рассматриваемая сходимость является локально равномерной. Обратно, предположим, что сходимость локально равномерна, и пусть K — некоторый компакт E . Для любого a из K существует окрестность \mathcal{U}_a точки a , на которой сходимость

равномерна. Компакт K покрывается конечным числом окрестностей \mathcal{U}_a ; следовательно, сходимость является равномерной на K .

Впрочем, все то, что было сказано по поводу простой или равномерной сходимости *последовательности* функций, распространяется на сходимость множества функций в смысле, указанном на стр. 61. Например, если f_t для любого $t \in \mathbb{R}$ является отображением E в метрическое пространство F : $x \rightarrow f_t(x)$, то можно говорить о простой или равномерной сходимости отображения f_t к отображению f из E в F , когда t стремится к 0 по положительным значениям или стремится к $+\infty$ и т. д.

Пространства, порожденные структурами пространств E и F

До настоящего времени мы вводили функциональные пространства, исходя только из алгебраических или топологических структур на F . Но если E и F уже наделены такими структурами, то можно определить новые пространства. Например, если E и F оба являются векторными пространствами над одним и тем же полем \mathbb{K} , то можно рассмотреть пространство линейных отображений E в F . Это векторное подпространство пространства F^E всех отображений E в F . Если E и F — топологические пространства, то можно ввести пространство $(F^E)_c$ непрерывных отображений E в F . Если F — метрическое пространство, то это пространство не является подпространством $(F^E)_b$, ибо непрерывное отображение не обязательно ограничено. Однако можно рассмотреть подпространство $(F^E)_{bc}$ пространства $(F^E)_b$, образованное непрерывными ограниченными отображениями E в F . Предположим, наконец, что E и F являются нормированными векторными пространствами. Тогда можно ввести, как мы это делали на стр. 114, пространство $\mathcal{L}(E; F)$ линейных непрерывных отображений E в F . Оно не является подпространством $(F^E)_{bc}$, так как линейное отображение, если только оно не является тождественно нулевым, ограниченным быть не может¹⁾. Обозначим через E_0 единичный шар из E . Если линейное отображение E в F задано на шаре E_0 , то в силу гомотетии $u(\lambda \vec{x}) = \lambda u(\vec{x})$ оно определено всюду на E . Впрочем, если задано некоторое отображение единичного шара E_0 в F , то можно выяснить, является ли оно сужением линейного отображения E в F . В самом деле, достаточно продолжить его на E , полагая $u(\vec{x}) = \|\vec{x}\| u(\vec{x}/\|\vec{x}\|)$ для $\|\vec{x}\| > 1$, и прове-

¹⁾ В самом деле, если \vec{a} есть некоторый вектор E , такой, что $u(\vec{a}) \neq \vec{0}$, то последовательность $u(n\vec{a}) = nu(\vec{a})$ не ограничена, ибо $\|nu(\vec{a})\| = n\|u(\vec{a})\|$ стремится к $+\infty$ при n , стремящемся к $+\infty$.

рить, является ли полученное отображение линейным. Согласно теореме 47, для того чтобы линейное отображение i из E в F было непрерывным, необходимо и достаточно, чтобы образ единичного шара E_0 при отображении i был ограниченным в F . Мы видим, следовательно, что пространство линейных отображений E в F можно отождествить с некоторым подпространством пространства всех отображений E_0 в F , а именно с подпространством отображений, являющихся сужением линейных отображений E в F . Пространство линейных непрерывных отображений E в F , т. е. $\mathcal{L}(E; F)$, можно отождествить с некоторым подпространством пространства $(F^{E_0})_b$ ограниченных отображений шара E_0 в пространство F . Учитывая определение, которое мы дали для нормы линейного непрерывного отображения E в F и для нормы ограниченного отображения шара E_0 в F , можно пространство $\mathcal{L}(E; F)$ отождествить также с некоторым нормированным векторным подпространством нормированного векторного пространства $(F^{E_0})_b$.

Легко видеть, что $\mathcal{L}(E; F)$ отождествляется с некоторым замкнутым подпространством $(F^{E_0})_b$. Если F полно, то, по теореме 64, $(F^{E_0})_b$ также полно. Поэтому $\mathcal{L}(E; F)$ как замкнутое подпространство полного пространства полно (теорема 43). Мы получили новое доказательство теоремы 50.

Непрерывность локально равномерного предела последовательности непрерывных функций

Теорема 65. Пусть E и F — два метрических пространства и $f_0, f_1, f_2, \dots, f_n, \dots$ — последовательность отображений E в F , локально равномерно сходящаяся к f . Если все функции f_n непрерывны в точке a из E , то и предельная функция непрерывна в точке a . Если функции f_n непрерывны всюду, то f также непрерывна всюду. Если сходимость равномерна на E и все f_n равномерно непрерывны на E , то f также равномерно непрерывна на E .

Доказательство. Пусть сходимость равномерна в окрестности \mathcal{U}_a точки a . По определению, для заданного $\varepsilon > 0$ существует такое m , что для всех x из \mathcal{U}_a имеет место неравенство

$$d(f_m(x), f(x)) \leq \frac{\varepsilon}{3}. \quad (\text{II}, 15; 16)$$

Зафиксировав m , заметим, что функция f_m предполагалась непрерывной в точке a . Следовательно, существует окрестность $\mathcal{V} \subset \mathcal{U}_a$ точки a , такая, что для всех $x \in \mathcal{V}$ имеет место неравенство

$$d(f_m(x), f_m(a)) \leq \frac{\varepsilon}{3}. \quad (\text{II}, 15; 17)$$

Но тогда для любого $x \in \mathcal{V}$

$$d(f(x), f(a)) \leq d(f(x), f_m(x)) + d(f_m(x), f_m(a)) + d(f_m(a), f(a)) \leq \varepsilon, \quad (\text{II}, 15; 18)$$

т. е. функция f непрерывна в точке a .

Из доказательства следует, что если f_n всюду непрерывны, то f также всюду непрерывна¹⁾.

Предположим теперь, что f_n равномерно непрерывны, а сходимость равномерна на E . Зададим $\varepsilon > 0$ и выберем m так, чтобы для всех x из E выполнялось неравенство (II, 15; 16). Функция f_m равномерно непрерывна. Следовательно, существует такое число $\eta > 0$, при котором из $d(x', x'') \leq \eta$ следует неравенство $d(f_m(x'), f_m(x'')) \leq \varepsilon/3$. Для $d(x', x'') \leq \eta$ при этом имеем

$$\begin{aligned} d(f(x'), f(x'')) &\leq d(f(x'), f_m(x')) + d(f_m(x'), f_m(x'')) + \\ &\quad + d(f_m(x''), f(x'')) \leq \varepsilon, \quad (\text{II}, 15; 19) \end{aligned}$$

что и доказывает равномерную непрерывность f .

Замечание. Напротив, последовательность непрерывных функций f_n вполне может сходиться просто к разрывной функции f . Так, например, если взять $f_n(x) = x^n$ для $0 \leq x \leq 1$, то f_n непрерывны для всех n и сходятся к разрывной функции f , равной 0 для $0 \leq x < 1$ и 1 для $x = 1$.

Следствие 1. В пространстве $(F^E)_b$ ограниченных отображений E в F , снабженном метрикой, определенной формулой (II, 15; 3), подпространство $(F^E)_{bc}$ ограниченных непрерывных отображений E в F замкнуто.

Это утверждение является другой формулировкой теоремы.

Следствие 2. Если F полно, то пространство $(F^E)_{bc}$ ограниченных непрерывных отображений E в F , снабженное метрикой (II, 15; 3), является полным метрическим пространством.

В самом деле, так как оно замкнуто в полном метрическом пространстве $(F^E)_b$ (теорема 64), то для доказательства достаточно применить теорему 43.

Следствие 3. Если F является пространством Банаха, то пространство $(F^E)_{bc}$ непрерывных ограниченных отображений E в F , снабженное нормой, определяемой по формуле (II, 15; 6), является пространством Банаха.

Особенно важен следующий частный случай. Возьмем в качестве F поле скаляров \mathbb{K} . Тогда пространство вещественных

¹⁾ Что касается первых двух результатов, то E может быть топологическим не обязательно метризуемым пространством.

или комплексных функций, ограниченных и непрерывных на метрическом пространстве E , является пространством Банаха.

В частности, пространство вещественных или комплексных непрерывных и ограниченных функций одной вещественной переменной является пространством Банаха.

Если E компактно (в большинстве практических приложений это будет замкнутый или открытый интервал вещественной прямой), то, согласно теореме 29, получаем, что всякая векторная функция, непрерывная на E , заведомо ограничена¹⁾). Тогда пространство $(F^E)_c$ функций, непрерывных на некотором компакте E , со значениями в банаевом пространстве F и с нормой (II, 15; 6) является пространством Банаха.

Некоторые контрпримеры

Мы теперь в состоянии дать контрпримеры, о которых говорилось в предыдущих параграфах.

1°) Обозначим через $\mathcal{C}([a, b])$ векторное (бесконечномерное) пространство вещественных или комплексных функций, непрерывных на замкнутом ограниченном интервале $[a, b]$, $a < b$, множества \mathbb{R}^2). Покажем, как в этом пространстве можно определить не эквивалентные друг другу нормы. Об этом мы говорили после теоремы 13. В качестве первой нормы возьмем ту, которая была определена формулой (II, 15; 6). Вторую норму определим, используя понятие интеграла:

$$N(f) = \int_a^b |f(x)| dx^3. \quad (\text{II, 15; 20})$$

Эти нормы не эквивалентны. В самом деле, с одной стороны, мы имеем оценку

$$N(f) \leq (b - a) \|f\|. \quad (\text{II, 15; 21})$$

С другой стороны, очевидно, не существует оценки вида $\|f\| \leq kN(f)$, где k — не зависящая от непрерывной функции f .

¹⁾ Мы это утверждаем только для вещественной функции. Однако если \vec{f} является векторной непрерывной функцией, то ее норма $\|\vec{f}\|$ является непрерывной вещественной, а, значит, и ограниченной функцией.

²⁾ $\mathcal{C}([a, b])$ можно записать также в виде $(K^{[a, b]})_c$, где K — поле вещественных или комплексных чисел.

³⁾ Это число, очевидно, ≥ 0 . Неравенство треугольника и соотношение $N(\lambda f) = |\lambda| N(f)$ очевидны. Если $f \neq 0$, т. е. если $f \neq 0$, то существует по крайней мере одна точка c на $[a, b]$, в которой $f(c) \neq 0$. Тогда, в силу непрерывности, существует некоторый интервал, окружающий c , где $|f(x)| > 0$, а, значит, и $N(f) > 0$.

постоянная. В самом деле, рассмотрим случай $[a, b] = [0, 1]$ и последовательность функций, указанных в примере 2 на стр. 151. Для этих функций имеем следующие нормы:

$$\|f_n\| = n^a, \quad N(f_n) = \frac{1}{2} n^{a-1}, \quad \text{откуда} \quad \frac{\|f_n\|}{N(f_n)} = 2n, \quad (\text{II, 15; 22})$$

что и доказывает наше утверждение. Если, например, $0 < \alpha < 1$, то f_n стремится к 0 при n , стремящемся к $+\infty$, по норме N , но не по норме $\|\cdot\|$. Легко доказывается, что в векторном пространстве $\mathcal{C}([a, b])$ существует бесконечное множество других попарно не эквивалентных норм.

2°) Покажем теперь, что в пространстве $\mathcal{C}([0, 1])$, снабженном нормой $\|\cdot\|$, единичный шар не компактен. Для этого достаточно рассмотреть последовательность функций f_n , определенных формулой $f_n(x) = x^n$ (см. замечание на стр. 156). Имеем:

$$f_n(x) = x^n, \quad \|f_n\| = 1 \quad \text{для } n \geq 1. \quad (\text{II, 15; 23})$$

Поскольку эта последовательность просто сходится к функции, равной 0 для $0 \leq x < 1$ и 1 для $x = 1$, то любая ее равномерно сходящаяся подпоследовательность должна сходиться к этому же пределу, но этого не может быть, так как пределом является разрывная функция. Мы нашли некоторую последовательность, принадлежащую единичному шару в $\mathcal{C}([0, 1])$, никакая подпоследовательность которой не сходится. Согласно теореме Больцано — Вейерштрасса (теорема 25), единичный шар не компактен. Итак, пространство $\mathcal{C}([0, 1])$ не является локально компактным. Как мы видели (теорема 45₂), это верно для любого бесконечномерного нормированного векторного пространства.

Теорема 66. Пусть E и F — два метрических пространства, A — некоторая часть E и $f_0, f_1, \dots, f_n, \dots$ — последовательность отображений A в F , равномерно сходящаяся к f . Пусть $a \in E$ — точка приложения A ¹⁾. Если для каждого n функции $f_n(x)$ имеют предел при x , стремящемся к a по значениям в A , и если F полно, то $f(x)$ имеет предел при x , стремящемся к a по значениям в A , и, кроме того,

$$\lim_{\substack{x \rightarrow a \\ x \in A}} f(x) = \lim_{n \rightarrow \infty} [\lim_{\substack{x \rightarrow a \\ x \in A}} f_n(x)]. \quad (\text{II, 15; 24})$$

Доказательство. Рассмотрим для каждого n функцию \tilde{f}_n , определенную на объединении $A \cup \{a\}$ следующим образом: $\tilde{f}_n(x) = f_n(x)$ для $x \neq a$ и $\tilde{f}_n(a) = \lim_{\substack{x \rightarrow a \\ x \in A}} f_n(x)$. По определению,

¹⁾ Никакого предположения относительно принадлежности точки a к множеству A не делается.

функция \tilde{f}_n непрерывна в точке a . С другой стороны, каким бы ни было $\varepsilon > 0$, существует такое целое p , что из неравенств $m \geq p$ и $n \geq p$ следует неравенство $d(f_m(x), f_n(x)) \leq \varepsilon$, а также неравенство $d(f(x), f_n(x)) \leq \varepsilon$ для $x \in A$, $x \neq a$. Устремляя x к a , в пределе получаем $d(\tilde{f}_m(a), \tilde{f}_n(a)) \leq \varepsilon$. Это означает, что $\tilde{f}_n(a)$ образуют последовательность Коши в F , а так как F по предположению полно, то она имеет предел, который мы обозначим через $\tilde{f}(a)$. Функция \tilde{f} , определенная формулой $\tilde{f}(x) = f(x)$ для $x \in A$, $x \neq a$, и принимающая только что определенное значение $\tilde{f}(a)$ в точке a , является функцией, определенной на множестве $A \cup \{a\}$, со значениями в F . При n , стремящемся к $+\infty$, $f_n(x)$ стремится к $f(x)$ для $x \neq a$. Но $\tilde{f}_n(a)$ стремится к $\tilde{f}(a)$ по определению $\tilde{f}(a)$, а, следовательно, \tilde{f}_n сходится просто к \tilde{f} . Однако, \tilde{f}_n сходится, кроме того, и равномерно, ибо для $m \geq p$, $n \geq p$ при $x \neq a$ имеем $d(\tilde{f}(x), \tilde{f}_n(x)) = d(f(x), f_n(x)) \leq \varepsilon$, а также $d(\tilde{f}_m(a), \tilde{f}_n(a)) \leq \varepsilon$. Поэтому, устремляя m к $+\infty$ и переходя к пределу, получаем неравенство $d(\tilde{f}(a), \tilde{f}_n(a)) \leq \varepsilon$, т. е. окончательно $d(\tilde{f}, \tilde{f}_n) \leq \varepsilon$.

Так как \tilde{f}_n непрерывны в точке a , то из теоремы 65 следует, что \tilde{f} также непрерывна в точке a , т. е. $f(x)$ имеет предел при x , стремящемся к a по значениям из A , и что этот предел равен $\tilde{f}(a) = \lim_{n \rightarrow \infty} \tilde{f}_n(a) = \lim_{n \rightarrow \infty} [\lim_{x \in A} f_n(x)]$. Теорема доказана.

Пример. Если f_n и f являются вещественными непрерывными функциями на вещественной прямой \mathbb{R} , если f_n сходится равномерно к f при n , стремящемся к $+\infty$, и если при любом n существует $\lim_{x \rightarrow +\infty} f_n(x) = l_n$, то существует $\lim_{x \rightarrow +\infty} f(x) = l$ и l_n сходятся к l .

Ряды функций со значениями в нормированном векторном пространстве

Пусть $\vec{u}_0, \vec{u}_1, \dots, \vec{u}_n, \dots$ — отображения множества E в нормированное векторное пространство F . Тогда можно рассмотреть ряд $\sum_{n=0}^{\infty} \vec{u}_n$. Говорят, что этот ряд сходится и имеет сумму \vec{S} (отображение E в F), если для каждого x из E ряд векторов $\sum_{n=0}^{\infty} \vec{u}_n(x)$ из F сходится в нормированном векторном пространстве F к сумме $\vec{S}(x)$. Это означает еще, что последовательность частных сумм $\vec{S}_n = \vec{u}_0 + \vec{u}_1 + \dots + \vec{u}_n$ является последовательностью функций на E со значениями в F , сходя-

щейся просто к S . Говорят, что ряд равномерно сходится, если последовательность частных сумм \vec{S}_n является равномерно сходящейся.

Если F является пространством Банаха, то при использовании понятий, связанных с абсолютной сходимостью или нормальной сходимостью, можно допустить некоторую неточность. Когда речь идет о рядах из векторов нормированного векторного пространства, мы можем не различать термины «ряд абсолютно сходится» или «ряд нормально сходится». Однако этого нельзя делать в случае, когда речь идет о функциональных рядах.

Говорят, что ряд просто *абсолютно сходится*, если для каждого x ряд из норм в банаевом пространстве F , т. е. ряд

$\sum_{n=0}^{\infty} \|\vec{u}_n(x)\|$, сходится. Это означает, что ряд $\sum_{n=0}^{\infty} \|\vec{u}_n\|^1$, составленный из функций, определенных на E , с вещественными положительными значениями просто сходится.

Говорят, что ряд *нормально сходится*, если ряд из норм в нормированном векторном пространстве $(F^E)_b$, а именно $\sum_{n=0}^{\infty} \|\vec{u}_n\|$, является сходящимся рядом с положительными членами.

Последнее понятие, очевидно, более сильное. Всякий нормально сходящийся ряд является и просто абсолютно сходящимся. Обратное же, вообще говоря, не верно. В каждом из этих случаев ряд просто сходится (F предполагается полным!). Кроме того, если ряд сходится нормально, то он сходится и равномерно. Нормальная сходимость является наиболее важным критерием равномерной сходимости рядов векторных функций. Этот факт часто выражают следующим образом.

Пусть задан ряд $\sum_{n=0}^{\infty} \vec{u}_n$ функций, определенных на E , со значениями в пространстве Банаха. Если существует такой ряд, составленный из положительных вещественных чисел $a_n \geq 0$, что

$\|\vec{u}_n(x)\| \leq a_n$ для любого x из E и $\sum_{n=0}^{\infty} a_n < \infty$, то данный ряд нормально, а, значит, и равномерно сходится. Этот критерий может быть полезен также и как критерий равномерной условной сходимости. В общем случае функционального ряда $\sum_{n=0}^{\infty} \vec{u}_n$ доказательство его равномерной сходимости можно начать с до-

¹⁾ Напомним, что $\|\vec{u}_n\|$ является функцией $x \rightarrow \|\vec{u}_n(x)\|$ и что $\|\vec{u}_n\| = \sup_{x \in E} \|\vec{u}_n(x)\|$.

казательства сходимости для каждого x из E . Затем вычисляется остаток $\vec{R}_m = \vec{u}_{m+1} + \vec{u}_{m+2} + \dots$ ряда и доказывается, что последовательность функций \vec{R}_m равномерно сходится к 0 при m , стремящемся к $+\infty$. С этой целью используется подходящая оценка остатка.

Приведем пример, представляющий интерес с точки зрения теории рядов Тейлора.

Теорема 67. Если ряд Тейлора (с комплексными коэффициентами) $\sum_{n=0}^{\infty} a_n x^n$ сходится в точке $x = R$ своего круга сходимости, то он равномерно сходится на всем интервале $[0, R]$. В частности, его сумма является непрерывной функцией на интервале $[0, R]$.

Доказательство. В самом деле, имеет место формула $a_n x^n = (a_n R^n) (x/R)^n$. Поэтому для каждого x из $[0, R]$ мы можем применить критерий сходимости Абеля¹⁾. Действительно, если положить $u_n = (x/R)^n$, $v_n = a_n R^n$ и за билинейное отображение B принять произведение, то мы увидим, что эти величины удовлетворяют условиям Абеля и формула (II, 14; 31) дает оценку остатка:

$$U_{m+1}(x) = \left| \left(\frac{x}{R} \right)^{m+1} - \left(\frac{x}{R} \right)^{m+2} \right| + \left| \left(\frac{x}{R} \right)^{m+2} - \left(\frac{x}{R} \right)^{m+3} \right| + \dots = \\ = \left(\left(\frac{x}{R} \right)^{m+1} \left(1 - \frac{x}{R} \right) \right) \left(1 + \frac{x}{R} + \left(\frac{x}{R} \right)^2 + \dots \right) = \left(\frac{x}{R} \right)^{m+1}, \quad (\text{II, 15; 26})$$

$$V_{m+1}(x) = V_{m+1} = \sup_{n \geq m+1} |a_{m+1} R^{m+1} + \dots + a_n R^n|, \\ R_{m+1}(x) \leq \left(\frac{x}{R} \right)^{m+1} \sup_{n \geq m+1} |a_{m+1} R^{m+1} + \dots + a_n R^n|. \quad (\text{II, 15; 27})$$

Тот же самый критерий Абеля не применим при $x = R$, так как последовательность $\left(\frac{x}{R} \right)^n = 1$, имеющая постоянную ограниченную вариацию, к нулю не сходится. Однако ряд предполагался сходящимся, и оценка остатка сохраняется по определению остатка. В результате получаем, что эта оценка справедлива для всех x из интервала $[0, R]$. Поскольку при m , стремящемся к $+\infty$, величина V_{m+1} стремится к 0, в силу критерия

¹⁾ Может показаться бессмысленным применять критерий, относящийся к условно сходящимся рядам, в области, где степенной ряд, мажорируемый геометрической прогрессией, абсолютно сходится! Но мы хотим доказать равномерную сходимость, и, поскольку при $x = R$ ряд не предполагается абсолютно сходящимся, другими методами мы не располагаем.

Коши сходимости числового ряда $\sum_{n=0}^{\infty} a_n R^n$, и $U_{m+1} \leq 1$, то видно, что остаток R_{m+1} сходится равномерно к 0 при m , стремящемся к $+\infty$. Тем самым доказана равномерная сходимость ряда Тейлора. Непрерывность суммы теперь вытекает из теоремы 65.

Доказанная теорема часто используется в различных областях математики. Рассмотрим, например, разложение Тейлора

$$\ln(1+x) = \frac{x}{1} - \frac{x^2}{2} + \dots + (-1)^{n-1} \frac{x^n}{n} + \dots \quad \text{для } |x| < 1. \quad (\text{II, 15; 28})$$

При $x = 1$, согласно теореме о знакопеременных рядах, этот ряд сходится. Следовательно, он представляет некоторую непрерывную функцию в интервале $[0, 1]$, а его сумма равна $\lim_{\substack{x \rightarrow 1 \\ x < 1}} \ln(1+x) = \ln 2$. При $x = 1$ отсюда получаем следующую формулу:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \dots + \frac{(-1)^{n-1}}{n} + \dots ^1). \quad (\text{II, 15; 29})$$

Точно так же, рассматривая разложение в ряд функции

$$\arctg x = \frac{x}{1} - \frac{x^3}{3} + \frac{x^5}{5} - \dots + (-1)^n \frac{x^{2n+1}}{2n+1} + \dots$$

и проводя те же самые рассуждения для $x = 1$, получаем формулу:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \dots + (-1)^n \frac{1}{2n+1} + \dots \quad (\text{II, 15; 30})$$

§ 16. БЕСКОНЕЧНЫЕ ПРОИЗВЕДЕНИЯ ВЕЩЕСТВЕННЫХ ИЛИ КОМПЛЕКСНЫХ ЧИСЕЛ И ФУНКЦИЙ

Определение. Пусть u_0, u_1, u_2, \dots — последовательность вещественных или комплексных чисел. Говорят, что **бесконечное произведение** $\prod_{n=0}^{\infty} u_n$ **сходится**, если при n , стремящемся к $+\infty$, сходится к некоторому конечному числу $\neq 0$ последовательность частных произведений $\Pi_n = \prod_{0 \leq m \leq n} u_m$.

¹⁾ В этом частном случае можно непосредственно воспользоваться теоремой о знакопеременных рядах. Для каждого $x \in [0, 1]$ эта теорема применима; значит, остаток $R_{m+1}(x)$ мажорируется по модулю первым отброшенным членом $\frac{x^{m+1}}{m+1} \leq \frac{1}{m+1}$ и потому действительно сходится равномерно к 0.

Во всех других случаях произведение называется *расходящимся*. Кажется довольно странным считать расходящимся произведение, в котором частные произведения Π_n сходятся к нулю. Но это вызвано, как мы увидим позже, многими причинами. Если хотя бы один из сомножителей u_n является нулем, то произведение, конечно, расходится. Если все числа u_n вещественны и ≥ 1 , то Π_n образуют возрастающую последовательность, а, значит, мы имеем предел, конечный или равный $+\infty$.

В том случае, когда произведение расходится, пишут $\prod_{n=0}^{\infty} u_n = +\infty$. Точно так же каждый раз, когда Π_n стремится к 0, для расходящегося произведения пишут $\prod_{n=0}^{\infty} u_n = 0$.

Теорема 68. Для того чтобы бесконечное произведение $\prod_{n=0}^{\infty} u_n$ было сходящимся, необходимо, чтобы его общий член u_n стремился к 1 при n , стремящемся к $+\infty$.

Доказательство. В самом деле, если бесконечное произведение сходится и его значением является $\Pi \neq 0$, то произведения Π_n и Π_{n-1} оба сходятся к Π , откуда следует, что их отношение u_n при n , стремящемся к $+\infty$, сходится к $\Pi/\Pi = 1$. Заметим сразу, что это свойство не сохраняется для бесконечных произведений, частные произведения которых Π_n сходятся к 0. Например, если мы рассмотрим бесконечное произведение $(1)(1/2)(1/3)\dots(1/n)\dots$, частные произведения которого, очевидно, стремятся к 0, то его общий член $1/n$ к 1 не сходится.

Замечание. В случае сходящегося бесконечного произведения остатком R_m называют произведение $\prod_{n \geq m+1} u_n$. Остаток R_m стремится к 1 при m , стремящемся к $+\infty$.

Для того чтобы бесконечное произведение было сходящимся, необходимо и достаточно, чтобы выполнялся критерий Коши: отношение Π_n/Π_m должно стремиться к 1, когда m и n стремятся к $+\infty$ (все члены u_n предполагаются $\neq 0$). Сходящееся произведение, очевидно, удовлетворяет этому критерию. Обратно, пусть бесконечное произведение удовлетворяет критерию Коши. Тогда существует целое p , такое, что при $n \geq p$ выполняется неравенство $|\Pi_p - \Pi_n| \leq |\Pi_p|$. Из неравенства $|\Pi_n| \leq |\Pi_n - \Pi_p| + |\Pi_p| \leq 2|\Pi_p|$ следует, что все Π_n ограничены. Пусть M — их точная верхняя грань. Тогда для заданного $\varepsilon > 0$, $\varepsilon \leq M/2$, можно определить q так, чтобы при любых $m \geq q$ и $n \geq q$ выполнялось неравенство $|\Pi_m - \Pi_n| \leq \frac{\varepsilon}{M} |\Pi_m| \leq \varepsilon$. По-

следовательность чисел Π_n является, следовательно, последовательностью Коши в поле комплексных чисел C . Посколько это поле полно, то Π_n имеет предел Π . Устремляя в последнем неравенстве n к $+\infty$, получаем

$$|\Pi_m - \Pi| \leq \frac{\epsilon}{M} |\Pi_m| \leq \frac{1}{2} |\Pi_m|,$$

т. е. $\Pi \neq 0$, а значит, бесконечное произведение сходится.

Бесконечные произведения и логарифмические ряды

При изучении сходимости или расходимости бесконечного произведения представляется заманчивым вычислить логарифм его членов с тем, чтобы заменить произведение рядом. Если члены вещественны, то логарифмы можно вычислить, начиная с того момента, когда все $u_n > 0$.

Если такого номера, начиная с которого все $u_n > 0$, не существует, то это означает, что общий член u_n при n , стремящемся к $+\infty$, к 1 не сходится. Но тогда сразу ясно, что произведение расходится, и его изучение заканчивается.

Если же при n , стремящемся к $+\infty$, общий член u_n стремится к 1, то, начиная с некоторого номера, все u_n строго положительны, и можно вычислять их логарифм.

Предположим, что u_n комплексны. Известно, как сложно брать логарифм комплексного числа, ибо каждое комплексное число имеет бесконечное множество логарифмов. Если положить $z = re^{i\theta}$, то мы получим общую формулу: $\ln z = \ln r + i\theta$, где θ определено с точностью до сомножителя, кратного 2π . Предположим, однако, что z изменяется в полуплоскости $x = \operatorname{Re} z > 0$. Тогда можно выбрать его аргумент между $-\pi/2$ и $\pi/2$ и затем определить по соответствующей формуле его логарифм. Так определенный логарифм является непрерывной функцией. Его называют главным значением логарифма и обозначают через \ln^1 .

В частности, если v — такое вещественное число, что $|v| < 1$, то $1 + v$ находится в предыдущей полуплоскости, и ранее определенный логарифм можно записать в виде разложения Тейлора

$$\ln(1 + v) = \frac{v}{1} - \frac{v^2}{2} + \dots + (-1)^{n-1} \frac{v^n}{n} + \dots \quad (\text{II}, 16; 1)$$

Пусть теперь $\prod_{n=0}^{\infty} u_n$ — бесконечное произведение комплексных чисел. Если u_n не стремится к 1 при n , стремящемся к $+\infty$,

¹⁾ Можно также определить главное значение логарифма в дополнении к полупрямой ≤ 0 в плоскости. Тогда $-\pi < \arg z < +\pi$ и $-i\pi < \operatorname{Im}(\ln z) < i\pi$.

то произведение расходится. Если же u_n стремится к 1 при n , стремящемся к $+\infty$, то, начиная с некоторого n , числа u_n находятся в полуплоскости $\operatorname{Re} u_n > 0$ и, опустив конечное число членов, можно взять логарифмы. При этом мы сразу же получаем такую теорему:

Теорема 69. Для того чтобы бесконечное произведение $\prod_{n=0}^{\infty} u_n$, для всех членов которого $\operatorname{Re} u_n > 0$, было сходящимся, необходимо и достаточно, чтобы сходился ряд $\sum_{n=0}^{\infty} \ln u_n$.

Предположим сначала, что бесконечное произведение сходится. Обозначим через Π_n частные произведения и через S_n частные суммы ряда логарифмов. Не обязательно $\ln \Pi_n = S_n$, и нельзя даже утверждать, что $\operatorname{Re} \Pi_n > 0$. Однако, согласно критерию Коши, существует такое целое число p , что при $n \geq p$ имеем $|\Pi_n/\Pi_p - 1| < 1$, а, следовательно, $\operatorname{Re}(\Pi_n/\Pi_p) > 0$, а также $\operatorname{Re}(\Pi/\Pi_p) > 0$. Зафиксируем такое p . Тогда $\ln(\Pi_n/\Pi_p) = S_n - S_p + 2k_n i\pi$. Поскольку Π_n/Π_p сходится к Π/Π_p , то из непрерывности главного значения логарифма следует, что $S_n - S_p + 2k_n i\pi$ при n , стремящемся к $+\infty$, сходится к пределу, а, значит, ряд с общим членом $\ln u_n + 2(k_n - k_{n-1})i\pi$ является сходящимся и его общий член стремится к нулю. Однако, так как u_n в силу того, что произведение сходится, стремится к 1, то $\ln u_n$ также сходится к 0. Начиная с некоторого значения n , мы получаем $k_{n-1} = k_n$, а, значит, ряд с общим членом $\ln u_n$ является также сходящимся.

Обратно, если ряд сходится, то S_n сходится к некоторому пределу S . В силу непрерывности показательной функции получаем, что числа $\Pi_n = e^{S_n}$ сходятся к $\Pi = e^S \neq 0$. Теперь видно, почему в доказательстве было существенным предположение $\Pi \neq 0$. Впрочем, если u_n вещественны > 0 и Π_n стремятся к 0, то $\ln u_n$ будет являться общим членом расходящегося ряда с суммой $-\infty$. Говорят, что *бесконечное произведение, которое по предположению сходится, абсолютно сходится (соответственно условно сходится)*, если ряд из логарифмов его членов (который будет определенным начиная с некоторого места) является абсолютно сходящимся (соответственно условно сходящимся). Естественно, что здесь используется весьма вольный оборот речи. Сказать, что произведение абсолютно сходится, вовсе не значит сказать, что произведение $\prod |u_n|$ сходится. (Поскольку $\prod_{n=0}^{\infty} |u_n| = \left| \prod_{n=0}^{\infty} u_n \right|$, то произведение абсолютно сходится, если оно сходится.) Если все u_n вещественны и ≥ 1 или заключены

между 0 и 1, то сходимость является синонимом абсолютной сходимости.

Теорема 70. Для того чтобы бесконечное произведение $\prod_{n=0}^{\infty} (1 + v_n)$, $v_n \neq -1$ ¹⁾, было абсолютно сходящимся, необходимо и достаточно, чтобы был сходящимся ряд $\sum_{n=0}^{\infty} |v_n|$.

Доказательство. Если произведение абсолютно сходится, то оно сходится просто, а, следовательно, $1 + v_n$ стремится к 1, и для достаточно больших $n \geq p$ можно вычислять логарифм членов произведения. Ряд $\sum_{n \geq p} |\ln(1 + v_n)|$ по предположению сходится. В сходящихся же рядах с положительными членами при исследовании сходимости общий член ряда можно заменять на эквивалентный ему. При n , стремящемся к $+\infty$, v_n стремится к 0, а тогда $\ln(1 + v_n) \sim v_n$ и, следовательно, ряд $\sum_{n \geq p} |v_n|$ также сходится, а вместе с ним и ряд $\sum_{n=0}^{\infty} |v_n|$.

Обратно, предположим, что ряд $\sum_{n=0}^{\infty} |v_n|$ сходится. Тогда v_n стремится к 0, а, следовательно, для $n \geq p$, при достаточно большом p , $\operatorname{Re}(1 + v_n) > 0$ и можно вычислить логарифмы чисел $1 + v_n$. Так как при n , стремящемся к $+\infty$, $|\ln(1 + v_n)| \sim |v_n|$, то ряд $\sum_{n \geq p} |\ln(1 + v_n)|$ сходится. Следовательно, сходится ряд $\sum_{n \geq p} \ln(1 + v_n)$, а, значит, согласно теореме 69, сходится произведение $\prod_{n \geq p} (1 + v_n)$. Поскольку все $v_n \neq -1$, то произведение $\prod_{n=0}^{\infty} (1 + v_n)$, не имеющее нулевых членов, также сходится.

Поскольку $\prod_{n=0}^{\infty} (1 + v_n)$ сходится и сходится ряд $\sum_{n \geq p} |\ln(1 + v_n)|$, рассматриваемое произведение сходится абсолютно.

Пример. Бесконечное произведение

$$\prod_{n \geq 1} \left(1 + \frac{1}{n^a}\right) \quad \text{или} \quad \prod_{n \geq 2} \left(1 - \frac{1}{n^a}\right) \quad (\text{II}, 16; 2)$$

¹⁾ Если не налагать ограничения $v_n \neq -1$, то может случиться, что $\sum |v_n| < +\infty$, а бесконечное произведение имеет нулевой член, т. е. расходится.

сходится, если $\alpha > 1$, и расходится, если $\alpha \leq 1$. В частности, оно расходится при $\alpha = 1$. Заметим, что последний случай практически очевиден, ибо здесь легко вычисляются частные произведения

$$\left(1 + \frac{1}{1}\right)\left(1 + \frac{1}{2}\right)\left(1 + \frac{1}{3}\right) \dots \left(1 + \frac{1}{n}\right) = n + 1, \quad (\text{II}, 16; 3)$$

$$\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{4}\right) \dots \left(1 - \frac{1}{n}\right) = \frac{1}{n}.$$

Можно даже сказать, что поскольку расходимость произведения $\prod_{n=1}^{\infty} \left(1 + \frac{1}{n}\right)$ или $\prod_{n=2}^{\infty} \left(1 - \frac{1}{n}\right)$ более очевидна, чем расходимость ряда $\sum_{n=1}^{\infty} \frac{1}{n}$, то с помощью теоремы 70 можно доказать расходимость этого ряда!

Бесконечные произведения вещественных или комплексных функций

Пусть $u_0, u_1, \dots, u_n, \dots$ — последовательность функций, определенных на множестве E с вещественными или комплексными значениями. Бесконечное произведение $\prod_{n=0}^{\infty} u_n$ называется просто сходящимся, если для любого x из E сходится бесконечное произведение $\prod_{n=0}^{\infty} u_n(x)$ вещественных или комплексных чисел. Это означает, что последовательность функций $\Pi_n = \prod_{0 \leq m \leq n} u_m$ просто сходится к нигде не равной нулю предельной функции.

Произведение называется абсолютно просто сходящимся, если для любого x из E числовое произведение $\prod_{n=0}^{\infty} u_n(x)$ абсолютно сходится.

Выражение «равномерная сходимость» для произведения функций не ясно. Фраза о том, что произведение комплексных функций $\prod_{n=0}^{\infty} u_n$ на E сходится равномерно к функции Π , может означать, что Π_n и Π нигде на E в нуль не обращаются и что функции Π_n равномерно сходятся к функции Π . Та же самая фраза может также означать, что Π_n и Π нигде на E в нуль не обращаются и что отношение Π_n/Π равномерно сходится

к 1. Как легко видеть, эти два понятия не обязательно совпадают. Однако они совпадают, если предел Π имеет на E равномерные оценки снизу и сверху такого вида: $0 < a \leq \Pi(x) \leq b < +\infty$. В самом деле, в этом случае из неравенства $|\Pi_n - \Pi| \leq \varepsilon$ следует неравенство $|\Pi_n/\Pi - 1| \leq \varepsilon/a$, а из неравенства $|\Pi_n/\Pi - 1| \leq \varepsilon$ следует неравенство $|\Pi_n - \Pi| \leq b\varepsilon$. Только в этом случае мы можем позволить себе говорить о равномерной сходимости бесконечного произведения функций. Однако всегда можно говорить о локальной равномерной сходимости бесконечного произведения функций, непрерывных на некотором топологическом пространстве E (и тогда предел Π будет также непрерывным). В самом деле, если функции Π_n сходятся локально равномерно к функции Π , то, согласно теореме 65, функция Π будет непрерывной. Так как она везде отлична от 0, то для каждой точки a существует окрестность \mathcal{U}'_a , в которой $|\Pi|$ ограничена сверху и снизу некоторыми фиксированными положительными числами. Но тогда во всякой окрестности $\mathcal{U}_a \subset \mathcal{U}'_a$, в которой функции Π_n сходятся равномерно к функции Π , отношение Π_n/Π равномерно сходится к 1. Обратно, предположим, что Π_n/Π локально равномерно сходится к 1. Тогда для каждого a из E существуют окрестность \mathcal{U}'_a и число n , такие, что $|\Pi_n(x)/\Pi(x) - 1| \leq 1/2$ для $x \in \mathcal{U}'_a$. Отсюда следует $1/2 \leq |\Pi_n(x)/\Pi(x)| \leq 3/2$, а значит, $\frac{2}{3}|\Pi_n(x)| \leq |\Pi(x)| \leq 2|\Pi_n(x)|$. Поскольку функции Π_n непрерывны и $\Pi_n(a) \neq 0$, то существует окрестность $\mathcal{U}''_a \subset \mathcal{U}'_a$ точки a , в которой функции $|\Pi_n|$ ограничены сверху и снизу положительными постоянными. При этом функция $|\Pi(x)|$ также будет ограниченной. Если теперь $\mathcal{U}_a \subset \mathcal{U}''_a$ является окрестностью, на которой Π_n/Π сходится равномерно к 1, то Π_n будут на ней равномерно сходиться к Π .

Применение к функции ζ Римана

Функция Римана ζ определяется формулой

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \quad (\text{II}, 16; 4)$$

Пусть $s = \sigma + it$ и δ — вещественное число > 0 . Тогда данный ряд, рассматриваемый как ряд функций, определенный в области $\sigma \geq 1 + \delta$ комплексной плоскости, является нормально сходящимся. В самом деле,

$$\left| \frac{1}{n^{\sigma+it}} \right| = \frac{1}{n^\sigma} \leq \frac{1}{n^{1+\delta}}. \quad (\text{II}, 16; 5)$$

Так как для любого n функция $s \rightarrow 1/n^s$ непрерывна в полу-плоскости $\sigma \geq 1 + \delta$, то видно, что сумма, т. е. функция ζ , также непрерывна в этой полуплоскости, а поскольку это верно для любого $\delta > 0$, функция ζ непрерывна во всей полуплоскости $\sigma > 1$.

Рассмотрим теперь бесконечное произведение, в котором p пробегает множество всех простых чисел

$$G(s) = \prod_p \left(\frac{1}{1 - 1/p^s} \right). \quad (\text{II, 16; 6})$$

Любой член этого произведения всегда $\neq 0$. Впрочем, знаменатель $1 - 1/p^s \neq 0$ для $\sigma > 0$. Кроме того, в этом случае при $\sigma > 0$ модулем числа $1/p^s$ является число $1/p^\sigma < 1$ и, следовательно, применима теорема 70.

Бесконечное произведение абсолютно сходится тогда и только тогда, когда сходится ряд $\sum_p 1/p^\sigma$. Последнее же, конечно, будет иметь место для $\sigma > 1$, поскольку сумма этого ряда мажорируется суммой сходящегося ряда $\sum_{n=1}^{\infty} 1/n^\sigma$.

Теорема 71. При $\sigma > 1$ имеет место равенство $G(s) = \zeta(s)$.

Доказательство. При доказательстве этой теоремы естественно считать s раз и навсегда фиксированным. Поскольку ряд $\sum_{n=1}^{\infty} 1/n^\sigma$ и произведение G сходятся, то для любого $\varepsilon > 0$ можно найти целое число m , обладающее следующими свойствами:

- остаток $\sum_{n>m} 1/n^\sigma$ мажорируется числом $\varepsilon/2$;
- если через $G_m(s)$ обозначить частные произведения, образованные из m первых множителей бесконечного произведения, то $|G_m(s) - G(s)| \leq \varepsilon/2$.

Для каждого простого числа p имеет место разложение в абсолютно сходящийся геометрический ряд:

$$\frac{1}{1 - 1/p^s} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \dots = \sum_{k=0}^{\infty} \frac{1}{p^{ks}}. \quad (\text{II, 16; 7})$$

В силу правила относительно произведения нескольких абсолютно сходящихся рядов (теорема 61), можно записать:

$$G_m(s) = \sum_{k_1, k_2, \dots, k_m} \left(\frac{1}{p_1} \right)^{k_1 s} \dots \left(\frac{1}{p_m} \right)^{k_m s}, \quad (\text{II, 16; 8})$$

где $p_1 = 2, p_2 = 3, p_3 = 5, \dots, p_m$ суть m первых простых чисел и k_1, k_2, \dots, k_m — целые числа ≥ 0 . Отсюда следует, что

$$G_m(s) = \sum_v \frac{1}{v^s}, \quad (\text{II, 16; 9})$$

где v пробегает последовательность всех целых чисел, которые в разложении на простейшие множители содержат только простые числа p_1, p_2, \dots, p_m . Рассмотрим теперь разность $G_m(s) - \zeta(s)$; она состоит из части членов ряда $\sum_n \frac{1}{n^s}$, соответствующих индексам $n > m$, и потому

$$|G_m(s) - \zeta(s)| \leq \sum_{n>m} \frac{1}{n^\sigma} \leq \frac{\epsilon}{2}, \quad (\text{II, 16; 10})$$

откуда вытекает неравенство

$$|G(s) - \zeta(s)| \leq |G(s) - G_m(s)| + |G_m(s) - \zeta(s)| \leq \epsilon. \quad (\text{II, 16; 11})$$

Поскольку ϵ произвольно, то получаем, что $G(s) = \zeta(s)$.

Следствие. При $\sigma > 1$ функция ζ в нуль никогда не обращается.

Это очевидно, так как она равна значению сходящегося бесконечного произведения. Предыдущие результаты, очевидно, не верны для $\sigma = 1$. В частности, расходится ряд $\sum_{n=1}^{\infty} 1/n$. Докажем точно так же следующее утверждение:

Теорема 72. Бесконечное произведение $\prod_p \frac{1}{1-1/p}$ расходится.

Заметим для этого, что для произвольного числа $A > 0$ можно найти такое целое m , что

$$1 + \frac{1}{2} + \dots + \frac{1}{m} \geq A. \quad (\text{II, 16; 12})$$

Рассмотрим теперь частные произведения G_m . Примененные выше разложения в геометрический ряд имеют смысл и, следовательно, G_m является суммой $\sum_v v^{-1}$, в которой v пробегает все целые числа, разложение которых на сомножители состоит только из простых чисел p_1, p_2, \dots, p_m . Отсюда, в частности, вытекает, что имеет место неравенство

$$G_m(1) = \sum_v \frac{1}{v} \geq 1 + \frac{1}{2} + \dots + \frac{1}{m} \geq A. \quad (\text{II, 16; 13})$$

Поскольку A произвольно, из этого неравенства следует, что рассматриваемое бесконечное произведение (сомножители которого > 1) расходится: $G(1) = +\infty$.

Следствие. Множество простых чисел бесконечно; кроме того, ряд $\sum_p 1/p$, составленный из простых чисел, расходится.

В самом деле, расходимость этого ряда эквивалентна расходимости бесконечного произведения $\prod_p (1 - 1/p) = 1/G(1)$.

Замечание. Рассмотрим знакопеременный ряд

$$\zeta_a(s) = \frac{1}{1^s} - \frac{1}{2^s} + \dots + (-1)^{n-1} \frac{1}{n^s} + \dots \quad (\text{II, 16; 14})$$

Как мы только что видели, этот ряд сходится для $\sigma > 0$. Мы сейчас покажем, что он даже равномерно сходится на каждом компакте открытой полуплоскости $\sigma > 0$ комплексной плоскости. Пусть K — такой компакт. Заметим прежде всего, что на K функция $|s|$ в силу ее непрерывности ограничена сверху некоторым числом S . Точно так же непрерывная всюду положительная функция σ на K ограничена снизу некоторым числом $\delta > 0$.

Применим теперь теорему Абеля (теорема 63). Имеем:

$$\frac{(-1)^{n-1}}{n^s} = u_n v_n, \quad \text{где } v_n = (-1)^{n-1}, \quad u_n = \frac{1}{n^s}.$$

Величины $|\sigma_{m,n}|$ мажорируются числом 1. Покажем, что последовательность чисел u_n имеет ограниченную вариацию. Имеем:

$$\frac{1}{n^s} - \frac{1}{(n+1)^s} = s \int_n^{n+1} \frac{dx}{x^{s+1}}, \quad (\text{II, 16; 15})$$

откуда

$$\left| \frac{1}{n^s} - \frac{1}{(n+1)^s} \right| \leq |s| \int_n^{n+1} \frac{dx}{x^{s+1}} \quad (\text{II, 16; 16})$$

и

$$\left| \frac{1}{n^s} - \frac{1}{(n+1)^s} \right| + \left| \frac{1}{(n+1)^s} - \frac{1}{(n+2)^s} \right| + \dots \leq |s| \int_n^{+\infty} \frac{dx}{x^{s+1}} = \frac{|s|}{\sigma n^\sigma}. \quad (\text{II, 16; 17})$$

¹⁾ Это общий метод для оценки разности; полагают $f(n+1) - f(n) = \int_n^{n+1} f'(t) dt$.

Следовательно, ряд сходится, а формула (II, 14; 31) дает для остатка следующую оценку:

$$\left| \sum_{n \geq m+1} \frac{(-1)^{n-1}}{n^s} \right| \leq \frac{|s|}{\sigma(m+1)^\sigma}. \quad (\text{II, 16; 18})$$

Так как при этом $|R_m| \leq S/[\delta(m+1)^\delta]$, где правая часть не зависит от s и стремится к 0 при m , стремящемся к $+\infty$, то сходимость ряда равномерна на K .

Если считать, что $\sigma > 1$, то можно указать простую связь между функциями ζ и ζ_a . В самом деле, из формулы

$$\zeta_a(s) = \zeta(s) - 2 \sum_{n=1}^{\infty} \left(\frac{1}{2n} \right)^s$$

получаем:

$$\zeta_a(s) = \zeta(s) \left(1 - \frac{1}{2^{s-1}} \right), \quad \text{или} \quad \zeta(s) = \frac{\zeta_a(s)}{1 - \frac{1}{2^{s-1}}}. \quad (\text{II, 16; 19})$$

Свойство равномерной сходимости, доказанное для ζ_a , показывает, что эта функция непрерывна на каждом компакте K открытой полуплоскости $\sigma > 0$. Следовательно, она непрерывна всюду в этой полуплоскости. В частности, при s , стремящемся к 1, $\zeta_a(s)$ стремится к $\zeta_a(1) = \ln 2$. Тогда из формулы (II, 16; 19) следует, что при s , стремящемся к 1, $\zeta(s)$ эквивалентна

$$\frac{\ln 2}{1 - e^{(1-s)\ln 2}} \sim \frac{\ln 2}{(s-1)\ln 2} = \frac{1}{s-1}.$$

В результате формула (II, 16; 19) позволяет продолжить функцию ζ в полуплоскость $\sigma > 0$ ¹⁾.

Другие способы продолжения позволяют определить функцию ζ во всей комплексной плоскости и показать, что это *голоморфная* комплексной переменной s , т. е. непрерывная функция с непрерывной первой производной по отношению к этой комплексной переменной в дополнении к точке $s = 1$ комплексной плоскости. Эта точка $s = 1$ является полюсом, $\zeta(1) = \infty$. Продолженная функция обращается в нуль в точках $s = -2, -4, -6, \dots$. Исследование этой функции дает сведения о распределении простых чисел, как мы это уже видели на простом примере. Риман высказал гипотезу, до сих пор еще не доказанную, о том, что продолженная функция ζ имеет все нули,

¹⁾ Это продолжение, естественно, не имеет больше ничего общего с суммой ряда $\sum_{n=1}^{\infty} 1/n^s$, не имеющей смысла для $s \leq 1$.

кроме предыдущих, на вертикальной полупрямой $\sigma = 1/2$. Доказательство этого утверждения дало бы исключительно важные сведения о распределении последовательности простых чисел. Во всяком случае, уже известные в настоящее время свойства функции ζ позволяют показать, что n -е простое число эквивалентно при n , стремящемся к бесконечности, числу $n \ln n$, или что число простых чисел, заключенных между 1 и N , эквивалентно при N , стремящемся к $+\infty$, числу $N/\ln N$.

Теория простых чисел является одной из самых интересных, но и самых трудных математических теорий.

Дифференциальное исчисление

§ 1. АФФИННЫЕ ПРОСТРАНСТВА

Более или менее интуитивное представление о пространстве дается еще в средней школе. В элементарной геометрии пространство рассматривается как совокупность его элементов, называемых точками. Исходя из точек — элементов этого пространства, — можно с помощью отношения эквивалентности ввести понятие свободных векторов, как это было сделано на стр. 19. Пространство свободных векторов является векторным пространством, обладающим обычными свойствами.

Однако пользоваться далее этими, недостаточно изученными понятиями невозможно, так как пространство в элементарной геометрии не было определено строго. В математике принимается такая схема.

На основе общей теории множеств корректно определяется понятие множества \mathbb{N} целых чисел $\geqslant 0$, а затем множество \mathbb{Z} целых чисел произвольного знака. Далее, исходя из отношений эквивалентности, вводится поле \mathbb{Q} рациональных чисел, как об этом говорилось на стр. 19. В полных математических курсах, опираясь на понятие сечения и исходя из поля \mathbb{Q} рациональных чисел, со всей строгостью определяется поле \mathbb{R} вещественных чисел и, наконец, поле \mathbb{C} комплексных чисел. После этого можно ввести общее понятие абстрактного векторного пространства над некоторым полем (которое будет чаще всего полем вещественных или полем комплексных чисел).

Понятие векторного пространства и его свойства уже были получены нами ранее. Исходя из этого понятия, введем со всей строгостью пространство элементарной геометрии, которое является аффинным евклидовым пространством. Оно очень близко к векторному пространству, но его нуль не обладает никакими свойствами, выделяющими его среди других элементов пространства. В дальнейшем элементы векторного пространства мы будем всегда обозначать буквами с проведенной над ними стрелкой и будем называть их векторами.

Определение. Аффинным пространством E над полем \mathbb{K} вещественных или комплексных чисел называется непустое множество (его элементы называются точками), которому, с одной

стороны, сопоставлено векторное пространство \vec{E} над полем вещественных или комплексных чисел, называемое векторным пространством, присоединенным к E (элементы которого мы будем называть векторами), а, с другой стороны, — некоторое отображение множества $E \times E$ в пространство \vec{E} (его свойства мы уточним ниже).

Если a и b — две точки E , то элементом, соответствующим паре $(a, b) \in E \times E$ при рассматриваемом отображении, является некоторый вектор из \vec{E} . Его обозначают через \vec{ab} и называют вектором с началом в a и концом в b . Отображение $E \times E$ в \vec{E} должно обладать следующими свойствами:

1°) Соотношение Шаля: каковы бы ни были точки a, b, c множества E , должно выполняться равенство

$$\vec{ab} + \vec{bc} + \vec{ca} = \vec{0}, \text{ где } \vec{0} — \text{нулевой вектор } \vec{E}. \quad (\text{III}, 1; 1)$$

Из этого свойства вытекает, в частности (если считать все три точки совпадающими с a), что, какова бы ни была точка $a \in E$, вектор \vec{aa} является нулевым вектором пространства \vec{E} . Если положить $b = c$, то мы получим, что \vec{ab} и \vec{ba} являются противоположными векторами.

2°) Какой бы ни была фиксированная точка a , отображение $x \rightarrow \vec{ax}$ должно быть биекцией E на \vec{E} . Мы примем следующие обозначения.

Прежде всего, вектор \vec{ab} с началом в a и концом в b может быть записан в виде $\vec{b} - \vec{a}$. Соотношение Шаля при этом примет такой вид: $\vec{b} - \vec{a} + \vec{c} - \vec{b} + \vec{a} - \vec{c} = \vec{0}$. Принятое обозначение согласуется с обычными свойствами вычитания. С другой стороны, если a — точка E и \vec{h} — вектор пространства \vec{E} , то из свойства биекции, определенной в 2°), следует, что в E существует, и притом единственная, точка b , такая, что $\vec{b} - \vec{a} = \vec{h}$. Эту точку удобно обозначать через $a + \vec{h}$. При этом, в силу соотношения Шаля, получаем $a + (\vec{h} + \vec{k}) = (a + \vec{h}) + \vec{k}$. Таким образом, новое понятие согласуется с привычными свойствами сложения. Размерность присоединенного векторного пространства называется также размерностью аффинного пространства E . Пустое множество можно рассматривать как аффинное пространство, не имеющее присоединенного векторного пространства.

Векторное пространство является частным случаем аффинного. Достаточно считать оба пространства совпадающими и поставить в соответствие любым двум элементам \vec{a} и \vec{b} векторного пространства вектор $\vec{ab} = \vec{b} - \vec{a}$. В частности, поле скаляров \mathbb{K} само есть аффинное пространство размерности 1.

Общей, или аффинной, системой координат аффинного конечномерного пространства E называется система, образованная точкой 0 (началом) пространства E и некоторым базисом $(\vec{e}_i)_{i \in I}$ ¹⁾ присоединенного векторного пространства \vec{E} . Рассмотрим теперь произвольную точку x из E . Вектор $\overrightarrow{x-0}$ имеет координаты $(x_i)_{i \in I}$, в выбранном базисе \vec{E} . Эти координаты называются координатами точки x в выбранной системе координат, при этом имеет место формула

$$x = 0 + \sum_{i \in I} x_i \vec{e}_i. \quad (\text{III}, 1;2)$$

Аффинные многообразия

Пусть F — непустое подмножество аффинного пространства E , обладающее следующим свойством.

Существует такое векторное подпространство \vec{F} векторного пространства \vec{E} , что для каждой пары (a, b) из $F \times F$ вектор $b - a$ принадлежит \vec{F} и для любой пары (a, \vec{h}) из $F \times \vec{F}$ точка $a + \vec{h}$ принадлежит F . Если такое векторное подпространство \vec{F} существует, то оно единственno, поскольку является множеством всех векторов $b - a$ для всевозможных пар (a, b) из $F \times F$. При этих условиях говорят, что F является *аффинным подпространством* или *аффинным многообразием* (часто даже *линейным многообразием пространства E*) и что \vec{F} является его *присоединенным векторным подпространством*.

Пространство F обладает структурой аффинного пространства, имеющего \vec{F} в качестве присоединенного векторного пространства, а в качестве отображения $F \times F$ в \vec{F} — сужение за-

¹⁾ Если \vec{E} имеет размерность n , то I является множеством «индексов», состоящим из некоторых n элементов, а $(\vec{e}_i)_{i \in I}$ является «семейством» n векторов из \vec{E} . Чаще всего $I = \{1, 2, \dots, n\}$ — это множество n первых целых положительных чисел, а базисом является «последовательность» n векторов из \vec{E} .

данного отображения $E \times E$ в \vec{E} . Условимся считать пустую часть множества E аффинным многообразием, не имеющим присоединенного векторного пространства. Пространство E само по себе является аффинным многообразием. Точка является аффинным многообразием размерности 0. Прямой называется аффинное многообразие размерности 1, плоскостью — аффинное многообразие размерности 2. Векторное подпространство \vec{F} пространства \vec{E} называется гиперплоскостью, если его дополнительные векторные подпространства имеют размерность 1. Аффинное многообразие F из E называется гиперплоскостью, если его присоединенное векторное подпространство является гиперплоскостью. Если E конечномерно размерности n , то гиперплоскость представляет собой аффинное многообразие размерности $n - 1$.

Два аффинных многообразия из E одной и той же размерности называются параллельными, если они имеют одно и то же присоединенное векторное пространство¹⁾. В частности, два совпадающих многообразия параллельны. Легко видеть, что при таком построении теории аффинных пространств так называемый постулат Евклида является теоремой, к тому же очевидной: через каждую точку пространства можно провести аффинное многообразие, параллельное заданному многообразию, и притом единственное. Это означает, что если \vec{F} — векторное подпространство \vec{E} и a — точка из E , то существует, и притом единственное, аффинное многообразие, состоящее из точки a и присоединенного векторного пространства \vec{F} : это множество точек вида $a + \vec{h}$, где $\vec{h} \in \vec{F}$.

Пересечение конечного или бесконечного числа аффинных многообразий аффинного пространства является аффинным многообразием. Отсюда вытекает, что если A является некоторой частью аффинного пространства, то существует наименьшее аффинное многообразие, содержащее A . Это — пересечение всех аффинных многообразий, содержащих A . Его называют аффинным многообразием, порожденным множеством A . Если размерности двух аффинных многообразий равны p и q , через i обозначена размерность их пересечения, а через s размерность аффинного многообразия, порожденного их объединением, то легко доказать, что в случае когда пересечение не пусто, имеет место формула: $p + q = i + s$.

¹⁾ Если E является плоскостью (аффинное пространство двух измерений), то две различные прямые параллельны тогда и только тогда, когда они не пересекаются.

Произведение $E_1 \times E_2$ двух аффинных пространств обладает, очевидно, структурой аффинного пространства, имеющего $\vec{E}_1 \times \vec{E}_2$ в качестве присоединенного векторного пространства, если положить $\overrightarrow{(b_1, b_2) - (a_1, a_2)} = \overrightarrow{(b_1 - a_1, b_2 - a_2)}$.

Линейные отображения. Аффинные отображения

Пусть E и F — два аффинных пространства. Отображение u пространства E в пространство F называется аффинным, если существует такое линейное отображение \vec{u} из \vec{E} в \vec{F} , что

$$\overrightarrow{u(b) - u(a)} = \vec{u}(\vec{b} - \vec{a}). \quad (\text{III}, 1; 3)$$

Отображение \vec{u} , очевидно, единствено, поскольку его значение на произвольном векторе из \vec{E} известно¹⁾. Пусть E и F — два конечномерных аффинных пространства с выбранной в них системой координат, т. е. известно начало a в E и базис $(\vec{e}_j)_{j \in J}$ в \vec{E} и начало b в F и базис $(\vec{f}_i)_{i \in I}$ в \vec{F} . Аффинное отображение u будет полностью определено, если известны координаты $(c_i)_{i \in I}$ точки $u(a)$ в аффинном пространстве F и координаты u_{ij} , $i \in I$, каждого вектора $\vec{u}(\vec{e}_j)$, $j \in J$, в векторном пространстве \vec{F} . Отображение u при этом определяется формулой

$$\begin{aligned} u(a) &= b + \sum_{i \in I} c_i \vec{f}_i, \\ \vec{u}(\vec{e}_j) &= \sum_{i \in I} u_{ij} \vec{f}_i. \end{aligned} \quad (\text{III}, 1; 4)$$

Можно также сказать, что каждой точке $x = a + \sum_{j \in J} x_j \vec{e}_j$ из E с координатами $(x_j)_{j \in J}$ ставится в соответствие точка $y = b + \sum_{i \in I} y_i \vec{f}_i$ из F с координатами $(y_i)_{i \in I}$ по формуле

$$y_i = c_i + \sum_{j \in J} u_{ij} x_j, \quad i \in I. \quad (\text{III}, 1; 5)$$

Если, в частности, F является полем вещественных или полем комплексных чисел, снабженным своей канонической системой координат, состоящей из нуля и единичных векторов,

¹⁾ Очень часто стрелка над линейным присоединенным отображением не ставится и через u одновременно обозначается как аффинное отображение, так и присоединенное к нему линейное. Обозначение \vec{u} может привести к ошибочному представлению о том, что u является вектором из \vec{E} или \vec{F} .

то можно говорить о вещественной или комплексной аффинной функции. Это функция, которая (если в E выбрана система координат) каждой точке x с координатами $(x_j)_{j \in J}$ ставит в соответствие вещественное или комплексное число:

$$u(x) = c + \sum_{j \in J} u_j x_j, \quad c = u(a), \quad u_j = \vec{u}(e_j). \quad (\text{III}, 1; 6)$$

Z Как видно, то, что часто называют линейной функцией, следует называть аффинной функцией, а то, что часто называют линейной однородной функцией (\vec{E} — векторное пространство, $a = \vec{0}$, $c = \vec{0}$), следует называть просто линейной функцией. Функция $y = ax + b$ является аффинной функцией, а $y = ax$ — ее линейной присоединенной функцией.

Если \vec{h} является вектором из векторного пространства, присоединенного к аффинному пространству E , то биекция $x \rightarrow x + \vec{h}$ пространства E на себя называется переносом на вектор \vec{h} . Это, очевидно, аффинное преобразование, присоединенное линейное отображение которого является тождественным преобразованием. Обратно, каждое аффинное преобразование, имеющее в качестве присоединенного линейного тождественное преобразование, является переносом. В самом деле, $\overrightarrow{u(b)} - \overrightarrow{u(a)} = \overrightarrow{b} - \overrightarrow{a}$; следовательно, $\overrightarrow{u(b)} - \overrightarrow{b} = \overrightarrow{u(a)} - \overrightarrow{a}$, т. е. $\overrightarrow{u(x)} - \overrightarrow{x}$ не зависит от x . Обозначив через \vec{h} значение этой разности, получим $u(x) = x + \vec{h}$.

Аффинные нормированные пространства

Z Говорят, что аффинное пространство нормировано, если нормировано его присоединенное векторное пространство. Подчеркнем, что норма является функцией, определенной в присоединенном векторном пространстве, а не в самом аффинном пространстве. Можно говорить о норме вектора, а не о норме точки.

Аффинное нормированное пространство обладает метрикой, определенной естественной функцией расстояния: $d(x, y) = \overrightarrow{\|x - y\|}$. Эта функция расстояния инвариантна относительно переноса в том смысле, что $d(x + \vec{h}, y + \vec{h}) = d(x, y)$. С другой стороны, при гомотетии с центром в нуле и отношением λ (определенной отображением $x \rightarrow x' = 0 + \lambda(\overrightarrow{x} - \overrightarrow{0})$) простран-

ства E в себя) расстояние умножается на $|\lambda|$ (в том смысле, что $d(x', y') = |\lambda|d(x, y)$). Если E является аффинным нормированным пространством, то отображение $(x, y) \rightarrow y - x$ из $E \times E$ в \vec{E} и отображение $(x, \vec{h}) \rightarrow x + \vec{h}$ из $E \times \vec{E}$ в E непрерывны.

Если E и F — аффинные нормированные пространства и E конечномерно, то любое аффинное отображение E в F всегда непрерывно. Для доказательства достаточно повторить рассуждения, проведенные на стр. 111. Напротив, если E и F бесконечномерны, то это утверждение не всегда справедливо. Из изложенного на стр. 111 следует, что линейное отображение одного векторного нормированного пространства в другое может не быть непрерывным.

Теорема 1. 1°) *Аффинное нормированное пространство полно тогда и только тогда, когда полно его присоединенное векторное подпространство.*

2°) *Аффинное подпространство F аффинного пространства E замкнуто тогда и только тогда, когда его присоединенное векторное подпространство \vec{F} замкнуто в \vec{E} (в частности, конечномерное аффинное подпространство всегда замкнуто).*

3°) *Для того чтобы аффинное отображение и аффинного нормированного пространства E в аффинное нормированное пространство F было непрерывным, необходимо и достаточно, чтобы присоединенное линейное отображение было непрерывным¹⁾. В этом случае и равномерно непрерывно.*

Доказательство. 1°) Пусть a — точка, выбранная за начало в E . Отображение $\vec{X} \rightarrow a + \vec{X}$ является биекцией \vec{E} на E , сохраняющей расстояние, т. е. сохраняющей структуру метрического пространства. Но тогда E будет полным одновременно с \vec{E} .

2°) Пусть $a \in F$. Тогда $\vec{X} \rightarrow a + \vec{X}$ является гомеоморфизмом \vec{E} на E , а образом \vec{F} является F . Значит, F замкнуто в E тогда и только тогда, когда \vec{F} замкнуто в \vec{E} .

3°) Если u является аффинным непрерывным отображением E в F , то линейное присоединенное отображение \vec{u} опре-

¹⁾ Напомним, что линейное отображение непрерывно тогда и только тогда, когда оно непрерывно в нуле пространства (теорема 47 гл. II).

деляется формулой

$$\overrightarrow{u}(\vec{X}) = \overrightarrow{u(a + \vec{X})} - \overrightarrow{u(a)}, \quad (\text{III}, 1; 7)$$

где a — некоторая фиксированная точка. Это отображение, очевидно, непрерывно. Обратно, если u непрерывно, то из неравенства

$$\|\overrightarrow{u(x) - u(y)}\| = \|\overrightarrow{u(x - y)}\| \leq \|\overrightarrow{u}\| \|\vec{x} - \vec{y}\| \quad (\text{III}, 1; 7_2)$$

следует, что отображение u удовлетворяет условию Липшица и, следовательно, равномерно непрерывно.

Теорема 2. Пусть E — аффинное нормированное пространство, f — непостоянная скалярная аффинная функция над E . Тогда уравнение $f(x) = 0$ определяет некоторую аффинную гиперплоскость H .

Обратно, любая аффинная гиперплоскость определяется бесконечным множеством уравнений этого вида. Все соответствующие аффинные функции пропорциональны одна другой.

Каждая аффинная гиперплоскость H замкнута или плотна в E . Она замкнута тогда и только тогда, когда аффинные функции f , определяющие уравнение этой плоскости, непрерывны.

Доказательство. 1°) Пусть f — непостоянная скалярная аффинная функция, а f_0 — линейная форма на присоединенном пространстве \vec{E}^1 .

Пространство E содержит по крайней мере две точки a и b , такие, что $f(a) \neq f(b)$. Формула

$$f(a + t(\vec{b} - \vec{a})) = f(a) + tf_0(\vec{b} - \vec{a}), \quad f_0(\vec{b} - \vec{a}) \neq 0,$$

показывает, что можно всегда выбрать t так, чтобы f обращалась в нуль хотя бы в одной точке c из E . Выбирая c в качестве начала, мы установим биекцию $\vec{X} \rightarrow c + \vec{X}$ из \vec{E} на E , которая позволит нам проводить рассуждения над пространством \vec{E} вместо пространства E . Пусть \vec{H} — множество, определенное уравнением $f_0(\vec{X}) = 0$, или, иначе, множество $f_0^{-1}(\{0\})$, или, иначе, ядро линейной формы f_0 . Пусть, с другой стороны, H — множество из E , определяемое уравнением $f(x) = 0$.

¹⁾ См. примечание на стр. 178. Если мы будем писать \vec{f} , то можно подумать, что $\vec{f}(\vec{X})$ является вектором, в то время как это лишь скаляр.

Поскольку $f(c) = 0$, то $x \in H$ тогда и только тогда, когда $\overrightarrow{x-c} \in \vec{H}$; другими словами, мы имеем $H = c + \vec{H}$.

Если мы покажем, что \vec{H} является векторной гиперплоскостью, то тем самым докажем, что H является аффинной гиперплоскостью. Поскольку f не постоянна, то f_0 не равна тождественно нулю. Следовательно, существует такой элемент \vec{e} из \vec{E} , что $f_0(\vec{e}) \neq 0$, а, заменяя при необходимости вектор \vec{e} на кратный ему вектор, мы можем всегда считать, что $f_0(\vec{e}) = 1$. Любой элемент \vec{X} из \vec{E} тогда запишется единственным образом в виде

$$\vec{X} = \vec{Y} + \lambda \vec{e}, \quad \lambda \in K, \quad \vec{Y} \in \vec{H} \quad (\text{где } f_0(\vec{Y}) = 0). \quad (\text{III}, 1; 7_3)$$

Эта формула эквивалентна равенству

$$\lambda = f_0(\vec{X}), \quad \vec{Y} = \vec{X} - f_0(\vec{X}) \vec{e}.$$

Это равенство показывает, что векторное подпространство \vec{H} и векторное одномерное подпространство, порожденное вектором \vec{e} , дополнительны в \vec{E} , т. е. что \vec{H} является гиперплоскостью.

2°) Обратно, пусть H — гиперплоскость в E . Ее присоединенным векторным подпространством \vec{H} является по определению некоторая гиперплоскость в \vec{E} .

Пусть теперь \vec{e} — некоторый вектор, дополнительный к \vec{H} в \vec{E} ($\vec{e} \notin \vec{H}$). Всякий вектор \vec{X} из \vec{E} допускает тогда единственное разложение вида (III, 1; 7₃). Скаляр λ зависит от \vec{X} , и мы можем обозначить его через $\lambda = f_0(\vec{X})$. Функция $f_0: \vec{X} \rightarrow f_0(\vec{X})$ является линейной формой на \vec{E} . Она не является тождественно нулевой (например, $f_0(\vec{e}) = 1$), а множество ее нулей совпадает с \vec{H} .

Всякая линейная форма g_0 , такая, что уравнение $g_0(\vec{X}) = 0$ определяет гиперплоскость \vec{H} , пропорциональна f . В самом деле, если $k = g_0(\vec{e})$, то $g_0(\vec{X}) = kf_0(\vec{X})$ для $\vec{X} = \vec{e}$ (по определению) и для $\vec{X} \in \vec{H}$, поскольку в этом случае справа и слева стоят нули, а, следовательно, $g_0(\vec{X}) = kf_0(\vec{X})$ для произвольного $\vec{X} \in \vec{E}$. Если теперь взять произвольную точку c

из H , то множество H будет определено одним из уравнений вида $g(x) = 0$, где $\overrightarrow{g(x)} = g_0(x - c)$. Мы видим, что все соответствующие аффинные функции g пропорциональны одна другой.

3°) Если f непрерывна, то множество H , прообраз замкнутого множества $\{0\}$ поля скаляров относительно отображения f , замкнуто в E . Пусть теперь функция f , а вместе с ней и функция f_0 разрывны. Согласно теореме 47 гл. II, для каждого целого n можно найти такой вектор \vec{a}_n из \vec{E} , что $|f_0(\vec{a}_n)| \geqslant n \|\vec{a}_n\|$. Умножая \vec{a}_n на соответствующий скаляр, можно добиться того, чтобы $\|\vec{a}_n\| \leqslant 1/n$ и одновременно $f_0(\vec{a}_n) = 1$. Пусть x — любая точка из E . Рассмотрим последовательность точек $x_n = x - f(x)\vec{a}_n$. Так как $f(x_n) = f(x) - f(x)f_0(\vec{a}_n) = 0$, то точки x_n принадлежат H . Поскольку $f(x)\vec{a}_n$ сходится к 0, последовательность x_n сходится к x , когда n стремится к бесконечности. Мы видим, что каждая точка x из E является точкой сгущения H , т. е. что H плотно в E , чем и заканчивается доказательство теоремы.

Эта теорема выявляет в некотором смысле удивительный факт, непривычный для тех, кто занимается изучением конечномерных пространств: может случиться, что некоторая гиперплоскость плотна в рассматриваемом пространстве. В дальнейшем нам представится случай встретиться с плотными векторными подпространствами векторных нормированных пространств (знаменитая теорема Вейерштрасса говорит, что подпространство полиномов плотно в пространстве $C[0, 1]$ непрерывных комплексных функций на $[0, 1]$).

Выпуклые множества в аффинных пространствах

Пусть a и b — две точки аффинного пространства E над полем \mathbb{K} вещественных или комплексных чисел. Множество точек, которые можно записать в виде $a + tb - a$, где t вещественное число, $0 \leqslant t \leqslant 1$, называется *отрезком с концами a и b* . Этот отрезок обозначают через $[a, b]$ и называют также замкнутым отрезком. Через $[a, b]$ (соответственно через $]a, b]$ и через $]a, b[$) обозначают тот же самый отрезок без точки b (соответственно без точки a и без обеих точек a и b). Здесь применяются выражения «полуоткрытый отрезок» или «открытый отрезок», хотя они не являются открытыми множествами, если аффинное пространство нормировано.

Часть аффинного пространства E называется *выпуклой*, если вместе с любыми двумя своими точками она целиком содержит весь отрезок с концами в этих точках.

В аффинном пространстве без труда определяется понятие *барицентра*¹⁾. Часть A аффинного пространства E выпукла тогда и только тогда, когда вместе с конечным числом точек она содержит их барицентр при любой системе масс ≥ 0 . Это легко проверить, заметив, что барицентр нескольких точек может быть построен с помощью последовательного построения барицентров двух точек и что барицентры системы двух точек с произвольными массами ≥ 0 находятся на соединяющем их отрезке.

Пустое множество, множество, сводящееся к одной точке, само аффинное пространство E и, более общо, любое аффинное многообразие из E являются выпуклыми. Если E нормировано, то всякий открытый или замкнутый шар является выпуклым. Любое пересечение конечного или бесконечного числа выпуклых частей является выпуклой частью. Можно доказать, что внутренняя часть и замыкание выпуклой части аффинного нормированного пространства выпуклы.

Евклидовы векторные и евклидовы аффинные пространства

Пусть \vec{E} — векторное пространство над полем вещественных чисел \mathbb{R} . *Евклидовым скалярным произведением* в \vec{E} называется *билинейная форма* $(\vec{X}, \vec{Y}) \rightarrow B(\vec{X}, \vec{Y}) \in \mathbb{R}$, которая яв-

¹⁾ Пусть A_1, A_2, \dots, A_n — точки аффинного пространства E , снабженные такими коэффициентами $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$, что $\sum_{i=1}^n \alpha_i \neq 0$. Тогда точка $G =$

$= O' + \left(\sum_{j=1}^n \alpha_j \right)^{-1} \sum_{i=1}^n \alpha_i \vec{O'A}_i$ пространства E , где O' — некоторая точка E , называется *барицентром* системы точек A_i . Если числа $\alpha_i \geq 0$, то их называют *массами* точек A_i , а барицентр G — *центром тяжести* системы точек A_i .

Точка G не зависит от выбора точки O' из E . Необходимым и достаточным условием того, что G является барицентром системы точек A_i , снабженных коэффициентами $\left(\alpha_i, \sum_{i=1}^n \alpha_i \neq 0 \right)$, является выполнение равенства

$\sum_{i=1}^n \alpha_i \vec{GA}_i = 0$. Барицентр системы точек не меняется, если часть этих точек заменить их барицентром (если он существует), снабженным коэффициентом, равным сумме коэффициентов соответствующих точек. — Прим. ред.

ляется симметричной, т. е. $B(\vec{X}, \vec{Y}) = B(\vec{Y}, \vec{X})$, и положительно определенной, т. е.

$$B(\vec{X}, \vec{X}) > 0 \quad \text{для } \vec{X} \neq 0. \quad (\text{III, 1; 8})$$

Два вектора \vec{X} и \vec{Y} из \vec{E} называются ортогональными относительно скалярного произведения B , если $B(\vec{X}, \vec{Y}) = 0$.

Векторное пространство, снабженное евклидовым скалярным произведением, называется *евклидовым векторным пространством*. Аффинное пространство, присоединенное векторное пространство которого евклидово, называется *евклидовым аффинным пространством*. То, что в элементарной геометрии называют трехмерным пространством, является просто евклидовым аффинным пространством трех измерений¹⁾.

Теорема 2. Пусть \vec{E} — векторное пространство над \mathbb{R} , B — некоторая билинейная симметричная форма на $\vec{E} \times \vec{E}$, такая, что $B(\vec{X}, \vec{X}) \geqslant 0$ для всех \vec{X} из \vec{E} . Тогда имеет место неравенство Коши — Шварца

$$|B(\vec{X}, \vec{Y})| \leqslant \sqrt{B(\vec{X}, \vec{X})} \cdot \sqrt{B(\vec{Y}, \vec{Y})} \quad (\text{III, 1; 9})$$

и неравенство Минковского

$$\sqrt{B(\vec{X} + \vec{Y}, \vec{X} + \vec{Y})} \leqslant \sqrt{B(\vec{X}, \vec{X})} + \sqrt{B(\vec{Y}, \vec{Y})}; \quad (\text{III, 1; 10})$$

кроме того, если B — положительно определенная форма, т. е. $B(\vec{X}, \vec{X}) > 0$ для $\vec{X} \neq \vec{0}$, то в обеих формулах стоит всегда знак строгого неравенства $<$, кроме тех случаев, когда в формуле (III, 1; 9) векторы \vec{X} и \vec{Y} пропорциональны и когда в формуле (III, 1; 10) эти векторы пропорциональны с положительным коэффициентом пропорциональности²⁾.

Доказательство. Для любых векторов \vec{X}, \vec{Y} и числа λ имеем

$$B(\vec{X} + \lambda\vec{Y}, \vec{X} + \lambda\vec{Y}) \geqslant 0, \quad (\text{III, 1; 11})$$

¹⁾ Это не совсем точно: в элементарной геометрии единица длины не обязательно фиксирована. Пространство становится аффинным евклидовым лишь после выбора в нем единицы длины.

²⁾ Если \vec{X} или \vec{Y} являются нулевыми векторами, то можно считать, что они пропорциональны с коэффициентами пропорциональности $\geqslant 0$.

или

$$B(\vec{X}, \vec{X}) + 2\lambda B(\vec{X}, \vec{Y}) + \lambda^2 B(\vec{Y}, \vec{Y}) \geqslant 0. \quad (\text{III}, 1; 12)$$

Вещественный трехчлен неотрицателен лишь в том случае, когда его дискриминант $\leqslant 0$, откуда и получается неравенство (III, 1; 9). Пусть теперь билинейная форма B положительно определена. Если $\vec{Y} \neq \vec{0}$, то $B(\vec{Y}, \vec{Y}) > 0$ и трехчлен не вырожден. Если \vec{X} не пропорционален \vec{Y} , то вектор $\vec{X} + \lambda \vec{Y}$ отличен от $\vec{0}$ для любого вещественного λ . Следовательно, левая часть соотношения (III, 1; 11) всегда > 0 . Но тогда трехчлен (III, 1; 12) больше нуля при любом λ , т. е. не имеет вещественных корней. Это означает, что его дискриминант < 0 и соотношение (III, 1; 9) имеет место со знаком $<$. Таким образом, в (III, 1; 9) всегда имеет место знак $<$, кроме случая, когда $\vec{Y} = \vec{0}$ или $\vec{Y} \neq \vec{0}$, но $\vec{X} = \lambda_0 \vec{Y}$, т. е. если \vec{X} и \vec{Y} пропорциональны. Что же касается неравенства Минковского, то оно эквивалентно неравенству

$$B(\vec{X} + \vec{Y}, \vec{X} + \vec{Y}) \leqslant B(\vec{X}, \vec{X}) + B(\vec{Y}, \vec{Y}) + 2\sqrt{B(\vec{X}, \vec{X}) \cdot B(\vec{Y}, \vec{Y})}, \quad (\text{III}, 1; 13)$$

или

$$\begin{aligned} B(\vec{X}, \vec{X}) + 2B(\vec{X}, \vec{Y}) + B(\vec{Y}, \vec{Y}) &\leqslant \\ &\leqslant B(\vec{X}, \vec{X}) + B(\vec{Y}, \vec{Y}) + 2\sqrt{B(\vec{X}, \vec{X}) \cdot B(\vec{Y}, \vec{Y})}. \end{aligned} \quad (\text{III}, 1; 14)$$

Последнее же неравенство следует из (III, 1; 9). Если B положительно определена, то равенство справедливо только в том случае, когда $B(\vec{X}, \vec{Y}) \geqslant 0$ и когда (III, 1; 9) имеет место со знаком $=$, т. е. если \vec{X} и \vec{Y} пропорциональны с неотрицательным коэффициентом пропорциональности.

Отсюда вытекает, что если евклидово пространство E евклидово, то функция $\vec{X} \rightarrow \sqrt{B(\vec{X}, \vec{X})}$ является нормой в E . Именно она служит для определения расстояния в элементарной геометрии. Евклидово пространство, векторное или аффинное, является нормированным. Скалярное произведение в евклидовом векторном пространстве принято обозначать через $(\vec{X} | \vec{Y})$, а через $\sqrt{(\vec{X} | \vec{X})}$ — норму $\|\vec{X}\|$.

Эрмитовы векторные и эрмитовы аффинные пространства¹⁾

Пусть \vec{E} и \vec{F} — векторные пространства над полем комплексных чисел \mathbb{C} . Отображение u из \vec{E} в \vec{F} называется *полулинейным*, если оно удовлетворяет следующим условиям:

$$\begin{aligned} u(\vec{X} + \vec{Y}) &= u(\vec{X}) + u(\vec{Y}) \quad \text{для } \vec{X} \in \vec{E}, \vec{Y} \in \vec{E}, \\ u(\lambda \vec{X}) &= \bar{\lambda} u(\vec{X}) \quad \text{для } \vec{X} \in \vec{E}, \lambda \in \mathbb{C}. \end{aligned} \quad (\text{III, 1; 15})$$

Если $\vec{F} = \mathbb{C}$, то u является *полулинейной формой*.

Полуторалинейной формой²⁾ на $\vec{E} \times \vec{E}$ называется функция $(\vec{X}, \vec{Y}) \rightarrow B(\vec{X}, \vec{Y}) \in \mathbb{C}$, линейная по \vec{X} при фиксированном \vec{Y} и полулинейная по \vec{Y} при фиксированном \vec{X} . Другими словами,

$$\begin{aligned} B(\vec{X}_1 + \vec{X}_2, \vec{Y}) &= B(\vec{X}_1, \vec{Y}) + B(\vec{X}_2, \vec{Y}), \\ B(\vec{X}, \vec{Y}_1 + \vec{Y}_2) &= B(\vec{X}, \vec{Y}_1) + B(\vec{X}, \vec{Y}_2), \\ B(\lambda \vec{X}, \vec{Y}) &= \bar{\lambda} B(\vec{X}, \vec{Y}), \lambda \in \mathbb{C}, \\ B(\vec{X}, \mu \vec{Y}) &= \bar{\mu} B(\vec{X}, \vec{Y}), \mu \in \mathbb{C}. \end{aligned} \quad (\text{III, 1; 16})$$

Эрмитовым скалярным произведением в векторном пространстве \vec{E} над полем комплексных чисел называется *полуторалинейная форма* $(\vec{X}, \vec{Y}) \rightarrow B(\vec{X}, \vec{Y})$, которая является эрмитовой, т. е. такой, что $B(\vec{Y}, \vec{X}) = B(\vec{X}, \vec{Y})$, и положительно определенной, т. е. удовлетворяющей неравенству (III, 1; 8). Векторное пространство над полем \mathbb{C} , снабженное эрмитовым скалярным произведением, называется *эрмитовым векторным пространством*. Аффинное пространство над полем \mathbb{C} , присоединенное векторное пространство которого эрмитово, называется *эрмитовым аффинным пространством*.

Теорема 2₃. *Заключения теоремы 2₂ сохраняются, если в ее условии заменить \mathbb{R} на \mathbb{C} и «билинейная симметрична» на «полуторалинейная эрмитова».*

¹⁾ В русской литературе вместо слова «эрмитово» принято использовать термин «унитарное». — Прим. ред.

²⁾ Полтора = один с половиной; полуторалинейная = линейная + полулинейная.

Доказательство. В самом деле, вместо (III, 1; 12) на этот раз для любого комплексного λ получаем:

$$B(\vec{X}, \vec{X}) + \lambda B(\vec{Y}, \vec{X}) + \bar{\lambda} B(\vec{X}, \vec{Y}) + \lambda \bar{\lambda} B(\vec{Y}, \vec{Y}) \geq 0. \quad (\text{III, 1; 17})$$

Пусть $B(\vec{X}, \vec{Y}) = \rho e^{i\theta}$, где $\rho = |B(\vec{X}, \vec{Y})|$. Положим $\lambda = te^{i\theta}$, где t — вещественное число (не обязательно ≥ 0). Из (III, 1; 17) следует, что для любого вещественного t имеет место неравенство

$$B(\vec{X}, \vec{X}) + 2|B(\vec{X}, \vec{Y})|t + B(\vec{Y}, \vec{Y})t^2 \geq 0, \quad (\text{III, 1; 18})$$

из которого снова следует (III, 1; 9).

Неравенство (III, 1; 10) эквивалентно неравенству

$$B(\vec{X}, \vec{Y}) + B(\vec{Y}, \vec{X}) \leq 2\sqrt{B(\vec{X}, \vec{X})B(\vec{Y}, \vec{Y})}, \quad (\text{III, 1; 18}_2)$$

или

$$\operatorname{Re}(B(\vec{X}, \vec{Y}))^1 \leq \sqrt{B(\vec{X}, \vec{X})B(\vec{Y}, \vec{Y})}, \quad (\text{III, 1; 18}_3)$$

что вытекает из (III, 1; 9).

Если B положительно определена, $\vec{Y} \neq \vec{0}$ и \vec{X} не пропорционален \vec{Y} , то (III, 1; 18) будет >0 для любого вещественного t ; следовательно, неравенство (III, 1; 9) будет справедливым со знаком $<$. Равенство возможно лишь в том случае, когда \vec{X} и \vec{Y} пропорциональны. Равенство в соотношении (III, 1; 10) может иметь место только в том случае, если, кроме того, $\operatorname{Re} B(\vec{X}, \vec{Y}) = |B(\vec{X}, \vec{Y})|$, т. е. если $B(\vec{X}, \vec{Y})$ вещественно и ≥ 0 , а, значит, если коэффициент пропорциональности ≥ 0 .

Таким образом, эрмитово векторное или аффинное пространство нормировано.

Приведем теперь результаты, пригодные как для евклидовых пространств над полем \mathbb{R} , так и для эрмитовых пространств над полем \mathbb{C} . Мы их будем излагать для эрмитовых пространств. Конечно, в евклидовом случае скаляры будут считаться вещественными, полулинейность будет означать линейность, полуторалинейность — билинейность, а $\bar{\lambda}$ будет заменяться на λ .

¹⁾ Re — вещественная часть.

Изоморфизм (или полуизоморфизм) конечномерного евклидова (или эрмитова) пространства и его сопряженного пространства

Через $\overset{\leftarrow}{E'}$ будем обозначать пространство, сопряженное к некоторому векторному пространству \vec{E} над полем K^1 . Если $\overset{\leftarrow}{\alpha}$ является элементом из $\overset{\leftarrow}{E'}$, то это линейная форма на \vec{E} , т. е. отображение $\vec{X} \rightarrow \overset{\leftarrow}{\alpha}(\vec{X})$. Вместо $\overset{\leftarrow}{\alpha}(\vec{X})$ удобнее писать $\overset{\leftarrow}{\alpha} \cdot \vec{X}$ или $\langle \overset{\leftarrow}{\alpha}, \vec{X} \rangle$. Известно, что $\overset{\leftarrow}{E'}$ также является векторным пространством, а функция $(\overset{\leftarrow}{\alpha}, \vec{X}) \rightarrow \langle \overset{\leftarrow}{\alpha}, \vec{X} \rangle \in K$ — некоторой билинейной формой на $\overset{\leftarrow}{E'} \times \vec{E}$, которую называют фундаментальной билинейной формой. Ее называют также скалярным произведением $\overset{\leftarrow}{\alpha} \in \overset{\leftarrow}{E'}$ и $\vec{X} \in \vec{E}$, но это скалярное произведение не имеет никакого отношения к скалярному произведению в евклидовом или эрмитовом пространствах, ибо $\overset{\leftarrow}{\alpha}$ и \vec{X} не принадлежат одному и тому же векторному пространству²⁾. При этом следует строго различать обозначения $\langle \overset{\leftarrow}{\alpha}, \vec{X} \rangle$ и $(\vec{X} | \vec{Y})$. Однако если \vec{E} является евклидовым или эрмитовым пространством, то оба скалярных произведения существуют одновременно.

Для фиксированного \vec{Y} функция $\vec{X} \rightarrow (\vec{X} | \vec{Y})$ является линейной формой на \vec{E} . Следовательно, ей соответствует некоторый элемент $\overset{\leftarrow}{\gamma}$ сопряженного пространства, такой, что

$$\langle \overset{\leftarrow}{\gamma}, \vec{X} \rangle = (\vec{X} | \vec{Y}) \quad \text{для любого } \vec{X} \text{ из } \vec{E}. \quad (\text{III}, 1; 19)$$

Этот элемент $\overset{\leftarrow}{\gamma}$ зависит от \vec{Y} , и поэтому мы будем обозначать его через $\overset{\leftarrow}{\gamma}_{\vec{Y}}$ или $\overset{\leftarrow}{\gamma}(\vec{Y})$. Тогда полулинейность правой части равенства (III, 1; 19) по отношению к \vec{Y} для каждого

1) Сопряженное пространство и его элементы удобно обозначать символами, снабженными сверху противоположно направленной стрелкой (тем более, что векторы из $\overset{\leftarrow}{E}$ называются ковекторами).

2) Кроме того, скалярное произведение \langle , \rangle всегда билинейно, а не полуторалинейно.

фиксированного \vec{X} означает, что $\overset{\leftarrow}{\gamma}_{\vec{Y}}$ полулинейно зависит от \vec{Y} :

$$\begin{aligned}\overset{\leftarrow}{\gamma}_{\vec{Y}_1 + \vec{Y}_2} &= \overset{\leftarrow}{\gamma}_{\vec{Y}_1} + \overset{\leftarrow}{\gamma}_{\vec{Y}_2}, \\ \overset{\leftarrow}{\gamma}_{\lambda \vec{Y}} &= \bar{\lambda} \overset{\leftarrow}{\gamma}_{\vec{Y}}.\end{aligned}\tag{III, 1; 20}$$

Таким образом, $\vec{Y} \rightarrow \overset{\leftarrow}{\gamma}_{\vec{Y}}$ является полулинейным отображением γ пространства \vec{E} в его сопряженное пространство \vec{E}' . Это отображение инъективно, ибо равенство $\overset{\leftarrow}{\gamma}_{\vec{Y}} = \vec{0}$ означает, что $\langle \overset{\leftarrow}{\gamma}_{\vec{Y}}, \vec{X} \rangle = (\vec{X} | \vec{Y}) = 0$, каким бы ни был вектор \vec{X} из \vec{E} . Полагая $\vec{X} = \vec{Y}$, получаем $(\vec{Y} | \vec{Y}) = 0$, откуда $\vec{Y} = \vec{0}$, что доказывает инъективность отображения γ . Если E конечномерно, то \vec{E} и \vec{E}' имеют одну и ту же размерность, а это означает, что γ является биекцией \vec{E} на \vec{E}' . Отсюда следует, что задание евклидовой структуры определяет изоморфизм между пространством и его сопряженным; задание эрмитовой структуры определяет полуизоморфизм между ними.

Воспользуемся только тем фактом, что γ является сюръекцией. Если α является элементом сопряженного пространства, т. е. некоторой линейной формой на \vec{E} , то существует, и притом единственный, элемент \vec{Y} из \vec{E} , такой, что

$$(\vec{X} | \vec{Y}) = \langle \overset{\leftarrow}{\alpha}, \vec{X} \rangle \text{ для всех } \vec{X} \text{ из } \vec{E}.$$

Элемент \vec{Y} является не чем иным, как $\overset{\longrightarrow}{\gamma^{-1}}(\overset{\leftarrow}{\alpha})$. Таким образом, получаем следующий результат:

Теорема 24. Если α является линейной формой над евклидовым или эрмитовым конечномерным пространством \vec{E} , то существует, и притом единственный, вектор \vec{Y} , такой, что форма α является скалярным произведением $\vec{X} \rightarrow (\vec{X} | \vec{Y})$.

Ортонормированные базисы

Ортонормированным называется базис $(\vec{e}_i)_{i \in I}$ евклидова или эрмитова конечномерного пространства \vec{E} , элементы которого попарно ортогональны и имеют длину 1.

Теорема 25. Любое евклидово или эрмитово конечномерное векторное пространство имеет ортонормированные базисы.

Доказательство. Это очевидно, если размерность n равна 1¹⁾, ибо если $\{\vec{f}_1\}$ является базисом, то $\{\vec{e}_1 = \vec{f}_1 / \| \vec{f}_1 \| \}$ будет ортонормированным базисом. Предположим теперь, что существование ортонормированного базиса доказано для любого евклидова или эрмитова пространства размерности $\leq n - 1$, и докажем его существование в пространстве \vec{E} размерности n . Пусть \vec{f}_1 — некоторый вектор $\neq \vec{0}$ и $\vec{e}_1 = \vec{f}_1 / \| \vec{f}_1 \|$. Множество векторов, ортогональных \vec{e}_1 , является гиперплоскостью \vec{H}_1 в \vec{E} . В самом деле, это множество векторов \vec{X} , удовлетворяющих линейному уравнению $\gamma_{\vec{e}_1}(\vec{X}) = (\vec{X} | \vec{e}_1) = 0$, где линейная форма $\gamma_{\vec{e}_1}$ не равна тождественно нулю, поскольку $\vec{e}_1 \neq \vec{0}$, и, кроме того, γ инъективна. Так как $(\vec{e}_1 | \vec{e}_1) > 0$, то эта гиперплоскость не содержит \vec{e}_1 , а, следовательно, \vec{H}_1 и прямая, порожденная вектором \vec{e}_1 , являются дополнительными; \vec{H}_1 называется гиперплоскостью, ортогональной \vec{e}_1 , и является евклидовым или эрмитовым пространством размерности $n - 1$. По предположению индукции она должна иметь хотя бы один ортонормированный базис $\vec{e}_2, \dots, \vec{e}_n$. При этом система векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ образует ортонормированный базис пространства \vec{E} , что и требовалось доказать.

Если $(\vec{e}_i)_{i \in I}$ является некоторым базисом в \vec{E} , то скалярное произведение двух векторов $\vec{X} = \sum_{i \in I} X_i \vec{e}_i$ и $\vec{Y} = \sum_{i \in I} Y_i \vec{e}_i$ записывается в виде

$$(\vec{X} | \vec{Y}) = \sum_{i, j \in I} (\vec{e}_i | \vec{e}_j) X_i \bar{Y}_j = \sum_{i, j \in I} g_{i, j} X_i \bar{Y}_j, \quad (\text{III}, 1; 21)$$

где $g_{i, j} = (\vec{e}_i | \vec{e}_j)$. Здесь $g_{i, i} = \bar{g}_{i, i}$, а неравенство $(\vec{X} | \vec{X}) > 0$ для $\vec{X} \neq \vec{0}$ записывается в виде

$$\sum_{i, j \in I} g_{i, j} X_i \bar{X}_j > 0, \quad (\text{III}, 1; 22)$$

1) Индукцию можно даже начинать с $n = 0$. В векторном пространстве размерности 0 базис пуст и его можно считать ортонормированным!

кроме случая, когда $X_i = 0$ для всех $i \in I$. Матрицу, составленную из элементов $g_{i,j}$, называют *эрмитовой положительно определенной матрицей*.

Рассматриваемый базис ортонормирован тогда и только тогда, когда матрица из $g_{i,j}$ является единичной:

$$\langle \vec{X} | \vec{Y} \rangle = \sum_{i \in I} X_i \bar{Y}_i, \quad \| \vec{X} \|^2 = \sum_{i \in I} |X_i|^2. \quad (\text{III}, 1; 23)$$

Обобщенные евклидовы или эрмитовы пространства ¹⁾

В специальной теории относительности возникает необходимость в обобщении евклидова или эрмитова пространства, пространстве Лоренца или пространстве Минковского. Это — *конечномерное* векторное пространство \vec{E} (в физике это четырехмерное пространство над полем \mathbb{R}), снабженное обобщенным скалярным произведением, т. е. полуторалинейной *не положительно определенной* эрмитовой формой. Будем ее, как обычно, обозначать через $\langle \vec{X} | \vec{Y} \rangle$. Для $\vec{X} \neq \vec{0}$ не обязательно должно выполняться неравенство $\langle \vec{X} | \vec{X} \rangle > 0$. Может случиться, что для некоторых векторов $\vec{X} \neq \vec{0}$ произведение $\langle \vec{X}, \vec{X} \rangle = 0$. Эти векторы называются *изотропными*, а их объединение — *изотропным конусом*²⁾. При этом вводится существенное условие: *полуторалинейная форма не вырождена*; иначе говоря, *не существует вполне изотропного вектора* $\vec{X} \neq \vec{0}$, т. е. ортогонального ко всем векторам пространства. Разумеется, здесь не выполняются неравенства Коши — Шварца и Минковского и такая структура *не определяет норму*. Однако отображение γ пространства \vec{E} в \vec{E}' всегда полулинейно и инъективно, ибо равенство $\gamma_{\vec{Y}} = \vec{0}$ означает, что вектор \vec{Y} ортогонален всем векторам \vec{X} из \vec{E} , т. е. вполне изотропен, а, значит, равен нулю. Отображение γ биективно. Линейной форме α над \vec{E} соответствует, как и ранее, некоторый вектор \vec{Y} из \vec{E} , и притом единственный, такой, что $\langle \alpha, \vec{X} \rangle = \langle \vec{X} | \vec{Y} \rangle$ для любого $\vec{X} \in \vec{E}$ — это вектор $\gamma^{-1}(\alpha)$.

¹⁾ Эти пространства называют также псевдоевклидовыми и псевдоэрмитовыми. То, что мы здесь называем ортонормированным базисом, носит также название псевдоортонормированного базиса.

²⁾ Это понятие отличается от принятого в физике: там рассматривается векторное пространство над \mathbb{R} и вещественное скалярное произведение.

Ортонормированным базисом обобщенного евклидова или эрмитова пространства называется базис $(\vec{e}_i)_{i \in I}$, элементы которого попарно ортогональны и скалярные квадраты которых равны ± 1 . Такие базисы всегда существуют. Для доказательства можно применить, только с некоторой осторожностью, построения по индукции, проведенные при доказательстве теоремы 2₅.

При $n = 1$ исходят из произвольного вектора $\vec{f}_1 \neq \vec{0}$. При $n = 1$ все векторы пространства пропорциональны вектору \vec{f}_1 , поэтому утверждение « \vec{f}_1 не является вполне изотропным» равносильно тому, что « \vec{f}_1 не является просто изотропным». Если взять теперь $\vec{e}_1 = \vec{f}_1 / \sqrt{|\langle \vec{f}_1, \vec{f}_1 \rangle|}$, то получим $\langle \vec{e}_1 | \vec{e}_1 \rangle = \pm 1$. Однако при переходе от $n - 1$ к n требуются более тонкие рассуждения. Покажем прежде всего, что существует по крайней мере один вектор \vec{f}_1 , не являющийся изотропным. Если бы каждый вектор был изотропным, т. е. ортогональным самому себе, то из равенства

$$(\vec{X} + \vec{Y} | \vec{X} + \vec{Y}) = (\vec{X} | \vec{Y}) + (\vec{Y} | \vec{Y}) + 2 \operatorname{Re}(\vec{X} | \vec{Y}) \quad (\text{III}, 1; 24)$$

следовало бы, что $\operatorname{Re}(\vec{X} | \vec{Y})$ является нулем для любых \vec{X} и \vec{Y} . Так как в случае $K = \mathbb{C}$ имеет место равенство $\operatorname{Re}(\vec{X} | i\vec{Y}) = -\operatorname{Im}(\vec{X} | \vec{Y})$, то мы получили бы $(\vec{X} | \vec{Y}) = 0$ для любых \vec{X} и \vec{Y} . Каждый вектор оказался вполне изотропным, что противоречит нашему предположению. Следовательно, всегда можно найти не изотропный вектор \vec{f}_1 и положить $\vec{e}_1 = \vec{f}_1 / \sqrt{|\langle \vec{f}_1, \vec{f}_1 \rangle|}$. Гиперплоскость \vec{H}_1 , ортогональная \vec{e}_1 , не содержит \vec{e}_1 , поскольку \vec{e}_1 не изотропен. Следовательно, \vec{H}_1 является дополнением к прямой, порождаемой вектором \vec{e}_1 . С другой стороны, скалярное произведение на \vec{H}_1 не вырожденно. В самом деле, если бы в \vec{H}_1 существовал вектор $\neq \vec{0}$, ортогональный всем векторам из \vec{H}_1 , то поскольку он ортогонален также и \vec{e}_1 , он был бы ортогонален всем векторам из \vec{E} , т. е. был бы вполне изотропным, что невозможно. Следовательно, \vec{H}_1 является обобщенным евклидовым или эрмитовым пространством размерности $n - 1$ и, значит, имеет ортонормальный базис $\vec{e}_2, \dots, \vec{e}_n$.

Так как $\vec{e}_1 \notin \vec{H}_1$, то $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ является ортонормированным базисом в \vec{E} .

Пусть теперь $(\vec{e}_i)_{i \in I}$ является ортонормированным базисом в \vec{E} . Если положить $(\vec{e}_i | \vec{e}_i) = \varepsilon_i = \pm 1$, то скалярное произведение векторов $\vec{X} = \sum_{i \in I} X_i \vec{e}_i$ и $\vec{Y} = \sum_{i \in I} Y_i \vec{e}_i$ запишется в виде

$$(\vec{X} | \vec{Y}) = \sum_{i \in I} \varepsilon_i X_i \bar{Y}_i \quad \text{и} \quad (\vec{X} | \vec{X}) = \sum_{i \in I} \varepsilon_i |X_i|^2. \quad (\text{III}, 1; 25)$$

Можно доказать, кроме того, важную теорему, называемую *законом инерции*: *число p чисел $\varepsilon_i > 0$ и число q чисел $\varepsilon_i < 0$ не зависят от выбора ортонормированного базиса*. В самом деле, рассмотрим фиксированный базис $(\vec{e}_i)_{i \in I}$, и пусть J (соответственно K) — подмножество в I , образованное теми i , для которых $\varepsilon_i = +1$ (соответственно -1). Пусть p (соответственно q) — число элементов в J (соответственно в K). Существует, по крайней мере одно, векторное подпространство \vec{F} из \vec{E} размерности p , на котором скалярное произведение является положительно определенным, а именно то подпространство, которое порождено векторами \vec{e}_i , $i \in J$. Векторного подпространства \vec{G} размерности $>p$, обладающего тем же свойством, не существует. В самом деле, если \vec{G} — любое векторное подпространство размерности $>p$, то оно пересекается с подпространством размерности $n-p$, порожденным векторами \vec{e}_i , $i \in K$, по некоторому не сводящемуся к $\vec{0}$ подпространству, поскольку сумма размерностей этих подпространств $>n$. Если $\vec{X} \neq \vec{0}$ — вектор этого пересечения, то необходимо имеем $(\vec{X} | \vec{X}) < 0$, а, значит, скалярное произведение на G не может быть положительно определенным. Итак, числа p и $q = n - p$ определяются независимо от первоначально выбранного базиса. Число p (соответственно q) является максимальной размерностью векторных подпространств из \vec{E} , на которых скалярное произведение положительно определено (соответственно отрицательно определено).

В итоге мы получаем такой результат:

Теорема 26. *Каждое обобщенное евклидово или эрмитово пространство имеет ортонормированные базисы. Число векторов такого базиса, скалярный квадрат которых равен $+1$ (соответственно -1), не зависит от выбора базиса. Это — максималь-*

ная размерность векторных подпространств, на которых скалярное произведение является положительно определенным (соответственно отрицательно определенным).

В физике, в специальной теории относительности, физическая вселенная пространства-времени является аффинным четырехмерным пространством E_4 над полем вещественных чисел. Его присоединенное векторное пространство \vec{E}_4 снабжено скалярным произведением, для которого $p = 3$ и $q = 1$. Галилеева система координат в E_4 является системой координат, образованной началом пространства E_4 и четырьмя векторами $\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_0$, образующими ортонормированный базис \vec{E}_4 и такими, что $e_1 = e_2 = e_3 = 1$ и $e_0 = -1$. Вектор $\vec{X} \in \vec{E}_4$ пространства-времени имеет 4 координаты $X_1, X_2, X_3, X_0 = cT$, где T — координата времени в рассматриваемой галилеевой системе координат и c — скорость света. Скалярный квадрат этого вектора равен $X_1^2 + X_2^2 + X_3^2 - X_0^2 = L^2 - c^2 T^2$, где L — пространственная длина вектора в рассматриваемой галилеевой системе координат.

§ 2. ВЕЩЕСТВЕННЫЕ ФУНКЦИИ ВЕЩЕСТВЕННОЙ ПЕРЕМЕННОЙ. НЕПРЕРЫВНОСТЬ СПРАВА И СЛЕВА

Пусть Ω — некоторая часть вещественной прямой \mathbb{R} и F — произвольное топологическое пространство. Говорят, что отображение f части Ω в F непрерывно справа в точке $a \in \Omega$, если $f(x)$ стремится к $f(a)$ при x , стремящемся к a в Ω по значениям $\geqslant a$. Если точка a изолирована справа в Ω , т. е. если существует такое число $\eta > 0$, что интервал $[a, a + \eta[$ не содержит ни одной точки Ω , то любое отображение Ω в F непрерывно справа в a . Отображение f непрерывно справа в точке a тогда и только тогда, когда сужение f на часть $x \geqslant a$ множество Ω непрерывно в точке a .

Такое же определение дается для непрерывности слева. Отображение f непрерывно в точке a тогда и только тогда, когда оно одновременно непрерывно справа и слева. Если точка a изолирована слева, то ее непрерывность в точке a эквивалентна непрерывности справа в этой точке. В дальнейшем все утверждения сделаны в предположении, что Ω является открытым множеством в \mathbb{R} . Такое предположение оставляет в стороне важный для практики случай, когда Ω является полуоткрытым или замкнутым интервалом. Большинство теорем будут справедливыми, вообще говоря, и в этом случае. Небольшие необходимые изменения читатель сделает сам.

Разрывы первого рода. Правильные функции

Говорят, что точка a является точкой разрыва первого рода функции $f(x)$, если при x , стремящемся к a по значениям, строго большим a , существует предел $f(x)$, обозначаемый через $f(a+0)$, при x , стремящемся к a по значениям, строго меньшим a , существует предел $f(x)$, обозначаемый через $f(a-0)$, и эти пределы одновременно не равны значению функции f в точке a . Если f непрерывна или имеет разрыв первого рода в точке a , а пространство F метрическое, то величина

$$\max [d(f(a), f(a-0)), d(f(a), f(a+0)), d(f(a-0), f(a+0))]$$

называется колебанием функции f в точке a . Оно равно нулю тогда и только тогда, когда f непрерывна в точке a . Если F является аффинным пространством, то можно вычислить разность $\overrightarrow{f(a+0) - f(a-0)}$, которая является элементом присоединенного векторного пространства \vec{F} . Эта разность называется скачком функции f в точке a . В определение скачка не входит значение $f(a)$ функции f в самой точке a . Поэтому он может быть нулем даже в том случае, когда f разрывна: значения $f(a+0)$ и $f(a-0)$ могут быть равны между собой, но не равны $f(a)$.

Функция f , разрывная в точке a , естественно, не обязана иметь в этой точке разрыв первого рода в том смысле, что пределы $\lim_{x \rightarrow a, x \rightarrow a} f(x)$ и $\lim_{x \leftarrow a, x \rightarrow a} f(x)$ не обязательно должны существовать. Например, функция, равная $\sin 1/x$ для $x \neq 0$ и 0 для $x = 0$, разрывна в начале координат, но не имеет в этой точке разрыва первого рода.

Теорема 3. *Если отображение f некоторого открытого множества Ω вещественной прямой \mathbb{R} в метрическое пространство F имеет лишь точки непрерывности и точки разрыва первого рода, то оно непрерывно всюду, кроме не более чем счетного множества точек множества Ω .*

Доказательство. Ограничимся доказательством для случая $\Omega = \mathbb{R}$. Пусть c — произвольная точка \mathbb{R} . Колебание $\omega(c)$ в точке c может быть каким угодно, но колебание $\omega(x)$ в точке x стремится к 0, когда x стремится к c по значениям, отличным от c . В самом деле, по определению $f(c+0)$, для любого $\varepsilon > 0$ существует такое $\eta > 0$, что из неравенства $c < x < c + \eta$ следует неравенство $d(f(x), f(c+0)) \leq \varepsilon/2$. Тогда для любых y и z из интервала $[c, c + \eta]$ имеем $d(f(y), f(z)) \leq d(f(y), f(c+0)) + d(f(c+0), f(z)) \leq \varepsilon$. Полагая

$y = x$ и устремляя z к x по значениям, строго большим x , а затем по значениям, строго меньшим x , получим $d(f(x+0), f(x)) \leq \epsilon$ и $d(f(x-0), f(x)) \leq \epsilon$. Устремляя $y < x$ и $z > x$ к x , получим, что $d(f(x-0), f(x+0)) \leq \epsilon$, откуда окончательно находим $\omega(x) \leq \epsilon$ для $c < x < c + \eta$. Действуя точно так же слева от точки c , можно убедиться, что утверждение верно и в этом случае. Рассмотрим теперь интервал $[-n, +n]$. Множество точек этого интервала, в которых колебание $\geq 1/k$, необходимо конечно. В самом деле, если бы это было не так, то можно было бы найти бесконечную последовательность точек разрыва в этом интервале, в которых колебание было бы $\geq 1/k$. Поскольку интервал компактен, из рассматриваемой последовательности можно извлечь такую подпоследовательность, которая сходилась бы к некоторой точке c этого интервала и состояла бы из точек, отличных от точки c (теорема 25 гл. II). Согласно только что доказанному, колебание в этих точках всюду $\geq 1/k$ и должно стремиться к 0, а это невозможно. Таким образом, множество точек интервала $[-n, +n]$, в которых колебание $\geq 1/k$, заведомо конечно. Если взять объединение этих особых точек для $k = 1, 2, 3, \dots$, то мы увидим, что множество точек x интервала $[-n, n]$, в которых колебание $\omega(x) > 0$, не более чем счетно и, следовательно, на всей числовой прямой \mathbb{R} множество таких точек, являясь объединением счетного множества не более чем счетных множеств, не более чем счетно. В каждой другой точке x , не принадлежащей этому не более чем счетному множеству, $\omega(x) = 0$ и f непрерывна.

Функцию, определенную на части прямой \mathbb{R} со значениями в топологическом пространстве F и имеющую только точки непрерывности или точки разрыва первого рода, мы будем называть *правильной*. Согласно доказанному, если пространство F метризуемо, то такая функция непрерывна всюду за исключением, быть может, не более чем счетного множества особых точек.

Замечание. Это бесконечное счетное множество точек разрыва может встретиться в действительности. Рассмотрим, например, вещественную функцию f вещественной переменной, определенную следующим образом:

$$f(x) = 0, \text{ если } x \text{ иррациональна,}$$

$$f\left(\frac{p}{q}\right) = \frac{1}{q}, \text{ если } x = \frac{p}{q} \text{ — несократимая дробь } (q > 0).$$

Легко видеть, что если x стремится к a по значениям, отличным от a , то $f(x)$ стремится к нулю. В самом деле, зададим $\epsilon > 0$. Пусть q_0 — целое число $\geq 1/\epsilon$. На $[a-1, a+1]$

имеется лишь конечное число рациональных чисел, знаменатели которых $< q_0$. Следовательно, существует такое $\eta > 0$, что каждое рациональное число из интервала $[a - \eta, a + \eta]$ имеет знаменатель $\geq q_0$, кроме, быть может, самой точки a , если она рациональна. Но тогда из неравенства $|x - a| \leq \eta, x \neq a$, следует, что $f(x) \leq \varepsilon$, — тем самым наше утверждение доказано. Мы видим, что f непрерывна в точке a тогда и только тогда, когда a иррациональна. Если a рациональна, то f имеет в точке a разрыв первого рода, причем $f(a+0) = f(a-0) = 0$. Следовательно, рассматриваемая функция правильная и имеет счетное (и плотное) множество точек разрыва. Ее скачок (но не колебание!) повсюду равен нулю.

Если F является топологическим не метризуемым пространством, то функция, определенная на \mathbb{R} со значениями в F , может быть правильной и всюду разрывной.

Производная вещественной функции вещественной переменной

Пусть f — отображение открытого множества Ω из \mathbb{R} в $F = \mathbb{R}$. Производной отображения f в точке $a \in \Omega$ называется предел, если он существует, отношения $[f(a+h) - f(a)]/h$ при h , стремящемся к 0 по значениям $\neq 0$, таким, что $a+h$ принадлежит Ω . Поскольку Ω является открытым множеством, то $a+h$ при достаточно малом $|h|$ принадлежит Ω . Если этот предел существует только при h , стремящемся к 0 по значениям > 0 , то его называют *производной отображения f справа*. Можно точно также говорить о *производной отображения f в точке a слева*. Производная отображения f в точке a существует тогда и только тогда, когда существуют производные справа и слева и обе они равны между собой¹⁾.

Функция, имеющая производную в точке a , непрерывна в этой точке. Если она имеет производную в a справа, то она непрерывна справа в a . Существование и значение производной функции f в $a \in \Omega$ зависят от значения f в окрестности a . Производная f в точке a обозначается обычно через $f'(a)$, или $\frac{df}{dx}(a)$, или $Df(a)$. Если производная существует всюду в Ω , то функция $x \rightarrow f'(x)$ называется *производной функцией*.

¹⁾ Если вместо открытого множества в качестве Ω берут замкнутый интервал $[a, b]$, то говорят также, что f имеет производную в точке a (соответственно в точке b), если она имеет производную справа (соответственно слева) в этой точке. Впрочем, это полностью соответствует общему определению производной, поскольку рассматриваются значения h , при которых $a+h$ (или $b+h$) принадлежат Ω .

ции f , или просто производной. Ее обозначают через f' , или $\frac{df}{dx}$, или Df . В свою очередь можно исследовать, имеет ли сама эта функция производную, что приводит к понятию второй производной, третьей производной и т. д. Производная порядка m в точке a обозначается через $f^{(m)}(a)$, или $\frac{d^m f}{dx^m}(a)$, или $\left(\frac{d}{dx}\right)^m f(a)$, или $D^m f(a)$. Производная функции порядка m обозначается через $f^{(m)}$, или $\frac{d^m f}{dx^m}$, или $\left(\frac{d}{dx}\right)^m f$, или $D^m f$.

Важно заметить, что о производной m -го порядка в точке $a \in \Omega$ можно говорить лишь в том случае, когда существуют все производные порядков $\leq m-1$ в множестве Ω или, по крайней мере, в некоторой окрестности точки a в Ω .

Говорят, что f дифференцируема m раз в Ω , если она имеет производную порядка m в каждой точке множества Ω . При этом она заведомо k раз дифференцируема для всех $k \leq m$, и ее производные порядков $\leq m-1$ непрерывны в Ω . Говорят, что f непрерывно дифференцируема m раз, или принадлежит классу C^m , если она имеет непрерывные производные до m -го порядка включительно. При этом она заведомо принадлежит классу C^k для $k \leq m$. Функция, дифференцируемая m раз, наверняка принадлежит классу C^{m-1} . Говорят, что f бесконечно дифференцируема или принадлежит классу C^∞ , если она имеет последовательные производные всех порядков. Все они непрерывны.

Однако в некоторых случаях принятые обозначения могут привести к ошибкам. Что означает, например, обозначение $f'(2x)$? Означает ли оно производную функции $x \rightarrow f(2x)$ или же значение производной f' функции f в точке с абсциссой $2x$? Мы будем понимать эти обозначения лишь во втором смысле. Если мы пожелаем записать производную функции $x \rightarrow f(2x)$ в точке a , то мы будем писать $(f(2x))'_{x=a}$, а производную функции функции $x \rightarrow f(2x)$ будем записывать в виде $(f(2x))'$. Таким образом, имеем: $(f(2x))' = 2f'(2x)$, $(f(2x))^{(m)} = 2^m f^{(m)}(2x)$. Следует отличать также $D^m f(2x)$, или $(D^m f)(2x) = f^{(m)}(2x)$, от $D^m(f(2x)) = (f(2x))^{(m)}$.

Можно, естественно, говорить о производной, принимающей значения $+\infty$ или $-\infty$, и то же самое для производной слева или производной справа. Однако, если это специально не оговорено, под производной мы всегда будем понимать производную, принимающую конечное значение.

Пусть f — некоторая вещественная функция, определенная на интервале $[a, b]$ прямой \mathbb{R} . Будем говорить, что эта функция кусочно m раз непрерывно дифференцируема, или кусочно принадлежит классу C^m , если она правильна и существуют

такие точки $a_0 = a, a_1, \dots, a_{n-1}, a_n = b$, $a_i < a_{i+1}$, что функции f_i , равные f в $]a_i, a_{i+1}[$, $f(a_i + 0)$ в a_i и $f(a_{i+1} - 0)$ в a_{i+1} , принадлежат классу C^m на интервале $[a_i, a_{i+1}]$. Например, если график функции $y = f(x)$ представляет собой ломаную линию, то f непрерывна, не дифференцируема, но кусочно принадлежит классу C^∞ .

Теорема 4 (Ролля). *Пусть вещественная непрерывная на замкнутом ограниченном интервале $[a, b]$ функция f имеет в каждой точке открытого интервала $]a, b[$ производную, конечную, или равную $+\infty$, или равную $-\infty$. Если при этом $f(a) = f(b) = 0$, то существует по крайней мере одна точка c в $]a, b[$, такая, что $f'(c) = 0$.*

Доказательство. Если функция f тождественно равна нулю, то результат очевиден. Если же эта функция не является тождественным нулем, то она принимает по крайней мере одно значение > 0 или хотя бы одно значение < 0 . Рассмотрим, например, первый случай. Так как f непрерывна на компакте $[a, b]$, то существует хотя бы одна точка c из $]a, b[$, в которой она достигает своего максимума. В этой точке производная справа необходимо ≤ 0 , в то время как производная слева ≥ 0 . Поскольку функция в любой точке $]a, b[$ предполагалась дифференцируемой, то эти производные справа и слева в точке c равны между собой, и, следовательно, производная $f'(c)$ равна нулю.

Замечание. Хорошо видно, почему нет необходимости ни предполагать f дифференцируемой в точках a и b , ни считать производную конечной. Напротив, существование производной, а не только производной справа и производной слева является существенным. Функция $f(x) = 1 - |x|$ на интервале $[-1, +1]$ непрерывна, равна нулю на его концах и в каждой точке имеет производную справа и слева. В точке $x = 0$, где

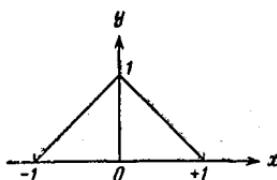


Рис. 5. $y = 1 - |x|$, $|x| \leq 1$.

f достигает своего максимума, производная справа равна -1 , а производная слева равна $+1$. Ни в одной точке ни производная справа, ни производная слева не обращаются в нуль. Если предполагать только, что f имеет в интервале $]a, b[$ про-

изводную справа f'_n , то можно доказать существование такой точки c_1 из $]a, b[$, в которой $f'_n(c_1) \leq 0$, и существование такой точки c_2 из того же интервала $]a, b[$, в которой $f'_n(c_2) \geq 0$ ¹⁾.

Теорема 5 (формула конечных приращений). *Пусть вещественная непрерывная на замкнутом интервале $[a, b]$ функция имеет в каждой точке открытого интервала $]a, b[$ производную, конечную или равную $+\infty$, или равную $-\infty$. Тогда на $]a, b[$ существует по крайней мере одна точка c , такая, что*

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \quad (\text{III}, 2; 1)$$

Формула конечных приращений часто записывается в следующем виде

$$\frac{f(x + h) - f(x)}{h} = f'(x + \theta h), \quad \text{где } \theta \in]0, 1[. \quad (\text{III}, 2; 2)$$

Замечание. Если предполагать только существование производной справа f'_n , то можно доказать существование такой точки c_1 , что $f'_n(c_1) \leq [f(b) - f(a)]/(b - a)$, и такой точки c_2 , что $f'_n(c_2) \geq [f(b) - f(a)]/(b - a)$.

Теорема 6 (формула Тейлора). *Пусть функция f является m раз непрерывно дифференцируемой в интервале $[a, b]$ и в каждой точке открытого интервала $]a, b[$ имеет производную $(m+1)$ -го порядка, конечную, или равную $+\infty$, или равную $-\infty$. Тогда существует хотя бы одна точка c в $]a, b[$, такая, что*

$$\begin{aligned} f(b) - f(a) - (b - a)f'(a) - \frac{(b - a)^2}{2!}f''(a) - \dots - \frac{(b - a)^m}{m!}f^{(m)}(a) = \\ = \frac{(b - a)^{m+1}}{(m+1)!}f^{(m+1)}(c). \end{aligned} \quad (\text{III}, 2; 3)$$

Эта формула часто записывается в следующем виде:

$$\begin{aligned} f(x + h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \dots + \frac{h^m}{m!}f^{(m)}(x) + \\ + \frac{h^{m+1}}{(m+1)!}f^{(m+1)}(x + \theta h), \end{aligned} \quad (\text{III}, 2; 4)$$

где θ — некоторое число из $]0, 1[$.

Формула конечных приращений является, очевидно, частным случаем формулы Тейлора. Формула же Тейлора доказывается

¹⁾ Доказательство этого утверждения более тонкое. См. также замечание, следующее за леммой теоремы 13.

следующим образом: вводится вспомогательная функция

$$\begin{aligned} g(x) = & f(b) - f(x) - (b - x)f'(x) - \dots \\ & \dots - \frac{(b - x)^m}{m!} f^{(m)}(x) - \frac{(b - x)^{m+1}}{(m+1)!} \lambda, \quad (\text{III}, 2; 5) \end{aligned}$$

где число λ определено так, что $g(a) = 0$.

Поскольку $g(b) = 0$, то к этой функции применима теорема Ролля. Это дает некоторую точку c из $]a, b[$, в которой $g'(c) = 0$, а, значит, $f^{(m+1)}(c) = \lambda$. Отсюда и вытекает нужный результат.

Монотонные функции

Говорят, что вещественная функция, определенная на открытом множестве $\Omega \subset \mathbb{R}^1$, монотонна, если она является возрастающей или убывающей на этом множестве. Монотонная функция не обязательно непрерывна, однако, согласно известной теореме математического анализа, она заведомо обладает пределом справа $f(a+0)$ и пределом слева $f(a-0)$ в каждой точке $a \in \Omega$ ¹). Оба эти предела не обязательно равны между собой или равны $f(a)$.

Монотонная функция, следовательно, является правильной. Применяя теорему 3, получаем, что такая функция всюду непрерывна, за исключением не более, чем счетного множества точек.

Небесполезно привести простой пример строго возрастающей функции, имеющей счетное плотное множество точек разрыва. Пусть h — некоторая функция > 0 , определенная на множестве рациональных чисел \mathbb{Q} и такая, что $\sum_{r \in \mathbb{Q}} h(r)$ конечна (речь идет о сумме счетного множества положительных чисел, не заданных в определенном порядке, смысл которой был определен на стр. 131). Рассмотрим теперь функцию f , определенную следующим образом:

$$f(x) = \sum_{r \in \mathbb{Q}, r < x} h(r). \quad (\text{III}, 2; 6)$$

¹) Как было сказано на стр. 195, Ω также может быть неоткрытым интервалом.

²) Если, например, f возрастает, то $f(a+0) = \inf_{x > a} (f(x))$ и $f(a-0) = \sup_{x < a} (f(x))$. При этом $f(a-0) \leq f(a) \leq f(a+0)$; скачок f в точке a не отрицателен.

Она всюду определена, >0 или возрастает. Это даже строго возрастающая функция, поскольку имеет место формула

$$f(x) - f(y) = \sum_{r \in Q, y \leq r < x} h(r) \quad \text{для } y < x. \quad (\text{III}, 2; 7)$$

Пусть a — произвольное вещественное число. Зададим $\epsilon > 0$. Поскольку сумма $\sum_{r \in Q} h(r)$ сходится, то существует конечное число рациональных точек r_0, r_1, \dots, r_n , таких, что

$$\sum_{r \in Q, r \neq r_0, r_1, \dots, r_n} h(r) \leq \epsilon. \quad (\text{III}, 2; 8)$$

При этом существует такое число $\eta > 0$, что

$$x \neq r_0, r_1, \dots, r_n \quad \text{для } a - \eta \leq x \leq a + \eta, \quad x \neq a. \quad (\text{III}, 2; 9)$$

В силу (III, 2; 7) имеют место следующие два неравенства:

$$\begin{aligned} f(a) &\geq f(x) \geq f(a) - \epsilon \quad \text{для } a - \eta \leq x < a, \\ f(a) + h(a) &\leq f(x) \leq f(a) + h(a) + \epsilon \quad \text{для } a \leq x \leq a + \eta \end{aligned} \quad (\text{III}, 2; 10)$$

при условии, что $h(a) = 0$ в случае иррационального a . Из этих неравенств следует, что $f(a - 0) = f(a)$ и $f(a + 0) = f(a) + h(a)$.

Функция f , таким образом, всюду непрерывна слева. Она непрерывна справа и, следовательно, просто непрерывна в каждой иррациональной точке. В каждой рациональной точке она разрывна и имеет скачок, равный $h(a)$. Она имеет счетное плотное множество точек разрыва первого рода, а именно множество рациональных точек.

Возрастающая функция не обязательно дифференцируема, однако для дифференцируемых функций имеется очень важный критерий возрастания.

Теорема 7. Пусть f — вещественная функция, определенная на некотором интервале (открытом, замкнутом или полуоткрытом) прямой \mathbb{R} и имеющая всюду производную, конечную или равную $+\infty$ или $-\infty$. Для того чтобы она была возрастающей, необходимо и достаточно, чтобы ее производная была всюду ≥ 0 .

Доказательство. Если функция возрастает, то, очевидно, $[f(x + h) - f(x)]/h \geq 0$ и, следовательно, предел этого отношения, который по предположению существует, также ≥ 0 .

Обратно, предположим, что f непрерывна, дифференцируема и имеет производную (конечную или нет) всюду ≥ 0 . Тогда, согласно теореме о конечных приращениях, всюду имеем

$[f(x+h) - f(x)]/h = f'(x + \theta h) \geqslant 0$, а это говорит о том, что f — возрастающая функция.

Замечания. 1°) Точно такое же рассуждение показывает, что если производная $f' > 0$, то f строго возрастает. Однако может случиться, что f строго возрастает, а ее производная не будет всюду > 0 . Такой будет, например, функция $f(x) = x^3$, производная которой повсюду $\geqslant 0$ и равна нулю в начале координат.

2°) Используя замечания, приведенные после теорем 4 и 5, можно доказать такую же теорему с производной справа (или производной слева) вместо обычной производной.

Дифференцируемые функции и теоремы о промежуточных значениях

Производная функции не обязательно непрерывна. В этом можно убедиться на примере функции, определенной следующим образом: $f(x) = x^2 \sin 1/x$ для $x \neq 0$ и $f(0) = 0$. Она всюду непрерывна и в любой точке $x \neq 0$ имеет производную $f'(x) = 2x \sin 1/x - \cos 1/x$. Производная этой функции в начале координат равна нулю. Таким образом, ее производная определена всюду в \mathbb{R} , но разрывна в точке $x = 0$, ибо $\cos 1/x$ не стремится к 0, когда $x \neq 0$ стремится к 0. Однако, если некоторая функция определена на интервале и дифференцируема, то ее производная функция обладает, как и все непрерывные функции, тем свойством, что вместе с любыми двумя своими значениями она принимает и все промежуточные значения. Для простоты предположим, что функция определена на \mathbb{R} . Пусть $f'(a) = \alpha$, $f'(b) = \beta$ и γ — число, лежащее строго между α и β . Тогда, выбирая достаточно малым число h , получим $[f(a+h) - f(a)]/h = \alpha' < \gamma$ и $[f(b+h) - f(b)]/h = \beta' > \gamma$. Если зафиксировать одно из таких значений h , то мы получим непрерывную функцию $x \rightarrow [f(x+h) - f(x)]/h$. В силу непрерывности эта функция вместе со значениями α' и β' принимает и все промежуточные значения и, в частности, значение γ , т. е. существует такая точка x , что $[f(x+h) - f(x)]/h = \gamma$. Если теперь к интервалу $[x, x+h]$ применить формулу конечных приращений, то мы придем к тому, что существует такая точка c , в которой $f'(c) = \gamma$.

Выпуклые функции

Вещественная функция f , определенная на интервале I вещественной прямой \mathbb{R} , называется *выпуклой*, если она обладает следующим свойством:

каковы бы ни были точки $A_1 = (a_1, f(a_1))$ и $A_2 = (a_2, f(a_2))$ графика функции f на плоскости \mathbb{R}^2 , хорда A_1A_2 лежит над дугой графика кривой f , соединяющей точки A_1 и A_2 . Это можно записать в виде следующего неравенства:

$$f(ta_1 + (1-t)a_2) \leq tf(a_1) + (1-t)f(a_2), \quad 0 \leq t \leq 1. \quad (\text{III}, 2; 6)$$

Этот же факт эквивалентен тому, что множество точек (x, y) , удовлетворяющих неравенству $y \geq f(x)$, является выпуклым множеством плоскости \mathbb{R}^2 (см. определение выпуклых множеств на стр. 184).

Учитывая сказанное ранее о барицентрических свойствах выпуклых множеств, из соотношения (III, 2; 6) можно получить более общее соотношение

$$f\left(\frac{\alpha_1 x_1 + \dots + \alpha_n x_n}{\alpha_1 + \dots + \alpha_n}\right) \leq \frac{\alpha_1 f(x_1) + \dots + \alpha_n f(x_n)}{\alpha_1 + \dots + \alpha_n} \quad (\text{III}, 2; 7)$$

для $x_1, x_2, \dots, x_n \in I$ и $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$.

Теорема 7₂. Для того чтобы вещественная функция f , определенная на интервале $I \subset \mathbb{R}$, была выпуклой, необходимо и достаточно, чтобы она обладала следующими свойствами:

1°) Функция f непрерывна во внутренности $\overset{\circ}{I}$ интервала I , и, кроме того, если I полуоткрыт или замкнут, она удовлетворяет в соответствующих концах I соотношению

$$\lim_{x \in \overset{\circ}{I}, x \rightarrow l} f(x) \leq f(l)^1). \quad (\text{III}, 2; 8)$$

2°) В каждой точке x из $\overset{\circ}{I}$ функция f имеет производную слева f'_l и производную справа f'_r . Эти производные совпадают всюду, кроме не более чем счетного множества точек из $\overset{\circ}{I}$, при этом имеют место такие неравенства:

$$f'_l \leq f'_r, \quad f'_r(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'_r(x_2) \quad \text{для } x_1 \leq x_2. \quad (\text{III}, 2; 9)$$

Отсюда, в частности, вытекает, что если f является дифференцируемой функцией, то она выпукла тогда и только тогда, когда ее первая производная является возрастающей функцией. Если же функция f дважды дифференцируема, то она выпукла тогда и только тогда, когда ее вторая производная $f''(x) \geq 0$.

Доказательство. 1°) Предположим вначале, что f выпукла. Если $0 < h < k$ и $[x, x+k] \subset I$, то точка $(x+h, f(x+h))$

¹⁾ То есть f полунепрерывна на концах интервала I .

лежит вне отрезка прямой $[(x, f(x)), (x + k, f(x + k))]$. Отсюда получаем следующее неравенство:

$$\frac{f(x + h) - f(x)}{h} \leq \frac{f(x + k) - f(x)}{k}. \quad (\text{III}, 2; 10)$$

Это неравенство показывает, что функция $h \rightarrow [f(x + h) - f(x)]/h$ является возрастающей при $h > 0$. Следовательно, она имеет предел при h , стремящемся к 0. Другими словами, f в каждой точке \dot{I} имеет производную справа, конечную или равную $-\infty$.

Кроме того,

$$f'_n(x) \leq \frac{f(x + k) - f(x)}{k}, \quad (\text{III}, 2; 11)$$

что дает левую часть второй системы неравенств (III, 2; 9) при $x = x_1$, $x + k = x_2$.

Аналогичное рассуждение можно провести с числами h и $k < 0$ и проверить, что f в каждой точке \dot{I} имеет производную слева f'_n , конечную или равную $+\infty$, и что справедлива правая часть второй системы неравенств (III, 2; 9).

Если в (III, 2; 10) положить $x + h = y$, $x + k = y + l$, то мы получим

$$\frac{f(y) - f(y - h)}{h} \leq \frac{f(y + l) - f(y - h)}{h + l}, \quad (\text{III}, 2; 12)$$

и аналогичное рассуждение, проведенное с отрицательными приращениями, дает

$$\frac{f(y + l) - f(y - h)}{h + l} \leq \frac{f(y + l) - f(y)}{l}. \quad (\text{III}, 2; 13)$$

Заметим только, что

$$\frac{f(y - h) - f(y)}{-h} \leq \frac{f(y + l) - f(y)}{l}. \quad (\text{III}, 2; 14)$$

Устремляя h и l к 0, получаем $f'_n(y) \leq f'_n(y)$, чем и заканчивается доказательство неравенств (III, 2; 9). Заметим, кроме того, что обе эти производные конечны.

Функция f'_n является, таким образом, возрастающей в \dot{I} , а, значит, имеет не более чем счетное множество точек разрыва. Если x не является точкой разрыва функции f'_n , то при $h > 0$ необходимо имеем: $f'_n(x - h) \leq f'_n(x) \leq f'_n(x)$, откуда, устремляя h

к нулю, находим $f'_n(x) = f'(x)$. Таким образом, f имеет производную всюду, кроме не более чем счетного множества точек.

Далее, функция f , будучи дифференцируемой справа и слева в каждой точке, непрерывна справа и слева, а, следовательно, непрерывна всюду во внутренности $\overset{\circ}{I}$ интервала I . Поскольку f'_n возрастает, то она либо везде ≤ 0 на I , либо всюду ≥ 0 , либо ≤ 0 строго слева от некоторой точки c и ≥ 0 строго справа от c . Поэтому функция f либо возрастает в $\overset{\circ}{I}$, либо убывает в $\overset{\circ}{I}$ или же убывает слева от c и возрастает справа от c . Так как эта функция непрерывна, то в точке c она достигает своего минимума. Во всех случаях она необходимо имеет предел справа на левом конце интервала I и предел слева на правом его конце.

Рассмотрим, например, случай правого конца b и предположим, что $b \in I$. Если $b' < x < b$, то точка $(x, f(x))$ будет лежать под хордой $[(b', f(b')), (b, f(b))]$. Переходя к пределу при x , стремящемся к b , мы получим то же соотношение в точке $(b, f(b-0))$, а это означает, что имеет место неравенство $f(b-0) \leq f(b)$. Тем самым заканчивается доказательство всех свойств функции f , указанных в теореме.

2°) Предположим, обратно, что f обладает всеми указанными в утверждении теоремы свойствами, но вместо второй системы неравенств (III, 2; 9) нам достаточно будет иметь неравенство $f'_n(x_1) \leq f'_n(x_2)$. Докажем, что она выпукла во внутренности $\overset{\circ}{I}$ интервала I , а тогда из неравенства (III, 2; 8) будет следовать, что она такова же во всем интервале I , если он полуоткрыт или замкнут. Итак, пусть a, b — две точки из $\overset{\circ}{I}$, $a < b$. Функция g , определенная формулой $g(x) = f(x) - f(a) - (x-a) \frac{f(b) - f(a)}{b - a}$, удовлетворяет всем перечисленным условиям в $\overset{\circ}{I}$ и, кроме того, $g(a) = g(b) = 0$. Для доказательства выпуклости f нам достаточно показать, что $g(x) \leq 0$ для $a \leq x \leq b$. Если бы это было не так, то функция g имела бы в интервале $[a, b]$ максимум > 0 в некоторой точке c . Рассуждения, проведенные в теореме 4 (Ролля), показывают, что $g'_n(c) \geq 0$ и $g'_n(c) \leq 0$. Из неравенства (III, 2; 9) (1-я система) следует, что обе эти величины равны нулю. Однако, поскольку g'_n , согласно (III, 2; 9), возрастает, она заведомо ≥ 0 в интервале $[c, b]$. Максимум $g(c) > 0$ должен быть, следовательно, $\leq g(b) = 0$, а это невозможно. Тем самым выпуклость функции f доказана.

Если f дифференцируема в I , то предыдущие условия сводятся к тому, что f' является возрастающей, а если же она дважды дифференцируема, то к тому, что $f'' \geq 0$.

§ 3. ПРОИЗВОДНАЯ ОТОБРАЖЕНИЯ ОДНОГО АФФИННОГО ПРОСТРАНСТВА В ДРУГОЕ.

ПРОИЗВОДНЫЙ ВЕКТОР ФУНКЦИИ СКАЛЯРНОЙ ПЕРЕМЕННОЙ

Рассмотрим отображение f открытого множества поля скаляров K в аффинное нормированное пространство F^1). Для $a \in \Omega$ можно придать смысл формуле

$$\overrightarrow{f'(a)} = \lim_{h \neq 0, h \rightarrow 0, a+h \in \Omega} \frac{\overrightarrow{f(a+h) - f(a)}}{h} \in \vec{F}. \quad (\text{III}, 3; 1)$$

В правой части мы имеем, прежде всего, разность $\overrightarrow{f(a+h) - f(a)}$ двух точек из F , которая является некоторым вектором присоединенного векторного пространства \vec{F} . Этот вектор можно делить на скаляр $h \neq 0$ и искать предел полученного вектора в \vec{F} , когда h стремится к 0, поскольку векторное пространство \vec{F} предполагалось нормированным. Если $\overrightarrow{f'(a)}$ существует, то его называют производным вектором отображения f в точке a . Существование производной и ее значение зависят не от нормы, а только от топологии пространства F , поскольку понятие производной связано с понятием предела. Если $K = \mathbb{R}$, то можно точно так же говорить о производной слева и производной справа. Если производная существует всюду в Ω , то можно рассматривать производную функцию $\vec{f}: x \rightarrow \overrightarrow{f'(x)}$. Это — некоторое отображение Ω в векторное нормированное пространство \vec{F} . Можно затем вычислить последующие производные при тех же условиях, что и в § 2. Они обозначаются так же, как и в случае вещественных функций, а именно: \vec{f}'' , \vec{f}''' , ..., $\vec{f}^{(m)}$, ..., и т. д. Все эти производные, если только они существуют, являются отображениями Ω в \vec{F} . Заметим, что f принимает свои значения в аффинном пространстве F , а ее производные \vec{f}' , \vec{f}'' , ... принимают свои значения в присоединенном векторном пространстве \vec{F} . Если положить $E = F = \mathbb{R}$, то мы вернемся к обычной производной вещественной функции вещественной переменной.

Теперь можно говорить о классах C^1 , C^2 , ..., C^m , ..., C^∞ функций, непрерывно дифференцируемых 1 раз, 2 раза, ..., m раз или бесконечно дифференцируемых со значениями в F .

¹⁾ В качестве K берется \mathbb{R} или C . Пространство предполагается аффинным над полем K . Если $K = \mathbb{R}$ и F задано как аффинное пространство над полем C , то ограничиваются рассмотрением F как аффинного пространства над полем \mathbb{R} .

Здесь, как и в § 2, дифференцируемая функция непрерывна, и о второй производной в точке $a \in \Omega$ говорят лишь в том случае, когда первая производная определена всюду в Ω или, по крайней мере, в некоторой окрестности точки a . Важный пример такого отображения имеется в механике. В этом случае $\mathbb{K} = \mathbb{R}$, переменная x есть переменная времени t , а пространство F является обычным трехмерным аффинным пространством. Движение частицы в этом пространстве задается функцией времени $t \rightarrow M(t)$ — функцией вещественной переменной t со значениями в F . Ее первая производная $d\vec{M}/dt$ является вектором скорости, а вторая производная $d^2\vec{M}/dt^2$ — вектором ускорения. Обе они принадлежат \vec{F} .

Если пространство F конечномерно и в нем выбрана система координат, состоящая из начала b и некоторого базиса $(\vec{f}_i)_{i \in I}$ из \vec{F} , то положение каждой точки пространства F определяется через ее координаты $(y_i)_{i \in I}$, а задание функции f , определенной на $\Omega \subset \mathbb{K}$ со значениями в F , эквивалентно заданию скалярных функций $(F_i)_{i \in I}$ по формуле:

$$f(x) = b + \sum_{i \in I} F_i(x) \vec{f}_i, \quad \text{где } y_i = F_i(x). \quad (\text{III}, 3; 2)$$

При этих условиях производная функции задается формулой:

$$\overrightarrow{f'}(x) = \sum_{i \in I} F'_i(x) \vec{f}_i. \quad (\text{III}, 3; 3)$$

Для того чтобы функция со значениями в конечномерном аффинном нормированном пространстве была дифференцируемой, необходимо и достаточно, чтобы ее компоненты в некоторой системе координат были скалярными дифференцируемыми функциями. Компоненты производной являются производными соответствующих компонент.

Общий случай. Частная производная вдоль вектора

Пусть теперь f — отображение открытого множества Ω аффинного нормированного пространства E в аффинное пространство F . Очевидно, теперь нельзя говорить о производной в предыдущем смысле. Поэтому мы вначале введем понятие производной или частной производной вдоль вектора \vec{X} из \vec{E} . Пусть a — некоторая точка Ω : Производной f в точке a вдоль вектора \vec{X} называется производная, если она существует, функции

$t \rightarrow f(a + t\vec{X})$ при $t = 0$. Эта производная обозначается через $\overrightarrow{D_{\vec{X}}f}(a) \in \vec{F}$. Таким образом, имеем:

$$\overrightarrow{D_{\vec{X}}f}(a) = \left(\frac{d}{dt} (f(a + t\vec{X})) \right)_{t=0} = \lim_{\substack{t \neq 0, t \rightarrow 0 \\ a + t\vec{X} \in \Omega}} \frac{\overrightarrow{f(a + t\vec{X})} - \overrightarrow{f(a)}}{t}. \quad (\text{III}, 3; 4)$$

Здесь t — скаляр, пробегающий множество $K_{a, \vec{X}}$ элементов из K , для которых $a + t\vec{X} \in \Omega$. Функция $t \rightarrow f(a + t\vec{X})$ является отображением $K_{a, \vec{X}}$ в F . Множество $K_{a, \vec{X}}$ является прообразом открытого множества Ω при непрерывном отображении $t \rightarrow a + t\vec{X}$ множества K в E . Это — открытое подмножество K , содержащее нуль, что позволяет вычислить производную в точке $t = 0$.

Существование и значение производной вдоль \vec{X} в точке a зависит исключительно от топологии пространства \vec{F} , а не от его нормы, поскольку понятие производной связано лишь с понятием предела. Если $\vec{X} = \vec{0}$, то производная существует в каждой точке Ω и равна нулю. Если $K = \mathbb{R}$ и если задано направление полупрямой, то существует вектор \vec{X} , и притом единственный, с нормой, равной 1, имеющий это направление. Производная вдоль такого вектора \vec{X} называется производной по рассматриваемому направлению. Она зависит, очевидно, от нормы в пространстве \vec{E} . Например, если E — евклидово конечномерное пространство, S — регулярная гиперплоскость, содержащаяся в Ω , a — некоторая точка из S , \vec{v} — ориентированная нормаль к S в точке a , то нормальная производная $d/d\vec{v}$ в точке a является производной вдоль единичного вектора \vec{v} .

Если E — поле скаляров и \vec{X} — единица этого поля, то производная вдоль \vec{X} является обычной производной в смысле определения (III, 3; 1): $\overrightarrow{D_{\vec{X}}f}(a) = \overrightarrow{f'(a)}$. Если $\overrightarrow{D_{\vec{X}}f}(x)$ существует для всех x , то можно рассматривать производную функцию вдоль \vec{X} функции f или просто производную функции f вдоль \vec{X} . Это будет функция $\overrightarrow{D_{\vec{X}}f}: x \rightarrow \overrightarrow{D_{\vec{X}}f}(x)$. Для фиксированного \vec{X} это — отображение Ω в \vec{F} . Затем можно, в свою очередь, искать производную (если она существует) функции $\overrightarrow{D_{\vec{X}}f}$ в точке a .

вдоль вектора \vec{Y} (отличного или нет от вектора \vec{X}). Ее обозначают через $\overrightarrow{D_{\vec{Y}} D_{\vec{X}} \vec{f}}(a)$ и называют частной производной второго порядка. Точно так же определяются производные высших порядков.

Матрица Якоби. Якобиан

Если пространство F конечномерно и $b, (\vec{f}_i)_{i \in I}$ образуют систему координат в F , то имеют место формулы:

$$\vec{f}(x) = b + \sum_{i \in I} F_i(x) \vec{f}_i, \quad \overrightarrow{D_{\vec{X}} \vec{f}}(x) = \sum_{i \in I} D_{\vec{X}} F_i(x) \vec{f}_i^1. \quad (\text{III}, 3; 5)$$

Предположим теперь, что E также конечномерно и $a, (\vec{e}_j)_{j \in J}$ — система координат в E . Тогда производные по векторам \vec{e}_j базиса E являются тем, что обычно называют частными производными функции f . Другими словами, по определению:

$$\overrightarrow{\partial_j f}(x) = \frac{\overrightarrow{\partial f}}{\partial x_j}(x) = \overrightarrow{D_{\vec{e}_j} f}(x) = \lim_{t \neq 0, t \rightarrow 0} \frac{\overrightarrow{f}(x + t \vec{e}_j) - \overrightarrow{f}(x)}{t}. \quad (\text{III}, 3; 6)$$

Наконец, если E и F — конечномерные пространства и если в E выбрана система координат, то производная $\overrightarrow{\partial_j f} = \frac{\overrightarrow{\partial f}}{\partial x_j}$ выражается следующим образом:

$$\frac{\overrightarrow{\partial f}}{\partial x_j} = \sum_{i \in I} \frac{\partial F_i}{\partial x_j} \vec{f}_i. \quad (\text{III}, 3; 7)$$

Матрица, составленная из элементов $\frac{\partial F_i}{\partial x_j}(x)$, в частном случае при $I = \{1, 2, \dots, m\}, J = \{1, 2, \dots, n\}$

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x) & \frac{\partial F_1}{\partial x_2}(x) & \dots & \frac{\partial F_1}{\partial x_n}(x) \\ \vdots & \ddots & \ddots & \ddots \\ \frac{\partial F_m}{\partial x_1}(x) & \frac{\partial F_m}{\partial x_2}(x) & \dots & \frac{\partial F_m}{\partial x_n}(x) \end{pmatrix}, \quad (\text{III}, 3; 8)$$

называется *производной матрицей*, или *матрицей Якоби*, функции f в точке $x \in \Omega$. Если $m = n$, то ее определитель называется

¹⁾ Поскольку F_i — скалярные функции, $D_{\vec{X}} F_i(x)$ являются скалярами.

якобианом функции f в точке x по отношению к рассматриваемой системе координат. Часто через $\frac{D(y_1, \dots, y_n)}{D(x_1, \dots, x_n)}$ обозначают якобиан функции $y = f(x)$, определенной в заданной системе координат скалярными функциями $y_i = F_i(x) = F_i(x_1, \dots, x_n)$, $i = 1, 2, \dots, n$. Говорят также, что это якобиан n функций $y_i = F_i$ от n переменных x_1, x_2, \dots, x_n . Для $n = 1$ якобиан сводится к обычной производной.

Недостатки понятия производной вдоль вектора

Понятие производной вдоль вектора обладает рядом недостатков. В самом деле:

1°) Функция может в каждой точке иметь производную вдоль любого вектора и в то же время не быть непрерывной. Рассмотрим, например, скалярную функцию, определенную на \mathbb{R}^2 формулой

$$f(x, y) = \begin{cases} 0 & \text{для } (x, y) = (0, 0), \\ \frac{x^5}{(y - x^2)^2 + x^8} & \text{для } (x, y) \neq (0, 0). \end{cases} \quad (\text{III}, 3; 9)$$

Эта функция является отношением двух полиномов. Правила дифференцирования суммы, произведения и частного для производной вдоль вектора те же, что и для производной функции скалярной переменной, поскольку она сводится к производной некоторой функции от t при $t = 0$ (формула (III, 3; 4)). Так как знаменатель функции f в каждой точке $\neq (0, 0)$ в нуль не обращается, то эта функция дифференцируема в каждой точке. Вычислим ее производную в начале координат вдоль вектора (X, Y) . Если $Y \neq 0$, то для $t \neq 0$ имеем

$$f(tX, tY) = \frac{t^5 X^5}{t^2 Y^2 + \dots} = \frac{X^5}{Y^2} t^3 + \dots \quad (\text{III}, 3; 10)$$

Поскольку функция f в начале координат равна нулю, то при $t = 0$ ее производная в начале координат вдоль рассматриваемого вектора равна нулю. Если $Y = 0$, $X \neq 0$, то для $t \neq 0$

$$f(tX, tY) = \frac{t^5 X^5}{t^4 X^4 + \dots} = tX + \dots \quad (\text{III}, 3; 11)$$

Так как f в начале координат равна нулю, то при $t = 0$ ее производная вдоль вектора $(X, 0)$ равна X . Производная же вдоль вектора $(0, 0)$ всегда равна нулю. Таким образом, функция f дифференцируема вдоль любого вектора. Однако она разрывна в начале: на параболе $y - x^2 = 0$ имеем $f(x, y) = x^5/x^8 = 1/x^3$ для $x \neq 0$, и это выражение стремится к ∞ , когда $x \neq 0$ стремится к 0.

2°) Если не делать никаких предположений о непрерывности частных производных, то может оказаться, что никакой связи между производными вдоль различных векторов из \vec{E} в одной и той же точке области Ω не существует. Конечно, если $\overrightarrow{D_X}f(a)$ существует, то для скалярного λ будет существовать $\overrightarrow{D_{\lambda X}}f(a)$, и при этом

$$\overrightarrow{D_{\lambda X}}f(a) = \lambda \overrightarrow{D_X}f(a). \quad (\text{III}, 3; 12)$$

В самом деле, это очевидно, если $\lambda = 0$. Если же $\lambda \neq 0$, то при $t \in \mathbb{K}$ и $a + t\lambda \vec{X} \in \Omega$ отношение $[f(a + t\lambda \vec{X}) - f(a)]/t$ имеет вид $\lambda [f(a + s\vec{X}) - f(a)]/s$, где $s = t\lambda \in \mathbb{K}$, $a + s\vec{X} \in \Omega$ и $s \neq 0$ стремится к 0, если $t \neq 0$ стремится к 0. Отсюда мы получаем (III, 3; 12). Однако нам бы хотелось, чтобы для фиксированного a производная $\overrightarrow{D_X}f(a)$ линейно зависела от \vec{X} . Но это не так, как показывает тот же пример (III, 3; 9), в котором производная в начале координат вдоль вектора $(X, 0)$ равна X , тогда как производная вдоль вектора (X, Y) , $\vec{Y} \neq 0$, равна нулю.

Недостатки понятия частной производной вдоль вектора вытекают из того, что при ее вычислении в точке a используется поведение функции только на прямых аффинного пространства E , исходящих из той точки, тогда как было бы полезнее учесть глобальное поведение функции f во всей окрестности точки a . Поэтому мы введем новое понятие производного отображения.

Полная производная, или производное отображение

Пусть f — отображение открытого множества Ω аффинного нормированного пространства E в аффинное нормированное пространство F . Говорят, что f имеет в точке $a \in \Omega$ производное отображение, или полную производную L , или дифференциал, или полный дифференциал $\vec{L}\vec{h}$, если L является линейным непрерывным отображением \vec{E} в \vec{F} и если для $a + \vec{h} \in \Omega$ имеет место равенство

$$f(a + \vec{h}) = f(a) + L\vec{h} + \varphi(\vec{h})\|\vec{h}\|, \quad (\text{III}, 3; 13)$$

где $\varphi(\vec{h})$ стремится к 0 при $\vec{h} \neq 0$, стремящемся к 0. Это равносильно утверждению, что приращение $\vec{\Delta}f = f(a + \vec{h}) - f(a)$ может быть представлено в виде линейного непрерывного

отображения $L \cdot \vec{h}$ с погрешностью, которая является бесконечно малой более высокого порядка, чем $\|\vec{h}\|$, когда \vec{h} стремится к $\vec{0}$ в \vec{E} . Заметим, что $\varphi(\vec{0})$ может выбираться произвольно и не представляет никакого интереса, но $\varphi(\vec{h})$ для $\vec{h} \neq \vec{0}$ определяется единственным образом по формуле

$$\varphi(\vec{h}) = \overrightarrow{[f(a + \vec{h}) - f(a) - L \cdot \vec{h}]} \| \vec{h} \|.$$

То, что L является производной отображения f в точке a , означает, что вектор $\varphi(\vec{h})$, определенный для таких $\vec{h} \neq \vec{0}$, что $a + \vec{h} \in \Omega$, стремится к $\vec{0}$, когда \vec{h} стремится к $\vec{0}$. Обычно договариваются считать $\varphi(\vec{0}) = \vec{0}$, чтобы функция φ была бы непрерывной в начале пространства \vec{F} и чтобы можно было снять ограничение $\vec{h} \neq \vec{0}$.

В силу теоремы 12 гл. II, существование и значение производного отображения зависят только от топологии пространств E и F , а не от их норм.

Теорема 8. *Если отображение f имеет производную в точке a , то эта производная единственна. В этом случае f непрерывно в a . Кроме того, отображение f имеет в точке a производную вдоль любого вектора \vec{X} из E , а отображение $\vec{X} \rightarrow \overrightarrow{D_{\vec{X}} f}(a)$ есть линейное непрерывное отображение \vec{E} в \vec{F} , являющееся не чем иным, как самим L , т. е.*

$$\overrightarrow{D_{\vec{X}} f}(a) = L \cdot \vec{X}. \quad (\text{III}, 3; 14)$$

Таким образом, существование производного отображения L в точке a влечет за собой существование производной в точке a вдоль любого вектора \vec{X} из \vec{E} .

Доказательство. Поскольку отображение L предполагалось непрерывным, то, когда \vec{h} стремится к $\vec{0}$, $L\vec{h}$ также стремится к $\vec{0}$. То же самое будет верно и для $\varphi(\vec{h}) \| \vec{h} \|$, а потому отображение f непрерывно в точке a . Пусть $\vec{X} \in \vec{E}$. Полагая в (III, 3; 13) $\vec{h} = t\vec{X}$ и замечая, что для достаточно малого $|t|$ точка $a + t\vec{X}$ лежит в открытом множестве Ω , получаем следующую формулу:

$$\overrightarrow{\frac{f(a + t\vec{X}) - f(a)}{t}} = L \cdot \vec{X} + \frac{|t|}{t} \varphi(t\vec{X}) \| \vec{X} \|. \quad (\text{III}, 3; 15)$$

Так как выражение $L \cdot \vec{X}$ определено для любого вектора \vec{X} из \vec{E} , то, устремляя $t \neq 0$ к нулю, из (III, 3; 15) получаем (III, 3; 14) и одновременно устанавливаем единственность производной. Если через f обозначено отображение $\Omega \subset E$ в F , то через $f'(a)$, или $\frac{df}{dx}(a)$, можно обозначить производное отображение функции f в точке a , так что $f'(a) \in \mathcal{L}(\vec{E}; \vec{F})$. Через $f'(a)\vec{X}$, или $\frac{df}{dx}(a) \cdot \vec{X} \in \vec{F}$, можно будет обозначить значение этого отображения на векторе \vec{X} из \vec{E} .

Таким образом, имеем формулу:

$$\overrightarrow{D}_{\vec{X}} f(a) = f'(a) \vec{X} \in \vec{F}. \quad (\text{III, 3; 14}_2)$$

Замечания. 1°) Пусть E и F — аффинные пространства над полем \mathbb{C} комплексных чисел. Тогда они заведомо являются аффинными над вещественным полем \mathbb{R} . Отображение L пространства \vec{E} в \vec{F} линейно, когда \vec{E} и \vec{F} рассматриваются как векторные пространства над полем \mathbb{C} , и *a fortiori* обладает этим свойством, когда \vec{E} и \vec{F} рассматриваются как векторные пространства над полем \mathbb{R} . Значит, отображение f множества $\Omega \subset E$ в F имеет производную $L \in \mathcal{L}(\vec{E}; \vec{F})$ в точке $a \in \Omega$, когда E и F рассматриваются как аффинные пространства над полем \mathbb{C} , и *a fortiori* имеет отображение L в качестве производной, когда E и F рассматриваются как аффинные пространства над полем \mathbb{R} . Обратное, вообще говоря, не верно, в чем мы убедимся позже в теории аналитических функций комплексной переменной.

2°) Из определения производной вытекает, что $f'(a)$ не есть вектор пространства \vec{F} , а линейное непрерывное отображение пространства \vec{E} в пространство \vec{F} . Это $f'(a) \cdot \vec{X}$ для $\vec{X} \in \vec{E}$ является вектором из \vec{F} . Однако предположим, что \vec{E} является полем скаляров \mathbb{K} . Тогда мы можем определить производный вектор $\overrightarrow{f'(a)} \in \vec{F}$ по формуле (III, 3; 1) и производное отображение $f'(a) \in \mathcal{L}(\mathbb{K}; \vec{F})$ по формуле (III, 3; 13). Между этими двумя понятиями имеется простая связь. Если существует производный вектор, то существует производное отображение, и наоборот; при этом

$$\overrightarrow{f'(a)} = f'(a) \cdot \vec{1}, \quad \vec{1} \in \mathbb{K}. \quad (\text{III, 3; 15}_2)$$

В самом деле, если существует полная производная $f'(a) \in \mathcal{L}(\mathbb{K}; \vec{F})$, то из теоремы 8 следует, что существует производный вектор $\overrightarrow{f'(a)} = \overrightarrow{D_1 f}(a)$, равный $f'(a) \cdot \vec{1}$. Обратно, если существует производный вектор $\overrightarrow{f'(a)}$, то для каждого $h \in K$

$$\vec{f}(a+h) = \vec{f}(a) + h\overrightarrow{f'(a)} + \vec{a}|h|,$$

где \vec{a} стремится к $\vec{0}$, когда h стремится к 0. Это означает, что существует производная $f'(a) \in \mathcal{L}(\mathbb{K}; \vec{F})$, являющаяся отображением $X \rightarrow \overrightarrow{f'(a)} \cdot X$. Впрочем, так как $X \in \mathbb{K}$, мы можем писать или $f'(a) \cdot X$, или $\overrightarrow{f'(a)} X$.

Z Слово производная неоднозначно, поскольку оно может обозначать как производное отображение $f'(a)$, так и производный вектор $\overrightarrow{f'(a)} = f'(a) \cdot \vec{1}$. Эта неоднозначность почти не мешает на практике. Если $E = F = \mathbb{K}$, то производный вектор является обычной производной $f'(a) \in \mathbb{K}$, а производное отображение является гомотетией $X \rightarrow f'(a)X$ из \mathbb{K} в \mathbb{K} .

3°) Если пространство E конечномерно и в нем выбрана система координат a , $(\vec{e}_j)_{j \in J}$, то производная $f'(x) \in \mathcal{L}(\vec{E}; \vec{F})$ связана с частными производными $f'(x) \cdot \vec{e}_j = \frac{\partial \vec{f}}{\partial x_j}(x)$ следующим образом:

$$f'(x) \cdot \vec{X} = f'(x) \cdot \left(\sum_{j \in J} X_j \vec{e}_j \right) = \sum_{j \in J} X_j f'(x) \vec{e}_j = \sum_{j \in J} X_j \frac{\partial \vec{f}}{\partial x_j}(x).$$
(III, 3; 15₃)

Пусть теперь оба пространства E и F конечномерны и в каждом из них выбрана система координат, а именно: a , $(\vec{e}_j)_{j \in J}$ в E и b , $(\vec{f}_i)_{i \in I}$ в F . Тогда каждая точка из E и F полностью определяется своими координатами и отображение f из Ω в F может быть определено системой m функций от n скалярных переменных, а именно: $y_i = f_i((x_j)_{j \in J})$, $i \in I$. (Если $I = \{1, 2, \dots, m\}$ и $J = \{1, 2, \dots, n\}$, то это будут функции $y_i = f_i(x_1, x_2, \dots, x_n)$ для $i = 1, 2, \dots, m$.) В этом случае производное отображение в точке x , если оно существует, определяется следующим образом: полагают $\vec{X} = \sum_{j \in J} X_j \vec{e}_j$ и обозна-

чают через $\vec{Y} = \sum_{i \in I} Y_i \vec{f}_i$ его образ при производном отображении. Тогда

$$\sum_{i \in I} Y_i \vec{f}_i = f'(x) \cdot \vec{X} = \sum_{j \in J} \frac{\partial f}{\partial x_j}(x) X_j = \sum_{i, j} \frac{\partial F_i}{\partial x_j}(x) X_j \vec{f}_i, \quad (\text{III}, 3; 16)$$

откуда

$$Y_i = \sum_{j \in J} \frac{\partial F_i}{\partial x_j}(x) X_j, \quad i \in I.$$

Полученная формула показывает, что *матрица производного отображения $f'(x)$ по отношению к рассматриваемым системам координат является матрицей Якоби* (III, 3; 8). Как всегда, столбцы матрицы линейного преобразования представляют собой векторы, являющиеся образами векторов базиса \vec{E} при этом преобразовании; другими словами, $f'(x) \cdot \vec{e}_j = \frac{\partial \vec{f}}{\partial x_j}(x)$. Существование производного отображения $f'(x)$ влечет существование $\frac{\partial \vec{f}}{\partial x_j}$, а, следовательно, существование $\frac{\partial F_i}{\partial x_j}$ и существование матрицы Якоби, но обратное не верно, как показывает пример (III, 3; 9). Если E и F имеют одинаковую размерность, то *якобиан в x является определителем отображения $f'(x) \in \mathcal{L}(\vec{E}; \vec{F})$ по отношению к рассматриваемым системам координат*.

Напомним, что можно также говорить об определителе линейного отображения конечномерного пространства \vec{E} в себя, не выделяя при этом какого-либо базиса (поскольку определитель отображения, вычисленный в некотором базисе, не зависит от этого базиса). Таким образом, можно говорить о *якобиане f в точке x , или определителе отображения $f'(x)$, для отображения f открытого множества из E в E без выделения какого-либо базиса в пространстве E* .

Понятие дифференциала

Вместо того чтобы через x, y и т. д. обозначать точки пространств E и F , а через \vec{X}, \vec{Y} и т. д. — векторы пространств \vec{E} и \vec{F} , часто оказывается удобным, оставляя для точек пространств E и F обозначение x и y , векторы пространств \vec{E} и \vec{F} обозначать через \vec{dx}, \vec{dy} . Тогда выражение для производной отображения f множества Ω в F -запишется в виде

$$\vec{dx} \rightarrow \vec{dy} = f'(x) \cdot \vec{dx}. \quad (\text{III}, 3; 17)$$

В частности, если пространство E конечномерно и в нем выбрана система координат, то эта формула запишется следующим образом:

$$\vec{dx} = \sum_{j \in J} dx_j \vec{e}_j, \quad \vec{dy} = f'(x) \cdot \vec{dx} = \sum_{j \in J} \frac{\partial f}{\partial x_j} dx_j. \quad (\text{III}, 3; 18)$$

Когда f является функцией над полем \mathbb{K}^2 , т. е. функцией двух скалярных переменных x, y , то ее часто обозначают через \vec{z} , а через \vec{p} и \vec{q} обозначают ее частные производные по x и y . Полная производная в дифференциальных обозначениях запишется в виде

$$(dx, dy) \rightarrow \vec{dz} = \vec{p} dx + \vec{q} dy. \quad (\text{III}, 3; 19)$$

Последняя формула из (III, 3; 16) после замены X_j, Y_i на dx_j, dy_i примет такой вид

$$dy_i = \sum_{j \in J} \frac{\partial F_i}{\partial x_j} (x) dx_j, \quad i \in I. \quad (\text{III}, 3; 19A)$$

Геометрическая интерпретация производного отображения: дифференцируемое многообразие и линейное касательное многообразие

Пусть f — отображение открытого множества Ω пространства E в пространство F . Его графиком называется множество \mathcal{E} точек $(x, f(x))$, $x \in \Omega$, из $E \times F$. Если f дифференцируемо в каждой точке Ω , то говорят, что \mathcal{E} является *дифференцируемым многообразием* с уравнением $y=f(x)$. Если, например, $E=\mathbb{R}^2$, $F=\mathbb{R}$, то \mathcal{E} является поверхностью в \mathbb{R}^3 , задаваемой уравнением $z=f(x, y)$. Если $E=\mathbb{R}$, $F=\mathbb{R}^2$, то \mathcal{E} есть кривая в \mathbb{R}^3 , описываемая уравнениями $y=g(x)$, $z=h(x)$ ¹⁾.

Выясним, какой смысл следует придать понятию линейного многообразия, касательного к графику \mathcal{E} в точке $A=(a, f(a))$?

Рассмотрим произвольное множество \mathcal{E} аффинного нормированного пространства G . Пусть A — точка из \mathcal{E} . Говорят, что вектор \vec{X} из \vec{G} является *касательным в точке A к множеству \mathcal{E}* , если существуют последовательность точек $M_0, M_1, \dots, M_n, \dots$ из \mathcal{E} , стремящаяся к A при n , стремящимся к бесконечности, и последовательность вещественных скаляров $\lambda_0,$

¹⁾ Вместо x и y мы пишем x, y, z , поскольку $E \times F = \mathbb{R}^3$. Конечно, фраза «уравнением множества \mathcal{E} является $y=f(x)$ » ничего не означает. Она является очевидным сокращением фразы: « \mathcal{E} есть множество пар (x, y) из $E \times F$, удовлетворяющих равенству $y=f(x)$ », или « $\mathcal{E} = \{(x, y) : x \in \Omega, y=f(x)\}$ ».

$\lambda_1, \dots, \lambda_n, \dots \geq 0$, такие, что векторы $\lambda_n \overrightarrow{AM}_n$ стремятся к \vec{X} при n , стремящемся к бесконечности. Вектор $\vec{0}$ всегда является касательным. Если \vec{X} — касательный вектор, то при вещественном $\lambda \geq 0$ вектор $\lambda \vec{X}$ тоже является касательным. Если $\vec{X} \neq 0$ — касательный вектор, то λ_n заведомо стремится к $+\infty$. Как и само понятие предела, касательные векторы в точке A к множеству \mathcal{E} зависят лишь от топологии, а не от нормы пространства \tilde{G} . Множество касательных векторов в точке A к множеству \mathcal{E} называется *векторной контингенцией множества \mathcal{E} в точке A* . Множество точек $A + \vec{X}$, где \vec{X} пробегает векторную контингенцию, называется *аффинной контингенцией множества \mathcal{E} в точке A* .

Теорема 8А. Пусть f — отображение открытого множества Ω из E в F , дифференцируемое в точке $a \in \Omega$. Векторной (соответственно аффинной) контингенцией в точке $A = (a, f(a) = b)$ множества \mathcal{E} , определяемого уравнением $y = f(x)$ в $E \times F$, является векторное подпространство в $\tilde{E} \times \tilde{F}$, определяемое уравнением

$$\vec{Y} = f'(a) \vec{X} \quad (\text{III, 3; 19}_2)$$

(соответственно аффинное подпространство из $E \times F$, определяемое уравнением

$$\vec{y} - b = f'(a) \cdot \vec{x} - \vec{a}. \quad (\text{III, 3; 19}_3)$$

Доказательство. Второе утверждение тождественно первому, поэтому достаточно доказать первое. Пусть $x_0, x_1, \dots, x_n, \dots$ — последовательность точек из Ω , сходящаяся к a , и $\lambda_0, \lambda_1, \dots, \lambda_n, \dots$ — последовательность вещественных скаляров ≥ 0 , таких, что $\lambda_n(\overrightarrow{x_n - a})$ стремится к некоторому пределу \vec{X} . По определению (III, 3; 13) производного отображения имеем:

$$\overrightarrow{f(x_n) - b} = f'(a) \cdot \overrightarrow{x_n - a} + \vec{a}_n \parallel \overrightarrow{x_n - a} \parallel, \quad (\text{III, 3; 19}_4)$$

где \vec{a}_n стремится к $\vec{0}$ при n , стремящемся к бесконечности. Отсюда

$$\lambda_n(\overrightarrow{f(x_n) - b}) = f'(a) \cdot \lambda_n(\overrightarrow{x_n - a}) + \vec{a}_n \lambda_n \parallel \overrightarrow{x_n - a} \parallel. \quad (\text{III, 3; 19}_5)$$

Так как отображение $f'(a)$ по условию непрерывно, а $\lambda_n(\overrightarrow{x_n - a})$ сходится к \vec{X} , то при n , стремящемся к бесконечности, первый член правой части последнего равенства стремится

к $f'(a) \cdot \vec{X}$. Поскольку $\|\lambda_n(\overrightarrow{x_n - a})\|$ сходится к $\|\vec{X}\|$, а $\|\vec{\alpha}_n\|$ стремится к 0, то второй член сходится к $\vec{0}$. Таким образом, получаем, что левая часть этого равенства при n , стремящемся к бесконечности, сходится к $\vec{Y} = f'(a) \cdot \vec{X}$. Отсюда вытекает утверждение теоремы. В самом деле,

1°) Пусть вектор (\vec{X}, \vec{Y}) касателен в точке A к множеству \mathcal{E} , и пусть $M_0, M_1, \dots, M_n, \dots$ — последовательность точек $(x_n, f(x_n))$ из \mathcal{E} , сходящихся к $A \in E \times F$, а $\lambda_0, \lambda_1, \dots, \lambda_n, \dots$ — последовательность скаляров ≥ 0 , таких, что $\lambda_n \overrightarrow{AM_n}$ имеет своим пределом (\vec{X}, \vec{Y}) в $E \times F$. Последнее означает, что $\lambda_n(\overrightarrow{x_n - a})$ сходится к \vec{X} , а $\lambda_n(\overrightarrow{f(x_n) - b})$ сходится к \vec{Y} . Мы только что видели, что тогда имеет место равенство (III, 3; 19₂).

2°) Обратно, пусть (\vec{X}, \vec{Y}) — некоторый вектор из $E \times F$, удовлетворяющий условию (III, 3; 19₂). Рассмотрим последовательность точек x_n из Ω , определенных равенством $x_n = a + t_n \vec{X}$, где числа t_n вещественны, положительны и стремятся к 0 при n , стремящемся к бесконечности, и последовательность λ_n , определенную равенством $\lambda_n = 1/t_n$. Тогда $\lambda_n(\overrightarrow{x_n - a}) = \vec{X}$. Следовательно, как мы видели выше, $\lambda_n(\overrightarrow{f(x_n) - b})$ стремится к $f'(a) \cdot \vec{X}$, т. е. к \vec{Y} . Последовательность $M_n = (x_n, f(x_n))$ принадлежит к \mathcal{E} и стремится к A , а $\lambda_n \overrightarrow{AM_n}$ сходится к вектору (\vec{X}, \vec{Y}) , являющемуся касательным вектором к множеству \mathcal{E} в точке A .

Векторное подпространство пространства $E \times F$, определяемое уравнением (III, 3; 19₂) (соответственно аффинное подпространство из $E \times F$, определяемое уравнением (III, 3; 19₃)) называется *векторным подпространством, касательным к многообразию \mathcal{E} в точке A* (соответственно линейным многообразием, касательным к многообразию \mathcal{E} в точке A). Аффинная функция (аффинное отображение из E в F) $x \rightarrow b + f'(a) \cdot (\overrightarrow{x - a})$ называется *аффинной функцией, касательной в точке a к функции f* . Итак, аффинное линейное многообразие, касательное в точке $A = (a, f(a) = b)$ к дифференцируемому многообразию, определяемому уравнением $y = f(x)$ в $E \times F$, получается заменой \vec{dx} на $\overrightarrow{x - a}$ и \vec{dy} на $\overrightarrow{y - b}$ в дифференциале $\vec{dy} = f'(a) \cdot \vec{dx}$.

Так, например, если $E = \mathbb{R}^2$, $F = \mathbb{R}$, то касательная плоскость в точке (a, b, c) к поверхности, определяемой уравнением $z =$

$=f(x, y)$, имеет уравнение

$$z - c = \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b) = p(x - a) + q(y - b). \quad (\text{III}, 3; 19_6)$$

Если $E = \mathbb{R}$, $F = \mathbb{R}^2$, то касательная в точке (a, b, c) к кривой, определяемой уравнением $y = g(x)$, $z = h(x)$, имеет уравнение

$$\begin{aligned} y - b &= g'(a)(x - a), \\ z - c &= h'(a)(x - a). \end{aligned} \quad (\text{III}, 3; 19_7)$$

Градиент вещественной функции в евклидовом пространстве

Пусть E — конечномерное аффинное евклидово пространство над полем вещественных чисел. Пусть f — вещественная функция, определенная в открытом множестве Ω пространства E . Если в некоторой точке x из E функция f имеет производную $\vec{f}'(x)$, то эта производная является линейным отображением \vec{E} в \mathbb{R} , т. е. линейной формой или элементом из \vec{E}' . Согласно сказанному в теореме 24, существует вектор \vec{Y} из \vec{E} , и притом единственный, такой, что

$$\vec{f}'(x) \cdot \vec{X} = (\vec{X} | \vec{Y}) \quad \text{для любого } \vec{X} \text{ из } \vec{E}. \quad (\text{III}, 3; 20)$$

Этот вектор называется градиентом f в точке x и обозначается через $\overrightarrow{\text{grad}} f(x)$. Заменяя \vec{X} на \vec{dx} и $\vec{f}'(x) \cdot \vec{dx}$ на $df \in \mathbb{R}$, согласно определению дифференциала, получаем

$$\begin{aligned} \overrightarrow{D_{\vec{X}} f}(x) &= (\overrightarrow{\text{grad}} f(x) | \vec{X}), \\ df &= (\overrightarrow{\text{grad}} f(x) | \vec{dx}). \end{aligned} \quad (\text{III}, 3; 21)$$

Пусть $(\vec{e}_i)_{i \in I}$ — ортонормированный базис в \vec{E} . Тогда, учитывая выражение скалярного произведения в ортонормированном базисе, формулу (III, 3; 21) можно записать в виде

$$df = \sum_{i \in I} Y_i dx_i, \quad (\text{III}, 3; 22)$$

где Y_i — компоненты градиента. Учитывая (III, 3; 18), имеем:

$$Y_i = (\overrightarrow{\text{grad}} f(x))_i = \frac{\partial f}{\partial x_i}. \quad (\text{III}, 3; 23)$$

Координатами градиента f в ортонормированном базисе являются частные производные $\partial f / \partial x_i$.

Это же утверждение остается справедливым в случае обобщенного евклидова пространства, но только тогда следует воспользоваться формулой (III, 1; 25). При этом (III, 3; 22) заменится на

$$d\vec{f} = \sum_{i \in I} \varepsilon_i Y_i dx_i, \quad (\text{III, 3; 24})$$

а, следовательно, координатами градиента будут $\varepsilon_i \frac{\partial \vec{f}}{\partial x_i}(x)$.

В частности, в специальной теории относительности это будут величины

$$\frac{\partial \vec{f}}{\partial x_1}, \quad \frac{\partial \vec{f}}{\partial x_2}, \quad \frac{\partial \vec{f}}{\partial x_3}, \quad - \frac{\partial \vec{f}}{\partial x_0} = - \frac{1}{c} \frac{\partial \vec{f}}{\partial t}.$$

Если градиент определен на всем Ω , то функция $x \rightarrow \overrightarrow{\text{grad}} f(x)$ называется *векторным полем* над Ω , или отображением Ω в \vec{E} . Это отображение обозначают через $\overrightarrow{\text{grad}} \vec{f}$.

Теорема 8₂ (производная постоянной, производная аффинной функции). *Постоянное отображение дифференцируемо и производная является присоединенным линейным отображением f из E в F дифференцируемо в каждой точке a из E , и его производная является присоединенным линейным отображением: $f'(a) = \vec{f} \in \vec{\mathcal{L}}(\vec{E}; \vec{F})$.*

Производная функция (которую мы определили на стр. 213) является, следовательно, постоянным отображением E в $\vec{\mathcal{L}}(\vec{E}; \vec{F}): x \rightarrow f'(x) = \vec{f}$. Этот результат обобщает тот факт, что производная аффинной функции $y = \alpha x + \beta$ является постоянной.

Доказательство очевидно.

Теорема 8₃ (дифференцирование является линейной операцией). *Если f — отображение из $\Omega \subset E$ в F и g — отображение из Ω в \vec{F} и если f и g имеют производные в точке $a \in \Omega$, то функция $\vec{f} + \vec{g}: x \rightarrow \vec{f}(x) + \vec{g}(x)$ имеет производную в a , равную сумме производных:*

$$(\vec{f} + \vec{g})'(a) = \vec{f}'(a) + \vec{g}'(a). \quad (\text{III, 3; 25})$$

Если λ — постоянный скаляр, то функция $\lambda \vec{g}: x \rightarrow \lambda \vec{g}(x)$ имеет производную в a , равную $\lambda g'(a)$.

Доказательство очевидно.

Случай, когда F является произведением аффинных пространств

Предположим, что f является отображением множества $\Omega \subset E$ в произведение $F_1 \times F_2 \times \dots \times F_m$ аффинных нормированных пространств. Оно определяется заданием отображений f_i пространства E в F_i , $i = 1, 2, \dots, m$ (см. стр. 13). Известно, что отображение f непрерывно тогда и только тогда, когда непрерывны функции f_i (теорема 17 гл. II и следующие за ней). Отображение f аффинно тогда и только тогда, когда аффинны f_i .

Теорема 84. *Пусть E и F_i , $i=1, 2, \dots, m$, — аффинные нормированные пространства. Пусть f — отображение открытого множества $\Omega \subset E$ в $F = F_1 \times F_2 \times \dots \times F_m$, определенное отображениями f_i множества Ω в F_i . Для того чтобы f было дифференцируемым в точке $a \in \Omega$, необходимо и достаточно, чтобы f_i были дифференцируемы в a , и тогда $f'(a)$ является линейным непрерывным отображением пространства \vec{E} в $\vec{F} = \vec{F}_1 \times \vec{F}_2 \times \dots \times \vec{F}_m$, определенным линейными отображениями $f'_i(a)$ из \vec{E} в \vec{F}_i .*

Другими словами, компоненты производной являются производными компонент, и при этом

$$f'(a) = (f'_1(a), f'_2(a), \dots, f'_m(a)), \quad \vec{df} = (\vec{df}_1, \vec{df}_2, \dots, \vec{df}_m). \quad (\text{III}, 3; 26)$$

Доказательство. Предположим, что отображения f_i дифференцируемы в точке a . Для приращения \vec{dx} аргумента x имеем приращения $\vec{\Delta y}_i$ функций $y_i = f_i(x)$, такие, что

$$\vec{\Delta y}_i = f'_i(a) \vec{dx} + \vec{\alpha}_i \parallel \vec{dx} \parallel, \quad (\text{III}, 3; 27)$$

где $\vec{\alpha}_i$ стремятся к $\vec{0}$ вместе с \vec{dx} . По определению аффинного произведения

$$\begin{aligned} \overrightarrow{\Delta(y_1, y_2, \dots, y_m)} &= (\vec{\Delta y}_1, \vec{\Delta y}_2, \dots, \vec{\Delta y}_m) = \\ &= (f'_1(a) \vec{dx}, f'_2(a) \vec{dx}, \dots, f'_m(a) \vec{dx}) + \\ &\quad + (\vec{\alpha}_1 \parallel \vec{dx} \parallel, \vec{\alpha}_2 \parallel \vec{dx} \parallel, \dots, \vec{\alpha}_m \parallel \vec{dx} \parallel) = \\ &= (f'_1(a), f'_2(a), \dots, f'_m(a)) \vec{dx} + (\vec{\alpha}_1, \vec{\alpha}_2, \dots, \vec{\alpha}_m) \parallel \vec{dx} \parallel, \quad (\text{III}, 3; 28) \end{aligned}$$

где, согласно определению топологического произведения, $(\vec{\alpha}_1, \vec{\alpha}_2, \dots, \vec{\alpha}_m)$ стремится к $\vec{0}$ в $\vec{F}_1 \times \vec{F}_2 \times \dots \times \vec{F}_m$, когда \vec{dx}

стремится к $\vec{0}$. Это говорит о том, что f дифференцируема в a и что $f'(a) = (f'_1(a), f'_2(a), \dots, f'_m(a))$.

Обратное утверждение доказывается аналогично, но в обратном порядке.

Случай, когда E является произведением аффинных пространств. Частные производные отображения

Если E является произведением аффинных пространств $E_1 \times E_2$, то отображение f множества $\Omega \subset E_1 \times E_2$ в F становится функцией двух переменных. Мы будем записывать ее в виде $y = f(x_1, x_2)$. При фиксированном x_1 в точке a_1 можно рассматривать частное отображение $f_{a_1}: x_2 \rightarrow f(x_1, x_2)$ и искать, если оно существует, производное отображение в точке a_2 . Если это возможно, то полученное производное отображение является отображением из E_2 в F . Его называют *частным производным отображения f по x_2 в точке (a_1, a_2)* и обозначают одним из следующих символов: $\partial_2 f(a_1, a_2)$ или $\frac{\partial f}{\partial x_2}(a_1, a_2)$. Точно так же можно рассматривать частное отображение f_{a_2} и соответствующее ему частное производное отображение $\frac{\partial f}{\partial x_1}(a_1, a_2)$ в той же самой точке (a_1, a_2) . Как и ранее (пример (III, 3; 9)), тот факт, что f имеет частные производные в некоторой точке, не влечет за собой с необходимостью того, что оно имеет полное производное отображение в этой точке (и даже того, что оно там непрерывно)¹⁾. Верным будет лишь обратное утверждение. Точнее, имеет место

Теорема 9. *Если E является произведением двух аффинных пространств $E = E_1 \times E_2$ и f имеет производное отображение $f'(a_1, a_2)$ в некоторой точке $a = (a_1, a_2)$ открытого подмножества Ω произведения $E_1 \times E_2$, то оно имеет в этой точке частные производные отображения и его полное производное отображение задается формулой:*

$$\frac{df}{dx}(a_1, a_2)(\vec{X}_1, \vec{X}_2) = \frac{\partial f}{\partial x_1}(a_1, a_2)\vec{X}_1 + \frac{\partial f}{\partial x_2}(a_1, a_2)\vec{X}_2. \quad (\text{III}, 3; 29)$$

Доказательство. Из теоремы 51 гл. II мы видим, что линейное непрерывное отображение $\frac{df}{dx}(a_1, a_2)$ из $\vec{E}_1 \times \vec{E}_2$ в \vec{F} определяет линейные непрерывные отображения L_1 и L_2 про-

1) Однако из теоремы 15 будет видно, что существование непрерывных частных производных влечет за собой существование полной производной функции, которая в этом случае будет непрерывной,

пространств \vec{E}_1 и \vec{E}_2 в \vec{F} и что оно задается формулой

$$\frac{df}{dx}(a_1, a_2)(\vec{X}_1, \vec{X}_2) = L_1 \vec{X}_1 + L_2 \vec{X}_2. \quad (\text{III}, 3; 30)$$

Если теперь переменной придать приращение $\vec{h} = (\vec{0}, \vec{h}_2)$, то, согласно определению производного отображения, приращение выразится формулой

$$\overrightarrow{f(a_1, a_2 + \vec{h}_2)} - \overrightarrow{f(a_1, a_2)} = L_2 \cdot \vec{h}_2 + \overrightarrow{\varphi(\vec{0}, \vec{h}_2)} \parallel \vec{h}_2 \parallel, \quad (\text{III}, 3; 31)$$

где $\overrightarrow{\varphi(\vec{0}, \vec{h}_2)}$ стремится к $\vec{0}$, когда \vec{h}_2 стремится к $\vec{0}$. Но это как раз означает, что L_2 является частным производным отображением отображения f по второй переменной в точке (a_1, a_2) .

Мы предоставляем читателю перенести результат на произведение $E_1 \times E_2 \times \dots \times E_n$ аффинных пространств. В дифференциальных обозначениях получим, что

$$\overrightarrow{df} = f'(x) \overrightarrow{dx} = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) \overrightarrow{dx}_j, \quad (\text{III}, 3; 32)$$

где на этот раз $\frac{\partial f}{\partial x_j}(x) \in \mathcal{L}(\vec{E}_j; \vec{F})$, $\overrightarrow{dx}_j \in \vec{E}_j$ и $\overrightarrow{dx} = (\overrightarrow{dx}_1, \overrightarrow{dx}_2, \dots, \overrightarrow{dx}_n) \in \vec{E} = \vec{E}_1 \times \vec{E}_2 \times \dots \times \vec{E}_n$.

Если множители \vec{E}_j совпадают с полем скаляров, то мы возвращаемся к формуле (III, 3; 18) (с отождествлением производного вектора и производного отображения, указанным в замечании 2°) на стр. 215).

Производная билинейного непрерывного отображения

Классическая формула $d(xy) = y dx + x dy$ для дифференциала произведения обобщается следующим образом:

Теорема 9₂. Пусть B — билинейное непрерывное отображение $\vec{E}_1 \times \vec{E}_2$ в \vec{F} . Тогда B дифференцируемо в каждой точке (\vec{a}_1, \vec{a}_2) из $\vec{E}_1 \times \vec{E}_2$. Его частная производная $\frac{\partial B}{\partial x_1}(\vec{a}_1, \vec{a}_2)$ является частным отображением $B_{\vec{a}_2} \in \mathcal{L}(\vec{E}_1; \vec{F})$, определенным формулой (см. стр. 122)

$$B_{\vec{a}_2}(\vec{X}_1) = B(\vec{X}_1, \vec{a}_2) \in \vec{F}. \quad (\text{III}, 3; 33)$$

Точно так же, его частная производная $\frac{\partial B}{\partial x_2}(\vec{a}_1, \vec{a}_2)$ является частным отображением $B_{\vec{a}_1} \in \mathcal{L}(\vec{E}_2; \vec{F})$, определенным формулой

$$B_{\vec{a}_1}(\vec{X}_2) = B(\vec{a}_1, \vec{X}_2) \in \vec{F}. \quad (\text{III}, 3; 34)$$

Его полная производная равна

$$B'(\vec{a}_1, \vec{a}_2) \cdot (\vec{X}_1, \vec{X}_2) = B(\vec{X}_1, \vec{a}_2) + B(\vec{a}_1, \vec{X}_2), \quad (\text{III}, 3; 35)$$

или в дифференциальных обозначениях:

$$dB = B(d\vec{x}_1, \vec{x}_2) + B(\vec{x}_1, d\vec{x}_2). \quad (\text{III}, 3; 36)$$

Доказательство. Результаты, относящиеся к частным производным, очевидны: частное отображение $B_{\vec{a}_2}: \vec{x}_1 \rightarrow \vec{B}(\vec{x}_1, \vec{a}_2)$ является линейным и непрерывным, а, следовательно, его производная в \vec{a}_1 совпадает с самим отображением (теорема 8₂), т. е. с отображением $\vec{X}_1 \rightarrow \vec{B}(\vec{X}_1, \vec{a}_2)$. Если отображение B дифференцируемо, то формула (III, 3; 35) вытекает из формулы (III, 3; 30). Однако частная дифференцируемость отображения B не влечет за собой его полной дифференцируемости и, следовательно, ее надо доказывать¹⁾.

Дадим \vec{x}_1, \vec{x}_2 приращения $d\vec{x}_1, d\vec{x}_2$, и пусть $\vec{\Delta B}$ является приращением отображения B . Тогда

$$\begin{aligned} \vec{\Delta B} &= B(\vec{a}_1 + d\vec{x}_1, \vec{a}_2 + d\vec{x}_2) - B(\vec{a}_1, \vec{a}_2) = \\ &= B(\vec{a}_1 + d\vec{x}_1, d\vec{x}_2) + B(d\vec{x}_1, \vec{a}_2). \end{aligned} \quad (\text{III}, 3; 37)$$

Теперь имеем

$$\vec{\Delta B} = B(d\vec{x}_1, \vec{a}_2) + B(\vec{a}_1, d\vec{x}_2) = B(d\vec{x}_1, d\vec{x}_2). \quad (\text{III}, 3; 38)$$

Правая часть оценивается по норме величиной

$$\|B\| \|d\vec{x}_1\| \|d\vec{x}_2\| \leq \|B\| (\|d\vec{x}_1\| + \|d\vec{x}_2\|)^2, \quad (\text{III}, 3; 39)$$

бесконечно малой более высокого порядка, чем $\|d\vec{x}_1\| + \|d\vec{x}_2\|$, когда $(d\vec{x}_1, d\vec{x}_2)$ стремится к $\vec{0}$ в $\vec{E}_1 \times \vec{E}_2$. Если в $\vec{E}_1 \times \vec{E}_2$ взять норму $\|\cdot\|_1 + \|\cdot\|_2$, то, согласно определению производного отображения (III, 3; 13), производная B в (\vec{a}_1, \vec{a}_2) существует

¹⁾ Теорема 15 позволила бы нам от этого избавиться, ибо частные производные непрерывны.

и задается формулой (III, 3; 35), чем и заканчивается доказательство теоремы.

Оставим читателю заботу о распространении этого результата на произведение n нормированных векторных пространств. На этот раз

$$\begin{aligned} \vec{dB} = & B(\vec{dx}_1, \vec{x}_2, \dots, \vec{x}_n) + B(\vec{x}_1, \vec{dx}_2, \dots, \vec{x}_n) + \dots \\ & \dots + B(\vec{x}_1, \vec{x}_2, \dots, \vec{dx}_n). \quad (\text{III}, 3; 40) \end{aligned}$$

Замечание. Отображение B' , ставящее вектору $\vec{x} = (\vec{x}_1, \vec{x}_2) \in \vec{E}_1 \times \vec{E}_2$ в соответствие производную $B'(\vec{x}) = B'(\vec{x}_1, \vec{x}_2) \in \mathcal{L}(\vec{E}_1 \times \vec{E}_2; \vec{F})$, определенную по формуле (III, 3; 35), является линейным и непрерывным отображением $\vec{E}_1 \times \vec{E}_2$ в $\mathcal{L}(\vec{E}_1 \times \vec{E}_2; \vec{F})$. Оно очевидным образом линейно, и имеет место оценка

$$\begin{aligned} \|B'(\vec{x}) \cdot \vec{X}\| = & \|B'(\vec{x}_1, \vec{x}_2) \cdot (\vec{X}_1, \vec{X}_2)\| \leq \|B(\vec{X}_1, \vec{x}_2)\| + \\ & + \|B(\vec{x}_1, \vec{X}_2)\| \leq 2\|B\|\|\vec{x}\|\|\vec{X}\|, \quad (\text{III}, 3; 40_2) \end{aligned}$$

а, следовательно,

$$\|B'(\vec{x})\| = \sup_{\|\vec{X}\|=1} \|B'(\vec{x}) \cdot \vec{X}\| \leq 2\|B\|\|\vec{x}\|, \quad (\text{III}, 3; 40_3)$$

откуда и вытекает, что B' непрерывно; кроме того, $B' \in \mathcal{L}(\vec{E}_1 \times \vec{E}_2; \mathcal{L}(\vec{E}_1 \times \vec{E}_2; \vec{F}))$ и

$$\|B'\| \leq 2\|B\|. \quad (\text{III}, 3; 40_4)$$

Дифференцируемые функции.

Непрерывно дифференцируемые функции

Если f имеет производную $f'(x)$ в каждой точке из Ω , то она называется *дифференцируемой* в Ω . Отображение $x \rightarrow f'(x)$ является ее производной функцией. Оно обозначается через f' , или df/dx , или Df и является отображением Ω в $\mathcal{L}(\vec{E}; \vec{F})$. Функция f называется *непрерывно дифференцируемой*, или принадлежащей классу C^1 , если производное отображение f' множества Ω в $\mathcal{L}(\vec{E}; \vec{F})$ непрерывно.

Теорема 10. *Если f является непрерывно дифференцируемым отображением Ω в F , то отображение $(x, \vec{X}) \rightarrow f'(x) \cdot \vec{X}$ из $\Omega \times \vec{E}$ в \vec{F} непрерывно. Обратно, если это отображение непрерывно*

1) Если считать $\|\vec{x}\| = \|\vec{x}_1\|_1 + \|\vec{x}_2\|_2$ и $\|\vec{X}\| = \|\vec{X}_1\|_1 + \|\vec{X}_2\|_2$, то коэффициент 2 в формулах (III, 3; 40₂)—(III, 3; 40₄) не появляется. — Прим. ред.

и если E конечномерно, то отображение f непрерывно дифференцируемо.

Доказательство. Предположим, что f' является непрерывным отображением Ω в $\mathcal{L}(\vec{E}; \vec{F})$. Пусть x_0, x_1, x_2, \dots — последовательность, сходящаяся к x в Ω , и $\vec{X}_0, \vec{X}_1, \vec{X}_2, \dots$ — последовательность, сходящаяся к \vec{X} в \vec{E} . Тогда, в силу непрерывности f' , последовательность $f'(x_n)$ сходится к $f'(x)$ в $\mathcal{L}(\vec{E}; \vec{F})$. Далее, в силу теоремы 17 гл. II, последовательность $(f'(x_n), \vec{X}_n)$ сходится к $(f'(x), \vec{X})$ в $\mathcal{L}(\vec{E}; \vec{F}) \times \vec{E}$. Поскольку каноническое отображение $(u, \vec{X}) \rightarrow u \cdot \vec{X}$ пространства $\mathcal{L}(\vec{E}; \vec{F}) \times \vec{E}$ в пространство \vec{F} непрерывно (теорема 54 гл. II), то последовательность $f'(x_n) \cdot \vec{X}_n$ сходится к $f'(x) \cdot \vec{X}$ в \vec{F} , что и доказывает непрерывность отображения $(x, \vec{X}) \rightarrow f'(x) \cdot \vec{X}$ множества $\Omega \times \vec{E}$ в \vec{F} .

Обратно, предположим, что это отображение непрерывно. Выберем какой-либо базис $(\vec{e}_j)_{j \in J}$ в \vec{E} . Высказанное предположение тем более влечет за собой непрерывность каждой частной производной $x \rightarrow f'(x) \cdot \vec{e}_j = \frac{\partial f}{\partial x_j}(x)$. Позднее мы увидим (теорема 15), что даже без предположения дифференцируемости f существование и непрерывность частных производных $\frac{\partial f}{\partial x_j}$ влечет за собой существование и непрерывность полной производной. Этим будет доказано обратное утверждение (легко убедиться, что следующие далее теоремы, вплоть до теоремы 15, не опираются на теорему 10).

Примеры непрерывно дифференцируемых функций

Поскольку производная аффинной функции постоянна (теорема 8₂), такая функция непрерывно дифференцируема. Билинейная непрерывная функция также непрерывно дифференцируема, поскольку ее производная является непрерывной линейной функцией (см. замечание, следующее за теоремой 9₂).

Отображение множества Ω в произведение $F_1 \times F_2 \times \dots \times F_m$ непрерывно дифференцируемо тогда и только тогда, когда каждая из его компонент непрерывно дифференцируема (теорема 8₄).

В дальнейшем нам встретится много других примеров таких функций (теорема 15).

Пространства дифференцируемых функций

В гл. II на стр. 146 мы видели, что если E является произвольным множеством, а \vec{F} — векторным пространством, то множество \vec{F}^E всевозможных отображений E в \vec{F} является векторным пространством. Если же F является аффинным пространством, то F^E также является аффинным пространством с присоединенным векторным пространством \vec{F}^E . Если f и g — два отображения E в F , то формулой $\overrightarrow{g(x) - f(x)}$ определяется отображение $\overrightarrow{g - f}$ множества E в \vec{F}^1). Если F является нормированным аффинным пространством, то пространство $(F^E)_b$ ограниченных отображений E в F является нормированным аффинным пространством с присоединенным векторным пространством $(\vec{F}^E)_b$. Это пространство полно, если полно F (теорема 64 гл. II). То же самое имеет место для $(F^E)_{cb}$ и $(\vec{F}^E)_{cb}$, если E является топологическим пространством.

Пусть, далее, E , F — аффинные нормированные пространства и Ω — открытое множество в E . Договоримся через $(F^\Omega)_{b;1}$ (и соответственно $(F^\Omega)_{cb;1}$) обозначать пространство функций, заданных в Ω , со значениями в F , дифференцируемых и ограниченных вместе со своими производными (соответственно непрерывно дифференцируемых, ограниченных вместе со своими производными)¹⁾. Это аффинное пространство с присоединенным векторным пространством $(\vec{F}^\Omega)_{b;1}$ (соответственно $(\vec{F}^\Omega)_{cb;1}$).

Образуем теперь нормированное аффинное пространство, вводя в присоединенном векторном пространстве норму

$$\|\overrightarrow{f}\|_1 = \max_{x \in \Omega} (\|\vec{f}(x)\|, \|f'(x)\|) = \max (\|\vec{f}\|, \|\vec{f}'\|) \quad (\text{III}, 3; 41)$$

(где $\|\vec{f}(x)\|$ — норма в F , а $\|f'(x)\|$ — норма в $\mathcal{L}(\vec{E}; \vec{F})$).

Позже мы увидим (теорема 113 гл. IV), что эти пространства полны, если полно F . Сказать, что последовательность дифференцируемых функций f_n сходится при n , стремящемся к бесконечности, к некоторой дифференцируемой функции f в смысле $(F^\Omega)_{b;1}$, означает сказать, что функции f_n сходятся равномерно к f , а функции f'_n сходятся равномерно к f' . Эта сходимость более сильная, чем сходимость в $(F^\Omega)_b$. Если E является полем

¹⁾ Если E является двухэлементным множеством $\{1, 2\}$, то можно убедиться в том, что $F^2 = F \times F$ является аффинным пространством с присоединенным векторным пространством $\vec{F}^2 = \vec{F} \times \vec{F}$.

²⁾ Причина, по которой приписывается индекс 1, выяснится на стр. 267. Речь идет о пространстве функций, один раз дифференцируемых.

вещественных чисел \mathbb{R} , F — полем \mathbb{R} или полем \mathbb{C} , а Ω — компактным интервалом $[a, b]^1)$ (соответственно всей вещественной прямой \mathbb{R}), то пространство $(F^\Omega)_{cb;1}$ является векторным пространством непрерывно дифференцируемых и ограниченных вместе со своей производной функций на $[a, b]$ (соответственно на \mathbb{R}) с вещественными или комплексными значениями. Это пространство очень важно для приложений.

Теорема 10₂. *Отображение, ставящее в соответствие каждой функции f ее производную f' , является линейным и непрерывным отображением с нормой $\leqslant 1$ пространства $(\vec{F}^\Omega)_{cb;1}$ (соответственно $(\vec{F}^\Omega)_{cb;1}$) в пространство $((\mathcal{L}(\vec{E}; \vec{F}))^\Omega)_b$ (соответственно в $((\mathcal{L}(\vec{E}; \vec{F}))^\Omega)_{cb}$).*

Доказательство очевидно. (Напомним, что если f является функцией на Ω со значениями в \vec{F} , то f' является функцией на Ω со значениями в $\mathcal{L}(\vec{E}; \vec{F})$ и, кроме того, $\|f'\| \leq \|f\|_1$.)

§ 4. ТЕОРЕМА О СЛОЖНОЙ ФУНКЦИИ

Теорема 11. *Пусть E, F, G — нормированные аффинные пространства, Ω — открытое множество из E , Ω' — открытое множество из F . Пусть f — отображение Ω в Ω' , а g — отображение Ω' в G . Если отображение f имеет в точке $a \in \Omega$ производную $f'(a) \in \mathcal{L}(\vec{E}; \vec{F})$, а отображение g имеет в точке $b = f(a)$ из Ω' производную $g'(b) \in \mathcal{L}(\vec{F}; \vec{G})$, то композиция отображений $h = g \circ f$ имеет производную в точке a и эта производная является композицией производных отображений*

$$h'(a) = g'(b) \circ f'(a) = g'(f(a)) \circ f'(a). \quad (\text{III}, 4; 1)$$

Доказательство. Перед доказательством заметим, что если положить $E=F=G=\mathbb{K}$, то производные отображения будут умножением на числа, являющиеся обычными производными, а написанная формула является формулой дифференцирования сложной функции в том виде, в каком она обычно пишется: $(g(f(x)))' = g'(f(x))f'(x)$. Удобство принятых нами общих обозначений заключается в том, что для аффинных пространств произвольного числа измерений, конечно- или бесконечномерных, они дают нам тот же самый формальный аппарат, что и для вещественной функции вещественной переменной.

Выберем приращение \vec{dx} таким образом, чтобы точка $a + \vec{dx}$ принадлежала Ω . Соответствующие приращения переменных y

¹⁾ Интервал $[a, b]$ не открыт; см. стр. 198, начало § 2.

и z будем обозначать через $\vec{\Delta}y$ и $\vec{\Delta}z$. Тогда можно написать формулы

$$\vec{\Delta}y = \overrightarrow{f(a + \vec{dx}) - f(a)} = f'(a) \cdot \vec{dx} + \vec{\alpha} \parallel \vec{dx} \parallel, \quad (\text{III}, 4; 2)$$

где $\parallel \vec{\alpha} \parallel$ стремится к 0 при \vec{dx} , стремящемся к $\vec{0}$, и

$$\vec{\Delta}z = \overrightarrow{g(b + \vec{\Delta}y) - g(b)} = g'(b) \cdot \vec{\Delta}y + \vec{\beta} \parallel \vec{\Delta}y \parallel, \quad (\text{III}, 4; 3)$$

где $\parallel \vec{\beta} \parallel$ стремится к 0, когда $\vec{\Delta}y$ стремится к $\vec{0}$. Из написанных равенств получаем:

$$\vec{\Delta}z = g'(b) \cdot f'(a) \cdot \vec{dx} + g'(b) \cdot \vec{\alpha} \parallel \vec{dx} \parallel + \vec{\beta} \parallel \vec{\Delta}y \parallel. \quad (\text{III}, 4; 4)$$

Разность

$$\vec{\Delta}z - g'(b) \cdot f'(a) \cdot \vec{dx} = g'(b) \cdot \vec{\alpha} \parallel \vec{dx} \parallel + \vec{\beta} \parallel \vec{\Delta}y \parallel \quad (\text{III}, 4; 5)$$

по норме не превосходит

$$\begin{aligned} \|g'(b)\| \|\vec{\alpha}\| \|\vec{dx}\| + \|\vec{\beta}\| (\|f'(a)\| \|\vec{dx}\| + \|\vec{\alpha}\| \|\vec{dx}\|) = \\ = \|\vec{dx}\| (\|g'(b)\| \|\vec{\alpha}\| + \|f'(a)\| \|\vec{\beta}\| + \|\vec{\alpha}\| \|\vec{\beta}\|). \end{aligned} \quad (\text{III}, 4; 6)$$

Так как отображение f дифференцируемо в точке a , то оно непрерывно в a . Поэтому если \vec{dx} стремится к $\vec{0}$, то вместе с ним и $\vec{\Delta}y$ стремится к $\vec{0}$, а, значит, $\vec{\alpha}$ и $\vec{\beta}$ стремятся к $\vec{0}$, и мы заключаем, что последняя скобка стремится к 0. Поскольку отображение $g'(b) \circ f'(a)$ непрерывно (как композиция двух непрерывных отображений), полученный результат означает, что отображение $h = g \circ f$ дифференцируемо в a и его производная равна $g'(b) \circ f'(a)$.

В дифференциальных обозначениях предыдущий результат выражается следующим образом. Дифференциал f задается выражением

$$\vec{dy} = f'(x) \cdot \vec{dx}. \quad (\text{III}, 4; 7)$$

В свою очередь дифференциал g задается формулой

$$\vec{dz} = g'(y) \cdot \vec{dy}. \quad (\text{III}, 4; 8)$$

Теперь дифференциал сложной функции получается из дифференциала функции g , если заменить y на $f(x)$, а \vec{dy} на дифференциал $f'(x) \cdot \vec{dx}$. Таким образом, непосредственно получаем:

$$\vec{dz} = h'(x) \cdot \vec{dx} = g'(f(x)) \cdot (f'(x) \cdot \vec{dx}). \quad (\text{III}, 4; 9)$$

Следствие 1 (перестановочность дифференцирования и непрерывного линейного отображения). *Если \vec{f} является отображением $\Omega \subset E$ в F , имеющим производную $\vec{f}'(a)$ в точке $a \in \Omega$, и если g — непрерывное аффинное отображение F в G , то сложное отображение $g \circ \vec{f}$ из Ω в G имеет производную в a , определенную формулой:*

$$(g \circ \vec{f})'(a) = \vec{g} \circ \vec{f}'(a). \quad (\text{III}, 4; 10)$$

Это утверждение вытекает из теорем 11 и 8₂ (впрочем, оно очевидно).

Предположим, в частности, что E является аффинным пространством над полем вещественных чисел, а \vec{F} — векторным пространством над полем комплексных чисел. Во всей этой теории следует оба их рассматривать как аффинные пространства над вещественным полем. Тем не менее, умножение на $\lambda \in \mathbb{C}$ сохраняет смысл в \vec{F} и его можно рассматривать как линейное непрерывное отображение \vec{F} в себя. Поэтому, если \vec{f} есть некоторое дифференцируемое отображение Ω в \vec{F} , то же самое можно сказать и о $\lambda \vec{f}$ и при этом $(\lambda \vec{f})' = \lambda(\vec{f}')$. Например, из того, что в \mathbb{R} функция $\cos x$ является производной функции $\sin x$, вытекает, что производной функции $i \sin x$ является $i \cos x$.

Следствие 2. *Пусть E , F , G — конечномерные аффинные пространства с выбранными в них системами координат. Тогда матрица Якоби отображения $h = g \circ \vec{f}$ в точке a равна произведению матрицы Якоби отображения g в точке $b = \vec{f}(a)$ на матрицу Якоби отображения \vec{f} в точке a .*

Это утверждение вытекает непосредственно из того факта, что матрица композиции двух линейных отображений равна произведению матриц. Мы это запишем в следующей форме. Пусть точка в E (соответственно в F , в G) определяется своими координатами $(x_k)_{k \in K}$, $(y_j)_{j \in J}$, $(z_i)_{i \in I}$. Для простоты возьмем $I = \{1, 2, \dots, l\}$, $J = \{1, 2, \dots, m\}$, $K = \{1, 2, \dots, n\}$. Тогда \vec{f} определяется с помощью m функций от n переменных: $y_j = F_j(x_1, x_2, \dots, x_n)$, $j = 1, 2, \dots, m$, g определяется с помощью l функций от m переменных: $z_i = G_i(y_1, y_2, \dots, y_m)$, $i = 1, 2, \dots, l$, а $h = g \circ \vec{f}$ определяется с помощью l функций от n переменных: $z_i = H_i(x_1, x_2, \dots, x_n)$, $i = 1, 2, \dots, l$. Кроме того, можно написать

$$\begin{aligned} H_i(x_1, x_2, \dots, x_n) &= \\ &= G_i(F_1(x_1, x_2, \dots, x_n), F_2(x_1, x_2, \dots, x_n), \dots, F_m(x_1, x_2, \dots, x_n)). \end{aligned} \quad (\text{III}, 4; 11)$$

Тогда имеем:

$$\begin{aligned}
 & \left| \begin{array}{cccc} \frac{\partial H_1}{\partial x_1}(a) & \frac{\partial H_1}{\partial x_2}(a) & \dots & \frac{\partial H_1}{\partial x_n}(a) \\ \frac{\partial H_2}{\partial x_1}(a) & \frac{\partial H_2}{\partial x_2}(a) & \dots & \frac{\partial H_2}{\partial x_n}(a) \\ \dots & \dots & \dots & \dots \\ \frac{\partial H_l}{\partial x_1}(a) & \frac{\partial H_l}{\partial x_2}(a) & \dots & \frac{\partial H_l}{\partial x_n}(a) \end{array} \right| = \\
 & = \left| \begin{array}{cccc} \frac{\partial G_1}{\partial y_1}(b) & \frac{\partial G_1}{\partial y_2}(b) & \dots & \frac{\partial G_1}{\partial y_m}(b) \\ \frac{\partial G_2}{\partial y_1}(b) & \frac{\partial G_2}{\partial y_2}(b) & \dots & \frac{\partial G_2}{\partial y_m}(b) \\ \dots & \dots & \dots & \dots \\ \frac{\partial G_l}{\partial y_1}(b) & \frac{\partial G_l}{\partial y_2}(b) & \dots & \frac{\partial G_l}{\partial y_m}(b) \end{array} \right| \times \\
 & \quad \times \left| \begin{array}{cccc} \frac{\partial F_1}{\partial x_1}(a) & \frac{\partial F_1}{\partial x_2}(a) & \dots & \frac{\partial F_1}{\partial x_n}(a) \\ \frac{\partial F_2}{\partial x_1}(a) & \frac{\partial F_2}{\partial x_2}(a) & \dots & \frac{\partial F_2}{\partial x_n}(a) \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_m}{\partial x_1}(a) & \frac{\partial F_m}{\partial x_2}(a) & \dots & \frac{\partial F_m}{\partial x_n}(a) \end{array} \right|, \quad (\text{III}, 4; 12)
 \end{aligned}$$

откуда

$$\frac{\partial H_i}{\partial x_k}(a) = \sum_{j \in J} \frac{\partial G_i}{\partial y_j}(b) \frac{\partial F_j}{\partial x_k}(a), \quad i \in I, \quad k \in K. \quad (\text{III}, 4; 13)$$

Точно так же имеем

$$\frac{\overrightarrow{\partial h}}{\partial x_k}(a) = \sum_{i \in J} \frac{\overrightarrow{\partial g}}{\partial y_i}(b) \frac{\partial F_i}{\partial x_k}(a), \quad k \in K. \quad (\text{III}, 4; 13_2)$$

Последнюю формулу можно записать короче, введя не совсем корректные, но очень удобные обозначения. Отождествим прежде всего h и g , считая, что речь идет об одной функции g , выражающейся либо через переменную x , либо через переменную y . Затем будем писать частные производные, не уточняя, в какой точке они вычислены. Конечно, при этом предполагается, что производная f вычисляется в точке a , а производная g вычисляется в точке $f(a) = b$. Далее, если не указывать функцию f , то вместо ее частных производных $\partial f_j / \partial x_k$ можно писать

частные производные $\frac{\partial y_j}{\partial x_k}$. Тогда рассматриваемая формула примет широко применяемую форму:

$$\frac{\partial G_i}{\partial x_k} = \sum_{j \in J} \frac{\partial G_i}{\partial y_j} \frac{\partial y_j}{\partial x_k}, \quad i \in J, k \in K. \quad (\text{III, 4; 14})$$

Можно даже позволить себе заменить g на текущую переменную z пространства H или полностью удалить g . При этом получим:

$$\frac{\partial z_i}{\partial x_k} = \sum_{j \in J} \frac{\partial z_i}{\partial y_j} \frac{\partial y_j}{\partial x_k}, \quad i \in J, k \in K, \quad (\text{III, 4; 15})$$

или

$$\frac{\partial}{\partial x_k} = \sum_{j \in J} \frac{\partial y_j}{\partial x_k} \frac{\partial}{\partial y_j}, \quad k \in K. \quad (\text{III, 4; 16})$$

Следствие 3. Если в условиях следствия 2 пространства E, F, G имеют одну и ту же размерность n , то якобиан сложного отображения h в точке a равен произведению якобиана отображения f в точке a на якобиан отображения g в точке $f(a) = b$, так что имеет место формула

$$\frac{D(z_1, z_2, \dots, z_n)}{D(x_1, x_2, \dots, x_n)} = \frac{D(z_1, z_2, \dots, z_n)}{D(y_1, y_2, \dots, y_n)} \frac{D(y_1, y_2, \dots, y_n)}{D(x_1, x_2, \dots, x_n)}. \quad (\text{III, 4; 16}_2)$$

По существу, именно эта формула оправдывает применение обозначения $\frac{D(y_1, y_2, \dots, y_n)}{D(x_1, x_2, \dots, x_n)}$ для якобиана функции $y = f(x)$ в выбранной системе координат.

Следствие 4. Пусть Ω — открытое множество пространства E , Ω' — открытое множество пространства F , а f — некоторая биекция Ω на Ω' , дифференцируемая в каждой точке вместе со своей обратной биекцией f^{-1} . Тогда производное отображение $f'(a)$ является биекцией \tilde{E} на \tilde{F} , а его обратная биекция является производной в точке $b = f(a)$ обратной биекции f^{-1} . Другими словами, имеет место формула

$$(f'(a))^{-1} = (f^{-1})'(b). \quad (\text{III, 4; 17})$$

В частности, если E и F конечномерны, то их размерности одинаковы¹⁾.

¹⁾ Из линейной алгебры известно, что если существует линейная биекция конечномерного векторного пространства на другое пространство, то они имеют одинаковую размерность. Мы здесь получили нечто более сильное: если существует биекция некоторого открытого множества конечномерного аффинного пространства E на открытое множество конечномерного аффинного пространства F , дифференцируемая вместе со своей обратной биекцией, то E

В самом деле, применим теорему о сложной функции к отображениям f и f^{-1} , удовлетворяющим равенствам: $f \circ f^{-1} = I_F$, $f^{-1} \circ f = I_E$. Поскольку производное отображение тождественного отображения является тождественным отображением, получаем: $f'(a) \circ (f^{-1})'(b) = I_{\vec{F}}$ и $(f^{-1})'(b) \circ f'(a) = I_{\vec{E}}$, чем и доказывается тот факт, что $f'(a)$ является биекцией и что ее обратной биекцией будет $(f^{-1})'(b)$. В частности, если E и F конечномерны, то их размерности одинаковы, и если в них выбраны системы координат, то матрица Якоби обратной биекции f^{-1} в точке $b = f(a)$ является обратной матрицей матрицы Якоби отображения f в точке a . Якобиан обратной биекции f^{-1} в точке b обратен якобиану f в точке a . Последнее записывается в сжатой форме:

$$\frac{D(y_1, y_2, \dots, y_n)}{D(x_1, x_2, \dots, x_n)} = \left(\frac{D(x_1, x_2, \dots, x_n)}{D(y_1, y_2, \dots, y_n)} \right)^{-1}. \quad (\text{III}, 4; 18)$$

Следствие 5. Композиция двух непрерывно дифференцируемых отображений является непрерывно дифференцируемым отображением.

Надо доказать, что отображение $x \rightarrow h'(x) = g'(f(x)) \circ f'(x)$ является непрерывным отображением $\Omega \subset E$ в $\mathcal{L}(\vec{E}; \vec{G})$. Отображение $x \rightarrow f'(x)$ является, по условию, непрерывной функцией f' . Отображение $x \rightarrow g'(f(x))$ есть функция $g' \circ f$. Здесь g' непрерывна по предположению, а f непрерывна в силу дифференцируемости. Следовательно, согласно теореме о сложной функции, функция $g' \circ f$ непрерывна (теорема 10 гл. II). Отсюда следует, что $x \rightarrow (f'(x), g'(f(x)))$ является непрерывным отображением Ω в произведение $(\mathcal{L}(\vec{E}; \vec{F}) \times \mathcal{L}(\vec{F}; \vec{G}))$ (замечание, следующее за теоремой 17 гл. II). Но отображение $x \rightarrow h'(x)$ является композицией рассмотренного отображения и билинейного непрерывного отображения $(u, v) \rightarrow v \circ u$ пространства $\mathcal{L}(\vec{E}; \vec{F}) \times \mathcal{L}(\vec{F}; \vec{G})$ в пространство $\mathcal{L}(\vec{E}; \vec{G})$. Из той же самой теоремы 10 гл. II следует, что это отображение непрерывно.

Примеры вычисления обычных производных

Пример 1. Пусть u — отображение открытого множества Ω нормированного аффинного пространства E на прямую \mathbb{R} . Применим после него отображение $x \rightarrow \arg \operatorname{tg} x$ из \mathbb{R} в \mathbb{R} . Тем самым мы определим новое отображение Ω в \mathbb{R} , которое

и F имеют одинаковую размерность. С помощью методов, напоминающих примененные на стр. 98, можно доказать большее: если существует гомеоморфизм некоторого открытого множества E на открытое множество F , то E и F имеют одну и ту же размерность. Здесь не требуется никакого предположения о дифференцируемости.

обозначается просто через $\operatorname{arc} \operatorname{tg} u$. Если отображение u предполагать дифференцируемым, то, учитывая, что функция $\operatorname{arc} \operatorname{tg} x$ дифференцируема и что ее производная равна $\frac{1}{1+x^2}$ (напомним, что для вещественных функций вещественной переменной производное отображение является умножением на скаляр — обычную производную; замечание 2, стр. 215), согласно теореме о сложной функции получаем, что отображение $\operatorname{arc} \operatorname{tg} u$ имеет производное отображение, определяемое формулой

$$(\operatorname{arc} \operatorname{tg} u)'(a) = \frac{u'(a)}{1+(u(a))^2}, \text{ или } (\operatorname{arc} \operatorname{tg} u)' = \frac{u'}{1+u^2}. \quad (\text{III}, 4; 19)$$

или в дифференциальных обозначениях:

$$d(\operatorname{arc} \operatorname{tg} u) = \frac{du}{1+u^2}. \quad (\text{III}, 4; 20)$$

Точно такие же рассуждения можно провести, если вместо $\operatorname{arc} \operatorname{tg}$ рассматриваются другие функции. Взяв, например, \ln , получаем (считая $u(x) > 0$ для $x \in \Omega$) следующие формулы¹⁾:

$$(\ln u)'(a) = \frac{u'(a)}{u(a)}, \text{ или } (\ln u)' = \frac{u'}{u}, \text{ или } d(\ln u) = \frac{du}{u}. \quad (\text{III}, 4; 21)$$

Пример 2. Производная произведения. Классическая формула производной произведения функций $(uv)' = u'v + uv'$ является частным случаем производной билинейного непрерывного отображения.

Теорема 12. Пусть \vec{u}, \vec{v} — дифференцируемые (соответственно непрерывно дифференцируемые) отображения открытого множества Ω нормированного аффинного пространства E в нормированные векторные пространства \vec{F}_1 и \vec{F}_2 , и пусть B — билинейное непрерывное отображение $\vec{F}_1 \times \vec{F}_2$ в \vec{G} . Тогда отображение $B(\vec{u}_1, \vec{u}_2): x \rightarrow B(\vec{u}_1(x), \vec{u}_2(x))$ из E в \vec{G} дифференцируемо (соответственно непрерывно дифференцируемо) и его производное отображение для $\vec{x} \in \vec{E}$ определяется следующей формулой:

$$(B(\vec{u}_1, \vec{u}_2))'(a) \cdot \vec{X} = B(u'_1(a) \cdot \vec{X}; \vec{u}_2(a)) + B(\vec{u}_1(a); u'_2(a) \cdot \vec{X}), \quad (\text{III}, 4; 22)$$

или

$$\overrightarrow{D_{\vec{X}}} B(\vec{u}_1, \vec{u}_2)(a) = B(\overrightarrow{D_{\vec{X}}} \vec{u}_1(a), \vec{u}_2(a)) + B(\vec{u}_1(a), \overrightarrow{D_{\vec{X}}} \vec{u}_2(a)), \quad (\text{III}, 4; 23)$$

¹⁾ Во второй формуле (III, 4; 19) или (III, 4; 21) правые части являются линейными непрерывными отображениями \vec{E} в \mathbb{R} , т. е. элементами сопряженного пространства \vec{E}' .

или в дифференциальных обозначениях

$$\begin{aligned} d(B(\vec{u}_1, \vec{u}_2)) &= B(u'_1(a) \cdot \vec{dx}, \vec{u}_2(a)) + B(\vec{u}_1(a), u'_2(a) \cdot \vec{dx}) = \\ &= B(\vec{du}_1, \vec{u}_2) + B(\vec{u}_1, \vec{du}_2). \quad (\text{III}, 4; 24)^1) \end{aligned}$$

Доказательство. Рассматриваемое отображение является композицией двух отображений: сначала берется отображение пространства E в $\vec{F}_1 \times \vec{F}_2$, определяемое формулой $x \rightarrow (\vec{u}_1(x), \vec{u}_2(x))$, а затем — отображение B пространства $\vec{F}_1 \times \vec{F}_2$ в \vec{G} . Каждое из этих двух отображений дифференцируемо (соответственно непрерывно дифференцируемо) (теоремы 8₄ и 9₂), и остается лишь применить теорему о сложной функции (соответственно ее следствие 5).

В качестве непосредственного применения предположим, что в некоторой механической задаче u и v представляют собой векторы ориентированного трехмерного евклидова аффинного пространства, являющиеся дифференцируемыми функциями времени t с производными $\vec{u}'(t)$ и $\vec{v}'(t)$. Тогда скалярное произведение и векторное произведение этих двух векторов также являются дифференцируемыми функциями времени и имеют место формулы²⁾

$$\begin{aligned} (\vec{u} | \vec{v})' &= (\vec{u}' | \vec{v}) + (\vec{u} | \vec{v}') \quad (\text{скалярное произведение}), \\ (\vec{u} \wedge \vec{v})' &= \vec{u}' \wedge \vec{v} + \vec{u} \wedge \vec{v}' \quad (\text{векторное произведение}). \end{aligned} \quad (\text{III}, 4; 32)$$

То, что нами было получено для билинейного отображения, естественно, годится и для мультилинейного отображения: если задано m дифференцируемых отображений $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m$ открытого множества $\Omega \subset E$ в нормированные векторные пространства $\vec{F}_1, \vec{F}_2, \dots, \vec{F}_m$ и если B есть m -линейное непрерывное отображение $\vec{F}_1 \times \vec{F}_2 \times \dots \times \vec{F}_m$ в \vec{G} , то функция $B^* = B(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m)$: $x \rightarrow B(\vec{u}_1(x), \vec{u}_2(x), \dots, \vec{u}_m(x))$ дифференцируема и ее производное отображение задается формулой

$$B^{**}(a) \cdot \vec{X} = \sum_{j=1}^m B(\overrightarrow{u_1(a)}, \dots, \overrightarrow{u_{j-1}(a)}, u'_j(a) \cdot \vec{X}, \overrightarrow{u_{j+1}(a)}, \dots, \overrightarrow{u_m(a)}). \quad (\text{III}, 4; 33)$$

¹⁾ Здесь и далее нумерация формул соответствует оригиналу. — Прим. ред.

²⁾ Поскольку t пробегает вещественную прямую \mathbb{R} , вместо производных отображений в смысле (III, 3; 13) мы имеем производные векторы в смысле (III, 3; 1). Для этого достаточно в (III, 4; 22) положить $X = 1 \in \mathbb{R}$.

Например, если три вектора $\vec{u}, \vec{v}, \vec{w}$ ориентированного трехмерного евклидова аффинного пространства — дифференцируемые функции времени, то их смешанное произведение $(\vec{u}, \vec{v}, \vec{w})$ является дифференцируемой скалярной функцией времени и его производная вычисляется по формуле

$$(\vec{u}, \vec{v}, \vec{w})' = (\vec{u}', \vec{v}, \vec{w}) + (\vec{u}, \vec{v}', \vec{w}) + (\vec{u}, \vec{v}, \vec{w}'). \quad (\text{III, 4; 34})$$

Пусть \vec{H} — нормированное векторное пространство. Рассмотрим отображение $u \rightarrow u \circ u \circ \dots \circ u = u^m$ пространства $\mathcal{L}(\vec{H}; \vec{H})$ в себя. Это отображение является композицией отображения $u \rightarrow (u, u, \dots, u)$ из $\vec{E} = \mathcal{L}(\vec{H}; \vec{H})$ в пространство $\vec{F}_1 \times \vec{F}_2 \times \dots \times \vec{F}_m = \mathcal{L}(\vec{H}; \vec{H}) \times \mathcal{L}(\vec{H}; \vec{H}) \times \dots \times \mathcal{L}(\vec{H}; \vec{H})$ и мультилинейного непрерывного отображения $(u_1, u_2, \dots, u_m) \rightarrow u_1 \circ u_2 \circ \dots \circ u_m$ пространства $\vec{F}_1 \times \vec{F}_2 \times \dots \times \vec{F}_m$ в пространство $\vec{G} = \mathcal{L}(\vec{H}; \vec{H})$ (теорема 54 гл. II). Применяя теорему 12, получаем

$$\begin{aligned} d(u^m) &= du \circ u \circ \dots \circ u + u \circ du \circ u \circ \dots \circ u + \dots + u \circ u \circ \dots \circ du = \\ &= du \circ u^{m-1} + u \circ du \circ u^{m-2} + \dots + u^{m-1} \circ du. \end{aligned} \quad (\text{III, 4; 35})$$

Если \vec{H} является полем скаляров \mathbb{K} , то $\mathcal{L}(\vec{H}; \vec{H})$ — коммутативное поле скаляров и мы возвращаемся к классическому случаю дифференцирования степенной функции:

$$d(x^m) = mx^{m-1} dx.$$

Пример 3. Производная отношения. Пусть u есть отображение открытого множества Ω аффинного нормированного пространства E в векторное нормированное пространство \vec{F} , и пусть v — скалярная функция, определенная на Ω и всюду $\neq 0$. Тогда можно определить отношение $\vec{u}/v: x \rightarrow \vec{u}(x)/v(x)$. Если u и v — дифференцируемые функции, то это отношение также дифференцируемо и его производная определяется формулой

$$d\left(\frac{\vec{u}}{v}\right) = \frac{v \vec{du} - \vec{dv} \vec{u}}{v^2}. \quad (\text{III, 4; 36})$$

Доказательство очевидно. Отображение \vec{u}/v является композицией отображения $x \rightarrow (\vec{u}(x), 1/v(x))$ множества Ω в $\vec{F} \times \mathbb{K}$ и билинейного отображения $(\vec{u}, \lambda) \rightarrow \lambda \vec{u}$ пространства $\vec{F} \times \mathbb{K}$ в \vec{F} .

Поскольку, согласно примеру 1, $d\left(\frac{1}{v}\right) = -\frac{dv}{v^2} = -\frac{v'(x) dx}{v^2(x)}$, то по формуле (III, 4; 24) получаем (III, 4; 36).

Применение к замене переменных. Пусть g — некоторая функция переменной x , т. е. некоторое отображение открытого множества Ω нормированного аффинного пространства E в нормированное аффинное пространство G . Практически эти пространства обычно конечномерны и в них заданы системы координат. Предположим теперь, что g удовлетворяет некоторому уравнению в частных производных или системе уравнений в частных производных, и мы хотим сделать замену переменной $y = f(x)$. Если речь идет о некоторой биекции, то g станет функцией переменной y , т. е. отображением некоторого открытого множества из F в пространство G . Теперь g будет удовлетворять новому уравнению или новой системе уравнений в частных производных относительно y , и наша задача состоит в том, чтобы найти эту систему уравнений. Такая постановка предполагает заранее известным, что дифференцируемость g по x эквивалентна ее дифференцируемости по y . Предполагается также, что f дифференцируема вместе с обратной биекцией f^{-1} . Будем называть x старой переменной, y — новой переменной и будем искать новое уравнение в частных производных исходя из старого. Поскольку старое уравнение содержит производную dg/dx , достаточно выразить ее через производную dg/dy , что делается по формуле $\frac{dg}{dx} = \frac{dg}{dy} \circ f'(x)$, а затем подставить результат в рассматриваемое уравнение в частных производных. Так получается новое уравнение, содержащее только производную dg/dy , являющуюся искомым уравнением, конечно, при условии, что оно содержит лишь переменную y , т. е. что повсюду переменная x заменена на переменную y через ее выражение $x = f^{-1}(y)$, требующее введения обратной биекции f^{-1} .

Пример 1. Уравнение колебания струны. Так называют уравнение¹⁾

$$\frac{\partial^2 \vec{U}}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 \vec{U}}{\partial t^2} = \vec{0}, \quad (\text{III, 4; 37})$$

где x — пространственная вещественная переменная, t — временная вещественная переменная и \vec{U} — векторная функция двух переменных x и t , т. е. отображение $\mathbb{R} \times \mathbb{R}$ в \vec{F} .

¹⁾ Мы должны были бы дать этот пример лишь после изучения производных порядка > 1 . Однако они уже прежде изучались. Необходимое обоснование будет дано в § 6; см. теорему 19.

Это уравнение имеет следующее физическое происхождение. Рассмотрим однородную струну с закрепленными концами A и B , способную совершать поперечные колебания в окрестности своего положения равновесия. Точка струны, занимавшая в равновесии положение M_0 , во время движения занимает положение M . Ее перемещение в каждый момент времени определяется вектором $\vec{U} = \overrightarrow{M_0 M}$. Если точка M_0 струны имеет на ней абсциссу x , то \vec{U} является функцией двух переменных x и t . Существенно заметить, что речь идет о *малых колебаниях* в некоторой окрестности положения равновесия, т. е. струна предполагается сильно натянутой, почти прямолинейной между A и B . Без этих предположений уравнение движения не будет таким простым, как указанное выше. Функция \vec{U} принимает свои значения в двумерном векторном подпространстве, перпендикулярном \vec{AB} . Величина v в уравнении — скорость распространения поперечных колебаний вдоль струны, определяемая по формуле

$$v = \sqrt{\frac{T}{\rho}}, \quad (\text{III}, 4; 38)$$

где T — среднее натяжение струны и ρ — линейная плотность, т. е. масса единицы длины струны. Заметим, что уравнение физически однородно, ибо размерность величины v , вычисляемой по последней формуле, в единицах массы, длины и времени равна $\sqrt{\frac{m l t^{-2}}{m l^{-1}}} = l t^{-1}$, т. е. является размерностью скорости¹⁾.

Выполним теперь замену переменных

$$\xi = x + vt, \quad \eta = x - vt. \quad (\text{III}, 4; 39)$$

Здесь речь идет о некоторой биекции, дифференцируемой вместе со своей обратной биекцией, определяемой линейными дифференцируемыми функциями

$$x = \frac{\xi + \eta}{2}, \quad t = \frac{\xi - \eta}{2v}. \quad (\text{III}, 4; 40)$$

Если в функции \vec{U} мы заменим x и t их значениями, то получим сложную функцию \vec{U}^* , определенную соотношениями

$$\xi, \eta \rightarrow \vec{U}^*(\xi, \eta) = \vec{U}\left(\frac{\xi + \eta}{2}, \frac{\xi - \eta}{2v}\right). \quad (\text{III}, 4; 41)$$

Коротко ее можно обозначить через \vec{U} , «считая выраженной через ξ и η ».

¹⁾ Разумеется, предполагается, что во время колебаний длина струны изменяется пренебрежимо мало.

Обратно, имеет место формула

$$\vec{U}(x, t) = \vec{U}^*(x + vt, x - vt). \quad (\text{III}, 4; 42)$$

Вычисляя старые производные с помощью новых, получаем формулы

$$\frac{\partial \vec{U}}{\partial x} = \frac{\partial \vec{U}}{\partial \xi} + \frac{\partial \vec{U}}{\partial \eta}, \quad \frac{\partial \vec{U}}{\partial t} = v \left(\frac{\partial \vec{U}}{\partial \xi} - \frac{\partial \vec{U}}{\partial \eta} \right), \quad (\text{III}, 4; 43)$$

которые позволяют найти формулы для частных производных второго порядка

$$\frac{\partial^2 \vec{U}}{\partial x^2} = \frac{\partial^2 \vec{U}}{\partial \xi^2} + 2 \frac{\partial^2 \vec{U}}{\partial \xi \partial \eta} + \frac{\partial^2 \vec{U}}{\partial \eta^2}, \quad \frac{1}{v^2} \frac{\partial^2 \vec{U}}{\partial t^2} = \frac{\partial^2 \vec{U}}{\partial \xi^2} - 2 \frac{\partial^2 \vec{U}}{\partial \xi \partial \eta} + \frac{\partial^2 \vec{U}}{\partial \eta^2}. \quad (\text{III}, 4; 44)$$

Отсюда теперь получаем, что

$$\frac{\partial^2 \vec{U}}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 \vec{U}}{\partial t^2} = 4 \frac{\partial^2 \vec{U}}{\partial \xi \partial \eta}. \quad (\text{III}, 4; 45)$$

В этой записи смешаны функции \vec{U} и \vec{U}^* . В действительности эту формулу надо было бы записать так¹⁾:

$$\frac{\partial^2 \vec{U}}{\partial x^2}(x, t) - \frac{1}{v^2} \frac{\partial^2 \vec{U}}{\partial t^2}(x, t) = 4 \frac{\partial^2 \vec{U}^*}{\partial \xi \partial \eta}(x + vt, x - vt). \quad (\text{III}, 4; 46)$$

Полученное новое уравнение относительно переменной \vec{U} , выраженной как функция от ξ и η (в действительности относительно \vec{U}^*), является значительно более простым уравнением в частных производных

$$\frac{\partial^2 \vec{U}}{\partial \xi \partial \eta} = 0. \quad (\text{III}, 4; 47)$$

Нетрудно решить это уравнение. В самом деле, если мы рассмотрим частную производную $\frac{\partial \vec{U}}{\partial \xi}$, то заметим, что ее частная производная по η равна нулю. Отсюда следует, что она постоянна по η , т. е. является «произвольной» функцией ξ . Но теперь видно, что \vec{U} является суммой какой-либо первообразной этой функции ξ , которая сама является «произвольной»

¹⁾ Правые части равенства являются функциями от x и t . Как обычно, $\frac{\partial^2 \vec{U}^*}{\partial \xi \partial \eta}(x + vt, x - vt)$ — это значение $\frac{\partial^2 \vec{U}^*}{\partial \xi \partial \eta}$ в точке $(x + vt, x - vt)$.

функцией ξ , и некоторой постоянной относительно ξ , т. е. «произвольной» функции η . Окончательно получаем формулу

$$\vec{U} = \vec{f}(\xi) + \vec{g}(\eta), \quad (\text{III}, 4; 47_2)$$

дающую искомое решение уравнения колебания струны:

$$\vec{U}(x, t) = \vec{f}(x + vt) + \vec{g}(x - vt), \quad (\text{III}, 4; 48)$$

где \vec{f} и \vec{g} — «произвольные» функции вещественной переменной. Выражение «произвольные функции» не точно. Надо, чтобы были обоснованы все вычисления при замене переменных. (см. теорему 19), для чего необходимо, чтобы функция \vec{U} имела полную производную 2-го порядка, а это в свою очередь означает, что функции \vec{f} и \vec{g} дважды дифференцируемы.

Упражнение. Решить тем же самым методом уравнение колебания струны с правой частью:

$$\frac{\partial^2 \vec{U}}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 \vec{U}}{\partial t^2} = \vec{g}(x, t), \quad (\text{III}, 4; 49)$$

где правая часть является заданной непрерывной функцией x и t . Это уравнение возникает в том случае, когда струна, кроме натяжения, подвергается действию других сил, поперечно направленных и зависящих только от x и t .

Пример 2. Частные производные в декартовых и полярных координатах на плоскости.

1°) Выражение производных в полярной системе координат через производные в декартовой системе координат.

По формулам замены переменных

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad (\text{III}, 4; 50)$$

пользуясь указанными ранее краткими обозначениями, получаем формулы:

$$\frac{\partial}{\partial r} = \cos \varphi \frac{\partial}{\partial x} + \sin \varphi \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial \varphi} = -r \sin \varphi \frac{\partial}{\partial x} + r \cos \varphi \frac{\partial}{\partial y}. \quad (\text{III}, 4; 51)$$

Каков же точный смысл этих формул?

Пусть g — отображение плоскости \mathbb{R}^2 в аффинное пространство над полем вещественных чисел. Его можно записать в виде $(x, y) \rightarrow g(x, y)$.

Пусть, с другой стороны, P — отображение плоскости \mathbb{R}^2 в себя, определенное формулами

$$(r, \varphi) \rightarrow (x, y), \quad x = r \cos \varphi, \quad y = r \sin \varphi. \quad (\text{III}, 4; 52)$$

Z (Внимание! Через (r, φ) обозначена точка плоскости \mathbb{R}^2 , для которой r и φ — ее обычные декартовы координаты. Ее образ $P(r, \varphi)$ имеет декартовы координаты x, y , определяемые по формулам (III, 4; 52), а, следовательно, (r, φ) являются полярными координатами этого образа.)

Отображение P дифференцируемо, и его матрица Якоби определяется формулой

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}^1). \quad (\text{III}, 4; 53)$$

Отсюда вытекает, что если g является дифференцируемой функцией, то сложное отображение $g^* = g \circ P$ также дифференцируемо и имеют место формулы

$$\frac{\vec{\partial} g^*}{\partial r}(r, \varphi) = \frac{\vec{\partial} g}{\partial x}(r \cos \varphi, r \sin \varphi) \cos \varphi + \frac{\vec{\partial} g}{\partial y}(r \cos \varphi, r \sin \varphi) \sin \varphi, \quad (\text{III}, 4; 54)$$

$$\frac{\vec{\partial} g^*}{\partial \varphi}(r, \varphi) = -\frac{\vec{\partial} g}{\partial x}(r \cos \varphi, r \sin \varphi) r \sin \varphi + \frac{\vec{\partial} g}{\partial y}(r \cos \varphi, r \sin \varphi) r \cos \varphi,$$

которые, пользуясь принятymi краткими обозначениями, можно записать в виде.

$$\frac{\vec{\partial} g}{\partial r} = \cos \varphi \frac{\vec{\partial} g}{\partial x} + \sin \varphi \frac{\vec{\partial} g}{\partial y}, \quad \frac{\vec{\partial} g}{\partial \varphi} = -r \sin \varphi \frac{\vec{\partial} g}{\partial x} + r \cos \varphi \frac{\vec{\partial} g}{\partial y} \quad (\text{III}, 4; 55)$$

(как обычно, g^* и g отождествлены). Отсюда получаются формулы (III, 4; 51).

2°) *Выражение производных в декартовых координатах через производные в полярных координатах.*

Можно попытаться выразить частные производные по x и y , исходя из частных производных по r и φ . Однако следует заметить, что отображение P не является биекцией. Отображение P сюръективно, но не инъективно, ибо точка (x, y) имеет бесконечное множество полярных координат. Поэтому следует ожидать, что производное отображение P не будет линейной биекцией. Тем не менее в окрестности точки (r, φ) ,

¹⁾ Из теоремы 84 следует, что если обе скалярные функции $(r, \varphi) \rightarrow r \cos \varphi$, $(r, \varphi) \rightarrow r \sin \varphi$ дифференцируемы, то отображение P также дифференцируемо. Каждая из этих двух функций дифференцируема, если дифференцируемы сомножители. Итак, остается показать, что дифференцируемы функции $(r, \varphi) \rightarrow r$ и $(r, \varphi) \rightarrow \cos \varphi$ или $\sin \varphi$. Далее, функция $(r, \varphi) \rightarrow r$ линейна, а функция $(r, \varphi) \rightarrow \cos \varphi$ является композицией отображений $(r, \varphi) \rightarrow \varphi$ и $\varphi \rightarrow \cos \varphi$, из которых первое линейно, а второе известно как дифференцируемая вещественная функция вещественной переменной.

Можно также применить следствие 2 теоремы 15: $(r, \varphi) \rightarrow r \cos \varphi$ и $(r, \varphi) \rightarrow r \sin \varphi$ имеют непрерывные частные производные.

в которой якобиан преобразования P : $\begin{vmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{vmatrix} = r$ отличен от 0, т. е. в точке $(x, y) \neq (0, 0)$, уравнения (III, 4; 54) разрешимы¹⁾, и, исходя из частных производных $\vec{\partial g^*}/\partial r$, $\vec{\partial g^*}/\partial \varphi$, можно вычислить частные производные $\vec{\partial g}/\partial x$, $\vec{\partial g}/\partial y$. Непосредственными вычислениями получаем:

$$\begin{aligned} \frac{\vec{\partial g}}{\partial x}(r \cos \varphi, r \sin \varphi) &= \cos \varphi \frac{\vec{\partial g^*}}{\partial r}(r, \varphi) - \frac{\sin \varphi}{r} \frac{\vec{\partial g^*}}{\partial \varphi}(r, \varphi), \\ \frac{\vec{\partial g}}{\partial y}(r \cos \varphi, r \sin \varphi) &= \sin \varphi \frac{\vec{\partial g^*}}{\partial r}(r, \varphi) + \frac{\cos \varphi}{r} \frac{\vec{\partial g^*}}{\partial \varphi}(r, \varphi). \end{aligned} \quad (\text{III}, 4; 56)$$

Заметим, что в этой формуле можно заменить r и φ соответствующими функциями от x и y .

Можно также посмотреть на вышеизложенное с другой точки зрения. Пусть (r_0, φ_0) — некоторая точка \mathbb{R}^2 , $r_0 > 0$, и пусть (x_0, y_0) ее образ при отображении P . Если P сузить на открытое множество Ω из \mathbb{R}^2 , определенное неравенствами $r > 0$, $\varphi_0 - \pi < \varphi < \varphi_0 + \pi$, то получим некоторую биекцию P_Ω (и даже гомеоморфизм) множества Ω на $P(\Omega)$, имеющую обратную биекцию P^{-1} . Когда точка (x, y) изменяется в $P(\Omega)$, ей можно отнести единственную непрерывно изменяющуюся систему полярных координат, т. е. единственный прообраз в Ω , и мы будем иметь: $g = g^* P_\Omega^{-1}$. При этом формулы (III, 4; 56), в которых r и φ заменены их выражениями в виде функций от x и y , будут иметь тот же смысл, что и формулы (III, 4; 54). Запишем эти формулы:

$$\begin{aligned} \frac{\vec{\partial g}}{\partial x}(x, y) &= \frac{x}{\sqrt{x^2 + y^2}} \frac{\vec{\partial g^*}}{\partial r}(\sqrt{x^2 + y^2}, \varphi(x, y)) - \\ &\quad - \frac{y}{x^2 + y^2} \frac{\vec{\partial g^*}}{\partial \varphi}(\sqrt{x^2 + y^2}, \varphi(x, y)), \\ \frac{\vec{\partial g}}{\partial y}(x, y) &= \frac{y}{\sqrt{x^2 + y^2}} \frac{\vec{\partial g^*}}{\partial r}(\sqrt{x^2 + y^2}, \varphi(x, y)) + \\ &\quad + \frac{x}{x^2 + y^2} \frac{\vec{\partial g^*}}{\partial \varphi}(\sqrt{x^2 + y^2}, \varphi(x, y)). \end{aligned} \quad (\text{III}, 4; 56_2)$$

Аналогичные формулы можно получить для $r_0 < 0$, если взять $r < 0$ и $r = -\sqrt{x^2 + y^2}$. Что же касается непрерывной функции φ : $(x, y) \rightarrow \varphi(x, y)$, то у нее нет простого аналитиче-

¹⁾ Мы применили здесь непосредственно очевидный частный случай теоремы о неявной функции (§ 8; см., в частности, замечание на стр. 311).

ского выражения. Очевидно, $\varphi(x, y) = \arctg y/x + k\pi$, где k зависит от x, y и не всегда одно и то же для $(x, y) \in P(\Omega)$.

Можно было бы для решения обеих задач применить дифференциальные обозначения. В самом деле, запишем дифференциал в виде

$$\vec{dg} = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy, \quad (\text{III}, 4; 57)$$

а затем воспользуемся формулами

$$dx = \cos \varphi dr - r \sin \varphi d\varphi, \quad dy = \sin \varphi dr + r \cos \varphi d\varphi, \quad (\text{III}, 4; 58)$$

которые дадут новое выражение дифференциала

$$\vec{dg^*} = \left(\cos \varphi \frac{\partial g}{\partial x} + \sin \varphi \frac{\partial g}{\partial y} \right) dr + \left(-r \sin \varphi \frac{\partial g}{\partial x} + r \cos \varphi \frac{\partial g}{\partial y} \right) d\varphi. \quad (\text{III}, 4; 59)$$

Коэффициенты при dr и $d\varphi$ являются частными производными $\partial g^*/\partial r$ и $\partial g^*/\partial \varphi$, что снова дает решение 1-й задачи и формулу (III, 4; 54).

Для того чтобы снова получить выражение (III, 4; 56), можно провести то же самое рассуждение, исходя из формулы

$$\vec{dg^*} = \frac{\partial g^*}{\partial r} dr + \frac{\partial g^*}{\partial \varphi} d\varphi. \quad (\text{III}, 4; 60)$$

Затем dr и $d\varphi$ следует выразить как функции от dx и dy путем решения системы (III, 4; 58), т. е. положить

$$dr = \cos \varphi dx + \sin \varphi dy \quad \text{и} \quad d\varphi = -\frac{1}{r} \sin \varphi dx + \frac{1}{r} \cos \varphi dy, \quad (\text{III}, 4; 61)$$

откуда подстановкой в (III, 4; 60) получается:

$$\vec{dg} = \left(\cos \varphi \frac{\partial g^*}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial g^*}{\partial \varphi} \right) dx + \left(\sin \varphi \frac{\partial g^*}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial g^*}{\partial \varphi} \right) dy. \quad (\text{III}, 4; 62)$$

Отсюда снова находится решение 2-й задачи и формула (III, 4; 56).

В качестве применения вычислим, например, лапласиан в полярных координатах.

Лапласиан функции g переменных x и y определяется формулой

$$\vec{\Delta g} = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}. \quad (\text{III}, 4; 63)$$

Для частных производных 1-го порядка уже имеются формулы (III, 4; 56). Вычисляя частные производные 2-го порядка

как производные от производных 1-го порядка:

$$\frac{\partial^2 g}{\partial x^2} = \left(\cos \varphi \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi} \right) \left(\cos \varphi \frac{\partial g^*}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial g^*}{\partial \varphi} \right), \quad (\text{III}, 4; 64)$$

получаем выражение вида

$$\vec{\alpha} \cos^2 \varphi + \vec{\beta} \sin \varphi \cos \varphi + \vec{\gamma} \sin^2 \varphi. \quad (\text{III}, 4; 65)$$

От частных производных по x к частным производным по y можно перейти, заменив φ на $\varphi - \frac{\pi}{2}$, т. е. $\cos \varphi$ и $\sin \varphi$ на $\sin \varphi$ и $-\cos \varphi$ соответственно. Выражение для $\frac{\partial^2 g}{\partial y^2}$ при этом получит вид $\vec{\alpha} \sin^2 \varphi - \vec{\beta} \sin \varphi \cos \varphi + \vec{\gamma} \cos^2 \varphi$, а лапласиан будет равен $\vec{\alpha} + \vec{\gamma}$, т. е. будет суммой коэффициентов при $\cos^2 \varphi$ и $\sin^2 \varphi$ в выражении (III, 4; 65). Непосредственное вычисление дает

$$\vec{\Delta}g(r \cos \varphi, r \sin \varphi) = \frac{\partial^2 g^*}{\partial r^2}(r, \varphi) + \frac{1}{r} \frac{\partial g^*}{\partial r}(r, \varphi) + \frac{1}{r^2} \frac{\partial^2 g^*}{\partial \varphi^2}(r, \varphi), \quad (\text{III}, 4; 66)$$

что можно коротко записать в виде:

$$\vec{\Delta}g = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r} \frac{\partial g}{\partial r} + \frac{1}{r^2} \frac{\partial^2 g}{\partial \varphi^2}, \quad (\text{III}, 4; 67)$$

или

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2}.$$

Замечания. 1°) Эта формула удовлетворяет правилу однородности. Если величины x, y, r имеют размерность длины l , а φ безразмерна, то размерность $\vec{\Delta}g$ равна gl^{-2} и то же самое имеет место для членов правой части.

2°) Во всех вычислениях предполагается, что g имеет полную производную 2-го порядка (см. § 6, теорема 19).

Пример 3. Вычисление лапласиана от функции, определенной в \mathbb{R}^n и зависящей только от расстояния до начала координат.

Пусть g — отображение полупрямой \mathbb{R}_+ (множество вещественных чисел ≥ 0) в нормированное аффинное пространство G . Рассмотрим отображение r пространства \mathbb{R}^n в полупрямую \mathbb{R}_+ , определенное по формуле

$$r(x_1, x_2, \dots, x_n) = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}. \quad (\text{III}, 4; 68)$$

Сложное отображение $g^* = g \circ r$, которое будет кратко записываться в виде $g(r)$, является отображением \mathbb{R}^n в G , зави-

сящим только от расстояния исходной точки до начала координат. Если g дважды дифференцируемо и $(x_1, x_2, \dots, x_n) \neq (0, 0, \dots, 0)$ (для того, чтобы r также было дважды дифференцируемым), то для отображения g^* в точке (x_1, x_2, \dots, x_n) существует лапласиан $\vec{\Delta}g^* = \sum_{i=1}^n \frac{\partial^2 \vec{g}^*}{\partial x_i^2}$. Этот лапласиан может быть вычислен предыдущими методами. Прежде всего

$$\begin{aligned}\frac{\partial \vec{g}^*}{\partial x_i} &= \vec{g}' \frac{\partial r}{\partial x_i} = \vec{g}' \frac{x_i}{r}, \\ \frac{\partial^2 \vec{g}^*}{\partial x_i^2} &= \vec{g}'' \frac{x_i^2}{r^2} + \vec{g}' \frac{1}{r} - \vec{g}' \frac{x_i^2}{r^3}.\end{aligned}\quad (\text{III, 4; 69})$$

Складывая затем результаты, полученные для $i = 1, 2, \dots, n$, находим:

$$\begin{aligned}\vec{\Delta}g^*(x_1, x_2, \dots, x_n) &= \vec{g}''(\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}) + \\ &+ \frac{n-1}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}} \vec{g}'(\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}),\end{aligned}\quad (\text{III, 4; 70})$$

что можно коротко записать в виде

$$\vec{\Delta}g = \vec{g}'' + \frac{n-1}{r} \vec{g}' = \frac{\partial^2 \vec{g}}{\partial r^2} + \frac{n-1}{r} \frac{\partial \vec{g}}{\partial r}. \quad (\text{III, 4; 71})$$

При $n = 2$ эта формула является частным случаем формулы (III, 4; 66), в которой g и g^* поменялись ролями. (Если в (III, 4; 66) предположить, что g^* зависит только от r , то $\frac{\partial^2 \vec{g}^*}{\partial \vec{r}^2} = 0$ и мы получим $\vec{\Delta}g = \frac{\partial^2 \vec{g}^*}{\partial r^2} + \frac{1}{r} \frac{\partial \vec{g}^*}{\partial r}$, что эквивалентно (III, 4; 71) при $n = 2$.) Легко проверяется правильность множителя $n-1$: при $n = 1$ получаем $\vec{\Delta}g^* = \vec{g}''$. Можно заметить также, что формула удовлетворяет правилу однородности: правая и левая части имеют размерность gl^{-2} (см. конец примера 2°).)

§ 5. ФОРМУЛА КОНЕЧНЫХ ПРИРАЩЕНИЙ

Предположение, что f имеет производную в точке $a \in \Omega$, означает, что приращение $f(a + \vec{h}) - f(a)$ для бесконечно малого \vec{h} выражается по формуле (III, 3; 13), и в этом смысле формулу (III, 3; 13) можно было бы назвать формулой

бесконечно малых приращений. Однако формула (III, 3; 13) не дает никакой оценки для приращения $\vec{f}(a + \vec{h}) - \vec{f}(a)$ функции f при определенном приращении аргумента \vec{h} . Этой цели будет служить формула конечных приращений, использующая производную функции f в точках, близких к точке a .

Формула (III, 2; 2) непосредственно обобщается на отображения E в F в том случае, когда F является полем вещественных чисел.

Теорема 13А. Пусть f — вещественная непрерывная функция, заданная на открытом множестве Ω аффинного нормированного пространства E над полем вещественных чисел. Если отрезок $[x, x + \vec{h}]$ целиком содержится в Ω и если функция f имеет производное отображение в каждой точке открытого интервала $[x, x + \vec{h}]$, то

$$f(x + \vec{h}) - f(x) = f'(x + \theta \vec{h}) \vec{h} \in \mathbb{R}, \text{ где } 0 < \theta < 1. \quad (\text{III, 5; 0})$$

Доказательство. Рассмотрим отображение $\Phi: t \rightarrow \vec{f}(x + t\vec{h})$ вещественного интервала $[0, 1]$ в $F = \mathbb{R}$. Это отображение является композицией отображения $t \rightarrow x + t\vec{h}$ отрезка $[0, 1]$ в множество Ω и отображения f множества Ω в \mathbb{R} . Согласно теореме о сложной функции, это отображение непрерывно на отрезке $[0, 1]$ (теорема 10 гл. II) и дифференцируемо на интервале $[0, 1]$ (теорема 11 гл. III) и его производная вычисляется по формуле

$$\Phi'(t) = f'(x + t\vec{h}) \vec{h} \in \mathbb{R}. \quad (\text{III, 5; 0}_2)$$

Поскольку $\Phi(1) - \Phi(0) = f(x + \vec{h}) - f(x)$, для получения необходимого результата достаточно к функции Φ на отрезке $[0, 1]$ применить формулу (III, 2; 2).

Напротив, формула (III, 2; 2) в том виде, какой она имеет, не может быть обобщена, если размерность F больше 1 или же если речь идет об аффинных пространствах над комплексным полем. Она не верна уже для комплексной функции вещественной переменной¹⁾. Рассмотрим, например, комплексную функцию $x \rightarrow e^{2ix}$, определенную на отрезке $[0, 1]$ из \mathbb{R} . Эта функция принимает одно и то же значение при $x = 0$ и $x = 1$. Если бы формула конечных приращений была верна в виде (III, 2; 2), то должна была бы существовать точка c интервала $[0, 1]$, в которой производная рассматриваемой комплексной функции

¹⁾ То есть отображения прямой \mathbb{R} в комплексную плоскость \mathbb{C} , где \mathbb{R} и \mathbb{C} рассматриваются как векторные пространства размерности 1 и 2 над полем вещественных чисел.

обращалась бы в нуль. Однако ее производная — функция $x \rightarrow \rightarrow 2i\pi e^{2i\pi x}$ — нигде в нуль не обращается.

Приведем формулу конечных приращений для общего случая в форме, несколько отличной от предыдущей:

Теорема 13. Пусть f — непрерывное отображение открытого множества Ω аффинного нормированного пространства E в аффинное нормированное пространство F . Тогда, если отрезок $[x, x + \vec{h}]$ целиком принадлежит Ω и если f имеет производное отображение в каждой точке открытого интервала $[x, x + \vec{h}]$ и его норма не превосходит M , то имеет место неравенство

$$\overrightarrow{\|f(x + \vec{h}) - f(x)\|} \leq M \|\vec{h}\|. \quad (\text{III}, 5; 1)$$

Доказательство. Предварительно докажем одну лемму.

Лемма. Пусть f — отображение отрезка $[0, 1]$ вещественной прямой \mathbb{R} в аффинное нормированное пространство F и g — вещественная функция на отрезке $[0, 1]$. Предположим, что функции f и g непрерывны на замкнутом отрезке $[0, 1]$ и дифференцируемы на открытом интервале $(0, 1)$. Если при этих условиях имеет место неравенство

$$\overrightarrow{\|f'(x)\|} \leq g'(x) \quad \text{для } 0 < x < 1, \quad (\text{III}, 5; 2)$$

то справедливо также неравенство

$$\overrightarrow{\|f(1) - f(0)\|} \leq g(1) - g(0). \quad (\text{III}, 5; 3)$$

Докажем эту лемму. Пусть задано $\varepsilon > 0$. Обозначим через A_ε множество точек x интервала $[0, 1]$, для которых

$$\overrightarrow{\|f(x) - f(0)\|} \leq g(x) - g(0) + \varepsilon x + \varepsilon. \quad (\text{III}, 5; 4)$$

Функция

$$x \rightarrow \overrightarrow{\|f(x) - f(0)\|} - g(x) + g(0) - \varepsilon x - \varepsilon \quad (\text{III}, 5; 4_2)$$

непрерывна. Множество A_ε точек, в которых эта функция ≤ 0 , замкнуто. В частности, A_ε содержит свою точную верхнюю грань β , которая тем самым является максимумом. Заметим, что β не может быть $= 0$. В самом деле, точка $x = 0$ принадлежит множеству A_ε , так как функция (III, 5; 4₂) при $x = 0$ принимает значение $-\varepsilon < 0$. Из непрерывности f и g следует, что A_ε содержит некоторую окрестность нуля. Однако неравенство $0 < \beta < 1$ также невозможно. Действительно, по определению производной (формула (III, 3; 13)), если $0 < \beta < 1$, то

существует число $\delta > 0$, при котором справедливы неравенства:

$$\begin{aligned} \overrightarrow{\|f(\beta + \delta) - f(\beta)\|} &\leq \overrightarrow{\|f'(\beta)\|} \delta + \frac{\epsilon}{2} \delta \leq g'(\beta) \delta + \frac{\epsilon}{2} \delta, \\ g(\beta + \delta) - g(\beta) &\geq g'(\beta) \delta - \frac{\epsilon}{2} \delta, \end{aligned} \quad (\text{III}, 5; 5)$$

откуда

$$\overrightarrow{\|f(\beta + \delta) - f(\beta)\|} \leq g(\beta + \delta) - g(\beta) + \epsilon \delta. \quad (\text{III}, 5; 6)$$

Однако, поскольку $\beta \in A_\epsilon$, то

$$\overrightarrow{\|f(\beta) - f(0)\|} \leq g(\beta) - g(0) + \epsilon \beta + \epsilon, \quad (\text{III}, 5; 6_2)$$

а, следовательно, после сложения,

$$\overrightarrow{\|f(\beta + \delta) - f(0)\|} \leq g(\beta + \delta) - g(0) + \epsilon(\beta + \delta) + \epsilon. \quad (\text{III}, 5; 6_3)$$

Отсюда следует, что $\beta + \delta \in A_\epsilon$, что невозможно, поскольку β является максимумом для множества A_ϵ .

Таким образом, $\beta = 1$. При этом (III, 5; 6₂) запишется в виде

$$\overrightarrow{\|f(1) - f(0)\|} \leq g(1) - g(0) + 2\epsilon. \quad (\text{III}, 5; 6_4)$$

Так как это неравенство справедливо при любом $\hat{\epsilon}$, отсюда следует справедливость неравенства (III, 5; 3), и лемма полностью доказана.

Замечания к лемме. 1°) Вместо того чтобы предполагать функции f и g дифференцируемыми, можно ограничиться дифференцируемостью справа и соответственно условием

$\overrightarrow{\|f'_n\|} \leq g'_n$. Это видно непосредственно из доказательства. Если заменить $f(x)$, $g(x)$ на $-f(1-x)$, $-g(1-x)$, то можно обойтись дифференцируемостью слева.

2°) Предположим сверх того, что в интервале $]0, 1[$ имеется хотя бы одна точка c , в которой выполняется строгое неравенство $\overrightarrow{\|f'_n(c)\|} < g'_n(c)$. Тогда $\overrightarrow{\|f(1) - f(0)\|} < g(1) - g(0)$. В самом деле, если $\delta = g'_n(c) - \overrightarrow{\|f'_n(c)\|} > 0$, то для достаточно малого $h > 0$ имеем:

$$\overrightarrow{\|f(c+h) - f(c)\|} \leq \overrightarrow{\|f'_n(c)\|} h + \frac{1}{3} h = g'_n(c) - \frac{2}{3} \delta h,$$

$$g(c+h) - g(c) \geq g'_n(c) - \frac{\delta}{3} h,$$

откуда

$$\overrightarrow{\|f(c+h) - f(c)\|} \leq g(c+h) - g(c) - \frac{\delta}{3}h < g(c+h) - g(c). \quad (\text{III}, 5; 6_5)$$

Однако лемма, примененная к интервалам $[0, c]$, $[c+h, 1]$ вместо интервала $[0, 1]$, дает

$$\begin{aligned} \overrightarrow{\|f(c) - f(0)\|} &\leq g(c) - g(0), \\ \overrightarrow{\|f(1) - f(c+h)\|} &\leq g(1) - g(c+h). \end{aligned} \quad (\text{III}, 5; 6_6)$$

Результат получается сложением $(\text{III}, 5; 6_5)$ и $(\text{III}, 5; 6_6)$.

3°) Если функция f также вещественна, то можно предполагать, что она всюду на $[0, 1]$ имеет производную, конечную или равную $-\infty$, и что имеется только неравенство вида $f'_n \leq g'_n$. Тогда справедливо неравенство $f(1) - f(0) \leq g(1) - g(0)$, а если хотя бы в одной точке $c \in]0, 1[$ имеет место строгое неравенство $f'_n(c) < g'_n(0)$, то $f(1) - f(0) < g(1) - g(0)$.

4°) Отсюда вытекают обобщения, указанные в замечаниях после теорем 4, 5, 7. Рассмотрим теорему 4 (Ролля), из которой следуют остальные. Неравенство $f'_n(x) < 0$ не может выполняться всюду. В самом деле, если положить $g = 0$, то $f(1) - f(0) < 0$, что противоречит условию. По той же причине не может всюду выполняться и неравенство $f'_n(x) > 0$. Следовательно, существует точка c_1 , в которой $f'_n(c_1) \leq 0$, и точка c_2 , где $f'_n(c_2) \geq 0$.

Вернемся теперь к доказательству теоремы. Рассмотрим отображение $\Phi: t \rightarrow f(x + th)$ отрезка $[0, 1]$ в F , введенное при доказательстве теоремы 13А. Его производная записывается в виде

$$\overrightarrow{\Phi'(t)} = f'(x + th) \vec{h}. \quad (\text{III}, 5; 7)$$

Отсюда вытекает, что в условиях теоремы норма этой производной не превосходит $M\|\vec{h}\|$. Если теперь применить лемму, выбирая в качестве g линейную функцию $t \rightarrow M\|\vec{h}\|t$, то получим неравенство $(\text{III}, 5; 1)$.

Теорема о конечных приращениях допускает следующее интересное дополнение:

Следствие 1. Пусть L — линейное непрерывное отображение \vec{E} в \vec{F} . Тогда при условиях теоремы 13 имеет место

неравенство

$$\overrightarrow{\|f(x + \vec{h}) - f(x) - L\vec{h}\|} \leq \omega \|\vec{h}\|, \quad (\text{III}, 5; 8)$$

где ω — точная верхняя грань величины $\|f'(\xi) - L\|$ при ξ , изменяющемся в интервале $[x, x + h]$. Часто предполагают, что функция f имеет производную в x , и тогда полагают $L = f'(x)$.

Доказательство. Достаточно применить теорему к функции $\xi \rightarrow f(\xi) - L \cdot (\xi - x)$, производная которой в точке ξ равна $f'(\xi) - L$.

Следствие 2. Если Ω является открытым выпуклым¹⁾ множеством аффинного нормированного пространства E и если f есть дифференцируемое отображение Ω в F , производная которого в каждой точке Ω по норме не превосходит некоторой постоянной M , то функция f удовлетворяет условию Липшица и, следовательно, равномерно непрерывна.

Доказательство. Из выпуклости множества Ω следует, что если x' и x'' — любые две точки Ω , то весь отрезок $[x', x'']$ принадлежит Ω . Следовательно, к нему можно применить формулу конечных приращений, и мы получим неравенство

$$\overrightarrow{\|f(x') - f(x'')\|} \leq M \|x' - x''\|, \quad (\text{III}, 5; 9)$$

чем и доказывается теорема.

Следствие 3. Пусть f — дифференцируемое отображение аффинного нормированного пространства E в аффинное нормированное пространство F . Предположим, кроме того, что его производная f' является равномерно непрерывным отображением пространства E в пространство $\mathcal{L}(E; F)$. При этих усло-

виях функция $x \rightarrow \frac{\overrightarrow{f(x + \vec{h}) - f(x) - f'(x)\vec{h}}}{\|\vec{h}\|}$ равномерно сходится к $\vec{0}$ при $\vec{h} \neq \vec{0}$, стремящемся к $\vec{0}$.

Доказательство. В силу предположения о равномерной непрерывности производной f' , по заданному $\varepsilon > 0$ можно указать такое $\eta > 0$, что из неравенства $\|x' - x''\| \leq \eta$ будет следовать неравенство $\|f'(x') - f'(x'')\| \leq \varepsilon$. Но тогда, как только $\|\vec{h}\| \leq \eta$, величина ω , входящая в формулу (III, 5; 8), соответствующую $L = f'(x)$, будет по норме не больше ε , что и доказывает высказанное утверждение.

¹⁾ Условие выпуклости весьма существенно.

Замечания. 1°) Пусть E — поле скаляров. Тогда, согласно следствию 3, функция $x \rightarrow (f(x + \vec{h}) - f(x)) / \|\vec{h}\|$ сходится равномерно к производной функции f' при $\vec{h} \neq \vec{0}$, стремящемся к $\vec{0}$. В этом случае говорят, что f *равномерно дифференцируема*.

2°) Эту теорему часто применяют в следующем случае. Предположим, что f определена лишь в открытом множестве Ω из E (возможно, на полуоткрытом или замкнутом интервале, если

$E = \mathbb{R}$). Тогда функция $x \rightarrow \frac{f(x + \vec{h}) - f(x) - f'(x)\vec{h}}{\|\vec{h}\|}$, очевидно,

не определена на всем множестве Ω . Ограничимся рассмотрением множества значений \vec{h} , норма которых не превосходит некоторого фиксированного числа $\delta > 0$. Тогда все рассматриваемые функции будут определены на одном и том же открытом множестве Ω_δ точек из E , расстояние которых до Ω больше δ . Если производная f' равномерно непрерывна, то отсюда полу-

чаем, что на множестве Ω_δ функция $x \rightarrow \frac{f(x + \vec{h}) - f(x) - f'(x)\vec{h}}{\|\vec{h}\|}$ равномерно сходится к $\vec{0}$, когда $\vec{h} \neq \vec{0}$ стремится к $\vec{0}$.

Теорема 14. Пусть Ω — открытое множество аффинного нормированного пространства E . Пусть c — некоторая точка Ω и Ω_0 — дополнение c в Ω . Пусть f — непрерывное отображение Ω в F , всюду дифференцируемое в Ω_0 . Если производная $f'(x)$ стремится к некоторому пределу L в $\mathcal{L}(E; F)$ при $x \in \Omega_0$, стремящемся к c , то f дифференцируема в точке c и ее производная в точке c является линейным непрерывным отображением L .

Доказательство. Достаточно применить к $x = c$ неравенство (III, 5; 8). Результат вытекает из того, что, в силу высказанных предположений, \vec{h} стремится к $\vec{0}$, когда \vec{h} стремится к $\vec{0}$. Теперь видно, почему в теореме 13 и ее следствии 1 было существенным не предполагать f дифференцируемой на концах интервала.

Полная дифференцируемость и частная дифференцируемость

Теорема 15. Пусть E_1, E_2, F — аффинные нормированные пространства, f — отображение открытого множества Ω пространства $E_1 \times E_2$ в F . Для того чтобы отображение f было

непрерывно дифференцируемым в Ω , необходимо и достаточно, чтобы оно имело частные производные $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}$, непрерывные в Ω .

Доказательство. Если отображение f дифференцируемо в множестве Ω , то из теоремы 9 следует существование частных производных и формула (III, 3; 29). Кроме того, неравенства (II, 13; 28) и (II, 13; 30) говорят о том, что если линейное непрерывное отображение u пространства $\vec{E}_1 \times \vec{E}_2$ в \vec{F} представимо по формуле (II, 13; 26) в виде суммы двух линейных непрерывных отображений u_1, u_2 пространств \vec{E}_1, \vec{E}_2 соответственно в \vec{F} , то

$$\begin{aligned} \|u_1\| &\leq \|u\|, \quad \|u_2\| \leq \|u\|, \\ \|u\| &\leq \|u_1\| + \|u_2\|. \end{aligned} \tag{III, 5; 10}$$

Если отображение f непрерывно дифференцируемо, то при x , стремящемся к x_0 , норма $\|f'(x) - f'(x_0)\|$ стремится к 0 и из первого неравенства (III, 5; 10) следует, что $\left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(x_0) \right\|$ ($i=1, 2$) стремится к 0, а это означает, что частные производные f непрерывны.

Обратно, пусть существуют частные производные функции f . Мы уже знаем, что отсюда еще не следует существование ее полной производной¹⁾ (пример (III, 3; 9)). Однако существование непрерывных частных производных влечет за собой существование и непрерывность полной производной (а значит, и непрерывность f).

В самом деле, пусть заданы $\vec{h}_1 \in \vec{E}_1$ и $\vec{h}_2 \in \vec{E}_2$. Если $a = (a_1, a_2) \in \Omega$, то, поскольку множество Ω открыто, существует такое число $\rho > 0$, что при $\|\vec{h}_1\| \leq \rho, \|\vec{h}_2\| \leq \rho$ точка $(a_1 + \vec{h}_1, a_2 + \vec{h}_2) \in \Omega$. При этом

$$\begin{aligned} \overrightarrow{f(a_1 + \vec{h}_1, a_2 + \vec{h}_2) - f(a_1, a_2)} &= \\ = \overrightarrow{f(a_1 + \vec{h}_1, a_2 + \vec{h}_2) - f(a_1 + \vec{h}_1, a_2)} &+ \overrightarrow{f(a_1 + \vec{h}_1, a_2) - f(a_1, a_2)}. \end{aligned} \tag{III, 5; 11}$$

¹⁾ И даже непрерывность f .

Применим следствие 1 теоремы 13 к непрерывно дифференцируемой функции $x_2 \rightarrow f(a_1 + \vec{h}_1, x_2)$ с $L = \frac{\partial f}{\partial x_2}(a_1, a_2)$:

$$\overrightarrow{f(a_1 + \vec{h}_1, a_2 + \vec{h}_2) - f(a_1 + \vec{h}_1, a_2)} = \frac{\partial f}{\partial x_2}(a_1, a_2) \vec{h}_2 + \vec{\omega} \|\vec{h}_2\|, \quad (\text{III}, 5; 12)$$

где

$$\|\vec{\omega}\| \leq \sup_{0 < \theta < 1} \left\| \frac{\partial f}{\partial x_2}(a_1 + \vec{h}_1, a_2 + \theta \vec{h}_2) - \frac{\partial f}{\partial x_2}(a_1, a_2) \right\|. \quad (\text{III}, 5; 13)$$

Если $\vec{h} = (\vec{h}_1, \vec{h}_2)$ стремится к $\vec{0}$, то в силу непрерывности $\frac{\partial f}{\partial x_2}$ в точке (a_1, a_2) последнее выражение стремится к 0.

С другой стороны, по самому определению производного отображения (III, 3; 13)

$$\overrightarrow{f(a_1 + \vec{h}_1, a_2) - f(a_1, a_2)} = \frac{\partial f}{\partial x_1}(a_1, a_2) \vec{h}_1 + \vec{\alpha} \|\vec{h}_1\|, \quad (\text{III}, 5; 14)$$

где $\vec{\alpha}$ стремится к $\vec{0}$ вместе с \vec{h}_1 .

Окончательно получаем

$$\begin{aligned} &\overrightarrow{f(a_1 + \vec{h}_1, a_2 + \vec{h}_2) - f(a_1, a_2)} = \\ &= \left(\frac{\partial f}{\partial x_1}(a_1, a_2) \vec{h}_1 + \frac{\partial f}{\partial x_2}(a_1, a_2) \vec{h}_2 \right) + (\vec{\alpha} \|\vec{h}_1\| + \vec{\omega} \|\vec{h}_2\|), \end{aligned} \quad (\text{III}, 5; 15)$$

где $\vec{\alpha}$ и $\vec{\omega}$ стремятся к $\vec{0}$ вместе с \vec{h} .

Первая скобка определяет линейное непрерывное отображение u пространства $\vec{E}_1 \times \vec{E}_2$ в пространство \vec{F} с помощью линейных отображений u_1 и u_2 пространств \vec{E}_1 и \vec{E}_2 в \vec{F} . Вторая скобка по норме не превосходит $(\|\vec{\alpha}\| + \|\vec{\omega}\|) \|\vec{h}\|$, где $\|\vec{\alpha}\| + \|\vec{\omega}\|$ стремится к 0 вместе с $\|\vec{h}\|$. Следовательно, f имеет полную производную $f'(a) \in \mathcal{L}(\vec{E}_1 \times \vec{E}_2; \vec{F})$ в точке a , определяемую с помощью $\frac{\partial f}{\partial x_1}(a_1, a_2) \in \mathcal{L}(\vec{E}_1; \vec{F})$ и $\frac{\partial f}{\partial x_2}(a_1, a_2) \in \mathcal{L}(\vec{E}_2; \vec{F})$. Покажем, что из непрерывности $\frac{\partial f}{\partial x_1}$ и $\frac{\partial f}{\partial x_2}$ следует непрерывность f' . Если x стремится к x_0 в Ω , то $\left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(x_0) \right\|$ стремится к 0 ($i = 1, 2$), и тогда из второго неравенства (III, 5; 10) следует, что $\|f'(x) - f'(x_0)\|$ стремится к 0, чем и заканчивается доказательство.

Замечания. 1°) Непрерывность $\frac{\partial f}{\partial x_2}$ и только существование $\frac{\partial f}{\partial x_1}$ достаточны для существования f' , как это видно непосредственно из доказательства (но не достаточны для непрерывности f').

2°) Теорема, доказанная для произведения двух пространств E_1 и E_2 , очевидно, справедлива для произведения n пространств E_1, E_2, \dots, E_n (доказать это можно или непосредственно, или по индукции). Кроме того, непрерывность всех частных производных, кроме, может быть, одной, и только существование этой производной достаточны для того, чтобы повлечь за собой существование полной производной.

3°) Пользуясь теми же самыми неравенствами (III, 5; 10), можно показать, что, если функция f зависит от параметра λ , пробегающего некоторое топологическое пространство Λ , и если $\frac{\partial f}{\partial x_1}(x_1, x_2, \lambda)$, $\frac{\partial f}{\partial x_2}(x_1, x_2, \lambda)$ непрерывно зависят от x_1, x_2, λ , то тоже самое верно и для $f'(x_1, x_2, \lambda)$.

Следствие 1. Пусть E, F — аффинные пространства, причем пространство E конечномерно и в нем выбрана система координат $0, (\vec{e}_j)_{j \in J}$. Для того чтобы отображение f открытого множества Ω из E в F было непрерывно дифференцируемым, необходимо и достаточно, чтобы оно имело частные производные $\frac{\partial f}{\partial x_i}$, непрерывные в Ω .

Результат вытекает из того, что введение системы координат отождествляет E с K^J .

Сказанное выше объясняет, почему в математических курсах для применения теоремы о сложной функции (теорема 11) всегда требуют, чтобы рассматриваемые функции имели непрерывные частные производные. В действительности такое предположение слишком сильно, поскольку оно влечет за собой существование и непрерывность полной производной, в то время как требуется установить лишь ее существование. Однако непрерывность частных производных обеспечивает существование полной производной, а одно лишь существование их этого обеспечить не может.

Следствие 2. Если, кроме того, в условиях следствия 1 пространство F конечномерно и в нем выбрана система координат $0, (\vec{f}_i)_{i \in I}$, то отображение f , определенное функциями F_i ($i \in I$) переменных x_j ($j \in J$)¹⁾, непрерывно дифференци-

¹⁾ Обозначения формулы (III, 3; 16).

руемо тогда и только тогда, когда функции F_i имеют непрерывные обычные частные производные $\frac{\partial F_i}{\partial x_j}$.

§ 6. ПРОИЗВОДНЫЕ ВЫСШИХ ПОРЯДКОВ

Пусть f — отображение открытого множества Ω аффинного нормированного пространства E в аффинное нормированное пространство F . Если это отображение всюду дифференцируемо, то его производная функция f' является отображением Ω в векторное нормированное пространство $\mathcal{L}(\vec{E}; \vec{F})$. Естественно изучить, будет ли в свою очередь дифференцируемым это отображение. Если это так, то его производная в точке a , обозначаемая $f''(a)$, будет элементом пространства $\mathcal{L}(\vec{E}; \mathcal{L}(\vec{E}; \vec{F}))$, и если такая производная существует всюду в Ω , то производная функция $f'': x \rightarrow f''(x)$ является отображением Ω в $\mathcal{L}(\vec{E}; \mathcal{L}(\vec{E}; \vec{F}))$.

Пусть теперь \vec{X} — некоторый вектор из \vec{E} . По определению $f''(a)$, $f''(a)\vec{X}$ является элементом $\mathcal{L}(\vec{E}; \vec{F})$, т. е. линейным непрерывным отображением \vec{E} в \vec{F} . Если взять другой вектор \vec{Y} из \vec{E} , то можно говорить о $(f''(a)\vec{X}) \cdot \vec{Y}$, являющемся вектором из \vec{F} . Дадим сейчас иную интерпретацию этому вектору.

Для фиксированного \vec{Y} отображение $u \rightarrow u\vec{Y}$ является линейным непрерывным отображением пространства $\mathcal{L}(\vec{E}; \vec{F})$ в \vec{F} (теорема 54, гл. II). Следовательно, согласно следствию 1 из теоремы 11, это отображение перестановочно с частной производной $D_{\vec{X}}$ вдоль вектора \vec{X} . Иначе говоря:

$$((D_{\vec{X}} f')(a)) \vec{Y} = (D_{\vec{X}} (f'(x) \vec{Y}))_{x=a} \in \vec{F}, \quad (\text{III}, 6; 1)$$

где правая часть существует каждый раз, когда существует левая часть равенства. Однако если существует $f''(a)$, то $D_{\vec{X}} f'(a)$ является не чем иным, как $f''(a)\vec{X}$, а, следовательно, левая часть является рассмотренным выше вектором $(f''(a)\vec{X}) \cdot \vec{Y}$. С другой стороны, $f'(x)\vec{Y}$ — это $\overrightarrow{D_{\vec{Y}} f}(x)$, поэтому правая часть существует и равна $(\overrightarrow{D_{\vec{X}} (D_{\vec{Y}} f)})(a)$.

Окончательно, если существует $f''(a)$, то существует $\overrightarrow{D_X}(D_{\vec{Y}}f)(a)$ и

$$(f''(a)\vec{X}) \cdot \vec{Y} = \overrightarrow{D_X}(D_{\vec{Y}}f)(a). \quad (\text{III}, 6; 2)$$

Правую часть называют также частной производной второго порядка $\overrightarrow{D_X}D_{\vec{Y}}f(a) = \overrightarrow{D_{\vec{X}, \vec{Y}}}^2 f(a)$ (конечно, при частном дифференцировании $D_{\vec{X}}D_{\vec{Y}}$ сначала вычисляют частную производную $D_{\vec{Y}}$, а затем частную производную $D_{\vec{X}}$).

Впрочем, согласно теореме 54₂ гл. II, линейное непрерывное отображение \vec{E} в $\mathcal{L}(\vec{E}; \vec{F})$ можно отождествить с билинейным непрерывным отображением $\vec{E} \times \vec{E}$ в \vec{F} . Согласно определению этого отождествления, это означает, что линейному непрерывному отображению U пространства \vec{E} в $\mathcal{L}(\vec{E}; \vec{F})$ ставится в соответствие билинейное непрерывное отображение u пространства $\vec{E} \times \vec{E}$ в \vec{F} , определенное равенством

$$u(\vec{X}, \vec{Y}) = (U\vec{X}) \cdot \vec{Y}^1. \quad (\text{III}, 6; 3)$$

Вторая производная часто рассматривается именно в этой форме:

В этом случае выражение $(f''(a)\vec{X}) \cdot \vec{Y}$ может быть записано в виде $f''(a)(\vec{X}, \vec{Y})$, $f''(a)$ становится элементом пространства $\mathcal{L}_2(\vec{E}, \vec{E}; \vec{F})$ или $\mathcal{L}_2(\vec{E}^2; \vec{F})$ билинейных непрерывных отображений $\vec{E} \times \vec{E} = \vec{E}^2$ в \vec{F} , и если $f''(x)$ существует для всех $x \in \Omega$, то $f'': x \rightarrow f''(x)$ является функцией, определенной на $\Omega \subset E$, со значениями в $\mathcal{L}_2(\vec{E}^2; \vec{F})$. При этом (III, 6; 2) принимает вид

$$f''(a)(\vec{X}, \vec{Y}) = \overrightarrow{D_X}D_{\vec{Y}}f(a), \quad (\text{III}, 6; 3_2)$$

где правая часть существует, если только существует левая, т. е. существует $f''(a)$ ²⁾.

1) Если $\vec{E}_1 \neq \vec{E}_2$, то имеется лишь единственный возможный способ отождествления $\mathcal{L}(\vec{E}_1; \mathcal{L}(\vec{E}_2; \vec{F}))$ с $\mathcal{L}_2(\vec{E}_1, \vec{E}_2; \vec{F})$. Поскольку здесь $\vec{E}_1 = \vec{E}_2 = \vec{E}$, то возможны два варианта отождествления, и мы выбрали один из них.

2) Напомним еще раз, что существование $f''(a)$ влечет за собой существование f' если и не во всем Ω , то по крайней мере в некоторой окрестности точки a .

Определение. Билинейное отображение u произведения $\vec{E} \times \vec{E}$ в векторное пространство \vec{F} называется *симметричным*, если, каковы бы ни были элементы \vec{X} и \vec{Y} из \vec{E} , мы имеем:

$$u(\vec{X}, \vec{Y}) = u(\vec{Y}, \vec{X}). \quad (\text{III, 6; 4})$$

Теорема 16. Пусть f — отображение открытого множества Ω аффинного нормированного пространства E в аффинное нормированное пространство F . Предположим, что оно имеет вторую производную $f''(a)$ в точке $a \in \Omega$. Тогда $f''(a)$ является билинейным непрерывным симметричным отображением $\vec{E} \times \vec{E}$ в \vec{F} :

$$f''(a)(\vec{X}, \vec{Y}) = f''(a)(\vec{Y}, \vec{X}). \quad (\text{III, 6; 5})$$

Если $E = \mathbb{K}^n$, то, выбирая в качестве \vec{X} и \vec{Y} векторы базиса, получаем известное соотношение

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Доказательство. Рассмотрим выражение

$$\vec{A} = \overrightarrow{f(a + t(\vec{X} + \vec{Y})) - f(a + t\vec{X}) - f(a + t\vec{Y}) + f(a)} \in \vec{F}. \quad (\text{III, 6; 5}_2)$$

Предположим, что $a \in \Omega$, $\vec{X} \in \vec{E}$ и $\vec{Y} \in \vec{E}$ фиксированы, и заставим скаляр $t \neq 0$ стремиться к 0. Предположим также, что $\vec{X} \neq \vec{0}$ и $\vec{Y} \neq \vec{0}$, так как в противном случае (III, 6; 5) очевидно. Поскольку множество Ω открыто, оно содержит некоторый шар с центром в точке a некоторого радиуса $r > 0$. Мы можем ограничиться такими значениями t , $|t| < \frac{r}{\|\vec{X}\| + \|\vec{Y}\|}$, при которых

точки a , $a + t\vec{X}$, $a + t\vec{Y}$, $a + t(\vec{X} + \vec{Y})$ находятся в этом шаре, а, следовательно, и в Ω . В силу выпуклости шара, множеству Ω будет принадлежать весь отрезок, соединяющий эти точки, и мы сможем применить формулу конечных приращений (теорема 13).

Выражение \vec{A} можно записать в виде разности

$$\vec{A} = \vec{g}(\vec{X}) - \vec{g}(\vec{0}), \quad (\text{III, 6; 6})$$

где \vec{g} — отображение из \vec{E} в \vec{F} , определенное в шаре $\{\vec{\xi}; \|\vec{\xi}\| < \|\vec{X}\|\}$ по формуле

$$\vec{g}(\vec{\xi}) = \overrightarrow{f(a + t(\vec{\xi} + \vec{Y})) - f(a + t\vec{\xi})} \in \vec{F} \quad (a, \vec{Y}, t \text{ фиксированы}). \quad (\text{III}, 6; 7)$$

Функция g дифференцируема, и ее производная имеет вид

$$g'(\vec{\xi}) = tf'(a + t(\vec{\xi} + \vec{Y})) - tf'(a + t\vec{\xi}). \quad (\text{III}, 6; 8)$$

К ней можно применить следствие 1 теоремы 13. Надо только выбрать подходящим образом L . Рассуждения будут носить интуитивный характер. Так как

$$\vec{g}(\vec{X}) - \vec{g}(\vec{0}) \approx g'(0) \vec{X} \approx t(f'(a + t\vec{Y}) - f'(a)) \vec{X} \approx t^2(f''(a) \vec{Y}) \cdot \vec{X}, \quad (\text{III}, 6; 9)$$

то положим

$$L = t^2(f''(a) \vec{Y}) \in \mathcal{L}(\vec{E}; \vec{F}). \quad (\text{III}, 6; 10)$$

Тогда формула (III, 5; 8) дает

$$\vec{A} = (\vec{g}(\vec{X}) - \vec{g}(\vec{0})) = L\vec{X} + \vec{\omega} \|\vec{X}\|, \quad (\text{III}, 6; 11)$$

где $\vec{\omega} \in \vec{F}$ оценивается по норме следующим образом:

$$\begin{aligned} \|\vec{\omega}\| &\leq \sup_{\vec{\xi} \in \vec{J}^0, \vec{X}} \|g'(\vec{\xi}) - L\| = \\ &= \sup_{\vec{\xi} \in \vec{J}^0, \vec{X}} \|tf'(a + t(\vec{\xi} + \vec{Y})) - tf'(a + t\vec{\xi}) - t^2(f''(a) \vec{Y})\|. \quad (\text{III}, 6; 12) \end{aligned}$$

Применим к f' определение производной (III, 3; 13):

$$\begin{aligned} f'(a + t(\vec{\xi} + \vec{Y})) &= f'(a) + f''(a)(t(\vec{\xi} + \vec{Y})) + \alpha \|t(\vec{\xi} + \vec{Y})\|, \\ f'(a + t\vec{\xi}) &= f'(a) + f''(a)t\vec{\xi} + \beta \|t\vec{\xi}\|, \end{aligned} \quad (\text{III}, 6; 13)$$

где при фиксированных a, \vec{X}, \vec{Y} и $\vec{\xi} \in \vec{J}^0, \vec{X}$ функции $\alpha = \alpha(t, \vec{\xi})$ и $\beta = \beta(t, \vec{\xi})$ стремятся равномерно к 0 в $\mathcal{L}(\vec{E}; \vec{F})$ при t , стремящемся к 0. Подставляя в (III, 6; 12), получаем

$$\begin{aligned} \|\vec{\omega}\| &\leq t^2 \sup_{\vec{\xi} \in \vec{J}^0, \vec{X}} \|\alpha(\vec{\xi} + \vec{Y}) + \beta\vec{\xi}\| \leq \\ &\leq \sup_{\vec{\xi} \in \vec{J}^0, \vec{X}} t^2 [(\|\alpha\| + \|\beta\|)(\|\vec{X}\| + \|\vec{Y}\|)], \quad (\text{III}, 6; 14) \end{aligned}$$

где последняя скобка стремится к 0 вместе с t . Учитывая значение L , из (III, 6; 11) теперь получаем, что $\frac{\vec{A}}{t^2}$ стремится к $(f''(a)\vec{Y}) \cdot \vec{X} = f''(a)(\vec{Y}, \vec{X})$ при t , стремящемся к 0.

Так как \vec{A} симметрично относительно \vec{X} и \vec{Y} , то таким же будет предел $\frac{\vec{A}}{t^2}$, чем и заканчивается доказательство теоремы.

Замечания. 1°) В видоизмененной форме

$$\overrightarrow{D_{\vec{X}} D_{\vec{Y}} f}(a) = \overrightarrow{D_{\vec{Y}} D_{\vec{X}} f}(a) \quad (\text{III, 6; 17})$$

получаем соотношение, в которое входит лишь сужение f на пересечение множества Ω с линейным многообразием, проходящим через a параллельно векторам \vec{X}, \vec{Y} . Поэтому нет необходимости предполагать само f дважды дифференцируемым. Достаточно предполагать, что сужение f на это многообразие имеет вторую производную в точке a .

2°) Имеется вариант этой теоремы, в котором существование $f''(a)$ не предполагается, но в котором говорится, что если $\overrightarrow{D_{\vec{X}} D_{\vec{Y}} f}$ и $\overrightarrow{D_{\vec{Y}} D_{\vec{X}} f}$ существуют в каждой окрестности точки a и непрерывны в a , то они равны в a .

Такая формулировка не сильнее и не слабее того, что мы привели, но имеется различие: если мы возьмем для простоты $\vec{E} = \mathbb{R}^2$, то существование полной второй производной в точке влечет за собой существование $\frac{\partial^2 f}{\partial x \partial y}$ и $\frac{\partial^2 f}{\partial y \partial x}$ в этой точке, но из него не следует ни их существование в ее окрестности, ни их непрерывность. Обратно, непрерывность этих частных производных недостаточна для того, чтобы обеспечить существование второй полной производной (включающей в себя еще частные производные $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}$). Известно, что одно только предположение о существовании производных $\frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}$ недостаточно для того, чтобы обеспечить их равенство.

Последовательные производные

Точно таким же образом определяются производные более высоких порядков. Производная порядка m в точке $a \in \Omega$ (которая определена только тогда, когда все производные порядков $\leq m - 1$ существуют во всей окрестности точки a) может

быть отождествлена с некоторым m -линейным отображением \vec{E}^m в \vec{F} . Если $f^{(m)}(x)$ существует для каждого $x \in \Omega$, то $f^{(m)}: x \rightarrow f^{(m)}(x)$ является отображением множества Ω в пространство $\mathcal{L}_m(\vec{E}^m; \vec{F})$ этих m -линейных непрерывных отображений. Точнее, $f^{(m)}$ определяется следующим образом. Пусть $f^{(m-1)}$ — производная порядка $m-1$, отождествляемая с некоторой функцией на Ω со значениями в $\mathcal{L}_{m-1}(\vec{E}^{m-1}; \vec{F})$. Тогда ее производная в точке a : $(f^{(m-1)})'(a) = f^{(m)}(a)$ является линейным непрерывным отображением \vec{E} в $\mathcal{L}_{m-1}(\vec{E}^{m-1}; \vec{F})$; оно отождествляется с некоторым элементом из $\mathcal{L}_m(\vec{E}^m; \vec{F})$, обозначаемым через $f^{(m)}(a)$, так что

$$f^{(m)}(a)(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) = ((f^{(m-1)})'(a)) \cdot (\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m). \quad (\text{III}, 6; 18)$$

По индукции теперь доказывается, что:

1°) Если $p \leq m$ и если рассматривать $f^{(p)}$ как функцию на Ω со значениями в $\mathcal{L}_p(\vec{E}^p; \vec{F})$, то ее производная $(m-p)$ -го порядка в a , если она существует, является элементом пространства $\mathcal{L}_{m-p}(\vec{E}^{m-p}; \mathcal{L}_p(\vec{E}^p; \vec{F}))$. Этот элемент можно отождествлять с некоторым элементом $f^{(m)}(a)$ из $\mathcal{L}_m(\vec{E}^m; \vec{F})$, если он существует. Оба этих элемента существуют одновременно, и при этом

$$\begin{aligned} (f^{(p)})^{(m-p)}(a) &= f^{(m)}(a) \in \mathcal{L}_m(\vec{E}^m; \vec{F}), \\ [((f^{(p)})^{(m-p)}(a))(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{m-p})] \cdot (\vec{X}_{m-p+1}, \dots, \vec{X}_m) &= \\ &= f^{(m)}(a)(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) \in \vec{F}. \quad (\text{III}, 6; 19) \end{aligned}$$

2°) Если существует $f^{(m)}(a)$, то существует

$$\overrightarrow{D_{\vec{X}_1} D_{\vec{X}_2} \dots D_{\vec{X}_m}} f(a) \in \vec{F} \text{ и}$$

$$f^{(m)}(a)(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) = \overrightarrow{D_{\vec{X}_1} D_{\vec{X}_2} \dots D_{\vec{X}_m}} f(a) \in \vec{F}. \quad (\text{III}, 6; 20)$$

Определение. m -линейное отображение u пространства $\vec{E} \times \vec{E} \times \dots \times \vec{E} = \vec{E}^m$ в пространство \vec{F} называется *симметричным*, если при любой перестановке $\sigma: k \rightarrow \sigma_k$ множества $\{1, 2, \dots, m\}$ и любой системы векторов $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m$ из \vec{E} имеет место формула

$$u(\vec{X}_{\sigma_1}, \vec{X}_{\sigma_2}, \dots, \vec{X}_{\sigma_m}) = u(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m). \quad (\text{III}, 6; 21)$$

Теорема 16 теперь может быть обобщена на производные произвольных порядков:

Теорема 16₂. Если отображение f множества $\Omega \subset E$ в пространство F имеет производную порядка m в точке $a \in \Omega$, то эта производная $f^{(m)}(a)$ является m -линейным непрерывным симметричным отображением \vec{E}^m в пространство \vec{F} , т. е. для любой перестановки σ множества $\{1, 2, \dots, m\}$ и любой системы векторов $(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) \in \vec{E}^m$ имеют место равенства

$$f^{(m)}(a)(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) = f^{(m)}(a)(\vec{X}_{\sigma_1}, \vec{X}_{\sigma_2}, \dots, \vec{X}_{\sigma_m}),$$

$$\overrightarrow{D_{\vec{X}_1} D_{\vec{X}_2} \dots D_{\vec{X}_m} f}(a) = \overrightarrow{D_{\vec{X}_{\sigma_1}} D_{\vec{X}_{\sigma_2}} \dots D_{\vec{X}_{\sigma_m}} f}(a). \quad (\text{III}, 6; 22)$$

Доказательство. Эта теорема была доказана для производных второго порядка. Предположим, что она справедлива для производных порядка $\leq m - 1$, и докажем ее для производных порядка $m \geq 3$. Предположим прежде всего, что $\sigma_1 = 1$ и что σ переставляет только числа $2, 3, \dots, m$. Тогда функции $\overrightarrow{D_{\vec{X}_2} \dots D_{\vec{X}_m} f}$ и $\overrightarrow{D_{\vec{X}_{\sigma_2}} \dots D_{\vec{X}_{\sigma_m}} f}$, определенные в Ω , со значениями в \vec{F} по предположению индукции совпадают. Следовательно, их производные в a вдоль вектора $\vec{X}_1 = \vec{X}_{\sigma_1}$ одинаковы и мы имеем равенство (III, 6; 22).

Предположим теперь, что $\sigma_1 = 2, \sigma_2 = 1$ и что σ сохраняет каждое из чисел $3, 4, \dots, m$. Тогда $\overrightarrow{D_{\vec{X}_3} \dots D_{\vec{X}_m} f} = \overrightarrow{D_{\vec{X}_{\sigma_3}} \dots D_{\vec{X}_{\sigma_m}} f}$ является функцией, определенной в Ω и имеющей первую производную всюду и вторую производную в точке a . Следовательно, к ней можно применить теорему 16 относительно производных вдоль \vec{X}_1 и \vec{X}_2 и снова получить (III, 6; 22).

Так как каждая перестановка чисел $\{1, 2, \dots, m\}$ является композицией перестановок этих двух типов (это очевидно, если $\sigma_1 = 1$; в любой перестановке можно от $1, 2, \dots, m$ перейти к $1, \sigma_1, \dots, m$, затем от $1, \sigma_1, \dots, m$ к $\sigma_1, 1, \dots, m$ и, наконец, от $\sigma_1, 1, \dots, m$ к $\sigma_1, \sigma_2, \dots, \sigma_m$), то теорема будет справедливой в общем случае.

Упражнение. Шаг за шагом можно доказать следующее. Пусть L есть m -линейное непрерывно симметричное отображение \vec{E}^m в \vec{F} . Оно определяет некоторое p -линейное

непрерывное симметричное отображение L_p пространства \vec{E}^p в пространство $\mathcal{L}_{m-p}(\vec{E}^{m-p}; \vec{F})$.

Тогда k -я производная функция «одночленной» функции $\vec{x} \rightarrow L\vec{x}^m$, действующей из \vec{E} в \vec{F} , является функцией вида¹⁾

$$\vec{x} \rightarrow m(m-1) \dots (m-k+1) L_{m-k} \vec{x}^{m-k} \in \mathcal{L}_k(\vec{E}^k; \vec{F}). \quad (\text{III}, 6; 22_2)$$

Полагая $E = F = \mathbb{K}$, получаем элементарную формулу дифференцирования одночленов.

З а м е ч а н и е. Теорема 8₃ (о линейности производной), теорема 8₄ (о дифференировании функции со значениями в произведении пространств) и следствие 1 теоремы 11 (о перестановочности производной и линейного непрерывного отображения) с помощью индукции по m непосредственно обобщаются на производные порядка m .

Случай, когда $E = \mathbb{K}^n$. Рассмотрим наиболее часто встречающийся на практике случай, когда пространство E конечномерно и в нем выбрана система координат $0, (\vec{e}_i)_{i \in I}$. Тогда для второй полной производной при $\vec{X} = \sum_{i \in I} X_i \vec{e}_i$ и $\vec{Y} = \sum_{i \in I} Y_i \vec{e}_i$ имеем следующую формулу:

$$f''(a)(\vec{X}, \vec{Y}) = \sum_{i, j \in I} X_i Y_j f''(a)(e_i, e_j) = \sum_{i, j \in I} X_i Y_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \in \vec{F}. \quad (\text{III}, 6; 23)$$

Для третьей производной получаем

$$f'''(a)(\vec{X}, \vec{Y}, \vec{Z}) = \sum_{i, j, k \in I} X_i Y_j Z_k \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a). \quad (\text{III}, 6; 24)$$

Более общо: если $I = \{1, 2, \dots, m\}$, то

$$\begin{aligned} f^{(m)}(a)(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) &= \\ &= \sum_j \frac{\partial^m f}{\partial x_{j_1} \partial x_{j_2} \dots \partial x_{j_m}}(a) X_{1, j_1} X_{2, j_2} \dots X_{m, j_m} \in \vec{F}, \end{aligned} \quad (\text{III}, 6; 25)$$

¹⁾ L_m и L_0 по условию совпадают с L , а \vec{x}^0 есть $1 \in \mathbb{K}$. Через $L_p \vec{x}^p$ будем обозначать $L_p(\underbrace{\vec{x}, \vec{x}, \dots, \vec{x}}_p)$.

где $X_{\alpha, \beta}$ является β -й координатой вектора \vec{X}_{α} ($\alpha = 1, 2, \dots, m$; $\beta = 1, 2, \dots, n$) и где j пробегает множество всевозможных отображений $\alpha \rightarrow j_{\alpha}$ множества $\{1, 2, \dots, m\}$ в множество $\{1, 2, \dots, n\}$.

Если $E = \mathbb{K}$ — поле скаляров, то производный вектор порядка m , $\overrightarrow{f^{(m)}}(a) \in \vec{F}$, определенный на стр. 208, связан с производным отображением порядка m , $f^{(m)}(a) \in \mathcal{L}_m(\mathbb{K}^m; \vec{F})$, формулой

$$\overrightarrow{f^{(m)}}(a) = f^{(m)}(a)(1, 1, \dots, 1). \quad (\text{III}, 6; 25_2)$$

Для частных производных можно ввести сокращенные обозначения, объединив одинаковые производные $\frac{\partial}{\partial x_k}$. Пусть \vec{p}^1 — некоторый элемент из \mathbb{N}^n , т. е. система n целых чисел $p_1, p_2, \dots, p_n \geq 0$. Тогда через \vec{D}^p или $\left(\frac{\partial}{\partial x}\right)^{\vec{p}}$ будем обозначать частную производную

$$D^{\vec{p}} = \left(\frac{\partial}{\partial x_1}\right)^{p_1} \left(\frac{\partial}{\partial x_2}\right)^{p_2} \dots \left(\frac{\partial}{\partial x_m}\right)^{p_m}. \quad (\text{III}, 6; 26)$$

Например, если $n = 3$, $\vec{p} = (1, 0, 2)$, то $D^{\vec{p}}$ будет означать производную $\frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_3}\right)^2$. Если $\vec{p} = \vec{0} = (0, 0, \dots, 0)$, то $D^{\vec{p}}f = f$. Величина $|\vec{p}| = p_1 + p_2 + \dots + p_n$ называется порядком \vec{p} или порядком частной производной. Если \vec{q} — некоторый другой элемент из \mathbb{N}^n , то через $\vec{p} + \vec{q}$ обозначается элемент $(p_1 + q_1, p_2 + q_2, \dots, p_n + q_n)$, так что $\overrightarrow{D^{p+q}f} = \overrightarrow{D^p}(\overrightarrow{D^q}f)$.

Говорят, что $\vec{p} \geq \vec{q}$, если $p_1 \geq q_1, p_2 \geq q_2, \dots, p_n \geq q_n$ (это некоторое отношение порядка в \mathbb{N}^n). В этом случае через $\vec{p} - \vec{q}$ обозначается элемент $(p_1 - q_1, p_2 - q_2, \dots, p_n - q_n)$ и $\vec{p} = \vec{q} + (\vec{p} - \vec{q})$.

¹⁾ Хотя \mathbb{N}^n не является векторным пространством, удобнее писать \vec{p} вместо p для того, чтобы напомнить, что это не целое число, а система n целых чисел.

В дальнейшем нам понадобятся также следующие обозначения:

$$\begin{aligned} \vec{p}! &= p_1! p_2! \dots p_n!, \\ \left(\begin{array}{c} \vec{p} \\ \vec{q} \end{array}\right) &= \left(\begin{array}{c} p_1 \\ q_1 \end{array}\right) \left(\begin{array}{c} p_2 \\ q_2 \end{array}\right) \dots \left(\begin{array}{c} p_n \\ q_n \end{array}\right)^1) \quad \text{для } \vec{q} \leqslant \vec{p}, \\ \vec{X}^{\vec{p}} &= X_1^{p_1} X_2^{p_2} \dots X_n^{p_n} \in K \quad \text{для } \vec{X} \in K^n, \vec{p} \in N^n. \end{aligned} \quad (\text{III}, 6; 27)$$

Определение. Говорят, что функция f является t раз непрерывно дифференцируемой, или принадлежит классу C^m , в Ω , если она имеет производные порядков $\leqslant m$, непрерывные в Ω . Каждое t раз дифференцируемое в Ω отображение принадлежит по крайней мере классу C^{m-1} .

Случай произведения пространств.

Полная и частная дифференцируемость

Теорема 15 может быть обобщена следующим образом:

Теорема 17. Для того чтобы отображение f некоторого открытого множества Ω произведения $E_1 \times E_2 \times \dots \times E_n$ в F принадлежало классу C^m , необходимо и достаточно, чтобы оно имело непрерывные в Ω частные производные порядков $\leqslant m$.

Доказательство. 1°) Поскольку каждый элемент u из $\mathcal{L}(E; F)$ эквивалентен системе элементов u_i пространств $\mathcal{L}(E_i; F)$, $i = 1, 2, \dots, n$, то тем самым определяется взаимно однозначное соответствие между $\mathcal{L}(E; F)$ и произведением пространств $\mathcal{L}(E_i; F)$, сохраняющее векторную структуру.

С другой стороны, из неравенств $\|u_i\| \leqslant \|u\| \leqslant \sum_{j=1}^n \|u_j\|$, $i = 1, 2, \dots, n$, следует, что норме в пространстве $\mathcal{L}(E; F)$ соответствует в произведении пространств $\mathcal{L}(E_i; F)$ некоторая норма, эквивалентная одной из их естественных норм. Поэтому, согласно замечанию на стр. 264 (обобщение теоремы 84), отображение множества $\Omega \subset E$ в пространство $\mathcal{L}(\vec{E}; \vec{F})$ принадлежит классу C^k тогда и только тогда, когда отображения множества Ω в пространства $\mathcal{L}(\vec{E}_i; \vec{F})$, определяющие это отображение, принадлежат классу C^k .

2°) Теорема 17 верна для $m = 1$ (теорема 15). Докажем ее справедливость для любого m методом математической ин-

¹⁾ Напомним, что $\binom{r}{s} = C_r^s = \frac{r!}{s!(r-s)!}$.

дукции. Предположим, что она верна для $\dot{m} \geq 2$. Пусть f принадлежит классу C^m . Тогда f' принадлежит классу C^{m-1} . Поскольку f' является отображением Ω в $\mathcal{L}(\vec{E}; \vec{F})$, определенным отображениями $\partial f / \partial x_i$ множества Ω в $\mathcal{L}(\vec{E}_i; \vec{F})$, каждое отображение $\partial f / \partial x_i$, согласно 1°), принадлежит классу C^{m-1} . По индуктивному предположению отображение $\partial f / \partial x_i$ имеет непрерывные в Ω частные производные порядка $\leq m-1$ и, следовательно, отображение f имеет непрерывные в Ω частные производные порядка $\leq m$. Обратно, если отображение f имеет частные производные порядка $\leq m$, непрерывные в Ω , то отображения $\partial f / \partial x_i$ существуют и имеют частные производные порядка $\leq m-1$, непрерывные в Ω . Поэтому, согласно индуктивному предположению, они принадлежат классу C^{m-1} .

В частности, эти отображения непрерывны, а, следовательно, согласно теореме 15, существует f' . Это — функция в Ω со значениями в $\mathcal{L}(\vec{E}; \vec{F})$, определенная при помощи функций $\partial f / \partial x_i$ на Ω со значениями в $\mathcal{L}(\vec{E}_i; \vec{F})$. Поскольку $\partial f / \partial x_i$ принадлежат классу C^{m-1} , то f' , согласно 1°), принадлежит классу C^{m-1} и, следовательно, отображение f принадлежит классу C^m .

Пространства m раз дифференцируемых функций

Обобщая понятия, введенные в конце § 3, мы будем обозначать через $(F^\Omega)_{b; m}$ (соответственно через $(F^\Omega)_{cb; m}$) пространство m раз дифференцируемых (соответственно m раз непрерывно дифференцируемых) ограниченных вместе с каждой производной порядка $\leq m$ отображений множества Ω в F .

Это пространство является аффинным, имеющим присоединенное векторное пространство $(\vec{F}^\Omega)_{b; m}$ (соответственно $(\vec{F}^\Omega)_{cb; m}$). Его можно считать аффинным нормированным пространством, если в присоединенном векторном пространстве ввести норму

$$\|\vec{f}\|_m = \sup_{\substack{x \in \Omega \\ 0 \leq k \leq m}} (\|f^{(k)}(x)\|), \quad (\text{III}, 6; 27_2)$$

где $\|f^{(k)}(x)\|$ — норма в $\mathcal{L}_k(\vec{E}^k; \vec{F})$.

Отсюда видно, зачем введен индекс 1 в обозначениях $(F^\Omega)_{b; 1}$ и $(F^\Omega)_{cb; 1}$. В этих обозначениях $(F^\Omega)_b$ (соответственно $(F^\Omega)_{cb}$) можно записать в виде $(F^\Omega)_{b; 0}$ (соответственно $(F^\Omega)_{cb; 0}$) и $\|\vec{f}\|$ в виде $\|\vec{f}\|_0$.

Теорема 17₂. Отображение, ставящее в соответствие каждой функции f ее производную функцию порядка $k \leq m$, есть линейное непрерывное отображение пространства $(\vec{F}^\Omega)_{b; m}$ (соответственно $(\vec{F}^\Omega)_{cb; m}$) в пространство $((\mathcal{L}_k(\vec{E}^k; \vec{F}))^\Omega)_{b; m-k}$ соответственно $((\mathcal{L}_k(\vec{E}^k; \vec{F}))^\Omega)_{cb; m-k}$ с нормой ≤ 1 .

Несмотря на несколько пугающую формулировку, утверждение это очевидно.

Очень часто в предыдущих пространствах рассматривают другие эквивалентные нормы, например определяют $\|\vec{f}\|_m$ как $\sum_{k=0}^m (\sup_{x \in \Omega} \|f^{(k)}(x)\|)$. Если пространство E n -мерно и если в нем выбрана система координат, то в качестве нормы часто берут числа

$$\|\vec{f}\|_m = \sup_{\substack{x \in \Omega \\ |\vec{p}| \leq m}} \overrightarrow{\|D^{\vec{p}}f(x)\|} \quad \text{или} \quad \|\vec{f}\|_m = \sum_{|\vec{p}| \leq m} (\sup_{x \in \Omega} \|D^{\vec{p}}f(x)\|).$$

Производная произведения (формула Лейбница)

Теорема 18. Пусть E — аффинное нормированное пространство, $\vec{F}_1, \vec{F}_2, \vec{G}$ — векторные нормированные пространства. Пусть B — билинейное непрерывное отображение $F_1 \times \vec{F}_2$ в \vec{G} . Если \vec{u}_1 (соответственно \vec{u}_2) является t раз дифференцируемым или t раз непрерывно дифференцируемым отображением открытого множества $\Omega \subset E$ в F_1 (соответственно в F_2), то таким же будет отображение $B(\vec{u}_1, \vec{u}_2)$: $x \rightarrow B(\vec{u}_1(x), \vec{u}_2(x))$ множества Ω в G . Производные этой функции вычисляются по формуле Лейбница.

Доказательство. Воспользуемся обозначениями теоремы 9₂. Тогда формула (III, 4; 22) запишется в виде

$$(B(u_1, u_2))'(x) \vec{X} = B_{\vec{u}_2(x)}(u'_1(x) \vec{X}) + B_{\vec{u}_1(x)}(u'_2(x) \vec{X}) \in \vec{G}, \quad (\text{III, 6; 28})$$

и, следовательно,

$$(B(u_1, u_2))'(x) = B_{\vec{u}_2(x)} \circ u'_1(x) + B_{\vec{u}_1(x)} \circ u'_2(x) \in \mathcal{L}(\vec{E}; \vec{G}). \quad (\text{III, 6; 29})$$

Воспользуемся теперь индукцией по m . При $m = 1$ теорема верна (теорема 12). Предполагая ее справедливой для производной порядка $m - 1$, докажем, что она верна для производной порядка m .

Пусть \vec{u}_1 и \vec{u}_2 m раз дифференцируемы или m раз непрерывно дифференцируемы. Тогда $x \rightarrow \vec{u}'_1(x)$ есть $m-1$ раз дифференцируемая или непрерывно дифференцируемая функция со значениями в $\mathcal{L}(\vec{E}; \vec{F})$. С другой стороны, функция $x \rightarrow \vec{B}_{\vec{u}_2(x)}$, определенная на Ω , со значениями в $\mathcal{L}(\vec{F}_1; \vec{G})$ является композицией m раз дифференцируемой или непрерывно дифференцируемой функции $x \rightarrow \vec{u}_2(x)$ из множества Ω в пространство \vec{F}_2 и линейного непрерывного отображения $\vec{v}_2 \rightarrow \vec{B}_{\vec{v}_2}$ пространства \vec{F}_2 в пространство $\mathcal{L}(\vec{F}_1; \vec{G})$. В силу перестановочности производной и линейного непрерывного отображения (следствие 1 теоремы 11), функция $x \rightarrow \vec{B}_{\vec{u}_2(x)}$ со значениями в $\mathcal{L}(\vec{F}_1; \vec{G})$ является m раз дифференцируемой или непрерывно дифференцируемой. Поскольку $(U_1, U_2) \rightarrow U_2 \circ U_1$ является билinearным непрерывным отображением пространства $\mathcal{L}(\vec{E}; \vec{F}_1) \times \mathcal{L}(\vec{F}_1; \vec{G})$ в пространство $\mathcal{L}(\vec{E}; \vec{G})$, индуктивное предположение позволяет утверждать, что $x \rightarrow \vec{B}_{\vec{u}_2(x)} \circ \vec{u}'_1(x)$ является $m-1$ раз дифференцируемым или непрерывно дифференцируемым отображением Ω в $\mathcal{L}(\vec{E}; \vec{G})$. Поскольку то же самое имеет место для отображения $x \rightarrow \vec{B}_{\vec{u}_1(x)} \circ \vec{u}'_2(x)$, функция $(B(u_1, u_2))'$ является $m-1$ раз дифференцируемой или непрерывно дифференцируемой, а, значит, m раз дифференцируема или непрерывно дифференцируема функция $B(u_1, u_2)$ ¹⁾.

В общем случае формула Лейбница является бесполезным усложнением полученного результата; поэтому мы ее приведем

¹⁾ Обычно E конечномерно, и в нем можно выбрать некоторую систему координат. По теореме 17, для того чтобы показать, что $B(u_1, u_2)$ принадлежит классу C^m , когда u_1 и u_2 принадлежат этому классу, достаточно доказать, что их частные производные по x_1, x_2, \dots, x_n порядка $\leq m$ непрерывны. Но это легко проверяется по индукции. Утверждение верно при $m = 1$. Предполагая его справедливость для $m - 1$, докажем, что оно верно для

$m \geq 2$. Имеем: $\frac{\partial}{\partial x_i} (B(\vec{u}_1, \vec{u}_2)) = B\left(\frac{\partial \vec{u}_1}{\partial x_i}, \vec{u}_2\right) + B\left(\vec{u}_1, \frac{\partial \vec{u}_2}{\partial x_i}\right)$. Здесь $\frac{\partial \vec{u}_1}{\partial x_i}$ и \vec{u}_2 имеют непрерывные частные производные порядка $\leq m-1$, а, следовательно, по индуктивному предположению такой же будет $B\left(\frac{\partial \vec{u}_1}{\partial x_i}, \vec{u}_2\right)$.

То же самое можно сказать и о функции $B\left(\vec{u}_1, \frac{\partial \vec{u}_2}{\partial x_i}\right)$. Но тогда $B(u_1, u_2)$ имеет непрерывные частные производные порядка $\leq m$.

лишь для частного случая конечномерного пространства E с системой координат $0, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. Если воспользоваться обозначением (III, 6; 26), то формулу Лейбница можно записать в виде

$$D^{\vec{p}}(B(u_1, u_2)) = \sum_{\substack{\vec{q} \leqslant \vec{p} \\ q \leqslant p}} \binom{\vec{p}}{\vec{q}} B(\overrightarrow{D^q u_1}, \overrightarrow{D^{p-q} u_2}). \quad (\text{III, 6; 36})$$

Наиболее важен тот случай, когда $\vec{F}_1 = \vec{F}_2 = \vec{G} = \mathbb{K}$ — поле скаляров и $B(u_1, u_2) = u_1 u_2$ — обычное произведение. Тогда

$$D^{\vec{p}} u_1 u_2 = \sum_{\substack{\vec{q} \leqslant \vec{p} \\ q \leqslant p}} \binom{\vec{p}}{\vec{q}} D^{\vec{q}} u_1 D^{\vec{p}-\vec{q}} u_2. \quad (\text{III, 6; 36}_2)$$

Применяемые здесь обозначения таковы, что в случае n -мерного пространства формула сохраняет точно такой же вид, как и в случае функции одной переменной.

Доказать эту формулу очень просто. В случае одной переменной ($n = 1$) она записывается в виде

$$(u_1 u_2)^{(m)} = \sum_{k \leqslant m} \binom{m}{k} u_1^{(k)} u_2^{(m-k)}. \quad (\text{III, 6; 36}_3)$$

Эта формула известна для $m = 1$ и по индукции легко доказывается для любого m . В самом деле, предположим, что она верна для $m - 1$. Тогда

$$\begin{aligned} (u_1 u_2)^{(m)} &= ((u_1 u_2)^{(m-1)})' = \left(\sum_{k' \leqslant m-1} \binom{m-1}{k'} u_1^{(k')} u_2^{(m-k'-1)} \right)' = \\ &= \sum_{k' \leqslant m-1} \left[\binom{m-1}{k'} u_1^{(k'+1)} u_2^{(m-k'-1)} + \binom{m-1}{k'} u_1^{(k')} u_2^{(m-k')} \right] = \\ &= \sum_{k \leqslant m} u_1^{(k)} u_2^{(m-k)} \left(\binom{m-1}{k-1} + \binom{m-1}{k} \right) = \sum_{k \leqslant m} \binom{m}{k} u_1^{(k)} u_2^{(m-k)}. \end{aligned}$$

Теперь проведем индукцию по числу n переменных. Предположим, что формула доказана для $n - 1$ переменных, и докажем ее справедливость для n . Имеем:

$$D^{\vec{p}}(u_1 u_2) = \left(\frac{\partial}{\partial x_1} \right)^{p_1} D^{\vec{p}'}(u_1 u_2), \quad \text{где } \vec{p}' = (p_2, \dots, p_n) \in \mathbb{N}^{n-1}.$$

Функция $D^{\vec{p}'}(u_1 u_2)$ для фиксированного x_1 является производной по $n - 1$ переменным x_2, x_3, \dots, x_n , в то время как $\left(\frac{\partial}{\partial x_1}\right)^{p_1}$ — производная по первой переменной. Поэтому

$$\begin{aligned} D^{\vec{p}}(u_1 u_2) &= \left(\frac{\partial}{\partial x_1}\right)^{p_1} \sum_{q' \leqslant p'} \binom{\vec{p}'}{\vec{q}'} D^{\vec{q}'} u_1 D^{\vec{p}' - \vec{q}'} u_2 = \\ &= \sum_{q_1 \leqslant p_1} \sum_{q' \leqslant p'} \binom{p_1}{q_1} \binom{\vec{p}'}{\vec{q}'} \left(\frac{\partial}{\partial x_1}\right)^{q_1} D^{\vec{q}'} u_1 \left(\frac{\partial}{\partial x_1}\right)^{p_1 - q_1} D^{\vec{p}' - \vec{q}'} u_2 = \\ &= \sum_{q \leqslant p} \binom{\vec{p}}{\vec{q}} D^{\vec{q}} u_1 D^{\vec{p} - \vec{q}} u_2. \end{aligned}$$

Теорема 19. Пусть E, F, G — три аффинных нормированных пространства, Ω и Ω' — открытые множества в E и F соответственно. Пусть f — отображение Ω в Ω' , а g — отображение Ω' в G . Если f и g являются m раз дифференцируемыми или m раз непрерывно дифференцируемыми, то таким же будет сложное отображение $h = g \circ f$.

Доказательство. Теорема верна для производной 1-го порядка (теорема 11 и ее следствие 5). Предполагая ее справедливой для производных порядка $\leqslant m - 1$, докажем, что она верна для производных порядка m , $m \geqslant 2$.

Перепишем формулу (III, 4; 1):

$$h'(x) = g'(f(x)) \circ f'(x). \quad (\text{III, 6; 37})$$

Согласно предположению, функция f дифференцируема m раз, а, следовательно, функция f' дифференцируема $m - 1$ раз. С другой стороны, так как функции f и g' обе $m - 1$ раз дифференцируемы, то функция $g' \cdot f: x \rightarrow g'(f(x))$ также $m - 1$ раз дифференцируема. Но тогда отображение $x \rightarrow f'(x)$ множества Ω в пространство $\mathcal{L}(\vec{E}; \vec{F})$ и отображение $x \rightarrow g'(f(x))$ множества Ω в пространство $\mathcal{L}(\vec{F}; \vec{G})$ дифференцируемы $m - 1$ раз, а $(u, v) \rightarrow v \circ u$ является билинейным непрерывным отображением $\mathcal{L}(\vec{E}; \vec{F}) \times \mathcal{L}(\vec{F}; \vec{G})$ в $\mathcal{L}(\vec{E}; \vec{G})$ (теорема 54 гл. II). Из теоремы 18 следует, что $x \rightarrow g'(f(x)) \circ f'(x)$ также $m - 1$ раз дифференцируема. Таким образом, отображение h' дифференцируемо $m - 1$ раз, а это означает, что отображение h дифференцируемо m раз.

Замечания. 1°) Теоремы 18 и 19 можно было бы доказать, проводя индукцию одновременно для обеих теорем.

В самом деле, теорема 19 для числа m опирается на утверждение теоремы 18 только для $m = 1$. С другой стороны, при доказательстве теоремы 18 для $m = 1$ следует учесть, что функция $x \rightarrow B(u_1(x), u_2(x))$ является композицией функции $x \rightarrow (u_1(x), u_2(x))$ класса C^{m-1} и билинейной непрерывной, а, следовательно, бесконечно дифференцируемой функции $(u_1, u_2) \rightarrow B(u_1, u_2)$, и, значит, можно воспользоваться теоремой 19 для $m = 1$.

Далее, теоремы 18 и 19 доказаны для $m = 1$. Отсюда можно получить теорему 19 для $m = 2$, а, следовательно, теорему 18 для $m = 2$, затем теорему 19 для $m = 3$, а, следовательно, теорему 18 для $m = 3$, и т. д.

2°) Вычислять последовательные производные функции $h = g \circ f$ можно по явным формулам, но они довольно сложны. Так, например, если \vec{X} и \vec{Y} — векторы из \vec{E} , то

$$\begin{aligned} h''(x)(\vec{X}, \vec{Y}) &= g''(f(x))(f'(x)\vec{X}, f'(x)\vec{Y}) + \\ &\quad + g'(f(x))(f''(x)(\vec{X}, \vec{Y})) \in \vec{G}. \end{aligned} \quad (\text{III}, 6; 38)$$

Если E, F, G — поля скаляров \mathbb{K} и если выбрано $X = Y = 1$, то получается элементарная формула

$$h''(x) = g''(f(x))f'^2(x) + g'(f(x))f''(x). \quad (\text{III}, 6; 39)$$

Иногда полезно иметь выражение производной порядка m для функции $h = g \circ f$ в частном случае $E = F = G = \mathbb{K}$. Мы его приведем в теореме 21.

§ 7. ФОРМУЛА ТЕЙЛОРА. МАКСИМУМ И МИНИМУМ

Так как производная $f^{(m)}(a)$ порядка m является m -линейным отображением, то можно вычислить ее значение на системе векторов $(\vec{X}, \vec{X}, \dots, \vec{X}) \in \vec{E}^m$, определенной одним вектором \vec{X} из \vec{E} . Выражение $f^{(m)}(a)(\vec{X}, \vec{X}, \dots, \vec{X})$ удобно записывать кратко в виде

$$f^{(m)}(a)\vec{X}^m = f^{(m)}(a)(\vec{X}, \vec{X}, \dots, \vec{X}). \quad (\text{III}, 7; 1)$$

Имеется столько формул Тейлора, сколько существует формул конечных приращений.

Теореме 13А соответствует формула Тейлора, обобщающая теорему 6:

Теорема 20А. Пусть f — вещественная m раз дифференцируемая функция, определенная на открытом множестве $\Omega \subset E$. Предположим, что замкнутый отрезок $[x, x + \vec{h}]$ пол-

нностью лежит в Ω и что f имеет производную $(m+1)$ -го порядка во всех точках открытого отрезка $[x, x + \vec{h}]$.

Тогда имеет место формула

$$\begin{aligned} f(x + \vec{h}) &= f(x) + \frac{f'(x)}{1!} \vec{h} + \frac{f''(x)}{2!} \vec{h}^2 + \dots \\ &\dots + \frac{f^{(m)}(x)}{m!} \vec{h}^m + \frac{f^{(m+1)}(x + \theta \vec{h})}{(m+1)!} \vec{h}^{m+1}, \quad (\text{III}, 7; 1_2) \end{aligned}$$

где θ — вещественное число и $0 < \theta < 1$.

Доказательство. Обозначим через Φ вещественную функцию на $[0, 1]$, определенную формулой $\Phi: t \rightarrow f(x + t\vec{h})$. По индукции проверяется, что на $[0, 1]$ существует производная порядка $k \leq m$ и на $[0, 1]$ — производная порядка $m+1$, причем

$$\Phi^{(k)}(t) = f^{(k)}(x + t\vec{h}) \vec{h}^k. \quad (\text{III}, 7; 1_3)$$

Теперь достаточно к этой функции применить теорему 6 (формула Тейлора для вещественной функции от одной вещественной переменной) на отрезке $[0, 1]$, чтобы получить требуемый результат.

Далее, теореме 13 соответствует следующая формула Тейлора:

Теорема 20. Пусть f есть m раз дифференцируемое отображение открытого множества $\Omega \subset E$ в F . Предположим, что замкнутый отрезок $[x, x + \vec{h}]$ полностью лежит в Ω и что f имеет производную порядка $m+1$ в каждой точке открытого отрезка $[x, x + \vec{h}]$, не превосходящую по норме числа M . Тогда имеет место формула

$$\begin{aligned} \|\vec{A}\| &= \left\| \overrightarrow{f(x + \vec{h}) - f(x) - f'(x) \vec{h} - \frac{f''(x)}{2!} \vec{h}^2 - \dots} \right. \\ &\quad \left. - \frac{f^{(m)}(x)}{m!} \vec{h}^m \right\| \leq \frac{M \|\vec{h}\|^{m+1}}{(m+1)!}. \quad (\text{III}, 7; 2) \end{aligned}$$

Доказательство. Это утверждение при $m = 0$ было уже доказано в теореме 13. Поэтому мы можем воспользоваться индукцией. Предположим, что это неравенство доказано для $m-1$, и докажем его справедливость для m , $m \geq 1$.

Рассмотрим функцию \vec{g} , определенную формулой

$$\vec{g}(t) = \overrightarrow{f(x + t\vec{h}) - f(x) - f'(x) \vec{h} - \dots - \frac{f^{(m)}(x)}{m!} (t\vec{h})^m}. \quad (\text{III}, 7; 3)$$

Это отображение отрезка $[0, 1]$ из \mathbb{R} в \vec{F} , и $\vec{A} = \overrightarrow{g(1) - g(0)}$. Оно непрерывно на этом отрезке и, кроме того, имеет непрерывную производную на всем отрезке, определенную формулой

$$\overrightarrow{g'(t)} = f'(x + t\vec{h}) \vec{h} - f'(x) \vec{h} - \dots - \frac{t^{m-1}}{(m-1)!} f^{(m)}(x) \vec{h}^m. \quad (\text{III}, 7; 4)$$

Согласно 2-й формуле (III, 6; 19),

$$f^{(k)}(x) \vec{h}^k = ((f')^{(k-1)}(x) \vec{h}^{k-1}) \cdot \vec{h}. \quad (\text{III}, 7; 5)$$

Поэтому (III, 7; 4) можно записать в виде

$$g'(t) = \left[f'(x + t\vec{h}) - f'(x) - \dots - \frac{(f')^{(m-1)}}{(m-1)!}(x) (t\vec{h})^{m-1} \right] \vec{h} \in \vec{F}, \quad (\text{III}, 7; 6)$$

где выражение, стоящее в квадратных скобках, является элементом из $\mathcal{L}(\vec{E}; \vec{F})$. Будучи примененным к $\vec{h} \in \vec{E}$, оно дает элемент пространства \vec{F} .

Найдем теперь оценку нормы этой производной. Мы видим, что выражение в скобках является аналогом выражения \vec{A} , но относящимся к приращению $t\vec{h}$, порядку $m-1$ и функции f' , определенной на Ω , со значениями в $\mathcal{L}(\vec{E}; \vec{F})$. В силу индуктивного предположения, учитывая, что $(f')^{(m)} = f^{(m+1)}$, мы сможем написать следующую оценку, пригодную для $0 \leq t \leq 1$:

$$\|\overrightarrow{g'(t)}\| \leq M \frac{\|\vec{h}\|^m}{m!} \|\vec{h}\| = M \|\vec{h}\|^{m+1} \frac{t^m}{m!}. \quad (\text{III}, 7; 7)$$

Пользуясь теперь леммой, приведенной в связи с теоремой 13 (заменив там $f(x)$ на $\overrightarrow{g(x)}$ и $g(x)$ на $M \|\vec{h}\|^{m+1} \frac{x^{m+1}}{(m+1)!}$), мы можем убедиться, что для функции \overrightarrow{g} на отрезке $[0, 1]$ выполняется неравенство

$$\|\overrightarrow{g(1) - g(0)}\| \leq M \frac{\|\vec{h}\|^{m+1}}{(m+1)!}, \quad (\text{III}, 7; 8)$$

и тем самым справедлива формула (III, 7; 2).

Следствие. Пусть L есть $(m+1)$ -линейное непрерывное отображение \vec{E}^{m+1} в \vec{F} . Тогда при условиях теоремы имеет ме-

сто формула

$$\overrightarrow{f(x + \vec{h}) - f(x) - f'(x)\vec{h} - \frac{f''(x)}{2!}\vec{h}^2} = \dots \\ \dots - \frac{f^{(m)}(x)}{m!}\vec{h}^m - \frac{L}{(m+1)!}\vec{h}^{m+1} \leq \omega \frac{\|\vec{h}\|^{m+1}}{(m+1)!}, \quad (\text{III}, 7; 9)$$

где ω является точной верхней границей $\|f^{(m+1)}(\xi) - L\|$, когда ξ пробегает интервал $[x, x + \vec{h}]$.

Доказательство. Утверждение легко сводится к случаю, когда L является $(m+1)$ -линейной симметричной формой. Впрочем, это единственный случай, представляющий интерес. Затем теорему следует применить к функции $\xi \rightarrow f(\xi) - \frac{L}{(m+1)!}(\xi - x)^{m+1}$, производные которой вычисляются по формуле (III, 6; 22₂).

В заключение заметим, что существует формула Тейлора, соответствующая определению производной (III, 3; 13), которую можно назвать формулой Тейлора для бесконечно малых приращений. Здесь только удобнее заменить $m+1$ на m .

Теорема 21. Если f имеет в Ω производные порядка $\leq m-1$ и производную порядка m в точке a , то

$$f(a + \vec{h}) = f(a) + f'(a)\vec{h} + \dots + \frac{f^{(m)}(a)}{m!}\vec{h}^m + \vec{\alpha} \|\vec{h}\|^m, \quad (\text{III}, 7; 10)$$

где $\vec{\alpha}$ стремится к $\vec{0}$ вместе с \vec{h} .

Будем доказывать эту формулу снова индукцией по m . Она верна для $m=1$. Предполагая, что она верна для целого $m-1$, докажем ее справедливость для $m \geq 2$. Функция

$$\overrightarrow{g(\vec{\xi})} = \overrightarrow{f(a + \vec{\xi}) - f(a) - f'(a)\vec{\xi} - \dots - \frac{f^{(m)}(a)}{m!}\vec{\xi}^m}, \quad (\text{III}, 7; 11)$$

определенная в некоторой окрестности точки $\vec{0}$ в \vec{E} , со значениями в \vec{F} дифференцируема. Согласно (III, 6; 22₂), ее производная имеет вид

$$g'(\vec{\xi}) = f'(a + \vec{\xi}) - f'(a) - \dots - \frac{(f')^{(m-1)}(a)}{(m-1)!}\vec{\xi}^{m-1} \in \mathcal{L}(\vec{E}; \vec{F}). \\ (\text{III}, 7; 12)$$

Применим к g' формулу (III, 7; 10) для $m - 1$:

$$g'(\vec{\xi}) = \overrightarrow{\beta(\vec{\xi})} \parallel \vec{\xi} \parallel^{m-1}, \quad (\text{III, 7; 12}_2)$$

где $\overrightarrow{\beta(\vec{\xi})}$ стремится к $\vec{0}$ вместе с $\vec{\xi}$.

Тогда формула конечных приращений (теорема 13) дает соотношения

$$\begin{aligned} \parallel \vec{g}(\vec{h}) \parallel &= \parallel \vec{g}(\vec{h}) - \vec{g}(\vec{0}) \parallel \leqslant \parallel \vec{h} \parallel \sup_{\vec{\xi} \in [\vec{0}, \vec{h}]} \parallel \vec{g}'(\vec{\xi}) \parallel \leqslant \\ &\leqslant \parallel \vec{h} \parallel \sup_{\vec{\xi} \in [\vec{0}, \vec{h}]} \parallel \overrightarrow{\beta(\vec{\xi})} \parallel \parallel \vec{\xi} \parallel^{m-1} \leqslant \varepsilon \parallel \vec{h} \parallel^m, \quad (\text{III, 7; 12}_3) \end{aligned}$$

где ε стремится к 0 вместе с $\parallel \vec{h} \parallel$, что и дает требуемый результат для m .

Применение формулы Тейлора для вычисления производных

Иногда легче получить разложение Тейлора, чем вычислить последовательные производные. Тогда эти производные можно получить из готового разложения. Другими словами, имеет место теорема, обратная к теореме 21:

Теорема 21₂. Пусть f есть m раз дифференцируемое отображение открытого множества $\Omega \subset E$ в F . Если найдены k -линейные непрерывные симметричные отображения L_k пространств \vec{E}^k в \vec{F} , $k = 1, 2, \dots, m$, и некоторый элемент L_0 из F , такие, что

$$f(a + \vec{h}) = L_0 + L_1 \vec{h} + \frac{L_2}{2!} \vec{h}^2 + \dots + \frac{L_m}{m!} \vec{h}^m + \vec{\alpha} \parallel \vec{h} \parallel^m, \quad (\text{III, 7; 12}_4)$$

где $\vec{\alpha} = \vec{\alpha}(\vec{h})$ стремится к $\vec{0}$ вместе с \vec{h} , то всегда имеют место равенства

$$L_k = f^{(k)}(a), \quad k = 1, 2, \dots, m, \quad \text{и} \quad L_0 = f(a).$$

Доказательство. Из теоремы 21 вытекает, что если мы положим

$$A_k = f^{(k)}(a) - L_k \in \mathcal{L}_k(\vec{E}^k; \vec{F}), \quad \vec{A}_0 = \overrightarrow{f(a) - L_0} \in \vec{F}, \quad (\text{III, 7; 12}_5)$$

то

$$\sum_{k=1}^m \frac{A_k}{k!} \vec{h}^k = \overrightarrow{\beta} \parallel \vec{h} \parallel^m, \quad (\text{III, 7; 12}_6)$$

где $\overrightarrow{\beta}$ стремится к $\vec{0}$ вместе с \vec{h} .

Устремляя \vec{h} к $\vec{0}$, мы увидим, что \vec{A}_0 равно нулю. Предположим, что $A_0, A_1, A_2, \dots, A_{k-1}$ — нули для $k \leq m$, и докажем, что тогда A_k также будет нулем. Отсюда будет вытекать, что все $A_k, k = 0, 1, 2, \dots, m$, — нули, и теорема будет доказана.

Положим $\vec{h} = t\vec{X}$, где \vec{X} фиксирован, и устремим скаляр t к 0. Тогда величина

$$t^k \left(\frac{A_k}{k!} \vec{X}^k + t \sum_{l=k+1}^m \frac{A_l}{l!} t^{l-k-1} \vec{X}^l \right) \quad (\text{III}, 7; 12_7)$$

является бесконечно малой более высокого порядка, чем t^m , и тем более чем t^k , когда t стремится к 0. Это означает, что ее отношение к t^k стремится к 0, а, следовательно, $A_k \vec{X}^k = 0$ для любого $\vec{X} \in \vec{E}$. Доказательство по индукции будет закончено, если мы докажем такую лемму:

Лемма. Пусть A есть k -линейное симметрическое отображение \vec{E}^k в \vec{F} . Если для любого \vec{X} из \vec{E}

$$A\vec{X}^k = \vec{0} \in \vec{F}, \quad (\text{III}, 7; 13)$$

то A равно нулю.

Эта лемма хорошо известна для $m = 2$, ибо тогда для любых $\vec{X} \in \vec{E}$ и $\vec{Y} \in \vec{E}$ в силу симметричности A имеем:

$$A(\vec{X} + \vec{Y}, \vec{X} + \vec{Y}) - A(\vec{X}, \vec{X}) - A(\vec{Y}, \vec{Y}) = 4A(\vec{X}, \vec{Y}),$$

и условие (III, 7; 13) влечет за собой обращение в нуль левой части равенства, а значит, и его правой части, откуда следует, что A равно нулю. Распространим это доказательство на общий случай.

Рассмотрим для фиксированных $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_k$ функцию h от $t = (t_1, t_2, \dots, t_k) \in \mathbb{R}^k$:

$$h(t) = A(t_1 \vec{X}_1 + t_2 \vec{X}_2 + \dots + t_k \vec{X}_k)^k. \quad (\text{III}, 7; 13_2)$$

По условию она тождественно равна нулю. Эта функция является полиномом от t_1, t_2, \dots, t_k . Следовательно, каждый коэффициент этого полинома равен нулю. В силу симметричности A коэффициент при t_1, t_2, \dots, t_k равен $k! A(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_k)$, откуда и следует утверждение леммы.

Замечание. Мы должны были предполагать функцию f дифференцируемой до порядка m . Это существенно. *Тот факт, что в окрестности a функция f имеет разложение вида*

(III, 7; 12₄), вовсе не влечет за собой существование производных порядка >1 . Например, вещественная функция f вещественной переменной, определенная формулами $f(0) = 0$, $f(x) = x^{m+1} \sin \frac{1}{x^{m+1}}$ для $x \neq 0$, имеет рассматриваемое разложение

в начале координат с $L_h = 0$. Она имеет производные всех порядков в дополнении к началу координат. Поскольку $|f(x) - f(0)| = |f(x)| \leq |x^{m+1}|$, в начале координат ее первая производная равна нулю. Но эта первая производная, равная $(m+1)x^m \sin \frac{1}{x^{m+1}} - \frac{m+1}{x} \cos \frac{1}{x^{m+1}}$, в точках $x \neq 0$

принимает сколь угодно большие значения в любой окрестности 0. Следовательно, она разрывна в начале координат и, значит, рассматриваемая функция не имеет там второй производной.

Применим теорему 21₂ к вычислению последовательных производных сложной функции.

Теорема 21₃. Если g и f — дифференцируемые m раз скалярные функции скалярной переменной, то производная порядка m функции $h = g \circ f$ в точке a задается следующей формулой, в которой $f^{(p)}$ означает $f^{(p)}(a)$ и $g^{(q)}$ означает $g^{(q)}(f(a))$:

$$h^{(m)}(a) = \sum_{k_1+2k_2+\dots+mk_m=m} \frac{m!}{k_1! k_2! \dots k_m! (1!)^{k_1} (2!)^{k_2} \dots (m!)^{k_m}} \times \\ \times g^{(k_1+k_2+\dots+k_m)}(f')^{k_1} (f'')^{k_2} \dots (f^{(m)})^{k_m}. \quad (\text{III}, 7; 13_3)$$

Доказательство. Воспользуемся разложением Тейлора

$$h(x) = g(f(x)) = \sum_{0 \leq l \leq m} \frac{(f(x) - f(a))^l}{l!} g^{(l)} + R_m. \quad (\text{III}, 7; 13_4)$$

Так как

$$f(x) - f(a) = \sum_{1 \leq k \leq m} \frac{(x-a)^k}{k!} f^{(k)} + \rho_m, \quad (\text{III}, 7; 13_5)$$

то с точностью до членов высшего порядка малости по сравнению с $(x-a)^m$, когда x стремится к a , имеем:

$$(f(x) - f(a))^l \approx \left(\sum_{1 \leq k \leq m} \frac{(x-a)^k}{k!} f^{(k)} \right)^l. \quad (\text{III}, 7; 13_6)$$

Можно написать формулу возвведения в степень суммы l слагаемых, обобщающую формулу бинома Ньютона (см. ниже)

формулу (III, 7; 26)). Согласно этой формуле,

$$(f(x) - f(a))^l = \sum_{k_1+k_2+\dots+k_m=l} \frac{l!}{k_1! k_2! \dots k_m!} \frac{(x-a)^{k_1+2k_2+\dots+mk_m}}{(1!)^{k_1} (2!)^{k_2} \dots (m!)^{k_m}} \times \\ \times (f')^{k_1} (f'')^{k_2} \dots (f^{(m)})^{k_m}. \quad (\text{III}, 7; 13_7)$$

По этой формуле вычисляется $\frac{(f(x) - f(a))^l}{l!}$, а затем, с точностью до бесконечно малых более высокого порядка, чем $(x-a)^m$, при x , стремящемся к a , находится выражение для $h(x)$:

$$h(x) \approx \sum_{\substack{0 \leq i \leq m \\ k_1+k_2+\dots+k_m=l}} \frac{(x-a)^{k_1+2k_2+\dots+mk_m}}{k_1! k_2! \dots k_m! (1!)^{k_1} (2!)^{k_2} \dots (m!)^{k_m}} \times \\ \times (f')^{k_1} (f'')^{k_2} \dots (f^{(m)})^{k_m} g^{(l)}. \quad (\text{III}, 7; 13_8)$$

Поскольку известно, что h дифференцируема m раз (теорема 19), величина $h^{(m)}(a)$ должна равняться произведению $m!$ и коэффициента при $(x-a)^m$ в предыдущем разложении. Отсюда и вытекает высказанное утверждение.

На этом примере видно, какую большую роль может сыграть разложение Тейлора в вычислении производных порядка >1 при условии, что существование искомых производных доказано другими путями.

Формула Тейлора относительно некоторой системы координат

Теорема 214. Если пространство E конечномерно, в нем выбрана система координат и вектор \vec{h} имеет координаты h_1, h_2, \dots, h_n , то, пользуясь обозначениями (III, 6; 26—27), выражение $\sum_{k=0}^m \frac{f^{(k)}(x)}{k!} \vec{h}^k$, входящее в правую часть формулы Тейлора, можно записать в виде

$$\sum_{k=0}^m \frac{f^{(k)}(x)}{k!} \vec{h}^k = \sum_{|\vec{p}| \leq m} \frac{\overrightarrow{D^p f}(x)}{\vec{p}!} \vec{h}^{\vec{p}} \in F. \quad (\text{III}, 7; 14)$$

Доказательство. Согласно формуле (III, 6; 25),

$$\frac{f^{(k)}(x)}{k!} \vec{h}^k = \frac{1}{k!} \sum_{i_1, i_2, \dots, i_k} \frac{\overrightarrow{\partial^k f}}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} (x) h_{i_1} h_{i_2} \dots h_{i_k} \in \vec{F}^1 \quad (\text{III}, 7; 14_2)$$

¹⁾ В формулу (III, 6; 25) входили векторы $\vec{X}_1, \vec{X}_2, \dots$, которые здесь равны h .

В этой формуле j_1, j_2, \dots, j_k — индексы, принимающие независимо друг от друга всевозможные значения из множества $\{1, 2, \dots, n\}$. Объединим между собой все конечные последовательности j_1, j_2, \dots, j_k , у которых p_1 индексов равны 1, p_2 индексов равны 2, ..., p_n индексов равны n . В сумме (III, 7; 14₂) они дадут элемент, который запишется в обозначениях (III, 6; 25 и 27) в виде

$$\frac{\overrightarrow{D^p f}(x) \vec{h}^p}{k!} = \overline{\left(\frac{\partial}{\partial x_1} \right)^{p_1} \left(\frac{\partial}{\partial x_2} \right)^{p_2} \cdots \left(\frac{\partial}{\partial x_n} \right)^{p_n}} f(x) \frac{h_1^{p_1} h_2^{p_2} \cdots h_n^{p_n}}{k!} \in \vec{F}. \quad (\text{III}, 7; 14_3)$$

Членов этого вида имеется столько, сколько существует отображений j множества $\{1, 2, \dots, k\}$ в множество $\{1, 2, \dots, n\}$, где $\{1\}$ имеет прообраз, состоящий из p_1 элементов, $\{2\}$ — из p_2 элементов, ..., $\{n\}$ — из p_n элементов, $|j| = p_1 + p_2 + \dots + p_n = k$. Это число $\gamma_{p_1, p_2, \dots, p_n} = \gamma_p$ называется числом «перестановок с повторениями» n объектов $\{1, 2, \dots, n\}$, где объект 1 берется p_1 раз, объект 2 берется p_2 раз, ..., объект n берется p_n раз. Существуют разные способы вычисления этого числа. Для того чтобы определить интересующее нас отображение j , нам надо сначала выбрать часть из $\{1, 2, \dots, k\}$, состоящую из p_1 элементов, которая и будет составлять $j^{-1}\{1\}$. Число возможных выборов равно числу подмножеств, содержащих по p_1 элементов из $\{1, 2, \dots, k\}$, т. е. $\frac{k!}{(p_1)! (k - p_1)!}$. Сделав этот выбор, надо в дополнении $\{1, 2, \dots, k\} - j^{-1}\{1\}$, состоящем из $k - p_1$ элементов, выбрать произвольно p_2 элементов, составляющих $j^{-1}\{2\}$. При этом число возможных выборов равно $\frac{(k - p_1)!}{p_2! (k - p_1 - p_2)!}$. Так можно продолжать далее. Число таких отображений j равно

$$\gamma_p = \frac{k!}{p_1! (k - p_1)!} \cdot \frac{(k - p_1)!}{p_2! (k - p_1 - p_2)!} \cdots \frac{(k - p_1 - p_2 - \dots - p_{n-1})!}{p_n! (k - p_1 - p_2 - \dots - p_n)!} = \\ = \frac{k!}{p_1! p_2! \cdots p_n!} = \frac{k!}{p!}, \quad (\text{III}, 7; 15)$$

а, следовательно,

$$\frac{\gamma_p}{k!} = \frac{1}{p!}, \quad (\text{III}, 7; 16)$$

и теорема доказана.

В принятых обозначениях полученная формула совпадает с формулой, соответствующей функциям одной переменной. В левой части формулы Тейлора $f(x + \vec{h})$ лежит в аффинном пространстве F , все члены правой ее части принадлежат вектор-

торному пространству \vec{F} , кроме первого $f(x)$, соответствующего $\vec{p} = \vec{0}$ и лежащего в аффинном пространстве F .

Следствие. Если при выполнении условий теоремы найдены такие элементы $\vec{c}_{\vec{p}} (\vec{c}_{\vec{p}} \in \vec{F} \text{ для } \vec{p} \neq \vec{0}, \vec{c}_{\vec{0}} \in F), |\vec{p}| \leq m$, что

$$f(a + \vec{h}) = \sum_{|\vec{p}| \leq m} \vec{c}_{\vec{p}} \frac{\vec{h}^{\vec{p}}}{\vec{p}!} + \vec{a} \parallel \vec{h} \parallel^m, \quad (\text{III, 7; 17})$$

где $\vec{a} \in \vec{F}$ стремится к $\vec{0}$ вместе с \vec{h} , и если известно, что f дифференцируема m раз в a , то имеют место равенства

$$\vec{c}_{\vec{p}} = D^{\vec{p}} f(a). \quad (\text{III, 7; 18})$$

Доказательство. Положим $\vec{a}_{\vec{p}} = D^{\vec{p}} f(a) - \vec{c}_{\vec{p}}$ (здесь также $\vec{a}_{\vec{0}} \in \vec{F}$). Тогда, составляя разность, можно убедиться, что норма «полинома»¹⁾ (с векторными коэффициентами) относительно h_1, h_2, \dots, h_n :

$$\sum_{|\vec{p}| \leq m} \vec{a}_{\vec{p}} \vec{h}^{\vec{p}} = \sum \vec{a}_{p_1, p_2, \dots, p_n} h_1^{p_1} h_2^{p_2} \dots h_n^{p_n} \quad (\text{III, 7; 19})$$

является бесконечно малой более высокого порядка, чем $\parallel \vec{h} \parallel^m$, при \vec{h} , стремящемся к $\vec{0}$. Полагая $\vec{h} = t \vec{X}$ и устремляя t к 0, с помощью индукции, аналогичной той, которая применялась при доказательстве теоремы 21₂ (но без использования леммы), можно убедиться в том, что однородные части полинома $\sum_{|\vec{p}| \leq m} \vec{a}_{\vec{p}} \vec{X}^{\vec{p}}$ степеней 0, 1, 2, ..., m тождественно равны нулю. Но у тождественно равного нулю полинома все коэффициенты равны нулю²⁾, чем и заканчивается доказательство следствия.

Упражнение. Применяя следствие, получите новое доказательство формулы Лейбница (III, 6; 36) в духе теоремы 21₃.

Теорема 22. Для того чтобы отображение f открытого связного множества $\Omega \subset E$ в пространство F имело производ-

¹⁾ См. далее замечание о полиномах к теореме 22.

²⁾ Это обычно доказывается при помощи индукции по числу n переменных X_1, X_2, \dots, X_n только для полиномов со скалярными коэффициентами. Но этот результат остается верным для векторных коэффициентов, причем сохраняется то же самое доказательство.

ную порядка $m+1$, равную нулю в Ω , необходимо и достаточно, чтобы отображение f было полиномом степени $\leq m$.

Доказательство. Пусть сначала $m=0$. Мы уже видели, что производная постоянной равна нулю, даже если Ω не связно. Нам надо доказать, что в случае связного множества Ω отображение f множества Ω в F , имеющее всюду равную нулю производную, постоянно.

Пусть x — произвольная точка Ω . Существует такое число $\rho_x > 0$, что открытый шар B_x с центром в x радиуса ρ_x целиком лежит в Ω . При $\|\vec{h}\| < \rho_x$ можно на отрезке $[x; x + \vec{h}]$, лежащем в Ω , применить теорему о конечных приращениях и получить неравенство

$$\overrightarrow{\|f(x + \vec{h}) - f(x)\|} \leq \sup_{\vec{\xi} \in [x, x + \vec{h}]} \|f'(\vec{\xi})\| = 0, \quad (\text{III}, 7; 20)$$

из которого следует, что $f(x)$ постоянна в B_x .

Пусть теперь A — произвольная фиксированная точка множества Ω . Обозначим через A множество точек ξ из Ω , таких, что $f(\xi) = f(A)$. Так как A является прообразом множества $\{f(A)\} \subset F$ при непрерывном отображении f , то оно замкнуто. Это множество также и открыто, ибо для $x \in A$ любой указанный выше шар принадлежит A . Поскольку Ω связно и A непусто, A совпадает со всем Ω , что и доказывает теорему для $m=0$.

Предположим теперь, что $m \geq 1$. Прежде всего надо дать определение полинома на E со значениями в F . Мы ограничимся частным случаем, когда E n -мерно.

Если выбрана система координат $0, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$, то полином относительно x_1, x_2, \dots, x_n степени $\leq n$ со значениями в F определяется как функция вида

$$f(x) = \sum_{|\vec{p}| \leq m} \vec{a}_{\vec{p}} \vec{x}^{\vec{p}} = \sum_{p_1 + p_2 + \dots + p_n \leq m} a_{p_1, p_2, \dots, p_n} x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}, \quad (\text{III}, 7; 21)$$

где $\vec{a}_{\vec{p}}$ принадлежат F , кроме \vec{a}_0 , лежащего в F .

Эта функция имеет тот же вид в любой другой системе координат и называется полиномом на E степени $\leq m$ со значениями в F . Ее производная порядка $m+1$ равна нулю независимо от того, связно или нет множество Ω , поскольку, согласно (III, 7; 21), ее частные производные порядка $m+1$ равны нулю.

Обратно, пусть Ω связно, и пусть f — отображение множества Ω в пространство F , $(m+1)$ -я производная которого

равна нулю. Пусть $a \in \Omega$. Формула Тейлора порядка m относительно любой системы координат имеет нулевой остаточный член \tilde{R}_m и, значит, отображение f в указанном шаре B_a является полиномом P степени $\leq m$. Обозначим теперь через k такое целое число, что $(f - P)^{(k)} = 0$ в Ω . Такое целое k существует, например $k = m + 1$. Тогда, если $k \geq 1$, то $(f - P)^{(k-1)}$ является функцией, первая производная которой равна нулю в связном множестве Ω . Значит, она постоянна в Ω . Но в B_a она равна нулю и, следовательно, она равна нулю в Ω . Таким образом, $k - 1$ обладает тем же свойством, что и k . Шаг за шагом спускаемся до $k = 0$ и получаем $f = P$ в Ω , что и завершает доказательство теоремы.

Замечания. 1°) Конечно, это заключение не верно, если множество Ω не связно. Достаточно взять в качестве Ω дополнение к началу координат на вещественной прямой \mathbb{R} и положить $m = 0$. Вещественное значение функция, определенная на Ω , имеющая всюду равную нулю производную, не обязательно постоянна в Ω . Она может быть равной некоторой постоянной на полуправой $]-\infty, 0[$ и другой постоянной на полуправой $]0, +\infty[$.

2°) Предположим, что у нас нет формулы (III, 7; 15), определяющей γ_p . Ее можно очень просто получить следующим образом.

Не считая γ_p известным, напишем формулу Маклорена для полинома степени m :

$$\overrightarrow{f(x)} = \sum_{|\vec{p}| \leq m} \gamma_{\vec{p}} \xrightarrow{\longrightarrow} \frac{D^{\vec{p}} f(0)}{k!} \vec{x}^{\vec{p}} = \sum_{|\vec{p}| \leq m} c_{\vec{p}} \xrightarrow{\longrightarrow} \frac{D^{\vec{p}} f(0)}{\vec{p}!} \vec{x}^{\vec{p}}, \quad (\text{III, 7; 22})$$

где $c_{\vec{p}} = \frac{\vec{p}!}{k!} \gamma_{\vec{p}}$, и покажем, что $c_{\vec{p}} = 1$.

Пусть $\vec{q} \in \mathbb{N}^n$, $\vec{q} \leq \vec{p}$. Тогда

$$\begin{aligned} D^{\vec{q}} \xrightarrow{\vec{p}} \frac{\vec{x}^{\vec{p}}}{\vec{p}!} &= \left(\left(\frac{\partial}{\partial x_1} \right)^{q_1} \frac{x_1^{p_1}}{p_1!} \right) \left(\left(\frac{\partial}{\partial x_2} \right)^{q_2} \frac{x_2^{p_2}}{p_2!} \right) \dots \left(\left(\frac{\partial}{\partial x_n} \right)^{q_n} \frac{x_n^{p_n}}{p_n!} \right) = \\ &= \begin{cases} \frac{\vec{x}^{\vec{p}-\vec{q}}}{(\vec{p}-\vec{q})!}, & \text{если } \vec{q} \leq \vec{p}^1, \\ 0, & \text{если } \vec{q} \not\leq \vec{p}. \end{cases} \end{aligned} \quad (\text{III, 7; 23})$$

¹⁾ Отношение порядка $\vec{q} \leq \vec{p}$ в \mathbb{N}^n является частичным, а не полным: $\vec{q} \leq \vec{p}$ вовсе не означает, что $\vec{q} > \vec{p}$. Так, например, при $n = 2$ мы имеем $(1, 1) \leq (0, 2)$, но при этом не выполняется и неравенство $(1, 1) > (0, 2)$.

Теперь дифференцирование $D^{\vec{q}}$ выражения (III, 7; 22) для $|\vec{q}| \leq m$ дает:

$$\overrightarrow{D^{\vec{q}}f(x)} = \sum_{\substack{|\vec{p}| \leq m \\ \vec{p} \geq \vec{q}}} c_{\vec{p}} \frac{\overrightarrow{D^{\vec{p}}f(0)}}{(\vec{p} - \vec{q})!} \vec{x}^{\vec{p} - \vec{q}}. \quad (\text{III, 7; 24})$$

Положим $\vec{x} = \vec{0}$. Тогда $\vec{x}^{\vec{p} - \vec{q}} = 0$ для $\vec{p} \neq \vec{q}$ и $\vec{x}^{\vec{p} - \vec{q}} = 1$ для $\vec{p} = \vec{q}$. Поэтому

$$\overrightarrow{D^{\vec{q}}f(0)} = c_{\vec{q}} \overrightarrow{D^{\vec{q}}f(0)}, \quad (\text{III, 7; 25})$$

откуда $c_{\vec{q}} = 1$, что и требовалось доказать.

Положим, в частности, $f(x) = (x_1 + x_2 + \dots + x_n)^k$. Теорема о сложной функции говорит о том, что любая частная производная порядка l этой функции равна $k(k-1)\dots(k-l+1)(x_1 + x_2 + \dots + x_n)^{k-l}$. При $x_1 = x_2 = \dots = x_n = 0$ эта производная равна нулю, кроме случая $l = k$, когда она равна $k!$. Поэтому формула Маклорена дает:

$$(x_1 + x_2 + \dots + x_n)^k = \sum_{|\vec{p}|=k} \gamma_{\vec{p}} \vec{x}^{\vec{p}} = \sum_{|\vec{p}|=k} \frac{k!}{\vec{p}!} \vec{x}^{\vec{p}}. \quad (\text{III, 7; 26})$$

Итак, $\gamma_{p_1, p_2, \dots, p_n} = \frac{(p_1 + p_2 + \dots + p_n)!}{p_1! p_2! \dots p_n!}$ есть коэффициент при $x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}$ в разложении $(x_1 + x_2 + \dots + x_n)^{p_1 + p_2 + \dots + p_n}$, что является обобщением формулы бинома Ньютона. Это, впрочем, очевидно, если воспользоваться комбинаторным определением $\gamma_{\vec{p}}$, данным после (III, 7; 14).

Коэффициенты бинома имеют очень простой вид, что позволяет сразу написать формулу Маклорена для функции двух скалярных переменных x и y . Полагая

$$\vec{p} = \frac{\vec{\partial f}}{\vec{\partial x}}(0, 0), \quad \vec{q} = \frac{\vec{\partial f}}{\vec{\partial y}}(0, 0),$$

$$\vec{r} = \frac{\vec{\partial^2 f}}{\vec{\partial x^2}}(0, 0), \quad \vec{s} = \frac{\vec{\partial^2 f}}{\vec{\partial x \partial y}}(0, 0), \quad \vec{t} = \frac{\vec{\partial^2 f}}{\vec{\partial y^2}}(0, 0), \quad (\text{III, 7; 26}_2)$$

$$\vec{\alpha} = \frac{\vec{\partial^3 f}}{\vec{\partial x^3}}(0, 0), \quad \vec{\beta} = \frac{\vec{\partial^3 f}}{\vec{\partial x^2 \partial y}}(0, 0), \quad \vec{\gamma} = \frac{\vec{\partial^3 f}}{\vec{\partial x \partial y^2}}(0, 0), \quad \vec{\delta} = \frac{\vec{\partial^3 f}}{\vec{\partial y^3}}(0, 0),$$

получим:

$$\begin{aligned} f(x, y) = & f(0, 0) + \vec{px} + \vec{qy} + \frac{1}{2} (\vec{rx^2} + 2\vec{sxy} + \vec{ty^2}) + \\ & + \frac{1}{6} (\vec{\alpha x^3} + 3\vec{\beta x^2y} + 3\vec{\gamma xy^2} + \vec{\delta y^3}) + \dots \quad (\text{III}, 7; 26_3) \end{aligned}$$

Применение к изучению максимумов и минимумов. Определения

Пусть f — вещественнозначная функция, определенная на топологическом пространстве E . Говорят, что f имеет в точке $a \in E$ относительный максимум¹⁾, если существует некоторая окрестность \mathcal{U} точки $a \in E$, такая, что сужение f на эту окрестность имеет в точке a максимум, т. е. если для любой точки x из \mathcal{U} имеет место неравенство $f(x) \leq f(a)$. Относительный максимум называют строгим, если можно выбрать \mathcal{U} таким образом, чтобы для $x \neq a$ из \mathcal{U} имело место строгое неравенство $f(x) < f(a)$.

Аналогичное определение дается для относительного минимума и строгого относительного минимума.

Конечно, максимум функции f , т. е. ее абсолютный максимум, является одновременно и относительным максимумом, в то время как противоположное утверждение, вообще говоря, не верно.

Говорят, что функция f имеет в точке a экстремум, если она имеет в этой точке относительный максимум или относительный минимум.

Необходимые условия экстремума

Теорема 23. Пусть f — вещественнозначная функция, определенная на открытом множестве Ω аффинного нормированного пространства E и дифференцируемая на нем. Необходимым условием существования в точке $a \in \Omega$ относительного максимума или минимума является обращение в нуль производного отображения $f'(a) \in \mathcal{L}(E; \mathbb{R}) = \tilde{E}$.

Доказательство. Заметим прежде всего, что если E конечномерно и в нем выбрана система координат $0, (e_i)_{i=1}^n$, то это необходимое условие выражается в общезвестной форме: для того чтобы f в точке a имела относительный максимум или минимум, необходимо, чтобы частные производные $\partial f / \partial x_i$

¹⁾ Вместо того чтобы говорить, что функция f имеет в точке a максимум или относительный максимум и т. д., иногда говорят, что точка a является для f максимумом или относительным максимумом и т. д. Это не верно, так как максимум равен $f(a)$, а не $a!$

в точке a были равны нулю или чтобы дифференциал в точке a : $\sum_{i \in I} \frac{\partial f}{\partial x_i}(a) dx_i$ был тождественно равен нулю.

Докажем теперь теорему в общем случае. Если f в точке a имеет относительный максимум или минимум, то функция $t \rightarrow f(a + t\vec{X})$ при фиксированном \vec{X} в \vec{E} определена для $t \in \mathbb{R}$ при достаточно малом $|t|$ (поскольку Ω — открытое множество) и имеет в точке $t = 0$ относительный максимум или минимум. Тогда доказательство, аналогичное доказательству теоремы Ролля (теорема 4), показывает, что ее производная в точке t равна нулю при $t = 0$. Эта производная равна $f'(a)\vec{X}$. Так как $f'(a)\vec{X} = 0$ для любого $\vec{X} \in \vec{E}$, то $f'(a) = 0$.

Замечания. По многим причинам теорема 23 не решает задачи об отыскании экстремума функции f .

1°) Чаще всего вещественная функция f задается не на открытом, а на замкнутом множестве F (например, на сегменте $[a, b]$ из \mathbb{R} или в замкнутом шаре из E). Она может и не иметь никакого экстремума, если множество, на котором она определена, не компактно (например, $f(x) = x$ не имеет экстремума на \mathbb{R}). Если она и имеет экстремум, то он может не удовлетворять теореме, которая применима только к *открытым множествам*. Естественно, если f дифференцируема во внутренности $\Omega = F$ множества F , то каждая точка экстремума a , находящаяся в Ω , будет найдена из условия $f'(a) = 0$, но экстремальные значения могут достигаться на границе множества F (например, если f есть функция x^2 , определенная на отрезке $[-1, 1]$, то ее минимум при $x = 0$ находится приравниванием нулю производной, поскольку 0 лежит внутри отрезка $[-1, 1]$, но точки максимума $x = \pm 1$ теореме не удовлетворяют, поскольку они находятся на концах этого интервала, где производная в нуль не обращается) ¹⁾.

2°) Обратно, уравнение $f'(a) = 0$ зачастую определяет точки, не являющиеся точками экстремума. Другими словами, необходимое условие $f'(a) = 0$ далеко не достаточно. Чтобы узнать, как ведет себя функция в точке a , надо изучить ее разложение Тейлора в окрестности этой точки.

1) В начале § 2 на стр. 195 мы говорили, что *большинство* теорем, сформулированных для открытого множества Ω из E , применимы также и к другим множествам, в особенности к неоткрытым интервалам \mathbb{R} . Теорема 23 — одна из тех, которые в этих условиях неприменимы. Ее доказательство опирается на доказательство теоремы Ролля (теорема 4), в котором существенно используется тот факт, что экстремальное значение достигается во внутренней точке интервала.

Исследуем подробнее эту проблему. Иногда факт обращения в нуль производной $f'(a)$ более важен, чем свойство экстремальности. Например, если рассматривается гиперповерхность $y = f(x)$ в пространстве $\Omega \times \mathbb{R}$, то условие $f'(a) = 0$ означает, что касательная гиперплоскость, проведенная к поверхности в точке $A = (a, f(a))$, горизонтальна. Это необходимое условие экстремума ординаты геометрически более важно, чем само свойство экстремальности. Поэтому *независимо от того, имеется или нет в точке a экстремум функции, мы будем называть точку a стационарной точкой функции f , если в ней $f'(a) = 0$.*

Нахождение необходимых и достаточных условий экстремума функции

Теорема 24. Пусть f — вещественная функция, определенная на открытом множестве Ω аффинного нормированного пространства E и m раз дифференцируемая. Если производные функции f порядка $1, 2, \dots, m-1$ равны нулю в точке $a \in \Omega$, а $f^{(m)}(a) \neq 0$ и если в точке a функция f имеет относительный максимум, то m четно и справедливо неравенство

$$f^{(m)}(a) \vec{X}^m \leqslant 0 \text{ для любого } \vec{X} \in \vec{E}. \quad (\text{III}, 7; 27)$$

Обратно, если f дифференцируема m раз в Ω , все производные f порядка $1, 2, \dots, m-1$ равны нулю в точке $a \in \Omega$ и если $f^{(m)}(a) \vec{X}^m$ ограничено сверху некоторым числом $-\delta < 0$ для всех \vec{X} из единичной сферы $\|\vec{X}\| = 1$ пространства \vec{E} , то f имеет в точке a строгий относительный максимум.

Доказательство. 1°) Докажем сначала первую часть теоремы, предполагая, что f имеет в точке $a \in \Omega$ относительный максимум. В этом случае функция $g: t \rightarrow f(a + t\vec{X})$, определенная в окрестности точки $t = 0$ на вещественной прямой \mathbb{R} , имеет относительный максимум в точке $t = 0$. Ее производная порядка k при $t = 0$ имеет вид $f^{(k)}(a) \vec{X}^k$. При $t = 0$ производные порядков $\leqslant m-1$ равны нулю, а производная порядка m равна $f^{(m)}(a) \vec{X}^m$. Поэтому для t , близких к нулю, имеем:

$$g(t) - g(0) = (f^{(m)}(a) \vec{X}^m) \frac{t^m}{m!} + \varepsilon_t t^m, \quad (\text{III}, 7; 28)$$

где ε_t стремится к 0 вместе с $|t|$. Отсюда следует, что либо $f^{(m)}(a) \vec{X}^m = 0$, либо разность $g(t) - g(0)$ в окрестности точки $t = 0$ имеет тот же знак, что и $(f^{(m)}(a) \vec{X}^m) t^m$. Поэтому,

окончательно, при любом $\vec{X} \in \vec{E}$ и вещественном t имеет место неравенство

$$(f^{(m)}(a) \vec{X}^m) t^m \leq 0. \quad (\text{III}, 7; 29)$$

При четном m это неравенство эквивалентно (III, 7; 27). Для доказательства первой части остается показать, что m не может быть нечетным. В самом деле, при нечетном m , полагая последовательно $t = +1$ и $t = -1$, из (III, 7; 29) получаем

$$f^{(m)}(a) \vec{X}^m = 0 \quad \text{для любого } \vec{X} \in \vec{E}. \quad (\text{III}, 7; 30)$$

Согласно лемме, относящейся к теореме 21₂, это соотношение влечет за собой равенство

$$f^{(m)}(a)(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) = 0 \quad \text{для любых } (\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) \in \vec{E}^m, \quad (\text{III}, 7; 31)$$

или $f^{(m)}(a) = 0$, что противоречит исходному предположению.

Замечание. Перед тем как перейти ко второй части теоремы, заметим следующее: может показаться, что если f имеет в точке a строгий относительный максимум, то форма порядка m : $f^{(m)}(a) \vec{X}^m$ является отрицательно определенной, т. е. имеет место неравенство

$$f^{(m)}(a) \vec{X}^m < 0 \quad \text{для } \vec{X} \neq \vec{0}. \quad (\text{III}, 7; 35)$$

Но это не так (за исключением, очевидно, случая, когда E одномерно). В самом деле, рассмотрим вещественную функцию двух переменных x, y :

$$f(x, y) = -(x^2 + y^4). \quad (\text{III}, 7; 36)$$

В начале координат она имеет строгий максимум. Поскольку это полином, его разложение по степеням x и y является разложением Маклорена, так что, обозначая через (\vec{X}, Y) вектор из $\vec{E} = \mathbb{R}^2$, получаем, что

$$\frac{1}{2!} f''(0, 0) (\vec{X}, Y)^2 = -X^2 \quad (\text{III}, 7; 37)$$

всегда ≤ 0 , но не является отрицательно определенной (это выражение обращается в нуль на векторах $(0, Y) \neq (0, 0)$).

2°) Докажем теперь вторую часть теоремы. Заметим прежде всего, что *условия второй части более сильные, чем заключение первой*. Иначе не может и быть. Если предположить, что $f^{(k)}(a) = 0$ при $k = 1, 2, \dots, m-1$, $f^{(m)}(a) \neq 0$ и что выполняется неравенство (III, 7; 27), то этого вовсе не достаточно для

того, чтобы обеспечить функции f относительный максимум в точке a .

В самом деле, рассмотрим функцию f двух вещественных переменных x, y :

$$f(x, y) = -x^2 + y^4. \quad (\text{III}, 7; 38)$$

Как мы видели для (III, 7; 36), имеет место соотношение (III, 7; 37), т. е. неравенство (III, 7; 27). Однако $f(0, 0) = 0$ и $f(0, y) > 0$ для $y \neq 0$, а, следовательно, f в начале координат относительного максимума не имеет. Поэтому мы ставим более сильные условия, указанные в формулировке теоремы:

$f^{(m)}(a) \vec{X}^m$ мажорируется на единичной сфере числом $-\delta < 0$. Это влечет за собой, в силу однородности, неравенство

$$f^{(m)}(a) \vec{X}^m \leq -\delta \|\vec{X}\|^m. \quad (\text{III}, 7; 39)$$

Последнее, в свою очередь, влечет за собой (III, 7; 35): $f^{(m)}(a)$ отрицательно определена. Если E конечномерно, то неравенство (III, 7; 39) эквивалентно неравенству (III, 7; 35). Действительно, в этом случае единичная сфера компактна и непрерывная функция $\vec{X} \rightarrow f^{(m)}(a) \vec{X}^m$, отрицательная в силу (III, 7; 35) на этой сфере, достигает на ней максимума, являющегося некоторым числом $-\delta < 0$. Однако для бесконечномерного пространства E это условие более сильное, чем (III, 7; 35).

Запишем разложение Тейлора порядка m в виде (III, 7; 10):

$$f(a + \vec{h}) = f(a) + \frac{f^{(m)}(a)}{m!} \vec{h}^m + \alpha(\vec{h}) \|\vec{h}\|^m, \quad (\text{III}, 7; 40)$$

где $\alpha(\vec{h})$ стремится к $\vec{0}$ вместе с \vec{h} . Тогда (III, 7; 39) дает

$$f(a + \vec{h}) \leq f(a) - \delta \|\vec{h}\|^m + \alpha(\vec{h}) \|\vec{h}\|^m. \quad (\text{III}, 7; 41)$$

Так как α стремится к $\vec{0}$, то для $\vec{h} \neq \vec{0}$ с достаточно малой нормой $f(a + \vec{h}) - f(a) < 0$ и, следовательно, в точке a функция f имеет строгий относительный максимум.

Естественно, аналогичная теорема имеет место для минимума. Отсюда получается правило, позволяющее выяснить, имеет ли f в точке a относительный максимум или минимум.

Правило. Прежде всего необходимо, чтобы $f'(a) = 0$. Затем ищется первое целое число m , такое, что $f^{(m)}(a) \neq 0$. Если m нечетно, то в точке a нет ни относительного максимума, ни относительного минимума. Если m четно, то рассматривается знак формы степени m : $\vec{X} \rightarrow f^{(m)}(a) \vec{X}^m$. Если она может принимать значения разных знаков, то в точке a нет ни относительного

максимума, ни относительного минимума¹⁾). Если она всегда ≤ 0 (соответственно ≥ 0), не будучи отрицательно определенной (соответственно положительно определенной), то без изучения производных порядка $>m$ ничего определенного заключить нельзя. Если же она отрицательно определенная (соответственно >0) и E конечномерно или же если имеет место более сильное условие $\sup_{\|\vec{X}\|=1} (f^{(m)}(a) \vec{X}^m) < 0$ (соответственно

$\inf_{\|\vec{X}\|=1} (f^{(m)}(a) \vec{X}^m) > 0$) для бесконечномерного E , то a является точкой строгого относительного максимума (соответственно минимума).

Из всего изложенного следует, что не существует простого условия, одновременно необходимого и достаточного для того, чтобы вещественная функция, определенная на $\Omega \subset E$, имела в точке a относительный максимум или минимум.

З а м е ч а н и я. 1°) Если функция f имеет в точке a все последовательные производные, равные нулю, и не является тождественно нулевой функцией, то изучение разложения Тейлора не позволяет выяснить вопроса о существовании в точке a относительного максимума или относительного минимума.

2°) Предположим, что в точке a функция имеет седло. Тогда можно попытаться узнать, в какой части области Ω в окрестности точки a имеет место неравенство $f(x) \geq f(a)$ и в какой части $f(x) \leq f(a)$. Здесь применимы те же правила. Если на множестве B единичной сферы $\|\vec{X}\|=1$ величина $f^{(m)}(a) \vec{X}^m$ не превосходит $-\delta < 0$, то во всем конусе $\{a + \lambda \vec{X}; \vec{X} \in B, \lambda > 0$ вещественно} в окрестности точки a имеет место неравенство $f(x) < f(a)$.

Частный случай вещественной функции f двух вещественных переменных x, y

При исследовании такой функции на экстремум надо прежде всего найти точки, в которых частные производные $p = \frac{\partial f}{\partial x}$ и $q = \frac{\partial f}{\partial y}$ равны нулю. Пусть (a, b) — такая точка. Рассмотр-

¹⁾ В этом случае говорят, что в точке a функция f имеет седло. Типичный пример такого седла дает функция двух вещественных переменных $f(x, y) = xy$. Ее графиком в \mathbb{R}^3 служит гиперболический параболоид $z = xy$. В начале координат функция f имеет седло. Сужением f на прямую $y = kx$ является функция $x \rightarrow kx^2$, имеющая строгий максимум в начале координат при $k < 0$ и строгий минимум при $k > 0$. Можно также говорить о седле и в том случае, когда m нечетно.

рим теперь разложение функции f по формуле Тейлора по степеням $x-a=X$ и $y-b=Y$. Это разложение начинается с

$$\frac{1}{2}(rX^2 + 2sXY + tY^2). \quad (\text{III}, 7; 42)$$

1-й случай: $rt - s^2 > 0$, r (и t) < 0 . Квадратичная форма (*III*, 7; 42) в этом случае отрицательно определенная, а, значит, f имеет в точке (a, b) строгий относительный максимум.

2-й случай: $rt - s^2 > 0$, r (и t) > 0 . Функция f имеет в точке (a, b) строгий относительный минимум.

3-й случай: $rt - s^2 < 0$. Квадратичная форма (*III*, 7; 42) принимает значения разных знаков, и, следовательно, мы имеем дело с *седлом*. Уравнение $rX^2 + 2sXY + tY^2 = 0$ задает две прямые D и D' на плоскости \mathbb{R}^2 , определяющие 4 попарно вертикальных угла. Обозначим их по порядку обхода через (1), (2), (3), (4). Предположим, например, что рассматриваемая квадратичная форма ≥ 0 в углах (1) и (3) и ≤ 0 в углах (2) и (4).

Пусть угол (1) определен в полярных координатах неравенствами $\rho \geq 0$ и $\alpha' \leq \varphi \leq \beta'$. Тогда каждый угол $\alpha' \leq \varphi \leq \beta'$, где $\alpha' > \alpha$ и $\beta' < \beta$, пересекает «единичную сферу» из \mathbb{R}^2 , т. е. окружность, по *компактной* дуге, на которой квадратичная форма имеет минимум $\delta > 0$. Следовательно, существует такое число $\varepsilon > 0$, что если положить $x=a+\rho \cos \varphi$, $y=b+\rho \sin \varphi$, $\rho > 0$, $\alpha' \leq \varphi \leq \beta'$, то для $0 < \rho \leq \varepsilon$ имеет место неравенство $f(x, y) > f(a, b)$.

Однако это число ε зависит от α' и β' . Поэтому невозможно, не привлекая к изучению производные f порядков > 2 в точке (a, b) , делать какие-либо выводы при ρ , стремящемся к 0, если одновременно с этим φ стремится к α или β . Тот же результат имеет место в угле (3). В углах (2) и (4) имеет место аналогичный результат, но с заменой неравенства $f(x, y) > f(a, b)$ на $f(x, y) < f(a, b)$.

Области плоскости \mathbb{R}^2 , в которых $f(x, y) > f(a, b)$ и $f(x, y) < f(a, b)$, разделены кривой $f(x, y) = f(a, b)$. Эта кривая в точке (a, b) имеет две ветви, касательные соответственно к D и D' . Однако, не привлекая к рассмотрению производные функции f порядков > 2 в точке (a, b) , невозможно сказать, в каком из углов находятся эти кривые.

4-й случай: $rt - s^2 = 0$, r или $t < 0$. Квадратичная форма равна квадрату некоторой линейной формы $\neq 0$, взятому с противоположным знаком. Она < 0 , но обращается в нуль на некоторой прямой D . Пусть $\varphi = \varphi_0 + k\pi$ — полярный угол этой прямой. Если $x=a+\rho \cos \varphi$, $y=b+\rho \sin \varphi$, $\rho \geq 0$, $\alpha' \leq \varphi \leq \beta'$, где α' и β' фиксированы, $\alpha' > \varphi_0$ и $\beta' < \varphi_0 + \pi$, то существует такое число $\varepsilon > 0$, что для $0 < \rho \leq \varepsilon$ имеет место неравенство

$f(x, y) < f(a, b)$. Аналогичный результат справедлив для $\alpha' > \varphi_0 + \pi$, $\beta' < \varphi_0 + 2\pi$.

К сказанному ничего добавить нельзя. В частности, без привлечения производных порядков > 2 нельзя сказать, имеет или нет функция f в точке (a, b) относительный максимум.

5-й случай: $rt - s^2 = 0$, r или $t > 0$. Результат аналогичен предыдущему, надо только заменить $f(x, y) < f(a, b)$ на $f(x, y) > f(a, b)$.

6-й случай: $r = s = t = 0$. Без привлечения производных функции f в точке (a, b) порядка > 2 никакого вывода о поведении функции f в точке (a, b) сделать невозможно.

Заключение. Проведенные исследования показывают, что различных вариантов поведения функции более одной переменной значительно больше, чем это имеет место для функции одной переменной.

Применение формулы Тейлора к изучению расположения гиперповерхности по отношению к касательной гиперплоскости

Пусть f — дифференцируемая вещественная функция, определенная на $\Omega \subset E$. Тогда $y = f(x)$ является уравнением дифференцируемого многообразия пространства $E \times \mathbb{R}$. Такое многообразие называют *дифференцируемой гиперповерхностью*. Ее касательная гиперплоскость в точке $A = (a, f(a))$ определяется уравнением (III, 3; 19₃). Расположение гиперповерхности по отношению к касательной гиперплоскости в окрестности точки A определяется знаком разности

$$\overrightarrow{f(x) - f(a)} - \overrightarrow{f'(a)(x - a)} \quad (\text{III}, 7; 43)$$

для x , близких к a .

Но $x \rightarrow \overrightarrow{f(x) - f(a)} - \overrightarrow{f'(a)(x - a)}$ является дифференцируемой вещественной функцией, производная которой в точке a равна нулю. Все сводится к тому, чтобы выяснить, имеет ли эта функция в точке a экстремум какого-нибудь типа или седло. Тем самым мы возвращаемся к теореме 24. В силу приведенного выше заключения, наличие седла является обстоятельством частным, а не исключительным и, следовательно, кроме случая одномерного E , гиперповерхность обычно пересекает свою касательную гиперплоскость, даже если $f''(a) \neq 0$. В частном случае функции f двух переменных x, y поверхность локально расположена с одной стороны от касательной плоскости, если $rt - s^2 > 0$, и пересекает ее, если $rt - s^2 \leq 0$.

§ 8. ТЕОРЕМА О НЕЯВНОЙ ФУНКЦИИ. ПОСТАНОВКА ЗАДАЧИ

Пусть f — отображение множества E в множество F . Тогда, если b — некоторая точка F , можно поставить задачу отыскания множества тех x из E , для которых $f(x) = b$, т. е. задачу отыскания прообраза $f^{-1}(\{b\})$. Поставленная задача называется задачей о *решении уравнения* $f(x) = b$. Пусть теперь f — отображение произведения множеств $E \times F$ в множество G , и пусть c — некоторая точка множества G .

Рассмотрим уравнение

$$f(x, y) = c. \quad (\text{III}, 8; 1)$$

Может случиться, что для заданного x уравнение $f(x, y) = c$ относительно y имеет решение, и притом единственное, при любом заданном значении x . В таком случае это уравнение определяет y как некоторую функцию $g(x)$ переменной x . Эта функция называется *неявной функцией*, определенной предыдущим уравнением. Она характеризуется следующим свойством:

$$f(x, g(x)) = c. \quad (\text{III}, 8; 2)$$

Можно еще сказать, что для $(x, y) \in E \times F$ соотношение (III, 8; 2) равносильно соотношению $y = g(x)$.

Естественно, что такие счастливые совпадения случаются редко. Чаще всего для некоторых значений x решения относительно y не существует, для других же значений x имеется несколько решений, а то и бесконечное множество их. Мы изучим следующий частный случай: E, F, G — топологические пространства. Имеется *частное решение* рассматриваемого уравнения: $x = a, y = b$. Предлагается выяснить, существует ли и притом единственны ли решение данного уравнения относительно y для x , *достаточно близких* к a , при условии, что это решение достаточно близко к b . Если это так, то можно будет определить неявную функцию $y = g(x)$, исходя из данного уравнения, по *крайней мере в некоторой окрестности точки* (a, b) . Геометрический смысл рассматриваемой задачи станет особенно понятным, если мы предположим, что каждое из пространств E, F и G является полем вещественных чисел. В этом случае уравнение $f(x, y) = 0$ определяет некоторую кривую в \mathbb{R}^2 , и мы хотим выразить эту кривую в обычной форме, представив y как функцию от x по крайней мере в некоторой окрестности точки (a, b) ¹.

¹⁾ Легко видеть, почему такие сужения на окрестности неизбежны: кривая $f(x, y) = 0$, вообще говоря, не может быть полностью задана в виде $y = g(x)$.

Существование неявной функции

Теорема 25. Пусть E — топологическое пространство, F и G — аффинные нормированные пространства, Ω — открытое множество из $E \times F$ и (a, b) — точка Ω . Пусть f — непрерывное отображение Ω в G и $f(a, b) = c$. Предположим, что при каждом фиксированном x функция f имеет частную производную $\frac{\partial f}{\partial y}(x, y) \in \mathcal{L}(\vec{F}; \vec{G})$ и что частная производная функция $\frac{\partial f}{\partial y}$ является непрерывным отображением Ω в $\mathcal{L}(\vec{F}; \vec{G})$. Предположим, кроме того, что $Q = \frac{\partial f}{\partial y}(a, b)$ является обратимым отображением пространства \vec{F} на пространство \vec{G} , т. е. является биекцией, имеющей обратную биекцию Q^{-1} , линейную и непрерывную. Предположим, наконец, что F полно. Тогда существуют такие открытые множества A и B в пространствах E и F , содержащие a и b соответственно, что для любого x из A уравнение (III, 8; 1) относительно y имеет решение в B , и притом единственное. Это решение y является функцией x , и так определенная функция $y = g(x)$ является непрерывным отображением A в B .

Доказательство. Напомним сначала бегло условия теоремы. Ищется решение уравнения относительно y , налагаются некоторые ограничения на частную производную по y , пространство F переменной y предполагается полным. Подвергнем уравнение трем последовательным преобразованиям:

1°) Пусть $\vec{\Omega}_1$ — множество точек $(x, \overrightarrow{y - b})$, где $(x, y) \in \Omega$. Положим $y = b + \vec{Y}$ и обозначим через \vec{f}_1 отображение множества $\vec{\Omega}_1$ в пространство \vec{G} , определенное формулой

$$\overrightarrow{f_1(x, \vec{Y})} = \overrightarrow{f(x, b + \vec{Y})} - \vec{c}^{-1}. \quad (\text{III}, 8; 3)$$

Уравнение (III, 8; 1) эквивалентно уравнению

$$\overrightarrow{f_1(x, \vec{Y})} = \vec{0} \quad (\text{III}, 8; 4)$$

в окрестности частного решения $(a, \vec{0})$.

1) На практике F не обязательно должно быть векторным пространством, в то время как G — почти всегда векторное пространство. Обычно рассматривается уравнение $\vec{f}(x, y) = \vec{0}$. Мы считаем G аффинным пространством, а c произвольным для того, чтобы сохранить симметрию между F и G . Во всяком случае, можно всегда перейти к случаю векторных F и G с $\vec{b} = \vec{0}$ и $\vec{c} = \vec{0}$.

2°) Рассмотрим функцию \vec{f}_2 , определенную формулой

$$Q^{-1} \circ \vec{f}_1: (x, \vec{Y}) \rightarrow Q^{-1}(f_1(x, \vec{Y})). \quad (\text{III}, 8; 5)$$

Это некоторое отображение $\vec{\Omega}_1$ в \vec{F} , обладающее свойствами, аналогичными свойствам отображения f . Его частная производная по второй переменной в начале координат, в силу следствия теоремы 11, определяется формулой

$$\frac{\partial f_2}{\partial \vec{Y}}(\vec{a}, \vec{0}) = Q^{-1} \circ Q = I \in \mathcal{L}(\vec{F}; \vec{F}), \quad (\text{III}, 8; 6)$$

где I — тождественное отображение. Уравнение (III, 8; 1) эквивалентно новому уравнению

$$\vec{f}_2(x, \vec{Y}) = \vec{0}, \quad (\text{III}, 8; 7)$$

ибо от (III, 8; 4) к (III, 8; 7) можно перейти с помощью Q^{-1} , а от (III, 8; 7) к (III, 8; 4) с помощью Q .

3°) Рассмотрим, наконец, функцию $\vec{\Phi}$, определенную формулой

$$\vec{\Phi}(\lambda, \vec{Y}) = \vec{Y} - f_2(\lambda, \vec{Y}); \quad (\text{III}, 8; 8)$$

$\vec{\Phi}$ является отображением $\vec{\Omega}_1$ в \vec{F} , обладающим свойствами, аналогичными свойствам f всюду, кроме одной особой точки: его частное производное отображение по второй переменной в начале координат равно нулю, так как $I - I = 0$. Уравнение (III, 8; 1) теперь эквивалентно уравнению

$$\vec{\Phi}(\lambda, \vec{Y}) = \vec{Y}, \quad \lambda = x, \quad \vec{Y} = \overrightarrow{y - b}, \quad (\text{III}, 8; 9)$$

которое нужно решать в окрестности $\lambda = a, \vec{Y} = \vec{0}$.

Теперь видно, что мы находимся в условиях, напоминающих условия теорем 46 и 46_2 гл. II (*теоремы о неподвижной точке*). Пусть A_1 и B_1 — такие окрестности точек a и $\vec{0}$ в E и \vec{F} соответственно, что $A_1 \times B_1 \subset \vec{\Omega}_1$. Если пространства E и Λ из теоремы 46_2 заменить соответственно на B_1 и A_1 , то функция $\vec{\Phi}$ будет непрерывной и тем более раздельно непрерывной на $A_1 \times B_1$. Однако для применения этих теорем нужны условия, которые здесь нельзя непосредственно проверить и которые требуют еще некоторого изменения исходных данных задачи.

1°) Отображение Φ должно быть отображением $A_1 \times \vec{B}_1$ в \vec{B}_1 ; однако оно отображает $A_1 \times \vec{B}_1$ в \vec{F} и не обязательно в \vec{B}_1 . Именно это вынуждает нас заменить A_1 и \vec{B}_1 на две меньшие окрестности A_2 и \vec{B}_2 точек a и $\vec{0}$ с тем, чтобы сужение Φ на $A_2 \times \vec{B}_2$ отображало это множество в \vec{B}_2 .

2°) Теперь надо проверить, что для фиксированного λ в A_2 отображение $\vec{\Phi}_\lambda: \vec{Y} \rightarrow \vec{\Phi}(\lambda, \vec{Y})$ окрестности \vec{B}_2 в себя является сжатием с коэффициентом сжатия $k < 1$, не зависящим от выбора λ в A_2 .

3°) Надо, чтобы метрическое пространство \vec{B}_2 было полным.

Займемся сначала условием 2°). Поскольку частная производная $\frac{\partial \vec{\Phi}}{\partial \vec{Y}}$ непрерывна и равна нулю в $(a, \vec{0})$, можно найти такую окрестность A'_2 точки a в A_1 и такой замкнутый шар¹⁾ \vec{B}_2 в \vec{B}_1 с центром в нуле радиуса $\beta > 0$, что из соотношений $\lambda \in A'_2$ и $\vec{Y} \in \vec{B}_2$ следует неравенство

$$\left\| \frac{\partial \vec{\Phi}}{\partial \vec{Y}} (\lambda, \vec{Y}) \right\| \leq \frac{1}{2}. \quad (\text{III}, 8; 10)$$

Применим теперь теорему о конечных приращениях. Это возможно, поскольку шар \vec{B}_2 выпуклый, и, следовательно, если \vec{Y}' и \vec{Y}'' лежат в этом шаре, то в нем лежит и весь отрезок $[\vec{Y}', \vec{Y}'']$. Получим неравенство

$$\|\vec{\Phi}(\lambda, \vec{Y}') - \vec{\Phi}(\lambda, \vec{Y}'')\| \leq \frac{1}{2} \|\vec{Y}' - \vec{Y}''\|. \quad (\text{III}, 8; 11)$$

Таким образом, условие 2°) выполняется с $k = 1/2$ для шаров A'_2 и \vec{B}_2 или для меньших множеств.

Займемся теперь условием 1°). Нам надо выбрать A_2 и \vec{B}_2 так, чтобы $\vec{\Phi}$ было отображением $A_2 \times \vec{B}_2$ в \vec{B}_2 . Окрестности A'_2 и \vec{B}_2 уже выбраны. Учитывая непрерывность функции Φ и тот факт, что $\vec{\Phi}(a, \vec{0}) = \vec{0}$, мы можем найти такую окрестность $A_2 \subset A'_2$ точки a , в которой имеет место неравенство

$$\|\vec{\Phi}(\lambda, \vec{0})\| \leq \frac{\beta}{2} \quad \text{для } \lambda \in A_2. \quad (\text{III}, 8; 12)$$

¹⁾ Мы увидим несколько ниже, почему B_2 должно быть замкнутым.

При этих условиях для $\lambda \in A_2$ и $\|\vec{Y}\| \leq \beta$ имеет место неравенство

$$\|\vec{\Phi}(\lambda, \vec{Y})\| \leq \|\vec{\Phi}(\lambda, \vec{0})\| + \|\vec{\Phi}(\lambda, \vec{Y}) - \vec{\Phi}(\lambda, \vec{0})\| \leq \frac{\beta}{2} + \frac{1}{2} \|\vec{Y}\| \leq \beta. \quad (\text{III}, 8; 13)$$

Это неравенство доказывает, что функция $\vec{\Phi}$ отображает $A_2 \times \vec{B}_2$ в \vec{B}_2 . Наконец, условие 3°) теперь выполнено, поскольку, в силу теоремы 43 гл. II, \vec{B}_2 , как замкнутая часть полного метрического пространства F , является полным метрическим пространством. Итак, мы находимся в условиях применимости теорем 46 и 46₂. Следовательно, для заданного λ в A_2 имеется, и притом единственное, решение уравнения (III, 8; 9) относительно \vec{Y} , такое, что $\|\vec{Y}\| \leq \beta$ и \vec{Y} является непрерывной функцией λ , $\vec{Y} = \vec{\Psi}(\lambda)$.

Уравнение (III, 8; 1) эквивалентно уравнению (III, 8; 9) и, следовательно, имеет, и притом единственное, решение относительно y в шаре B_2 пространства F с центром в b радиуса β для любого заданного x из A_2 . Кроме того, если мы обозначим через g отображение, которое каждой точке x из A_2 ставит в соответствие единственное решение y из B_2 , $y = g(x)$, то $g(x) = b + \vec{\Psi}(x)$ и, следовательно, g является непрерывным отображением A_2 в B_2 .

Заметим, что определение окрестностей A_2 и B_2 проводилось в два приема. Сначала были найдены достаточно малые A'_2 и B'_2 . Естественно, их можно по желанию еще уменьшить. Однако, как только B_2 фиксировано, достаточно малое множество A_2 выбирается уже в зависимости от B_2 . Результат a posteriori представляется очевидным, поскольку функция g непрерывна. Из непрерывности функции g следует также, что если B является внутренностью $\overset{\circ}{B}_2$ множества B_2 , то существует открытое множество¹⁾ в E , содержащее a , такое, что из $x \in A$ следует $g(x) \in B$, и теорема доказана.

З а м е ч а н и е. Рассмотрим уравнение более общего вида:

$$f(x, y) = z, \quad (\text{III}, 8; 14)$$

в котором ищется y , близкое к b , как функция x и z , близких к токам a и c , таким, что $f(a, b) = c$.

¹⁾ Для проведения доказательства мы были вынуждены брать замкнутый шар B_2 , ибо нам нужна была замкнутость шара \vec{B}_2 для того, чтобы он был полным! Но удобнее формулировать теорему для открытых множеств A и B .

Это уравнение сводится к уравнению теоремы 25:

$$\vec{f}_1((x, z), y) = \vec{0}, \text{ где } \vec{f}_1((x, z), y) = \overrightarrow{f(x, y) - z}; \quad (\text{III}, 8; 15)$$

f_1 является отображением некоторого открытого множества из $(E \times G) \times F$ в \vec{G} и ищется y , близкое к b , как функция точки (x, z) , близкой к точке (a, c) , такой, что $\vec{f}_1((a, c), b) = \vec{0}$.

Дифференцируемость неявной функции

Предположим теперь, что функция f дифференцируема, и выясним, влечет ли это за собой дифференцируемость неявной функции g , определенной заданной функцией f . При этом мы несколько изменим условия теоремы: мы будем предполагать, что неявная функция существует и непрерывна, даже если существование и непрерывность не следуют непосредственно из теоремы 25 (мы предполагаем существование производной только в точке (a, b)). Кроме того, пространство E будет предполагаться аффинным для того, чтобы можно было говорить о дифференцируемости функции g .

Теорема 26. Пусть E, F, G — три аффинных нормированных пространства, Ω — открытое множество из $E \times F$, f — отображение Ω в G . Пусть A и B — открытые множества пространств E и F соответственно, $A \times B \subset \Omega$ и g является отображением A в B , тождественно удовлетворяющим соотношению (III, 8; 2). Пусть, далее, отображение f дифференцируемо в точке (a, b) , $b = g(a)$, и его частные производные P и Q в рассматриваемой точке являются линейными непрерывными отображениями \vec{E} и \vec{F} в \vec{G} соответственно, причем отображение Q обратимо. Тогда, если отображение g непрерывно в точке a , то оно дифференцируемо в точке a и его производное отображение задается формулой

$$g'(a) = -Q^{-1} \circ P = -\left(\frac{\partial f}{\partial y}(a, b)\right)^{-1} \circ \left(\frac{\partial f}{\partial x}(a, b)\right). \quad (\text{III}, 8; 16)$$

Доказательство. Если заранее предположить существование производной отображения g , то формула (III, 8; 16) получается сразу. В самом деле, мы можем продифференцировать тождество (III, 8; 2), пользуясь теоремой о сложной функции, и получить уравнение

$$\frac{\partial f}{\partial x}(a, b) + \frac{\partial f}{\partial y}(a, b) \circ g'(a) = 0 \in \mathcal{L}(\vec{E}; \vec{G}), \quad (\text{III}, 8; 17)$$

$$\text{или } P + Q \circ g'(a) = 0,$$

откуда путем композиции слева с Q^{-1} вытекает (III, 8; 16). Но доказательство получается более сложным, так как мы *не знаем заранее*, дифференцируема ли функция g в точке a . Пусть $\overrightarrow{dx} = \overrightarrow{x - a}$ — некоторое приращение x , а $\overrightarrow{\Delta y}$ — соответствующее приращение $y = g(x)$. Из того факта, что g является неявной функцией, определенной рассматриваемым уравнением, и что функция f дифференцируема в (a, b) , следует формула

$$\overrightarrow{0} = \overrightarrow{\Delta f} = P \overrightarrow{dx} + Q \overrightarrow{\Delta y} + \overrightarrow{\alpha} (\|\overrightarrow{dx}\| + \|\overrightarrow{\Delta y}\|). \quad (\text{III, 8; 18})$$

Вектор $\overrightarrow{\alpha}$ зависит от \overrightarrow{dx} и $\overrightarrow{\Delta y}$, т. е. в конечном счете только от \overrightarrow{dx} . Он стремится к $\overrightarrow{0}$, если \overrightarrow{dx} и $\overrightarrow{\Delta y}$ стремятся к $\overrightarrow{0}$, т. е. при \overrightarrow{dx} , стремящемся к $\overrightarrow{0}$, ибо, в силу непрерывности g в точке a , $\overrightarrow{\Delta y}$ стремится к $\overrightarrow{0}$ вместе с \overrightarrow{dx} . Отсюда вытекает следующий результат для $\overrightarrow{\Delta y}$:

$$Q \overrightarrow{\Delta y} = -P \overrightarrow{dx} - \overrightarrow{\alpha} (\|\overrightarrow{dx}\| + \|\overrightarrow{\Delta y}\|),$$

или

$$\overrightarrow{\Delta y} = -(Q^{-1} \circ P) \overrightarrow{dx} - (Q^{-1} \overrightarrow{\alpha}) (\|\overrightarrow{dx}\| + \|\overrightarrow{\Delta y}\|). \quad (\text{III, 8; 19})$$

Теперь необходимо избавиться от $\overrightarrow{\Delta y}$ в правой части. Это можно сделать с помощью некоторых оценок. С этой целью напишем сначала неравенство

$$\|\overrightarrow{\Delta y}\| \leq \|Q^{-1}\| \|P\| \|\overrightarrow{dx}\| + \|Q^{-1}\| \|\overrightarrow{\alpha}\| \|\overrightarrow{dx}\| + \|Q^{-1}\| \|\overrightarrow{\alpha}\| \|\overrightarrow{\Delta y}\|. \quad (\text{III, 8; 20})$$

Мы видели, что при \overrightarrow{dx} , стремящемся к $\overrightarrow{0}$, $\overrightarrow{\alpha}$ также стремится к $\overrightarrow{0}$. Следовательно, можно выбрать \overrightarrow{dx} настолько малым, чтобы коэффициент $\|Q^{-1}\| \|\overrightarrow{\alpha}\|$ не превосходил $1/2$. Перенося затем налево последний член, получим следующее неравенство, имеющее место для всех достаточно малых $\|\overrightarrow{dx}\|$:

$$\frac{1}{2} \|\overrightarrow{\Delta y}\| \leq \left(\|Q^{-1}\| \|P\| + \frac{1}{2} \right) \|\overrightarrow{dx}\|,$$

или

$$\|\overrightarrow{\Delta y}\| \leq (2 \|Q^{-1}\| \|P\| + 1) \|\overrightarrow{dx}\|, \quad (\text{III, 8; 21})$$

или

$$\|\overrightarrow{dx}\| + \|\overrightarrow{\Delta y}\| \leq 2 (\|Q^{-1}\| \|P\| + 1) \|\overrightarrow{dx}\| = k \|\overrightarrow{dx}\|,$$

где k — некоторая постоянная. Используя теперь эту оценку для последнего члена правой части равенства (III, 8; 19), окончательно получаем

$$\vec{\Delta y} = -(Q^{-1} \circ P) \vec{dx} + \vec{\beta} \|\vec{dx}\|, \text{ где } \|\vec{\beta}\| \leq k \|Q^{-1}\| \|\vec{a}\|. \quad (\text{III}, 8; 22)$$

Здесь $\vec{\beta}$ стремится к $\vec{0}$, когда \vec{dx} стремится к $\vec{0}$, чем и заканчивается доказательство теоремы.

Замечание. В дифференциальных обозначениях без труда приходим к следующему правилу:

Правило. Для нахождения дифференциала функции u как функции x , неявно определяемой уравнением $f(x, u) = c$, следует продифференцировать это уравнение, что дает

$$P \vec{dx} + Q \vec{dy} = 0, \quad (\text{III}, 8; 23)$$

а затем разрешить полученное соотношение относительно \vec{dy} :

$$\vec{dy} = -(Q^{-1} \circ P) \vec{dx}. \quad (\text{III}, 8; 24)$$

Дифференцируемость функции $u \rightarrow u^{-1}$ на $\mathcal{L}(\vec{F}; \vec{G})$

Для того чтобы иметь возможность в полной мере использовать теоремы 25 и 26, надо знать, как изменяется $\left(\frac{\partial f}{\partial u}(x, y)\right)^{-1}$ при изменении (x, y) . Точно так же, как при обосновании дифференцируемости сложной функции (теорема 19) была использована дифференцируемость произведения (теорема 18), для обоснования дифференцируемости неявной функции будет нужна производная обратной функции.

Теорема 27. Пусть \vec{F}, \vec{G} — векторные нормированные пространства, и пусть \mathcal{U} (соответственно \mathcal{U}^{-1}) — множество обратимых элементов пространства $\mathcal{L}(\vec{F}; \vec{G})$ (соответственно $\mathcal{L}(\vec{G}; \vec{F})$). Тогда:

1°) Биекция $u \rightarrow u^{-1}$ множества \mathcal{U} на множество \mathcal{U}^{-1} является гомеоморфизмом.

2°) Если пространства \vec{F} и \vec{G} полны, то множества \mathcal{U} и \mathcal{U}^{-1} открыты, а биекция $u \rightarrow u^{-1}$ дифференцируема вместе со своей обратной биекцией и ее производная в точке $u \in \mathcal{L}(\vec{F}; \vec{G})$ является отображением

$$du \rightarrow -u^{-1} \circ du \circ u^{-1} \quad (\text{III}, 8; 25)$$

пространства $\mathcal{L}(\vec{F}; \vec{G})$ в пространство $\mathcal{L}(\vec{G}; \vec{F})$.

Первое доказательство. Для простоты во всех случаях будем предполагать пространства \vec{F} и \vec{G} полными. Тогда, как мы видели в теореме 62 гл. II, если $u_0 \in \mathcal{U}$, то всякий элемент из $\mathcal{L}(\vec{F}; \vec{G})$, принадлежащий открытому шару с центром в u_0 радиуса $\|u_0^{-1}\|^{-1}$, будет находиться в \mathcal{U} , а, следовательно, \mathcal{U} открыто. Точно такие же рассуждения можно провести для множества \mathcal{U}^{-1} , меняя ролями \vec{F} и \vec{G} .

Покажем теперь, что отображение $u \rightarrow u^{-1}$ дифференцируемо в точке $u_0 \in \mathcal{U}$. Это отображение непрерывно и, поскольку обратное к нему отображение также непрерывно (стоит лишь поменять ролями \vec{F} и \vec{G}), оно является гомеоморфизмом, дифференцируемым вместе с обратным гомеоморфизмом. Формула (II, 14; 26) дает (для $\|du\| < \|u_0^{-1}\|^{-1}$)

$$(u_0 + du)^{-1} = u_0^{-1} - u_0^{-1} du u_0^{-1} + u_0^{-1} du u_0^{-1} du u_0^{-1} - \dots \quad (\text{III, 8; 26})$$

Отсюда получается неравенство

$$\begin{aligned} &\|(u_0 + du)^{-1} - u_0^{-1}\| + u_0^{-1} du u_0^{-1} \| \leq \\ &\leq \|u_0^{-1}\|^3 \|du\|^2 (1 + \|u_0^{-1}\| \|du\| + (\|u_0^{-1}\| \|du\|)^2 + \dots) = \\ &= \frac{\|u_0^{-1}\|^3 \|du\|^2}{1 - \|u_0^{-1}\| \|du\|} = \frac{\|u_0^{-1}\|^2 \|du\|^2}{\|u_0^{-1}\|^{-1} - \|du\|}. \end{aligned} \quad (\text{III, 8; 27})$$

При $\|du\|$, стремящемся к 0, это выражение мажорируется постоянной, умноженной на $\|du\|^2$, и, следовательно, является бесконечно малой более высокого порядка, чем $\|du\|$. Поскольку $du \rightarrow -u_0^{-1} du u_0^{-1}$ является линейным непрерывным отображением $\mathcal{L}(\vec{F}; \vec{G})$ в $\mathcal{L}(\vec{G}; \vec{F})$, оно, в силу определения (III, 3; 13)¹⁾, является производной отображения $u \rightarrow u^{-1}$ в точке u_0 .

Второе доказательство. Ограничимся снова случаем, когда \vec{F} и \vec{G} полны. Тогда такими же будут пространства

1) Заметим, что если \vec{F} и \vec{G} произвольны, то \mathcal{U} и \mathcal{U}^{-1} могут оказаться пустыми! Например, если \vec{F} и \vec{G} конечномерны, то обратные отображения \vec{F} на \vec{G} существуют лишь в том случае, когда \vec{F} и \vec{G} имеют одинаковую размерность.

Заметим также, что если \mathcal{U} не пусто, то F не может быть полным, если не полно G , и обратно, ибо каждый элемент из \mathcal{U} устанавливает взаимно однозначное соответствие между \vec{F} и \vec{G} , сохраняющее сходимость и последовательности Коши.

$\mathcal{L}(\vec{F}; \vec{G})$ и $\mathcal{L}(\vec{G}; \vec{F})$ (теорема 50 гл. II). Пусть $u_0 \in \mathcal{U}$. Элемент v из $\mathcal{L}(\vec{G}; \vec{F})$ называется *обратным справа* к элементу $u \in \mathcal{L}(\vec{F}; \vec{G})$, если $u \circ v = I$. Такой элемент v существует не всегда (например, для $u=0$ не существует обратного справа элемента). Может существовать бесконечное множество таких элементов. Однако, если элемент u обратим, то существует лишь один такой элемент, а именно u^{-1} , ибо из $u \circ v = I$ умножением слева на u^{-1} получаем $v = u^{-1}$.

Поиски элемента, обратного справа к элементу u , приводят к решению уравнения $uv = I$ относительно v . Мы можем, следовательно, применить теорему о неявной функции. Отображение $(u, v) \rightarrow uv$, как билинейное непрерывное отображение, непрерывно дифференцируемо (теорема 9₂). Его частная производная по v в точке (u_0, v_0) , $v_0 = u_0^{-1}$, является линейным непрерывным отображением $V \rightarrow W = u_0 V$ ¹⁾ пространства $\mathcal{L}(\vec{G}; \vec{F})$ в пространство $\mathcal{L}(\vec{G}; \vec{G})$ (формула (II, 3; 34)). Это отображение обратимо, и обратное к нему отображение является непрерывным отображением $W \rightarrow V = u_0^{-1} W$ пространства $\mathcal{L}(\vec{G}; \vec{G})$ в пространство $\mathcal{L}(\vec{G}; \vec{F})$. Теорема 25 показывает, что тогда существует открытая окрестность $(\mathcal{U}_0)_n$ элемента u_0 в пространстве $\mathcal{L}(\vec{F}; \vec{G})$ и открытая окрестность $(\mathcal{U}_0)_n^{-1}$ элемента u_0^{-1} в пространстве $\mathcal{L}(\vec{G}; \vec{F})$, такие, что каждый элемент u окрестности $(\mathcal{U}_0)_n$ имеет единственный обратный справа элемент u_n^{-1} , лежащий в $(\mathcal{U}_0)_n^{-1}$. Кроме того, согласно теореме 26, отображение $u \rightarrow u_n^{-1}$ является непрерывным отображением $(\mathcal{U}_0)_n$ в $(\mathcal{U}_0)_n^{-1}$, дифференцируемым в точке u_0 . То же самое рассуждение можно провести для обратного слева элемента $w \in \mathcal{L}(\vec{G}; \vec{F})$ к элементу $u \in \mathcal{L}(\vec{F}; \vec{G})$, удовлетворяющего условию $wu = I$. Точно так же определяются открытые множества $(\mathcal{U}_0)_l$ и $(\mathcal{U}_0)_l^{-1}$ и отображение $u \rightarrow u_l^{-1}$ множества $(\mathcal{U}_0)_l$ в $(\mathcal{U}_0)_l^{-1}$. Обозначим теперь $\mathcal{U}_0 = (\mathcal{U}_0)_n \cap (\mathcal{U}_0)_l$. Это некоторая открытая окрестность точки u_0 . Если $u \in \mathcal{U}_0$, то одновременно существует обратный слева элемент u_l^{-1} в $(\mathcal{U}_0)_l^{-1}$ и обратный справа u_n^{-1} в $(\mathcal{U}_0)_n^{-1}$. Однако, если некоторый элемент имеет одновременно обратный справа и обратный слева, то они совпадают и

¹⁾ В дифференциальных обозначениях $d(uv) = du v_0 + u_0 dv$, и частная производная по v в точке (u_0, v_0) есть отображение $dv \rightarrow u_0 dv$.

рассматриваемый элемент обратим, ибо из $uu_{\pi}^{-1}=I$ и $u_{\pi}^{-1}u=I$ следует $u_{\pi}^{-1}uu_{\pi}^{-1}=(u_{\pi}^{-1}u)u_{\pi}^{-1}=u_{\pi}^{-1}$ и точно так же $u_{\pi}^{-1}(uu_{\pi}^{-1})=u_{\pi}^{-1}$.

Мы видим теперь, что каждый элемент окрестности \mathcal{U}_0 обратим. Отсюда следует, что \mathcal{U} является окрестностью каждой своей точки, т. е. является открытым множеством. Меняя ролями \vec{F} и \vec{G} , получаем тот же результат для множества \mathcal{U}^{-1} . Далее, отображение $u \rightarrow u^{-1}$, совпадающее в окрестности \mathcal{U}_0 точки u_0 с отображением $u \rightarrow u_{\pi}^{-1}$, непрерывно в точке u_0 , а, следовательно, и всюду на \mathcal{U} . Его обратное отображение обладает тем же свойством (в чем легко убедиться, меняя ролями \vec{F} и \vec{G}), а, значит, мы имеем гомеоморфизм \mathcal{U} на \mathcal{U}^{-1} . Наконец, по той же самой причине, это отображение дифференцируемо вместе с обратным отображением.

Вычисление производной выполняется по правилу (III, 8; 23). Дифференцируя $uv=I$, получаем

$$du \circ v + u \circ dv = 0. \quad (\text{III, 8; 31})$$

Умножая слева на u^{-1} , находим

$$dv = -u^{-1} \circ du \circ v = -u^{-1} \circ du \circ u^{-1}. \quad (\text{III, 8; 32})$$

Третье доказательство. Это доказательство пригодно лишь в случае конечномерных пространств \vec{F} и \vec{G} . Тогда их следует считать пространствами одинаковой размерности $n^1)$. Выберем в каждом из них систему координат. Тогда их можно отождествить с пространством \mathbb{K}^n , а $\mathcal{L}(\vec{F}; \vec{G})$ и $\mathcal{L}(\vec{G}; \vec{F})$ отождествить с векторным пространством K^{n^2} квадратных матриц из n строк и n столбцов. Пусть \mathfrak{M} — такая матрица, а m_{ij} , $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$, — ее элементы. Функция $\mathfrak{M} \rightarrow f(\mathfrak{M})$ принадлежит классу C^m тогда и только тогда, когда она является функцией n^2 скалярных переменных m_{ij} , принадлежащей классу C^m . Определитель $\mathfrak{M} \rightarrow \det \mathfrak{M}$ является полиномом относительно m_{ij} и, следовательно, является скалярной бесконечно дифференцируемой функцией от \mathfrak{M} . Множество \mathcal{U} обратимых матриц является прообразом множества \mathbb{C}^0 при этом непрерывном отображении, ибо матрица обратима тогда и только тогда, когда ее определитель $\neq 0$. Так как множество \mathbb{C}^0 открыто в \mathbb{K} , то \mathcal{U} также открыто.

Матрица \mathfrak{M}^{-1} , обратная к матрице \mathfrak{M} , состоит из элементов $\frac{M_{ji}}{\det \mathfrak{M}}$, где M_{ji} — алгебраические дополнения матрицы \mathfrak{M} ; M_{ji}

¹⁾ См. примечание на стр. 301.

являются полиномами относительно m_{ij} , а, следовательно, $\frac{M_{ji}}{\det \mathfrak{M}}$ является скалярной бесконечно дифференцируемой функцией на \mathcal{U} . Соответствие $\mathfrak{M} \rightarrow \mathfrak{M}^{-1}$ является отображением множества обратимых матриц \mathcal{U} в \mathbb{K}^n , составляющие которого являются скалярными функциями класса C^∞ , и, следовательно, оно само принадлежит классу C^∞ . Поскольку оно совпадает со своим обратным отображением, мы имеем гомеоморфизм класса C^∞ . Тем самым мы доказали нечто большее, чем утверждалось в теореме: по этому поводу см. замечание к приведенному ниже следствию.

Для того чтобы подсчитать производную отображения $u \rightarrow u^{-1}$, лучше всего воспользоваться концом второго доказательства. Заметим, что в *конечномерном случае*, если u имеет обратный справа элемент u_n^{-1} , то, переходя к матрицам, получим $\mathfrak{M}\mathfrak{M}_n^{-1} = I$. Отсюда следует, что $(\det \mathfrak{M})(\det \mathfrak{M}_n^{-1}) = 1$, а, значит, $\det \mathfrak{M} \neq 0$, и мы получаем, что элемент u обратим (и его обратным является элемент u_n^{-1}).

Следствие. Пусть $x \rightarrow u(x)$ — дифференцируемое отображение открытого множества Ω диффинного пространства E в пространство $\mathcal{L}(\vec{F}; \vec{G})$ (\vec{F} и \vec{G} — пространства Банаха).

Если для каждого $x \in \Omega$ отображение $u(x)$ является обратимым элементом пространства $\mathcal{L}(\vec{F}; \vec{G})$, то $x \rightarrow u^{-1}(x)$ является дифференцируемым отображением Ω в $\mathcal{L}(\vec{G}; \vec{F})$ и его производная для $\vec{X} \in \vec{E}$ задается формулой

$$(u^{-1})'(x) \cdot \vec{X} = -u^{-1}(x) \circ (u'(x) \cdot \vec{X}) \circ u^{-1}(x) \in \mathcal{L}(\vec{G}; \vec{F}). \quad (\text{III}, 8; 33)$$

В частности, если \mathfrak{M} является квадратной матрицей из n строк и n столбцов, зависящей от скалярного параметра t и дифференцируемой по t , и если она обратима для каждого t , то производная обратной матрицы \mathfrak{M}^{-1} может быть вычислена по формуле

$$\frac{d\mathfrak{M}^{-1}}{dt} = -\mathfrak{M}^{-1} \frac{d\mathfrak{M}}{dt} \mathfrak{M}^{-1}. \quad (\text{III}, 8; 34)$$

Для доказательства достаточно применить следствие к $E = \mathbb{K}$, $\vec{F} = \vec{G} = \mathbb{K}^n$. При $n=1$ мы возвращаемся к производной обратной функции для заданной скалярной функции f , нигде не обращающейся в нуль:

$$\left(\frac{1}{f}\right)' = -\frac{f'}{f^2}. \quad (\text{III}, 8; 35)$$

Теперь можно применить полученные результаты к дифференцированию неявных функций.

Теорема 28 (теорема о неявной функции). *Пусть E, F, G — три аффинных нормированных пространства, f — непрерывно дифференцируемое отображение открытого множества $\Omega \subset E \times F$ в G .*

1°) *Пусть A и B — открытые множества в пространствах E и F соответственно, $A \times B \subset \Omega$, и пусть g — непрерывное отображение A в B , удовлетворяющее условию (III, 8; 2). Если для каждого x из A частная производная $\frac{\partial f}{\partial y}(x, g(x))$ является обратимым элементом пространства $\mathcal{L}(\vec{F}; \vec{G})$, то g является непрерывно дифференцируемым отображением A в B .*

2°) *Пусть (a, b) — некоторая точка Ω , $f(a, b) = c$. Если $\frac{\partial f}{\partial y}(a, b)$ является обратимым элементом пространства $\mathcal{L}(\vec{F}; \vec{G})$ и пространство F полно, то существуют открытые множества A и B в пространствах E и F , содержащие соответственно a и b , $A \times B \subset \Omega$, такие, что для каждого $x \in A$ уравнение (III, 8; 1) относительно y имеет, и при том единственное, решение в B и определяемая им функция $y = g(x)$ является непрерывно дифференцируемым отображением A в B , а ее производная задается формулой (III, 8; 16).*

Доказательство. 1°) Из теоремы 26 вытекает, что g дифференцируема в каждой точке A и что ее производная в точке $x \in A$ является элементом $\mathcal{L}(\vec{E}; \vec{F})$, задаваемым формулой

$$g'(x) = - \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \circ \frac{\partial f}{\partial x}(x, g(x)). \quad (\text{III, 8; 37})$$

Отображение $x \rightarrow g(x)$ предполагается непрерывным в A , а функции $\frac{\partial f}{\partial x}$ и $\frac{\partial f}{\partial y}$ — непрерывными в Ω . Следовательно, $x \rightarrow \frac{\partial f}{\partial x}(x, g(x))$ и $x \rightarrow \frac{\partial f}{\partial y}(x, g(x))$ являются непрерывными отображениями A в пространства $\mathcal{L}(\vec{E}; \vec{G})$ и $\mathcal{L}(\vec{F}; \vec{G})$ соответственно. Так как $u \rightarrow u^{-1}$ является непрерывным отображением \mathcal{U} в \mathcal{U}^{-1} (теорема 27), то функция $x \rightarrow \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1}$ является непрерывным отображением A в $\mathcal{L}(\vec{G}; \vec{F})$.

Поскольку сложная функция $(u, v) \rightarrow v \circ u$ является непрерывным отображением $\mathcal{L}(\vec{E}; \vec{G}) \times \mathcal{L}(\vec{G}; \vec{F})$ в $\mathcal{L}(\vec{E}; \vec{F})$ (теорема 54 гл. II), g' является непрерывным отображением множества A в пространство $\mathcal{L}(\vec{E}; \vec{F})$, а, значит, функция g непрерывно дифференцируема на множестве A .

2°) Поскольку на этот раз пространства F и G полны (см. примечание на стр. 301: G полно, так как известен обратимый элемент $\frac{\partial f}{\partial y}(a, b)$ из $\mathcal{L}(\vec{F}; \vec{G})$), то множество \mathcal{U} обратимых элементов из $\mathcal{L}(\vec{F}; \vec{G})$ открыто. Так как $(x, y) \rightarrow \frac{\partial f}{\partial y}(x, y)$ предполагается непрерывным отображением Ω в $\mathcal{L}(\vec{F}; \vec{G})$, то прообраз \mathcal{U} при этом отображении является открытым подмножеством Ω_1 множества Ω , содержащим (a, b) . Теорема 25 применима к сужению f на Ω_1 (поскольку F полно). Поэтому можно определить такие открытые множества A и B в E и F , содержащие соответственно a и b , $A \times B \subset \Omega_1$, что для каждого $x \in A$ найдется, и притом единственный, элемент $y \in B$, являющийся решением уравнения (III, 8; 1) и такой, что определяемая при этом функция $y = g(x)$ из A в B непрерывна. Поскольку тогда $(x, g(x)) \in A \times B \subset \Omega_1$ для каждого x из A , то мы находимся в условиях пункта 1°), а, следовательно, функция g непрерывно дифференцируема, и по теореме 26 получаем (III, 8; 16).

Следствие. Если \vec{F} и \vec{G} — пространства Банаха, \mathcal{U} — открытое множество обратимых элементов пространства $\mathcal{L}(\vec{F}; \vec{G})$, то отображение $u \rightarrow u^{-1}$ множества \mathcal{U} на \mathcal{U}^{-1} непрерывно дифференцируемо вместе со своим обратным отображением.

Для доказательства достаточно применить теорему 28, п. 1°) к уравнению $uv = I$ (см. теорему 27, второе доказательство). Можно также непосредственно доказать непрерывность производной в теореме 27.

Частный случай, когда $E = F = G = \mathbb{K}$ — скалярное поле

Напомним, что если через $q = \frac{\partial f}{\partial y}(a, b)$ обозначить обычную частную производную скалярной функции f двух скалярных переменных, то ее частным производным отображением Q будет умножение на q . Следовательно, оно обратимо тогда и только тогда, когда частная производная q отлична от нуля. Отсюда вытекает, что если функция f имеет непрерывные обычные частные производные $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ (тогда, согласно теореме 15, функция f дифференцируема) и $q = \frac{\partial f}{\partial y}(a, b) \neq 0$, то мы находимся в условиях применимости теоремы 28 существования, единственности и дифференцируемости неявной функции. Именно при этих условиях дается обычно теорема о неявной функции

в математическом анализе. Формула (III, 8; 16) принимает вид

$$g'(a) = - \frac{\frac{\partial f}{\partial x}(a, b)}{\frac{\partial f}{\partial y}(a, b)}. \quad (\text{III}, 8; 38)$$

Предположим теперь, что $\frac{\partial f}{\partial y}(a, b) = 0$, но $\frac{\partial f}{\partial x}(a, b) \neq 0$.

Тогда в окрестности точки (a, b) можно выразить x как функцию y , $x=h(y)$. Поскольку $h'(b)=0$, кривая, определяемая уравнением $f(x, y)=c$, имеет в точке (a, b) вертикальную касательную, и невозможность выразить y как функцию x становится очевидной.

Предположим, наконец, что $\frac{\partial f}{\partial x}(a, b) = \frac{\partial f}{\partial y}(a, b) = 0$. Кривая имеет в (a, b) особую точку. Эта точка может быть изолированной особой точкой (например, если $K=\mathbb{R}$, $(a, b)=(0, 0)$, $f(x, y)=x^2+y^2$). Тогда кривой в окрестности (a, b) не существует. Однако, если кривая имеет, например, «двойную точку с различными касательными», то можно в окрестности (a, b) попытаться разделить кривую на две различные ветви и для каждой из них вычислить y как функцию от x или наоборот. Для простоты будем считать, что f принадлежит классу C^3 и $a=b=c=0$. Вместо переменных x и y введем переменные x и $m=m=y/x$. Данное уравнение примет вид $F(x, m)=f(x, mx)=0$. Посмотрим теперь, позволяет ли это уравнение найти m как функцию x . При $x=0$ функция F равна нулю при любом m и даже в том случае, когда $\partial f/\partial x$ и $\partial f/\partial y$ не обращаются одновременно в нуль в начале координат. Здесь, кроме того, производная $\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + m \frac{\partial f}{\partial y}$ равна нулю при $x=0$ и любом m . Позже (теорема 119 гл. IV) мы убедимся, что отсюда вытекает принадлежность к классу C^1 функции $G(x, m) = \frac{F(x, m)}{x^2}$ (значение этой функции при $x=0$ определяется переходом к пределу).

Предположим, что разложение Тейлора функции f в начале координат имеет вид

$$f(x, y) = \frac{1}{2}(rx^2 + 2sxy + ty^2) + \dots . \quad (\text{III}, 8; 38_2)$$

Тогда в точке $x=0, m=m_0$ имеем

$$G(0, m_0) = \lim_{x \rightarrow 0} \frac{F(x, m_0 x)}{x^2} = \frac{1}{2}(r + 2sm_0 + tm_0^2). \quad (\text{III}, 8; 38_3)$$

Если трехчлен $r + 2sm + tm^2$ имеет два различных корня $m=m_1$ и $m=m_2$ в K , то $G(0, m_1)=0$, $G(0, m_2)=0$. Для

уравнения $G(x, m) = 0$ мы имеем два частных решения $(0, m_1)$ и $(0, m_2)$. Если $\frac{\partial G}{\partial m}(0, m_1)$ и $\frac{\partial G}{\partial m}(0, m_2)$ отличны от нуля, то теорема о неявной функции позволяет нам разрешить уравнение $G(x, m) = 0$ в окрестности этих точек и получить две функции $m = \mu_1(x)$ и $m = \mu_2(x)$ класса C^1 . Две ветви искомой кривой будут определены уравнениями класса C^1 : $y = x\mu_1(x)$, $y = x\mu_2(x)$. Касательные в начале координат к этим ветвям имеют угловые коэффициенты m_1 и m_2 .

Производные $\frac{\partial G}{\partial m}(0, m)$ в корнях m_1 и m_2 трехчлена $G(0, m)$ не могут быть равными нулю, так как по условию эти корни различны, а

$$\frac{\partial G}{\partial m}(0, m_0) = s + m_0 t. \quad (\text{III}, 8; 38_4)$$

Итак, теорема о неявной функции может применяться даже в тех случаях, когда ее условия непосредственно не удовлетворяются.

Случай, когда E, F, G конечномерны

Пусть E, F, G — аффинные конечномерные пространства. Пусть в присоединенных векторных пространствах $\vec{E}, \vec{F}, \vec{G}$ выбраны базисы $(\vec{e}_k)_{k \in K}, (\vec{f}_j)_{j \in J}, (g_i)_{i \in I}$, а в качестве начал координат взяты начала пространств E, F, G . Тогда отображение f из $\Omega \subset E \times F$ в G определяется формулами

$$z_i = F_i((x_k)_{k \in K}, (y_j)_{j \in J}), \quad i \in I. \quad (\text{III}, 8; 39)$$

Непрерывность f выражается в том, что скалярные функции F_i являются непрерывными функциями своих аргументов x_k и y_j . Сказать, что f имеет непрерывную частную производную по второй переменной, означает сказать, что F_i имеют непрерывные частные производные первого порядка по переменным y_j (теорема 15). Обратимость частного производного отображения Q в точке (a, b, c) , где $c = f(a, b)$, означает следующее. Прежде всего, G и F должны иметь одинаковую размерность, т. е. должно быть $I = J$. Тогда рассматриваемая обратимость будет иметь место тогда и только тогда, когда якобиан, состоящий из элементов $\frac{\partial F_i}{\partial y_j}(a, b)$, являющийся определителем матрицы линейного отображения Q , отличен от нуля. Если эти условия выполнены, то в окрестности точки (a, b) можно (теорема 25) определить y как функцию $g(x)$, т. е. выразить y_j в виде

$$y_j = G_j((x_k)_{k \in K}), \quad j \in J = I. \quad (\text{III}, 8; 40)$$

Если теперь функции F_i имеют частные производные не только по переменным y_j , но также и по переменным x_k , непрерыв-

ные в Ω , то можно утверждать, что отображение f принадлежит классу C^1 . Тогда можно применить теорему 28 и формулу (III, 8; 16) и найти производную матрицу отображения g , т. е. матрицу, составленную из частных производных функций y_j по переменным x_k , в виде отношения двух матриц. Если $I=J=\{1, 2, \dots, m\}$, $K=\{1, 2, \dots, n\}$, то

$$\begin{aligned} & \left(\begin{array}{cccc} \frac{\partial G_1}{\partial x_1} & \frac{\partial G_1}{\partial x_2} & \cdots & \frac{\partial G_1}{\partial x_n} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial G_m}{\partial x_1} & \frac{\partial G_m}{\partial x_2} & \cdots & \frac{\partial G_m}{\partial x_n} \end{array} \right) = \\ & = - \left(\begin{array}{cccc} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \frac{\partial F_m}{\partial y_2} & \cdots & \frac{\partial F_m}{\partial y_m} \end{array} \right)^{-1} \left(\begin{array}{cccc} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \frac{\partial F_m}{\partial x_2} & \cdots & \frac{\partial F_m}{\partial x_n} \end{array} \right). \quad (\text{III}, 8; 41) \end{aligned}$$

Вместо того чтобы производить вычисления с матрицами, можно также найти частные производные $\frac{\partial G_j}{\partial x_k}$ для фиксированного k и $j \in J$, придавая x приращение \vec{dx} , определенное условиями $dx_{k'} = dx_k$, если $k' = k$, и 0, если $k' \neq k$, и вычисляя дифференциалы dy_j , $j \in J$, путем решения системы m линейных уравнений с m неизвестными и отличным от нуля определителем (ибо он является якобианом):

$$\frac{\partial F_i}{\partial x_k} dx_k + \sum_{j \in J} \frac{\partial F_i}{\partial y_j} dy_j = 0, \quad i \in I. \quad (\text{III}, 8; 42)$$

Обратная функция как неявная функция

Найти обратную функцию $x=g(y)$ для некоторого гомеоморфизма $y=f(x)$ означает решить относительно x уравнение $f(x)=y$ для каждого заданного y . Эта задача является задачей отыскания неявной функции.

Теорема 29. Пусть E и F — аффинные полные нормированные пространства и f — непрерывно дифференцируемое отображение некоторого открытого множества $\Omega \subset E$ в F .

1°) Если a — точка множества Ω и $f'(a)$ — обратимый элемент пространства $\mathcal{L}(E; F)$, то образ каждой окрестности точки a при отображении f является окрестностью точки $b=f(a)$ и образ каждого открытого шара с центром в точке a при этом отображении содержит некоторый открытый шар с центром в точке b . Существуют такие открытые множества A и B в

пространствах E и F , содержащие соответственно a и b , что f является гомеоморфизмом A на B , непрерывно дифференцируемым вместе со своим обратным гомеоморфизмом.

Кроме того, имеем:

$$(f^{-1})'(b) = (f'(a))^{-1}. \quad (\text{III}, 8; 43)$$

2°) Если для каждого $x \in \Omega$ производная $f'(x)$ является обратимым элементом из $\mathcal{L}(\vec{E}; \vec{F})$, то образ каждого открытого множества из Ω при отображении f является открытым множеством и, в частности, $f(\Omega)$ является открытым множеством в F . Если, кроме того, отображение f инъективно, то оно является гомеоморфизмом множества Ω на $f(\Omega)$, непрерывно дифференцируемым вместе со своим обратным гомеоморфизмом.

Доказательство. Мы видели, что обратное отображение $g = f^{-1}$ является неявной функцией $x = g(y)$, определяемой уравнением $\vec{f}_1(x, y) = \vec{f}(x) - y = \vec{0}$, где \vec{f}_1 — отображение $\Omega \times F$ в \vec{F} . Применим предыдущие теоремы с неизбежным изменением обозначений: мы будем искать x как функцию y , а не наоборот. Поэтому необходимые предположения будут относиться к частной производной относительно x : $\frac{\partial \vec{f}_1}{\partial x} = \frac{\partial \vec{f}}{\partial x}$. Здесь E полно, а \vec{f}_1 непрерывно дифференцируемо на $\Omega \times F$, поскольку его частные производные $\frac{\partial \vec{f}_1}{\partial x}(x, y) = f'(x)$ и $\frac{\partial \vec{f}_1}{\partial y} = -I$ непрерывны (теорема 15).

Докажем утверждение 1°). Условия, необходимые для применения теоремы 28, 2°), выполнены в (a, b) . Следовательно, можно найти такие открытые множества A_0 и B , $a \in A_0 \subset E$, $b \in B \subset F$, что уравнение $f(x) = y$ для $y \in B$ имеет в A_0 единственное решение x и таким образом определенная функция $x = g(y)$ есть непрерывно дифференцируемая функция из B в A_0 . Однако эти открытые множества A_0 и B не отвечают непосредственно требуемым условиям. Образ $f(A_0)$ «покрывает» B один и только один раз, но он слишком велик: $f(A_0) \supset B$. Нам надо взять $A = A_0 \cap f^{-1}(B)$. Тогда f будет отображать A в B . Впрочем, для $y \in B$ точка $g(y)$ множества A_0 заведомо лежит в прообразе $f^{-1}(B)$ (поскольку $f(g(y)) = y \in B$), а, следовательно, в A . Таким образом, g отображает B в A . Итак, f и g — два взаимно обратных отображения A в B и B в A , где B открыто по построению и A открыто как пересечение двух открытых множеств. Отображение f непрерывно дифференцируемо по определению, а g по теореме о неявной функции. Следовательно, мы имеем два непрерывно дифференцируемых взаимно обратных гомеоморфизма. Пусть теперь \mathscr{U} — некоторая окрестность точки a в Ω . Тогда $\mathscr{U} \cap A$ является окрестностью точки a в A ,

и так как f — гомеоморфизм A на B , то множество $f(\mathcal{V} \cap A)$ является окрестностью точки b в B , а, значит, и в F , поскольку B — открытое множество в F . Множество $f(\mathcal{V})$ тем более является окрестностью точки b в пространстве F . Множество \mathcal{V} может быть, в частности, некоторым открытым шаром с центром в точке a . Множество $f(\mathcal{V})$, как окрестность точки b , содержит некоторый открытый шар с центром в точке b . Таким образом, утверждение 1°) доказано. Производная функции g в точке b вычисляется по правилу (III, 8; 23): надо решить относительно \vec{dx} уравнение

$$f'(a) \vec{dx} - \vec{dy} = \vec{0}, \quad (\text{III, 8; 44})$$

что дает

$$\vec{dx} = (f'(a))^{-1} \vec{dy}, \quad (\text{III, 8; 45})$$

откуда и вытекает (III, 8; 43).

Докажем теперь утверждение 2°). Пусть \mathcal{O} — открытое множество из Ω . Множество \mathcal{O} является окрестностью для каждой своей точки x . Так как $f'(x)$ обратима, то, согласно 1°), $f(\mathcal{O})$ является окрестностью точки $f(x)$, а, следовательно, множество $f(\mathcal{O})$ открыто. В частности, открыто множество $f(\Omega)$. Если, кроме того, отображение f инъективно, то оно является биективным и непрерывным отображением множества Ω на множество $f(\Omega)$, и образ каждого открытого множества открыт. Согласно теореме 11 гл. II, отображение f является гомеоморфизмом. Если a — произвольная точка Ω , то в открытом множестве B , определенном в 1°), функция f^{-1} совпадает с функцией g , и, следовательно, непрерывно дифференцируема в B , а, значит, во всем множестве $f(\Omega)$.

Замечание. Теорема 29, 1°) приводит к той же формуле, что и следствие 4 теоремы 11. Но две эти формулировки до некоторой степени обратны одна другой, поскольку, с одной стороны, для обратимости $f'(a)$ нужны два дифференцируемых взаимно обратных гомеоморфизма, а, с другой стороны, для нахождения двух взаимно обратных дифференцируемых гомеоморфизмов нужна обратимость $f'(a)$. Очевидно также, что условие 29, 2°) носит более тонкий характер.

Если E и F конечномерны и если в них выбрана система координат, то обратимость $f'(a)$ означает, что E и F имеют одну и ту же размерность и что якобиан функции f в точке a в выбранной системе координат отличен от нуля.

Теорема 29 приводит к следующему определению:

Определение. Пусть Ω и F — два топологических пространства и f — непрерывное отображение Ω в F . Говорят, что f является *открытым отображением*, если образ каждого

открытого множества из Ω при отображении f является открытым множеством в F . Говорят, что f является локальным гомеоморфизмом пространства Ω в пространство F , если для любой точки $a \in \Omega$ существует открытое множество A в Ω , содержащее a , и открытое множество B в F , содержащее b , такие, что f является гомеоморфизмом A на B .

Теперь легко проверяются следующие утверждения:

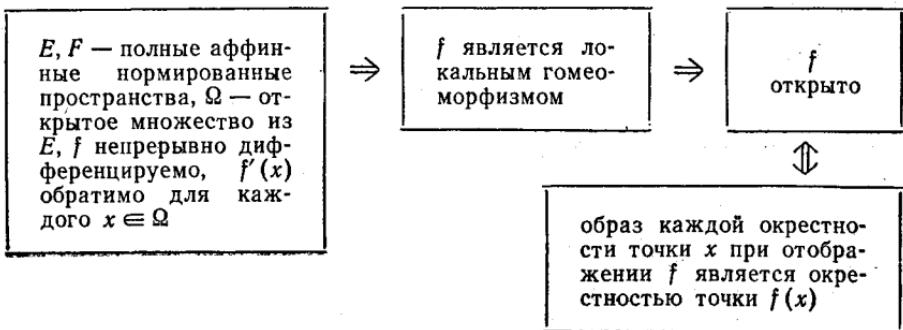
А) Для того чтобы отображение f было открытым, необходимо и достаточно, чтобы для каждой точки $a \in \Omega$ образ любой окрестности точки a при отображении f был окрестностью точки $f(a)$, или чтобы для каждой точки a образ любого открытого шара с центром в a при отображении f содержал некоторый открытый шар с центром в точке $f(a)$. (Достаточность была установлена при доказательстве теоремы 29, 2°), а необходимость очевидна.)

В) Всякий локальный гомеоморфизм является открытым отображением. Этот факт мы установили при доказательстве теоремы 29, 1°): когда f — локальный гомеоморфизм, для любой точки $x \in \Omega$ и любой окрестности \mathcal{U} точки x множество $f(\mathcal{U})$ является окрестностью точки $f(x)$, а, следовательно, согласно А), отображение f открыто.

С) Если f — непрерывно дифференцируемое отображение некоторого открытого множества Ω полного аффинного нормированного пространства E в аффинное нормированное пространство F и если для каждого $x \in \Omega$ производная $f'(x)$ обратима, то f является локальным гомеоморфизмом Ω в F . Это устанавливает теорема 29. В частности, f является открытым отображением, а $f(\Omega)$ является открытым множеством в F .

Д) Открытое инъективное отображение f является гомеоморфизмом Ω на образ $f(\Omega)$. Этот факт установлен нами при доказательстве теоремы 29, 2°).

Следующая схема подводит итог полученным результатам. Предполагается, что f является непрерывным отображением топологического пространства Ω в топологическое пространство F .



Необходимо иметь в виду, что здесь нельзя заходить слишком далеко и что, в частности, обратные утверждения к указанным выше следствиям не верны.

Z Рассмотрим С). Если f является непрерывно дифференцируемым локальным гомеоморфизмом Ω в F , то его производная $f'(x)$ не обязательно обратима для каждого $x \in \Omega$. Например, $x \rightarrow x^3$ является дифференцируемым гомеоморфизмом \mathbb{R} на \mathbb{R} , но его производная равна нулю в начале координат a , следовательно, его обратный гомеоморфизм $y \rightarrow \sqrt[3]{y}$ не дифференцируем в начале координат.

Z Рассмотрим В). Открытое отображение не обязательно является локальным гомеоморфизмом. Например, всякая проекция произведения топологических пространств на пространства-сомножители открыта, как это следует из определения окрестностей, но, вообще говоря, не является локальным гомеоморфизмом (проекция $(x, y) \rightarrow x$ пространства \mathbb{R}^2 на \mathbb{R} не является локальным гомеоморфизмом!).

Наконец, в D) условие «отображение f инъективно» необходимо точно так же, как в теореме 29, 2°). Например, отображение $z \rightarrow z^2$ пространства \mathbb{C} в \mathbb{C} , суженное на открытое множество Ω , дополнительное к началу координат, удовлетворяет условию « $f'(z) = 2z \neq 0$ обратимо для каждого $z \in \Omega$ ». Значит, оно является локальным гомеоморфизмом и открытым отображением, но не является гомеоморфизмом, поскольку оно не инъективно.

Вернемся к теореме 29. Предполагая $f'(a)$ обратимой, мы доказали свойство гомеоморфности. Отсюда как следствие вытекает менее сильное свойство открытости отображения. Это последнее свойство связано с менее сильным свойством, чем обратимость $f'(a)$, а именно с сюръективностью:

Теорема 30. Пусть f — непрерывно дифференцируемое отображение открытого множества Ω нормированного аффинного пространства E в конечномерное аффинное пространство F .

1°) Если a — точка Ω и если $f'(a)$ — сюръективное отображение \vec{E} на \vec{F} , то образ любой окрестности точки a при отображении f является окрестностью точки $f(a)$, а образ каждого открытого шара с центром в точке a содержит некоторый открытый шар с центром в $f(a)$.

2°) Если $f'(x)$ для любого $x \in \Omega$ является сюръективным отображением \vec{E} на \vec{F} , то образ каждого открытого множества из Ω при отображении f является открытым множеством в F . В частности, $f(\Omega)$ является открытым множеством. (Другими словами, f является открытым отображением.)

Предположение о том, что F конечномерно, очевидно, слишком ограничительно. От него можно избавиться, предполагая E и F полными.

Доказательство. Докажем утверждение 1°). Пусть (\vec{f}_i) , $i = 1, 2, \dots, m$, — некоторый базис в пространстве \vec{F} . Поскольку отображение $f'(a)$ сюръективно, существуют векторы $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m$, имеющие образами векторы \vec{f}_i . Векторы \vec{X}_i линейно независимы, поскольку линейно независимы их образы при отображении $f'(a)$; поэтому сужение $f'(a)$ на m -мерное векторное подпространство \vec{E}_0 , порожденное векторами \vec{X}_i , является сюръективным отображением \vec{E}_0 на \vec{F} . Но так как \vec{E}_0 и \vec{F} имеют одинаковую размерность, то $f'(a)$ является биективным отображением \vec{E}_0 на \vec{F} ; другими словами, рассматриваемое как линейное отображение \vec{E}_0 в \vec{F} , оно обратимо. Пространство \vec{E}_0 , в силу конечномерности, полно (теорема 41 гл. II). Сужение отображения f на аффинное подпространство E_0 , проходящее через a параллельно \vec{E}_0 , имеет в точке a обратимую производную, а само E_0 полно. Применяя к нему теорему 29, получаем утверждение 1°). Мы видим, что если \mathcal{U} является окрестностью точки a , то не только образ $f(\mathcal{U})$ является окрестностью $f(a)$, но даже $f(\mathcal{U} \cap E_0)$ есть окрестность точки $f(a)$.

Переход к 2°) производится точно так же, как при доказательстве теоремы 29, т. е. с использованием критерия А) для открытых отображений. Естественно, здесь не возникает вопроса о том, является ли f локальным гомеоморфизмом, так как $f'(x)$ не предполагалось инъективным.

Вычисление производных высшего порядка неявной функции

Теорема 31. Если при условиях теоремы 28 функция f непрерывно дифференцируема m раз в Ω , то неявная функция g непрерывно дифференцируема m раз в A .

Доказательство. Поскольку эта теорема уже доказана для $m=1$ (теорема 28), рассуждения будем проводить по индукции: предположим, что она доказана для производных порядка $m-1$, и докажем ее для производных порядка $m \geq 2$. Заметим сначала, что уравнение $uv=I$, рассматриваемое во втором доказательстве теоремы 27, входит сюда в виде уже изученного случая, ибо функция $(u, v) \rightarrow uv$ билинейна и непрерывна, а, следовательно, бесконечно дифференцируема. Отображение $u \rightarrow u^{-1}$, в силу индуктивного предположения, принадлежит классу C^{m-1} на \mathcal{U} . Рассмотрим теперь формулу

(III, 8; 37). Функция g непрерывна, поскольку, в силу индуктивного предположения, она принадлежит классу C^{m-1} . Функция $\partial f/\partial x$ принадлежит классу C^{m-1} , так как f принадлежит C^m . Поэтому, согласно теореме 19, отображение $x \rightarrow \frac{\partial f}{\partial x}(x, g(x))$ принадлежит классу C^{m-1} . То же самое имеет место для отображения $x \rightarrow \frac{\partial f}{\partial y}(x, g(x))$. Поскольку, как мы только что видели, функция $u \rightarrow u^{-1}$ также принадлежит классу C^{m-1} , из теоремы 19 следует, что функция $x \rightarrow \left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1}$ принадлежит классу C^{m-1} . Наконец, $(u, v) \rightarrow v \circ u$ является билинейным непрерывным отображением $\mathcal{L}(\vec{E}; \vec{G}) \times \mathcal{L}(\vec{G}; \vec{F})$ в $\mathcal{L}(\vec{E}; \vec{G})$ (теорема 54). Из теоремы 18 следует, что g' принадлежит C^{m-1} , а, значит, g принадлежит классу C^m .

Следствие 1. Если \vec{F} и \vec{G} — пространства Банаха, а \mathcal{U} (соответственно \mathcal{U}^{-1}) — открытое множество обратимых элементов из $\mathcal{L}(\vec{F}; \vec{G})$ (соответственно из $\mathcal{L}(\vec{G}; \vec{F})$), то $u \rightarrow u^{-1}$ является гомеоморфизмом \mathcal{U} на \mathcal{U}^{-1} , бесконечно дифференцируемым вместе со своим обратным гомеоморфизмом.

Это дополняет теорему 27. По существу мы уже доказали это следствие по индукции во время доказательства теоремы 30.

Определение. Если биекция вместе со своей обратной биекцией принадлежит классу C^m , но она называется C^m -диффеоморфизмом. Так, например, отображение $u \rightarrow u^{-1}$ является C^∞ -дiffeоморфизмом \mathcal{U} на \mathcal{U}^{-1} .

Следствие 1₂. Если u является t раз дифференцируемым отображением (соответственно принадлежит классу C^m) открытого множества Ω аффинного нормированного пространства E в пространство $\mathcal{L}(\vec{F}; \vec{G})$ и если $\vec{u}(x)$ для каждого $x \in \Omega$ является обратимым элементом пространства $\mathcal{L}(\vec{F}; \vec{G})$, то $x \rightarrow u^{-1}(x)$ является t раз дифференцируемым (соответственно принадлежащим классу C^m) отображением Ω в $\mathcal{L}(\vec{G}; \vec{F})$.

Это вытекает из следствия 1 и теоремы 19, примененной к композиции отображений $x \rightarrow u(x)$ и $u \rightarrow u^{-1}$.

Отметим важный частный случай: функция $1/f$, обратная к скалярной функции f ($\vec{F} = \vec{G} = \mathbb{K} = \mathcal{L}(\vec{F}; \vec{G})$), t раз дифференцируемой или принадлежащей классу C^m на $\Omega \subset E$ и всюду $\neq 0$, также t раз дифференцируема или принадлежит классу C^m .

Если $t \rightarrow \mathfrak{M}(t)$ является дважды дифференцируемой функцией, определенной на \mathbb{K} , со значениями в пространстве квадратных матриц и если $\mathfrak{M}(t)$ обратима для каждого t , то вторая производная $\frac{d^2}{dt^2}(\mathfrak{M}^{-1})$ получается непосредственным дифференцированием (III, 8; 34):

$$\frac{d^2}{dt^2}(\mathfrak{M}^{-1}) = 2\mathfrak{M}^{-1} \frac{d\mathfrak{M}}{dt} \mathfrak{M}^{-1} \frac{d\mathfrak{M}}{dt} \mathfrak{M}^{-1} - \mathfrak{M}^{-1} \frac{d^2\mathfrak{M}}{dt^2} \mathfrak{M}^{-1}.$$

Следствие 2. Если в условиях теоремы 29, 1°) отображение f принадлежит классу C^m , то обратная биекция f^{-1} отображения f является отображением B в A класса C^m : f является C^m -диффеоморфизмом A на B .

Вычисление в самом общем случае производной функции g порядка m , исходя из частных производных функции f порядков $\leq m$, приводит к чрезвычайно сложным алгебраическим выражениям.

Мы ограничимся рассмотрением частного примера. Пусть f — функция трех вещественных переменных. Будем считать, что уравнение

$$f(x, y, z) = 0 \quad (\text{III, 8; 46})$$

определяет z как функцию переменных x и y в окрестности точки, в которой частная производная $\partial f / \partial z$ отлична от нуля. Вычислим в этой точке частные производные p, q, r, s, t функции z , исходя из частных производных функции f . В (III, 8; 16) мы уже получили формулы

$$p = \frac{\partial z}{\partial x} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}}, \quad q = \frac{\partial z}{\partial y} = -\frac{\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial z}}. \quad (\text{III, 8; 47})$$

Остается теперь с помощью теоремы о сложной функции продифференцировать предыдущие выражения по x и y и получить следующие выражения:

$$r = \frac{\partial^2 z}{\partial x^2} = \frac{-\frac{\partial^2 f}{\partial x^2} \left(\frac{\partial f}{\partial z} \right)^2 + 2 \frac{\partial^2 f}{\partial x \partial z} \frac{\partial f}{\partial x} \frac{\partial f}{\partial z} - \frac{\partial^2 f}{\partial z^2} \left(\frac{\partial f}{\partial x} \right)^2}{\left(\frac{\partial f}{\partial z} \right)^3},$$

$$s = \frac{\partial^2 z}{\partial x \partial y} = \frac{-\frac{\partial^2 f}{\partial x \partial y} \left(\frac{\partial f}{\partial z} \right)^2 + \frac{\partial^2 f}{\partial x \partial z} \frac{\partial f}{\partial y} \frac{\partial f}{\partial z} + \frac{\partial^2 f}{\partial y \partial z} \frac{\partial f}{\partial x} \frac{\partial f}{\partial z} - \frac{\partial^2 f}{\partial z^2} \frac{\partial f}{\partial x} \frac{\partial f}{\partial y}}{\left(\frac{\partial f}{\partial z} \right)^3},$$

$$t = \frac{\partial^2 z}{\partial y^2} = \frac{-\frac{\partial^2 f}{\partial y^2} \left(\frac{\partial f}{\partial z} \right)^2 + 2 \frac{\partial^2 f}{\partial y \partial z} \frac{\partial f}{\partial y} \frac{\partial f}{\partial z} - \frac{\partial^2 f}{\partial z^2} \left(\frac{\partial f}{\partial y} \right)^2}{\left(\frac{\partial f}{\partial z} \right)^3}. \quad (\text{III, 8; 48})$$

Можно поступить иначе. Соотношение, дающее p , записывается в исходном виде

$$\frac{\partial f}{\partial x} + p \frac{\partial f}{\partial z} = 0. \quad (\text{III}, 8; 49)$$

Затем полученное равенство дифференцируется по x и y с одновременной заменой dz , dp , dq на $p dx + q dy$, $r dx + s dy$, $s dx + t dy$. Таким путем получается соотношение

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} dx + \frac{\partial^2 f}{\partial x \partial y} dy + \frac{\partial^2 f}{\partial x \partial z} (p dx + q dy) + \frac{\partial f}{\partial z} (r dx + s dy) + \\ + p \frac{\partial^2 f}{\partial x \partial z} dx + p \frac{\partial^2 f}{\partial y \partial z} dy + p \frac{\partial^2 f}{\partial z^2} (p dx + q dy) = 0. \quad (\text{III}, 8; 50) \end{aligned}$$

Приравнивая затем нулю коэффициенты при dx и dy , получаем некоторые равенства, из которых находятся частные производные r и s . Аналогичные вычисления проводятся исходя из соотношения $\frac{\partial f}{\partial y} + q \frac{\partial f}{\partial z} = 0$ для нахождения s и t .

Существует другой метод, использующий разложение Тейлора (см. теоремы 21₂ и 21₃).

Запишем разложение Тейлора функции f в окрестности точки x, y, z до 2-го порядка:

$$\begin{aligned} 0 = f(x + X, y + Y, z + Z) - f(x, y, z) = \frac{\partial f}{\partial x} X + \frac{\partial f}{\partial y} Y + \frac{\partial f}{\partial z} Z + \\ + \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2} X^2 + \frac{\partial^2 f}{\partial y^2} Y^2 + \frac{\partial^2 f}{\partial z^2} Z^2 + 2 \frac{\partial^2 f}{\partial x \partial z} YZ + \right. \\ \left. + 2 \frac{\partial^2 f}{\partial z \partial x} ZX + 2 \frac{\partial^2 f}{\partial z \partial y} XY \right] + \dots. \quad (\text{III}, 8; 51) \end{aligned}$$

Если в этой формуле Z заменить его разложением Тейлора в точке X, Y , т. е. положить

$$Z = pX + qY + \frac{1}{2} (rX^2 + 2sXY + tY^2) + \dots, \quad (\text{III}, 8; 52)$$

то получится тождественный нуль. Приравнивая нулю коэффициенты при X, Y, X^2, XY, Y^2 , получим 5 уравнений, дающих p , q , r , s , t . Прежде всего, приравнивая нулю коэффициенты при X и Y , получаем соотношения для определения p и q . Приравнивая затем нулю коэффициенты при X^2 , XY , Y^2 , находим

$$\begin{aligned} \frac{\partial f}{\partial z} r + \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial z^2} p^2 + 2 \frac{\partial^2 f}{\partial x \partial z} p = 0, \\ \frac{\partial f}{\partial z} s + \frac{\partial^2 f}{\partial z^2} pq + \frac{\partial^2 f}{\partial y \partial z} p + \frac{\partial^2 f}{\partial x \partial z} q + \frac{\partial^2 f}{\partial x \partial y} q = 0, \quad (\text{III}, 8; 53) \\ \frac{\partial f}{\partial z} t + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} q^2 + 2 \frac{\partial^2 f}{\partial y \partial z} q = 0, \end{aligned}$$

откуда и получаются выражения (III, 8; 48).

В этом методе, использующем разложение Тейлора, можно провести другой вариант преобразований: вычислить первые производные p и q , перейти к вычислению вторых, приравнивая нулю соотношение (III, 8; 51) и выражая в нем Z как функцию других величин, т. е. $X, Y, X^2, XY, Y^2, XZ, YZ, Z^2$. При этом получится формула

$$\frac{\partial f}{\partial z} Z = - \frac{\partial f}{\partial x} X - \frac{\partial f}{\partial y} Y - \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2} X^2 + \frac{\partial^2 f}{\partial y^2} Y^2 + \frac{\partial^2 f}{\partial z^2} Z^2 + \right. \\ \left. + 2 \frac{\partial^2 f}{\partial y \partial z} YZ + 2 \frac{\partial^2 f}{\partial z \partial x} ZX + 2 \frac{\partial^2 f}{\partial x \partial y} XY \right] + \dots, \quad (\text{III, 8; 54})$$

в которой переменная Z выражена через XZ, YZ и Z^2 . Однако для того чтобы иметь приближение второго порядка, можно удовлетвориться заменой в этом выражении Z на $pX + qY$, где p и q были вычислены ранее. Таким путем получим выражение

$$\frac{\partial f}{\partial z} Z = - \frac{\partial f}{\partial x} X - \frac{\partial f}{\partial y} Y - \\ - \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2} X^2 + 2 \frac{\partial^2 f}{\partial x \partial y} XY + \frac{\partial^2 f}{\partial y^2} Y^2 + 2 \frac{\partial^2 f}{\partial x \partial z} X(pX + qY) + \right. \\ \left. + 2 \frac{\partial^2 f}{\partial y \partial z} Y(pX + qY) + \frac{\partial^2 f}{\partial z^2} (pX + qY)^2 \right] + \dots \quad (\text{III, 8; 55})$$

Вычисляя коэффициенты при $X^2, 2XY, Y^2$ в разложении Z , снова получим $\frac{1}{2}r, \frac{1}{2}s, \frac{1}{2}t$.

Техника замены переменных и замены функций

С заменой переменных мы встречались ранее до теоремы о неявной функции, но теперь эта теорема позволяет производить значительно более важные замены.

Предположим, что задано уравнение в частных производных второго порядка относительно скалярной функции z двух скалярных переменных x и y :

$$F(x, y, z, p, q, r, s, t) = 0. \quad (\text{III, 8; 56})$$

Необходимо выполнить в нем одновременно замену переменных и замену функции по формулам

$$X = X(x, y, z), \quad Y = Y(x, y, z), \quad Z = Z(x, y, z)^1). \quad (\text{III, 8; 57})$$

Мы хотим узнать, как преобразуется рассматриваемое уравнение в частных производных при такой замене переменных. Она законна лишь в том случае, когда (III, 8; 57) определяет C^2 -диффеоморфизм некоторого открытого множества из \mathbb{K}^3 на открытое множество в \mathbb{K}^3 и поверхность $z = z(x, y)$, определяе-

¹⁾ X, Y, Z имеют, очевидно, не тот смысл, какой им придавался ранее.

мая уравнением (III, 8; 56), преобразуется в поверхность $Z = Z(X, Y)$ так, что обе написанные функции принадлежат классу C^2 . Следуя обычному методу, надо выразить p, q, r, s, t как функции новых частных производных P, Q, R, S, T . Можно начать с производных первого порядка и затем перейти к вычислению производных второго порядка. Для производных первого порядка можно начать с соотношения $dZ = PdX + QdY$, заменяя dX, dY, dZ их дифференциальными выражениями, найденными из (III, 8; 57), и получая таким путем соотношение между dx, dy, dz . Это соотношение, представленное в виде $dz = pdx + qdy$, позволит выразить p и q через x, y, z, P, Q . Дифференцируя полученное соотношение, затем находим dp как функцию от dx, dy, dz, dP, dQ . Заменяя dz, dP, dQ на $pdx + qdy, Rdx + Sdy, SdX + TdY$, затем p и q их значениями, найденными прежде, dX и dY — их значениями, полученными из (III, 8; 57) (где dz должно быть заменено на $pdx + qdy$ с уже найденными p и q), мы выражаем dp как комбинацию dx и dy , которая, будучи представленной в виде $dp = rdx + sdy$, дает r и s как функции от x, y, z, P, Q, R, S, T .

Точно так же вычисляется t . Полученные результаты подставляются в (III, 8; 56).

Полученное уравнение относительно P, Q, R, S, T содержит в действительности величины x, y, z , а не X, Y, Z . Поэтому, если эти переменные входят в уравнение, следует еще выразить (x, y, z) через (X, Y, Z) , решая систему уравнений (III, 8; 57)¹⁾.

§ 9. ДИФФЕРЕНЦИРУЕМЫЕ МНОГООБРАЗИЯ

В теореме 8А мы рассматривали множество \mathcal{E} в произведении аффинных пространств $E \times F$, определенное уравнением $y = f(x)$. Оно называется *дифференцируемым многообразием*, если f является дифференцируемой функцией. Это множество называется *t раз дифференцируемым многообразием*, если функция f дифференцируема t раз; *непрерывно дифференцируемым многообразием* или *t раз непрерывно дифференцируемым многообразием*, если f непрерывно дифференцируема или t раз непрерывно дифференцируема; *бесконечно дифференцируемым многообразием*, если f бесконечно дифференцируема. Вместо того чтобы говорить, что многообразие t раз непрерывно

¹⁾ Вместо (III, 8; 57) могут быть заданы старые координаты как функции новых или заданы три неразрешенных соотношения между старыми и новыми переменными. Имеется много возможных вариантов. Вместо того чтобы начать сразу применять стандартную замену, надо сначала подумать! В математическом анализе можно встретить, например, такую замену переменных, когда кривизна кривой $y = y(x)$, выраженная через y' и y'' в декартовой системе координат, вычисляется как функция от r, r', r'' для кривой $r = r(\phi)$, заданной в полярных координатах.

дифференцируемо, говорят, что оно принадлежит классу C^m . Вместо того чтобы говорить, что многообразие бесконечно дифференцируемо, говорят, что оно принадлежит классу C^∞ . Чаще всего встречаются многообразия, принадлежащие классам C^m и C^∞ . Если $m = 0$, то многообразие называется *топологическим*. Оно определяется уравнением $y = f(x)$, где функция f непрерывна. Топологические многообразия играют очень важную роль, но мы ограничимся изучением многообразий класса C^m при $m \geq 1$.

Очевидно, что такое определение многообразия не является самым общим. Так, например, в евклидовом аффинном пространстве размерности N естественно рассматривать сферу как бесконечно дифференцируемое многообразие размерности $N-1$. Однако если выбрать в нем систему координат, то невозможно описать полностью *всю сферу*, выражая одну из координат как бесконечно дифференцируемую функцию остальных¹⁾. Придется разделить сферу на достаточно малые области, так чтобы в каждой из этих областей можно было одну из координат выразить как бесконечно дифференцируемую функцию остальных. Это обстоятельство приводит к общему понятию многообразия класса C^m .

Определение многообразия путем выражения некоторых координат как функции остальных. Пусть E — аффинное пространство размерности N над полем \mathbb{K} вещественных или комплексных чисел и V — некоторое подмножество E . Множество V называется *многообразием размерности n , принадлежащим классу C^m* , если для каждой точки a множества V существуют: система координат $0, e_1, e_2, \dots, e_n$ пространства E , открытое множество \mathcal{B} в подпространстве, натянутом на n первых координатных осей, система $N-n$ скалярных функций C_k , $k = 1, 2, \dots, N-n$, класса C^m , определенных на \mathcal{B} , открытое множество \mathcal{V} пространства E , содержащее a , проекция которого на подпространство, натянутое на n первых осей, совпадает с \mathcal{B} , такие, что пересечение $V \cap \mathcal{V}$ является в точности множеством точек $x = (x_1, \dots, x_N)$ из E , удовлетворяющих уравнениям

$$x_{n+k} = G_k(x_1, \dots, x_N)^2, \quad k = 1, 2, \dots, N-n. \quad (\text{III}, 9; 1)$$

¹⁾ Уравнение сферы с центром в начале координат радиуса R в \mathbb{R}^n имеет вид $x_n = \pm \sqrt{R^2 - x_1^2 - \dots - x_{n-1}^2}$. Из-за наличия знаков \pm это не одна функция (и даже если выбрать знак $+$, то не получится функция, дифференцируемая при $x_n = 0$).

²⁾ Согласно указанному свойству проекции множества \mathcal{V} , точка (x_1, x_2, \dots, x_n) лежит в \mathcal{B} , где определены функции G_k .

Итак, в \mathcal{U} последние $N - n$ координат x_{n+1}, \dots, x_N на многообразии V выражаются как функции класса C^m первых координат x_1, x_2, \dots, x_n , которые можно выбирать в \mathcal{V} произвольно. (В том, что первые координаты играют особую роль, нет ничего удивительного, поскольку для рассматриваемой точки a из V мы выбираем систему координат в E . Если эту особую роль играют другие координаты, то изменением порядка векторов в системе координат можно добиться того, чтобы это были первые n координат.)

Позже мы дадим (следствие теоремы 32) эквивалентное определение, в котором первые n координат не играют никакой особой роли, поскольку раз и навсегда выбирается определенная система координат в E .

Можно еще сказать, что *разделение системы из N координат на n первых и $N - n$ остальных* равносильно представлению E в виде произведения $\mathbb{K}^n \times \mathbb{K}^{N-n}$ и что, если обозначить через g отображение открытого множества \mathcal{V} из \mathbb{K}^n в \mathbb{K}^{N-n} , определенное формулами (III, 9; 1), то g принадлежит классу C^m , а множество $V \cap \mathcal{U}$ является в точности множеством точек $x = (y, z)$ из $E = \mathbb{K}^n \times \mathbb{K}^{N-n}$, удовлетворяющих уравнению $z = g(y)$.

Таким образом, мы вернулись к частному случаю, рассмотренному в теореме 8А, но для того, чтобы это стало возможным, мы в окрестности точки a специальным образом выбрали систему координат.

Определение многообразия при помощи его параметрического представления

Пусть V — множество N -мерного аффинного пространства E . Истинным¹⁾ параметрическим представлением множества V размерности n , принадлежащим классу C^m , называется отображение $\Phi: u \rightarrow \Phi(u)$ открытого множества \mathcal{O} из \mathbb{K}^n в E , обладающее следующими свойствами:

- 1°) Φ является гомеоморфизмом \mathcal{O} на V ;
- 2°) Φ является отображением множества \mathcal{O} в пространство E , принадлежащим классу C^m ;
- 3°) в каждой точке $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathcal{O}$ производное отображение $\Phi'(\alpha) \in \mathcal{L}(\mathbb{K}^n; \vec{E})$ имеет ранг n . Иначе говоря, образ векторного пространства \mathbb{K}^n при этом отображении является векторным подпространством в E размерности n , т. е. частные производные векторы $\frac{\partial \Phi}{\partial u_j}(a)$, $j = 1, 2, \dots, n$, являются линейно независимыми векторами в E , т. е. хотя бы один из

¹⁾ В дальнейшем слово «истинное» часто опускается. — Прим. ред.

определителей ранга n производной матрицы (матрицы, составленной из элементов $\frac{\partial \Phi_i}{\partial u_j}(a)$) отличен от нуля.

Можно подумать, что многообразия размерности n класса C^m из E будут единственными множествами, допускающими такое параметрическое представление. С другой стороны, совершенно очевидно, что многообразие, вообще говоря, может и не иметь такого параметрического представления. В самом деле, для параметрического представления необходимо, чтобы рассматриваемое многообразие было гомеоморфно некоторому открытому множеству из \mathbb{K}^n . Пример сферы, являющейся компактным множеством, показывает, что многообразие может не быть, вообще говоря, гомеоморфным открытому множеству из \mathbb{K}^n . Мы снова убеждаемся, что только достаточно малые области многообразия могут иметь такие параметрические представления.

Конечно, в задание Φ входит задание открытого множества \mathcal{O} из \mathbb{K}^n и его образа $\Phi(\mathcal{O})$ в V , но их обычно указывают, когда пишут $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$.

Теорема 32. Для того чтобы множество V некоторого аффинного пространства E размерности N было многообразием класса C^m размерности n , необходимо и достаточно, чтобы для каждой точки $a \in V$ существовала такая открытая окрестность \mathcal{U} точки a в E , чтобы пересечение $V \cap \mathcal{U}$ допускало параметрическое представление размерности n , принадлежащее классу C^m .

Доказательство. 1°) Предположим, что V является многообразием размерности n , принадлежащим классу C^m . Пусть a — некоторая точка V . Определим открытые множества \mathcal{B} и \mathcal{U} , как это было сделано в определении. Тогда для того, чтобы получить параметрическое представление искомого вида, достаточно положить $\mathcal{O} = \mathcal{B}$ и при помощи функций G_k из (III, 9; 1) определить отображение $u \rightarrow x = \Phi(u)$ формулами

$$x_1 = u_1, \dots, x_n = u_n, \quad x_{n+1} = G_1(u_1, \dots, u_n), \dots, x_N = G_{N-n}(u_1, \dots, u_n). \quad (\text{III}, 9; 2)$$

Здесь якобиан $\frac{D(x_1, \dots, x_n)}{D(u_1, \dots, u_n)}$ отличен от нуля¹⁾.

2°) Предположим теперь, что V есть некоторое множество пространства E , обладающее указанными свойствами. Дока-

¹⁾ То, что мы здесь проделали, напоминает «хитрый» прием первокурсника, который для исследования кривой $y = f(x)$ предпочтает записывать ее в виде $x = t$, $y = f(t)$.

жем, что тогда V является многообразием размерности m , принадлежащим классу C^m .

Пусть a — некоторая точка V и Φ — истинное параметрическое представление, т. е. гомеоморфизм некоторого открытого множества \mathcal{O} из \mathbb{K}^n на окрестность a в V . Выберем произвольно систему координат в E . Тогда отображение Φ будет иметь в точке $\alpha = \Phi^{-1}(a)$ некоторую производную матрицу по отношению к системам координат в \mathbb{K}^n и E . Эта производная матрица по предположению имеет ранг n , и по крайней мере один из ее определителей n -го порядка $\neq 0$. Изменяя при необходимости порядок векторов базиса E , мы можем предполагать, что этот определитель является якобианом $\frac{D(x_1, \dots, x_n)}{D(u_1, \dots, u_n)}$. Таким путем построенная система координат в E отождествляет это пространство с пространством \mathbb{K}^n , а разделение на систему из n первых координат и $N - n$ остальных отождествляет его с $\mathbb{K}^n \times \mathbb{K}^{N-n}$. Поэтому каждую точку x из E мы будем представлять как пару (y, z) , $y \in \mathbb{K}^n$, $z \in \mathbb{K}^{N-n}$. Пусть $a = (b, c)$. Отображение Φ является отображением открытого множества \mathcal{O} в произведение $\mathbb{K}^n \times \mathbb{K}^{N-n}$ и, следовательно, разлагается на систему двух отображений X и Ψ множества \mathcal{O} соответственно в пространства \mathbb{K}^n и \mathbb{K}^{N-n} , а именно: $y = X(u)$, $z = \Psi(u)$. Но тогда якобиан отображения X множества $\mathcal{O} \subset \mathbb{K}^n$ в точке α отличен от нуля. Следовательно, согласно теореме 29, существует открытое множество \mathcal{A} в множестве \mathcal{O} , содержащее точку α , и открытое множество \mathcal{B} в пространстве \mathbb{K}^n , содержащее точку b , такие, что X является гомеоморфизмом \mathcal{A} на \mathcal{B} , непрерывно дифференцируемым вместе со своим обратным гомеоморфизмом. Кроме того, отображение X принадлежит классу C^m , а следствие 2 теоремы 31 говорит о том, что его обратный гомеоморфизм также принадлежит классу C^m . Обозначим этот обратный гомеоморфизм \mathcal{B} на \mathcal{A} через X^{-1} . Образ \mathcal{A} при отображении Φ является открытым множеством в V , поскольку Φ есть гомеоморфизм множества \mathcal{O} на его образ, являющийся открытым множеством в V . Следовательно, существует такая открытая окрестность \mathcal{U} точки a в E , что $\Phi(\mathcal{A}) = V \cap \mathcal{U}$. Поскольку $\Phi(\mathcal{A})$ проектируется в пространство \mathbb{K}^n на множество \mathcal{B} , можно считать, что множество \mathcal{U} также проектируется на \mathcal{B} (в противном случае можно взять его пересечение с открытым множеством $\{x; (x_1, x_2, \dots, x_n) \in \mathcal{B}\}$ пространства \mathbb{K}^n). Для того чтобы точка $x = (y, z)$ принадлежала пересечению $V \cap \mathcal{U}$, необходимо и достаточно, чтобы она была образом некоторой точки из \mathcal{A} при отображении Φ , т. е. чтобы точка y принадлежала \mathcal{B} , а точка z была связана с точкой y формулой $z = \Psi(X^{-1}(y))$. Эта композиция $g = \Psi \circ X^{-1}$ двух отображений

класса C^m , по теореме 19, сама является отображением класса C^m . Итак, нам удалось найти окрестность \mathcal{V} точки a в пространстве E и некоторую систему координат в E , такие, что пересечение $V \cap \mathcal{V}$ представляется с помощью уравнения $z = g(y)$, где g — отображение \mathcal{B} в E класса C^m . Поскольку это верно для любой точки a из V , то V является многообразием размерности n класса C^m .

Одновременно мы доказали, что если заранее выбрать определенную систему координат в E , то для каждой точки a из V достаточно соответствующим образом изменить порядок векторов базиса для того, чтобы получить систему координат, обладающую свойствами, указанными в определении многообразия. Это изменение, очевидно, зависит от выбора точки a .

Таким образом, можно сформулировать

Следствие 1. Пусть V — некоторое множество аффинного пространства E размерности N . Для того чтобы V было многообразием размерности n класса C^m , необходимо и достаточно, чтобы при произвольно выбранной системе координат в E для каждой точки $a \in V$ выполнялись следующие условия: найдется такая перестановка $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ множества $1, 2, \dots, N$, что существуют некоторое открытое множество \mathcal{B} в подпространстве, порожденном началом координат и базисными векторами $e_{\sigma_1}, e_{\sigma_2}, \dots, e_{\sigma_n}$, и открытое множество $\mathcal{V} \subset E$, содержащее a , проекция которого на предыдущее подпространство совпадает с \mathcal{B} ; кроме того, существует система $N - n$ скалярных функций G_k , $k = 1, 2, \dots, N - n$, класса C^m , заданных на \mathcal{B} и таких, что пересечение $V \cap \mathcal{V}$ определяется уравнениями

$$x_{\sigma_{n+k}} = G_k(x_{\sigma_1}, x_{\sigma_2}, \dots, x_{\sigma_n}), \quad k = 1, 2, \dots, N - n. \quad (\text{III}, 9; 3)$$

Следствие 2. Пусть E и F — два конечномерных аффинных пространства одной и той же размерности, H — диффеоморфизм класса C^m некоторого открытого множества \mathcal{U} из E на открытое множество $H(\mathcal{U})$ из F . Тогда если V является многообразием размерности n класса C^m в множестве \mathcal{U} , то его образ $H(V)$ является многообразием размерности n класса C^m в $H(\mathcal{U})$.

Это следствие говорит о том, что многообразия класса C^m сохраняются при C^m -диффеоморфизмах.

Доказательство. Пусть $\Phi : \mathcal{O} \rightarrow \Phi(\mathcal{O})$ — параметрическое представление открытого множества $\Phi(\mathcal{O})$ из V . Тогда отображение $H \circ \Phi$ является гомеоморфизмом \mathcal{O} на $H(\Phi(\mathcal{O}))$, т. е. отображением класса C^m множества \mathcal{O} в пространство F . Пусть $a \in \mathcal{O}$. Так как ранг отображения $\Phi'(\alpha)$ равен n , а ото-

бражение $H'(\Phi(\alpha))$ является линейной биекцией, то ранг $(H \circ \Phi)'(\alpha) = H'(\Phi(\alpha)) \circ \Phi'(\alpha)$ также равен n (следствие 4 теоремы 11, примененное к отображению H). Значит, отображение $H \circ \Phi$ является параметрическим представлением размерности n класса C^m . Можно выбрать семейство представлений, подобных Φ , так, чтобы их образы покрывали V ; тогда образы $H \circ \Phi$ покрывают $H(V)$, чем и заканчивается доказательство следствия.

Следствие 2₂. Для того чтобы множество V аффинного пространства E размерности N было многообразием размерности n класса C^m , необходимо и достаточно, чтобы для каждой точки $a \in V$ существовал C^m -диффеоморфизм $\tilde{\Phi}$ открытого множества $\tilde{O} \subset \mathbb{K}^N$ на такую окрестность U точки a в E , что $V \cap U$ является образом при отображении $\tilde{\Phi}$ пересечения \tilde{O} и векторного подпространства \mathbb{K}^n , порожденного первыми n векторами базиса \mathbb{K}^N .

Это следствие говорит о том, что с помощью C^m -диффеоморфизма можно локально свести многообразие класса C^m размерности n в E к векторному подпространству размерности n в пространстве \mathbb{K}^N .

Доказательство. 1^o) Предположим сначала, что множество V обладает требуемым свойством. В каждой точке α из $\tilde{\Phi}^{-1}(V)$ производная $\tilde{\Phi}'(\alpha)$ является линейной биекцией \mathbb{K}^n на E (следствие 4 теоремы 11). Следовательно, она вместе со своим сужением на рассматриваемое подпространство \mathbb{K}^n является инъекцией, и, значит, это сужение имеет ранг n . Сужение Φ отображения $\tilde{\Phi}$ на $O = \tilde{O} \cap \mathbb{K}^n$ является параметрическим представлением размерности n класса C^m , и, поскольку каждая точка V имеет некоторую окрестность, допускающую такое параметрическое представление, V действительно является многообразием размерности n класса C^m .

2^o) Обратно, пусть V является многообразием размерности n класса C^m . Согласно определению, если $a \in V$, то существует система координат в E и некоторая окрестность U точки a в E , в которой V может быть определено формулами (III, 9; 1) с $(x_1, x_2, \dots, x_n) \in \mathcal{B}$. Обозначим через \mathcal{B}^* открытое множество из \mathbb{K}^N , образованное точками, первые n координат которых определяют точку из \mathcal{B} : $\mathcal{B}^* = \{x; (x_1, x_2, \dots, x_n) \in \mathcal{B}\}$. Тогда формулы

$$\begin{aligned} x_j &= u_j, \quad j = 1, 2, \dots, n, \\ x_{n+k} &= u_{n+k} + G_k(u_1, u_2, \dots, u_n), \quad k = 1, 2, \dots, N - n, \end{aligned} \quad (\text{III, 9; 3}_2)$$

определят некоторое отображение $\tilde{\Phi}$ множества \mathcal{B}^* в \mathbb{K}^N , совпадающее с E , причем $\tilde{\Phi}(\mathcal{B}^*) = \mathcal{B}$. Отображение $\tilde{\Phi}$ является

биекцией с обратной биекцией $\tilde{\Phi}^{-1}$, определяемой равенствами
 $u_j = x_j, j = 1, 2, \dots, n,$
 $u_{n+k} = x_{n+k} - G_k(x_1, x_2, \dots, x_n), k = 1, 2, \dots, N - n.$ (III, 9; 3₃)

Так как G_k принадлежат классу C^m , то оба отображения $\tilde{\Phi}$ и $\tilde{\Phi}^{-1}$ принадлежат классу C^m . Следовательно, $\tilde{\Phi}$ является гомеоморфизмом класса C^m вместе со своим обратным гомеоморфизмом. Если через \tilde{O} обозначить прообраз $\tilde{\Phi}^{-1}(\mathcal{V})$, то \tilde{O} является открытым множеством в \mathbb{K}^N , а сужение $\tilde{\Phi}$ на \tilde{O} обладает теми же свойствами. Но тогда $V \cap \mathcal{V}$ определяется уравнениями (III, 9; 1) и, следовательно, является образом подпространства $u_{n+k} = 0, k = 1, 2, \dots, N - n$, при отображении $\tilde{\Phi}$, чем и заканчивается доказательство следствия.

Если V является многообразием размерности n класса C^m , то всякое параметрическое представление Φ класса C^m открытого множества $V \cap \mathcal{V}$ этого многообразия называется также *локальной картой класса C^m* , или просто *картой*, если нет опасности путаницы. Множество $\Phi(O) = V \cap \mathcal{V}$ называется *образом этой карты*, а Φ называется также *картой множества $V \cap \mathcal{V}$* . Если $\Phi(O)$ содержит точку $a \in V$, то говорят также, что *карта покрывает эту точку*. *Атласом многообразия V* называется множество карт открытых множеств из V , образы которых покрывают V . *Универсальным атласом V* называется множество всех карт (размерности n класса C^m) открытых множеств из V^1).

Слова «карта» и «атлас», очевидно, взяты из картографии. Земная поверхность приближенно представляет собой сферу, и ее можно рассматривать как некоторое многообразие V размерности 2. Глобального параметрического представления этого компактного многообразия, заданного в некотором открытом множестве из \mathbb{R}^2 , не существует. Однако существует покрытие V системой открытых множеств, каждое из которых является образом некоторой карты. Поскольку V компактно, существует, естественно, *конечный атлас* для V , т. е. атлас, содержащий конечное число карт²). Каждая «страница» атласа является прямоугольником, который мы будем рассматривать как *открытый* прямоугольник из \mathbb{R}^2 . Около тех или иных точек этого прямоугольника отмечается название соответствующего места на Земле, т. е. некоторых точек множества V . На идеальной с математической точки зрения карте

¹⁾ Универсальный атлас есть множество карт, имеющее мощность континума.

²⁾ Географические атласы содержат конечное число карт. Универсальный атлас по причине, указанной в предыдущем примечании, был бы очень громоздким и обошелся бы слишком дорого.

каждой точке прямоугольника ставится в соответствие некоторая точка Земли, т. е. определяется отображение Φ открытого прямоугольника на некоторое открытое множество многообразия V . Из того факта, что карты являются открытыми множествами, немедленно вытекает, что некоторые точки V покрываются несколькими картами атласа (мы требуем, чтобы каждый город мира был представлен внутренней точкой хотя бы одной из прямоугольных карт). Если бы каждая точка покрывалась не более чем один раз, то V было бы объединением конечного числа >1 открытых непересекающихся множеств, т. е. было бы несвязным. Однако хотим мы этого или нет, Земля, как и всякая сфера, связна.

Примеры многообразий. 1°) Каждое линейное многообразие размерности n является бесконечно дифференцируемым многообразием. Оно содержит атлас, состоящий из одной карты, получаемой при выборе системы координат в этом многообразии, — системы координат, которая определяет некоторую линейную биекцию Φ пространства \mathbb{K}^n на это многообразие, обладающую всеми требуемыми свойствами. В частности, пространство E само является многообразием.

2°) Мы неявно предполагали, что $n \geq 1$. Можно рассматривать как многообразие размерности 0 каждое множество V изолированных точек E . Картой некоторой окрестности точки $a \in V$ является отображение $0 \rightarrow a$ множества $\mathbb{K}^0 = \{0\}$ в множество $\{a\}$.

Можно также рассматривать пустое множество \emptyset как многообразие размерности -1 , вовсе не имеющее карт.

Многообразие размерности 1 называется *кривой*, многообразие размерности 2 — *поверхностью*, а многообразие размерности $N - 1$ называется *гиперповерхностью*. Однако эти слова использовались и используются в столь различных смыслах, что надо быть осторожными при их применении. Приято рассматривать лемнискату Бернулли как «кривую»; однако из-за своей особой точки она не является многообразием. Говоря «кривая класса C^m », следует уточнить, что речь идет о «многообразии размерности 1 класса C^m », и лемниската будет исключена.

3°) Многообразие V размерности N аффинного пространства E размерности N — это просто некоторое открытое множество из E . (В самом деле, здесь $N - n = 0$. Если обратиться к определению, можно убедиться, что уравнение (III, 9; 1) отсутствует, а $V \cap \mathcal{V} = \mathcal{V} = \mathcal{B}$ является открытым множеством в E , содержащим a . Определение выражает просто тот факт, что каждая точка $a \in V$ имеет окрестность \mathcal{V} в E ,

принадлежащую V , т. е. что V открыто.) Каждое такое открытое множество является многообразием класса C^∞ и имеет атлас из одной карты, определяемой тождественным отображением $V \subset E$ в V , где E отождествляется с \mathbb{K}^n с помощью некоторой системы координат.

Каждая часть многообразия V из E , являющаяся открытым множеством в V , будет многообразием того же класса и той же размерности.

Обратно, если W и V — некоторые многообразия из E одной и той же размерности n и $W \subset V$, то W является просто некоторым открытым подмножеством из V . В самом деле, пусть $a \in W$. Согласно следствию 2₂ теоремы 32, существует C^m -дiffeоморфизм $\tilde{\Phi}$ некоторого открытого множества $\tilde{O} \subset \subset \mathbb{K}^n$ на открытое множество из E , содержащее точку a , такой, что множество $\tilde{\Phi}^{-1}(V)$ является пересечением \tilde{O} с некоторым векторным подпространством \mathbb{K}^n пространства \mathbb{K}^n . Следствие 2 показывает в свою очередь, что $\tilde{\Phi}^{-1}(W)$ является также многообразием из \mathbb{K}^n размерности n , заведомо содержащемся в \mathbb{K}^n . Из того что мы видели в начале пункта 3°), следует, что $\tilde{\Phi}^{-1}(W)$ является открытым множеством в пространстве \mathbb{K}^n и, значит, $\tilde{\Phi}^{-1}(W)$ является окрестностью точки $\tilde{\Phi}^{-1}(a)$ в множестве $\tilde{\Phi}^{-1}(V)$, а, следовательно, W является окрестностью a в V . Так как это утверждение справедливо для любой точки a из W , то W является открытым множеством в V .

4°) Покажем, что в аффинном евклидовом N -мерном пространстве над \mathbb{R} сфера является $(N - 1)$ -мерным бесконечно дифференцируемым многообразием. Для простоты выберем некоторую систему координат и в качестве сферы S возьмем единичную сферу с центром в начале. Обозначим через $\mathcal{V}_{j,+}$ открытое множество из E , определенное неравенством $x_j > 0$, а через \mathcal{B}_j — открытое множество из \mathbb{K}^{N-1} , определенное неравенством $x_1^2 + x_2^2 + \dots + x_{j-1}^2 + x_{j+1}^2 + \dots + x_N^2 < 1$. Тогда в $\mathcal{V}_{j,+}$ сфера S будет определяться уравнением

$$(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_N) \in \mathcal{B}_j, \\ x_j = \sqrt{1 - x_1^2 - \dots - x_{j-1}^2 - x_{j+1}^2 - \dots - x_N^2}. \quad (\text{III}, 9; 4)$$

Точно так же можно рассуждать для открытого множества $\mathcal{V}_{j,-}$, определенного неравенством $x_j < 0$. Так как все рассуждения можно провести для $j = 1, 2, \dots, N$, то исходное утверждение доказано полностью. Кроме того, мы видим, что сфера S покрывается атласом, состоящим из $2N$ карт класса C^∞ размерности $N - 1$.

Рассмотрим классическое параметрическое представление двумерной сферы в \mathbb{R}^3 :

$$\begin{aligned}x &= \sin \theta \cos \varphi, \\y &= \sin \theta \sin \varphi, \\z &= \cos \theta.\end{aligned}\tag{III, 9; 5}$$

Здесь в качестве \mathcal{O} можно взять, например, открытое множество $0 < \varphi < 2\pi$, $0 < \theta < \pi$ пространства \mathbb{R}^2 . Отображение (III, 9; 5) дает карту открытого множества сферы, определяемого неравенствами $y \neq 0$ или $y = 0$, $x < 0$ (дополнение к замкнутому полумеридиану $\varphi = 0$, $0 \leq \theta \leq \pi$). Обычно предыдущим представлением пользуются для всей сферы, допуская значения $0 \leq \varphi \leq 2\pi$, $0 \leq \theta \leq \pi$, но тогда это представление не будет истинным параметрическим представлением. Другими словами, одна эта карта, как мы уже видели, не может составлять атлас всей сферы. С помощью стереографической проекции легко показать, что существует атлас сферы, состоящий только из двух карт. В самом деле, рассмотрим обратное отображение к стереографической проекции, т. е. отображение \mathbb{R}^{N-1} на сферу, определяемое формулами

$$x_i = \frac{2u_i}{\sum_{j=1}^{N-1} u_j^2 + 1}, \quad i = 1, 2, \dots, N-1, \quad x_N = \varepsilon \frac{\sum_{j=1}^{N-1} u_j^2 - 1}{\sum_{j=1}^{N-1} u_j^2 + 1}. \tag{III, 9; 10}$$

При $\varepsilon = +1$ (соответственно -1) эти формулы определяют карту класса C^∞ , образом которой служит открытое множество сферы, дополнительное к северному полюсу (соответственно к южному полюсу) (т. е. к точке $0, 0, \dots, \varepsilon^1$). Мы имеем, таким образом, атлас сферы, состоящий из двух карт.

Из доказательства теоремы 32 вытекает, что отображение $\Theta: (y, z) \rightarrow X^{-1}(y)$ является отображением \mathcal{V} в \mathcal{A} , сужение которого на многообразие $V \cap \mathcal{V}$ является также сужением обратного гомеоморфизма Φ^{-1} к Φ . Таким образом, получаем такую теорему:

Теорема 33. *Если $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ является параметрическим представлением размерности n класса C^m открытого множества многообразия V из E и если $\Phi(\alpha) = a$, то существуют открытое множество \mathcal{A} в \mathcal{O} , содержащее α , открытое множество \mathcal{V} в E , содержащее a , такие, что $V \cap \mathcal{V} = \Phi(\mathcal{A})$, и, наконец, существует отображение Θ множества \mathcal{V} в \mathcal{A} класса C^m ,*

¹⁾ В качестве упражнения можно проверить выполнение всех необходимых свойств карты.

сужение которого на $V \cap \mathcal{U}$ совпадает с сужением обратного гомеоморфизма Φ^{-1} . В частности, в \mathcal{A} имеем: $\Theta \circ \Phi = I$.

Однако равенство $\Phi \circ \Theta = I$ не выполняется, поскольку если исходить из точки \mathcal{U} , не принадлежащей V , то ее образ при отображении Θ лежит в \mathcal{A} и потому образ при отображении Φ находится в V , а, значит, отличен от исходной точки. Напротив, можно сказать, что отображение $\Phi \circ \Theta$, будучи отображением $(y, z) \rightarrow (y, g(y))$ множества \mathcal{U} на $V \cap \mathcal{U}$, является своего рода «проекцией \mathcal{U} на $V \cap \mathcal{U}$ ».

Следствие 1. Пусть Φ_1 и Φ_2 — два параметрических представления класса C^m одного и того же открытого множества $V \cap \mathcal{U}$ многообразия V из E . Пусть \mathcal{O}_1 и \mathcal{O}_2 — открытые области определения функций Φ_1 и Φ_2 , такие, что $\Phi_1(\mathcal{O}_1) = \Phi_2(\mathcal{O}_2)$. Если через Φ_2^{-1} обозначить обратный гомеоморфизм к Φ_2 , являющийся отображением $V \cap \mathcal{U}$ в \mathcal{O}_2 , то отображение $\Phi_2^{-1} \circ \Phi_1$ множества \mathcal{O}_1 в \mathcal{O}_2 является гомеоморфизмом класса C^m , так же как и его обратный гомеоморфизм $\Phi_1^{-1} \circ \Phi_2$.

Доказательство. Пусть $\alpha_1 \in \mathcal{O}_1$ и $\Phi_1(\alpha_1) = a$. Рассмотрим отображения Θ_1 и Θ_2 , определенные, согласно теореме, на двух окрестностях \mathcal{U}_1 и \mathcal{U}_2 точки a в E . Пусть $\mathcal{U}_0 = \mathcal{U}_1 \cap \mathcal{U}_2$. Обозначим через \mathcal{A}'_1 и \mathcal{A}'_2 прообразы \mathcal{U}_0 при отображениях Φ_1 и Φ_2 . Тогда на множестве \mathcal{A}'_1 отображение $\Phi_2^{-1} \circ \Phi_1$ совпадает с отображением $\Theta_2 \circ \Phi_1$, поскольку $\Phi_1(u)$ для $u \in \mathcal{A}'_1$ лежит в $V \cap \mathcal{U}_0$ и на $V \cap \mathcal{U}_0 \subset V \cap \mathcal{U}_2$ отображение Φ_2^{-1} совпадает с Θ_2 . Следовательно, согласно теореме 19, эта композиция отображений класса C^m сама принадлежит классу C^{m+1}). Таким образом, каждая точка α_1 множества \mathcal{O}_1 имеет некоторую открытую окрестность, на которой $\Phi_2^{-1} \circ \Phi_1$ принадлежит классу C^m , а, следовательно, это отображение принадлежит классу C^m на \mathcal{O}_1 .

Следствие 2. Если V является многообразием размерности n , то оно не может быть многообразием размерности $n' \neq n$.

Доказательство. В самом деле, предположим, что для некоторой точки $a \in V$ имеется карта размерности n некоторой открытой окрестности точки a и карта размерности n' дру-

¹⁾ Проводить то же самое рассуждение непосредственно для $\Phi_2^{-1} \circ \Phi_1$ нельзя, ибо Φ_2^{-1} является отображением некоторого открытого множества из V , а не открытого множества аффинного пространства в \mathbb{K}^n , и, следовательно, утверждение, что Φ_2^{-1} принадлежит классу C^m , не имеет смысла. Позднее на стр. 336 мы узнаем, какой смысл можно ему придать.

гой открытой окрестности этой точки. Тогда для пересечения этих окрестностей будут существовать две различные карты размерностей n и n' . В рассуждениях, проведенных в предыдущем следствии с параметрическими представлениями, не предполагалось, что размерности n и n' должны быть равными. Однако $\Phi_2^{-1} \circ \Phi_1$ должно быть C^m -диффеоморфизмом множества \mathcal{O}_1 на \mathcal{O}_2 ; поэтому из следствия 4 теоремы 11 вытекает, что размерности карт должны быть равными.

В противоположность этому многообразие класса C^m тем более принадлежит классу C^k для любого $k \leq m$.

Определение многообразия с помощью неявных уравнений

Теорема 33₂. Для того чтобы множество V аффинного пространства E размерности N было многообразием размерности n класса C^m , необходимо и достаточно, чтобы для каждой точки $a \in V$ в E существовала окрестность \mathcal{U} этой точки и система $N-n$ скалярных функций F_k , $k = 0, 1, 2, \dots, N-n$, определенных на \mathcal{U} , класса C^m и обладающих следующими свойствами:

1°) производные $\overset{\leftarrow}{F'_k}(a)$ ¹⁾ образуют $N-n$ линейно независимых форм над \vec{E} ;

2°) пересечение $V \cap \mathcal{U}$ в точности определяется уравнениями $F_k(x) = 0$.

Говорят, что такая система уравнений является нормальной системой $N-n$ уравнений множества V в окрестности точки a .

Заметим, что если положить $n=N$, то уравнений не будет, а V будет некоторым открытым множеством из E . Если же положить $n=0$, то получится N уравнений, и в некоторой окрестности точки a других решений, кроме самой точки a , не будет. В этом случае a является изолированной точкой множества V , а V будет множеством изолированных точек E , или многообразием размерности 0.

Доказательство. 1°) Предположим, что V является многообразием. По определению это многообразие задается уравнениями $F_k(x) = 0$, где F_k определяются равенствами

$$F_k(x_1, x_2, \dots, x_N) = x_{n+k} - G_k(x_1, x_2, \dots, x_n). \quad (\text{III}, 9; 11)$$

Поскольку в точке a (или в любой точке из \mathcal{U}) dF_k содержат dx_{n+k} с коэффициентом 1 и не содержат ни одного из dx_{n+j} для $j = 1, 2, \dots, N-n$, то dF_k линейно независимы.

¹⁾ $F'_k(a) \equiv \mathcal{L}(\vec{E}; K) = \overset{\leftarrow}{E'}$.

2°) Предположим теперь, что V является множеством, обладающим указанными в теореме свойствами. Утверждение

о том, что система линейных форм $F'_k(a)$ независима, означает, что если в E выбрать какую-либо систему координат, то по крайней мере один из миноров порядка $N-n$ матрицы, составленной из частных производных функций F_k , отличен от нуля. Предположим для определенности, что отличен от нуля якобиан $\frac{D(F_1, F_2, \dots, F_{N-n})}{D(x_{n+1}, x_{n+2}, \dots, x_N)}(a)$ (к этому случаю можно всегда перейти, изменив при необходимости порядок нумерации векторов базиса E). Положим, как и ранее, $y = (x_1, x_2, \dots, x_n)$ и $z = (x_{n+1}, x_{n+2}, \dots, x_N)$, так чтобы точку $x \in E$ можно было отождествить с парой (y, z) пространства $\mathbb{K}^n \times \mathbb{K}^{N-n}$. Теперь можно считать, что множество функций F_k определяет функцию \vec{f} на $\mathbb{K}^n \times \mathbb{K}^{N-n}$ со значениями в \mathbb{K}^{N-n} , такую, что $\vec{f}(a) = \vec{f}(b, c) = \vec{0}$.

Предположение относительно якобиана означает, что частная производная $\frac{\partial f}{\partial z}(a) = \frac{\partial f}{\partial z}(b, c) \in \mathcal{L}(\mathbb{K}^{N-n}; \mathbb{K}^{N-n})$ этой функции обратима. Из теоремы о неявной функции (теорема 29) следует, что существуют окрестность \mathcal{B} точки b в \mathbb{K}^n и окрестность \mathcal{C} точки c в \mathbb{K}^{N-n} , такие, что $\mathcal{B} \times \mathcal{C} \subset \mathcal{V}$, а уравнение $\vec{f}(y, z) = \vec{0}$ имеет, и притом единственное, решение в \mathcal{C} для точек y , заданных в окрестности \mathcal{B} . Кроме того, определяемая таким способом неявная функция $z = g(y)$ является непрерывным отображением \mathcal{B} в \mathcal{C} . Из теоремы 31 следует, что если \vec{f} принадлежит классу C^m , то g также будет принадлежать классу C^m . Мы видим, что пересечение V с $\mathcal{B} \times \mathcal{C}$ определяется в точности уравнением $z = g(y)$ и что, поскольку это утверждение справедливо для любой точки $a \in V$, множество V является многообразием размерности n класса C^m .

Приведем другое доказательство утверждения 2°). Обозначим через Ψ отображение множества \mathcal{V} в пространство \mathbb{K}^n , определенное соотношением $(x_1, x_2, \dots, x_N) \rightarrow (u_1, u_2, \dots, u_n)$, где

$$\begin{aligned} u_1 &= x_1, & u_2 &= x_2, \dots, & u_n &= x_n, \\ u_{n+k} &= F_k(x_1, x_2, \dots, x_N), & k &= 1, 2, \dots, N-n, \end{aligned} \quad (\text{III}, 9; 11_2)$$

и будем исходить, как и в первом доказательстве, из случая $\frac{D(F_1, F_2, \dots, F_{N-n})}{D(x_{n+1}, \dots, x_N)}(a) \neq 0$. Поскольку этот якобиан является якобианом отображения Ψ в точке a и отличен от нуля, сущ-

стует такое открытое множество $\mathcal{V}_1 \subset \mathcal{V}$, содержащее a , что сужение Ψ на \mathcal{V}_1 является гомеоморфизмом класса C^m , точно так же как и его обратный гомеоморфизм Φ множества \mathcal{V}_1 на некоторое открытое множество \mathcal{A}_1 , содержащее $\Psi(a) = a$ (теоремы 29 и 31). Но тогда Φ является C^m -дiffeоморфизмом множества \mathcal{A}_1 на \mathcal{V}_1 , а $V \cap \mathcal{V}_1$ есть образ векторного подпространства \mathbb{K}^n пространства \mathbb{K}^N при отображении $\Phi: u_{n+k} = 0, k = 1, 2, \dots, N - n$. Из следствия 2₂ теоремы 32 теперь вытекает, что V — многообразие размерности n класса C^m .

Следствие 1. Для того чтобы множество V пространства E было гиперповерхностью (т. е. многообразием размерности $N - 1$) пространства E класса C^m , необходимо и достаточно, чтобы для каждой точки $a \in V$ существовали окрестность \mathcal{U} точки a в E и скалярная функция F , определенная в \mathcal{U} , класса C^m , производная которой $\overleftarrow{F'(a)} \neq \overleftarrow{0}$ ¹⁾, такие, что множество $V \cap \mathcal{U}$ в точности определяется уравнением $F(x) = 0$.

Теперь становится очевидным, что сфера евклидова пространства является многообразием, поскольку сфера с центром в точке a радиуса R определяется в E уравнением $\overrightarrow{(x-a)} \overrightarrow{|x-a|} = R^2$.

Этот пример показывает, что в то время как, вообще говоря, многообразие нельзя полностью задать с помощью разрешенных уравнений или же с помощью параметрического представления, возможность определить его целиком с помощью неявных уравнений указанного в теореме вида значительно более вероятна.

Естественно, условие, что $\overleftarrow{F'_k(a)}$ независимы или, в случае одного уравнения, что производная $\overleftarrow{F'(a)}$ отлична от нуля, абсолютно необходимо. Например, конус второго порядка, определенный в пространстве \mathbb{R}^3 уравнением $x^2 + y^2 - z^2 = 0$, не является многообразием. Он не удовлетворяет ни одному из определений из-за особой точки в начале координат.

Вещественные и комплексные многообразия

До сих пор в качестве поля \mathbb{K} можно было брать поле вещественных или поле комплексных чисел. В соответствии с первым или вторым случаем говорят, что V является *вещественным* многообразием или *комплексным* многообразием.

1) Напомним, что $\overleftarrow{F'(a)} \neq \overleftarrow{0}$ в некоторой системе координат означает, что $\frac{\partial F}{\partial x_i}(a), i = 1, 2, \dots, N$, не обращаются в нуль одновременно.

Любое аффинное пространство над полем комплексных чисел является также аффинным пространством удвоенной размерности над полем вещественных чисел. Точно так же \mathbb{C}^n , рассматриваемое как векторное пространство над полем вещественных чисел, может быть отождествлено с \mathbb{R}^{2n} . Так как каждое отображение, дифференцируемое по отношению к полю комплексных чисел, тем более дифференцируемое по отношению к полю вещественных чисел, то каждое многообразие V размерности n класса C^m по отношению к полю комплексных чисел можно рассматривать как многообразие размерности $2n$ того же класса C^m по отношению к полю вещественных чисел. Если нет специальных указаний, будем считать размерностью многообразия его размерность по отношению к полю вещественных чисел.

Абстрактные многообразия¹⁾

Для определения многообразия нет необходимости вести рассмотрение только в аффинном пространстве. Например, множество страниц географического атласа дает полное описание поверхности Земли, и при этом нет необходимости представлять себе Землю как двумерное многообразие, погруженное в аффинное пространство трех измерений. Можно ввести понятие абстрактного многообразия.

Определение. Абстрактным многообразием V размерности n класса C^m (не обязательно лежащим в некотором аффинном пространстве) называется топологическое пространство V , снаженное некоторым атласом или системой карт, обладающих следующими свойствами:

1°) Каждая карта является гомеоморфизмом Φ некоторого открытого множества \mathcal{O} пространства \mathbb{K}^n на открытое множество $\Phi(\mathcal{O})$ пространства V . (Говорить, что этот гомеоморфизм принадлежит классу C^m , не приходится, поскольку речь идет об отображении открытого множества аффинного пространства в топологическое пространство, не лежащее в аффинном пространстве.)

2°) Пусть Φ_1 и Φ_2 — две произвольные карты атласа множества V . Если образы $\Phi_1(\mathcal{O}_1)$ и $\Phi_2(\mathcal{O}_2)$ имеют непустое пересечение Ω , то прообразы Ω_1 и Ω_2 множества Ω при отображениях Φ_1 и Φ_2 являются открытыми множествами в \mathcal{O}_1 и \mathcal{O}_2 . Сужения Φ_1 и Φ_2 на Ω_1 и Ω_2 имеют один и тот же образ. Они должны обладать следующим свойством, аналогичным утверждению следствия 1 теоремы 33: отображение $\Phi_2^{-1} \circ \Phi_1$ откры-

¹⁾ Здесь мы приводим лишь беглый набросок теории. Абстрактные многообразия подробно изучаются в других курсах.

того множества $\Omega_1 \subset \mathbb{K}^n$ на открытое множество $\Omega_2 \subset \mathbb{K}^n$ принадлежит классу C^m . Эта фраза имеет определенный смысл, ибо речь идет об отображении открытого множества аффинного пространства в аффинное пространство. Именно этим условием определяется принадлежность многообразия V классу C^m .

3°) *Множество образов карт должно составлять покрытие пространства V^1 .*

Изучение абстрактных многообразий играет большую роль в математике и во многих разделах физики. Рассмотрим, например, механическую систему, имеющую конечное число степеней свободы, например гироскоп с одной закрепленной точкой оси вращения. Положение этой системы «может быть определено с помощью конечного числа вещественных параметров q_1, q_2, \dots, q_n ». В действительности эта система параметров весьма произвольна, и очень редко удается все без исключения положения механической системы эффективно выразить через значения конечного числа параметров. Так, например, если положение гироскопа определяется углами Эйлера ψ, θ, φ , то множество всех положений гироскопа не имеет параметрического представления.

В действительности *множество положений механической системы может быть определено как абстрактное многообразие, которое в случае гироскопа является трехмерным многообразием V класса C^∞* . Но это многообразие является абстрактным и, естественно, не погружено в аффинное пространство. Когда говорят, что 3 угла Эйлера выбираются для определения положения гироскопа, то это означает, что рассматривается частная карта, представляющая только некоторое открытое множество многообразия, являющегося множеством положений гироскопа²⁾.

Тем не менее теория абстрактных многообразий приводит в некоторых вопросах к значительно большим усложнениям, чем теория многообразий, погруженных в аффинные пространства. Если это не будет приводить ни к каким усложнениям, мы будем рассматривать абстрактные многообразия. Каждый

¹⁾ Разрешается, вообще говоря, дополнить первоначально заданный атлас. Для этого добавляются все карты, не нарушающие принадлежность классу C^m , т. е. все гомеоморфизмы Φ открытых множеств \mathcal{O} из \mathbb{K}^n на открытые множества из V , обладающие следующим свойством: для каждой карты Φ_i заданного атласа, такой, что $\Phi(\mathcal{O}) \cap \Phi_i(\mathcal{O}_i) = \Omega \neq \emptyset$, отображение $\Phi^{-1} \circ \Phi_i$ является C^m -дiffeоморфизмом $\Phi_i^{-1}(\Omega)$ на $\Phi^{-1}(\Omega)$.

²⁾ Для получения параметрического представления надо ограничиться, например, открытым множеством $0 < \psi < 2\pi, 0 < \theta < \pi, 0 < \varphi < 2\pi$, которое не исчерпывает множества всех положений гироскопа. Впрочем, многообразие здесь компактно и для полного представления достаточно конечного числа (>1) карт.

раз, когда это может привести к упрощению, мы будем предполагать, что многообразия лежат в аффинных пространствах.

Определение. Пусть V и W — многообразия класса C^m размерностей p и q над одним и тем же полем \mathbb{K} , и пусть H — непрерывное отображение V в W . Говорят, что H принадлежит классу C^k , если, каковы бы ни были точка $a \in V$, $H(a) = b$, карта Φ окрестности \mathcal{V} точки a в V и карта Ψ окрестности \mathcal{W} точки b в W , сложное отображение $\Psi^{-1} \circ H \circ \Phi$, определенное на открытом множестве $\Phi^{-1}(\mathcal{V}) \cap \Psi^{-1}(H^{-1}(\mathcal{W}))$ пространства \mathbb{K}^p , со значениями в \mathbb{K}^q принадлежит классу C^k в обычном смысле. Очевидно, если многообразие принадлежит классу C^m , то определить отображение класса $C^{m'}$ для $m' > m$ невозможно. Очевидно также, что любое отображение класса C^k принадлежит классу $C^{k'}$ для $k' \leq k$. Всегда можно ограничиться случаем $k = m$, поскольку если V и W принадлежат классу C^m , $m \geq k$, то они тем более принадлежат классу C^k .

Если U , V , W — три многообразия класса C^m , H — отображение класса C^m множества U в V , K — отображение класса C^m множества V в W , то $K \circ H$ является отображением U в W класса C^m . В самом деле, если Φ (соответственно Ψ) является картой некоторой окрестности \mathcal{U} точки a (соответственно \mathcal{W} точки $c = (K \circ H)(a)$) в множестве U (соответственно W) и если через X мы обозначим карту некоторой окрестности \mathcal{V} точки $b = H(a)$ в V , то на множестве $\Phi^{-1}(\mathcal{U}) \cap \Psi^{-1}(H^{-1}(\mathcal{W})) \cap \Phi^{-1}(H^{-1}(K^{-1}(\mathcal{W})))$

$$\Psi^{-1} \circ K \circ H \circ \Phi = (\Psi^{-1} \circ K \circ X) \circ (X^{-1} \circ H \circ \Phi); \quad (\text{III}, 9; 12)$$

каждое из двух отображений правой части принадлежит, по предположению, классу C^m , а, значит, согласно теореме 19, их композиция также принадлежит классу C^m . Гомеоморфизм V на W , принадлежащий классу C^m вместе со своим обратным гомеоморфизмом, называется C^m -дiffeоморфизмом. Если такой диффеоморфизм существует, то, согласно следствию теоремы 11, V и W имеют одинаковую размерность.

Теорема 33₃. Пусть V и W — многообразия класса C^m и H — непрерывное отображение V в W . Для того чтобы H принадлежало классу C^m , необходимо и достаточно, чтобы для каждой точки $a \in V$ существовали хотя бы одна карта Φ_a некоторой окрестности точки a в V и хотя бы одна карта Ψ_b некоторой окрестности точки $b = H(a)$ в W , такие, что отображение $\Psi_b^{-1} \circ H \circ \Phi_a$ принадлежит классу C^m .

Доказательство. Необходимость этого условия очевидна. Надо доказать лишь его достаточность. Для этого мы должны показать, что если Φ и Ψ — произвольные карты окрестностей точек a и b в V и W , то $\Psi^{-1} \circ H \circ \Phi$ принадлежит классу C^m . В окрестности точки $a = \Phi^{-1}(b)$ имеем

$$\Psi^{-1} \circ H \circ \Phi = (\Psi^{-1} \circ \Psi_a) \circ (\Psi_a^{-1} \circ H \circ \Phi_a) \circ (\Phi_a^{-1} \circ \Phi). \quad (\text{III}, 9; 13)$$

Средний член правой части предполагается принадлежащим классу C^m . Крайние члены, согласно аксиоме 2°) на стр. 334 о картах многообразия, также принадлежат классу C^m . Это и доказывает теорему, которая имеет многочисленные следствия. Приведем наиболее важные из них:

Следствие 1. Если V и W являются открытыми множествами аффинных пространств E и F , то отображение H множества V в W принадлежит классу C^m тогда и только тогда, когда оно принадлежит классу C^m в смысле, определенном на стр. 266.

В самом деле, возьмем для V и W тождественные карты. Тогда $\Psi^{-1} \circ H \circ \Phi$ сводится к H .

Следствие 2. Пусть V — многообразие класса C^m аффинного пространства E . Тогда каноническая инъекция H многообразия V в $W = E$ также принадлежит классу C^m .

Для доказательства достаточно взять произвольную карту Φ для V и тождественную карту для $W = E$. Тогда $\Psi^{-1} \circ H \circ \Phi = \Phi$, где Φ — отображение класса C^m .

Следствие 3. Пусть V лежит в аффинном пространстве E . Тогда любое отображение H множества V в W , являющееся сужением на V некоторого отображения \tilde{H} класса C^m открытого множества $\Omega \subset E$ в W , является отображением V в W класса C^m .

В самом деле, в этом случае H является композицией $\tilde{H} \circ J$, где J — каноническая инъекция V в E , которая, согласно следствию 2, принадлежит классу C^m .

Следствие 4. Пусть V — многообразие размерности n класса C^m , а Φ — карта некоторого открытого множества \mathcal{V} многообразия V . Тогда Φ является C^m -дiffeоморфизмом множества $\Phi^{-1}(\mathcal{V})$ на \mathcal{V} .

В самом деле, возьмем для множества $\Phi^{-1}(\mathcal{V}) \subset \mathbb{K}^m$ тождественную карту, а для множества \mathcal{V} карту Φ . Для $H = \Phi$ или Φ^{-1} отображение вида $\Psi_1^{-1} \circ H \circ \Phi_1$ является тождественным отображением множества $\Phi^{-1}(\mathcal{V})$, принадлежащим тем самым

классу C^m . Следствие 1 теоремы 33 очевидным образом вытекает теперь из следствия 4, однако мы уже неоднократно использовали его для получения предыдущих результатов.

Следствие 5. Предположим, что многообразие W размерности q погружено в аффинное пространство F . Для того чтобы отображение H многообразия V в W было класса C^m , необходимо и достаточно, чтобы для любой карты Φ для V отображение $H \circ \Phi$ принадлежало классу C^m или чтобы для любой точки $a \in V$ существовала карта Φ_a некоторой окрестности точки a в V , такая, что отображение $H \circ \Phi_a$ принадлежит классу C^m .

В самом деле, пусть Ψ — карта некоторой окрестности точки $b = H(a) \in W$. Уменьшая при необходимости эту окрестность, можно сделать так, чтобы отображение Ψ^{-1} продолжалось в отображение Θ класса C^m некоторой окрестности точки $b \in F$ в \mathbb{K}^q . Теперь мы должны показать, что для каждой карты Φ (или для некоторой частной карты) $\Psi^{-1} \circ H \circ \Phi$ принадлежит классу C^m тогда и только тогда, когда $H \circ \Phi$ принадлежит классу C^m . Так как $H \circ \Phi = \Psi \circ (\Psi^{-1} \circ H \circ \Phi)$ и $\Psi^{-1} \circ H \circ \Phi = \Theta \circ (H \circ \Phi)$, где отображения Ψ и Θ , определенные на открытых множествах аффинных пространств, принадлежат классу C^m , то результат вытекает из теоремы 19.

Это следствие можно сформулировать следующим образом: если $W \subset F$, то отображение H множества V в W принадлежит классу C^m тогда и только тогда, когда это отображение, рассматриваемое как отображение V в F , принадлежит классу C^m .

Векторное пространство, касательное в точке к многообразию аффинного пространства E размерности N

Теорема 334. Пусть V — многообразие размерности n класса C^1 в аффинном пространстве E . В каждой точке $a \in V$ векторная (соответственно аффинная) контингенция в смысле теоремы 8А является векторным подпространством в \vec{E} (соответственно аффинным подпространством в E) размерности n . Это подпространство называется векторным (соответственно аффинным) пространством, касательным в точке a к многообразию V , и обозначается через $\vec{T}(a; V)$ (соответственно $T(a; V)$).

Если $n = N$, то векторным (соответственно аффинным) касательным пространством является само пространство \vec{E} (соответственно E). Если $n = 0$, то им будет $\{\vec{0}\}$ (соответственно $\{a\}$).

Доказательство. Для доказательства достаточно обратиться к определению многообразия. Выбирая соответствующую систему координат в E , это пространство можно отождествить с пространством $\mathbb{K}^n \times \mathbb{K}^{N-n}$ и определить многообразие в окрестности точки a уравнением $z = g(y)$. Мы приходим теперь к теореме 8А и, учитывая формулу (III, 3; 19₂), видим, что векторное подпространство, касательное в точке $a = (b, c)$, определяется в выбранной системе координат уравнением

$$\vec{Z} = g'(b) \vec{Y} \quad (\text{III}, 9; 14)$$

или что линейное многообразие, касательное в точке a , определяется (в той же системе координат) уравнением

$$\overrightarrow{z - c} = g'(b) \overrightarrow{(y - b)}. \quad (\text{III}, 9; 15)$$

Следствие 1. Пусть V — многообразие класса C^1 размерности n аффинного пространства E . Пусть Φ — некоторая карта, т. е. отображение открытого множества $\mathcal{O} \subset \mathbb{K}^n$ на открытое множество из V . Пусть α — некоторая точка из \mathcal{O} и $a = \Phi(\alpha)$. Производное отображение $\Phi'(\alpha)$ является линейной биекцией \mathbb{K}^n на векторное пространство, касательное в точке a к многообразию V .

Доказательство. Предположим сначала, что Φ — отображение класса C^1 множества \mathcal{O} в V , которое может и не быть истинным параметрическим представлением ($\Phi'(\alpha)$ не обязательно имеет ранг n). Пусть \vec{X} — некоторый вектор из \mathbb{K}^n . Рассмотрим последовательность точек $x_n = \Phi(\alpha + t_n \vec{X})$, где t_n образуют последовательность положительных вещественных чисел, стремящуюся к нулю при n , стремящимся к бесконечности. Точки x_n принадлежат V и стремятся к a . С другой стороны,

векторы $\frac{1}{t_n} (\overrightarrow{x_n - a})$, которые можно также записать в виде

$\overrightarrow{\Phi(\alpha + t_n \vec{X}) - \Phi(\alpha)} / t_n$, сходятся к производной отображения Φ в

точке α по вектору \vec{X} , т. е. к $\Phi'(\alpha) \vec{X}$. Этот вектор принадлежит векторному пространству, касательному к V в точке a . Отсюда вытекает, что образ \mathbb{K}^n при отображении $\Phi'(\alpha)$ содержится в касательном векторном пространстве. Если теперь мы учтем, что по предположению $\Phi'(\alpha)$ имеет ранг n , то получим, что этот образ является векторным подпространством размерности n и, следовательно, совпадает со всем касательным пространством в точке a , а отображение $\Phi'(\alpha)$ есть линейная биекция пространства \mathbb{K}^n на это касательное векторное пространство.

Отсюда вытекает, что векторы $\frac{\partial \Phi}{\partial u_j}(a)$ образуют базис векторного пространства, касательного к V в точке a , и что линейное многообразие, касательное к V в точке a , задается параметрическим уравнением

$$x = a + \sum_{j=1}^n t_j \frac{\partial \Phi}{\partial u_j}(a), \quad t_j \in \mathbb{K}^1. \quad (\text{III}, 9; 16)$$

Следствие 1₂. Пусть V и W — два многообразия класса C^1 в двух аффинных пространствах E и F . Пусть H — такое отображение E в F класса C^1 , что $H(V) \subset W$. Тогда для каждого $a \in V$ при $b = H(a) \in W$ образ касательного векторного пространства $\vec{T}(a; V)$ при отображении $H'(a) \in \mathcal{L}(E; F)$ содержится в касательном векторном пространстве $\vec{T}(b; W)$: $H'(a) \vec{T}(a; V) \subset \vec{T}(b; W)$.

Доказательство. Пусть Φ — карта ($\mathcal{O} \subset \mathbb{K}^n \rightarrow \Phi(\mathcal{O})$) некоторой окрестности $\Phi(\mathcal{O})$ точки a в V и $\Phi(a) = b$. Тогда $H \circ \Phi$ является отображением класса C^1 множества \mathcal{O} в $W \subset F$. В начале доказательства предыдущего следствия мы видели, что $(H \circ \Phi)'(a) \mathbb{K}^n \subset \vec{T}(b; W)$. Согласно теореме о сложной функции, $(H \circ \Phi)'(a) = H'(a) \circ \Phi'(a)$, а по следствию 1, $\Phi'(a) \mathbb{K}^n = \vec{T}(a; V)$, откуда и вытекает доказываемое утверждение.

Следствие 2. Пусть выполнены условия следствия 1. Если \mathcal{O} — такое открытое множество пространства \mathbb{K}^n , содержащее a , что сужение Φ на \mathcal{O} допускает обратное слева отображение Θ , удовлетворяющее условиям теоремы 33, то линейная биекция $\Phi'(a)$ пространства \mathbb{K}^n на $\vec{T}(a; V)$ имеет в качестве обратной биекции сужение на $\vec{T}(a; V)$ производного отображения $\Theta'(a)$.

Доказательство. По теореме 11 о сложной функции из тождества $\Theta \circ \Phi = I$ получается тождество $\Theta'(a) \circ \Phi'(a) = I$, из которого следует, что сужение $\Theta'(a)$ на образ \mathbb{K}^n при отображении $\Phi'(a)$, т. е. на $\vec{T}(a; V)$, является обратной биекцией для $\Phi'(a)$.

¹⁾ В весьма частном случае, когда V — поверхность пространства \mathbb{R}^3 , заданная параметрическим представлением $(u, v) \rightarrow M(u, v)$, касательная плоскость в некоторой точке порождается частными производными $\partial M / \partial u$, $\partial M / \partial v$ в этой точке.

Следствие 3. Пусть Φ_1 и Φ_2 — отображения открытых множеств \mathcal{O}_1 и \mathcal{O}_2 пространства \mathbb{K}^n в V , определяющие карты с одним и тем же образом в многообразии V , погруженном в E . Пусть \vec{X} — вектор, касательный к V в точке a , $a = \Phi_1(\alpha_1) = \Phi_2(\alpha_2)$. Если $\vec{\xi}_1$ и $\vec{\xi}_2$ — два таких вектора пространства \mathbb{K}^n , для которых вектор \vec{X} является образом при отображениях $\Phi'_1(\alpha_1)$ и $\Phi'_2(\alpha_2)$, то $\vec{\xi}_2$ является образом $\vec{\xi}_1$ относительно производной отображения $\Phi_2^{-1} \circ \Phi_1$ класса C^1 в точке α_1 .

В самом деле, достаточно заметить, что

$$\vec{\xi}_2 = ((\Phi'_2(\alpha_2))^{-1} \circ \Phi'_1(\alpha_1)) \vec{\xi}_1. \quad (\text{III}, 9; 17)$$

Точнее, согласно предыдущему следствию, обратное отображение к $\Phi'_2(\alpha_2)$ является также сужением отображения $\Theta'_2(a)$ на векторное подпространство в точке a . Поэтому можно записать зависимость между $\vec{\xi}_2$ и $\vec{\xi}_1$ в виде

$$\vec{\xi}_2 = (\Theta'_2(a) \circ \Phi'_1(\alpha_1)) \vec{\xi}_1. \quad (\text{III}, 9; 18)$$

На этот раз Φ_1 и Θ_2 являются отображениями открытого множества некоторого аффинного пространства в аффинное пространство. Поэтому можно применить теорему о сложной функции (теорему 11) и заменить предыдущее равенство равенством

$$\vec{\xi}_2 = (\Theta_2 \circ \Phi_1)'(\alpha_1) \vec{\xi}_1. \quad (\text{III}, 9; 19)$$

Наконец, из свойств отображения Θ_2 следует, что $\Theta_2 \circ \Phi_1$ совпадает с $\Phi_2^{-1} \circ \Phi_1$ и, следовательно, можно написать

$$\vec{\xi}_2 = (\Phi_2^{-1} \circ \Phi_1)'(\alpha_1) \vec{\xi}_1, \quad (\text{III}, 9; 20)$$

чём и заканчивается доказательство следствия.

Следствие 4. Пусть V — многообразие класса C^1 размерности n аффинного пространства E . Пусть a — некоторая точка V , и пусть $F_k(x) = 0$, $k = 1, 2, \dots, N - n$, — нормальная система уравнений многообразия V в некоторой окрестности точки a в E . Тогда векторное подпространство, касательное к V в точке a , определяется в E уравнениями

$$F'_k(a) \vec{X} = 0, \quad k = 1, 2, \dots, N - n, \quad (\text{III}, 9; 21)$$

а линейное многообразие, касательное к V в точке a , определяется в E уравнениями

$$F'_k(a) \overrightarrow{(x - a)} = 0, \quad k = 1, 2, \dots, N - n. \quad (\text{III}, 9; 22)$$

Доказательство. Пусть x_n — последовательность точек многообразия V , стремящаяся к a , и λ_n — такая последовательность положительных вещественных чисел, что последовательность векторов $\lambda_n(\overrightarrow{x_n - a})$ стремится к вектору \vec{X} из \vec{E} при n , стремящемся к бесконечности. Поскольку тогда $F_k(x_n) = 0$ и $F_k(a) = 0$, из определения производной (формула (III, 3; 13)) следует, что

$$\overrightarrow{F'_k(a)(x_n - a)} + \overrightarrow{\alpha_n(x_n - a)} = 0, \quad (\text{III}, 9; 23)$$

где $\|\overrightarrow{\alpha_n}\|$ стремится к 0 при n , стремящемся к бесконечности. Отсюда, умножая на λ_n , получаем:

$$\overrightarrow{F'_k(a)\lambda_n(x_n - a)} + \overrightarrow{\alpha_n\lambda_n(x_n - a)} = 0. \quad (\text{III}, 9; 24)$$

Поскольку $\lambda_n(\overrightarrow{x_n - a})$ сходится к \vec{X} , а $\overrightarrow{F'_k(a)}$ является непрерывной линейной формой, то каждый вектор \vec{X} , касательный в точке a к многообразию V , удовлетворяет $N-n$ линейным уравнениям (III, 9; 21). Поскольку эти $N-n$ линейных уравнений предполагаются независимыми, они определяют некоторое n -мерное векторное подпространство пространства \vec{E} . Это подпространство является векторным пространством, касательным к V в точке a . Отсюда непосредственно вытекает утверждение, относящееся к касательному линейному многообразию.

Следствие 5. Пусть E — аффинное евклидово пространство размерности N над полем вещественных чисел и V — гиперповерхность, определяемая уравнением $F(x) = 0$, где F — вещественная функция класса C^1 , градиент которой в точке a отличен от 0. Этот градиент нормален к V в точке a , а вектор $\frac{\overrightarrow{\text{grad } F(a)}}{\|\overrightarrow{\text{grad } F(a)}\|}$ является единичным вектором нормали к V в точке a ¹⁾.

¹⁾ Если в E выбрать ортонормальный базис, то

$$\begin{aligned} & \frac{\overrightarrow{\text{grad } F(a)}}{\|\overrightarrow{\text{grad } F(a)}\|} = \\ & = \left[\frac{\frac{\partial F}{\partial x_1}(a)}{\sqrt{\sum_{i=1}^N \left(\frac{\partial F}{\partial x_i}(a) \right)^2}}, \frac{\frac{\partial F}{\partial x_2}(a)}{\sqrt{\sum_{i=1}^N \left(\frac{\partial F}{\partial x_i}(a) \right)^2}}, \dots, \frac{\frac{\partial F}{\partial x_N}(a)}{\sqrt{\sum_{i=1}^N \left(\frac{\partial F}{\partial x_i}(a) \right)^2}} \right]. \end{aligned}$$

Следствие 6. Пусть V — многообразие размерности n класса C^1 аффинного пространства E . Пусть Φ — карта, отображающая открытое множество $\mathcal{O} \subset \mathbb{K}^n$ на открытое множество многообразия V . В этом случае отображение $(u, \vec{U}) \rightarrow (\Phi(u), \Phi'(u) \vec{U})$, ставящее в соответствие каждой паре (u, \vec{U}) , где u — точка, а \vec{U} — вектор из \mathbb{K}^n , пару, составленную из образа $\Phi(u)$, являющегося точкой V , и вектора $\Phi'(u) \vec{U}$, касательного к многообразию V в этой точке, является гомеоморфизмом $\mathcal{O} \times \mathbb{K}^n$ на свой образ.

Доказательство. Рассматриваемое отображение непрерывно (теорема 10) и биективно. Для доказательства того, что это отображение является гомеоморфизмом, достаточно выяснить, что его обратное отображение непрерывно. Впрочем, поскольку непрерывность является локальным свойством, достаточно ограничиться сужением обратного отображения на множество пар (x, \vec{X}) , для которых x пробегает некоторую окрестность точки a в V . Пользуясь теперь теоремой 33 и выбирая эту окрестность так, как указывается в этой теореме, можно убедиться (следствие 2), что обратное отображение является сужением отображения $(x, \vec{X}) \rightarrow (\Theta(x), \Theta'(x) \vec{X})$ из $\mathcal{O} \times \vec{E}$ в $\mathbb{K}^n \times \mathbb{K}^n$. Это отображение непрерывно также и в силу теоремы 10.

Векторное пространство, касательное к абстрактному многообразию в точке¹⁾

Пусть V — абстрактное многообразие размерности n класса C^1 и $a \in V$. Как определить касательный вектор к V в точке a ? Пусть $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ — карта некоторого открытого множества $\Phi(\mathcal{O})$ из V , содержащего точку a , и пусть $\alpha \in \mathbb{K}^n$ — такая точка, что $\Phi(\alpha) = a$. Интуитивно кажется ясным, что если $\vec{\xi}$ — произвольный вектор \mathbb{K}^n , то можно с помощью отображения Φ поставить ему в соответствие вектор \vec{X} , касательный к многообразию V в точке a . Пусть $\Phi_1: \mathcal{O}_1 \rightarrow \Phi_1(\mathcal{O}_1)$ и $\Phi_2: \mathcal{O}_2 \rightarrow \Phi_2(\mathcal{O}_2)$ — две карты открытых множеств многообразия V , содержащих точку a . Пусть $\vec{\xi}_1$ и $\vec{\xi}_2$ — два вектора из \mathbb{K}^n . При каком условии им будет соответствовать при отображениях Φ_1 и Φ_2 один и тот же вектор \vec{X} , касательный к V в точке a ? Пусть Ω — пересечение $\Phi_1(\mathcal{O}_1) \cap \Phi_2(\mathcal{O}_2)$, и пусть $\Omega_1 = \Phi_1^{-1}(\Omega)$ и $\Omega_2 = \Phi_2^{-1}(\Omega)$. Тогда $\Phi_2^{-1} \circ \Phi_1$ будет C^m -дiffeоморфизмом Ω_1 на Ω_2 . В точке

¹⁾ См. примечание на стр. 334.

$\alpha_1 \in \Omega_1$ он имеет производное отображение $(\Phi_2^{-1} \circ \Phi_1)'(\alpha_1)$. Мы условимся, что векторы $\vec{\xi}_1$ и $\vec{\xi}_2$ представляют при отображениях Φ_1 и Φ_2 один и тот же вектор \vec{X} , касательный к многообразию V в точке a , если имеет место соотношение

$$\vec{\xi}_2 = (\Phi_2^{-1} \circ \Phi_1)'(\alpha_1) \vec{\xi}_1. \quad (\text{III}, 9; 24_2)$$

Такое определение оправдывается следствием 3 теоремы 334.

Мы приходим к следующему определению: *касательным вектором к многообразию V в точке a* называется такое семейство $(\vec{\xi}_i)_{i \in I}$ векторов из \mathbb{K}^n (где I — множество индексов всех возможных карт Φ_i многообразия V , для которых $\Phi_i(\mathcal{O}_i)$ содержит a), что выполняются соотношения $(\Phi_i^{-1} \circ \Phi_j)'(a_i) \vec{\xi}_i = \vec{\xi}_j$, $i, j \in I$. Однозначно определяется сумма двух касательных векторов $(\vec{\xi}_i)_{i \in I}$, $(\vec{\eta}_i)_{i \in I}$ по формуле $(\vec{\xi}_i + \vec{\eta}_i)_{i \in I}$. Точно так же определяется произведение касательного вектора $(\vec{\xi}_i)_{i \in I}$ на $\lambda \in \mathbb{K}$: $(\lambda \vec{\xi}_i)_{i \in I}$. Векторы, касательные к многообразию V в точке a , образуют векторное пространство, имеющее размерность n , т. е. размерность самого многообразия V . Через $\vec{T}(a; V)$ обозначается векторное пространство, касательное к многообразию V в точке a . Следует заметить, что если a и b — две различные точки V , то между касательными векторными пространствами $\vec{T}(a; V)$ и $\vec{T}(b; V)$ нет никакой простой связи. Эти векторные пространства не являются подпространствами одного и того же заранее заданного векторного пространства, как в случае многообразия в аффинном пространстве.

Пусть V и W — два многообразия класса C^1 размерностей соответственно p и q . Пусть H — отображение V в W класса C^1 и $b = H(a)$. Примем обозначения, введенные в определении на стр. 336: Φ является картой некоторой окрестности точки a в V , $\Phi(\alpha) = a$, а Ψ является картой окрестности b в W . Пусть $\vec{\xi}$ — некоторый вектор пространства \mathbb{K}^p , определяющий с помощью отображения Φ вектор \vec{X} , касательный к V в точке a . Его образ при производном отображении $(\Psi^{-1} \circ H \circ \Phi)'(\alpha)$ является некоторым вектором $\vec{\eta}$ из \mathbb{K}^q , определяющим с помощью отображения Ψ вектор \vec{Y} , касательный к W в точке b . Легко доказать, что \vec{Y} зависит только от \vec{X} и a и не зависит от выбранных карт Φ и Ψ . Можно сказать, что задание H определяет линейное отображение $\vec{T}(a; V)$ в $\vec{T}(b; W)$. Это отображение назы-

вают производным отображением отображения H в точке a и обозначают через $H'(a)$. Это производное отображение является обобщением рассмотренных выше производных отображений для отображений класса C^1 аффинных пространств. В самом деле, если V и W — некоторые открытые множества пространств \mathbb{K}^p и \mathbb{K}^q , то в качестве Φ и Ψ можно взять тождественные отображения. Векторное пространство, касательное к V в точке a , тогда совпадает с \mathbb{K}^p , а векторное пространство, касательное к W в точке a , совпадает с \mathbb{K}^q ; при этом $\vec{X} = \vec{\xi}$ и $\vec{Y} = \vec{\eta}$. Определенное выше производное отображение является отображением $H'(a)$ из \mathbb{K}^p в \mathbb{K}^q . Без труда проверяется, что так определенное обобщенное производное отображение удовлетворяет теореме о сложной функции (теорема 11). Впрочем, это определение отображения $H'(a) \in \mathcal{L}(\vec{T}(a; V); \vec{T}(b; W))$ является естественным результатом следствия 1₂ теоремы 33₄.

С другой стороны, если V есть многообразие, содержащееся в аффинном пространстве E , то в качестве отображения H можно взять инъекцию V в E . Производное отображение $H'(a)$ определяет тогда линейное отображение $\vec{T}(a; V)$ в $\vec{T}(a; E) = \vec{E}$. Можно проверить, что это линейное отображение является инъекцией. Оно дает возможность отождествить $\vec{T}(a; V)$ с его образом, т. е. с некоторым векторным подпространством из \vec{E} , и это подпространство есть найденное нами ранее векторное подпространство, касательное к многообразию V в точке a , лежащему в аффинном пространстве E .

Мы не настаиваем на том, чтобы читатель сразу усвоил это довольно сложное понятие векторного пространства, касательного к абстрактному многообразию. Оно приведено здесь лишь для иллюстрации того, что построения, проведенные нами ранее в случае аффинных пространств, допускают обобщение на произвольные дифференцируемые многообразия.

Теорема о постоянном ранге

Теорема 33₅. Пусть \mathcal{O} — открытое множество пространства \mathbb{K}^n , и пусть $\Phi: u \rightarrow u = \Phi(u)$ — такое отображение класса C^m множества \mathcal{O} в аффинное пространство E , что в каждой точке $\alpha \in \mathcal{O}$ производное отображение $\Phi'(\alpha)$ имеет один и тот же ранг $l \leq n$. Тогда для каждой точки $\alpha \in \mathcal{O}$ существует открытое множество \mathcal{A} , содержащееся в \mathcal{O} и содержащее α , образ которого при отображении Φ является многообразием $V \subset E$ размерности l , принадлежащим классу C^m .

Доказательство. Выберем в E некоторую систему координат. Сказать, что ранг $\Phi'(\alpha)$ равен l , означает прежде

всего, что по крайней мере один из определителей с l строками и l столбцами производной матрицы отображения Φ в точке α отличен от нуля. Изменяя при необходимости порядок векторов базиса E , мы можем предположить, что этот определитель является якобианом $\frac{D(x_1, x_2, \dots, x_l)}{D(u_1, u_2, \dots, u_l)}(a)$.

Так как Φ принадлежит классу C^1 якобиан $\frac{D(x_1, x_2, \dots, x_l)}{D(u_1, u_2, \dots, u_l)}$ является непрерывной скалярной функцией на \mathcal{O} и существует окрестность \mathcal{A}_0 точки a , в которой этот определитель не обращается в нуль. Поскольку в \mathcal{O} и, следовательно, в \mathcal{A}_0 ранг производного отображения всюду равен l , в \mathcal{A}_0 каждый минор производной матрицы, содержащий предыдущий определитель, равен нулю. Разложим \mathbb{K}^n в произведение $\mathbb{K}^l \times \mathbb{K}^{n-l}$. Точку $u \in \mathbb{K}^n$ мы будем обозначать через (v, w) , где $v = (u_1, \dots, u_l)$ и $w = (u_{l+1}, \dots, u_n)$. Пусть $\alpha = (\beta, \gamma)$. Тем же самым способом разложим E в его системе координат в произведение $\mathbb{K}^l \times \mathbb{K}^{n-l}$ и положим $x = (y, z)$, где $y = (x_1, x_2, \dots, x_l)$ и $z = (x_{l+1}, \dots, x_N)$. Пусть $a = \Phi(\alpha) = (b, c)$, и пусть $y = X(v, w)$, $z = \Psi(v, w)$ определяют $x = \Phi(u)$ при этом разложении. Согласно предположению, сделанному относительно якобиана, из теоремы о неявной функции (теоремы 28 и 31)¹⁾ следует, что существуют открытые множества \mathcal{A}' и \mathcal{A}'' в пространствах \mathbb{K}^l и \mathbb{K}^{n-l} , содержащие соответственно β и γ , $\mathcal{A}' \times \mathcal{A}'' \subset \mathcal{A}_0$, и открытое множество \mathcal{B} в пространстве \mathbb{K}^l , содержащее точку b , такие, что для произвольно заданных точек w и y в \mathcal{A}'' и \mathcal{B} существует единственный элемент $v \in \mathcal{A}'$, для которого $X(v, w) = y$. Кроме того, таким образом определенная функция $v = \Lambda(y, w)$, действующая из $\mathcal{B} \times \mathcal{A}''$ в \mathcal{A}' , принадлежит классу C^m . Поскольку, в силу непрерывности Λ , всегда можно, не изменяя \mathcal{A}' , заменить \mathcal{B} и \mathcal{A}'' меньшими окрестностями, мы можем взять в качестве \mathcal{A}'' некоторый шар с центром в точке γ . Обозначим через \mathcal{A}_1 произведение $\mathcal{A}' \times \mathcal{A}''$, а через W — его образ при отображении Φ . Если $x = (y, z)$ принадлежит W и если, кроме того, y принадлежит \mathcal{B} , то $v = \Lambda(y, w)$ и, следовательно, $z = \Psi(\Lambda(y, w), w) = g(y, w)$. Покажем, что в действительности g не зависит от w . Для этого вычислим дифференциал dg в некоторой точке (y, w) .

Для выполнения этих вычислений надо начинать с дифференцирования отображения X по формуле

$$\vec{dy} = \frac{\partial X}{\partial v}(v, w) \vec{dv} + \frac{\partial X}{\partial w}(v, w) \vec{dw}. \quad (\text{III}, 9; 25)$$

¹⁾ Здесь решается уравнение $X(v, w) = y$ относительно v при заданных w и y .

Тот факт, что якобиан $\frac{D(x_1, x_2, \dots, x_l)}{D(u_1, u_2, \dots, u_l)} \neq 0$, означает, что частная производная $\frac{\partial X}{\partial v}(v, w) \in \mathcal{L}(\mathbb{K}^l; \mathbb{K}^l)$ обратима. Предыдущее уравнение можно разрешить относительно \vec{dv} :

$$\begin{aligned}\vec{dv} &= \left(\frac{\partial X}{\partial v}(v, w) \right)^{-1} \vec{dy} - \left(\left(\frac{\partial X}{\partial v}(v, w) \right)^{-1} \cdot \frac{\partial X}{\partial w}(v, w) \right) \vec{dw} = \\ &= \frac{\partial \Lambda}{\partial y}(y, w) \vec{dy} + \frac{\partial \Lambda}{\partial w}(y, w) \vec{dw}. \quad (\text{III}, 9; 26)\end{aligned}$$

Подставляя найденный результат в дифференциал Ψ :

$$\vec{dz} = \frac{\partial \Psi}{\partial v}(v, w) \cdot \vec{dv} + \frac{\partial \Psi}{\partial w}(v, w) \cdot \vec{dw}, \quad (\text{III}, 9; 27)$$

получаем искомый дифференциал. Тот факт, что все миноры ранга $>l$ производной матрицы отображения Φ равны нулю в $\mathcal{A}' \times \mathcal{A}''$, означает, что dx_{l+k} , $k = 1, 2, \dots, N-l$, зависят только от dx_1, dx_2, \dots, dx_l и не зависят от du_{l+1}, \dots, du_n . Другими словами, в развернутом выражении

$$\vec{dz} = \frac{\partial \Psi}{\partial v} \cdot \left[\left(\frac{\partial X}{\partial v} \right)^{-1} \cdot \vec{dy} - \left(\frac{\partial X}{\partial v} \right)^{-1} \cdot \frac{\partial X}{\partial w} \cdot \vec{dw} \right] + \frac{\partial \Psi}{\partial w} \cdot \vec{dw} \quad (\text{III}, 9; 28)$$

величина \vec{dz} в действительности зависит только от \vec{dy} и не зависит от \vec{dw} . Иначе говоря, функция g , определенная на $\mathcal{B} \times \mathcal{A}''$, со значениями в \mathcal{A}' имеет тождественно равную нулю частную производную dg/dw . Поскольку в качестве \mathcal{A}'' мы выбрали шар, т. е. связное множество, для фиксированного y можно к функции $w \rightarrow g(y, w)$, определенной на \mathcal{A}'' , применить теорему 22 и убедиться в том, что при фиксированном y величина $g(y, w)$ постоянна, т. е. что g не зависит от w . Это — некоторая функция только переменной y , определенная на \mathcal{B} . Отсюда вытекает, что множество точек (y, z) , принадлежащих W , для которых y лежит в \mathcal{B} , является многообразием V размерности l класса C^m , определенным уравнением $z = g(y)$. Это множество V точек E является образом при отображении Φ пересечения \mathcal{A}_1 и множества $X^{-1}(\mathcal{B})$ точек u , таких, что $X(u) \in \mathcal{B}$, т. е. образом пересечения двух открытых множеств и, следовательно, образом некоторого открытого множества \mathcal{A} из \mathcal{A}_1 , содержащего α . В силу сказанного, образ $\Phi(\mathcal{A})$ является многообразием V размерности l класса C^m , и теорема доказана.

Замечания. 1°) Если даже отображение Φ является биекцией и $l = n$, будет ошибкой считать, что образ при отображении Φ всего открытого множества \mathcal{O} окажется многообразием в E .

Рассмотрим, например, на плоскости \mathbb{R}^2 лемнискату Бернули. Выберем на этой лемнискате направление обхода 1, 2, 3, 4, указанное на рис. 6 стрелками. Всегда можно найти такую биекцию $\Phi: u \rightarrow \Phi(u)$ класса C^∞ вещественной прямой \mathbb{R} в плоскость \mathbb{R}^2 , чтобы $\Phi(\mathbb{R})$ была в точности этой лемнискатой, $\Phi(0)$ — двойной точкой лемнискаты и чтобы при u , стремящемся к $-\infty$ (соответственно $+\infty$), $\Phi(u)$ стремилась к двойной точке по ветви 1 (соответственно 4):

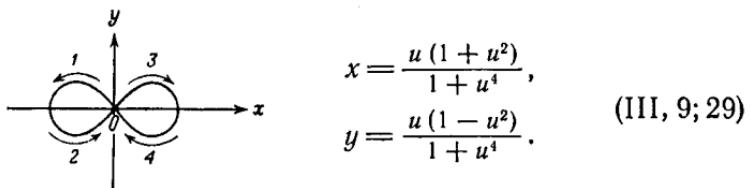


Рис. 6.

Производная $\Phi'(\alpha)$ имеет ранг, равный 1, в каждой точке $\alpha \in \mathbb{R}$. Другими словами, $x'(u)$ и $y'(u)$ одновременно в нуль не обращаются. Кроме того, Φ является биекцией¹⁾. Тем не менее Φ не является, очевидно, гомеоморфизмом, ибо можно найти последовательность точек лемнискаты, сходящуюся к двойной точке, для которой последовательность u , вместо того чтобы сходиться к 0, сходится к $+\infty$ или $-\infty$. Впрочем, это видно и из того, что лемниската является компактным множеством, в то время как \mathbb{R} этим свойством не обладает. Итак, образ $\Phi(\mathbb{R})$ не является многообразием, так как он имеет особую точку в начале координат. Однако для точки $\alpha = 0$ можно найти такую окрестность, например интервал $]-A, A[$, где A — произвольное положительное число, что образ этой окрестности при отображении Φ является многообразием.

2°) Предположим, что $n \geq N$ и $l = N$. Предположение о том, что $\Phi'(\alpha)$ для любого $\alpha \in \mathcal{O}$ имеет ранг, равный $l = N$, означает просто, что $\Phi'(\alpha)$ является сюръекцией \mathbb{K}^n на \vec{E} . Но тогда $\Phi(\mathcal{A})$, как многообразие размерности N аффинного пространства размерности N , есть открытое множество в \vec{E} . Поэтому если Ω — некоторое открытое множество из \mathcal{O} , то $\Phi(\Omega)$ будет окрестностью каждой из своих точек, т. е. открытым множеством, и мы снова убеждаемся, что Φ — открытое отображение (теорема 30). Отсюда вытекает также первая часть теоремы 30. Если известно лишь, что $\Phi'(\alpha_0)$ имеет ранг, равный N , в некоторой точке α_0 , то имеется по крайней мере один минор ранга N производной матрицы (в некоторой системе координат пространства E), отличный от нуля в точке α_0 . В силу его

¹⁾ Двойная точка получается лишь при $u = 0$.

непрерывности, он отличен от нуля во всех точках a некоторого открытого множества $\mathcal{O}_1 \subset \mathcal{O}$, содержащего α_0 . Но Φ является открытым отображением множества \mathcal{O}_1 в E , и, следовательно, образ любой окрестности точки α_0 является окрестностью точки $a_0 = \Phi(\alpha_0)$ в E .

3°) В случае $l < n$ отображение Φ , естественно, не инъективно и каждая из точек многообразия $V = \Phi(\mathcal{A})$ является образом бесконечного множества точек из \mathcal{A} . В самом деле, если $x_0 = (y_0, z_0)$ — некоторая точка V , то можно произвольно выбрать точку w в \mathcal{A}'' , затем вычислить $v = \Lambda(y_0, w)$, и тогда рассматриваемая точка x_0 будет образом точки $u = (v, w) = (\Lambda(y_0, w), w)$. Если рассматривать сужение отображения Φ на \mathcal{A} , то прообразом x_0 будет множество точек (v, w) из \mathcal{A} , для которых $w \in \mathcal{A}''$ и $v = \Lambda(y_0, w)$. Это множество является многообразием класса C^m размерности $n - l$ в пространстве \mathbb{K}^n , определенным предыдущим явным соотношением. Напротив, если $l = n$, то Φ — инъекция (для достаточно малого \mathcal{A}).

4°) Из последней теоремы снова вытекает теорема 32. В самом деле, если для каждой точки $a \in V$ существует параметрическое представление $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$, $\Phi(a) = a$, всюду имеющее ранг $l = n$, то из теоремы 33₅ следует существование такой открытой окрестности \mathcal{A} точки a в \mathcal{O} , что $\Phi(\mathcal{A}) = W$ является многообразием. Но Φ , кроме того, является гомеоморфизмом \mathcal{O} на открытое множество в V . Поэтому $\Phi(\mathcal{A})$ является открытым множеством в V , и значит, существует такая открытая окрестность \mathcal{V}_1 точки a в E , что $W = V \cap \mathcal{V}_1$. Но если для каждой точки $a \in V$ найдется такая открытая окрестность \mathcal{V}_1 в E , что $V \cap \mathcal{V}_1$ — многообразие, то по определению V само является многообразием.

5°) Если ранг отображения Φ не постоянен, то *не существует никакой теоремы*, позволяющей утверждать, что образ окрестности точки α при отображении Φ имеет простую структуру. Рассмотрим, например, отображение Φ плоскости \mathbb{R}^2 в пространство \mathbb{R}^3 , определяемое формулами $(r, \varphi) \rightarrow (x, y, z)$, где

$$\begin{aligned} x &= r \cos \varphi, \\ y &= r \sin \varphi, \\ z &= r. \end{aligned} \tag{III, 9; 30}$$

Так как в точке $\alpha = (0, \varphi_0)$ производная $\partial\Phi/\partial\varphi$ равна нулю, то ранг производного отображения равен 1. Однако в близких точках этот ранг равен 1 или 2. Образом \mathbb{R}^2 при отображении Φ является конус в \mathbb{R}^3 , определяемый уравнением $x^2 + y^2 - z^2 = 0$, для которого отображение Φ не определяет параметрического представления. Образ каждой окрестности рассматриваемой

точки α при отображении Φ является окрестностью вершины конуса, которая не является многообразием из-за наличия особой точки в вершине.

Сформулируем по этому поводу следующее свойство:

Если Φ является отображением класса C^1 открытого множества \mathcal{O} пространства \mathbb{K}^n в E , то ранг $l(\alpha)$ производного отображения $\Phi'(\alpha)$ является полунепрерывной снизу функцией α .

В самом деле, покажем, что если $\alpha_0 \in \mathcal{O}$, то существует такое открытое множество \mathcal{A}_0 из \mathcal{O} , содержащее α_0 , что $l(\alpha) \geq l(\alpha_0)$ для каждого $\alpha \in \mathcal{A}_0$. Действительно, выберем в E некоторую систему координат. Согласно определению ранга, существует по крайней мере один минор порядка $l(\alpha_0)$ производной матрицы отображения Φ в точке α_0 , отличный от нуля. Однако, поскольку Φ принадлежит классу C^1 , этот минор является непрерывной функцией α , а, следовательно, существует окрестность \mathcal{A}_0 точки α_0 , во всех точках которой этот минор $\neq 0$. Значит, в каждой точке $\alpha \in \mathcal{A}_0$ ранг $\Phi'(\alpha) \geq l(\alpha_0)$.

Зависимые и независимые функции

Пусть $\Phi_1, \Phi_2, \dots, \Phi_N$ суть N скалярных функций класса C^1 от n переменных u_1, u_2, \dots, u_n . Они определяют отображение $\Phi: u \rightarrow x = \Phi(u)$ класса C^1 открытого множества $\mathcal{O} \subset \mathbb{K}^n$ в \mathbb{K}^N .

Посмотрим, в каких случаях эти функции являются зависимыми или независимыми в окрестности точки $\alpha \in \mathcal{O}$. Пусть $a = \Phi(\alpha)$.

1-й случай. Независимость. Предположим, что ранг отображения $\Phi'(\alpha)$ равен N — числу функций. Другими словами, один из миноров ранга N производной матрицы, состоящей из элементов $\frac{\partial \Phi_i}{\partial u_j}(\alpha)$, $i = 1, 2, \dots, N$, $j = 1, 2, \dots, n$, отличен от нуля. Для этого, очевидно, необходимо, чтобы $n \geq N$, т. е. чтобы число переменных было не меньше числа функций. Тогда теорема 30 (или 33₅) говорит нам о том, что образ всякой окрестности α при отображении Φ является окрестностью точки a . Функции $\Phi_1(u), \Phi_2(u), \dots, \Phi_N(u)$ могут принимать произвольные значения, лишь бы только эти значения были достаточно близкими к a_1, a_2, \dots, a_n . Следовательно, если между функциями Φ_i имеется соотношение вида

$$R(\Phi_1(u_1, u_2, \dots, u_n), \Phi_2(u_1, u_2, \dots, u_n), \dots, \Phi_N(u_1, u_2, \dots, u_n)) \equiv 0, \quad (\text{III}, 9; 31)$$

где R — некоторая функция N переменных, то R должна быть тождественным нулем в окрестности точки (a_1, a_2, \dots, a_n) , т. е. по существу никакого соотношения нет.

В этом случае естественно говорить, что N функций Φ_i независимы в окрестности точки α .

2-й случай. Зависимость. Предположим теперь, что $\Phi'(\alpha)$ для каждой точки α имеет один и тот же ранг $l < N$. Тогда из теоремы 33₅ следует, что существует окрестность точки α , образ которой при отображении Φ является многообразием V размерности l класса C^1 пространства \mathbb{K}^N . Изменяя при необходимости порядок заданных функций Φ_i , можно прийти к случаю, когда это многообразие задается уравнениями вида

$$x_{l+k} = G_k(x_1, x_2, \dots, x_l), \quad k = 1, 2, \dots, N-l. \quad (\text{III}, 9; 32)$$

Тогда в окрестности α функции Φ_i удовлетворяют $N - l$ нетривиальным соотношениям

$$\begin{aligned} \Phi_{l+k}(u_1, \dots, u_n) &\equiv \\ &\equiv G_k(\Phi_1(u_1, \dots, u_n), \Phi_2(u_1, \dots, u_n), \dots, \Phi_l(u_1, \dots, u_n)), \\ &\quad k = 1, 2, \dots, N-l. \end{aligned} \quad (\text{III}, 9; 33)$$

На этот раз это настоящие соотношения, поскольку при заданных значениях $\Phi_1, \Phi_2, \dots, \Phi_l$ они определяют функции Φ_{l+k} . Последние $N - l$ соотношений независимы, так как функции $x \rightarrow x_{l+k} = G_k(x_1, x_2, \dots, x_l)$ независимы в смысле 1-го случая.

В самом деле, их дифференциалы $dx_{l+k} - dG_k$ таковы, что dx_{l+k} входит с ненулевым коэффициентом лишь в k -й из них. Следовательно, для фиксированного x не существует линейного соотношения между ними с постоянными коэффициентами. Ранг системы этих $N - l$ дифференциалов равен $N - l$ для любого x , близкого к точке a . Естественно говорить в этом случае, что функции $\Phi_1, \Phi_2, \dots, \Phi_N$ зависимы и удовлетворяют $N - l$ независимым соотношениям.

Существуют, конечно, и другие, не рассматриваемые нами случаи, когда ранг производной не постоянен в окрестности точки α . В этих случаях результаты будут гораздо более сложные, чем те, которые мы получили выше. Поэтому утверждение такого рода: N функций $\Phi_1, \Phi_2, \dots, \Phi_N$ от n переменных u_1, u_2, \dots, u_n с непрерывными частными производными первого порядка независимы тогда и только тогда, когда производная матрица функций Φ_i имеет ранг N , и зависимы тогда и только тогда, когда производная матрица имеет ранг $< N$, — не является точным и в конечном счете неверно.

Рассмотрим, например, представление (III, 9; 30) конуса $x^2 + y^2 - z^2 = 0$. Три функции F, G, H двух вещественных переменных r, φ удовлетворяют соотношению $F^2 + G^2 - H^2 = 0$. Но в окрестности точки $(0, \varphi_0)$, представляющей начало координат $x=y=z=0$, из этого соотношения невозможно выразить

одну из этих трех функций как непрерывно дифференцируемую функцию двух других, как это делалось во 2-м случае.

Можно столкнуться со значительно более сложными фигурами, чем конус, для которых невозможно придать никакого «утилитарного» смысла понятию зависимости или независимости. Поэтому разумно ограничиться только рассмотренными выше случаями.

Особые, или параметрические, многообразия

Вводимое здесь новое понятие является обобщением понятия пути, введенного в гл. II на стр. 90.

Особым, или параметрическим, многообразием размерности n класса C^m в аффинном пространстве E размерности N называется отображение Φ класса C^m некоторого многообразия V размерности n класса C^m (абстрактного или погруженного в аффинное пространство) в пространство E . Образ $\Phi(V) = W$ называется *образом параметрического многообразия*. Не следует путать его с самим параметрическим многообразием, которое является отображением Φ многообразия V в W . Например, если Φ постоянно: $\Phi(V) = \{a\}$, $a \in E$, то образом является точка, но параметрическое многообразие — это постоянное отображение Φ . Лемниската Бернулли в \mathbb{R}^2 не является многообразием, но она есть образ некоторого параметрического многообразия при $V = \mathbb{R}$ (см. формулу (III, 9; 29)); в этом примере отображение Φ инъективно. Она является также образом параметрического многообразия, для которого V — окружность, причем можно построить параметрическое многообразие так, что, когда V пробегается в некотором заданном направлении, лемниската будет пробегаться в направлении 1, 2, 3, 4, указан-

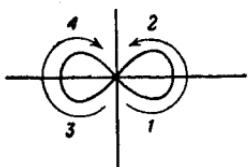


Рис. 7.

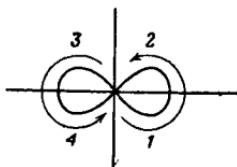


Рис. 8.

ном на рис. 7, а можно так, что она будет пробегаться в направлении 1, 2, 3, 4 на рис. 8. Наконец, она является образом параметрического многообразия, для которого V представляет собой систему двух отдельных окружностей, которые при отображении Φ переходят в две петли лемнискаты. Все это — различные параметрические многообразия.

Если вместо V взять замкнутый отрезок $[a, b]$ из \mathbb{R} , то отображение Φ класса C^m отрезка $[a, b]$ в E называется *кривой класса C^m из E* с началом $\Phi(a)$ и концом $\Phi(b)$. Если $\Phi(a) = \Phi(b)$, то говорят, что это *замкнутая кривая* (хотя это никак не связано с топологическим понятием замкнутого множества).

Два параметрических многообразия $\Phi_1: V_1 \rightarrow E$, $\Phi_2: V_2 \rightarrow E$ называют *C^m -эквивалентными*, если существует C^m -дiffeоморфизм H многообразия V_1 на V_2 , такой, что $\Phi_1 = \Phi_2 \circ H$ (тогда $\Phi_2 = \Phi_1 \circ H^{-1}$).

Такие многообразия имеют одни и те же образы. Это соответствует тому, что называется в математике *изменением параметрического представления* кривой или поверхности. Рассмотренные выше различные параметрические представления лемнискаты не являются эквивалентными параметрическими многообразиями.

§ 10. УСЛОВНЫЕ МАКСИМУМЫ И МИНИМУМЫ

Пусть Ω — открытое множество аффинного нормированного пространства E над полем \mathbb{R} , и пусть f, g_1, g_2, \dots, g_m суть $m+1$ непрерывно дифференцируемых вещественных функций на Ω . Займемся изучением экстремумов, т. е. максимумов или минимумов, функций f на подмножестве $A \subset \Omega$, определенном уравнениями

$$g_1(x) = 0, g_2(x) = 0, \dots, g_m(x) = 0. \quad (\text{III}, 10; 1)$$

Мы будем искать, следовательно, максимумы или минимумы функции f не на Ω , а только на некотором замкнутом, а не открытом подмножестве множества Ω . В этом случае говорят об *условном максимуме* или *условном минимуме*.

Теорема 34. Пусть f, g_1, g_2, \dots, g_m — вещественные непрерывно дифференцируемые функции, определенные на открытом множестве Ω аффинного нормированного пространства E . Пусть a — некоторая точка Ω , удовлетворяющая уравнениям $g_1(a) = g_2(a) = \dots = g_m(a) = 0$. Если линейные дифференциальные формы $g'_1(a), g'_2(a), \dots, g'_m(a) \in \overset{\leftarrow}{E'}$ независимы, то необходимым условием того, что точка a является точкой *условного максимума* или *минимума* функции f на подмножестве A множества Ω , определенном уравнениями $g_1(x) = g_2(x) = \dots = g_m(x) = 0$, является наличие таких m вещественных постоянных $\lambda_1, \lambda_2, \dots, \lambda_m$, при которых имеет место соотношение

$$f'(a) = \lambda_1 g'_1(a) + \lambda_2 g'_2(a) + \dots + \lambda_m g'_m(a). \quad (\text{III}, 10; 2)$$

Величины λ_i называются *множителями Лагранжа*, соответствующими *экстремальной точке a* .

Перед доказательством этой теоремы укажем ее геометрический смысл. В открытом множестве Ω уравнения $g_i(x) = 0$ определяют в окрестности точки a некоторое дифференцируемое многообразие. Уравнения $\overrightarrow{g'_i}(a)(x - a) = 0$, $i = 1, 2, \dots, m$, являются уравнениями касательного линейного многообразия. Соотношения (III, 10; 2) выражают тот факт, что гиперплоскость $\overrightarrow{f'(a)} \cdot (x - a)$, касательная к гиперповерхности $f(x) = f(a)$ в точке a , содержит это линейное многообразие.

Доказательство. Проведем сначала доказательство для конечномерного пространства E . В этом случае можно выбрать некоторую систему координат $0, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ и выразить точку $x \in E$ в этой системе через ее координаты x_1, x_2, \dots, x_n .

Предположение, что в точке a дифференциалы dg_i являются независимыми линейными формами, означает, что существует m координат, например x_1, x_2, \dots, x_m , таких, что якобиан $\frac{D(g_1, g_2, \dots, g_m)}{D(x_1, x_2, \dots, x_m)}(a) \neq 0$. Для этого, очевидно, необходимо, чтобы $n \geq m$. Тогда в окрестности этой точки a , согласно теореме 28, переменные x_1, x_2, \dots, x_m можно выразить как функции от $g_1, g_2, \dots, g_m, x_{m+1}, \dots, x_n$. Найденные значения можно подставить в f и получить новую непрерывно дифференцируемую функцию $F(g_1, g_2, \dots, g_m, x_{m+1}, \dots, x_n)$.

Запишем теперь, что функция F имеет экстремум в точке a при условии, что x_{m+1}, \dots, x_n изменяются, в то время как g_1, g_2, \dots, g_m остаются нулями. Необходимым условием этого является равенство нулю частных производных функции F по x_{m+1}, \dots, x_n в точке a . Это означает, что дифференциал dF в точке a имеет вид

$$dF = \lambda_1 dg_1 + \lambda_2 dg_2 + \dots + \lambda_m dg_m, \quad (\text{III, 10; 3})$$

где $\lambda_1, \lambda_2, \dots, \lambda_m$ — вещественные постоянные. Поскольку дифференциал df , выраженный через dx_1, dx_2, \dots, dx_n , получается заменой в дифференциале dF , записанном относительно переменных $dg_1, dg_2, \dots, dg_m, dx_{m+1}, \dots, dx_n$, дифференциалов dg_i их выражениями через dx_1, dx_2, \dots, dx_n , этот дифференциал df удовлетворяет соотношению

$$df = \lambda_1 dg_1 + \lambda_2 dg_2 + \dots + \lambda_m dg_m, \quad (\text{III, 10; 4})$$

что приводит к (III, 10; 2), и теорема доказана.

Предположим теперь, что пространство E бесконечномерно. Поскольку $g'_i(a)$ по предположению независимы, они могут принимать произвольные заданные значения. Для каждого i можно найти такой вектор $\vec{X}_i \in \vec{E}$, что $g'_i(a) \cdot \vec{X}_i = \delta_{ij} = 0$ для $j \neq i$ и

1 для $j = i$. Построенные так векторы $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m$, очевидно, независимы, ибо если положить $\sum_{i=1}^m \lambda_i \vec{X}_i = \vec{0}$, то мы получим $g'_j(a) \cdot \sum_{i=1}^m \lambda_i \vec{X}_i = \lambda_j = 0$. Положим теперь $\lambda_i = f'(a) \cdot \vec{X}_i$. Линейная форма $f'(a) - \sum_{i=1}^m \lambda_i g'_i(a)$ равна нулю на векторах $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m$, а, следовательно, на порождаемом ими векторном пространстве. Пусть \vec{X} — произвольный вектор пространства \vec{E} . Обозначим через \vec{E}_0 подпространство пространства \vec{E} , порожденное векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m, \vec{X}$. На аффинном пространстве E_0 , проходящем через a параллельно \vec{E}_0 , функция f должна иметь экстремум в точке a на множестве $A_0 = A \cap E_0$, определенном уравнениями $g_i(x) = 0$. Однако пространство E_0 имеет конечную размерность m или $m + 1$. Следовательно, существуют такие постоянные $\mu_1, \mu_2, \dots, \mu_m$, что производная $f'(a) - \sum_{i=1}^m \mu_i g'_i(a)$ равна нулю на \vec{E}_0 . Если записать, что она обращается в нуль на $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m$, то мы получим, что $\mu_i = \lambda_i$. Отсюда следует, что $f'(a) - \sum_{i=1}^m \lambda_i g'_i(a)$ равна нулю на произвольном векторе \vec{X} пространства \vec{E} , а, значит, обращается в нуль на всем \vec{E} , чем и заканчивается доказательство теоремы.

Из доказательства вытекает, что множители Лагранжа λ_i определяются единственным образом, что, впрочем, ясно было и заранее, поскольку $g'_i(a)$ предполагались независимыми.

Замечание. Пусть g — некоторая функция класса C^1 , определенная на $\Omega \subset E$, со значениями в t -мерном аффинном пространстве G , такая, что ее производная $g'(a) \in \mathcal{L}(\vec{E}; \vec{G})$ имеет ранг t . Для того чтобы f имела в точке a экстремум на подмножестве A множества Ω , определенном уравнением $g(x) = g(a)$, необходимо, чтобы существовал такой элемент λ из $\vec{G}' = \mathcal{L}(\vec{G}; \mathbb{R})$, при котором имеет место равенство

$$f'(a) = \lambda \circ g'(a) \in \vec{E}' = \mathcal{L}(\vec{E}; \mathbb{R}). \quad (\text{III}, 10; 4_2)$$

В самом деле, выберем в G некоторую систему координат с началом в точке $g(a)$. Таким путем G отождествляется с \mathbb{R}^m . Теперь составляющие функции g — это t скалярных функций

$g_i, i = 1, 2, \dots, m$, и множество A определяется уравнениями $g_i(x) = 0$. То, что $g'(a) \in \mathcal{L}(\vec{E}; \vec{G})$ имеет ранг m , означает, что $g'_i(a) \in \mathcal{L}(\vec{E}; \mathbb{R})$ независимы. Мы находимся в условиях применимости теоремы. Существование таких множителей Лагранжа $\lambda_i, i = 1, 2, \dots, m$, при которых имеет место соотношение (III, 10; 2), эквивалентно существованию такого элемента $\lambda \in \vec{G}'$, при котором имеет место соотношение (III, 10; 4₂). Однако результат, записанный в виде (III, 10; 4₂), вовсе не зависит от выбора системы координат в G . Заметим, что в (III, 10; 4₂) производная $g''(a)$ является элементом пространства $\mathcal{L}(\vec{G}; \mathbb{R})$, а λ есть некоторый элемент пространства $\mathcal{L}(\vec{G}; \mathbb{R})$, и, следовательно, $\lambda \circ g'(a)$ является элементом пространства $\mathcal{L}(\vec{E}; \mathbb{R})$.

Практический способ вычисления условного максимума или минимума

Предположим для определенности, что E имеет размерность n , что в нем выбрана некоторая система координат и что каждая точка $x \in E$ задана своими координатами x_1, x_2, \dots, x_n . Будем искать точку $a = (a_1, a_2, \dots, a_n)$ условного экстремума функции f при условии $g_i(x) = 0, i = 1, 2, \dots, m$. Прежде всего должно удовлетворяться уравнение (III, 10; 2), эквивалентное системе n уравнений

$$\frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x_j}(a), \quad j = 1, 2, \dots, n. \quad (\text{III}, 10; 5)$$

Если сюда добавить систему m уравнений (III, 10; 1) (где x должно быть заменено на a), то получится система $m + n$ уравнений относительно $m + n$ неизвестных, а именно, относительно координат a_j точки a и множителей Лагранжа λ_i . Нет необходимости полностью решать эту систему. Можно при желании исключить множители Лагранжа λ_i и свести данную систему к системе из n уравнений относительно n неизвестных a_j .

Замечание. Хотя практически предыдущие уравнения записывают, не проверяя, что в рассматриваемой точке a производные функций g_i независимы, это является существенным условием применимости теоремы.

Предположим, например, что f — некоторая вещественная непрерывно дифференцируемая функция n вещественных переменных x_1, x_2, \dots, x_n и g есть функция $x_1^2 + x_2^2 + \dots + x_{n-1}^2$.

Что же можно сказать о точке a , в которой функция f достигает своего максимального или минимального значения на подмножестве A , определяемом уравнением $g(x) = 0$ в \mathbb{R}^n ? Это уравнение эквивалентно системе уравнений $x_1 = x_2 = \dots = x_{n-1} = 0$, и, следовательно, рассматриваемое подмножество A является осью x_n . Для того чтобы функция f имела в точке a оси x_n условный максимум или минимум, необходимо лишь, чтобы ее частная производная $\frac{\partial f}{\partial x_i}$ равнялась нулю в точке a без каких-либо условий, относящихся к другим частным производным. Если же, напротив, учесть существование множителя Лагранжа λ , при котором $\frac{\partial f}{\partial x_i}(a) = \lambda \frac{\partial g}{\partial x_i}(a)$, $i = 1, 2, \dots, n$, то мы получим, что все частные производные функции f в точке a равны нулю! Но в этом случае теорема 34 неприменима, так как в точке a оси x_n дифференциал функции g равен нулю.

Заметим, кроме того, что, как и для безусловных максимумов и минимумов, этот метод не позволяет утверждать, что данная функция имеет максимум или минимум в точке a , поскольку иами использованы лишь члены первого порядка в разложении Тейлора функции f в точке a . Заметим также, что этот метод совсем не затрагивает условные экстремумы функций, которые могут соответствовать граничным точкам множества E , если оно не является открытой частью E .

Этим методом можно воспользоваться для того, чтобы определить оси некоторой поверхности 2-го порядка в евклидовом n -мерном аффинном пространстве, отыскивая точки этой поверхности, в которых расстояние от центра экстремально. Точно так же можно искать оси сечения такой поверхности плоскостью. В общем случае ищется точка x гиперповерхности, заданной уравнением

$$g(x_1, x_2, \dots, x_n) = 0 \quad (g \text{ непрерывна}) \quad (\text{III}, 10; 6)$$

в пространстве \mathbb{R}^n , евклидово расстояние от которой до заданной точки $b = (b_1, b_2, \dots, b_n)$ минимально. Поскольку поверхность является замкнутым множеством (как прообраз нуля при непрерывном отображении g пространства \mathbb{R}^n в \mathbb{R}), то, согласно изложенному на стр. 83, существует хотя бы одна точка этой поверхности, расстояние от которой до точки b минимально. Мы должны искать точку x , реализующую минимум величины

$$f(x) = \sum_{i=1}^n (x_i - b_i)^2, \quad (\text{III}, 10; 7)$$

среди всех точек, удовлетворяющих уравнению поверхности. Если x удовлетворяет этому условию и если в этой точке дифференциал dg не равен нулю, т. е. n частных производных

$\frac{\partial g}{\partial x_i}$ не обращаются одновременно в нуль, то существует такой множитель Лагранжа λ , при котором имеют место равенства

$$x_i - b_i = \lambda \frac{\partial g}{\partial x_i}(x), \quad i = 1, 2, \dots, n. \quad (\text{III}, 10; 8)$$

Исключая отсюда λ , можно будет определить искомую точку x , выписывая n уравнений с n неизвестными:

$$\frac{\frac{\partial g}{\partial x_1}(x)}{x_1 - b_1} = \frac{\frac{\partial g}{\partial x_2}(x)}{x_2 - b_2} = \dots = \frac{\frac{\partial g}{\partial x_n}(x)}{x_n - b_n}, \quad g(x) = 0. \quad (\text{III}, 10; 9)$$

Эти уравнения выражают тот факт, что искомая точка, если она не особая, является основанием нормали к гиперповерхности, проведенной из точки b . Это согласуется с геометрической интерпретацией, данной после формулировки теоремы 34: в искомой точке x гиперповерхность касательна к сфере с центром b , проходящей через эту точку. При таком решении задачи необходимо иметь в виду следующее:

1°) Точка поверхности, являющаяся основанием нормали, проведенной через точку b , не обязательно определяет минимум расстояния. Она может определять относительный минимум, относительный максимум или седло.

2°) Может случиться, что основания нормалей не дают эффективно абсолютный минимум расстояния и что такой минимум реализуется в особой точке поверхности, т. е. в точке, где все n частных производных $\frac{\partial g}{\partial x_i}$ одновременно равны нулю и где, следовательно, техника множителей Лагранжа не пригодна. Впрочем, заметим, что если для получения (III, 10; 9) исключается параметр λ , то тем самым вводятся эти новые решения. Если поверхность не имеет особых точек, то, как мы видели, минимум обязательно существует. Отсюда следует, что через каждую точку b можно провести к этой поверхности *по крайней мере одну нормаль*. Если, сверх того, поверхность компактна, то расстояние от точки поверхности до точки b , будучи непрерывной функцией на компакте, имеет максимум, и в этом случае, считая, что на поверхности нет особых точек, можно из точки b провести еще по крайней мере одну нормаль, основание которой даст максимум расстояния от точек поверхности до точки b .

Применение теории условных максимумов. Неравенства Гёльдера и Минковского

Пусть c_1, c_2, \dots, c_n — n заданных и фиксированных положительных чисел. Пусть задано, кроме того, некоторое $p > 0$. Для каждой точки $\vec{X} = (X_1, X_2, \dots, X_n)$ из \mathbb{C}^n положим

$$\|\vec{X}\|_p = (c_1|X_1|^p + c_2|X_2|^p + \dots + c_n|X_n|^p)^{1/p} \quad (\text{III}, 10; 10)$$

и назовем эту величину *нормой* \vec{X} *порядка* p *относительно коэффициентов* $c_i > 0$. Очевидно, это число $\geqslant 0$. С другой стороны, тривиально проверяется, что $\|\lambda \vec{X}\| = |\lambda| \|\vec{X}\|$. Наконец, равенство $\|\vec{X}\|=0$ эквивалентно равенству $\vec{X}=\vec{0}$. Для того чтобы обосновать название нормы, остается доказать неравенство выпуклости.

Позже мы увидим, что это неравенство выполняется только для $p \geqslant 1$. Только в этом случае будет оправдано название нормы. Тем не менее мы будем продолжать говорить о *норме порядка* p и обозначать ее тем же самым образом, даже если $p < 1$.

Конечно, отображение $(\vec{X}, p) \rightarrow \|\vec{X}\|_p$ является непрерывным отображением $\mathbb{C}^n \times \mathbb{R}_+$ в \mathbb{R} , где \mathbb{R}_+ есть множество вещественных чисел > 0 . Посмотрим, что происходит с этой нормой, когда p стремится к $+\infty$. Пусть $M = \sup(|X_1|, |X_2|, \dots, |X_n|)$, и предположим для определенности, что i является тем индексом, при котором $|X_i| = M$. Тогда имеют место неравенства

$$\|\vec{X}\|_p \leq \left(\sum_{i=1}^n c_i \right)^{1/p} M \quad \text{и} \quad \|\vec{X}\|_p \geq c_i^{1/p} M. \quad (\text{III}, 10; 11)$$

Правые части этих неравенств при p , стремящемся к $+\infty$, стремятся к M . Отсюда следует, что для фиксированного \vec{X} при p , стремящемся к бесконечности, $\|\vec{X}\|_p$ стремится к M .

Поэтому обычно полагают

$$\|\vec{X}\|_\infty = \sup(|X_1|, |X_2|, \dots, |X_n|). \quad (\text{III}, 10; 12)$$

Заметим теперь, что рассмотренные на стр. 43 три нормы над \mathbb{C}^n являются нормами $\|\vec{X}\|_\infty$, $\|\vec{X}\|_1$, $\|\vec{X}\|_2$. Посмотрим, что происходит с $\|\vec{X}\|_p$ при p , стремящемся к 0. В общем случае характер поведения нормы не представляет интереса. Это интересно лишь в частном случае, когда сумма $\sum_{i=1}^n c_i$ равна 1. В этом случае норма называется *средним порядка* p *чисел* $|X_1|, |X_2|, \dots, |X_n|$ *по отношению к коэффициентам* c_i .

Здесь речь идет о *взвешенном среднем*.

Обычное *среднее арифметическое* соответствует $c_i = 1/n$ и $p = 1$. Для $c_i = 1/n$ и произвольного p средним порядка p является число, p -я степень которого есть среднее арифметическое p -х степеней $|X_i|$.

Положим $\sum_{i=1}^n c_i = 1$ и будем искать предел $\|\vec{X}\|_p$ при p , стремящемся к 0. Предположим сначала, что ни одно из чисел X_i не равно нулю. В этом случае можно написать разложение

$$|X_i|^p = e^{p \ln |X_i|} = 1 + p \ln |X_i| + \dots, \quad (\text{III}, 10; 13)$$

откуда

$$\begin{aligned} \sum_{i=1}^n c_i |X_i|^p &= \sum_{i=1}^n c_i + p \sum_{i=1}^n c_i \ln |X_i| + \dots = \\ &= 1 + p \sum_{i=1}^n c_i \ln |X_i| + \dots; \end{aligned} \quad (\text{III}, 10; 14)$$

учитывая, что $\sum_{i=1}^n c_i = 1$, получаем разложение

$$\ln \left(\sum_{i=1}^n c_i |X_i|^p \right) = p \sum_{i=1}^n c_i \ln |X_i| + \dots. \quad (\text{III}, 10; 15)$$

Отсюда видно, что при p , стремящемся к 0, $\ln \|\vec{X}\|_p = -\frac{1}{p} \ln \left(\sum_{i=1}^n c_i |X_i|^p \right)$ стремится к $\sum_{i=1}^n c_i \ln |X_i|$, и, следовательно, $\|\vec{X}\|_p$ стремится к

$$\prod_{i=1}^n |X_i|^{c_i}. \quad (\text{III}, 10; 16)$$

Эта величина является *взвешенным средним геометрическим* чисел $|X_i|$ с весами c_i .

Если все c_i равны $1/n$, то мы получаем обычное среднее геометрическое $\sqrt[n]{|X_1 X_2 \dots X_n|}$. Легко проверить, что результат остается верным, если некоторые X_i равны нулю. Предел $\|\vec{X}\|_p$ будет тогда равен нулю.

Мы приходим к тому, чтобы при $\sum_{i=1}^n c_i = 1$ положить

$$\|\vec{X}\|_0 = \prod_{i=1}^n |X_i|^{c_i}. \quad (\text{III}, 10; 17)$$

Теорема 35 (неравенство Минковского). Для $p \geq 1$ ¹⁾ имеет место неравенство выпуклости:

$$\|\vec{X} + \vec{Y}\|_p \leq \|\vec{X}\|_p + \|\vec{Y}\|_p, \quad (\text{III}, 10; 18)$$

причем для $1 < p < \infty$ имеет место строгое неравенство, за исключением случая, когда векторы \vec{X} и \vec{Y} пропорциональны с некоторым коэффициентом пропорциональности ≥ 0 .

Доказательство. Неравенство очевидно для $p=1$ и $p=+\infty$. Остается, следовательно, изучить случай конечного $p > 1$ и при этом, естественно, провести доказательство для случая, когда оба вектора \vec{X} и \vec{Y} отличны от нуля. При этом можно ограничиться случаем, когда X_i, Y_i вещественны и ≥ 0 . Однако множество точек \mathbb{R}^n , координаты которых ≥ 0 , не является открытым множеством в векторном нормированном пространстве и, следовательно, теоремы об экстремумах не применимы. Поэтому мы докажем неравенство для всего векторного пространства \mathbb{R}^n . Положим $\|\vec{X}\|_p = \alpha$, $\|\vec{Y}\|_p = \beta$ и $\|\vec{X} + \vec{Y}\|_p = \gamma$. Рассмотрим множество A точек (\vec{X}, \vec{Y}) из $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$, для которых α и β принимают заданные отличные от нуля значения α_0 и β_0 . Это, очевидно, некоторая компактная часть пространства \mathbb{R}^{2n} . В самом деле, это множество A замкнуто как прообраз точки $(\alpha_0, \beta_0) \in \mathbb{R} \times \mathbb{R}$ при непрерывном отображении $(\vec{X}, \vec{Y}) \rightarrow (\|\vec{X}\|_p, \|\vec{Y}\|_p)$ пространства $\mathbb{R}^n \times \mathbb{R}^n$ в $\mathbb{R} \times \mathbb{R}$. Оно ограничено, ибо задание α_0 и β_0 определяет мажоранту для всех координат векторов \vec{X} и \vec{Y} . Согласно теореме 23 гл. II, это множество компактно. Поскольку γ является непрерывной функцией точки (\vec{X}, \vec{Y}) из $\mathbb{R}^n \times \mathbb{R}^n$, она принимает на компакте A максимум или минимум. Обозначим через (\vec{X}_*, \vec{Y}_*) точку, в которой этот минимум или максимум достигается. Это будет точка условного экстремума. В самом деле, в этой точке непрерывная на $\mathbb{R}^n \times \mathbb{R}^n$ функция γ^p достигает своего максимума или

¹⁾ Для $p < 1$ результат не имеет места. Возьмем, например, $p = 1/2$, $n = 2$, $c_1 = c_2 = 1$, $X_1, X_2, Y_1, Y_2 > 0$. Тогда (III, 10; 8) означает, что

$$(\sqrt{X_1 + Y_1} + \sqrt{X_2 + Y_2})^2 \leq (\sqrt{X_1} + \sqrt{X_2})^2 + (\sqrt{Y_1} + \sqrt{Y_2})^2,$$

$$\text{или } \sqrt{(X_1 + Y_1)(X_2 + Y_2)} \leq \sqrt{X_1 X_2} + \sqrt{Y_1 Y_2},$$

$$\text{или } (X_1 + Y_1)(X_2 + Y_2) \leq X_1 X_2 + Y_1 Y_2 + 2 \sqrt{X_1 X_2 Y_1 Y_2},$$

$$\text{или } X_1 Y_2 + X_2 Y_1 \leq 2 \sqrt{X_1 X_2 Y_1 Y_2},$$

в то время как в силу неравенства $a + b \geq 2\sqrt{ab}$ должно иметь место неравенство вида \geq .

минимума среди точек, связанных двумя соотношениями: $\alpha^p(\vec{X}, \vec{Y}) = \alpha^p(\vec{X}) = \alpha_0^p$ и $\beta^p(\vec{X}, \vec{Y}) = \beta^p(\vec{Y}) = \beta_0^p$.

Все три функции α^p , β^p , γ^p непрерывно дифференцируемы.

Рассмотрим, например, функцию α^p . Она имеет частную производную по X_i в открытом множестве $X_i \neq 0$, равную

$$\frac{\partial}{\partial X_i} (\alpha^p) = pc_i \varepsilon_i |X_i|^{p-1}, \quad \varepsilon_i = \operatorname{sign} X_i. \quad (\text{III, 10; 19})$$

При X_i , стремящемся к 0, эта производная имеет равный нулю предел. Согласно теореме 14, примененной к функции α^p , если изменяется только X_i , существует также частная производная на гиперплоскости $X_i = 0$ и эта производная равна нулю¹⁾. Следовательно, производная α^p по X_i всюду существует. Производная функция определена соотношением (III, 10; 19). Эта функция непрерывна во всем $\mathbb{R}^n \times \mathbb{R}^n$, и, следовательно, α^p , имеющая непрерывные частные производные первого порядка, согласно теореме 15, непрерывно дифференцируема. То же самое имеет место для функций β^p и γ^p , производные которых имеют вид

$$\frac{\partial}{\partial Y_i} (\beta^p) = pc_i \eta_i |Y_i|^{p-1}, \quad \eta_i = \operatorname{sign} Y_i, \quad (\text{III, 10; 20})$$

и

$$\frac{\partial}{\partial X_i} (\gamma^p) = \frac{\partial}{\partial Y_i} (\gamma^p) = pc_i \zeta_i |X_i + Y_i|^{p-1}, \quad \zeta_i = \operatorname{sign} (X_i + Y_i). \quad (\text{III, 10; 21})$$

(Очевидно, предыдущие результаты не были правильными для $p \leq 1$.) Теперь в той точке (\vec{X}, \vec{Y}) , где реализуется либо максимум, либо минимум, существуют два множителя Лагранжа λ и μ , такие, что²⁾

$$d(\gamma^p) = \lambda d(\alpha^p) + \mu d(\beta^p), \quad (\text{III, 10; 22})$$

или

$$pc_i \xi_i |X_i + Y_i|^{p-1} = \lambda pc_i \varepsilon_i |X_i|^{p-1}, \quad i = 1, 2, \dots, n. \quad (\text{III, 10; 23})$$

$$pc_i \xi_i |X_i + Y_i|^{p-1} = \mu pc_i \eta_i |Y_i|^{p-1}, \quad i = 1, 2, \dots, n.$$

Если для некоторого индекса i координата X_i равна нулю, то первое уравнение из (III, 10; 23) для этого индекса превращается в равенство $X_i + Y_i = 0$, откуда в свою очередь $Y_i = 0$.

1) Непосредственная проверка показывает, что в тех точках, где $X_i = 0$, производная $\partial/\partial X_i$ равна нулю!

2) При условии, что дифференциалы $d(\alpha^p)$ и $d(\beta^p)$ независимы. Однако один из них зависит только от dX_i , а другой только от dY_i . Они могут быть зависимыми только тогда, когда один из них равен нулю, т. е. когда $\vec{X} = \vec{0}$ или $\vec{Y} = \vec{0}$, что мы исключили, положив $\alpha_0 \neq 0$ и $\beta_0 \neq 0$.

Точно так же, если $Y_i=0$, то второе уравнение даст $X_i=0$. Следовательно, в той точке, в которой имеет место максимум или минимум, если одна из координат X_i , Y_i обратилась в нуль, то и другая также обратится в нуль.

Если ни одна из координат не равна нулю, уравнения (III, 10; 23) для индекса i эквивалентны уравнениям

$$\zeta_i \left| 1 + \frac{Y_i}{X_i} \right|^{p-1} = \varepsilon_i \lambda, \quad \zeta_i \left| 1 + \frac{X_i}{Y_i} \right|^{p-1} = \eta_i \mu. \quad (\text{III, 10; 24})$$

Во всяком случае множество этих уравнений, после исключения λ и μ , можно записать в эквивалентной форме

$$\frac{Y_1}{X_1} = \frac{Y_2}{X_2} = \dots = \frac{Y_n}{X_n}. \quad (\text{III, 10; 25})$$

Некоторые из отношений могут оказаться неопределенными, вида $0/0$ (но не все, поскольку \vec{X} и \vec{Y} не нули). Пусть s — общее значение этих отношений. Тогда $\beta_0 = |s| \alpha_0$ и, следовательно,

$$\gamma = \left(\sum_{i=1}^n c_i |1 + s|^p |X_i|^p \right)^{1/p} = |1 + s| \alpha_0. \quad (\text{III, 10; 26})$$

Отсюда с необходимостью вытекает, что максимум γ равен $\alpha_0 + \beta_0$ и достигается при $s > 0$ и что его минимум равен $|\alpha_0 - \beta_0|$ и достигается при $s < 0$. Теперь получаем, что всегда имеет место неравенство $\gamma \leqslant \alpha + \beta$, являющееся неравенством выпуклости (III, 10; 18). Кроме тех точек, в которых реализуется максимум, везде имеет место строгое неравенство $\gamma < \alpha + \beta$. В точках максимума имеет место (III, 10; 25) с общим значением отношений $s > 0$. Иначе говоря, векторы \vec{X} и \vec{Y} пропорциональны с коэффициентом пропорциональности $\geqslant 0$ ¹⁾. Конечно, это доказывает вторую часть теоремы лишь в том случае, когда векторы \vec{X} и \vec{Y} имеют вещественные координаты. Однако точно так же в общем случае комплексных координат получается последовательность неравенств

$$\|\vec{X} + \vec{Y}\|_p \leqslant \left(\sum_{i=1}^n c_i (|X_i| + |Y_i|)^p \right)^{1/p} \leqslant \|\vec{X}\|_p + \|\vec{Y}\|_p. \quad (\text{III, 10; 27})$$

Первое из них является строгим неравенством, кроме случая, когда все отношения X_i/Y_i вещественны и $\geqslant 0$, чем и доказывается теорема в общем случае.

1) Мы пишем $\geqslant 0$ вместо > 0 для того, чтобы вернуться к отложенному ранее очевидному случаю, когда по крайней мере из двух векторов \vec{X} и \vec{Y} равен нулю.

Теорема 36 (неравенство Гёльдера). *Если при произвольных числах p и q и таком числе r , что $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, обозначить через \vec{XY} точку с координатами $X_1 Y_1, X_2 Y_2, \dots, X_n Y_n$, то*

$$\|\vec{XY}\|_r \leq \|\vec{X}\|_p \|\vec{Y}\|_q. \quad (\text{III}, 10; 28)$$

Кроме того, если $0 < p < +\infty$, $0 < q < +\infty$, то всегда имеет место строгое неравенство, за исключением случая, когда

$$\frac{|Y_1|^q}{|X_1|^p} = \frac{|Y_2|^q}{|X_2|^p} = \dots = \frac{|Y_n|^q}{|X_n|^p}. \quad (\text{III}, 10; 29)$$

Докажем прежде всего теорему для случая, когда все три числа p , q , r больше 1 и конечны. Как и в теореме 35, мы воспользуемся методом множителей Лагранжа. Положим

$$\alpha = \|\vec{X}\|_p, \quad \beta = \|\vec{Y}\|_q, \quad \gamma = \|\vec{XY}\|_r.$$

Если мы рассмотрим точку, в которой достигается либо максимум, либо минимум функции γ , когда α и β принимают заданные значения α_0 и $\beta_0 \neq 0$, то существуют два таких множителя λ и μ , что $d(\gamma^r) = \lambda d(\alpha^p) + \mu d(\beta^q)$, или

$$\begin{aligned} r c_i \varepsilon_i |X_i|^{r-1} |Y_i|^r &= \lambda p c_i \varepsilon_i |X_i|^{p-1}, \quad i = 1, 2, \dots, n, \\ r c_i \eta_i |X_i|^r |Y_i|^{r-1} &= \mu q c_i \eta_i |Y_i|^{q-1}, \quad i = 1, 2, \dots, n. \end{aligned} \quad (\text{III}, 10; 30)$$

Отыскание минимума производится без труда. В самом деле, если для каждого i хотя бы одна из координат X_i , Y_i равна нулю, но таким образом, что одновременно $\alpha = \alpha_0$ и $\beta = \beta_0$ (например, если $X_1 = \frac{\alpha_0}{c_1^{1/p}}$ и $X_2 = X_3 = \dots = X_n = 0$ и далее $Y_1 = 0, Y_2 = \frac{\beta_0}{c_2^{1/q}}, Y_3 = \dots = Y_n = 0$), то \vec{XY} равняется нулю и минимум γ также равен нулю. Мы не будем касаться этого случая и будем искать только максимум.

Если для некоторого индекса i координата X_i равна нулю, то второе уравнение для этого индекса дает $Y_i = 0$ или $\mu = 0$. Однако если $\mu = 0$, то при любом i хотя бы одна из двух координат X_i , Y_i равна нулю, и мы приходим к случаю минимума, уже исключенному из рассмотрения. Таким образом, $\mu \neq 0$ и, следовательно, из $X_i = 0$ следует $Y_i = 0$. Точно так же из $Y_i = 0$ следует $X_i = 0$. Если X_i и Y_i обе $\neq 0$, то уравнения (III, 10; 30) дают соотношения

$$p\lambda |X_i|^p = q\mu |Y_i|^q, \quad i = 1, 2, \dots, n, \quad (\text{III}, 10; 31)$$

а, следовательно, (III, 10; 29).

Некоторые из этих соотношений, но не все, могут представиться в неопределенной форме 0/0.

Положим $|Y_i|^q = s|X_i|^p$. Тогда $\beta_0^q = s\alpha_0^p$. Отсюда получаем

$$\gamma^r = s^{r/q} \sum_{i=1}^n c_i |X_i|^{r+\frac{pr}{q}} = s^{r/q} \sum_{i=1}^n c_i |X_i|^p = s^{r/q} \alpha_0^p, \quad (\text{III}, 10; 32)$$

что дает $\gamma^r = \alpha_0^r \beta_0^r$, или $\gamma = \alpha_0 \beta_0$. Отсюда вытекает, что для каждой точки $(\vec{X}, \vec{Y}) \in \mathbb{R}^n \times \mathbb{R}^n$ имеет место неравенство $\gamma \leq \alpha_0 \beta_0$, что совпадает с неравенством (III, 10; 28).

Кроме того, везде имеет место строгое неравенство $\gamma < \alpha_0 \beta_0$, за исключением отдельных точек, где реализуется максимум, и в этих точках имеют место равенства пропорциональности (III, 10; 29).

Теорема была доказана лишь для конечных p, q, r , больших 1. Однако если мы будем считать теорему доказанной для индексов p, q и числа r , определенного равенством $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, и если мы положим $|X_i| = |\xi_i|^k$, $|Y_i| = |\eta_i|^k$, $0 < k < +\infty$, то будет иметь место неравенство

$$\|\vec{\xi}\vec{\eta}\|_{kr}^k \leq \|\vec{\xi}\|_{kp}^k \|\vec{\eta}\|_{kq}^k, \quad (\text{III}, 10; 33)$$

откуда следует, если возвести обе части равенства в степень $1/k$, что теорема доказана для показателей kp, kq, kr , таких, что

$$\frac{1}{kp} + \frac{1}{kq} = \frac{1}{kr}.$$

Отсюда вытекает, что доказательство теоремы для произвольных конечных показателей >0 можно свести к уже изученному случаю показателей вида $p/k, q/k$, выбирая k таким образом, чтобы числа $p/k, q/k, r/k$ были >1 . Это, однако, не дает возможности доказать теорему, если показатели p, q, r равны

$+\infty$ или 0 при $\sum_{i=1}^n c_i = 1$. Если, например, $p = +\infty$, то $q = r$ и неравенство (III, 10; 28) запишется в виде

$$\|\vec{XY}\|_q \leq \|\vec{X}\|_\infty \|\vec{Y}\|_q. \quad (\text{III}, 10; 34)$$

Это неравенство тривиально, даже если q также равно $+\infty$.

Во всяком случае оставшиеся неравенства могут быть доказаны предельным переходом.

Следствие 1. Два показателя $p, p' \geq 1$ называются сопряженными, если имеет место соотношение

$$\frac{1}{p} + \frac{1}{p'} = 1, \quad \text{или} \quad p' = \frac{p}{p-1}. \quad (\text{III}, 10; 35)$$

В этом частном случае неравенство Гёльдера имеет вид

$$\left| \sum_{i=1}^n c_i X_i Y_i \right| \leq \| \vec{X} \|_p \| \vec{Y} \|_{p'}. \quad (\text{III}, 10; 36)$$

Кроме того, для $p \neq 1$ и $\neq \infty$ имеет место строгое неравенство, за исключением случая, когда выполняется соотношение пропорциональности

$$\frac{|Y_1|^{p'}}{|X_1|^p} = \frac{|Y_2|^{p'}}{|X_2|^p} = \dots = \frac{|Y_n|^{p'}}{|X_n|^p}, \quad (\text{III}, 10; 37)$$

где все произведения $X_i Y_i$ имеют один и тот же аргумент¹⁾.

Для доказательства достаточно применить общее неравенство, в котором $r=1$. Из него тогда получится соотношение $\| \vec{XY} \|_1 \leq \| \vec{X} \|_p \| \vec{Y} \|_{p'}$ и, следовательно, тем более соотношение (III, 10; 36). Однако

$$\left| \sum_{i=1}^n c_i X_i Y_i \right| < \| \vec{XY} \|_1, \quad (\text{III}, 10; 38)$$

кроме случая, когда все комплексные числа $X_i Y_i$ имеют один и тот же аргумент. Очевидно, что, кроме случая, указанного в условии, неравенство (III, 10; 36) будет строгим.

Следствие 1₂ (неравенство Коши — Шварца). Имеет место неравенство

$$\left| \sum_{i=1}^n c_i X_i \bar{Y}_i \right| \leq \| \vec{X} \|_2 \| \vec{Y} \|_2 \quad (\text{III}, 10; 39)$$

со знаком строгого неравенства $<$, кроме того случая, когда векторы \vec{X} и \vec{Y} пропорциональны.

Достаточно применить следствие 1 к сопряженным показателям $p=p'=2$. Аргумент $X_i \bar{Y}_i$ совпадает с аргументом X_i / Y_i , а все числа Y_i / X_i имеют общий модуль и общий аргумент тогда и только тогда, когда векторы \vec{X} и \vec{Y} пропорциональны. Мы давали ранее другое доказательство (формула (III, 1; 11)). Для получения (III, 10; 39) достаточно применить формулу (III, 1; 11) к положительно определенной полуторалинейной эрмитовой форме

$$B(\vec{X}, \vec{Y}) = \sum_{i=1}^n c_i X_i \bar{Y}_i. \quad (\text{III}, 10; 40)$$

¹⁾ Если все X_i отличны от нуля, то соотношение (III, 10; 37) означает, что существует такая постоянная k , что $|Y_i| = k |X_i|^{p-1}$.

Следствие 2. Если сумма весов $\sum_{i=1}^n c_i$ равна 1, то для фиксированного \vec{X} среднее порядка p , $\|\vec{X}\|_p$ является непрерывной возрастающей функцией p . Эта функция строго возрастает, кроме случая, когда $|X_1|=|X_2|=\dots=|X_n|$, в котором она постоянна.

В самом деле, достаточно применить неравенство Гёльдера к случаю, когда $Y_1=Y_2=\dots=Y_n=1$. Если учесть, что $\sum_{i=1}^n c_i = 1$, то $\|\vec{Y}\|_q = 1$. Это дает при любом $q \geq 0$ и, следовательно, при любом $r \leq p$ неравенство

$$\|\vec{X}\|_r \leq \|\vec{X}\|_p. \quad (\text{III}, 10; 41)$$

Более того, здесь будет всегда иметь место строгое неравенство, кроме случая, когда

$$\frac{1}{|X_1|^p} = \frac{1}{|X_2|^p} = \dots = \frac{1}{|X_n|^p}, \quad (\text{III}, 10; 42)$$

т. е. когда все $|X_i|$ равны. При этом среднее очевидным образом будет постоянной, равной общему значению этих чисел.

Следствие 2₂. Среднее геометрическое конечного числа неотрицательных чисел не превосходит их среднего арифметического и всегда строго меньше его, за исключением случая, когда все эти числа равны между собой.

Для доказательства достаточно применить (III, 10; 41) при $r=0$ и $p=1$.

Доказательство неравенств Гёльдера и Минковского с помощью теории выпуклых функций. Функция $x \rightarrow x^s$, $s \geq 1$, выпукла для $x > 0$. В самом деле, она непрерывна, а ее вторая производная для $x > 0$ существует и равна $s(s-1)x^{s-2} > 0$ (теорема 7₂). Следовательно, для $z_1, z_2, \dots, z_n \geq 0$ и $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$ имеем:

$$\left(\frac{\alpha_1 z_1 + \dots + \alpha_n z_n}{\alpha_1 + \dots + \alpha_n} \right)^s \leq \frac{\alpha_1 z_1^s + \dots + \alpha_n z_n^s}{\alpha_1 + \dots + \alpha_n}, \quad (\text{III}, 10; 43)$$

или, умножая на $(\alpha_1 + \dots + \alpha_n)^s$ и возводя в степень $1/s$:

$$(\alpha_1 z_1 + \dots + \alpha_n z_n) \leq (\alpha_1 z_1^s + \dots + \alpha_n z_n^s)^{1/s} (\alpha_1 + \dots + \alpha_n)^{1-1/s}. \quad (\text{III}, 10; 44)$$

Положим теперь для $X_i, Y_i \geq 0$, $c_i \geq 0$, $p, q > 0$, $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$

$$\alpha_i = c_i Y_i^q, \quad z_i = X_i' Y_i^{r-q}, \quad s = \frac{p}{r}. \quad (\text{III}, 10; 45)$$

Тогда

$$\frac{1}{s} = \frac{r}{p}, \quad 1 - \frac{1}{s} = \frac{r}{q} \quad (\text{ибо } \frac{r}{p} + \frac{r}{q} = 1),$$

$$\alpha_i z_i = c_i (X_i Y_i)' , \quad (\text{III, 10; 46})$$

$$\alpha_i z_i^s = c_i X_i^s Y_i^{q+s(r-q)} = c_i X_i^p Y_i^{p+q-\frac{pq}{r}} = c_i X_i^p.$$

Возводя теперь (III, 10; 44) в степень $1/r$, получаем неравенство Гёльдера

$$(c_1 (X_1 Y_1)^r + \dots + c_n (X_n Y_n)^r)^{1/r} \leq$$

$$\leq (c_1 X_1^p + \dots + c_n X_n^p)^{1/p} (c_1 Y_1^q + \dots + c_n Y_n^q)^{1/q}. \quad (\text{III, 10; 47})$$

Функция $x \rightarrow (1 - x^{1/p})^p$, $p \geq 1$, выпукла на $[0, 1]$, ибо ее вторая производная на $]0, 1[$ равна

$$(1 - x^{1/p})^{p-2} \left(1 - \frac{1}{p}\right) x^{1/p-2} \geq 0.$$

Следовательно, для $0 \leq z_1, z_2, \dots, z_n \leq 1$ и $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$ имеем

$$\left(1 - \left(\frac{\alpha_1 z_1 + \dots + \alpha_n z_n}{\alpha_1 + \dots + \alpha_n}\right)^{1/p}\right)^p \leq \frac{\alpha_1 (1 - z_1^{1/p})^p + \dots + \alpha_n (1 - z_n^{1/p})^p}{\alpha_1 + \dots + \alpha_n}. \quad (\text{III, 10; 48})$$

Возводя в степень $1/p$, а затем умножая на $(\alpha_1 + \dots + \alpha_n)^{1/p}$, получаем

$$(\alpha_1 + \dots + \alpha_n)^{1/p} \leq (\alpha_1 z_1 + \dots + \alpha_n z_n)^{1/p} +$$

$$+ (\alpha_1 (1 - z_1^{1/p})^p + \dots + \alpha_n (1 - z_n^{1/p})^p)^{1/p}. \quad (\text{III, 10; 49})$$

Для $c_i, X_i, Y_i \geq 0$ положим

$$\alpha_i = c_i (X_i + Y_i)^p, \quad \alpha_i \geq 0,$$

$$z_i = \left(\frac{X_i}{X_i + Y_i}\right)^p, \quad 0 \leq z_i \leq 1. \quad (\text{III, 10; 50})$$

Отсюда

$$\alpha_i z_i = c_i X_i^p,$$

$$\alpha_i (1 - z_i^{1/p})^p = c_i Y_i^p. \quad (\text{III, 10; 51})$$

Теперь (III, 10; 49) принимает вид неравенства Минковского:

$$(c_1 (X_1 + Y_1)^p + \dots + c_n (X_n + Y_n)^p)^{1/p} \leq$$

$$\leq (c_1 X_1^p + \dots + c_n X_n^p)^{1/p} (c_1 Y_1^p + \dots + c_n Y_n^p)^{1/p}. \quad (\text{III, 10; 52})$$

Неравенство Минковского можно также получить из неравенства Гёльдера. В самом деле,

$$\sum_{i=1}^n (X_i + Y_i)^p = \sum_{i=1}^n (X_i + Y_i)^{p-1} X_i + \sum_{i=1}^n (X_i + Y_i)^{p-1} Y_i. \quad (\text{III}, 10; 53)$$

Согласно неравенству Гёльдера, написанному относительно показателей p , p' , $\frac{1}{p} + \frac{1}{p'} = 1$ ($p' = \frac{p}{p-1}$), имеем:

$$\begin{aligned} \sum_{i=1}^n (X_i + Y_i)^p &\leq \left(\sum_{i=1}^n (X_i + Y_i)^{(p-1)p'} \right)^{1/p'} \left(\sum_{i=1}^n X_i^p \right)^{1/p} + \\ &+ \left(\sum_{i=1}^n (X_i + Y_i)^{(p-1)p'} \right)^{1/p'} \left(\sum_{i=1}^n Y_i^p \right)^{1/p}. \end{aligned} \quad (\text{III}, 10; 54)$$

Поскольку $(p-1)p' = p$, достаточно обе части равенства умножить на $\left(\sum_{i=1}^n (X_i + Y_i)^p \right)^{-1/p'}$ и получить

$$\left(\sum_{i=1}^n (X_i + Y_i)^p \right)^{1-(1/p')} \leq \left(\sum_{i=1}^n X_i^p \right)^{1/p} + \left(\sum_{i=1}^n Y_i^p \right)^{1/p}, \quad (\text{III}, 10; 55)$$

что совпадает с неравенством Минковского, так как $1 - \frac{1}{p'} = \frac{1}{p}$.

§ 11. ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ

Постановка задачи

Рассмотрим следующую задачу механики:

Пусть в трехмерном пространстве заданы две точки A и B и некоторая кривая \mathcal{C} класса C^1 , соединяющая эти точки. По этой кривой без трения и под влиянием своего веса свободно скользит некоторая материальная точка.

Выходя из точки A с нулевой начальной скоростью, она попадает в точку B через время t . При заданных точках A и B это время является функцией рассматриваемой кривой. Спрашивается, как следует выбрать эту кривую, чтобы время t было минимальным? Кривая, обладающая таким свойством, называется *брахистохроной* («кривой наискорейшего спуска»). Будем пока считать очевидным, что кривая, реализующая этот минимум, располагается в вертикальной плоскости, проходящей через отрезок AB . Выберем в этой плоскости горизонтальную ось x и вертикальную ось z , направленную вниз. Пусть $z=a$, $x=\alpha$ и $z=b$, $x=\beta$ — координаты точек A и B . Классическая теория падения тел говорит, что в заданной точке кривой при

отсутствии трения скорость v материальной точки задается формулой

$$v^2 = 2g(z - a). \quad (\text{III}, 11; 1)$$

Отсюда получается дифференциал времени как функция дифференциала криволинейной абсциссы: $dt = \frac{ds}{\sqrt{2g(z-a)}}$, так что окончательно время, необходимое для перехода от A к B , задается интегралом

$$\int_a^b \frac{\sqrt{1+x'^2}}{\sqrt{2g(z-a)}} dz = \int_a^b \frac{\sqrt{1+f'^2(z)}}{\sqrt{2g(z-a)}} dz, \quad (\text{III}, 11; 2)$$

в котором предполагается, что кривая определяется с помощью x как функции z : $x=f(z)$ ¹⁾. Мы приходим к следующей математической задаче: найти такую вещественную функцию f вещественной переменной z , для которой интеграл (III, 11; 2) при заданных $a, b, g, f(a)=\alpha, f(b)=\beta$ принимает возможно меньшее значение.

Можно поставить общую задачу следующим образом. Пусть $[a, b]$ — вещественный интервал, F — аффинное нормированное пространство над полем вещественных чисел, \mathcal{U} — открытое множество из $F \times \vec{F}$. Пусть, с другой стороны, задана непрерывная вещественная функция L : $(x, (y, \vec{Y})) \rightarrow L(x, y, \vec{Y})$, определенная на $[a, b] \times \mathcal{U}$. Пусть, наконец, $f: x \rightarrow y=f(x)$ — непрерывно дифференцируемая функция, определенная на $[a, b]$ со значениями в F , такая, что для $x \in [a, b]$ все $(f(x), \vec{f}'(x))$ лежат в \mathcal{U} ²⁾.

При этих условиях сложная функция $x \rightarrow L(x, f(x), \vec{f}'(x))$ является непрерывной функцией на $[a, b]$ со значениями в \mathbb{R} . При этом можно вычислить интеграл

$$J(f) = \int_a^b L(x, f(x), \vec{f}'(x)) dx, \quad (\text{III}, 11; 3)$$

являющийся некоторым вещественным числом.

¹⁾ Возможность выразить x как функцию z спорна: см. рассуждения, приведенные на стр. 386.

²⁾ Поскольку речь идет о функции на \mathbb{R} , $\vec{f}'(x)$ является производным вектором в точке x , т. е. элементом \vec{F} , определенным с помощью формулы (III, 3; 1). Он равен $f'(x) \cdot 1$, если $f'(x) \in \mathcal{L}(\mathbb{R}; \vec{F})$ является производным отображением. С другой стороны, $[a, b]$ не является открытым множеством в \mathbb{R} ; см. по этому поводу начало § 2, стр. 195.

Это вещественное число зависит от выбора функции f . Тем самым определяется некоторая функция $J: f \rightarrow J(f)$. Требуется при фиксированных a, b и L среди всех функций f , принимающих фиксированные значения α, β на концах a, b , найти ту, которая обеспечивает интегралу $J(f)$ максимум или минимум. Если F является m -мерным пространством, то введением системы координат его можно отождествить с \mathbb{R}^m , и тогда функция, определенная на $[a, b]$ со значениями в F , является системой вещественных функций $f_i, i=1, 2, \dots, m$, определенных на $[a, b]$, а производная функция является системой m производных функций f'_i . В этом случае L есть вещественная функция $2m+1$ независимых переменных $x, y_i, Y_i, i=1, 2, \dots, m$, и среди всех систем непрерывно дифференцируемых на $[a, b]$ и принимающих заданные значения в a и b функций $f_i, i=1, 2, \dots, m$, отыскиваются те, которые обеспечивают экстремум интегралу

$$J(f_1, \dots, f_m) = \int_a^b L(x, f_1(x), \dots, f_m(x), f'_1(x), \dots, f'_m(x)) dx. \quad (\text{III}, 11; 4)$$

Напомним, что пространство $E = (F^{[a, b]})_{cb}$, непрерывно дифференцируемых функций, определенных на $[a, b]$ со значениями в F , является аффинным нормированным пространством с присоединенным векторным нормированным пространством $(\vec{F}^{[a, b]})_{cb}$ ¹⁾. Расстояние между двумя элементами f и g аффинного пространства E задается формулой (III, 3; 41):

$$\|\overrightarrow{f - g}\|_1 = \sup_{a \leq x \leq b} (\|f(x) - g(x)\|, \|f'(x) - g'(x)\|). \quad (\text{III}, 11; 5)$$

Мы рассмотрим здесь сначала некоторое подмножество аффинного пространства E — множество Ω , образованное такими функциями f , что образ при отображении (f, \vec{f}) интервала $[a, b]$ содержится в открытом множестве \mathcal{U} из $F \times \vec{F}$.

Теорема 37. Пусть: $[a, b]$ — отрезок в \mathbb{R} , F — аффинное нормированное пространство над полем вещественных чисел и \mathcal{U} — открытое множество из $F \times \vec{F}$. Множество Ω функций f , определенных на $[a, b]$ со значениями в F , таких, что $(f(x), \vec{f}(x))$ для всех $x \in [a, b]$ лежит в \mathcal{U} , является открытым множеством в пространстве $E = (F^{[a, b]})_{cb}$, непрерывно дифференцируемых функций, определенных на $[a, b]$ со значениями в F .

¹⁾ Поскольку $[a, b]$ компактно, непрерывная на $[a, b]$ функция ограничена и символ cb , означающий непрерывность и ограниченность, может быть заменен символом c , обозначающим непрерывность.

В самом деле, пусть f_0 — некоторая точка этого множества. Образ компактного интервала $[a, b]$ при отображении (f_0, \vec{f}_0) является компактом K в \mathcal{U} (теорема 28 гл. II). Согласно показанному на стр. 84, расстояние от K до $C\mathcal{U}$ является некоторым числом $\delta > 0$. Рассмотрим теперь в E открытый шар с центром f_0 радиуса δ , образованный всеми функциями \vec{f} , удовлетворяющими неравенству $\|\vec{f} - f_0\| < \delta$. Мы видим, что для каждой такой функции \vec{f} отображение (f, \vec{f}) принимает все свои значения в \mathcal{U} . Иначе говоря, каждый такой шар принадлежит множеству Ω и множество Ω открыто в \vec{E} .

Множество функций $f \in E$, удовлетворяющих заданным равенствам $f(a) = a$, $f(b) = b$, является некоторым аффинным подпространством E_0 из E . Его присоединенное векторное подпространство является множеством \vec{E}_0 функций из \vec{E} , обращающихся в нуль в точках a и b . Множество Ω отсекает на E_0 некоторое открытое множество Ω_0 из E_0 . Далее, $J: f \rightarrow J(f)$ является вещественной функцией, определенной на открытом множестве Ω_0 аффинного нормированного пространства E_0 , и мы ищем максимум или минимум этой функции. Теорема 23 должна нам дать необходимое условие. Мы видим здесь, что теорема 23 и все теоремы дифференциального исчисления представляют интерес для случая бесконечномерных нормированных пространств. В самом деле, если даже F конечномерно, если даже это, например, вещественная прямая \mathbb{R} , входящее сюда пространство E бесконечномерно.

Для облегчения формулировок мы не будем повторять данные $[a, b] \subset \mathbb{R}$, F , \mathcal{U} , Ω , L , J , если в этом не будет особой необходимости.

Дифференцируемость J

Для изучения дифференцируемости функции J мы будем рассматривать общий случай аффинного пространства E и рассматривать производную $J'(f_0)$ на элементе f_0 . Сужение J на E_0 имеет производной в $f_0 \in E_0$ сужение производной $J'(f_0)$ на \vec{E}_0 . Обозначим через $\vec{\delta f}$ заданное приращение f . Это некоторый элемент из \vec{E} , т. е. функция $x \rightarrow \vec{\delta f}(x)$, определенная на $[a, b]$ со значениями в \vec{F} . Мы будем искать «главную часть» $\vec{\delta J}$ приращения ΔJ функции J , когда f_0 заменяется на $f_0 + \vec{\delta f}$. Обозначение $\vec{\delta f}$, которое в том случае, когда f является отображением $[a, b]$ в F , означает дифференциал этого отображения:

$\vec{df} = \vec{f}'(x) dx \in \vec{F}$, применять опасно. Поэтому мы применяем обозначение $\vec{\delta f}$ и по аналогии δJ .

Теорема 38. Вещественная функция J , определенная на Ω формулой (III, 11; 3), непрерывно дифференцируема, и ее производная $\overset{\leftarrow}{J'}(f_0) \in \mathcal{L}(\vec{E}; \mathbb{R}) = \vec{E}'$ в точке f_0 определяется формулой

$$\begin{aligned}\vec{\delta f} \rightarrow \delta J &= J'(f_0) \cdot \vec{\delta f}_0 = \\ &= \int_a^b (L'(x, f_0(x), \vec{f}'_0(x)) : (0, \vec{\delta f}(x), \vec{\delta f}'(x))) dx = \\ &= \int_a^b (\overset{\leftarrow}{\partial_2 L}(x, f_0(x), \vec{f}'_0(x)) \cdot \vec{\delta f}(x) + \overset{\leftarrow}{\partial_3 L}(x, f_0(x), \vec{f}'_0(x)) \cdot \vec{\delta f}'(x)) dx,\end{aligned}\quad (\text{III, 11; 6})$$

где L' — полная производная, а $\overset{\leftarrow}{\partial_2 L}$ и $\overset{\leftarrow}{\partial_3 L}$ — частные производные функции L (определенной на открытом множестве в $[a, b] \times F \times \vec{F}$) по 2-й и 3-й переменным.

Заметим, что вместо $\overset{\leftarrow}{\partial_2 L}$ и $\overset{\leftarrow}{\partial_3 L}$ можно было бы писать $\overset{\leftarrow}{\partial L}/\partial y$ и $\overset{\leftarrow}{\partial L}/\partial \vec{Y}$, поскольку мы через y и \vec{Y} обозначили переменные в пространствах F и \vec{F} .

Вместо $L(x, f(x), \vec{f}'(x))$ коротко пишут также $L(x, y, \vec{y}')$; поэтому вместо $\overset{\leftarrow}{\partial_2 L}$ и $\overset{\leftarrow}{\partial_3 L}$ можно также писать $\overset{\leftarrow}{\partial L}/\partial y$ и $\overset{\leftarrow}{\partial L}/\partial \vec{y}'$. Это производные функции $(x, y, \vec{y}') \rightarrow L(x, y, \vec{y}')$, когда $x \in [a, b]$, $y \in F$, $\vec{y}' \in \vec{F}$ рассматриваются как независимые переменные. После того как они вычислены, следует заменить y на $f_0(x)$ и \vec{y}' на $\vec{f}'_0(x)$.

Так как $y \rightarrow L(x, y, \vec{Y})$ является отображением пространства F в \mathbb{R} , то $\overset{\leftarrow}{\partial_2 L}(x, y, \vec{Y})$ является элементом пространства $\mathcal{L}(\vec{F}, \mathbb{R}) = \vec{F}'$, сопряженного к \vec{F} , откуда следует выбор стрелки $\overset{\leftarrow}{\partial_2 L}$. То же самое имеет место и для $\overset{\leftarrow}{\partial_3 L}$. Далее, для каждого $x \in [a, b]$ имеем $\overset{\leftarrow}{\partial_2 L}(y, f_0(x), \vec{f}'_0(x)) \in \vec{F}'$ и $\vec{\delta f}(x) \in \vec{F}$, откуда следует, что $\overset{\leftarrow}{\partial_2 L}(x, f_0(x), \vec{f}'_0(x)) \cdot \vec{\delta f}(x)$ в (III, 11; 6) является вещественным числом; (III, 11; 6) является интегралом от некоторой вещественной функции на $[a, b]$.

Доказательство. Вариация ΔJ интеграла, соответствующая вариации $\delta \vec{f}$ функции f , задается формулой

$$\Delta J = \int_a^b [L(x, f_0(x) + \delta \vec{f}(x), \vec{f}'_0(x) + \delta \vec{f}'(x)) - L(x, f_0(x), \vec{f}'_0(x))] dx. \quad (\text{III}, 11; 7)$$

Формула конечных приращений (следствие 1 теоремы 13) для фиксированного x дает

$$\begin{aligned} L(x, f_0(x) + \delta \vec{f}(x), \vec{f}'_0(x) + \delta \vec{f}'(x)) - L(x, f_0(x), \vec{f}'_0(x)) &= \\ &= L'(x, f_0(x), \vec{f}'_0(x)) \cdot (0, \delta \vec{f}(x), \delta \vec{f}'(x)) + R(x) = \\ &= \overleftarrow{\partial_2} L(x, f_0(x), \vec{f}'_0(x)) \cdot \delta \vec{f}(x) + \overleftarrow{\partial_3} L(x, f_0(x), \vec{f}'_0(x)) \cdot \delta \vec{f}'(x) + R(x), \end{aligned} \quad (\text{III}, 11; 8)$$

где $R(x)$ допускает оценку

$$\begin{aligned} |R(x)| &\leq \sup_{\substack{\|\xi\| \leq \|\delta \vec{f}(x)\| \\ \|\xi'\| \leq \|\delta \vec{f}'(x)\|}} \|L'(x, f_0(x) + \xi, \vec{f}'_0(x) + \xi') - \\ &\quad - L(x, f_0(x), \vec{f}'_0(x))\| \cdot \|(0, \delta \vec{f}(x); \delta \vec{f}'(x))\|^1. \end{aligned} \quad (\text{III}, 11; 9)$$

Воспользуемся теперь усиленной теоремой (теорема 31 гл. II) о равномерной непрерывности²⁾. Функция L' предполагалась непрерывной на $[a, b] \times \mathcal{U}$. Множество точек $(x, f_0(x), \vec{f}'_0(x))$, $x \in [a, b]$, содержится в компакте $[a, b] \times K$ множества $[a, b] \times \mathcal{U}$ (в обозначениях, принятых при доказательстве теоремы 37).

¹⁾ На $\mathbb{R} \times \vec{F} \times \vec{F}$ мы выбрали норму $\|(X, \vec{Y}, \vec{Z})\| = \sup(|X|, \|\vec{Y}\|, \|\vec{Z}\|)$.

²⁾ Речь идет о следующей теореме: Если f является непрерывным отображением метрического пространства E в метрическое пространство F и если K является компактом в E , то, каково бы ни было $\varepsilon > 0$, существует такое $\eta > 0$, что из $d(x', x'') \leq \eta$, $x' \in K$, следует $d(f(x'), f(x'')) \leq \varepsilon$. Теорема 31 относится к $E = K$. Доказательство остается прежним. Если бы это было не так, то можно было бы найти две такие последовательности $x'_n \in K$, $x''_n \in E$, что $d(x'_n, x''_n) \leq 1/n$ и $d(f(x'_n), f(x''_n)) > \varepsilon$. Поскольку K является компактом, можно из x'_n извлечь подпоследовательность, сходящуюся к некоторому элементу $c \in K$. Тогда x''_n тоже стремились бы к c в E . Так как f непрерывна в c , то $f(x'_n)$ и $f(x''_n)$ стремятся к $f(c)$, а, значит, $d(f(x'_n), f(x''_n))$ стремится к 0, и мы получили противоречие.

Итак, при заданном $\epsilon > 0$ существует число η , $0 < \eta < \delta$, такое, что из $x \in [a, b]$, $\|\vec{\xi}\| \leq \eta$, $\|\vec{\xi}'\| \leq \eta$ следует

$$\|L'(x, f_0(x) + \vec{\xi}, \vec{f}'_0(x) + \vec{\xi}') - L'(x, f_0(x), \vec{f}'_0(x))\| \leq \frac{\epsilon}{b-a}. \quad (\text{III, 11; 10})$$

Далее, из $\|\vec{\delta f}\|_1 \leq \eta$ следует, что

$$|R(x)| \leq \frac{\epsilon}{b-a} \|\vec{\delta f}\|_1. \quad (\text{III, 11; 11})$$

Таким образом, имеем:

$$\Delta J = \int_a^b (\overset{\leftarrow}{\partial_2 L} \cdot \vec{\delta f} + \overset{\leftarrow}{\partial_3 L} \cdot \vec{\delta f}') dx + \int_a^b R(x) dx, \quad (\text{III, 11; 11}_2)$$

где каждая из функций $x \rightarrow \overset{\leftarrow}{\partial_2 L}(x, f_0(x), \vec{f}'_0(x)) \vec{\delta f}(x)$, $x \rightarrow \overset{\leftarrow}{\partial_3 L}(x, f_0(x), \vec{f}'_0(x)) \vec{\delta f}'(x)$, а, следовательно, также и $x \rightarrow R(x)$ непрерывны. В самом деле, $\overset{\leftarrow}{\partial_2 L}$ является непрерывным отображением $[a, b] \times \mathcal{U}$ в \vec{F}' , поскольку L непрерывно дифференцируема. Функции $f_0, \vec{f}'_0, \vec{\delta f}$ непрерывны как функции из $[a, b]$ в F, \vec{F}, \vec{F} соответственно. Следовательно, $x \rightarrow \overset{\leftarrow}{\partial_2 L}(x, f_0(x), \vec{f}'_0(x))$ является непрерывной функцией из $[a, b]$ в \vec{F}' и, поскольку $(\vec{u}, \vec{X}) \rightarrow \vec{u} \cdot \vec{X} = \langle \vec{u}, \vec{X} \rangle$ есть непрерывное отображение из $(\vec{F}' = \mathcal{L}(\vec{F}; \mathbb{R})) \times \vec{F}$ в \mathbb{R} , то $x \rightarrow \overset{\leftarrow}{\partial_2 L}(x, f_0(x), \vec{f}'_0(x)) \cdot \vec{\delta f}(x)$ есть непрерывная функция из $[a, b]$ в \mathbb{R} .

Второй интеграл при $\|\vec{\delta f}\|_1 \leq \eta$ мажорируется величиной $\epsilon \|\vec{\delta f}\|_1$. Следовательно, он является бесконечно малым по сравнению с $\|\vec{\delta f}\|_1$, когда $\|\vec{\delta f}\|_1$ стремится к 0.

Первый интеграл определяет некоторую линейную форму, непрерывную на \vec{E} . В самом деле,

$$\begin{aligned} \left| \int_a^b (\overset{\leftarrow}{\partial_2 L} \cdot \vec{\delta f} + \overset{\leftarrow}{\partial_3 L} \cdot \vec{\delta f}') dx \right| &= \\ &= \left| \int_a^b (L'(x, f_0(x), \vec{f}'_0(x)) \cdot (0, \vec{\delta f}(x), \vec{\delta f}'(x))) dx \right| \leq \\ &\leq (b-a) M \|\vec{\delta f}\|_1, \quad \text{где } M = \sup_{a \leq x \leq b} \|L'(x, f_0(x), \vec{f}'_0(x))\|. \end{aligned} \quad (\text{III, 11; 11}_3)$$

Тем самым мы доказали, что функция J дифференцируема в $f_0 \in \Omega$ и что ее производная задается формулой (III, 11; 6).

Покажем теперь, что функция J непрерывно дифференцируема на Ω . Рассмотрим элементы f_0 и \tilde{f} из Ω . Имеем:

$$\begin{aligned} & \overleftarrow{J'(f)} \cdot \delta \tilde{f} - \overleftarrow{J'(f_0)} \cdot \delta \tilde{f} = \\ & = \int_a^b (L'(x, f(x), \vec{f'}(x)) - L'(x, f_0(x), \vec{f'_0}(x))) \cdot (0, \vec{\delta f}(x), \vec{\delta f'}(x)) dx. \end{aligned} \quad (\text{III, 11; 11}_4)$$

Согласно (III, 11; 10), из $\|\overrightarrow{f - f_0}\|_1 \leq \eta$ следует, что

$$|(\overleftarrow{J'(f)} - \overleftarrow{J'(f_0)}) \cdot \vec{\delta f}| \leq \varepsilon \|\vec{\delta f}\|_1 \quad (\text{III, 11; 12})$$

и

$$\|\overleftarrow{J'(f)} - \overleftarrow{J'(f_0)}\| = \sup_{\|\vec{\delta f}\|_1 \leq 1} |(\overleftarrow{J'(f)} - \overleftarrow{J'(f_0)}) \cdot \vec{\delta f}| \leq \varepsilon. \quad (\text{III, 11; 12}_2)$$

Итак, $\overleftarrow{J'(f)} - \overleftarrow{J'(f_0)}$ сходится к $\overleftarrow{0}$ в \tilde{E}' , когда f стремится к f_0 в E , и, следовательно, J принадлежит классу C^1 .

Пример. Небесполезно бегло рассмотреть простой случай $F = \mathbb{R}$ и $\mathcal{U} = [a, b] \times \mathbb{R} \times \mathbb{R}$. Будем предполагать, следовательно, что L является функцией класса C^1 трех вещественных переменных x, y, y' , $x \in [a, b]$, $y \in \mathbb{R}$, $y' \in \mathbb{R}$. Рассматривается функция J , определенная на $(\mathbb{R}^{[a, b]})_{cb, 1}$ формулой

$$J(f) = \int_a^b L(x, f(x), f'(x)) dx. \quad (\text{III, 11; 12}_3)$$

Функции f дается приращение δf , являющееся также некоторой функцией класса C^1 на $[a, b]$. Приращение ΔJ функционала J в этом случае записывается в виде

$$\Delta J = \int_a^b (L(x, f_0 + \delta f, f'_0 + \delta f') - L(x, f_0, f'_0)) dx. \quad (\text{III, 11; 12}_4)$$

Его «главной частью» является дифференциал

$$\delta J = \int_a^b \left(\frac{\partial L}{\partial y}(x, f_0(x), f'_0(x)) \delta f(x) + \frac{\partial L}{\partial y'}(x, f_0(x), f'_0(x)) \delta f'(x) \right) dx. \quad (\text{III, 11; 12}_5)$$

Обоснование этой формулы вытекает из оценки вида

$$|\delta J - \Delta J| \leq \varepsilon \|\delta f\|_1 \quad \text{для } \|\delta f\|_1 \leq \eta \text{ (стр. 375)}, \quad (\text{III, 11; 12}_6)$$

получающейся применением к L формулы конечных приращений.

Преобразуем теперь (III, 11; 6), вводя для L и элемента $f_0 \in \Omega$, в котором мы вычисляем производную J , упрощающее предположение: L и f_0 взяты из класса C^2 .

Теорема 39. Если в условиях теоремы 38 функция L принадлежит классу C^2 на $[a, b] \times \mathcal{U}$ и если $f_0 \in \Omega$ принадлежит классу C^2 на $[a, b]$, то производная $J'(f_0)$ может быть записана в виде

$$\delta J = \overleftarrow{J'}(f_0) \cdot \vec{\delta f} = [\overleftarrow{\partial_3 L}(x, f_0(x), \vec{f}'_0(x)) \cdot \vec{\delta f}(x)]_{x=a}^{x=b} + \\ + \int_a^b [\overleftarrow{\partial_2 L}(x, f_0(x), \vec{f}'_0(x)) - \frac{d}{dx}(\overleftarrow{\partial_3 L}(x, f_0(x), \vec{f}'_0(x))] \cdot \vec{\delta f}(x) dx. \quad (\text{III, 11; 13})$$

Квадратная скобка $[]_{x=a}^{x=b}$ означает разность значений для $x = b$ и $x = a$.

Доказательство. Достаточно в формуле (III, 11; 6) применить интегрирование по частям. В этой формуле d/dx является производной отображения $x \rightarrow \overleftarrow{\partial_3 L}(x, f_0(x), \vec{f}'_0(x))$ отрезка $[a, b]$ в \vec{F}' . Это, следовательно, не частная производная $\partial_1 L = \partial/\partial x$, а полная производная, которая учитывает, что f_0 и \vec{f}'_0 являются функциями x . Выполняемое здесь интегрирование по частям является более общим, чем то, которое обычно употребляется в математическом анализе. Оно производится по формуле

$$\int_a^b B(\vec{u}(x), \vec{v}'(x)) dx = [B(\vec{u}(x), \vec{v}(x))]_{x=a}^{x=b} - \int_a^b B(\vec{u}'(x), \vec{v}'(x)) dx, \quad (\text{III, 11; 14})$$

где B — непрерывная билинейная форма. Эта формула непосредственно вытекает из того факта, что производной функции $x \rightarrow B(\vec{u}(x), \vec{v}(x))$ является функция $x \rightarrow B(\vec{u}'(x), \vec{v}(x)) + [B(\vec{u}(x), \vec{v}'(x))]$ (теорема 12). Здесь речь идет (относительно B) о непрерывном билинейном отображении $(\vec{u}, \vec{v}) \rightarrow \vec{u} \cdot \vec{v}$ или $\langle \vec{u}, \vec{v} \rangle$ из $\vec{F}' \times \vec{F}$ в \mathbb{R} . Естественно, это интегрирование по частям можно выполнить только в том случае, когда рассматриваемые функции u и v непрерывно дифференцируемы. Поэтому мы предполагали, что f_0 принадлежит классу C^2 . Поскольку \vec{L} принадлежит классу C^2 , $\partial_3 L$ принадлежит классу C^1 . Так как

f_0, \vec{f}'_0 принадлежат классу C^1 , то $u: x \rightarrow \overset{\leftarrow}{\partial}_3 L(x, f_0(x), \vec{f}'_0(x))$, согласно следствию 5 теоремы 11, также принадлежит классу C^1 .

Пример. Если $F = \mathbb{R}$, $\mathcal{U} = \mathbb{R} \times \mathbb{R}$, то формула (III, 11; 12₅), полученная в соответствии с теоремой 39, дает частный случай формулы (III, 11; 13):

$$\delta J = \left[\frac{\partial L}{\partial y'}(x, f_0(x), f'_0(x)) \delta f(x) \right]_{x=a}^{x=b} + \int_a^b \left[\frac{\partial L}{\partial y'}(x, f_0(x), f'_0(x)) - \frac{d}{dx} \left(\frac{\partial L}{\partial y'}(x, f_0(x), f'_0(x)) \right) \right] \delta f(x) dx.$$

Необходимые условия экстремума

Вернемся к первоначально поставленной проблеме экстремума. Необходимое условие того, что точка $f_0 \in \Omega_0$ (т. е. $f_0 \in \Omega$ при заданных $f_0(a) = \alpha$ и $f_0(b) = \beta$) доставляет экстремум функции J на открытом множестве Ω_0 из E_0 , заключается в том, что $J'(f_0)$ должна равняться нулю на \vec{E}_0 (теорема 23).

Теорема 40. Пусть $[a, b]$ — некоторый отрезок в \mathbb{R} , F — аффинное нормированное пространство, \mathcal{U} — открытое множество из $F \times \vec{F}$, Ω — множество непрерывно дифференцируемых отображений f из $[a, b]$ в F , таких, что $(f(x), f'(x)) \in \mathcal{U}$ для всех $x \in [a, b]$, Ω_0 — подмножество Ω , образованное функциями f , удовлетворяющими заданным условиям $f(a) = \alpha$, $f(b) = \beta$, и L — функция класса C^2 на $[a, b] \times \mathcal{U}$.

Для того чтобы функция $f_0 \in \Omega_0$ класса C^2 доставляла максимум или минимум на множестве Ω_0 интегралу J , определяемому формулой (III, 11; 3), или, более общо, являлась для него стационарной точкой, необходимо, чтобы функция f_0 была решением дифференциального уравнения 2-го порядка

$$\overset{\leftarrow}{\partial}^2 L(x, f_0(x), \vec{f}'_0(x)) - \frac{d}{dx} \left(\overset{\leftarrow}{\partial}_3 L(x, f_0(x), \vec{f}'_0(x)) \right) = \overset{\leftarrow}{0}, \quad (\text{III, 11; 15})$$

или

$$\frac{\partial L}{\partial y} - \frac{d}{dx} \left(\frac{\partial L}{\partial y'} \right) = \overset{\leftarrow}{0}. \quad (\text{III, 11; 16})$$

Это уравнение называется уравнением Эйлера, а его решение называется экстремалью интеграла J или функции L .

Доказательство. Если $\overset{\leftarrow}{\delta} f \in \vec{E}_0$, то $\overset{\leftarrow}{\delta} f(a) = \overset{\leftarrow}{\delta} f(b) = \overset{\leftarrow}{0}$. Поэтому в равенстве (III, 11; 13) скобка $[\]_{x=a}^{x=b}$ равна нулю:

Необходимое условие экстремума $\delta J = 0$ теперь говорит о том, что *интеграл*, фигурирующий в правой части равенства (III, 11; 13), равен нулю, какова бы ни была функция $\delta \vec{f} \in \vec{E}_0$. Для того чтобы отсюда получить, что функция \vec{f}_0 удовлетворяет дифференциальному уравнению (III, 11; 15), достаточно применить к функции Φ :

$$x \rightarrow \overset{\leftarrow}{\Phi}(x) = \overset{\leftarrow}{\partial_2 L}(x, f_0(x), \overset{\rightarrow}{f'_0}(x)) - \frac{d}{dx}(\overset{\leftarrow}{\partial_3 L}(x, f_0(x), \overset{\rightarrow}{f'_0}(x))) \quad (\text{III, 11; 17})$$

лемму Хаара.

Лемма Хаара

Если $\overset{\leftarrow}{\Phi}$ является непрерывной функцией на интервале $[a, b]$ со значениями в сопряженном пространстве $\overset{\leftarrow}{F}'$ некоторого векторного нормированного пространства \vec{F} и если для каждой непрерывно дифференцируемой функции η , определенной на $[a, b]$ со значениями в \vec{F} , обращающейся в нуль на концах этого интервала, имеет место равенство

$$\int_a^b \langle \overset{\leftarrow}{\Phi}(x), \overset{\rightarrow}{\eta}(x) \rangle dx = 0, \quad (\text{III, 11; 18})$$

то функция $\overset{\leftarrow}{\Phi}$ тождественно равна нулю на $[a, b]$.

Предположим, что $\overset{\leftarrow}{\Phi}$ не является тождественным нулем, и покажем, что тогда мы придем к противоречию. Пусть x_0 — точка, в которой $\overset{\leftarrow}{\Phi}(x_0) \neq 0$. Всегда можно считать точку x_0 отличной от концов a и b , ибо если непрерывная функция $\overset{\leftarrow}{\Phi}$ будет тождественно равна нулю в открытом интервале $]a, b[$, то она будет тождественным нулем и в замкнутом интервале. Утверждение «непрерывная линейная форма $\overset{\leftarrow}{\Phi}(x_0)$ отлична от 0 на \vec{F}' » равносильно существованию такой точки $\vec{e} \in \vec{F}$, что значение этой линейной формы на \vec{e} отлично от 0. При этом \vec{e} можно выбрать так, чтобы значение рассматриваемой линейной формы на \vec{e} было > 0 (в противном случае можно заменить \vec{e} на $-\vec{e}$). В силу непрерывности $\overset{\leftarrow}{\Phi}$, существует настолько малое число α , что интервал $[x_0 - \alpha, x_0 + \alpha]$ содержится в $[a, b]$ и непрерывная скалярная функция $x \rightarrow \langle \overset{\leftarrow}{\Phi}(x), \vec{e} \rangle$ на нем положительна

Пусть теперь φ — непрерывно дифференцируемая скалярная функция, положительная в $]x_0 - \alpha, x_0 + \alpha[$ и равная нулю вне этого интервала. Можно взять, например, функцию (рис. 9)

$$\varphi(x) = \begin{cases} 0, & \text{если } x \notin]x_0 - \alpha, x_0 + \alpha[, \\ \cos^2\left(\frac{\pi}{2\alpha}(x - x_0)\right) \text{ или } (\alpha^2 - (x - x_0)^2)^2, & \text{если } x \in]x_0 - \alpha, x_0 + \alpha[. \end{cases} \quad (\text{III}, 11; 19)$$

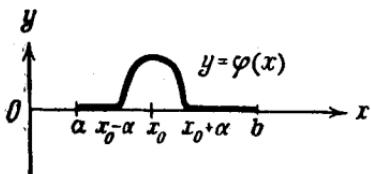


Рис. 9.

Возьмем в качестве $\vec{\eta}$ функцию $\vec{\eta}(x) = \vec{e}\varphi(x)$. Она, очевидно, непрерывно дифференцируема в $[a, b]$ и обращается в нуль на обоих концах. Функция $x \rightarrow \langle \vec{\Phi}(x), \vec{\eta}(x) \rangle$ непрерывна, >0 в интервале $]x_0 - \alpha, x_0 + \alpha[$ и равна нулю вне этого интервала. Ее интеграл, очевидно, >0 .

Мы пришли к противоречию: $\vec{\Phi}$ должна быть тождественно равной нулю, чем и заканчивается доказательство леммы.

Пример. Для того чтобы функция f_0 доставляла экстремум интегралу J среди всех функций $f \in \mathbb{R}^{[a, b]}$, принадлежащих классу C^1 (см. пример, приведенный после теоремы 38) и удовлетворяющих заданным граничным условиям $f(a) = \alpha, f(b) = \beta$, необходимо и достаточно, чтобы, если f_0 принадлежит классу C^2 , для любой функции $\delta f \in \mathbb{R}^{[a, b]}$ класса C^1 , удовлетворяющей условиям $\delta f(a) = \delta f(b) = 0$, имело место равенство

$$\int_a^b \left[\frac{\partial L}{\partial y'}(x, f_0(x), f'_0(x)) - \frac{d}{dx} \left(\frac{\partial L}{\partial y'}(x, f_0(x), f'_0(x)) \right) \right] \delta f(x) dx = 0. \quad (\text{III}, 11; 19a)$$

Лемма Хаара утверждает, что для этого необходимо и достаточно, чтобы функция f_0 удовлетворяла равенствам $f_0(a) = \alpha, f_0(b) = \beta$ и являлась решением дифференциального уравнения Эйлера

$$\frac{\partial L}{\partial y} - \frac{d}{dx} \left(\frac{\partial L}{\partial y'} \right) = 0. \quad (\text{III}, 11; 19b)$$

Следствие. Пусть f_0 — некоторая функция, определенная на $[a, b]$ и доставляющая при условиях теоремы 40 стационарное значение функционалу J . Тогда при $a_1 \geqslant a, b_1 \leqslant b$ функция

f_0 доставляет также стационарное значение интегралу

$$J_1(f) = \int_{a_1}^{b_1} L(x, f(x), \vec{f}'(x)) dx \quad (\text{III}, 11; 19_2)$$

на множестве всех функций f класса C^1 , определенных на $[a_1, b_1]$ со значениями в F , таких, что пара (f, \vec{f}') отображает $[a_1, b_1]$ в \mathcal{U} и удовлетворяет условиям $f(a_1) = \alpha_1 = f_0(a_1)$, $f(b_1) = \beta_1 = f_0(b_1)$.

В самом деле, уравнение Эйлера не использует концы интервала.

Этот результат не очевиден даже для случая абсолютного максимума или минимума. Предположим, например, что f_0 является точкой минимума для J . Тогда для заданной на $[a_1, b_1]$ функции f с описанными выше свойствами обозначим через \tilde{f} функцию, равную f в $[a_1, b_1]$ и f_0 в $[a, a_1[$ и $]b_1, b]$. Эта функция непрерывна, поскольку $f(a_1) = f_0(a_1)$, $f(b_1) = f_0(b_1)$. Если она принадлежит классу C^1 , то, в силу свойства минимальности, $J(\tilde{f}) \geq J(f_0)$, а, следовательно, вычитая из двух интегралов их

общую часть $\int_a^{a_1} + \int_{b_1}^b$, получим $J_1(f) \geq J_1(f_0)$ и f_0 по-преж-

нему будет доставлять минимум функционалу J_1 . Однако \tilde{f} в точках a_1 и b_1 имеет не обязательно равные правые и левые производные, а, значит, \tilde{f} не принадлежит классу C^1 и проведенное рассуждение не применимо. Для того чтобы сделать его пригодным, необходимо воспользоваться приближением \tilde{f} с помощью функций класса C^1 и выполнить предельный переход, который мы здесь не уточняем, но который показывает, независимо от уравнения Эйлера, что f_0 остается точкой минимума для J_1 .

Замечания. 1°) Заранее мы не можем быть уверенными в том, что функция f_0 , для которой интеграл достигает своего максимума или минимума, действительно существует. В самом деле, Ω_0 — открытое некомпактное множество (см. замечание 1°) к теореме 23).

2°) Если даже можно доказать, что элемент f_0 рассматриваемого открытого множества Ω_0 , на котором интеграл достигает своего максимального или минимального значения, существует, то это еще не означает, что этот элемент *дважды* непрерывно дифференцируем, а вторая производная может не быть связана с рассматриваемой задачей. В то время как на функцию L , входящую в условие задачи, можно наложить ограничения, относительно функции f_0 это сделать нельзя, поскольку она является *неизвестной*. Несколько более сложное

доказательство, чем проведенное нами выше, позволяет обойти эту трудность и показать в большинстве случаев, что функция класса C^1 , доставляющая функционалу J стационарное значение, *заведомо принадлежит классу C^2* .

3°) Практически обычно пишут дифференциальное уравнение (III, 11; 16) и ищут решение этого уравнения, удовлетворяющее *краевым условиям* $f(a) = \alpha$, $f(b) = \beta$. Относительно f получается дифференциальное уравнение второго порядка. Вместо того чтобы писать, как обычно, что функция f и ее производная f' принимают в некоторой заданной точке заданные начальные значения, пишут два условия другого рода, а именно что функция f принимает в двух заданных точках заданные значения. *Ничто не доказывает в общем случае ни существования решения, обладающего этими свойствами, ни его единственности.*

Возвратимся к формуле (III, 11; 4). Здесь δf является системой m приращений ($\delta f_1, \delta f_2, \dots, \delta f_m$). Частная производная $\frac{\partial L}{\partial y}(x, y, \vec{Y})$, как вектор из $\vec{F}' = \mathbb{R}^m$, является системой m производных $\frac{\partial L}{\partial y_i}(x, y_1, y_2, \dots, y_m, Y_1, Y_2, \dots, Y_m)$, $i = 1, \dots, m$. В свою очередь $\frac{\partial L}{\partial y'}(x, y, \vec{Y})$ есть система m производных $\frac{\partial L}{\partial y'_i}$ или $\frac{\partial L}{\partial Y_i}(x, y_1, \dots, y_m, Y_1, \dots, Y_m)$, $i = 1, 2, \dots, m$.

Наконец, скалярное произведение $\langle \vec{a}, \vec{X} \rangle$ или $\vec{a} \cdot \vec{X}$ для $\vec{a} = (a_1, a_2, \dots, a_m)$, $\vec{X} = (X_1, X_2, \dots, X_m)$ равно $a_1 X_1 + a_2 X_2 + \dots + a_m X_m$. Уравнение Эйлера, относящееся к функции $f = (f_1, f_2, \dots, f_m)$, эквивалентно системе m составляющих уравнений:

$$\frac{\partial L}{\partial y_i} - \frac{d}{dx} \left(\frac{\partial L}{\partial y'_i} \right) = 0, \quad i = 1, 2, \dots, m, \quad (\text{III, 11; 20})$$

или

$$\frac{\partial L}{\partial y_i} - \frac{\partial^2 L}{\partial y'_i \partial x} - \sum_{j=1}^m \frac{\partial^2 L}{\partial y'_i \partial y_j} y'_j - \sum_{j=1}^m \frac{\partial^2 L}{\partial y'_i \partial y'_j} y''_j = 0. \quad (\text{III, 11; 21})$$

Решение этой системы зависит от $2m$ произвольных постоянных. Они должны удовлетворять $2m$ условиям $f_i(a) = \alpha_i$, $f_i(b) = \beta_i$, $i = 1, 2, \dots, m$. Полезно переписать выражение для δJ при $\delta f \in \vec{E}_0$ в виде

$$\delta J = \int_a^b \left(\sum_{i=1}^m \left(\frac{\partial L}{\partial y_i} - \frac{d}{dx} \left(\frac{\partial L}{\partial y'_i} \right) \right) \delta f_i \right) dx. \quad (\text{III, 11; 21}_2)$$

4°) Мы записали условия для вариации первого порядка, не налагая никаких условий на вариации высших порядков (формула Тейлора). Поэтому остается неизвестным, идет ли речь о максимуме, минимуме или седле (см. замечание 2°) к теореме 23).

Простые случаи интегрируемости уравнений Эйлера

1°) Предположим, что L не зависит от y . Тогда $\frac{\partial L}{\partial y} = 0$. Уравнение сводится непосредственно к дифференциальному уравнению 1-го порядка

$$\overleftarrow{\frac{\partial L}{\partial y'}}(x, y') = \text{постоянная } c \in \overleftarrow{F'} \quad (\text{III, 11; 21}_3)$$

(если $F = \mathbb{R}$, то $\frac{\partial L}{\partial y'}(x, y') = c$ — вещественная постоянная).

2°) Предположим, что L не зависит от x . Уравнения Эйлера в частном случае (III, 11; 21) запишутся в виде

$$\frac{\partial L}{\partial y_i} - \sum_{j=1}^m \frac{\partial^2 L}{\partial y_j \partial y_i} y'_j - \sum_{j=1}^m \frac{\partial^2 L}{\partial y'_i \partial y'_j} y''_j = 0, \quad i = 1, 2, \dots, m. \quad (\text{III, 11; 21}_4)$$

Умножая i -е уравнение на y'_i и суммируя, получаем:

$$\sum_{i=1}^m \frac{\partial L}{\partial y_i} y'_i - \sum_{i, j=1}^m \frac{\partial^2 L}{\partial y_j \partial y'_i} y'_i y'_j - \sum_{i, j=1}^m \frac{\partial^2 L}{\partial y'_i \partial y'_j} y'_i y''_j = 0, \quad (\text{III, 11; 21}_5)$$

или

$$\frac{d}{dx} \left(L - \sum_{i=1}^m \frac{\partial L}{\partial y'_i} y'_i \right) = 0. \quad (\text{III, 11; 21}_6)$$

Отсюда находим первый интеграл

$$L - \sum_{i=1}^m \frac{\partial L}{\partial y'_i} y'_i = c, \quad \text{или} \quad L - \overleftarrow{\frac{\partial L}{\partial y'}} \vec{y}' = c, \quad (\text{III, 11; 21}_7)$$

который, как мы увидим ниже, имеет интересную интерпретацию в задачах механики (формула (III, 11; 96)).

В частности, если $m = 1$, то мы приходим к уравнению первого порядка

$$L - \frac{\partial L}{\partial y'} y' = c, \quad (\text{III, 11; 22})$$

в котором из-за умножения на y' введено постороннее решение $y = \text{const}$.

Пример. *Брахистохрона.* Вернемся к примеру, рассмотренному в начале этого параграфа. Требуется найти минимум интеграла (III, 11; 2). В нем z — независимая переменная, $x = f(z)$ — вещественная функция вещественной переменной z , а x' — ее производная. Уравнение Эйлера здесь записывается в соответствующих обозначениях в виде (формула (III, 11; 21₃)):

$$\frac{\partial L}{\partial x'} = \frac{x'}{\sqrt{1+x'^2}} \cdot \frac{1}{\sqrt{2g(z-a)}} = \text{const} = \frac{\pm 1}{\sqrt{4gc}}, \quad (\text{III}, 11; 23)$$

где $c > 0$ — некоторая произвольная постоянная. Мы отбросим случай постоянной, равной 0, которую невозможно записать в виде $\frac{\pm 1}{\sqrt{4gc}}$. Такая постоянная дает $x = \text{const} = a$. Это соответствует случаю, когда B лежит на одной вертикали с A и, конечно, решение дается вертикальным отрезком AB .

Рассматриваемое уравнение можно записать в виде

$$x'^2 = \frac{z-a}{2c} (1+x'^2), \text{ или } x'^2 \left(1 - \frac{z-a}{2c}\right) = \frac{z-a}{2c}. \quad (\text{III}, 11; 24)$$

Если сделать замену переменной (законную в силу неравенства $0 \leq \frac{z-a}{c} \leq 2$)

$$z = a + c(1 - \cos u), \quad (\text{III}, 11; 25)$$

то мы получим новое уравнение

$$\begin{aligned} \frac{dx}{dz} &= \pm \sqrt{\frac{z-a}{2c-(z-a)}} = \pm \sqrt{\frac{1-\cos u}{1+\cos u}}, \\ \frac{dx}{du} &= \pm c \sin u \sqrt{\frac{1-\cos u}{1+\cos u}} = \pm 2c \sin^2 \frac{u}{2} = \pm c(1-\cos u), \end{aligned} \quad (\text{III}, 11; 26)$$

для которого без труда найдем решение

$$z = a + c(1 - \cos u), \quad x = \pm c(u - \sin u) + \text{const}. \quad (\text{III}, 11; 27)$$

Постоянная c и дополнительная постоянная в (III, 11; 27) должны быть выбраны так, чтобы кривая прошла через A и B . Возьмем для простоты точку A в начале координат. Тогда для $u = 2k\pi$ имеем $z = a = 0$. В силу периодичности (замена u на $u - 2k\pi$) можно считать, что это условие выполняется при $u = 0$; при этом мы получим $x = 0$, если const из (III, 11; 27) возьмем равной нулю. В соотношениях (III, 11; 27) можно

взять знак +, ибо знак — может быть скомпенсирован заменой u на $-u$.

Итак, мы видим, что брахистохроной является циклоида Γ , имеющая в исходной точке A точку возврата с вертикальной касательной. Постоянная c должна быть определена так, чтобы кривая проходила через точку B . Поскольку c изменяется, мы имеем семейство цикloid Γ_c , гомотетичных кривой Γ_1 , соответствующей $c = 1$. Постоянная c геометрически может быть определена следующим образом: прямая AB пересекает Γ_1 в конечном числе точек C_1, C_2, \dots (кроме случая, когда AB горизонтальна и имеется бесчисленное множество решений). Если C_i — одна из них, то гомотетия кривой Γ_1 с центром гомотетии A и коэффициентом AB/AC_i проходит через B и является решением поставленной задачи (рис. 10).

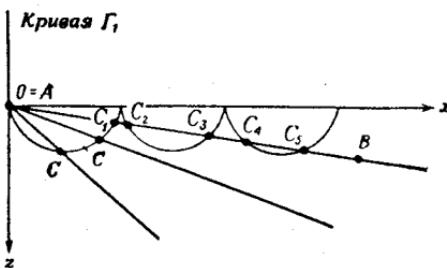


Рис. 10.

В этом примере сосредоточены почти все трудности вариационного исчисления.

1°) Когда мы ставили задачу в самом начале параграфа, мы предположили, что кривая находится в вертикальной плоскости, проходящей через A и B . Это ограничение может быть легко устранено. Выберем произвольно три координаты в \mathbb{R}^3 , считая ось z направленной вертикально вниз. Будем искать брахистохрону как дугу параметрической кривой класса C^2 , представленной тремя функциями $x(w)$, $y(w)$, $z(w)$. Будем предполагать, что точки A и B соответствуют фиксированным значениям $w = w_1$ и $w = w_2$. Тогда время, необходимое для движения от точки A к B , выразится в виде

$$t = \int_{w_1}^{w_2} \frac{ds}{\sqrt{2g(z-a)}} = \int_{w_1}^{w_2} \frac{\sqrt{x'^2 + y'^2 + z'^2}}{\sqrt{2g(z-a)}} dw, \quad (\text{III}, 11; 28)$$

где функции x , y , z принимают заданные значения (координаты точек A и B) при $w = w_1$, $w = w_2$.

Поскольку имеется 3 неизвестные функции, то мы имеем 3 уравнения Эйлера. Двумя первыми, в силу (III, 11; 21₃),

являются

$$\frac{x'}{\sqrt{x'^2 + y'^2 + z'^2}} \frac{1}{\sqrt{2g(z-a)}} = \text{const},$$

$$\frac{y'}{\sqrt{x'^2 + y'^2 + z'^2}} \frac{1}{\sqrt{2g(z-a)}} = \text{const}. \quad (\text{III}, 11; 29)$$

Отсюда получается соотношение с постоянными коэффициентами $\lambda x' + \mu y' = 0$, а затем $\lambda x + \mu y + v = 0$. Таким образом, кривая заведомо находится в вертикальной плоскости, проходящей, конечно, через A и B , и выбор двух новых координат в этой плоскости вполне закончен.

Впрочем, ничего не зная о вариационном исчислении, в этом результате можно было убедиться непосредственно. Возьмем произвольную кривую Γ , но систему координат выберем так, чтобы плоскость xz была вертикальной плоскостью, проходящей через A и B . Ортогональная проекция Γ_0 кривой на эту плоскость определится функциями $x(w)$, 0 , $z(w)$. Она пройдет также через A и B . Время падения t_0 вдоль этой проекции задается тем же самым интегралом (III, 11; 28), что и t , если заменить y' на 0 : это время строго меньше t , если $y'(w)$ не равно тождественно нулю, т. е. если кривая Γ не лежит в плоскости $y = 0$. Таким образом, минимум t может быть реализован только на кривой, лежащей в вертикальной плоскости, проходящей через A и B .

2°) С самого начала мы допустили, что брахистохона может быть представлена при помощи функции x переменной z . Это предположение ничем не оправдано, как это видно из окончательного результата. В самом деле, получив уравнение Эйлера, мы преобразовали его, совершив соответствующую замену переменной (III, 11; 25). Решениями преобразованного уравнения оказались циклоиды, в которых x не обязательно выражается как функция от z ! Если повторить построение, проведенное на стр. 385, то AB может пересечь Γ_1 в единственной точке C , расположенной на первой поднимающейся дуге циклоиды. При этом имеется единственное решение, соответствующее точкам A и B , и для этого решения x не выражается как функция от z . Как же исправить это положение? Мы можем действовать, как в п. 1°), отыскивая Γ в вертикальной плоскости, проходящей через A и B , как параметрическую кривую, принадлежащую классу C^2 .

Получается два уравнения Эйлера относительно x и z :

$$X = - \frac{d}{dw} \left(\frac{x}{\sqrt{x'^2 + z'^2}} \frac{1}{\sqrt{2g(z-a)}} \right) = 0,$$

или

$$\frac{x'}{\sqrt{x'^2 + z'^2}} \cdot \frac{1}{\sqrt{2g(z-a)}} = \text{const}, \quad (\text{III, 11; 30})$$

$$Z = -\frac{1}{2} \frac{\sqrt{x'^2 + z'^2}}{\sqrt{2g(z-a)}^{3/2}} - \frac{d}{dw} \left(\frac{z'}{\sqrt{x'^2 + z'^2}} \frac{1}{\sqrt{2g(z-a)}} \right) = 0.$$

При заданных A и B существует бесконечное множество решений этих уравнений, ибо *уравнения (III, 11; 30) не являются независимыми*. В самом деле, мы находимся в условиях п. 2°) на стр. 383 и существует комбинация $x'X + z'Z$ левых частей уравнений (III, 11; 30), приводящая к уравнению (III, 11; 21_b):

$$\frac{d}{dw} \left(L - x' \frac{\partial L}{\partial x'} - z' \frac{\partial L}{\partial z'} \right) = 0.$$

Поскольку функция L однородна и первой степени относительно x' , z' , то тождество Эйлера $x' \frac{\partial L}{\partial x'} + z' \frac{\partial L}{\partial z'} \equiv L$ показывает, что эта комбинация тождественно равна нулю. Впрочем, это и не удивительно: главное — это сама кривая Γ , а не ее параметризация. Все параметризации, эквивалентные одной и той же кривой Γ , дадут одно и то же время падения, и если одна из кривых экстремальна, то такими же будут и остальные. Это неудобство можно устранить следующим образом. Считать, что w представляет криволинейную абсциссу s для всех кривых невозможно, поскольку концы A и B всех этих кривых должны соответствовать одним и тем же значениям w_1 и w_2 переменной w , а кривые могут иметь разную длину. Однако удобно считать, что *параметр w является криволинейной абсциссой с искомой экстремальной кривой Γ* . Тогда, не имея возможности заменить в (III, 11; 28) $\sqrt{x'^2 + z'^2}$ на 1, это можно сделать в (III, 11; 30). Получится два упрощенных уравнения. Каждое решение первого уравнения обращает также в нуль комбинацию $x'X + z'Z$, а значит, и $z'Z$ и, следовательно, удовлетворяет также и второму уравнению, за исключением того случая, когда $z = \text{const}$ ($z = \text{const}$, $x' = 1$ является решением первого, а не второго уравнения). Итак, исключая постороннее решение $z = \text{const}$, получаем, что (III, 11; 30) эквивалентно уравнению

$$\frac{dx/ds}{\sqrt{2g(z-a)}} = \text{const}, \quad (\text{III, 11; 31})$$

что совпадает с (III, 11; 23) и снова дает, после соответствующей замены переменной, циклоиды. Таким образом, *несмотря на ложное допущение в самом начале, окончательный результат оказывается верным*. Замена переменных (III, 11; 25) восстанавливает положение.

Считая x функцией от z , мы только отбрасываем кривые $z = \text{const}$, которые не являются решениями задачи.

3°) Вспомним замечание 1°) к теореме 40. Мы говорили, что существование максимума или минимума может быть не обеспечено. В самом деле, рассмотрим задачу минимизации интеграла (III, 11; 2) функцией $x = f(z)$ класса C^1 . Единственные возможные решения — это найденные циклоиды. Если в условиях, указанных в п. 2°) на стр. 386, через A и B не проходит ни одна циклоида, позволяющая выразить x как функцию z , то искомый минимум не существует. Интеграл (III, 11; 2) на изученном множестве $\Omega_0 \subset E_0$ будет иметь точную нижнюю грань > 0 , но не минимум.

4°) В замечании 3°), стр. 382, мы говорили, что краевые условия $y(a) = \alpha$, $y(b) = \beta$ не обеспечивают единственности решения уравнения Эйлера. Возможно, что решений нет вовсе (в этом мы только что убедились) или этих решений имеется конечное или бесконечное (стр. 387) множество.

5°) Для брахистохроны не выполняются условия, при которых можно применить теорему 40, поскольку рассматриваемая функция $L = \sqrt{1+x'^2}/\sqrt{2g(z-a)}$ имеет особенность в точке $z = a$, являющейся одной из границ интегрирования! К несчастью, с подобными фактами часто приходится сталкиваться в математике. Доказываются прекрасные общие теоремы, которые почти не применимы к встречающимся частным случаям.

6°) Выше мы говорили, что если кривая удовлетворяет дифференциальному уравнению Эйлера, то она является экстремальной для интеграла с совершенно произвольными пределами. Это кажется противоречащим найденным результатам, а именно тому, что кривая заведомо имеет вертикальную касательную в начальной точке (это, впрочем, видно непосредственно из дифференциального уравнения: для $z = a$ имеем $x' = 0$). Объясняется это просто: кривая действительно является экстремальной для интеграла

$$\int_{a_1}^{b_1} \frac{\sqrt{1+x'^2}}{\sqrt{2g(z-a)}} dz, \quad (\text{III}, 11; 33)$$

каким бы ни были начало a_1 и конец b_1 , но всегда с одной и той же подинтегральной функцией, т. е. с одной и той же постоянной a в выражении $\sqrt{2g(z-a)}$.

Не обязательно рассматривать поставленную механическую задачу только при условии, что начальная скорость равна нулю. Та же циклоида с концами в A_1 и B_1 , где A_1 уже не является точкой возврата с вертикальной касательной, минимизирует интеграл (III, 11; 28), определяющий время падения точки вдоль

кривой, когда начальная скорость в точке A_1 равна $\sqrt{2g(a_1 - a)}$. $a_1 > a$. Эта величина соответствует скорости 0 в точке A . Эти замечания не должны обескуражить читателя. Они показывают только, что, рассматривая уравнения Эйлера, можно решить лишь небольшую часть поставленной проблемы о максимуме или минимуме. Чтобы справиться со всеми трудностями в полной общности, необходимо привлечь очень мощные методы, использующие теорию пространств Гильберта, теорему 30₂ гл. II, алгебраическую топологию, теорию дифференциальных уравнений и т. д.

Уравнение геодезических на поверхности

Пусть V_n — многообразие размерности n над полем вещественных чисел в аффинном евклидовом пространстве E_N конечной размерности N . Пусть \mathcal{C} — параметрическая кривая на V класса C^1 . Говорят, что она является *геодезической между точками A и B* , если длина дуги AB кривой \mathcal{C} минимальна или по крайней мере стационарна среди всех кривых класса C^1 , соединяющих A и B на многообразии V ¹⁾. Согласно следствию из теоремы 40, такая кривая остается геодезической между любыми двумя точками A_1, B_1 дуги AB и поэтому окончательно геодезической можно назвать кривую, сохраняющую это свойство для любых двух своих точек. Она является экстремальной относительно длины, т. е. является решением некоторого уравнения Эйлера.

Пусть $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ — карта многообразия V . Здесь \mathcal{O} — некоторое открытое множество из \mathbb{R}^n . Обозначим через $u \rightarrow M(u)$ отображение, определяющее эту карту. Будем предполагать, что V — многообразие класса C^3 , а, следовательно, M также принадлежит классу C^3 . Несколько позже мы покажем, что элемент дуги на V можно записать в таком виде:

$$ds = \|d\vec{M}\| = \left\| \sum_{i=1}^n \frac{\partial \vec{M}}{\partial u_i} du_i \right\| = \\ = \sqrt{\left(\sum_{i=1}^n \frac{\partial \vec{M}}{\partial u_i} du_i \middle| \sum_{i=1}^n \frac{\partial \vec{M}}{\partial u_j} du_j \right)} = \sqrt{\sum_{i,j=1}^n g_{ij}(u) du_i du_j}, \quad (\text{III, 11; 34})$$

где

$$g_{ij}(u) = \left(\frac{\partial \vec{M}}{\partial u_i}(u) \middle| \frac{\partial \vec{M}}{\partial u_j}(u) \right) \quad (\text{III, 11, 35})$$

¹⁾ Максимум длин кривых, соединяющих две точки, не существует: точная верхняя грань этих длин равна $+\infty$.

является вещественной функцией класса C^2 , определенной на множестве \mathcal{O} . Параметрическая кривая класса C^1 на V представляется функциями $w \rightarrow u_i(w)$. Так как функция $L = \sqrt{\sum_{i,j} g_{ij}(u) u'_i \cdot u'_j}$ принадлежит классу C^2 , то применима теорема 40, а уравнениями Эйлера будут являться уравнения

$$\frac{\partial}{\partial u_k} \left(\sqrt{\sum_{i,j} g_{ij}(u) u'_i u'_j} \right) - \frac{d}{dw} \left(\frac{\partial}{\partial u'_k} \sqrt{\sum_{i,j} g_{ij}(u) u'_i u'_j} \right) = 0. \quad (\text{III}, 11; 36)$$

В частном случае поверхности в евклидовом пространстве трех измерений после замены u_1 и u_2 на u и v получаем:

$$ds = \sqrt{E du^2 + 2F du dv + G dv^2},$$

где

$$E = \left\| \frac{\partial \vec{M}}{\partial u} \right\|^2, \quad F = \left(\frac{\partial \vec{M}}{\partial u} \mid \frac{\partial \vec{M}}{\partial v} \right), \quad G = \left\| \frac{\partial \vec{M}}{\partial v} \right\|^2. \quad (\text{III}, 11; 37)$$

Если искать кривую V , представляя v как функцию u , то уравнение Эйлера запишется в виде

$$\begin{aligned} \frac{\partial}{\partial v} \sqrt{E(u, v) + 2F(u, v)v' + G(u, v)v'^2} - \\ - \frac{d}{du} \left(\frac{F(u, v) + G(u, v)v'}{\sqrt{E(u, v) + 2F(u, v)v' + G(u, v)v'^2}} \right) = 0. \quad (\text{III}, 11; 38) \end{aligned}$$

Это — дифференциальное уравнение 2-го порядка с решением, зависящим от двух произвольных постоянных. Если ищется дуга геодезической кривой, соединяющей две точки A и B , то для определения этих двух постоянных имеются два уравнения.

Может случиться, что через две точки A и B не проходит никакой геодезической кривой. Так, например, если в \mathbb{R}^2 многообразие V представляет собой открытое множество $x^2 + y^2 > 1$, то через точки $A = (2, 0)$ и $B = (-2, 0)$ не проходит ни одной геодезической в этом многообразии. (Если заменить V его замыканием \bar{V} , представляющим собой множество точек $x^2 + y^2 \geqslant 1$, то найдется параметрическая дуга класса C^1 , соединяющая A и B в \bar{V} и реализующая минимум длины, и даже имеются две дуги, симметричные относительно AB и дающие ту же минимальную длину. Каждая из этих дуг составлена из отрезка касательной AC , проведенной из точки A к окружности $x^2 + y^2 = 1$, дуги CD этой окружности и отрезка касательной DB , проведенной из точки B . Однако эти кривые лежат в \bar{V} , а не в V , а \bar{V} не является многообразием. По этой причине теорема 40 не применима, так как мы имеем дело не с открытым

множеством E_0 , а с замкнутым. Дуга CD окружности уравнению Эйлера не удовлетворяет, ибо его решениями в данном случае являются прямые на плоскости \mathbb{R}^2 .) Можно показать, что если многообразие V компактно или, более общо, полно по евклидовой метрике, то любые две его точки могут быть всегда соединены хотя бы одной геодезической, обеспечивающей абсолютный минимум длины. Вообще говоря, имеется бесконечное множество геодезических, соединяющих A и B , и можно показать, что те из них, которые не дают абсолютного минимума длины, дают по крайней мере относительный минимум.

Рассмотрим, например, цилиндр вращения в \mathbb{R}^3 , определяемый уравнением

$$x^2 + y^2 = a^2. \quad (\text{III}, 11; 39)$$

Его можно «развернуть» на плоскость, т. е. можно найти параметрическое представление цилиндра, сохраняющее длины. Это — отображение H плоскости \mathbb{R}^2 на цилиндр: $(u, v) \rightarrow (x, y, z)$, где

$$\begin{aligned} x &= a \cos \frac{u}{a}, \\ y &= a \sin \frac{u}{a}, \\ z &= v. \end{aligned} \quad (\text{III}, 11; 40)$$

Отображение H сохраняет длины в том смысле, что $ds^2 = dx^2 + dy^2 + dz^2 = du^2 + dv^2$. Поэтому длина дуги кривой класса C^1 на плоскости \mathbb{R}^2 равна длине ее образа при отображении H , а геодезические цилиндры — это образы геодезических в \mathbb{R}^2 при отображении H , т. е. прямых на плоскости \mathbb{R}^2 .

Однако H не является биекцией. Если A — некоторая точка цилиндра, то ее прообразом является множество точек $A_n = A_0 + 2\pi an \vec{i}$, где \vec{i} — единичный вектор оси $0x$ в \mathbb{R}^2 . Если A и B — две точки цилиндра, то каждый отрезок прямой A_pB_q , $p \in \mathbb{Z}$, $q \in \mathbb{Z}$, будет иметь образом геодезическую дугу, соединяющую точки A и B . Поскольку A_pB_q и $A_{p+k}B_{q+k}$ имеют один и тот же образ, можно будет ограничиться выбором отрезков прямой A_0B_n , $n \in \mathbb{Z}$. Это также дает бесконечное множество геодезических дуг класса C^1 , соединяющих A и B . Только одна из них (за исключением двух симметричных) дает минимум длины. Все другие дают относительный минимум. В частности, если $B = A$, определяется бесконечное множество геодезических с началом и концом в A . Одна из них длины 0 дает минимум длины. Та же, которая является образом прямой A_0A_n , имеет длину, равную $2\pi|n|$. Она определяет относительный минимум и даже абсолютный минимум длин кривых

класса C^1 , соединяющих A с самой собой, причем алгебраическое число оборотов этих кривых вокруг оси $0z$ равно n (это термин из алгебраической топологии).

В качестве примера рассмотрим решение уравнения Эйлера для задачи о геодезических.

Пусть в \mathbb{R}^3 определена поверхность, образованная вращением вокруг оси $0z$ кривой, заданной уравнением $Z = F(r)$ в полу-полярных координатах. Элемент дуги на этой поверхности определяется формулой

$$ds = \sqrt{dr^2 + r^2 d\phi^2 + F'^2(r) dr^2}. \quad (\text{III}, 11; 41)$$

Если при этих условиях ищут геодезическую, выбирая ϕ как функцию r , то длина дуги кривой AB записывается в виде

$$\int_{r(A)}^{r(B)} \sqrt{1 + F'^2(r) + r^2 \phi'^2} dr. \quad (\text{III}, 11; 42)$$

Дифференциальное уравнение Эйлера имеет вид

$$\frac{r^2 \phi'}{\sqrt{1 + F'^2(r) + r^2 \phi'^2}} = \text{const} = k, \quad (\text{III}, 11; 43)$$

а его решение находится в квадратурах:

$$\phi = \int \frac{k \sqrt{1 + F'^2(r)}}{r \sqrt{r^2 - k^2}} dr + \phi_0. \quad (\text{III}, 11; 44)$$

Уравнение (III, 11; 43) можно записать также в виде

$$r^2 \frac{d\phi}{ds} = k, \quad \text{или} \quad r \left(\frac{r d\phi}{ds} \right) = k, \quad \text{или} \quad r \cos \Phi = k, \quad (\text{III}, 11; 45)$$

где Φ — угол между касательной к параллели (ориентированной в сторону возрастающих значений ϕ) и касательной к геодезической (ориентированной в направлении возрастания длины дуги). Этим формулам можно придать следующий геометрический смысл. Пусть точка описывает геодезическую кривую с единичной скоростью. Тогда производная по времени момента ее скорости по отношению к оси $0z$ равна моменту ее ускорения. Поскольку скорость равна 1, то ускорение направлено по главной нормали к кривой, т. е., как это мы увидим позже в силу геометрического свойства геодезических, по нормали к поверхности. Это значит, что вектор ускорения пересекает ось $0z$ и, следовательно, его момент относительно оси $0z$ равен нулю. Значит, момент скорости относительно оси $0z$ постоянен, что и выражается формулой (III, 11; 43).

Относительный экстремум

Теорема 41. Пусть в обозначениях теоремы 40 $M_i, i = 1, 2, \dots, m$, — функции, обладающие теми же свойствами, что и функция L , а k_i — заданные вещественные числа. Пусть K_i — функции, заданные на Ω_0 соотношениями:

$$f \rightarrow K_i(f) = \int_a^b M_i(x, f(x), \vec{f}'(x)) dx. \quad (\text{III}, 11; 46)$$

Предположим, что в точке $f_0 \in \Omega_0$ производные $K'_i(f_0)$ независимы, а $K_i(f_0) = k_i$. Для того чтобы функция f_0 обеспечивала максимум или минимум функции J на Ω_0 в множестве таких функций f из Ω_0 , что $K_i(f) = k_i, i = 1, 2, \dots, m$, необходимо и достаточно, чтобы существовали такие вещественные числа $\lambda_i, i = 1, 2, \dots, m$, при которых справедливы уравнения Эйлера

$$\frac{\overleftarrow{\delta L}}{\overleftarrow{dy}} - \frac{d}{dx} \left(\frac{\overleftarrow{\delta L}}{\overleftarrow{dy'}} \right) = \sum_{i=1}^m \lambda_i \left(\frac{\overleftarrow{\delta M_i}}{\overleftarrow{dy}} - \frac{d}{dx} \left(\frac{\overleftarrow{\delta M_i}}{\overleftarrow{dy'}} \right) \right). \quad (\text{III}, 11; 47)$$

Для доказательства достаточно применить теорию множителей Лагранжа (теорема 34).

Если F является n -мерным, то мы получаем систему n дифференциальных уравнений 2-го порядка, решения которых при известных λ_i зависят от $2n$ произвольных постоянных. Однако λ_i не известны. Поэтому общее решение этой системы зависит от $2n + m$ произвольных постоянных, которые нужно выбрать так, чтобы удовлетворить $2n$ краевым условиям и m условиям $K_i = k_i, i = 1, 2, \dots, m$.

Пример. Среди всех кривых класса C^1 , соединяющих точки A и B и имеющих заданную длину l , найти такую кривую, которая вместе с отрезком AB ограничивает максимальную площадь. Будем искать эту кривую, выражая y как функцию x и выбирая \vec{AB} в качестве оси x . Тогда длина выразится интегралом

$$l = \int_{x(A)}^{x(B)} \sqrt{1 + y'^2} dx, \quad (\text{III}, 11; 48)$$

а площадь¹⁾ будет вычисляться с помощью интеграла

$$S = \int_{x(A)}^{x(B)} y dx. \quad (\text{III}, 11; 49)$$

¹⁾ Так называемая «алгебраическая площадь». Для получения обычной площади можно ограничиться случаем $y \geqslant 0$.

Согласно сказанному выше, должен существовать такой множитель Лагранжа λ , при котором справедливо уравнение Эйлера

$$-\lambda \frac{d}{dx} \left(\frac{y'}{\sqrt{1+y'^2}} \right) = 1, \quad \text{или} \quad -\frac{\lambda y''}{(1+y'^2)^{3/2}} = 1. \quad (\text{III}, 11; 50)$$

Мы показали, что кривизна кривой должна быть постоянной, т. е. кривая является дугой окружности. Ее центр и радиус не известны, а следовательно, она зависит от трех произвольных постоянных. Если считать, что дуга искомой окружности проходит через точки A и B , находится над осью x ($y \geq 0$) и имеет заданную длину l , то окружность определится однозначно. Если окажется, что $l > \frac{\pi}{2} |AB|$, то искомая дуга окружности будет больше полуокружности и она не может быть явно представлена уравнением, выражающим y как функцию x . См. по этому поводу изложенное на стр. 386. Можно, естественно, предположить, что A и B совпадают, и тогда мы придем к следующему утверждению:

Среди всех замкнутых кривых заданной длины l класса C^1 максимальную площадь ограничивает окружность.

Радиус этой окружности равен, очевидно, $l/2\pi$, и, следовательно, ее площадь определяется формулой $S = l^2/4\pi$, что позволяет утверждать следующее:

Любая замкнутая кривая длины l класса C^1 ограничивает площадь, удовлетворяющую неравенству

$$S \leq \frac{l^2}{4\pi}, \quad (\text{III}, 11; 51)$$

при этом всегда имеет место строгое неравенство, за исключением того случая, когда кривая является окружностью.

Обратно, если некоторая замкнутая кривая класса C^1 ограничивает площадь S , то ее длина удовлетворяет неравенству

$$l \geq 2\sqrt{\pi S}, \quad (\text{III}, 11; 52)$$

причем всегда имеет место строгое неравенство, за исключением того случая, когда кривая является окружностью.

Повторяем еще раз, что мы не дали строгого решения задачи. Так, например, неравенства (III, 11; 51) и (III, 11; 52) полностью не обоснованы.

Замена переменных

Пусть F_1 — аффинное нормированное пространство, h — отображение $[a, b] \times F_1$ в F , принадлежащее классу C^1 . Если теперь f_1 — функция класса C^1 , определенная на $[a, b]$ со значе-

ниями в F_1 , то сложная функция f

$$x \rightarrow f(x) = h(x, f_1(x)) \quad (\text{III}, 11; 53)$$

является функцией класса C^1 , определенной на $[a, b]$ со значениями в F . Таким образом, h определяет отображение $f_1 \rightarrow f$ пространства $E_1 = (F_1^{[a, b]})_{cb; 1}$ в пространство $E = (F^{[a, b]})_{cb; 1}$. Мы сейчас убедимся в том, что если h принадлежит классу C^2 , то это отображение дифференцируемо.

Теорема 42. Отображение $f_1 \rightarrow f$ пространства $E_1 = (F_1^{[a, b]})_{cb; 1}$ в пространство $E = (F^{[a, b]})_{cb; 1}$, определенное формулой (III, 11; 53), в которой h — отображение $[a, b] \times F_1$ в F класса C^2 , дифференцируемо, а его производная в точке f_1 пространства $(F_1^{[a, b]})_{cb; 1}$ определяется по формуле:

$$\delta f_1 \rightarrow \delta f, \quad (\text{III}, 11; 54)$$

где

$$\vec{\delta f}(x) = \partial_2 h(x, f_1(x)) \cdot \vec{\delta f}_1(x).$$

Доказательство. Следует доказать два утверждения.

1°) Формула (III, 11; 54) для фиксированного f_1 определяет некоторое линейное непрерывное отображение $\vec{\delta f}_1 \rightarrow \vec{\delta f}$ пространства \vec{E}_1 в пространство \vec{E} . Поскольку h взято из класса C^2 , то $\partial_2 h$ принадлежит классу C^1 , и из следствия 5 теоремы 11, с одной стороны, и теоремы 12, с другой, сразу же вытекает, что если $\vec{\delta f}_1$ лежит в \vec{E}_1 , то $\vec{\delta f}$ принадлежит \vec{E} . Отображение $\vec{\delta f}_1 \rightarrow \vec{\delta f}$ очевидным образом линейно, а его непрерывность следует из оценок

$$\begin{aligned} \|\vec{\delta f}\|_0 &\leq (\sup_{a \leq x \leq b} \|\partial_2 h(x, f_1(x))\|) \|\vec{\delta f}_1\|_0, \\ \|\vec{\delta f}'\|_0 &\leq \sup_{a \leq x \leq b} \|\partial_1 \partial_2 h(x, f_1(x)) + \partial_2^2 h(x, f_1(x)) \cdot \vec{f}'_1(x)\| \|\vec{\delta f}_1\|_0 + \\ &\quad + \sup_{a \leq x \leq b} \|\partial_2 h(x, f_1(x))\| \|\vec{\delta f}'_1\|_0, \end{aligned} \quad (\text{III}, 11; 55)$$

дающих неравенство

$$\|\vec{\delta f}\|_1 \leq \text{const} \|\vec{\delta f}_1\|_1. \quad (\text{III}, 11; 56)$$

2°) Для приращения $\vec{\delta f}_1$ функции f_1 приращение $\vec{\Delta f}$ функции f определяется формулой:

$$\vec{\Delta f}(x) = h(x, f_1(x) + \vec{\delta f}_1(x)) - h(x, f_1(x)). \quad (\text{III}, 11; 57)$$

Естественно аппроксимировать его приращением $\vec{\delta f}$, определяемым формулой (III, 11; 54). Считая приращение $\vec{\delta f}$ определенным формулой (III, 11; 54), можно рассмотреть $\vec{\Delta f} - \vec{\delta f}$, а также $\vec{\Delta f}' - \vec{\delta f}'$. Для каждого из этих двух выражений можно применить формулу конечных приращений для фиксированного x , как это делалось при доказательстве теоремы 38, и одновременно применить усиленную теорему о равномерной непрерывности (см. примечание¹) на стр. 374). Можно показать также, что при заданном $\varepsilon > 0$ можно найти такое $\eta > 0$, что из $\|\vec{\delta f}_1\|_1 \leq \eta$ следует $\|\vec{\Delta f} - \vec{\delta f}\|_1 \leq \varepsilon \|\vec{\delta f}\|_1$. Это говорит о том, что $\|\vec{\Delta f} - \vec{\delta f}\|_1$ является бесконечно малой более высокого порядка по сравнению с нормой $\|\vec{\delta f}_1\|_1$, когда последняя стремится к нулю, а это означает, что $\vec{\delta f}$ является искомым дифференциалом.

Доказанная теорема обобщается следующим образом:

Если h принадлежит классу C^{n+h} на $[a, b] \times F_1$, то отображение $f_1 \rightarrow f$ пространства $(F_1^{[a, b]})_{cb; m}$ в пространство $(F^{[a, b]})_{cb; m}$ принадлежит классу C^k .

Приложение к задаче о геодезических

Пусть V_n — многообразие размерности n евклидова аффинного пространства E_N размерности N . Рассмотрим параметрические кривые класса C^1 , соединяющие точки A и B и расположенные или не расположенные на V . Будем предполагать, что такая кривая определяется отображением $w \rightarrow M(w)$ фиксированного отрезка $[a, b] \subset \mathbb{R}$ в E_N . Если через x_i обозначить координаты точки M в некотором ортонормированном базисе E_N , то x_i будут функциями w , а длина дуги AB кривой \mathcal{C} будет определяться интегралом

$$l = \int_a^b \sqrt{\sum x_i'^2} dw. \quad (\text{III, 11; 58})$$

Определим некоторую кривую \mathcal{C}_0 функцией $M(w)$, а соседнюю кривую \mathcal{C} функцией $M(w) + \vec{\delta M}(w)$. Дифференциал δl длины в условиях теоремы 39 определяется выражением

$$\delta l = \sum_{i=1}^N \int_a^b -\frac{d}{dw} \left(\sqrt{\sum_i x_i'^2} \right) \delta x_i(w) dw. \quad (\text{III, 11; 59})$$

Здесь величины $x'_i / \sqrt{\sum_i x_i^2}$ являются направляющими косинусами полукасательной к \mathcal{C}_0 в направлении возрастания дуг (или возрастания w). Пусть $\vec{t}(w)$ — единичный вектор этой полукасательной. Формула (III, 11; 59) может быть записана в виде

$$\delta l = - \int_a^b \left(\frac{d\vec{t}}{dw} \mid \vec{\delta M}(w) \right) dw. \quad (\text{III, 11; 60})$$

Согласно формуле Френе, на \mathcal{C}_0

$$\frac{d\vec{t}}{ds} = \frac{\vec{n}}{R}. \quad (\text{III, 11; 61})$$

Эта формула, доказываемая в соответствующих разделах дифференциальной геометрии в аффинном евклидовом трехмерном пространстве, пригодна и для произвольного числа измерений N : из равенства $(\vec{t} \mid \vec{t}) = 1$ дифференцированием получаем $(\frac{d\vec{t}}{ds} \mid \vec{t}) = 0$, а, следовательно, $d\vec{t}/ds$ является вектором нормали к кривой, а, значит, его можно обозначить через \vec{n}/R . Здесь \vec{n} — единичный вектор направления, называемого главной нормалью, а $R > 0$, по определению, — радиус кривизны. Этот радиус можно рассматривать как коэффициент растяжения при сферическом отображении; плоскость (\vec{t}, \vec{n}) (двумерная) обычно называется соприкасающейся.

Итак,

$$\delta l = - \int_a^b \left(\frac{\vec{n}}{R} \mid \vec{\delta M} \right) \frac{ds}{dw} dw = - \int_{s(a)}^{s(b)} \left(\frac{\vec{n}}{R} \mid \vec{\delta M} \right) ds. \quad (\text{III, 11; 62})$$

Здесь кривая \mathcal{C} расположена на многообразии V , а, следовательно, $\vec{\delta M}$ не является произвольной вариацией функции $w \rightarrow M(w)$. Рассмотрим настолько малую дугу геодезической кривой, чтобы она находилась в образе карты $\mathcal{O} \xrightarrow{\Phi} \Phi(\mathcal{O})$ многообразия V (в обозначениях, принятых на стр. 389). Кривая на \mathcal{O} будет определена отображением $w \rightarrow u(w)$ отрезка $[a, b]$ в \mathcal{O} , а образ кривой в V будет определяться сложной функцией $w \rightarrow M(u(w))$. Теорема 42 утверждает, что если V принадлежит классу C^2 , то дифференциал $\vec{\delta M}$, соответствующий вариации δu кривой, проведенной в \mathcal{O} , задается равенством

$$\vec{\delta M}(w) = M'(u(w)) \cdot \vec{\delta u}(w) = \sum_{j=1}^n \frac{\partial \vec{M}}{\partial u_j}(u(w)) \delta u_j(w). \quad (\text{III, 11; 63})$$

Но тогда теорема о сложной функции (теорема 11) утверждает, что дифференциал δl по отношению к приращениям δu_j получается из дифференциала (III, 11; 62) заменой в нем $\vec{\delta M}$ на дифференциал (III, 11; 63) по отношению к $\vec{\delta u}_j$. Поэтому искомый дифференциал запишется в виде

$$\delta l = - \int_{s(a)}^{s(b)} \sum_{j=1}^n \left(\vec{n} \left| \frac{\partial \vec{M}}{\partial u_j} \right. \right) \delta u_j \, ds, \quad (\text{III}, 11; 64)$$

где на этот раз δu_j являются произвольными скалярными функциями $w \rightarrow \delta u_j(w)$, определенными на $[a, b]$ класса C^1 и принимающими нулевые значения в точках a и b . Другими словами, $\vec{\delta u} = (\delta u_1, \delta u_2, \dots, \delta u_n)$ является произвольной функцией, определенной на $[a, b]$, со значениями в \mathbb{R}^n класса C^1 , обращающейся в нуль в точках a и b .

Теорема 40 (или лемма Хаара) теперь показывает, что геодезическая класса C^2 на V характеризуется уравнениями Эйлера

$$\left(\vec{n} \left| \frac{\partial \vec{M}}{\partial u_j} \right. \right) = 0, \quad j = 1, 2, \dots, n, \quad (\text{III}, 11; 65)$$

выражающими тот факт, что главная нормаль к геодезической кривой в каждой из своих точек является нормалью к многообразию или что соприкасающаяся плоскость перпендикулярна к многообразию.

Для произвольной кривой \mathcal{C}_0 на V дифференциал часто оставляют с небольшими видоизменениями в виде (III, 11; 62). Обозначим через γ угол между главной нормалью к \mathcal{C}_0 и нормалью к поверхности, отсчитываемый от 0 до π . Если теперь через \vec{n}_γ обозначить единичный вектор ортогональной проекции главной нормали на касательную гиперплоскость, то проекция вектора \vec{n} будет равна $\vec{n}_\gamma \sin \gamma$. Поскольку $\vec{\delta M}$ в каждой точке кривой \mathcal{C}_0 является вектором касательной гиперплоскости $\sum_{j=1}^n \frac{\partial \vec{M}}{\partial u_j} \delta u_j$, то его скалярное произведение на \vec{n} равно скалярному произведению на $\vec{n}_\gamma \sin \gamma$. Отсюда

$$\delta l = - \int_{s(a)}^{s(b)} \left(\vec{n}_\gamma \left| \vec{\delta M} \right. \right) ds, \quad (\text{III}, 11; 66)$$

где $R_\gamma = R/\sin \gamma$ является радиусом геодезической кривизны кривой \mathcal{C}_0 . Это — радиус кривизны ортогональной проекции кривой на касательную гиперплоскость. Геодезические линии — это кривые с бесконечным радиусом геодезической кривизны, или кривые с нулевой геодезической кривизной: $1/R_\gamma = 0$.

Поставим теперь такую задачу: на поверхности V_2 аффинного евклидова трехмерного пространства найти кривую \mathcal{C} заданной длины l , соединяющую две заданные точки A, B и вместе с заданной кривой Γ_0 , проходящей через эти точки, ограничивающую участок поверхности наибольшей площади.

Пусть кривая \mathcal{C} представляется функцией параметра w , принимающего в точках A и B заданные значения. Тогда можно определить дифференциал dl при переходе от кривой \mathcal{C}_0 к соседней кривой \mathcal{C} , соответствующей приращению $\vec{\delta M}(w)$. Дифференциал δS рассматриваемой площади является главной линейной частью «алгебраической» площади, заключенной между \mathcal{C}_0 и \mathcal{C} . Произведем интуитивную оценку δS . Пусть M_0 и P_0 — две соседние точки кривой \mathcal{C}_0 , соответствующие параметрам w_0 и $w_0 + dw$, с дугой M_0P_0 длины ds . В результате перемещения кривой они перейдут в точки M и P : $M = M_0 + \vec{\delta M}(w_0)$, $P = P_0 + \vec{\delta M}(w_0 + dw)$.

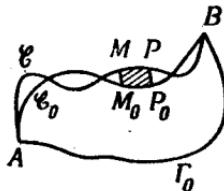


Рис. 11.

Если считать M_0P_0PM «параллелограммом», то его площадь будет равна произведению ds проекции $\vec{\delta M}$ на нормаль в точке M_0 к \mathcal{C}_0 , лежащую в плоскости, касательной к рассматриваемой поверхности. Единичный вектор этой нормали, с точностью до знака, совпадает с вектором, обозначенным нами через \vec{n}_γ .

Площадь параллелограмма равна, следовательно, $\pm (\vec{n}_\gamma | \vec{\delta M}) ds$, а вариация искомой площади δS есть

$$\delta S = \int_{s(a)}^{s(b)} \pm (\vec{n}_\gamma | \vec{\delta M}) ds, \quad (\text{III}, 11; 67)$$

где знак + или — в каждой точке \mathcal{C}_0 выбирается в зависимости от того, повернута ли кривая выпуклостью к Γ_0 или нет (нормаль \vec{n}_γ направлена всегда в сторону вогнутости).

Согласно теории множителей Лагранжа (теорема 41), должна существовать такая вещественная постоянная λ , что $\delta S - \lambda \delta l$ равна нулю на \mathcal{C}_0 , где

$$\int_{s(a)}^{s(b)} \left(\pm 1 - \frac{\lambda}{R_\gamma} \right) (\vec{n}_\gamma | \vec{\delta M}) ds = 0. \quad (\text{III}, 11; 68)$$

Естественно, $\vec{\delta M}$ не является произвольным приращением. Если (u, v) — параметры, определяющие карту поверхности, то согласно формуле (III, 11; 63)

$$\vec{\delta M} = \frac{\partial \vec{M}}{\partial u} \delta u + \frac{\partial \vec{M}}{\partial v} \delta v, \quad (\text{III}, 11; 69)$$

где δu и δv — произвольные приращения, равные нулю в A и B . Уравнениями Эйлера (см. лемму Хаара) будут

$$\left(\pm 1 - \frac{\lambda}{R_y} \right) \left(\vec{n}_y \mid \frac{\partial \vec{M}}{\partial u} \right) = 0, \quad \left(\pm 1 - \frac{\lambda}{R_y} \right) \left(\vec{n}_y \mid \frac{\partial \vec{M}}{\partial v} \right) = 0. \quad (\text{III}, 11; 70)$$

Поскольку вектор \vec{n}_y не может быть одновременно ортогональным к двум независимым векторам $\partial \vec{M}/\partial u$ и $\partial \vec{M}/\partial v$, эти уравнения эквивалентны уравнению

$$\pm 1 - \frac{\lambda}{R_y} = 0. \quad (\text{III}, 11; 71)$$

Так как $R_y \geq 0$, а λ постоянно, то сохраняется один из знаков $+$ или $-$. Таким образом, \mathcal{C}_0 является кривой постоянной геодезической кривизны, что обобщает результаты, полученные на стр. 394, и следующие за ними. Замкнутая кривая заданной длины, охватывающая максимальную площадь, имеет постоянную геодезическую кривизну.

Переменные концы. Условия трансверсальности

Предположим, что вместо того чтобы искать кривую класса C^1 , соединяющую точки A и B многообразия V некоторого аффинного евклидова пространства E и имеющую наименьшую длину (отыскание геодезической), мы бы начали искать кривую минимальной длины, соединяющую заданные точки подмногообразий \mathcal{A} и \mathcal{B} многообразия V . Эта задача является более общей, чем предыдущие, поскольку концы a и b интервала интегрирования и значения α и β функции f в этих точках в этом случае не заданы, а должны лишь удовлетворять некоторым соотношениям. Функция J теперь является функцией не только f , но также и точек a и b , относительно которых мы будем предполагать, что они описывают некоторый заданный компактный отрезок \mathbb{R}_1 множества \mathbb{R} .

Теорема 43. Пусть \mathbb{R}_1 — некоторый компактный отрезок из \mathbb{R} , F — аффинное нормированное пространство над полем вещественных чисел, \mathcal{U} — открытое множество из $F \times \tilde{F}$, L — вещественная функция на $\mathbb{R}_1 \times \mathcal{U}$ класса C^2 . Если Ω является

открытым множеством из $E = (F^{\mathbb{R}_1})_{cb; 1}$, образованным такими функциями f , при которых (f, \vec{f}') отображает \mathbb{R}_1 в \mathcal{U} , то функция J , определенная на $\Omega \times \mathbb{R}_1 \times \mathbb{R}_1$ по формуле (III, 11; 3), непрерывно дифференцируема и ее производная в f_0, a_0, b_0 , где f_0 принадлежит классу C^2 , определяется соотношением

$$\begin{aligned} (\vec{f}, \delta a, \delta b)^1) \rightarrow \delta J = & \int_a^b \left(\frac{\overleftarrow{\delta L}}{\partial y} - \frac{d}{dx} \left(\frac{\overleftarrow{\delta L}}{\partial y'} \right) \right) \cdot \vec{\delta f} dx + \\ & + \left(\frac{\overleftarrow{\delta L}}{\partial y'} (b, f_0(b), \vec{f}'_0(b)) \cdot \vec{\delta f}(b) + L(b, f_0(b), \vec{f}'_0(b)) \delta b \right) - \\ & - \left(\frac{\overleftarrow{\delta L}}{\partial y'} (a, f_0(a), \vec{f}'_0(a)) \cdot \vec{\delta f}(a) + L(a, f_0(a), \vec{f}'_0(a)) \delta a \right). \quad (\text{III, 11; 72}) \end{aligned}$$

Доказательство. Функция J в точке f имеет частную производную $\partial J / \partial f$, определяемую формулой (III, 11; 6). Значит,

$$\begin{aligned} \frac{\overleftarrow{\delta J}}{\partial y} (f_0, a_0, b_0) \cdot \vec{\delta f} = & \\ = & \int_a^b \left(\frac{\overleftarrow{\delta L}}{\partial y} (x, f_0(x), \vec{f}'_0(x)) \cdot \vec{\delta f}(x) + \frac{\overleftarrow{\delta L}}{\partial y'} (x, f_0(x), \vec{f}'_0(x)) \cdot \vec{\delta f}'(x) \right) dx. \quad (\text{III, 11; 73}) \end{aligned}$$

Частные производные J по a и b известны: это частные производные определенного интеграла по границам интервала интегрирования:

$$\begin{aligned} \frac{\partial J}{\partial a} (f_0, a_0, b_0) &= -L(a, f_0(a), \vec{f}'_0(a)), \\ \frac{\partial J}{\partial b} (f_0, a_0, b_0) &= L(b, f_0(b), \vec{f}'_0(b)^2). \quad (\text{III, 11; 74}) \end{aligned}$$

Для того чтобы получить отсюда, что J непрерывно дифференцируема на $\Omega \times \mathbb{R}_1 \times \mathbb{R}_1$, достаточно применить теорему 15. Частная непрерывность $\partial J / \partial f$ по f для фиксированных a и b была доказана в теореме 38. Дополняя соответствующим образом доказательство, легко убедиться (мы будем это считать доказанным), что $\partial J / \partial f$ непрерывна по f, a, b .

¹⁾ δa и δb являются вещественными числами и можно было бы их обозначить через da и db . Но поскольку по причинам, указанным на стр. 372, вариацию f мы обозначили через $\vec{\delta f}$ и, по аналогии, дифференциал J через δJ , мы будем также обозначать через δa и δb приращения для a и b .

²⁾ Здесь $a, b, J(f, a, b)$ — вещественные числа. Поэтому $\partial J / \partial a$ и $\partial J / \partial b$ — обычные производные вещественной функции вещественной переменной.

Другие типы непрерывности очевидны. Если f, a, b стремятся в $E, \mathbb{R}_1, \mathbb{R}_1$ соответственно к f_0, a_0, b_0 , то

$$\overrightarrow{f(b) - f_0(b_0)} = \overrightarrow{f(b) - f_0(b)} + \overrightarrow{f_0(b) - f_0(b_0)} \quad (\text{III}, 11; 75)$$

стремится к $\vec{0}$ в \vec{F} в силу равномерной сходимости f к f_0 в первом члене правой части (вытекающей из сходимости f к f_0 в пространстве E) и непрерывности f_0 в точке b_0 во втором члене. Точно так же, в силу равномерной сходимости \vec{f}' к \vec{f}'_0 и непрерывности \vec{f}'_0 в b_0 , разность $\vec{f}'(b) - \vec{f}'_0(b_0)$ стремится к $\vec{0}$. Поскольку L непрерывна в точке $(b_0, f_0(b_0), \vec{f}'_0(b_0))$ пространства $\mathbb{R}_1 \times F \times \vec{F}$, то разность $L(b, f(b), \vec{f}'(b)) - L(b_0, f_0(b_0), \vec{f}'_0(b_0))$ стремится к 0, чем и доказывается непрерывность $\partial J / \partial b$ в точке $(f_0, a_0, b_0) \in \Omega \times \mathbb{R}_1 \times \mathbb{R}_1$. Непрерывность $\partial J / \partial a$ доказывается аналогично.

Полный дифференциал δJ в (f_0, a_0, b_0) вычисляется теперь по формуле

$$\begin{aligned} \delta J = & \int_{a_0}^{b_0} \left(\frac{\partial L}{\partial y} \cdot \overrightarrow{\delta f} + \frac{\partial L}{\partial y'} \cdot \overrightarrow{\delta f}' \right) dx + \\ & + L(b_0, f_0(b_0), \vec{f}'_0(b_0)) \delta b - L(a_0, f_0(a_0), \vec{f}'_0(a_0)) \delta a. \end{aligned} \quad (\text{III}, 11; 76)$$

Если теперь в интеграле правой части полученного равенства выполнить интегрирование по частям, как это делалось в (III, 11; 13), то, поскольку $\overrightarrow{\delta f}$ не обязательно обращается в нуль в точках a_0 и b_0 , мы получим формулу (III, 11; 72).

Изменим это выражение, введя вместо $\overrightarrow{\delta f}(a_0), \overrightarrow{\delta f}(b_0)$ дифференциалы $\overrightarrow{\delta a}, \overrightarrow{\delta b}$ значений f в точках a и b .

Теорема 44. Пусть \mathbb{R}_1 — некоторый интервал \mathbb{R} , F — аффинное нормированное пространство, ξ — отображение $E = (F^{\mathbb{R}_1})_{cb; 1} \times \mathbb{R}_1$ в F , определенное равенством

$$\xi(f, x) = f(x) \in F, \quad (\text{III}, 11; 77)$$

где $f(x)$ — значение f в x^1). Это отображение непрерывно дифференцируемо, и его производная определяется по формуле

$$\xi'(f_0, x_0) \cdot (\overrightarrow{\delta f}, \overrightarrow{\delta x_0}) = \delta(f(x)) = \overrightarrow{\delta f}(x_0) + \vec{f}'(x_0) \cdot \delta x_0 \equiv \vec{F}. \quad (\text{III}, 11; 78)$$

¹⁾ Функция ξ подобна билинейному каноническому отображению $(u, \vec{X}) \rightarrow u \cdot \vec{X}$ пространства $\mathcal{L}(\vec{E} \times \vec{F}) \times \vec{E}$ в \vec{F} .

Доказательство. Для фиксированного x отображение $f \rightarrow f(x) \in F$ является аффинным непрерывным отображением E в F , а его линейным присоединенным отображением является $\vec{f} \rightarrow \vec{f}(x)$. Следовательно (теорема 8₂), ξ имеет частную производную по f , которая определяется формулой

$$\frac{\partial \xi}{\partial f}(f, x) \cdot \vec{\delta f} = \vec{\delta f}(x) \in \vec{F}. \quad (\text{III}, 11; 79)$$

Эта частная производная непрерывно зависит от f и x . В самом деле, из формулы конечных приращений (теорема 13)

$$\|\vec{\delta f}(x) - \vec{\delta f}(x_0)\| \leq \|\vec{\delta f}'\|_0 |x - x_0| \leq \|\vec{\delta f}'\|_1 |x - x_0| \quad (\text{III}, 11; 80)$$

следует неравенство

$$\begin{aligned} \left\| \frac{\partial \xi}{\partial f}(f, x) - \frac{\partial \xi}{\partial f}(f_0, x_0) \right\| &= \sup_{\|\vec{\delta f}\|_1 \leq 1} \left\| \left(\frac{\partial \xi}{\partial f}(f, x) - \frac{\partial \xi}{\partial f}(f_0, x_0) \right) \cdot \vec{\delta f} \right\| = \\ &= \sup_{\|\vec{\delta f}\|_1 \leq 1} \|\vec{\delta f}(x) - \vec{\delta f}(x_0)\| \leq |x - x_0|, \end{aligned} \quad (\text{III}, 11; 81)$$

доказывающее требуемую непрерывность. Для фиксированного f отображение $\xi: x \rightarrow f(x)$ является самой функцией f . Она дифференцируема, поскольку f принадлежит классу C^1 . Значит, отображение ξ имеет частную производную по x , определяемую формулой

$$\frac{\partial \xi}{\partial x}(f, x) = \vec{f}'(x) \in \vec{F}^1. \quad (\text{III}, 11; 82)$$

Производная функция $\vec{\partial \xi}/\partial x$ непрерывно зависит от f и x , поскольку при f , стремящемся к f_0 в $(F^R)_{cb, 1}$, и x , стремящемся к x_0 , $f'(x)$ стремится к $f'_0(x_0)$ (это рассуждение уже было проведено на стр. 402 с b, b_0 вместо x, x_0). Таким образом, ξ имеет непрерывные частные производные, и требуемый результат вытекает из теоремы 15.

Согласно (III, 11; 78), формулу (III, 11; 72) можно изменить, заменяя в ней $\vec{\delta f}(b)$ на $\vec{\delta \alpha} - \vec{f}'_0(b_0) \delta b$ и $\vec{\delta f}(a)$ на $\vec{\delta \alpha} - \vec{f}'_0(a_0) \delta a$. Тогда мы получим такой результат:

¹⁾ Поскольку x пробегает поле вещественных чисел, то производный вектор можно взять в виде $\frac{\partial \xi}{\partial x} = \frac{\partial \xi}{\partial x} \cdot 1$. См. замечание 2^o) после теоремы 8.

Следствие 1. Если в условиях теоремы 43 значения f в точках a и b обозначить через α и β , то производная J в точке (f_0, a_0, b_0) определяется формулой

$$\delta J = \int_{a_0}^{b_0} \left[\frac{\overleftarrow{dL}}{\partial y}(x, f_0(x), \dot{f}_0(x)) - \frac{d}{dx} \left(\frac{\overleftarrow{dL}}{\partial y'}(x, f_0(x), \dot{f}_0(x)) \right) \right] \cdot \overrightarrow{\delta f}(x) dx + \\ + \left[\frac{\overleftarrow{dL}}{\partial y'}(b_0, f_0(b_0), \dot{f}_0(b_0)) \cdot (\overrightarrow{\delta \beta} - \dot{f}_0(b_0) \delta b) + L(b_0, f_0(b_0), \dot{f}_0(b_0)) \delta b \right] - \\ - \left[\frac{\overleftarrow{dL}}{\partial y'}(a_0, f_0(a_0), \dot{f}_0(a_0)) \cdot (\overrightarrow{\delta \alpha} - \dot{f}_0(a_0) \delta a) + L(a_0, f_0(a_0), \dot{f}_0(a_0)) \delta a \right]. \quad (\text{III}, 11; 84)$$

Следствие 2. Предположим, что выполнены условия теоремы 43. Пусть a, b, α, β — заданные функции параметра t , пробегающего некоторое открытое множество аффинного нормированного пространства T , со значениями в $\mathbb{R}_1, \mathbb{R}_1, F, F$ соответственно. Для того чтобы интеграл $J(f, a(t), b(t))$ среди всех значений $t \in T$ и всех $f \in \Omega$, таких, что $f(a(t)) = \alpha(t), f(b(t)) = \beta(t)$, достигал своего максимального или минимального значения при f_0, t_0 , где $f_0(a(t_0)) = \alpha(t_0)$ и $f_0(b(t_0)) = \beta(t_0)$, необходимо, чтобы функция была решением уравнения Эйлера (III, 11; 16), а точка t_0 , кроме того, удовлетворяла так называемым условиям трансверсальности, т. е. обращала в нуль обе квадратные скобки выражения (III, 11; 84), в которых $\delta a, \delta b, \overrightarrow{\delta \alpha}, \overrightarrow{\delta \beta}$ являются дифференциалами функций a, b, α, β , соответствующими произвольному приращению δt параметра t в точке t_0 .

Доказательство. Проведем беглое нестрогое рассуждение. Прежде всего, среди всех функций $f \in \Omega$, удовлетворяющих условиям $f(a(t_0)) = \alpha(t_0)$ и $f(b(t_0)) = \beta(t_0)$, функция f_0 обеспечивает максимум или минимум значения функции $J(f, a(t_0), b(t_0))$, соответствующего фиксированным пределам $a(t_0)$ и $b(t_0)$. Следовательно, по теореме 14, она должна удовлетворять уравнению Эйлера (III, 11; 16). При этом дифференциал δJ в (III, 11; 84) не содержит интеграла и сводится к выражению, заключенному в квадратные скобки. Надо, кроме того, заметить, что в силу соотношений $f(a(t)) = \alpha(t), f(b(t)) = \beta(t)$, переменные f и t не являются независимыми. Однако мы можем рассматривать t как независимую переменную. Для фиксированного t функция f подчиняется некоторым соотношениям и потому остается в некотором аффинном подпространстве E_t пространства E (которое меняется вместе с t). Поскольку t является независимой переменной, а дифференциал δJ в (III, 11; 84) зависит только от приращения δt (через величины

$\delta a = \overleftarrow{a'}(t_0) \cdot \overrightarrow{\delta t}, \delta b, \overrightarrow{\delta \alpha}, \overrightarrow{\delta \beta}$ и не зависит от $\overrightarrow{\delta f}$ (поскольку f_0 удовлетворяет уравнению Эйлера), то необходимое условие экстремума сводится к тому, что $\delta J = 0$ при любом $\overrightarrow{\delta t}$, откуда и вытекает утверждение следствия.

Применение к геодезическим кривым

Пусть $(w, t) \rightarrow M(w, t)$ — отображение класса C^2 множества $[0, 1] \times T$ в многообразие V класса C^3 евклидова аффинного пространства E_N конечной размерности N . Здесь $[0, 1]$ — отрезок прямой \mathbb{R} , T — открытое множество аффинного нормированного пространства (практически $T = \mathbb{R}$ и $t \in T$ — время). Для каждого $t \in T$ частное отображение $w \rightarrow M(w, t)$ определяет параметрическую кривую \mathcal{C}_t класса C^2 на V с концами $A_t = M(0, t)$ и $B_t = M(1, t)$.

Выясним, как изменяется длина l_t кривой в зависимости от t . Дифференциал δl , соответствующий приращению $\overrightarrow{\delta t}$ в точке t_0 , задается формулой (III, 11; 84). Интеграл же сводится к (III, 11; 66). С другой стороны, здесь $a(t) \equiv 0$, $b(t) \equiv 1$, а, следовательно, $\delta a = \delta b = 0$, $a(t) = A_t$ и $b(t) = B_t$. Положим $\mathcal{C}_0 = \mathcal{C}_{t_0}$, $A_0 = A_{t_0}$, $B_0 = B_{t_0}$. Известно, что если в E_N выбрать некоторый ортонормированный базис, то компонентами $\overleftarrow{\partial L} / \partial y^i$

в этом базисе будут $x'_i(w) / \sqrt{\sum_{i=1}^N x_i'^2(w)}$, $i = 1, 2, \dots, N$, т. е.

направляющие косинусы $\cos \Phi_i(w)$ полукасательной \overrightarrow{t} (в направлении возрастающих w) к кривой \mathcal{C}_0 ¹) в точке w . Если через α_i , β_i мы обозначим координаты точек A и B , то первое выражение, заключенное в квадратные скобки в формуле (III, 11; 84), будет равно

$$\sum_{i=1}^N (\cos \Phi_i)_{w=1} \cdot \delta \beta_i = (\overrightarrow{t}(B_0) | \overrightarrow{\delta B}). \quad (\text{III, 11; 85})$$

Аналогичный результат имеет место в точке A . Поэтому окончательно получаем:

$$\delta l = l'(t_0) \cdot \overrightarrow{\delta t} = - \int_{s(A)}^{s(B)} \left(\frac{\overrightarrow{n_y}}{R_y} | \overrightarrow{\delta M} \right) ds + (\overrightarrow{t}(B_0) | \overrightarrow{\delta B}) - (\overrightarrow{t}(A_0) | \overrightarrow{\delta A}). \quad (\text{III, 11; 86})$$

¹⁾ Не следует путать параметр t с единичным вектором \overrightarrow{t} касательной к \mathcal{C}_0 .

В этой формуле $\vec{\delta M}$ равен $\frac{\partial M}{\partial t}(w, t_0) \cdot \vec{\delta t}$, в то время как $\vec{\delta A}, \vec{\delta B}$ соответствуют частным значениям $w = 0, w = 1$.

В частности, если кривая является геодезической кривой из V , зависящей от параметра t , или, проще, если для $t = t_0$ кривая \mathcal{C}_0 является геодезической кривой из V , то имеет место соотношение

$$\delta l = (\vec{t}(B_0) | (\vec{\delta B})) - (\vec{t}(A_0) | (\vec{\delta A})). \quad (\text{III, 11; 87})$$

Нами получен замечательный результат:

Вариация длины геодезической удовлетворяет той же самой формуле, что и вариация длины отрезка прямой в евклидовом пространстве.

Пусть, например, на V отыскивается кривая минимальной длины, соединяющая точку A из \mathcal{A} и точку B из \mathcal{B} , где \mathcal{A} и \mathcal{B} — заданные подмногообразия V (см. стр. 400). Многообразия \mathcal{A} и \mathcal{B} можно представить параметрически, по крайней мере локально, с помощью двух параметров ξ, η , пробегающих открытые множества аффинных пространств Ξ и H . Тогда T будет представлять собой множество $\Xi \times H$, а t — пару (ξ, η) . Согласно следствию 2, кривая, прежде всего, должна быть геодезической, а ее концы A и B должны быть выбраны на \mathcal{A} и \mathcal{B} так, чтобы (III, 11; 87) обращалось в нуль при любом $\vec{\delta t}$, т. е. при любых $\vec{\delta \xi}, \vec{\delta \eta}$. Это означает, что *геодезическая кривая должна быть нормалью к \mathcal{A} в точке A и нормалью к \mathcal{B} в точке B* .

Канонические уравнения Гамильтона

Рассмотрим такие задачи вариационного исчисления, в которых F является пространством \mathbb{R}^m , так что функция f эквивалентна системе m функций f_1, f_2, \dots, f_m переменной x . Вводя вспомогательные функции $z_i = y'_i$, можно уравнения Эйлера записать в виде $2m$ дифференциальных уравнений *первого порядка*. Рассматривая L как заданную функцию $2m + 1$ переменных x, y_i, z_i , получаем:

$$\begin{aligned} y'_i &= z_i, \quad i = 1, 2, \dots, m, \\ \frac{\partial L}{\partial y_i} - \frac{d}{dx} \left(\frac{\partial L}{\partial z_i} \right) &= 0, \quad i = 1, 2, \dots, m. \end{aligned} \quad (\text{III, 11; 88})$$

Произведем теперь следующую замену неизвестных функций: вместо функций y_i и z_i введем новые неизвестные функции q_i и p_i , где $q_i = y_i$, а p_i заданы как функции от x, y_i и

z_i формулой

$$p_i = \frac{\partial L}{\partial z_i}. \quad (\text{III}, 11; 89)$$

Будем считать, что здесь речь идет о настоящей замене функций, т. е. формула (III, 11; 89) позволяет выразить каждую из переменных z_i как функцию x, y_j и p_j . С другой стороны, вместо заданной функции L переменных x, y_i и z_i введем функцию H по формуле

$$H(x, q_1, q_2, \dots, q_m, p_1, p_2, \dots, p_m) = \sum_{i=1}^m p_i z_i - L. \quad (\text{III}, 11; 90)$$

Функцию H переменных x, q_i и p_i называют *функцией Гамильтона*¹⁾. Ее дифференциал выражается следующим образом (частные производные L берутся по x, y_i, z_i):

$$\begin{aligned} dH = & \sum_{i=1}^m (p_i dz_i + z_i dp_i) - \frac{\partial L}{\partial x} dx - \sum_{i=1}^m \frac{\partial L}{\partial y_i} dy_i - \sum_{i=1}^m \frac{\partial L}{\partial z_i} dz_i = \\ & = \sum_{i=1}^m z_i dp_i - \sum_{i=1}^m \frac{\partial L}{\partial y_i} dq_i - \frac{\partial L}{\partial x} dx. \end{aligned} \quad (\text{III}, 11; 91)$$

Эта формула верна как в старой системе переменных (x, y_i, z_i) , так и в новой (x, q_i, p_i) . Это показывает, что частные производные H по x, q_i, p_i определяются равенствами:

$$\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}, \quad \frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial y_i}, \quad \frac{\partial H}{\partial p_i} = z_i. \quad (\text{III}, 11; 92)$$

При этих условиях дифференциальные уравнения Эйлера (III, 11; 88) выражаются для функций q_i и p_i переменной x в следующей простой форме:

$$q'_i = \frac{\partial H}{\partial p_i}, \quad p'_i = -\frac{\partial H}{\partial q_i}. \quad (\text{III}, 11; 93)$$

Итак, имеет место

Теорема 45. Если в условиях теоремы 40 F является пространством \mathbb{R}^m , произведена замена переменных и функций $(x, y_i, z_i) \rightarrow (x, q_i, p_i)$ по формулам (III, 11; 89) и введена функция Гамильтона H по формуле (III, 11; 90), то, предполагая, что H является функцией x, q_i, p_i , уравнения Эйлера для функций

¹⁾ В то время как функция L , выраженная через x, y_i, z_i , называется функцией Лагранжа.

²⁾ Заметим, что dL/dx является частной производной в системе переменных (x, y_i, z_i) , в то время как dH/dx является частной производной в системе переменных (x, q_i, p_i) .

q_i и p_i переменной x можно записать в виде (III, 11; 93). Эти уравнения называются уравнениями Гамильтона относительно функции H .

Замечание. Мы положили $F = \mathbb{R}^m$ только ради простоты. Предположим, что F конечномерно. Всегда можно вместо функций y и $z = \vec{y}'$, определенных на $[a, b]$, со значениями в F и \vec{F} соответственно взять новые неизвестные функции $q = y$ и $p = \vec{\partial L}/\partial z$, определенные на $[a, b]$, со значениями соответственно в F и \vec{F}' . Функция Гамильтона H является вещественной функцией на $[a, b] \times F \times \vec{F}'$ и определяется равенством

$$H(x, q, \overset{\leftarrow}{p}) = \langle \overset{\leftarrow}{p}, \vec{z} \rangle - L(x, y, \vec{z}), \quad (\text{III, 11; 94})$$

где $\langle \overset{\leftarrow}{p}, \vec{z} \rangle \in \mathbb{R}$ — скалярное произведение $\overset{\leftarrow}{p} \in \vec{F}'$ на $\vec{z} \in \vec{F}$. При этом предполагается, что y и \vec{z} заменены их значениями как функции q и p .

Для неизвестных функций q и $\overset{\leftarrow}{p}$ переменной x уравнения Гамильтона будут иметь вид:

$$\frac{\overset{\rightarrow}{dq}}{dx} = \frac{\partial H}{\partial p}, \quad \frac{\overset{\leftarrow}{dp}}{dx} = - \frac{\overset{\leftarrow}{\partial H}}{\partial q}. \quad (\text{III, 11; 95})$$

При фиксированном $\overset{\leftarrow}{p}$, H является вещественной функцией $\overset{\leftarrow}{q} \in \vec{F}$, а, следовательно, $-\frac{\partial H}{\partial \vec{q}}(x, q, \overset{\leftarrow}{p}) \in \mathcal{L}(\vec{F}; \mathbb{R}) = (\vec{F}')'$; при этом требуется, чтобы выражение было равным $\overset{\leftarrow}{dp}/dx$. Точно так же $\frac{\overset{\leftarrow}{\partial H}}{\partial p}(x, q, \overset{\leftarrow}{p}) \in \mathcal{L}(\vec{F}', \mathbb{R}) = (\vec{F}')'$. Так как пространство $(\vec{F}')'$, сопряженное к \vec{F}' , совпадает с \vec{F} , то $\frac{\overset{\leftarrow}{\partial H}}{\partial p}(x, q, \overset{\leftarrow}{p}) \in \vec{F}$ и требуется, чтобы последнее выражение было равным $\overset{\rightarrow}{dq}/dx$. Функция $\overset{\leftarrow}{p}$ называется «сопряженным моментом» функции q . Это название объясняется тем, что q принимает значения в F , а $\overset{\leftarrow}{p}$ — в \vec{F}' .

Предположим, в частности, что L явно не зависит от x . Тогда H не будет явно зависеть от x , и предыдущие уравнения означают, что величина

$$\frac{dH}{dx} = \sum \frac{\partial H}{\partial q_i} q'_i + \sum \frac{\partial H}{\partial p_i} p'_i \quad (\text{III, 11; 96})$$

тождественно равна нулю. Другими словами, H является первым интегралом для экстремальных кривых. Вдоль любой экстремальной кривой функция H постоянна. На это мы указывали в (III, 1; 21₇).

Применения к механике

Уравнения Гамильтона имеют исключительно важные приложения как в механике, так и в теоретической физике.

Рассмотрим для простоты механическую задачу о движении системы материальных точек с наложенными на нее определенными идеальными связями, находящейся под действием поля потенциальных сил, не зависящих от времени. В этой задаче положение системы может быть, например, описано с помощью конечного числа параметров q_1, q_2, \dots, q_m . Для этой системы точек можно подсчитать потенциальную энергию U , являющуюся известной функцией параметров. С другой стороны, всегда возможно вычислить кинетическую энергию $T = \sum \frac{mv^2}{2}$, являющуюся квадратичной формой относительно первых производных q'_1, q'_2, \dots, q'_m ($q'_i = dq_i/dt$), коэффициенты которой сами зависят от q_1, q_2, \dots, q_m . Решение этой задачи дает «траектории», каждая из которых определяется некоторыми функциями t : $t \rightarrow q_i(t)$. Траектории рассматриваемой задачи являются решением некоторой экстремальной задачи. Для двух определенных моментов t_1 и t_2 вдоль действительной траектории или произвольных «фиктивных» траекторий можно рассмотреть интеграл

$$\int_{t_1}^{t_2} L(q_i, q'_i) dt = \int_{t_1}^{t_2} (T(q_i, q'_i) - U(q_i)) dt \text{ } ^1). \quad (\text{III, 11; 97})$$

Нетрудно доказать, что среди всех траекторий, проходящих в заданные начальный и конечный моменты t_1 и t_2 через одни и те же точки $q_i(t_1), q_i(t_2)$, реальной траекторией является та, при которой интеграл (III, 11; 97) достигает своего стационарного значения.

Другими словами, каждая траектория является некоторой экстремалью, а уравнение, определяющее траектории рассматриваемой механической задачи, является системой уравнений Эйлера:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial q'_i} \right) = 0, \quad i = 1, 2, \dots, m, \quad (\text{III, 11; 98})$$

¹⁾ $L(q_i, q'_i)$ здесь означает $L(q_1, q_2, \dots, q_m, q'_1, q'_2, \dots, q'_m)$.

или

$$\frac{\partial T}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial T}{\partial q'_i} \right) = \frac{\partial U}{\partial q_i}.$$

В этом случае они называются *уравнениями Лагранжа*, а функция $L = T - U$ — *функцией Лагранжа*. Переход к переменным Гамильтона определяется формулами

$$p_i = \frac{\partial L}{\partial q'_i} = \frac{\partial T}{\partial q'_i}. \quad (\text{III, 11; 99})$$

Функция Гамильтона H принимает вид

$$H(q_i, p_i) = \sum_{i=1}^m q'_i \frac{\partial T}{\partial q'_i} - L. \quad (\text{III, 11; 100})$$

Однако если учесть тот факт, что T является квадратичной формой относительно q'_i , то тождество Эйлера относительно однородных функций приведет к формуле

$$\sum_{i=1}^m q'_i \frac{\partial T}{\partial q'_i} = 2T. \quad (\text{III, 11; 101})$$

Если теперь выразить H в виде

$$H = 2T - (T - U) = T + U, \quad (\text{III, 11; 102})$$

то мы увидим, что H является *энергией системы* (сумма потенциальной и кинетической энергий), выраженной как функция q_i и p_i . Если H выразить как функцию q_i и p_i , то уравнения Гамильтона механической задачи будут являться системой уравнений (III, 11; 93).

Функция Гамильтона H не зависит от времени и, следовательно, является первым интегралом системы. Иначе говоря, *вдоль траектории системы энергия H , сумма кинетической и потенциальной энергий, остается постоянной* — это один из основных законов классической механики.

Вариационное исчисление для кратных интегралов

Изучим следующую задачу. Рассмотрим компактную кривую \mathcal{C} класса C^1 в аффинном евклидовом трехмерном пространстве над полем вещественных чисел и среди всех поверхностей \mathcal{S} класса C^1 , ограниченных этой кривой¹⁾, будем искать поверхность

¹⁾ Мы здесь не будем пытаться уточнять смысл выражения «поверхность \mathcal{S} , ограниченная кривой \mathcal{C} ». Оно определено в гл. VI. Скажем только, что \mathcal{S} и \mathcal{C} являются многообразиями (в смысле § 9) размерностей 2 и 1, что у них нет общих точек и что $\mathcal{S} = \mathcal{P} \cup \mathcal{C}$.

наименьшей площади (задача на минимальную площадь). Если координаты точек пространства обозначить через (x, y, z) и считать, что искомая поверхность выражается в виде функции z от x и y , то площадь рассматриваемой поверхности выразится в виде двойного интеграла:

$$S = \iint_{\Sigma} V \sqrt{1 + p^2 + q^2} dx dy, \quad p = \frac{\partial z}{\partial x}, \quad q = \frac{\partial z}{\partial y}. \quad (\text{III, 11; 103})$$

распространенного на область Σ плоскости xy , ограниченную кривой Γ , где Σ — проекция на плоскость xy поверхности \mathcal{S} , а Γ — проекция кривой \mathcal{C} . Мы пришли к задаче об отыскании функции $z = f(x, y)$ класса C^1 , принимающей заданные значения вдоль контура Γ плоскости \mathbb{R}^2 (для того чтобы поверхность \mathcal{S} проходила через кривую \mathcal{C}) и такой, что выписанный выше интеграл принимает минимальное значение.

Рассмотрим в пространстве \mathbb{R}^n открытую область \mathcal{O} , ограниченную компактной гиперповерхностью Γ класса C^1 . Пусть F — аффинное нормированное пространство и \mathcal{U} — открытое множество из $F \times \vec{F}^n$. Пусть теперь L — вещественная функция, определенная на $\bar{\mathcal{O}} \times \mathcal{U}$ класса $C^{2,1}$. Мы будем обозначать ее через

$$L(x_1, x_2, \dots, x_n, z, \vec{p}_1, \vec{p}_2, \dots, \vec{p}_n).$$

Если f является отображением класса C^1 области $\bar{\mathcal{O}}$ в F , то оно имеет частные производные $\vec{p}_i = \vec{\partial f} / \partial x_i$, являющиеся функциями, определенными на $\bar{\mathcal{O}}$, со значениями в \vec{F} . Если образ области $\bar{\mathcal{O}}$ при отображении $(f, \vec{p}_1, \dots, \vec{p}_n)$ принадлежит открытому множеству \mathcal{U} в $F \times \vec{F}^n$, то можно рассмотреть кратный интеграл

$$\begin{aligned} J(f) &= \iint_{\mathcal{O}} \dots \int L(x_1, \dots, x_n, f(x_1, \dots, x_n), \vec{p}_1(x_1, \dots, x_n), \dots \\ &\quad \dots, \vec{p}_n(x_1, \dots, x_n)) dx_1 \dots dx_n = \\ &= \iint_{\mathcal{O}} \dots \int L(x_i, f, \vec{p}_i) dx_1 \dots dx_n \quad (\text{III, 11; 104}) \end{aligned}$$

¹⁾ $\bar{\mathcal{O}}$ является некоторым компактом, а не открытым множеством \mathbb{R}^n . В начале § 2, стр. 195, определялась производная функции на компактном в \mathbb{R} интервале $[a, b]$. Точно так же это можно сделать для функции, заданной на компакте $\bar{\mathcal{O}} \subset \mathbb{R}^n$. В дальнейшем в теореме 28 гл. VI мы определим смысл выражения «открытое множество \mathcal{O} , ограниченное компактной гиперповерхностью Γ класса C^1 ».

и поставить следующую задачу: среди всевозможных функций f , принимающих вдоль контура Γ заданные значения, найти ту, при которой интеграл J принимает наименьшее или наибольшее значение.

Для решения этой задачи достаточно провести рассуждения, весьма близкие к тем, которые проводились в случае простого интеграла. При этом получается следующий результат:

Теорема 46. Функция $J: f \rightarrow Jf$ принадлежит классу C^1 , а ее дифференциал задается формулой

$$\delta J = \int \int \dots \int \left(\frac{\overleftarrow{\delta L}}{\partial z} \cdot \overrightarrow{\delta f} + \sum_{i=1}^n \frac{\overleftarrow{\delta L}}{\partial p_i} \cdot \overrightarrow{\delta p_i} \right) dx_1 \dots dx_n, \quad (\text{III}, 11; 105)$$

$$\delta p_i = \frac{\partial}{\partial x_i} \delta f.$$

Если L и f_0 принадлежат классу C^2 и рассматривается подмножество функций f , принимающих заданные значения на контуре Γ , то этот дифференциал записывается в виде

$$\delta J = \int \int \dots \int \left(\frac{\overleftarrow{\delta L}}{\partial z} - \sum_{i=1}^n \frac{d}{dx_i} \left(\frac{\overleftarrow{\delta L}}{\partial p_i} \right) \right) \overrightarrow{\delta f} dx_1 \dots dx_n, \quad (\text{III}, 11; 106)$$

где $\frac{d}{dx_i} \left(\frac{\overleftarrow{\delta L}}{\partial p_i} \right)$ означает частную производную по x_i сложной функции

$(x_1, \dots, x_n) \rightarrow$

$$\rightarrow \left(\frac{\overleftarrow{\delta L}}{\partial p_i} \right) (x_1, \dots, x_n, f_0(x_1, \dots, x_n), \frac{\partial f}{\partial x_1}(x_1, \dots, x_n), \dots, \frac{\partial f}{\partial x_n}(x_1, \dots, x_n)).$$

Для того чтобы среди всех функций f класса C^1 , принимающих на контуре заданные значения, функция f_0 класса C^2 минимизировала или максимизировала интеграл J , необходимо, чтобы f_0 удовлетворяла дифференциальным уравнениям Эйлера в частных производных 2-го порядка:

$$\frac{\overleftarrow{\delta L}}{\partial z} - \sum_{i=1}^n \frac{d}{dx_i} \left(\frac{\overleftarrow{\delta L}}{\partial p_i} \right) = 0. \quad (\text{III}, 11; 107)$$

Доказательство. Единственное место в доказательстве, не являющееся полностью аналогичным тому, что было сделано для простых интегралов, — это переход от формулы (III, 11; 105) к формуле (III, 11; 106).

Запишем (III, 11; 105) в виде

$$\delta J = \int \int \dots \int \left(\frac{\overleftarrow{\partial L}}{\partial z} \cdot \overrightarrow{\delta f} + \sum_{i=1}^n \frac{d}{dx_i} \left(\frac{\overleftarrow{\partial L}}{\partial p_i} \cdot \overrightarrow{\delta f} \right) - \sum_{i=1}^n \frac{d}{dx_i} \left(\frac{\overleftarrow{\partial L}}{\partial p_i} \right) \cdot \overrightarrow{\delta f} \right) dx_1 \dots dx_n. \quad (\text{III}, 11; 107_2)$$

Эта формула совпадает с (III, 11; 106), если только

$$\int \int \dots \int \sum_{i=1}^n \frac{d}{dx_i} \left(\frac{\overleftarrow{\partial L}}{\partial p_i} \cdot \overrightarrow{\delta f} \right) dx_1 \dots dx_n = 0. \quad (\text{III}, 11; 107_3)$$

Если воспользоваться формулой Остроградского (VI, 7; 53), которая будет доказана позже, то интеграл по объему (III, 11; 107₃) можно заменить интегралом по поверхности

$$\int \dots \int \sum_{i=1}^n \left(\frac{\overleftarrow{\partial L}}{\partial p_i} \cdot \overrightarrow{\delta f} \cos \alpha_i \right) dS. \quad (\text{III}, 11; 107_4)$$

Поскольку рассматривается подмножество функций f , принимающих заданные значения на Γ , то $\overrightarrow{\delta f}$ равен нулю на Γ , а, следовательно, последний интеграл действительно равен нулю.

Переход от формулы (III, 11; 106) к уравнению Эйлера (III, 11; 107) требует применения леммы Хаара для случая многих переменных, доказательство которой не является существенно более сложным, чем то, которое нами было проделано для функции одной переменной.

Вернемся, например, к задаче об отыскании минимальных поверхностей. Минимальная поверхность, в которой z выражено в виде функции от x и y в \mathbb{R}^3 , удовлетворяет уравнению в частных производных:

$$\frac{d}{dx} \left(\frac{p}{\sqrt{1+p^2+q^2}} \right) + \frac{d}{dy} \left(\frac{q}{\sqrt{1+p^2+q^2}} \right) = 0, \quad (\text{III}, 11; 108)$$

которое может быть записано в виде

$$\frac{r(1+p^2+q^2)-p(pr+qs)}{(1+p^2+q^2)^{3/2}} + \frac{t(1+p^2+q^2)-q(ps+qt)}{(1+p^2+q^2)^{3/2}} = 0 \quad (\text{III}, 11; 109)$$

или

$$(r+t)(1+p^2+q^2) - (rp^2+2spq+tq^2) = 0. \quad (\text{III}, 11; 110)$$

Если ввести понятие средней кривизны, то можно увидеть, что предыдущее уравнение эквивалентно равенству $R_1 + R_2 = 0$: минимальная поверхность является поверхностью нулевой

средней кривизны, т. е. поверхностью, в каждой точке которой главные радиусы кривизны имеют противоположные значения. Конечно, при строгом решении этой задачи вариационного исчисления мы встретимся с теми же трудностями, с которыми нам приходилось сталкиваться до сих пор, но они значительно более сложные, поскольку речь идет о задаче с кратными интегралами. Мы знаем, что для поверхностей \mathcal{S} , ограниченных контуром \mathcal{C} , существует точная нижняя грань их площади > 0 , но не знаем, достигается ли она, и тем более не знаем (в случае когда этот минимум существует), достигается ли он поверхностью, для которой переменная z может быть представлена в виде функции x и y класса C^2 . В этом заключаются основные трудности экстремальных проблем, связанных с кратными интегралами. Предположим даже, что выполнимость перечисленных выше условий проверена и что функция z удовлетворяет уравнениям в частных производных (III, 11; 110). Тогда остается найти еще то решение уравнения в частных производных, для которого поверхность \mathcal{S} проходит через заданный контур \mathcal{C} , т. е. является такой поверхностью, для которой z принимает заданные значения вдоль проекции Γ контура \mathcal{C} на плоскость xy . Задача отыскания минимальной поверхности, ограниченной заданной кривой, называется задачей Плато. Она была решена Дугласом. Ее решение слишком сложно, и потому мы не будем на нем останавливаться.

Рассмотрим другую задачу. Задано открытое множество $\mathcal{O} \subset \mathbb{R}^n$, ограниченное гиперповерхностью Γ класса C^1 . Требуется найти вещественную функцию f на $\bar{\mathcal{O}}$ класса C^1 , принимающую заданные значения на контуре Γ и минимизирующую значение кратного интеграла (называемого интегралом Дирихле):

$$\int_{\mathcal{O}} \int \dots \int \left(\sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right|^2 \right) dx_1 \dots dx_n = \int_{\mathcal{O}} \int \dots \int \left(\sum_{i=1}^n p_i^2 \right) dx_1 \dots dx_n. \quad (\text{III}, 11; 111)$$

Здесь f должна быть (если она принадлежит классу C^2) решением уравнения Эйлера

$$\sum_{i=1}^n \frac{\partial p_i}{\partial x_i} = 0, \quad (\text{III}, 11; 112)$$

т. е. решением уравнения Лапласа

$$\Delta f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} = 0. \quad (\text{III}, 11; 113)$$

Говорят также, что f должна быть гармонической функцией. Таким образом, мы приходим к необходимости решения задачи Дирихле: в классе C^2 функций, определенных на $\bar{\Omega}$, найти гармоническую функцию, принимающую заданные значения на контуре Γ . Существует много физических проблем, приводящих к необходимости решения задачи Дирихле. В этих проблемах интеграл (III, 11; 111) представляет собой некоторую энергию, и задача заключается в отыскании функции, реализующей некоторое равновесное состояние, которое дает минимум энергии. Риман в XIX веке решал задачу Дирихле этим методом, предварительно доказав существование минимума интеграла. К несчастью, решение Римана содержало ошибки. С одной стороны, как это мы указывали в п. 1°) на стр. 381, нет никакой гарантии, что минимум существует, а дополнительные замечания Римана, на которых строилось доказательство существования, были основаны на неверной теореме компактности¹⁾. С другой стороны, как мы видели в п. 2°) на стр. 381, ничто не доказывает, что этот минимум, если он существует, реализуется функцией f класса C^2 . В решении этих задач имеется много существенных трудностей, как мы это уже не раз отмечали. В дальнейшем при изучении гармонических функций будет показано, как можно устраниТЬ затруднения, возникающие в методе Римана, и как методами этого типа эффективно решается задача Дирихле.

¹⁾ Метод Римана основывался на том предположении, что единичный шар в бесконечномерном банаховом пространстве компактен, как и в случае конечномерного пространства. Однако мы видели, что это не верно (теорема 45₂ гл. II).

ГЛАВА IV

Интегральное исчисление

§ 1. ИНТЕГРАЛ РИМАНА НА ПРЯМОЙ

Пусть заданы пространство Банаха \vec{F} над полем K вещественных или комплексных чисел и функция \vec{f} со значениями в \vec{F} на интервале $[a, b]$ вещественной прямой R . Мы намереваемся определить интеграл

$$\int_{[a, b]} \vec{f}(x) dx \in \vec{F}. \quad (\text{IV}, 1; 1)$$

Мы видим, почему необходимо предположение о том, что функция \vec{f} должна принимать значения в некотором банаховом пространстве. Ведь, как известно, интеграл можно рассматривать как предел конечных сумм вида

$$\sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{f}(\xi_i), \quad (\text{IV}, 1; 2)$$

где c_0, c_1, \dots, c_n — возрастающая последовательность $n + 1$ точек интервала $[a, b]$, $c_0 = a$, $c_n = b$ и $\xi_i \in [c_i, c_{i+1}]$. Однако для того чтобы можно было рассматривать такую сумму, надо прежде всего иметь возможность определять каждый член $(c_{i+1} - c_i)\vec{f}(\xi_i)$ этой суммы, а для этого надо уметь строить произведение элемента $\vec{f}(\xi_i)$ из \vec{F} на вещественный скаляр $c_{i+1} - c_i$. Затем надо иметь возможность рассматривать сумму таких элементов из \vec{F} , а, следовательно, \vec{F} должно быть векторным пространством над полем вещественных чисел. Кроме того, интеграл — не сумма, а предел сумм. Поэтому надо иметь право рассматривать в пространстве \vec{F} пределы, а значит, естественно, предполагать, что \vec{F} является векторным нормированным пространством над полем вещественных чисел. Теоретически этого должно быть достаточно для определения интеграла. Однако если не предполагать полноту \vec{F} , то невозможно будет найти практически применимые критерии существования этого интеграла, ибо только в этом случае можно будет доказать существование предела, не зная его заранее.

Поэтому в дальнейшем, если не оговорено противное, мы будем предполагать, что \vec{F} — банахово пространство. Естественно, можно также рассматривать банаховы пространства над полем комплексных чисел, поскольку их можно трактовать как банаховы пространства над полем вещественных чисел.

Функцию \vec{f} мы будем считать всегда ограниченной. Другими словами, при x , пробегающем $[a, b]$, будем считать $\|\vec{f}(x)\|$ не превосходящей некоторого фиксированного числа. Впрочем, как это делалось на стр. 148, через $\|\vec{f}\|$ мы будем обозначать функцию ≥ 0 : $x \rightarrow \|\vec{f}(x)\|$, определенную на $[a, b]$, в то время как $\|\vec{f}(x)\|$ будет представлять собой точную верхнюю грань этой функции¹⁾.

Кроме того, мы будем рассматривать функции \vec{f} , определенные на всей числовой прямой, и интегрировать их на всей прямой, считая всегда, что эти функции равны нулю вне некоторого ограниченного интервала.

Таким образом, мы будем писать выражения

$$\int \vec{f}(x) dx, \text{ или просто } \int \vec{f} dx, \text{ или } \int \vec{f}. \quad (\text{IV}, 1; 3)$$

Тогда, по определению, символ (IV, 1; 1) является интегралом на всей прямой от функции \vec{f} , равной \vec{f} в интервале $[a, b]$

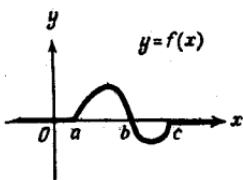


Рис. 12.

и $\vec{0}$ в дополнении к этому интервалу. Замыкание точек x топологического пространства X , в которых функция \vec{f} , определенная в пространстве X , со значениями в векторном пространстве \vec{F} отлична от $\vec{0}$, называется *носителем* функции \vec{f} . По определению, *носитель всегда замкнут*. Носитель функции \vec{f} является наименьшим замкнутым множеством X , на дополнении к которому функция \vec{f} тождественно равна $\vec{0}$. Например, для вещественной функции, график которой изображен на рис. 12,

¹⁾ Если \vec{F} — поле скаляров, то $\|\cdot\|$ заменится на $|\cdot|$ и тогда $||| \cdot |||$ можно будет заменить на $\|\cdot\|$.

множество точек, в которых $f(x) \neq 0$, представляет собой открытый интервал $[a, c]$ с удаленной точкой b . Однако носителем этой функции будет замкнутый интервал $[a, c]$.

Если f — вещественная функция, равная 0 во всех иррациональных точках \mathbb{R} и 1 во всех рациональных точках, то множество точек, в которых она отлична от нуля, является множеством рациональных точек \mathbb{Q} , в то время как ее носителем будет вся вещественная прямая \mathbb{R} .

Если \vec{f} и \vec{g} — две функции, определенные на X , со значениями в векторном пространстве \vec{F} , то носитель функции $\vec{f} + \vec{g}$ содержится, очевидно, в объединении носителей \vec{f} и \vec{g} . В самом деле, точка x , в которой $\vec{f}(x) + \vec{g}(x) \neq \vec{0}$, принадлежит либо множеству точек A , в которых $\vec{f} \neq \vec{0}$, либо множеству точек B , в которых $\vec{g} \neq \vec{0}$. Обе эти возможности не исключают друг друга. Но тогда эта точка принадлежит объединению $A \cup B$ и, следовательно, носитель содержится в замыкании $\overline{A \cup B}$, которое содержится в объединении $\overline{A} \cup \overline{B}$ носителей. Мы будем интегрировать на \mathbb{R} функции f , определенные на \mathbb{R} и имеющие на нем компактные носители.

Ступенчатые функции

Функция f , определенная на вещественной прямой \mathbb{R} , со значениями в произвольном множестве F называется *ступенчатой*, если существует такая конечная возрастающая последовательность точек $c_0, c_1, c_2, \dots, c_n$ из \mathbb{R} , что на каждом из открытых интервалов $(-\infty, c_0], [c_0, c_1], \dots, [c_{n-1}, c_n], [c_n, +\infty)$ функция f постоянна. (Никаких дополнительных предположений относительно значений функции в самих точках c_i не требуется.) Такая последовательность точек $c_0, c_1, c_2, \dots, c_n$ называется *разбиением* Δ прямой \mathbb{R} , допустимым для ступенчатой функции f . Естественно, существуют «наилучшие» разбиения, связанные с f , т. е. такие, при которых число точек c_i будет наименьшим. Для одной и той же ступенчатой функции f существует бесконечное число допустимых разбиений \mathbb{R} . Разбиение Δ' прямой \mathbb{R} называется *измельчением* разбиения Δ , если последовательность c'_i содержит последовательность c_i . Всякое измельчение допустимого разбиения для функции f является допустимым разбиением для этой функции. Существует по крайней мере одно разбиение, являющееся измельчением двух заданных разбиений Δ' и Δ'' прямой \mathbb{R} . Оно может быть получено объединением точек разбиений, относящихся к Δ' и Δ'' , и перенумерацией их в порядке возрастания. Если значения ступенчатой функции

изменяется в конечном числе точек, то она останется ступенчатой функцией. Если множество \vec{F} является векторным пространством, то произведение ступенчатой функции на скаляр и сумма двух ступенчатых функций будет ступенчатой функцией. Рассмотрим внимательнее последнее утверждение. Если \vec{f} и \vec{g} — две ступенчатые функции, то им соответствуют некоторые разбиения Δ' и Δ'' прямой \mathbb{R} , а, значит, и измельчение Δ этих разбиений. Разбиение Δ является допустимым разбиением \mathbb{R} для каждой из функций \vec{f} и \vec{g} , а потому и для суммы $\vec{f} + \vec{g}$ этих функций.

Определим теперь *интеграл Римана ступенчатой функции с компактным носителем* и значениями в векторном пространстве \vec{F} над полем K (не обязательно нормированном) по формуле

$$\int \vec{f} = \sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{f}(\xi_i), \quad \xi_i \in [c_i, c_{i+1}], \quad (\text{IV}, 1; 4)$$

где c_0, c_1, \dots, c_n — допустимое разбиение для данной ступенчатой функции¹⁾.

Следует проверить, что введенное определение корректно, т. е. что правая часть равенства (IV, 1; 4) не зависит от выбранного разбиения. Пусть Δ' и Δ'' — два допустимых разбиения для функции \vec{f} и Δ — измельчение каждого из этих разбиений. Тогда непосредственно видно, что выражения (IV, 1; 4), относящиеся к Δ' и Δ'' , оба дадут тот же результат, что и (IV, 1; 4) относительно Δ . В самом деле, если $[c'_i, c'_{i+1}]$ — интервал разбиения Δ' , то в разбиении Δ он будет разбит на некоторое конечное число интервалов $[d_j, d_{j+1}], [d_{j+1}, d_{j+2}], \dots, [d_{k-1}, d_k]$. В каждом из этих интервалов, как и в исходном интервале $[c'_i, c'_{i+1}]$, функция \vec{f} является одной и той же постоянной \vec{f}_i , и, следовательно, имеет место равенство

$$\sum_{l=j, j+1, \dots, k-1} (d_{l+1} - d_l) \vec{f}_i = (c'_{i+1} - c'_i) \vec{f}_i, \quad (\text{IV}, 1; 5)$$

откуда следует тождественность сумм, соответствующих Δ' и Δ .

Если λ — произвольный скаляр и если \vec{f} и \vec{g} — ступенчатые функции с компактным носителем, то имеют место очевидные

¹⁾ Поскольку предполагалось, что функция \vec{f} имеет компактный носитель, она необходимо равна нулю в $]-\infty, c_0]$ и в $[c_n, +\infty[$. Заметим, что значения функции \vec{f} в точках c_i в рассуждениях не принимаются во внимание.

формулы

$$\int \lambda \vec{f} = \lambda \int \vec{f} \quad \text{и} \quad \int (\vec{f} + \vec{g}) = \int \vec{f} + \int \vec{g}. \quad (\text{IV}, 1; 6)$$

Это очевидно для умножения на скаляр. Что же касается суммы функций, то достаточно выбрать разбиение Δ , общее для \vec{f} и \vec{g} и, следовательно, для суммы $\vec{f} + \vec{g}$.

Таким образом, получаем следующую теорему:

Теорема 1. *Множество ступенчатых функций с компактным носителем, определенных на \mathbb{R} , со значениями в векторном пространстве \vec{F} над полем вещественных или комплексных чисел \mathbb{K} само является векторным пространством над полем \mathbb{K} , а интеграл $\vec{f} \rightarrow \int \vec{f}$ является линейным отображением этого векторного пространства в \vec{F} . Если $\vec{F} = \mathbb{R}$ и $f \geq 0$, то $\int f \geq 0$.*

Если, кроме того, \vec{F} является векторным нормированным пространством, то имеют место следующие неравенства:

$$\left\| \int \vec{f} \right\| \leq \int \|\vec{f}\| = \int \|\vec{f}(x)\| dx, \quad \left\| \int_{[a, b]} \vec{f} \right\| \leq (b - a) \sup_{a \leq x \leq b} \|\vec{f}(x)\|. \quad (\text{IV}, 1; 7)$$

Если значение ступенчатой функции с компактным носителем изменить в конечном числе точек прямой \mathbb{R} , то она остается ступенчатой функцией с компактным носителем, а ее интеграл не изменится.

Конец теоремы был доказан перед ее формулировкой. Неравенства (IV, 1; 7) очевидны на разбиении Δ , допустимом для функции \vec{f} :

$$\left\| \sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{f}_i \right\| \leq \sum_{i=0}^{n-1} (c_{i+1} - c_i) \|\vec{f}_i\| \leq (b - a) \sup_i \|\vec{f}_i\|$$

при $c_0 = a$ и $c_n = b$.

Пусть теперь \vec{f} и \vec{g} — две ступенчатые функции с компактными носителями, отличающиеся лишь на конечном числе точек \mathbb{R} . Эти точки необходимо находятся среди точек c_i разбиения Δ , допустимого одновременно для функций \vec{f} и \vec{g} , а в интеграл не входят значения функции в самих точках c_i разбиения Δ .

Для определения интеграла от произвольной функции придется воспользоваться предельным переходом.

Верхний интеграл Римана от ограниченной функции $f > 0$ с компактным носителем

Определение. Пусть $f \geq 0$ — вещественная ограниченная функция с компактным носителем, определенная на прямой \mathbb{R} .

Верхним интегралом Римана $\int^* f$ функции f называется точная нижняя грань интегралов от ступенчатых функций с компактным носителем $f_1 \geq f$.

Таким образом, по определению

$$\int^* f = \inf_{\substack{f_1 - \text{ступенчатая} \\ f_1 \geq f}} \left(\int f_1 \right) \geq 0. \quad (\text{IV}, 1; 8)$$

Легко понять причины предыдущих ограничений, в силу которых функция f предполагается ограниченной и имеющей компактный носитель. Если бы этих ограничений не было, то не существовало бы ступенчатой функции с компактным носителем, мажорирующей функцию f . Конечно, в том случае, когда функция f сама ступенчатая, в качестве f_1 можно взять ее саму и тогда $\int^* f = \int f$.

Теорема 2. Если $f \leq g$, то $\int^* f \leq \int^* g$. Для положительных ограниченных функций f и g с компактными носителями и неотрицательного скаляра λ имеют место формулы

$$\int^* \lambda f = \lambda \int^* f, \quad \int^* (f + g) \leq \int^* f + \int^* g. \quad (\text{IV}, 1; 9)$$

Говорят также, что верхний интеграл положительных функций обладает свойством выпуклости. Верхний интеграл функции не изменяется при изменении значений этой функции в конечном числе точек из \mathbb{R} .

Доказательство. Высказанные утверждения почти очевидны. Докажем, например, свойство выпуклости.

Пусть f_1 и g_1 — ступенчатые ограниченные функции с компактным носителем, мажорирующие функции f и g соответственно. Тогда $f + g \leq f_1 + g_1$, и мы имеем

$$\int^* (f + g) \leq \int (f_1 + g_1) = \int f_1 + \int g_1. \quad (\text{IV}, 1; 10)$$

Левая часть этого соотношения не превосходит точной нижней грани его правой части, равной $\int^* f + \int^* g$, откуда и следует (IV, 1; 9).

Пусть теперь f и g совпадают всюду, кроме конечного числа точек \mathbb{R} , и $f_1 \geq f$ — ступенчатая функция с компактным носителем. Всегда можно найти такую ступенчатую функцию $g_1 \geq g$

с компактным носителем, которая всюду совпадает с f_1 , кроме конечного числа точек. При этом $\int^* g \leq \int^* g_1 = \int^* f_1$. Переходя справа к точной нижней грани, получим неравенство $\int^* g \leq \int^* f$. Заменяя в предыдущих рассуждениях f на g и g на f , точно так же получим $\int^* f \leq \int^* g$, откуда и следует равенство $\int^* f = \int^* g$.

З а м е ч а н и я. 1°) Можно было бы точно так же определить нижний интеграл Римана $\int_* f$ как точную верхнюю грань интегралов ступенчатых функций, не превосходящих f . Однако этот нижний интеграл обладал бы тогда свойством вогнутости, а именно:

$$\int_* (f+g) \geq \int_* f + \int_* g, \quad (\text{IV}, 1; 11)$$

практически почти не применяемым.

2°) Можно было бы подумать, что имеет место равенство $\int^* (f+g) = \int^* f + \int^* g$. Однако это не так. В самом деле, рассмотрим функцию f , равную нулю в дополнении к $[0, 1]$ и во всех иррациональных точках отрезка $[0, 1]$ и равную 1 во всех рациональных точках этого отрезка. Пусть f_1 — ступенчатая функция $\geq f$. В каждом интервале $[c_i, c_{i+1}]$ любого допустимого разложения Δ для функции f_1 имеются рациональные точки. Следовательно, постоянные значения f_1 , наверное ≥ 1 . Но тогда f_1 не меньше характеристической функции отрезка $[0, 1]$. Обратно, характеристическая функция является ступенчатой и мажорирует f . Поэтому $\int^* f = 1$. Если через g мы обозначим функцию, полученную аналогично функции f , но со взаимной заменой значений в рациональных и иррациональных точках, то мы также получим $\int^* g = 1$. Однако $f+g$ является характеристической функцией отрезка $[0, 1]$ и $\int^* (f+g) = \int (f+g) = 1 < 1 + 1$.

Из этого примера видно, что если ввести нижние интегралы, указанные в замечании 1°), то мы получим: $\int_* f = \int_* g = 0$ и $\int_* (f+g) = \int (f+g) = 1 > 0 + 0$.

3°) Изменение функции на счетном множестве точек \mathbb{R} может привести к изменению его верхнего интеграла. Так, напри-

мер, функция f из замечания 2°) отлична от нуля в рациональных точках, а ее интеграл равен 1, а не 0.

4°) Условие выпуклости естественным образом обобщается на сумму конечного числа функций, но не обобщается на сумму счетного числа их. Другими словами, если $f_0, f_1, \dots, f_n, \dots$ — последовательность неотрицательных ограниченных функций с носителями в одном и том же отрезке $[a, b]$ прямой \mathbb{R} и если ряд $\sum_{n=0}^{\infty} f_n(x)$ при каждом x сходится к пределу $f(x)$, где функция f ограничена (а ее носитель, очевидно, лежит в $[a, b]$), то неравенство $\int^* f \leq \sum_{n=0}^{\infty} \int^* f_n$ может не выполняться.

Рассмотрим, например, множество рациональных чисел отрезка $[0, 1]$. Их можно расположить в последовательность a_0, a_1, a_2, \dots . Обозначим через f_n характеристическую функцию множества $\{a_n\}$. Тогда $f = \sum_{n=0}^{\infty} f_n$ является функцией, рассмотренной в замечании 2°), причем $\int^* f = 1 > \sum_{n=0}^{\infty} \int^* f_n = 0$.

В теории интеграла Лебега мы введем также понятие верхнего интеграла, который будет обладать свойством выпуклости и в случае счетных сумм $\int^ \sum_{n=0}^{\infty} f_n \leq \sum_{n=0}^{\infty} \int^* f_n$, и это будет существенным превосходством интеграла Лебега над интегралом Римана.*

Интегрируемые функции со значениями в пространстве Банаха

Определение. Пусть \vec{f} — функция, определенная на вещественной прямой \mathbb{R} , со значениями в банаховом пространстве \vec{F} . Говорят, что функция \vec{f} интегрируема по Риману, если она ограничена, имеет компактный носитель и если для любого $\varepsilon > 0$ существует такая ступенчатая функция \vec{g} с компактным носителем и значениями в \vec{F} , что

$$\int^* \|\vec{f} - \vec{g}\| \leq \varepsilon. \quad (\text{IV}, 1; 12)$$

Это эквивалентно существованию такой последовательности ступенчатых функций $\vec{f}_0, \vec{f}_1, \dots, \vec{f}_n, \dots$ с компактными носителями и значениями в \vec{F} , для которых верхние интегралы $\int^* \|f - f_n\|$

сходятся к 0 при n , стремящемся к бесконечности. Такая последовательность функций называется *аппроксимирующей последовательностью* функции \vec{f} для интеграла Римана¹⁾.

Если две функции отличаются только в конечном числе точек и если одна из них интегрируема, то другая также будет интегрируемой (и будет иметь ту же самую аппроксимирующую последовательность).

Замечания. 1°) Пусть f — ограниченная функция, определенная на \mathbb{R} , с компактным носителем и значениями в базаховом пространстве \vec{F} . Если, каково бы ни было $\epsilon > 0$, существует такая интегрируемая по Риману функция g , что $\int^* \|\vec{f} - \vec{g}\| \leq \epsilon$, то *сама функция \vec{f} интегрируема*.

В самом деле, при заданном $\epsilon > 0$ мы можем сначала выбрать такую интегрируемую функцию \vec{g} , что $\int^* \|\vec{f} - \vec{g}\| \leq \epsilon/2$. В силу интегрируемости \vec{g} можно найти такую ступенчатую функцию \vec{h} с компактным носителем, что $\int^* \|\vec{g} - \vec{h}\| \leq \epsilon/2$. Тогда, в силу неравенства выпуклости (IV, 1; 9), $\int^* \|\vec{f} - \vec{h}\| \leq \epsilon$, а это, по определению, означает, что \vec{f} интегрируема.

Точно так же, если задана такая последовательность интегрируемых функций $\vec{f}_0, \vec{f}_1, \dots, \vec{f}_n, \dots$, что $\int^* \|\vec{f} - \vec{f}_n\|$ стремится к 0 (в частности, если функции \vec{f}_n сохраняют свои носители на одном и том же компакте и равномерно сходятся к \vec{f}), то \vec{f} интегрируема. В этом случае говорят, что последовательность \vec{f}_n функций (не обязательно ступенчатых) является *аппроксимирующей последовательностью* в интегральном смысле для функции \vec{f} .

2°) Если \vec{f} — интегрируемая функция, то, каково бы ни было $\epsilon > 0$, существует ступенчатая функция \vec{g} с компактным носителем и ступенчатая функция \vec{h} с компактным носителем и вещественными значениями ≥ 0 , такие, что $\|\vec{f} - \vec{g}\| \leq h$ и $\int h \leq \epsilon$, и обратно.

¹⁾ Не следует думать, что такая последовательность \vec{f}_n сходится просто к \vec{f} . На стр. 550 мы приведем пример, в котором $\vec{f}_n(x)$ не сходится к $\vec{f}(x)$ ни при каком значении x .

Обратное утверждение очевидно; поэтому мы докажем лишь прямое утверждение.

Прежде всего найдем ступенчатую функцию \vec{g} с компактным носителем, такую, что $\int^* \|\vec{f} - \vec{g}\| \leq \varepsilon/2$. Тогда, по определению верхнего интеграла, найдется такая неотрицательная вещественная ступенчатая функция h с компактным носителем, что $\|\vec{f} - \vec{g}\| \leq h$ и $\int h \leq \varepsilon$.

3°) Если f — вещественная интегрируемая по Риману функция, определенная на \mathbb{R} , то, каково бы ни было $\varepsilon > 0$, можно найти две такие ступенчатые функции g_1 и g_2 с компактными носителями, что $g_1 \leq f \leq g_2$ и $\int g_2 - g_1 \leq \varepsilon$, и обратно.

Обратное очевидно; поэтому докажем лишь прямое утверждение. Пусть f интегрируема. Определим, согласно замечанию 2°), вещественную ступенчатую функцию g и ступенчатую функцию $h \geq 0$ с компактными носителями так, чтобы $|f - g| \leq h$ и $\int h \leq \varepsilon/2$. Если теперь положить $g_1 = g - h$ и $g_2 = g + h$, то g_1 и g_2 будут удовлетворять указанным условиям.

Интеграл от интегрируемой функции

Теорема 3. Пусть \vec{f} — интегрируемая по Риману функция со значениями в банаховом пространстве \vec{F} , и пусть $\vec{f}_0, \vec{f}_1, \vec{f}_2, \dots$ — аппроксимирующая функцию \vec{f} последовательность ступенчатых функций. Тогда величины $\int \vec{f}_n \in \vec{F}$ сходятся к некоторому пределу в \vec{F} и этот предел не зависит от рассматриваемой аппроксимирующей последовательности.

Доказательство. Из неравенства

$$\left\| \int \vec{f}_m - \int \vec{f}_n \right\| \leq \int \|\vec{f}_m - \vec{f}_n\| \leq \int^* \|\vec{f} - \vec{f}_m\| + \int^* \|\vec{f} - \vec{f}_n\| \quad (\text{IV}, 1; 13)$$

следует, что при m и n , стремящихся к бесконечности, $\left\| \int \vec{f}_m - \int \vec{f}_n \right\|$ стремится к 0 . Это означает, что последовательность $\int \vec{f}_n$ является последовательностью Коши в \vec{F} . Поскольку \vec{F} полно¹⁾, эта последовательность имеет предел \vec{L} . Остается убедиться, что этот предел не зависит от выбора

¹⁾ Именно здесь существует тот факт, что \vec{F} полно.

аппроксимирующей последовательности. Если мы рассмотрим теперь две такие аппроксимирующие последовательности $\vec{f}_0, \vec{f}_1, \vec{f}_2, \dots$ и $\vec{g}_0, \vec{g}_1, \vec{g}_2, \dots$, то из них можно образовать последовательность $\vec{f}_0, \vec{g}_0, \vec{f}_1, \vec{g}_1, \vec{f}_2, \vec{g}_2, \dots$, которая также является аппроксимирующей для \vec{f} . Это значит, что последовательность величин $\int \vec{f}_0, \int \vec{g}_0, \int \vec{f}_1, \int \vec{g}_1, \int \vec{f}_2, \int \vec{g}_2, \dots$ должна иметь предел; другими словами, последовательности $\int \vec{f}_n$ и $\int \vec{g}_n$ действительно имеют один и тот же предел.

Определение. Значение предела, указанного в теореме 3, называется *интегралом Римана интегрируемой функции \vec{f}* и обозначается через $\int \vec{f}(x) dx$ или просто через $\int \vec{f} \in \vec{F}$.

Если две интегрируемые функции \vec{f} и \vec{g} отличаются лишь в конечном числе точек \mathbb{R} , то они имеют один и тот же интеграл, поскольку они имеют одни и те же аппроксимирующие последовательности.

Говорят, что \vec{f} интегрируема на отрезке $[a, b]$ из \mathbb{R} , если функция \vec{f} , равная \vec{f} на $[a, b]$ и $\vec{0}$ вне его, интегрируема, и в этом случае полагают

$$\int_{[a, b]} \vec{f} = \int \vec{f}. \quad (\text{IV}, 1; 13_2)$$

Замечание. Если \vec{f} является ступенчатой функцией, то данное сейчас определение $\int \vec{f}$ соответствует первоначальному определению (IV, 1; 4). В самом деле, в качестве аппроксимирующей последовательности для функции \vec{f} надо взять последовательность $\vec{f}_1, \vec{f}_2, \vec{f}_3, \dots, \vec{f}_n, \dots$, и мы придем к исходному интегралу.

Теорема 4. *Если λ — скаляр, а \vec{f} и \vec{g} — две интегрируемые функции со значениями в \vec{F} , то $\lambda \vec{f}$ и $\vec{f} + \vec{g}$ интегрируемы и имеют место формулы*

$$\int \lambda \vec{f} = \lambda \int \vec{f}, \quad \int (\vec{f} + \vec{g}) = \int \vec{f} + \int \vec{g}. \quad (\text{IV}, 1; 14)$$

Иначе говоря, множество интегрируемых по Риману функций (необходимо ограниченных с компактным носителем) является векторным пространством над полем \mathbb{K} , а интеграл является линейным отображением этого векторного пространства в \vec{F} .

Доказательство. Утверждение, относящееся к умножению на скаляр, очевидно; поэтому достаточно рассмотреть случай суммы. Пусть \vec{f}_n и \vec{g}_n — две аппроксимирующие последовательности для функций \vec{f} и \vec{g} . Тогда имеет место неравенство

$$\int^* \|(\vec{f} + \vec{g}) - (\vec{f}_n + \vec{g}_n)\| \leq \int^* \|\vec{f} - \vec{f}_n\| + \int^* \|\vec{g} - \vec{g}_n\|, \quad (\text{IV}, 1; 15)$$

из которого следует, что последовательность $\vec{f}_n + \vec{g}_n$ является аппроксимирующей для функции $\vec{f} + \vec{g}$, а, следовательно, функция $\vec{f} + \vec{g}$ интегрируема. Это неравенство означает, что $\int (\vec{f} + \vec{g})$ является пределом величин $\int (\vec{f}_n + \vec{g}_n) = \int \vec{f}_n + \int \vec{g}_n$, т. е. равняется $\int \vec{f} + \int \vec{g}$.

Следствие. Пусть $a, b, c, a \leq b \leq c$, — три точки прямой \mathbb{R} . Пусть \vec{f} — функция, определенная на $[a, c]$ со значениями в \vec{F} . Для того чтобы она была интегрируемой на $[a, c]$, необходимо и достаточно, чтобы она была интегрируемой на $[a, b]$ и на $[b, c]$, и в этом случае

$$\int_{[a, c]} \vec{f} = \int_{[a, b]} \vec{f} + \int_{[b, c]} \vec{f}. \quad (\text{IV}, 1; 16)$$

Доказательство. Пусть \vec{f} — функция, определенная на \mathbb{R} , равная \vec{f} на $[a, c]$ и $\vec{0}$ вне этого отрезка. Пусть $\Phi_{[a, b]}$, $\Phi_{[b, c]}$ и $\Phi_{[a, c]}$ — характеристические функции интервалов $[a, b]$, $[b, c]$ и $[a, c]$. Если функция \vec{f} интегрируема на $[a, b]$ и $[b, c]$, то интегрируемыми будут функции $\vec{f} \cdot \Phi_{[a, b]}$ и $\vec{f} \cdot \Phi_{[b, c]}$, а значит, и $\vec{f} \cdot \Phi_{[a, c]}$. Но тогда интегрируема сумма $\vec{f} \cdot (\Phi_{[a, b]} + \Phi_{[b, c]}) = \vec{f}$, а, значит, интегрируема на $[a, c]$ функция \vec{f} и из (IV, 1, 14) следует (IV, 1; 16).

Обратно, если \vec{f} интегрируема на $[a, c]$ и если \vec{f}_n является аппроксимирующей последовательностью функции \vec{f} для интеграла на $[a, c]$, то интеграл $\int_{[a, c]}^* \|\vec{f} - \vec{f}_n\|$ стремится к нулю.

Вместе с ним стремятся к нулю интегралы $\int_{[a, b]}^* \|\vec{f} - \vec{f}_n\|$ и $\int_{[b, c]}^* \|\vec{f} - \vec{f}_n\|$. Это означает, что \vec{f}_n является аппроксимирующей последовательностью функции \vec{f} для интегралов, распространенных на $[a, b]$ и $[b, c]$, а \vec{f} интегрируема на этих интервалах.

Теорема 5. Если функция \vec{f} , определенная на \mathbb{R} , со значениями в \vec{F} интегрируема, то функция $\|\vec{f}\|: x \rightarrow \|\vec{f}(x)\|$ является вещественной интегрируемой неотрицательной функцией и имеет место неравенство

$$\left\| \int \vec{f} \right\| \leqslant \int \|\vec{f}\|. \quad (\text{IV}, 1; 17)$$

Доказательство. В самом деле, пусть \vec{f}_n является аппроксимирующей последовательностью для функции \vec{f} , составленной из ступенчатых функций. Тогда имеет место неравенство

$$\|\vec{f}\| - \|\vec{f}_n\| \leqslant \|\vec{f} - \vec{f}_n\|, \quad (\text{IV}, 1; 18)$$

из которого следует, что $\int^* \|\vec{f}\| - \|\vec{f}_n\|$ сходится к нулю, и, следовательно, последовательность ступенчатых функций $\|\vec{f}_n\|$ является аппроксимирующей последовательностью для функции $\|\vec{f}\|$. Эта функция, таким образом, интегрируема и $\int \|\vec{f}_n\|$ стремится к $\int \|\vec{f}\|$. При этом для любого n имеет место неравенство $\left\| \int \vec{f}_n \right\| \leqslant \int \|\vec{f}_n\|$. При n , стремящемся к бесконечности, $\int \vec{f}_n$ стремится к $\int \vec{f}$ по определению, а, значит, в силу непрерывности нормы (теорема 9 гл. II) $\left\| \int \vec{f}_n \right\|$ стремится к $\left\| \int \vec{f} \right\|$, а $\int \|\vec{f}_n\|$ стремится к $\int \|\vec{f}\|$, откуда и следует неравенство (IV, 1; 17).

Следствие 1. Интеграл вещественной интегрируемой неотрицательной функции f неотрицателен.

В самом деле, так как $|f| = f$, то, применяя (IV, 1; 17), получаем $\int f \geqslant \left| \int f \right| \geqslant 0$.

Следствие 2. Если f и g — такие вещественные интегрируемые функции, что $f \leqslant g$, то $\int f \leqslant \int g$.

Действительно, достаточно к положительной функции $g - f$ применить следствие 1 и учесть, что $\int (g - f) = \int g - \int f$ (теорема 4).

Следствие 3. Для интегрируемой на $[a, b]$ функции \vec{f} имеет место неравенство

$$\left\| \int_{[a, b]} \vec{f} \right\| \leqslant (b - a) \sup_{a \leqslant x \leqslant b} \|\vec{f}(x)\|. \quad (\text{IV}, 1; 19)$$

В самом деле, прежде всего имеет место неравенство (IV, 1; 17) относительно интеграла $\int_{[a, b]} : \left\| \int_{[a, b]} \vec{f} \right\| \leq \int_{[a, b]} \| \vec{f} \|$. Так как функция $\|\vec{f}\|$ на $[a, b]$ не превосходит постоянной $M = \sup_{a \leq x \leq b} \|\vec{f}(x)\|$, то, согласно следствию 2, имеет место неравенство $\int_{[a, b]} \|\vec{f}\| \leq \leq \int_{[a, b]} M = M(b - a)$.

Следствие 3₂. Если f интегрируема на $[a, b]$, вещественна и $m \leq f(x) \leq M$, то

$$m(b - a) \leq \int_{[a, b]} f \leq M(b - a). \quad (\text{IV, 1; 20})$$

Это утверждение непосредственно вытекает из следствия 2.

Следствие 4. Интеграл вещественной неотрицательной интегрируемой функции f совпадает с ее верхним интегралом.

Действительно, пусть f_n — аппроксимирующая функцию f последовательность, состоящая из вещественных ступенчатых функций с компактными носителями. Можно считать, что $f_n \geq 0$. В противном случае f_n можно заменить функциями $|f_n|$, образующими аппроксимирующую последовательность. Тогда

$$f \leq f_n + |f - f_n|,$$

откуда (согласно теореме 2)

$$\int^* f \leq \int f_n + \int^* |f - f_n| \quad (\text{IV, 1; 20}_2)$$

и точно так же

$$f_n \leq f + |f_n - f|, \quad (\text{IV, 1; 20}_3)$$

а потому

$$\int f_n \leq \int^* f + \int^* |f - f_n|,$$

откуда

$$\left| \int^* f - \int f_n \right| \leq \int^* |f - f_n|. \quad (\text{IV, 1; 20}_4)$$

Правая часть стремится к нулю при n , стремящемся к бесконечности, а $\int f_n$ стремится к $\int f$, откуда и вытекает требуемый результат.

Следствие 5. Пусть \vec{f}_n является последовательностью интегрируемых функций, \vec{f} — ограниченная функция с компактным носителем, а $\int^* \|\vec{f} - \vec{f}_n\|$ при n , стремящемся к бесконечности,

стремится к нулю (это будет иметь место, в частности, если \vec{f}_n равномерно сходятся к функции \vec{f} и сохраняют свой носитель на фиксированном компакте). Тогда функция \vec{f} интегрируема, а $\int \vec{f}_n$ сходятся к $\int \vec{f}$.

Мы уже видели ранее (замечание 1°) на стр. 424), что функция \vec{f} интегрируема, а из неравенства

$$\left\| \int \vec{f} - \int \vec{f}_n \right\| \leq \int \| \vec{f} - \vec{f}_n \| = \int^* \| \vec{f} - \vec{f}_n \| \quad (\text{IV}, 1; 21)$$

следует, что $\int \vec{f}_n$ сходится к $\int \vec{f}$.

Теорема 6 (о перестановочности интеграла и линейного непрерывного отображения). Пусть L — линейное непрерывное отображение банахова пространства \vec{F} в банахово пространство \vec{G} . Пусть \vec{f} — интегрируемая по Риману функция, определенная на прямой \mathbb{R} , со значениями в пространстве \vec{F} . Тогда функция $L \circ \vec{f}$, определенная на \mathbb{R} , со значениями в \vec{G} интегрируема по Риману и имеет место соотношение

$$\int L \circ \vec{f} = L \left(\int \vec{f} \right) \in \vec{G}. \quad (\text{IV}, 1; 22)$$

Доказательство. Рассмотрим сначала ступенчатую функцию \vec{g} с компактным носителем, определенную в \mathbb{R} , со значениями в \vec{F} . Функция $L \circ \vec{g}$ является ступенчатой функцией с компактным носителем, и имеет место соотношение

$$\int L \circ \vec{g} = \sum_{i=0}^n (c_{i+1} - c_i) L \circ \vec{g}_i = L \left(\sum_{i=0}^n (c_{i+1} - c_i) \vec{g}_i \right) = L \left(\int \vec{g} \right). \quad (\text{IV}, 1; 23)$$

Пусть теперь \vec{f} — произвольная интегрируемая по Риману функция. Пусть $\vec{f}_0, \vec{f}_1, \dots, \vec{f}_n, \dots$ — аппроксимирующая \vec{f} последовательность ступенчатых функций с компактными носителями. Покажем, что функции $L \circ \vec{f}_n$ составляют аппроксимирующую последовательность для функции $L \circ \vec{f}$. В самом деле, так как

$$\| (L \circ \vec{f} - L \circ \vec{f}_n) \| \leq \| L \| \| \vec{f} - \vec{f}_n \|,$$

то

$$\int^* \| (L \circ \vec{f}) - (L \circ \vec{f}_n) \| \leq \| L \| \int^* \| \vec{f} - \vec{f}_n \|.$$
(IV, 1; 24)

Отсюда видно, что левая часть последнего соотношения стремится к нулю при n , стремящемся к бесконечности, что и доказывает интегрируемость функции $L \circ \vec{f}$.

Соотношение (IV, 1; 23) имеет место для $\vec{g} = \vec{f}_n$. Так как $L \circ \vec{f}_n$ является аппроксимирующей последовательностью для функции $L \circ \vec{f}$, то при n , стремящемся к бесконечности, левая часть этого соотношения стремится к $\int L \circ \vec{f}$.

С другой стороны, $\int \vec{f}_n$ в правой части равенства стремится к $\int \vec{f}$ и, в силу непрерывности отображения L , эта правая часть стремится к $L\left(\int \vec{f}\right)$, что и доказывает соотношение (IV, 1; 23) для случая $\vec{g} = \vec{f}_n$.

Теорема 7. Пусть \vec{f} и \vec{g} — две интегрируемые по Риману функции со значениями в банаховых пространствах \vec{F} и \vec{G} . Пусть B — билинейное непрерывное отображение пространства $\vec{F} \times \vec{G}$ в банахово пространство \vec{H} . Тогда функция $B(\vec{f}, \vec{g})$: $x \rightarrow B(\vec{f}(x), \vec{g}(x))$ интегрируема по Риману и имеет место неравенство

$$\left\| \int B(\vec{f}, \vec{g}) \right\| \leq \|B\| \left\| \vec{f} \right\| \left\| \int \vec{g} \right\|. \quad (\text{IV, 1; 24}_2)$$

Доказательство. Пусть $\vec{f}_0, \vec{f}_1, \vec{f}_2, \dots$ и $\vec{g}_0, \vec{g}_1, \vec{g}_2, \dots$ — последовательности ступенчатых функций с компактными носителями, аппроксимирующие функции \vec{f} и \vec{g} . Заметим, прежде всего, что всегда можно предполагать выполненным неравенство $\left\| \vec{g}_n \right\| \leq 3 \left\| \vec{g} \right\|$. (В самом деле, предположим, что в одном из интервалов разбиения функция \vec{g}_n равна некоторому (постоянному) вектору из \vec{G} с нормой $> 3 \left\| \vec{g} \right\|$. Тогда во всем этом интервале разность $\left\| \vec{g} - \vec{g}_n \right\|$ не меньше постоянной $2 \left\| \vec{g} \right\|$. Заменим в каждом интервале такого вида постоянный вектор \vec{g}_n на пропорциональный ему постоянный вектор \vec{g}'_n с нормой, равной $\left\| \vec{g} \right\|$. Так как $\left\| \vec{g} - \vec{g}'_n \right\| \leq \left\| \vec{g} \right\| + \left\| \vec{g}'_n \right\| \leq 2 \left\| \vec{g} \right\|$, а $2 \left\| \vec{g} \right\| \leq \left\| \vec{g} - \vec{g}_n \right\|$, то $\left\| \vec{g} - \vec{g}'_n \right\| \leq \left\| \vec{g} - \vec{g}_n \right\|$. Функции \vec{g}'_n , заменяющие функции \vec{g}_n , удовлетворяют неравенству $\int^* \left\| \vec{g} - \vec{g}'_n \right\| \leq \int^* \left\| \vec{g} - \vec{g}_n \right\|$,

а, следовательно, последовательность \vec{g}'_n также является аппроксимирующей для функции \vec{g} . На этот раз все аппроксимирующие функции ограничены числом $3 \|\vec{g}\|$.) При этих условиях функция $B(\vec{f}_n, \vec{g}_n)$ является ступенчатой функцией с компактным носителем и имеет место неравенство

$$\begin{aligned} \|B(\vec{f}_n, \vec{g}_n) - B(\vec{f}, \vec{g})\| &\leq \|B(\vec{f}_n - \vec{f}, \vec{g}_n)\| + \|B(\vec{f}, \vec{g}_n - \vec{g})\| \leq \\ &\leq \|B\| \|\vec{f}_n - \vec{f}\| 3 \|\vec{g}\| + \|B\| \|\vec{f}\| \|\vec{g}_n - \vec{g}\|, \quad (\text{IV, 1; 25}) \end{aligned}$$

а, значит,

$$\begin{aligned} \int \|B(\vec{f}_n, \vec{g}_n) - B(\vec{f}, \vec{g})\| &\leq \\ &\leq \|B\| (3 \|\vec{g}\| \int \|\vec{f}_n - \vec{f}\| + \|\vec{f}\| \int \|\vec{g}_n - \vec{g}\|). \quad (\text{IV, 1; 26}) \end{aligned}$$

Отсюда следует, что функции $B(\vec{f}_n, \vec{g}_n)$ образуют аппроксимирующую последовательность для функции $B(\vec{f}, \vec{g})$, а, значит, эта функция интегрируема по Риману. Поскольку

$$\|B(\vec{f}, \vec{g})\| \leq \|B\| \|\vec{f}\| \|\vec{g}\|, \quad (\text{IV, 1; 27})$$

то после интегрирования получаем неравенство (IV, 1; 24).

Следствие 1. Если f и g — интегрируемые по Риману комплексные функции, то функция-произведение fg интегрируема по Риману и имеет место неравенство

$$\left| \int fg \right| \leq \|f\| \int |g|. \quad (\text{IV, 1; 28})$$

Следствие 2. Если f и g — вещественные интегрируемые по Риману функции, то функция fg интегрируема по Риману, а в том случае, когда для всех x функция $g \geq 0$ и $m \leq f(x) \leq M$, имеет место неравенство

$$m \int g \leq \int fg \leq M \int g. \quad (\text{IV, 1; 29})$$

В самом деле, так как fg интегрируема по Риману и для всех x выполняются неравенства

$$mg(x) \leq f(x)g(x) \leq Mg(x), \quad (\text{IV, 1; 30})$$

то, интегрируя эти неравенства, получаем (IV, 1; 29)¹⁾.

Замечание. Если f является непрерывной функцией, то любая промежуточная величина между минимумом и максимумом

¹⁾ Конечно, эти неравенства не имеют места, если для функции $g(x)$ не выполняется условие $g(x) \geq 0$. Пусть, например, $f(x) = g(x) = x$, а интегрирование ведется на $[-1, +1]$. Здесь $M = 1$. Однако неравенство

мумом функции f может быть записана в виде $f(\xi)$, где $\xi \in \mathbb{R}$ (теорема о промежуточных значениях, являющаяся следствием теоремы 33 гл. II). Значит, в этом случае при $g \geq 0$, а также при $g \leq 0$ имеет место соотношение

$$\int f(x) g(x) dx = f(\xi) \int g(x) dx, \quad (\text{IV}, 1; 31)$$

называемое *теоремой о среднем*.

Примеры интегрируемых по Риману функций

Основные примеры вытекают из следствия 5 теоремы 5.

Теорема 8. Для того чтобы функция, определенная на замкнутом интервале $[a, b]$ прямой \mathbb{R} , со значениями в полном¹⁾ метрическом пространстве F была бы правильной, необходимо и достаточно, чтобы она была пределом равномерно сходящейся последовательности ступенчатых функций, определенных на $[a, b]$, со значениями в F .

Доказательство. 1°) Пусть f_0, f_1, f_2, \dots — последовательность ступенчатых функций, определенных на $[a, b]$, равномерно сходящаяся к функции f . Пусть c — точка $[a, b]$.

Так как f_n являются ступенчатыми функциями, то при x , стремящемся к c строго сверху, каждая из этих функций имеет предел. Поскольку пространство F предполагалось полным, то из теоремы 66 гл. II следует, что функция f обладает тем же самым свойством. Аналогичное рассуждение показывает, что функция f имеет предел при x , стремящемся к c строго снизу. Отсюда видно, что любая точка c является или точкой разрыва первого рода или точкой непрерывности рассматриваемой функции f , т. е. что функция f правильная.

2°) Обратно, пусть f — правильная функция со значениями в метрическом пространстве F . Здесь нет необходимости предполагать его полным. Покажем, что, каково бы ни было число $\varepsilon > 0$, можно найти такую ступенчатую функцию, определенную на $[a, b]$, со значениями в F , что $d(f, g) \leq \varepsilon$. Полагая затем последовательно $\varepsilon = 1, 1/2, 1/3, 1/4, \dots$, мы сможем образовать нужную нам последовательность ступенчатых функций.

Пусть x — произвольная точка $[a, b]$. Если точка x отлична от точки a , то, учитывая, что $f(y)$ имеет предел при y , стремящемся

$2/3 = \int_{-1}^1 x^2 dx \leq 1 \int_{-1}^1 x dx = 0$ не справедливо. Опыт показывает, что, несмотря на предостережения, эта ошибка очень живучая.

¹⁾ Для того чтобы выполнялось необходимое условие, не нужно предполагать пространство полным.

к x по значениям $\langle x$, можно найти такой интервал $[x - \alpha, x]$, $\alpha > 0$, для любых двух точек которого ξ и η , отличных от x , имеет место неравенство $d(f(\xi), f(\eta)) \leq \varepsilon$.

Точно так же, если точка x отлична от точки b , можно найти интервал $[x, x + \beta]$, обладающий таким же свойством. Объединяя полученные результаты, получаем, что на отрезке $[a, b]$ всегда можно найти такой открытый в $[a, b]^1$ интервал, содержащий x , в котором для двух точек ξ и η , отличных от x и расположенных по одну сторону от x , имеет место неравенство $d(f(\xi), f(\eta)) \leq \varepsilon$.

Поскольку отрезок $[a, b]$ компактен, то по свойству Гейне—Бореля—Лебега его можно покрыть конечным числом таких открытых множеств. Обозначим через $c_0 = a, c_1, c_2, \dots, c_n = b$ множество точек, образованное из всех начал и концов этих открытых интервалов, и включим в него точки x , для которых были построены эти покрытия. Теперь мы можем следующим образом определить ступенчатую функцию g : в каждой точке c_i она равна значению f в этой точке, а на каждом из интервалов $[c_i, c_{i+1}]$ она равна постоянной $f(\xi_i)$, где ξ_i — произвольная точка этого интервала. Тогда, учитывая способ выбора интервалов, для любой точки x получаем неравенство $d(f(x), g(x)) \leq \varepsilon$, чем и заканчивается доказательство теоремы.

Следствие 1. Любая непрерывная функция, определенная на замкнутом интервале $[a, b]$ прямой \mathbb{R} , со значениями в метрическом пространстве F является пределом равномерно сходящейся последовательности ступенчатых функций.

Заметим, что доказательство для случая непрерывной функции значительно проще, чем для случая произвольной правильной функции. Достаточно заметить, что f равномерно непрерывна на компакте $[a, b]$ (теорема 31 гл. II). Тогда для любого заданного $\varepsilon > 0$ найдется такое $\eta > 0$, что из $d(x', x'') \leq \eta$ будет следовать неравенство $d(f(x'), f(x'')) \leq \varepsilon$.

Если теперь ввести на отрезке $[a, b]$ такое разбиение $c_0 = a, c_1, c_2, \dots, c_n = b$, при котором $c_{i+1} - c_i$ были бы $\leq \eta$, и если выбрать функцию g равной $f(\xi_i)$ в $[c_i, c_{i+1}]$, где $\xi_i \in [c_i, c_{i+1}]$ и $g(b) = f(b)$, то мы получим $d(f, g) \leq \varepsilon$.

Следствие 2. Всякая правильная функция, определенная на прямой \mathbb{R} , с компактным носителем и значениями в банаховом пространстве \tilde{F} интегрируема по Риману. В частности, инте-

¹⁾ Для $x \neq a$ и $x \neq b$ это будет множество $[x - \alpha, x + \beta]$. При $x = a$ получим множество $[a, a + \beta]$, а для $x = b$ — множество $[b - \alpha, b]$. Во всех случаях это открытые множества топологического пространства $[a, b]$. В то же время это и интервалы. Поэтому мы их предполагаем открытыми в $[a, b]$ интервалаами. Конечно, они не обязательно открыты в смысле определения, данного на стр. 26.

грируема по Риману любая непрерывная функция с компактным носителем.

Следствие 3. Любая вещественная функция, определенная на интервале $[a, b]$ прямой \mathbb{R} и монотонная на этом интервале, интегрируема.

Это справедливо, ибо монотонная функция правильна.

Конечно, имеются неправильные интегрируемые функции. Например, функция, равная $\sin 1/x$ на интервале $[0, 1]$ и 0 вне его, интегрируема по Риману, хотя ее разрыв в начале координат не является разрывом первого рода.

Теорема 9. Ограниченнaя функция с компактным носителем и значениями в банаевом пространстве \vec{F} , непрерывная всюду, кроме конечного числа точек, интегрируема по Риману¹⁾.

Доказательство. Для доказательства теоремы достаточно воспользоваться замечанием 1°) на стр. 424 и показать, что для любого $\varepsilon \geq 0$ найдется такая интегрируемая функция \vec{g} , что $\int^* \|\vec{f} - \vec{g}\| \leq \varepsilon$.

Пусть M — максимум $\|\vec{f}\|$, а N — число точек разрыва функции \vec{f} . Рассмотрим функцию \vec{g} , равную \vec{f} вне объединения интервалов $[c_i - \eta, c_i + \eta]$, где c_i — точки разрыва \vec{f} , $\eta = \varepsilon/(2NM)$, и равную нуль в этих интервалах. Тогда \vec{g} будет всюду непрерывной функцией, кроме конечного числа точек, в которых она имеет лишь разрывы 1-го рода, а, следовательно, согласно следствию 2 из теоремы 8, она интегрируема. Функция $\|\vec{f} - \vec{g}\|$ не превосходит ступенчатой функции, равной постоянной M в каждом из N интервалов $[c_i - \eta, c_i + \eta]$ и равной 0 вне их, так что имеет место неравенство $\int^* \|\vec{f} - \vec{g}\| \leq M \cdot 2N\eta = \varepsilon$.

Вычисление интеграла функции с помощью сумм Коши — Римана

Пусть \vec{f} — определенная на интервале $[a, b]$ прямой \mathbb{R} со значениями в банаевом пространстве \vec{F} и интегрируемая по Риману функция. Предлагается следующий способ вычисления ее интеграла.

¹⁾ Не следует думать, что функцию, обладающую этим свойством, можно сделать непрерывной, изменения ее значения в конечном числе точек \mathbb{R} . Это видно из предыдущего примера.

С помощью точек $c_0 = a, c_1, c_2, \dots, c_n = b$ производится разбиение Δ интервала $[a, b]$ на частичные интервалы. Для каждого i через $\vec{\theta}_i$ обозначается произвольный вектор пространства \vec{F} , принадлежащий замыканию множества значений \vec{f} в замкнутом интервале $[c_i, c_{i+1}]$. Затем строится сумма Коши — Римана $\sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{\theta}_i$. В какой же мере эта сумма является приближением к интегралу?

Покажем, что если взять произвольную последовательность разбиений $\Delta_0, \Delta_1, \Delta_2, \dots$, такую, чтобы длина наибольшего из интервалов разбиения Δ_n стремилась к нулю при n , стремящемся к бесконечности, то предыдущая сумма (где $\vec{\theta}_i$ произвольны в каждом интервале разбиения Δ_n) будет стремиться к $\int \vec{f}$ при n , стремящемся к бесконечности.

Теорема 10. Пусть \vec{f} — функция, определенная на ограниченном интервале $[a, b]$ прямой \mathbb{R} , со значениями в банаховом пространстве \vec{F} и интегрируемая по Риману. Тогда по любому $\varepsilon > 0$ можно найти число $\eta > 0$, обладающее следующим свойством: для любого разбиения Δ интервала $[a, b]$, определенного такой возрастающей последовательностью точек $c_0 = a, c_1, c_2, \dots, c_n = b$, что наибольший из интервалов $[c_i, c_{i+1}]$ имеет длину $\leq \eta$, при любом выборе вектора $\vec{\theta}_i$ в замыкании множества значений функции \vec{f} на интервале $[c_i, c_{i+1}]$ имеет место неравенство

$$\left\| \int_{[a, b]} \vec{f} - \sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{\theta}_i \right\| \leq \varepsilon. \quad (\text{IV}, 1; 32)$$

Доказательство. Прежде всего, согласно замечанию 2^o) на стр. 424, можно найти такую ступенчатую функцию \vec{g} , определенную на $[a, b]$ со значениями в \vec{F} , и такую ступенчатую функцию \vec{h} с вещественными неотрицательными значениями, для которых $\|\vec{f} - \vec{g}\| \leq h$ и $\int h \leq \varepsilon/4$. Обозначим через $d_0 = a, d_1, d_2, \dots, d_N = b$ точки разбиения, являющиеся общими для обеих ступенчатых функций \vec{g} и \vec{h} . Покажем теперь, что $\eta = \varepsilon/(8MN)$, где $M = \|\vec{f}\|$, обладает необходимым свойством. В самом деле, пусть Δ — произвольное разбиение отрезка $[a, b]$ точками $c_i, i = 1, 2, \dots, n$, такое, что наибольшая из длин отрезков $[c_i, c_{i+1}]$ не превосходит η . Выберем $\vec{\theta}_i$, как было сказано выше, и обозна-

чим через $\vec{\theta}$ ступенчатую функцию, имеющую c_i точками разбиения, которая равна $\vec{\theta}_i$ в каждом из интервалов $[c_i, c_{i+1}[$ и равна $\vec{f}(b)$ для $x = b$. Сумма Римана $\sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{\theta}_i$ в этом случае будет равна $\int_{[a, b]} \vec{\theta}$. Найдем теперь оценку функции $\|\vec{f}(x) - \vec{\theta}(x)\|$ на каждом из интервалов $[c_i, c_{i+1}[$.

1°) Предположим прежде всего, что интервал $[c_i, c_{i+1}[$ содержится в одном из интервалов $]d_j, d_{j+1}[$. В этом интервале имеет место неравенство

$$\|\vec{f}(x) - \vec{\theta}(x)\| = \|\vec{f}(x) - \vec{\theta}_i\| \leq \|\vec{f}(x) - \vec{g}(x)\| + \|\vec{g}(x) - \vec{\theta}_i\|. \quad (\text{IV}, 1; 33)$$

Первая норма здесь не превосходит $h(x)$. Что же касается второй нормы, то по определению $\vec{\theta}_i$ она не превосходит

$$\sup_{c_i \leq \xi \leq c_{i+1}} \|\vec{g}(x) - \vec{f}(\xi)\|. \quad (\text{IV}, 1; 34)$$

Однако поскольку на интервале $]d_j, d_{j+1}[$ функция \vec{g} постоянна, то в этой формуле можно заменить $\vec{g}(x)$ на $\vec{g}(\xi)$. Следовательно, $\|\vec{g}(x) - \vec{\theta}_i\|$ не превосходит величины $\sup_{c_i \leq \xi \leq c_{i+1}} h(\xi)$, которая в силу постоянства h на интервале $]d_j, d_{j+1}[$ равна $h(x)$. Окончательно в рассматриваемом интервале имеет место оценка:

$$\|\vec{f}(x) - \vec{\theta}(x)\| \leq 2h(x). \quad (\text{IV}, 1; 35)$$

Сумма интегралов от функции $\|\vec{f} - \vec{\theta}\|$ по всем интервалам не превосходит $2 \int h \leq \varepsilon/2$.

2°) Предположим теперь, что интервал $[c_i, c_{i+1}[$ содержит одну из точек разбиения d_j . Поскольку $\vec{\theta}_i$ является точкой прикосновения множества значений \vec{f} , то можно утверждать, что $\|\vec{f}(x) - \vec{\theta}(x)\|$ не превосходит в этом интервале постоянной $2M$. Поскольку таких интервалов имеется не более $2N$, а длина каждого из них не больше η , то сумма интегралов от функции $\|\vec{f} - \vec{\theta}\|$ по всем этим интервалам не превосходит $4MN\eta = \varepsilon/2$.

Окончательно получаем неравенство

$$\int_{[a, b]} \|\vec{f}(x) - \vec{\theta}(x)\| dx \leq \varepsilon, \quad (\text{IV}, 1; 36)$$

и теорема доказана.

Следствие. Пусть f — интегрируемая по Риману вещественная функция, определенная на ограниченном интервале $[a, b]$ прямой \mathbb{R} . Тогда, каково бы ни было число $\varepsilon > 0$, существует число $\eta > 0$, обладающее следующим свойством: если для каждого такого разбиения Δ отрезка $[a, b]$ возрастающей последовательностью точек $c_0 = a, c_1, c_2, \dots, c_n = b$, при котором наибольшее из чисел $c_{i+1} - c_i$ не превосходит η , положить $M_i = \sup_{c_i \leq x \leq c_{i+1}} f(x)$ и $m_i = \inf_{c_i \leq x \leq c_{i+1}} f(x)$, то будут иметь место неравенства

$$\sum_{i=0}^{n-1} (c_{i+1} - c_i) m_i \leq \int_{[a, b]} f(x) dx \leq \sum_{i=0}^{n-1} (c_{i+1} - c_i) M_i \quad (\text{IV}, 1; 37)$$

и

$$\sum_{i=0}^{n-1} (c_{i+1} - c_i) (M_i - m_i) \leq \varepsilon. \quad (\text{IV}, 1; 38)$$

Это утверждение непосредственно вытекает из теоремы, поскольку на $[c_i, c_{i+1}]$ имеют место неравенства $m_i \leq f(x) \leq M_i$, где M_i и m_i являются точками прикосновения множества значений функции f .

Замечание. Следствие уточняет замечание 3°) на стр. 425. В качестве функций g_1 и g_2 можно взять функции, принимающие в интервалах $[c_i, c_{i+1}]$ соответственно значения m_i и M_i и значение $f(b)$ в точке b .

Среднее значение функции на интервале

Пусть \vec{f} — функция, определенная на интервале $[a, b]$ прямой \mathbb{R} , со значениями в банаховом пространстве \vec{F} , интегрируемая по Риману на $[a, b]$. Средним значением \vec{f} на интервале $[a, b]$ называется величина

$$\mathfrak{M}(\vec{f}; [a, b]) = \frac{1}{b-a} \int_{[a, b]} \vec{f}(x) dx \in \vec{F}^1. \quad (\text{IV}, 1; 39)$$

¹⁾ Если $a = b$, то договариваются считать средним значением величину $\vec{f}(a)$.

Из неравенств (IV, 1; 19) и (IV, 1; 20) вытекает, что это среднее не превосходит $\|\vec{f}\|$ и что в вещественном случае оно заключено между точной нижней и точной верхней гранями функции f . Это среднее значение может быть вычислено с помощью предела по следующей формуле:

$$\mathfrak{M}(\vec{f}; [a, b]) = \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=0}^n f\left(a + k \left(\frac{b-a}{n}\right)\right). \quad (\text{IV, 1; 40})$$

В самом деле, рассмотрим разбиение интервала $[a, b]$ на n интервалов длины $(b-a)/n$ и вычислим сумму Коши — Римана:

$$\sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{f}(c_i) = \sum_{k=0}^{n-1} \frac{b-a}{n} \vec{f}\left(a + k \left(\frac{b-a}{n}\right)\right). \quad (\text{IV, 1; 41})$$

Из теоремы 10 вытекает, что, поскольку \vec{f} интегрируема по Риману, эта сумма необходимо сходится к $\int_a^b \vec{f}$.

Соотношение (IV, 1; 40) получается из (IV, 1; 41) после умножения его на выражение $\frac{n}{n+1} \cdot \frac{1}{b-a}$, стремящееся к $\frac{1}{b-a}$ при n , стремящемся к бесконечности, и добавления слагаемого $\frac{1}{n+1} \vec{f}(b)$, стремящегося к $\vec{0}$.

Если f является вещественной непрерывной функцией, то из теоремы о промежуточных значениях (следствие теоремы 33 гл. II) следует, что $\mathfrak{M}(f; [a, b])$ равна $f(c)$, где c — некоторая точка отрезка $[a, b]$.

§ 2. МЕРЫ РАДОНА НА ЛОКАЛЬНО КОМПАКТНОМ ПРОСТРАНСТВЕ

Меры Радона на компактном пространстве

Пусть X — компактное топологическое пространство. Обозначим через $\mathcal{C}(X)$ пространство непрерывных на X функций со скалярными значениями, обозначенное нами ранее через $(K^X)_c$ ¹.

Согласно теореме о максимуме (теорема 29 гл. II), непрерывная скалярная функция, определенная на компакте X , ограничена. Как отмечалось ранее, функция $\varphi \rightarrow \|\varphi\| = \sup_{x \in X} |\varphi(x)|$

¹) Это изменение обозначений может показаться странным. Но в дальнейшем нам придется рассматривать случай, когда X не является компактным, и мы тогда введем пространство $\mathcal{C}(X)$, не совпадающее с $(K^X)_c$. Почти всегда в качестве K будет браться поле комплексных чисел C .

является нормой в $\mathcal{C}(X)$. Сходимость в смысле этой нормы является равномерной сходимостью функций φ . В следствии теоремы 64 гл. II говорится о том, что $\mathcal{C}(X)$ является банаевым пространством.

Пространство, сопряженное к этому пространству, в соответствии с общими обозначениями записывается в виде $\mathcal{C}'(X)$. Элементы этого сопряженного пространства называются *мерами Радона в пространстве X*. Таким образом, если μ есть мера Радона, то она определяет некоторое отображение, которое каждому элементу φ из $\mathcal{C}(X)$, т. е. каждой непрерывной скалярной функции φ на X ставит в соответствие число $\mu(\varphi)$.

Определенное таким образом отображение линейно в том смысле, что

$$\mu(\varphi_1 + \varphi_2) = \mu(\varphi_1) + \mu(\varphi_2), \quad \mu(k\varphi) = k\mu(\varphi) \quad \text{для } k \in \mathbb{K}. \quad (\text{IV}, 2; 1)$$

Кроме того, оно непрерывно. Этот факт может выражаться различными способами: либо он означает, что для каждой последовательности скалярных непрерывных функций $\varphi_0, \varphi_1, \varphi_2, \dots$, равномерно сходящейся на X к 0, числовая последовательность $\mu(\varphi_n)$ сходится к нулю, либо он означает, что мера μ обладает такой нормой $\|\mu\|$, при которой справедливы соотношения

$$|\mu(\varphi)| \leq \|\mu\| \|\varphi\| \quad \text{и} \quad \|\mu\| = \sup_{\|\varphi\| \leq 1} |\mu(\varphi)|. \quad (\text{IV}, 2; 2)$$

С другой стороны, пространство $\mathcal{C}'(X)$, будучи сопряженным к банаеву пространству, само является банаевым пространством (теорема 50 гл. II). Его векторная структура определяется следующим образом.

Если μ и ν — две меры, то мера $\mu + \nu$ определяется по формуле

$$(\mu + \nu)(\varphi) = \mu(\varphi) + \nu(\varphi). \quad (\text{IV}, 2; 3)$$

Если k — скаляр, то мера $k\mu$ определяется равенством

$$(k\mu)(\varphi) = k\mu(\varphi). \quad (\text{IV}, 2; 4)$$

Норма меры μ определяется по формуле (IV, 2; 2).

Согласно общим обозначениям, выражение $\mu(\varphi)$ может быть также записано в виде $\mu \cdot \varphi$ или $\langle \mu, \varphi \rangle$.

Примеры мер Радона

1-й пример. Мерой Дирака в точке $a \in X$ называется мера, определенная формулой

$$\delta(a)(\varphi) = \varphi(a). \quad (\text{IV}, 2; 5)$$

Говорят, что эта мера определена единичной массой в точке а множества X . Если X является векторным конечномерным пространством; то мерой Дирака (без уточнения точки a) называют меру, относящуюся к нулю пространства X .

Пусть теперь заданы счетная последовательность точек a_0, a_1, a_2, \dots пространства X и счетная последовательность комплексных чисел c_0, c_1, c_2, \dots . Тогда через $\sum_v c_v \delta_{(a_v)}$ обозначается мера μ , определяемая формулой

$$\left(\sum_v c_v \delta_{(a_v)} \right) \cdot \varphi = \sum_v c_v \varphi(a_v). \quad (\text{IV}, 2; 6)$$

Это выражение, наверное, будет иметь смысл, если сходится ряд $\sum_v |c_v|$, что мы будем всегда предполагать выполненным.

Эту меру называют атомической и говорят, что она определена счетным множеством масс c_v , расположенных в точках a_v пространства X .

Докажем, что в том случае, когда все a_v различны, норма этой меры равна сумме числового ряда

$$\left\| \sum_v c_v \delta_{(a_v)} \right\| = \sum_v |c_v|. \quad (\text{IV}, 2; 7)$$

Прежде всего, если функция φ по модулю не превосходит 1, то

$$|\mu(\varphi)| \leq \sum_v |c_v|, \quad (\text{IV}, 2; 7_2)$$

откуда вытекает неравенство $\|\mu\| \leq \sum_v |c_v|$. С другой стороны, по определению суммы сходящегося ряда с положительными членами существует такое конечное подмножество J множества индексов, при котором сумма $\sum_{v \in J} |c_v|$ будет $\geq \left(\sum_v |c_v| \right) - \frac{\epsilon}{2}$.

Обозначим теперь через φ непрерывную комплексную функцию над X , не превосходящую по модулю 1 и принимающую в каждой из точек $a_v, v \in J$, значение $\bar{c}_v / |c_v|$. Теперь видно, что

$$\mu(\varphi) = \sum_{v \in J} |c_v| + \sum_{v \notin J} c_v \varphi(a_v), \quad (\text{IV}, 2; 8)$$

¹⁾ Пусть X , например, является интервалом $[a, b]$ прямой \mathbb{R} . Тогда, очевидно, можно найти комплексную непрерывную функцию φ , не превосходящую по модулю 1 и принимающую в конечном числе точек из X данные значения, равные по модулю 1. Пусть d_1, d_2, \dots, d_l — эти точки. Тогда достаточно взять функцию φ аффинной в каждом из интервалов $[a, d_1], [d_1, d_2], \dots, [d_l, b]$. Эта возможность вовсе не очевидна, если X является произвольным компактом. Однако это существенно в нашем доказательстве. В дальнейшем мы убедимся, что это возможно (следствие 3 теоремы 11).

откуда следует оценка снизу

$$|\mu(\varphi)| \geq \left(\sum_v |c_v| - \frac{\varepsilon}{2} \right) - \sum_{v \notin J} |c_v| \geq \sum_v |c_v| - \varepsilon. \quad (\text{IV}, 2; 9)$$

Так как $\|\varphi\| = 1$, то

$$\|\mu\| \geq |\mu(\varphi)| \geq \sum_v |c_v| - \varepsilon. \quad (\text{IV}, 2; 10)$$

Поскольку $\varepsilon > 0$ произвольно, получаем $\|\mu\| \geq \sum_v |c_v|$, откуда и вытекает требуемый результат.

Заметим также, что обозначение Σ , написанное ранее без обоснования, может быть теперь строго обосновано: речь идет о *нормально сходящемся ряде векторов* *банахова пространства* $\mathcal{C}'(X)$. В самом деле, так как $\|c_v \delta_{(a_v)}\| = |c_v|$ и $\sum_v |c_v| < +\infty$, то ряд $\sum_v c_v \delta_{(a_v)}$ нормально сходится к некоторому элементу из $\mathcal{C}'(X)$.

Пусть μ_1 — этот элемент. Поскольку для фиксированного φ функция $v \rightarrow v(\varphi)$ является линейной непрерывной формой над $\mathcal{C}'(X)$, то из теоремы 60 гл. II получаем

$$\mu_1(\varphi) = \left(\sum_v c_v \delta_{(a_v)} \right) \cdot \varphi = \sum_v (c_v \delta_{(a_v)} \cdot \varphi) = \sum_v c_v \varphi(a_v) = \mu(\varphi)^1. \quad (\text{IV}, 2; 11)$$

Следовательно, $\mu = \mu_1$ можно записать в виде $\sum_v c_v \delta_{(a_v)}$.

2-й пример. Возьмем в качестве X компактный интервал $[a, b]$ прямой \mathbb{R} . Тогда интеграл $\varphi \rightarrow \int_{[a, b]} \varphi(x) dx$ определяет некоторую меру Радона над X . Из неравенства (IV, 1; 19) видно, что норма этой меры $\leq b - a$. Впрочем, если вычислить значение этой меры на функции $\varphi \equiv 1$, то мы получим неравенство $\|\mu\| \geq \left| \int_{[a, b]} \varphi \right| = b - a$. Итак, норма такой меры равна $b - a$.

Более общо, если p является интегрируемой по Риману комплекснозначной функцией, определенной на $[a, b]$, то она определяет меру по формуле

$$\mu(\varphi) = \int_{[a, b]} p(x) \varphi(x) dx. \quad (\text{IV}, 2; 12)$$

¹⁾ По определению.

Норма этой меры, очевидно, $\leq \int |p(x)| dx$. Позже мы покажем, что норма в точности равна этому интегралу.

Определенные таким образом меры, в противоположность атомическим мерам, называются *рассеянными* в том смысле, что они не содержат точечной массы.

Говорят также, что эта мера является мерой с плотностью p по отношению к мере dx , и записывают ее через $p dx$ или $p(x) dx^1$.

В силу обозначения dx , принятого для меры $\Phi \rightarrow \int_{[a, b]} \Phi$, произ-

вольная мера вместо μ обозначается через $d\mu$, а выражение $\mu(\Phi)$ называется *интегралом от функции Φ по мере μ или $d\mu$* и вместо $\mu(\Phi)$ часто обозначается через

$$\int_X \Phi(x) d\mu(x), \text{ или } \int_X \Phi d\mu, \text{ или } \int_X \Phi \mu. \quad (\text{IV}, 2; 13)$$

Естественным образом можно образовать меру как сумму двух мер, изученных в примерах 1 и 2, т. е. определить ее по формуле

$$\mu(\Phi) = \sum_v c_v \Phi(a_v) + \int_{[a, b]} p(x) \Phi(x) dx. \quad (\text{IV}, 2; 14)$$

Говорят, что эта мера является суммой точечных масс c_v , расположенных в точках a_v интервала $[a, b]$, и рассеянной меры плотности p по отношению к мере dx . Ее обозначают через

$$\mu = \sum_v c_v \delta_{(a_v)} + p dx. \quad (\text{IV}, 2; 15)$$

¹⁾ Обозначение \int для меры $\Phi \rightarrow \int_{[a, b]} \Phi(x) dx$ совершенно корректно. Обо-

значение dx логически противовесственно, ибо x фигурирует в интеграле как *переменная интегрирования*, которую можно обозначить или y , или t , или другим символом. Точно так же более корректно обозначать меру через $p(x)dx$, а не в виде $p dx$, подчеркивая этим двойным введением переменной x ее промежуточную роль. При этом $p(x) dx$ может быть заменена на $p(y) dy$ или $p(t) dt$. Эта некорректность аналогична той, которая отмечалась при употреблении выражения «функция x^2 » вместо «функция $x \rightarrow x^2$ » (см. стр. 11). Однако, как и в случае функций, часто допускают вольность речи. Если возникает сомнение, то говорят «мера $1(x)dx$ » вместо «мера dx », поскольку здесь речь идет о мере $p(x)dx$, где $p(x) = 1$ для всех x , т. е. $p(x) \equiv 1$. Можно так же для dx , как и для $p dx$, писать полностью «мера

$\Phi \rightarrow \int_{[a, b]} \Phi(x) p(x) dx$.

Норма этой меры, очевидно, оценивается сверху числом

$$\sum_v |c_v| + \int_{[a, b]} |p(x)| dx. \quad (\text{IV}, 2; 16)$$

Покажем, что если все a_v различны, то норма в точности равна этой величине.

Зададим $\varepsilon > 0$ и определим конечное множество индексов J так, чтобы $\sum_{v \in J} |c_v| \leq \varepsilon/6$. Поскольку p интегрируема по Риману, то можно найти такую ступенчатую на $[a, b]$ функцию q , для которой имеет место неравенство

$$\int_{[a, b]} |p(x) - q(x)| dx \leq \frac{\varepsilon}{6}. \quad (\text{IV}, 2; 16_2)$$

Обозначим теперь через $d_0 = a, d_1, d_2, \dots, d_n = b$ множество точек разбиения, соответствующего функции q , и точек $a_v, v \in J$. Для каждого $i = 0, 1, 2, \dots, n-1$ можно теперь найти такой интервал $[d'_i, d''_{i+1}] \subset]d_i, d_{i+1}[$, при котором имело бы место неравенство

$$\sum_{i=0}^{n-1} |q_i| ((d'_i - d_i) + (d_{i+1} - d''_{i+1})) \leq \frac{\varepsilon}{6}, \quad (q_i = q \text{ в }]d_i, d_{i+1}[). \quad (\text{IV}, 2; 16_3)$$

Определим теперь функцию φ следующим образом. В каждом из интервалов $[d'_i, d''_{i+1}]$ будем считать ее равной постоянной $\bar{q}_i/|q_i|$. В каждой из точек $d_i, i = 1, 2, \dots, n$, совпадающей с точкой $a_v, v \in J$, положим ее равной числу $\bar{c}_v/|c_v|$. В интервалах, где она еще не определена, определим ее как аффинную функцию. Очевидно, модуль φ всюду не превосходит 1.

С другой стороны, имеют место неравенства

$$\left| \sum_v c_v \varphi(a_v) - \sum_v |c_v| \right| \leq \left| \sum_v c_v \varphi(a_v) - \sum_{v \in J} c_v \varphi(a_v) \right| + \left| \sum_{v \in J} c_v \varphi(a_v) - \sum_{v \in J} |c_v| \right| + \left| \sum_{v \in J} |c_v| - \sum_v |c_v| \right|. \quad (\text{IV}, 2; 16_4)$$

Второй член правой части равен нулю, поскольку $\varphi(a_v) = \frac{\bar{c}_v}{|c_v|}$ для $v \in J$. Первый и третий члены не превосходят $\varepsilon/6$. Поэтому

$$\left| \sum_v c_v \varphi(a_v) - \sum_v |c_v| \right| \leq \frac{2\varepsilon}{6} = \frac{\varepsilon}{3}. \quad (\text{IV}, 2; 16_5)$$

Далее,

$$\begin{aligned} \left| \int_{[a, b]} p(x) \varphi(x) dx - \int_{[a, b]} |p(x)| dx \right| &\leq \\ &\leq \left| \int_{[a, b]} p(x) \varphi(x) dx - \int_{[a, b]} q(x) \varphi(x) dx \right| + \\ &+ \left| \int_{[a, b]} q(x) \varphi(x) dx - \int_{[a, b]} |q(x)| dx \right| + \\ &+ \left| \int_{[a, b]} |q(x)| dx - \int_{[a, b]} |p(x)| dx \right|. \quad (\text{IV}, 2; 16_6) \end{aligned}$$

Согласно (IV, 2; 16₂), первый и третий члены правой части не превосходят $\varepsilon/6$. Поскольку в интервале $[d'_i, d'_{i+1}]$ функция φ равна $q_i/|q_i|$, то, в силу (IV, 2; 16₃), второй член не больше $2\varepsilon/6 = \varepsilon/3$. Поэтому

$$\left| \int_{[a, b]} p(x) \varphi(x) dx - \int_{[a, b]} |p(x)| dx \right| \leq \frac{2\varepsilon}{6} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3}. \quad (\text{IV}, 2; 16_7)$$

Добавляя сюда (IV, 2; 16₅) и учитывая (III, 2; 14), для построенной функции φ получаем неравенство

$$\left| \mu(\varphi) - \left(\int_{[a, b]} |p(x)| dx + \sum_v c_v \right) \right| \leq \varepsilon, \quad (\text{IV}, 2; 16_8)$$

доказывающее наше утверждение, а именно то, что $\|\mu\|$ равна (IV, 2; 16).

Важно подчеркнуть, что *мера с плотностью p не изменяется, если функция p меняется в конечном числе точек*.

Плотность определяет меру, но различные плотности могут дать одну и ту же меру.

З-й пример. Хотя мы еще не знакомились с кратными интегралами, мы можем, однако, заметить, что если X является замкнутым параллелепипедом в \mathbb{R}^n , т. е. множеством, определенным неравенствами $a_i \leq x_i \leq b_i$, $i = 1, 2, \dots, n$, то на X можно определить меру $dx = dx_1 dx_2 \dots dx_n$ по формуле

$$\Phi \rightarrow \int \int \dots \int \Phi(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n. \quad (\text{IV}, 2; 17)$$

$$\prod_{i=1}^n [a_i, b_i]$$

Норма этой меры равна объему параллелепипеда X , т. е.

$$\prod_{i=1}^n (b_i - a_i).$$

Меры на локально компактном пространстве

Множество $X = \mathbb{R}$ не компактно. Тем не менее линейные формы $\varphi \rightarrow \varphi(a)$ и $\varphi \rightarrow \int \varphi(x) dx$ продолжают сохранять смысл по крайней мере в том случае, когда на φ налагаются некоторые ограничения, например условие, заключающееся в том, что φ имеет компактный носитель. Поэтому имеет смысл дать следующее определение.

Пусть X — локально компактное топологическое пространство. Обозначим через $\mathcal{C}(X)$ векторное пространство непрерывных скалярнозначных функций φ , определенных на X и имеющих компактный носитель¹⁾ (этот носитель, естественно, не уточняется; каждая из функций φ этого пространства имеет компактный носитель K , но K зависит от φ). Заметим, впрочем, что *ни в коем случае не следует смешивать функцию φ , определенную и непрерывную на всем X и имеющую компактный носитель, с функцией, определенной и непрерывной на некотором компакте из X .*

Например, функция, изображенная на рис. 13, определена и непрерывна на интервале $[a, b]$. Эта функция не является функцией, определенной на всей числовой прямой \mathbb{R} и имеющей $[a, b]$ в качестве носителя. Если ее продолжить на \mathbb{R} , положив вне $[a, b]$ равной нулю, то она не будет непрерывной на \mathbb{R} , поскольку точки a и b являются ее точками разрыва.

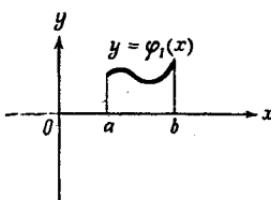


Рис. 13.

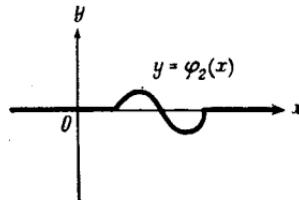


Рис. 14.

Напротив, функция, изображенная на рис. 14, определена на \mathbb{R} , непрерывна, имеет компактный носитель, а, следовательно, принадлежит пространству $\mathcal{C}(\mathbb{R})$.

Если K является компактом X , то через $\mathcal{C}_K(X)$ будем обозначать векторное подпространство пространства $\mathcal{C}(X)$, образованное функциями φ , носители которых лежат в компакте K . Функция φ из $\mathcal{C}(X)$ принадлежит $\mathcal{C}_K(X)$ тогда и только тогда, когда она равна нулю на $\mathbb{C}K$. Отсюда следует, что она равна

¹⁾ $\mathcal{C}(X)$ является подпространством $(\mathbb{K}^X)_{cb}$; см. примечание на стр. 439.

нулю не только на $C\bar{K}$, но также и на границе компакта K , поскольку каждая точка этой границы является пределом точек множества $C\bar{K}$, а функция φ непрерывна. Не следует смешивать $\mathcal{C}_k(X)$ и $\mathcal{C}(K)$. Функция φ_1 (см. рис. 13) принадлежит $\mathcal{C}([a, b])$. Функция φ_2 (см. рис. 14) принадлежит пространству $\mathcal{C}_{[a, b]}(\mathbb{R})$.

Пространство $\mathcal{C}(X)$ является объединением векторных пространств $\mathcal{C}_k(X)$, когда множество K пробегает все компактные подпространства X .

Если в пространстве $\mathcal{C}_k(X)$ ввести норму $\|\varphi\| = \sup_{x \in X} |\varphi(x)|$,

то оно станет банаевым пространством. В самом деле, $(K^x)_{cb}$ является в этой норме банаевым пространством (следствие 3 теоремы 65 гл. II), а $\mathcal{C}_k(X)$ является в нем замкнутым векторным подпространством. (В самом деле, если при n , стремящемся к бесконечности, последовательность $\varphi_n \in \mathcal{C}_k(X)$ сходится к φ в $(K^x)_{cb}$, т. е. сходится равномерно и тем более просто, то в силу равенства $\varphi_n(x) = 0$ для $x \in C\bar{K}$ получим также $\varphi(x) = 0$, откуда $\varphi \in \mathcal{C}_k(X)$.) Из теоремы 43 гл. II теперь следует, что $\mathcal{C}_k(X)$ является банаевым пространством. Естественно, такую же норму можно ввести и на пространстве $\mathcal{C}(X)$. Однако можно показать, что тогда оно не будет банаевым. Во всяком случае это не та норма, которая нас будет интересовать в дальнейшем.

Определение. *Мерой Радона на локально компактном пространстве X называется линейная форма μ , определенная на векторном пространстве $\mathcal{C}(X)$, сужение которой на каждое подпространство $\mathcal{C}_k(X)$, где K является компактом X , непрерывно.*

Выясним точный смысл этого определения.

Мера μ определяет некоторое отображение, которое каждой функции φ , принадлежащей пространству $\mathcal{C}(X)$, т. е. каждой скалярной непрерывной функции φ , определенной на X , имеющей компактный носитель, ставит в соответствие скаляр $\mu(\varphi)$. Это соответствие должно быть линейным в смысле (IV, 2; 1). Однако мы не говорим, что это соответствие непрерывно на $\mathcal{C}(X)$, поскольку в этом векторном пространстве мы не вводим ни топологию, ни норму. Мы можем сказать только, что сужение этого отображения на $\mathcal{C}_k(X)$ должно быть непрерывным. Теперь, каким бы ни был компакт K из X и какова бы ни была последовательность непрерывных функций $\varphi_0, \varphi_1, \varphi_2, \dots$ с носителями в K , равномерно сходящаяся к 0 при n , стремящемся к $+\infty$, последовательность комплексных чисел $\mu(\varphi_n)$ сходится к нулю. Мы получаем, что каждому компакту K из X можно поставить в соответствие норму меры относительно K , а именно

такую норму, при которой имеет место формула

$$|\mu(\varphi)| \leq \|\mu\|_K \|\varphi\| \quad \text{для } \varphi \in \mathcal{C}_K(X), \quad (\text{IV}, 2; 18)$$

где $\|\mu\|_K = \sup_{\substack{\varphi \in \mathcal{C}_K(X) \\ \|\varphi\| \leq 1}} |\mu(\varphi)|$.

Примеры мер Радона

Пример 1. Мера $\delta_{(a)}$, определенная по формуле (IV, 2; 5), также является некоторой мерой Радона на X . Аналогично, сумма $\sum_v c_v \delta_{(a_v)}$ будет некоторой мерой Радона при условии, что ряд $\sum_v |c_v|$ является «локально сходящимся», т. е. частичная сумма $\sum_{a_v \in K} |c_v|$ сходится для любого компакта K из X . Формула (IV, 2; 6) в этом случае определяет линейную форму на $\mathcal{C}(X)$. С другой стороны, для каждого компакта K для любой $\varphi \in \mathcal{C}_K(X)$ имеет место неравенство $\left| \sum_v c_v \varphi(a_v) \right| \leq \left(\sum_{a_v \in K} |c_v| \right) \|\varphi\|$, из которого вытекает непрерывность рассматриваемого отображения. Отсюда следует также, что $\|\mu\|_K \leq \sum_{a_v \in K} |c_v|$.

Однако можно получить более точное утверждение об этой норме. В самом деле, в каждой точке a_v границы \dot{K} множества K функция φ из $\mathcal{C}_K(X)$ необходимо равна нулю (каждая окрестность этой граничной точки содержит точки дополнения к K , где функция φ равна нулю и, кроме того, φ непрерывна (см. по этому поводу стр. 447)). Поэтому при оценке суммы $\sum c_v (a_v)$ точки a_v границы \dot{K} можно не учитывать, так что имеет место неравенство $\|\mu\|_K \leq \sum_{a_v \in \dot{K}} |c_v|$, где \dot{K} — внутренность K .

Теперь, пользуясь методом, примененным на стр. 443, нетрудно доказать, что $\|\mu\|_K$ в точности равна $\sum_{a_v \in \dot{K}} |c_v|$.

Пример 2. Мера \int или dx на \mathbb{R} , как и в случае компактного интервала $[a, b]$ прямой \mathbb{R} , определяется точно так же с помощью отображения $\varphi \rightarrow \int \varphi$. Ее называют *канонической мерой прямой \mathbb{R}* или *мерой Лебега на \mathbb{R}* . Для каждого интервала $[a, b]$ ее норма $\|\mu\|_{[a, b]}$ равна $b - a$.

Если ρ является функцией, определенной на \mathbb{R} , комплексно-значной и локально интегрируемой по Риману, т. е. интегрируемой на каждом ограниченном интервале, то она определяет меру Радона μ по формуле (IV, 2; 12); мы обозначаем ее через

pdx или $p(x)dx$ и называем p плотностью по отношению к мере Лебега.

Норма относительно компактного интервала $[a, b]$ равна при этом $\|\mu\|_{[a, b]} = \int_{[a, b]} |p(x)|dx$. Заметим, что это утверждение не вытекает непосредственно из утверждения, полученного из формулы (IV, 2; 16). В самом деле, функция φ из $\mathcal{C}_{[a, b]}(\mathbb{R})$ в противоположность функциям из $\mathcal{C}([a, b])$ (см. стр. 447) равна нулю в точках a и b .

Если же выбрать a', b' , $a < a' < b' < b$, так, чтобы $\int_{[a, a']} |p(x)|dx + \int_{[b', b]} |p(x)|dx \leq \frac{\varepsilon}{3}$, и если непрерывная на интервале $[a', b']$ функция φ определена так, что

$$\left| \int_{[a', b']} p(x) \varphi(x) dx - \int_{[a', b']} |\dot{p}(x)| dx \right| \leq \frac{\varepsilon}{3},$$

согласно тому, как это было сделано на стр. 444, то можно продолжить функцию φ на интервал $[a, b]$ таким образом, чтобы она была аффинной в интервалах $[a, a']$, $[b', b]$ и принимала нулевое значение в точках a и b . Представляя теперь интервал интегрирования $[a, b]$ в виде объединения трех интервалов $[a, a']$, $[a', b']$, $[b', b]$, получаем следующее неравенство:

$$\left| \int_{[a, b]} p(x) \varphi(x) dx - \int_{[a, b]} |p(x)| dx \right| \leq \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

доказывающее предыдущее утверждение. (Затем можно, естественно, ввести меры вида $\sum_v c_v \delta_{(a_v)} + p dx$.)

В силу того что функции φ имеют компактные носители, мы смогли ввести меры предыдущего типа, где функция p может возрастать на бесконечности с произвольной скоростью. Например, p может быть функцией e^{x^2} или функцией $e^{e^{x^2}}$, и она вполне определит меру Радона pdx .

Нормой меры μ называется точная верхняя грань чисел $\|\mu\|_K$ по всем компактам K из X . Эта норма, обозначаемая через $\|\mu\|$, неотрицательна, конечна или равна $+\infty$ (и, следовательно, несмотря на свое название, не является нормой). Очевидно, ее можно определить по формуле

$$\|\mu\| = \sup_{\substack{\varphi \in \mathcal{C}(X) \\ \|\varphi\|_1 \leq 1}} |\mu(\varphi)|. \quad (\text{IV, 2; 19})$$

Хотя пространство мер невозможно рассматривать как сопряженное к пространству $\mathcal{C}(X)$, его, однако, обозначают через $\mathcal{C}'(X)$. Это множество является, естественно, векторным пространством с тем же самым определением суммы и умноже-

ния на скаляр, что и в том случае, когда X компактно. Однако оно не нормировано. Напротив, подпространство пространства $\mathcal{C}'(X)$, образованное из таких μ , для которых $\|\mu\|$ конечна, является векторным нормированным пространством. Легко видеть, что это пространство является банаевым.

Применения к механике и физике

В механике часто приходится рассматривать распределение масс в пространстве. В свою очередь в физике сталкиваются с распределением электрических зарядов. Меры Радона над аффинным евклидовым трехмерным пространством E дают хорошие модели для этих распределений масс и распределений зарядов. То, что в механике называется точечной массой m , помещенной в точке a , или в физике — электрическим зарядом e , расположенным в точке a , является не чем иным, как мерой Радона $m\delta_{(a)}$ или $e\delta_{(a)}$.

Если в физике рассматривается распределение масс или зарядов, определенных с помощью плотности $p(x)$, то это означает, что рассматривается мера Радона $p(x)dx$, где dx является мерой трехмерного объема в E . В случае прямоугольной системы координат это будет мера $p(x, y, z) dx dy dz$.

В механике и физике часто рассматривается распределение масс или зарядов, расположенных на замкнутой поверхности Σ , принадлежащей классу C^1 , в трехмерном евклидовом пространстве с «поверхностной плотностью» $p(x)$. Эта величина является мерой Радона, определенной равенством

$$\mu(\phi) = \int_{\Sigma} \phi(x) p(x) dS^1. \quad (\text{IV}, 2; 20)$$

Ее удобнее будет обозначать через $p dS$. Точно так же распределение зарядов или масс, расположенных на замкнутой кривой L класса C^1 , линейной плотности p относительно дуги ds определяется по формуле

$$\mu(\phi) = \int_L \phi(x) p(x) ds \quad (\text{IV}, 2; 21)$$

и может быть обозначено через $p ds$.

Векторные меры

Рассмотренные выше меры были скалярными.

Пусть теперь \vec{E} — векторное нормированное пространство. Мерой $\vec{\mu}$ на X со значениями в \vec{E} будем называть линейное ото-

¹⁾ Такие «поверхностные интегралы» будут определены позже в § 10.

бражение $\mathcal{C}(X)$ в \vec{E} , сужение которого на каждое подпространство $\mathcal{C}_K(X)$, где K является компактом X , непрерывно. Таким образом, здесь для $\varphi \in \mathcal{C}(X)$ величина $\vec{\mu}(\varphi)$ или $\vec{\mu} \cdot \varphi$ является элементом \vec{E} . Как прежде, определяются нормы $\|\vec{\mu}\|_K$ и $\|\vec{\mu}\|$. В качестве примера мер при полном \vec{E} можно взять $\vec{\mu} = \sum_v \vec{c}_v \delta_{(a_v)}$, где $\vec{c}_v \in \vec{E}$, или $\vec{d}\mu = \vec{p}(x) dx$, где \vec{p} является функцией, определенной на \mathbb{R} , со значениями в \vec{E} и локально интегрируемой по Риману. При этом $\sum_v \vec{c}_v \varphi(a_v) \in \vec{E}$ (предполагая, как ранее, что на любом компакте K множества X $\sum_{a_v \in K} \|\vec{c}_v\| < \infty$, получим, согласно теореме 55 гл. II, что ряд $\sum_v \vec{c}_v \varphi(a_v)$ сходится) и $\int \vec{p}(x) \varphi(x) dx \in \vec{E}$. Для норм имеют место неравенства

$$\left\| \sum_v \vec{c}_v \delta_{(a_v)} \right\|_K \leq \sum_{a_v \in K} \|\vec{c}_v\| \quad \text{и} \quad \|\vec{p} dx\|_{[a, b]} \leq \int_{[a, b]} \|\vec{p}(x)\| dx^1.$$

¹⁾ Мы всегда писали $\mathcal{C}(X)$ без уточнения поля скаляров. Пусть \vec{E} — векторное нормированное пространство над комплексным полем. Тогда \mathbb{C} -мера $\vec{\mu}$ со значениями в \vec{E} для поля скаляров $\mathbb{K} = \mathbb{C}$ является отображением пространства $\mathcal{C}(X; \mathbb{C})$ комплексных непрерывных функций φ с компактным носителем в пространство \vec{E} , линейным относительно \mathbb{C} . \mathbb{R} -мера $\vec{\mu}$ со значениями в \vec{E} для поля скаляров $\mathbb{K} = \mathbb{R}$ является отображением пространства $\mathcal{C}(X; \mathbb{R})$ вещественных непрерывных функций φ с компактным носителем в \vec{E} , линейным относительно \mathbb{R} . Мера $\vec{\mu}$ для поля \mathbb{C} определяет меру $\mu_{\mathbb{R}}$ для поля \mathbb{R} . Мера $\mu_{\mathbb{R}}$ является лишь сужением на $\mathcal{C}(X; \mathbb{R})$ функции $\vec{\mu}: \mathcal{C}(X; \mathbb{C}) \rightarrow \vec{E}$.

Так, например, если $\vec{E} = \mathbb{C}$, то $\vec{\mu}$ является скалярной мерой, однако $\mu_{\mathbb{R}}$ является уже векторной мерой со значениями в векторном пространстве \mathbb{C} размерности 2 над \mathbb{R} .

Пространство скалярных \mathbb{K} -мер (или \mathbb{K} —сопряженное пространство к $\mathcal{C}(X; \mathbb{K})$) обозначается через $\mathcal{C}'(X; \mathbb{K})$.

Рассмотрим теперь при тех же условиях векторное нормированное пространство \vec{E} над полем \mathbb{C} , и пусть \vec{v} является некоторой мерой относительно \mathbb{R} , т. е. является некоторым \mathbb{R} -линейным отображением пространства $\mathcal{C}(X; \mathbb{R})$ в \vec{E} . Тогда возможно единственным образом продолжить ее на некоторую меру $\vec{v}_{\mathbb{C}}$ относительно \mathbb{C} , полагая $\vec{v}_{\mathbb{C}}(\varphi_1 + i\varphi_2) = \vec{v}(\varphi_1) + i\vec{v}(\varphi_2)$. При этом $(\vec{v}_{\mathbb{C}})_{\mathbb{R}} = \vec{v}$ и $(\mu_{\mathbb{R}})_{\mathbb{C}} = \vec{\mu}$.

Заметим, что нормы $\vec{\mu}$ и $\mu_{\mathbb{R}}$ (или \vec{v} и $\vec{v}_{\mathbb{C}}$) могут быть различными. Рассмотрим, например, скалярную меру $\mu = \delta_{(a)} + i\delta_{(b)}$. Ее норма относительно

Разложение единицы

Приведем одну теорему относительно вещественных непрерывных функций на локально компактном пространстве X , которой мы будем пользоваться в теории мер.

Теорема 11. Пусть X — локально компактное пространство, «счетное в бесконечности», т. е. являющееся объединением счетного множества компактов. Пусть, кроме того, $(\Omega_i)_{i \in I}$ — покрытие (конечное или нет) пространства X открытыми множествами. Тогда существует система вещественных непрерывных функций α_i , зависящих от того же множества индексов I , такая, что $0 \leq \alpha_i \leq 1$, α_i имеет носитель в Ω_i , каждая точка X имеет окрестность, в которой лишь конечное число функций α_i не равно тождественно нулю, сумма $\sum_{i \in I} \alpha_i$ тождественно равна 1 на X . Если X является многообразием класса C^m (т. е. конечно или $= +\infty$) над полем вещественных чисел \mathbb{R} , то функции α_i можно брать из класса C^{m+1} .

Замечания. 1°) По-видимому, вместе с образованием суммы $\sum_{i \in I} \alpha_i$ возникает задача обоснования ее сходимости, тем более досадная, что I может не быть счетным. Однако в условии указывается, что в каждой точке $x \in X$ и даже во всех точках окрестности x эта сумма конечна, поскольку все ее члены, кроме конечного числа их, равны нулю.

При локально компактном X мы приходим к тому, что на каждом компакте из X все α_i , кроме конечного числа их, равны тождественно нулю. В этом случае говорят, что система α_i локально конечна.

2°) Хорошо видно, почему функции α_i определяют разложение единицы: имеет место разложение 1 в сумму функций α_i с носителями в заданных открытых множествах Ω_i .

Если Ω_i компакты, то α_i имеют компактные носители и, следовательно, принадлежат $\mathcal{C}(X)$. Говорят, что разложение единицы $(\alpha_i)_{i \in I}$ подчинено покрытию $(\Omega_i)_{i \in I}$.

скалярного поля \mathbb{C} равна $\sup|\varphi(a) + i\varphi(b)|$, где \sup берется по всем комплексным функциям φ , не превосходящим по модулю 1. Величина эта равна 2. Норма этой же меры в поле скаляров \mathbb{R} вычисляется как тот же \sup , взятый по вещественным функциям φ , не превосходящим по модулю 1, и равна $\sqrt{2}$. Почти всегда берут $K = C$.

1) Это улучшение (переход в класс C^m) в интегральном исчислении не играет никакой роли. Однако оно нам понадобится в теории распределений.

Существенным здесь является то, что рассматривается многообразие C^m по отношению к полю вещественных чисел. Утверждение неверно в случае поля комплексных чисел. Далее мы увидим, что при доказательстве леммы 1 будет применяться поле вещественных чисел.

3°) Если X локально компактно, но не является счетным в бесконечности, то результат не верен. Однако имеет место более слабое утверждение, достаточное для некоторых приложений, которое мы здесь приводить не будем (оно основывается на леммах 1 и 2, верных для произвольного локально компактного пространства X).

Нормированное векторное конечномерное пространство локально компактно и счетно в бесконечности, поскольку оно является объединением шаров $\|x\| \leq n, n \in \mathbb{N}$.

4°) Пространство X , для которого справедлива теорема о разложении единицы, называется *паракомпактным* пространством. Локально компактное, счетное в бесконечности пространство паракомпактно. Можно показать, что метризуемое пространство паракомпактно.

Для доказательства мы предварительно рассмотрим три леммы.

Лемма 1. Пусть Ω — открытое множество локально компактного пространства X и a — точка Ω . Тогда существует непрерывная функция γ , удовлетворяющая условию $0 \leq \gamma \leq 1$, с компактным носителем в Ω , такая, что $\gamma(a) > 0$. Если X является многообразием класса C^m , то γ можно брать в классе C^m .

Естественно, что лемма тем интереснее, чем меньше окрестность Ω точки a .

1°) Пусть X — произвольное локально компактное пространство. Теорема верна в общем случае, но мы ее докажем только для того случая, когда локально компактное пространство X метризуемо. Итак, предположим, что в нем введена некоторая метрика. Поскольку Ω является открытым множеством, содержащим a , а X — локально компактно, то найдется компактный шар $B(a, \eta)$ с центром в точке a и радиусом $\eta > 0$, содержащийся в Ω ¹⁾. Обозначим теперь через F вещественную функцию, непрерывную на полупрямой \mathbb{R}_+ (вещественные числа ≥ 0), удовлетворяющую неравенствам $0 \leq F \leq 1$ и такую, что $F(0) > 0$ и $F(t) = 0$ для $t \geq \eta$. Можно будет, например, взять функцию, определенную по формуле

$$F(t) = \begin{cases} 0 & \text{для } t > \eta, \\ \eta - t & \text{для } 0 \leq t \leq \eta. \end{cases} \quad (\text{IV}, 2; 22)$$

Функция $x \rightarrow F(d(a, x))$ обладает требуемыми свойствами. Она непрерывна как суперпозиция двух непрерывных функций.

¹⁾ В самом деле, a имеет компактную окрестность V . Так как $\Omega \cap V$ является окрестностью точки a , то существует такое $\eta > 0$, что замкнутый шар $B(a, \eta)$ лежит в $\Omega \cap V$. Этот шар лежит в Ω и, будучи замкнутым в компакте V , компактен.

Поскольку $d(a, a) = 0$ и $F(0) = \eta$, то при $x = a$ она принимает значение η . Она равна нулю вне шара $B(a, \eta)$, поскольку вне этого шара $d(a, x) > \eta$, и $F(t) = 0$ при $t \geq \eta$. Отсюда следует, что ее носитель содержится в этом шаре, являющемя компактом в Ω . Понятно, почему метризуемость пространства является *существенным упрощением* задачи. В самом деле, метрика определяет бесконечное множество вещественных непрерывных функций в метрическом пространстве, а именно, функции $x \rightarrow d(a, x)$. Напротив, если мы знаем только, что X является топологическим пространством, то, кроме постоянных функций, мы не знаем *a priori* других вещественных непрерывных функций на X .

2°) Предположим теперь, что X является многообразием V размерности n над полем вещественных чисел класса C^m (возможно, $m = +\infty$)¹⁾. Пусть Φ является некоторой картой V с образом, содержащим точку a . Кarta Φ является гомеоморфизмом класса C^m некоторого открытого множества \mathcal{O}_1 из \mathbb{R}^n на $\Phi(\mathcal{O}_1) \subset V$. Пусть $\Omega_1 = \Phi^{-1}(\Omega)$ и $a_1 = \Phi^{-1}(a)$.

Предположим, что лемма доказана в \mathbb{R}^n для функции класса C^m . Другими словами, предположим, что найдена некоторая функция γ_1 класса C^m на \mathbb{R}^n с компактным носителем K_1 , содержащимся в Ω_1 , удовлетворяющая неравенству $0 \leq \gamma_1 \leq 1$ и такая, что $\gamma_1(a_1) > 0$. Тогда лемма будет доказанной также и в V . В самом деле, определим γ на V следующим образом: $\gamma = \gamma_1 \circ \Phi^{-1}$ в множестве $\Phi(\mathcal{O}_1) \subset \Omega$ и $\gamma = 0$ вне его.

Функция Φ^{-1} является обратной биекцией для Φ , т. е. отображением множества $\Phi(\mathcal{O}_1)$ на \mathcal{O}_1 . Функция γ принадлежит классу C^m . Действительно, пусть $b \in V$. Если $b \in \Phi(\Omega_1)$, то γ равна $\gamma_1 \circ \Phi^{-1}$, и поскольку Φ является картой класса C^m , а γ_1 — функцией класса C^m на \mathcal{O}_1 , то γ принадлежит классу C^m в окрестности точки b (теорема Зз гл. II). Если $b \notin \Phi(\Omega_1)$, то b лежит в открытом множестве $C_{V \setminus \Phi(K_1)}$ многообразия V , дополнительном к компакту $\Phi(K_1)$. Функция γ равна нулю в этом множестве и, следовательно, принадлежит классу C^m в окрестности точки b . Мы только что видели, что γ равна нулю в дополнении к компакту $\Phi(K_1)$, а, следовательно, ее носитель лежит в компакте $\Phi(K_1) \subset \Phi(\Omega_1) \subset \Omega$. Эта функция удовлетворяет неравенствам $0 \leq \gamma \leq 1$ и $\gamma(a) = \gamma_1(a_1) > 0$.

Остается решить задачу для функции γ_1 , принадлежащей классу C^m на \mathbb{R}^n . Поскольку число m может быть произвольным, то мы должны найти функцию γ_1 класса C^∞ . Введем в \mathbb{R}^n обычную евклидову метрику. Тогда функция $x_1 \rightarrow r = d(a_1, x_1)$ не-

1) Поскольку каждая точка многообразия имеет окрестность, гомеоморфную некоторому открытому множеству из \mathbb{R}^n , то существует также меньшая компактная окрестность. Многообразие является локально компактным пространством.

прерывна, но не дифференцируема в точке a_1 , однако ее квадрат r^2 , являясь полиномом 2-й степени относительно своих координат, принадлежит классу C^∞ .

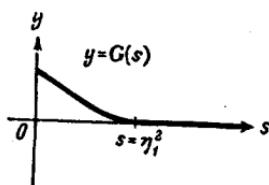
Пусть $\eta_1 > 0$ — такое число, что компактный шар $B(a_1, \eta_1)$ лежит в множестве Ω_1 . Если вместо F рассмотреть вещественную функцию G , определенную на \mathbb{R}_+ по формуле

$$G(s) = \begin{cases} 0 & \text{для } s \geq \eta_1^2, \\ -\frac{1}{\eta_1^2 - s} & \text{для } s < \eta_1^2, \\ e^{-1/(\eta_1^2 - s)} & \text{для } s < \eta_1^2, \end{cases} \quad (\text{IV}, 2; 23)$$

то получим, что она удовлетворяет неравенствам $0 \leq G \leq 1$, что $G(0) > 0$, что $G(s)$ равна нулю для $s \geq \eta_1^2$ и что G непрерывна (ибо, когда $s \rightarrow \eta_1^2$, $s < \eta_1^2$, $\lim_{s \rightarrow \eta_1^2} e^{-1/(\eta_1^2 - s)} = e^{-\infty} = 0$). Более того, G принадлежит классу C^∞ . В самом деле, последовательным дифференцированием можно убедиться, что для $s < \eta_1^2$ производные G имеют вид

$$G^{(k)}(s) = P_k \left(\frac{1}{\eta_1^2 - s} \right) e^{-\frac{1}{\eta_1^2 - s}}, \quad (\text{IV}, 2; 24)$$

где P_k — полином от одной переменной. При s , стремящемся к η_1^2 по значениям $< \eta_1^2$, функция $G^{(k)}(s)$ стремится к нулю (поскольку показательная функция растет быстрее любого полинома). Для $s > \eta_1^2$ имеем $G^{(k)}(s) \equiv 0$. Следовательно, $G^{(k)}(s)$ стремится к 0 при s , стремящемся к η_1^2 по произвольным значениям $\neq \eta_1^2$. Теорема 14 гл. III, примененная шаг за шагом к последовательным производным, дает, что G бесконечно дифференцируема, а все ее производные при $s = \eta_1^2$ равны нулю. Вот график этой функции G :



Р и с. 15

(Разложение Тейлора порядка m для функции G в окрестности точки $s = \eta_1^2$ сводится к остаточному члену. Ряд Тейлора функции G по степеням $s - \eta_1^2$ сходится, поскольку все его члены равны нулю, но он не представляет функцию для $s < \eta_1^2$.)

Функция $\gamma_1(x_1) = G((d(a_1, x_1))^2) = G(r^2)$ принадлежит классу C^∞ как композиция двух функций G и r^2 класса C^∞ и $\gamma_1(a_1) = G(0) > 0$.

Так как $0 \leq G \leq 1$, то $0 \leq \gamma \leq 1$. Наконец, для $d(a_1, x_1) \geq \eta_1$ функция γ_1 равна нулю, поскольку $G(s)$ равна нулю для $s \geq \eta_1^2$. Лемма 1 доказана.

Лемма 2. Пусть X — локально компактное пространство и $(\Omega_i)_{i \in I}$ — открытое покрытие X . Пусть K — некоторый компакт пространства X . Тогда существует система непрерывных функций $\beta_i \geq 0$, зависящих от того же множества индексов I , равных нулю, кроме конечного их числа, имеющих компактные носители, содержащиеся в Ω_i , и таких, что сумма $\sum_{i \in I} \beta_i$ положительна на K . Если X является многообразием класса C^m , то функции β_i можно выбирать из класса C^m .

Пусть ξ — некоторая точка K . Точка ξ может принадлежать нескольким открытым множествам Ω_i . Выберем произвольно одно из них и обозначим через $i(\xi)$ соответствующий ему индекс, такой, чтобы $\xi \in \Omega_{i(\xi)}$. Пусть теперь $\gamma_\xi: x \rightarrow \gamma_\xi(x)$ — функция, удовлетворяющая условиям леммы 1 относительно открытого множества $\Omega_{i(\xi)}$ и точки ξ этого множества. Обозначим через ω_ξ множество точек x , в которых $\gamma_\xi(x) > 0$. При ξ , пробегающем K , множества ω_ξ образуют некоторое открытое покрытие K . В самом деле, они являются открытыми множествами и $\xi \in \omega_\xi$. Последнее означает, что эти открытые множества полностью покрывают K . Поскольку K является компактом, для него существует конечное подпокрытие. Иначе говоря, существует конечное число точек $\xi_1, \xi_2, \dots, \xi_l$, таких, что ω_{ξ_v} , при $v = 1, 2, \dots, l$ образуют открытое покрытие K . Обозначим теперь через β_i функцию $\sum_{i(\xi_v)=i} \gamma_{\xi_v}$. Функция β_i непрерывна и ≥ 0 .

Если X является многообразием класса C^m , то эта функция принадлежит классу C^m . Ее носитель содержитя в объединении носителей функций γ_{ξ_v} , $i(\xi_v) = i$, т. е. является компактом и содержитя в Ω_i . Для каждой точки $x \in K$ существует хотя бы одна из точек ξ_v , такая, что $x \in \omega_{\xi_v}$, ибо ω_{ξ_v} образуют покрытие K . Соответствующая функция γ_{ξ_v} в точке x положительна. Следовательно, соответствующая функция β_i ($i = i(\xi_v)$) положительна в точке x . Отсюда вытекает, что функция $\sum_{i \in I} \beta_i$ положительна на K , чем и завершается доказательство леммы¹⁾.

¹⁾ Носитель функции $\sum_{i \in I} \beta_i = \beta$, содержащийся в объединении носителей функций β_i , является некоторым компактом H , содержащим открытое

Лемма 3. Пусть X — локально компактное, счетное в бесконечности пространство. Тогда существует последовательность компактов K_n и открытых множеств U_n с компактными замыканиями, таких, что $K_n \subset U_n$, K_n образуют покрытие X и любой компакт пересекается лишь с конечным числом \bar{U}_n .

Так как X счетно в бесконечности, существует последовательность компактов A_n , объединение которых дает X . Каждый из A_n имеет компактную окрестность B_n (см. примечание¹⁾). Внутренности $B_n \supset A_n$ в объединении дают X . Положим $C_n = B_0 \cup B_1 \cup B_2 \cup \dots \cup B_n$. Множества C_n образуют возрастающую последовательность компактов, и множества $\mathring{C}_n \supset A_n$ в объединении дают X . Кроме того, для каждого компакта из X существует такое целое число n , что $\mathring{C}_n \supset K$. В самом деле, последовательность $\mathring{C}_n \cap K$ является возрастающей последовательностью открытых множеств на компакте K , объединение которых равно K . Следовательно, одно из них совпадает с K .

Положим затем $D_n = C_n - \mathring{C}_{n-1} = C_n \cap C \mathring{C}_{n-1}$. Множество D_n замкнуто (как пересечение двух замкнутых множеств), содержитя в компакте C_n и, следовательно, компактно. Объединение D_n снова дает X . Действительно, если $x \in X$ и если n является таким первым числом, при котором $x \in C_n$ и $x \notin C_{n-1}$, то $x \notin \mathring{C}_{n-1}$ и, следовательно, $x \in D_n$.

Пусть теперь $p(n)$ — наименьшее целое число p , такое, что $\mathring{C}_p \supset C_n$. Пусть $q(n)$ — наибольшее целое число q , такое, что $C_q \subset \mathring{C}_{n-1}$. Положим $E_n = C_{p(n)} - \mathring{C}_{q(n)}$. Множество E_n компактно по той же причине, что и рассмотренное выше множество D_n ; E_n является окрестностью D_n . В самом деле, если $x \in D_n$, то $x \in C_n$, а, следовательно, $x \in \mathring{C}_{p(n)}$. Далее, так как $x \notin \mathring{C}_{n-1}$, то $x \notin C_{q(n)}$, а значит, $x \in \mathring{C}_{p(n)} - C_{q(n)}$. Поэтому $\mathring{C}_{p(n)} - C_{q(n)} =$

множество U точек, в которых $\beta > 0$, и содержащим само K . Отсюда следует, что H является компактной окрестностью K . Можно доказать непосредственно следующее свойство: в локально компактном пространстве X любой компакт K имеет компактную окрестность H и даже фундаментальную систему компактных окрестностей.

В самом деле, пусть U — некоторая окрестность K . Поскольку каждая точка $x \in K$ имеет фундаментальную систему компактных окрестностей в X (стр. 75), то существует некоторое открытое множество Ω_x в X , содержащее x , с компактным замыканием $\bar{\Omega}_x$ в U . Множества Ω_x образуют некоторое открытое покрытие компакта K . Из него можно выбрать конечное подпокрытие $\Omega_1, \Omega_2, \dots, \Omega_n$. Множество $\bar{\Omega}_1 \cup \bar{\Omega}_2 \cup \dots \cup \bar{\Omega}_n$ является компактной окрестностью K , содержащейся в U .

$= \dot{C}_{p(n)} \cap CC_{q(n)}$ является открытым множеством, содержащим D_n и, конечно, содержащимся в E_n . Наконец, последовательность E_n локально конечна. В самом деле, пусть K — некоторый компакт. Тогда существует такое q , что $C_q \supset K$ (и даже $\dot{C}_q \supset K$). Затем существует такое n , что $\dot{C}_{n-1} \supset C_q$. Тогда $q(n) \geq q$, а значит, для $m \geq n$ имеем: $q(m) \geq q$ и $\dot{C}_{q(m)} \supset C_q \supset K$. Отсюда вытекает, что E_m с K не пересекаются.

Очевидно, $K_n = D_n$ и $U_n = \dot{E}_n$ удовлетворяют условиям леммы 3¹⁾.

Докажем теперь теорему 11, используя множества K_n и U_n .

Для каждого n рассмотрим покрытие компакта K_n из X открытыми множествами $\Omega_i \cap U_n$. К нему можно применить лемму 2 и найти непрерывные функции $\beta_{i,n} \geq 0$ класса C^m (если только X является многообразием класса C^m), равные нулю, кроме конечного их числа, и такие, что $\beta_{i,n}$ имеют носители в $\Omega_i \cap U_n$ (компактные, поскольку U_n компактно), а сумма $\sum_{i \in I} \beta_{i,n}$ положительна на K_n .

Обозначим теперь через β_i функцию $\sum_{n=0}^{\infty} \beta_{i,n}$. Хотя эта сумма

и бесконечна, она локально конечна. В самом деле, если K является некоторым компактом X , то существует лишь конечное число n , при котором K пересекается с U_n . Для каждого из таких n имеется только конечное число отличных от нуля функций $\beta_{i,n}$. Для всех других n функции $\beta_{i,n}$ с носителем в U_n равны нулю на K . Отсюда следует, что функция β_i непрерывна на X , а в случае, когда X является многообразием класса C^m , функция β_i принадлежит классу C^m . Действительно, свойство непрерывности и дифференцируемости в точке множества X может быть доказано, если вести рассмотрения только в некоторой окрестности этой точки, которая может быть выбрана компактной. Тогда сумма, определяющая β_i , конечна в этой окрестности. Носитель функции β_i содержится в замыкании объединения носителей функций $\beta_{i,n}$, $n = 0, 1, 2, \dots$. Это объединение локально конечно, т. е. лишь конечное число носителей пересекается с любым компактом пространства X . Отсюда непосредственно вытекает, что объединение этих носителей замкнуто и что

¹⁾ Это построение лучше понять на примере векторного нормированного пространства X . Возьмем в качестве A_n шар $\|x\| \leq n$. В качестве B_n можно будет взять шар $\|x\| \leq n+1$, а вместо C_n взять B_n . Множество $D_n = K_n$ будет представлять собой шаровой слой $n \leq \|x\| \leq n+1$; для любого n $p(n) = n+1$ и $q(n) = n-2$. Множество E_n теперь будет шаровым слоем, а U_n является открытым множеством $n-1 < \|x\| < n+1$. Все построение в этом случае становится очевидным.

оно является носителем β_i ¹⁾). Этот носитель, следовательно, содержится в Ω_i .

Положим теперь $\beta = \sum_{i \in I} \beta_i = \sum_{i, n} \beta_{i,n}$. То же самое рассуждение показывает, что β непрерывна и принадлежит классу C^m , если X является многообразием класса C^m . Каждая точка X принадлежит хотя бы одному K_n , и, следовательно, в этой точке $\sum_{i \in I} \beta_{i,n} > 0$, так как больше нуля по крайней мере одна из функций $\beta_{i,n}$. Последнее означает, что $\beta > 0$ всюду в X . Но тогда функция $a_i = \beta_i/\beta$ непрерывна на X , принадлежит классу C^m , если X является многообразием класса C^m , и имеет носитель в Ω_i . Так как система β_i локально конечна, то система $a_i \geq 0$ также локально конечна. Далее, сумма $\sum_{i \in I} a_i$ равна $\sum_{i \in I} \beta_i/\beta = 1$, чем и заканчивается доказательство теоремы.

Следствие 1. Пусть X — локально компактное, счетное в бесконечности пространство, F — замкнутое множество из X , а Ω — открытое множество, содержащее F . Тогда существует непрерывная вещественная функция α на X , принадлежащая классу C^m , если X — многообразие класса C^m , имеющая носитель в Ω , удовлетворяющая неравенству $0 \leq \alpha \leq 1$ и равная 1 на некоторой окрестности множества F .

В самом деле, для доказательства достаточно применить теорему к покрытию X , образованному множествами Ω и CF . Если через α и β обозначить функции, соответствующие этому покрытию, то $\alpha + \beta = 1$. Носитель β лежит в CF , его дополнение является открытым множеством, содержащим F , на котором β равна нулю, а следовательно, $\alpha = 1$.

Заметим, что если F компактно, то, согласно примечанию на стр. 456, можно взять Ω с компактным замыканием и тогда α имеет компактный носитель. Впрочем, в этом случае нет необходимости предполагать X счетным в бесконечности (см. замечание 3°) после формулировки теоремы 11).

Следствие 2. Пусть X — локально компактное, счетное в бесконечности пространство, A и B — два замкнутых непересекающихся подмножества X . Тогда существует вещественная непрерывная функция α , принадлежащая классу C^m , если X —

¹⁾ Объединение F бесконечного семейства замкнутых множеств F_n не замкнуто. Однако оно замкнуто, если это семейство локально конечно. В самом деле, пусть $x \in X$ принадлежит замыканию множества F , а \mathcal{V} — компактная окрестность x . Пусть n — такое целое число, что $F_m \cap \mathcal{V} = \emptyset$ для $m > n$. Тогда x является точкой замыкания множества $\bigcup_{m \leq n} F_m$. Это множество замкнуто как объединение конечного числа замкнутых множеств. Отсюда следует, что $x \in \bigcup_{m \leq n} F_m \subset F$, т. е. следует, что F замкнуто.

многообразие класса C^m , равная 1 всюду на некоторой окрестности A и равная 0 всюду на некоторой окрестности B .

Действительно, для доказательства достаточно к пространству X , замкнутому множеству A и открытому множеству $\Omega = CB$ применить следствие 1. Тот факт, что функция α имеет носитель в CB означает, что она равна нулю всюду на некоторой окрестности B .

Следствие 3. Пусть X — локально компактное, счетное в бесконечности пространство, F_1, F_2, \dots, F_n — семейство замкнутых попарно непересекающихся частей X , \vec{E} — банахово пространство и $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ — произвольные векторы из \vec{E} . Существует непрерывная на X функция $\vec{f}(x)$ со значениями в \vec{E} , равная e_i в окрестности множества F_i .

Доказательство. Пусть $I = \{1, 2, \dots, n\}$. Обозначим через F'_i объединение множеств F_j , $j \neq i$, а через Ω_i — дополнение множества F'_i . Поскольку пересечение F'_i пусто, то объединение Ω_i равно X . Множества Ω_i образуют некоторое открытое покрытие X . Пусть (α_i) , $i \in I$, — соответствующее разложение единицы. Тогда носитель функции α_i лежит в Ω_i , а значит, α_i равна нулю в некоторой окрестности F'_i . Поскольку сумма $\sum_i \alpha_i$ равна 1, то мы получаем, что функция α_i равна 1 в некоторой окрестности F_i . Функция

$$\vec{f}(x) = \sum_{i \in I} \vec{e}_i \alpha_i(x) \quad \text{или} \quad \vec{f} = \sum_{i \in I} \vec{e}_i \alpha_i \quad (\text{IV}, 2; 25)$$

удовлетворяет неравенствам $\|\vec{f}(x)\| \leq (\max_{i \in I} \|\vec{e}_i\|) \sum_{i \in I} \alpha_i(x) = \max_{i \in I} \|\vec{e}_i\|$. Отсюда $\|\vec{f}\| \leq \max_{i \in I} \|\vec{e}_i\|$. Однако поскольку на F_i функция \vec{f} равна \vec{e}_i , то имеют место обратные неравенства, а, значит, $\|\vec{f}\| = \max_{i \in I} \|\vec{e}_i\|$.

Следствие 4. Пусть X — локально компактное, счетное в бесконечности пространство, F — замкнутое подмножество X . Пусть \vec{g} — непрерывная функция на F со значениями в банаховом пространстве \vec{E} . Предположим, что каждая точка $a \in F$ имеет такую открытую окрестность \mathcal{U}_a в X , что сужение \vec{g} на $F \cap \mathcal{U}_a$ можно продолжить на \mathcal{U}_a до некоторой непрерывной функции \vec{G}_a (соответственно принадлежащей классу C^m , если X является многообразием класса C^m). Тогда функцию \vec{g} , определенную на F , можно продолжить на X до непрерывной функ-

ции \vec{G} (соответственно принадлежащей классу C^m) со значениями в \vec{E} .

Доказательство. Множество всех \mathcal{V}_a и CF образует открытое покрытие X . Ему можно поставить в соответствие разложение единицы, образованное функциями α_a с носителем в \mathcal{V}_a и функцией α_0 с носителем в CF : $(\sum_{a \in F} \alpha_a) + \alpha_0 = 1$.

Рассмотрим теперь функцию $\alpha_a \vec{G}_a$. Эта функция определена только в \mathcal{V}_a . Однако ее можно определить на всем X , придавая ей значение 0 вне \mathcal{V}_a . При этом она останется непрерывной (соответственно принадлежащей классу C^m). В открытом множестве \mathcal{V}_a она совпадает с функцией $\alpha_a \vec{G}_a$, которая непрерывна (соответственно принадлежит классу C^m). Если $x \notin \mathcal{V}_a$, то она совпадает с 0 (в открытом дополнении носителя α_a). Рассмотрим теперь сумму $\vec{G} = \sum_{a \in F} \alpha_a \vec{G}_a$. Так как эта сумма локально конечна, то она определяет на X некоторую непрерывную функцию (соответственно функцию класса C^m).

В каждой точке $x \in F$ для любого $a \in F$ имеет место равенство $\vec{G}_a(x) = \vec{g}(x)$, ибо каждое \vec{G}_a является продолжением \vec{g} . Поэтому для каждого $x \in F$ имеем: $\sum_{a \in F} \alpha_a \vec{G}_a(x) = \vec{g}(x) \sum_{a \in F} \alpha_a(x)$. Поскольку $\alpha_0(x)$ равна нулю на F , сумма равна 1, а, следовательно, $\vec{G}(x) = \vec{g}(x)$, так что функция \vec{G} является продолжением функции \vec{g} .

Заметим, что это следствие содержит как частные случаи все предыдущие следствия. Если, например, мы обратимся к следствию 3, то увидим, что в нем на X продолжают лишь функцию \vec{g} , определенную на $\bigcup_i F_i$ и равную постоянной e_i на каждом F_i .

Следствие 5. Пусть X является многообразием класса C^m , $m \geq 1$, размерности N и V — замкнутое подмногообразие класса C^m многообразия X размерности n . Пусть \vec{g} — функция, определенная на V , со значениями в пространстве \vec{E} класса C^m . Тогда эту функцию можно продолжить в X до функции \vec{G} класса C^m со значениями в \vec{E} .

Доказательство. Достаточно обратиться к условиям предыдущего следствия. Для точки $a \in V$ можно найти откры-

тую окрестность \mathcal{V}_a точки a в X и некоторую карту Φ_a , т. е. C^m -диффеоморфизм некоторого открытого множества \mathcal{O}_a из \mathbb{R}^n на \mathcal{V}_a . Кроме того, если записать \mathbb{R}^n в виде произведения $\mathbb{R}^n \times \mathbb{R}^{N-n}$ и представить каждую из его точек x как пару (y, z) , $y \in \mathbb{R}^n$, $z \in \mathbb{R}^{N-n}$, то можно считать, что отображение Φ_a переводит пересечение $\mathcal{O}_a \cap \mathbb{R}^n$ на пересечение $\mathcal{V}_a \cap V$ (следствие 2₂ теоремы 32 гл. III)¹⁾.

Задача продолжения \vec{g} из $V \cap \mathcal{V}_a$ на \mathcal{V}_a сводится к задаче продолжения функции $\vec{h} = \vec{g} \circ \Phi_a$, определенной в $\mathcal{O}_a \cap \mathbb{R}^n$, на открытое множество \mathcal{O}_a или, по крайней мере, на некоторую окрестность точки $a = \Phi_a^{-1}(a)$ в этом открытом множестве. Такое продолжение в \mathbb{R}^N очевидно. Оно определено функцией $(y, z) \rightarrow \vec{h}(y)$, что и требовалось доказать.

Следствие 6. Пусть X — локально компактное, счетное в бесконечности пространство и F_1, F_2, \dots, F_n — замкнутые попарно непересекающиеся подмножества. Тогда существуют попарно непересекающиеся окрестности подмножеств F_1, F_2, \dots, F_n .

В самом деле, построим, согласно следствию 3, функцию G с вещественными значениями, соответствующую значениям $e_i = \pm i$, $i = 1, 2, \dots, n$. Открытые окрестности $\mathcal{V}_i = \{x \in X; |i - 1/2| < G(x) < i + 1/2\}$ множеств F_i удовлетворяют утверждению леммы.

Пространство, обладающее этим свойством, называется *нормальным*. Локально компактное счетное в бесконечности пространство нормально. Более общо, паракомпактное пространство нормально. Однако обратное утверждение не верно.

Следствие 7. Пусть X — локально компактное счетное в бесконечности пространство. Пусть $(\Omega_i)_{i \in I}$ является открытым покрытием X . Тогда найдется локально конечное покрытие $(\Omega'_i)_{i \in I}$, зависящее от того же множества индексов и такое, что $\overline{\Omega'_i} \subset \Omega_i$.

В самом деле, для непосредственного применения теоремы 11 достаточно в качестве Ω'_i взять множество таких x , чтобы $\alpha_i(x) > 0$; Ω'_i являются открытыми множествами. Так как $\overline{\Omega_i}$ являются носителями α_i , то они содержатся в Ω_i . В каждой точке

¹⁾ Это следствие было доказано для подмногообразия V размерности n некоторого аффинного пространства размерности N , например \mathbb{R}^N . Здесь X является абстрактным многообразием размерности N , V — некоторым подмногообразием X размерности n . Однако с помощью карты окрестности a в X приходят к открытому множеству из \mathbb{R}^n и к подмногообразию этого открытого множества. Следствие применено именно в этих условиях.

$x \in X$ сумма $\sum_{i \in I} a_i(x)$ равна 1, а, следовательно, одна из $a_i(x)$ положительна и x лежит в соответствующем множестве Ω'_i . Тем самым множества Ω'_i образуют некоторое покрытие. Это покрытие локально конечно, поскольку локально конечно система a_i .

Укажем, наконец, без доказательства результат, частично обобщающий следствие 5 (но относящийся к непрерывным функциям, а не к функциям класса C^n). Перед этим введем новое понятие. Пусть \vec{F} — банахово пространство и \vec{f} — непрерывная функция с компактным носителем в X и со значениями в \vec{F} . Говорят, что эта функция *разложима*, если она выражается в виде

$$\vec{f}(x) = \sum_{i=1}^N \vec{g}_i \varphi_i(x) \quad \text{или} \quad \vec{f} = \sum_{i=1}^N \vec{g}_i \varphi_i,$$

где \vec{g}_i — постоянные векторы из \vec{F} , а φ_i — скалярные непрерывные функции с компактным носителем. Множество значений такой функции лежит в векторном конечномерном подпространстве пространства \vec{F} , порожденном векторами \vec{g}_i . Обратно, пусть \vec{f} — непрерывная функция с компактным носителем, принимающая свои значения в векторном конечномерном подпространстве \vec{G} пространства \vec{F} . Пусть $\vec{g}_1, \vec{g}_2, \dots, \vec{g}_N$ — базис \vec{G} . Для каждого $x \in X$, $\vec{f}(x)$ является вектором \vec{G} , а следовательно, его можно единственным образом записать в виде $\vec{f}(x) = \sum_{i=1}^N \vec{g}_i \varphi_i(x)$,

и, значит, функция \vec{f} является суммой $\sum_{i=1}^N \vec{g}_i \varphi_i$. Функция $\varphi_i: X \rightarrow \mathbb{C}$ является композицией функции $\vec{f}: X \rightarrow \vec{G}$ и непрерывного отображения $\vec{G} \rightarrow \mathbb{C}$, которое каждому вектору из \vec{G} ставит в соответствие его i -ю координату в выбранном базисе. Такая функция непрерывна. Ее носитель, поскольку он содержится в носителе \vec{f} , компактен, а сама функция \vec{f} разложима. Таким образом, непрерывные разложимые функции с компактным носителем можно назвать непрерывными функциями с компактным носителем *конечного ранга*. Их множество является подпространством векторного пространства $\mathcal{C}(X; \vec{F})$ непрерывных функций с компактным носителем и значениями в \vec{F} . Можно показать, что это множество изоморфно тензорному произведению

$\mathcal{C}(X)$ на \vec{F} , которое кратко обозначают через $\mathcal{C}(X) \otimes \vec{F}$. Это множество совпадает с $\mathcal{C}(X; \vec{F})$, если \vec{F} конечномерно.

Следствие 8. Пусть K является некоторым компактом X и f — непрерывная функция на K со значениями в банаевом пространстве \vec{F} . Тогда, каково бы ни было $\varepsilon > 0$ и какова бы ни была окрестность \mathcal{U} множества K , существует на X такая непрерывная разложимая функция \vec{f}_ε с компактным носителем $\subset \mathcal{U}$, что $\|\vec{f} - \vec{f}_\varepsilon\| \leq \varepsilon$. Можно даже эту функцию выбрать в виде $\vec{f}_\varepsilon = \sum_{i=1}^N \vec{g}_i \Phi_i$, где $\Phi_i \geq 0$, и так, чтобы было справедливым неравенство

$$\left\| \vec{f} - \sum_{i=1}^N \vec{g}_i \Phi_i \right\| \leq \varepsilon.$$

Доказательство. Для любой точки $a \in K$ существует такая открытая окрестность $\mathcal{U}_a \subset \mathcal{U}$ точки a в X , что для каждой точки x из этой окрестности имеет место неравенство: $\|\vec{f}(x) - \vec{f}(a)\| \leq \varepsilon$.

Множество K покрывается конечным числом окрестностей $\mathcal{U}_a : \mathcal{U}_{a_1}, \mathcal{U}_{a_2}, \dots, \mathcal{U}_{a_N}$. Пусть $\alpha_1, \alpha_2, \dots, \alpha_N, \alpha_0$ — некоторое разбиение единицы в X относительно покрытий \mathcal{U}_{a_i} и $C K$. Тогда разложимая функция $\vec{f}_\varepsilon = \sum_{i=1}^N \vec{f}(a_i) \alpha_i$ удовлетворяет условию следствия. В самом деле, поскольку $\alpha_0(x)$ отлична от нуля только в $C K$, где $\vec{f}(x)$ равна нулю; то $\vec{f}(x) \alpha_0(x)$ равна нулю, а, следовательно, $\vec{f}(x) = \sum_{i=1}^N \alpha_i(x) \vec{f}(x)$.

Теперь имеем

$$\left\| \vec{f}(x) - \sum_{i=1}^N \vec{f}(a_i) \alpha_i(x) \right\| = \left\| \sum_{i=1}^N (\vec{f}(x) - \vec{f}(a_i)) \alpha_i(x) \right\| \leq \sum_{i=1}^N \varepsilon \alpha_i(x) \leq \varepsilon^1.$$

Точно так же

$$\begin{aligned} \left| \|\vec{f}(x)\| - \sum \|\vec{f}(a_i)\| \alpha_i(x) \right| &\leq \\ &\leq \left| \sum_{i=1}^N (\|\vec{f}(x)\| - \|\vec{f}(a_i)\|) \alpha_i(x) \right| \leq \sum_{i=1}^n \varepsilon \alpha_i(x) \leq \varepsilon. \end{aligned}$$

¹⁾ $\|\vec{f}(x) - \vec{f}(a_i)\| \leq \varepsilon$ для $x \in \mathcal{U}_{a_i}$ и $\alpha_i(x) = 0$ для $x \notin \mathcal{U}_{a_i}$.

Теорема 12. Пусть X — произвольное метрическое пространство, Y — замкнутое подмножество X , \vec{g} — непрерывная функция на Y со значениями в нормированном конечномерном векторном пространстве \vec{E} . Тогда существует на X непрерывная функция \vec{G} со значениями в \vec{E} , являющаяся продолжением \vec{g} (т. е. совпадающая с \vec{g} на Y) и удовлетворяющая условию:

$$\sup_{x \in X} \|\vec{G}(x)\| = \sup_{y \in Y} \|\vec{g}(y)\| \leq +\infty. \quad (\text{IV}, 2; 26)$$

Если g вещественна, то и G можно брать вещественной и, сверх того,

$$\sup_{x \in X} G(x) = \sup_{y \in Y} g(y), \quad \inf_{x \in X} G(x) = \inf_{y \in Y} g(y). \quad (\text{IV}, 2; 27)$$

(Эти величины могут быть равны $\pm \infty$.)

Если X локально компактно и является объединением счетного множества компактных подмножеств, аналогичный результат имеет место, если \vec{E} является бесконечномерным банаховым пространством.

Носитель меры Радона

Интуитивно под носителем распределения масс в механике или распределением электрических зарядов в физике понимают наименьшее замкнутое множество, содержащее всю массу или весь заряд.

Определение. Пусть X — локально компактное пространство, μ — мера Радона на X со значениями в векторном нормированном пространстве \vec{E} , Ω — некоторое открытое множество X . Говорят, что μ равна нулю в Ω , если для каждой функции φ из $\mathcal{C}(X)$ с носителем в Ω $\mu(\varphi)$ равна нулю¹.

Интуитивно это значит, что в открытом множестве Ω нет массы или заряда.

Точно так же говорят, что две меры μ_1 и μ_2 на X совпадают или равны в Ω , если их разность $\mu_1 - \mu_2$ равна нулю в Ω .

Теорема 13. Пусть X — локально компактное пространство и $(\Omega_t)_{t \in I}$ — произвольное семейство (конечное или нет)

¹) Здесь надо обратить внимание на то, что носитель φ является компактом, а Ω открыто. В дальнейшем мы увидим, почему надо выбирать Ω открытым.

открытых множеств этого пространства. Тогда, если мера Радона μ на X равна нулю на каждом из множеств Ω_i , то она равна нулю в их объединении Ω .

Доказательство. Пусть φ является некоторой функцией из $\mathcal{E}(X)$ с компактным носителем K в Ω . Нам надо доказать, что $\mu(\varphi)$ равна нулю. Поскольку K компактно, то его можно покрыть конечным числом множеств Ω_i . Пусть это будут $(\Omega_i)_{i \in J}$. Пусть $(a_i)_{i \in J}$ и a_0 — разложение единицы относительно покрытия $(\Omega_i)_{i \in J}$, СК пространства X . Поскольку $\sum_{i \in J} a_i = 1$ на K , то на всем X имеет место тождество $\varphi = \sum_{i \in J} a_i \varphi^1$.

Так как μ линейна, то $\mu(\varphi) = \sum_{i \in J} \mu(a_i \varphi)$. Каждый из членов правой части этого равенства равен нулю, ибо $a_i \varphi$ имеет носитель в Ω_i , а μ предполагалась равной нулю в Ω_i , и, следовательно, $\mu(a_i \varphi)$ также равна нулю. Теперь мы видим, для чего нужно разложение единицы, — чтобы разложить φ , имеющую носитель в объединении Ω множеств Ω_i , в сумму конечного числа функций $\varphi_i = a_i \varphi$, имеющих носители в Ω_i .

Следствие. Пусть μ — мера Радона на локально компактном пространстве X . Тогда среди всех открытых множеств существует наибольшее, в котором мера μ равна нулю.

В самом деле, рассмотрим все открытые множества, в которых мера μ равна нулю. Объединение этих множеств открыто, а, согласно теореме, μ на нем равна нулю. Это объединение, очевидно, является наибольшим открытым множеством, в котором μ равна нулю.

Определение. Замкнутое множество F , являющееся дополнением к наибольшему открытому множеству Ω , в котором мера Радона μ на X равна нулю, называется *носителем* μ . Пустой носитель имеет единственная мера, равная нулю ($\mu(\varphi) = 0 \in \vec{E}$ для любой функции φ).

Из этого определения вытекает, что если φ является функцией из $\mathcal{E}(X)$, μ — мерой Радона и если носитель φ и носи-

¹⁾ Это верно как для любой точки $x \in K$ (поскольку тогда $\sum_{i \in J} a_i(x) = 1$) так и для $x \notin K$ (поскольку тогда оба члена равны нулю).

тель μ не имеют общих точек, т. е. если φ обращается в нуль на некоторой окрестности носителя μ , то $\mu(\varphi) = \vec{0}$. В самом деле, носитель функции φ в этом случае лежит в дополнении к носителю μ , т. е. по определению носителя лежит в открытом множестве, в котором μ равна нулю.

Отсюда следует, что если две функции φ_1 и φ_2 из $\mathcal{C}(X)$ равны на всей окрестности носителя μ , то $\mu(\varphi_1)$ и $\mu(\varphi_2)$ равны между собой. В самом деле, функция $\varphi_1 - \varphi_2$ равна нулю на открытом множестве, содержащем носитель μ , а, значит, носитель $\varphi_1 - \varphi_2$ и носитель μ не пересекаются, откуда и следует наше утверждение. Однако можно доказать более точный результат:

Теорема 13₂. *Если φ обращается в нуль на носителе μ , то $\mu(\varphi) = 0$. Если φ_1 и φ_2 равны на носителе μ , то $\mu(\varphi_1) = \mu(\varphi_2)$.*

Доказательство. Пусть F — носитель μ и K — носитель φ . Зададим $\varepsilon > 0$. Поскольку мера μ непрерывна на $\mathcal{C}_K(X)$, то найдется такое $\eta > 0$, что для функции $\psi \in \mathcal{C}_K(X)$, такой, что $\|\psi\| \leq \eta$, имеет место неравенство $\|\mu(\psi)\| \leq \varepsilon$. Поскольку φ непрерывна и равна нулю на F , существует такая окрестность \mathcal{U} множества F , в которой $|\varphi| \leq \eta$.

Пусть α — непрерывная вещественная функция на X , $0 \leq \alpha \leq 1$, имеющая компактный носитель $\subset \mathcal{U}$, равная 1 в окрестности компакта $K \cap F$ (следствие 1 теоремы 11).

Рассмотрим носитель функции $\alpha\varphi - \varphi = (\alpha - 1)\varphi$. Прежде всего, он содержится в носителе K функции φ . Поскольку $\alpha - 1$ равна нулю в окрестности множества $K \cap F$, то он содержится в $C(K \cap F) = CK \cup CF$. Поскольку этот носитель лежит в K , то он находится в CF , т. е. в дополнении к носителю μ . Поэтому $\mu(\alpha\varphi - \varphi) = \vec{0}$ или $\mu(\alpha\varphi) = \mu(\varphi)$.

Теперь $\alpha\varphi \in \mathcal{C}_K(X)$ и $\|\alpha\varphi\| \leq \eta$ (поскольку $|\varphi| \leq \eta$ на \mathcal{U} и носитель α лежит в \mathcal{U}). Значит, $\|\mu(\alpha\varphi)\| \leq \varepsilon$ и, следовательно, $\|\mu(\varphi)\| \leq \varepsilon$.

Поскольку ε произвольно, отсюда следует, что $\mu(\varphi) = \vec{0}$.

Следствие. Для того чтобы мера μ на X имела носитель, лежащий в множестве A , образованном конечным числом точек a_1, a_2, \dots, a_n , необходимо и достаточно, чтобы она была линейной комбинацией мер Дирака:

$$\vec{\mu} = \vec{c}_1 \delta_{(a_1)} + \vec{c}_2 \delta_{(a_2)} + \dots + \vec{c}_n \delta_{(a_n)}. \quad (\text{IV}, 2; 27_2)$$

Доказательство. Легко видеть прежде всего, что если мера μ имеет указанный вид, то она обязательно имеет носитель в A , поскольку на дополнении к A она равна нулю. Нам остается доказать обратное утверждение. Пусть μ — мера, имеющая носитель в A . Пусть θ_i — функция из $\mathcal{C}(X)$, равная 1 в точке a_i и равная 0 в точках a_j , $j \neq i$ (следствие 3 теоремы 11). Положим теперь $\mu(\theta_i) = c_i$ для всех i . Тогда для любой функции φ имеет место соотношение

$$\varphi = \sum_{i=1}^n \varphi(a_i) \theta_i + \chi, \quad (\text{IV}, 2; 27_3)$$

где χ является функцией, равной нулю на A . Из теоремы получаем, что $\mu(\chi) = 0$, так что

$$\vec{\mu}(\varphi) = \sum_{i=1}^n \varphi(a_i) \vec{\mu}(\theta_i) = \sum_{i=1}^n c_i \vec{\mu}(a_i), \quad (\text{IV}, 2; 27_4)$$

чём и заканчивается доказательство следствия.

Теорема 14. Пусть $\vec{\mu}$ — мера Радона на локально компактном пространстве X . Точка $a \in X$ принадлежит носителю F меры μ тогда и только тогда, когда для любого открытого множества ω пространства X , содержащего a , мера μ в этом множестве отлична от нуля.

Точка a не принадлежит носителю F тогда и только тогда, когда существует открытое множество ω , содержащее a , в котором мера μ равна нулю.

Доказательство. Из эквивалентности соотношений $A \Rightarrow B$ и $\neg B \Rightarrow \neg A$ (теорема 11 гл. I) следует, что обе части теоремы эквивалентны. Поэтому достаточно доказать, например, вторую часть. Точка a не принадлежит F тогда и только тогда, когда она лежит в $\Omega = \overline{CF}$. Поскольку Ω является наибольшим открытым множеством, в котором $\mu = 0$, то точка a не принадлежит F тогда и только тогда, когда существует такое открытое множество ω , содержащее a , в котором $\mu = 0$.

Пример 1. Пусть $\vec{\mu} = \sum_v \vec{c}_v \delta_{(a_v)}$ — атомическая мера Радона. Точки a_v по условию различны, а сумма $\sum_{a_v \in K} \|\vec{c}_v\|$ для каждого компакта K из X конечна (см. стр. 448). Носитель этой меры Радона является замыканием объединения a_v .

Доказательство. Обозначим через A объединение a_v , через \bar{A} его замыкание и через Ω — дополнение к \bar{A} . Очевидно, что если носитель φ лежит в Ω , то $\vec{\mu}(\varphi) = \vec{0}$. Следовательно, носитель F меры μ содержится в \bar{A} . Остается показать, что каждая точка a_j объединения A принадлежит носителю F . Тогда $A \subset F$, или, в силу замкнутости носителя F , $\bar{A} \subset F$, или, окончательно, $\bar{A} = F$. Итак, пусть точка $a_j \in A$ и \mathcal{V} — компактная окрестность a_j . Ряд $\sum_{a_v \in \mathcal{V}} \|\vec{c}_v\|$ сходится. Значит, существует конечное подмножество J множества индексов I , такое, что

$$\sum_{a_v \in \mathcal{V} \setminus J} \|\vec{c}_v\| \leq \frac{1}{2} \|\vec{c}_j\|. \quad \text{Поскольку } J \text{ конечно, существует открыт}$$

тое множество ω , содержащее a_j , лежащее в \mathcal{V} и не содержащее никаких других точек a_k с $k \in J$. При этом имеет место следующая оценка:

$$\sum_{\substack{a_v \in \omega \\ v \neq j}} \|\vec{c}_v\| \leq \frac{1}{2} \|\vec{c}_j\|. \quad (\text{IV, 2; 28})$$

Пусть теперь Ω — произвольная открытая окрестность точки a_j . Покажем, что μ не может обратиться в нуль в Ω , а это будет свидетельствовать о том, что a_j принадлежит носителю F меры μ . Согласно лемме 1 к теореме 11, существует вещественная непрерывная функция φ с компактным носителем в $\omega \cap \Omega$, удовлетворяющая неравенствам $0 \leq \varphi \leq 1$ и такая, что $\varphi(a_j) = 1$. При этом имеют место соотношения

$$\vec{\mu}(\varphi) = \vec{c}_j + \sum_{v \neq j} \vec{c}_v \varphi(a_v), \quad (\text{IV, 2; 29})$$

откуда

$$\|\vec{\mu}(\varphi)\| \geq \|\vec{c}_j\| - \sum_{\substack{a_v \in \omega \\ v \neq j}} \|\vec{c}_v\| \geq \frac{1}{2} \|\vec{c}_j\|,$$

а, значит, $\vec{\mu}(\varphi) \neq \vec{0}$, чем и заканчивается доказательство утверждения.

Пример 2. Пусть \vec{r} — интегрируемая по Риману, определенная на \mathbb{R} функция со значениями в банаховом пространстве \vec{E} . Тогда носитель меры $\vec{r} dx$ плотности \vec{r} относительно меры dx лежит в носителе функции \vec{r} и в случае непрерывности \vec{r} совпадает с носителем \vec{r} .

Это свойство устанавливает связь между понятиями носителя функции и носителя меры.

Прежде всего, в силу предыдущей теоремы, носитель μ содержится в носителе \vec{p} . В самом деле, если носитель φ находится в дополнении носителя \vec{p} , то интеграл $\int \varphi \vec{p} dx$, очевидно, равен нулю. Покажем, что в случае непрерывности \vec{p} носитель μ совпадает с носителем \vec{p} . Пусть a — такая точка, что $\vec{p}(a) \neq 0$. Покажем, что a принадлежит носителю μ . Отсюда будет вытекать, что носитель μ содержит все точки x , в которых $\vec{p}(x) \neq 0$, и, следовательно, содержит также замыкание множества этих точек, т. е. носитель функции \vec{p} . Отсюда мы получим, что носитель μ совпадает с носителем \vec{p} . Поскольку функция \vec{p} непрерывна в точке a , то существует такое открытое множество ω , содержащее a , что для любого $x \in \omega$ имеет место неравенство

$$\|\vec{p}(x) - \vec{p}(a)\| \leq \frac{1}{2} \|\vec{p}(a)\|. \quad (\text{IV}, 2; 30)$$

Теперь видно, что обращение в нуль меры μ в любом открытом множестве Ω , содержащем a , невозможно. В самом деле, выберем произвольную функцию $\varphi \geq 0$ из $\mathcal{C}(X)$ с носителем в $\omega \cap \Omega$. Тогда будут справедливы соотношения

$$\left\| \int (\vec{p}(x) - \vec{p}(a)) \varphi(x) dx \right\| \leq \frac{1}{2} \|\vec{p}(a)\| \int \varphi. \quad (\text{IV}, 2; 31)$$

Следовательно,

$$\begin{aligned} \left\| \int \vec{p}(x) \varphi(x) dx \right\| &\geq \left\| \int \vec{p}(a) \varphi(x) dx \right\| - \frac{1}{2} \|\vec{p}(a)\| \int \varphi = \\ &= \frac{1}{2} \|\vec{p}(a)\| \int \varphi > 0, \end{aligned} \quad (\text{IV}, 2; 32)$$

откуда следует, что $\mu(\varphi) \neq 0$ и, значит, мера μ не равна нулю в Ω .

З а м е ч а н и е. Если \vec{p} не непрерывна, то носитель меры $\vec{p}dx$ не обязательно является носителем \vec{p} . Например, если \vec{p} всюду, кроме точки $a \in \mathbb{R}$, равна нулю, то носителем \vec{p} является $\{a\}$. Однако мера $\vec{p}dx$ равна нулю и ее носитель пуст.

Пример 3. Пусть $\vec{\mu}$ является мерой Радона, определенной на \mathbb{R} по формуле $\vec{\mu} = \sum_v \vec{c}_v \delta_{(a_v)} + \vec{p} dx$, где все точки a_v различны, сумма $\sum_v \|\vec{c}_v\|$ локально сходится, а функция \vec{p} непрерывна. Тогда носитель $\vec{\mu}$ является объединением носителя A , функции \vec{p} и замыкания множества a_v .

Этот результат сразу же вытекает из предыдущего. Действительно, легко видеть, что носитель $\vec{\mu}$ содержится в рассматриваемом множестве. Поэтому достаточно последовательно показать, что этот носитель содержит каждую из точек a_v и, кроме того, каждую такую точку a , не принадлежащую замыканию a_v , в которой $\vec{p}(a) \neq \vec{0}$. Рассмотрим точку a_j . Повторяя построение из примера 1, выберем в качестве ω открытый интервал, такой, что

$$\int_{\omega} \|\vec{p}(x)\| dx \leq \frac{1}{4} \|\vec{c}_j\|.$$

Построим затем для каждой окрестности Ω точки a_j такую же функцию φ , что и в примере 1, для которой на этот раз имеет место соотношение

$$\vec{\mu}(\varphi) = \vec{c}_j + \sum_{v \neq j} \vec{c}_v \varphi(a_v) + \int_{\omega} \vec{p}(x) \varphi(x) dx.$$

Отсюда

$$\|\vec{\mu}(\varphi)\| \geq \|\vec{c}_j\| - \frac{1}{2} \|\vec{c}_j\| - \frac{1}{4} \|\vec{c}_j\| = \frac{1}{4} \|\vec{c}_j\|,$$

чем и доказывается первое утверждение.

Рассмотрим теперь точку a , не принадлежащую замыканию множества точек a_v и такую, что $\vec{p}(a) \neq \vec{0}$. Построим такое же открытое множество ω , что и в примере 2, но так, чтобы оно не содержало точек a_v . После этого рассуждения, проведенные при доказательстве примера 2, повторяются без изменения.

Объединяя результаты, указанные на стр. 443, с доказанными в примерах 1—3, получим следующую теорему:

Теорема 15. Пусть a_v — различные точки X , \vec{c}_v — такие векторы банахова пространства \vec{E} , при которых для любого компакта K из X сумма $\sum_{a_v \in K} \|\vec{c}_v\|$ конечна. Тогда можно определить на X меру $\vec{\mu} = \sum_v \vec{c}_v \delta_{(a_v)}$ со значениями в \vec{E}

по формуле

$$\left(\sum_v \vec{c}_v \delta_{(a_v)} \right) \cdot \varphi = \sum_v \vec{c}_v \varphi(a_v). \quad (\text{IV}, 2; 32_2)$$

Для каждого компакта K имеет место неравенство

$$\left\| \sum_v \vec{c}_v \delta_{(a_v)} \right\|_K \leq \sum_{\substack{a_v \in K \\ a_v \in K}} \| \vec{c}_v \| \quad (\text{IV}, 2; 32_3)$$

и, кроме того,

$$\left\| \sum_v \vec{c}_v \delta_{(a_v)} \right\| \leq \sum_v \| \vec{c}_v \|. \quad (\text{IV}, 2; 32_4)$$

(Если \vec{E} является полем скаляров, то эти неравенства переходят в равенства¹⁾ .)

Пусть X является прямой \mathbb{R} и \vec{p} — функция на \mathbb{R} со значениями в \vec{E} , локально интегрируемая по Риману. Тогда эта функция определяет меру $\vec{p} dx$ со значениями в \vec{E} по формуле

$$\vec{p} dx \cdot \varphi = \int \vec{p}(x) \varphi(x) dx. \quad (\text{IV}, 2; 32_5)$$

Для каждого интервала $[a, b]$ прямой \mathbb{R}

$$\| \vec{p} dx \|_{[a, b]} \leq \int_{[a, b]} \| \vec{p}(x) \| dx, \quad (\text{IV}, 2; 32_6)$$

где неравенство переходит в равенство, если \vec{E} является полем скаляров.

¹⁾ Однако ничего такого не получится, если \vec{E} является векторным пространством размерности $n \geq 2$. Возьмем, например, $\vec{E} = \mathbb{R}^n$ с введенной в нем нормой Минковского $\| (x_1, x_2, \dots, x_n) \|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$, $1 < p \leq +\infty$. Пусть a_1, a_2, \dots, a_n суть n точек компакта X , и пусть \vec{c}_v есть точка \mathbb{R}^n , v -я координата которой равна 1, а все остальные равны нулю (\vec{c}_v образуют канонический базис \mathbb{R}^n). Тогда $\| \vec{\mu}(\varphi) \| = \left\| \sum_{v=1}^n \varphi(a_v) \vec{c}_v \right\| = \left(\sum_{v=1}^n | \varphi(a_v) |^p \right)^{1/p}$. Точная верхняя грань этого выражения для $\| \varphi \| \leq 1$

достигается при $\varphi = 1$. Она равна $n^{1/p}$. Итак, $\sum_{v=1}^n \| \vec{c}_v \| = n$ и для $n \geq 2$ $n^{1/p} < n$.

Для суммы мер на всем компактном интервале $[a, b]$ прямой \mathbb{R} имеет место неравенство

$$\left\| \sum_v \vec{c}_v \delta_{(a_v)} + \vec{p} dx \right\|_{[a, b]} \leq \sum_{a_v \in [a, b]} \|\vec{c}_v\| + \int_{[a, b]} \|\vec{p}(x)\| dx, \quad (\text{IV}, 2; 32)$$

переходящее в равенство, если \vec{E} является полем скаляров.

Носитель меры $\sum_v \vec{c}_v \delta_{(a_v)}$ является замыканием объединения a_v . Носитель меры $\sum_v \vec{c}_v \delta_{(a_v)} + \vec{p} dx$ при непрерывном \vec{p} является объединением замыкания множества a_v и носителя \vec{p} .

Продолжение меры на непрерывные функции φ с некомпактным носителем

Интуитивно кажется возможным придать смысл выражению $\vec{\mu}(\varphi)$ в том случае, когда $\vec{\mu}$ является мерой Радона с носителем A , а φ — скалярной непрерывной функцией с носителем B даже в том случае, когда B некомпактно, лишь бы только A и B имели компактное пересечение K .

Действительно, пусть α является некоторой функцией из $\mathcal{C}(X)$, равной 1 на всей окрестности компакта K (согласно следствию 1 теоремы 11, такие функции существуют). Тогда носитель функции $\alpha\varphi$, поскольку он находится в пересечении носителей функций α и φ , компактен.

Положим по определению

$$\vec{\mu}(\varphi) = \vec{\mu}(\alpha\varphi) \quad (\text{IV}, 2; 33)$$

— соотношение, имеющее смысл, поскольку $\alpha\varphi \in \mathcal{C}(X)$. Для обоснования этого определения достаточно показать, что правая часть не зависит от выбора α . Итак, пусть α_1 и α_2 — две функции, равные 1 в некоторой окрестности компакта K . Функция $\alpha_1\varphi - \alpha_2\varphi$ и мера $\vec{\mu}$ имеют непересекающиеся носители.

(Если некоторая точка принадлежит носителю $\vec{\mu}$, то она лежит в A . Если же она принадлежит носителю функции $(\alpha_1 - \alpha_2)\varphi$, то она содержится в носителе функции φ , т. е. в B и в носителе функции $\alpha_1 - \alpha_2$, а значит, лежит в C_K . Рассматриваемое пересечение содержится в $A \cap B \cap C_K = K \cap C_K = \emptyset$.) Тогда

$$\vec{\mu}(\alpha_1\varphi - \alpha_2\varphi) = \vec{0} \text{ или } \vec{\mu}(\alpha_1\varphi) = \vec{\mu}(\alpha_2\varphi).$$

В частности, если $\vec{\mu}$ имеет компактный носитель, то для непрерывной функции φ с произвольным носителем можно

определить $\vec{\mu}(\varphi)$, полагая $\vec{\mu}(\varphi) = \vec{\mu}(\alpha\varphi)$, где $\alpha \in \mathcal{C}(X)$ и равна 1 в некоторой окрестности носителя μ .

Например, если $\mu = \delta_{(a)}$, то для непрерывной скалярной функции φ с произвольным носителем получаем формулу (IV, 2; 5). Отсюда следует, что в случае $\mu = \delta_{(a)}$ можно вычислить $\mu(\varphi)$ даже для разрывных функций φ , что мы и сделаем в § 3.

Пусть μ — мера с компактным носителем. Вектор $\vec{\mu}(1)$ из \vec{E} называется «полной массой». Например, если \vec{E} полно и $\vec{\mu} = \sum_v \vec{c}_v \delta_{(a_v)}$, где a_v содержатся в одном компакте и $\sum_v \|c_v\| < \infty$, то полная масса равна $\sum_v \vec{c}_v \in \vec{E}$. Если $X = \mathbb{R}$ и $d\vec{\mu} = \vec{p} dx$, где \vec{p} имеет компактный носитель, то полная масса равна $\int \vec{p}(x) dx \in \vec{E}$. Легко видеть, откуда возникло выражение «полная масса». Если μ принимает вещественные значения и представляет собой распределение зарядов (в физике), то рассматриваемое выражение действительно даёт «полный заряд» (естественно, заряд алгебраический — полная масса $\delta_{(a)} - \delta_{(b)}$ равна нулю).

Определенное только что продолжение $\vec{\mu}$ обладает следующими очевидными свойствами.

Теорема 16. Пусть A и B — замкнутые множества из X с компактным пересечением K . Для меры Радона $\vec{\mu}$ на X со значениями в векторном нормированном пространстве \vec{E} и носителем в A и для непрерывной скалярной функции φ с носителем в B можно определить $\vec{\mu}(\varphi) \in \vec{E}$ с помощью формулы (IV, 2; 33), где α — функция из $\mathcal{C}(X)$, равная 1 на некоторой окрестности пересечения K . Результат не зависит от выбора α . Если φ_1 и φ_2 обладают теми же свойствами, что и φ , и если k является некоторым скаляром, то $\vec{\mu}(\varphi_1 + \varphi_2) = \vec{\mu}(\varphi_1) + \vec{\mu}(\varphi_2)$ и $\vec{\mu}(k\varphi) = k\vec{\mu}(\varphi)$. Если непрерывные функции φ_n с носителем в B равномерно сходятся к 0 при n , стремящемся к бесконечности, то $\vec{\mu}(\varphi_n) \in \vec{E}$ сходятся к 0. Кроме того, имеет место оценка

$$\|\vec{\mu}(\varphi)\| \leq \|\vec{\mu}\| \sup_{x \in X} |\varphi(x)|. \quad (\text{IV, 2; 33}_2)$$

Замечание. В условии теоремы достаточно было считать, что существует окрестность компакта K , на которой φ_n равномерно сходятся к 0 при $n \rightarrow \infty$. Отсюда уже следует, что $\vec{\mu}(\varphi_n)$ сходятся к 0.

Принцип кусочной склейки мер

Теорема 17. Пусть X — локально компактное пространство и $(\Omega_i)_{i \in I}$ — произвольное семейство открытых множеств, образующих покрытие X . Пусть $(\mu_i)_{i \in I}$ — семейство мер, зависящих от того же семейства индексов. Предположим, что μ_i является мерой, определенной на локально компактном пространстве Ω_i ¹⁾ со значениями в одном и том же нормированном пространстве \vec{E} , и что в каждом из пересечений $\Omega_i \cap \Omega_j$ двух открытых множеств меры μ_i и μ_j равны. Тогда на X существует, и притом единственная, мера μ со значениями в \vec{E} , которая в каждом открытом множестве Ω_i равна мере μ_i .

Доказательство. Единственность μ очевидна: если μ_1 и μ_2 — две меры, обладающие указанными свойствами, то $\mu_1 = \mu_2$ равна нулю в каждом Ω_i , а, следовательно, по теореме 13 равна нулю в их объединении, которое совпадает с X .

Значит, $\mu_1 = \mu_2$.

Доказательство существования можно было бы провести с помощью модифицированной теоремы 11 (см. замечание 3°)) без какого-либо предположения о счетности X . Однако для простоты мы предположим, что X является объединением счетного множества компактов, и применим разложение единицы.

Итак, пусть $(a_i)_{i \in I}$ — непрерывное разложение единицы, подчиненное покрытию $(\Omega_i)_{i \in I}$ пространства X . Пусть φ — некоторая функция из $\mathcal{C}(X)$. Функция $a_i \varphi$ принадлежит $\mathcal{C}(\Omega_i)$. Поэтому к ней можно применить меры μ_i и рассматривать $\mu_i(a_i \varphi)$. Определим теперь меру μ по формуле

$$\mu(\varphi) = \sum_{i \in I} \mu_i(a_i \varphi). \quad (\text{IV}, 2; 33_3)$$

Эта формула имеет определенный смысл. В самом деле, носитель функции φ компактен, и на этом компакте все функции a_i , кроме конечного их числа, равны нулю, так что предыдущая сумма конечна. Она, очевидно, определяет линейное отображение пространства $\mathcal{C}(X)$ в банахово пространство \vec{E} . Из неравенства

$$\|\mu(\varphi)\| \leq \sum_{i \in I} \|\mu_i\|_K \|a_i \varphi\| \leq \left(\sum_{i \in I} \|\mu_i\|_K \right) \|\varphi\|, \quad (\text{IV}, 2; 33_4)$$

¹⁾ Открытое множество локально компактного пространства локально компактно.

где J — такое конечное множество индексов, при которых α_i не равны тождественно нулю на K , следует, что это отображение непрерывно на $\mathcal{C}_K(X)$.

Введенное отображение определяет, очевидно, некоторую меру на X со значениями в \vec{E} . Остается доказать, что в множествах Ω_j имеет место равенство $\overset{\rightarrow}{\mu} = \overset{\rightarrow}{\mu_j}$. Если компактный носитель функции φ лежит в Ω_j , то для каждого i носитель $\alpha_i \varphi$ содержится в $\Omega_j \cap \Omega_i$. Поскольку в этом открытом множестве $\overset{\rightarrow}{\mu_j} = \overset{\rightarrow}{\mu_i}$, то в формуле (IV, 2; 33₃) можно заменить $\overset{\rightarrow}{\mu_i}(\alpha_i \varphi)$ на $\overset{\rightarrow}{\mu_j}(\alpha_i \varphi)$ и получить равенство

$$\overset{\rightarrow}{\mu}(\varphi) = \sum_{i \in I} \overset{\rightarrow}{\mu_i}(\alpha_i \varphi) = \overset{\rightarrow}{\mu_j} \left(\sum_{i \in I} \alpha_i \varphi \right) = \overset{\rightarrow}{\mu_j}(\varphi), \quad (\text{IV, 2; 33}_5)$$

из которого следует утверждение теоремы.

Комплексные и вещественные меры

Определение. До сих пор мы не уточняли характер поля скаляров \mathbb{K} , которое можно было отождествлять либо с \mathbb{R} , либо с \mathbb{C} . Через $\mathcal{C}(X)$ мы обозначали пространство непрерывных скалярных функций на X с компактным носителем. Оба возможных случая поля \mathbb{K} рассматривались нами одновременно. Однако очевидно, что в некоторых задачах целесообразно уточнить, о каком поле скаляров идет речь.

Пусть $\mathcal{C}(X; \mathbb{C}) = \mathcal{C}(\bar{X})$ — пространство комплексных функций, и пусть μ — комплексная мера на X , $\mu \in \mathcal{C}'(X)$.

Говорят, что мера μ *вещественна*, если $\mu(\varphi)$ — вещественна¹⁾ для любой вещественной функции φ .

Для того чтобы мера Радона μ была вещественной, необходимо и достаточно, чтобы для любой комплексной функции φ из $\mathcal{C}(X)$ имело место равенство

$$\mu(\bar{\varphi}) = \overline{\mu(\varphi)}, \quad (\text{IV, 2; 34})$$

другими словами, чтобы $\mu(\varphi)$ и $\mu(\bar{\varphi})$ были комплексно сопряженными.

В самом деле, пусть μ является вещественной мерой. Любая комплексная функция φ может быть представлена единственным

¹⁾ Вот другая интерпретация. Пусть μ — комплексная мера, т. е. \mathbb{C} -линейное отображение $\mathcal{C}(X; \mathbb{C})$ в поле скаляров \mathbb{C} . Оно определяет (см. примечание на стр. 451) некоторую меру μ_R относительно \mathbb{R} со значениями в \mathbb{C} , т. е. определяет некоторое \mathbb{R} -линейное отображение $\mathcal{C}(X; \mathbb{R})$ в двумерное векторное пространство \mathbb{C} над \mathbb{R} . При этом μ вещественна тогда и только тогда, когда μ_R будет принимать свои значения в подпространстве \mathbb{R} пространства \mathbb{C} .

образом в виде $\varphi = \varphi_1 + i\varphi_2$, где φ_1 и φ_2 — вещественные функции. Эти функции необходимо непрерывны и имеют компактный носитель. Так как $\mu(\varphi) = \mu(\varphi_1) + i\mu(\varphi_2)$ и $\mu(\bar{\varphi}) = \mu(\varphi_1) - i\mu(\varphi_2)$, а $\mu(\varphi_1)$, $\mu(\varphi_2)$ вещественны, то получаем соотношение (IV, 2; 34). Обратно, если это соотношение справедливо, то это говорит о том, что для каждой вещественной функции φ число $\mu(\varphi)$ совпадает со своим сопряженным, т. е. вещественно.

Пусть теперь μ — произвольная мера Радона. Обозначим через $\bar{\mu}$ комплексно сопряженную меру Радона, определяемую формулой

$$\bar{\mu}(\varphi) = \overline{\mu(\bar{\varphi})} \quad \text{для } \varphi \in \mathcal{C}(X), \text{ или } \overline{\mu(\varphi)} = \bar{\mu}(\bar{\varphi}). \quad (\text{IV}, 2; 35)$$

Прежде всего, это настоящая мера Радона. В самом деле, формула определяет некоторую комплексную функцию на $\mathcal{C}(X)$. Функция эта линейна, поскольку

$$\begin{aligned} \bar{\mu}(\varphi_1 + \varphi_2) &= \overline{\mu(\bar{\varphi}_1 + \bar{\varphi}_2)} = \overline{\mu(\bar{\varphi}_1) + \mu(\bar{\varphi}_2)} = \overline{\mu(\bar{\varphi}_1)} + \overline{\mu(\bar{\varphi}_2)} = \\ &= \bar{\mu}(\varphi_1) + \bar{\mu}(\varphi_2), \end{aligned} \quad (\text{IV}, 2; 36)$$

$$\bar{\mu}(k\varphi) = \overline{\mu(k\bar{\varphi})} = \overline{k\mu(\bar{\varphi})} = k\overline{\mu(\bar{\varphi})} = k\bar{\mu}(\varphi)^1, \quad (\text{IV}, 2; 36)$$

и она непрерывна на каждом $\mathcal{C}_K(X)$, где K — компакт из X .

Конечно, комплексно сопряженной мерой для $\bar{\mu}$ будет μ , т. е. имеет место обратимость, согласно которой можно говорить, что меры μ и $\bar{\mu}$ комплексно сопряжены друг другу. Две меры μ и $\bar{\mu}$ комплексно сопряжены тогда и только тогда, когда их значения на двух комплексно сопряженных функциях φ и $\bar{\varphi}$ соответственно комплексно сопряжены (вторая формула (IV, 2; 35)), или, иначе, когда, для любой вещественной функции φ , $\mu(\varphi)$ и $\bar{\mu}(\varphi)$ комплексно сопряжены. Отображение $\mu \rightarrow \bar{\mu}$ является полулинейной биекцией пространства $\mathcal{C}'(X)$ на себя, т. е.

$$\begin{aligned} \overline{\mu_1 + \mu_2} &= \bar{\mu}_1 + \bar{\mu}_2, \\ \overline{k\mu} &= \bar{k}\bar{\mu}. \end{aligned} \quad (\text{IV}, 2; 37)$$

Мера вещественна тогда и только тогда, когда она совпадает со своей комплексно сопряженной (формула (IV, 2; 34)). Соотношение между μ и $\bar{\mu}$ можно еще выразить в следующей форме:

любая комплексная мера Радона μ может быть записана, и притом единственным образом, в виде $\mu = \mu_1 + i\mu_2$, где μ_1 и μ_2 — вещественные меры Радона; при этом $\bar{\mu} = \mu_1 - i\mu_2$.

¹⁾ Следует внимательно рассмотреть эту последовательность равенств; здесь существенно используется тот факт, что правая часть равенства (IV, 2; 35) дважды содержит символ сопряжения. Функция $\varphi \rightarrow \mu(\varphi)$ полулинейна и не является линейной точно так же, как и функция $\varphi \rightarrow \mu(\bar{\varphi})$.

В самом деле, если такое разложение возможно, то, учитывая (IV, 2; 37) и то, что вещественная мера совпадает со своей сопряженной, получаем $\bar{\mu} = \mu_1 - i\mu_2$ и, следовательно, имеет место единственное представление

$$\mu_1 = \frac{\mu + \bar{\mu}}{2}, \quad \mu_2 = \frac{\mu - \bar{\mu}}{2i}. \quad (\text{IV, 2; 38})$$

Обратно, обе меры, определенные этой формулой, совпадают со своими сопряженными и являются, следовательно, вещественными. При этом $\mu = \mu_1 + i\mu_2$.

Будет полезным собрать все полученные результаты в общую таблицу:

$$\varphi = \varphi_1 + i\varphi_2, \quad \bar{\varphi} = \varphi_1 - i\varphi_2; \quad \varphi_1 = \frac{\varphi + \bar{\varphi}}{2}, \quad \varphi_2 = \frac{\varphi - \bar{\varphi}}{2i};$$

$$\bar{\mu}(\varphi) = \overline{\mu(\bar{\varphi})}, \quad \overline{\mu(\varphi)} = \bar{\mu}(\bar{\varphi}); \quad (\text{IV, 2; 39})$$

$$\mu = \mu_1 + i\mu_2, \quad \bar{\mu} = \mu_1 - i\mu_2; \quad \mu_1 = \frac{\mu + \bar{\mu}}{2}, \quad \mu_2 = \frac{\mu - \bar{\mu}}{2i}.$$

Вещественные положительные меры

Говорят, что \mathbb{C} -мера Радона μ на X положительна, если для любой неотрицательной функции $\varphi \in \mathcal{C}(X)$ имеет место неравенство $\mu(\varphi) \geqslant 0$. Это записывается так: $\mu \geqslant 0$. Очевидно, для такой меры из $\varphi_1 \leqslant \varphi_2$ следует $\mu(\varphi_1) \leqslant \mu(\varphi_2)$.

Из определения следует, что положительная мера обязательна вещественна.

Мера Радона μ называется отрицательной (в обозначениях: $\mu \leqslant 0$), если $\mu(\varphi) \leqslant 0$ для любой функции $\varphi \geqslant 0$.

Для того чтобы мера была отрицательной, необходимо и достаточно, чтобы она была противоположной по знаку к некоторой положительной мере.

Если μ и ν — две вещественные меры на X , то говорят, что $\mu \geqslant \nu$, если $\mu - \nu \geqslant 0$. Это означает, что для любой функции $\varphi \geqslant 0$ из $\mathcal{C}(X)$ имеет место неравенство $\mu(\varphi) \geqslant \nu(\varphi)$.

Если $\mu \geqslant \nu \geqslant 0$ и $\varphi \geqslant \psi \geqslant 0$, то $\mu(\varphi) \geqslant \nu(\psi) \geqslant 0$.

З а м е ч а н и е. Пусть μ — некоторая мера $\geqslant 0$. Для любой вещественной функции φ имеет место неравенство $-|\varphi| \leqslant \varphi \leqslant |\varphi|$, а, значит, $-\mu(|\varphi|) \leqslant \mu(|\varphi|) \leqslant \mu(|\varphi|)$, или

$$|\mu(\varphi)| \leqslant \mu(|\varphi|). \quad (\text{IV, 2; 39}_2)$$

Это соотношение сохраняется и для комплексной функции φ , однако его доказательство в этом случае значительно сложнее. Пусть K — носитель φ и \mathcal{V} — компактная окрестность K . Пусть ψ — функция, построенная по функции φ согласно следствию 8

теоремы 11 и такая, что

$$\psi = \sum_{i=1}^N g_i \psi_i, \quad g_i \in \mathbb{C}, \quad \psi_i \in \mathcal{C}(X), \quad (\text{IV}, 2; 39_3)$$

где $\psi_i \geqslant 0$ и имеют носитель в \mathcal{V} . Тогда

$$\left| \varphi(x) - \sum_{i=1}^N g_i \psi_i(x) \right| \leq \frac{\varepsilon}{2 \|\mu\|_{\mathcal{V}}}, \quad (\text{IV}, 2; 39_4)$$

$$\left| |\varphi(x)| - \sum_{i=1}^N |g_i| \psi_i(x) \right| \leq \frac{\varepsilon}{2 \|\mu\|_{\mathcal{V}}}. \quad (\text{IV}, 2; 39_5)$$

Первое и второе неравенства последовательно дают

$$|\mu(\varphi)| \leq \sum_{i=1}^N |g_i| \mu(\psi_i) + \frac{\varepsilon}{2} \leq \mu(|\varphi|) + \varepsilon. \quad (\text{IV}, 2; 39_6)$$

Отсюда вытекает справедливость неравенства $(\text{IV}, 2; 39_2)$, поскольку ε произвольно.

Позже мы укажем другой метод, использующий теорию продолжения Лебега. Будем рассматривать μ как некоторую \mathbb{R} -меру (\mathbb{R} -линейное отображение $\mathcal{C}(X; \mathbb{R})$ в \mathbb{R}). Функция $\varphi \in \mathcal{C}(X; \mathbb{C})$ является функцией с векторными значениями в двумерном пространстве \mathbb{C} над \mathbb{R} . Поэтому $(\text{IV}, 2; 39_2)$ перейдет в формулу $(\text{IV}, 3; 27)$.

Из $(\text{IV}, 2; 39_2)$ следует, что для вычисления $\|\mu\|_K$ можно ограничиться рассмотрением вещественных неотрицательных функций φ , при этом

$$\|\mu\|_K = \sup_{\substack{\varphi \in \mathcal{C}_K(X) \\ 0 \leq \varphi \leq 1}} \mu(\varphi) \quad \text{и} \quad \|\mu\| = \sup_{\substack{\varphi \in \mathcal{C}(X) \\ 0 \leq \varphi \leq 1}} \mu(\varphi). \quad (\text{IV}, 2; 39_7)$$

Используя понятия, введенные в примечании на стр. 451, можно сказать, что если μ является вещественной положительной \mathbb{C} -мерой, то меры μ и μ_K имеют одинаковую норму. Позже (см. замечание на стр. 488) мы увидим, что это утверждение остается справедливым, если μ является произвольной вещественной \mathbb{C} -мерой. Но это, конечно, неверно, когда μ является произвольной комплексной \mathbb{C} -мерой (это видно из примера в примечании на стр. 451). Можно также сказать, что в формуле $\|\mu\|_K = \sup_{|\varphi| \leq 1} |\mu(\varphi)|$ следует брать всевозможные комплексные

функции φ . Если μ вещественна, то можно ограничиться вещественными функциями φ . Если $\mu \geqslant 0$, то достаточно брать $\varphi \geqslant 0$.

Бинарное отношение $\mu \leq v$ определяет на множестве вещественных мер на X некоторое отношение порядка. Для обоснования этого утверждения достаточно доказать единственное нетривиальное свойство: из неравенств $\mu \leq v$ и $\mu \geq v$ всегда следует равенство $\mu = v$. Для $\varphi \geq 0$ из этих неравенств следует равенство $\mu(\varphi) = v(\varphi)$. Но всякая вещественная функция из $\mathcal{C}(X)$ может быть представлена в виде разности двух неотрицательных функций из $\mathcal{C}(X)$ (например, $\varphi = |\varphi| - (|\varphi| - \varphi)$), а, следовательно, равенство справедливо для любой вещественной функции φ , а значит, и для комплексной функции $\varphi = \varphi_1 + i\varphi_2$, откуда $\mu = v$. Отношение порядка «совместимо с \mathbb{R} -векторной структурой пространства вещественных мер» в том смысле, что:

$$\left. \begin{array}{l} \text{из } \mu_1 \leq v_1 \text{ и } \mu_2 \leq v_2 \text{ следует } \mu_1 + \mu_2 \leq v_1 + v_2, \\ \text{из } \mu \leq v \text{ при } k \geq 0 \text{ следует } k\mu \leq kv. \end{array} \right\} (\text{IV}, 2; 39_8)$$

Векторное пространство над \mathbb{R} , снабженное структурой порядка, обладающей этими свойствами, называется *полуупорядоченным векторным пространством*. Первое соотношение выражает тот факт, что $\mu \geq v$ эквивалентно неравенству $\mu - v \geq 0$ и что $\mu \geq 0, v \geq 0$ влечет за собой $\mu + v \geq 0$. Второе означает, что при $k \geq 0$ из $\mu \geq 0$ следует $k\mu \geq 0$. Заметим, что в этом случае из $\mu \leq v$ следует $-v \leq -\mu$, т. е. следует неравенство $k\mu \geq kv$ при $k \leq 0$.

Пример. Рассмотрим меру, определенную на \mathbb{R} формулой (IV, 2; 15). Ее сопряженная мера равна

$$\sum_v \bar{c}_v \delta_{(a_v)} + \bar{p} dx. \quad (\text{IV}, 2; 40)$$

Мера μ вещественна тогда и только тогда, когда мера $\sum_v (c_v - \bar{c}_v) \delta_{(a_v)} + (p - \bar{p}) dx$ равна нулю. Эта мера будет равна нулю, если все числа c_v вещественны и вещественна функция p . Обратно, если эта мера равна нулю, то ее носитель пуст. Отсюда по теореме 15₂ в случае непрерывной функции p можно получить, что все величины $c_v - \bar{c}_v$ равны нулю и что равна нулю функция $p - \bar{p}$, иначе говоря, что вещественны c_v и вещественна функция p . Этот результат, естественно, может быть неверным, если функция p не является непрерывной. Конечно, c_v вещественны, но функция p не обязательно вещественна. Однако можно утверждать, и это существенно, что не обязательно равные p и $\operatorname{Re} p$ определяют одну и ту же меру $p dx = \operatorname{Re} p dx$ [поскольку $(p - \operatorname{Re} p) dx = \left(\frac{p - \bar{p}}{2}\right) dx = 0$] и что, следовательно, в том случае, когда p не вещественна, можно заменить ее вещественной частью $\operatorname{Re} p$ без изменения меры.

Мера μ , определенная формулой (IV, 2; 15), при $c_v \geq 0$ и $p \geq 0$ положительна. В случае непрерывной функции p можно доказать и обратное утверждение¹⁾.

В самом деле, если точки a_v считать различными и один из коэффициентов c_v отрицательным или если в некоторой точке a непрерывная функция p принимает значение $p(a) < 0$, то без труда доказывается, что функция $\varphi \geq 0$, построенная в примерах 1—3 на стр. 468—471, удовлетворяет неравенству $\mu(\varphi) < 0$, и мы получили противоречие.

Решетки

Полуупорядоченное множество E называется *решеткой*, если любая его часть, образованная из двух элементов, имеет точную верхнюю грань (т. е. наименьшую из мажорант) и точную нижнюю грань. То же самое будет и для любого конечного непустого подмножества E . Например, вещественная прямая \mathbb{R} является решеткой. Полуупорядоченное множество \mathbb{R}^x функций на X с вещественными значениями (см. стр. 23) есть решетка так же, как и множество $(\mathbb{R}^x)_c$ вещественных непрерывных функций на топологическом пространстве X . Напротив, множество полиномов на \mathbb{R} с вещественными коэффициентами относительно обычного отношения порядка, очевидно, не является решеткой. Если E — полуупорядоченное векторное пространство (что имело место в предыдущих примерах), то для того чтобы оно было решеткой, достаточно, чтобы для любого $e \in E$ существовал элемент $e^+ = \sup(e, 0)$. В самом деле, пусть e и f — два произвольных элемента из E . Тогда элемент $e + (f - e)^+$ мажорирует элемент $e = e + 0$ и $e + (f - e) = f$. Если g мажорирует e и f , то $g - e$ мажорирует $e - e = 0$ и $f - e$ т. е. мажорирует $(f - e)^+$. Значит, g мажорирует элемент $e + (f - e)^+$, являющийся, таким образом, точной верхней гранью e и f . С другой стороны, e и f имеют точную нижнюю грань, а именно $-\sup(-e, -f)$. Наконец, всегда имеет место равенство $\sup(e, f) + \inf(e, f) = e + f$. Действительно, очевидно, что $\sup(e - h, f - h) = \sup(e, f) - h$. Полагая $h = e + f$, получаем нужную формулу. Положим $e^- = \sup(-e, 0) = (-e)^+$. Тогда $e = e + 0 = \sup(e, 0) + \inf(e, 0) = e^+ + e^-$. Элементы e^+ и e^- называются положительной и отрицательной частью e . Несмотря на свое название, $e^- \geq 0$, а не ≤ 0 , и скорее $-e^-$ заслуживает названия отрицательной части²⁾. Теперь видно, что каждый

¹⁾ Если p не является непрерывной, то на те же вопросы дают полный ответ теоремы 52—54.

²⁾ Точно так же a и b называются вещественной и мнимой частью комплексного числа $a + bi$, хотя скорее bi заслуживает названия мнимой части.

элемент e из E является разностью двух неотрицательных элементов: e^+ и e^- . Такое представление возможно бесконечным числом способов, поскольку можно положить $e = e_1 - e_2$, где $e_1 = e^+ + a$ и $e_2 = e^- + a$, $a \geq 0$. Однако e^+ и e^- являются наименьшими элементами ≥ 0 , разность которых дает e . В самом деле, если $e = e_1 - e_2$, e_1 и $e_2 \geq 0$, то e_1 мажорирует e и 0, а, следовательно, и e^+ , e_2 мажорирует $-e$ и 0, а значит, и e^- . Мы возвращаемся к прежней форме: $e_1 = e^+ + a$, $e_2 = e^- + a$, $a \geq 0$. Элементы e^+ и e^- являются «дизъюнктными» в том смысле, что единственным элементом ≥ 0 , минорирующим оба элемента, является 0. Мало того, каждый элемент g , минорирующий одновременно e^+ и e^- , не больше нуля. В самом деле, высказанное утверждение означает, что $\inf(e^+, e^-) = 0$. Далее, $e^+ + \inf(0, e^- - e^+) = e^+ + \inf(0, -e) = e^+ - \sup(0, e) = e^+ - e^+ = 0$. Наконец, абсолютной величиной e называют точную верхнюю грань элементов e и $-e$: $|e| = \sup(e, -e)$.

Так как $|e| + e = \sup(2e, 0) = 2e^+$, то $|e| = 2e^+ - e = 2e^+ - (e^+ - e^-) = e^+ + e^-$. Поскольку $\inf(e^+, e^-) = 0$, то $e^+ + e^-$ также равен $\sup(e^+, e^-)$.

Говорят, что полуупорядоченное множество является полной решеткой, если любая его непустая мажорируемая часть имеет точную верхнюю грань, а любая непустая минорируемая часть имеет точную нижнюю грань (из первого условия, впрочем, следует второе, ибо если некоторая непустая часть A минорируема, то множество ее минорант мажорируемо и не пусто, а значит, имеет точную верхнюю грань, являющуюся точной нижней гранью части A). Вещественная прямая \mathbb{R} является полной решеткой точно так же, как и множество \mathbb{R}^X функций на X с вещественными значениями. Однако если X является топологическим пространством, то множество $(\mathbb{R}^X)_c$ вещественных непрерывных на X функций является решеткой, но не полной.

Например, пусть X — вещественная прямая \mathbb{R} и A — часть множества $(\mathbb{R}^{\mathbb{R}})_c$, образованная из непрерывных функций, мажорируемых характеристической функцией отрезка $[-1, 1]$. Точная верхняя грань множества A в полуупорядоченном множестве $\mathbb{R}^{\mathbb{R}}$ всевозможных вещественных функций на \mathbb{R} является этой характеристической функцией. Так как характеристическая функция отрезка $[-1, 1]$ разрывна, то A не имеет точной верхней грани в множестве $(\mathbb{R}^{\mathbb{R}})_c$ непрерывных функций.

Теорема 18. Полуупорядоченное векторное пространство вещественных мер Радона на локально компактном пространстве X является полной решеткой. Если μ является некоторой вещественной мерой, $\mu^+ = \sup(\mu, 0)$, $\mu^- = \sup(-\mu, 0)$, $|\mu| = \sup(\mu, -\mu) = \sup(\mu^+, \mu^-) = \mu^+ + \mu^-$, то имеют место сле-

дующие формулы:

$$\mu^+(\varphi) = \sup_{\substack{0 \leq \psi \leq \varphi \\ \psi \in \mathcal{C}(X)}} \mu(\psi) \quad \text{для } \varphi \geq 0, \quad (\text{IV}, 2; 41)$$

$$\|\mu\| = \|\mu^+\| + \|\mu^-\|, \quad (\text{IV}, 2; 41_2)$$

где нормы вычислены на X или на некотором компакте K пространства X ;

$$|\mu|(\varphi) = \sup_{\substack{|\psi| \leq |\varphi| \\ \psi \in \mathcal{C}(X)}} \mu(\psi) \quad \text{для } \varphi \geq 0 \quad (\text{IV}, 2; 41_3)$$

и

$$\|\mu\| = \|\mu\| = \|\mu^+\| + \|\mu^-\|$$

на X или на некотором компакте K из X .

Таким образом, для $\varphi \geq 0$, $\mu(\varphi) \leq \mu^+(\varphi)$, $-\mu(\varphi) \leq \mu^-(\varphi)$ и, для произвольного φ , $|\mu(\varphi)| \leq |\mu|(|\varphi|)$.

Доказательство. Обозначим через $\mathcal{C}^+(X)$ множество функций $\varphi \geq 0$ из $\mathcal{C}(X)$. Определим на $\mathcal{C}^+(X)$ функцию μ^+ со значениями в \mathbb{R}_+ по формуле (IV, 2; 41).

Она дает конечное значение, поскольку если K является носителем функции φ , то имеет место неравенство $|\mu(\varphi)| \leq \|\mu\|_K \|\varphi\| \leq \|\mu\|_K \|\varphi\|$. Неравенство $\mu^+(\varphi) \geq 0$ следует из того, что среди функций ψ имеется функция 0. Покажем, что эта функция μ^+ на $\mathcal{C}^+(X)$ обладает следующими свойствами «линейности»:

$$\begin{aligned} \mu^+(\varphi_1 + \varphi_2) &= \mu^+(\varphi_1) + \mu^+(\varphi_2), \\ \mu^+(\lambda\varphi) &= \lambda\mu^+(\varphi) \quad \text{при } \lambda \geq 0. \end{aligned} \quad (\text{IV}, 2; 41_4)$$

Второе свойство очевидно. Поэтому достаточно доказать лишь первое. Каковы бы ни были функции ψ_1 и ψ_2 , такие, что $0 \leq \psi_1 \leq \varphi_1$ и $0 \leq \psi_2 \leq \varphi_2$, имеют место неравенства $0 \leq \psi_1 + \psi_2 \leq \varphi_1 + \varphi_2$, откуда $\mu(\psi_1) + \mu(\psi_2) = \mu(\psi_1 + \psi_2) \leq \mu^+(\varphi_1 + \varphi_2)$. Переходя к точной верхней грани относительно ψ_1 и ψ_2 , отсюда непосредственно получаем:

$$\mu^+(\varphi_1) + \mu^+(\varphi_2) \leq \mu^+(\varphi_1 + \varphi_2).$$

Впрочем, если $\psi \in \mathcal{C}^+(X)$, $0 \leq \psi \leq \varphi_1 + \varphi_2$, то можно найти такие две функции ψ_1 и ψ_2 , которые удовлетворяли бы соотношениям:

$$0 \leq \psi_1 \leq \varphi_1, \quad 0 \leq \psi_2 \leq \varphi_2 \quad \text{и} \quad \psi_1 + \psi_2 = \psi.$$

Для этого достаточно взять

$$\psi_1 = \inf(\psi, \varphi_1) \quad \text{и} \quad \psi_2 = \psi - \psi_1.$$

Из определения следует, что имеют место соотношения $0 \leq \psi_1 \leq \varphi_1$, $0 \leq \psi_2$ и $\psi_1 + \psi_2 = \psi$. Пусть теперь $x \in X$. Если $\psi(x) \leq \varphi_1(x)$, то $\psi_1(x) = \psi(x)$ и, следовательно, $\psi_2(x) = 0 \leq \varphi_2(x)$.

Если же, напротив, $\psi(x) > \varphi_1(x)$, то $\psi_1(x) = \varphi_1(x)$, откуда $\psi_2(x) = \psi(x) - \varphi_1(x) \leqslant \varphi_1(x) + \varphi_2(x) - \varphi_1(x) = \varphi_2(x)$.

Во всех случаях получаем $0 \leqslant \psi_2 \leqslant \varphi_2$.

Теперь будет справедливо неравенство

$$\mu(\psi) = \mu(\psi_1) + \mu(\psi_2) \leqslant \mu^+(\varphi_1) + \mu^+(\varphi_2),$$

откуда, переходя к точной верхней грани относительно ψ , получаем неравенство $\mu^+(\varphi_1 + \varphi_2) \leqslant \mu^+(\varphi_1) + \mu^+(\varphi_2)$, заканчивающее доказательство соотношения (IV, 2; 41₄).

Пусть теперь задана вещественная функция φ из $\mathcal{C}(X)$ произвольного знака. Ее можно бесконечным числом способов записать в виде разности $\varphi = \varphi_1 - \varphi_2$ двух функций $\geqslant 0$ из $\mathcal{C}(X)$. Достаточно, например, положить $\varphi = \varphi^+ - \varphi^-$, поскольку обе непрерывные функции φ^+ и φ^- имеют компактный носитель.

Определим теперь $\mu^+(\varphi)$ по формуле

$$\mu^+(\varphi) = \mu^+(\varphi_1) - \mu^+(\varphi_2).$$

Такое определение будет обоснованным лишь в том случае, когда мы докажем независимость $\mu^+(\varphi)$ относительно разложения $\varphi = \varphi_1 - \varphi_2$ ¹⁾. Итак, если $\varphi = \varphi_3 - \varphi_4$ — другое разложение, то $\varphi_1 + \varphi_4 = \varphi_2 + \varphi_3$, откуда, согласно (IV, 2; 41₄), $\mu^+(\varphi_1) + \mu^+(\varphi_4) = \mu^+(\varphi_2) + \mu^+(\varphi_3)$ и, следовательно, $\mu^+(\varphi_1) - \mu^+(\varphi_2) = \mu^+(\varphi_3) - \mu^+(\varphi_4)$, что полностью доказывает эту независимость. Мера μ^+ теперь определена как некоторая функция на $\mathcal{C}(X; \mathbb{R})$ со значениями в \mathbb{R} . Очевидно, кроме того, что эта функция μ^+ линейна на этом векторном пространстве. В самом деле, формула $\mu^+(\lambda\varphi) = \lambda\mu^+(\varphi)$ очевидна и достаточно лишь доказать аддитивность. Если φ и ψ — две вещественные функции из $\mathcal{C}(X)$, то можно их записать в виде $\varphi = \varphi_1 - \varphi_2$ и $\psi = \psi_1 - \psi_2$, где $\varphi_1, \varphi_2, \psi_1, \psi_2 \geqslant 0$. Теперь можно записать разложение $\varphi + \psi = (\varphi_1 + \psi_1) - (\varphi_2 + \psi_2)$, из которого с помощью (IV, 2; 41₄) следует формула

$$\begin{aligned} \mu^+(\varphi + \psi) &= \mu^+(\varphi_1 + \psi_1) - \mu^+(\varphi_2 + \psi_2) = \\ &= \mu^+(\varphi_1) + \mu^+(\psi_1) - \mu^+(\varphi_2) - \mu^+(\psi_2) = \mu^+(\varphi) + \mu^+(\psi), \end{aligned}$$

заканчивающая доказательство линейности μ^+ .

Проверим теперь свойство непрерывности, которое нам позволит утверждать, что μ^+ является некоторой скалярной \mathbb{R} -мерой.

Пусть K — некоторый компакт и α — функция $\geqslant 0$ из $\mathcal{C}(X)$, равная 1 в некоторой окрестности компакта K (следствие теоремы 11). Положим $\mu^+(\alpha) = k$. Тогда для вещественной функ-

1) Если пользоваться только φ^+ и φ^- , то это доказательство не нужно. Однако тогда мы не сможем доказать линейности μ^+ , так как $(\varphi + \psi)^+ \neq \varphi^+ + \psi^+$.

ции φ имеют место неравенства

$$-\|\varphi\| \alpha \leqslant \varphi \leqslant \|\varphi\| \alpha, \quad (\text{IV}, 2; 41_5)$$

а, следовательно,

$$-\|\varphi\| \mu^+(\alpha) \leqslant \mu^+(\varphi) \leqslant \|\varphi\| \mu^+(\alpha) \text{ или } |\mu^+(\varphi)| \leqslant k \|\varphi\|,$$

из которых следует непрерывность μ^+ на $\mathcal{C}(X; \mathbb{R})$ и справедливость неравенства $\|\mu^+\|_K \leqslant k = \mu^+(\alpha)$ ¹⁾. Этим заканчивается доказательство того, что μ^+ является некоторой \mathbb{R} -мерой. Мера эта, очевидно, $\geqslant 0$.

По определению, $\mu^+ \geqslant \mu$. Кроме того, каждая мера $\nu \geqslant 0$, мажорирующая меру μ , т. е. мажорирующую μ и 0, для любой функции $\varphi \geqslant 0$ и $0 \leqslant \psi \leqslant \varphi$ удовлетворяет неравенствам $\nu(\varphi) \geqslant \nu(\psi) \geqslant \mu(\psi)$, а, следовательно, после перехода к точной верхней грани относительно ψ , — неравенству $\nu(\varphi) \geqslant \mu^+(\varphi)$, из которого вытекает соотношение $\nu \geqslant \mu^+$. Последнее означает, что μ^+ является точной верхней гранью μ и 0 в полуупорядоченном векторном пространстве вещественных мер на X . Значит, это пространство является решеткой.

Для произвольного компакта K имеет место неравенство $\|\mu\|_K \leqslant \|\mu^+\|_K + \|\mu^-\|_K$. Покажем, что на самом деле здесь имеет место равенство. Каково бы ни было $\varepsilon > 0$, существует такая функция $\varphi_1 \in \mathcal{C}_K(X)$, при которой имеет место неравенство

$$\|\mu^+\| \geqslant \mu^+(\varphi_1) \geqslant \|\mu^+\|_K - \frac{\varepsilon}{4}, \quad 0 \leqslant \varphi_1 \leqslant 1.$$

(Можно взять $\varphi_1 \geqslant 0$, поскольку $\mu^+ \geqslant 0$. См. формулу (IV, 2; 39₇).)

Существует теперь функция $\psi_1 \in \mathcal{C}_K(X)$, $0 \leqslant \psi_1 \leqslant \varphi_1 \leqslant 1$, такая, что

$$\mu(\psi_1) \geqslant \mu^+(\varphi_1) - \frac{\varepsilon}{4} \geqslant \|\mu^+\|_K - \frac{\varepsilon}{2}.$$

Точно так же существует такая функция $\psi_2 \in \mathcal{C}_K(X)$, $0 \leqslant \psi_2 \leqslant 1$, при которой

$$-\mu(\psi_2) \geqslant \|\mu^-\|_K - \frac{\varepsilon}{2}.$$

Функция $\theta = \psi_1 - \psi_2$ лежит в $\mathcal{C}_K(X)$ и удовлетворяет условиям $|\theta| \leqslant 1$ и $\mu(\theta) \geqslant \|\mu^+\|_K + \|\mu^-\|_K - \varepsilon$. Наше утверждение следует из того, что ε произвольно. Тот же результат имеет место для норм, рассматриваемых на пространстве X .

Рассмотрим теперь меру $|\mu| = \mu^+ + \mu^-$ и покажем, что она может быть определена по формуле (IV, 2; 41₃). Прежде всего, так как для $\psi = \psi^+ - \psi^-$; $\mu(\psi) = \mu(\psi^+) - \mu(\psi^-) \leqslant \mu^+(\psi^+) +$

¹⁾ Мимоходом мы доказали общее свойство: любая линейная форма на $\mathcal{C}(X) \geqslant 0$ над $\mathcal{C}^+(X)$ является мерой, т. е. непрерывна на каждом $\mathcal{C}_K(X)$, где K — компакт, содержащийся в X .

$+\mu^-(\psi^-) \leq \mu^+(\varphi) + \mu^-(\varphi) = |\mu|(\varphi)$, то левая часть формулы (IV, 2; 41₃) не меньше правой ее части. Однако существуют такие функции $\psi_1, \psi_2, 0 \leq \psi_1 \leq \varphi, 0 \leq \psi_2 \leq \varphi$, для которых

$$\mu(\psi_1) \geq \mu^+(\varphi) - \frac{\varepsilon}{2}, \quad -\mu(\psi_2) \geq \mu^-(\varphi) - \frac{\varepsilon}{2}.$$

Функция $0 = \psi_1 - \psi_2$ удовлетворяет неравенству $|\theta| \leq \varphi$. При этом $\mu(\theta) \geq \mu^+(\varphi) + \mu^-(\varphi) - \varepsilon = |\mu|(\varphi) - \varepsilon$. В силу произвольности $\varepsilon > 0$ отсюда получаем соотношение (IV, 2; 41₃), которое говорит о том, что

$$\|\mu\|_K = \sup_{\substack{0 \leq \varphi \leq 1 \\ \varphi \in \mathcal{C}_K(X)}} |\mu|(\varphi) = \sup_{\substack{|\psi| \leq 1 \\ \psi \in \mathcal{C}_K(X)}} |\mu(\psi)| = \|\mu\|_K,$$

а, следовательно, $\|\mu\|_K = \|\mu^+\|_K + \|\mu^-\|_K$.

Мы определили μ^+ , μ^- , $|\mu|$ как скалярные \mathbb{R} -меры (\mathbb{R} -линейные отображения $\mathcal{C}(X; \mathbb{R})$ в \mathbb{R}). Полагая $\mu^+(\varphi_1 + i\varphi_2) = \mu^+(\varphi_1) + i\mu^+(\varphi_2)$ и т. д., можно определить их как скалярные \mathbb{C} -меры, вещественные, ≥ 0 в смысле, указанном на стр. 476—478.

Согласно формуле (IV, 2; 39₇), для \mathbb{C} -меры ≥ 0 норма вычисляется, исходя из ее значения на функциях ≥ 0 . Откуда получается, что формула $\|\mu\| = \|\mu\| = \|\mu^+\| + \|\mu^-\|$ сохраняется для вещественных \mathbb{C} -мер. Кроме того, для комплексных функций ψ (и, конечно, $\varphi \geq 0$) сохраняется формула (IV, 2; 41₃) вместе со своим следствием (для комплексной φ) $|\mu(\varphi)| \leq \|\mu\|(|\varphi|)$ (поскольку $|\mu(\varphi)| \leq |\mu^+(\varphi)| + |\mu^-(\varphi)| \leq \|\mu^+(\varphi)\| + \|\mu^-(\varphi)\| = \|\mu\|(|\varphi|))$.

Остается, наконец, показать, что пространство вещественных мер является полной решеткой. (Это замечательно! Напомним, что пространство $\mathcal{C}(X; \mathbb{R})$ не является полной решеткой.) Итак, пусть $(\mu_i)_{i \in I}$ — произвольное мажорируемое семейство вещественных мер. Для него существует такая мера v , что для всех i $\mu_i \leq v$. Для каждого конечного непустого подмножества J множества I обозначим через μ_J точную верхнюю грань мер μ_i , $i \in J$. Очевидно, $\mu_K \geq \mu_J$, если $K \supset J$. Надо лишь доказать, что множество μ_J имеет точную верхнюю грань. Пусть $\varphi \geq 0$. Положим

$$\mu_0(\varphi) = \sup_J \mu_J(\varphi). \quad (\text{IV, 2; 41}_6)$$

(Внимание! Написанное вовсе не одно и то же, что $\sup_{i \in I} \mu_i(\varphi)$.)

Это дает совсем другой результат, ибо уже для двух мер μ_1 и μ_2 равенство $(\sup(\mu_1, \mu_2))(\varphi) = \sup(\mu_1(\varphi), \mu_2(\varphi))$ не справедливо: Например, для μ и 0 равенство $\mu^+(\varphi) = \sup(\mu(\varphi), 0) = (\mu(\varphi))^+$ места не имеет.) Эта величина конечна, так как она мажорируется величиной $v(\varphi)$.

Таким путем определяется некоторая функция на $\mathcal{C}^+(X)$. Докажем, что эта функция «линейна» в смысле формулы (IV, 2; 41₄). Относительно умножения на скаляр это очевидно. Поэтому достаточно рассмотреть свойство аддитивности. Из неравенства

$$\mu_J(\varphi_1 + \varphi_2) = \mu_J(\varphi_1) + \mu_J(\varphi_2) \leq \mu_0(\varphi_1) + \mu_0(\varphi_2)$$

переходом к точной верхней грани относительно J получаем $\mu_0(\varphi_1 + \varphi_2) \leq \mu_0(\varphi_1) + \mu_0(\varphi_2)$. Однако существует такое конечное множество $J \subset I$, при котором $\mu_J(\varphi_1) \geq \mu_0(\varphi_1) - \varepsilon/2$, и такое конечное множество K , что $\mu_K(\varphi_2) \geq \mu_0(\varphi_2) - \varepsilon/2$. Поэтому, обозначая $L = J \cup K$, получим:

$$\begin{aligned} \mu_0(\varphi_1 + \varphi_2) &\geq \mu_L(\varphi_1 + \varphi_2) = \mu_L(\varphi_1) + \mu_L(\varphi_2) \geq \mu_J(\varphi_1) + \mu_K(\varphi_2) \geq \\ &\geq \mu_0(\varphi_1) + \mu_0(\varphi_2) - \varepsilon, \end{aligned}$$

откуда и вытекает искомый результат. Затем можно определить μ_0 на $\mathcal{C}(X; \mathbb{R})$ тем же самым способом, какой был использован при определении μ^+ из μ в начале доказательства этой теоремы. Мера μ_0 будет представлять собой \mathbb{R} -линейную форму на $\mathcal{C}(X; \mathbb{R})$. Надо лишь доказать ее непрерывность на $\mathcal{C}_K(X; \mathbb{R})$. Применим способ, использованный прежде для μ^+ . Выберем произвольный индекс i . Тогда для $i \in J$ из неравенства $\mu_J \leq v$ следует $\mu_J^+ \leq v^+$, а из неравенства $-\mu_J \leq -\mu_i$ следует $\mu_J^- \leq \mu_i^-$. Поэтому для любой вещественной функции φ

$$\mu_J(\varphi) = \mu_J(\varphi^+) - \mu_J(\varphi^-) \leq v^+(\varphi^+) + \mu_i^-(\varphi^-) \leq v^+(|\varphi|) + \mu_i^-(|\varphi|).$$

Точно так же

$$\begin{aligned} -\mu_J(\varphi) &= -\mu_J(\varphi^+) + \mu_J(\varphi^-) \leq \\ &\leq \mu_i^-(\varphi^+) + v^+(\varphi^-) \leq v^+(|\varphi|) + \mu_i^-(|\varphi|). \end{aligned}$$

Окончательно: $|\mu_J(\varphi)| \leq (v^+ + \mu_i^-)(|\varphi|)$, или, переходя к точной верхней грани относительно J , $|\mu_0(\varphi)| \leq (v^+ + \mu_i^-)(|\varphi|)$. Отсюда сразу же вытекает непрерывность μ_0 и, кроме того, неравенство

$$\|\mu_0\|_K \leq \|v^+\|_K + \|\mu_i^-\|_K.$$

Теперь μ_0 является \mathbb{R} -мерой (ее можно, очевидно, продолжить в $\mathcal{C}(X; \mathbb{C})$ до некоторой вещественной \mathbb{C} -меры). Она очевидным образом мажорирует все меры μ_J и по определению (IV, 2; 41₆) является наименьшей мерой, обладающей этим свойством. Следовательно, μ_0 является точной верхней гранью мер μ_J и, значит, множество вещественных мер на X является полной решеткой.

Замечание. Мы видели (формула (IV, 2; 39₇)), что норма \mathbb{C} -меры $\mu \geq 0$ может быть вычислена с помощью $\mu(\varphi)$, где φ — вещественные функции ≥ 0 . Теперь мы можем сказать, что норма вещественной \mathbb{C} -меры μ может быть вычислена с помощью $\mu(\varphi)$ с вещественными функциями φ (и что она, следовательно, равна норме μ_R ; см. примечание на стр. 451). Это как раз то, что мы видели во время доказательства теоремы. Конечно, формула (IV, 2; 39₇) не верна. Однако, в силу соотношения (IV, 2; 41₃), для вещественной или комплексной φ всегда имеет место неравенство $|\mu(\varphi)| \leq |\mu|(|\varphi|)$.

Говорят, что векторная мера μ на X *абсолютно мажорируется* вещественной мерой $\lambda \geq 0$, если для любой вещественной неотрицательной функции $\varphi \in \mathcal{C}(X)$ имеет место неравенство

$$\overrightarrow{\mu}(\varphi) \leq \lambda(\varphi). \quad (\text{IV}, 2; 42)$$

Если φ — вещественная функция произвольного знака, то

$$\|\overrightarrow{\mu}(\varphi)\| = \|\overrightarrow{\mu}(\varphi^+) - \overrightarrow{\mu}(\varphi^-)\| \leq \lambda(\varphi^+) + \lambda(\varphi^-) = \lambda(|\varphi|). \quad (\text{IV}, 2; 42_2)$$

С помощью разложения единицы (см. доказательство соотношения (IV, 2; 39₂)) можно показать, что это неравенство сохраняется и для комплексной функции φ . Если μ является вещественной \mathbb{C} -мерой, то из формулы (IV, 2; 41₃) следует, что наименьшей мерой, мажорирующей по модулю меру μ , является мера $|\mu|$.

Пусть \vec{E} — конечномерное векторное пространство, и пусть e_1, e_2, \dots, e_n — базис \vec{E} как векторного пространства над полем \mathbb{R} . Для каждой вещественной функции φ можно написать разложение

$$\overrightarrow{\mu}(\varphi) = \sum_{i=1}^n \mu_i(\varphi) \vec{e}_i,$$

где $\mu_i: \varphi \mapsto \mu_i(\varphi)$ являются \mathbb{R} -линейными формами на $\mathcal{C}(X; \mathbb{R})$, непрерывными на $\mathcal{C}_K(X; \mathbb{R})$ (K — компакт пространства X), т. е. скалярными \mathbb{R} -мерами на X . Так как для каждого $\varphi \geq 0$

$$\|\overrightarrow{\mu}(\varphi)\| \leq \max_{i=1, 2, \dots, n} \|\vec{e}_i\| \sum_{i=1}^n |\mu_i|(\varphi),$$

то $\overrightarrow{\mu}$ абсолютно мажорируется положительной мерой $\max_{i=1, 2, \dots, n} \|\vec{e}_i\| \sum_{i=1}^n |\mu_i|$. Но тогда она имеет наименьшую абсолютную мажоранту. Более общо, любая мера со значениями в банаховом пространстве \vec{E} , обладающая абсолютной мажо-

рантой, имеет наименьшую абсолютную мажоранту, а именно: точную нижнюю грань минорируемого нулем непустого множества абсолютных мажорант (множество вещественных мер является полной решеткой).

Однако мера со значениями в бесконечномерном банаховом пространстве, вообще говоря, не имеет абсолютной мажоранты. Рассмотрим, например, случай компактного пространства X . Пусть $\vec{E} = \mathcal{C}(X)$ — пространство с обычной нормой. Тождественное отображение $\varphi \rightarrow \varphi$ определяет на X меру μ со значениями в $\vec{E} = \mathcal{C}(X)$. У этой меры нет абсолютной мажоранты (кроме случая, когда X конечно). В самом деле, пусть λ — ее абсолютная мажоранта. Тогда для любой функции $\varphi \geq 0$ $\|\varphi\| = \|\mu(\varphi)\| \leq \lambda(\varphi)$. Если X — бесконечное множество, то для любого n можно найти n попарно не пересекающихся открытых множеств (следствие 6 теоремы 11, примененное к n замкнутым множествам, сводящимся к отдельным точкам). Для каждого из них можно найти непрерывные функции a_i , $0 \leq a_i \leq 1$, с носителями в этих множествах, принимающие значение 1 не менее чем в одной точке (лемма 1 к теореме 11). Так как

$$1 = \|a_i\| \leq \lambda(a_i), \text{ то } n \leq \sum_{i=1}^n \lambda(a_i) = \lambda\left(\sum_{i=1}^n a_i\right) \leq \lambda(1), \text{ что невозможno, поскольку } \lambda(1) \text{ конечно. Мы получаем, что } \mu \text{ абсолютной мажоранты не имеет.}$$

Замечание. Мы определили $|\mu|$ для вещественной \mathbb{C} -меры μ , но не для комплексной \mathbb{C} -меры. Теперь возможно сделать и это. Более общо, если μ является мерой со значениями в конечномерном векторном пространстве \vec{E} над полем \mathbb{R} или \mathbb{C} , то в качестве $|\mu|$ можно взять ее наименьшую абсолютную мажоранту. Эта мажоранта является некоторой мерой ≥ 0 и $\|\mu(\varphi)\| \leq |\mu|(|\varphi|)$. В соотношении $\|\mu\| \leq \|\mu\|$ в случае скалярной меры (вещественной \mathbb{R} -меры или комплексной \mathbb{C} -меры) имеет место равенство. Равенство не будет выполняться в общем случае векторной меры (см. примечание к теореме 15), а, значит в частности, для \mathbb{R} -меры со значением в \mathbb{C} .

§ 3. ПРОДОЛЖЕНИЕ ПОЛОЖИТЕЛЬНОЙ МЕРЫ. ТЕОРИЯ ЛЕБЕГА

Пусть μ — некоторая мера ≥ 0 на локально компактном пространстве X . Постараемся определить меру $\mu(\varphi)$ для функций φ , отличных от непрерывных функций с компактным носителем. Точно так же в случае меры $\mu = dx$ с помощью теории

Римана мы могли определить $\int \phi dx$ для функций ϕ , отличных от непрерывных функций, имеющих компактные носители. Впрочем, если $\mu = \delta_{(a)}$, то совершенно естественно положить $\mu(\phi) = \delta_{(a)}(\phi) = \phi(a)$ для любой комплексной функции ϕ на X .

Излагаемая здесь теория Лебега является значительно более общей, чем теория Римана. Даже в случае меры $\mu = dx$ множество интегрируемых по Лебегу функций шире, чем множество функций, интегрируемых по Риману. Однако эта теория сложнее, тоньше и зачастую скучнее. Поэтому мы позволим себе принять некоторые из теорем без доказательства¹⁾.

Раз и навсегда предположим, что X является локально компактным, счетным в бесконечности пространством, т. е. пространством, являющимся объединением счетного множества компактов²⁾. Это будет, например, случай конечномерных векторных пространств, поскольку их всегда можно рассматривать как счетное объединение замкнутых шаров, т. е. компактных частей. Все локально компактные пространства, обычно встречающиеся в анализе, обладают этим свойством. Поэтому вводимое ограничение не является особо существенным. Оно не является необходимым для всей теории, но потребуется для нескольких фундаментальных теорем. С другой стороны, мы всегда будем считать, что на X задана мера Радона $\mu \geq 0$. Говорят, что X является пространством с мерой, если X локально компактно, счетно в бесконечности и снабжено мерой $\mu \geq 0$. Тот факт, что $\mu \geq 0$, существен.

Внешние меры открытых множеств

Пусть \mathcal{O} — открытая часть X . Внешней мерой открытого множества \mathcal{O} относительно меры μ называется норма $\|\mu\|_{\mathcal{O}}$ сужения μ на это открытое множество, т. е. точная верхняя грань $\mu(\phi)$ на всех функциях ϕ из $\mathcal{C}(X)$ с носителем в \mathcal{O} , удовлетворяющих неравенству $0 \leq \phi \leq 1$. Эта мера будет обозначаться через $\mu^*(\mathcal{O})$. Число $\mu^*(\mathcal{O})$ неотрицательно, конечно или бесконечно.

Для мер открытых множеств можно доказать ряд свойств. Некоторые из этих свойств очевидны.

1) Приводимый здесь метод изложения теории Лебега не является ни самым кратким, ни самым лучшим. Однако: 1) будет введено минимальное число новых понятий, а именно: не будет необходимости в изучении полу-непрерывных снизу функций; 2) одновременно получатся интегралы от векторных функций; 3) все, что здесь вводится, во всяком случае постоянно применяется в анализе.

2) Мы повторим это условие, если оно будет существенным для справедливости теоремы. Такое ограничение уже вводилось в теореме 11.

Например, если \mathcal{O}_1 и \mathcal{O}_2 — два открытых множества и если $\mathcal{O}_1 \subset \mathcal{O}_2$, то в силу положительности μ получаем: $\mu^*(\mathcal{O}_1) \leq \mu^*(\mathcal{O}_2)$.

С другой стороны, $\mu^*(\emptyset) = 0$ и $\mu^*(X) = \|\mu\|$.

Приведем два примера не столь элементарных свойств.

1°) Пусть $\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2, \dots$ — возрастающая последовательность открытых множеств X . Их объединение \mathcal{O} является открытым множеством, и имеет место формула

$$\mu^*(\mathcal{O}) = \mu^*\left(\bigcup_n \mathcal{O}_n\right) = \lim_{n \rightarrow \infty} \mu^*(\mathcal{O}_n). \quad (\text{IV}, 3; 1)$$

Докажем эту формулу. Прежде всего, имеет место очевидное неравенство $\mu^*(\mathcal{O}) \geq \mu^*(\mathcal{O}_n)$.

Пусть теперь M — такое произвольное конечное число, что $M < \mu^*(\mathcal{O})$. Нам достаточно доказать, что существует такое целое число n , что $\mu^*(\mathcal{O}_n) \geq M$. Согласно определению $\mu^*(\mathcal{O})$, существует такая функция $\varphi \in \mathcal{C}(X)$ с носителем в \mathcal{O} , что $0 \leq \varphi \leq 1$ и $\mu(\varphi) \geq M$. Открытые множества \mathcal{O}_n пересекают компактный носитель K функции φ по открытым множествам этого компакта. Тем самым определяется возрастающая последовательность открытых множеств этого компакта. Поскольку объединение \mathcal{O} множеств \mathcal{O}_n покрывает K , то существует такое целое число n , при котором \mathcal{O}_n покрывает компакт K . Функция φ принадлежит $\mathcal{C}(X)$, имеет носитель в \mathcal{O}_n и удовлетворяет неравенству $0 \leq \varphi \leq 1$, так что, по определению меры $\mu^*(\mathcal{O}_n)$, имеет место соотношение $\mu(\varphi) \leq \mu^*(\mathcal{O}_n)$. Отсюда получаем неравенство $\mu^*(\mathcal{O}_n) \geq M$, доказывающее наше утверждение.

2°) Пусть \mathcal{O} является объединением последовательности открытых множеств $\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2, \dots$. Тогда

$$\mu^*(\mathcal{O}) \leq \sum_{n=0}^{\infty} \mu^*(\mathcal{O}_n). \quad (\text{IV}, 3; 2)$$

В самом деле, пусть $\varphi \in \mathcal{C}(X)$ имеет носитель K в \mathcal{O} и $0 \leq \varphi \leq 1$. Множества \mathcal{O}_n образуют открытое покрытие K . Поскольку K является компактом, то достаточно конечного числа этих множеств $(\mathcal{O}_i)_{i \in I}$ для покрытия K . Пусть $(a_i)_{i \in I}$ является разложением единицы, подчиненным покрытию $\{\mathcal{O}_i\}_{i \in I}$ компакта K . Тогда $a_i \varphi$ принадлежит $\mathcal{C}(X)$, имеет носитель в \mathcal{O}_i и $0 \leq a_i \varphi \leq 1$, а, значит, $\mu(a_i \varphi) \leq \mu^*(\mathcal{O}_i)$. Отсюда

$$\mu(\varphi) = \sum_{i \in I} \mu(a_i \varphi) \leq \sum_{i \in I} \mu^*(\mathcal{O}_i) \leq \sum_{n=0}^{\infty} \mu^*(\mathcal{O}_n). \quad (\text{IV}, 3; 3)$$

Переходя в левой части к точной верхней грани по всем рассматриваемым функциям φ , получаем соотношение (IV, 3; 2).

Замечания. 1°) Предположим, что μ является атомической мерой $\sum_v c_v \delta_{(a_v)}$, где все $c_v \geq 0$. Из теоремы 15 вытекает, что

$$\mu^*(\mathcal{O}) = \|\mu\|_{\mathcal{O}} = \sum_{a_v \in \mathcal{O}} c_v \leq +\infty. \quad (\text{IV}, 3; 4)$$

2°) Пусть $X = \mathbb{R}$ и μ является мерой $p dx$, где $p \geq 0$ — локально интегрируемая по Риману функция. Если, кроме того, \mathcal{O} является открытым ограниченным интервалом $[a, b]$, то

$$\mu^*([a, b]) = \int_{[a, b]} p(x) dx. \quad (\text{IV}, 3; 5)$$

В самом деле, неравенство $\mu^*([a, b]) \leq \int_{[a, b]} p(x) dx$ очевидно. Однако легко найти a' и b' , $a < a' < b' < b$, такие, что

$$\int_{[a, a']} p(x) dx + \int_{[b', b]} p(x) dx \leq ((a' - a) + (b' - b)) \sup_{a \leq x \leq b} |p(x)| \leq \varepsilon.$$

Если теперь взять функцию φ , равную 1 в $[a', b']$, 0 в $C[a'', b'']$, где $a'' = (a + a')/2$, $b'' = (b + b')/2$, и аффинную там, где она еще не определена, то получим:

$$\mu(\varphi) \geq \int_{[a', b']} p(x) dx \geq \int_{[a, b]} p(x) dx - \varepsilon.$$

Искомый результат получается в силу произвольности ε .

3°) Если мера μ равна нулю в открытом множестве \mathcal{O} (см. определение на стр. 465), то $\mu^*(\mathcal{O}) = 0$, и обратно. Мера дополнения к носителю меры μ равна нулю.

4°) Если замыкание открытого множества \mathcal{O} компактно, то мера $\mu^*(\mathcal{O})$ конечна. В самом деле, найдем, согласно следствию 1 теоремы 11, функцию $\theta \in \mathcal{C}(X)$, ≥ 0 и $\equiv 1$ на $\bar{\mathcal{O}}$. Тогда для всех функций φ , служащих для определения $\mu^*(\mathcal{O})$, необходимо имеет место неравенство $\mu(\varphi) \leq \mu(\theta)$ и, следовательно, $\mu^*(\mathcal{O}) \leq \mu(\theta)$.

Внутренняя мера компакта

Пусть K — некоторый компакт из X . Внутренней мерой компакта K относительно меры μ называется точная нижняя грань $\mu(\varphi)$ по всевозможным функциям $\varphi \geq 0$ из $\mathcal{C}(X)$ и ≥ 1 на окрестностях компакта K . Эта мера обозначается через $\mu_*(K)$. Она всегда конечна, поскольку, согласно следствию 1 теоремы 11, существует хотя бы одна функция φ_0 , обладающая указанными

свойствами и, следовательно, искомая точная нижняя грань $\leq \mu(\varphi_0)$.

Предположим, что K является некоторым компактом, содержащимся в открытом множестве \mathcal{O} . Покажем, что тогда $\mu_*(K) \leq \mu^*(\mathcal{O})$.

По следствию 1 теоремы 11, существует функция $\varphi \in \mathcal{C}(X)$, удовлетворяющая неравенствам $0 \leq \varphi \leq 1$, равная 1 в некоторой окрестности K и имеющая носитель в \mathcal{O} . Теперь, по определению $\mu_*(K)$ и $\mu^*(\mathcal{O})$, имеют место неравенства $\mu_*(K) \leq \mu(\varphi)$ и $\mu(\varphi) \leq \mu^*(\mathcal{O})$, из которых вытекает наше утверждение.

Примеры. 1°) Если μ является атомической мерой $\sum_v c_v \delta_{(a_v)}$, где $c_v \geq 0$, то

$$\mu_*(K) = \sum_{a_v \in K} c_v. \quad (\text{IV}, 3; 5_2)$$

В самом деле, для любой функции $\varphi \geq 0$, такой, что $\varphi \geq 1$ в некоторой окрестности K , имеет место неравенство $\sum_v c_v \varphi(a_v) \geq \sum_v c_v$, а значит, и $\mu_*(K) = \sum_{a_v \in K} c_v$. С другой стороны, пусть \mathcal{O} — некоторая окрестность K с компактным замыканием (см. примечание на стр. 456). Тогда $\sum_{a_v \in \mathcal{O}} c_v < +\infty$. Следовательно, существует такое конечное множество J , что $\sum_{\substack{v \notin J \\ a_v \in \mathcal{O}}} c_v \leq \varepsilon$.

Обозначим через \mathcal{O}' открытое множество, полученное из \mathcal{O} удалением точек a_v , $v \in J$, $a_v \notin K$. Тогда

$$\mu_*(K) \leq \mu^*(\mathcal{O}') = \sum_{a_v \in \mathcal{O}'} c_v \leq \sum_{a_v \in K} c_v + \varepsilon. \quad (\text{IV}, 3; 5_3)$$

Поскольку ε произвольно, $\mu_*(K) \leq \sum_{a_v \in K} c_v$, а, значит, $\mu_*(K) = \sum_{a_v \in K} c_v$.

2°) Пусть μ является мерой $p dx$ на \mathbb{R} , где $p \geq 0$, и пусть K — ограниченный интервал $[a, b]$. Тогда

$$\mu_*([a, b]) = \int_{[a, b]} p(x) dx. \quad (\text{IV}, 3; 5_4)$$

Рассуждения остаются теми же, что и в примере 1°). Прежде всего непосредственно видно, что $\mu_*([a, b]) \geq \int_{[a, b]} p(x) dx$. Выберем затем такой открытый интервал $[a', b']$, $a' < a < b < b'$,

чтобы $\int_{[a, a']} p(x) dx + \int_{[b, b']} p(x) dx \leq \varepsilon$. Тогда

$$\mu_*([a, b]) \leq \mu^*([a', b']) = \int_{[a', b']} p(x) dx \leq \int_{[a, b]} p(x) dx + \varepsilon,$$

откуда и вытекает требуемый результат.

Измеримые множества. Мера множеств

Определение. Внешней мерой подмножества A множества X относительно μ называется точная нижняя грань внешних мер открытых множеств, содержащих A . Эта мера обозначается через $\mu^*(A)$.

Внутренней мерой подмножества A относительно μ называется точная верхняя грань внутренних мер компактов, содержащихся в A . Внутренняя мера обозначается через $\mu_*(A)$. Поскольку для компакта K , содержащегося в открытом множестве \mathcal{O} , справедливо неравенство $\mu_*(K) \leq \mu^*(\mathcal{O})$, то $\mu_*(A) \leq \mu^*(A)$.

Если $A \subset B$, то $\mu^*(A) \leq \mu^*(B)$ и $\mu_*(A) \leq \mu_*(B)$.

Если A является объединением конечного или счетного семейства множества A_n , $n = 0, 1, 2, \dots$, то имеет место неравенство

$$\mu^*(A) \leq \sum_{n \geq 0} \mu^*(A_n). \quad (\text{IV}, 3; 6)$$

Это неравенство очевидно, если одно из чисел $\mu^*(A_n)$ равно $+\infty$. Предположим поэтому, что все они конечны.

Для каждого n можно найти такое открытое множество $\mathcal{O}_n \supset A_n$, что $\mu^*(\mathcal{O}_n) \leq \mu^*(A_n) + \frac{\varepsilon}{2^{n+1}}$. Если $\mathcal{O} = \bigcup_n \mathcal{O}_n$, то получаем

$$\mu^*(A) \leq \mu^*(\mathcal{O}) \leq \sum_n \mu^*(\mathcal{O}_n) \leq \sum_n \mu^*(A_n) + \varepsilon^1. \quad (\text{IV}, 3; 6_2)$$

Поскольку ε произвольно, мы получаем отсюда (IV, 3; 6).

Предположим теперь, что множества A_n не пересекаются. В этом случае неравенство (IV, 3; 6) *вовсе не обязано перейти в равенство* (оно станет равенством, если A_n измеримы; см. формулу (IV, 3; 7₃)). [Подчеркнем, однако, что неравенство перейдет в равенство тогда, когда не только A_n , но и \bar{A}_n не будут пересекаться. Мы воспользуемся этим только после теоремы 19, а докажем с помощью этой теоремы].

Для заданного N можно найти открытые непересекающиеся множества $\mathcal{O}_n \supset \bar{A}_n$, $n = 0, 1, 2, \dots, N$ (следствие 6 теоремы 11). Для каждого открытого множества $\mathcal{O} \supset A$ имеет место включе-

¹⁾ Согласно (IV, 3; 2).

ние $\mathcal{O} \supset \bigcup_{n=0}^N (\mathcal{O} \cap \mathcal{O}_n)$. Так как множества $\mathcal{O} \cap \mathcal{O}_n$ открыты, а значит, измеримы и не пересекаются, то по теореме 19 получаем

$$\mu(\mathcal{O}) \geq \sum_{n=0}^N \mu(\mathcal{O} \cap \mathcal{O}_n) \geq \sum_{n=0}^N \mu^*(A_n).$$

Переходя к точной нижней грани по всем \mathcal{O} , получаем неравенство $\mu^*(A) \geq \sum_{n=0}^N \mu^*(A_n)$, откуда, устремляя N к $+\infty$, получаем $\mu^*(A) \geq \sum_{n=0}^{\infty} \mu^*(A_n)$ и, следовательно, $\mu^*(A) = \sum_{n=0}^{\infty} \mu^*(A_n)$. Замыкания A_n не пересекаются.] С другой стороны, при непересекающихся A_n для внутренних мер имеют место неравенства. Ограничимся, как и ранее, случаем конечных $\mu_*(A_n)$, так как в противном случае рассматриваемое свойство очевидно. Для каждого n можно найти такой компакт $K_n \subset A_n$, при котором $\mu_*(K_n) \geq \mu_*(A_n) - \varepsilon/2^{n+2}$. Затем для заданного N можно найти непрерывные функции α_n , $n=0, 1, 2, \dots, N$, $0 \leq \alpha_n \leq 1$, принимающие значение 1 в некоторой окрестности K_n и имеющие попарно пересекающиеся носители (пользуясь следствием 6 теоремы 11, мы можем взять попарно не пересекающиеся открытые множества \mathcal{U}_i , содержащие K_i , $i=0, 1, 2, \dots, N$, и применить следствие 1 этой теоремы к каждой паре $K_i \subset \mathcal{U}_i$). Пусть теперь задана функция $\varphi \in \mathcal{C}(X)$, $\varphi \geq 0$, равная 1 в окрестности множества $K_0 \cup K_1 \cup \dots \cup K_N$ и такая, что $\mu(\varphi) \leq \mu_*(K_0 \cup K_1 \cup \dots \cup K_n) + \varepsilon/2$. Так как $\alpha_n \varphi \geq 0$ и равна 1 на окрестности множества K_n , то $\mu(\alpha_n \varphi) \geq \mu_*(K_n)$. Однако поскольку носители функций α_n не пересекаются, то $\varphi \geq \alpha_0 \varphi + \dots + \alpha_N \varphi$, а, следовательно, $\mu(\varphi) \geq \sum_{n=0}^N \mu(\alpha_n \varphi) \geq \sum_{n=0}^N \mu_*(K_n)$, что дает неравенство

$$\mu_*\left(\bigcup_{n=0}^N K_n\right) \geq \sum_{n=0}^N \mu_*(K_n) - \frac{\varepsilon}{2}.$$

Окончательно,

$$\mu_*(A) \geq \mu_*\left(\sum_{n=0}^N A_n\right) \geq \mu_*\left(\bigcup_{n=0}^N K_n\right) \geq \sum_{n=0}^N \mu_*(K_n) - \frac{\varepsilon}{2} \geq \sum_{n=0}^N \mu_*(A_n) - \varepsilon. \quad (\text{IV}, 3; 7)$$

Поскольку это неравенство справедливо при любом $\varepsilon > 0$, то $\mu_*(A) \geq \sum_{n=0}^N \mu_*(A_n)$, и поскольку последнее соотношение верно при любом N , то для попарно не пересекающихся A_n , дающих

в объединении множество A , получаем

$$\mu_*(A) \geq \sum_{n=0}^{\infty} \mu_*(A_n). \quad (\text{IV}, 3; 7_2)$$

Говорят, что подмножество A множества X псевдоизмеримо по μ , если его внутренние и внешние меры (конечные или нет) равны между собой. Их общее значение называется мерой A и обозначается через $\mu(A)$. Мы ввели название псевдоизмеримого, а не измеримого множества, поскольку, как мы увидим позднее, это определение слишком широко: измеримым следует назвать часть псевдоизмеримых множеств, для которых справедливо свойство аддитивности.

Докажем некоторые свойства так определенной меры множеств.

1°) *Открытые множества псевдоизмеримы.* В самом деле, пусть \mathcal{O} — некоторое открытое множество и M — такое конечное число, что $\mu^*(\mathcal{O}) > M$. Тогда существует такая функция $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, с носителем в \mathcal{O} , что $\mu(\varphi) \geq M$.

Пусть K — компактный носитель φ . Для каждой функции $\psi \in \mathcal{C}(X)$, ≥ 0 и ≥ 1 на некоторой окрестности K справедливо неравенство $\mu(\psi) \geq \mu(\varphi) \geq M$. Переходя к точной нижней грани по всевозможным таким ψ , получим $\mu_*(K) \geq M$. Поскольку это справедливо для любого M , то $\mu_*(\mathcal{O}) = \sup_{K \subset \mathcal{O}} \mu_*(K) = \mu^*(\mathcal{O})$,

что означает псевдоизмеримость открытого множества \mathcal{O} .

2°) *Компактные множества псевдоизмеримы.* Пусть K — произвольный компакт. Пусть φ — функция из $\mathcal{C}(X)$, ≥ 0 и ≥ 1 в окрестности K и такая, что $\mu(\varphi) \leq \mu_*(K) + \varepsilon$. Пусть \mathcal{O} — открытое множество $\supset K$, на котором $\varphi(x) \geq 1$. Так как для любой функции $\psi \in \mathcal{C}(X)$, $0 \leq \psi \leq 1$, с носителем в \mathcal{O} имеет место неравенство $\psi \leq \varphi$, то $\mu(\psi) \leq \mu(\varphi)$. Переходя к точной верхней грани по всем ψ , получим неравенства $\mu^*(\mathcal{O}) \leq \mu(\varphi) \leq \mu_*(K) + \varepsilon$. Отсюда, поскольку ε произвольно, получаем равенство $\mu^*(K) = \inf_{\mathcal{O} \supset K} \mu^*(\mathcal{O}) = \mu_*(K)$, означающее псевдоизмеримость компакта K .

3°) Пусть A является объединением конечного или счетного числа псевдоизмеримых непересекающихся множеств A_n . Из неравенств (IV, 3; 6) и (IV, 3; 7₂) непосредственно следует, что A псевдоизмеримо и что

$$\mu(A) = \sum_{n \geq 0} \mu(A_n). \quad (\text{IV}, 3; 7_3)$$

4°) Пусть C и D — два непересекающихся подмножества X . Мы уже знаем, что

$$\mu^*(C \cup D) \leq \mu^*(C) + \mu^*(D), \quad (\text{IV}, 3; 7_4)$$

Пусть \mathcal{O} — открытое множество $\supset (C \cup D)$ и K — некоторый компакт $\subset C$. Тогда $\mathcal{O} - K$ является некоторым открытым множеством, содержащим D , а, значит, $\mu(\mathcal{O} - K) \geq \mu^*(D)$. С другой стороны, \mathcal{O} является объединением двух псевдоизмеримых непересекающихся подмножеств K (компакт) и $\mathcal{O} - K$ (открытое множество). Из 3°) следует, что

$$\mu(\mathcal{O}) = \mu(K) + \mu(\mathcal{O} - K) \geq \mu(K) + \mu^*(D).$$

Переходя теперь по всевозможным K при фиксированном \mathcal{O} к точной верхней грани, получаем: $\mu(\mathcal{O}) \geq \mu_*(C) + \mu^*(D)$. Точная верхняя грань по всевозможным \mathcal{O} дает замечательное неравенство:

$$\mu^*(C \cup D) \geq \mu_*(C) + \mu^*(D). \quad (\text{IV}, 3; 7_5)$$

(Здесь, очевидно, можно переставлять множества C и D .)

В частности, если C псевдоизмеримо, то из (IV, 3; 7₄) и (IV, 3; 7₅) можно получить замечательное равенство (замечательное потому, что редко удается получать равенства, в которых участвуют не псевдоизмеримые множества):

$$\mu^*(C \cup D) = \mu(C) + \mu^*(D). \quad (\text{IV}, 3; 7_6)$$

Что же касается внутренних мер, то, согласно (IV, 3; 7₂), прежде всего имеем:

$$\mu_*(C \cup D) \geq \mu_*(C) + \mu_*(D). \quad (\text{IV}, 3; 7_7)$$

Пусть K — некоторый компакт $\subset (C \cup D)$ и \mathcal{O} — открытое множество $\supset D$. Компакт $K \cap C \mathcal{O}$ содержится в C , а, следовательно, его мера $\leq \mu_*(C)$. Так как множество $K \cup \mathcal{O}$ является объединением двух псевдоизмеримых непересекающихся множеств \mathcal{O} (открытое множество) и $K \cap C \mathcal{O}$ (компакт), то оно псевдоизмеримо и из 3°) получаем

$$\mu(K) \leq \mu(K \cup \mathcal{O}) = \mu(\mathcal{O}) + \mu(K \cap C \mathcal{O}) \leq \mu(\mathcal{O}) + \mu_*(C).$$

Переходя последовательно к точной верхней грани по всем K , а затем к точной нижней грани по всевозможным \mathcal{O} , получаем:

$$\mu_*(C \cup D) \leq \mu_*(C) + \mu^*(D). \quad (\text{IV}, 3; 7_8)$$

(Здесь можно, очевидно, переставлять C и D .)

Меняя ролями C и D , затем применяя соотношения (IV, 3; 7₇) и (IV, 3; 7₈), для псевдоизмеримого C можно получить замечательное равенство:

$$\mu_*(C \cup D) = \mu(C) + \mu_*(D). \quad (\text{IV}, 3; 7_9)$$

Из соотношений (IV, 3; 7₆) и (IV, 3; 7₉) следует, что если множества C и D не пересекаются, множества C и $C \cup D$ псевдоизмеримы, а множество C имеет конечную меру, то

множество D псевдоизмеримо, поскольку

$$\mu^*(D) = \mu_*(D) = \mu(C \cup D) - \mu(C).$$

В том случае, когда $\mu(C) = \mu(C \cup D) = +\infty$, этот результат не сохраняется, так как разность при этом не определена. Впрочем, легко привести и контрпример. Позже (стр. 503) мы приведем пример не псевдоизмеримого по $\mu = dx$ множества D , содержащегося в ограниченном интервале $[a, b]$ прямой \mathbb{R} . Пусть $C = C([a, b])$. Тогда оно открыто и имеет меру, равную $+\infty$. Поскольку $(C \cup D) \supset C$, это множество псевдоизмеримо, имеет меру $+\infty$, но тем не менее множество D не является псевдоизмеримым.

Если положить $A = C \cup D$ и $B = C$, то полученный результат можно записать в другом эквивалентном виде. Если A и B псевдоизмеримы, $A \supset B$ и если B имеет конечную меру, то множество $A - B = A \cap CB$ псевдоизмеримо и $\mu(A - B) = \mu(A) - \mu(B)$.

Полученное соотношение не сохраняется, если $\mu(B) = +\infty$. Полагая $A = X$, получаем, что если B — псевдоизмеримо и имеет конечную меру, то CB также псевдоизмеримо (вообще говоря, это неверно, если $\mu(B) = +\infty$).

5°) Пусть A является объединением конечного числа псевдоизмеримых множеств: $A = \bigcup_{n=0}^N A_n$. Тогда A псевдоизмеримо (и

в силу (IV, 3; 6) имеет место неравенство $\mu(A) \leq \sum_{n=0}^N \mu(A_n)$). Следует подчеркнуть, что п. 5°) отличается от п. 3°), так как множества A_n здесь не предполагаются непересекающимися. Это свойство очевидно, если одно из множеств A_n имеет бесконечную меру, так как тогда и множество A будет иметь бесконечную меру. Предположим поэтому, что все меры $\mu(A_n)$ конечны. Пусть \mathcal{O}_n — открытое множество, содержащее A_n и такое, что $\mu(\mathcal{O}_n) \leq \mu(A_n) + \varepsilon/[2(N+1)]$, и пусть K_n — такой компакт $\subset A_n$,

что $\mu(K_n) \geq \mu(A_n) - \varepsilon/[2(N+1)]$. Тогда $\mathcal{O} = \bigcup_{n=0}^N \mathcal{O}_n$ является открытым множеством, содержащим A , и $K = \bigcup_{n=0}^N K_n$ является компактом $\subset A$. Из неравенств $\mu^*(A) - \mu_*(A) \leq \mu(\mathcal{O}) - \mu(K) = \mu(\mathcal{O} - K) \leq \sum_{n=0}^N \mu(\mathcal{O}_n - K_n)$ (ибо $(\mathcal{O} - K) \subset \bigcup_{n=0}^N (\mathcal{O}_n - K_n)$) $\leq \sum_{n=0}^N (\mu(\mathcal{O}_n) - \mu(K_n)) \leq \varepsilon$, так как ε произвольно, следует, что A псевдоизмеримо.

6°) Пусть A является объединением счетного множества псевдоизмеримых множеств A_n . Пусть $B_n = A_0 \cup A_1 \cup \dots \cup A_n$. Согласно 5°), множества B_n псевдоизмеримы, и A_n является объединением возрастающей последовательности B_n . Если B_n имеют конечные меры, то, согласно 4°), множества $C_n = B_n - B_{n-1}$ также псевдоизмеримы.

Наконец, согласно 3°), множество A , являющееся объединением непересекающихся множеств $B_0, C_1, C_2, C_3, \dots$, псевдоизмеримо. Результат сохраняется и в том случае, когда одно из множеств A_n имеет бесконечную меру, так как тогда бесконечную меру будет иметь множество A . Таким образом, мы доказали, что любое конечное или счетное объединение псевдоизмеримых подмножеств псевдоизмеримо. Если B_n является возрастающей последовательностью псевдоизмеримых множеств, объединение которых равно B (здесь $B = A$), то (согласно предыдущему, B псевдоизмеримо) имеет место равенство: $\mu(B) = \lim_{n \rightarrow \infty} \mu(B_n)$. В самом деле, согласно 3°), $\mu(B) = \mu(B_0) + \mu(C_1) + \mu(C_2) + \dots$ и, согласно 4°), $\mu(C_n) = \mu(B_n) - \mu(B_{n-1})$, если только эти величины конечны. Отсюда в случае конечных величин $\mu(B_n)$ получаем $\mu(B) = \lim_{n \rightarrow \infty} \mu(B_n)$. Для случая, когда одна из этих величин бесконечна, этот результат очевиден.

7°) Пусть A является конечным пересечением псевдоизмеримых подмножеств A_n с конечными мерами: $A = \bigcap_{n=0}^N A_n$. Тогда

A будет псевдоизмеримым.

Для доказательства будем рассуждать, как в п. 5°), с теми же обозначениями, но полагая на этот раз $\mathcal{O} = \bigcap_{n=0}^N \mathcal{O}_n$ и $K = \bigcap_{n=0}^N K_n$.

Тогда $\mu^*(A) - \mu_*(A) \leq \mu(\mathcal{O}) - \mu(K) = \mu(\mathcal{O} - K) \leq \sum_{n=0}^N \mu(\mathcal{O}_n - K_n)$
 $\left(\text{так как } \mathcal{O} - K \subset \bigcup_{n=0}^N (\mathcal{O}_n - K_n) \right) = \sum_{n=0}^N (\mu(\mathcal{O}_n) - \mu(K_n)) \leq \varepsilon$, что

в силу произвольности ε доказывает наше утверждение. Однако здесь, в противоположность п. 5°), результат не сохраняется, если мера одного из множеств A_n бесконечна. В самом деле, пусть D — не псевдоизмеримое множество, содержащееся в ограниченном интервале $[a, b]$ прямой \mathbb{R} (см. стр. 503). Так как $A = [a, b]$ псевдоизмеримо, то множество $B = D \cup C[a, b]$, содержащее $C[a, b]$, псевдоизмеримо и имеет бесконечную меру. Здесь $A \cap B = D$ не является псевдоизмеримым.

8°) Пересечение конечного или счетного числа псевдоизмеримых множеств конечной меры псевдоизмеримо.

В самом деле, пусть $A = \bigcap_{n=0}^{\infty} A_n$ и $B_n = A_0 \cap A_1 \cap \dots \cap A_n$.

Тогда, согласно 7°), каждое множество B_n является псевдоизмеримым, последовательность B_n убывающая и ее пересечением является множество A . Далее, $B_n = B_0 - (B_0 - B_n)$. Так как последовательность $B_0 - B_n$ является возрастающей, то к ней можно применить 6°), откуда вытекает, что $B_0 - (B_0 - A) = A$ псевдоизмеримо и, кроме того, $\mu(B_0 - A) = \lim_{n \rightarrow \infty} \mu(B_0 - B_n)$. Таким образом, если B_n является убывающей последовательностью псевдоизмеримых множеств конечной меры с пересечением, равным B (здесь A), то $\mu(B) = \lim_{n \rightarrow \infty} \mu(B_n)$. Полученные результаты можно было бы считать весьма удивительными, если бы не имели места следующие факты:

Если A и B псевдоизмеримы и $A \supset B$, но $\mu(B) = +\infty$, то $A - B$ может оказаться не псевдоизмеримым. Пересечение двух псевдоизмеримых множеств может оказаться не псевдоизмеримым, если одно из этих множеств имеет бесконечную меру.

Эти недостатки можно устранить, если ввести более сильное определение.

Множество A называется измеримым, если для любого компакта K пересечение $A \cap K$ псевдоизмеримо (оно необходимо имеет конечную меру). Измеримое множество псевдоизмеримо, т. е. свойство измеримости сильнее свойства псевдоизмеримости. Действительно, множество X можно представить как объединение последовательности компактов K_n . Следовательно, если $A \cap K_n$ псевдоизмеримы, то, согласно 6°), A также будет псевдоизмеримым. С другой стороны, всякое псевдоизмеримое множество конечной меры измеримо. Это следует из того, что в этом случае, согласно 7°), псевдоизмеримо множество $A \cap K$. Однако псевдоизмеримое множество бесконечной меры не обязательно измеримо. Например, если D не псевдоизмеримо и содержитя в $[a, b] \subset \mathbb{R}$, то множество $D \cup C[a, b]$ псевдоизмеримо, имеет бесконечную меру, но не измеримо, так как его пересечение с компактом $[a, b]$ не является псевдоизмеримым. Очевидно также, что если A и B измеримы, $A \supset B$, то $A - B$ всегда измеримо, поскольку $(A - B) \cap K = (A \cap K) - (B \cap K)$. Однако формула $\mu(A - B) = \mu(A) - \mu(B)$ теряет смысл при $\mu(B) = +\infty$.

С другой стороны, объединение или пересечение A конечного или счетного множества измеримых множеств A_n измеримо, поскольку в этом случае $A \cap K$ является объединением или пересечением множеств $A_n \cap K$. Впрочем, дополнение C_A измеримого множества A измеримо и от объединения к пересечению можно перейти через дополнение. Напротив, формула $\mu(B) = \lim_{n \rightarrow \infty} \mu(B_n)$ для убывающей последовательности измеримых

множеств B_n , пересечение которых равно B , всегда предполагает, что B_n имеют конечные меры (это видно из доказательства этой формулы). Если, например, мы возьмем меру dx на \mathbb{R} , то последовательность множеств $B_n = \{x \in \mathbb{R}; |x| \geq n\}$ является убывающей, каждое из множеств имеет бесконечную меру, их пересечение пусто, а, значит, имеет меру 0.

Добавим еще, что каждое замкнутое множество F измеримо, поскольку в том случае, когда X является объединением компактов K_n , это множество является объединением компактов $F \cap K_n$. Переходом к дополнению отсюда легко получить, что каждое открытое множество измеримо даже в том случае, когда оно имеет бесконечную меру.

Повторим определения, введенные в процессе изложения материала.

Определение. Множество $A \subset X$ псевдоизмеримо, если $\mu^*(A) = \mu_*(A)$. Это общее значение называется мерой множества и обозначается через $\mu(A)$. Множество A называется измеримым, если оно псевдоизмеримо и если, кроме того, для каждого компакта K множество $A \cap K$ псевдоизмеримо.

Нами была доказана следующая теорема:

Теорема 19. 1°) Все открытые и замкнутые множества измеримы.

2°) Если A и B — два измеримых подмножества X и $A \supset B$, то множество $A - B$ измеримо и $\mu(A - B) = \mu(A) - \mu(B)$, если только $\mu(A)$ и $\mu(B)$ не равны $+\infty$. В частности, если A измеримо, то $C(A)$ также измеримо.

3°) Если A является объединением конечного или счетного семейства измеримых множеств, $A = \bigcup_n A_n$, то A измеримо и, кроме того, имеет место неравенство

$$\mu(A) \leq \sum_n \mu(A_n), \quad (\text{IV}, 3; 8)$$

переходящее в равенство, если множества A_n попарно не пересекаются.

3°₂) Пересечение конечного или счетного семейства измеримых множеств измеримо.

4°) Пусть A_0, A_1, A_2, \dots — возрастающая последовательность измеримых множеств. Тогда их объединение A измеримо и, кроме того, имеет место формула

$$\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n). \quad (\text{IV}, 3; 9)$$

5°) Если A_0, A_1, A_2, \dots — убывающая последовательность измеримых множеств, то их пересечение A измеримо. Кроме того, если все множества A_n имеют конечную меру, то имеет место

формула (IV, 3; 9). Эта формула не верна, если все $\mu(A_n)$ бесконечны.

Замечание 1. Позже мы получим новую формулу для внешних мер: если A_n является *возрастающей* последовательностью произвольных подмножеств X , объединение которых равно A , то $\mu^*(A) = \lim_{n \rightarrow \infty} \mu^*(A_n)$ (см. следствие 0 теоремы 36). Однако сейчас

надо будет доказать это равенство в частном случае, который нам вскоре понадобится: для любого $A \subset X$ $\mu^*(A) = \sup \mu^*(A \cap K)$, где \sup берется по всевозможным компактам K множества X . Это действительно частный случай высказанного выше общего утверждения, поскольку X является объединением возрастающей последовательности компактов K_n . Предположим прежде всего, что $\mu^*(A) < +\infty$. Тогда существует открытое множество $\mathcal{O} \supset A$, имеющее конечную меру. Пересечения $\mathcal{O} \cap CK_n$ образуют убывающую последовательность непересекающихся измеримых множеств, имеющих конечные меры. Согласно п. 5°) теоремы 19, при n , стремящемся к бесконечности, $\mu(\mathcal{O} \cap CK_n)$ стремится к 0. Так как $(A \cap CK_n) \subset (\mathcal{O} \cap CK_n)$, то $\mu^*(A \cap CK_n)$ также стремится к 0. Поскольку $A = (A \cap K_n) \cup (A \cap \mathcal{O} \cap CK_n)$, то $\mu^*(A) \leq \mu^*(A \cap K_n) + \mu^*(A \cap \mathcal{O} \cap CK_n)$, а, следовательно, $\mu^*(A \cap K_n) \leq \mu^*(A)$ стремится к $\mu^*(A)$ при n , стремящемся к бесконечности.

Предположим теперь, что $\mu^*(A) = +\infty$. Допустим для простоты, что X метризуемо и что рассматриваются замкнутые компактные шары (в противном случае можно провести рассуждение, аналогичное проведенному на стр. 457).

Положим $A_n = A \cap \{x \in X; n \leq d(x, a) \leq n+1\}$. Положим затем $B_0 = A_0 \cup A_2 \cup A_4 \cup \dots$ и $B_1 = A_1 \cup A_3 \cup A_5 \cup \dots$. Тогда $A = B_0 \cup B_1$; поэтому $\mu^*(A) \leq \mu^*(B_0) + \mu^*(B_1)$, а, значит, хотя бы одно из слагаемых, например $\mu^*(B_0)$, равно $+\infty$. Так как B_0 является объединением непересекающихся замкнутых множеств, то (см. стр. 494) $+\infty = \mu^*(B_0) = \mu^*(A_0) + \mu^*(A_2) + \mu^*(A_4) + \dots$. Полученный ряд оказался расходящимся. Но тогда

$$\begin{aligned} \mu^*(A \cap \{x \in X; d(x, a) \leq 2N+1\}) &\geq \mu^*(A_0 \cup A_2 \cup \dots \cup A_{2N}) = \\ &= \mu^*(A_0) + \mu^*(A_2) + \dots + \mu^*(A_{2N}) \end{aligned}$$

стремится к $+\infty = \mu^*(A)$ при N , стремящемся к бесконечности.

Замечание 2. Не может быть и речи о распространении теоремы 19 на объединение или пересечение *несчетных* семейств подмножеств. Например, по мере dx любое множество $A \subset \mathbb{R}$ (измеримое или нет, конечной меры или нет) является объединением своих точек, которые представляют собой измеримые множества нулевой меры.

Различные примеры. 1°) Пусть μ — атомическая мера $\sum_v c_v \delta_{(a_v)}$, где все $c_v \geq 0$. Из замечаний на стр. 492—493 следует, что все подмножества A множества X измеримы и имеет место формула

$$\mu(A) = \sum_{a_v \in A} c_v \leq +\infty. \quad (\text{IV}, 3; 10)$$

2°) Для меры $p dx$ на \mathbb{R} всегда существуют неизмеримые подмножества, кроме того случая, когда мера $p dx$ равна нулю.

Приведем пример неизмеримого множества. Пусть X представляет собой тригонометрическую окружность. На ней можно определить естественную меру Радона $\mu = d\theta$, называемую угловой мерой. Эту меру можно определить, например, следующим образом.

Поскольку каждая точка X определяется некоторым вещественным числом с точностью до 2π , то пространство $\mathcal{C}(X)$ непрерывных функций на X (так как X компактно, то они имеют компактный носитель) можно отождествить с пространством непрерывных периодических функций на вещественной прямой с периодом 2π . Обозначим через $\mathcal{P}_{2\pi}(\mathbb{R})$ пространство этих функций, нормированное по формуле $\|\varphi\| = \max_{x \in \mathbb{R}} |\varphi(x)|$.

Определять меру на X или определять линейную непрерывную форму на $\mathcal{P}_{2\pi}(\mathbb{R})$ — это, в сущности, равносильно. Поэтому мера $\mu = d\theta$ может быть задана формулой

$$\mu(\varphi) = \int_{[a, a+2\pi]} \varphi(x) dx, \quad \varphi \in \mathcal{P}_{2\pi}(\mathbb{R}), \quad (\text{IV}, 3; 11)$$

где точку a можно выбрать произвольно. Результат от выбора a не зависит, поскольку все функции φ периодичны с периодом 2π ¹⁾. Эта мера инвариантна относительно вращения (с центром в начале координат) множества X . Это означает, что при любом α имеет место формула

$$\mu(\varphi(x)) = \mu(\varphi(x - \alpha)), \quad \varphi \in \mathcal{P}_{2\pi}(\mathbb{R}). \quad (\text{IV}, 3; 12)$$

Отсюда также естественно вытекает, что μ -мера множеств X инвариантна относительно вращения. Если A является некоторым подмножеством X и если A_α есть множество, полученное из A вращением на угол α , то справедлива формула $\mu^*(A_\alpha) = \mu^*(A)$, $\mu_*(A_\alpha) = \mu_*(A)$. Кроме того, A_α измеримо тогда и

¹⁾ Мы привели эту меру к известному понятию интеграла Римана на \mathbb{R} . Однако не представляет никакого труда построить теорию интеграла Римана $\int_X \varphi(\theta) d\theta$ на тригонометрической окружности X .

только тогда, когда измеримо множество A , и при этом $\mu(A_\alpha) = \mu(A)$.

Обозначим теперь через α число, неизмеримое с π и выбранное раз и навсегда. Будем говорить, что точки x, y множества X конгруэнты, если $y - x$ целое кратное α . Тем самым в множестве X мы определим некоторое отношение эквивалентности. Пусть \dot{X} — фактормножество. Следует ожидать, что фактормножество \dot{X} , как и само множество X , имеет мощность континуума. В самом деле, с одной стороны, существует сюръекция X на \dot{X} , а, значит, $\text{card } X \geq \text{card } \dot{X}$. С другой стороны, каждый класс эквивалентности счетен и потому $\text{card } X \leq v \text{ card } \dot{X}$, и, поскольку $\text{card } \dot{X} \geq v$ (противоположное означает, что существует лишь конечное число классов; так как каждый из них счетен, то и X будет счетным, что невозможно), то, согласно теореме 5 гл. I, $\text{card } X = \text{card } \dot{X} = v$ — мощность континуума. Пусть теперь A является подмножеством X , содержащим по точке из каждого класса эквивалентности. (Очевидно, в классе эквивалентности ни одна из точек не выделяется среди других и выбор точки не регламентируется никаким правилом. Мы вынуждены для этого континуального числа выборов пользоваться аксиомой выбора Цермело (см. примечание на стр. 28).)

Докажем теперь, что множество A не является μ -измеримым. В самом деле, так как множество A содержит не более одной точки из каждого класса эквивалентности, то множества $A_{n\alpha}, n \in \mathbb{Z}$, попарно не пересекаются. С другой стороны, поскольку A содержит не менее одной точки из каждого класса эквивалентности, то объединение множеств $A_{n\alpha}$ дает все множество X . Впрочем, A и $A_{n\alpha}$ могут быть преобразованы одно в другое с помощью вращения $n\alpha$, и, следовательно, их внешние и внутренние меры совпадают. Если A измеримо, то множества $A_{n\alpha}$ измеримы и имеют ту же самую меру; при этом выполняется равенство

$$2\pi = \mu(X) = \sum_{n=-\infty}^{+\infty} \mu(A_{n\alpha}) = \mu(A) \sum_{n=-\infty}^{+\infty} 1.$$

Однако это равенство невозможно. В самом деле, если A имеет нулевую меру, то отсюда следует, что X также будет иметь нулевую меру, в то время как его мера равна 2π . Если мера A больше нуля, то X имеет бесконечную меру.

Итак, A неизмеримо. Кроме того, из неравенства (IV, 3; 7₂) следует, что

$$2\pi = \mu_*(X) \geq \sum_{n=-\infty}^{+\infty} \mu_*(A_{n\alpha}) = \mu_*(A) \sum_{n=-\infty}^{+\infty} 1,$$

т. е. что нижняя мера множества A равна нулю. Однако это множество неизмеримо, и, следовательно, его внешняя мера > 0 . Приведенный здесь пример кажется неестественным. Можно отметить общий факт относительно всех ненулевых мер вида pdx на прямой \mathbb{R} или $p\theta$ на тригонометрической окружности (p — интегрируемая по Риману функция): *неизвестно, как явно образовать неизмеримые множества. Можно лишь с помощью аксиомы выбора доказать их существование.* Это означает, что все множества, которые встречаются в приложениях и которые обычно всегда определены явно, измеримы. Открытый, полуоткрытый или замкнутый интервал $[a, b]$ измерим, и его мера равна $\int p(x) dx$.

Z 3°) Теорема 19 позволяет указать очень широкую категорию измеримых множеств.

Назовем σ -алгеброй частей множества X совокупность \mathcal{T} его частей, обладающую следующими свойствами:

a) если $A \in \mathcal{T}$, то $CA \in \mathcal{T}$;

b) любое объединение или пересечение конечного или счетного множества частей, принадлежащих \mathcal{T} , принадлежит \mathcal{T}^1 .

Вот некоторые примеры σ -алгебр: прежде всего это пустая σ -алгебра, не содержащая ни одного множества; затем σ -алгебра, состоящая лишь из некоторого подмножества и его дополнения, и, наконец, σ -алгебра всех частей множества X .

Если \mathcal{T}_1 и \mathcal{T}_2 — две σ -алгебры, то части X , принадлежащие одновременно обеим σ -алгебрам, образуют некоторую σ -алгебру, которая называется пересечением σ -алгебр $\mathcal{T}_1 \cap \mathcal{T}_2$. Точно также определяется пересечение $\bigcap_{i \in I} \mathcal{T}_i$ конечного или бесконечного

семейства σ -алгебр. Отсюда, в частности, вытекает, что если \mathcal{S} является произвольным множеством частей X , то существует наименьшая σ -алгебра, содержащая это множество, а именно — пересечение всех содержащих его σ -алгебр. Пусть теперь X — топологическое пространство. Тогда борелевской называется наименьшая σ -алгебра, содержащая все открытые в смысле топологии X части (и, следовательно, содержащая также все замкнутые части). Часть множества X называется борелевской, если она принадлежит борелевской σ -алгебре. Борелевская σ -алгебра из X зависит только от топологии X . Например, пересечение счетного множества открытых частей (которое не обязательно открыто) и объединение счетного множества замкнутых частей (которое не обязательно замкнуто) являются борелевскими. Имеются и другие примеры. Борелевская часть «универсально

¹⁾ Автор вместо термина σ -алгебры использует термин «tribu» (племя). — Прим. ред.

измерима» в том смысле, что она измерима по любой из мер Радона μ на X .

В самом деле, по теореме 19 множество всех μ -измеримых частей X является σ -алгеброй, зависящей, естественно, от μ . Это множество заведомо содержит открытые подмножества и, следовательно, содержит борелевскую σ -алгебру. Весьма тонкие теоремы позволяют показать, что существуют и другие универсально измеримые части X , отличные от борелевских.

Множества нулевой меры

Теорема 20. Для того чтобы часть A множества X имела нулевую меру относительно μ , необходимо и достаточно, чтобы, каково бы ни было $\varepsilon > 0$, существовало такое открытое множество \mathcal{O} из X , содержащее A , что $\mu(\mathcal{O}) \leq \varepsilon$. Любое объединение конечной или счетной совокупности множеств меры нуль имеет нулевую меру.

Доказательство. Первое условие означает, что внешняя мера $\mu^*(A)$ равна нулю. Поскольку тогда меньшая внутренняя мера неизбежно также равна нулю, рассматриваемое множество A измеримо и имеет нулевую меру. То, что касается конечной или счетной совокупности множеств, вытекает из теоремы 19 и из того факта, что сумма ряда с нулевыми членами равна нулю.

Следствие. Всякое счетное множество на вещественной прямой \mathbb{R} по мере dx имеет нулевую меру.

Следствие вытекает из теоремы 20 и из того факта, что множество, состоящее из одной точки, имеет нулевую меру.

Замечания. 1°) Мера всей вещественной прямой \mathbb{R} или отрезка с несовпадающими концами не равна нулю по мере dx . Отсюда вытекает новое доказательство того факта, что прямая или отрезок не являются счетными множествами. Эти примеры еще раз показывают, что объединение бесконечной несчетной совокупности множеств нулевой меры, в частности совокупность точек, не обязательно должно иметь меру, равную нулю.

2°) Глядя на следствие, можно было бы подумать, что на вещественной прямой \mathbb{R} не существует других множеств нулевой меры по мере dx , кроме счетных множеств. Однако это неверно. Мы сейчас построим на интервале $[0, 1]$ множество, называемое совершенным множеством Кантора, имеющее мощность континуума и нулевую меру. Рассмотрим сначала множество E_1 , являющееся объединением интервалов $[0, \frac{1}{3}]$ и $[\frac{2}{3}, 1]$. Разделим каждый из этих интервалов на три равные части, оставим из них лишь первую и третью части и обозначим образован-

ное таким образом множество через E_2 . Множество E_2 является объединением 4 интервалов $\left[0, \frac{1}{9}\right]$, $\left[\frac{2}{9}, \frac{3}{9}\right]$, $\left[\frac{6}{9}, \frac{7}{9}\right]$, $\left[\frac{8}{9}, 1\right]$. Разделим снова каждый из этих 4 интервалов, составляющих E_2 , на 3 равных интервала, выбросим средний из них и полученное множество, являющееся объединением 8 интервалов, обозначим через E_3 . Продолжая так далее, мы получим убывающую последовательность замкнутых множеств $E_1, E_2, \dots, E_n, \dots$. Мера множества E_n равна, очевидно, $\left(\frac{2}{3}\right)^n$, а, значит, пересечение E всех множеств E_n имеет нулевую меру. Докажем, однако, что множество E имеет мощность континуума. В самом деле, любая точка E может быть определена единственным образом с помощью произвольной последовательности $a_1, a_2, \dots, a_n, \dots$, где a_i является одним из двух слов: «первый» или «второй».

Пусть $x \in E$. Слово a_1 будет первым или вторым в зависимости от того, находится ли x в первом интервале $\left[0, \frac{1}{3}\right]$ или во втором $\left[\frac{2}{3}, 1\right]$, объединение которых составляет множество E_1 . Пусть, например, $a_1 = \text{«первый»}$. Слово a_2 будет первым или вторым в зависимости от того, находится ли x в первом подинтервале $\left[0, \frac{1}{9}\right]$ или же во втором подинтервале $\left[\frac{2}{9}, \frac{3}{9}\right]$ разбиения интервала $\left[0, \frac{1}{3}\right]$, входящего в множество E_2 . Так можно продолжать неограниченно. Положим теперь $b_n = 0$, если $a_n = \text{«первый»}$ и $b_n = 2$, если $a_n = \text{«второй»}$. Непосредственно видно, что соответствие между x и последовательностью $b_n = 0$ или 2 полностью определяется равенством $x = \sum_{n=1}^{\infty} \frac{b_n}{3^n}$ и что эта формула определяет некоторую биекцию между E и множеством $\{0, 2\}^{\mathbb{N}}$ отображений \mathbb{N} в $\{0, 2\}$ или множеством последовательностей элементов 0 или 2. Из результатов гл. I следует, что $\text{card } E = \text{card } \mathfrak{P}(\mathbb{N}) = 2^{\mathbb{N}} = \gamma$, т. е. мощность континуума.

Впрочем, множество E можно определить, воспользовавшись представлением точек интервала $[0, 1]$ не в десятичной, а в троичной системе счисления, т. е. системе счисления по основанию 3. Множество E тогда будет множеством точек, в троичном разложении которых применяются только числа 0 и 2 и не используется число 1 (элемент, имеющий два троичных разложения принадлежит E тогда и только тогда, когда одно из этих разложений может быть записано только с помощью чисел 0 и 2). Множество E , описанное впервые в прошлом столетии Кантором, обладает, кроме того, следующими свойствами: оно *совершенно*. Это означает, что оно замкнуто в \mathbb{R} и что ни одна из его точек

не является изолированной. Любая окрестность каждой из его точек содержит даже, как это легко доказать, бесконечное множество точек E мощности континуум.

Приведем еще один поучительный пример. Пусть a_0, a_1, a_2, \dots — плотная последовательность точек прямой \mathbb{R} (например, множество рациональных чисел, упорядоченных в последовательность). Пусть c_0, c_1, c_2, \dots — такая последовательность положительных чисел, что $\sum_{n=0}^{\infty} c_n < +\infty$. Обозначим через E_ε множество

$E_\varepsilon = \bigcup_{n=0}^{\infty} \{x \in \mathbb{R}; |x - a_n| < \varepsilon c_n\}$. Это множество открыто (как объединение открытых множеств), и его мера естественным образом мажорируется числом $2\varepsilon \sum_{n=0}^{\infty} c_n$. Обозначим теперь

через E пересечение множеств E_ε , соответствующих всевозможным $\varepsilon > 0$. Мера этого множества равна нулю. Можно было бы подумать, что оно счетно и сводится к последовательности точек a_n . Однако мы рекомендуем удивленному читателю самому доказать, что в действительности это множество имеет мощность континуума.

Если в качестве меры на \mathbb{R}^n взять меру dx_1, dx_2, \dots, dx_n , то мера множества будет соответствовать его объему, и, следовательно, как мы увидим позже, дифференцируемое многообразие размерности $< n$ будет необходимо иметь нулевую меру. Здесь мы получаем очень простые примеры множеств нулевой меры, имеющих мощность континуума.

3°) Множество $C\{a\}$ по мере Дирака $\delta_{(a)}$ имеет нулевую меру.

4°) По мере, тождественно равной 0, любое множество имеет меру, равную нулю.

Свойства, выполняющиеся почти всюду

Пусть P — свойство, относящееся к точкам локально компактного пространства X , снабженного мерой Радона $\mu \geqslant 0$ ¹⁾. Это свойство определяется заданием множества A точек X , обладающих данным свойством.

Говорят, что это свойство выполняется μ -почти всюду, или μ -почти все точки X обладают этим свойством, если дополнение $C\bar{A}$ множества A имеет нулевую меру относительно меры μ . Можно было бы ограничиться словами «почти всюду», если бы

¹⁾ Свойство P относительно точек X можно определить как отображение множества X в двухэлементное множество {да, нет}. Говорят, что $x \in X$ обладает свойством P , если $P(x) = \text{да}$.

мера μ была задана раз навсегда и если бы это не приводило к недоразумениям. Заметим, однако, что при мере $\mu = dx$ понятие «почти всюду» вполне соответствует нашей интуиции. Это не так для других мер. Например, для меры $\mu = \delta_{(a)}$ — единичная масса в точке a — некоторое свойство выполняется почти всюду, если оно выполняется в точке a . Для нулевой меры μ любое свойство выполняется почти всюду (так же точно, впрочем, как и отрицание этого свойства). Вот что означает для функции f , определенной на метрическом пространстве X со значениями в метрическом пространстве F , свойство быть *почти всюду непрерывной*:

$$(\exists A \in \mathfrak{P}(X), \mu(CA) = 0) (\forall a \in A) (\forall \varepsilon > 0)$$

$$(\exists \eta > 0) (\forall x \in X, d(a, x) \leq \eta) : (d(f(a), f(x)) \leq \varepsilon),$$

или иначе: множество точек, в которых f разрывна, имеет относительно μ меру, равную нулю. Если \vec{F} является векторным пространством, то утверждение «некоторое отображение \vec{f} пространства X в \vec{F} почти всюду равно нулю» означает, что множество точек $x \in X$, в которых $\vec{f}(x) \neq \vec{0}$, имеет меру, равную нулю.

На вещественной прямой по мере $\mu = dx$ почти все точки являются иррациональными и даже трансцендентными.

Теорема 21. Пусть P_0, P_1, P_2, \dots — конечное или счетное множество свойств, относящихся к точкам X . Если каждое из этих свойств выполняется почти всюду, то свойство $P = P_0 \wedge P_1 \wedge P_2 \wedge P_3 \wedge \dots$, состоящее в том, что одновременно должны выполняться все свойства P_n , само выполняется почти всюду.

Это фундаментальная теорема теории вероятностей.

Доказательство. Пусть A_n — множество точек, обладающих свойством P_n . Тогда $A = \bigcap_n A_n$ будет множеством точек, обладающих свойством P . Его дополнение $\bigcup_n (CA_n)$ имеет нулевую меру как объединение счетного множества множеств нулевой меры.

Конечно, аналогичного утверждения для несчетного множества свойств не существует. Например, для меры $\mu = dx$ на \mathbb{R} свойство « x обладает свойством P_a , если $x \neq a$ » верно почти всюду. Здесь $P = \bigwedge_a P_a$ является следующим свойством: « x не совпадает ни с одной точкой \mathbb{R} ». Этим свойством не обладает ни одна точка.

μ -измеримые функции со значениями в метризуемом сепарабельном пространстве

Предварительное определение. Пусть F — метризуемое пространство. Говорят, что это пространство *сепарабельно*, если в F существует счетное плотное множество.

Пусть f — отображение локально компактного пространства X , снабженного мерой $\mu \geq 0$, в метризуемое сепарабельное пространство F . Говорят, что f *μ -измеримо* или *просто измеримо* (когда это не приводит к недоразумению), если прообраз при отображении f каждой открытой части F является μ -измеримой частью X .

В этом определении причина, из-за которой мы предположили F сепарабельным, непосредственно не видна. Можно определить измеримые функции со значениями в произвольном топологическом пространстве, но тогда надо будет дать более сложное определение, эквивалентное предыдущему в том случае, когда F метризуемо и сепарабельно¹⁾. Из этого определения вытекает, что каждое непрерывное отображение X в F измеримо, ибо прообраз каждого открытого множества из F является открытым множеством, а, следовательно, согласно п. 1°) теоремы 19, измерим.

Естественно, переходом к дополнению можно было бы в определении заменить открытые подмножества на замкнутые. Впрочем, множество частей F , прообраз которых при отображении f μ -измерим, является некоторой σ -алгеброй (см. стр. 505). Если f измеримо, то эта σ -алгебра содержит открытые, а, значит, борелевские множества, и прообраз каждого борелевского множества из F является μ -измеримым.

Пусть A — некоторая часть X и φ — ее характеристическая функция. Эта функция, рассматриваемая как отображение X в топологическое сепарабельное пространство \mathbb{R} ²⁾ или дискретное пространство $\{0, 1\}$, измерима тогда и только тогда, когда A измеримо, поскольку прообраз каждого открытого множества при этом отображении равен либо \emptyset , либо X , либо A , либо $\complement A$.

Теорема 21₂. *Пусть f и g — два отображения X в F . Если эти отображения почти всюду равны и одно из них измеримо, то другое также измеримо.*

Доказательство. Пусть A — множество точек x , в которых $f(x) = g(x)$. По условию, $\complement A$ является множеством меры

¹⁾ См. следствие теоремы 33.

Мы здесь не повторяем, вообще говоря, что F сепарабельно, кроме тех случаев, когда это существенно для справедливости теоремы.

²⁾ \mathbb{R} сепарабельно, поскольку счетное множество \mathbb{Q} рациональных чисел плотно в \mathbb{R} .

нуль. Следовательно, все множества, входящие в CA , имеют нулевую меру, а, значит, измеримы. Предположим теперь, что отображение f измеримо, и покажем, что тогда отображение g также измеримо. Пусть \mathcal{O} — открытое подмножество F . Перейдем от $f^{-1}(\mathcal{O})$ к $g^{-1}(\mathcal{O})$ с помощью двух следующих операций: добавим сначала множество элементов, принадлежащих $g^{-1}(\mathcal{O})$, но не принадлежащих $f^{-1}(\mathcal{O})$. Это множество является частью CA , а, значит, представляет собой измеримое множество. Далее, вычтем множество точек, принадлежащих $f^{-1}(\mathcal{O})$, но не принадлежащих $g^{-1}(\mathcal{O})$. Это множество снова является частью CA и, значит, измеримо. Это означает, что $g^{-1}(\mathcal{O})$ измеримо и что, следовательно, функция g измерима.

Теорема 22. Пусть F и G — два метризуемых пространства, f — измеримое отображение X в F и g — непрерывное отображение F в G . Тогда сложное отображение $g \circ f$ множества X в G измеримо.

Доказательство. В самом деле, пусть \mathcal{O} — открытое множество из G . Тогда прообраз $(g \circ f)^{-1}(\mathcal{O}) = f^{-1}(g^{-1}(\mathcal{O}))$ является прообразом при отображении f открытого множества $g^{-1}(\mathcal{O})$ и, следовательно, измерим.

Теорема 23. Пусть X — локально компактное пространство, снабженное мерой $\mu \geq 0$, и пусть F — метризуемое пространство. Тогда любое отображение f пространства X в F , являющееся μ -почти всюду пределом последовательности f_0, f_1, f_2, \dots μ -измеримых отображений X в F , измеримо. В частности, каждое отображение, являющееся почти всюду пределом последовательности непрерывных отображений, измеримо.

Когда мы говорим, что f является почти всюду пределом последовательности f_n , то хотим этим сказать, что эта последовательность является почти всюду сходящейся и имеет пределом функцию f . Другими словами, существует множество A , дополнение к которому по мере μ имеет нулевую меру, такое, что для каждого $x \in A$ последовательность $f_n(x)$ сходится к $f(x)$ при n , стремящемся к $+\infty$.

Доказательство. Предположим прежде всего, что последовательность f_n просто сходится всюду к f . Пусть B — произвольное подмножество F . Обозначим через A_n прообраз $f_n^{-1}(B)$, а затем через \mathcal{A}_m пересечение $\bigcap_{n \geq m} A_n$. Множество \mathcal{A}_m представляет собой множество таких точек x , для которых все $f_n(x)$ при $n \geq m$ лежат в B . Множества \mathcal{A}_m образуют некоторую возрастающую последовательность множеств, объединение которых мы обозначим через \mathcal{A} .

Поскольку каждое из множеств $f_n(x)$ измеримо, то для открытого или замкнутого подмножества B каждое из A_n измеримо и, следовательно, множества \mathcal{A}_m и \mathcal{A} измеримы.

Предположим, что B открыто, и покажем, что в этом случае множество \mathcal{A} заключено между $f^{-1}(B)$ и $f^{-1}(\bar{B})$:

$$f^{-1}(B) \subset \mathcal{A} \subset f^{-1}(\bar{B}). \quad (\text{IV}, 3; 13)$$

Прежде всего очевидно включение $\mathcal{A} \subset f^{-1}(\bar{B})$. В самом деле, $x \in \mathcal{A}$ означает, что x принадлежит одному из \mathcal{A}_m с подходящим m . Значит, $f_n(x) \in B$ для $n \geq m$, а, следовательно, $f(x) \in \bar{B}$ (теорема 15 гл. II). В рассуждении не требуется, чтобы множество B было открыто.

Пусть теперь B открыто; покажем, что $f^{-1}(B) \subset \mathcal{A}$. Если $x \in f^{-1}(B)$, то это говорит о том, что $f(x)$, равный пределу $f_n(x)$, является элементом y из B . Поскольку B открыто, существует некоторая окрестность элемента y , лежащая в B . Из определения предела вытекает, что существует такое целое число m , что для любого $n \geq m$ элемент $f_n(x)$ принадлежит B . Но тогда x принадлежит A_n для всех $n \geq m$ и, следовательно, принадлежит \mathcal{A}_m и тем более \mathcal{A} . Таким образом, мы не знаем, измеримы ли прообразы $f^{-1}(B)$ или $f^{-1}(\bar{B})$, но мы знаем, что если B открыто, то существует измеримое подмножество \mathcal{A} , заключенное между ними.

Пусть теперь B — произвольное замкнутое подмножество F . Обозначим через B_k объединение всех шаров¹⁾ радиуса $1/k$ с центрами в B . Это множество как объединение открытых шаров открыто. Последовательность B_k , $k = 1, 2, 3, \dots$, является убывающей последовательностью, а пересечение B_k равно B . В самом деле, очевидно, что это пересечение содержит B . С другой стороны, если x принадлежит этому пересечению, то расстояние от x до замкнутого множества B не больше $1/k$ при любом k , т. е. равно нулю, а это означает, что $x \in B$. В свою очередь пересечение замыканий \bar{B}_k совпадает с B . В самом деле, с одной стороны, оно содержит B , но с другой, если точка принадлежит замыканию \bar{B}_k , она является пределом точек из B_k и, в силу непрерывности функции расстояния (см. гл. II, стр. 83), расстояние от x до $\bar{B} \leq 1/k$. Таким образом, если x лежит в пересечении \bar{B}_k , то его расстояние до B равно нулю, т. е. x лежит в B . Мы образовали две убывающие последовательности множеств: открытые множества B_k и множества \bar{B}_k , пересечение которых равно B . Поскольку B_k открыто, то из сказанного выше вытекает, что между множествами $f^{-1}(B_k)$ и $f^{-1}(\bar{B}_k)$ можно вставить измеримое множество $\mathcal{A}^{(k)}$. Пересечение измеримых множеств $\mathcal{A}^{(k)}$, являющееся измеримым, должно

¹⁾ Надо сначала выбрать в F какую-нибудь метрику.

содержать пересечение множеств $f^{-1}(B_k)$ и содержаться в пересечении множеств $f^{-1}(\bar{B}_k)$. Тем самым обязательно измеримо множество $f^{-1}(B)$, чем и доказывается теорема для этого случая.

Перейдем теперь к общему случаю, при котором последовательность f_n только почти всюду сходится к f . Обозначим через A множество тех x , для которых последовательность $f_n(x)$ сходится к $f(x)$. Изменим все функции f_n , полагая $f_n(x) = c$ при x , принадлежащем $\mathbf{C}A$, где c — произвольный фиксированный элемент F . Измененные функции \tilde{f}_n почти всюду равны исходным функциям f_n и, следовательно, каждая из них измерима (теорема 21). Однако теперь последовательность измененных функций \tilde{f}_n сходится *всюду* к функции, равной f во всех точках множества A и равной постоянной c во всех точках множества $\mathbf{C}A$. Эта функция, согласно изложенному выше, измерима, и поскольку она почти всюду равна f , то f также измерима. Эта теорема допускает следующее очень важное обращение.

μ -этажные функции

Этажные функции по отношению к мере Радона $\mu \geq 0$ на локально компактном пространстве X обобщают ступенчатые функции, введенные для определения интеграла Римана.

Пусть F — произвольное множество. Отображение f пространства X в множестве F называется μ -этажным или просто этажным (когда это не приводит к недоразумению), если существует разбиение X на конечное число непересекающихся подмножеств X_1, X_2, \dots, X_n , таких, что каждое из них μ -измеримо и на каждом из них отображение f постоянно. Такое разбиение называется допустимым разбиением для отображения f . Естественно, что для данного отображения f может существовать бесконечное множество таких разбиений точно так же, как это было для ступенчатых функций. Среди них существует наилучшее в том смысле, что оно содержит наименьшее возможное число частей, но оно нас не интересует. Если \vec{f} и \vec{g} — две этажные функции со значениями в векторном пространстве \vec{F} , то сумма $\vec{f} + \vec{g}$ и произведение \vec{f} на постоянный скаляр являются этажными функциями.

Последнее утверждение очевидно. Первое же вытекает из того, что если $(X_i)_{i \in I}$ является допустимым разбиением X для \vec{f} и $(X_j)_{j \in J}$ является допустимым разбиением для \vec{g} , то каждое из множеств $X_i \cap X_j$ измеримо и на каждом из них функция \vec{f} и функция \vec{g} постоянны, а вместе с ними постоянна функция $\vec{f} + \vec{g}$. Другими словами, *этажные функции на X со значениями*

в векторном пространстве \vec{F} образуют векторное подпространство \vec{F}^x всех функций, заданных на X со значениями в \vec{F} .

Теорема 23₂. Для того чтобы отображение локально компактного счетного в бесконечности пространства X в метризуемое пространство F было μ -измеримым, необходимо и достаточно, чтобы оно было пределом μ -почти всюду некоторой последовательности μ -этажных функций.

Доказательство. Каждая этажная функция измерима, поскольку прообраз каждой части при отображении, определяемом этой функцией, есть объединение измеримых частей X , на которых функция постоянна. Согласно теореме 23, этим свойством будет обладать функция, являющаяся пределом почти всюду последовательности этажных функций.

Обратно, пусть f — измеримое отображение X в метрическое сепарабельное пространство F . Пусть a_0, a_1, a_2, \dots — некоторая последовательность, плотная в F . Зададим произвольное $\varepsilon > 0$. Объединение всех шаров $B(a_n, \varepsilon)$ равно F . Но эти шары не являются непересекающимися. Образуем последовательность множеств $C_{n, \varepsilon} = B(a_n, \varepsilon) \cap C B(a_0, \varepsilon) \cap C B(a_1, \varepsilon) \cap \dots \cap C B(a_{n-1}, \varepsilon)$. Множества $C_{n, \varepsilon}$ не пересекаются, а их объединение равно F . Множества $D_{n, \varepsilon} = f^{-1}(C_{n, \varepsilon})$ не пересекаются, и их объединение равно X . Кроме того, $C_{n, \varepsilon}$ являются борелевскими множествами, а, следовательно, $D_{n, \varepsilon}$ измеримы. Пусть K — некоторый компакт X . Множества $D_{n, \varepsilon} \cap K$ не пересекаются, измеримы, а их объединение равно K . Следовательно, ряд $\sum_{n=0}^{\infty} \mu(D_{n, \varepsilon} \cap K)$ сходится, а его сумма равна $\mu(K)$. Для каждого $\delta > 0$ можно определить целое число $N(\varepsilon, K, \delta)$, зависящее от ε , K , δ и такое, чтобы объединение $D_{k, \varepsilon} \cap K$ для всех $k > N$ имело меру $\leq \delta$. Пусть $E(\varepsilon, K, \delta)$ является объединением множеств $D_{k, \varepsilon}$ при $k > N(\varepsilon, K, \delta)$. Тогда $\mu(E(\varepsilon, K, \delta) \cap K) \leq \delta$.

Определим теперь следующим образом функции f_n . Пусть K_n — такая возрастающая последовательность компактов, что X является объединением внутренностей \hat{K}_n (см. в лемме 3 к теореме 11 последовательность B_n).

Известно, что для любого компакта K множества X существует такое целое число m , что $K \subset K_n$ для $n \geq m$. Положим

$$f_n(x) = \begin{cases} a_k, & \text{если } x \in D_{k, 1/n}, k \leq N(1/n, K_n, 1/2^{n+1}), \\ a & \text{в других случаях.} \end{cases}$$

Множества $D_{k, 1/n}$ для $k = 0, 1, 2, \dots$ не пересекаются, измеримы и в объединении дают все множество X . Значит, f_n является по построению этажной функцией. Если $x \in D_{k, 1/n}$,

$k \leq N(1/n, K_n, 1/2^{n+1})$, то $f_n(x) = a_k$, и так как $f(D_{k, 1/n}) \subset C_{k, 1/n} \subset B(a_k, 1/n)$, то имеем $f(x) \in B(a_k, 1/n)$. Но тогда $d(f(x), f_n(x)) \leq 1/n$. Следовательно, $f_n(x)$ не сходится к $f(x)$ при n , стремящемся к бесконечности, только в том случае, когда для бесконечного числа значений n точка x находится в объединении множеств $D_{k, 1/n}$, $k > N(1/n, K_n, 1/2^{n+1})$, т. е. в $E(1/n, K_n, 1/2^{n+1})$. Множество E точек X , в которых $f_n(x)$ не сходится к $f(x)$ при n , стремящемся к бесконечности, удовлетворяет для бесконечного множества n соотношению $E \subset E(1/n, K_n, 1/2^{n+1})$, а, значит, $E \subset \bigcup_{n \geq m} E(1/n, K_n, 1/2^{n+1})$ для любого m .

Пусть K — произвольный компакт из X и m — такое целое число, что $K_m \supset K$. Тогда $E \cap K \subset \bigcup_{n \geq m} (E(1/n, K_n, 1/2^{n+1}) \cap K_n)$, а, следовательно, $\mu(E \cap K) \leq \sum_{n \geq m} 1/2^{n+1} = 1/2^m$.

Так как в качестве m можно взять сколь угодно большое целое число, то $\mu(E \cap K) = 0$. Поскольку это справедливо для каждого компакта K из X , то $\mu(E) = 0$, и, значит, f_n сходится к f почти всюду при n , стремящемся к бесконечности.

Следствие 1. Пусть F_1, F_2, \dots, F_m — некоторые метрические пространства. Для того чтобы отображение f пространства X в произведение $F_1 \times F_2 \times \dots \times F_m$, определенное отображениями f_i из X в F_i , было измеримым, необходимо и достаточно, чтобы каждое из составляющих отображений f_i было измеримым. В частности, для того чтобы отображение X в аффинное конечномерное пространство, снабженное системой координат, было измеримым, необходимо и достаточно, чтобы каждая из его составляющих была измеримой.

Доказательство. Совершенно ясно, что если отображение f измеримо, то каждое из его составляющих отображений f_i измеримо. В самом деле, каждая составляющая получается как композиция f и проекции F на F_i , являющейся непрерывным отображением. Остается лишь применить теорему 22.

Обратное утверждение носит более глубокий характер. Оно опирается на теорему 23₂. По предположению, f_i является измеримым отображением. Следовательно, существует последовательность этажных отображений $(f_i)_n$ множества X в F_i , сходящаяся к f_i , при n , стремящемся к бесконечности, всюду, кроме точек множеств $B_i \subset X$ меры нуль. Объединение B множеств B_i также имеет нулевую меру. По определению топологического произведения, в каждой точке $x \in CB$ последовательность $f_n(x) = ((f_1)_n(x), (f_2)_n(x), \dots, (f_m)_n(x))$ сходится к $f(x)$. Следовательно, f , являясь пределом почти всюду последовательности этажных функций, измерима.

Следствие 2. Если \vec{F} является векторным нормированным пространством, \vec{f} и \vec{g} — измеримыми отображениями X в \vec{F} и k — скалярной измеримой функцией, то отображения $\vec{f} + \vec{g}$ и $k\vec{f}$ измеримы, т. е. μ -измеримые на X функции со значениями в \vec{F} образуют некоторое векторное пространство.

Доказательство. Отображение $x \rightarrow \vec{f}(x) + \vec{g}(x)$ является композицией отображения $x \rightarrow (\vec{f}(x), \vec{g}(x))$ множества X в измеримое, по следствию 1, множество $\vec{F} \times \vec{F}$ и непрерывного отображения $(\vec{f}, \vec{g}) \rightarrow \vec{f} + \vec{g}$ пространства $\vec{F} \times \vec{F}$ в пространство \vec{F} . Согласно теореме 22, это отображение измеримо. Точно так же отображение $x \rightarrow k(x)\vec{f}(x)$ является композицией измеримого отображения $x \rightarrow (k(x), \vec{f}(x))$ множества X в $\mathbb{K} \times \vec{F}$ и непрерывного отображения $(k, \vec{f}) \rightarrow k\vec{f}$ множества $\mathbb{K} \times \vec{F}$ в \vec{F} .

Следствие 3. Пусть $\vec{F}, \vec{G}, \vec{H}$ — векторные нормированные пространства, \vec{f} и \vec{g} — измеримые отображения X в \vec{F} и \vec{G} соответственно, и пусть B — билинейное непрерывное отображение $\vec{F} \times \vec{G}$ в \vec{H} . Тогда отображение $B(\vec{f}, \vec{g}): x \rightarrow B(\vec{f}(x), \vec{g}(x))$ из X в \vec{H} измеримо.

Доказательство. Рассматриваемое отображение является композицией измеримого отображения $x \rightarrow (\vec{f}(x), \vec{g}(x))$ множества X в пространство $\vec{F} \times \vec{G}$ (следствие 1) и непрерывного отображения B пространства $\vec{F} \times \vec{G}$ в пространство \vec{H} , а, значит, согласно теореме 22, является измеримым отображением^{1).}

Из предыдущих теорем можно сделать вывод, что все функции, встречающиеся на практике, измеримы. Для случая, когда X — прямая и $\mu = dx$, никто не может привести явный пример неизмеримой вещественной функции. Конечно, пользуясь аксиомой выбора, можно доказать существование вещественной неизмеримой функции, например, характеристической функции множества A , построенного на стр. 504.

¹⁾ Предположение о том, что B билинейно, излишне. Мы приводим это следствие потому, что оно полезно на практике.

Борелевские функции

Отображение f топологического пространства X в топологическое пространство F называется *борелевским*, если прообраз при отображении f каждой открытой части F является борелевским множеством. Так как части F , прообразы которых при отображении f являются борелевскими, образуют, очевидно, σ -алгебру частей из F , то эта σ -алгебра может содержать открытые множества (или замкнутые множества) лишь тогда, когда она содержит борелевские. Поэтому в определении можно заменить открытые части на замкнутые или борелевские части из F . Это понятие зависит только от топологии X и F . Например, часть из X будет борелевской тогда и только тогда, когда ее характеристическая функция является борелевской. Мы видим, что если X локально компактно и счетно в бесконечности, а метризуемое пространство F сепарабельно, то борелевская функция на X со значениями в F универсально измерима, т. е. μ -измерима по любой мере Радона $\mu \geq 0$ на X .

Борелевские функции обладают следующими свойствами.

1°) *Композиция двух борелевских отображений является борелевским отображением*: если f является борелевским отображением X в Y , а g — борелевским отображением Y в Z , то $g \circ f$ является борелевским отображением X в Z , поскольку прообраз борелевского множества является борелевским множеством.

2°) *Предел последовательности борелевских отображений X в метризуемое пространство F является борелевским отображением*. Доказательство то же самое, что и доказательство теоремы 23; следует лишь заменить в нем везде слово «измеримый» на «борелевский».

3°) Не ясно, является ли каждая борелевская функция пределом последовательности этажных борелевских функций (под этим понимается функция, принимающая лишь конечное число значений на некоторых борелевских множествах; все это без какой-либо связи с мерой Радона). Воспользуемся обозначениями, введенными при доказательстве теоремы 23₂. Пусть u_n — функция, равная a_n в множестве $D_{n, 1/n}$. Так как $B(a_k, 1/n)$, а, значит, $C_{k, 1/n}$ и, следовательно, $D_{k, 1/n}$ являются борелевскими множествами и $d(f, u_n) \leq 1/n$, то функция u_n является борелевской. Однако она имеет счетное множество этажей $D_{k, 1/n}$, следовательно, не является этажно-борелевской. Функция u_n является простым пределом при m , стремящемся к бесконечности, этажно-борелевских функций $v_{n, m}$, где $v_{n, m} = a_k$ в $D_{k, 1/n}$ для $k \leq m$ и $v_{n, m} = a \in F$ вне этих множеств. Таким образом,

каждая борелевская функция f со значениями в метризуемом сепарабельном пространстве F является равномерным пределом последовательности борелевских функций u_n , каждая из которых является простым пределом последовательности этажно-борелевских функций $v_{n,m}$.

4°) Для того чтобы отображение f пространства X в произведение $F_1 \times F_2 \times \dots \times F_m$ было борелевским, необходимо и достаточно в случае метрических сепарабельных пространств F_i , чтобы каждая из составляющих f_i отображения f была борелевским отображением X в F_i . В самом деле, каждое отображение f_i является композицией f и непрерывной проекции F на F_i . Так как f является борелевским отображением, то f_i также будут борелевскими отображениями. Обратно, предположим, что f_i — борелевские отображения и F_i — метризуемые сепарабельные пространства. Теперь можно воспроизвести доказательство следствия 1 теоремы 23₂, но с двумя последовательными переходами к пределу. Пользуясь свойством 3°), получаем, что отображение f борелевское.

Замечание. Теперь можно существенно улучшить результат теоремы 23₂.

1°) Пусть сначала A — измеримое множество конечной меры. Тогда существует убывающая последовательность открытых множеств $\mathcal{O}_n \supset A$, таких, что $\mu(\mathcal{O}_n) \leq \mu(A) + 1/n$. Пересечение A^* множеств \mathcal{O}_n является борелевским множеством, содержащим A , и $\mu(A^* - A) = 0$. Заменяя открытые множества $\mathcal{O}_n \supset A$ на компакты $K_n \subset A$, можно построить борелевское множество $A_* \subset A$ так, что $\mu(A - A_*) = 0$. Если мера множества A не конечна, то его можно представить в виде объединения множеств $A \cap K_n$, где K_n — возрастающая последовательность компактов, объединение которых дает X , и, значит, во всех случаях можно найти два борелевских множества A^* и A_* , такие, что $A_* \subset A \subset A^*$ и $\mu(A^* - A_*) = 0$.

Можно сказать, что с точностью до множества нулевой μ -меры всякое μ -измеримое множество является борелевским. Это обстоятельство весьма интересно, поскольку борелевские множества универсально измеримы и зависят только от топологии X . Можно еще сказать, что σ -алгебра μ -измеримых подмножеств X (зависящая от μ) является наименьшей σ -алгеброй, содержащей все борелевские множества (зависящие лишь от топологии X) и подмножества нулевой μ -меры.

2°) Пусть теперь f является μ -измеримой функцией со значениями в метрическом сепарабельном пространстве F . При построении этажных функций f_n в теореме 23₂ вместо того, чтобы брать $f_n = a_k$ на D_k , мы положим $f_n = b_k$, где b_k — произ-

вольная точка множества $B(a_k, 1/n)$. Теперь можно b_k выбирать в образе $f(X) \subset F$ (ибо, если $D_{k, 1/n}$ не пусто, то $f(X) \cap B(a_k, 1/n)$ не пусто). Таким образом, можно считать, что все функции f_n принимают значения в образе $f(X)$.

3°) Пусть A — этажи функции f_n . Каждый этаж является объединением некоторого борелевского множества $B_{i,n}$ и некоторого множества нулевой меры. Функция g_n , равная f_n на множествах $B_{i,n}$ и любой фиксированной точке множества $f(X)$ вне $B_{i,n}$, является борелевской и почти всюду равна f_n . Поэтому функции g_n могут полностью заменить функции f_n . Эти функции g_n сходятся к f только почти всюду. Пусть N — множество точек, в которых они не сходятся. Пусть N^* — борелевское множество нулевой меры, содержащее N . Пусть h_n — функция, равная g_n на CN^* и фиксированному элементу множества $f(X)$ на N^* . Тогда h_n будут борелевскими функциями со значениями в $f(X)$. Однако теперь они сходятся *всюду* к некоторому пределу h , почти всюду равному f и принимающему значения в $f(X)$. Предел h последовательности борелевских функций является борелевской функцией. Поэтому окончательно получаем следующее:

Каждая измеримая функция f , определенная на X , со значениями в F почти всюду равна некоторой борелевской функции, принимающей значения в $f(X)$, которая сама является почти всюду пределом последовательности этажно-борелевских функций, значения которых лежат в $f(X)$.

4°) Приведем улучшение другого рода. Предположим, что F компактно, или просто, что $f(X)$ содержится в компакте множества F . Тогда для каждого n найдется конечное число шаров $B(a_k, 1/n)$, покрывающих это множество. Для построения функций f_n вместо числа $N(1/n, K_n, 1/2^{n+1})$ в общем случае можно взять конечное число этих шаров. Так как при этом $d(f_n, f) \leqslant 1/n$, то этажные функции f_n сходятся *всюду* и *равномерно* к функции f . Можно даже в этом случае выбрать их так, чтобы $f_n(X) \subset f(X)$. Однако их невозможно взять борелевскими (за исключением того случая, когда сама функция f является борелевской; тогда по самому построению они будут борелевскими).

Предположим, например, что $F = \mathbb{C}$ и что f является ограниченной борелевской функцией с компактным носителем K . Свои значения она принимает в некотором компакте пространства \mathbb{C} , а, следовательно, является равномерным пределом некоторой последовательности этажных борелевских функций, ограниченных в их совокупности и имеющих носители в фиксированном компакте (поскольку всегда можно умножить их на характеристическую функцию компакта K).

Интеграл от векторной этажной функции

Пусть \vec{f} есть μ -этажная функция с компактным носителем. Тогда интегралом от этажной функции \vec{f} по мере μ называют величину

$$\int \vec{f} d\mu = \sum_{i \in I} \vec{f}_i \mu(X_i) \quad (\text{IV}, 3; 14)$$

для допустимого разложения функции \vec{f} . Здесь \vec{f}_i — постоянные значения функции \vec{f} на X_i . Непосредственно видно, как мы в этом убедились в случае интеграла Римана, что такая величина не зависит от выбранного разложения функции \vec{f} . Интеграл от характеристической функции измеримого множества с компактным замыканием является его мерой. Таким образом определенный интеграл является линейным отображением пространства этажных функций с компактным носителем и значениями в \vec{F} в векторное пространство \vec{F} . С другой стороны, если \vec{f} является этажной функцией, равной нулю в дополнении к измеримому множеству A , то мы сразу же получаем оценку

$$\left\| \int \vec{f} d\mu \right\| \leq \int \| \vec{f} \| d\mu \leq \| \vec{f} \| \mu(A). \quad (\text{IV}, 3; 15)$$

Если \vec{f} и \vec{g} являются этажными почти всюду равными функциями, то они имеют один и тот же интеграл.

Верхний интеграл от вещественной неотрицательной функции

Определение. Пусть X — локально компактное пространство, счетное в бесконечности, снабженное мерой Радона $\mu \geq 0$. Пусть f — вещественная функция ≥ 0 , определенная на X , с конечными или бесконечными значениями¹⁾. Если f ограничена и имеет компактный носитель, то *верхним интегралом от функции f относительно меры μ* называется точная нижняя грань интегралов относительно μ от этажных функций g с компактным носителем $\geq f$. Пусть теперь f — произвольная функция. Тогда верхним интегралом от f относительно μ называется точная верхняя грань, конечная или равная $+\infty$, верхних интегралов от ограниченных функций ≥ 0 с компактным носителем, не превосходящих f .

¹⁾ В этом случае выражение «вещественная функция» не верно, так как функция принимает значения не в \mathbb{R} , а в $\bar{\mathbb{R}}$.

Верхний интеграл от f относительно μ обозначается через $\int^* f d\mu$ ¹⁾, или $\int^* f(x) d\mu(x)$, или $\int^* f$, если это не приводит к недоразумению.

З а м е ч а н и е. Пусть M — некоторое число $\geqslant 0$ и K — некоторый компакт множества X . Обозначим через $f_{M,K}$ функцию, равную $f(x)$ в точках x компакта K , в которых $f(x) \leqslant M$, равную M в точках компакта K , в которых $f(x) > M$, и равную 0 вне K . Тогда $f_{M,K}$ является ограниченной функцией с компактным носителем, а ее верхний интеграл описывается согласно первой части определения. Вторая часть эквивалентна следующему утверждению: верхний интеграл от f является точной верхней гранью верхних интегралов от функций $f_{M,K}$. Можно также сказать, что если M_0, M_1, M_2, \dots является числовой последовательностью, стремящейся к $+\infty$ при n , стремящемся к $+\infty$, и если K_0, K_1, K_2, \dots является возрастающей последовательностью компактов, дающих в объединении X (такая последовательность всегда существует, поскольку X предполагается счетным в бесконечности), то верхний интеграл от f является пределом при n , стремящемся к $+\infty$, верхних интегралов от функций f_{M_n, K_n} :

Теорема 24. *Если функции f и g почти всюду равны, то их верхние интегралы равны.*

Д о к а з а т е л ь с т в о. Предположим сначала, что обе почти всюду равные функции f и g ограничены числом $M \geqslant 0$ и имеют носители, лежащие в компакте K . Пусть B — множество нулевой меры точек x , в которых $f(x) \neq g(x)$.

Пусть теперь f_1 — этажная функция $\geqslant f$. Если через g_1 обозначить верхнюю огибающую функции f_1 и произведения M на характеристическую функцию B , то g_1 будет этажной, $g_1 \geqslant g$, и поскольку f_1 и g_1 почти всюду равны, то $\int^* g \leqslant \int^* g_1 = \int f_1$. Если взять теперь точную нижнюю грань интегралов от f_1 , то

¹⁾ Существенная разница между верхним интегралом Римана (относительно dx) и верхним интегралом Лебега (относительно произвольной меры $\mu \geqslant 0$) состоит в том, что первый получается одним переходом к пределу, в то время как второй требует двух предельных переходов при вычислении точной верхней грани и точной нижней грани. Здесь мы произвели четыре перехода к пределу: два для того, чтобы определить измеримые множества, и два теперь. Однако это всего лишь видимость, вызванная методом изложения, не являющимся, как мы об этом уже говорили, наилучшим (см. примечание на стр. 490). В действительности определение верхнего интеграла можно свести к двум предельным переходам, но не к одному.

Позже мы получим следующее соотношение между верхним интегралом Римана и верхним интегралом Лебега: для ограниченной с компактным носителем функции $f \geqslant 0$ на \mathbb{R} при $\mu = dx$ верхний интеграл Римана \geqslant верхнего интеграла Лебега.

мы увидим, что $\int^* g \leqslant \int^* f$. Такое же рассуждение показывает, что $\int^* f \leqslant \int^* g$. Предположим теперь, что f и g — произвольные функции. Пусть теперь f_2 — некоторая ограниченная функция $\leqslant f$ с компактным носителем. Если через g_2 мы обозначим нижнюю огибающую функций g и f_2 , то получим $g_2 \leqslant g$. Кроме того, так как f_2 и g_2 почти всюду равны, ограничены и имеют компактный носитель, то $\int^* g \geqslant \int^* g_2 = \int^* f_2$. Переходя к точной верхней грани в верхних интегралах от f_2 , получаем: $\int^* g \geqslant \int^* f$. Аналогичное рассуждение с перестановкой f и g окончательно приводит к равенству $\int^* g = \int^* f$.

Следствие. Если функция $f \geqslant 0$ почти всюду равна нулю, то ее верхний интеграл равен нулю.

Теорема 25. Пусть f и g — две функции $\geqslant 0$ на X и k — скалярная постоянная. Тогда если $f \leqslant g$, то $\int^* f \leqslant \int^* g$. Кроме того,

$$\int^* (f + g) \leqslant \int^* f + \int^* g \quad \text{и} \quad \int^* kf = k \int^* f. \quad (\text{IV}, 3; 16)$$

Первая формула является неравенством выпуклости. В противоположность тому, что имело место для интеграла Римана, эта формула распространяется на счетное неравенство выпуклости¹⁾, в чем проявляется существенное преимущество интеграла Лебега над интегралом Римана.

Доказательство. Свойства эти очевидны. Докажем, например, первую из формул (IV, 3; 16).

Предположим сначала, что f и g ограничены и имеют компактный носитель. Тогда, с точностью до очевидных изменений, доказательство совпадет с доказательством, приведенным для интеграла Римана в теореме 2. Рассмотрим теперь общий случай, когда f и g произвольны. Если ввести определенные выше функции $f_{M, K}$, то мы получим неравенство

$$(f + g)_{M, K} \leqslant f_{M, K} + g_{M, K}, \quad (\text{IV}, 3; 17)$$

откуда

$$\int^* (f + g)_{M, K} \leqslant \int^* f_{M, K} + \int^* g_{M, K} \leqslant \int^* f + \int^* g.$$

¹⁾ Следствие 3 теоремы 36.

Переходя в интеграле $\int^*(f+g)_{M,K}$ к точной верхней грани по всем значениям M и всем компактам $K \subset X$, получаем требуемый результат.

Легко видеть, что если A является некоторой частью X и φ_A — ее характеристическая функция, то

$$\int^* \varphi_A = \mu^*(A)^1). \quad (\text{IV, 3; 18})$$

В частности, $\int^* 1$ является мерой X , т. е. нормой μ .

Теорема 26. *Если верхний интеграл от $f \geq 0$ конечен, то функция f почти всюду конечна. Если верхний интеграл от f равен нулю, то и функция f почти всюду равна нулю.*

Доказательство. Пусть сначала $\int^* f < +\infty$. Обозначим через A_n множество точек x , в которых $f(x) \geq n$. Если φ_n является его характеристической функцией, то $f \geq n\varphi_n$, откуда вытекает неравенство

$$\int^* f \geq n \mu^*(A_n). \quad (\text{IV, 3; 19})$$

Если через A обозначить множество точек x , в которых $f(x) = +\infty$, то при любом n будет справедливым неравенство

$$\int^* f \geq n \mu^*(A), \quad (\text{IV, 3; 20})$$

из которого следует, что $\mu^*(A) = 0$, т. е. f почти всюду принимает конечные значения.

Предположим теперь, что $\int^* f = 0$. Обозначим на этот раз через B_k множество точек x , в которых $f(x) \geq 1/k$, а через φ_k — его характеристическую функцию. Так как $f \geq \frac{1}{k}\varphi_k$, то имеет место неравенство

$$0 = \int^* f \geq \frac{1}{k} \mu^*(B_k), \quad (\text{IV, 3; 21})$$

из которого следует, что $\mu^*(B_k) = 0$. Множество B точек x , в которых $f(x) > 0$, являясь объединением счетной совокупности множеств B_k с нулевой мерой, само имеет нулевую меру, а, следовательно, функция f почти всюду равна нулю.

¹⁾ Это очевидно, если \bar{A} компактно. Если же A произвольно, то, согласно замечанию 1° к теореме 19, результат можно получить с помощью формулы $\mu^*(A) = \sup_{K \subset A} \mu^*(A \cap K)$ (K — компакты).

Интегрируемость функций с векторными значениями

Определение. Пусть X — локально компактное пространство, снабженное мерой Радона $\mu \geq 0$ и \vec{F} — векторное нормированное пространство над полем вещественных или комплексных чисел. Говорят, что функция \vec{f} , определенная на X , со значениями в \vec{F} μ -интегрируема, или интегрируема относительно μ , или просто интегрируема, когда это не приводит к недоразумению, если, каково бы ни было $\varepsilon > 0$, существует такая функция \vec{g} , определенная на X , со значениями в \vec{F} , μ -этажная и с компактным носителем, для которой имеет место неравенство:

$$\int^* \| \vec{f} - \vec{g} \| d\mu \leq \varepsilon. \quad (\text{IV}, 3; 22)$$

Сказанное эквивалентно следующему определению: существует такая последовательность $\vec{f}_1, \vec{f}_2, \vec{f}_3, \dots$ этажных функций с компактными носителями, определенных на X , со значениями в \vec{F} , что величина $\int^* \| \vec{f} - \vec{f}_n \|$ стремится к нулю при n , стремящемся к бесконечности. Такая последовательность называется аппроксимирующей последовательностью функции \vec{f} для интеграла Лебега относительно μ . Вместо интегрируемая функция говорят также суммируемая.

Если две функции, определенные на X , со значениями в \vec{F} почти всюду равны и если одна из них интегрируема, то интегрируема и другая функция, которая будет иметь ту же аппроксимирующую последовательность (теорема 24).

Замечания. Замечания 1°)–3°) на стр. 424–425 могут быть здесь повторены с заменой слов «интегрируема по Риману относительно dx на \mathbb{R} » словами «имеет компактный носитель и интегрируема по Лебегу относительно μ на X » (ограничение «с компактным носителем» не нужно для замечания 1°)) и слова «ступенчатая» на слово « μ -этажная». Если нам надо будет сделать ссылку на это замечание, то мы будем говорить: см. замечания 1°), 2°) или 3°) на стр. 524.

Интеграл Лебега от функций с векторными значениями

Теорема 27. Пусть \vec{f} — отображение X в банаево пространство \vec{F} . Если \vec{f} интегрируема и если \vec{f}_n образуют аппроксимирующую последовательность этажных функций с ком-

пактными носителями для \vec{f} , то интегралы $\int \vec{f}_n$ имеют предел, не зависящий от выбора аппроксимирующей последовательности.

Доказательство аналогично доказательству теоремы 3.

Определение. В условиях предыдущей теоремы общий предел интегралов $\int \vec{f}_n$ по всем аппроксимирующими последовательностям называется *интегралом от функции \vec{f} относительно μ* . Интеграл этот обозначается различными способами:

$$\int \vec{f} d\mu, \quad \int \vec{f} \mu, \quad \int \vec{f}(x) d\mu(x), \quad \int \vec{f}^1. \quad (\text{IV}, 3; 23)$$

Если две интегрируемые функции почти всюду равны, то они имеют один и тот же интеграл, поскольку у них одни и те же аппроксимирующие последовательности.

Если \vec{f} является отображением некоторой части Y множества X в \vec{F} , то говорят, что \vec{f} интегрируема на Y , если интегрируемой будет функция \vec{f} , определенная на X , равная \vec{f} на Y и $\vec{0}$ вне Y . При этих условиях полагают

$$\int_Y \vec{f} = \int \vec{f}^2. \quad (\text{IV}, 3; 24)$$

Замечание. Естественно, каждая этажная функция с компактным носителем интегрируема и ее интеграл определяется по формуле (IV, 3; 14).

Теорема 28. Если \vec{f} и \vec{g} — две интегрируемые на X функции со значениями в \vec{F} и k — скалярная постоянная, то $\vec{f} + \vec{g}$ и $k\vec{f}$ также интегрируемы; кроме того, имеют место следующие равенства:

$$\int k\vec{f} = k \int \vec{f} \quad \text{и} \quad \int (\vec{f} + \vec{g}) = \int \vec{f} + \int \vec{g}. \quad (\text{IV}, 3; 25)$$

Доказательство аналогично доказательству теоремы 4.

Эта теорема говорит о том, что μ -интегрируемые на X функции со значениями в \vec{F} образуют векторное пространство по

1) Имеется такое множество различных обозначений для интеграла от \vec{f} относительно μ , что остается лишь сказать: всякая формула, в которой присутствует векторная функция \vec{f} и мера $\mu \geq 0$, означает интеграл от \vec{f} относительно μ .

2) Если \vec{f} уже определена всюду на X , то \vec{f} является произведением $\vec{f}\varphi_Y$ функции \vec{f} на характеристическую функцию φ_Y множества Y .

отношению к полю скаляров \mathbb{K} и что интеграл является линейным отображением этого векторного пространства в \vec{F} .

Следствие 1. Если Y и Z — два подмножества X , то функция \vec{f} будет интегрируемой на их объединении тогда и только тогда, когда она интегрируема на каждом из этих подмножеств; при этом

$$\int_{Y \cup Z} \vec{f} = \int_Y \vec{f} + \int_Z \vec{f}. \quad (\text{IV, 3; 26})$$

Доказательство аналогично доказательству следствия теоремы 4.

Следствие 2. Если \vec{f} является этажной функцией, $\vec{f} = \sum_i \vec{f}_i \Phi_{X_i}$, где \vec{f}_i — постоянные, а Φ_{X_i} — характеристические функции измеримых множеств X_i конечной меры (не обязательно имеющих компактное замыкание), то \vec{f} интегрируема и имеет место соотношение (IV, 3; 14). В частности, если A является измеримым множеством конечной меры, то $\int \Phi_A = \mu(A)$.

Доказательство. В силу теоремы 28, достаточно доказать только последнее утверждение. Пусть K_n — возрастающая последовательность компактов, объединение которых равно X . По теореме 19, $\mu(A) = \lim_{n \rightarrow \infty} \mu(A \cap K_n)$ и $\lim_{n \rightarrow \infty} \mu(A - A \cap K_n) = 0$.

Характеристические функции $\Phi_{A \cap K_n}$ являются этажными с компактным носителем. Поэтому $\int \Phi_{A \cap K_n} = \mu(A \cap K_n)$ и (согласно (IV, 3; 18)) $\int^*(\Phi_A - \Phi_{A \cap K_n}) = \mu^*(A - A \cap K_n) = \mu(A - A \cap K_n)$ стремится к нулю. Таким образом, $\Phi_{A \cap K_n}$ образуют аппроксимирующую последовательность для Φ_A , которая поэтому оказывается интегрируемой и $\int \Phi_A = \lim_{n \rightarrow \infty} \int \Phi_{A \cap K_n} = \lim_{n \rightarrow \infty} \mu(A \cap K_n) = \mu(A)$.

Теорема 29. Если функция \vec{f} , определенная на X , со значениями в \vec{F} интегрируема, то функция $\|\vec{f}\|: x \rightarrow \|\vec{f}(x)\|$ также интегрируема и имеет место неравенство

$$\left\| \int \vec{f} \right\| \leq \int \|\vec{f}\|. \quad (\text{IV, 3; 27})$$

Доказательство аналогично доказательству теоремы 5.

Следствия. Здесь можно повторить следствия теоремы 5. В следствиях 3 и 3_2 надо заменить $[a, b]$ на произвольную измеримую часть Y множества X и $(b - a)$ на $\mu(Y)$. Ссылаясь на эти результаты, мы будем говорить: следствия 1, 2, 3, 3_2 , 4, 5 теоремы 29.

Теорема 30. Если φ является функцией с комплексными значениями, принадлежащей множеству $\mathcal{C}(X)$, то она интегрируема относительно μ и ее интеграл $\int \varphi$ совпадает с $\mu(\varphi)$.

Эта теорема весьма существенна. Она говорит о том, что довольно сложное построение, выполненное нами для интегрирования, дает некоторое продолжение исходной линейной формы $\varphi \rightarrow \mu(\varphi)$, определенной на $\mathcal{C}(X)$.

Доказательство. Докажем прежде всего, что каждая функция $\varphi \in \mathcal{C}(X)$ интегрируема и, более общо, что каждая непрерывная с компактным носителем K функция \vec{f} на X со значениями в \vec{F} интегрируема. Пусть \mathcal{V} — некоторая компактная окрестность K , и пусть задано произвольное число $\varepsilon > 0$. Для каждой точки $a \in K$ выберем такую ее окрестность $\mathcal{V}_a \subset \mathcal{V}$, чтобы для $x \in \mathcal{V}_a$ выполнялось неравенство $\|\vec{f}(x) - \vec{f}(a)\| \leq \varepsilon / \mu(\mathcal{V})$. Конечное число таких окрестностей покрывает все множество K . Пусть это будут окрестности $\mathcal{V}_{a_1}, \mathcal{V}_{a_2}, \dots, \mathcal{V}_{a_n}$. Положим $C_1 = \mathcal{V}_{a_1}, C_2 = \mathcal{V}_{a_2} \cap C \mathcal{V}_{a_1}, \dots, C_n = \mathcal{V}_{a_n} \cap (C(\mathcal{V}_{a_1} \cup \dots \cup \mathcal{V}_{a_2} \cup \dots \cup \mathcal{V}_{a_{n-1}}))$. Каждое из множеств C_i является борелевским, а значит, измеримым. Обозначим через \vec{f}_e функцию, равную $\vec{f}(a_i)$ в C_i и равную $\vec{0}$ в дополнении к объединению \mathcal{V}_{a_i} . Функция \vec{f}_e — этажная, имеет компактный носитель и $\|\vec{f} - \vec{f}_e\| \leq \frac{\varepsilon}{\mu(\mathcal{V})}$, поэтому $\int^* \|\vec{f} - \vec{f}_e\| \leq \varepsilon$. Поскольку ε произвольно, \vec{f} интегрируема.

Остается теперь доказать, что для любой функции $\varphi \in \mathcal{C}(X)$ $\mu(\varphi) = \int \varphi d\mu$. Для доказательства нам понадобятся две леммы.

Лемма 1. Пусть A — измеримая часть X , φ — непрерывная функция с компактным носителем, мажорируемая характеристической функцией χ_A множества A . Тогда $\mu(\varphi) \leq \mu(A)$.

Доказательство леммы 1. Очевидно, $\int \varphi \leq \int \chi_A = \mu(A)$. Однако нас интересует не $\int \varphi$, а $\mu(\varphi)$, и мы еще не знаем, равны ли они между собой.

Пусть \mathcal{O} — некоторое открытое множество $\supset A$. Если мы покажем, что $\mu(\varphi) \leq \mu(\mathcal{O})$, то, переходя к точной нижней грани по всем \mathcal{O} , мы получим нужный нам результат. Все сводится к тому, чтобы исключить A и заменить его на открытое множество \mathcal{O} . Затем можно предположить, что $\bar{\mathcal{O}}$ компактно. Действительно, если \mathcal{U} является компактной окрестностью носителя (по предположению компактного) функции φ , то можно будет заменить множество \mathcal{O} на множество $\mathcal{O} \cap \mathcal{U}$, имеющее компактное замыкание. Наконец, можно считать функцию $\varphi \geq 0$ или в противном случае заменить ее на φ^+ . Пусть теперь K_n — компакт $\{x \in X; \varphi(x) \geq 1/n\}$, $n \geq 1$. Пусть α_n — непрерывная функция, $0 \leq \alpha_n \leq 1$, с носителем в \mathcal{O} , равная 1 на K_n (следствие 1 теоремы 11).

Положим $\varphi_n = \alpha_n \varphi$. Функция φ_n непрерывна, имеет носитель в \mathcal{O} , $0 \leq \varphi_n \leq 1$, а, значит, $\mu(\varphi_n) \leq \mu(\mathcal{O})$ (см. определение на стр. 490). Однако $|\varphi - \varphi_n| = (1 - \alpha_n)\varphi \leq 1/n$ и носитель этой функции лежит в $\bar{\mathcal{O}}$. Следовательно, $\mu(\varphi - \varphi_n) \leq \frac{1}{n} \|\mu\|_{\bar{\mathcal{O}}}$ и, значит,

$$\mu(\varphi) \leq \mu(\varphi_n) + \frac{1}{n} \|\mu\|_{\bar{\mathcal{O}}} \leq \mu(\mathcal{O}) + \frac{1}{n} \|\mu\|_{\bar{\mathcal{O}}}.$$

Так как n произвольно, то лемма доказана.

Лемма 2. Пусть \vec{g} — векторная этажная функция с компактным носителем, мажорируемая по норме числом M , причем число множеств, на которые разбито пространство X , равно N . Тогда, каково бы ни было число $\delta > 0$, существует такая расположимая¹⁾ непрерывная функция $\vec{\gamma}$ с компактным носителем, что:

- 1°) $\|\vec{g} - \vec{\gamma}\| \leq NM$;
- 2°) $\vec{\gamma} = \vec{g}$ всюду, кроме множества с мерой $\leq \delta$;
- 3°) если функция g является скалярной, а, следовательно, $\gamma \in \mathcal{C}(X)$, то $\int |\gamma - \mu(\gamma)| \leq \delta$;
- 4°) если $g \geq 0$, то $\gamma \geq 0$.

Доказательство леммы 2. Пусть $\vec{g} = \sum_{i=1}^N \vec{g}_i \chi_{A_i}$, где χ_{A_i} — характеристическая функция этажа A_i функций \vec{g} , на котором \vec{g} принимает постоянное значение $\vec{g}_i \neq \vec{0}$. Пусть K_i — компакт и \mathcal{O}_i — открытое множество, такие, что $K_i \subset A_i \subset \mathcal{O}_i$, $\mu(\mathcal{O}_i) - \mu(K_i) \leq \delta/NM'$, где $M' = \max(M, 1)$. Пусть γ_i — не-

¹⁾ Это понятие было введено в следствии 8 теоремы 11,

прерывная функция, $0 \leqslant \gamma_i \leqslant 1$, с носителем в \mathcal{O}_i , равная 1 на некоторой окрестности компакта K_i . Положим $\gamma = \sum_{i=1}^N \vec{g}_i \gamma_i$. Тогда:

1°) $0 \leqslant \chi_{A_i} \leqslant 1$, $0 \leqslant \gamma_i \leqslant 1$, а, следовательно $|\gamma_i - \chi_{A_i}| \leqslant 1$ и $\|\vec{g} - \vec{\gamma}\| \leqslant \sum_{i=1}^N \|g_i\| |\chi_{A_i} - \gamma_i| \leqslant NM$.

2°) На K_i $\gamma_i = \chi_{A_i} = 1$, на $\mathcal{O}_i - K_i$ $\gamma_i = \chi_{A_i} = 0$, а, следовательно, всюду, кроме множества $\mathcal{O}_i - K_i$ меры $\leqslant \delta/NM' \leqslant \delta/N$. Поэтому $\vec{g} = \vec{\gamma}$ всюду, кроме некоторого множества, содержащегося в $\bigcup_{i=1}^N (\mathcal{O}_i - K_i)$, меры $\leqslant N(\delta/N) = \delta$.

Замечание. Из 1°) и 2°) вытекает, что $\int \|\vec{g} - \vec{\gamma}\| \leqslant \delta NM$.

3°) Пусть $\vec{F} = \mathbb{C}$, а $\int \gamma_i$ и $\mu(\gamma_i)$ оба заключены между $\mu(K_i)$ и $\mu(\mathcal{O}_i)$, а, следовательно, $|\int \gamma_i - \mu(\gamma_i)| \leqslant \mu(\mathcal{O}_i) - \mu(K_i) \leqslant \frac{\delta}{NM'} \leqslant \frac{\delta}{NM}$. Тогда

$$\left| \int \gamma - \mu(\gamma) \right| \leqslant \sum_{i=1}^N |g_i| \left| \int \gamma_i - \mu(\gamma_i) \right| \leqslant M \frac{\delta}{NM} N = \delta. \quad (\text{IV}, 3; 28)$$

4°) Если $g_i \geqslant 0$, то $\gamma \geqslant 0$.

Доказательство теоремы. Пусть $\varphi \in \mathcal{C}(X)$. Зададим $\epsilon > 0$. Тогда найдутся такие этажные функции g и $h \geqslant 0$ с компактным носителем, что $|\varphi - g| \leqslant h$ и $\int h \leqslant \epsilon/2$ (см. замечание 2°) на стр. 524). Пусть M — максимум модулей величин 1, φ , g , h и N — максимальное число этажей функций g и h .

Применим последовательно к g и $h \geqslant 0$ лемму 2 при $\delta = \epsilon/[2(1 + 5MN)]$. Согласно этой лемме, найдутся две непрерывные функции α и β с компактными носителями, обладающие свойствами 1°) — 4°). Прежде всего,

$$\left| \int \varphi - \mu(\varphi) \right| \leqslant \left| \int \varphi - \int \alpha \right| + \left| \int \alpha - \mu(\alpha) \right| + |\mu(\alpha) - \mu(\varphi)|. \quad (\text{IV}, 3; 29)$$

Далее,

$$1°) \quad |\varphi - \alpha| \leqslant |\varphi - g| + |g - \alpha| \leqslant h + |g - \alpha|,$$

где $\int h \leqslant \epsilon/2$. Функция $g - \alpha$ равна нулю всюду, кроме множества меры $\leqslant \delta$, на котором она мажорируется по модулю

числом NM , а, значит, $\int |g - \alpha| \leq \delta NM$. Окончательно получаем

$$\left| \int \varphi - \int \alpha \right| \leq \frac{\varepsilon}{2} + \delta NM. \quad (\text{IV}, 3; 30)$$

$$2^{\circ}) \quad \left| \int \alpha - \mu(\alpha) \right| \leq \delta. \quad (\text{IV}, 3; 31)$$

$$3^{\circ}) \quad |\mu(\alpha) - \mu(\varphi)| = |\mu(\alpha - \varphi)| \leq \mu(|\alpha - \varphi|).$$

Из 1^o) следует, что $|\alpha - \varphi| \leq h + |g - \alpha| \leq \beta + |h - \beta| + |g - \alpha|$. Далее, функция $|g - \alpha|$ равна нулю всюду, кроме множества меры δ , где она мажорируется числом MN ; функция $|h - \beta|$ также равна нулю всюду, кроме множества меры δ , где она мажорируется числом NM . Поэтому $|\alpha - \varphi| - \beta$ является непрерывной функцией с компактным носителем, равной нулю всюду, кроме некоторого измеримого множества меры $\leq 2\delta$, на котором она мажорируется числом $2NM$.

Применяя к $(|\varphi - \alpha| - \beta)/2MN$ лемму 1, получаем:

$$|\mu(\alpha) - \mu(\varphi)| \leq 2\delta \cdot 2NM = 4NM\delta. \quad (\text{IV}, 3; 32)$$

Из (IV, 3; 30) — (IV, 3; 32), в силу соответствующего выбора δ , получаем:

$$\left| \int \varphi - \mu(\varphi) \right| \leq \frac{\varepsilon}{2} + \delta NM + \delta + 4\delta NM \leq \varepsilon. \quad (\text{IV}, 3; 33)$$

Так как ε произвольно, то $\int \varphi = \mu(\varphi)$.

З а м е ч а н и е. Смешно проводить доказательство для такого очевидного утверждения, но это пришлось сделать в качестве неизбежного наказания за избранное изложение теории Лебега (см. примечание на стр. 490).

Следствие 1. *Мера Радона $\mu \geq 0$ на X будет известной, если будут известны меры открытых множеств (или компактных множеств) ¹⁾.*

Прежде всего, формулы $\mu(K) = \inf_{\sigma \supset K} \mu(\sigma)$ и $\mu(\sigma) = \sup_{K \subset \sigma} \mu(K)$ говорят о том, что если известны меры открытых множеств, то известны меры компактов, и наоборот. Но тогда будут известны меры всех множеств, а, значит, интегралы всех функций. Из равенства $\mu(\varphi) = \int \varphi d\mu$ вытекает, что $\mu(\varphi)$ известна для $\varphi \in \mathcal{C}(X)$ и потому известна мера μ .

¹⁾ Мы не утверждаем, что можно произвольно выбирать меры открытых множеств. Мы говорим лишь, что две меры Радона ≥ 0 , дающие для открытых множеств одни и те же меры, совпадают.

Следствие 2. *Мера Радона $\mu \geq 0$ на вещественной прямой \mathbb{R} будет известной, если будут известны меры открытых интервалов (или замкнутых интервалов).*

Прежде всего из формул $\mu([a, b]) = \lim_{\epsilon \rightarrow 0} \mu([a - \epsilon, b + \epsilon])$ и $\mu([a, b]) = \lim_{\epsilon \rightarrow 0} \mu([a + \epsilon, b - \epsilon])$ вытекает, что один из этих случаев следует из другого.

Затем легко показать, что каждое открытое множество из \mathbb{R} является объединением счетного множества непересекающихся интервалов. Тогда меры открытых множеств известны, и остается лишь применить следствие 1.

Например, теория интеграла Римана (§ 1) позволила построить меру Радона dx на \mathbb{R} , которая для каждого интервала $[a, b]$ дает меру $b - a$. Настоящее следствие говорит о том, что в этом случае достаточно знать только это свойство.

Замечание. Пусть теперь ϕ — скалярная непрерывная функция с некомпактным носителем; но такая, что носитель μ и носитель ϕ имеют компактное пересечение K . Мы определили ранее $\mu(\phi)$ (теорема 16) даже в том случае, когда μ является векторной мерой. Легко видеть, что если $\mu \geq 0$, то функция ϕ является μ -измеримой и $\mu(\phi)$ совпадает с $\int \phi d\mu$.

В самом деле, если $\alpha \in \mathcal{C}(X)$ равна 1 на некоторой окрестности K , то $\mu(\phi)$ нами определялась как $\mu(\alpha\phi)$. Функция $\alpha\phi$ лежит в $\mathcal{C}(X)$, а, следовательно, интегрируема и, согласно теореме 30, $\mu(\alpha\phi) = \int (\alpha\phi) d\mu$. Функции $\alpha\phi$ и ϕ равны на носителе μ^1 , поэтому они μ -почти всюду равны между собой; следовательно, ϕ интегрируема и $\int \phi d\mu = \int (\alpha\phi) d\mu = \mu(\alpha\phi) = \mu(\phi)$.

Теорема 31. 1°) *Если f является ограниченной вещественной функцией ≥ 0 с компактным носителем на \mathbb{R} , то ее верхний интеграл Римана по dx не меньше ее верхнего интеграла Лебега.*

2°) *Если \vec{f} является функцией, определенной на вещественной прямой, со значениями в \vec{F} , интегрируемой по Риману относительно меры dx , то она тем более интегрируема по Лебегу и ее интеграл в смысле Лебега равен интегралу в смысле Римана.*

Доказательство. Докажем сначала 1°). Обозначим через $\int^{*R} f$ (соответственно $\int^{*L} f$) верхний интеграл в смысле

¹⁾ В каждой точке носителя μ либо $\alpha=1$ (если эта точка лежит в K), либо $\alpha=0$; поэтому всюду $\alpha\phi=\phi$. С другой стороны, мы видели на стр. 492 в замечании 3°), что дополнение к носителю имеет нулевую меру.

Римана (соответственно в смысле Лебега). Поскольку каждая ступенчатая функция является этажной функцией, то, вспоминая определение этих двух верхних интегралов, получаем:

$$\int^* L f \leq \int^* R f.$$

Z Неравенство не обязательно может перейти в равенство. Может иметь место знак $<$. Рассмотрим, например, функцию f , равную 1 во всех рациональных точках отрезка $[0, 1]$ и 0 в остальных точках. На стр. 492 мы видели, что $\int^* R f = 1$. Однако так как множество рациональных чисел имеет меру, равную нулю, то $\int^* L f = 0$.

Пусть теперь \vec{f} — некоторая интегрируемая по Риману функция и \vec{f}_n — аппроксимирующую последовательность, образованная из ступенчатых функций с компактными носителями, а, значит, тем более и из этажных функций. Тогда $\int^* L |\vec{f} - \vec{f}_n| \leq \int^* R |\vec{f} - \vec{f}_n|$ стремится к 0. Это значит, что \vec{f}_n образуют аппроксимирующую последовательность для интеграла Лебега, а, следовательно, \vec{f} интегрируема по Лебегу, и ее интеграл $\int^* L \vec{f}$, являющийся пределом интегралов $\int^* L \vec{f}_n = \int^* R \vec{f}_n$, равен $\int^* R \vec{f}$.

Z Напротив, указанная выше функция $f \geq 0$ интегрируема по Лебегу, имеет интеграл, равный нулю (поскольку она почти всюду равна нулю), однако не интегрируема по Риману.

Теорема 32. Каждая непрерывная на X функция со значениями в \vec{F} и компактным носителем интегрируема. Пусть \vec{f} — функция, определенная на X , со значениями в \vec{F} . Для того чтобы \vec{f} была интегрируемой, необходимо и достаточно, чтобы, каково бы ни было $\varepsilon > 0$, существовала такая функция \vec{g} , определенная на X , со значениями в \vec{F} , непрерывная и с компактным носителем, что $\int^* |\vec{f} - \vec{g}| \leq \varepsilon$, или необходимо и достаточно, чтобы существовала аппроксимирующая последовательность функции \vec{f} , образованная из непрерывных функций с компактным носителем. Эти функции могут даже быть выбраны разложимыми¹⁾.

¹⁾ См. определение в следствии 8 теоремы 11.

Доказательство. Интегрируемость непрерывных функций с компактным носителем была установлена в начале доказательства теоремы 30. Следовательно, всякая функция, имеющая аппроксимирующую последовательность, образованную из непрерывных функций с компактным носителем, интегрируема.

Пусть теперь \vec{f} — произвольная интегрируемая функция. Тогда существует такая этажная функция \vec{g} с компактным носителем, что $\int \|\vec{f} - \vec{g}\| \leq \varepsilon/2$. Затем, согласно лемме 2, из доказательства теоремы 30 следует существование такой разложимой непрерывной функции $\vec{\gamma}$ с компактным носителем, что $\int \|\vec{g} - \vec{\gamma}\| \leq \delta NM$. Выбирая $\delta = \varepsilon/2NM$, получим $\int \|\vec{g} - \vec{\gamma}\| \leq \varepsilon/2$, откуда $\int \|\vec{f} - \vec{\gamma}\| \leq \varepsilon$, что и требовалось доказать.

Замечание. Эта теорема в особенности удобна, когда \vec{F} является полем комплексных чисел, так как тогда она предстает в виде $\int f$ как предел величин $\mu(g)$, получающихся непосредственно из определения μ как линейной формы, заданной на пространстве $\mathcal{C}(X)$.

Интегрируемость и интегралы от функций, определенных почти всюду

Пусть сначала f является вещественной функцией ≥ 0 , определенной почти всюду на X . Продолжим ее произвольным образом до некоторой функции $\vec{f} \geq 0$, определенной всюду на X . Интеграл $\int^* \vec{f}$ не зависит от выбранного продолжения, поскольку любые два таких продолжения почти всюду равны между собой. Общее их значение называется верхним интегралом от f и обозначается через $\int^* f$. Точно так же, если \vec{f} является функцией, определенной почти всюду на X , со значениями в \vec{F} и \vec{f} — произвольное продолжение этой функции до функции, определенной на всем X , со значениями в \vec{F} , то интегрируемость \vec{f} и величина ее интеграла не зависят от выбранного продолжения.

Если \vec{f} интегрируема, то говорят, что \vec{f} интегрируема и общее значение интегралов $\int \vec{f}$ называется интегралом от \vec{f} и обозначается через $\int \vec{f}$.

§ 4. ТЕОРЕМА ЛЕБЕГА О СХОДИМОСТИ. ПРОСТРАНСТВО L^1

Теорема о сходимости, которую мы сейчас докажем, значительно сильнее всего того, что излагалось в связи с интегралом Римана, и именно благодаря этой теореме интеграл Лебега обладает существенным преимуществом по сравнению с интегралом Римана. Условия, лишь немногим более ограничительные, чем обычная сходимость функций, оказываются достаточными для обеспечения сходимости интегралов.

Вначале мы приведем теорему, показывающую, как можно перейти от простой сходимости к тому, что напоминает равномерную сходимость.

Теорема 33 (Егоров). *Пусть f_n — последовательность измеримых отображений X в метрическое пространство F . Предположим еще, что эта последовательность просто сходится μ -почти всюду к функции f . Тогда, каким бы ни был компакт K из X и каково бы ни было число $\delta > 0$, найдется такая компактная часть K_δ компакта K , что $\mu(K - K_\delta) \leq \delta$, при этом на K_δ отображения f_n равномерно сходятся к отображению f .*

Доказательство. Заранее ограничимся компактом K . Зададим фиксированное число ε' и обозначим через $A_{n, \varepsilon'}$ множество точек x из K , в которых $d(f_n(x), f(x)) > \varepsilon'$.

Докажем сначала, что это множество $A_{n, \varepsilon'}$ μ -измеримо. Поскольку каждая из функций f_n и f измерима (теорема 23), то отображение $x \rightarrow (f_n(x), f(x))$ является измеримым отображением K в произведение $F \times F$ (следствие 1 теоремы 23₂). Так как $(x, y) \rightarrow d(x, y)$ является непрерывным отображением $F \times F$ в \mathbb{R} (теорема 17₂ гл. II), то из теоремы 22 следует, что вещественная функция $x \rightarrow d(f_n(x), f(x))$ измерима. Множество $A_{n, \varepsilon'}$ является пересечением K и прообраза открытого множества $\{\varepsilon'\}, +\infty]$ из \mathbb{R} при этом отображении. По определению, это множество измеримо. Обозначим теперь через $\mathcal{A}_{m, \varepsilon'}$ объединение $\bigcup_{n \geq m} A_{n, \varepsilon'}$. Множество $\mathcal{A}_{m, \varepsilon'}$ измеримо. Точка $x \in K$ принадлежит этому объединению тогда и только тогда, когда существует такое целое $n \geq m$, что $d(f_n(x), f(x)) > \varepsilon'$. Последовательность $\mathcal{A}_{m, \varepsilon'}$ является убывающей последовательностью измеримых множеств. Если точка x принадлежит пересечению

$\bigcap_{m=0}^{\infty} \mathcal{A}_{m, \varepsilon'}$ множеств $\mathcal{A}_{m, \varepsilon'}$, то последовательность $f_n(x)$ не может сходиться к $f(x)$, а поскольку f_n сходятся почти всюду к f , то это пересечение имеет меру, равную нулю. Так как убывающая последовательность множеств $\mathcal{A}_{m, \varepsilon'}$ пересекается по множеству меры нуль, то из п. 5°) теоремы 19 следует, что для задан-

ногого $\delta' > 0$ существует такое целое m , что $\mu(\mathcal{A}_{m, \varepsilon'}) \leq \delta'^{-1}$. Обозначим через $\mathcal{A}_{\delta', \varepsilon'}$ это множество $\mathcal{A}_{m, \varepsilon'}$.

Таким образом, мы доказали следующий результат:

Каковы бы ни были числа $\delta' > 0$, существует такое целое число $m' = m'(\delta', \varepsilon')$ и такое множество $\mathcal{A}_{\delta', \varepsilon'} \subset K$ меры $\leq \delta'$, что для всех $n \geq m'(\delta', \varepsilon')$ и любого $x \in K - \mathcal{A}_{\delta', \varepsilon'}$ имеет место неравенство $d(f_n(x), f(x)) \leq \varepsilon'$.

Будем теперь брать последовательно $\delta' = \delta/2^{v+1}$, $\varepsilon' = 1/2^v$, $v = 1, 2, \dots$. Рассмотрим последовательность целых чисел $m_v = m'(\delta/2^{v+1}, 1/2^v)$ и множеств $\mathcal{A}_v = \mathcal{A}_{\delta/2^{v+1}, 1/2^v}$. Обозначим через \mathcal{A}_δ объединение множеств \mathcal{A}_v . Множество \mathcal{A}_δ является измеримым множеством меры $\leq \delta/2$, содержащимся в K . Каково бы ни было v , для $n \geq m_v$, теперь имеет место неравенство $d(f_n(x), f(x)) \leq 1/2^v$, лишь бы только x не принадлежал множеству \mathcal{A}_δ . Это означает, что на дополнении K'_δ этой части \mathcal{A}_δ по отношению в K последовательность f_n равномерно сходится к f . Имеем: $\mu(K'_\delta) = \mu(K - \mathcal{A}_\delta) \geq \mu(K) - \delta/2$. Теперь, по определению меры множеств, найдется такой компакт K_δ , содержащийся в K'_δ , что $\mu(K_\delta) \geq \mu(K) - \delta$ или $\mu(K - K_\delta) \leq \delta$ и на K_δ функции f_n равномерно сходятся к f , чем и заканчивается доказательство теоремы.

Замечания. 1°) Может показаться, что возможно найти такую часть K_δ компакта K , чтобы $\mu(K - K_\delta) = 0$ и чтобы на K_δ функции f_n сходились равномерно к f , но это не так. Рассмотрим, например, последовательность вещественных функций, определенных на \mathbb{R} по формуле

$$f_n(x) = \begin{cases} 1 & \text{для } 0 < x < 1/n, \\ 0 & \text{в остальных точках.} \end{cases} \quad (\text{IV, 4; 1})$$

Очевидно, при n , стремящемся к $+\infty$, f_n сходятся просто к 0. На компакте $[0, 1]$ при любом ε , $0 < \varepsilon < 1$, и любом n множество точек x , в которых $|f_n(x)| > \varepsilon$, имеет меру $1/n > 0$, что противоречит предположению.

2°) Если X не компактно, то заменить K на X невозможно. В самом деле, рассмотрим вещественные функции вещественной переменной и возьмем в качестве f_n функцию, определенную по формуле

$$f_n(x) = \begin{cases} 0 & \text{при } x < n, \\ 1 & \text{при } x \geq n. \end{cases} \quad (\text{IV, 4; 2})$$

¹⁾ Только здесь мы воспользовались тем фактом, что рассмотрение ведется на компакте K ; множества $\mathcal{A}_{m, \varepsilon'}$ имеют конечную меру, и теорема может быть применена.

Очевидно, что последовательность функций f_n просто сходится к функции 0. Однако, каково бы ни было ε , $0 < \varepsilon < 1$, и каково бы ни было n , множество точек x , в которых справедливо неравенство $|f_n(x)| > \varepsilon$, измеримо и имеет бесконечную меру.

Следствие (свойство Лузина). *Пусть f — отображение X в метризуемое сепарабельное пространство F . Если f μ -измеримо, то, каким бы ни был компакт K из X и каково бы ни было $\delta > 0$, найдется такой компакт $K_\delta \subset X$, что $\mu(K - K_\delta) \leq \delta$, а сужение f на K_δ является непрерывным отображением, и обратно.*

Доказательство. 1°) Предположим сначала, что f является этажной функцией. Эта функция принимает постоянные значения f_i на конечном числе непересекающихся множеств A_i , $i = 1, 2, \dots, N$, объединение которых равно X . Пусть K — некоторый компакт. Для каждого i обозначим через K_i такой компакт $\subset A_i \cap K$, что $\mu(K_i) \geq \mu(A_i \cap K) - \delta/N$. Тогда в качестве K_δ достаточно взять объединение K_i . Множество K_δ яв-

ляется компактом, $\mu(K - K_\delta) \leq \sum_{i=1}^N \frac{\delta}{N} = \frac{\delta}{N} N = \delta$ и, кроме

того, K_δ является объединением непересекающихся и замкнутых множеств K_i . Поскольку прообраз каждого замкнутого множества (и даже всего множества) из F при отображении f_δ является объединением K_i , а, следовательно, замкнут, то сужение f_δ функции f на K_δ непрерывно на K_δ .

Пусть теперь f — произвольное измеримое отображение. Согласно теореме 23₂, существует последовательность этажных функций f_n , сходящихся почти всюду к f при n , стремящимся к бесконечности. Согласно теореме Егорова, если мы выберем некоторую метрику на F , то найдется такой компакт K'_δ , $\mu(K - K'_\delta) \leq \delta/2$, на котором функции f_n сходятся равномерно к f . Поскольку f_n является этажной функцией, то из только что доказанного вытекает, что можно найти такой компакт $K_n \subset K'_\delta$, $\mu(K_n - K'_\delta) \leq \delta/2^{n+2}$, чтобы функции f_n были непрерывными на K_n . Положим теперь $K_\delta = \bigcap_{n=0}^{\infty} K_n$. Это множество является

некоторым компактом из K и $K - K_\delta \subset (K - K'_\delta) \cup \left(\bigcup_{n=0}^{\infty} (K'_\delta - K_n) \right)$, а, следовательно, $\mu(K - K_\delta) \leq \mu(K - K'_\delta) + \sum_{n=0}^{\infty} \mu(K'_\delta - K_n) \leq \frac{\delta}{2} +$

$+ \frac{\delta}{2} = \delta$. Кроме того, на K_δ все функции f_n непрерывны и равномерно сходятся к f , а, значит, функция f непрерывна на K_δ . Мы использовали теорему Егорова, в которой пространство F предполагается метрическим (для того чтобы иметь возможность говорить о равномерной сходимости). Однако окончательный результат (непрерывность f на K_δ), очевидно, не зависит от этой метрики.

2°) Обратно, предположим, что свойство Лузина для функции f выполнено. Пусть \mathcal{O} — некоторое открытое множество F и $B = f^{-1}(\mathcal{O})$. Поскольку f непрерывна на K_δ , то $B \cap K_\delta$ является открытым множеством компакта K_δ . Таким образом, это множество является пересечением некоторого открытого множества из X и компакта K_δ , а, значит, μ -измеримо. Далее,

$$\begin{aligned}\mu(B \cap K_\delta) &\leq \mu_*(B \cap K) \leq \mu^*(B \cap K) \leq \mu(B \cap K_\delta) + \mu^*(B \cap (K - K_\delta)) \leq \\ &\leq \mu(B \cap K_\delta) + \mu(K - K_\delta) \leq \mu(B \cap K_\delta) + \delta,\end{aligned}$$

а, следовательно, $\mu^*(B \cap K) - \mu_*(B \cap K) \leq \delta$. Поскольку δ произвольно, то $\mu^*(B \cap K) = \mu_*(B \cap K)$, что означает измеримость множества $B \cap K$. Так как полученное справедливо для любого компакта K и X является счетным объединением компактов, то $B = f^{-1}(\mathcal{O})$ измеримо, а f является измеримым отображением.

З а м е ч а н и е. Обратное утверждение 2°) справедливо для произвольного топологического пространства F (лишь бы, как всегда, X было счетным в бесконечности). Напротив, утверждение 1°) использовало теорему Егорова, из-за которой F предполагалось метризуемым, и теорему 23₂, в которой требуется, как это видно из ее доказательства, чтобы F было метризуемым и сепарабельным. Таким образом, свойство Лузина, вообще говоря, является более сильным, чем измеримость, однако оно ему эквивалентно, если метризуемое пространство F сепарабельно. Напомним, что на стр. 510 мы дали *предварительное определение измеримости*, предполагая метризуемое F сепарабельным. Теперь мы можем дать окончательное определение: *каково бы ни было топологическое пространство F , отображение f называется измеримым, если оно обладает свойством Лузина*. Теперь можно показать, что теоремы 21, 22 и следствия 1, 2, 3 теоремы 23₂ верны без каких-либо условий, налагаемых на F ; теоремы 23, 23₂ верны только для метризуемого пространства F ; теорема Егорова 33 предполагает F метрическим (для того чтобы можно было говорить о равномерной сходимости). Последнее следствие (Лузина) можно перефразировать, говоря, что если f измерима, то прообраз каждого борелевского множества

из F измерим, в то время как обратное утверждение справедливо лишь в том случае, когда метризуемое F сепарабельно. Учитывая сказанное в примечании на стр. 510, мы проявили осторожность, рассмотрев только предварительное определение, предполагающее F метризуемым и сепарабельным, для того чтобы высказать утверждение с предположениями, соответствующими окончательному определению.

Теорема 34 (частная теорема Лебега). Пусть X — локально компактное пространство, снабженное мерой Радона $\mu \geq 0$, и \vec{F} — пространство Банаха. Пусть \vec{f}_n — последовательность функций, определенных на X , со значениями в \vec{F} , имеющих носители в одном и том же компакте K и мажорируемых по норме одним и тем же числом $M \geq 0$. Если функции \vec{f}_n интегрируемы, измеримы¹⁾ и просто сходятся почти всюду к \vec{f} при n , стремящемся к $+\infty$, то \vec{f} интегрируема и измерима, $\int \vec{f}_n$ сходятся к $\int \vec{f}$ и даже $\int \|\vec{f} - \vec{f}_n\|$ сходится к 0.

Доказательство. Зададим $\varepsilon > 0$. Из предыдущей теоремы вытекает, что для числа $\delta = \varepsilon/4M$ можно найти такую измеримую часть K_δ компакта K , что $\mu(K - K_\delta) \leq \delta$, и такую, что на K_δ последовательность \vec{f}_n равномерно сходится к \vec{f} . Теперь можно выбрать такое целое число p , что для всех $n \geq p$ и любого $x \in K_\delta$ имеет место неравенство

$$\|\vec{f}(x) - \vec{f}_n(x)\| \leq \frac{\varepsilon}{2\mu(K)}. \quad (\text{IV}, 4; 3)$$

Отсюда следует, что для $n \geq p$ функция $\|\vec{f} - \vec{f}_n\|$ не превосходит функции, принимающей значение $\varepsilon/2\mu(K)$ на множестве K_δ , равной $2M$ на множестве $K - K_\delta$ и равной 0 на дополнении к компакту K . Поэтому для $n \geq p$ имеет место оценка

$$\int^* \|\vec{f} - \vec{f}_n\| \leq \frac{\varepsilon}{2\mu(K)} \mu(K) + 2M\delta = \varepsilon. \quad (\text{IV}, 4; 4)$$

Это неравенство показывает, что последовательность интегрируемых функций \vec{f}_n является аппроксимирующей последовательностью для функции \vec{f} . Отсюда следует, что \vec{f} интегрируема, $\int \vec{f}$ является пределом $\int \vec{f}_n$ и $\int \|\vec{f} - \vec{f}_n\|$ сходится к нулю.

¹⁾ Предположения излишни. В теореме 39, п. 1°) будет показано, что ограниченная функция с компактным носителем интегрируема тогда и только тогда, когда она измерима. Но мы пока еще этого не знаем.

Теорема 35 (общая теорема Лебега). Пусть X — локально компактное пространство, снабженное мерой Радона $\mu \geq 0$, и пусть \vec{F} — банахово пространство. Предположим, с одной стороны, что последовательность \vec{f}_n , состоящая из интегрируемых и измеримых функций¹⁾, просто сходится почти всюду к \vec{f} при n , стремящемся к $+\infty$, и, с другой стороны, что существует такая вещественная интегрируемая определенная на X функция $g \geq 0$, при которой имеет место неравенство $\|\vec{f}_n\| \leq g$. Тогда функция \vec{f} интегрируема и измерима, $\int \vec{f}_n$ сходится к $\int \vec{f}$ и $\int \|\vec{f} - \vec{f}_n\|$ сходится к нулю.

Естественно, общая теорема Лебега содержит частную теорему Лебега. Достаточно положить функцию g равной произведению M на характеристическую функцию компакта K . Кроме того, поскольку теорема 35 будет доказана независимо от теоремы 34, можно было бы привести теорему 34 как простое следствие теоремы 35. В целях простоты мы предпочли сначала доказать теорему 34.

Доказательство. Зададим $\varepsilon > 0$. Поскольку функция g интегрируема, то существует такая ограниченная по модулю числом $M > 0$ с компактным носителем K функция $\gamma \geq 0$, что имеет место неравенство

$$\int^* |g - \gamma| \leq \frac{\varepsilon}{8}. \quad (\text{IV}, 4; 5)$$

По теореме 33, для $\delta = \varepsilon/8M$ можно найти такое подмножество K_δ компакта K , что $\mu(K - K_\delta) \leq \delta$, на котором \vec{f}_n равномерно сходится к \vec{f} . Найдем теперь такое целое число p , чтобы при $n \geq p$ на K_δ имело место неравенство

$$\|\vec{f} - \vec{f}_n\| \leq \frac{\varepsilon}{2\mu(K)}. \quad (\text{IV}, 4; 6)$$

Очевидно, справедливо неравенство

$$\int^* \|\vec{f} - \vec{f}_n\| \leq \int_{K_\delta}^* \|\vec{f} - \vec{f}_n\| + \int_{CK_\delta}^* \|\vec{f} - \vec{f}_n\|. \quad (\text{IV}, 4; 7)$$

Первый интеграл справа является интегралом от функции, не превосходящей $\varepsilon/2\mu(K)$, и, следовательно, этот интеграл $\leq \varepsilon/2$ для всех $n \geq p$. Поскольку $\|\vec{f}_n\|$ мажорируется функцией g и

¹⁾ Здесь также требование измеримости чрезмерно. В теореме 39, п. 2° будет показано, что интегрируемая функция измерима.

$\vec{f}_n(x)$ стремится к $\vec{f}(x)$ почти для всех x , то $\|\vec{f}\|$ почти всюду мажорируется функцией g , а, значит, второй интеграл не пре-
восходит

$$\begin{aligned} \int_{CK_\delta} 2g &\leq \int_{CK_\delta} 2\gamma + \int^* 2|g - \gamma| = \\ &= \int_{K-K_\delta} 2\gamma + \int^* 2|g - \gamma| \leq 2M\delta + \frac{2\epsilon}{8} = \frac{\epsilon}{2}. \quad (\text{IV}, 4; 8) \end{aligned}$$

В итоге мы получаем неравенство

$$\int^* \|\vec{f} - \vec{f}_n\| \leq \epsilon \quad \text{для } n \geq p. \quad (\text{IV}, 4; 9)$$

Это неравенство говорит о том, что последовательность \vec{f}_n является аппроксимирующей последовательностью для функции \vec{f} , составленной из интегрируемых функций. Отсюда следует, что функция \vec{f} интегрируема, $\int \vec{f}_n$ сходится к $\int \vec{f}$ и $\int \|\vec{f} - \vec{f}_n\|$ сходится к нулю.

Примеры применений теоремы Лебега

1°) Рассмотрим интеграл

$$\int_{[0, 1]} x^n dx = \frac{1}{n+1}. \quad (\text{IV}, 4; 10)$$

При n , стремящемся к бесконечности, он сходится к нулю. Это можно было предвидеть с помощью теоремы Лебега. Действительно, x^n сходится к 0 при n , стремящемся к бесконечности, во всех точках полуоткрытого интервала $[0, 1]$, т. е. почти всюду на интервале интегрирования. Функции $|x^n|$, кроме того, мажорируются постоянной 1. Здесь мы можем воспользоваться частной теоремой Лебега. (Легко видеть, что x^n не сходятся равномерно к своему пределу.)

2°) Рассмотрим последовательность вещественных функций, определенных для $n \geq 1$ по формуле

$$f_n(x) = \begin{cases} n^\alpha & \text{для } 0 < x \leq \frac{1}{n}, \\ 0 & \text{вне этого интервала.} \end{cases} \quad (\text{IV}, 4; 11)$$

Эти функции просто сходятся к 0. Интеграл от f_n , очевидно, равен $n^{\alpha-1}$ и потому стремится к 0 при $\alpha < 1$.

Наименьшая мажоранта, общая для всех функций f_n , определяется формулой

$$g(x) = \begin{cases} \left[\frac{1}{x}\right]^\alpha & \text{для } 0 < x \leq 1^1), \\ 0 & \text{вне этого интервала.} \end{cases} \quad (\text{IV}, 4; 12)$$

При $\alpha < 1$ функция g мажорируется на $[0, 1]$ интегрируемой функцией $1/x^\alpha$ ²), а, следовательно, интегрируема. Это значит, что при $\alpha < 1$ применима теорема Лебега. Напротив, при $\alpha \geq 1$ теорему Лебега применить невозможно, поскольку мы имеем последовательность функций, просто сходящихся к 0, интеграл от которых к нулю не сходится. В самом деле, поскольку функция $\left[\frac{1}{x}\right]^\alpha \geq \frac{1}{x^\alpha} - 1$ не интегрируема на $[0, 1]$, то функции f_n не мажорируются одной и той же интегрируемой функцией g . Этот пример показывает, что ограничения, фигурирующие в формулировке теоремы Лебега, неизбежны. Однако имеется случай, для которого такие ограничения не нужны. Так будет тогда, когда функции f_n образуют возрастающую последовательность вещественных функций (см. теорему 36).

Замечание. Теорема Лебега сформулирована только для последовательности \vec{f}_n . Пусть теперь T — некоторое топологическое пространство, A — часть T и a — точка пространства T , принадлежащая замыканию A .

Предположим, что \vec{f}_t является интегрируемой на X функцией со значениями в \vec{F} , зависящей от параметра $t \in A$. Предположим еще, что при $t \in A$, стремящемся к a , функция \vec{f}_t просто сходится почти всюду к \vec{f} и что имеет место неравенство $\|\vec{f}_t\| \leq g$, где g — фиксированная интегрируемая функция ≥ 0 . Возникает вопрос: можно ли при этих условиях утверждать, что \vec{f} интегрируема и что $\int \vec{f}_t \in \vec{F}$ сходится к $\int \vec{f}$, когда t стремится к a , оставаясь в множестве A ? Это будет именно так, если T метризуемо. В самом деле, выберем в T некоторую метрику. Прежде всего, a , находясь в замыкании A , является пределом не менее одной последовательности точек из A (теорема 15 гл. II). При этом \vec{f}_{a_n} сходятся просто почти всюду к \vec{f} при n , стремящемся к бесконечности, и удовлетворяют неравенству

¹⁾ $[n]$ — целая часть n .

²⁾ Мы еще не доказали, что $1/x^\alpha$ интегрируема на $[0, 1]$ тогда и только тогда, когда $\alpha < 1$. Однако здесь речь идет только о примере, обоснование которого можно найти в обычных учебниках по математическому анализу.

$\|\vec{f}_{a_n}\| \leq g$. Из теоремы Лебега следует, что \vec{f} интегрируема. Кроме того, для любой последовательности этого вида $\int \vec{f}_{a_n}$ сходится к $\int \vec{f}$. Предположим, что $\int \vec{f}_t$ для значений $t \in A$, сходящихся к a , не сходится к $\int \vec{f}$. Тогда найдется такое $\varepsilon > 0$, что при любом $\eta > 0$ можно будет найти такую точку $t \in A$, для которой $d(t, a) \leq \eta$ и $\left\| \int \vec{f}_t - \int \vec{f} \right\| > \varepsilon$.

Задавая значения $\eta = 1, 1/2, 1/3, \dots, 1/n, \dots$, мы можем определить такую последовательность $a_1, a_2, a_3, \dots, a_n, \dots$ элементов из A , что $d(a_n, a) \leq 1/n$ и $\left\| \int \vec{f}_{a_n} - \int \vec{f} \right\| > \varepsilon$. Но это невозможно, так как $\int \vec{f}_{a_n}$ должен сходиться к $\int \vec{f}$, чем и доказывается наше утверждение.

Теорема 36 (теорема Фату). Пусть f_0, f_1, f_2, \dots — возрастающая последовательность функций ≥ 0 на X (с конечными или бесконечными значениями), т. е. $f_0 \leq f_1 \leq f_2 \leq \dots$. Если при этих условиях f является пределом функций f_n , то справедливо соотношение

$$\lim_{n \rightarrow \infty} \int^* f_n = \int^* f \leq +\infty. \quad (\text{IV}, 4; 13)$$

Доказательство. Очевидно, левая часть не превосходит правой; поэтому нам достаточно доказать противоположное неравенство. С этой целью мы сначала докажем лемму.

Лемма. Пусть f_0, f_1, f_2, \dots — возрастающая последовательность неотрицательных функций, определенных на X . Пусть g_0, g_1, g_2, \dots — такая последовательность интегрируемых функций (не обязательно возрастающая), мажорирующих f_0, f_1, f_2, \dots , что

$$\int g_n \leq \int^* f_n + \frac{\varepsilon}{2^{n+1}}. \quad (\text{IV}, 4; 14)$$

Тогда, если через γ_n обозначить верхнюю огибающую функций $g_0, g_1, g_2, \dots, g_n$, мы получим неравенство

$$\int \gamma_n \leq \int^* f_n + \varepsilon \left(1 - \frac{1}{2^{n+1}}\right). \quad (\text{IV}, 4; 15)$$

Это неравенство, очевидно, справедливо при $n = 0$. Докажем его справедливость при любых n по индукции, предполагая, что оно верно для $n - 1$.

Для любых вещественных чисел справедливо равенство

$$u + v = \sup(u, v) + \inf(u, v). \quad (\text{IV}, 4; 16)$$

Отсюда следует, что можно написать следующее равенство:

$$\begin{aligned} \int \gamma_n &= \int \sup(\gamma_{n-1}, g_n) = \\ &= \int \gamma_{n-1} + \int g_n - \int \inf(\gamma_{n-1}, g_n). \quad (\text{IV}, 4; 17) \end{aligned}$$

Рассмотрим члены правой части. В силу предположения индукции, первый интеграл не превосходит $\int^* f_{n-1} + \epsilon(1 - 1/2^n)$. Второй член, в силу неравенства (IV, 4; 14), не превосходит $\int^* f_n + \epsilon/2^{n+1}$.

С другой стороны, поскольку g_{n-1} мажорирует f_{n-1} , то γ_{n-1} мажорирует f_{n-1} . То же самое будет иметь место и для функции g_n , мажорирующей f_n и тем более f_{n-1} . Поэтому справедливо неравенство $\inf(\gamma_{n-1}, g_n) \geq f_{n-1}$, откуда

$$\int \inf(\gamma_{n-1}, g_n) \geq \int^* f_{n-1}. \quad (\text{IV}, 4; 18)$$

Окончательно получаем, что выражение $\int \gamma_n$ допускает такую оценку:

$$\begin{aligned} \int \gamma_n &\leq \int^* f_{n-1} + \epsilon \left(1 - \frac{1}{2^n}\right) + \\ &+ \int^* f_n + \frac{\epsilon}{2^{n+1}} - \int^* f_{n-1} = \int^* f_n + \epsilon \left(1 - \frac{1}{2^{n+1}}\right), \quad (\text{IV}, 4; 19) \end{aligned}$$

что позволяет закончить доказательство неравенства (IV, 4; 15) по индукции.

Перейдем теперь к доказательству теоремы Фату.

1°) Предположим сначала, что f ограничена сверху числом $M \geq 0$ и имеет компактный носитель K . Зададим $\epsilon > 0$. Для каждой функции f_n можно найти такую этажную ограниченную функцию $g_n \geq f_n$ с компактным носителем, для которой справедливо неравенство (IV, 4; 14). Вполне можно предполагать, что сама функция g_n ограничена числом M и имеет компактный носитель K . В противном случае ее можно заменить на 0 во всех точках x , не принадлежащих K , и на M во всех точках x , в которых $g_n(x) > M$. Образованная таким образом новая последовательность этажных функций будет удовлетворять неравенству (IV, 4; 14).

Обозначим через γ_n верхнюю огибающую функций g_0, g_1, \dots, g_n . Тогда последовательность γ_n будет возрастающей по-следовательностью этажных функций ≥ 0 . Они имеют предел γ ,

который не превосходит M и имеет носитель в K . Согласно частной теоремы Лебега (теорема 34), γ интегрируема и $\int \gamma_n$ сходится к $\int \gamma$. С другой стороны, согласно лемме, справедливо неравенство (IV, 4; 15).

Поскольку теперь, как это мы видели ранее, $\int \gamma_n$ сходится к $\int \gamma$, то имеет место неравенство

$$\int \gamma \leqslant \lim_{n \rightarrow \infty} \int^* f_n + \varepsilon, \quad (\text{IV}, 4; 20)$$

где предел существует, так как речь идет о некоторой возрастающей последовательности неотрицательных чисел. Однако поскольку для любого n справедливы неравенства $\gamma \geqslant \gamma_n \geqslant f_n$, то $\gamma \geqslant f$ и мы получаем соотношение

$$\int^* f \leqslant \lim_{n \rightarrow \infty} \int^* f_n + \varepsilon. \quad (\text{IV}, 4; 21)$$

Это соотношение верно при любом $\varepsilon > 0$; поэтому справедливо требуемое неравенство

$$\int^* f \leqslant \lim_{n \rightarrow \infty} \int^* f_n. \quad (\text{IV}, 4; 22)$$

2º) Предположим теперь, что $f \geqslant 0$ произвольна. Обозначим, как всегда, через $f_{M, K}$ функцию, равную f во всех точках x , принадлежащих K , в которых $f(x) \leqslant M$, равную нулю вне K и равную M в точках $x \in K$, в которых $f(x) > M$.

Построим такие же функции для f_n . Тогда последовательность функций $(f_n)_{M, K}$ является возрастающей последовательностью функций $\geqslant 0$, сходящейся к $f_{M, K}$. В силу только что доказанного,

$$\int^* f_{M, K} = \lim_{n \rightarrow \infty} \int^* (f_n)_{M, K} \leqslant \lim_{n \rightarrow \infty} \int^* f_n. \quad (\text{IV}, 4; 23)$$

Переходя теперь к точной верхней грани по всем значениям M и K , снова получаем (IV, 4; 22), чем и заканчивается доказательство теоремы Фату.

Следствие 0. Пусть A_0, A_1, A_2, \dots — некоторая возрастающая последовательность частей X , объединение которых дает A . Тогда

$$\mu^*(A) = \lim_{n \rightarrow \infty} \mu^*(A_n). \quad (\text{IV}, 4; 23_2)$$

Это соотношение вытекает непосредственно из теоремы Фату и соотношения (IV, 3; 18).

Следствие 1. Пусть f_n — последовательность интегрируемых функций произвольного знака, заданных на X , со значениями в $\bar{\mathbb{R}}$. Предположим, что эта последовательность монотонна (т. е. для всех n либо $f_n \geq f_{n-1}$, либо $f_n \leq f_{n-1}$) и сходится к f . Тогда

1°) если $\int^* |f| < +\infty$, то f интегрируема и $\int f_n$ сходится к $\int f$;

2°) если $\int^* |f| = +\infty$, то f не интегрируема и $\int f_n$ стремится к $\pm\infty$.

Доказательство. Докажем утверждение 1°). Предположим, для определенности, что последовательность f_n является возрастающей. Тогда имеет место неравенство $f_0 \leq f_n \leq f$. Отсюда следует, что $|f_n| \leq |f_0| + |f|$, а, значит,

$$\left| \int f_n \right| \leq \int^* |f| + \int |f_0|. \quad (\text{IV}, 4; 24)$$

Последовательность вещественных чисел $\int f_n$ оказалась возрастающей и ограниченной. Значит, она имеет предел. Поэтому для любого $\epsilon > 0$ можно найти такое целое число p , при котором для всех $m \geq n \geq p$ имеет место неравенство

$$\int f_m - \int f_n \leq \epsilon \quad \text{или} \quad \int (f_m - f_n) \leq \epsilon^1). \quad (\text{IV}, 4; 24_2)$$

Последовательность функций $f_m - f_n$ при фиксированном n и m , стремящемся к бесконечности, является возрастающей последовательностью неотрицательных функций. Поэтому мы можем применить теорему Фату и записать

$$\int^* (f - f_n) \leq \epsilon \quad \text{для } n \geq p. \quad (\text{IV}, 4; 25)$$

Поскольку речь идет о неотрицательной функции, то $\int^* (f - f_n) = \int^* |f - f_n|$. Отсюда получается, что последовательность f_n является аппроксимирующей последовательностью для f , состоящей из интегрируемых функций, чем и заканчивается доказательство нашего утверждения.

Для доказательства утверждения 2°) предположим, что $\int^* |f| = +\infty$. Если f интегрируема, то интегрируемой будет функция $|f|$, что противоречит равенству $\int^* |f| = +\infty$. Поэтому f не интегрируема.

¹⁾ Функция $f_m - f_n$ не определена в тех точках, где обе функции f_m и f_n равны бесконечности. Однако эти точки образуют множество нулевой меры, и на него можно не обращать внимания.

С другой стороны, если предположить, как ранее, что последовательность f_n возрастает, то мы получим, что $f_n - f_0$ образуют возрастающую последовательность неотрицательных функций, сходящихся к $f - f_0$. По теореме Фату в этом случае можно написать

$$\int^* (f - f_0) = \lim_{n \rightarrow \infty} \int f_n - \int f_0. \quad (\text{IV}, 4; 26)$$

Так как имеет место неравенство

$$|f| \leq |f_0| + |f - f_0|, \text{ то } +\infty = \int^* |f| \leq \int |f_0| + \int^* |f - f_0|, \quad (\text{IV}, 4; 27)$$

откуда $\int^* |f - f_0| = +\infty$, так что соотношение (IV, 4; 26) полностью доказывает наше утверждение.

Предыдущие результаты получены при некоторых предположениях относительно $\int^* |f|$. Можно наложить условия на величины $\int f_n$, и тогда мы получим следующие эквивалентные результаты.

Следствие 2. В условиях следствия 1:

1°) если величины $\left| \int f_n \right|$ ограничены, то f интегрируема и $\int f_n$ сходится к $\int f$;

2°) если величины $\left| \int f_n \right|$ не ограничены, то f не интегрируема и $\int f_n$ сходится к $\pm\infty$.

Из теоремы Фату как следствие вытекает теорема о счетной выпуклости, значение которой мы уже подчеркивали.

Следствие 3 (неравенство счетной выпуклости). Пусть $u_0, u_1, u_2, \dots, u_n, \dots$ — последовательность вещественных неотрицательных функций (с конечными или бесконечными значениями), определенных на X . Тогда имеет место неравенство счетной выпуклости:

$$\int^* \left(\sum_{n=0}^{\infty} u_n \right) \leq \sum_{n=0}^{\infty} \int^* u_n \leq +\infty. \quad (\text{IV}, 4; 28)$$

Доказательство. Для конечной суммы $\sum_{n=0}^m$ неравенство выпуклости справедливо (теорема 25). Правая часть при m , стремящемся к бесконечности, стремится к правой части

соотношения (IV, 4; 28) независимо от того, конечна она или нет.

С другой стороны, поскольку сумма $u_0 + u_1 + \dots + u_m$ возрастает вместе с m и сходится при m , стремящемся к бесконечности, к сумме $\sum_{n=0}^{\infty} u_n$, то по теореме Фату левая часть будет сходиться к левой части соотношения (IV, 4; 28), откуда и вытекает требуемый результат.

З а м е ч а н и е. Из теоремы следует, что если $\sum_{n=0}^{\infty} \int^* u_n < +\infty$, то $\int^* \left(\sum_{n=0}^{\infty} u_n \right) < +\infty$, а, следовательно, согласно теореме 26, ряд $\sum_{n=0}^{\infty} u_n$ почти всюду сходится.

Следствие 4 (о перестановочности знаков \sum и \int для рядов с положительными членами). *Если u_n образуют последовательность неотрицательных и интегрируемых функций (с конечными или бесконечными значениями) на X , то имеет место равенство*

$$\int^* \left(\sum_{n=0}^{\infty} u_n \right) = \sum_{n=0}^{\infty} \int u_n \leqslant +\infty. \quad (\text{IV, 4; 29})$$

Если же одна из частей этого равенства принимает конечное значение, то функция $S = \sum_{n=0}^{\infty} u_n$ интегрируема и

$$\int \sum_{n=0}^{\infty} u_n = \sum_{n=0}^{\infty} \int u_n. \quad (\text{IV, 4; 30})$$

В самом деле, равенство $\int S_m = \sum_{n=0}^m \int u_n$ (где $S_m = \sum_{n=0}^m u_n$) справедливо для любой конечной суммы. Остается лишь, учитывая теорему и следствия 1 и 2, перейти к пределу при m , стремящемся к бесконечности.

Мы видим, насколько полезна теорема Фату при интегрировании рядов, составленных из положительных функций. Напротив, теорема Лебега 35 в ее прямой форме мало приспособлена для операций над рядами, поскольку для того чтобы иметь возможность написать равенство (IV, 4; 30) для функций u_n в векторными значениями, необходимо знать, будет ли разность

$\vec{S} - \vec{S}_n$, т. е. остаток \vec{R}_n , мажорироваться по норме при произвольном n некоторой интегрируемой функцией $g \geq 0$, а это, вообще говоря, установить трудно. Тем не менее, имеет место весьма важная теорема.

Теорема 37. Пусть \vec{u}_n — последовательность функций, определенных μ -почти всюду на X , со значениями в банаховом пространстве \vec{F} . Тогда, если ряд $\sum_{n=0}^{\infty} \int^* \|\vec{u}_n\|$ сходится, то ряд $\sum_{n=0}^{\infty} \vec{u}_n(x)$ μ -почти всюду сходится абсолютно. Следовательно, этот ряд определяет μ -почти всюду некоторый вектор $\vec{S}(x)$. Если функции \vec{u}_n интегрируемы, то определенная таким путем μ -почти всюду на X функция \vec{S} со значениями в \vec{F} μ -интегрируема и имеет место равенство

$$\int \vec{S} = \sum_{n=0}^{\infty} \int \vec{u}_n, \quad \text{или} \quad \int \sum_{n=0}^{\infty} \vec{u}_n = \sum_{n=0}^{\infty} \int \vec{u}_n. \quad (\text{IV}, 4; 31)$$

Кроме того,

$$\left\| \int \vec{S} \right\| \leq \int \|\vec{S}\| \leq \sum_{n=0}^{\infty} \int \|\vec{u}_n\|, \quad (\text{IV}, 4; 32)$$

$$\lim_{m \rightarrow \infty} \int \|\vec{R}_m\| = 0, \quad (\text{IV}, 4; 33)$$

где \vec{R}_m — остаток $\sum_{n=m+1}^{\infty} \vec{u}_n$.

Доказательство. Сначала можно продолжить \vec{u}_n туда, где они не определены, и считать их определенными всюду.

Если мы предположим, что ряд $\sum_{n=0}^{\infty} \int^* \|\vec{u}_n\|$ сходящийся, и обозначим через f неотрицательную функцию, определенную равенством

$$f(x) = \sum_{n=0}^{\infty} \|u_n(x)\| \leq +\infty, \quad (\text{IV}, 4; 33_2)$$

то, согласно (IV, 4; 29), получим $\int^* f < +\infty$.

Теперь из теоремы 26 получаем, что функция f почти всюду конечна. Так как в точках x , в которых $f(x)$ конечна, ряд

$\sum_{n=0}^{\infty} \vec{u}_n(x)$ абсолютно сходится в \vec{F} , то, согласно теореме 55 гл. II, он тем более просто сходится. Обозначим через $\vec{S}(x)$ его сумму. Очевидно, \vec{S} является функцией, определенной почти всюду на X , со значениями в \vec{F} . Если теперь мы положим $\vec{S}_m = \sum_{n=0}^{\infty} \vec{u}_n$ и $\vec{R}_m = \vec{S} - \vec{S}_m$, то для \vec{R}_m получим оценку

$$\int^* \|\vec{R}_m\| \leq \int^* \left(\sum_{n=m+1}^{\infty} \|\vec{u}_n\| \right) = \sum_{n=m+1}^{\infty} \int^* \|\vec{u}_n\|. \quad (\text{IV}, 4; 34)$$

Поскольку ряд $\sum \int^* \|\vec{u}_n\|$ с неотрицательными членами сходится, то из последнего неравенства следует, что $\int^* \|\vec{R}_m\|$ сходится к 0 при m , стремящемся к бесконечности. Следовательно, если \vec{u}_n интегрируемы, то последовательность \vec{S}_m является аппроксимирующей последовательностью для функции \vec{S} , составленной из интегрируемых функций, а, значит, \vec{S} интегрируема и $\int \vec{S}_m$ стремится к $\int \vec{S}$, откуда и вытекают соотношения (IV, 4; 31). Нами даже доказано большее, поскольку мы знаем, что $\int \|\vec{S} - \vec{S}_m\| = \int \|\vec{R}_m\|$ стремится к 0 при m , стремящемся к бесконечности, т. е. справедливо равенство (IV, 4; 33). Соотношение (IV, 4; 32) является результатом применения соотношения (IV, 4; 34) к \vec{S} вместо \vec{R}_m .

Теорема 38. Пусть \vec{f} — функция, определенная почти всюду на X , со значениями в \vec{F} , и пусть \vec{f}_n — аппроксимирующая последовательность функции \vec{f} , т. е. такая последовательность функций, что $\int^* \|\vec{f} - \vec{f}_n\|$ сходится к 0 при n , стремящемся к бесконечности. Тогда из последовательности \vec{f}_n можно извлечь такую подпоследовательность, которая будет почти всюду сходиться к \vec{f} ¹).

В самом деле, шаг за шагом можно определить строго возрастающую последовательность целых чисел $p_0, p_1, p_2 \dots$,

1) Функции \vec{f}_n и \vec{f} не предполагались интегрируемыми, но именно этот случай будет наиболее интересным. Эта теорема (которую можно было бы доказать значительно раньше) дает интересное освещение поведения аппроксимирующей последовательности.

таких, что

$$\int^* \|\vec{f} - \vec{f}_{p_n}\| \leq \frac{1}{2^n}. \quad (\text{IV}, 4; 35)$$

При этом имеет место оценка

$$\int^* \|\vec{f}_{p_{n+1}} - \vec{f}_{p_n}\| \leq \frac{1}{2^{n-1}}, \quad (\text{IV}, 4; 36)$$

и, следовательно, если мы положим $\vec{u}_n = \vec{f}_{p_{n+1}} - \vec{f}_{p_n}$, то получим неравенство $\sum_{n=0}^{\infty} \int^* \|\vec{u}_n\| < +\infty$.

Из теоремы 37 теперь следует, что ряд $\sum_{n=0}^{\infty} \vec{u}_n$ является почти всюду сходящимся, другими словами, что подпоследовательность \vec{f}_{p_n} сходится почти всюду к некоторой предельной функции \vec{g} . Кроме того, из доказанного в теореме 37 следует, что $\int^* \|\vec{f}_{p_n} - \vec{g}\|$ сходится к 0, а поскольку $\int^* \|\vec{f}_{p_n} - \vec{f}\|$ также сходится к 0, то интеграл $\int^* \|\vec{f} - \vec{g}\|$ равен нулю. Значит, \vec{f} и \vec{g} почти всюду равны, а это означает, что последовательность \vec{f}_{p_n} почти всюду сходится к \vec{f} .

Замечание. Этот результат весьма важен. Он показывает, что из последовательности \vec{f}_n или из любой ее подпоследовательности можно извлечь частичную подпоследовательность, сходящуюся почти всюду к \vec{f} . Однако будет неверным считать, что сама последовательность \vec{f}_n необходимо сходится почти всюду к \vec{f} .

В самом деле, рассмотрим следующий пример. Возьмем в качестве X тригонометрическую окружность, а в качестве меры μ выберем меру $d\theta$. Рассмотрим последовательность точек a_n , определенных по формуле $a_n = 1 + 1/2 + 1/3 + \dots + 1/n$, $n \geq 1$, и обозначим через f_n функцию, равную 1 в интервале $[a_n, a_{n+1}]$ и равную 0 вне его. Так как $\int f_n = 1/(n+1)$, то $\int |f_n|$ сходится к 0 при n , стремящемся к бесконечности. Функции f_n образуют аппроксимирующую последовательность для функции 0. Однако, как легко видеть, последовательность f_n не сходится ни при каком значении x к 0. В самом деле, поскольку ряд $\sum_{n=1}^{\infty} 1/n$ расходится, то точки a_n проходят бесконечное число раз

по окружности и, следовательно, каждая точка x для бесконечного числа значений n содержится в интервале $[a_n, a_{n+1}]$, а значит, для бесконечного числа значений n $f_n(x) = 1$. Теорема же, напротив, утверждает, что из последовательности f_n можно выделить подпоследовательность, почти всюду сходящуюся к 0. В самом деле, определим последовательность целых чисел p_n таким образом, чтобы интервал $[a_{p_n}, a_{p_n+1}]$ содержал фиксированную точку c . Поскольку длина этого интервала стремится к 0, последовательность f_{p_n} сходится к нулю в каждой точке $\neq c$, т. е. всюду, кроме одной точки, а, значит, почти всюду.

Характеристика интегрируемых функций.

Интегрируемость и измеримость

Теорема 39. Пусть \tilde{f} — отображение локально компактного пространства X , снабженного мерой Радона $\mu \geq 0$, в банахово пространство \tilde{F} .

1°) Если функция \tilde{f} ограничена и имеет компактный носитель, то, для того чтобы она была интегрируемой, необходимо и достаточно, чтобы она была измеримой.

2°) Если \tilde{f} произвольна, то, для того чтобы она была интегрируемой, необходимо и достаточно, чтобы она была измеримой и чтобы верхний интеграл $\int^* \|\tilde{f}\|$ был конечен.

Доказательство. Заметим, прежде всего, что первая часть теоремы является частным случаем второй части.

Предположим сначала, что \tilde{f} интегрируема. Тогда, в силу теоремы 29, $\|\tilde{f}\|$ также интегрируема и, следовательно, согласно следствию 4 из теоремы 29, интеграл $\int^* \|\tilde{f}\| = \int \|\tilde{f}\|$ необходимо конечен. С другой стороны, если \tilde{f}_n является аппроксимирующей последовательностью для функции \tilde{f} , составленной из этажных функций, то из теоремы 38 следует, что из последовательности \tilde{f}_n можно извлечь подпоследовательность, сходящуюся почти всюду к \tilde{f} . Таким образом, функция \tilde{f} является почти всюду пределом некоторой последовательности измеримых функций и, следовательно, по теореме 23 измерима.

Докажем теперь обратное утверждение. Предположим, что \tilde{f} измерима и что интеграл $\int^* \|\tilde{f}\|$ конечен.

1°) Предположим сначала, что \tilde{f} ограничена по норме числом $M \geq 0$ и имеет компактный носитель K .

Мы знаем, что тогда \vec{f} будет почти всюду пределом некоторой последовательности этажных функций \vec{f}_n (теорема 23₂). Заменяя \vec{f}_n на $\vec{f}_n \frac{M}{\|\vec{f}_n\|}$ на тех этажах, где норма $> M$ (это не

нарушает почти всюду сходимости к \vec{f}), можно предполагать, что все $\|\vec{f}_n\|$ ограничены числом M . Кроме того, их носители можно считать лежащими в K , ибо в противном случае их можно умножить на характеристическую функцию компакта K . Из частной теоремы Лебега 34 теперь вытекает, что \vec{f} интегрируема, а \vec{f}_n образуют ее аппроксимирующую последовательность.

2°) Перейдем теперь к общему случаю произвольной функции \vec{f} . Пусть K_0, K_1, K_2, \dots — возрастающая последовательность компактов, объединение которых составляет X .

Обозначим через \vec{f}_n функцию, равную \vec{f} на множестве H_n точек $x \in K_n$, в которых $\|\vec{f}(x)\| \leq n$, и равную $\vec{0}$ вне его. Множество H_n является пересечением компакта K_n и прообраза интервала $[0, +n]$, определяемого измеримой функцией $\|\vec{f}\|$ (композиция функции \vec{f} и нормы; см. теорему 22). Следовательно, H_n как пересечение измеримых множеств измеримо. Функция \vec{f}_n является произведением \vec{f}_{H_n} и, следовательно, также измерима (следствие 3 теоремы 23₂). Поскольку она ограничена и имеет компактный носитель, она интегрируема. Мы знаем также, что функции $\|\vec{f}_n\|$ измеримы. Эти функции образуют возрастающую последовательность неотрицательных функций и их предел равен $\|\vec{f}\|$. В силу предположения $\int^* \|\vec{f}\| < +\infty$, из следствия 1 теоремы Фату получаем, что $\|\vec{f}\|$ интегрируема. Но тогда функции \vec{f}_n сходятся к \vec{f} при n , стремящемся к бесконечности, и мажорируются по норме интегрируемой функцией $\|\vec{f}\| \geq 0$. По теореме Лебега 35 функция \vec{f} интегрируема.

Следствие 1. *Всякая ограниченная измеримая функция \vec{f} , равная нулю вне некоторого множества Y конечной внешней меры, интегрируема.*

В самом деле,

$$\int^* \|\vec{f}\| \leq \|\vec{f}\| \mu^*(Y) < +\infty.$$

Следствие 2. *Если $\|\mu\| < +\infty$, то всякая μ -измеримая и*

ограниченная функция \vec{f} , и в частности всякая непрерывная и ограниченная функция, интегрируема.

В самом деле,

$$\int^* \|\vec{f}\| \leq \|\vec{f}\| \|\mu\|.$$

Следствие 3. Для того чтобы \vec{f} была интегрируемой, необходимо и достаточно, чтобы она была измеримой и верхние интегралы от функций $\|\vec{f}\|_{M, K}$, определенные на стр. 521, были ограниченными независимо от выбора M и K . Достаточно даже, чтобы ограниченность имела место для последовательности M_n , стремящейся к бесконечности, и для возрастающей подпоследовательности компактов K_n , объединение которых равно X .

Это утверждение вытекает из того, что при этих условиях $\int^* \|\vec{f}\|$ является пределом величин $\int^* \|\vec{f}\|_{M_n, K_n}$.

Следствие 4 (выпуклость и среднее значение). Пусть \vec{C} — выпуклое замкнутое множество из \vec{F} . Если \vec{f} интегрируема на измеримом множестве A конечной меры $\neq 0$ и если $\vec{f}(A) \subset \vec{C}$, то ее среднее значение на A : $\frac{1}{\mu(A)} \int_A \vec{f} d\mu$ принадлежит \vec{C} .

Доказательство. Предположим раз и навсегда, что $\vec{0} \in \vec{C}$. Будем продолжать функцию \vec{f} нулем в $C \setminus A$ так, чтобы имело место соотношение $\vec{f}(X) \subset \vec{C}$.

Предположим сначала, что \vec{f} — этажная функция. Ее можно записать в виде $\vec{f} = \sum_{i=1}^N \vec{g}_i \chi_{A_i}$, где $\vec{g}_i \in \vec{C}$ и χ_{A_i} — характеристические функции множеств A_i . В противоположность тому, как это делается в других случаях, мы будем включать слагаемое $\vec{0} \cdot \chi_{A_0}$ там, где \vec{f} принимает значение $\vec{0}$, с тем расчетом, чтобы можно было считать A_i непересекающимися и $\bigcup_{i=1}^n A_i = X$. Тогда, согласно барицентрическому свойству выпуклости, вектор

$$\frac{1}{\mu(A)} \int_A \vec{f} d\mu = \frac{1}{\mu(A)} \sum_{i=1}^N \vec{g}_i \mu(A_i \cap A)$$

лежит в \vec{C} , поскольку $\sum_{i=1}^N \mu(A_i \cap A) = \mu(A)$.

Предположим теперь, что \vec{f} ограничена и имеет компактный носитель. Часть 1°) доказательства теоремы 39 дает нам возможность с помощью теоремы 23₂ и ее замечания 2°) построить аппроксимирующую последовательность для \vec{f} , образованную из этажных функций \vec{f}_n со значениями в $\vec{f}(X) \subset \vec{C}$. Согласно предыдущему, $\frac{1}{\mu(A)} \int_A \vec{f}_n d\mu$ будет лежать в \vec{C} , а, значит, в силу замкнутости \vec{C} , там же будет находиться и $\frac{1}{\mu(A)} \int_A \vec{f} d\mu$.

Пусть, наконец, \vec{f} — произвольная функция. Часть 2°) доказательства теоремы 39 дает нам аппроксимирующую последовательность для функции \vec{f} , образованную из ограниченных с компактным носителем функций \vec{f}_n , принимающих свои значения в $\vec{C} \cup \{\vec{0}\} = \vec{C}$. Новый переход к пределу дает необходимый результат. Доказательство закончено, если $\vec{0} \in \vec{C}$.

Пусть, в общем случае, \vec{c} — некоторая точка \vec{C} . Тогда множество $\vec{C} - \vec{c}$ замкнуто, выпукло и содержит $\vec{0}$. Функция $\vec{f} - \vec{c}$ принимает значения в этом выпуклом множестве. Поэтому

$$\frac{1}{\mu(A)} \int_A \vec{f} d\mu = \vec{c} + \frac{1}{\mu(A)} \int_A (\vec{f} - \vec{c}) d\mu \in \vec{c} + (\vec{C} - \vec{c}) = \vec{C},$$

что и требовалось доказать.

Замечание. Ограничение $\mu(A) \neq 0$ можно убрать, если написать

$$\int_A \vec{f} d\mu \in \mu(A) \vec{C}.$$

Замечания. 1°) Как мы уже говорили на стр. 516, функции, встречающиеся на практике, обычно измеримы. В этих случаях интегрируемость \vec{f} эквивалентна неравенству $\int^* \|\vec{f}\| < \infty$.

2°) Пусть f — измеримая неотрицательная функция, определенная на X . Если $\int^* f$ конечен, то эта функция интегрируема (согласно следствию 4 теоремы 29) имеет место равенство $\int^* f = \int f$. Если $\int^* f = +\infty$, то $\int f$ смысла не имеет. Однако если f измерима ≥ 0 и $\int^* f = +\infty$, то теперь будет разумным

писать $\int f = \int^* f = +\infty$ (однако это не означает, что f интегрируем!). Таким образом, если f измерима и ≥ 0 , то во всех случаях $\int^* f = \int f$. Впрочем, для измеримого подмножества A множества X в случае $\mu^*(A) = +\infty$ мы уже писали $\mu(A) = +\infty$.

Теория интегрирования, основанная на свойствах непрерывных и полунепрерывных снизу функций

В предыдущем изложении теории продолжения Лебега мы отвели существенную роль мере множеств и этажным функциям. Этого можно было бы полностью избежать, воспользовавшись полунепрерывными снизу функциями со значениями в $\bar{\mathbb{R}}$ (см. определение на стр. 85).

Прежде всего, непосредственно видно, что верхняя или нижняя огибающего конечного числа полунепрерывных снизу функций также является полунепрерывной снизу функцией и что верхняя огибающая (но не нижняя огибающая) произвольного конечного или бесконечного семейства полунепрерывных снизу функций является полунепрерывной снизу функцией. В частности, верхняя огибающая семейства непрерывных функций является полунепрерывной снизу функцией. [Полунепрерывные сверху функции обладают аналогичными свойствами. Однако на этот раз полунепрерывной сверху является нижняя огибающая произвольного семейства полунепрерывных сверху функций (в частности, непрерывных функций). Впрочем, f полунепрерывна сверху тогда и только тогда, когда $-f$ полунепрерывана снизу.] Характеристическая функция некоторой части A множества X полунепрерывна снизу (соответственно сверху) тогда и только тогда, когда эта часть открыта (соответственно замкнута). Свойства открытых и замкнутых множеств, относящиеся к объединению или пересечению конечного или произвольного числа множеств, являются выражением свойств полунепрерывных снизу или сверху функций относительно верхних или нижних огибающих конечного или произвольного семейства функций. Сумма конечного числа полунепрерывных сверху или снизу неотрицательных функций со значениями в $\bar{\mathbb{R}}$ (но не в \mathbb{R} , поскольку тогда сумма может оказаться неопределенной или равной $+\infty, -\infty$) обладает тем же самым свойством. Сумма бесконечного семейства полунепрерывных снизу (в частности, непрерывных) неотрицательных функций полунепрерывна снизу, поскольку она является верхней огибающей конечных частных сумм.

Полунепрерывная снизу (или сверху) функция f является борелевской, а, значит, измеримой по любой мере Радона ≥ 0 .

В самом деле, при любом a множество $f^{-1}([a, +\infty])$ открыто. Поэтому, составляя разность, находим, что при любых a и b множество $f^{-1}([a, b])$ будет борелевским, а вместе с ним будет борелевским и множество $f^{-1}([a, b]) = \bigcup_{n=1,2,\dots} f^{-1}([a, b - 1/n]).$

Однако каждое открытое множество \mathcal{O} из $\bar{\mathbb{R}}$ является объединением счетного множества открытых непересекающихся интервалов. Поэтому $f^{-1}(\mathcal{O})$ будет борелевским множеством, а вместе с этим функция f будет борелевской.

Теорема 39₂. *Верхний интеграл от полунепрерывной снизу неотрицательной функции $f \geq 0$ (с конечными значениями или нет) является точной верхней границей интегралов от неотрицательных функций $\varphi \in \mathcal{C}(X)$, которые минорируют f .*

Доказательство. Пусть f — полунепрерывная снизу функция, принимающая значения в $\bar{\mathbb{R}}_+$. Для каждой функции $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq f$, имеет место неравенство $\mu(\varphi) = \int \varphi \leq \int^* f$.

Мы должны показать, что для любого $A < \int^* f$ можно найти такую функцию $\varphi \in \mathcal{C}(X)$, что $0 \leq \varphi \leq f$ и $\mu(\varphi) \geq A$.

Прежде всего, из определения верхнего интеграла следует, что если $A < B < \int^* f$, то найдутся такое число $M \geq 0$ и такой компакт $K \subset X$, что $\int^* f_{M,K} \geq B$ ¹⁾. Введем открытую окрестность \mathcal{V} компакта K , ее характеристическую функцию $\chi_{\mathcal{V}}$ и функцию $g = \inf(f, M\chi_{\mathcal{V}})$. Так как $f_{M,K} \leq g \leq f$, то $\int^* g \geq B$. Функция g , кроме того, полунепрерывна снизу как нижняя огибающая двух таких функций. Теперь функция g ограничена и имеет компактный носитель. Пусть $\chi_{k,n}$ — характеристическая функция множества $\mathcal{O}_{k,n} = g^{-1}([kM/n, M])$. Это множество открыто, поскольку g полунепрерывна снизу. Положим теперь

$$g_n = \sum_{k=1}^{n-1} \frac{M}{n} \chi_{k,n}.$$

Функция g_n является этажной функцией, принимающей значения kM/n на множествах $g^{-1}([kM/n, (k+1)M/n])$ и равной 0 вне этих множеств. Таким образом, $g_n \leq g$ и $g - g_n \leq M/n$. При n , стремящемся к бесконечности, функции g_n , имеющие носители в фиксированном компакте \mathcal{V} , сходятся равномерно к g , а, значит, $\int g_n$ сходятся к $\int g$. Следовательно, существует такое целое число n , что $\int g_n > A$. Зафиксировав

¹⁾ Смотри определение $f_{M,K}$ на стр. 521.

такое n , положим $\delta = \int g_n - A > 0$. Для каждого k найдем такие непрерывные функции $\varphi_k \geq 0$ с компактными носителями в $\mathcal{O}_{k,n}$, что $\mu(\varphi_k) \geq \mu(\mathcal{O}_{k,n}) - \delta/M$. Функция $\varphi = \sum_{k=1}^{n-1} \frac{M}{n} \varphi_k$ неотрицательна, имеет компактный носитель, мажорируема функцией g_n (а значит, мажорируема функцией g и, следовательно, функцией f) и такова, что

$$\mu(\varphi) \geq \sum_{k=1}^{n-1} \frac{M}{n} \mu(\mathcal{O}_{k,n}) - \sum_{k=1}^{n-1} \frac{M}{n} \frac{\delta}{M} = \int g_n - \frac{n-1}{n} \delta > A.$$

Замечание. Этот результат для полунепрерывных снизу неотрицательных функций является весьма частным. Если, например, взять характеристическую функцию f множества иррациональных чисел в \mathbb{R} , то ее верхний интеграл равен $+\infty$. Однако для любой функции $\varphi \in \mathcal{C}(X)$, такой, что $0 \leq \varphi \leq f$, имеет место равенство $\mu(\varphi) = 0$.

Для произвольной неотрицательной функции имеет место

Теорема 39₃. *Верхний интеграл неотрицательной функции является точной нижней гранью верхних интегралов от полунепрерывных снизу функций, мажорирующих данную функцию.*

Доказательство. Зададим $\varepsilon > 0$. Пусть вначале функция f ограничена и имеет компактный носитель. Пусть g — такая этажная функция с компактным носителем, что $g \geq f$ и $\int g \leq \int^* f + \varepsilon/2$. Функция g имеет вид $\sum_{i=1}^N g_i \chi_{A_i}$, где g_i — постоянные > 0 и A_i — измеримые множества конечной меры. Пусть теперь $M = \max_{1 \leq i \leq N} |g_i|$. Найдем для каждого i такое открытое множество $\mathcal{O}_i \supset A_i$, чтобы $\mu(\mathcal{O}_i) \leq \mu(A_i) + \varepsilon/(2NM)$, и положим $h = \sum_{i=1}^N g_i \chi_{\mathcal{O}_i}$. Тогда, так как \mathcal{O}_i — открытое множество, каждая функция $\chi_{\mathcal{O}_i}$ полунепрерывна снизу, а, следовательно, h также полунепрерывна снизу и имеют место неравенства:

$$f \leq g \leq h \quad \text{и} \quad \int h \leq \sum_{i=1}^N g_i \left(\mu(A_i) + \frac{\varepsilon}{2NM} \right) \leq \int g + \frac{\varepsilon}{2} \leq \int^* f + \varepsilon,$$

что и доказывает теорему для этого случая.

Пусть теперь функция f произвольна. Рассмотрим функцию f_{n,K_n} (K_n — возрастающая последовательность компактов,

объединение которых равно X ; определение f_M , к см. на стр. 521). Эта функция ограничена и имеет компактный носитель. Следовательно, существует полунепрерывная снизу функция g_n , такая, что

$$f_{n, K_n} \leq g_n \quad \text{и} \quad \int g_n \leq \int^* f_{n, K_n} + \frac{\epsilon}{2^{n+1}}.$$

Если теперь мы положим $\gamma_n = \sup(g_0, g_1, \dots, g_n)$, то из леммы, использованной при доказательстве теоремы Фату 36, вытекает неравенство

$$\int \gamma_n \leq \int^* f_{n, K_n} + \epsilon \left(1 - \frac{1}{2^{n+1}}\right).$$

Последовательность γ_n является возрастающей, каждая из функций γ_n полунепрерывна снизу, а их предел γ при n , стремящемся к бесконечности, являющийся также их верхней огибающей, полунепрерывен снизу. Для любого n имеет место неравенство $f_{n, K_n} \leq \gamma_n \leq \gamma$, а значит, $f \leq \gamma$. С другой стороны, по теореме Фату

$$\int^* \gamma = \lim_{n \rightarrow \infty} \int^* \gamma_n \leq \lim_{n \rightarrow \infty} \int^* f_{n, K_n} + \epsilon = \int^* f + \epsilon,$$

чем и заканчивается доказательство теоремы в общем случае.

Замечания. 1°) Заменить полунепрерывные снизу функции на непрерывные функции, очевидно, невозможно. Например, если f является характеристической функцией рациональных чисел на \mathbb{R} , то ее верхний интеграл равен нулю, однако любая мажорирующая ее функция ≥ 1 и имеет бесконечный интеграл.

2°) Как мы это видели в примечании на стр. 521, здесь мы получаем верхний интеграл от некоторой неотрицательной функции с помощью двух предельных переходов, исходя из функции $\mu(\varphi)$:

$$\int^* f = \inf_{(g \geq f, g \text{ полунепр. снизу})} (\sup_{\substack{0 \leq \varphi \leq g \\ \varphi \in \mathcal{C}(X)}} \mu(\varphi)). \quad (\text{IV}, 4; 40)$$

По-видимому, введение верхнего интеграла по этой формуле, взятой за *определение*, было бы наилучшим и наиболее коротким. Однако такой метод будет более абстрактным. Затем можно было бы *определить* непосредственно скалярные интегрируемые функции по теореме 32, т. е. с применением аппроксимирующей последовательности, составленной функциями из $\mathcal{C}(X)$. Тогда интеграл $\int \varphi$ был бы, *по определению*, равен $\mu(\varphi)$ (что позволило бы избавиться от «идиотской» теоремы 30). Меру множеств можно определить, после определения интеграла от функций, как интеграл от их характеристической функции. Ко-

нечно, фундаментальные теоремы останутся теми же самыми, а некоторые доказательства окажутся более сложными, чем в принятом нами изложении.

Для того чтобы определить интеграл от векторных функций, введем пространство $\mathcal{C}(X) \otimes \vec{F}$ непрерывных *разложимых* функций с компактным носителем, т. е. функций вида $\vec{g} = \sum_{i=1}^N \vec{g}_i \varphi_i$, где \vec{g}_i — постоянные из \vec{F} и $\varphi_i \in \mathcal{C}(X)$. Положим, по определению, $\int \vec{g} = \sum_{i=1}^N \vec{g}_i \mu(\varphi_i)$ (что полностью совпадает с соответствующим интегралом в принятой нами теории). Можно убедиться в том, что результат совершенно не зависит от частного представления непрерывной и разложимой функции \vec{g} (как мы уже это делали при введении интеграла от этажных функций). Так снова определяются те же самые интегрируемые функции, что и ранее, и тот же самый интеграл. В самом деле, каждая интегрируемая функция в этой новой теории имеет аппроксимирующую последовательность, образованную из интегрируемых функций в прежнем смысле (функций из $\mathcal{C}(X) \otimes \vec{F}$), и тот же самый интеграл. Обратное утверждение справедливо в силу теоремы 32.

Для полунепрерывных снизу неотрицательных функций имеется обобщение теоремы Фату, справедливое для *несчетных* семейств:

Теорема 39₄. Пусть $(f_i)_{i \in I}$ — произвольное семейство полуунепрерывных снизу неотрицательных функций, и пусть f — их верхняя огибающая. Предположим, кроме того, что семейство является «возрастающим» в следующем смысле: какими бы ни были индексы $i, j \in I$, существует такое $k \in I$, что $f_k \geq \sup (f_i, f_j)$. Тогда $\int^* f = \sup_{i \in I} \int^* f_i$.

Доказательство. Совершенно очевидно, что $\int^* f \geq \sup_{i \in I} \int^* f_i$. Пусть теперь A — произвольное число $< \int^* f$. Согласно теореме 39₂, можно найти такую функцию $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq f$, что $\int \varphi > A$. Пусть K — носитель функции φ и $\delta = (\int \varphi - A)/\mu(K)$. Для каждого $a \in K$ можно выбрать такой индекс $i_a \in I$, при котором $f_{i_a}(a) \geq f(a) - \delta/2 \geq \varphi(a) - \delta/2$. При этом, поскольку φ непрерывна и f_{i_a} полунепрерывна снизу,

найдется такая открытая окрестность \mathcal{U}_a точки a в X , что для всех $x \in \mathcal{U}_a$ справедливо неравенство $f_{t_a}(x) \geq \varphi(x) - \delta$. Множество K может быть покрыто конечным числом множеств \mathcal{U}_a . Пусть это будут множества $\mathcal{U}_{a_1}, \mathcal{U}_{a_2}, \dots, \mathcal{U}_{a_n}$. Обозначим теперь через k индекс, при котором $f_k \geq \sup(f_{t_{a_1}}, f_{t_{a_2}}, \dots, f_{t_{a_n}})$ (условие «возрастания»). Тогда на всем K имеем $f_k \geq \varphi - \delta$, и, следовательно, всюду $f_k \geq (\varphi - \delta)\chi_K$, поскольку вне K вторая функция равна нулю, а первая ≥ 0 . При этом $\int f_k \geq \int \varphi - \delta \mu(K) \geq A$, что и доказывает наше утверждение.

Это почти единственное свойство теории интегрирования, справедливое для предела несчетных семейств функций. Этот результат не верен для произвольных $f_i \geq 0$. Пусть, например, $f = 1$ на \mathbb{R} и $\mu = dx$. Возьмем в качестве f_i характеристические функции конечных множеств. Тогда $\int^* f_i = 0$ и $\int^* f = +\infty$.

Замечание. Существенно, что мы всегда предполагали X счетным в бесконечности. Вот пример того, что может случиться, если не делать этого предположения. Пусть пространство X является суммой несчетного числа прямых \mathbb{R} . Точнее, пусть I является несчетным множеством индексов и для каждого $i \in I$ множество \mathbb{R}_i изоморфно \mathbb{R} . Множества \mathbb{R}_i предполагаются непересекающимися в X , а их объединение считается равным X . Введем в X следующую топологию: подмножество X считается открытым, если открыто его пересечение с каждым множеством \mathbb{R}_i , $i \in I$. Множество X не является счетным в бесконечности. В самом деле, множества \mathbb{R}_i образуют открытое покрытие любого компакта K , из которого можно выбрать конечное подпокрытие. Другими словами, K содержится в объединении конечного числа множеств \mathbb{R}_i . Следовательно, счетное объединение компактов из X содержится в счетном объединении множеств \mathbb{R}_i и не может совпадать со всем X . Введем на каждом \mathbb{R}_i меру dx . Тогда будет определена мера Радона $\mu \geq 0$ на X . В самом деле, если $\varphi \in \mathcal{C}(X)$, то ее компактный носитель содержится в конечном объединении \mathbb{R}_i : $\bigcup_{i \in J} \mathbb{R}_i$, и можно будет положить

$$\mu(\varphi) = \sum_{i \in J} \int \varphi_i dx, \text{ где } \varphi_i \text{ — сужение } \varphi \text{ на } \mathbb{R}_i.$$

Рассмотрим теперь множество A , образованное из точек 0 всех множеств \mathbb{R}_i . Оно замкнуто, однако не является измеримым в смысле определения, данного на стр. 501. В самом деле, всякое открытое множество, содержащее A , должно содержать открытые непустые множества на всех \mathbb{R}_i , а, значит, его мера равна $+\infty$. Каждый компакт, содержащийся в A , содержит только нули конечного

числа множеств R_i . Следовательно, его мера равна нулю. Таким образом, $\mu^*(A) = +\infty$ и $\mu_*(A) = 0$. Поэтому, например, равенство $\mu^*(A) = \sup \mu^*(A \cap K)$ для каждого компакта K из X неверно. Пусть χ_A — характеристическая функция множества A . Она не измерима в смысле, указанном на стр. 510. Тем не менее она обладает свойством Лузина, которое в случае счетности в бесконечности влечет за собой измеримость. Ее верхний интеграл, определенный так, как это указано на стр. 520, является точной верхней гранью верхних интегралов от ограниченных функций с компактным носителем, минорирующих эту функцию. Далее, каждая неотрицательная функция с компактным носителем, ограничивающая снизу функцию χ_A , имеет интеграл, равный нулю, поэтому $\int^* \chi_A = 0 \neq \mu^*(A)$. Напротив, каждая полу-непрерывная снизу функция $\geq \chi_A$ имеет бесконечный интеграл. В теории интегрирования, где верхний интеграл определяется полунепрерывными снизу функциями, $\int^* \chi_A = \mu^*(A) = +\infty$. Значит, различные теории интегрирования не эквивалентны. Если попытаться в одну теорию включить случай множества X , не счетного в бесконечности, то интеграл нужно определять с помощью полунепрерывных снизу функций. Мера множеств определяется интегралом от их характеристических функций. Измеримые функции определяются по Лузину. Множество называется измеримым, если измерима его характеристическая функция, т. е. обладает свойством Лузина. В этом случае большая часть прежних теорем сохраняется, но не все. Предположение о том, что X счетно в бесконечности, мы повторяли в условиях теорем только тогда, когда это было необходимым, и даже в теории интегрирования, построение которой исходило из полунепрерывных снизу функций. Однако каждый человек, обладающий элементарной осторожностью, будет остерегаться даже думать о локально компактных пространствах X , не являющихся счетными в бесконечности, так как при этом возможны самые грубые ошибки

Пространства $\mathcal{L}^p(X, \mu; \tilde{F})$

Пусть f — вещественная неотрицательная функция с конечными или бесконечными значениями, определенная μ -почти всюду на X . Назовем нормой порядка p функции f , где p — вещественное конечное число ≥ 1 , число, определяемое по формуле

$$N_p(f) = N_p(X, \mu; f) = \left(\int^* f^p \right)^{1/p}. \quad (\text{IV}, 4; 41)$$

Теорема 40. Если $f \leq g$, то $N_p(f) \leq N_p(g)$. Имеют место соотношения

$$N_p(f+g) \leq N_p(f) + N_p(g), \quad N_p(kf) = |k| N_p(f) \quad (\text{IV}, 4; 42)$$

(неравенство выпуклости или неравенство Минковского). Более общо, имеет место неравенство счетной выпуклости:

$$N_p\left(\sum_{n=0}^{\infty} u_n\right) \leq \sum_{n=0}^{\infty} N_p(u_n), \quad (\text{IV}, 4; 43)$$

где u_n — неотрицательные функции, определенные на множестве X .

Доказательство. Соотношение для kf очевидно. Рассмотрим функции f и g . Предположим сначала, что они являются этажными функциями с компактным носителем и что имеют место разложения

$$f = \sum_{i=1}^n f_i \varphi_{X_i}, \quad g = \sum_{i=1}^n g_i \varphi_{X_i}, \quad (\text{IV}, 4; 44)$$

где f_i и g_i — постоянные и X_i — измеримые подмножества X .

Интеграл $N_p(f)$ является нормой порядка p : $\|\vec{Z}\|_p$, где \vec{Z} — вектор пространства \mathbb{C}^n с составляющими Z_1, Z_2, \dots, Z_n (где $Z_i = f_i$) относительно весов c_1, c_2, \dots, c_n , $c_i = \mu(X_i)$. Рассматривая точно так же функцию g , можно убедиться в том, что неравенство выпуклости является неравенством Минковского, доказанным в теореме 35 гл. III.

Предположим теперь, что f и g — произвольные, но ограниченные функции с компактным носителем. Обозначим через f_1 и g_1 этажные функции с компактным носителем, мажорирующие функции f и g . Тогда имеет место неравенство

$$N_p(f+g) \leq N_p(f_1+g_1) \leq N_p(f_1) + N_p(g_1). \quad (\text{IV}, 4; 45)$$

Переходя справа к точной нижней грани, получаем искомое неравенство выпуклости для функций f и g .

Перейдем теперь к случаю произвольных неотрицательных функций f и g . Если мы построим определенные ранее функции $f_M, k, g_M, k, (f+g)_M, k$, то получим неравенство

$$(f+g)_{M, k} \leq f_{M, k} + g_{M, k}, \quad (\text{IV}, 4; 46)$$

откуда вытекает неравенство

$$N_p((f+g)_{M, K}) \leq N_p(f_{M, K}) + N_p(g_{M, K}) \leq N_p(f) + N_p(g). \quad (\text{IV}, 4; 47)$$

Чтобы получить нужную формулу, достаточно слева перейти к точной верхней грани по всем M и K .

Докажем, наконец, неравенство счетной выпуклости. Для конечного числа m первых членов имеет место неравенство

$$N_p(u_0 + u_1 + \dots + u_m) \leq N_p(u_0) + N_p(u_1) + \dots + N_p(u_m). \quad (\text{IV}, 4; 48)$$

Для получения нужного результата надо, пользуясь теоремой Фату 36, перейти к пределу при m , стремящемся к бесконечности.

Определение. Пусть X — локально компактное пространство, счетное в бесконечности, снабженное мерой Радона $\mu \geq 0$, и пусть \vec{F} — некоторое банахово пространство. Пусть p — конечное вещественное число ≥ 1 .

Говорят, что функция \vec{f} , определенная на X , со значениями в \vec{F} , принадлежит пространству $\mathcal{L}^p(X, \mu; \vec{F})$, или просто $\mathcal{L}^p(\vec{F})$, если это не может привести к недоразумению¹⁾, или что она имеет интегрируемую p -ю степень, если, каково бы ни было число $\varepsilon > 0$, существует такая этажная функция \vec{g} с компактным носителем, определенная на X , и со значениями в \vec{F} , при которой имеет место неравенство $N_p(\|\vec{f} - \vec{g}\|) \leq \varepsilon$, или иначе: если существует последовательность $\vec{f}_0, \vec{f}_1, \vec{f}_2, \dots$ этажных функций с компактным носителем, для которых $N_p(\|\vec{f} - \vec{f}_n\|)$ сходится к 0 при n , стремящемся к бесконечности. В этом случае говорят, что последовательность \vec{f}_n является аппроксимирующей последовательностью для функции \vec{f} в пространстве $\mathcal{L}^p(X, \mu; \vec{F})$, составленной из этажных функций²⁾. Иногда, допуская вольность речи, говорят, что функция \vec{f} принадлежит пространству $\mathcal{L}^p(\vec{F})$, если она определена только почти всюду и если любое ее продолжение туда, где она не определена, принадлежит $\mathcal{L}^p(\vec{F})$. Пространство $\mathcal{L}^1(X, \mu; \vec{F})$ является пространством интегрируемых функций со значениями в \vec{F} . Таким образом, мы обобщаем понятие пространства инте-

¹⁾ Напротив, не следует в обозначении опускать \vec{F} . Обозначение $\mathcal{L}^p(X, \mu)$ или \mathcal{L}^p применяется в том случае, когда \vec{F} является скалярным полем.

²⁾ Если провести рассуждение, аналогичное доказательству теоремы 32, опираясь при этом на лемму 2 к теореме 30 с $\delta = (\varepsilon/2NM)^p$, то можно увидеть, что этажные функции можно заменить на непрерывные разложимые функции с компактным носителем.

Интегрируемых функций. Однако для $p \neq 1$ интегрируемость функции, принадлежащей пространству $\mathcal{L}^p(X, \mu; \vec{F})$, гарантирована.

Теорема 41. *Если f принадлежит пространству $\mathcal{L}^p(\vec{F})$, то неотрицательная функция $\|\vec{f}\|: x \rightarrow \|\vec{f}(x)\|$ принадлежит пространству \mathcal{L}^p .*

Доказательство проводится точно так же, как доказательство теоремы 29 для случая $p = 1$.

Теорема 42. *Пространство $\mathcal{L}^p(\vec{F})$ является векторным пространством, и в этом пространстве функция N_p , определенная формулой*

$$N_p(\vec{f}) = N_p(\|\vec{f}\|) = \left(\int^* \|\vec{f}\|^p \right)^{1/p}, \quad (\text{IV, 4; 49})$$

является полунормой.

Доказательство. Пусть \vec{f}_n и \vec{g}_n — две аппроксимирующие последовательности функций \vec{f} и \vec{g} в $\mathcal{L}^p(\vec{F})$. Тогда из неравенства Минковского следует, что сумма $\vec{f}_n + \vec{g}_n$ является аппроксимирующей последовательностью для $\vec{f} + \vec{g}$, принадлежащей поэтому к $\mathcal{L}^p(\vec{F})$. Аналогичное рассуждение можно провести относительно функции $k\vec{f}$, где k — скалярная постоянная. Полунормой в векторном пространстве называется неотрицательная скалярная функция N , удовлетворяющая соотношениям:

$$\begin{aligned} N(k\vec{X}) &= |k| N(\vec{X}), \\ N(\vec{X} + \vec{Y}) &\leq N(\vec{X}) + N(\vec{Y}). \end{aligned} \quad (\text{IV, 4; 50})$$

Если, кроме того, из $N(\vec{X}) = 0$ следует $\vec{X} = \vec{0}$, то мы получаем обычную норму. Тот факт, что N_p является полунормой, тривиальным образом вытекает из теоремы Минковского 40. Однако, как легко видеть, эта функция не является нормой. В самом деле, если $N_p(\vec{f}) = 0$, т. е. если $\int^* \|\vec{f}\|^p = 0$, то это означает лишь, что функция \vec{f} почти всюду равна нулю (теорема 26), и, значит, она не обязательно должна быть нулевым элементом векторного пространства $\mathcal{L}^p(\vec{F})$.

Теорема 43. *Пусть u_n является последовательностью функций, определенных почти всюду на X , со значениями в F ,*

принадлежащих $\mathcal{L}^p(\vec{F})$ и таких, что ряд $\sum_{n=0}^{\infty} N_p(\vec{u}_n)$ является сходящимся. Тогда ряд $\sum_{n=0}^{\infty} \vec{u}_n(x)$ почти всюду абсолютно сходится, и если через $\vec{S}(x)$ обозначить его сумму $\vec{S}(x) = \sum_{n=0}^{\infty} \vec{u}_n(x)$, то \vec{S} будет функцией, определенной почти всюду и принадлежащей пространству $\mathcal{L}^p(\vec{F})$. Кроме того, имеет место соотношение

$$N_p(\vec{S}) = N_p\left(\sum_{n=0}^{\infty} \vec{u}_n\right) \leq \sum_{n=0}^{\infty} N_p(\vec{u}_n) \quad (\text{IV}, 4; 51)$$

(неравенство счетной выпуклости). Величина $N_p(S - \vec{S}_m) = N_p(\vec{R}_m)$ стремится к 0 при m , стремящемся к бесконечности, а $N_p(\vec{S}_m)$ стремится к $N_p(\vec{S})$.

Доказательство аналогично доказательству, проведенному для $p = 1$ (теорема 37). Хотя N_p является полунормой, а не нормой, и, следовательно, $\mathcal{L}^p(\vec{F})$ является лишь полунормированным векторным пространством, часто применяется язык теории нормированных пространств. Например, если \vec{f}_n является последовательностью функций из $\mathcal{L}^p(\vec{F})$ и если $N_p(\vec{f} - \vec{f}_n)$ стремится к 0, то говорят, что « \vec{f}_n стремится к \vec{f} в $\mathcal{L}^p(\vec{F})$ ». Однако здесь предел не единственный. Если \vec{f}_n «стремится к» \vec{f} , то эта функция также стремится и к \vec{g} , лишь бы только эта функция была почти всюду равной \vec{f} . В теореме 43 утверждается, что \vec{S}_m стремится к \vec{S} в $\mathcal{L}^p(\vec{F})$, поскольку $N_p(\vec{S} - \vec{S}_m)$ стремится к нулю при m , стремящемся к бесконечности. В ней говорится также, что ряд $\sum_{n=0}^{\infty} \vec{u}_n$ является сходящимся в $\mathcal{L}^p(\vec{F})$ и даже абсолютно сходящимся. (Мы снова допускаем вольности речи, поскольку \vec{S} определена лишь почти всюду. Аналогичная вольность заключается уже в самой записи $\vec{S} \in \mathcal{L}^p(\vec{F})$.)

Следствие. Если последовательность функций \vec{f}_n сходится к функции \vec{f} в $\mathcal{L}^p(\vec{F})$ при n , стремящемся к бесконечности, то из рассматриваемой последовательности можно извлечь такую подпоследовательность, которая сходится почти всюду к \vec{f} .

Доказательство аналогично доказательству теоремы 38.

Теорема 44. Для того чтобы функция \vec{f} , определенная на X со значениями в банаховом пространстве \vec{F} , принадлежала пространству $\mathcal{L}^p(\vec{F})$, необходимо и достаточно, чтобы она была измеримой и чтобы интеграл $\int^* \|\vec{f}\|^p$ был конечным.

Доказательство остается тем же самым, что и в случае $p = 1$ (теорема 39).

Следствие. Если \vec{f} принадлежит пространству $\mathcal{L}^p(\vec{F})$, то функция $\|\vec{f}\|^p: x \rightarrow \|\vec{f}(x)\|^p$ интегрируема на X . Более общо, если $q \leq p$, то функция $\|\vec{f}\|^p$ принадлежит пространству $\mathcal{L}^{p/q}(X, \mu) = \mathcal{L}^{p/q}$.

Доказательство. Из сформулированного условия и теоремы вытекает, что функция \vec{f} измерима. В силу непрерывности нормы, то же самое имеет место для функции $\|\vec{f}\|^q$ (теорема 9 гл. II и теорема 22). Теперь нужный результат вытекает, в силу теоремы 44, из неравенства $\int^* (\|\vec{f}\|^q)^{p/q} = (N_p(\vec{f}))^p < +\infty$.

Теорема 45 (о применении линейного непрерывного отображения). Пусть L — линейное непрерывное отображение банахова пространства \vec{F} в банахово пространство \vec{G} . Если \vec{f} является функцией, определенной на X со значениями в \vec{F} , принадлежащей пространству $\mathcal{L}^p(\vec{F})$, то функция $L \circ \vec{f}$ принадлежит пространству $\mathcal{L}^p(\vec{G})$ и, кроме того, имеет место неравенство

$$N_p(L \circ \vec{f}) \leq \|L\| N_p(\vec{f}). \quad (\text{IV}, 4; 52)$$

Если, в частности, $p = 1$, то $L \circ \vec{f}$ интегрируема и

$$\int L \circ \vec{f} = L \left(\int \vec{f} \right). \quad (\text{IV}, 4; 53)$$

Так как $\|(L \circ \vec{f})x\| \leq \|L\| \|\vec{f}(x)\|$, то неравенство очевидно. Рассматривая аппроксимирующую последовательность, легко убедиться в справедливости заключительной части теоремы. Мы видим, что непрерывное линейное отображение L пространства \vec{F} в пространство \vec{G} определяет линейное непрерывное

отображение пространства $\mathcal{L}^p(\vec{F})$ в пространство $\mathcal{L}^p(\vec{G})$ с нормой $\leq \|L\|^1$.

Теорема 46 (о применении билинейного непрерывного отображения; неравенство Гёльдера). *Пусть \vec{f} принадлежит $\mathcal{L}^p(\vec{F})$ и \vec{g} принадлежит $\mathcal{L}^q(\vec{G})$, а B является непрерывным билинейным отображением $\vec{F} \times \vec{G}$ в \vec{H} . Если теперь $r \geq 1$ определено из равенства $1/r = 1/p + 1/q$, то функция $B(\vec{f}, \vec{g}): x \rightarrow B(\vec{f}(x), \vec{g}(x))$ принадлежит пространству $\mathcal{L}^r(\vec{H})$ и, кроме того, имеет место неравенство Гёльдера:*

$$N_r(B(\vec{f}, \vec{g})) \leq \|B\| N_p(\vec{f}) N_q(\vec{g}). \quad (\text{IV}, 4; 54)$$

Доказательство. Предположим сначала, что \vec{f} и \vec{g} являются вещественными неотрицательными функциями. Тогда неравенство Гёльдера относительно произведения $\vec{f}\vec{g}$ доказывается точно так же, как доказывалось неравенство Минковского (теорема 40). Только на этот раз надо свести его к теореме 36 гл. III. Затем неравенство (IV, 4; 54) можно получить и во всех остальных случаях, так как, в силу неравенства $\|B(\vec{f}(x), \vec{g}(x))\| \leq \|B\| \|\vec{f}(x)\| \|\vec{g}(x)\|$, задача сводится непосредственно к случаю произведения двух вещественных неотрицательных функций. Заметим теперь, что в силу предположений, наложенных на функции \vec{f} и \vec{g} , они измеримы, и поэтому из следствия 3 теоремы 23₂ вытекает, что $B(\vec{f}, \vec{g})$ является измеримым отображением X в \vec{H} . Для того чтобы убедиться, что эта функция принадлежит пространству $\mathcal{L}^r(\vec{H})$, теперь достаточно применить теорему 44.

Замечания. 1°) Если $r < 1$ (например, если $p = 1$, $q = 1$, а значит, $r = 1/2$), то никакого определенного заключения сделать невозможно.

2°) Содержание теоремы можно передать, сказав, что $(\vec{f}, \vec{g}) \rightarrow B(\vec{f}, \vec{g})$ является билинейным непрерывным отображением пространства $\mathcal{L}^p(\vec{F}) \times \mathcal{L}^q(\vec{G})$ в пространство $\mathcal{L}^r(\vec{H})$ (в упрощенном изложении, принятом на стр. 565).

Следствие 1. *Пусть p и p' — два сопряженных показателя, т. е. показателя, удовлетворяющих соотношению $1/p + 1/p' = 1$,*

¹⁾ Здесь опять допущена некоторая неточность языка, о которой мы говорили на стр. 565, ибо $\mathcal{L}^p(\vec{F})$ является всего лишь полуформированным.

или $p' = 1/(1-p)$. Если \vec{f} и \vec{g} принадлежат соответственно пространствам $\mathcal{L}^p(\vec{F})$ и $\mathcal{L}^{p'}(\vec{G})$, то функция $B(\vec{f}, \vec{g})$ интегрируема и имеет место неравенство:

$$\left\| \int B(\vec{f}, \vec{g}) \right\| \leq \|B\| N_p(\vec{f}) N_{p'}(\vec{g}). \quad (\text{IV, 4; 55})$$

Для доказательства достаточно пременить теорему, учитывая тот факт, что $\left\| \int B(\vec{f}, \vec{g}) \right\| \leq \int \|B(\vec{f}, \vec{g})\| = N_1(B(\vec{f}, \vec{g}))$.

Это следствие играет особо важную роль в том случае, когда рассматриваются функции с вещественными или комплексными значениями и когда билинейное отображение является обычным произведением. В этом случае имеет место неравенство

$$\left| \int fg \right| \leq N_p(f) N_{p'}(g). \quad (\text{IV, 4; 56})$$

Следствие 2. Если \vec{f} и \vec{g} принадлежат соответственно пространствам $\mathcal{L}^2(\vec{F})$ и $\mathcal{L}^2(\vec{G})$, то функция $B(\vec{f}, \vec{g})$ интегрируема и имеет место неравенство

$$\left\| \int B(\vec{f}, \vec{g}) \right\| \leq \|B\| N_2(\vec{f}) N_2(\vec{g}). \quad (\text{IV, 4; 57})$$

Эта теорема особенно важна в случае функций с комплексными значениями, для которых имеет место неравенство Коши—Шварца:

$$\left| \int f\bar{g} \right| \leq \left(\int |f|^2 \right)^{1/2} \left(\int |g|^2 \right)^{1/2}. \quad (\text{IV, 4; 58})$$

Общеизвестно прямое доказательство этого неравенства. С этой целью записывают очевидное неравенство

$$\int (f + \lambda g)(\bar{f} + \bar{\lambda}\bar{g}) = \int |f + \lambda g|^2 \geq 0 \quad (\text{IV, 4; 59})$$

и, разлагая левую часть по λ , получают, что трехчлен

$$\int |f|^2 + \lambda \int f\bar{g} + \bar{\lambda} \int f\bar{g} + |\lambda|^2 \int |g|^2 \quad (\text{IV, 4; 60})$$

при любом λ должен быть неотрицательным. Это означает, что его дискриминант должен быть неположительным, откуда следует неравенство (IV, 4; 58).

Совсем просто это получается из неравенства Коши — Шварца (теорема 2₃ гл. III), примененного к полуторалинейной форме $(f, g) \rightarrow \int f\bar{g}$, определенной в векторном пространстве \mathcal{L}^2 .

Однако для того, чтобы иметь право говорить о *векторном пространстве* \mathcal{L}^2 , определять указанную полуторалинейную форму и иметь возможность писать (IV, 4; 59) и (IV, 4; 60), надо показать, что если $f \in \mathcal{L}^2$ и $g \in \mathcal{L}^2$, то $\bar{f}\bar{g}$ интегрируема и $f+g \in \mathcal{L}^2$. Это можно сделать непосредственно следующим образом. Функции $f+g$ и $\bar{f}\bar{g}$ измеримы. Так как

$$|f+g|^2 \leq 2(|f|^2 + |g|^2) \quad \text{и} \quad |\bar{f}\bar{g}| \leq |f|^2 + |g|^2, \quad (\text{IV}, 4; 61)$$

то $N_2(f+\bar{g})$ и $N_1(\bar{f}\bar{g})$ конечны и результат вытекает из теоремы 44. Полезно знать, в каком случае в (IV, 4; 58) имеет место обобщенное неравенство \leq и в каком случае — строгое неравенство $<$. Если, с одной стороны, $N_2(\bar{g}) > 0$ и если, с другой, для каждого значения λ норма $N_2(\bar{f} + \lambda\bar{g}) > 0$, то трехчлен (IV, 4; 59) в нуль никогда не обратится, а его дискриминант будет отрицателен, а, следовательно, неравенство будет строгим. Мы видим, что всегда имеет место строгое неравенство, за исключением того случая, когда существуют две комплексные постоянные α и β , не обращающиеся одновременно в нуль и такие, что функция $\bar{\alpha}\bar{f} + \bar{\beta}\bar{g}$ равна почти всюду нулю.

Следствие 3 (критерий принадлежности различным пространствам $\mathcal{L}^p(\vec{F})$). *Если $\|\mu\| < +\infty$, то каждая функция, принадлежащая $\mathcal{L}^p(\vec{F})$, принадлежит всем пространствам $\mathcal{L}^r(\vec{F})$ для $r \leq p$ и, кроме того, имеет место неравенство*

$$N_r(\vec{f}) \leq \|\mu\|^{1/r - 1/p} N_p(\vec{f}). \quad (\text{IV}, 4; 61_2)$$

Доказательство. В самом деле, для доказательства достаточно к функции \vec{f} и к постоянной скалярной функции $g = 1$ применить теорему 46. Функция B в этом случае является билинейным отображением $(\vec{f}, k) \rightarrow k\vec{f}$ пространства $\vec{F} \times \mathbb{K}$ в \vec{F} . Число q определяется из равенства $1/q = 1/r - 1/p$ ¹.

Эта теорема особенно важна в различных приложениях. Если, например, X является ограниченным интервалом вещественной прямой и $\mu = dx$, то при $r \leq p$ принадлежность к \mathcal{L}^p влечет за собой принадлежность к \mathcal{L}^r . В частности, функция,

¹) Доказательство одной принадлежности $\vec{f} \in \mathcal{L}^r(\vec{F})$ без неравенства (IV, 4; 61) очевидно. В самом деле, \vec{f} измерима. Далее, $\|\vec{f}(x)\|^r \leq \|\vec{f}(x)\|^p$, если $\|\vec{f}(x)\| \geq 1$, и ≤ 1 , если $\|\vec{f}(x)\| \leq 1$. Значит, $\|\vec{f}(x)\|^r \leq 1 + \|\vec{f}(x)\|^p$ и верхний интеграл от $\|\vec{f}\|^r$ конечен (так как $\|\mu\| < +\infty$, то $\int 1$ конечен).

принадлежащая к какому-либо пространству \mathcal{L}^p , принадлежит к \mathcal{L}^1 и, значит, интегрируема. Напротив, как легко видеть:

1°) Между пространствами \mathcal{L}^p не существует отношений обратного порядка. Функция, принадлежащая \mathcal{L}^p , вовсе не обязана принадлежать \mathcal{L}^s при $s > p$. Пусть, например, на компактном интервале $[0, 1]$ прямой \mathbb{R} в качестве меры взято $d\mu = dx$ и рассматривается функция $f(x) = 1/x^\alpha$. Тогда для $\alpha = 1/2$ она принадлежит \mathcal{L}^1 , но не принадлежит \mathcal{L}^2 .

2°) Если норма меры μ бесконечна, то между различными пространствами \mathcal{L}^p не существует никаких отношений включения. Возьмем, например, $X = \mathbb{R}$ и $d\mu = dx$. Тогда пример, приведенный в п. 1°), показывает, что принадлежность функции к \mathcal{L}^p не влечет за собой ее принадлежности к \mathcal{L}^s , $s > p$. Обратно, функция $f(x) = 1/\sqrt{1+x^2}$ в этом случае принадлежит \mathcal{L}^2 , но не принадлежит \mathcal{L}^1 .

3°) В пространстве $(\mathbb{C}^{[0, 1]})_{cb}$ комплексных непрерывных функций, определенных на интервале $[0, 1]$ вещественной прямой \mathbb{R} , нормы N_p относительно меры dx при различных p не эквивалентны. Этот факт дает нам простой пример неэквивалентных норм в векторном бесконечномерном пространстве (см. на стр. 58 теорему 13 гл. II). В этом пространстве N_p является нормой, а не только полуформой. В самом деле, если $N_p(f) = 0$, то функция f должна быть почти всюду равной нулю. Эта функция непрерывна. Пусть c — произвольная точка $[0, 1]$. Для любого $\varepsilon > 0$ должно существовать такое число $\eta > 0$, что из неравенства $|x - c| \leq \eta$ следует неравенство $|f(x) - f(c)| \leq \varepsilon$. Множество таких точек x заполняет интервал длины $\geq \eta$, а функция f почти всюду равна нулю. Следовательно, должна существовать хотя бы одна точка x , такая, что $f(x) = 0$, а, значит, $|f(c)| \leq \varepsilon$. Так как ε произвольно, то $f(c) = 0$. Поскольку c произвольно, функция f тождественно равна нулю, а число N_p является нормой.

Рассмотрим для каждого n функцию $f_n(x)$ (см. рис. 16), определенную по формуле

$$f_n(x) = \begin{cases} n(1-nx) & \text{для } 0 \leq x \leq \frac{1}{n}, \\ 0 & \text{для } x > \frac{1}{n}. \end{cases} \quad (\text{IV, 4; 61}_2)$$

Эта функция непрерывна, и имеет место равенство

$$N_p(f_n) = n^{1-1/p} \left(\frac{1}{p+1}\right)^{1/p}. \quad (\text{IV, 4; 61}_3)$$

Устремим n к бесконечности. Тогда для $q > p$ невозможно найти такое фиксированное число $k > 0$, при котором

$N_q(f_n) \leq k N_p(f_n)$, и из теоремы 12 гл. II вытекает неэквивалентность N_q и N_p .

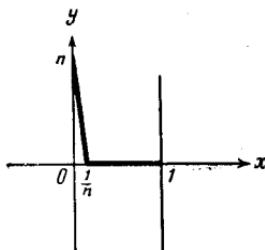


Рис. 16.

Следствие 4. Если функция \vec{f} одновременно принадлежит пространствам $\mathcal{L}^p(\vec{F})$ и $\mathcal{L}^q(\vec{F})$, то она принадлежит всем $\mathcal{L}^r(\vec{F})$, $p \leq r \leq q$, и имеет место неравенство

$$N_r(\vec{f}) \leq (N_p(\vec{f}))^\alpha (N_q(\vec{f}))^\beta, \quad (\text{IV}, 4; 61_4)$$

где

$$\alpha = \frac{\frac{1}{r} - \frac{1}{q}}{\frac{1}{p} - \frac{1}{q}}, \quad \beta = \frac{\frac{1}{p} - \frac{1}{r}}{\frac{1}{p} - \frac{1}{q}}.$$

Доказательство. Поскольку функция \vec{f} измерима, то достаточно (теорема 44) проверить соотношение (IV, 4; 61₄) для $\|\vec{f}\|$ или же для произвольной измеримой вещественной функции $g \geq 0$ ¹⁾.

Поскольку $p \leq r \leq q$, то $1/q \leq 1/r \leq 1/p$. Число $1/r$ определяет барицентр точек $1/p$ и $1/q$ с весами α и $\beta \geq 0$, $\alpha + \beta = 1$. Поэтому $1/r = \alpha/p + \beta/q$. Числа α и β равны числам, указанным в условии теоремы. Поскольку $\alpha + \beta = 1$, то $g = g^\alpha g^\beta$ и, кроме того, $g^\alpha \in \mathcal{L}^{p/\alpha}$, $g^\beta \in \mathcal{L}^{q/\beta}$ и $N_{p/\alpha}(g^\alpha) = (N_p(g))^\alpha$, $N_{q/\beta}(g^\beta) = (N_q(g))^\beta$.

1) Если проверяется лишь принадлежность \vec{f} пространству $\mathcal{L}^r(\vec{F})$, а не соотношение (IV, 4; 61₄), то этот факт очевиден. В самом деле, поскольку \vec{f} измерима, то достаточно доказать, что $N_r(\vec{f})$ конечна. Далее, $\|\vec{f}(x)\|^r \leq \|\vec{f}(x)\|^p$, если $\|\vec{f}(x)\| \leq 1$, и $\|\vec{f}(x)\|^q$, если $\|\vec{f}(x)\| \geq 1$. Поэтому $\|\vec{f}(x)\|^r \leq \|\vec{f}(x)\|^p + \|\vec{f}(x)\|^q$ и ее верхний интеграл, очевидно, конечен.

Если теперь к произведению $g^{\alpha}g^{\beta} = g$ применить неравенство Гёльдера (IV, 4; 54) с показателями p/α , q/β , $\frac{1}{p/\alpha} + \frac{1}{q/\beta} = \frac{1}{r}$, то мы получим соотношение (IV, 4; 61₄).

В качестве следствия получаем: множество конечных чисел $r \geq 1$, для которых заданная функция \vec{f} принадлежит $\mathcal{L}^p(\vec{F})$, является интервалом из $[1, +\infty[$ (открытым, полуоткрытым или замкнутым). Кроме того, неравенство (IV, 4; 61₄) показывает, что на этом интервале функция $\xi \rightarrow \ln N_{1/\xi}(f)$ выпукла (см. стр. 204).

Пространства $L^p(X, \mu; \vec{F})$. Теорема Фишера — Рисса

При изучении пространств $\mathcal{L}^p(\vec{F})$ постоянно приходится испытывать существенные затруднения, связанные с наличием полунорм вместо норм, не позволяющих пользоваться теоремами, относящимися к нормированным векторным пространствам.

Для того чтобы обойти эти затруднения, мы введем новое нормированное пространство. В пространстве $\mathcal{L}^p(\vec{F})$ множество функций \vec{f} , для которых $N_p(\vec{f}) = 0$, является некоторым векторным подпространством, а именно подпространством функций, почти всюду равных нулю. Мы можем теперь рассматривать факторпространство пространства $\mathcal{L}^p(\vec{F})$ по этому векторному подпространству. Это факторпространство называют пространством $L^p(X, \mu; \vec{F})$ или просто $L^p(\vec{F})$, когда это не приводит к недоразумениям. Элементом $L^p(\vec{F})$ является не функция, определенная на X со значениями в \vec{F} , а класс попарно почти всюду равных между собой функций. Две функции, определенные на X , называются μ -эквивалентными в смысле Лебега, или просто эквивалентными, когда это не приводит к недоразумениям, если они μ -почти всюду равны. Тогда $L^p(\vec{F})$ является «пространством функций, принадлежащих $\mathcal{L}^p(\vec{F})$ и определенных с точностью до эквивалентности».

Конечно, $L^p(\vec{F})$, будучи факторпространством некоторого векторного пространства по векторному подпространству, само является векторным пространством (гл. I, стр. 20). Кроме того, это векторное пространство нормировано. В самом деле, каждый элемент \vec{f}^* пространства $L^p(\vec{F})$ является множеством функций \vec{f} из $\mathcal{L}^p(\vec{F})$, имеющих одну и ту же полунорму $N_p(\vec{f})$. Этую общую полунорму называют нормой рассматриваемого эле-

мента \vec{f} пространства $L^p(\vec{F})$. То, что это норма, очевидно. Действительно, она тривиальным образом удовлетворяет неравенствам (IV, 4; 50). Кроме того, если $N_p(\vec{f}) = 0$, то \vec{f} является классом, составленным из функций, почти всюду равных нулю, т. е. нулевым элементом пространства $L^p(\vec{F})$.

Теорема 47 (Фишера — Рисса). *Векторное нормированное пространство $L^p(X, \mu; \vec{F})$, где \vec{F} — некоторое банахово пространство, само является банаховым пространством, т. е. это пространство полно.*

Доказательство. Для доказательства мы можем опереться на теорему 56 гл. II. Пусть \vec{u}_n — такая последовательность элементов из $L^p(\vec{F})$, при которой ряд $\sum_{n=0}^{\infty} N_p(\vec{u}_n)$ сходится. Выберем в каждом классе \vec{u}_n некоторый представитель, т. е. элемент \vec{u}_n из класса \vec{u}_n . Тогда будет иметь место соотношение $\sum_{n=0}^{\infty} N_p(\vec{u}_n) < +\infty$. Из теоремы 43 следует, что ряд $\sum_{n=0}^{\infty} \vec{u}_n$ сходится почти всюду к некоторому элементу \vec{S} из $\mathcal{L}^p(\vec{F})$ и что, кроме того, $N_p(\vec{S} - \vec{S}_m)$ сходится к 0 при m , стремящемся к бесконечности. Если через \vec{S} обозначить класс \vec{S} , то он будет некоторым элементом пространства $L^p(\vec{F})$, а ряд $\sum_{n=0}^{\infty} \vec{u}_n$ будет сходиться к \vec{S} в $L^p(\vec{F})$. Этим рассуждением заканчивается доказательство полноты рассматриваемого пространства.

Теорема 48. *Пусть \vec{f} — некоторый элемент пространства $L^1(\vec{F})$. Будем рассматривать $\int \vec{f}$ как значение $\int \vec{f}$ на произвольном элементе \vec{f} из класса \vec{f} . Тогда этот интеграл является линейным непрерывным отображением пространства $L^1(X, \mu; \vec{F})$ в пространство \vec{F} с нормой, равной 1. При этом имеет место*

$$\left\| \int \vec{f} \right\| \leqslant N_1(\vec{f}). \quad (\text{IV, 4; 62})$$

Доказательство тривиально, по крайней мере в той его части, которая утверждает, что норма $\leqslant 1$. То, что норма в точности равна 1, доказывается сложнее, и мы это утверждение примем без доказательства. Впрочем, оно очевидно в том случае, когда \vec{F} является скалярным полем, ибо тогда для неотри-

цательной функции f имеет место точное равенство $\int f = N_1(f)$.

Теорема 49. В пространстве $L^p(\vec{F})$ множество классов этажных функций с компактным носителем или множество классов непрерывных разложимых функций с компактным носителем плотны.

Утверждение, относящееся к этажным функциям, очевидно, поскольку, если $\vec{f} \in L^p(\vec{F})$ и $\vec{f} \in \mathcal{L}^p(\vec{F})$ является представителем класса \vec{f} , то для функции \vec{f} найдется аппроксимирующая последовательность, образованная из этажных функций \vec{f}_n . При этом элементы \vec{f}_n сходятся к элементу \vec{f} в $L^p(\vec{F})$. Что касается непрерывных функций, то для доказательства следует обратиться к примечанию²⁾ на стр. 563.

Замечание. Мы говорили (теорема 41 гл. II), что каждое нормированное конечномерное векторное пространство полно. Однако это неверно для бесконечномерных векторных нормированных пространств. Пространство $L^p([0, 1], dx)$ полно, но его векторное подпространство, составленное из классов непрерывных функций, плотно и, значит, не замкнуто, а следовательно, не полно (следствие 1 теоремы 42 гл. II). Отсюда легко получить, что пространство $(\mathbb{C}^{[0, 1]})_{cb}$, снабженное нормой N_p (см. стр. 570), также не полно.

Пространства $\mathcal{L}^\infty(\vec{F})$ и $L^\infty(\vec{F})$

Определение. Говорят, что μ -измеримая вещественная неотрицательная функция f , определенная на X , μ -существенно ограничена или просто существенно ограничена, когда это не приводит к путанице, если существует такое число M , что почти всюду функция f не превосходит M . Точная нижняя грань M_0 чисел M , обладающих этим свойством, называется существенной верхней гранью f и обозначается через $N_\infty(f)$. Это число характеризуется следующими двумя свойствами:

1°) Каково бы ни было число $\varepsilon > 0$, найдется такое множество значений x с мерой > 0 , что имеет место неравенство $f(x) > M - \varepsilon$.

2°) Неравенство $f(x) > M$ может иметь место только на множестве значений x нулевой меры.

Если функция f не является существенно ограниченной, то полагают $N_\infty(f) = +\infty$.

Нетрудно видеть, что $N_\infty(f+g) \leq N_\infty(f) + N_\infty(g)$ и что $N_\infty(kf) = kN_\infty(f)$ для $k \geq 0$.

Через $\mathcal{L}^\infty(X, \mu; \vec{F})$ или просто $\mathcal{L}^\infty(\vec{F})$, когда это не приводит к недоразумению, обозначается множество функций, определенных на X со значениями в \vec{F} , измеримых и существенно ограниченных, т. е. таких, что функция $\|\vec{f}\|$ является μ -существенно ограниченной. На этом пространстве с помощью равенства $N_\infty(\vec{f}) = N_\infty(\|\vec{f}\|)$ определяется полуформа $N_\infty(\vec{f})$.

Через $L^\infty(X, \mu; \vec{F})$ или просто $L^\infty(\vec{F})$, когда это не приводит к недоразумению, обозначается векторное факторпространство пространства $\mathcal{L}^\infty(X, \mu; \vec{F})$ по векторному подпространству функций, μ -почти всюду равных нулю. Это пространство является некоторым векторным нормированным пространством с нормой $N_\infty(\vec{f}) = N_\infty(\vec{f})$, где \vec{f} — произвольный элемент из класса \vec{f} . Без труда проверяется, что теоремы 40—47 остаются в силе, если некоторые из показателей будут равны $+\infty$.

Теорема 44 исчезает, так как она включается в определение $\mathcal{L}^\infty(\vec{F})$. Следствие теоремы 43 может быть значительно улучшено: если \vec{f}_n является такой последовательностью функций из $\mathcal{L}^\infty(\vec{F})$, что при n , стремящемся к бесконечности, $N_\infty(\vec{f} - \vec{f}_n)$ стремится к 0, то функции \vec{f}_n почти всюду сходятся к \vec{f} (нет необходимости извлекать сходящуюся подпоследовательность). В самом деле, пусть B_n — множество таких значений x , что $\|\vec{f}(x) - \vec{f}_n(x)\| > N_\infty(\vec{f} - \vec{f}_n)$. Это множество имеет нулевую меру. Объединение B множеств B_n также имеет нулевую меру. Пусть $A = C B$. Так как $\sup_{x \in A} \|\vec{f}(x) - \vec{f}_n(x)\| = N_p(\vec{f} - \vec{f}_n)$ стремится к 0 при n , стремящемся к бесконечности, то функции \vec{f}_n сходятся просто, и даже равномерно, к функции \vec{f} . Сходимость в $\mathcal{L}^\infty(\vec{F})$ является «с точностью до множества меры нуль» равномерной сходимостью функций.

Отсюда вытекает, что в рассматриваемом пространстве теорема 49 места не имеет! В самом деле, пусть $X = \mathbb{R}$ и $\mu = dx$. Если \vec{g} — непрерывная функция, то $N_\infty(\vec{g})$ совпадает с $\|\vec{g}\|$. Действительно, множество точек x , в которых $\|\vec{g}(x)\| > N_\infty(\vec{g})$, является открытым множеством и имеет меру, равную нулю, т. е. пусто, и, следовательно, $N_\infty(\vec{g})$ действительно является точной верхней гранью для $\|\vec{g}\|$, т. е. равна $\|\vec{g}\|$. Если $\vec{f} \in \mathcal{L}^\infty(\vec{F})$ является пределом непрерывных функций \vec{f}_n , то $N_\infty(\vec{f}_m - \vec{f}_n)$,

а вместе с ним и $\|\vec{f}_m - \vec{f}_n\|$, стремится к 0 при m и n , стремящихся к бесконечности. Функции \vec{f}_n образуют последовательность Коши в $(\vec{F}^R)_{cb}$, и, поскольку это пространство полно (следствие 2 теоремы 65 гл. II), они равномерно сходятся к непрерывной и ограниченной функции \vec{g} . При этом $\|\vec{f}_n - \vec{g}\|$, а вместе с ней и $N_\infty(\vec{f}_n - \vec{g})$, сходится к 0. Мы получаем, что $N_\infty(\vec{f} - \vec{g}) = 0$, т. е. \vec{f} и \vec{g} почти всюду равны. Таким образом, единственными пределами непрерывных функций могут быть функции из $\mathcal{L}^\infty(\vec{F})$, почти всюду равные непрерывным функциям. Характеристическая функция интервала $[a, b]$ этим свойством не обладает.

Теорема 50. Пусть \vec{f} — измеримая функция, определенная на X , со значениями в \vec{F} . Множество точек p интервала $[1, +\infty]$, для которых $\vec{f} \in \mathcal{L}^p(\vec{F})$, образует открытый, полуоткрытый или замкнутый интервал $[p_1, p_2]$. Функция $p \rightarrow N_p(\vec{f})$ является непрерывной функцией, определенной на $[p_1, p_2]$, со значениями ≥ 0 в $\bar{\mathbb{R}}$. Если, кроме того, $\|\mu\| = 1$, то эта функция является всегда возрастающей и даже строго возрастающей, кроме того случая, когда $\|\vec{f}\|$ почти всюду равна одной и той же постоянной c , а тогда она постоянна и равна c . Функция $\xi \rightarrow \ln N_{1/\xi}(\vec{f})$ всегда выпукла на $[1/p_2, 1/p_1]$.

Доказательство. Тот факт, что множество точек p , для которых $\vec{f} \in \mathcal{L}^p(\vec{F})$, образует некоторый интервал, вытекает из следствия 4 теоремы 46, распространенного на бесконечные показатели.

Докажем непрерывность функции N_p . При этом достаточно, очевидно, ограничиться случаем, когда \vec{f} вещественна и неотрицательна. Пусть сначала p_0 — внутренняя точка интервала. Пусть p'_1 и p'_2 таковы, что $p_1 < p'_1 < p_0 < p'_2 < p_2$. При p , стремящемся к p_0 , функция f^p сходится просто к f^{p_0} . В силу примечания на стр. 571, для $p'_1 \leq p \leq p'_2$ имеет место неравенство $f^p \leq f^{p'_1} + f^{p'_2}$, правая часть которого не зависит от p и интегрируема. Из теоремы Лебега 35 следует, что $\int f^p$ сходится к $\int f^{p_0}$. Следовательно, $N_p(f) = (\int f^p)^{1/p}$ сходится к $N_{p_0}(f) = (\int f^{p_0})^{1/p_0}$, и непрерывность функции N_p в точке p_0 доказана. Можно было бы также воспользоваться теоремой 7₂.

Пусть теперь $p_0 = p_1$. Если $p_1 \in [p_1, p_2]$, то надо будет взять такое p'_2 , чтобы $p_1 < p'_2 < p_2$. Затем достаточно заменить $[p'_1, p'_2]$ на $[p_1, p'_2]$ и провести то же самое рассуждение, что и в предыдущем случае, чтобы доказать требуемую непрерывность. Единственным новым случаем будет тот, когда $p_1 \notin [p_1, p_2]$, т. е. когда $N_{p_1}(f) = +\infty$. В этом случае надо показать, что $N_p(f)$ стремится к $+\infty$ при p , стремящемся к p_1 по значениям $> p_1$. Зададим $A > 0$; поскольку $N_{p_1}(f) = +\infty$, то существует такое число $M \geq 1$ и такой компакт K из X , что $\int (f_{M,K})^{p_1} \geq 2A$. Для p , стремящегося к p_1 так, что $p_1 \leq p \leq p'_2$, функция $(f_{M,K})^p$ сходится к $(f_{M,K})^{p_1}$, сохраняя свой носитель в K и оставаясь ограниченной числом $M^{p'_2}$. Следовательно, по частной теореме Лебега 34, при p , стремящемся к p_1 , $\int (f_{M,K})^p$ стремится к $\int (f_{M,K})^{p_1}$. Значит, существует такое $\eta > 0$, что из $p \leq p_1 + \eta$ вытекает неравенство $\int (f_{M,K})^p \geq A$ и тем более неравенство $\int f^p \geq A$, чем полностью доказывается непрерывность в точке p_1 . Непрерывность в точке p_2 доказывается точно так же, если только $p_2 < +\infty$.

Предположим теперь, что $p_2 = +\infty$, и положим $M = N_\infty(f) \leq +\infty$. Надо показать, что при p , стремящемся к $+\infty$, $N_p(f)$ стремится к M . Зададим $M' < M$ и выберем M'_1 так, чтобы $M' < M'_1 < M$. Тогда существует множество Y меры > 0 , на котором $f(x) \geq M'_1$, и $N_p(f) \geq M'_1 (\mu(Y))^{1/p}$. При p , стремящемся к $+\infty$, последняя величина стремится к M'_1 . Следовательно, существует такое число q' , что при $p \geq q'$ имеет место неравенство $N_p(f) \geq M'$.

Этого достаточно для доказательства непрерывности в точке $p_2 = +\infty$, если $N_\infty(f) = +\infty$. Если же $M = N_\infty(f) < +\infty$, то надо еще установить неравенство противоположного смысла. Пусть теперь $M'' > M$. В неравенстве (IV, 4; 61₄), в котором вместо p, q, r взяты p_1, ∞, p , имеем: $\alpha = (1/p)/(1/p_1)$, $\beta = (1/p_1 - 1/p)/(1/p_1)$. Очевидно, при p , стремящемся к бесконечности, α стремится к 0 и β стремится к 1. При этом правая часть неравенства (IV, 4; 61₄) стремится к $(N_{p_1}(f))^0 (N_\infty(f))^\Gamma = M$.

Таким образом, существует такое целое число q'' , что при $p \geq q''$ имеет место неравенство $N_p(f) \leq M''$. При этом, если $q = \sup(q', q'')$, то из неравенства $p \geq q$ следует неравенство $M' \leq N_p(f) \leq M''$, и непрерывность рассматриваемой функции в точке p_2 полностью доказана.

При $\|\mu\| = 1$ неравенство (IV, 4; 61) означает, что функция $p \rightarrow N_p(f)$ является возрастающей. Можно показать, что если f не равна почти всюду некоторой постоянной, то эта функция является строго возрастающей. Мы это примем без доказательства.

Продолжение мер, не обладающих свойством неотрицательности

Мы только что видели, что для вещественной меры Радона $\mu \geq 0$ функцию $\varphi \rightarrow \mu(\varphi)$, первоначально определенную на $\mathcal{C}(X)$, можно продолжить до функции $f \rightarrow \int f d\mu$, определенной в значительно более широком векторном пространстве $\mathcal{L}^1(X, \mu)$. Мы даже определили этот интеграл для векторных функций $\vec{f} \in \mathcal{L}^1(X, \mu; \vec{F})$.

Можно ли сделать что-либо подобное, когда мера μ вещественна, но не является ≥ 0 или когда она является комплексной или векторной?

Такая теория продолжения существует, но мы о ней в настоящий момент говорить не можем. Тем не менее мы сейчас изучим частное продолжение хотя бы для того случая, когда мера μ принимает значения в конечномерном векторном пространстве.

Обозначим через $\Gamma(X)$ векторное пространство комплексно-значных борелевских ограниченных функций с компактнымносителем на X . Пространство $\mathcal{C}(X)$ является векторным подпространством пространства $\Gamma(X)$. Мы не будем вводить топологию на $\Gamma(X)$, но введем некоторое понятие сходимости, которое мы назовем L -сходимостью (L — начальная буква фамилии Лебег). (Можно показать, что эта сходимость не соответствует сходимости в топологическом пространстве.)

Последовательность f_0, f_1, f_2, \dots из $\Gamma(X)$ называется L -сходящейся к f , если при n , стремящемся к бесконечности, функции f_n сходятся к f , оставаясь в совокупности ограниченными по модулю и сохраняя свой носитель в некотором компакте K множества X .

Пусть μ — мера Радона ≥ 0 . Ее лебеговское продолжение позволяет определить некоторую функцию $f \rightarrow \int f d\mu$, определенную на $\Gamma(X)$, принимающую комплексные значения и обладающую следующими свойствами:

1°) Эта функция линейна.

2°) Если последовательность f_n является L -сходящейся к f , то $\int f_n d\mu$ сходится к $\int f d\mu$.

3°) Пусть \mathcal{O} — открытое множество X с компактным замыканием и $\chi_{\mathcal{O}}$ — его характеристическая функция. Тогда, каково бы ни было $\varepsilon > 0$, существует такой компакт $K \subset \mathcal{O}$, что для каждой функции φ из $\mathcal{C}(X)$, $0 \leq \varphi \leq 1$, равной 1 в некоторой окрестности компакта K и имеющей носитель в \mathcal{O} , имеет место неравенство

$$\left| \int \chi_{\mathcal{O}} d\mu - \int \varphi d\mu \right| \leq \varepsilon.$$

4°) Эта функция положительна в том смысле, что если $f \geq 0$, то $\int f d\mu \geq 0$.

Если \vec{E} является банаховым пространством, то всякое отображение пространства $\Gamma(X)$ в \vec{E} , удовлетворяющее трем предыдущим условиям, будет называться *линейным L-непрерывным отображением* $\Gamma(X)$ в \vec{E} . Если $\vec{E} = \mathbb{C}$ и выполняется четвертое условие, то отображение будет называться положительным (≥ 0). Таким образом, продолжение Лебега некоторой меры Радона $\mu \geq 0$ определяет линейную *L-непрерывную форму* ≥ 0 на $\Gamma(X)$.

Условие 3°) эквивалентно аналогичному условию 3₂): каковы бы ни были компакт K из X и $\varepsilon > 0$, существует такое открытое множество $\mathcal{O} \supset K$, что для любой функции $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, равной 1 на некоторой окрестности K и имеющей носитель в \mathcal{O} , имеет место неравенство $\left| \int \chi_K d\mu - \mu(\varphi) \right| \leq \varepsilon$. В самом деле, пусть заданы компакт K , число ε и выполняется условие 3°). Выберем открытую окрестность \mathcal{U} компактного замыкания K и найдем компакт H открытого множества $\mathcal{U} - K$, существование которого утверждается в 3°). Тогда открытое множество $\mathcal{U} - H = \mathcal{O}$ удовлетворяет относительно компакта K условию 3₂). Аналогично проводится доказательство обратного утверждения.

Если X метризуемо, то каждое отображение $\Gamma(X)$ в \vec{E} , удовлетворяющее условию 2°), автоматически удовлетворяет условию 3°). В самом деле, рассмотрим последовательность компактов $K_n = \{x \in \mathcal{O}; d(x, \mathbb{C}\mathcal{O}) \geq 1/n\}$. Для заданного $\varepsilon > 0$ при достаточно большом n компакты K_n удовлетворяют условию 3°). Если бы это было не так, то для каждого n можно было бы найти функцию $\varphi_n \in \mathcal{C}(X)$, $0 \leq \varphi_n \leq 1$, равную 1 на окрестности компакта K_n , с носителем в \mathcal{O} и такую, что $\left| \int \chi_{\mathcal{O}} d\mu - \int \varphi_n d\mu \right| > \varepsilon$. Последовательность φ_n сходится к $\chi_{\mathcal{O}}$, а это противоречит условию *L-непрерывности*. Однако если X не метризуемо, то условие 3°), в котором используется своего

рода предел несчетного семейства, не будет получаться из свойства 2°), имеющего место только для последовательностей. Условие 3°) вводилось не ради удовольствия усложнять себе жизнь: оно будет нужно в теореме единственности 50₂.

Для каждого L -непрерывного линейного отображения μ пространства $\Gamma(X)$ в \tilde{E} будет удобным через $\mu(A)$ обозначать значение μ на характеристической функции некоторой борелевской части A пространства X с компактным замыканием.

Теорема 50₂. Пусть \tilde{E} — банахово пространство. Каждое линейное и L -непрерывное отображение пространства $\Gamma(X)$ в \tilde{E} , равное нулю на векторном подпространстве $\mathcal{C}(X)$, тождественно равно нулю.

Доказательство. Из условия непрерывности 3°) вытекает, что эта функция заведомо равна нулю на характеристических функциях открытых множеств с компактным замыканием. Поэтому нам потребуется следующая лемма.

Лемма 1. Пусть \mathcal{U} — топологическое пространство и \mathcal{F} — множество частей \mathcal{U} , обладающее следующими свойствами:

1°) Устойчивость относительно разности:

если $A \in \mathcal{F}$, $B \in \mathcal{F}$ и $B \subset A$, то $A - B \in \mathcal{F}$.

2°) Устойчивость относительно возрастающего предела:

если A_n образуют возрастающую последовательность частей, принадлежащих \mathcal{F} , то их объединение принадлежит \mathcal{F} .

3°) Устойчивость относительно убывающего предела:

если A_n образуют убывающую последовательность частей из \mathcal{F} , то их пересечение принадлежит \mathcal{F} ; если \mathcal{F} — наименьшее множество, обладающее свойствами 1°) — 3°), содержащее все открытые части \mathcal{U} , то \mathcal{F} является σ -алгеброй борелевских частей \mathcal{U} .

Докажем эту лемму. Заметим, что предыдущие свойства не позволяют сразу утверждать, что \mathcal{F} представляет собой σ -алгебру. Мы не знаем, принадлежит ли пересечение $A \cap B$ двух частей из \mathcal{F} множеству \mathcal{F} . Если мы это докажем, то тогда сможем утверждать, что \mathcal{F} есть σ -алгебра. В самом деле, устойчивость относительно пересечения двух частей влечет за собой устойчивость относительно пересечения конечного числа частей, а затем, в силу условия, наложенного на убывающие последовательности, устойчивость относительно счетных пересечений. С другой стороны, из устойчивости относительно разности следует, что дополнение некоторой части из \mathcal{F} принадлежит \mathcal{F} , поскольку \mathcal{U} , как открытое множество, принадлежит \mathcal{F} . Устойчивость

теперь будет иметь место для любых счетных объединений, и, следовательно, \mathcal{F} в действительности представляет собой σ -алгебру.

Рассмотрим теперь для открытой части \mathcal{O} пространства \mathcal{U} множество \mathcal{F}_σ всех таких частей A из \mathcal{U} , что $A \cap \mathcal{O}$ принадлежит \mathcal{F} . Очевидно, это множество частей устойчиво относительно разности, объединения возрастающих последовательностей и пересечения убывающих последовательностей. С другой стороны, это множество содержит все открытые множества, поскольку пересечение двух открытых множеств является открытым множеством, а \mathcal{F} содержит все открытые множества. Поскольку, по предположению, \mathcal{F} является наименьшим множеством частей \mathcal{U} , обладающим указанными тремя свойствами, то множество \mathcal{F}_σ содержит \mathcal{F} . Другими словами, для каждой части A множества \mathcal{F} и каждого открытого множества \mathcal{O} из \mathcal{U} множество $A \cap \mathcal{O}$ принадлежит \mathcal{F} . Пусть теперь A — некоторая часть \mathcal{F} . Рассмотрим множество \mathcal{F}_A частей B из \mathcal{U} , таких, что $A \cap B$ принадлежит \mathcal{F} . Это множество является семейством частей, устойчивым относительно разности, объединения возрастающих последовательностей и пересечения убывающих последовательностей. Согласно только что установленному, оно содержит все открытые части \mathcal{U} . Кроме того, оно необходимо содержит все семейство \mathcal{F} . Отсюда следует, что если A и B принадлежат \mathcal{F} , то $A \cap B$ также лежит в \mathcal{F} . Это означает, что \mathcal{F} является σ -алгеброй. Поскольку \mathcal{F} содержит все открытые множества, то \mathcal{F} содержит борелевскую σ -алгебру. Однако борелевская σ -алгебра обладает всеми тремя условиями устойчивости, и \mathcal{F} — наименьшее множество частей, обладающих этим свойством. Значит, \mathcal{F} совпадает с борелевской σ -алгеброй, и лемма 1 доказана.

Вернемся теперь к доказательству теоремы. Множество \mathfrak{M} борелевских частей A , содержащихся в открытом множестве \mathcal{U} пространства X с компактным замыканием, для которых $\mu(A) = 0$, очевидно, в силу линейности μ , устойчиво относительно разности, а в силу L -непрерывности μ , устойчиво относительно объединения возрастающих последовательностей и пересечения убывающих последовательностей. С другой стороны, согласно только что доказанному перед леммой 1, \mathfrak{M} содержит все открытые части \mathcal{U} . Лемма 1 тогда показывает, что \mathfrak{M} содержит все борелевские части \mathcal{U} . Обратим внимание на одно небольшое затруднение: у нас нет уверенности в том, что часть A множества \mathcal{U} , борелевская относительно \mathcal{U} , будет борелевской относительно X . По этому поводу мы докажем вторую лемму, распространяющую теоремы 5 и 6 гл. II на борелевские части.

Лемма 2. Пусть \mathcal{U} — некоторая часть X . Всякая часть A множества \mathcal{U} , борелевская, в X , является борелевской в \mathcal{U} , и обратно, если только множество \mathcal{U} само борелевское в X .

В самом деле, рассмотрим все части B множества X , пересечения которых с \mathcal{U} и $C\mathcal{U}$ являются борелевскими множествами в топологических пространствах \mathcal{U} и $C\mathcal{U}$. Очевидно, они образуют некоторую σ -алгебру частей X . Так как эта σ -алгебра содержит все открытые множества X (поскольку они пересекаются с \mathcal{U} по открытым множествам в \mathcal{U}), то она содержит все борелевские части X . Отсюда вытекает, что каждая часть B множества X , борелевская в X и содержащаяся в \mathcal{U} , борелевская в \mathcal{U} .

Рассмотрим теперь множество частей C множества \mathcal{U} , являющихся борелевскими, если их рассматривать в X . Непосредственно видно, что если \mathcal{U} является борелевским в X , то это множество является σ -алгеброй частей \mathcal{U} . (Все свойства, необходимые для образования σ -алгебры, выполняются очевидным образом, кроме, быть может, свойства, относящегося к дополнению. Однако если часть $A \subset \mathcal{U}$ борелевская в X , то часть $C_{\mathcal{U}}A = \mathcal{U} - A = X - C\mathcal{U} - A$ будет борелевской в X , если только часть $C\mathcal{U}$ борелевская в X , т. е. если часть \mathcal{U} борелевская в X .) Поскольку оно содержит все открытые части из \mathcal{U} (ибо открытая часть \mathcal{U} , будучи пересечением \mathcal{U} с некоторой открытой частью X , является борелевским множеством в X , поскольку \mathcal{U} и открытая часть борелевские в X), то оно содержит все борелевские части \mathcal{U} . Следовательно, если C является частью \mathcal{U} , борелевской в \mathcal{U} , то она будет борелевской в X , что и утверждается в лемме 2.

В силу этой второй леммы мы теперь знаем, что функция μ равна нулю на всех борелевских частях X с компактным замыканием (т. е. на всех их характеристических функциях). Доказательство будет закончено с помощью последней леммы.

Лемма 3. *Линейное L -непрерывное отображение μ пространства $\Gamma(X)$ в \vec{E} , равное нулю на характеристических функциях множеств, является нулевым. Линейная L -непрерывная форма ($\vec{E} = \mathbb{C}$) на $\Gamma(X)$, неотрицательная на характеристических функциях множеств, неотрицательна.*

Функция μ равна нулю (соответственно ≥ 0) на всех борелевских этажных функциях (соответственно этажных ≥ 0) с компактным носителем из $\Gamma(X)$.

Согласно замечанию, следующему за теоремой 23₂, каждая борелевская функция на X с компактным носителем является равномерным пределом некоторой последовательности ограниченных в совокупности борелевских этажных функций с носителем в фиксированном компакте и ≥ 0 , если она ≥ 0 . Приме-

нение L -непрерывности показывает, что μ равна нулю на всем $\Gamma(X)$ (соответственно ≥ 0 на $\Gamma_+(X)$).

Следствие 1. Если два линейных L -непрерывных отображения μ_1 и μ_2 пространства $\Gamma(X)$ в \vec{E} совпадают на $\mathcal{C}(X)$ или на характеристических функциях множеств, то они совпадают всюду.

Следствие 2. Если μ является линейной L -непрерывной формой на $\Gamma(X)$ и если, кроме того, она положительна на $\mathcal{C}_+(X)$ или на характеристических функциях множеств, то эта функция положительна на $\Gamma_+(X)$ и является продолжением Лебега в $\Gamma(X)$ некоторой меры Радона ≥ 0 .

Положительность на характеристических функциях, в силу леммы 3, влечет за собой положительность на $\Gamma_+(X)$, а, значит, на $\mathcal{C}_+(X)$. Поэтому при обоих предположениях имеет место положительность на $\mathcal{C}_+(X)$.

Если рассматривать μ как линейную форму на $\mathcal{C}(X)$, то из L -непрерывности вытекает, что она непрерывна на каждом $\mathcal{C}_k(X)$, т. е. что она является некоторой мерой Радона, и, следовательно, мера Радона ≥ 0 . Эта линейная форма и продолжение Лебега этой положительной меры являются линейными L -непрерывными формами на $\Gamma(X)$, совпадающими на $\mathcal{C}(X)$, а, следовательно, на $\Gamma(X)$.

Пусть теперь μ — некоторая мера Радона на X со значениями в банаховом пространстве \vec{E} . Говорят, что μ допускает борелевское продолжение, если существует линейное L -непрерывное отображение пространства $\Gamma(X)$ в \vec{E} , совпадающее с μ на $\mathcal{C}(X)$. Из теоремы 50₂ вытекает, что если такое продолжение существует, то оно единственno.

Легко видеть, что такое продолжение существует не всегда. В самом деле, имеется по крайней мере одно необходимое условие продолжения: если последовательность функций из $\mathcal{C}(X)$ L -сходится к 0, то последовательность $\mu(\varphi_n)$ должна сходиться к 0 в \vec{E} . Возьмем, к примеру, «тождественную меру», определенную на стр. 489, не допускающую абсолютной мажоранты: X — компакт, $\vec{E} = \mathcal{C}(X)$, $\mu(\varphi) = \varphi$ [$\varphi \in \mathcal{C}(X)$, $\mu(\varphi) \in \vec{E} = \mathcal{C}(X)$]. Здесь L -сходимость последовательности φ_n к 0 не влечет за собой, очевидно, сходимости $\mu(\varphi_n)$ в \vec{E} , т. е. равномерной сходимости φ_n . (Пример: $X = [0, 1]$, $\varphi_n(x) = 0$ для $x = 0$ и $x \geq 1/n$, $\varphi_n(x) = 1$ для $x = 1/(2n)$ и φ_n аффинна в каждом из интервалов $[0, \frac{1}{2n}]$, $[\frac{1}{2n}, \frac{1}{n}]$.)

Теорема 50₃. *Множество $\mathcal{M}(X; \vec{E})$ мер Радона, определенных на X со значениями в банаховом пространстве \vec{E} , допускающих борелевское продолжение, является векторным пространством, а отображение $(\vec{\mu}, f) \rightarrow \int f d\vec{\mu}$ является билинейным отображением множества $\mathcal{M}(X; \vec{E}) \times \Gamma(X)$ в \vec{E} . Если \vec{E} конечномерно, то каждая мера со значениями в \vec{E} допускает борелевское продолжение.*

Доказательство. Если $\vec{\mu}$ и \vec{v} являются двумя мерами, допускающими борелевское продолжение, то $\vec{\mu} + \vec{v}$, а также $k\vec{\mu}$ при скалярном k допускают борелевское продолжение. В самом деле, достаточно определить это продолжение по формулам

$$\int f d(\vec{\mu} + \vec{v}) = \int f d\vec{\mu} + \int f d\vec{v} \quad \text{и} \quad \int f d(k\vec{\mu}) = k \int f d\vec{\mu},$$

как тем самым будут определены линейные L -непрерывные отображения $\Gamma(X)$ в \vec{E} , совпадающие соответственно с $\vec{\mu} + \vec{v}$ и $k\vec{\mu}$ на $\mathcal{C}(X)$. Одновременно доказывается, что $\mathcal{M}(X; \vec{E})$ является некоторым векторным пространством и что рассматриваемое отображение билинейно. По теории Лебега $\mathcal{M}(X; \mathbb{C})$ содержит меры Радона ≥ 0 , а, следовательно, вещественные меры как разности неотрицательных мер (теорема 18), а значит, и комплексные меры ($\mu = \mu_1 + i\mu_2$, где μ_1 и μ_2 вещественны). Если теперь μ является скалярной мерой и \vec{e} — элемент из \vec{E} , то мера μe , заданная формулой $(\mu e)(\varphi) = \vec{e}\mu(\varphi)$, допускает продолжение, определенное по формуле $\int f d(\mu e) = \vec{e} \int f d\mu$. Если пространство \vec{E} конечномерно и $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ — некоторый его базис, то каждая мера $\vec{\mu}$ со значениями в \vec{E} запишется в виде $\vec{\mu} = \sum_{i=1}^n \vec{\mu}_i \vec{e}_i$, а, следовательно, она продолжима.

Теорема 50₄. *Пусть μ — вещественная мера на X . Формулы, доказанные в теореме 18 для μ^+ , μ^- и $|\mu|$, сохраняются для борелевских продолжений:*

$$\int f d\mu^+ = \sup_{\substack{g \in \Gamma(X) \\ 0 \leq g \leq f}} \int g d\mu \quad \text{для } f \in \Gamma(X), \quad f \geq 0,$$

$$\int f d|\mu| = \sup_{\substack{g \in \Gamma(X) \\ |g| \leq f}} \int g d\mu \quad \text{для } f \in \Gamma(X), \quad f \geq 0. \quad (\text{IV, 4; 63})$$

Для каждой борелевской части A множества X с компактным замыканием имеют место равенства

$$\mu^+(A) = \sup_{B \subset A} \mu(B), \quad \mu^-(A) = \sup_{B \subset A} (-\mu(B)). \quad (\text{IV}, 4; 64)$$

Доказательство. Во всех случаях ясно, что левая часть не меньше правой. Поэтому достаточно установить неравенства обратного смысла. Прежде всего, если \mathcal{O} является открытым множеством с компактным замыканием, то существует такая функция $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, с носителем в \mathcal{O} , что $\mu^+(\varphi) \geq \mu^+(\mathcal{O}) - \varepsilon/2$. Согласно формуле (IV, 2; 41), существует такая функция $\psi \in \mathcal{C}(X)$, $0 \leq \psi \leq \varphi \leq \chi_{\mathcal{O}}$, что $\mu(\psi) \geq \mu^+(\varphi) - \varepsilon/2$. Мы имеем, таким образом, неравенство $\mu(\psi) \geq \mu^+(\mathcal{O}) - \varepsilon$, которое доказывает первую формулу в случае, когда f является характеристической функцией некоторого открытого множества (и тогда, очевидно, можно даже ограничиться функцией g из $\mathcal{C}(X)$)¹.

Пусть теперь A является произвольным борелевским множеством с компактным замыканием и характеристической функцией χ_A . Пусть \mathcal{O} — такое открытое множество, что $|\mu|(\mathcal{O} - A) \leq \varepsilon/2$. Пусть ψ — функция, определенная предыдущим способом относительно открытого множества \mathcal{O} и такая, что $\mu(\psi) \geq \mu^+(\mathcal{O}) - \varepsilon/2$, а, значит, $\geq \mu^+(A) - \varepsilon/2$. Рассмотрим теперь функцию (разрывную) $g = \psi \chi_A$, $0 \leq g \leq \chi_A$. Из неравенства $\int \psi \chi_A d\mu = \mu(\psi) - \int \psi \chi_{\mathcal{O}-A} d\mu \geq \mu(\psi) - |\mu|(\mathcal{O} - A) \geq \mu(\psi) - \varepsilon/2 \geq \mu^+(A) - \varepsilon$ вытекает первая формула для случая, когда f представляет собой характеристическую функцию множества A . Используя соответствующие линейные комбинации, легко убедиться, что эта формула верна для любой этажной борелевской функции $f \geq 0$ с компактным носителем.

Пусть, наконец, $f \geq 0$ — произвольная функция из $\Gamma(X)$, M — ее точная верхняя грань и K — ее носитель. Эта функция является равномерным пределом последовательности этажных борелевских неотрицательных функций с носителем в K , ограниченных числом M . Каково бы ни было $\varepsilon > 0$, можно определить такую этажную борелевскую неотрицательную функцию h с носителем в K и ограниченную числом M , что $|f - h| \leq \delta = \varepsilon/(4\mu^+(K))$. Если мы положим $k = (h - \delta)^+$, то получим $0 \leq k \leq f$ и $f - k \leq 2\delta$. Отсюда $\int k d\mu^+ \geq \int f d\mu^+ - 2\delta\mu^+(K) = \int f d\mu^+ - \varepsilon/2$.

¹⁾ Если f полунепрерывна снизу, то можно ограничиться непрерывной функцией g , но только не тогда, когда f произвольна, даже если уже $\mu \geq 0$ (см. замечание после теоремы 39₂).

Только что мы видели, что существует такая функция $g \in \Gamma(X)$, $0 \leq g \leq k$, а, значит, $\leq f$, удовлетворяющая соотношению $\int g d\mu \geq \int k d\mu^+ - \varepsilon/2$ и, следовательно, $\geq \int f d\mu^+ - \varepsilon$.

Первая формула (IV, 4; 63) полностью доказана. Вторая формула (IV, 4; 63) доказывается точно так же. Остается доказать соотношения (IV, 4; 64).

Пусть A — борелевская часть множества X с компактным замыканием. Пусть задано $\varepsilon > 0$. Возьмем произвольные числа $\delta, \eta > 0$. Позже мы установим, что они зависят от ε . Согласно (IV, 4; 63), можно найти такую функцию $g \in \Gamma(X)$, $0 \leq g \leq \chi_A$, что $\mu^+(A) \geq \int g d\mu \geq \mu^+(A) - \delta$. Отсюда вытекает, что $\mu^+(A) \geq \geq \int g d\mu^+ \geq \mu^+(A) - \delta$, а, следовательно, $\int g d\mu^- = \int g d\mu^+ - \int g d\mu \leq \delta$.

Обозначим через B борелевское множество точек x , в которых $g(x) \geq \eta$. Так как функция $\eta\chi_B$ заключена между 0 и g , то $\eta\mu^-(B) = \int \eta\chi_B d\mu^- \leq \int g d\mu^- \leq \delta$, а, значит, $\mu^-(B) \leq \delta/\eta$. Если теперь рассмотреть функцию k , равную 1 на B , η на $A - B$ и 0 на $\complement A$, то она будет заключена между g и χ_A ; поэтому будут иметь место соотношения $\mu^+(B) + \eta\mu^+(A - B) = \int k d\mu^+ \geq \geq \int g d\mu^+ \geq \mu^+(A) - \delta$ и, следовательно, $\mu^+(B) \geq \mu^+(A) - \delta - \eta\mu^+(A - B) \geq \mu^+(A) - \delta - \eta\mu^+(A)$.

Окончательно получаем:

$$\mu(B) = \mu^+(B) - \mu^-(B) \geq (\mu^+(A) - \delta - \eta\mu^+(A)) - \frac{\delta}{\eta}.$$

Множество B зависит от выбора δ и η . Однако если мы выберем δ, η так, чтобы $\delta + \eta\mu^+(A) + \delta/\eta \leq \varepsilon$, то получим такое борелевское множество $B \subset A$, что $\mu(B) \geq \mu^+(A) - \varepsilon$, и формула (IV, 4; 64) будет доказанной. Такой выбор δ, η возможен. Надо сначала взять $\eta = \varepsilon/(2\mu^+(A))$, т. е. $\eta\mu^+(A) = \varepsilon/2$, а затем взять δ так, чтобы $\delta(1 + 1/\eta) = \varepsilon/2$, или $\delta = \varepsilon/[2(1 + 1/\eta)]$. Таким образом, теорема 50₄ полностью доказана.

В дальнейшем окажется полезным следующее понятие.

Пусть μ — некоторая вещественная мера. Говорят, что она *вполне положительна* (соответственно *вполне отрицательна*, *вполне нулевая*) на некоторой части A множества X , если для каждой борелевской части B множества A с компактным замыканием имеет место неравенство $\mu(B) \geq 0$ (соответственно $\leq 0, = 0$).

Тот факт, что μ вполне положительна (соответственно отри-

цательна) на A , записывается в виде: $\mu(A) \geq 0$ (соответственно $\mu(A) \leq 0$). Если μ вполне положительна на A , то она заведомо вполне положительна на любой части A . Если μ вполне положительна на счетном множестве борелевских частей, то она вполне положительна на их объединении A . (В самом деле, если $B \subset \bigcup_{n=0}^{\infty} A_n$, то $B = \bigcup_{n=0}^{\infty} B_n$, где B_n — непересекающиеся борелевские множества и $B_n \subset A_n$: $B_0 = B \cap A_0$, $B_1 = B \cap A_1 \cap CA_0$, ..., $B_n = B \cap A_n \cap CA_{n-1} \cap \dots \cap CA_0$. Теперь $\mu(B) = \sum_{n=0}^{\infty} \mu(B_n) \geq 0$.)

Для того чтобы $\mu(A)$ было ≥ 0 , необходимо и достаточно, чтобы A имело для μ^- нижнюю меру, равную нулю.

В самом деле, пусть A имеет для μ^- нулевую нижнюю меру. Для каждого борелевского множества $B \subset A$ с компактным замыканием будет справедливо равенство $\mu^-(B) = 0$, а, следовательно $\mu(B) = \mu^+(B) \geq 0$, и значит, μ вполне положительна на A . Обратно, предположим, что $\mu(A) \geq 0$. Тогда $\mu(K) \geq 0$ для каждого компактного подмножества $K \subset A$. Из формулы (IV, 4; 64), примененной к K и μ^- , получаем $\mu^-(K) = \sup_{B \subset K} (-\mu(B)) = 0$, а, следовательно, $\mu^-(A) = 0$.

Теорема 505. Если μ является вещественной мерой на локально компактном счетном в бесконечности пространстве X , то это пространство можно представить в виде объединения двух борелевских дополнительных множеств X^+ и X^- , таких, что мера μ будет вполне положительной на X^+ и вполне отрицательной на X^- .

Доказательство. Пусть K — некоторый компакт X . В предыдущей теореме мы видели, что $\mu^+(K) = \sup_{A \subset K} \mu(A)$. Значит, можно определить последовательность таких борелевских частей пространства X , содержащихся в K : A_0, A_1, A_2, \dots , что $\mu(A_n) \geq \mu^+(K) - 1/2^n$.

Положим теперь $B_n = \bigcup_{p \geq n} A_p$. Имеют место соотношения:

$$\begin{aligned} \mu^+(A_n) &\geq \mu^+(K) - \frac{1}{2^n}, \quad \text{а следовательно,} \\ \mu^+(B_n) &= \mu^+(K); \\ \mu^-(A_n) &= \mu^+(A_n) - \mu(A_n) \leq \frac{1}{2^n}, \quad \text{поэтому} \end{aligned} \tag{IV, 4; 65}$$

$$\mu^-(B_n) \leq \frac{1}{2^{n-1}}.$$

Если мы через K^+ обозначим пересечение $\bigcap_{n \geq 0} B_n$ и через K^- — множество $K - K^+$, то сразу получим равенства

$$\begin{aligned}\mu^+(K^+) &= \mu^+(K), & \mu^-(K^+) &= 0, \\ \mu^+(K^-) &= 0, & \mu^-(K^-) &= \mu^-(K).\end{aligned}\quad (\text{IV}, 4; 66)$$

Тем самым мы определили две искомые части для компакта K . Пусть теперь K_0, K_1, K_2, \dots — возрастающая последовательность компактов, объединение которых равно X . Для каждого из K_n по предыдущему способу можно определить части K_n^+ и K_n^- . Положим теперь $X^+ = \bigcup_{n \geq 0} K_n^+$ и $X'^- = \bigcup_{n \geq 0} K_n^-$. Поскольку

мера μ вполне положительна на K_n^+ , то она вполне положительна на их объединении X^+ , а поскольку она вполне отрицательна на K_n^- , то она вполне отрицательна на X'^- .

Нельзя быть уверенным, что пересечение X^+ и X'^- равно нулю. Однако если это пересечение обозначить через Y , то мера μ на Y будет одновременно вполне положительной и вполне отрицательной, т. е. будет вполне нулевой. Заменяя затем X^+ и X'^- на X^+ и $X^- = X'^- - Y$, получим множества, удовлетворяющие условиям теоремы.

Замечание. Интересно выяснить степень неопределенности только что найденной системы множеств X^+, X^- . Пусть Y^+, Y^- — другое решение этой же задачи. Множество $X^+ \cap Y^-$ точек, принадлежащих X^+ и не принадлежащих Y^+ , или множество $X^- \cap Y^+$ элементов, принадлежащих Y^+ и не принадлежащих X^+ , являются частями X , на которых μ одновременно вполне положительна и вполне отрицательна, т. е. вполне нулевая. Можно сказать, в некотором смысле, что X^+ и Y^+ отличаются на некоторое множество вполне нулевой меры по μ , т. е. нулевой меры для $|\mu|$. То же самое имеет место для X^- и Y^- . Обратно, совершенно очевидно, что если изменить X^+ и X^- на множества, на которых μ вполне нулевая, то их свойства не изменятся.

Следствие. Для того чтобы две меры $\mu, \nu \geq 0$ были дизъюнктными (т. е. такими, что $\inf(\mu, \nu) = 0$), необходимо и достаточно, чтобы можно было представить X в виде объединения таких двух множеств X_μ и X_ν , что $\nu(X_\mu) = 0$ и $\mu(X_\nu) = 0$, и тогда их можно выбрать борелевскими и дополнительными друг к другу.

Доказательство. Предположим, что указанное представление возможно. Тогда для каждой меры $\lambda \geq 0$, мажорируемой одновременно мерами μ и ν , справедливы неравенства $\lambda(X_\mu) \leq \nu(X_\mu) = 0$ и $\lambda(X_\nu) \leq \mu(X_\nu) = 0$, т. е. $\lambda(X) = 0$, а, зна-

чит, $\lambda = 0$. Так как $\inf(\mu, v) \geq 0$, то этот $\inf = 0$, и меры μ и v — дизъюнктны. Обратно, пусть μ и v — дизъюнктные меры. Положим $\lambda = \mu - v$. Из $\inf(\mu, v) = 0$ следует, что $\sup(\mu, v) = \mu + v$ (см. стр. 481). Но тогда $\lambda^+ = \sup(\mu - v, 0) = \sup(\mu, v) - v = \mu + v - v = \mu$ и точно так же $\lambda^- = v$. Нужный результат получается, если положить $X_\mu = X^+$ и $X_v = X^-$, где оба множества борелевские и дополнительные друг другу.

Замечание. Говорят, что некоторая мера $\mu \geq 0$ *сосредоточена на части A пространства X*, если C_A имеет μ -меру, равную нулю. Это определение не имеет прямой связи с носителем. Носитель является *наименьшим замкнутым множеством, на котором сосредоточена мера μ* . Однако часть A не обязательно замкнута. Например, если a_v являются рациональными числами из \mathbb{R} и если ряд $\sum_v c_v$, $c_v > 0$, сходится, то мера $\sum_v c_v \delta_{(a_v)}$ сосредоточена на \mathbb{Q} , но ее носителем является \mathbb{R} . Впрочем, не существует наименьшего множества, на котором сосредоточена мера μ , кроме случая атомических мер (стр. 468), ибо из A всегда можно удалить произвольное множество нулевой μ -меры. Например, меру dx на \mathbb{R} можно считать сосредоточенной на множестве иррациональных чисел. Теперь предыдущее следствие может быть записано в следующем виде:

Для того чтобы меры μ и v (≥ 0) были дизъюнктными, необходимо и достаточно, чтобы существовало два непересекающиеся множества из X , на которых сосредоточены соответственно меры μ и v : на дополнении CX_μ сосредоточена мера v , а на дополнении CX_v сосредоточена мера μ (т. е. если X_μ и X_v — дополнительные множества, то на X_μ сосредоточена мера μ , а на X_v сосредоточена мера v).

Например, если μ является некоторой атомической мерой ≥ 0 на X (т. е. имеет вид $\sum c_v \delta_{(a_v)}$, где $\sum_{a_v \in K} c_v < +\infty$ для каждого

компакта K ; здесь a_v образуют бесконечное счетное множество на K , а, значит, и на объединении X счетного множества компактов) и если v является некоторой рассеянной мерой (т. е. такой, для которой каждая точка, a , значит, и всякое счетное множество, имеет нулевую меру), то меры μ и v являются дизъюнктными, поскольку они сосредоточены соответственно на множестве точек a_v и его дополнении.

Можно также дать следующее эквивалентное условие, которое выглядит асимметричным.

Для того чтобы меры μ и v (≥ 0) были дизъюнктными, необходимо и достаточно, чтобы мера v была μ -сингулярной, т. е. сосредоточенной на множестве, μ -мера которого равна нулю.

В самом деле, сказать, что на множестве A нулевой μ -меры сосредоточена мера ν , означает сказать, что $\mu(A) = 0$ и $\nu(\overrightarrow{CA}) = 0$.

§ 5. УМНОЖЕНИЕ МЕРЫ НА ФУНКЦИЮ

Произведение векторной меры на непрерывную скалярную функцию

Пусть μ — мера на локально компактном пространстве X со значениями в векторном нормированном пространстве \vec{E} . Пусть, кроме того, p — непрерывная скалярная функция на X . *Произведением меры μ на функцию p* , обозначаемым через $\vec{\mu}p$, или $d(p\vec{\mu})$, или $p\vec{d\mu}$, называется *мера ν на X со значениями в \vec{E}* , определяемая по формуле

$$\vec{\nu}(\varphi) = \vec{\mu}(p\varphi) = \vec{\mu}(p\varphi), \quad \varphi \in \mathcal{C}(X). \quad (\text{IV}, 5; 1)$$

Эта формула действительно определяет некоторую меру. В самом деле, пусть φ принадлежит множеству $\mathcal{C}(X)$. Поскольку функция p непрерывна, то функция $p\varphi$ также непрерывна, а ее носитель, содержащийся в носителе функции φ , компактен. Поэтому $p\varphi$ также лежит в $\mathcal{C}(X)$, так что правая часть в (IV, 5; 1) имеет смысл. Эта часть линейно зависит от φ , и поэтому ν является линейным отображением $\mathcal{C}(X)$ в \vec{E} .

Остается доказать непрерывность этого отображения. Предположим, что носитель φ лежит в компакте K множества X . Тогда справедливо неравенство

$$\|\vec{\nu}(\varphi)\| = \|\vec{\mu}(p\varphi)\| \leq \|\vec{\mu}\|_K \sup_{x \in K} |p(x)| \|\varphi\|, \quad (\text{IV}, 5; 2)$$

доказывающее непрерывность функции ν и дающее одновременно оценку

$$\|\vec{\nu}\|_K \leq \|\vec{\mu}\|_K \sup_{x \in K} |p(x)|. \quad (\text{IV}, 5; 3)$$

Это определение показывает также, почему нельзя умножать меру на функцию, не являющуюся непрерывной скалярной функцией. В самом деле, правая часть в (IV, 5; 1) имеет смысл только в том случае, когда $p\varphi$ является непрерывной скалярной функцией с компактным носителем, а это верно лишь тогда, когда p является непрерывной скалярной функцией.

Пример 1. Если $\vec{\mu}$ является мерой $\sum_v \vec{c}_v \delta_{(a_v)}$, то имеет место формула

$$p \left(\sum_v \vec{c}_v \delta_{(a_v)} \right) = \sum_v \vec{c}_v p(a_v) \delta_{(a_v)}. \quad (\text{IV}, 5; 4)$$

В частности, $p\delta_{(a)} = p(a) \delta_{(a)}$.

Пример 2. Пусть $\vec{\mu}$ — мера, обозначавшаяся нами ранее через $\vec{p} dx$, и q — непрерывная скалярная функция. Тогда мера $q(\vec{p} dx)$ является уже известной мерой $(qp) dx$. В частности, если p является непрерывной скалярной функцией, то pdx является мерой-произведением непрерывной скалярной функции p на скалярную меру dx .

Элементарные свойства

Z Носитель меры $\vec{p}\vec{\mu}$, очевидно, содержится в пересечении носителя p и носителя $\vec{\mu}$. Он может быть строго меньше, чем это пересечение. Возьмем, например, в качестве $\vec{\mu}$ вещественную меру Дирака δ на \mathbb{R} , а в качестве p функцию x . Тогда, согласно (IV, 5; 4), $x\delta = 0$.

Здесь носителями являются соответственно нуль и вся прямая, и они пересекаются в нуле, в то время как их произведение, будучи равным нулю, имеет пустой носитель.

Отображение $(p, \vec{\mu}) \rightarrow \vec{p}\vec{\mu}$, ставящее в соответствие каждой паре, состоящей из непрерывной скалярной функции и меры со значениями в \vec{E} , некоторую меру со значением в \vec{E} , является, очевидно, билинейным.

Наконец, имеет место следующее правило ассоциативности: если p и q являются двумя непрерывными скалярными функциями, то

$$p(q\vec{\mu}) = (pq)\vec{\mu}. \quad (\text{IV}, 5; 5)$$

Случай, когда $\vec{\mu}$ — вещественная мера $\geqslant 0$

В том частном случае, когда $\vec{\mu}$ является вещественной мерой Радона $\geqslant 0$, известно значение $\vec{\mu}$ на функциях, не являющихся обязательно непрерывными с компактным носителем. Поэтому в данном случае произведение можно определить в более широком смысле и даже умножать $\vec{\mu}$ на функции с векторными значениями.

Определение. Пусть μ — вещественная мера Радона $\geqslant 0$ на локально компактном, счетном в бесконечности пространстве X . Пусть \vec{p} — функция, определенная μ -почти всюду на X , со значениями в банаховом пространстве \vec{E} и локально μ -интегрируемая, т. е. μ -интегрируемая на каждом компакте из X ¹⁾. Тогда *произведением* $\vec{p}\mu$ называется мера ν на X со значениями в \vec{E} , определенная по формуле

$$\vec{p}\mu(\varphi) = \int \vec{p}\varphi d\mu \in \vec{E}. \quad (\text{IV}, 5; 6)$$

Покажем, что эта формула определяет некоторую меру на X со значениями в \vec{E} . Если φ принадлежит $\mathcal{C}(X)$, то она, в частности, μ -измерима и ограничена, а ее носитель содержится в компакте K пространства X . Так как функция \vec{p} , по предположению, интегрируема на K , то функция $\vec{p}\varphi$ (следствие 2 теоремы 23₂) μ -измерима, а верхний интеграл $\int_K^* \|\vec{p}\varphi\| d\mu \leqslant \|\varphi\| \int_K \|\vec{p}\| d\mu$

конечен. Отсюда следует, что функция $\vec{p}\varphi$ интегрируема и, следовательно, правая часть соотношения (IV, 5; 6) имеет смысл.

Естественно, что эта правая часть линейно зависит от φ и тем самым определяет некоторое линейное отображение $\mathcal{C}(X)$ в \vec{E} . Что же касается непрерывности, то она очевидным образом следует из последнего неравенства, которое, кроме того, означает, что норма меры $\vec{p}\mu$ на компакте K допускает оценку

$$\|\vec{p}\mu\|_K \leqslant \int_K \|\vec{p}\| d\mu. \quad (\text{IV}, 5; 7)$$

Переходя к точной верхней грани по всевозможным компактам K , отсюда получаем:

$$\|\vec{p}\mu\| \leqslant \int \|\vec{p}\| d\mu. \quad (\text{IV}, 5; 7_2)$$

Можно показать, что неравенство переходит в равенство в том случае, когда \vec{p} принимает скалярные значения или лежит в одномерном векторном пространстве. Однако это, вообще го-

1) Функция \vec{p} локально μ -интегрируема тогда и только тогда, когда она μ -измерима и когда для каждого компакта K из X имеет место неравенство $\int_K^* \|\vec{p}\| dx < +\infty$.

воля, не так, если \hat{r} принимает значения в пространстве \vec{E} размерности $\geqslant 2$.

Если r является функцией с комплексными значениями (соответственно вещественными, соответственно вещественными $\geqslant 0$), то же самое будет иметь место и для произведения $r\mu$. Здесь носитель $r\mu$ содержится в пересечении носителей r и μ . Кроме того, мы полностью обосновали теперь применявшееся ранее обозначение pdx для некоторых мер на прямой.

Теорема 51. Пусть μ является вещественной мерой $\geqslant 0$ на локально компактном, счетном в бесконечности пространстве X , r — вещественная функция $\geqslant 0$, локально μ -интегрируемая на X , и \hat{f} — функция, определенная на X , со значениями в банаховом пространстве \vec{E} . Функция \hat{f} μ -интегрируема тогда и только тогда, когда μ -интегрируема функция \hat{rf} . Кроме того,

$$\int \hat{f} d(r\mu) = \int (p\hat{f}) d\mu^1 \quad \text{или} \quad \int \hat{f} (p d\mu) = \int (\hat{f} p) d\mu. \quad (\text{IV}, 5; 8)$$

Мы примем эту теорему без доказательства. Ее доказательство основывается на следующем равенстве, относящемся к верхним интегралам: если \hat{f} — произвольная функция $\geqslant 0$, то

$$\int^* \hat{f} (p d\mu) = \int^* (\hat{f} p) d\mu. \quad (\text{IV}, 5; 9)$$

Эта теорема, очевидно, особенно полезна, и ею постоянно пользуются на практике, зачастую этого и не замечая.

В частности, если для функций с вещественными значениями $\geqslant 0$ рассматривается выражение $\int fg d\mu$, то безразлично, будем ли мы рассматривать его как интеграл от fg по $d\mu$, или как интеграл от f по мере-произведению $g d\mu$, или как интеграл от g по мере $f d\mu$. Все это, конечно, при условии, что все эти меры-произведения имеют смысл, т. е. что функции f и g локально μ -интегрируемы.

Следствие. Если \hat{r} является функцией на X со значениями в \vec{E} , q — вещественной локально μ -интегрируемой функцией $\geqslant 0$, то функция \hat{r} локально $q\mu$ -интегрируема тогда и только тогда, когда $\hat{r}\hat{q}$ локально μ -интегрируема и меры-произведения $\hat{r}(q\mu)$ и $(\hat{r}\hat{q})\mu$ совпадают.

¹⁾ Легко видеть, почему рекомендуется применять обозначение $p\mu$. В отношении (IV, 5; 8) $d(r\mu)$ совпадает с $p\mu$. Мы уже писали pdx ,

В самом деле, они имеют одно и то же значение на каждой функции $\varphi \in \mathcal{C}(X)$. Говорят, что мера $r\mu$ имеет плотность r по отношению к вещественной мере $\mu \geq 0$.

Меры с базой μ . Меры с базой ≥ 0

Говорят, что мера v , определенная на X , со значениями в банаховом пространстве \vec{E} является мерой с базой μ , где μ — вещественная мера ≥ 0 , если она может быть записана в виде $r\mu$, где r — некоторая локально μ -интегрируемая функция на X со значениями в \vec{E} . Говорят, что v является мерой с базой ≥ 0 , если существует такая мера $\mu \geq 0$, что v есть мера с базой μ .

Теперь можно рассмотреть следующие вопросы:

1°) Сколькими способами может быть выражена одна и та же мера v в виде $r\mu$, где мера μ задана?

2°) Каков характер мер, которые могут выражаться в этом виде, т. е. мер с базой μ ?

Теоремы 52—53—54 (Лебега — Никодима и Данфорда — Петтиса)¹⁾. Пусть X — локально компактное счетное в бесконечности пространство и μ — некоторая мера Радона ≥ 0 на X .

1°) Пусть v — некоторая другая мера ≥ 0 на X . Если v имеет базу μ , то каждая μ -измеримая функция и μ -измеримое множество v -измеримы. Кроме того, каждое множество нулевой μ -меры также имеет нулевую v -меру. Обратно, если каждая борелевская часть с компактным замыканием и нулевой μ -мерой имеет также нулевую v -меру, то v является мерой с базой μ , $v = r\mu$, где r — единственная с точностью до значений на множестве нулевой μ -меры функция. (Если $r_1\mu = r_2\mu$, то r_1 и r_2 μ -почти всюду равны.)

Для того чтобы мера $v \geq 0$ была мажорируема мерой μ , необходимо и достаточно, чтобы $v = r\mu$, где $r \leq 1$ μ -почти всюду.

2°) Каждая мера v со значениями в банаховом пространстве \vec{E} с базой ≥ 0 допускает борелевское продолжение, абсолютно мажорируемое и имеющее наименьшую абсолютную мажоранту. Если предполагается, что v имеет базу ≥ 0 , то для того, чтобы она имела базу μ , необходимо и достаточно, чтобы каждое бо-

¹⁾ Мы собрали несколько теорем в одну. Часть 1°) принадлежит Лебегу и Никодиму, часть 2°) является частным случаем более общей теоремы Данфорда и Петтиса.

релевское подмножество с компактным замыканием и нулевой μ -мерой имело также нулевую ν -меру. При этих условиях $\nu = \vec{p}\mu$, где \vec{p} определяется с точностью до изменений на множестве нулевой μ -меры. Для того чтобы она была абсолютно мажорируема мерой μ , необходимо и достаточно, чтобы $\nu = \vec{p}\mu$ и $\|\vec{p}\| \leq 1$ μ -почти всюду. При этом для каждой ограниченной борелевской функции f с компактным носителем имеет место неравенство

$$\left\| \int f d\nu \right\| \leq \int |f| d\mu. \quad (\text{IV}, 5; 9_2)$$

Для того чтобы эта функция имела μ в качестве наименьшей абсолютной мажоранты, необходимо и достаточно, чтобы $\nu = \vec{p}\mu$, где $\|\vec{p}\| = 1$ μ -почти всюду.

3°) Если E конечномерно, то каждая мера ν со значениями в E является мерой с базой ≥ 0 . В частности, каждая скалярная мера имеет базу ≥ 0 .

4°) Если p и q являются вещественными, локально μ -интегрируемыми функциями на X , то $p\mu \leq q\mu$ тогда и только тогда, когда $p \leq q$ μ -почти всюду. Мера $(\sup_{\vec{p}}(p, q))\mu$ является точной верхней гранью мер $p\mu$ и $q\mu$. Если ν является мерой с базой ≥ 0 и со значениями в E , то для того, чтобы она была абсолютно мажорируема мерой $q\mu$, необходимо и достаточно, чтобы $\nu = p\mu$ и $\|\vec{p}\| \leq q$ μ -почти всюду.

Доказательство. А) Предположим сначала, что $\nu \geq 0$, $\nu = p\mu$, $p \geq 0$ локально μ -интегрируема. Пусть A есть μ -измеримая часть с компактным замыканием. Пусть \mathcal{O}_n — последовательность таких открытых множеств, содержащих множество A , что $\mu(\mathcal{O}_n) \leq \mu(A) + 1/n$, и пусть K_n — последовательность таких компактов, содержащихся в A , что $\mu(K_n) \geq \mu(A) - 1/n$. Пусть $A^* = \bigcap_n \mathcal{O}_n$ и $A_* = \bigcup_n K_n$. Множества A^* и A_* являются борелевскими, а, значит, измеримыми множествами при любой мере Радона, $A_* \subset A \subset A^*$ и $\mu(A^* - A_*) = 0$. Тогда $\nu(A^* - A_*) = \int_{A^* - A_*} p d\mu = 0$. Так как $\nu(A_*) \leq \nu(A) \leq \nu^*(A) \leq \nu(A^*)$, то

часть A также ν -измерима. Если часть A не имеет компактного замыкания, то такое заключение сохраняется в силу того, что измеримость A эквивалентна измеримости ее пересечений с компактами, при этом если $\mu(A) = 0$, то $\nu(A) = \int_A p d\mu = 0$.

Измеримость функций сводится к измеримости множеств. Каждая μ -измеримая функция является v -измеримой.

Обратно, предположим, что каждое борелевское множество с компактным замыканием и нулевой μ -мерой имеет также нулевую v -меру. Тогда мы сможем построить некоторую функцию $p \geq 0$, локально μ -интегрируемую и такую, что $v = p\mu$. При этом v будет мерой с базой μ . Если $v \leq \mu$, то из построения будет видно, что $p \leq 1$. Первая часть теоремы, однако, еще не будет доказана, поскольку нужно показать, что p определяется с точностью до произвольных изменений на множестве нулевой μ -меры.

Построим искомую функцию p . Пусть k и n — целые числа ≥ 0 . Если мы введем меру $v - \frac{k}{2^n}\mu$, то, согласно теореме 50₅, множество X можно представить в виде объединения двух таких непересекающихся борелевских множеств $X_{k,n}^+$ и $X_{k,n}^-$, что на первом мера $v - \frac{k}{2^n}\mu$ будет вполне положительной, а на втором — вполне отрицательной. Для $k = 0$ мы положим $X_{0,n}^+ = X$ и $X_{0,n}^- = \emptyset$.

Пусть теперь $A_{k,n} = X_{k,n}^+ \cap X_{k+1,n}^-$, и пусть $h < k$. Нельзя быть уверенным в том, что $X_{h,n}^- \subset X_{k,n}^-$. Однако если через $N_{h,k,n}$ обозначить борелевское множество $X_{h,n}^- \cap X_{k,n}^+$ точек, принадлежащих первому множеству, но не принадлежащих второму, то это множество будет иметь нулевую μ -меру. В самом деле, для каждой борелевской части B с компактным замыканием данного множества имеют место неравенства

$$v(B) \leq \frac{h}{2^n}\mu(B) \quad \text{и} \quad v(B) \geq \frac{k}{2^n}\mu(B),$$

что возможно только в том случае, когда $\mu(B) = 0$, а значит, $\mu_*(N_{h,k,n}) = 0$. Так как это множество борелевское, а значит, μ -измеримо, то $\mu(N_{h,k,n}) = 0$. Если мы обозначим через N_n объединение всех $N_{h,k,n}$, соответствующих всем значениям h, k , $h < k$, то получим некоторое множество нулевой μ -меры. Если x не принадлежит этому исключительному множеству N_n и если $x \in X_{h,n}^-$, то $x \in X_{k,n}^-$ для всех $k \geq h$. Отсюда вытекает, что пересечение $A_{h,n}$ и $A_{k,n}$, $h \neq k$, содержится в N_n , а следовательно, имеет нулевую μ -меру. (В самом деле, предположим, например, что $h < k$. Тогда $(A_{h,n} \cap A_{k,n}) \subset (X_{h+1,n}^- \cap X_{k,n}^+) \subset N_n$.)

Объединение всех $A_{k,n}$ при фиксированном n и $k = 0, 1, 2, \dots$ содержит почти все X . В самом деле, пусть $x \in X$. Предположим сначала, что существует хотя бы одно такое целое число h , что $x \in X_{h,n}^-$. Тогда найдется первое целое h , обладающее этим свойством (поскольку $X_{0,n}^- = \emptyset$, то $h \geq 1$). При этом $x \notin X_{h-1,n}^-$,

и, следовательно, $x \in X_{h-1, n}^+$, а потому $x \in A_{h-1, n}$. Однако если мы рассмотрим множество M_n точек x , не принадлежащих ни одному из $X_{k, n}^-$, $k = 0, 1, 2, \dots$, то это будет множество точек, принадлежащих всем $X_{k, n}^+$. Следовательно, для каждой борелевской части B с компактным носителем множества M_n при любом k имеет место неравенство $\nu(B) \geq \frac{k}{2^n} \mu(B)$, из которого следует, что $\mu(B) = 0$, так что нижняя μ -мера множества M_n равна нулю. Поскольку это множество борелевское, то его μ -мера также равна нулю. Все это можно подытожить следующим образом: для фиксированного n и $k=0, 1, 2, \dots$ множества $A_{k, n}$, с точностью до множеств нулевой μ -меры, попарно не пересекаются и в объединении дают множество X . Обозначим теперь через p_n функцию, равную $k/2^n$ на множестве $A_{k, n}$. Ее невозможно будет определить по этому правилу в точках пересечения двух множеств $A_{k, n}$ или в точках, не принадлежащих объединению этих множеств, т. е. на борелевском множестве $N_n \cup M_n$ нулевой μ -меры. Положим на этом множестве, например, $p_n = 0$. Тогда функция p_n будет борелевской. Кроме того, она будет и локально μ -интегрируемой, поскольку для каждого компакта K

$$\int_K^* p_n d\mu = \sum_{n \geq 0} \int_{K \cap A_{k, n}} \frac{k}{2^n} d\mu \leq \sum_{n \geq 0} \nu(K \cap A_{k, n}) \leq \nu(K) < +\infty.$$

Таким образом, мы можем полностью определить меру $p_n \mu$. На каждом из множеств $A_{k, n}$ имеют место неравенства $p_n \mu \ll \nu \ll (p_n + 1/2^n) \mu$. Эти же неравенства сохраняются на $\bigcup_{k \geq 0} A_{k, n}$.

Множества M_n имеют нулевую μ -меру, а, следовательно, в силу сделанного предположения относительно ν и μ , они имеют нулевую ν -меру (это используется только здесь). Значит, неравенства справедливы на M_n и, следовательно, справедливы на всем X .

Будем теперь изменять n . Выполнив построение для n , мы его можем использовать для $n+1$, полагая $X_{2k, n+1}^\pm = X_{k, n}^\pm$. Это возможно, поскольку на этих множествах должны выполняться одинаковые неравенства

$$\nu - \frac{k}{2^n} \mu = \nu - \frac{2k}{2^{n+1}} \mu \ll 0 \quad \text{или} \quad \gg 0.$$

Множества $A_{2k, n+1}$ и $A_{2k+1, n+1}$, с точностью до множеств μ -меры нуль, попарно не пересекаются и в объединении дают множество $A_{k, n}$. Поэтому μ -почти всюду имеем $p_n \leq p_{n+1} \leq p_n + 1/2^n$. Если мы изменим функции p_n , полагая их равными нулю на борелевском множестве нулевой μ -меры $M \cup N$, $M = \bigcup_{n \geq 0} M_n$,

$N = \bigcup_{n \geq 0} N_n$, то они будут по-прежнему обладать теми же свойствами, однако всюду будут иметь место неравенства $p_n \leq p_{n+1} \leq p_n + 1/2^n$. Последовательность функций p_n оказывается возрастающей и имеющей некоторый предел p , где $p_n \leq p \leq p_n + 1/2^n$. Функция p является борелевской и локально μ -интегрируемой, а, значит, определяющей меру $p\mu$. Для каждого φ из $\mathcal{C}(X)$, $\varphi \geq 0$, имеем

$$\int p_n \varphi d\mu \leq \int \varphi dv \leq \int p_n \varphi d\mu + \frac{1}{2^n} \int \varphi d\mu,$$

$$\int p_n \varphi d\mu \leq \int p \varphi d\mu \leq \int p_n \varphi d\mu + \frac{1}{2^n} \int \varphi d\mu,$$

откуда

$$\left| \int p \varphi d\mu - \int \varphi dv \right| \leq \frac{1}{2^n} \int \varphi d\mu.$$

Поскольку n произвольно, левая часть равна нулю. Значит, $p\mu$ и v совпадают на $\mathcal{C}_+(X)$ и, следовательно, на $\mathcal{C}(X)$, т. е. равны. Это как раз то, что мы хотели доказать. Если, кроме того, $v \leq \mu$, то можно взять $X_{2^n, n}^- = X$ и $X_{2^n, n}^+ = \emptyset$, а, значит, $p_n \leq 1$ и $p \leq 1$.

В) Пусть \vec{v} — мера со значениями в пространстве Банаха \vec{E} и с базой $\lambda \geq 0$, так что $\vec{v} = \vec{q}\lambda$, где \vec{q} локально λ -интегрируема. Эта мера тривиальным образом абсолютно мажорируется мерой $\|\vec{q}\| \lambda \geq 0$. Согласно теореме 18, она должна иметь наименьшую абсолютную мажоранту. Она также допускает борелевское продолжение, определяемое формулой

$$\int f d\vec{v} = \int \vec{f} \vec{q} d\lambda. \quad (\text{IV}, 5; 9_3)$$

Мы определили некоторое линейное отображение $\Gamma(X)$ в \vec{E} , продолжающее отображение, определенное мерой \vec{v} на $\mathcal{C}(X)$, поскольку если функция f борелевская, а значит, λ -измерима, то и функция $\vec{f} \vec{q}$ λ -измерима. Функция $|f|$ ограничена числом $M \geq 0$ и имеет компактный носитель K , а функция \vec{q} локально λ -интегрируема, поэтому $\chi_K \vec{q}$ λ -интегрируема. Значит, $\|\vec{f} \vec{q}\| \leq M \chi_K \|\vec{q}\|$ имеет конечный верхний λ -интеграл, а, следовательно, $\vec{f} \vec{q}$ λ -интегрируема. Остается доказать L -непрерывность этого отображения. Предположим, что функции f_n сходятся к f , оставаясь ограниченными по модулю числом M и сохраняя свои носители в компакте K . Функции $\vec{f}_n \vec{q}$ сходятся к $\vec{f} \vec{q}$, оставаясь мажорируемыми по норме λ -интегрируемой фиксированной функ-

цией $\chi_K M \|\vec{g}\|$. Из теоремы 35 о сходимости Лебега следует, что тогда $\int f_n \vec{g} d\lambda$ сходится к $\int \vec{f} \vec{q} d\lambda$. Однако надо еще проверить условие непрерывности 3°) (см. стр. 579). Итак, пусть \mathcal{O} — открытое множество с компактным замыканием. Функция $\|\vec{q}\|$ интегрируема на множестве \mathcal{O} . Поэтому существует такое число $M \geq 0$, что $\int_{\mathcal{O}} (\|\vec{q}\| - \|\vec{q}\|_M) d\lambda \leq \varepsilon/2$, и существует такой компакт $K \subset \mathcal{O}$, что $\lambda(\mathcal{O} - K) \leq \varepsilon/2M$. Компакт K удовлетворяет относительно v условию 3°). В самом деле, если $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, равна 1 в некоторой окрестности компакта K и имеет носитель в \mathcal{O} , то

$$\begin{aligned} \|v(\varphi) - v(\mathcal{O})\| &\leq \int_{\mathcal{O}-K} \|\vec{q}\| d\lambda \leq \int_{\mathcal{O}-K} \|\vec{q}\|_M d\lambda + \\ &+ \int_{\mathcal{O}-K} (\|\vec{q}\| - \|\vec{q}\|_M) d\lambda \leq M \cdot \frac{\varepsilon}{2M} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Замечание. На стр. 489, 583 мы привели пример не абсолютно мажорируемой и не имеющей борелевского продолжения меры. Вышеизложенное показывает, что она не является мерой с базой ≥ 0 . Можно также привести пример абсолютно мажорируемой меры (а, следовательно, имеющей наименьшую абсолютную мажоранту), но не имеющей базы ≥ 0 . Оба понятия не эквивалентны, одно из них сильнее, чем другое.

С) Мера вида $\vec{p}\mu$, где μ -почти всюду $\|\vec{p}\| \leq 1$, очевидно, абсолютно мажорируема мерой μ . Поэтому нам следует лишь доказать обратное утверждение. При доказательстве мы ограничимся случаем сепарабельного пространства \vec{E} . Сначала мы докажем лишь частичное обратное утверждение: если v имеет вид $\vec{p}\mu$ и если эта мера абсолютно мажорируема мерой μ , то $\|\vec{p}\| \leq 1$ μ -почти всюду. Предположим, что множество $\{x \in X; \|\vec{p}(x)\| > 1\}$ имеет μ -меру > 0 . Так как оно является объединением множеств $A_\varepsilon = \{x \in X; \|\vec{p}(x)\| \geq 1 + 2\varepsilon\}$ для $\varepsilon = 1, 1/2, 1/3, \dots$, то существует такое $\varepsilon > 0$, при котором это множество A_ε имеет μ -меру > 0 . Мы будем выбирать $\varepsilon < 1/2$ по причинам, которые выясняются позже.

Пусть a — некоторая точка единичной сферы S пространства \vec{E} , и пусть $B(a, \varepsilon)$ — шар с центром в точке a радиуса ε . Этот шар — замкнутое выпуклое множество. Обозначим через $K(a, \varepsilon)$ объединение шаров, полученных из шара $B(a, \varepsilon)$ путем гомотетии с центром в нуле пространства и с вещественным

коэффициентом подобия $\geqslant 1$. Непосредственно видно, что $K(a, \varepsilon)$ также является выпуклым замкнутым множеством. В самом деле, пусть $\overset{\rightarrow}{\alpha x}$ и $\overset{\rightarrow}{\beta y}$ — две точки этого множества, где $\alpha \geqslant 1$, $\beta \geqslant 1$, x и y лежат в шаре $B(a, \varepsilon)$. Тогда точку $\overset{\rightarrow}{t\alpha x + (1-t)\beta y}$ можно записать в виде $\overset{\rightarrow}{\gamma(sx + (1-s)y)}$, где $\gamma = t\alpha + (1-t)\beta \geqslant 1$, $s = t\alpha/[t\alpha + (1-t)\beta]$, $1-s = [(1-t)\beta]/[t\alpha + (1-t)\beta]$. Так как $\overset{\rightarrow}{sx + (1-s)y} \in B(a, \varepsilon)$, то рассматриваемая точка лежит в $K(a, \varepsilon)$, а это означает, что множество $K(a, \varepsilon)$ выпукло.

Предположим теперь, что некоторая последовательность точек $\overset{\rightarrow}{k_n x_n}$ этого выпуклого множества сходится к пределу \vec{y} из \vec{E} . Поскольку нормы векторов $\overset{\rightarrow}{x_n}$ не меньше $1 - \varepsilon$, то числа k_n ограничены сверху. Однако они также ограничены снизу числом 1, поэтому из последовательности k_n можно извлечь подпоследовательность (для которой мы сохраним прежнее обозначение k_n), сходящуюся к некоторому пределу $k \geqslant 1$. Но тогда точки $\overset{\rightarrow}{x_n}$ будут сходиться к пределу $\vec{x} = \vec{y}/k$, принадлежащему замкнутому множеству $B(a, \varepsilon)$. Здесь $\vec{y} = k\vec{x}$, $k \geqslant 1$, $\vec{x} \in B(a, \varepsilon)$, а поэтому \vec{y} принадлежит множеству $K(a, \varepsilon)$, которое тем самым оказывается замкнутым. В силу сепарабельности \vec{E} , найдется счетное множество точек a (пусть это будут a_1, a_2, a_3, \dots), таких, что шары $B(a_n, \varepsilon)$ покрывают сферу S^1). Выпуклые замкнутые множества $K(a_n, \varepsilon)$ покрывают множество всех точек с нормой $\geqslant 1$ из \vec{E} , а множества $(1+2\varepsilon)K(a_n, \varepsilon)$ покрывают множество всех точек из \vec{E} с нормой $\geqslant 1+2\varepsilon$.

Множество A_ε является объединением μ -измеримых подмножеств $A_n = A_\varepsilon \cap \overset{\rightarrow}{p^{-1}((1+2\varepsilon)K(a_n, \varepsilon))}$. По крайней мере одно из них имеет μ -меру > 0 . Пусть это будет, например, A_1 . Пусть $K_1 \subset A_1$ есть некоторый компакт μ -меры > 0 . Поскольку $\vec{p}(K_1) \subset (1+2\varepsilon)K(a_1, \varepsilon)$, то из неравенства выпуклости (следствие 4 теоремы 39) следует, что $\int_{K_1} \vec{p} d\mu \leqslant \mu(K_1)(1+2\varepsilon)K(a_1, \varepsilon)$, а, значит, $\left\| \int_{K_1} \vec{p} d\mu \right\| \geqslant \mu(K_1)(1+2\varepsilon)(1-\varepsilon) > \mu(K_1)$, поскольку $\mu(K_1) \neq 0$ и $\varepsilon < 1/2$.

¹⁾ Всякое подмножество S метрического сепарабельного пространства E само сепарабельно. В самом деле, пусть последовательность точек b_n плотна в E . Выберем a_n в S так, чтобы $d(a_n, b_n) \leqslant d(b_n, s) + 1/n$. Тогда последовательность точек a_n будет плотной в S .

Однако это невозможно, так как мера $\vec{r}\mu$ предполагалась абсолютно мажорируемой мерой μ . Таким образом, для каждой функции $\varphi \geq 0$ из $\mathcal{C}(X)$ имеем $\left\| \int \varphi \vec{r}d\mu \right\| \leq \mu(\varphi)$. Условие 3° L -непрерывности (см. стр. 579) означает, что $\left\| \int \chi_{K_1} \vec{r}d\mu \right\| \leq \mu(K_1)$ для каждого компакта K_1 . Таким образом, если мера $\vec{r}\mu$ абсолютно мажорируема мерой μ , то $\|\vec{r}\| \leq 1$ почти всюду.

D) Из предыдущего, в частности, вытекает, что если мера $\vec{r}\mu$ равна нулю, то, каково бы ни было вещественное число $k > 0$, мера $\vec{k}\vec{r}\mu$ абсолютно мажорируется мерой μ , а, следовательно, величина $k\|\vec{r}\|$ μ -почти всюду мажорируема числом 1, т. е. норма $\|\vec{r}\|$ μ -почти всюду мажорируема числом $1/k$. Устремляя теперь k к бесконечности, можно убедиться в том, что норма $\|\vec{r}\|$ μ -почти всюду равна нулю. Следовательно, если $\vec{r}_1\mu$ и $\vec{r}_2\mu$ представляют одну и ту же меру, то функции \vec{r}_1 и \vec{r}_2 μ -почти всюду равны. Это доказывает единственность, фигурирующую в утверждениях 1°) и 2°) теоремы. Таким образом, п. 1°) теоремы полностью доказан.

E) Предположим теперь, что $\vec{v} = \vec{r}\mu$ имеет меру μ в качестве наименьшей абсолютной мажоранты. Тогда $\|\vec{r}\| \leq 1$ μ -почти всюду. Однако мера $\vec{r}\mu$ всегда абсолютно мажорируема величиной $\|\vec{r}\|\mu \leq \mu$. Поэтому, если мера μ является наименьшей абсолютной мажорантой, то $\|\vec{r}\|\mu = \mu$, а, значит, норма $\|\vec{r}\|$ μ -почти всюду равна 1. Обратно, предположим, что $\|\vec{r}\|$ μ -почти всюду равна 1. Тогда мера $\vec{r}\mu$ абсолютно мажорируема мерой μ . Покажем, что мера μ является наименьшей абсолютной мажорантой.

В самом деле, пусть наименьшей абсолютной мажорантой является мера λ . Тогда $\lambda \leq \mu$. Предположим, что $\lambda < \mu$. Тогда найдутся такое число ϵ , $0 < \epsilon \leq 1/2$, и такое борелевское множество A с компактным замыканием, что $\mu(A) > (1 + 2\epsilon)\lambda(A)$. (В самом деле, если для каждого $\epsilon > 0$ имеет место неравенство $\mu(A) \leq (1 + 2\epsilon)\lambda(A)$, то, согласно следствию 2 теоремы 50₂, $\mu \leq (1 + 2\epsilon)\lambda$. Из последнего неравенства следует, что $\mu \leq \lambda$, что противоречит предположению.) Однако $\|\vec{r}\| = 1$ и единичная сфера из \vec{E} , которая по предположению сепарабельна, покрыта счетным множеством (замкнутых и выпуклых) шаров

$B(a_n, \varepsilon)$. Поэтому хотя бы для одного из множеств $A_n = A \cap p^{-1}(B(a_n, \varepsilon))$, например для $n=1$, $\mu(A_n) > (1+2\varepsilon)\lambda(A_n)$. Теперь можно найти такой компакт $K_1 \subset A_1$, что $\mu(K_1) > (1+2\varepsilon)\lambda(K_1)$. Так как на K_1 функция p принимает значения в выпуклом замкнутом множестве $B(a_1, \varepsilon)$, то имеет место соотношение выпуклости $\int_{K_1}^{\rightarrow} p d\mu \leq \mu(K_1) B(a_1, \varepsilon)$, или

$$\left\| \int_{K_1}^{\rightarrow} p d\mu \right\| \geq (1-\varepsilon) \mu(K_1) > (1-\varepsilon)(1+2\varepsilon)\lambda(K_1) \geq \lambda(K_1),$$

т. е. $\left\| \int_{K_1}^{\rightarrow} p d\mu \right\| > \lambda(K_1)$, что противоречит, как мы уже видели,

предположению, согласно которому $p\mu$ абсолютно мажорируема мерой λ . Следовательно, $p\mu$ имеет меру μ в качестве наименьшей абсолютной мажоранты тогда и только тогда, когда $\|p\| \mu$ -почти всюду равна 1.

Пусть теперь $\vec{p}\mu$ (где \vec{p} локально μ -интегрируема) есть некоторая мера. Ее можно записать в виде

$$\frac{\vec{p}}{\|p\|} (\|p\|\mu), \text{ где } \left\| \frac{\vec{p}}{\|p\|} \right\| = 1^1).$$

Ее наименьшей абсолютной мажорантой является мера $\|p\|\mu$.

F) Пусть v — мера с базой ≥ 0 , абсолютно мажорируемая мерой μ . Если мы знаем, что $v = p\mu$, то из C) следует, что $\|p\| \leq 1$ μ -почти всюду. Покажем, что v является мерой с базой μ . Она имеет базу ≥ 0 , а, следовательно, $v = \vec{q}\lambda$. Мы только что видели, что ее наименьшей абсолютной мажорантой является мера $\|q\|\lambda$. Поскольку она абсолютно мажорируема мерой μ , то $\|q\|\lambda \leq \mu$. Согласно A), мы можем теперь утверждать, что $\|q\|\lambda = g\mu$, $0 \leq g \leq 1$. Отсюда

$$v = \vec{q}\lambda = \frac{\vec{q}}{\|q\|} (\|q\|\lambda) = \frac{\vec{q}}{\|q\|} g\mu = \vec{p}\mu,$$

¹⁾ $\vec{p}/\|p\|$ не определена в точках, где $\vec{p} = \vec{0}$. Однако такие точки образуют множество точек, $\|p\|\mu$ -мера которых равна нулю. Мы можем значения $\vec{p}/\|p\|$ на этом множестве выбирать произвольно, но норма при этом должна быть равной 1, для того, чтобы эта норма была всюду равной 1.

что полностью доказывает наше утверждение. Точно так же, если мера ν имеет базу $\geqslant 0$ и имеет меру μ в качестве наименьшей абсолютной мажоранты, то для нее мера μ , в частности, является абсолютной мажорантой, а, значит, она имеет вид $p\mu$ и тогда, согласно Е), получаем, что $\|p\| = 1$ μ -почти всюду. Мы нашли следующее необходимое и достаточное условие того, чтобы мера ν с базой $\geqslant 0$ была абсолютно мажорируема мерой μ (соответственно того, чтобы мера ν имела μ в качестве наименьшей абсолютной мажоранты): $\nu = p\mu$, где $\|p\| \leqslant 1$ (соответственно $= 1$) μ -почти всюду.

Неравенство (IV, 5; 9₂), справедливое по предположению, когда функция f взята из $\mathcal{C}(X)$ (и уже использованное в два приема при доказательстве в случае, когда f являлась характеристической функцией некоторого компакта), теперь очевидным образом верно для любой функции f из $\Gamma(X)$, поскольку $\nu = p\mu$, где $\|p\| \leqslant 1$ μ -почти всюду, и имеет место равенство (IV, 5; 9₃).

G) Если мера ν имеет базой меру μ , то каждое борелевское множество с компактным замыканием нулевой μ -меры также имеет нулевую ν -меру. Докажем обратное, предполагая, что ν имеет базу $\geqslant 0$. Положим $\nu = q\lambda$, где $\|q\| = 1$, что возможно, если в качестве λ взята наименьшая абсолютная мажоранта.

Пусть A — борелевское множество с компактным замыканием и нулевой μ -мерой. Тогда это множество, как и все его борелевские подмножества, имеет нулевую $\nu = q\lambda$ -меру, а, следовательно, мера $\chi_A q\lambda$, равная нулю на всех борелевских множествах с компактным замыканием, является нулевой мерой. Отсюда вытекает, что функция $\chi_A q$, а, значит, и ее норма $\chi_A \lambda$ -почти всюду равны нулю. Иначе говоря, множество A имеет λ -меру, равную нулю. Согласно А), мера λ имеет базу μ , а, значит, такую же базу имеет $\nu = q\lambda$.

Нами полностью теперь доказаны п. 1°) и 2°) теоремы.

H) Множество мер на X со значениями в \vec{E} и базой $\geqslant 0$ является некоторым векторным пространством. В самом деле, если меры ν_1 и ν_2 имеют базы λ_1 и λ_2 , $\nu_1 = q_1\lambda_1$, $\nu_2 = q_2\lambda_2$, то меры λ_1 и λ_2 обе мажорируются мерой $\lambda = \lambda_1 + \lambda_2$, а, следовательно, $\lambda_1 = g_1\lambda$, $\lambda_2 = g_2\lambda$, где функции g_1 и g_2 заключены между 0 и 1. Отсюда следует, что мера $\nu_1 + \nu_2 = (q_1g_1 + q_2g_2)\lambda$ имеет базу

≥ 0 . Точно так же, если мера v имеет базу ≥ 0 и k — скаляр, то мера \vec{kv} имеет базу ≥ 0 . Всякая вещественная мера, являясь разностью двух мер ≥ 0 , будет мерой с базой ≥ 0 (точнее, мера v имеет базой меру $|v|$, являющуюся ее наименьшей абсолютной мажорантой). Каждая комплексная мера $v = v_1 + iv_2$ имеет базу ≥ 0 . Если теперь пространство \vec{E} конечномерно и если $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ — базис в \vec{E} , то $\vec{v} = \sum_{i=1}^n \vec{e}_i v_i$, где v_i — скалярные меры с базой ≥ 0 . Но тогда меры $\vec{e}_i v_i$, а вместе с ними и мера v будут иметь базы ≥ 0 . Итак, доказан п. 3°) теоремы.

I) Утверждение 4°) теоремы теперь тривиально. Пусть p — вещественная локально μ -интегрируемая функция. Известно, что $|p\mu|$ является наименьшей абсолютной мажорантой функции $p\mu$ (см. сказанное после теоремы 18). Кроме того, из окончания п. Е) следует, что $|p|\mu$ является этой наименьшей абсолютной мажорантой. Поэтому $|p\mu| = |p|\mu$. Но тогда

$$(p\mu)^+ = \frac{|p\mu| + p\mu}{2} = \frac{|p| + p}{2}\mu = p^+\mu$$

и аналогично $(p\mu)^- = p^-\mu$.

Если теперь функции p и q вещественны и локально μ -интегрируемы, то

$$\sup(p\mu, q\mu) = p\mu + ((q - p)\mu)^+ = (p + (q - p)^+)\mu = (\sup(p, q))\mu$$

и аналогично $\inf(p\mu, q\mu) = (\inf(p, q))\mu$. В частности, неравенство $p\mu \leq q\mu$ означает, что точной верхней гранью мер $p\mu$ и $q\mu$ является мера $q\mu$, а, следовательно, $\sup(p, q) = q$ μ -почти всюду или $p \leq q$ μ -почти всюду ($p\mu \geq 0$ эквивалентно: $p \geq 0$ μ -почти всюду).

Наконец, пусть значения \vec{p} лежат в \vec{E} , а значения q вещественны и ≥ 0 , и пусть мера $p\mu$ абсолютно мажорируема мерой $q\mu$. Тогда, так как, согласно Е), ее наименьшей абсолютной мажорантой является $\|\vec{p}\| \mu$, то $\|\vec{p}\| \mu \leq q\mu$, а, следовательно, $\|\vec{p}\| \leq q$ μ -почти всюду. Однако если мера v со значениями в \vec{E} и базой ≥ 0 абсолютно мажорируема мерой $q\mu$, то, согласно F), мера v имеет базой меру $q\mu$, а, следовательно, и меру μ . Это означает, что $v = \vec{p}\mu$ и тогда $\|\vec{p}\| \leq q$ μ -почти всюду. Этим заканчивается доказательство утверждения 4°) и, значит, всей теоремы.

Замечание 1. Во всех условиях, относящихся к мере v со значениями в \vec{E} , предполагалось, что v имеет базу ≥ 0 (напри-

мер, если мера v с базой \rightarrow ≥ 0 абсолютно мажорируема мерой μ , то $v = p\mu$, $\|p\| \leq 1$ μ -почти всюду). Контрпримеры показывают, что это условие неизбежно, за исключением, конечно, случая, когда пространство \vec{E} конечномерно, так как тогда любая мера со значениями в \vec{E} имеет базу ≥ 0 . Общая теорема Данфорда — Петтиса дает широкий класс пространств \vec{E} , для которых нет необходимости предполагать v имеющей базу ≥ 0 .

З а м е ч а н и е 2. Физики часто интерпретируют меру Дирака δ на прямой так, чтобы ее можно было записать в виде $\delta(x) dx$, т. е. в виде произведения dx на некоторую функцию. Мы теперь видим, что такая запись невозможна. В самом деле, начало координат является множеством меры нуль для dx , а не для δ . Поэтому δ не будет мерой с базой dx .

Следствие 1. Пусть μ и v — две меры ≥ 0 на X . Тогда существуют наибольшая мера v_1 с базой μ , мажорируемая мерой v и определяемая по формуле

$$v_1 = \sup_{n \in \mathbb{N}} (\inf(v, n\mu)), \quad (\text{IV}, 5; 9_4)$$

и μ -сингулярная¹⁾ наибольшая мера v_2 , мажорируемая мерой v . Мера v может быть записана в виде $v = v_1 + v_2$, и это единственное представление v в виде суммы двух мер ≥ 0 , таких, что первая мера имеет базой меру μ , а вторая μ -сингулярна. Кроме того, v_1 и v_2 — дизъюнктные меры.

Доказательство. Прежде всего, $\inf(v, n\mu)$ мажорируем мерой $n\mu$ и поэтому, согласно п. 1°) теоремы, имеет вид $p_n\mu$, $p_n \leq n$ μ -почти всюду. При возрастании n так определенные меры возрастают; поэтому $p_n\mu \leq p_{n+1}\mu$ или $p_n \leq p_{n+1}$ μ -почти всюду. Последовательность функций p_n с точностью до множеств μ -меры нуль оказывается возрастающей. Пусть p — ее предел, определенный μ -почти всюду. Так как p_n μ -измеримы, то функция p также μ -измерима. С другой стороны, для каждого boreлевского множества A с компактным замыканием $\int_A p_n d\mu \leq \infty$

$\leq v(A)$ и, следовательно, по теореме 36 Фату, $\int_A^* p d\mu \leq v(A)$,

а, значит, функция p локально μ -интегрируема и определяет некоторую меру $p\mu = v_1 \geq 0$ с базой μ , мажорируемую мерой v .

¹⁾ См. стр. 589. Это означает, что v_2 сосредоточена на некотором множестве μ -меры нуль или что v_2 и μ дизъюнктны.

Для каждой функции $\varphi \geq 0$ из $\mathcal{C}(X)$ имеет место соотношение (Фату или Лебега) $\int \varphi p d\mu = \lim_{n \rightarrow \infty} \int \varphi p_n d\mu$, а, значит, $= \sup_{n \in \mathbb{N}} \int \varphi p_n d\mu$. В силу формулы (IV, 2; 41₆), отсюда вытекает, что мера v_1 является точной верхней гранью для мер $p_n \mu$, а, значит, удовлетворяет формуле (IV, 5; 9₄). Если произвольная мера $\lambda \geq 0$ мажорируется мерой v и имеет базой меру μ , то она имеет вид $q\mu$. Если $q_n = \inf(q, n)$, то из п. 4°) теоремы следует, что $\inf(\lambda, n\mu) = \inf(q\mu, n\mu) = q_n\mu$. Однако из неравенства $\lambda \leq \leq v$ следует, что $q_n\mu = \inf(\lambda, n\mu) \leq \inf(v, n\mu) \leq v_1$. Кроме того, из неравенства $q_n\mu \leq v_1$, или $\int \varphi q_n d\mu \leq v_1(\varphi)$, имеющего место при любой функции $\varphi \geq 0$, $\varphi \in \mathcal{C}(X)$, согласно теоремам Фату и Лебега, получаем неравенство $\int \varphi q d\mu \leq v_1(\varphi)$, или $\lambda = q\mu \leq \leq v_1$. Значит, v_1 является наибольшей мерой ≥ 0 , мажорирующей мерой v и имеющей базой меру μ .

Мера $v_2 = v - v_1$ теперь ≥ 0 и мажорируется мерой v . Эта мера μ -сингулярная. В самом деле, если положить $\theta = = \inf(v_2, \mu) \geq 0$, то мы увидим, что $v_1 + \theta$ мажорируется мерой $v_1 + v_2 = v$ и имеет базой меру μ , поскольку $\theta \leq \mu$. Таким образом, $v_1 + \theta = v_1$ и $\theta = 0$, а, значит, v_2 и μ дизъюнктны. Согласно следствию теоремы 50₅, существует некоторое множество нулевой v_2 -меры, на котором сосредоточена мера μ . Однако на этом же множестве также сосредоточена мера $p\mu$, а, значит, меры v_2 и $p\mu = v_1$ также дизъюнктны (более общо, это рассуждение показывает, что если две неотрицательные меры λ_1 и λ_2 дизъюнктны, то две меры с базами соответственно λ_1 и λ_2 также дизъюнктны). Пусть теперь λ — некоторая мера ≥ 0 , мажорируемая мерой v и дизъюнктная с мерой μ . Как мы только что видели, она является дизъюнктной с мерой v_1 . Из $\inf(\lambda, v_1) = 0$ вытекает, что $\sup(\lambda, v_1) = \lambda + v_1$. Однако меры λ и v_1 мажорируются мерой v , а, значит, это также можно сказать о мере $\sup(\lambda, v_1)$, или, что равносильно, мере $\lambda + v_1$. Поэтому $\lambda \leq v - - v_1 = v_2$ и, значит, v_2 является наибольшей мерой ≥ 0 , мажорируемой мерой v и дизъюнктной с мерой μ .

Теперь легко видеть, что $v = v_1 + v_2$ является единственным представлением меры v в виде суммы двух мер ≥ 0 соответственно с базой μ и μ -сингулярной. В самом деле, если $v = \theta_1 + + \theta_2$ является таким разложением, то, согласно максимальному свойству мер v_1 и v_2 , имеем $\theta_2 \leq v_2$, а, следовательно, $\theta_1 \geq v_1$. Так как одновременно $\theta_1 \leq v_1$, то $\theta_1 = v_1$ и $\theta_2 = v_2$.

Пусть теперь \vec{v} — векторная мера с базой ≥ 0 . Мы будем говорить, что v сосредоточена на части A множества X , если $\vec{C}A$

имеет нулевую меру для наименьшей абсолютной мажоранты $\overrightarrow{|v|}$. Тогда \overrightarrow{v} будет называться μ -сингулярной или дизъюнктной с мерой μ , если она сосредоточена на некотором множестве нулевой μ -меры.

Следствие 2. Пусть μ — некоторая мера ≥ 0 на X , а \overrightarrow{v} — векторная мера с базой ≥ 0 . Эта мера допускает единственное разложение $\overrightarrow{v} = \overrightarrow{v}_1 + \overrightarrow{v}_2$, где \overrightarrow{v}_1 имеет базой меру μ , а \overrightarrow{v}_2 μ -сингулярна. Кроме того, разложение, соответствующее наименьшей абсолютной мажоранте $\overrightarrow{|v|}$, имеет вид $\overrightarrow{|v|} = \overrightarrow{|v_1|} + \overrightarrow{|v_2|}$.

Доказательство. Так как мера \overrightarrow{v} имеет базу ≥ 0 , то можно записать $\overrightarrow{v} = \overrightarrow{q}\lambda$. Пусть $\lambda = \lambda_1 + \lambda_2$ — такое разложение λ относительно μ , что λ_1 имеет базой меру μ , а λ_2 μ -сингулярна (следствие 1). Поскольку λ_1 и λ_2 не превосходят λ и функция \overrightarrow{q} локально λ -интегрируема, то она заведомо локально λ_1 -интегрируема и локально λ_2 -интегрируема, при этом $\overrightarrow{q}\lambda = \overrightarrow{q}\lambda_1 + \overrightarrow{q}\lambda_2$. Однако мера λ_1 , а, значит, и мера $\overrightarrow{q}\lambda_1$ имеют базой меру μ . Мера λ_2 дизъюнктна с мерой μ , а, значит, такой же будет и мера $\overrightarrow{q}\lambda_2$, как это мы видели при доказательстве предыдущего следствия. Тем самым мы получили разложение требуемого вида. Поскольку $\overrightarrow{|v|} = \overrightarrow{q}\|\lambda\|$, $\overrightarrow{|v_1|} = \overrightarrow{q}\|\lambda_1\|$ и $\overrightarrow{|v_2|} = \overrightarrow{q}\|\lambda_2\|$, то очевидным образом получаем, что $\overrightarrow{|v|} = \overrightarrow{|v_1|} + \overrightarrow{|v_2|}$. Что же касается единственности, то она сводится к следующему: если $\overrightarrow{v}_1 + \overrightarrow{v}_2 = \overrightarrow{0}$, где \overrightarrow{v}_1 имеет базой меру μ , а \overrightarrow{v}_2 μ -сингулярна, то \overrightarrow{v}_1 и \overrightarrow{v}_2 суть нули. Но это действительно так. Если A является множеством нулевой μ -меры, на котором сосредоточена мера v_2 , то, поскольку $\overrightarrow{|v_1|}$ имеет базой меру μ , множество A имеет нулевую $\overrightarrow{|v_1|}$ -меру. Так как мера v_2 сосредоточена на множестве A , то множество \overrightarrow{CA} также имеет нулевую $\overrightarrow{|v_1|} = \overrightarrow{|v_2|}$ -меру. Значит, X является множеством нулевой $\overrightarrow{|v_1|}$ -меры, а, следовательно, $\overrightarrow{v}_1 = \overrightarrow{0}$ и $\overrightarrow{v}_2 = \overrightarrow{0}$.

Следствие 3. Пусть \overrightarrow{v} — мера с базой ≥ 0 и μ — мера ≥ 0 . Для того чтобы мера v имела базой меру μ , необходимо и достаточно, чтобы для каждого компакта K и любого числа $\varepsilon > 0$ существовало такое число $\delta(K, \varepsilon) = \delta > 0$, что для любого борелевского множества $A \subset K$, такого, что $\mu(A) \leq \delta$, имело место неравенство $\|\overrightarrow{v}(A)\| \leq \varepsilon$.

Доказательство. Пусть сначала мера $\vec{v} = \vec{p}\mu$. Поскольку функция \vec{p} интегрируема на K , то существует такое число $M \geq 0$, что $\int_K (\|\vec{p}\| - \|\vec{p}\|_M) d\mu \leq \frac{\epsilon}{2}$, где $\|\vec{p}\|_M = \inf(\|\vec{p}\|, M)$ (теорема 35 Лебега, примененная к разности $\|\vec{p}\| - \|\vec{p}\|_M$ при M , стремящемся к $+\infty$). Теперь достаточно положить $\delta = \epsilon/2M$. В самом деле, если $\mu(A) \leq \delta$, то

$$\begin{aligned} \|\vec{v}(A)\| &= \left\| \int_A \vec{p} d\mu \right\| \leq \int_A \|\vec{p}\| d\mu = \\ &= \int_A \|\vec{p}\|_M d\mu + \int_A (\|\vec{p}\| - \|\vec{p}\|_M) d\mu \leq M\delta + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Обратно, предположим, что имеет место указанное свойство. Тогда каждое борелевское множество с компактным замыканием нулевой μ -меры имеет также нулевую \vec{v} -меру и, согласно теореме Лебега — Никодима, \vec{v} имеет базой меру μ .

Следствие 4. Пусть заданы мера \vec{v} с базой ≥ 0 и мера $\mu \geq 0$. Для того чтобы мера \vec{v} имела базой меру μ , необходимо и достаточно, чтобы для каждого открытого множества \mathcal{O} с компактным замыканием и любого $\epsilon > 0$ нашлось такое число $\eta(\mathcal{O}, \epsilon) = \eta > 0$, что для функции $\varphi \in \mathcal{C}(X)$ с носителем в \mathcal{O} , $0 \leq \varphi \leq 1$, из $\mu(\varphi) \leq \eta$ вытекает неравенство $\|\vec{v}(\varphi)\| \leq \epsilon$ ¹.

Смысл следствия заключается в том, что оно дает возможность выяснить, имеет ли мера \vec{v} базой меру μ , не используя продолжение Лебега. Обе меры рассматриваются только как функции, определенные на $\mathcal{C}_+(X)$.

Доказательство. Предположим, что $\vec{v} = \vec{p}\mu$. Пусть δ и M — числа, определенные так же, как и в предыдущем следствии, относительно множества $\bar{\mathcal{O}}$ и числа ϵ : $\int_{\bar{\mathcal{O}}} (\|\vec{p}\| - \|\vec{p}\|_M) d\mu \leq \epsilon/2$ и $\delta = \epsilon/2M$. Если теперь функция $\varphi \in \mathcal{C}(X)$ имеет носитель в \mathcal{O} , $0 \leq \varphi \leq 1$, и если $\mu(\varphi) \leq \delta$, то

$$\begin{aligned} \|\vec{v}(\varphi)\| &\leq \int \varphi \|\vec{p}\| d\mu = \\ &= \int \|\vec{p}\|_M \varphi d\mu + \int (\|\vec{p}\| - \|\vec{p}\|_M) \varphi d\mu \leq M\delta + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

¹) Внимание! Это неравенство не имеет места, если не предполагать, что $\varphi \leq 1$.

Число $\eta = \delta$ обладает требуемыми свойствами относительно \mathcal{O} и ε .

Обратно, предположим, что мера \vec{v} с базой ≥ 0 обладает свойством, сформулированным в условии следствия. Пусть сначала $v \geq 0$. Выберем такое множество A с компактным замыканием, что $\mu(A) = 0$, и пусть \mathcal{O} — некоторое открытое множество с компактным замыканием, содержащее A . Каждому $\varepsilon > 0$, согласно предположению, соответствует некоторое число $\eta = \eta(\mathcal{O}, \varepsilon)$, обладающее указанными свойствами. Выберем такое открытое множество \mathcal{A} , что $A \subset \mathcal{A} \subset \mathcal{O}$ и $\mu(\mathcal{A}) \leq \eta$. Для каждой функции $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, с носителем в \mathcal{A} имеют место неравенства $\mu(\varphi) \leq \mu(\mathcal{A}) \leq \eta$, а, значит, $v(\varphi) \leq \varepsilon$. Переходя к точной верхней грани по всем φ , получаем $v(\mathcal{A}) \leq \varepsilon$ и, следовательно, $v^*(A) \leq \varepsilon$. Так как ε произвольно, отсюда следует, что $v^*(A) = 0$. Значит, каждое борелевское множество с компактным замыканием нулевой μ -меры имеет также нулевую v -меру. Теорема Лебега — Никодима говорит, что в этом случае мера v имеет базой меру μ .

Пусть теперь мера v вещественна. Пусть $\eta = \eta(\mathcal{O}, \varepsilon)$ — число, соответствующее, согласно предположению, множеству \mathcal{O} и числу ε . Если функция $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, имеет носитель в \mathcal{O} и $\mu(\varphi) \leq \eta$, то для каждой функции $\psi \in \mathcal{C}(X)$, $0 \leq \psi \leq \varphi$, имеет место неравенство $\mu(\psi) \leq \eta$, а, значит, $|v(\psi)| \leq \varepsilon$. Переходя к точной верхней грани по всем ψ (формула (IV, 2; 41)), получим $v^+(\varphi) \leq \varepsilon$. Точно так же проверяется справедливость неравенства $v^-(\varphi) \leq \varepsilon$. Таким образом v^+ и v^- обладают свойствами, указанными в условии следствия, и ≥ 0 . Значит, они, а вместе с ними и мера v , имеют базой меру μ .

Пусть теперь v — мера со значениями в конечномерном пространстве \vec{E} . Пусть (\vec{e}_i) , $i = 1, 2, \dots, N$, — некоторый базис \vec{E} над полем \mathbb{R} . Тогда $\vec{v} = \sum_{i=1}^N v_i \vec{e}_i$.

Отображение « i -я координата» является \mathbb{R} -линейным, непрерывным отображением \vec{E} в \mathbb{R} . Значит, существует такое ε' , что из $\|v(\varphi)\| \leq \varepsilon'$ следует $|v_i(\varphi)| \leq \varepsilon$. Если теперь $\eta = \eta(\mathcal{O}, \varepsilon')$ — число, соответствующее по лемме v , \mathcal{O} и ε' , то из неравенства $\mu(\varphi) \leq \eta$ будет следовать неравенство $\|\vec{v}(\varphi)\| \leq \varepsilon'$, а, значит, $|v_i(\varphi)| \leq \varepsilon$. Поэтому каждая мера v_i имеет базой меру μ , и, следовательно, такую же базу имеет мера \vec{v} .

Мы здесь не приводим доказательства для случая произвольного банахового пространства \vec{E} .

Применение к продолжению меры с векторными значениями

До настоящего времени мы определяли интегралы $\int f d\mu$, где f не является функцией из $\mathcal{C}(X)$ только тогда, когда μ —вещественная мера $\geqslant 0$ или когда f —борелевская ограниченная функция с компактным носителем (стр. 578). Теперь можно дать определение для более общих случаев. Пусть $\vec{\mu}$ —некоторая мера на X со значениями в банаховом пространстве \vec{E} . Предположим, что она имеет базу $\geqslant 0$, т. е. что ее можно представить в виде $r\mu_0$, где μ_0 —вещественная мера $\geqslant 0$ и r —локально μ_0 -интегрируемая функция со значениями в \vec{E} . Попытаемся дать следующее определение.

Скалярная функция f , определенная на X , $\vec{\mu}$ -интегрируема тогда и только тогда, когда μ_0 -интегрируема функция rf , и тогда, по определению,

$$\int f d\vec{\mu} = \int (\vec{r}\vec{f}) d\mu_0. \quad (\text{IV}, 5; 10)$$

Точно так же, если $\vec{\mu}$ является скалярной мерой, то она может быть записана хотя бы одним способом в виде $r\mu_0$, где μ_0 —мера $\geqslant 0$, r —определенная на X вещественнонозначная локально μ_0 -интегрируемая функция. Теперь можно сказать, что *функция \vec{f} , определенная на X со значениями в банаховом пространстве \vec{F} , $\vec{\mu}$ -интегрируема тогда и только тогда, когда μ_0 -интегрируема функция rf , и положить*

$$\int \vec{f} d\vec{\mu} = \int (r\vec{f}) d\mu_0. \quad (\text{IV}, 5; 11)$$

Точно так же можно определить операцию, обобщающую две предыдущие. Пусть $\vec{\mu}$ —некоторая мера на X со значениями в \vec{E} , выражаемая в виде $r\mu_0$, где μ_0 —мера $\geqslant 0$ и r —локально μ_0 -интегрируемая функция, определенная на X , со значениями в \vec{F} . Пусть \vec{f} —некоторая функция, определенная на X , со значениями в \vec{F} , и пусть B —билинейное непрерывное отображение $\vec{E} \times F$ в банахово пространство \vec{G} . Посмотрим, какой смысл следует придать выражению $\int B(d\vec{\mu}, \vec{f})$.

Если \vec{f} есть $\vec{\mu}$ -этажная функция, т. е. функция, представимая в виде $\vec{f} = \sum_i \vec{f}_i \Phi_{X_i}$, где \vec{f}_i —некоторые постоянные, а Φ_{X_i} —ха-

рактеристические функции μ -измеримых множеств X_i конечной μ -меры (т. е. таких, что $\mu(X_i) = \int \Phi_{X_i} d\mu$ имеет смысл, согласно (IV, 5; 10)), то можно будет положить

$$\int B(d\mu, \vec{f}) = \sum_i B(\vec{\mu}(X_i), \vec{f}_i). \quad (\text{IV}, 5; 12)$$

Теперь для произвольной функции \vec{f} можно утверждать, что $\int B(d\mu, \vec{f})$ имеет смысл тогда и только тогда, когда функция $B(\vec{p}, \vec{f}) : x \rightarrow B(p(x), \vec{f}(x))$ μ_0 -интегрируема, и положить

$$\int B(d\mu, \vec{f}) = \int B(\vec{p}, \vec{f}) d\mu_0. \quad (\text{IV}, 5; 13)$$

Эти определения оправданы в силу следующих двух соображений. Во-первых, если мера μ сама ≥ 0 , то из теоремы 51 следует, что новое определение (IV, 5; 11) совпадает со старым; во-вторых, имеет место

Теорема 54₂. Если меру μ можно представить в виде $\vec{p}_1 \mu_1$ и в виде $\vec{p}_2 \mu_2$, где обе меры $\mu_1, \mu_2 \geq 0$ и где \vec{p}_1 (соответственно \vec{p}_2) локально μ_1 -интегрируема (соответственно μ_2 -интегрируема); то оба интеграла вида $\int B(d\mu, \vec{f})$, определенные по формуле (IV, 5; 13), совпадают. Другими словами, $\int B(\vec{p}_2, \vec{f}) d\mu_1$ имеет смысл тогда и только тогда, когда имеет смысл $\int B(\vec{p}_2, \vec{f}) d\mu_2$, и значения этих интегралов совпадают.

В самом деле, обозначим через θ меру-сумму $\mu_1 + \mu_2$. Тогда $\mu_1 \leq \theta$ и, согласно теореме 53, найдется такая вещественная функция $g_1 \geq 0$, что мера μ_1 будет равна $g_1 \theta$. Теперь, согласно следствию теоремы 51, мера $\vec{p}_1 \mu_1 = \vec{p}_1(g_1 \theta)$ равна мере $(\vec{p}_1 g_1) \theta$.

Затем, согласно (IV, 5; 13), интеграл $\int B(d\mu, \vec{f})$ имеет смысл относительно представления μ в виде $\vec{p}_1 \mu_1$ тогда и только тогда, когда имеет смысл $\int B(\vec{p}_1, \vec{f}) d\mu_1$, а, согласно теореме 51, последнее выражение имеет смысл тогда и только тогда, когда имеет смысл $\int (B(\vec{p}_1, \vec{f}) g_1) d\theta = \int B(\vec{p}_1 g_1, \vec{f}) d\theta$ и выполняется равенство

$$\int B(d\mu, \vec{f}) = \int B(\vec{p}_1 g_1, \vec{f}) d\theta. \quad (\text{IV}, 5; 14)$$

Если теперь левую часть определить исходя из представления μ в виде $\vec{p}_2\vec{\mu}_2$, $\vec{\mu}_2 = \vec{g}_2\theta$, то мы получим:

$$\int B(d\vec{\mu}, \vec{f}) = \int B(\vec{p}_2\vec{g}_2, \vec{f}) d\theta. \quad (\text{IV}, 5; 15)$$

Однако, поскольку $(\vec{p}_1\vec{g}_1)\theta$ и $(\vec{p}_2\vec{g}_2)\theta$ представляют одну и ту же меру μ , то из теоремы 52 вытекает, что $\vec{p}_1\vec{g}_1$ и $\vec{p}_2\vec{g}_2$ равны θ -почти всюду, так что два значения, полученные для $\int B(d\vec{\mu}, \vec{f})$, существуют одновременно и одинаковы.

Таким образом мы можем продолжать векторную меру μ каждый раз, когда она имеет базу ≥ 0 .

Следствие. Пусть \vec{E} — векторное конечномерное пространство, μ — некоторая мера на X со значениями в \vec{E} конечной нормы и \vec{f} — непрерывная и ограниченная функция, определенная на X , со значениями в банаховом пространстве \vec{F} . Тогда для билинейного непрерывного отображения B из $\vec{E} \times \vec{F}$ в банахово пространство \vec{G} интеграл $\int B(d\vec{\mu}, \vec{f})$ имеет смысл.

Доказательство. Согласно теореме 54, $\vec{\mu} = \vec{q}\mu_0$, где $\mu_0 = |\vec{\mu}|$ имеет конечную норму и $\|\vec{q}\| = 1$. Тогда отображение $B(\vec{q}, \vec{f})$ μ_0 -измеримо (следствие 3 теоремы 23₂) и ограничено, а, значит, μ_0 -интегрируемо, ибо μ_0 имеет конечную норму (следствие 2 теоремы 39).

Применение к интегрируемости функции по нескольким мерам

Теорема 54₃. Пусть μ и ν — две меры ≥ 0 на X . Если функция \vec{f} , определенная на X , со значениями в банаховом пространстве \vec{F} одновременно интегрируема по μ и ν , то она интегрируема по $\mu + \nu$, и при этом

$$\int \vec{f} d(\mu + \nu) = \int \vec{f} d\mu + \int \vec{f} d\nu. \quad (\text{IV}, 5; 16)$$

В самом деле, по теореме 53, $\mu = p(\mu + \nu)$ и $\nu = q(\mu + \nu)$, где p и q $(\mu + \nu)$ -измеримы и мажорируются числом 1. Кроме того, $\mu + \nu = (p + q)(\mu + \nu)$ и поэтому, согласно теореме 52, $p + q$ $(\mu + \nu)$ -почти всюду равна 1. Поскольку функция \vec{f} μ -интегрируема, то, в силу теоремы 51, функция $\vec{f}p$ $(\mu + \nu)$ -интегри-

руема и, кроме того,

$$\int \vec{f} d\mu = \int \vec{f} p d(\mu + v). \quad (\text{IV}, 5; 17)$$

Точно так же проверяется, что функция $\vec{f}q$ $(\mu + v)$ -интегрируема и

$$\int \vec{f} d\nu = \int \vec{f} q d(\mu + v). \quad (\text{IV}, 5; 18)$$

Следовательно, $(\mu + v)$ -интегрируема функция $\vec{f}(p + q)$, а поскольку $p + q = 1$ $(\mu + v)$ -почти всюду, то $(\mu + v)$ -интегрируемой будет сама функция \vec{f} ; складывая равенства (IV, 5; 17) и (IV, 5; 18), мы получим (IV, 5; 16).

Пример. Пусть \vec{E} — евклидово конечномерное пространство над \mathbb{R} и $(\vec{X} | \vec{Y})$ — скалярное произведение в \vec{E} , являющееся билинейной непрерывной формой. Если $\vec{\mu}$ — некоторая мера на X со значениями в \vec{E} и \vec{f} — функция на X со значениями в \vec{E} , то можно придать определенный смысл выражению $\int (d\vec{\mu} | \vec{f})$ и считать его равным вещественному числу. В самом деле, по отношению к ортонормированному базису e_1, e_2, \dots, e_n

$$\vec{\mu} = \sum_{i=1}^n \mu_i \vec{e}_i \quad \text{и} \quad \vec{f} = \sum_{i=1}^n f_i \vec{e}_i.$$

Если при этом

$$\sum_{i=1}^n \int f_i d\mu_i$$

имеет смысл, то можно говорить, что интеграл $\int (d\vec{\mu} | \vec{f})$ также имеет смысл и равен той же величине. Можно также утверждать, что $\int (d\vec{\mu} | \vec{f})$ имеет смысл, если $\vec{\mu}$ имеет конечную норму, а борелевская функция \vec{f} ограничена.

Замечание. Мы предполагали, что B является билинейным отображением, определенным на $\vec{E} \times \vec{F}$, и обозначали его через $B(\vec{X}, \vec{Y})$, $\vec{X} \in \vec{E}$, $\vec{Y} \in \vec{F}$. Соответствующим образом обозначается интеграл $\int B(d\vec{\mu}, \vec{f})$ для $\vec{\mu}$ со значениями в \vec{E} и \vec{f} со значениями в \vec{F} . Если B обозначается другими символами, то аналогично надо обозначать и интеграл. Мы учли это в случае скалярного произведения, которое мы обозначили через $(\vec{X} | \vec{Y})$, и

соответствующим образом обозначили интеграл $\int (\vec{d}\mu | \vec{f})$. Можно было бы то же самое скалярное произведение обозначить через $(\vec{Y} | \vec{X})$, и тогда тот же интеграл записать в виде $\int (\vec{f} | d\mu)$. Предположим теперь, что $\vec{E} = \vec{F} = \mathbb{R}^3$, и пусть B является векторным произведением, т. е. билинейным непрерывным отображением $(X, Y) \rightarrow X \wedge Y$ пространства $\mathbb{R}^3 \times \mathbb{R}^3$ в \mathbb{R}^3 . Тогда $(\vec{X}, \vec{Y}) \rightarrow \vec{Y} \wedge \vec{X}$ является противоположным билинейным отображением. В этом случае мы получим два противоположных вектора: $\int d\mu \wedge \vec{f}$ и $\int \vec{f} \wedge d\mu$. В физике постоянно приходится встречаться с интегралами такого типа.

Сопряженность пространств L^p и $L^{p'}$

Пусть X — локально компактное пространство, снабженное мерой Радона $\mu \geq 0$. В дальнейшем пространство $L^p(X, \mu)$ мы будем кратко обозначать через L^p . Точно так же, как в следствии 1 теоремы 46, для $1 \leq p \leq +\infty$ положим $p' = p/(p-1)$, так что $1/p + 1/p' = 1$. Пусть $\dot{h} \in L^{p'}$ (через \dot{h} мы обозначаем класс эквивалентности, т. е. множество функций $h \in L^{p'}$, попарно почти всюду равных между собой). Этот элемент определяет по формуле

$$\langle u_{\dot{h}}, \dot{f} \rangle = \int_X h(x) f(x) d\mu(x), \quad f \in \dot{f}, \quad (\text{IV, 5; 19})$$

некоторую линейную непрерывную форму на L^p . В самом деле, согласно следствию 1 теоремы 46, правая часть имеет смысл, так что $u_{\dot{h}}$ является линейной формой на L^p . Непрерывность этой формы вытекает из неравенства Гёльдера, причем

$$\|u_{\dot{h}}\| = \frac{\sup |\langle u_{\dot{h}}, \dot{f} \rangle|}{N_p(\dot{f})} \leq N_{p'}(\dot{h}). \quad (\text{IV, 5; 20})$$

Отображение $\dot{h} \rightarrow u_{\dot{h}}$ является теперь линейным отображением пространства $L^{p'}$ в пространство $(L^p)'$, сопряженное к пространству L^p .

Теорема 54. *Определенное выше отображение $\dot{h} \rightarrow u_{\dot{h}}$ является при $1 \leq p < +\infty$ изометрической биекцией $L^{p'}$ на $(L^p)'$. Если же $p = +\infty$, то оно является изометрической инъекцией $L^{p'}$ в $(L^p)'$, которое не может быть сюръекцией, если носитель μ содержит бесконечное множество точек.*

Перед доказательством теоремы уточним ее смысл: изометрия означает, что $\|u_{\dot{h}}\| = N_{p'}(\dot{h})$, или также, что $\dot{h} \rightarrow u_{\dot{h}}$ являет-

ся инъекцией, поскольку $u_{\dot{h}} = 0$ тогда и только тогда, когда $\dot{h} = 0$. С другой стороны, когда мы говорим, что при $p < +\infty$ имеет место сюръекция, то мы этим хотим сказать, что каждая линейная непрерывная форма на L^p может быть выражена (единственным образом) в виде $u_{\dot{h}}$, где $\dot{h} \in L^{p'}$. Поэтому при $p < +\infty$ можно отождествить пространство $(L^p)'$, сопряженное к пространству L^p , с пространством $L^{p'}$. Однако L^1 является, вообще говоря, только некоторым подпространством пространства, сопряженного к пространству L^∞ . (Следует иметь в виду, что сопряженным к L^1 является пространство L^∞ , однако пространство, сопряженное к L^∞ , не совпадает с пространством L^1 .)

Доказательство. Докажем, что при $p = +\infty$ отображение $\dot{h} \rightarrow u_{\dot{h}}$ изометрично (а, следовательно, инъективно). Обозначим через f функцию $\dot{h}/|\dot{h}|$ (давая ей, например, значение 1 там, где $\dot{h} = 0$). Очевидно, $f \in \mathcal{L}^\infty$ и $|f| \equiv 1$. Из неравенств

$$\|u_{\dot{h}}\| \geq \frac{\langle u_{\dot{h}}, f \rangle}{N_\infty(f)} = \frac{\int h \frac{\dot{h}}{|\dot{h}|} d\mu}{1} \geq \int |h| d\mu = N_1(\dot{h})$$

вытекает, что $\|u_{\dot{h}}\| \geq N_1(\dot{h})$, а, следовательно, $\|u_{\dot{h}}\| = N_1(\dot{h})$, чем и доказывается изометричность в случае $p = +\infty$. Можно было бы провести то же самое рассуждение для $p < +\infty$, но мы в этом случае будем одновременно доказывать и изометричность, и сюръективность. Пусть $1 \leq p < +\infty$, и пусть u — непрерывная линейная форма на L^p . Тогда

$$|\langle u, f \rangle| \leq \|u\| N_p(f). \quad (\text{IV}, 5; 21)$$

Продолжим линейную форму u на пространство \mathcal{L}^p , полагая ее равной нулю на всех μ -почти всюду равных нулю функциях. Тогда она будет определять линейную форму на $\mathcal{C}(X)$, очевидно, непрерывную на каждом $\mathcal{C}_K(X)$, где K — компакт из X . Действительно, если φ сходится к нулю в $\mathcal{C}_K(X)$, то она тем более сходится к нулю в \mathcal{L}^p . Значит, форма u определяет некоторую меру Радона v на X .

Если $p = 1, p' = +\infty$, то

$$|v(\varphi)| = |\langle u, \varphi \rangle| \leq \|u\| N_1(\varphi) = \|u\| \mu(|\varphi|). \quad (\text{IV}, 5; 22)$$

Отсюда следует, что мера v абсолютно мажорируется мерой $\|u\| \mu$. Согласно теореме 52—53—54, существует μ -измеримая функция h , $|h| \leq \|u\|$, а, следовательно, $N_\infty(h) \leq \|u\|$, такая, что $v = h\mu$. Если $1 < p < +\infty$, то для того, чтобы показать, что мера v имеет базой меру μ , можно применить теорему 52 Лебега—Никодима. Мера v допускает, как мы это отмечали на стр. 584, борелевское продолжение в $\Gamma(X)$. В свою очередь

форма u продолжима на \mathcal{L}^p и $\Gamma(X) \subset \mathcal{L}^p$. Естественно выяснить, является ли продолжение v на $\Gamma(X)$ сужением u на $\Gamma(X)$. Для этого достаточно доказать, что сужение u на $\Gamma(X)$, являющееся непрерывным в топологии, определяемой пространством \mathcal{L}^p , будет также L -непрерывным. Прежде всего, очевидно, что если некоторая последовательность функций $f_n \in \Gamma(X)$ сходится просто к 0, оставаясь ограниченной и сохраняя свой носитель в фиксированном компакте, то она сходится к 0 в \mathcal{L}^p (частная теорема Лебега 34). Так как форма u непрерывна в \mathcal{L}^p , то $\langle u, f_n \rangle$ сходится к 0. Пусть \mathcal{O} — открытое множество из X с компактным замыканием, и пусть задано $\varepsilon > 0$. Тогда можно найти такой компакт K , что $\mu(\mathcal{O} - K) \leq (\varepsilon/\|u\|)^p$. (Здесь существенно то, что $p < +\infty$.) При этих условиях для каждой функции φ из $\mathcal{C}(X)$, $0 \leq \varphi \leq 1$, равной 1 на некоторой окрестности компакта K и имеющей носитель в \mathcal{O} , имеют место неравенства

$$\begin{aligned} |\langle u, \chi_{\mathcal{O}} \rangle - \langle u, \varphi \rangle| &\leq \|u\| N_p(\chi_{\mathcal{O}} - \varphi) \leq \|u\| N_p(\chi_{\mathcal{O}-K}) = \\ &= \|u\| (\mu(\mathcal{O} - K))^{1/p} \leq \varepsilon, \end{aligned}$$

из которых следует L -непрерывность u на $\Gamma(X)$ и, следовательно, совпадение u и борелевского продолжения v на $\Gamma(X)$. Но тогда, если A является борелевским множеством с компактным замыканием и μ -мерой, равной нулю, то его характеристическая функция μ -почти всюду равна нулю и значение u на этой характеристической функции равно нулю. Но тогда также $v(A) = 0$. Согласно теореме Лебега — Никодима (теорема 52), v является мерой с базой μ . Значит, существует такая локально μ -интегрируемая функция h , что $v = h\mu$.

Пусть M — некоторое число ≥ 0 и K — компакт пространства X . Обозначим через $h_{M,K}$ функцию, определенную так, как на стр. 521: при $h \geq 0$ она равна h в каждой точке K , в которой $|h(x)| \leq M$, и равна 0 во всех остальных точках.

Пусть теперь f — функция $|h_{M,K}|^{p'-2} \bar{h}_{M,K}$, равная нулю там, где $h_{M,K} = 0$. Ее модуль равен $|h_{M,K}|^{p'-1}$ и $|f| = |h_{M,K}|^{p'}$. Мы видели, что каждая измеримая функция μ -почти всюду равна некоторой борелевской функции. Поэтому можно, изменяя при необходимости функцию f на множестве нулевой μ -меры, предполагать ее борелевской и, поскольку $|f| \leq M^{p-1}$, принадлежащей $\Gamma(X)$. Теперь имеем неравенства

$$\begin{aligned} \|u\| &\geq \frac{|\langle u, f \rangle|}{N_p(f)} = \frac{\left| \int f h d\mu \right|}{N_p(f)} = \frac{\int |h_{M,K}|^{p'} d\mu}{\left(\int |h_{M,K}|^{(p'-1)p} d\mu \right)^{1/p}} = \\ &= \left(\int |h_{M,K}|^{p'} d\mu \right)^{\frac{1}{p'}} = N_{p'}(h_{M,K}). \quad (\text{IV}, 5; 23) \end{aligned}$$

Так как эти неравенства справедливы при любых M и K , то, переходя к точной верхней грани, можно получить неравенство

$$N_{p'}(h) \leq \|u\| \text{ и, следовательно, } h \in \mathcal{L}^{p'}. \quad (\text{IV}, 5; 24)$$

Теперь линейная форма u и форма u_h , определенная, исходя из $h \in \mathcal{L}^{p'}$, совпадают на множестве $\mathcal{C}(X)$. Так как множество $\mathcal{C}(X)$ плотно в \mathcal{L}^p (еще раз в силу того, что $p < +\infty$; см. теорему 49), то эти формы совпадают на \mathcal{L}^p . Этим доказывается сюръективность, а из неравенств (IV, 5; 20) и (IV, 5; 24) одновременно следует изометричность $\|u_h\| = N_{p'}(h)$, чем и заканчивается доказательство позитивных частей теоремы. Что же касается ее негативной части, а именно того, что каждый раз, когда носитель μ содержит бесконечное множество точек, отображение $h \rightarrow u_h$ не сюръективно и, следовательно, сопряженное к пространству L^∞ не может быть отождествлено с L^1 , то эту часть мы примем без доказательства.

Замечание. Пусть теперь \vec{E} — произвольное $\overset{\leftarrow}{\text{банахово}}$ пространство. Тогда каждый элемент $\overset{\leftarrow}{h}$ из $L^{p'}(\vec{E})$ определяет на $L^p(\vec{E})$ некоторую линейную непрерывную форму $u_{\overset{\leftarrow}{h}}$ по формуле

$$\langle u_{\overset{\leftarrow}{h}}, \vec{f} \rangle = \int \langle \overset{\leftarrow}{h}, \vec{f} \rangle d\mu, \quad (\text{IV}, 5; 25)$$

и, кроме того, с помощью того же неравенства Гельдера мы получаем неравенство $\|u_{\overset{\leftarrow}{h}}\| \leq N_{p'}(\overset{\leftarrow}{h})$. Тем самым определено еще одно отображение $\overset{\leftarrow}{h} \rightarrow u_{\overset{\leftarrow}{h}}$ пространства $L^{p'}(\vec{E})$ в пространство $(L^p(\vec{E}))'$, сопряженное к пространству $L^p(\vec{E})$.

Можно показать, что если \vec{E} конечномерно и $p < +\infty$, то это отображение является изометрической биекцией, а это позволяет еще раз отождествить пространство, сопряженное к пространству $L^p(\vec{E})$, с пространством $L^{p'}(\vec{E})$. Этот факт мы примем без доказательства. Однако без дополнительных условий это утверждение не верно, если \vec{E} окажется бесконечномерным банаховым пространством.

§ 6. ОБРАЗ МЕРЫ ПРИ ОТОБРАЖЕНИИ

Если μ — мера, определенная на локально компактном пространстве X и принимающая значения в векторном нормированном пространстве \vec{E} , и если L — линейное непрерывное

отображение \vec{E} в векторное нормированное пространство \vec{F} , то, очевидно, можно определить новую меру $L \circ \mu$ по формуле

$$(L \circ \mu)(\varphi) = L(\vec{\mu}(\varphi)). \quad (\text{IV}, 6; 1)$$

В самом деле, $L \circ \mu$ является линейным отображением $\mathcal{C}(X)$ в \vec{F} , и если функция φ имеет носитель в компакте K , то имеет место оценка

$$\|(L \circ \mu)(\varphi)\| \leq \|L\| \|\vec{\mu}\|_K \|\varphi\|, \quad \text{откуда} \quad \|L \circ \mu\|_K \leq \|L\| \|\vec{\mu}\|_K. \quad (\text{IV}, 6; 2)$$

Эта мера называется *образом меры μ при линейном непрерывном отображении L* .

Однако мы хотим рассмотреть образ другого характера, играющий весьма важную роль в практике. Пусть μ — мера со значениями в \vec{E} , определенная на локально компактном пространстве X , и пусть H — отображение X в локально компактное пространство Y . Покажем, что если H обладает необходимыми свойствами, то существует образ меры, обозначаемый через $H\vec{\mu}$, являющейся мерой на локально компактном пространстве Y со значениями в \vec{E} .

Известно, что если функция φ принадлежит $\mathcal{C}(Y)$, то она имеет прообраз $H^*\varphi$, являющийся функцией на X , определенной по формуле

$$H^*\varphi = \varphi \circ H, \quad \text{или} \quad (H^*\varphi)(x) = \varphi(H(x)). \quad (\text{IV}, 6; 3)$$

Тем самым мы естественным путем пришли к тому, чтобы определить образ $H\vec{\mu}$ по формуле

$$(H\vec{\mu})(\varphi) = \vec{\mu}(H^*\varphi) = \int \varphi(H(x)) d\vec{\mu}(x). \quad (\text{IV}, 6; 4)$$

Однако это определение может иметь смысл только в том случае, когда имеет смысл правая часть равенства (IV, 6; 4) и когда, кроме того, она в самом деле определяет некоторую меру на Y со значениями в \vec{E} . Разберем два возможных случая.

1-й случай. H — непрерывное собственное отображение.

Определение. Непрерывное отображение H локально компактного пространства X в локально компактное пространство Y называется *собственным*, или *непрерывным в бесконечности*, если прообраз каждого компакта из Y при отображении H является некоторым компактом пространства X .

Если X компактно, то каждое непрерывное отображение H пространства X в Y является собственным. Действительно, каждый компакт из Y замкнут, а, следовательно, его прообраз при отображении H замкнут в X , и поскольку X компактно, то он компактен (теоремы 21 и 22 гл. II).

Если X не компактно, то постоянное отображение X в Y не может быть собственным. Ортогональная проекция на одну из координатных осей плоскости \mathbb{R}^2 не является собственным отображением.

Теорема 55. *Если X и Y — метрические пространства, в которых все замкнутые шары компактны, то непрерывное отображение H будет собственным тогда и только тогда, когда прообраз каждого ограниченного подмножества Y при отображении H является ограниченным подмножеством пространства X .*

В самом деле, так как отображение H непрерывно, то прообраз замкнутого множества замкнут, и поскольку замкнутые шары компактны, то компакты совпадают с замкнутыми ограниченными множествами¹⁾.

Следствие. *Если X и Y — такие метрические пространства, в которых замкнутые шары компактны, то непрерывное отображение H пространства X в Y будет собственным тогда и только тогда, когда образ при отображении H каждой неограниченно удаляющейся последовательности точек в X является неограничено удаляющейся последовательностью точек в Y .*

Говорят, что последовательность точек x_n метрического пространства неограниченно удаляется при n , стремящемся к бесконечности, если расстояние от точек x_n до некоторой фиксированной точки a стремится к $+\infty$. Так как $d(b, x_n) \geq d(a, x_n) - d(a, b)$, то в этом случае расстояние x_n до любой другой фиксированной точки b также стремится к $+\infty$.

Доказательство. 1°) Предположим сначала, что отображение H собственное. Тогда, если последовательность точек x_n неограниченно удаляется на X , то $H(x_n)$ неограниченно удаляется на Y . В противном случае можно было бы найти подпоследовательность, содержащуюся в некотором шаре, т. е. в некоторой ограниченной части Y , и тогда, в силу того что отображение H собственное, соответствующая подпоследовательность x_n должна, в силу теоремы, оставаться ограниченной, что противоречит предположению о том, что x_n неограниченно удаляется.

¹⁾ Компактная часть произвольного метрического пространства замкнута и ограничена. Обратно, в метрическом пространстве, в котором замкнутые шары компактны, замкнутая ограниченная часть содержитя в некотором замкнутом шаре, т. е. в некотором компакте, и так как она замкнута, то она компактна. Это — доказательство п. 1°) и 2°а) теоремы 23 гл. II.

2°) Обратно, пусть рассматриваемое свойство выполнено, т. е. пусть отображение H каждую неограниченно удаляющуюся последовательность в X преобразует в неограниченно удаляющуюся последовательность в Y . Тогда отображение H собственное. В самом деле, если B — некоторая ограниченная часть Y , то ее прообраз должен быть ограниченным. В противном случае в нем можно найти неограниченно удаляющуюся последовательность точек, образы которых будут также неограниченно удаляющимися, что противоречит тому, что эти образы лежат в ограниченной части B . Согласно теореме, отображение H собственное.

Пример. Отображения $x \rightarrow x^n$ при произвольном целом $n \geq 1$ являются собственными отображениями вещественной прямой \mathbb{R} на себя.

Название «непрерывное в бесконечности» для собственного отображения оправдывается доказанным следствием. Если к множеству X добавить «бесконечно удаленную точку» и то же самое сделать для множества Y , а затем отображение H продолжить, полагая, что образом бесконечно удаленной точки в X при отображении H является бесконечно удаленная точка Y , то отображение H будет собственным тогда и только тогда, когда таким способом продолженное отображение будет непрерывным в бесконечно удаленной точке.

Теорема 56. *Если H — собственное отображение локально компактного пространства X в локально компактное пространство Y , то образ каждой замкнутой части пространства X при отображении H является замкнутой частью пространства Y).*

Доказательство. Пусть A — замкнутая часть пространства X . Положим $B = H(A)$ и докажем, что B замкнуто. Надо показать, что каждая точка прикосновения $b \in Y$ множества B принадлежит B . Обозначим через β компактный шар с центром в точке b . Поскольку b является точкой прикосновения множества B , то этот шар пересекается с B и, кроме того, b является точкой прикосновения пересечения $\beta \cap B$ ²⁾. Так как отображение H , по предположению, собственное, то прообраз $H^{-1}(\beta)$ является компактной частью X . Он пересекается с A по некоторой компактной его части K . Если теперь точка x принадлежит K , то ее образ принадлежит одновременно β и B , а, следовательно, и их пересечению $\beta \cap B$. Обратно, если некоторая точка y принадле-

¹⁾ Если H непрерывно, то прообраз замкнутого множества замкнут и образ компакта компактен. Если, кроме того, H собственное, то прообраз компакта компактен и образ замкнутого множества есть замкнутое множество.

²⁾ Напомним, что b является точкой прикосновения множества, если каждый шар с центром в точке b пересекается с этим множеством.

жит пересечению $\beta \cap B$, то она принадлежит B и, следовательно, является образом по крайней мере одной точки $x_0 \in A$, а поскольку она принадлежит β , то $x_0 \in H^{-1}(\beta)$ т. е. $x_0 \in K$. Таким образом, образ $H(K)$ совпадает с пересечением $\beta \cap B$. Так как отображение H непрерывно, а K компактно, то отсюда следует, что множество $\beta \cap B$ компактно, а, следовательно, замкнуто. Значит, точка b , являющаяся его точкой прикосновения, принадлежит этому множеству и тем самым принадлежит множеству B , что и требовалось доказать.

Собственные непрерывные отображения дадут нам пример отображений, при которых образ $H\mu$ имеет смысл.

Теорема 57. *Если X и Y — локально компактные пространства, \vec{E} — нормированное векторное пространство, μ — некоторая мера на X со значениями в \vec{E} и носителем A и если H является непрерывным отображением X в Y , сужение которого на A есть собственное отображение A в Y , то формула (IV, 6; 4) определяет $H\mu$ как некоторую меру на Y , называемую образом μ при отображении H .*

Эти условия, естественно, будут выполнены, если отображение H является непрерывным и собственным отображением X в Y или если отображение H непрерывно и μ имеет компактный носитель.

Доказательство. Пусть ϕ — некоторая функция из $\mathcal{C}(Y)$ с компактным носителем K . Ее прообраз $H^*\phi = \phi \circ H$ непрерывен как композиция двух непрерывных отображений. Выясним, каким будет носитель этого отображения. Пусть Ω — множество таких точек y , что $\phi(y) \neq 0$. Согласно определению носителя функции ϕ , замыкание $\bar{\Omega} = K$. Множество точек x , в которых функция $H^*\phi$ отлична от нуля, является прообразом $H^{-1}(\Omega)$. Поскольку прообраз $H^{-1}(K)$ замкнут и содержит $H^{-1}(\Omega)$, носитель $H^*\phi$, очевидно, содержится в $H^{-1}(K)$. Естественно, этот прообраз $H^{-1}(K)$ не обязательно компактен. Он был бы компактным, если бы отображение H пространства X в пространство Y было бы собственным, но это предположение несколько ограничительно. Мы будем предполагать, что только сужение H на A является собственным. Это означает, что пересечение $H^{-1}(K)$ и A является компактной частью A . Но тогда, как мы видели в теореме 16, можно придать смысл выражению $\mu(H^*\phi)$, поскольку $H^*\phi$ — непрерывная скалярная функция и поскольку пересечение носителя μ и носителя $H^*\phi$ компактно. Теперь видно, что выражение, стоящее справа в (IV, 6; 4), имеет определенный смысл и

линейно зависит от φ . Если φ сохраняет свой носитель в фиксированном компакте K пространства Y и равномерно сходится к 0, то $H^*\varphi$ сохраняет свой носитель в замкнутом фиксированном множестве $H^{-1}(K)$, пересечение которого с A компактно, и сходится равномерно к 0. Из теоремы 16 следует, что $\mu(H^*\varphi)$ также сходится к 0. Этим доказывается, что правая часть соотношения (IV, 6; 4) определяет некоторую меру на Y со значениями на \vec{E} и, следовательно, $\vec{H}\mu$ существует¹⁾.

Теорема 58. 1°) В условиях теоремы 57 относительно норм имеет место неравенство:

$$\|\vec{H}\mu\| \leq \|\mu\| \leq +\infty. \quad (\text{IV, 6; 5})$$

2°) Если мера μ комплексна (соответственно вещественна, соответственно вещественна ≥ 0), то такой же будет и мера $\vec{H}\mu$.

3°) Носитель $\vec{H}\mu$ содержится в образе носителя μ при отображении H .

4°) Если носители мер μ_1 и μ_2 лежат в таком замкнутом множестве A , что сужение H на A является собственным, то

$$\begin{aligned} H(\vec{\mu}_1 + \vec{\mu}_2) &= \vec{H}\mu_1 + \vec{H}\mu_2, \\ H(k\vec{\mu}) &= k(\vec{H}\mu), \end{aligned} \quad (\text{IV, 6; 6})$$

иначе говоря, отображение $\mu \rightarrow \vec{H}\mu$ линейно²⁾.

Доказательство. 1°) Предположим, что функция $\varphi \in \mathcal{C}(Y)$ такова, что $|\varphi| \leq 1$. Тогда также $|H^*\varphi| \leq 1$. В силу (IV, 2; 33₂), отсюда получаем неравенство

$$\|\vec{H}\mu(\varphi)\| = \|\vec{\mu}(H^*\varphi)\| \leq \|\vec{\mu}\|, \quad (\text{IV, 6; 7})$$

из которого следует неравенство (IV, 6; 5).

2°) Очевидно.

3°) Пусть A — носитель μ в X , и пусть B — его образ $H(A)$. По теореме 56 множество B замкнуто. Пусть теперь φ — некоторая функция из $\mathcal{C}(Y)$, носитель которой K не имеет общих точек с B . Тогда носитель функции $H^*\varphi$, который содержится в прообразе $H^{-1}(K)$ носителя K функции φ при отображении H , не может пересечься с A . Следовательно, $\mu(H^*\varphi) = 0$, а, зна-

1) Примеры мы приведем позже на стр. 628. Однако полезно посмотреть их сейчас.

2) Заметим, что отображение H непрерывно и собственно, но не обладает свойством линейности, а локально компактные пространства X и Y не являются векторными. Однако, несмотря на это, отображение $\mu \rightarrow \vec{H}\mu$ линейно!

чит, и $(H\mu)\phi = 0$. Другими словами, мера $H\mu$ равна нулю в дополнении к B , а ее носитель лежит в B . Этот носитель может оказаться строго меньше B . Тем не менее легко показать, что если мера μ вещественна и ≥ 0 , то носитель меры $H\mu$ в точности равен B^1 .

4) Очевидно.

2-й случай. Мера μ вещественна и ≥ 0 .

Теорема 59. Пусть X и Y — локально компактные пространства, счетные в бесконечности, μ — вещественная мера ≥ 0 на X и H — такое μ -измеримое отображение X в Y , что прообраз при этом отображении каждого компакта из Y имеет конечную меру относительно μ . Тогда формула (IV, 6; 4) определяет $H\mu$ как некоторую меру ≥ 0 на Y . Кроме того, если K является некоторым компактом из Y , то имеет место оценка

$$\|H\mu\|_K \leq \int_{H^{-1}(K)} d\mu, \text{ откуда } \|H\mu\| \leq \|\mu\|. \quad (\text{IV, 6; 8})$$

Мы будем говорить, что отображение H μ -собственное, если оно удовлетворяет предыдущим условиям. Эти условия всегда выполняются, если H измеримо, а μ имеет конечную норму.

Доказательство. Последнее утверждение очевидно: если отображение H измеримо и мера μ имеет конечную норму, т. е. $\mu(X) < +\infty$, то для каждого компакта K множество $H^{-1}(K)$ измеримо и имеет конечную меру.

Пусть теперь задана некоторая функция $\phi \in \mathcal{C}(Y)$. Согласно теореме 22, ее прообраз $H^*\phi = \phi \circ H$ является комплексной μ -измеримой функцией на X . Эта функция ограничена величиной $\|\phi\|$. Поскольку множество точек, в которых эта функция $\neq 0$, очевидно, содержится в прообразе носителя K функции ϕ при отображении H и, значит, в измеримом множестве конечной меры относительно μ , то $\int |(H^*\phi)| d\mu$ конечен. Из теоремы 39 теперь вытекает, что $H^*\phi$ μ -интегрируема. Поэтому можно придать смысл правой части равенства (IV, 6; 4). Эта правая часть линейно зависит от ϕ и ≥ 0 , если $\phi \geq 0$.

С другой стороны, если ϕ имеет носитель в K , то $\|H^*\phi\|$ равна $\|\phi\|$, и имеет место оценка

$$|(H\mu)(\phi)| \leq \|\phi\| \int_{H^{-1}(K)} d\mu, \quad (\text{IV, 6; 9})$$

¹⁾ Это вытекает из следствия 1 теоремы 60. В самом деле, пусть $B_1 \subset B$ является носителем меры $H\mu$. Если $B_1 \neq B$, то CB_1 является открытым множеством $H\mu$ -меры нуль, пересекающимся с B . Следовательно, $H^{-1}(CB_1)$ будет открытым множеством нулевой μ -меры, пересекающимся с носителем A меры μ , что невозможно.

из которой следует, что $H\mu$ непрерывна на $\mathcal{C}_k(Y)$ и, значит, является некоторой мерой, и что имеет место неравенство (IV, 6; 8).

Теорема 60. Пусть в условиях теоремы 59 \vec{f} — некоторое отображение пространства Y в банахово пространство \vec{F} . Для того чтобы отображение \vec{f} было $H\mu$ -интегрируемым (соответственно $H\mu$ -измеримым), необходимо и достаточно, чтобы прообраз $H^*\vec{f} = \vec{f} \circ H$ был μ -интегрируемой функцией на X со значениями в \vec{F} (соответственно μ -измеримой), и тогда имеет место следующее равенство:

$$\int \vec{f} d(H\mu) = \int (H^*\vec{f}) d\mu = \int \vec{f}(H(x)) d\mu(x). \quad (\text{IV, 6; 10})$$

Эту теорему мы примем без доказательства, в котором следовало бы воспользоваться равенством

$$\int^* \vec{f} d(H\mu) = \int^* (H^*\vec{f}) d\mu \leq +\infty, \quad (\text{IV, 6; 11})$$

справедливым для вещественных функций ≥ 0 .

Легко получается

Следствие 1. Пусть выполнены условия теоремы 59 и B — некоторая часть Y . Часть B $H\mu$ -измерима тогда и только тогда, когда μ -измерим ее прообраз $H^{-1}(B)$, и в этом случае имеет место равенство

$$H\mu(B) = \mu(H^{-1}(B)) \leq +\infty. \quad (\text{IV, 6; 12})$$

Следствие 2. Справедливо равенство

$$\|H\mu\| = \|\mu\| \leq +\infty. \quad (\text{IV, 6; 13})$$

Для доказательства достаточно применить следствие 1 к множеству $B = Y$.

Теорема 61 (о транзитивности образов мер). Пусть X, Y, Z — три локально компактных пространства, H — отображение X в Y , K — отображение Y в Z и μ — мера на X со значением в \vec{E} .

1°) Если H и K непрерывны и если сужение сложного отображения $K \circ H$ пространства X в Z на носитель μ является собственным, то $K(H\mu)$ и $(K \circ H)\mu$, согласно теореме 57, имеют смысл и равны между собой.

2°) Если мера μ вещественна и ≥ 0 , X, Y, Z счетны в бесконечности и, в условиях теоремы 59, существует $K(H\mu)$, то, в силу этой теоремы, существует $(K \circ H)(\mu)$ и имеет место тождество $K(H\mu) = (K \circ H)\mu$.

Доказательство. 1-й случай. Пусть A — носитель меры μ и $B = H(A)$. Пусть C' — компактная часть множества Z . Тогда образ пересечения $H^{-1}(K^{-1}(C')) \cap A$ при отображении H равен пересечению $K^{-1}(C') \cap B$. Так как $K \circ H$ по условию является собственным на A , то первое из этих множеств компактно. Тогда второе множество также компактно, поскольку оно является образом компакта при непрерывном отображении, а это означает, что сужение K на B является собственным.

Пусть теперь B' — компактная часть пространства Y и C' — ее образ при отображении K . Тогда пересечение $H^{-1}(B') \cap A$ содержится в пересечении $H^{-1}(K^{-1}(C')) \cap A$, которое по предположению компактно. Таким образом, прообраз каждого компакта из Y при сужении H на A является замкнутым множеством, содержащимся в некотором компакте X и, значит, является компактом. Это говорит о том, что H является собственным отображением на носителе A меры μ . Теперь мера $H\mu$ имеет смысл, а ее носитель лежит в B (теорема 58, 3°). Следовательно, K , будучи собственным на B , является собственным на этом носителе. Согласно теореме 57, $K(H\mu)$ имеет смысл. Пусть теперь φ — некоторая функция из $\mathcal{C}(Z)$. Тогда справедливы равенства

$$((K \circ H) \mu)(\varphi) = \int \varphi(K(H(x))) d\mu(x), \quad (IV, 6; 14)$$

$$(K(H\mu))(\varphi) = \int \varphi(K(y)) d(H\mu(y)) = \int \varphi(K(H(x))) d\mu(x),$$

откуда $(K \circ H) \mu = K(H\mu)$.

2-й случай: $\mu \geqslant 0$. Пусть теперь отображение H μ -собственное, а отображение K $(H\mu)$ -собственное. Пусть C' — замкнутое подмножество Z . Так как отображение K $H\mu$ -измеримо, то прообраз $K^{-1}(C')$ является $H\mu$ -измеримым. Но тогда, согласно следствию 1 теоремы 60, в силу того, что H является μ -собственным, его прообраз $H^{-1}(K^{-1}(C'))$ необходимо μ -измерим, а это означает, что $K \circ H$ есть μ -измеримое отображение.

Кроме того, если C' компактно и отображение K $H\mu$ -собственное, то $K^{-1}(C')$ имеет конечную $H\mu$ -меру. Отсюда в силу того, что H μ -собственное, вытекает, что $H^{-1}(K^{-1}(C'))$ имеет конечную μ -меру. Это доказывает, что $K \circ H$ является μ -собственным. Пусть теперь φ — некоторый элемент $\mathcal{C}(Z)$.

По определению $(K \circ H) \mu$, имеем

$$((K \circ H) \mu)(\varphi) = \int \varphi(K(H(x))) d\mu(x). \quad (IV, 6; 15)$$

Согласно теореме 60, это выражение можно записать в виде

$$\int \varphi(K(y)) d(H\mu(y)) \text{ или } (K(H\mu))(\varphi), \quad (IV, 6; 16)$$

чем и доказывается равенство $(K \circ H) \mu = K(H \mu)$,

З а м е ч а н и е. Условия применимости обоих случаев являются *крайне различными*. В первом случае указанное условие предполагалось выполненным для $K \circ H$, а не последовательно для H и для K , тогда как во втором случае условия предполагались выполненными последовательно для H и для K , но не для $K \circ H$. Легко доказать, что условия такого рода не могут быть переставлены. Например, в первом случае может случиться, что H непрерывно и является собственным на носителе μ , K непрерывно и является собственным на носителе $H\mu$, и тем не менее $K \circ H$ не является собственным на носителе μ^1). Во втором случае вполне может оказаться, что $K \circ H$ μ -измеримо, хотя H не является μ -измеримым²⁾.

Случай, когда H является гомеоморфизмом X на Y

Если отображение H — гомеоморфизм, то оно, очевидно, удовлетворяет всем условиям, позволяющим определить $H\mu$ для меры μ , определенной на X . В этом случае существует обратное отображение H^{-1} , которое также является некоторым гомеоморфизмом. При этом, согласно теореме 61, $H^{-1}(H\mu) = (H^{-1} \circ H)\mu = \mu$. Впрочем, в этом случае, полагая $H^*\nu = H^{-1}\nu$, можно также определить прообраз при отображении H меры ν , определенной на Y , и образ функции $\varphi \in \mathcal{C}(X)$ при этом отображении, если положить $H\varphi = (H^{-1})^*\varphi = \varphi \circ H^{-1}$. Для каждой меры μ на X и для каждой функции $\varphi \in \mathcal{C}(X)$ имеют место соотношения

$$(H\varphi)(y) = \varphi(H^{-1}(y)); \quad H\varphi(H(x)) = \varphi(x); \quad x \in X, y \in Y; \quad (IV, 6; 17)$$

$$H\mu(\psi) = \mu(H^{-1}(\psi)); \quad H\mu(H\varphi) = \mu(\varphi); \quad \varphi \in \mathcal{C}(X), \quad \psi \in \mathcal{C}(Y).$$

Интеграл от Hf по $H\mu$ равен интегралу от f по μ , а мера множества $H(A)$ (если $A \subset X$) по отношению к $H\mu$ равна мере множества A по отношению к μ .

¹⁾ Пусть, например, $X = Y = Z = \mathbb{R}$. В качестве собственного отображения H возьмем $x \rightarrow x^2$. Мера μ равна $x dx$. Тогда $H\mu(\varphi) = \int \varphi(x^2) x dx = 0$, следовательно, $H\mu = 0$. Если отображение K постоянно, то оно собственное на носителе меры $H\mu$ (который пуст). Однако $K \circ H$ постоянно на \mathbb{R} и, следовательно, не собственное.

²⁾ Если в качестве H взять не μ -измеримое отображение, а в качестве K — постоянное отображение, то $K \circ H$ будет постоянным отображением, а значит, μ -измеримым!

Так как $\varphi(H^{-1}(y))$ является функцией $H\varphi$, то первая из формул (IV, 6; 17) побуждает использовать обозначение $d\mu(H^{-1}(y))$ для образа $H\mu$. При этом формула (IV, 6; 10) запишется в очень удобной форме:

$$\int \vec{f}(y) d\mu(H^{-1}(y)) = \int \vec{f}(H(x)) d\mu(x), \quad (\text{IV}, 6; 17_2)$$

как бы полученной «заменой переменной $y = H(x)$ ».

Кроме того, если $d\mu$ имеет вид $p d\lambda$, $d\mu(x) = p(x) d\lambda(x)$, то

$$d\mu(H^{-1}(y)) = p(H^{-1}(y)) d\lambda(H^{-1}(y)). \quad (\text{IV}, 6; 17_3)$$

В самом деле, для $\varphi \in \mathcal{C}(Y)$ имеем:

$$\begin{aligned} \int \varphi(y) d\mu(H^{-1}(y)) &= \int \varphi(H(x)) d\mu(x) = \int \varphi(H(x)) p(x) d\lambda(x) = \\ &= \int \varphi(H(x)) p(H^{-1}(H(x))) d\lambda(x) = \int \varphi(y) [p(H^{-1}(y)) d\lambda(H^{-1}(y))]. \end{aligned} \quad (\text{IV}, 6; 17_4)$$

Обобщение теоремы 59 на случай, когда μ не ≥ 0

Методы, рассмотренные в предыдущем пункте, позволяют с помощью теоремы 59 определить $H\mu$ даже в том случае, когда мера μ векторная, лишь бы только она имела базу ≥ 0 . В самом деле, перед теоремой 54₂ мы выяснили, что следует понимать под μ -интегрируемой скалярной функцией f , не принадлежащей $\mathcal{C}(X)$. Говорят, что $H\mu$ существует, если, какова бы ни была функция $\varphi \in \mathcal{C}(X)$, функция $H^*\varphi$ μ -интегрируема, а если φ равномерно сходится к 0 и сохраняет при этом носитель в фиксированном компакте, то интеграл $\int (H^*\varphi) d\mu$ сходится к 0. В этом случае полагают по определению

$$H\mu(\varphi) = \int (H^*\varphi) d\mu = \int \varphi(H(x)) d\mu(x). \quad (\text{IV}, 6; 18)$$

Это произойдет, в частности, тогда, когда μ имеет конечную норму, \vec{E} конечномерно и H непрерывно, так как, в силу непрерывности и ограниченности $H^*\varphi$, можно воспользоваться следствием теоремы 54₂. Согласно теореме 54, $\mu = q\mu_0$, где мера $\mu_0 \geq 0$, имеет конечную норму и $\|\vec{q}\| = 1$. При этом имеет место неравенство

$$\left\| \int (H^*\varphi) \vec{q} d\mu_0 \right\| \leq \|\mu_0\| \|\varphi\|, \quad (\text{IV}, 6; 18_2)$$

показывающее, что $\|\vec{H\mu}\| \leq \|\mu_0\| = \|\vec{|\mu|}\|$. В действительности можно даже доказать, что $\|\vec{H\mu}\| \leq \|\vec{\mu}\|$.

Различные примеры образов мер

1°) Предположим, что H — постоянное отображение X в Y . Образ X при этом отображении — это некоторая точка $b \in Y$. Здесь

$$H^*\varphi(x) = \varphi(H(x)) = \varphi(b), \quad (\text{IV}, 6; 19)$$

и, следовательно,

$$(H\vec{\mu})(\varphi) = \left(\int d\vec{\mu} \right) \varphi(b), \quad \text{или} \quad H\vec{\mu} = \left(\int d\vec{\mu} \right) \delta_{(b)}. \quad (\text{IV}, 6; 20)$$

Образ меры $\vec{\mu}$ является точечной массой в точке b , равной полной массе меры $\vec{\mu}$, $\int d\vec{\mu}$. Такой же результат ожидается и интуитивно: H переносит X целиком в точку b ; следовательно, это отображение переносит в точку b всю массу.

Этот результат предполагает, что рассмотрение ведется в сформулированных выше условиях, т. е. X компактно, например, или мера μ вещественна, ≥ 0 и имеет конечную норму. Здесь очень хорошо видно, почему надо рассматривать ограничения на H . Если, например, H — постоянное отображение вещественной прямой $X = \mathbb{R}$ в Y и если $d\mu = dx$ есть мера Лебега пространства \mathbb{R} , то отображение H не является μ -собственным. Если бы существовала мера $H\mu$, то она должна была бы иметь массу $+\infty$ в точке $b = H(\mathbb{R})$, что невозможно, ибо точка, будучи компактным подмножеством, не может иметь бесконечную меру.

2°) Предположим, что $\vec{\mu}$ — это линейная комбинация мер Дирака, а именно $\vec{\mu} = \sum_v \vec{c}_v \delta_{(a_v)}$. В этом случае мера $H\vec{\mu}$, если она определена, задается равенством

$$H\vec{\mu} = \sum_v \vec{c}_v \delta_{H(a_v)} \quad \text{и, в частности,} \quad H\delta_{(a)} = \delta_{(H(a))}. \quad (\text{IV}, 6; 21)$$

Этот пример показывает конкретный физический смысл образа меры. Если мера $\vec{\mu}$ составлена из некоторого числа масс, расположенных в некоторых точках, то ее образ получается «переносом» этих масс в соответствующие точки с помощью преобразования H . Априори ясно, что операция (IV, 6; 21) существует для меры, образованной из суммы конечного числа точечных масс. Смысл изложенного в этом пункте заключается в том,

чтобы показать, что перенос меры при отображении H является операцией, возможной в значительно более общих случаях¹⁾.

3°) Предположим теперь, что X — аффинное конечномерное пространство, и пусть \vec{h} — некоторый вектор из присоединенного векторного пространства \vec{X} . Возьмем в качестве преобразования H параллельный перенос $\tau_{\vec{h}}$, определяемый вектором \vec{h} с помощью формулы $x \rightarrow x + \vec{h}$. Это преобразование является гомеоморфизмом (и, следовательно, к нему можно применить различные формулы (IV, 6; 17)). Образ функции φ , определенной на X , задается по формуле

$$\tau_{\vec{h}}\varphi(x) = \varphi(\tau_{-\vec{h}}x) = \varphi(x - \vec{h})^2. \quad (\text{IV}, 6; 22)$$

Образ меры μ на X при переносе \vec{h} является мерой, определяемой по формуле

$$(\tau_{\vec{h}}\mu)(\varphi) = \mu(\tau_{\vec{h}}\varphi) = \mu(\tau_{-\vec{h}}\varphi) = \int \varphi(x + \vec{h}) d\mu(x). \quad (\text{IV}, 6; 23)$$

Согласно теореме 60, в случае, когда мера μ вещественна и ≥ 0 , функция \vec{f} со значениями в банаховом пространстве \vec{F} является $\tau_{\vec{h}}\mu$ -интегрируемой тогда и только тогда, когда функция $\tau_{\vec{h}}\vec{f}: x \rightarrow \vec{f}(x + \vec{h})$ является μ -интегрируемой, и тогда имеет место равенство

$$\int \vec{f} d(\tau_{\vec{h}}\mu) = \int \vec{f}(x + \vec{h}) d\mu(x). \quad (\text{IV}, 6; 24).$$

¹⁾ Отождествим каждую точку $a \in X$ с мерой $\delta_{(a)}$. Отождествим также X с некоторой частью множества $\mathcal{C}'(X)$, а Y отождествим с некоторой частью множества $\mathcal{C}'(Y)$. Тогда отображение $H: x \rightarrow H(x)$ станет отображением $\delta_{(x)} \rightarrow \delta_{H(x)}$, и отображение $\mu \rightarrow H\mu$ есть линейное продолжение отображения H в некоторое отображение $\mathcal{C}'(X)$ (или некоторой его части) в $\mathcal{C}'(Y)$.

²⁾ К этим формулам надо привыкнуть: параллельный перенос φ , определяемый вектором \vec{h} , задается формулой $x \rightarrow \varphi(x - \vec{h})$. Предположим, что X — вещественная прямая \mathbb{R} , и построим график функции $y = \varphi(x)$. Перенести функцию φ означает перенести график φ параллельно оси x . Значение $\tau_{\vec{h}}\varphi$ в точке x является значением φ в точке $\tau_{-\vec{h}}x = x - \vec{h}$.

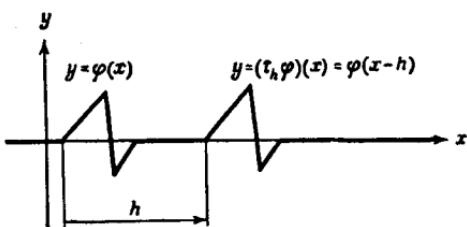


Рис. 17.

В обозначениях (IV, 6; 17₂) можно написать, что

$$\tau_{\vec{h}} \mu = d\mu(x - \vec{h}), \quad (\text{IV}, 6; 24_2)$$

и тогда (IV, 6; 24) запишется в виде

$$\int \vec{f}(x) d\mu(x - \vec{h}) = \int \vec{f}(x + \vec{h}) d\mu(x). \quad (\text{IV}, 6; 24_3)$$

Возьмем $\vec{X} = \mathbb{R}$. Если $d\mu$ имеет вид $p(x)dx$, то с помощью замены переменной $x + h = \xi$ (IV, 6; 24₃) запишется в виде $\int \vec{f}(\xi) p(\xi - h) d\xi$, так что

$$d\mu(x - h) = p(x - h) dx \text{ и } d(x - h) = dx \text{ для } p \equiv 1. \quad (\text{IV}, 6; 24_4)$$

4°) Возьмем теперь в качестве H отображение \mathbb{R} на \mathbb{R} , определяемое гомотетией с центром в начале координат и отношением k . Если $k = 0$, то отображение H постоянно. Этот случай уже был рассмотрен в (IV, 6; 20).

Предположим теперь, что $k \neq 0$. В этом случае H является гомеоморфизмом. По определению, имеет место формула

$$(H\varphi)(x) = \varphi\left(\frac{x}{k}\right). \quad (\text{IV}, 6; 25)$$

Здесь формула (IV, 6; 17₂) дает

$$\int \vec{f}(x) d\mu\left(\frac{x}{k}\right) = \int \vec{f}(kx) d\mu(x). \quad (\text{IV}, 6; 26)$$

Если, в частности, мера μ является мерой $p(x)dx$, где функция p локально интегрируема по dx , то, используя общую формулу замены переменной (IV, 9; 72) (которая будет доказана позже), получаем формулу

$$H\mu(\varphi) = \int \varphi(H(x)) d\mu(x) = \int \varphi(kx) p(x) dx = \int \varphi(y) p\left(\frac{y}{k}\right) \frac{dy}{|k|}, \quad (\text{IV}, 6; 27)$$

откуда

$$d\mu\left(\frac{x}{k}\right) = p\left(\frac{x}{k}\right) \frac{dx}{|k|} \quad \text{и} \quad d\left(\frac{x}{k}\right) = \frac{dx}{|k|}. \quad (\text{IV}, 6; 28)$$

§ 7. ШИРОКАЯ СХОДИМОСТЬ МЕР РАДОНА

Сходимость по норме. Локальная сходимость по норме

Выясним более тщательно, в каком случае следует говорить, что последовательность мер Радона $\vec{\mu}_n$ на X со значениями в векторном пространстве \vec{E} сходится к предельной мере $\vec{\mu}$. Можно было бы считать, что сходимость имеет место, если $\|\vec{\mu}_n - \vec{\mu}\|$

сходится к нулю, когда n стремится к бесконечности. Это — *сходимость по норме*. Но такое определение сходимости не отражает сходимости мер, поскольку норма меры может оказаться бесконечной, и, значит, несмотря на свое название, она не будет настоящей нормой¹⁾). Естественно дать следующее определение.

Говорят, что последовательность мер μ_n на X со значениями в \vec{E} *сходится к предельной мере* μ при n , стремящемся к бесконечности, если, каким бы ни был компакт K множества X , последовательность норм $\|\mu_n - \mu\|_K$ сходится к 0 при n , стремящемся к бесконечности.

Согласно тому, что мы отметили на стр. 149 относительно сходимости, определенной с помощью нормы на пространствах линейных непрерывных отображений одного векторного пространства в другое, это просто означает, что, каким бы ни был компакт K из X , функции μ_n на пространстве $\mathcal{C}(X)$ сходятся к функции μ равномерно относительно единичного шара из $\mathcal{C}_K(X)$. Такая сходимость является *локальной сходимостью по норме*. Говорят, что меры μ_n локально сходятся по норме к мере μ .

Пример. Пусть μ — вещественная фиксированная ≥ 0 мера Радона, и пусть p_n — последовательность функций со значениями в банаховом пространстве \vec{E} , локально μ -интегрируемых, просто μ -почти всюду сходящихся к предельной функции p при n , стремящемся к бесконечности, и таких, что имеет место неравенство $\|p_n(x)\| \leq g(x)$, где g — фиксированная неотрицательная локально μ -интегрируемая функция.

Теперь меры-произведения $p_n \mu$ будут локально сходиться по норме к мере $p \mu$ при n , стремящемся к бесконечности.

В самом деле, пусть K — некоторый компакт X . Тогда имеет место неравенство

$$\|\mu_n - \mu\|_K \leq \int_K \|p_n - p\| d\mu. \quad (\text{IV}, 7; 1)$$

Нормы $\|p_n - p\|$ сходятся просто μ -почти всюду к 0 и мажорируются функцией $2g \geq 0$, μ -интегрируемой на K . Согласно

¹⁾ Это обстоятельство уже было отмечено на стр. 154 относительно сходимости последовательности отображений множества E в метрическое пространство F . Однако предыдущее определение можно взять за определение сходимости на подмножествах, образованных мерами с конечной нормой.

теореме Лебега (теорема 35), правая часть сходится к 0 при n , стремящемся к бесконечности.

Несмотря на естественность этого определения, оно, вообще говоря, слишком жестко.

Рассмотрим, например, последовательность точек a_n из X , стремящихся к a при n , стремящемся к бесконечности. Было бы естественным надеяться, что меры Дирака $\delta_{(a_n)}$ сходятся к мере Дирака $\delta_{(a)}$ при n , стремящемся к бесконечности. Однако при предыдущем определении сходимости этого не будет. В самом деле, если, например, X компактно и если a_n отличны от a , то, согласно (IV, 2; 7), $\|\delta_{(a_n)} - \delta_{(a)}\| = 2$ и, следовательно, эта величина не стремится к 0 при n , стремящемся к бесконечности. Отсюда вытекает необходимость введения нового определения.

Широкая сходимость

Говорят, что последовательность мер Радона μ_n на X со значениями в \vec{E} широко сходится к предельной мере μ при n , стремящемся к бесконечности, если, какова бы ни была функция φ из $\mathcal{C}(X)$, последовательность векторов $\mu_n(\varphi)$ сходится в \vec{E} к вектору $\mu(\varphi)$ при n , стремящемся к бесконечности. Если \vec{E} является полем скаляров, то речь идет о сходимости чисел $\mu_n(\varphi)$ к числу $\mu(\varphi)$. Если μ_n вещественны (соответственно ≥ 0), то то же самое будет и для меры μ . Если меры рассматривать как функции на $\mathcal{C}(X)$, то широкая сходимость является простой сходимостью этих функций и потому ее часто называют простой сходимостью. Естественно, что локальная сходимость по норме влечет за собой широкую сходимость¹⁾.

Примеры. 1°) Если a_n является последовательностью точек из X , сходящейся к a при n , стремящемся к бесконечности, то меры Дирака $\delta_{(a_n)}$ широко сходятся к мере Дирака $\delta_{(a)}$.

В самом деле, какова бы ни была функция $\varphi \in \mathcal{C}(X)$, функции $\varphi(a_n)$ сходятся к $\varphi(a)$, ибо каждая из функций φ непрерывна.

2°) Предположим, что X — конечномерное аффинное пространство. Тогда для каждого вектора $\vec{h} \in \vec{X}$ можно определить параллельный перенос меры Радона $\tau_{\vec{h}} \mu$ (формула (IV, 6; 23)).

¹⁾ Возникла довольно любопытная ситуация. В § 15 гл. II мы видели, что понятие простой сходимости является слишком слабым, и, потому нам пришлось определить равномерную сходимость. Здесь же мы замечаем, что понятие равномерной сходимости является слишком сильным, и мы вынуждены возвратиться к простой сходимости!

Если теперь \vec{h}_n является последовательностью векторов из \vec{X} , сходящейся к \vec{h} , то перенесенные меры $\tau_{\vec{h}_n} \vec{\mu}$ широко сходятся к перенесенной мере $\tau_{\vec{h}} \vec{\mu}$. В самом деле, согласно определению,

$$(\tau_{\vec{h}_n} \vec{\mu} - \tau_{\vec{h}} \vec{\mu})(\varphi) = \int (\varphi(x + \vec{h}_n) - \varphi(x + \vec{h})) d\vec{\mu}(x). \quad (\text{IV}, 7; 2)$$

Наше утверждение будет доказано, если мы покажем, что функции $x \rightarrow \varphi(x + \vec{h}_n)$ сходятся равномерно к функции $x \rightarrow \varphi(x + \vec{h})$. В самом деле, эти функции сохраняют свои носители в фиксированном компакте K . (Пусть K_0 — носитель φ . Он компактен и, следовательно, если на X ввести некоторую норму, K_0 будет содержаться в некотором замкнутом шаре с центром a_0 радиуса R_0 . Пусть теперь r — точная верхняя грань всех $\|\vec{h}_n\|$. Тогда носитель $\tau_{-\vec{h}_n} \varphi$ будет содержаться в компактном шаре с центром a_0 радиуса $R_0 + r$.) Следовательно, они сходятся в $\mathcal{C}_K(X)$. Наше утверждение будет вытекать из того, что $\vec{\mu}$ является непрерывной функцией на пространстве $\mathcal{C}_K(X)$.

Докажем поэтому, что $\vec{\mu}$ имеет место равномерная сходимость. Так как функция φ непрерывна и имеет компактный носитель, то она равномерно непрерывна. В самом деле, пусть, как было сказано выше, $B_0(a_0, R_0)$ — некоторый шар, содержащий носитель K_0 функции φ . Функция φ на компакте $B(a_0, R_0 + 1)$ равномерно непрерывна (теорема 31 гл. II). Следовательно, при заданном $\varepsilon > 0$ найдется такое η , $0 < \eta \leq 1$, что из неравенства $\|x' - x''\| \leq \eta$ вытекает неравенство $|\varphi(x') - \varphi(x'')| \leq \varepsilon$ для любых x' и x'' из $B(a_0, R_0 + 1)$. Если x' и x'' одновременно не лежат в предыдущем шаре и удовлетворяют неравенству $\|x' - x''\| \leq \eta$, то они лежат вне шара $B(a_0, R_0)$ и тем более вне носителя K_0 функции φ . Поэтому $\varphi(x') = \varphi(x'') = 0$, так что окончательно во всех случаях имеет место неравенство $|\varphi(x') - \varphi(x'')| \leq \varepsilon$, означающее, что функция φ равномерно непрерывна на X .

Если теперь $\|\vec{h}_n - \vec{h}\| \leq \eta$, то для любой точки $x \in E$ справедливо неравенство $|\varphi(x + \vec{h}_n) - \varphi(x + \vec{h})| \leq \varepsilon$, из которого вытекает равномерная сходимость $\tau_{-\vec{h}_n} \varphi$ к $\tau_{-\vec{h}} \varphi$ при \vec{h}_n , стремящемся к \vec{h} ; мы тем самым доказали, что $\tau_{\vec{h}_n} \vec{\mu}$ широко сходится к $\tau_{\vec{h}} \vec{\mu}$, когда \vec{h}_n стремится к \vec{h} .

3°) Если в предыдущем примере μ является мерой $\geqslant 0$ (или, более общо, с базой $\geqslant 0$), то доказательство становится еще проще. В самом деле, равномерная сходимость функции $\varphi(x + \vec{h}_n)$ к $\varphi(x + \vec{h})$ не является необходимой. Имеет место простая сходимость; функции $\varphi(x + \vec{h}_n)$ ограничены по модулю числом $\|\varphi\|$ и сохраняют свой носитель в фиксированном компакте. Поэтому результат вытекает из теоремы 34 Лебега. Полученный результат можно обобщить следующим образом:

Пусть H_n — последовательность непрерывных отображений X в Y , просто сходящихся при n , стремящемся к бесконечности, к некоторому непрерывному отображению H . Пусть μ — некоторая мера $\geqslant 0$ на X с конечной нормой¹⁾. Тогда образы мер $H_n\mu$ широко сходятся к образу меры $H\mu$.

В самом деле, для $H_n\mu$, $H\mu$ и $\varphi \in \mathcal{C}(Y)$ имеет место формула (IV, 6; 4). При n , стремящемся к бесконечности, $\varphi(H_n(x))$ сходится просто к $\varphi(H(x))$ и при этом по-прежнему остается ограниченной постоянной $\|\varphi\|$. Эта постоянная интегрируема, поскольку μ имеет конечную норму. Результат вытекает из теоремы 35 Лебега.

4°) Если $X = [a, b]$ — ограниченный интервал \mathbb{R} , то мера

$$\mu_n = \frac{1}{n+1} \left(\delta_{(a)} + \delta_{\left(a + \frac{b-a}{n}\right)} + \dots + \delta_{\left(a + k \frac{b-a}{n}\right)} + \dots + \delta_{(b)} \right) \quad (\text{IV}, 7; 3)$$

при n , стремящемся к бесконечности, широко сходится к мере $dx/(b-a)$. В самом деле, если φ принадлежит $\mathcal{C}([a, b])$, то имеет место формула о среднем (IV, 1; 40), доказывающая наше утверждение.

Функции, μ -интегрируемые по Риману

Рассмотрим следующую задачу. Пусть $\vec{\mu}_n$ — последовательность мер Радона, широко сходящихся к предельной мере $\vec{\mu}$ при n , стремящемся к бесконечности. Можно ли тогда утверждать, что $\int \varphi d\vec{\mu}_n$ сходится к $\int \varphi d\vec{\mu}$ при n , стремящемся к бесконечности, для таких функций φ , которые не принадлежат $\mathcal{C}(X)$? Ясно, что это неверно при произвольной функции φ . Рассмотрим, например, снова последовательность точек a_n из X , сходящихся к a , при n , стремящемся к бесконечности. Мы видели, что $\delta_{(a_n)}$ широко сходится к $\delta_{(a)}$. Если теперь φ — некоторая функция,

¹⁾ Для простоты мы предположим, что μ имеет конечную норму. Можно было бы исходить из других предположений типа «собственных отображений».

разрывная в точке a , то, очевидно, $\varphi(a_n)$ не будет, вообще говоря, сходиться к $\varphi(a)$, что и является контрпримером для предыдущего предложения.

Следовательно, нужно рассматривать особые функции φ . Мы ограничимся случаем вещественных мер ≥ 0 .

Определение. Пусть μ — вещественная мера Радона ≥ 0 на X и f — вещественная функция ≥ 0 на X , ограниченная и имеющая компактный носитель. *Верхним интегралом Римана* функции f относительно меры μ , $\int^{*R} f d\mu$, называется точная нижняя грань мер $\mu(\varphi)$ по всевозможным функциям $\varphi \in \mathcal{C}(X)$, мажорирующим функцию f .

На первый взгляд это определение кажется странным, так как оно не соответствует определению верхнего интеграла Римана относительно dx , которое было дано в § 1. Однако легко показать, что когда мера μ является мерой dx на \mathbb{R} , то оба определения совпадают. Обозначим через \int^{*R_1} и \int^{*R_2} верхние интегралы Римана в смысле § 1 и в смысле введенного определения при $\mu = dx$. Пусть φ — некоторая функция из $\mathcal{C}(X)$, мажорирующая функцию f . Так как любая функция из $\mathcal{C}(X)$ интегрируема по Риману в смысле § 1, то $\int^{*R_1} f dx \leq \int^{*R_1} \varphi dx = \mu(\varphi)$, откуда, переходя к точной нижней грани по всевозможным функциям φ , получаем неравенство $\int^{*R_1} f dx \leq \int^{*R_2} f dx$. С другой стороны, можно найти такую ступенчатую функцию с компактным носителем $g \geq 0$, что $\int^{*R_1} f dx \geq \int g dx - \frac{\varepsilon}{2}$. Функция g имеет вид $\sum_{i=1}^N g_i \chi_{[a_i, b_i]}$, где g_i постоянны и $[a_i, b_i]$ суть N интервалов из \mathbb{R} . Положим $M = \max_{i=1}^N g_i$, $a' = a - \varepsilon/4MN$, $b' = b + \varepsilon/4MN$ и обозначим через φ_i функцию, равную 1 на $[a_i, b_i]$, 0 на $]-\infty, a'_i]$ и $[b'_i, +\infty[$ и аффинную в каждом из интервалов $[a'_i, a_i]$, $[b_i, b'_i]$. Обозначим через φ функцию $\sum_{i=1}^N g_i \varphi_i$. Тогда, с одной стороны, $f \leq \varphi$, а, с другой стороны, справедливо неравенство

$$\int g dx \geq \int \varphi dx - \frac{\varepsilon}{4MN} \cdot M \cdot 2N = \int \varphi dx - \frac{\varepsilon}{2},$$

из которого вытекает неравенство

$$\int^{*R_1} f dx \geq \int \varphi dx - \varepsilon \geq \int^{*R_2} f dx - \varepsilon.$$

Так как ε произвольно, то полученные соотношения показывают, что оба понятия верхнего интеграла Римана совпадают. В соответствии с тем, что мы видели в примере для dx (стр. 532), верхний интеграл Римана, который всегда не меньше верхнего интеграла Лебега, вообще говоря, строго больше его. Если μ — мера Радона на вещественной прямой \mathbb{R} , отличная от dx то имеется две возможности для определения верхнего интеграла Римана: одна — с помощью функций из $\mathcal{C}(X)$ и другая — с помощью ступенчатых функций с компактным носителем. В § 1 у нас не было другого выбора, ибо мы знали лишь интегралы от ступенчатых функций и не были знакомы с интегралами от непрерывной функции с компактным носителем. Только после введения в § 1 понятия интеграла Римана dx стала мерой Радона на \mathbb{R} . Напротив, для произвольной меры Радона μ на \mathbb{R} мы можем выбирать одну из двух возможностей. Эти два определения, вообще говоря, не совпадают. Если, например, мы будем рассматривать меру μ на \mathbb{R} как сумму меры dx и меры δ и если в качестве функции f возьмем характеристическую функцию интервала $[0, 1]$, являющуюся ступенчатой функцией, то получим равенство $\int_{-\infty}^{+\infty} f d\mu = \int f d\mu = 1$. Если же взять произвольную функцию $\varphi \in \mathcal{C}(X)$, мажорирующую функцию f , то получим: $\varphi(0) \geq 1$ и $\int \varphi d\mu \geq 2$, а, следовательно, $\int_{-\infty}^{+\infty} f d\mu \geq 2$ или на самом деле $= 2$.

Однако совершенно ясно, что наиболее интересным определением в общем случае будет определение, использующее функции из $\mathcal{C}(X)$. Это объясняется не только тем, что лишь это определение позволяет иметь дело с произвольным локально компактным пространством X , а не только с прямой \mathbb{R} , сколько наличием ряда важных свойств этого интеграла, о которых речь будет идти позже.

Если \tilde{f} — ограниченная функция с компактным носителем со значениями в банаевом пространстве \tilde{F} , то говорят, что эта функция μ -интегрируема по Риману, если она обладает аппроксимирующей последовательностью *относительно верхнего интеграла Римана*, образованной из непрерывных разложимых функций с компактным носителем: $\tilde{f}_n \in \mathcal{C}(X) \otimes \tilde{F}$ (см. рассуждения перед следствием 8 теоремы 11), и $\int_{-\infty}^{+\infty} \|\tilde{f} - \tilde{f}_n\| d\mu$ сходится к 0 при n , стремящемся к бесконечности. Теперь можно определить интеграл Римана от функции \tilde{f} . Функции \tilde{f}_n образуют аппроксимирующую последовательность для \tilde{f} в смысле ин-

тегрирования по Лебегу. Поэтому \vec{f} интегрируема по Лебегу и ее интеграл Римана совпадает с ее интегралом Лебега.

Каждая непрерывная функция с компактным носителем интегрируема по Риману, поскольку (см. следствие 8 теоремы 11) она является равномерным пределом разложимых функций с носителем в фиксированном компакте. На вещественной прямой \mathbb{R} правильная функция не обязательно интегрируема по Риману в противоположность тому, что утверждалось в следствии 2 теоремы 8 для $\mu = dx$: при $\mu = dx + \delta$ характеристическая функция интервала $[0, 1]$ не может быть интегрируема по Риману, так как в противном случае ее верхний интеграл Римана, равный 2, совпал бы с интегралом Римана, или, что то же самое, с интегралом Лебега, равным 1.

Теоремы 1 и 6 остаются справедливыми вместе с сопровождающими их замечаниями, а следствия (кроме того, что относятся к интегралам от характеристических функций интервалов) справедливы при замене ступенчатых функций непрерывными разложимыми функциями. Заметим, что так же, как и в § 1, в смысле Римана интегрируются только ограниченные функции с компактным носителем.

Теорема 62. Пусть μ_n — последовательность вещественных мер Радона ≥ 0 на X , широко сходящаяся к пределу μ при n , стремящемся к бесконечности. Пусть \vec{f} — ограниченная с компактным носителем функция, определенная на X , со значениями в \vec{F} , измеримая по каждой из мер μ_n и интегрируемая по Риману относительно меры μ . Тогда $\int \vec{f} d\mu_n$ сходится к $\int \vec{f} d\mu$.

Доказательство. Прежде всего, \vec{f} интегрируема по Лебегу по всем μ_n и μ . Согласно замечанию 2°) на стр. 424 (где ступенчатые функции заменены непрерывными разложимыми функциями), можно найти непрерывную с компактным носителем разложимую функцию \vec{g} со значениями в \vec{F} и вещественную ≥ 0 непрерывную функцию h с компактным носителем, такие, что $\|\vec{f} - \vec{g}\| \leq h$ и $\mu(h) \leq \varepsilon/8$. Поскольку μ_n широко сходятся к μ , можно найти такое целое число p , что для $n \geq p$ имеет место неравенство $|\mu_n(h) - \mu(h)| \leq \varepsilon/8$, а, следовательно, $\mu_n(h) \leq \varepsilon/4$.

С другой стороны, из равенства $\int \vec{g} d\mu = \sum_{i=1}^N \int \vec{g}_i \varphi_i d\mu = \sum_{i=1}^N g_i \nu(\varphi)$

следует, что сходимость μ_n к μ при любой непрерывной разложимой функции \vec{g} с компактным носителем влечет за собой

сходимость $\int \vec{g} d\mu_n$ к $\int \vec{g} d\mu$, и потому существует такое целое q , что для всех $n \geq q$ имеет место неравенство

$$\left\| \int \vec{g} d\mu_n - \int \vec{g} d\mu \right\| \leq \frac{\epsilon}{2}.$$

Отсюда для $n \geq \max(p, q)$ непосредственно следует неравенство

$$\begin{aligned} \left\| \int \vec{f} d\mu_n - \int \vec{f} d\mu \right\| &\leq \int \|\vec{f} - \vec{g}\| d\mu_n + \left\| \int \vec{g} d\mu_n - \int \vec{g} d\mu \right\| + \\ &+ \int \|\vec{g} - \vec{f}\| d\mu \leq \mu_n(h) + \frac{\epsilon}{2} + \mu(h) \leq \epsilon, \end{aligned}$$

из которого следует утверждение теоремы.

Замечание. Эта теорема очень интересна: относительно функции \vec{f} требуется всего лишь ее μ -интегрируемость по Риману по предельной мере μ и не делается никаких аналогичных предположений относительно интегрируемости по μ_n (конечно, предполагается, что \vec{f} μ_n -измерима, а, следовательно, μ_n -интегрируема по Лебегу).

Теорема 63. Для того чтобы ограниченная с компактным носителем функция \vec{f} , определенная на X , со значениями в \vec{F} была μ -интегрируемой по Риману, необходимо и достаточно, чтобы множество ее точек разрыва имело μ -меру Лебега, равную нулю.

Доказательство. 1°) Предположим, что \vec{f} μ -интегрируема по Риману, и пусть \vec{g} и h — функции, определенные так же, как в теореме 62: $\|\vec{f} - \vec{g}\| \leq h$, $\mu(h) \leq \epsilon$. Назовем колебанием функции \vec{f} на окрестности \mathcal{V} точки a величину $\omega(\vec{f}; \mathcal{V}) = \sup_{x, y \in \mathcal{V}} \|\vec{f}(x) - \vec{f}(y)\|$. Колебанием функции \vec{f} в точке a назовем число $\omega(\vec{f}; a) = \inf_{\mathcal{V}} \omega(\vec{f}; \mathcal{V})$, где \inf берется по всевозможным окрестностям \mathcal{V} точки a ¹⁾. Функция \vec{f} непрерывна в точке a тогда и только тогда, когда ее колебание в этой точке равно нулю. Так как функция \vec{g} непрерывна, то имеет место неравенство $\omega(\vec{f}; a) \leq \omega(\vec{g}; a) + \omega(\vec{f} - \vec{g}; a) = \omega(\vec{f} - \vec{g}; a)$. В каждой

¹⁾ Это определение совпадает с частным определением колебания, данным на стр. 196 для правильных функций.

окрестности \mathcal{V} точки a имеет место неравенство

$$\|(\vec{f}(x) - \vec{g}(x)) - (\vec{f}(y) - \vec{g}(y))\| \leq h(x) + h(y).$$

Поскольку функция h непрерывна, то мы получаем отсюда, что $\omega(\vec{f} - \vec{g}; a) \leq 2h(a)$.

Обозначим теперь через A_α множество $\{x \in X; \omega(\vec{f}, x) \geq \alpha\}$, $\alpha > 0$. Это множество содержит в множестве $\{x \in X; h(x) \geq \alpha/2\}$. Так как $\mu(h) \leq \varepsilon$, то его внешняя μ -мера мажорируется числом $2\varepsilon/\alpha$. Поскольку это справедливо для любого $\varepsilon > 0$, то μ -мера равна нулю. Так как этот результат имеет место при любом α , то множество тех точек, в которых колебание функции \vec{f} больше 0, т. е. множество точек разрыва функции \vec{f} , имеет нулевую μ -меру.

2°) Обратно, предположим, что множество точек разрыва функции \vec{f} имеет μ -меру, равную нулю. Покажем сначала, что отсюда будет следовать μ -измеримость функции \vec{f} (и, следовательно, поскольку она ограничена и имеет компактный носитель, μ -интегрируемость в смысле Лебега). Действительно, она очевидным образом удовлетворяет критерию Лузина (следствие из теоремы 33): если K — произвольный компакт и если множество A точек разрыва функции \vec{f} имеет μ -меру, равную нулю, то можно найти такое открытое множество \mathcal{O} , содержащее A , μ -меры $\leq \delta$, что для $K_\delta = K \cap \mathcal{O}$ имеет место неравенство $\mu(K - K_\delta) \leq \delta$. В каждой точке K_δ функция \vec{f} непрерывна на X и, следовательно, тем более непрерывно ее сужение на K_δ .

Поскольку функция \vec{f} μ -интегрируема по Лебегу, то, согласно теореме 32, существует такая разложимая непрерывная функция \vec{g} с компактным носителем, что $\|\vec{f} - \vec{g}\| = h$, где $\int h d\mu \leq \varepsilon/2$. Обозначим через h^* «верхнюю функцию», присоединенную к h , т. е. функцию, определенную следующим образом: для каждой окрестности \mathcal{V} точки a в X полагают $h^*(\mathcal{V}) = \sup_{x \in \mathcal{V}} h(x)$, а затем полагают $h^*(a) = \inf_{\mathcal{V}} h^*(\mathcal{V})$. Легко видеть, что h^* полу-непрерывна сверху. В самом деле, каждая точка a имеет такую открытую окрестность \mathcal{V} , что $h^*(\mathcal{V}) \leq h^*(a) + \varepsilon$, и поэтому для каждой точки $x \in \mathcal{V}$ имеют место неравенства $h^*(x) \leq h^*(\mathcal{V}) \leq h^*(a) + \varepsilon$. Впрочем, h^* совпадает с h в каждой точке, где h непрерывна, т. е. μ -почти всюду. Поэтому $\int h d\mu = \int h^* d\mu$.

Пусть k — произвольная непрерывная интегрируемая функция, мажорирующая функцию h^* (например, $k \in \mathcal{C}(X)$). Тогда $k - h^*$ полунепрерывна снизу, ≥ 0 и, значит, борелевская. Она

мажорируется функцией k и поэтому интегрируема. Согласно теореме 39₂, $\int (k - h^*) d\mu$ является точной верхней гранью для мер $\mu(\psi)$, $\psi \in \mathcal{C}(X)$, $0 \leq \psi \leq k - h^*$, а, значит, $\int h^* d\mu$ является точной нижней гранью мер $\mu(k - \psi) = \mu(\varphi)$ для всех непрерывных интегрируемых функций $\varphi \geq 0$, мажорирующих h^* (в пространстве функций $\varphi \in \mathcal{C}(X)$, если функция k взята также в $\mathcal{C}(X)$). Таким образом, можно найти такую функцию $\varphi \in \mathcal{C}(X)$, $\varphi \geq h^* \geq h$, что $\mu(\varphi) \leq \int h^* d\mu + \varepsilon/2 = \int h d\mu + \varepsilon/2 \leq \varepsilon$. Но тогда $\|\vec{f} - \vec{g}\| \leq \varphi$, где $\mu(\varphi) \leq \varepsilon$. Поскольку это справедливо для любого $\varepsilon > 0$, то функция \vec{f} μ -интегрируема по Риману.

Следствие 1. Для того чтобы часть A множества X с компактным замыканием имела μ -интегрируемую по Риману характеристическую функцию, необходимо и достаточно, чтобы ее граница имела μ -меру (Лебега), равную нулю.

Доказательство. Так как в окрестности каждой внутренней точки части A характеристическая функция постоянна и равна 1, а в окрестности каждой внешней точки этой части она постоянна и равна 0, то множество точек разрыва характеристической функции множества A является совокупностью граничных точек этого множества.

Следствие 2. Если меры $\mu_n \geq 0$ на X и широко сходятся к μ при n , стремящемся к бесконечности, и если A есть часть X с компактным замыканием, измеримая для всех μ_n , граница которой \hat{A} имеет нулевую меру для предельной меры μ , то $\mu_n(A)$ сходятся к $\mu(A)$.

Следствие 3. Если меры Радона μ_n на \mathbb{R} при n , стремящемся к бесконечности, широко сходятся к мере dx и если ограниченная с компактным носителем функция φ измерима при всех μ_n и интегрируема по Риману относительно dx , то $\int \varphi d\mu_n$ сходится к $\int \varphi dx$.

На прямой \mathbb{R} имеет место теорема, обратная к предыдущим теоремам, которая представляет и самостоятельный интерес:

Теорема 64. Пусть μ_n — меры Радона ≥ 0 на прямой \mathbb{R} . Если для каждого ограниченного интервала $[a, b]$, такого, что $\mu(\{a\}) = \mu(\{b\}) = 0$, при n , стремящемся к бесконечности, $\mu_n([a, b])$ сходятся к $\mu([a, b])$, то меры μ_n широко сходятся к мере μ .

Читателю предлагается доказать эту теорему в качестве упражнения.

Широкая сходимость и равномерная сходимость

Рассмотрим теперь задачи другого характера. Предположим, что меры μ_n широко сходятся к некоторому пределу μ . Выясним, можно ли для функций φ найти такие множества \mathcal{A} функций φ , чтобы $\mu_n(\varphi)$ сходились к $\mu(\varphi)$ равномерно на \mathcal{A} .

Как мы видели ранее, такая равномерная сходимость не может иметь места, если \mathcal{A} является единичным шаром из $C_k(X)$. Значит, нужны множества меньшие, чем единичный шар.

Здесь мы воспользуемся следующей общей теоремой функционального анализа:

Теорема 65 (Банаха — Штейнгауза). Пусть u_n — линейные непрерывные отображения банахова пространства \tilde{E} в векторное нормированное пространство \tilde{F} . Если функции u_n , определенные на \tilde{E} , со значениями в \tilde{F} при n , стремящемся к бесконечности, просто сходятся к некоторой предельной функции u , то нормы $\|u_n\|$ ограничены некоторым фиксированным числом M , функция u линейна и непрерывна, а последовательность u_n сходится к u равномерно на каждой компактной части \mathcal{A} пространства \tilde{E} .

Доказательство. Здесь существенным является тот факт, что \tilde{E} банахово, т. е. полно. Без этого предположения все результаты становятся неверными. Рассматриваемая теорема — одна из редких теорем, в которых исходное пространство надо считать полным. Теорема Банаха — Штейнгауза является одной из наиболее глубоких и плодотворных теорем современного анализа.

Докажем сначала, что нормы функций u_n ограничены одним и тем же числом $M \geqslant 0$, но для этого нам понадобится следующая лемма.

Лемма (теорема Бэра)¹⁾. Пусть E — полное метрическое пространство, и пусть $\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2, \dots$ — последовательность открытых множеств, плотных в E . Тогда их пересечение $A = \bigcap_{n \geqslant 0} \mathcal{O}_n$

(которое теперь не должно быть открытым) также плотно в E .

¹⁾ Мы ее назвали леммой, хотя это одна из наиболее важных теорем анализа, значение которой далеко выходит за те границы, которыми мы ограничили здесь ее применение.

Доказательство леммы. Пусть Ω — произвольное открытое множество из E . Нам надо доказать, что $\Omega \cap A$ не пусто. Поскольку множество \mathcal{O}_0 плотно, то существует по крайней мере одна точка пересечения $\Omega \cap \mathcal{O}_0$. Поскольку последнее множество открыто, существует даже некоторый замкнутый шар $B_0 = B(a_0, \delta_0)$, полностью содержащийся в $\Omega \cap \mathcal{O}_0$. Повторим теперь эту операцию, заменив \mathcal{O}_0 на \mathcal{O}_1 и Ω на открытый шар \dot{B}_0 . Тогда найдется некоторый замкнутый шар $B_1 = B(a_1, \delta_1)$, полностью лежащий в $\dot{B}_0 \cap \mathcal{O}_1$. При этом мы выберем $\delta_1 \leq 1$. Продолжая так далее, мы сможем построить последовательность точек a_n и последовательность чисел $\delta \leq 1/n$, такие, чтобы каждый замкнутый шар $B_n = B(a_n, \delta_n)$ содержался в открытом шаре $\dot{B}_{n-1} = \dot{B}_{n-1}(a_{n-1}, \delta_{n-1})$ и в \mathcal{O}_n . Все эти замкнутые шары лежат в Ω .

Последовательность центров шаров a_n является последовательностью Коши. В самом деле, все точки a_{n+p} при $p \geq 0$ содержатся в замкнутом шаре B_n и поэтому $d(a_{n+p}, a_n) \leq 1/n$. Поскольку пространство E предполагалось полным, последовательность точек a_n сходится к некоторой предельной точке a . Так как все точки a_{n+p} при $p \geq 0$ содержатся в шаре B_n , то точка a заведомо лежит в B_n , а, следовательно, и в $\Omega \cap \mathcal{O}_n$. Поскольку это верно для любого n , то $a \in \Omega \cap A$, чем и заканчивается доказательство леммы.

Часто оказывается удобным пользоваться этой леммой, применяя переход к дополнениям. Дополнение к открытому множеству замкнуто, а дополнение к плотному множеству, внешность которого пуста, является множеством с пустой внутренностью. Поэтому имеет место такое утверждение:

Пусть E — полное метрическое пространство и F_n — последовательность замкнутых множеств с пустой внутренностью. Тогда объединение $F = \bigcup_{n \geq 0} F_n$ (которое не обязано быть замкнутым) имеет пустую внутренность (в частности, $F \neq E$).

Совершенно ясно, что неполное метрическое пространство этим свойством не обладает. Например, поле \mathbb{Q} рациональных чисел с обычной метрикой является объединением счетного числа множеств, сводящихся к одной точке; каждая точка замкнута в \mathbb{Q} и имеет пустую внутренность.

Докажем теперь первую часть теоремы, а именно ограниченность норм $\|u_n\|$. Обозначим через F_k множество точек $x \in E$, в которых все $\overrightarrow{\|u_n(x)\|}$ не превосходят k . Так как функция u_n непрерывна, то множество $F_{n,k} = \{x \in E; \overrightarrow{\|u_n(x)\|} \leq k\}$ за-

мкнуто, а, следовательно, $F_k = \bigcap_{n=0}^k F_{n,k}$ также замкнуто. Объединение множеств F_k при $k = 0, 1, 2, \dots$ дает все пространство E . В самом деле, для каждой точки $x \in E$ по предположению функции $\overrightarrow{u_n(x)}$ имеют предел, а, значит, их значения ограничены, т. е. существует такое целое число k (зависящее от x), что все $\|\overrightarrow{u_n(x)}\| \leq k$ и, значит, $x \in F_k$. Поскольку E — полное метрическое пространство (банахово), то согласно лемме Бэра внутренность всех F_k не может быть пустой. Значит, существует такое k , при котором F_k имеет непустую внутренность, а, следовательно, имеется некоторый шар $B(a, \rho)$, содержащийся в F_k . Теперь для $\|x\| \leq \rho$ и для любого n имеем $\|\overrightarrow{u_n(x)}\| \leq \|\overrightarrow{u_n(x+a)}\| + \|\overrightarrow{u_n(a)}\| \leq 2k$, а, следовательно, $\|\overrightarrow{u_n}\| \leq 2k/\rho$, чем и завершается доказательство нашего утверждения. (Заметим, что в этом результате вовсе не требуется, чтобы функции u_n сходились. Надо лишь, чтобы для каждого $x \in E$ нормы $\|\overrightarrow{u_n(x)}\|$ были ограниченными некоторой константой, которая, вообще говоря, произвольно зависит от x . Однако из предыдущего доказательства следует, что $\|\overrightarrow{u_n}\|$ ограничены, т. е. существует такое число M , что эта константа для $\|\overrightarrow{u_n(x)}\|$ будет $\leq M \|x\|$.)

Очевидно, функция u линейна. В самом деле, так как $u_n(\vec{x} + \vec{y}) = u_n(\vec{x}) + u_n(\vec{y})$, то переходом к пределу получаем, что $u(\vec{x} + \vec{y}) = u(\vec{x}) + u(\vec{y})$. Точно так же проверяется, что $u(k\vec{x}) = ku(\vec{x})$. Из неравенства $\|\overrightarrow{u_n(\vec{x})}\| \leq M \|\vec{x}\|$ переходом к пределу можно получить неравенство $\|\overrightarrow{u(\vec{x})}\| \leq M \|\vec{x}\|$, из которого следует, что функция u непрерывна и ее норма $\leq M$.

Пусть теперь \mathcal{A} — некоторый компакт пространства \vec{E} . Зададим $\varepsilon > 0$. Тогда найдется конечное число открытых шаров с центрами в множестве \mathcal{A} радиуса $\varepsilon/3M$, покрывающих \mathcal{A} (в силу определения компактного множества). Пусть $(\vec{a}_i)_{i=1}^n$ — центры этих шаров. Согласно предположению о простой сходимости, для каждого из этих центров \vec{a}_i можно найти такое целое число p_i , что при $n \geq p_i$ имеет место неравенство

$$\|\overrightarrow{u_n(\vec{a}_i)} - \overrightarrow{u(\vec{a}_i)}\| \leq \frac{\varepsilon}{3}. \quad (\text{IV}, 7; 16)$$

Если теперь через p обозначить наибольшее из чисел p_i , то можно убедиться, что для $n \geq p$ и \vec{X} , находящихся от \vec{a}_i на

расстоянии $< \varepsilon/3M$, имеет место неравенство

$$\begin{aligned} \|u_n(\vec{x}) - u(\vec{x})\| &\leq \|u_n(\vec{x}) - u_n(\vec{a}_i)\| + \|u_n(\vec{a}_i) - u(\vec{a}_i)\| + \\ &+ \|u(\vec{a}_i) - u(\vec{x})\| \leq M \frac{\varepsilon}{3M} + \frac{\varepsilon}{3} + M \frac{\varepsilon}{3M} = \varepsilon. \quad (\text{IV}, 7; 17) \end{aligned}$$

Однако так как каждая точка \mathcal{A} находится на расстоянии $< \varepsilon/3M$ от одной из точек \vec{a}_i , то окончательно для любой точки $\vec{x} \in \mathcal{A}$ и $n \geq p$ получаем неравенство $\|u_n(\vec{x}) - u(\vec{x})\| \leq \varepsilon$. Значит, u_n сходится к u равномерно на множестве \mathcal{A} , и этим заканчивается доказательство теоремы.

Следствие 1. Пусть меры μ_n на X со значениями в векторном нормированном пространстве \vec{E} при n , стремящемся к бесконечности, широко сходятся к некоторой мере μ . Тогда для каждого компакта K нормы $\|\mu_n\|_K$ равномерно ограничены. Если \mathcal{A} — некоторый компакт пространства $\mathcal{C}_K(X)$, то $\mu_n(\vec{\varphi})$ сходятся к $\mu(\vec{\varphi})$ равномерно по $\vec{\varphi}$ в \mathcal{A}^1 .

Так как $\mathcal{C}_K(X)$ — банахово пространство (см. стр. 447), то это следствие непосредственно вытекает из сформулированной теоремы.

Следствие 2. Пусть μ_n — меры, широко сходящиеся к мере μ , и φ_n — функции из $\mathcal{C}(X)$, равномерно сходящиеся к некоторой функции φ и сохраняющие свои носители в одном и том же компакте K . Тогда $\mu_n(\varphi_n)$ сходятся к $\mu(\varphi)$.

Доказательство. По следствию 1 нормы $\|\mu_n\|_K$ ограничены одним и тем же числом $M \geq 0$. С другой стороны, φ_n сходятся в $\mathcal{C}_K(X)$ к φ . Далее,

$$\mu_n(\varphi_n) - \mu(\varphi) = \mu_n(\varphi_n - \varphi) + (\mu_n(\varphi) - \mu(\varphi)). \quad (\text{IV}, 7; 17_2)$$

В силу условия широкой сходимости второй член правой части сходится к 0 . Так как норма первого члена мажорируется величиной $M \|\varphi_n - \varphi\|_K$, то он также стремится к нулю, чем и заканчивается доказательство следствия.

1) Вообще говоря, $\mathcal{C}_K(X)$ — бесконечномерное пространство. Его единичный шар не компактен, и поэтому в качестве \mathcal{A} нельзя брать единичный шар, как мы это уже отмечали ранее.

Компактные подмножества пространства $\mathcal{C}_k(X)$

Для того чтобы иметь возможность применить следствие 1, нам надо уметь определять компактные множества в пространствах $\mathcal{C}_k(X)$. В дальнейшем мы займемся общим изучением компактов, а сейчас ограничимся частными случаями этих множеств.

Теорема 66. Пусть A, B, C — три топологических пространства, A — компактное, а C — метрическое. Пусть f — непрерывное отображение $A \times B$ в C . Тогда для каждого $y \in B$ можно определить частное отображение $f_y: x \rightarrow f(x, y)$ множества A в C , $f_y \in (C^A)_{cb}$.

Если y сходится к b в пространстве B , то частная функция f_y сходится к частной функции f_b равномерно на A . Кроме того, отображение $y \rightarrow f_y$, которое каждому элементу $y \in B$ ставит в соответствие частную функцию f_y , является непрерывным отображением B в $(C^A)_{cb}$.

Доказательство. Зафиксируем точку b в B и зададим произвольное $\varepsilon > 0$. Какова бы ни была точка a , функция f непрерывна в точке (a, b) . Следовательно, существует такая открытая окрестность \mathcal{U}_a точки a и такая окрестность \mathcal{V}_a точки b в множествах A и B соответственно, что для $x \in \mathcal{U}_a$ и $y \in \mathcal{V}_a$ имеет место неравенство

$$d(f(x, y), f(a, b)) \leq \frac{\varepsilon}{2}. \quad (\text{IV}, 7; 18)$$

Тем более, выполняется неравенство

$$d(f(x, b), f(a, b)) \leq \frac{\varepsilon}{2}. \quad (\text{IV}, 7; 19)$$

Отсюда можно вывести неравенство, имеющее место для $x \in \mathcal{U}_a$, $y \in \mathcal{V}_a$:

$$d(f(x, y), f(x, b)) \leq d(f(x, y), f(a, b)) + d(f(a, b), f(x, b)) \leq \varepsilon. \quad (\text{IV}, 7; 20)$$

Когда точка a меняется, окрестности \mathcal{U}_a образуют открытое покрытие компакта A . Следовательно, можно выбрать конечное покрытие $(\mathcal{U}_{a_i})_{i \in I}$ множества A . Этим окрестностям \mathcal{U}_{a_i} соответствуют окрестности \mathcal{V}_{a_i} . Обозначим через \mathcal{V} пересечение всех \mathcal{V}_{a_i} . Множество \mathcal{V} является некоторой окрестностью точки b в B . Тогда для x из \mathcal{U}_{a_i} и y из \mathcal{V} имеет место неравенство (IV, 7; 20). Поскольку окрестности \mathcal{U}_{a_i} образуют покрытие A , предыдущее неравенство имеет место для каждого $x \in A$ и каждого $y \in \mathcal{V}$, что можно записать так: $d(f_y, f_b) \leq \varepsilon$ для $y \in \mathcal{V}$, а это соотношение доказывает теорему.

Замечание. Если оба пространства A и B метрические и компактные, то доказательство становится значительно проще. В самом деле, в этом случае, согласно теореме 31 гл. II, функция f равномерно непрерывна. Значит, для любого $\varepsilon > 0$ можно найти такое $\eta > 0$, что из неравенств $d(x', x'') \leq \eta$, $d(y', y'') \leq \eta$ следует $d(f(x', y'), f(x'', y'')) \leq \varepsilon$. В частности, из $d(y, b) \leq \eta$ при любом x следует неравенство $d(f(x, y), f(x, b)) \leq \varepsilon$ и, следовательно, $d(f_y, f_b) \leq \varepsilon$, что и доказывает теорему.

Следствие 1. Пусть A и B — компактные пространства, C — метрическое пространство и f — непрерывное отображение $A \times B$ в C . Тогда множество частных отображений f_y из A в C , когда y пробегает пространство B , является компактным подмножеством пространства $(C^A)_{cb}$.

В самом деле, мы имеем здесь образ компакта B при непрерывном отображении $y \rightarrow f_y$.

Следствие 2. Пусть X — локально компактное пространство, T — компактное пространство. Пусть φ — функция с вещественными или комплексными значениями, непрерывная на $X \times T$. Тогда, каким бы ни был элемент $t \in T$, существует частное отображение φ_t , являющееся вещественной или комплексной функцией, непрерывной на X . Предположим, что, когда t пробегает T , функция φ_t имеет носитель в одном и том же компакте K пространства X . Тогда, если t пробегает T , множество функций φ_t образует компактную часть пространства $\mathcal{C}_K(X)$.

Для доказательства достаточно применить следствие 1 к $f = \varphi$, $A = K$, $B = T$, $C = \mathbb{R}$ или \mathbb{C} .

Следствие 3. Предположим, что выполнены условия следствия 2. Пусть μ_n — меры Радона на X со значениями в векторном нормированном пространстве \vec{E} , широко сходящиеся к предельной мере μ при n , стремящемся к бесконечности. Тогда интегралы $\int \varphi(x, t) d\mu_n(x)$ при n , стремящемся к бесконечности, сходятся к $\int \varphi(x, t) d\mu(x)$ равномерно по t в пространстве T .

Широкая сходимость последовательности мер к мере Дирака

Теорема 67. Пусть μ_n — последовательность мер ≥ 0 на локально компактном пространстве X . Предположим, что в каждой компактной окрестности U точки $a \in X$ последовательность интегралов $\int_U d\mu_n$ сходится к 1 при n , стремящемся

к бесконечности. Тогда меры μ_n широко сходятся к мере Дирака $\delta_{(a)}$.

Перед началом доказательства заметим, что в нашем предположении подразумевается, что, каким бы ни был компакт K пространства X , не содержащий a , $\int_K d\mu_n$ сходится к 0 при n , стремящемся к бесконечности.

В самом деле, дополнение к K является некоторым открытым множеством X , т. е. некоторым локально компактным пространством¹⁾. Пусть теперь \mathcal{V} — компактная окрестность точки a в этом пространстве, т. е. компактная окрестность точки a в X , не имеющая общих точек с K .

Тогда, с одной стороны, $\int_{\mathcal{V}} d\mu_n$ сходится к 1, а с другой — $\int_{X \setminus K} d\mu_n$ сходится к 1 и, значит, $\int_K d\mu_n$ сходится к 0.

Условия теоремы означают, что общая масса на конечном расстоянии стремится к 1 и стремится сконцентрироваться в точке a .

Доказательство. Пусть φ — некоторая функция из $C(X)$ с компактным носителем K . В силу непрерывности φ , при заданном $\varepsilon > 0$ существует такое открытое множество, содержащее a , что для любой точки x из этого множества имеет место неравенство $|\varphi(x) - \varphi(a)| \leq \varepsilon/4$. Если \mathcal{V} — компактная окрестность точки a в этом открытом множестве, то это неравенство тем более сохраняется на \mathcal{V} . Разложим теперь разность $\int \varphi d\mu_n - \varphi(a)$ в сумму трех членов, а именно:

$$\begin{aligned} \int \varphi d\mu_n - \varphi(a) = & \int_{C\mathcal{V}} \varphi d\mu_n + \int_{\mathcal{V}} (\varphi(x) - \varphi(a)) d\mu_n(x) + \\ & + \varphi(a) \left[\int_{\mathcal{V}} d\mu_n - 1 \right]. \quad (IV, 7; 21) \end{aligned}$$

Последовательно оценим сверху модули этих трех членов.

1°) Первое слагаемое правой части стремится к 0 при n , стремящемся к бесконечности. В самом деле, поскольку φ имеет носитель в K , интеграл на $C\mathcal{V}$ в действительности является интегралом на пересечении $K \cap C\mathcal{V}$. Поэтому он мажорируется числом $\|\varphi\| \int_A d\mu_n$, где A — замыкание $K \cap C\mathcal{V}$. Так как

это множество замкнуто и содержится в компакте K , то множество A компактно. С другой стороны, поскольку множество

¹⁾ Смотри примечания на стр. 453 и на стр. 475.

A содержится в \mathcal{U} и a не является точкой прикосновения множества A , то множество A не содержит точки a .

Согласно замечанию, предшествовавшему доказательству, $\int_A d\mu_n$ сходится к 0.

2°) Рассмотрим теперь третье слагаемое. По условию, величина, заключенная в квадратные скобки, стремится к 0, и, следовательно, третье слагаемое также сходится к 0 при n , стремящемся к бесконечности. Значит, существует такое целое число p' , что при $n \geq p'$ сумма абсолютных величин первого и третьего слагаемых не превосходит $\varepsilon/2$.

3°) Рассмотрим, наконец, второе слагаемое. Согласно выбору \mathcal{U} , оно мажорируется величиной $\frac{\varepsilon}{4} \int_{\mathcal{U}} d\mu_n$. Последний интеграл стремится к 1 при n , стремящемся к бесконечности. Следовательно, существует такое целое число p'' , что этот интеграл для $n \geq p''$ не больше 2. Но тогда для $n \geq p''$ второе слагаемое не превосходит по модулю $\varepsilon/2$. Если теперь через p обозначить наибольшее из целых чисел p' и p'' , то мы увидим, что для $n \geq p$ будет иметь место неравенство $\left| \int \varphi d\mu_n - \varphi(a) \right| \leq \varepsilon$, и теорема доказана.

Пример 1. Пусть μ_n — последовательность мер ≥ 0 , полная масса которых $\int d\mu_n$ сходится к 1 и носитель которых равномерно сходится к точке a . Под этим мы понимаем следующее: какова бы ни была окрестность \mathcal{U} точки a , носитель меры μ_n для достаточно больших n содержится в \mathcal{U} . Тогда меры μ_n широко сходятся к $\delta_{(a)}$ при n , стремящемся к бесконечности. Это очевидно, так как условия теоремы выполняются.

Пример 1₂. Как частный случай рассмотрим последовательность функций $p_n \geq 0$ на вещественной прямой \mathbb{R} , интегрируемых относительно меры dx и таких, что $\int p_n(x) dx$ сходится к 1 при n , стремящемся к бесконечности, причем носитель функции p_n равномерно сходится к точке a . Тогда меры $p_n dx$ широко сходятся к мере $\delta_{(a)}$.

Этот пример дает новую интерпретацию некоторых случаев, для которых не применима теорема о сходимости Лебега.

Рассмотрим, например, следующую последовательность функций p_n на \mathbb{R} :

$$p_n(x) = \begin{cases} n & \text{для } 0 < x \leq 1/n, \\ 0 & \text{в остальных точках.} \end{cases} \quad (\text{IV}, 7; 22)$$

Эти функции всюду сходятся к 0 при n , стремящемся к бесконечности. Однако они не мажорируются никакой функцией ≥ 0 , локально интегрируемой по мере dx (см. формулу (IV, 4; 11)). Следовательно, теорема Лебега здесь не применима, и лучшим доказательством этого факта служит то, что все интегралы $\int p_n dx$ равны 1 и, следовательно, не стремятся к 0.

Если теперь φ принадлежит $\mathcal{C}(\mathbb{R})$, то $\int \varphi p_n dx$ не сходится к 0. Мы отмечали этот результат в связи с теоремой Лебега. Теперь дадим ему другое истолкование: из сказанного в примере 1₂ вытекает, что меры $p_n dx$ широко сходятся не к мере 0, а к мере δ .

Интеграл $\int p_n \varphi dx$ сходится к $\varphi(0)$. Такое явление встречается довольно часто. Когда функции p_n просто сходятся почти всюду к предельной функции p , но таким образом, что нет сходимости интегралов, то это часто означает¹⁾, что меры $p_n dx$ сходятся к сумме меры $p dx$ и меры, представляющей некоторые точечные массы.

Пример 2. Пусть f — некоторая функция ≥ 0 в пространстве \mathbb{R}^n , интегрируемая по мере $dx = dx_1 dx_2 \dots dx_n$ с $\int f dx = 1$. Если для вещественного $k \neq 0$ положить

$$\mu_k = |k|^n f(kx) dx, \quad (\text{IV}, 7; 23)$$

то μ_k будет образом fdx при гомотетии $x \rightarrow x/k$ (см. формулу (IV, 6; 28) для $n = 1$, а затем теорему 102 для произвольного n). Меры μ_k широко сходятся к мере Дирака δ при k , стремящемся к бесконечности. Это можно увидеть, применив теорему 67, или непосредственно, используя замену переменных. Для $\varphi \in \mathcal{C}(\mathbb{R}^n)$ имеем:

$$\mu_k(\varphi) = \int \varphi(x) f(kx) |k|^n dx. \quad (\text{IV}, 7; 24)$$

Производя замену переменных $kx = \xi$, получаем (см. далее теорему 102):

$$\mu_k(\varphi) = \int \varphi\left(\frac{\xi}{k}\right) f(\xi) d\xi. \quad (\text{IV}, 7; 25)$$

При k , стремящемся к бесконечности, ξ/k стремится к 0. Так как функция φ непрерывна, то при этом $\varphi(\xi/k)$ сходится просто к $\varphi(0)$, а, следовательно, функция $\xi \rightarrow \varphi(\xi/k)f(\xi)$ сходится просто к $\varphi(0)f(\xi)$. Поскольку она остается ограниченной

¹⁾ Конечно, не всегда! Если вместо n , как в (IV, 7; 22), мы возьмем n^α , как в (IV, 4; 11), при $\alpha > 1$, то ничего подобного не получится.

неотрицательной функцией $\|\varphi\| f(\xi)$, интегрируемой по $d\xi$, то из теоремы 35 Лебега следует, что $\mu_k(\varphi)$ при k , стремящемся к бесконечности, стремится к $\varphi(0) \int f(\xi) d\xi = \varphi(0)$, откуда следует требуемое заключение.

Имеется и более прямое доказательство. Мы видим, что $\int f dx$ есть мера ≥ 0 , имеющая полную массу, равную 1. Гомотетия с центром в 0 и отношением $1/k$ при k , стремящемся к бесконечности, стягивает все массы к началу координат. Пример 3°) на стр. 634 теперь дает самое короткое доказательство: гомотетия H_k с отношением $1/k$ при k , стремящемся к бесконечности, сходится к гомотетии H с нулевым отношением, т. е. к постоянному отображению $x \rightarrow 0$, а, следовательно, $H_k \mu$ сходится к $H\mu$. Поскольку масса μ равна +1, то, согласно примеру 1°) на стр. 628, $H\mu = \delta$.

Вот два часто используемых примера.

а) Функция f на \mathbb{R} определена формулой

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (\text{IV}, 7; 26)$$

Она полностью удовлетворяет требуемым условиям, поскольку, как мы увидим позже, интеграл от f равен 1.

Последовательность мер $\sqrt{n} f(\sqrt{n} x) dx$:

$$\mu_n = \frac{1}{\sqrt{2\pi}} e^{-\frac{n x^2}{2}} \sqrt{n} dx \quad (\text{IV}, 7; 27)$$

широко сходится к δ при n , стремящемся к бесконечности.

На рис. 18 изображены графики двух из этих колоколообразных кривых, представляющих плотности мер μ_n по отношению к мере Лебега, т. е. функции $\sqrt{n} f(\sqrt{n} x)$.

Можно также взять

$$f(x) = e^{-\pi x^2}, \quad \sqrt{n} f(\sqrt{n} x) = \sqrt{n} e^{-\pi n x^2}. \quad (\text{IV}, 7; 28)$$

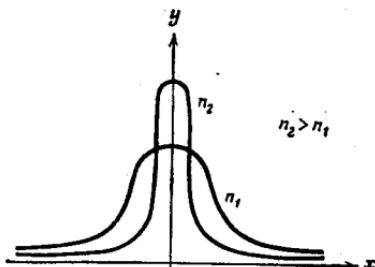


Рис. 18.

б) Функция f задана формулой

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}. \quad (\text{IV}, 7; 29)$$

Ее интеграл, очевидно, равен 1, и потому меры

$$\mu_n = \frac{n}{\pi} \frac{dx}{(1+n^2x^2)} \quad (\text{IV}, 7; 30)$$

при n , стремящемся к бесконечности, широко сходятся к мере δ .

Замечание. Эти свойства были бы невозможными без введения понятия широкой сходимости: если K — компактный интервал $[-A, A]$, то, согласно сказанному относительно формулы (IV, 2; 16),

$$\|\mu_k - \delta\|_K = \|f(kx)|k|^n dx - \delta\|_K \text{ равно } \int_K f(kx)|k|^n dx + 1$$

и стремится к 2 при k , стремящемся к бесконечности. Локальная сходимость по норме отсутствует!

Теорема 68. Существует такая последовательность полиномов $P_n \geq 0$ вещественной переменной, что меры P_ndx широко сходятся к δ при n , стремящемся к бесконечности.

Эта теорема имеет большое значение. Из нее в дальнейшем будет получена теорема 80₂ Вейерштрасса (о приближении непрерывных функций полиномами).

Доказательство. Пусть e_n — последовательность положительных чисел, стремящихся к $+\infty$ при n , стремящемся к бесконечности. Положим

$$P_n(x) = \frac{e_n}{\sqrt{2\pi}} \left(1 - \frac{e_n^2 x^2}{2n}\right)^n. \quad (\text{IV}, 7; 31)$$

Для четного n эта величина, очевидно, ≥ 0 .

Покажем, что P_ndx широко сходятся к δ при n , стремящемся к бесконечности, если только числа e_n/\sqrt{n} стремятся к нулю. Для этого можно было бы использовать теорему 67, но доказательство также просто можно провести и непосредственно. Пусть $\varphi \in \mathcal{C}(\mathbb{R})$. Тогда

$$(P_n dx)(\varphi) = \int \frac{e_n}{\sqrt{2\pi}} \left(1 - \frac{e_n^2 x^2}{2n}\right)^n \varphi(x) dx. \quad (\text{IV}, 7; 32)$$

Пусть $[-A, A]$ — интервал, содержащий носитель функции φ . Последний интеграл фактически вычисляется на $[-A, A]$.

Выполним замену переменных $e_n x = \xi$. Тогда

$$(P_n dx)(\varphi) = \frac{1}{\sqrt{2\pi}} \int \left(1 - \frac{\xi^2}{2n}\right)^n \varphi\left(\frac{\xi}{e_n}\right) d\xi, \quad (\text{IV}, 7; 33)$$

где интеграл в действительности вычисляется на $[-e_n A, +e_n A]$.

Для фиксированного ξ при n , стремящемся к бесконечности, функция $\varphi(\xi/e_n)$, в силу ее непрерывности в начале координат, стремится к $\varphi(0)$. Переменная $(1 - \xi^2/2n)^n$ стремится при этом к $e^{-\xi^2/2}$. Поэтому интегрируемая функция сходится просто к функции $\varphi(0)e^{-\xi^2/2}$. Если применима теорема 35 Лебега, то правая часть равенства (IV, 7; 33) сходится к $\frac{\varphi(0)}{\sqrt{2\pi}} \int e^{-\xi^2/2} d\xi = \varphi(0)$, и тогда теорема будет доказана.

Найдем оценку сверху для подинтегральной функции. Для $0 \leq u < 1$ из разложения логарифма в ряд следует, что $\ln(1-u) \leq -u$, откуда $1-u \leq e^{-u}$ и, значит, $(1-\xi^2/2n)^n \leq e^{-\xi^2/2}$ при $\xi^2/2n < 1$.

Подинтегральная функция равна $\varphi(\xi/e_n)(1 - \xi^2/2n)^n$, а, значит, для $|\xi| \leq e_n A$ мажорируется величиной $\|\varphi\|(1 - \xi^2/2n)^n$ и равна нулю для $|\xi| > e_n A$. Для $|\xi| \leq e_n A$ имеем: $\xi^2/2n \leq (A^2/2) \cdot (e_n^2/n)$. Поскольку e_n/\sqrt{n} стремится к 0 при n , стремящемся к бесконечности, существует такое целое число p , что $(A^2/2) \cdot (e_n^2/n) < 1$ при $n \geq p$.

Теперь для $n \geq p$ имеет место неравенство

$$\left| \left(1 - \frac{\xi^2}{2n}\right)^n \varphi\left(\frac{\xi}{e_n}\right) \right| \leq \|\varphi\| e^{-\xi^2/2}, \quad (\text{IV}, 7; 34)$$

правая часть которого интегрируема. Значит, применима теорема Лебега и потому $P_n dx$ действительно сходятся к δ .

Узкая сходимость последовательности мер конечной нормы

Пусть $\overset{\rightarrow}{\mu}$ — мера Радона на X со значениями в векторном нормированном конечномерном пространстве \vec{E} , имеющая *кoneчную норму*.

Согласно следствию из теоремы 54₂ в этом случае для каждой скалярной непрерывной и ограниченной функции φ на X можно определить $\int \varphi d\overset{\rightarrow}{\mu}$. Тем самым мы пришли к следующему определению, играющему фундаментальную роль в теории вероятностей:

Определение. Говорят, что последовательность мер $\overset{\rightarrow}{\mu}_n$ на X со значениями в конечномерном векторном нормирован-

ном пространстве \vec{E} , имеющих конечную норму¹⁾, узко сходится к мере μ с конечной нормой, если, какова бы ни была скалярная непрерывная и ограниченная функция φ на X при n , стремящемся к бесконечности, интеграл $\int \varphi d\mu_n$ сходится к $\int \varphi d\mu$.

Естественно, что узкая сходимость влечет за собой широкую сходимость. Однако узкая сходимость слабее сходимости по норме, которая означает, что $\|\mu_n - \mu\|$ сходится к нулю и, следовательно, $\int \varphi d\mu_n$ сходится к $\int \varphi d\mu$ равномерно, если функция φ остается ограниченной по норме на X .

Если меры μ_n узко сходятся к мере μ , то, полагая $\varphi \equiv 1$, получаем, что $\int d\mu_n$ сходится к $\int d\mu$. В частности, если меры μ_n и μ неотрицательны, то $\|\mu_n\|$ сходится к $\|\mu\|$.

Теорема 69. Если меры μ_n на X со значениями в конечномерном векторном нормированном пространстве \vec{E} сходятся узко к мере μ при n , стремящемся к бесконечности, то нормы $\|\mu_n\|$ ограничены в совокупности.

Доказательство. Каждая мера μ_n определяет некоторое линейное непрерывное отображение u_n пространства $(K^X)_{cb}$ скалярных непрерывных ограниченных функций на X в пространство \vec{E} , а именно: $\varphi \rightarrow \int \varphi d\mu_n$.

Из определения $\|\mu_n\|$, как точной верхней грани норм $\|\mu_n(\varphi)\|$ при $\varphi \in \mathcal{C}(X)$ и $\|\varphi\| \leq 1$, следует, что $\|u_n\| \geq \|\mu_n\|^2$. Так как $(K^X)_{cb}$ является банаховым пространством (следствие 3 теоремы 65 гл. II), то из теоремы Банаха — Штейнгауза следует, что u_n ограничены и тем более ограничены $\|\mu_n\|$. Теорема доказана.

Пусть μ — некоторая неотрицательная мера с конечной нормой. Будем называть *верхним интегралом Римана по мере μ* от ограниченной функции $f \geq 0$ (с произвольным носителем) точную нижнюю грань интегралов от непрерывных функций ≥ 0 , мажорирующих эту функцию. Если функция f имеет компактный носитель, то мы возвращаемся к определению,

¹⁾ Мы говорим, что нормы «конечны», а не ограничены. В теореме 69 мы увидим, что это условие влечет за собой ограниченность $\|\mu_n\|$ в их совокупности.

²⁾ Можно даже доказать, что $\|u_n\| = \|\mu_n\|$.

данному на стр. 635. В самом деле, если непрерывная ограниченная функция g мажорирует функцию f и если $\alpha \geq 0$ является непрерывной функцией с компактным носителем, равной 1 в некоторой окрестности носителя f , то αg по-прежнему мажорирует функцию f и $\int^{*R} f d\mu \leq \int \alpha g d\mu \leq \int g d\mu$.

Непрерывную ограниченную функцию, определенную на X , со значениями в \vec{F} будем называть *разложимой*, если ее можно представить в виде конечной суммы $\sum_{i=1}^N \vec{g}_i \varphi_i$, где \vec{g}_i — векторы из \vec{F} , а φ_i — скалярные непрерывные ограниченные функции. Говорят, что функция \vec{f} на X со значениями в \vec{F} μ -интегрируема по Риману, если она имеет аппроксимирующую (в смысле верхнего интеграла Римана) последовательность непрерывных ограниченных разложимых функций.

Теорема 69₂. *Если μ_n образуют последовательность мер ≥ 0 с конечными нормами, узко сходящихся к некоторой мере μ с конечной нормой при n , стремящемся к бесконечности, и если φ является вещественной или комплексной функцией, определенной на X , ограниченной, измеримой относительно всех μ_n и μ -интегрируемой по Риману, то $\int \varphi d\mu_n$ сходится к $\int \varphi d\mu$.*

Доказательство аналогично доказательству теоремы 62.

Теорема 70. *Ограниченнная функция \vec{f} , определенная на X и принимающая значения в \vec{F} , μ -интегрируема по Риману тогда и только тогда, когда множество ее точек разрыва имеет μ -меру (Лебега), равную нулю.*

Доказательство аналогично доказательству теоремы 63.

Следствие 1. *Для того чтобы часть A множества X имела характеристическую функцию, μ -интегрируемую по Риману, необходимо и достаточно, чтобы ее граница \dot{A} имела μ -меру (Лебега), равную нулю.*

Следствие 2. *Если меры $\mu_n \geq 0$ с конечными нормами узко сходятся к некоторой мере μ с конечной нормой и если A — часть X , измеримая по всем μ_n , граница которой имеет μ -меру, равную нулю, то $\mu_n(A)$ сходятся к $\mu(A)$ при n , стремящемся к бесконечности.*

Сходимость широкая и сходимость узкая

Теорема 71. *Для того чтобы последовательность мер $\mu_n \geq 0$ на X с конечными нормами сходилась узко к некоторой мере $\mu \geq 0$ с конечной нормой, необходимо и достаточно, чтобы*

эти нормы сходились к μ широко и чтобы $\|\mu_n\| = \int d\mu_n$ сходилась к $\|\mu\| = \int d\mu$.

Доказательство. Условие это, очевидно, необходимо, поэтому остается доказать лишь его достаточность.

Пусть φ — непрерывная ограниченная функция на X и $M = \|\varphi\|$. Зададим $\varepsilon > 0$. Согласно определению нормы меры μ , существует такая непрерывная функция α с компактным носителем, $0 \leq \alpha \leq 1$, что

$$\int \alpha d\mu \geq \int d\mu - \frac{\varepsilon}{4M}. \quad (\text{IV}, 7; 36)$$

Поскольку меры μ_n сходятся широко к μ при n , стремящемся к бесконечности, $\int \alpha d\mu_n$ сходится к $\int \alpha d\mu$. Так как, по предположению, $\int d\mu_n$ сходится к $\int d\mu$, найдется такое целое число p' , что для $n \geq p'$

$$\int \alpha d\mu_n \geq \int d\mu_n - \frac{\varepsilon}{3M}, \quad (\text{IV}, 7; 37)$$

или

$$\int (1 - \alpha) d\mu_n \leq \frac{\varepsilon}{3M} \quad \text{и} \quad \int (1 - \alpha) d\mu \leq \frac{\varepsilon}{4M}. \quad (\text{IV}, 7; 38)$$

Теперь для функции $\varphi \in \mathcal{C}(X)$ имеем следующее разложение:

$$\mu_n(\varphi) - \mu(\varphi) = (\mu_n(\alpha\varphi) - \mu(\alpha\varphi)) + \mu_n((1 - \alpha)\varphi) - \mu((1 - \alpha)\varphi). \quad (\text{IV}, 7; 39)$$

В силу предположения о широкой сходимости, первый член правой части сходится к 0. Значит, найдется такое целое число p'' , что для всех $n \geq p''$ будет иметь место неравенство

$$|\mu_n(\alpha\varphi) - \mu(\alpha\varphi)| \leq \frac{\varepsilon}{3}. \quad (\text{IV}, 7; 40)$$

В силу неравенства (IV, 7; 38) для двух других членов при $n \geq p = \sup(p', p'')$ имеет место неравенство:

$$|\mu_n(\varphi) - \mu(\varphi)| \leq \frac{\varepsilon}{3} + M \frac{\varepsilon}{3M} + M \frac{\varepsilon}{4M} < \varepsilon, \quad (\text{IV}, 7; 41)$$

чем и заканчивается доказательство теоремы.

Замечание. Если мы вернемся к примерам 1, 1₂ и 2 широкой сходимости к δ (стр. 648—649), то увидим, что в этом случае имеет место и узкая сходимость. То же самое можно сказать относительно примеров 1°), 3°), 4°) на стр. 632—634.

Если теперь мы снова вернемся к теореме 68, то уверенности в узкой сходимости у нас уже не будет, поскольку, если P_n — полиномы ≥ 0 , то $\int P_n(x) dx = +\infty$, откуда не следует, что нормы мер конечны.

Если a_n составляют последовательность неограниченно удаляющихся точек из X (когда X — метрическое пространство; см. определение в следствии из теоремы 55), то $\delta_{(a_n)}$ широко сходятся к 0, ибо при $\varphi \in \mathcal{C}(X)$ для достаточно больших значений n точки a_n находятся вне носителя функции φ (компактного, а потому ограниченного), и, следовательно, $\varphi(a_n) = 0$. Однако, так как $\|\delta_{(a_n)}\| = 1$, то $\delta_{(a_n)}$ не сходится узко к 0. Этот факт дает нам физическое истолкование узкой сходимости. Широкая сходимость учитывает лишь происходящее на конечном расстоянии. Последовательность мер $\mu_n \geq 0$ может сходиться широко к некоторой мере μ , и, однако, это не мешает значительной части масс μ_n «потеряться» в бесконечности. Узкая сходимость, напротив, мешает этому, поскольку $\|\mu_n\|$ стремится к $\|\mu\|$.

Теорема 72. Пусть X — локально компактное пространство и T — некоторое компактное пространство. Пусть φ — скалярная непрерывная и ограниченная функция на $X \times T$. Пусть, кроме того, задана последовательность μ_n мер Радона ≥ 0 на X с конечными нормами, узко сходящаяся при n , стремящаяся к бесконечности, к некоторой предельной мере μ .

Тогда интегралы $\int \varphi(x, t) d\mu_n$ сходятся к интегралу $\int \varphi(x, t) d\mu$ равномерно, когда t пробегает компакт T .

Доказательство. Теорема естественным образом обобщает следствие 3 теоремы 66 на случай узкой сходимости.

Положим $M = \sup_{x, t} |\varphi(x, t)|$. Если вернуться к обозначениям, принятым в доказательстве теоремы 71, то можно найти функцию α , зависящую от числа M и удовлетворяющую соотношению (IV, 7; 36), а следовательно, и (IV, 7; 38) для $n \geq p'$. Приведем теперь разложение (IV, 7; 39) относительно частной функции φ_t : $x \rightarrow \varphi(x, t)$. Первый член $\mu_n(\alpha \varphi_t)$ стремится к $\mu(\alpha \varphi_t)$ равномерно относительно t из T . В самом деле, поскольку мы имеем непрерывную скалярную функцию $(x, t) \rightarrow \alpha(x)\varphi(x, t)$ и для любого t функция $x \rightarrow \alpha(x)\varphi(x, t)$, определенная на X , имеет носитель в фиксированном компакте, а именно в носителе α , то мы находимся на этот раз в условиях следствия 3 теоремы 66. Следовательно, можно независимо от t определить p'' так, чтобы при $n \geq p''$ относительно φ_t имело место неравенство (IV, 7; 40). Что же касается вто-

рого и третьего членов, то они могут быть оценены независимо от t , исходя из соотношений (IV, 7; 38). Поэтому для $n \geq p = \sup(p', p'')$ и любого $t \in T$ имеет место неравенство (IV, 7; 41), чем и заканчивается доказательство теоремы.

§ 8. ТЕНЗОРНОЕ ПРОИЗВЕДЕНИЕ МЕР. КРАТНЫЕ ИНТЕГРАЛЫ

Постановка задачи

Пусть X и Y — локально компактные пространства. Из определения окрестностей непосредственно вытекает, что $X \times Y$ также компактно. Кроме того, если X и Y счетны в бесконечности, то таким же будет пространство $X \times Y$. В самом деле, если H_n составляют некоторую последовательность компактов, дающих в объединении пространство X , а K_n — последовательность компактов, объединение которых равно Y , то $H_n \times K_n$ является последовательностью компактов, объединение которых совпадает с пространством $X \times Y$.

Пусть теперь μ — некоторая мера Радона на X и ν — некоторая мера Радона на Y . Рассмотрим, можно ли, исходя из этих мер, построить новую меру Радона на $X \times Y$. При этом мы ограничимся скалярными мерами Радона¹⁾. В пространстве \mathbb{R}^n мера $dx_1 dx_2 \dots dx_n$, обычно служащая для определения кратных интегралов, является «произведением» мер dx_i , определенных на множителях \mathbb{R} произведения $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$.

Покажем, что если μ и ν — скалярные меры, то существует мера-произведение или тензорное произведение, которое мы обозначим через $\mu \otimes \nu$, или $\mu_x \otimes \nu_y$, или $d\mu(x) \otimes d\nu(y)$, или $d\mu(x) d\nu(y)$. Интеграл от функции \vec{f} , определенной на $X \times Y$, со значениями в банаховом пространстве \vec{F} по отношению к этой мере-произведению называется двойным интегралом и обозначается в обычных символах двойных интегралов:

$$\int \vec{f} d(\mu \otimes \nu) = \int \int \vec{f}(x, y) d\mu(x) d\nu(y)^2. \quad (\text{IV}, 8; 1)$$

¹⁾ Для мер со значениями в конечномерных векторных пространствах имеют место аналогичные результаты. Но они не верны для мер со значениями в нормированных бесконечномерных векторных пространствах.

²⁾ Обозначение $d\mu(x) d\nu(y)$ для меры-произведения некорректно, поскольку x и y не являются переменными! Напротив, в формуле (IV, 8; 1) это же обозначение вполне корректно, так как x и y фигурируют в ней дважды и могут быть заменены любыми другими символами. Точно так же будет вполне корректным для $\lambda = \mu \otimes \nu$ писать $d\lambda(x, y) = d\mu(x) \otimes d\nu(y)$ или $d\mu(x) d\nu(y)$.

Пусть u и v — функции, принадлежащие соответственно пространствам $\mathcal{C}(X)$ и $\mathcal{C}(Y)$. Рассмотрим теперь функцию φ , определенную на $X \times Y$ по формуле

$$\varphi(x, y) = u(x) \cdot v(y). \quad (\text{IV}, 8; 2)$$

Эту функцию обозначают также через $u \otimes v$. Функция $(x, y) \rightarrow x \rightarrow u(x)$ как композиция двух непрерывных функций непрерывна. Точно так же непрерывна функция $(x, y) \rightarrow v(y)$. Поэтому наша функция, являющаяся произведением двух непрерывных функций, также непрерывна. Кроме того, она имеет компактный носитель. Если H и K — носители функций u и v в X и Y соответственно, то ее носитель равен произведению $H \times K$. В самом деле, обозначим через A и B множества точек, в которых u и v отличны от нуля. Тогда $\varphi \neq 0$ в том и только том случае, когда u и $v \neq 0$, т. е. на $A \times B$. Носитель φ есть, следовательно, замыкание $\bar{A} \times \bar{B}$, которое равно произведению $\bar{A} \times \bar{B}$ замыканий, т. е. произведению $H \times K$.

В этом случае хотелось бы, очевидно, чтобы двойной интеграл выражался в виде

$$\int \int u(x) v(y) d\mu(x) d\nu(y) = \left(\int u(x) d\mu(x) \right) \left(\int v(y) d\nu(y) \right), \quad (\text{IV}, 8; 3)$$

т. е. чтобы имело место соотношение

$$(\mu \otimes \nu)(u \otimes v) = \mu(u) \nu(v). \quad (\text{IV}, 8; 4)$$

Мы сейчас увидим, что этого достаточно для того, чтобы единственным образом определить меру — тензорное произведение $\mu \otimes \nu$.

Существование и единственность тензорного произведения

Теорема 73. Пусть X и Y — два локально компактных пространства, а μ и ν — меры Радона, вещественные или комплексные, на X и Y соответственно. Тогда существует, и при том единственная, мера Радона λ на $X \times Y$, удовлетворяющая равенству

$$\lambda(u \otimes v) = \mu(u) \nu(v). \quad (\text{IV}, 8; 5)$$

Эта мера называется тензорным произведением мер μ и ν и обозначается через $\mu \otimes \nu$.

Доказательство этой теоремы мы проведем в несколько этапов.

Первый этап: единственность тензорного произведения. Пусть H и K — компактные множества из X и Y соответственно. Рассмотрим сначала все «разложимые» функции φ вида

$u \otimes v$, где u принадлежит пространству $\mathcal{C}_H(X)$, а v — пространству $\mathcal{C}_K(Y)$. Рассмотрим теперь все функции φ , являющиеся конечной суммой (но заранее не ограниченной) разложимых функций. Такие функции φ могут быть записаны в виде

$$\varphi(x, y) = \sum_{i \in I} u_i(x) v_i(y), \quad u_i \in \mathcal{C}_H(X), \quad v_i \in \mathcal{C}_K(Y). \quad (\text{IV}, 8; 6)$$

Множество таких функций образует некоторое векторное подпространство $\Gamma_{H \times K}$ пространства $\mathcal{C}(X \times Y)$. Имеет место следующая теорема:

Теорема 74 (теорема о плотности). *Пусть H_0 и K_0 — некоторые компакты пространств X и Y соответственно, а H и K — компактные окрестности множеств H_0 и K_0 ¹⁾. Тогда замыкание векторного подпространства $\Gamma_{H \times K}$ в пространстве $\mathcal{C}_{H \times K}(X \times Y)$ содержит подпространство $\mathcal{C}_{H_0 \times K_0}(X \times Y)$.*

Эта теорема означает, что функция φ , непрерывная на $X \times Y$, с носителем в $H_0 \times K_0$ является равномерным пределом некоторой последовательности функций, каждая из которых может быть записана в виде (IV, 8; 6), где носители функций u_i лежат в H , а носители функций v_i лежат в K .

Доказательство. Ограничимся случаем метрических пространств X и Y . В произведении $X \times Y$ мы введем одну из естественных метрик произведения.

Пусть φ — некоторый элемент пространства $\mathcal{C}_{H_0 \times K_0}(X \times Y)$. Эта функция равномерно непрерывна (см. стр. 633), т. е. каково бы ни было $\varepsilon > 0$, существует такое $\eta > 0$, что из неравенств $d(x', x'') \leq \eta$, $d(y', y'') \leq \eta$ вытекает неравенство $|\varphi(x', y') - \varphi(x'', y'')| \leq \varepsilon$. Каждая точка H_0 будет являться центром открытого шара, содержащегося в H , радиус которого можно считать $\leq \eta/2$. Множество этих открытых шаров образует покрытие H_0 , и, следовательно, из него можно извлечь конечное подпокрытие, образованное из шаров U_i , $i \in I$, радиуса $\leq \eta/2$, содержащихся в H . Точно так же можно образовать конечное покрытие K_0 шарами V_j , $j \in J$, с радиусами $\leq \eta/2$, целиком лежащими в K . Точки $U_i \times V_j$ образуют некоторое конечное покрытие множества $H_0 \times K_0$, целиком лежащее в $H \times K$.

Если $x' \in U_i$, $x'' \in U_i$, $y' \in V_j$, $y'' \in V_j$, то $d(x', x'') \leq \eta$ и $d(y', y'') \leq \eta$. Отсюда следует неравенство $|\varphi(x', y') - \varphi(x'', y'')| \leq \varepsilon$. Положим $I' = I \cup \{0\}$ и $J' = J \cup \{0\}$ и рассмотрим открытые

¹⁾ Такие окрестности существуют. Смотри примечание на стр. 457.

покрытия $(U_i)_{i \in I'}$ множества X и $(V_j)_{j \in J'}$ множества Y , где $U_0 = CH_0$ и $V_0 = CK_0$. Пусть теперь $(\alpha_i)_{i \in I'}$ и $(\beta_j)_{j \in J'}$ — разложения единицы, соответствующие этим покрытиям (это означает, что X и Y счетны в бесконечности; можно избавиться от этого условия, но мы этого делать не будем). Выберем в каждом из множеств U_i некоторую точку x_i , а в каждом V_j — точку y_j и определим функцию Φ на $X \times Y$ по формуле:

$$\Phi(x, y) = \sum_{i \in I', j \in J'} \alpha_i(x) \beta_j(y) \varphi(x_i, y_j). \quad (\text{IV}, 8; 7)$$

Поскольку при равных нулю i или j функция $\varphi(x_i, y_j) = 0$, то вместо $\sum_{i \in I', j \in J'}$ можно писать сумму $\sum_{i \in I, j \in J}$. Функция Φ имеет вид (IV, 8; 6) относительно множества индексов $I \times J$, где, например, можно считать $u_{ij}(x) = \alpha_i(x)$, $v_{ij}(y) = \varphi(x_i, y_j) \beta_j(y)$. Тогда для i и j , отличных от 0, функции u_{ij} имеют носитель в H , а функции v_{ij} — в K . Заметим, что произведения $\alpha_i \otimes \beta_j$, $(i, j) \in I' \times J'$, образуют некоторое разложение единицы для $X \times Y$, подчиненное покрытию его множествами $U_i \times V_j$, $(i, j) \in I' \times J'$. Поэтому имеет место равенство $\Phi(x, y) = \sum_{i, j} \alpha_i(x) \beta_j(y) \varphi(x_i, y_j)$, где суммирование может относиться как $I \times J$, так и $I' \times J'$. Поэтому

$$|\Phi(x, y) - \varphi(x, y)| \leq \sum_{i, j} |\alpha_i(x) \beta_j(y)| |\varphi(x_i, y_j) - \varphi(x, y)|. \quad (\text{IV}, 8; 8)$$

Члены правой части отличны от нуля только тогда, когда $i \neq 0, j \neq 0$, $(x, y) \in U_i \times V_j$. Но тогда $|\varphi(x_i, y_j) - \varphi(x, y)| \leq \varepsilon$ и, следовательно, правая часть не превосходит $\sum_{i, j} |\alpha_i(x) \beta_j(y)| \varepsilon = \varepsilon$; доказательство теоремы закончено.

Следствие. Если меры μ и ν заданы на X и Y , то существует только одна мера λ , удовлетворяющая условиям теоремы 73.

В самом деле, для функции $\varphi \in \mathcal{C}_{H_0 \times K_0}(X \times Y)$ существует последовательность функций $\psi_n \in \mathcal{C}_{H \times K}(X \times Y)$ вида (IV, 8; 6), сходящаяся равномерно к φ при n , стремящемся к бесконечности. Поэтому, если мера λ существует, $\lambda(\varphi) = \lim_{n \rightarrow \infty} \lambda(\varphi_n)$. Однако в этом случае, в силу равенства (IV, 8; 5), известно ее значение на разложимых функциях $u \otimes v$, а, следовательно, после суммирования — на всех функциях ψ_n ; предыдущий переход к пределу говорит о том, что $\lambda(\varphi)$ определена для любой

функции $\varphi \in \mathcal{C}(X \times Y)$, чем полностью доказывается единственность тензорного произведения¹⁾.

Создавшаяся ситуация подобна, очевидно, ситуации, имевшей место в теореме 49 гл. II (о продолжении линейных непрерывных функций, определенных на плотных подмножествах). Однако в данном случае имеется небольшая трудность, заключающаяся в том, что мы не рассматривали $\mathcal{C}(X \times Y)$ как топологическое пространство. Это обстоятельство вынуждает нас провести прямое доказательство следствия.

Второй этап: существование тензорного произведения.

Теорема 75. Пусть φ — некоторая функция из $\mathcal{C}(X \times Y)$. Тогда для каждого фиксированного $x \in X$ можно вычислить интеграл

$$\psi(x) = \int \varphi(x, y) d\nu(y). \quad (\text{IV}, 8; 9)$$

Функция $\psi: x \rightarrow \psi(x)$ непрерывна и имеет компактный носитель в X . Следовательно, можно вычислить

$$\lambda(\varphi) = \int \psi(x) d\mu(x) = \int \left[\int \varphi(x, y) d\nu(y) \right] d\mu(x). \quad (\text{IV}, 8; 10)$$

Линейная форма $\varphi \rightarrow \lambda(\varphi)$ является мерой Радона на $X \times Y$, удовлетворяющей соотношению (IV, 8; 5).

Доказательство. Обозначим через H и K проекции носителя φ на пространства X и Y соответственно. Эти множества, как непрерывные образы некоторого компакта, сами компакты. Тогда носитель частной функции $\varphi_x: y \rightarrow \varphi(x, y)$ при любом x лежит в K . Впрочем, $\varphi_x \equiv 0$, если $x \notin H$. В теореме 66 мы видели, что при x , стремящемся к a , частная функция φ_x стремится к частной функции φ_a равномерно на компакте K (а, значит, равномерно на Y , поскольку ее носитель всегда лежит в K) и, следовательно, φ_x стремится к φ_a в нормированном векторном пространстве $\mathcal{C}_K(Y)$. Из определения меры ν как линейной формы на $\mathcal{C}(Y)$, сужение которой на $\mathcal{C}_K(Y)$ непрерывно, следует, что $\psi(x) = \int \varphi_x d\nu$ сходится к $\psi(a) = \int \varphi_a d\nu$, т. е. что функция ψ непрерывна на X . Так как при $x \notin H$, $\varphi_x \equiv 0$, а, значит, $\psi(x) = 0$, носитель этой функции лежит в H . Следовательно, она принадлежит пространству $\mathcal{C}(X)$, и мы можем вычислить $\mu(\psi)$, а потому имеет смысл выражение (IV, 8; 10). Очевидно, эта функция линейно зависит от φ .

¹⁾ Независимо от нахождения тензорного произведения мы доказали большее: две меры λ_1 и λ_2 на $X \times Y$, принимающие одно и то же значение на каждой функции φ из $\mathcal{C}(X \times Y)$ вида $u \otimes v$, $u \in \mathcal{C}(X)$, $v \in \mathcal{C}(Y)$, принимают одно и то же значение для любой функции $\varphi \in \mathcal{C}(X, Y)$, т. е. равны между собой.

Если ϕ сохраняет носитель в фиксированном компакте $H \times K$, то ψ сохраняет носитель в H и удовлетворяет неравенству

$$|\psi(x)| \leq \|\phi\|_H \|v\|_K. \quad (\text{IV}, 8; 11)$$

Отсюда вытекает оценка

$$|\lambda(\phi)| \leq \|\mu\|_H \|\psi\| \leq \|\mu\|_H \|v\|_K \|\phi\|. \quad (\text{IV}, 8; 12)$$

Мы видели, что λ является мерой на $X \times Y$ и, кроме того, имеет место неравенство

$$\|\lambda\|_{H \times K} \leq \|\mu\|_H \|v\|_K. \quad (\text{IV}, 8; 13)$$

Остается лишь убедиться в том, что λ удовлетворяет соотношению (IV, 8; 5). Если мы положим $\phi(x, y) = u(x)v(y)$, то получим, что $\psi(x) = u(x) \int v(y) dv(y) = u(x)v(v)$, а, следовательно, $\lambda(\phi) = v(v) \int u(x) d\mu(x) = \mu(u)v(v)$, чем и заканчивается доказательство теоремы.

Следствие. Если заданы μ и v , то найдется хотя бы одна мера λ , удовлетворяющая условиям теоремы 73.

Таким образом, теорема 73 установлена путем доказательства двух не зависящих друг от друга утверждений — существования и единственности.

Замечание. Можно было бы начать с того, чтобы зафиксировать y и вычислить интеграл

$$\theta(y) = \int \phi(x, y) d\mu(x). \quad (\text{IV}, 8; 14)$$

Полученная функция θ была бы непрерывной функцией с компактным носителем в Y , и можно было бы вычислить

$$\lambda(\phi) = \int \theta(y) dv(y) = \int \left[\int \phi(x, y) d\mu(x) \right] dv(y). \quad (\text{IV}, 8; 15)$$

Таким путем можно определить другую меру Радсна λ , удовлетворяющую соотношению (IV, 8; 5). Из единственности, доказанной в следствии теоремы 74, следует, что мы получим тот же результат. Поэтому имеет место соотношение

$$\int \left[\int \phi(x, y) d\nu(y) \right] d\mu(x) = \int \left[\int \phi(x, y) d\mu(x) \right] d\nu(y). \quad (\text{IV}, 8; 16)$$

Последнее равенство составляет содержание теоремы о перестановке знаков интегрирования.

Примеры тензорных произведений

Если a является точкой X , а b — некоторой точкой Y , то между соответствующими мерами Дирака имеет место следующее очевидное соотношение:

$$\delta_{(a)} \otimes \delta_{(b)} = \delta_{(a, b)} \quad \text{или} \quad d\delta_{(a)}(x) d\delta_{(b)}(y) = d\delta_{(a, b)}(x, y), \quad (\text{IV}, 8; 17)$$

вытекающее из того, что его правая часть удовлетворяет равенству (IV, 8; 5).

Пусть теперь μ является мерой $\sum_i c_i \delta_{(a_i)}$, а ν — мерой $\sum_j d_j \delta_{(b_j)}$. Тогда имеет место формула

$$\sum_i c_i \delta_{(a_i)} \otimes \sum_j d_j \delta_{(b_j)} = \sum_{i,j} c_i d_j \delta_{(a_i, b_j)}. \quad (\text{IV}, 8; 18)$$

Элементарные свойства

Легко видеть, что отображение $(\mu, \nu) \rightarrow \mu \otimes \nu$ пространства $\mathcal{C}'(X) \times \mathcal{C}'(Y)$ в пространство $\mathcal{C}'(X \times Y)$ является *билинейным отображением*¹⁾, т. е. имеют место равенства:

$$(\mu_1 + \mu_2) \otimes (\nu_1 + \nu_2) = \mu_1 \otimes \nu_1 + \mu_1 \otimes \nu_2 + \mu_2 \otimes \nu_1 + \mu_2 \otimes \nu_2, \quad (\text{IV}, 8; 19)$$

$$h\mu \otimes k\nu = hk(\mu \otimes \nu), \text{ где } h \text{ и } k — \text{скаляры.}$$

Очевидно также, что если μ и ν вещественны, то функция $\mu \otimes \nu$ также вещественна и что, если μ и $\nu \geqslant 0$, то и $\mu \otimes \nu \geqslant 0$. В самом деле, для того чтобы проверить неравенство, $(\mu \otimes \nu)(\varphi) \geqslant 0$, достаточно при $\varphi \geqslant 0$ провести вычисления теоремы 75.

Наконец, соотношение (IV, 8; 13) показывает, что если μ и ν имеют конечные нормы, то конечную норму будет иметь произведение $\mu \otimes \nu$, и при этом

$$\|\mu \otimes \nu\| \leqslant \|\mu\| \|\nu\|. \quad (\text{IV}, 8; 19_2)$$

Носитель меры $\mu \otimes \nu$

Теорема 76. Если A — носитель меры μ , а B — носитель меры ν , то носитель меры $\mu \otimes \nu$ есть множество $A \times B$.

Доказательство. 1°) Так как мера μ равна нулю в C_A , то в открытом множестве $C_A \times Y$ произведение $\mu \otimes \nu$ равно нулю. Точно так же мера $\mu \otimes \nu$ равна нулю в $X \times C_B$, ибо ν обращается в нуль в C_B . Значит, мера $\mu \otimes \nu$ равна нулю в объединении этих двух открытых множеств (теорема 13). Это объединение равно дополнению множества $A \times B$. Значит, носитель меры $\mu \otimes \nu$ лежит в $A \times B$.

2°) Пусть теперь (a, b) — произвольная точка множества $A \times B$, и пусть \mathcal{A} (соответственно \mathcal{B}) является некоторой открытой окрестностью точки a в X (соответственно точки b

¹⁾ Этим и объясняется название — тензорное произведение.

в Y). Согласно свойству носителя (теорема 14), найдется такая функция $u \in \mathcal{C}(X)$ с носителем в \mathcal{A} , что $\mu(u) \neq 0$. Точно так же найдется такая функция $v \in \mathcal{C}(Y)$ с носителем в \mathcal{B} , что $v(v) \neq 0$. Если теперь мы положим $\varphi(x, y) = u(x)v(y)$, то получим $(\mu \otimes v)(\varphi) = \mu(u)v(v) \neq 0$. Этим доказано, что не существует такой окрестности точки (a, b) , в которой мера $\mu \otimes v$ была бы равна нулю, и что, следовательно, точка (a, b) принадлежит носителю меры $\mu \otimes v$. Значит, этот носитель совпадает с $A \times B$.

Вычисление двойного интеграла путем двух последовательных простых интегрирований

В теореме 75 предлагается способ вычисления двойного интеграла с помощью двух последовательных простых интегрирований. Однако эта теорема применима лишь в том случае, когда рассматривается интеграл от функции φ , принадлежащей пространству $\mathcal{C}(X \times Y)$. Для μ и $v \geq 0$ имеет место значительно более общая теорема.

Теорема 77. (Фубини — Лебега). Предположим, что X и Y — локально компактные пространства, счетные в бесконечности. Пусть μ и v — меры ≥ 0 на X и Y соответственно. Пусть \vec{F} — некоторое пространство Банаха, а \vec{f} — функция, определенная на $X \times Y$ со значениями в \vec{F} и интегрируемая (соответственно измеримая) относительно $\mu \otimes v$. Тогда для фиксированного $x \in X$ частная функция $\vec{f}_x: y \rightarrow \vec{f}(x, y)$ не всегда интегрируема (соответственно измерима) относительно v , однако она интегрируема (соответственно измерима) относительно v по крайней мере для μ -почти всех значений x из X . Можно также для μ -почти всех значений x определить¹⁾

$$\vec{g}(x) = \int \vec{f}(x, y) dv(y) \in \vec{F}. \quad (\text{IV}, 8; 20)$$

Функция $\vec{g}: x \rightarrow \vec{g}(x)$, определенная μ -почти всюду на X , со значениями в \vec{F} будет μ -интегрируемой, и имеет место формула

$$\int \int \vec{f}(x, y) d\mu(x) dv(y) = \int \vec{g}(x) d\mu(x) = \int \left[\int \vec{f}(x, y) dv(y) \right] d\mu(x). \quad (\text{IV}, 8; 21)$$

¹⁾ Мы предполагаем теперь, что функция \vec{f} интегрируема. Дальнейшие рассуждения больше не связаны со случаем «соответственно измерима».

Эту формулу обычно пишут в виде

$$\int \int \vec{f}(x, y) d\mu(x) dv(y) = \int d\mu(x) \int \vec{f}(x, y) dv(y). \quad (\text{IV}, 8; 22)$$

Это обозначение несколько необычно. Чтобы не писать скобок, мы записываем $\int \vec{g}(x) d\mu(x)$ в виде $\int d\mu(x) \vec{g}(x)$. К этим часто применяемым обозначениям следует привыкнуть.

Доказательство. Докажем сначала такое утверждение:

Лемма 1. Пусть f — функция ≥ 0 (с конечными или бесконечными значениями), полуунипрерывная снизу на $X \times Y$. Тогда имеет место формула

$$\int^* d\mu(x) \int^* f(x, y) dv(y) = \int \int^* f(x, y) d\mu(x) dv(y). \quad (\text{IV}, 8; 22_2)$$

Рассмотрим множество $\mathcal{C}_+(X \times Y; f)$ всех функций ≥ 0 из $\mathcal{C}(X \times Y)$, ограничивающих снизу функцию f . Функции этого множества обладают свойством возрастания в смысле теоремы 39₄. Кроме того, верхняя огибающая этих функций равна f . [В самом деле, эта огибающая, очевидно, $\leq f$. Однако если $a \in X$, то найдется такая открытая окрестность \mathcal{V} точки a , что $f(\mathcal{V}) \geq (f(a) - \varepsilon)^+ = \max(f(a) - \varepsilon, 0)$. Значит, существует некоторая компактная окрестность \mathcal{W} точки a в \mathcal{V} , а потому, согласно следствию 1 из теоремы 11, существует непрерывная функция ψ , $0 \leq \psi \leq 1$, с носителем в \mathcal{W} , равная 1 в \mathcal{W} . Имеем: $0 \leq (f(a) - \varepsilon)^+ \psi \leq f$ (это справедливо в \mathcal{V} , а вне \mathcal{V} обе функции равны нулю). Эта функция в точке a принимает значение $(f(a) - \varepsilon)^+$. Следовательно, точная верхняя грань величин $\psi(a)$ по всевозможным функциям ψ из $\mathcal{C}_+(X \times Y)$, мажорируемым функцией f , равна $f(a)$.] К множеству $\mathcal{C}_+(X \times Y; f)$ можно применить теорему 39₄. Это снова, впрочем, дает теорему 39₂, а именно:

$$\int \int^* f(x, y) d\mu(x) dv(y) = \sup_{\varphi \in \mathcal{C}_+(X \times Y; f)} (\mu \otimes v)(\varphi).$$

Рассмотрим, кроме того, множество соответствующих частных функций $\varphi_x: y \rightarrow \varphi(x, y)$. Это множество является множеством непрерывных функций ≥ 0 , удовлетворяющих условию возрастания теоремы 50₄, а его верхней огибающей является функция $f_x: y \rightarrow f(x, y)$. Из теоремы следует, что

$$\int^* f(x, y) dv(y) = \sup_{\varphi \in \mathcal{C}_+(X \times Y; f)} \int \varphi(x, y) dv(y).$$

Однако для каждой функции φ функция $J_\varphi : x \rightarrow \int \varphi(x, y) d\nu(y)$ непрерывна, $\geqslant 0$ и имеет компактный носитель на X (теорема 75). Значит, верхняя огибающая таких функций J_φ , т. е. функция $F: x \rightarrow \int^* f(x, y) d\nu(y)$, является полунепрерывной снизу функцией $\geqslant 0$. Так как J_φ и J_ψ мажорируются функцией $J_{\varphi+\psi}$, а, значит, $\sup(J_\varphi, J_\psi) \leqslant J_{\varphi+\psi}$, то множество J_φ удовлетворяет условию возрастания теоремы 394. Значит, можно снова применить эту теорему и тогда

$$\begin{aligned} \int^* d\mu(x) \int^* f(x, y) d\nu(y) &= \sup_{\varphi \in \mathcal{E}_+(X \times Y; f)} \int d\mu(x) \int \varphi(x, y) d\nu(y) = \\ &= \sup_{\varphi \in \mathcal{E}_+(X \times Y; f)} (\mu \otimes \nu)(\varphi)^1 = \int \int^* f(x, y) d\mu(x) d\nu(y), \end{aligned}$$

чем и доказывается лемма 1.

Лемма 2. Пусть $f \geqslant 0$ — произвольная функция на $X \times Y$. Тогда

$$\int^* d\mu(x) \int^* f(x, y) d\nu(y) \leqslant \int \int^* f(x, y) d\mu(x) d\nu(y)^2. \quad (\text{IV}, 8; 23)$$

В самом деле, пусть $g \geqslant f$ и полунепрерывна снизу. В силу леммы 1,

$$\begin{aligned} \int^* d\mu(x) \int^* f(x, y) d\nu(y) &\leqslant \int^* d\mu(x) \int^* g(x, y) d\nu(y) = \\ &= \int \int^* g(x, y) d\mu(x) d\nu(y). \end{aligned}$$

Переходя, согласно теореме 50₃, в последнем члене к точной нижней грани при переменной функции g , получим $\int \int^* f(x, y) d\mu(x) d\nu(y)$, откуда и вытекает (IV, 8; 23).

Перейдем теперь к доказательству теоремы. Пусть \vec{f} — некоторая $(\mu \otimes \nu)$ -интегрируемая функция на $X \times Y$ со значениями в \vec{F} . Эта функция имеет аппроксимирующую последовательность, образованную непрерывными разложимыми функциями \vec{f}_n с компактными носителями вида $\sum_i g_{in}\varphi_{in}$, $\varphi_{in} \in$

¹⁾ В силу теоремы 75.

²⁾ Результат леммы 1 может навести на мысль, что полунепрерывные снизу функции занимают весьма привилегированное положение среди неотрицательных функций в том смысле, что для них имеет место равенство вместо неравенства. Теорема 78 покажет, что равенство имеет место и для всех $(\mu \otimes \nu)$ -измеримых функций $\geqslant 0$ на $X \times Y$, в частности, для всех борелевских функций, а всякая полунепрерывная функция борелевская.

$\in \mathcal{C}(X \times Y)$. Для такой функции теорема Фубини — Лебега очевидна. Она сводится к случаю функций из $\mathcal{C}(X \times Y)$, а это в силу теоремы 75 является следствием определения тензорного произведения.

Величины $\int^* d\mu(x) \int^* \|\vec{f} - \vec{f}_n\| d\nu(y) \leq \int \int^* \|\vec{f} - \vec{f}_n\| d\mu(x) d\nu(y)$ (лемма 2) сходятся к 0. Но тогда в силу теоремы 38 можно выбрать такую подпоследовательность, которую мы будем по-прежнему обозначать через \vec{f}_n , для которой $\int^* \|\vec{f} - \vec{f}_n\| d\nu(y)$ сходится к 0 для $d\mu$ -почти всех x из X . Значит, существует такое множество A из X нулевой μ -меры, что для $x \notin A$ эта сходимость имеет место. Функция $(\vec{f}_n)_x: y \rightarrow \vec{f}_n(x, y)$ непрерывна, разложима, имеет компактный носитель и ν -интегрируема. Поэтому, для каждого $x \in A$, f_x ν -интегрируема и $(\vec{f}_n)_x$ образуют аппроксимирующую последовательность:

$$\vec{g}(x) = \int \vec{f}(x, y) d\nu(y) = \lim_{n \rightarrow \infty} \int \vec{f}_n(x, y) d\nu(y) = \lim_{n \rightarrow \infty} \vec{g}_n(x).$$

Каждая из функций \vec{g}_n непрерывна, разложима и имеет компактный носитель (теорема 75), а, следовательно, μ -интегрируема и $\int \vec{g}_n d\mu = \int \int \vec{f}_n d(\mu \otimes \nu)$. Функция \vec{g} определена μ -почти всюду и

$$\int^* \|\vec{g} - \vec{g}_n\| d\mu \leq \int^* d\mu(x) \int^* \|\vec{f} - \vec{f}_n\| d\nu(y) \leq \int \int^* \|\vec{f} - \vec{f}_n\| d(\mu \otimes \nu)$$

стремится к нулю. Следовательно, \vec{g} μ -интегрируема, а \vec{g}_n образуют аппроксимирующую последовательность:

$$\begin{aligned} \int d\mu(x) \int \vec{f}(x, y) d\nu(y) &= \int \vec{g} d\mu = \lim_{n \rightarrow \infty} \int \vec{g}_n d\mu = \\ &= \lim_{n \rightarrow \infty} \int \int \vec{f}_n d(\mu \otimes \nu) = \int \int \vec{f} d(\mu \otimes \nu). \end{aligned}$$

Теорема доказана.

Замечания. 1°) Естественно, что можно было бы идти в обратном порядке и получить формулу

$$\int \int \vec{f}(x, y) d\mu(x) d\nu(y) = \int d\nu(y) \int \vec{f}(x, y) d\mu(x). \quad (\text{IV}, 8; 24)$$

Отсюда, в частности, вытекает следующее: если заранее известно, что функция \vec{f} интегрируема по $\mu \otimes \nu$, то можно написать

формулу изменения порядка интегрирования

$$\int d\mu(x) \int \vec{f}(x, y) d\nu(y) = \int d\nu(y) \int \vec{f}(x, y) d\mu(x). \quad (\text{IV}, 8; 25)$$

Z 2°) Если \vec{f} не интегрируема по $\mu \otimes \nu$, то может случиться, что одна часть в равенстве (IV, 8; 25) имеет смысл, а другая — нет. Может даже случиться, что обе части имеют смысл, но значения их не совпадают.

Рассмотрим, например, на квадрате $[0, 1] \times [0, 1]$ пространства \mathbb{R}^2 меру-произведение $dx \otimes dy$ и вещественную функцию f , определенную по формуле

$$f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2} \quad \text{для } (x, y) \neq (0, 0). \quad (\text{IV}, 8; 26)$$

Мы не затрудняем себя определением $f(0, 0)$, поскольку при любом фиксированном x эта функция может быть не определена лишь при *одном* значении y и, наоборот, при любом фиксированном y — лишь при *одном* значении x . Однако множество, состоящее из одной точки, имеет меру нуль относительно dx или dy .

Покажем, что выражение

$$\int_{[0, 1]} dx \int_{[0, 1]} \frac{x^2 - y^2}{(x^2 + y^2)^2} dy \quad (\text{IV}, 8; 27)$$

имеет смысл и что оно равно $\pi/4$.

В самом деле, для фиксированного $x \neq 0$ первообразной по y для непрерывной функции $y \rightarrow \frac{x^2 - y^2}{(x^2 + y^2)^2}$ является функция $y \rightarrow \frac{y}{x^2 + y^2}$. Поэтому для $x \neq 0$ имеем:

$$\int_{[0, 1]} \frac{x^2 - y^2}{(x^2 + y^2)^2} dy = \frac{1}{x^2 + 1}. \quad (\text{IV}, 8; 28)$$

Для того чтобы применить формулу (IV, 8; 27), нет необходимости вычислять интеграл при $x = 0$, ибо множество $\{0\}$ имеет нулевую меру относительно dx . Окончательно получаем

$$\int_{[0, 1]} dx \int_{[0, 1]} \frac{x^2 - y^2}{(x^2 + y^2)^2} dy = \int_0^1 \frac{dx}{1+x^2} = \frac{\pi}{4}. \quad (\text{IV}, 8; 29)$$

Если теперь мы попытаемся вычислить этот интеграл в обратном порядке

$$\int_{[0, 1]} dy \int_{[0, 1]} \frac{x^2 - y^2}{(x^2 + y^2)^2} dx, \quad (\text{IV}, 8; 30)$$

то, переобозначая переменные интегрирования x и y через y и x , придем к интегралу

$$\int_{[0,1]} dx \int_{[0,1]} \frac{y^2 - x^2}{(x^2 + y^2)^2} dy, \quad (\text{IV}, 8; 31)$$

противоположному интегралу (IV, 8; 27) и, следовательно, равному $-\pi/4$. Таким образом, хотя оба интеграла (IV, 8; 27) и (IV, 8; 30) имеют смысл, их значения не совпадают. Это означает, что теорема Фубини в этом случае не применима, а, следовательно, функция f не интегрируема относительно меры-произведения $dx \otimes dy$. С помощью методов, которые будут подробно изложены позже, в этом можно убедиться непосредственно, выполнив предварительно замену переменных перехода к полярным координатам $x = r \cos \varphi$, $y = r \sin \varphi$ (формула (IV, 9; 92)), а затем применив следствие 4 из теоремы 108 при $\alpha = 2$.

Z 3°) Нельзя надеяться на то, что \vec{f}_x будет интегрируемой по y для всех значений x .

В самом деле, если мы вернемся к примеру с мерами dx и dy на \mathbb{R} , то увидим, что \vec{f} можно изменить на прямой $x = a$, параллельной оси y , не изменяя ее интегрируемости по отношению к мере $dx \otimes dy$, поскольку прямая $x = a$ является множеством меры нуль в \mathbb{R}^2 .

Это позволяет нам в качестве функции $y \rightarrow f(a, y)$ брать произвольную функцию от y , которая не обязана быть интегрируемой по y относительно dy . Однако это обстоятельство не играет никакой роли в том вопросе, который нас интересует.

Как только мы узнали, что функция g определена для μ -почти всех значений x и интегрируема относительно μ , нам этого достаточно для того, чтобы вычислить двойной интеграл путем двух последовательных простых интегрирований. В этом даже состоит некоторое удобство, которым надо еще уметь пользоваться: если для некоторых исключительных значений x , образующих множество μ -меры нуль, интеграл $\int \vec{f}(x, y) d\nu(y)$ или не существует, или сложен для вычисления, нет никакой необходимости заниматься им в этих точках.

4°) Теорему Фубини можно сравнить с теоремой суммирования по блокам в безусловно сходящихся рядах (теорема 59 гл. II). В этой теореме точно так же мы должны сначала предположить, что ряд безусловно сходится для того, чтобы иметь затем возможность суммировать его блоками. Мы видели, что существует, однако, случай, когда не нужно знать заранее, сходится ли ряд безусловно, а именно случай, когда члены ряда

вещественны и $\geqslant 0$. То же самое имеет место и здесь, как это видно из следующей теоремы и ее следствия 1.

Теорема 78 (Фубини — Фату). Пусть X, Y — локально компактные пространства, счетные в бесконечности. Пусть f — некоторая функция $\geqslant 0$ (с конечными или бесконечными значениями), определенная на $X \times Y$ и $(\mu \otimes \nu)$ -измеримая на нем. Тогда имеет место такое равенство:

$$0 \leqslant \int^* d\mu(x) \int^* f(x, y) d\nu(y) = \int \int^* f(x, y) d\mu(x) d\nu(y) \leqslant +\infty. \quad (\text{IV}, 8; 32)$$

Доказательство. Поскольку функция f $(\mu \otimes \nu)$ -измерима, она является пределом некоторой возрастающей последовательности измеримых ограниченных функций $f_n \geqslant 0$ с компактным носителем (например, функций $f_{m, k}$, указанных на стр. 521). Для каждой функции f_n , согласно теореме Фубини — Лебега (теорема 77),

$$\int d\mu(x) \int f_n(x, y) d\nu(y) = \int \int f_n(x, y) d\mu(x) d\nu(y),$$

где второй интеграл левой части равенства, $\int f_n(x, y) d\nu(y)$, имеет смысл для μ -почти всех значений x . То же самое равенство будет иметь место, если везде вместо интегралов \int взять имеющие смысл интегралы \int^* (с конечными значениями или нет). Теперь остается устремить n к бесконечности и применить два раза слева и один раз справа теорему 36 Фату.

Замечание. Можно утверждать несколько больше. Пусть A_n — множество нулевой μ -меры точек x , для которых функция $(f_n)_x$ не является ν -интегрируемой. Пусть $A = \bigcup A_n$; A также имеет μ -меру, равную нулю. Для $x \notin A$ все функции $(f_n)_x$ ν -измеримы, а, следовательно, измерим их предел f_x . Таким образом, частная функция f_x ν -измерима для μ -почти всех значений x (см. обобщение следствия 4).

Кроме того, $g_n(x) = \int f_n(x, y) d\nu(y)$ и $g(x) = \int^* f(x, y) d\nu(y)$ определены на C_A и μ -измеримы. Следовательно, g определена на C_A и μ -измерима.

Однако мы видели, что для измеримой функции $\geqslant 0$ вместо \int^* можно писать \int , если даже она принимает бесконечные значения (замечание 2 после следствия 3 теоремы 39). В наших условиях для $(\mu \otimes \nu)$ -измеримой функции $f \geqslant 0$ можно

написать такие соотношения:

$$0 \leqslant \int d\mu(x) \int f(x, y) d\nu(y) = \int \int f(x, y) d\mu(x) d\nu(y) \leqslant +\infty. \quad (\text{IV}, 8; 33)$$

Здесь второй интеграл в левой части равенства имеет смысл для μ -почти всех значений x как интеграл (конечный или бесконечный) от ν -измеримой функции $f_x \geqslant 0$. Первый интеграл левой части (конечный или бесконечный) является интегралом от функции $g \geqslant 0$, определенной μ -почти всюду и μ -измеримой. Двойной интеграл в правой части является интегралом от $(\mu \otimes \nu)$ -измеримой функции $f \geqslant 0$.

Следствие 1. Пусть X и Y — локально компактные пространства, счетные в бесконечности, μ и ν — меры Радона $\geqslant 0$ на X и Y соответственно, f — вещественная функция $\geqslant 0$, измеримая относительно меры $\mu \otimes \nu$. Пусть для μ -почти всех значений x частная функция $f_x : y \rightarrow f(x, y)$ интегрируема по ν , а функция g , определенная для μ -почти всех значений x по формуле

$$g(x) = \int f(x, y) d\nu(y), \quad (\text{IV}, 8; 33_2)$$

интегрируема по μ . Тогда функция f интегрируема по $\mu \otimes \nu$ и имеет место равенство (IV, 8; 22).

Доказательство очевидно. Условия теоремы означают, что левая часть равенства (IV, 8; 33₂) конечна, а, следовательно, будет конечной и его правая часть. Поскольку функция f $\mu \otimes \nu$ -измерима и ее верхний интеграл конечен, то, согласно теореме 39, она интегрируема.

Следствие 2. Если функция \vec{f} со значениями в банаховом пространстве \vec{F} определена на пространстве $X \times Y$ и $(\mu \otimes \nu)$ -измерима на нем, а хотя бы один из трех интегралов

$$\int \int^* \| \vec{f}(x, y) \| d\mu(x) d\nu(y), \quad \int^* d\mu(x) \int^* \| \vec{f}(x, y) \| d\nu(y),$$

$$(\text{IV}, 8; 34)$$

$$\int^* d\nu(y) \int^* \| \vec{f}(x, y) \| d\mu(x)$$

конечен, то функция \vec{f} интегрируема относительно $\mu \otimes \nu$ и имеют место формулы (IV, 8; 22) и (IV, 8; 24).

В самом деле, поскольку функция $\| \vec{f} \|$ измерима, то можно применить теорему Фубини — Фату в предыдущем виде, и тогда тот факт, что один из трех рассматриваемых интегралов конечен, будет означать, что конечны все три рассматрива-

мых интеграла, а, следовательно, в частности, конечен интеграл $\int \int^* \vec{f}(x, y) \|d(\mu \otimes v)\|$. Из теоремы 39 вытекает, что если функция \vec{f} $(\mu \otimes v)$ -измерима, то она и $(\mu \otimes v)$ -интегрируема. Мы попадаем тем самым в условия применимости теоремы Фубини — Лебега.

Замечание. Естественно, что теорему Фубини можно применить к интегралам вида $\int \int_A^* \vec{f}(x, y) d\mu(x) d\nu(y)$, где A является $(\mu \otimes v)$ -измеримым подмножеством пространства $X \times Y$, лишь бы только функция \vec{f} была интегрируемой на множестве A или была на нем $(\mu \otimes v)$ -измерима и ≥ 0 . Поскольку это утверждение сводится к вычислению интеграла $\int \int \vec{f}(x, y) d\mu(x) d\nu(y)$, где \vec{f} — функция, равная функции \vec{f} на A и $\vec{0}$ вне A , то новой теоремы мы не получаем. Пусть $A(a)$ при $a \in X$ — это сечение множества A вертикалью с абсциссой a , т. е. множество точек $y \in Y$, таких, что $(a, y) \in A$. Тогда

$$\int \int_A^* \vec{f}(x, y) d\mu(x) d\nu(y) = \int_X d\mu(x) \int_{A(X)}^* \vec{f}(x, y) d\nu(y). \quad (\text{IV}, 8; 35)$$

В частности, если $f = 1$ на A , то

$$(\mu \otimes v)(A) = \int v(A(x)) d\mu(x), \quad (\text{IV}, 8; 35_2)$$

где $v(A(x))$ является v -мерой сечения $A(x)$, определенной для μ -почти всех x (конечной или бесконечной).

Следствие 3. Пусть A есть $(\mu \otimes v)$ -измеримая часть пространства $X \times Y$. Для того чтобы она имела нулевую $(\mu \otimes v)$ -меру, необходимо и достаточно, чтобы для μ -почти каждого значения x ее сечение $A(x)$ имело v -меру, равную нулю.

В самом деле, $\int v(A(x)) d\mu(x)$, согласно теореме 26, равен нулю тогда и только тогда, когда подинтегральная функция μ -почти всюду равна нулю.

Замечание. В этой формулировке *существенно* предположение, что A $(\mu \otimes v)$ -измеримо. Можно (с помощью аксиомы Цермело) доказать существование такого подмножества A из $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, которое с каждой из прямых $x = \text{const}$ пересекается в единственной точке, но $(dx \otimes dy)$ -мера которого не равна нулю (потому что оно не $(dx \otimes dy)$ -измеримо).

Следствие 4. Пусть f — функция со значениями в метризуемом сепарабельном пространстве F . Если f $(\mu \otimes \nu)$ -измерима, то для μ -почти всех значений $x \in X$ частная функция $f_x: y \rightarrow f(x, y)$ является ν -измеримой.

Доказательство. Пусть A — борелевская часть пространства $X \times Y$. Тогда для каждой точки $x \in X$ сечение $A(x)$ борелевское. В самом деле, части $X \times Y$, сечения которых для данного x борелевские, очевидно, образуют σ -алгебру, содержащую все открытые множества из $X \times Y$, и, следовательно, все борелевские множества.

Если теперь f — борелевская функция на $X \times Y$, то все ее частные функции f_x борелевские. В самом деле, если \mathcal{O} — некоторое открытое множество из F , то $f_x^{-1}(\mathcal{O}) = (f^{-1}(\mathcal{O})) (x)$ является борелевским множеством как сечение, проходящее через точку x борелевского множества $f^{-1}(\mathcal{O})$.

Если теперь f $(\mu \otimes \nu)$ -измерима, то существует борелевская функция g , равная всюду функции f , кроме точек некоторого множества A $(\mu \otimes \nu)$ -меры нуль (см. замечание после теоремы 23₂). Сечение $A(x)$ для μ -почти всех значений x имеет ν -меру, равную нулю (следствие 3), и f_x совпадает с борелевской функцией g_x всюду, кроме множества $A(x)$, а, значит, эта функция ν -измерима.

Случай, когда интегрируемая функция является произведением функции от x и функции от y

Теорема 79. Пусть μ и ν — меры ≥ 0 на локально компактных пространствах, счетных в бесконечности. Пусть \vec{f} — функция, определенная на X , со значениями в банаховом пространстве \vec{F} , а \vec{g} — функция, определенная на Y , со значениями в банаховом пространстве \vec{G} , и пусть B — непрерывное билинейное отображение $\vec{F} \times \vec{G}$ в бацахово пространство \vec{H} .

Если \vec{f} μ -интегрируема (соответственно μ -измерима) и \vec{g} ν -интегрируема (соответственно ν -измерима), то функция $B(\vec{f}, \vec{g}): (x, y) \rightarrow B(\vec{f}(x), \vec{g}(y))$ является $(\mu \otimes \nu)$ -интегрируемой (соответственно $(\mu \otimes \nu)$ -измеримой)¹⁾ и имеет место формула

$$\int \int B(\vec{f}, \vec{g}) d(\mu \otimes \nu) = B \left(\int \vec{f} d\mu, \int \vec{g} d\nu \right). \quad (\text{IV, 8; 36})$$

1) В противоположность тому, что было сделано в теоремах Фубини, для получения измеримости по отношению к $\mu \otimes \nu$ здесь исходят из измеримости относительно μ и ν .

Обратно, если \vec{F} , \vec{G} и \vec{H} — скалярные поля, B — обычное произведение, а функция $\int g$ ($\mu \otimes v$)-интегрируема и ни одна из двух функций f и g не равна почти всюду нулю, то функция f μ -интегрируема, а функция g v -интегрируема.

Прежде чем проводить доказательство, заметим, что если $\vec{F} = \vec{G} = \vec{H}$ является скалярным полем и B есть обычное произведение, то (IV, 8; 36) записывается в виде

$$\int \int f(x) g(y) d\mu(x) dv(y) = \left(\int f(x) d\mu(x) \right) \left(\int g(y) dv(y) \right). \quad (\text{IV}, 8; 37)$$

З Этот результат непосредственно из определения (IV, 8; 4) меры $\mu \otimes v$ не вытекает, ибо f и g не являются непрерывными функциями с компактным носителем.

Доказательство. Докажем сначала обратное утверждение. Предположим, что fg интегрируема относительно $\mu \otimes v$. Тогда из теоремы Фубини — Лебега следует, что для μ -почти всех значений x функция $y \rightarrow f(x)g(y)$ является v -интегрируемой. Поскольку мы предположили, что f не является почти всюду равной нулю, найдется хотя бы одно значение x , для которого функция $y \rightarrow f(x)g(y)$ будет v -интегрируемой и $f(x) \neq 0$. Отсюда следует, что функция g v -интегрируема. Точно так же можно показать, что функция f μ -интегрируема¹⁾.

Для доказательства прямого утверждения нам понадобится

Лемма. Если часть A множества X имеет нулевую меру относительно μ , то множество $A \times Y$ имеет меру, равную нулю, относительно $\mu \otimes v$.

Эта лемма важна потому, что при $\mu(A) = 0$ мера $v(Y)$ может, вообще говоря, быть бесконечной. Мы уже видели, что в плоскости \mathbb{R}^2 прямая, параллельная одной из координатных осей, имеет нулевую меру по мере $dxdy$.

Доказательство. Зададим $\varepsilon > 0$. Пусть K — некоторый компакт пространства Y . Он имеет конечную меру и существует такое открытое множество $\mathcal{O}' \supset K$ из Y , что $v(\mathcal{O}') \leq \leq v(K) + 1$. Тогда найдется такое открытое множество $\mathcal{O} \supset A$ из X , что $\mu(\mathcal{O}) \leq \varepsilon / (v(K) + 1)$. Так как множество $\mathcal{O} \times \mathcal{O}'$ открыто в $X \times Y$, то оно $(\mu \otimes v)$ -измеримо, и мы можем к его характеристической функции применить формулу (IV, 8; 33), которая даст равенство

$$(\mu \otimes v)(\mathcal{O} \times \mathcal{O}') = \mu(\mathcal{O}) v(\mathcal{O}') \leq \varepsilon. \quad (\text{IV}, 8; 38)$$

¹⁾ За исключением рассмотренного частного случая этот результат, вообще говоря, не верен. Он, например, очевидным образом не верен для $B = 0$. Точно так же, если $f \equiv 0$, то функция $f(x)g(y) \equiv 0$ и, следовательно, $(\mu \otimes v)$ -интегрируема, хотя отсюда вывести v -интегрируемость функции g невозможно.

Значит, $(\mu \otimes v)^*(A \times K) \leq \varepsilon$. Так как ε произвольно, то $(\mu \otimes v)(A \times K) = 0$. Однако Y является объединением возрастающей последовательности компактов K_n ; поэтому $A \times Y$, как объединение последовательности множеств $A \times K_n$ ($\mu \otimes v$)-меры, равной нулю, имеет $(\mu \otimes v)$ -меру, равную нулю.

Окончание доказательства прямого утверждения теоремы

Сначала докажем, что функция $B(\vec{f}, \vec{g})$ является $(\mu \otimes v)$ -измеримой.

Поскольку \vec{f} (соответственно \vec{g}) измерима, то, согласно теореме 32, для \vec{f} (соответственно \vec{g}) найдется аппроксимирующая последовательность, образованная непрерывными функциями, из которой, в силу теоремы 38, можно извлечь подпоследовательность \vec{f}_n (соответственно \vec{g}_n), μ (соответственно v)-почти всюду сходящуюся к \vec{f} , т. е. сходящуюся к \vec{f} (соответственно \vec{g}) на множестве A (соответственно B), дополнение которого имеет μ -меру (соответственно v -меру), равную нулю. Функция $B(\vec{f}_n, \vec{g}_n)$ непрерывна и сходится к $B(\vec{f}, \vec{g})$ на множестве $A \times B$, дополнение которого $C(A \times B) = (CA \times Y) \cup (X \times CB)$, согласно лемме, имеет меру нуль. Значит, по теореме 23, функция $B(\vec{f}, \vec{g})$ измерима. Для того чтобы доказать ее интегрируемость, достаточно показать, что верхний интеграл от ее нормы конечен. Этот интеграл допускает оценку

$$\int \int^* \|B(\vec{f}, \vec{g})\| d(\mu \otimes v) \leq \|B\| \int \int^* \|\vec{f}(x)\| \|\vec{g}(y)\| d\mu(x) d\nu(y). \quad (\text{IV}, 8; 39)$$

Аналогично, функция $(x, y) \rightarrow \|\vec{f}(x)\| \|\vec{g}(y)\|$ ($\mu \otimes v$)-измерима, и, поскольку она ≥ 0 , к ней можно применить теорему Фубини — Фату. Эта теорема (при $\int^* = \int$) дает

$$\int \int \|\vec{f}(x)\| \|\vec{g}(y)\| d\mu(x) d\nu(y) = \int d\mu(x) \int \|\vec{f}(x)\| \|\vec{g}(y)\| d\nu(y). \quad (\text{IV}, 8; 40)$$

Поскольку \vec{g} по условию v -интегрируема, последний интеграл конечен и равен $\|\vec{f}(x)\| \int \|\vec{g}(y)\| d\nu(y)$. Следовательно, искомый верхний интеграл равен $\left(\int \|\vec{f}\| d\mu \right) \left(\int \|\vec{g}\| d\nu \right) < +\infty$ и, значит, функция $B(\vec{f}, \vec{g})$ является $(\mu \otimes v)$ -интегрируемой.

К ней можно теперь применить теорему Фубини — Лебега и записать, что

$$\int \int B(\vec{f}, \vec{g}) d(\mu \otimes \nu) = \int d\mu(x) \int B(\vec{f}(x), \vec{g}(y)) d\nu(y). \quad (\text{IV, 8; 41})$$

Функция \vec{g} ν -интегрируема и для фиксированного x отображение $\vec{g} \rightarrow B(\vec{f}(x), \vec{g})$ из \vec{G} в \vec{H} линейно и непрерывно. Теорема 45 (о перестановочности интеграла и линейно непрерывного отображения) дает равенство

$$\int B(\vec{f}(x), \vec{g}(y)) d\nu(y) = B\left(\vec{f}(x), \int \vec{g} d\nu\right), \quad (\text{VI, 8; 42})$$

откуда

$$\int \int B(\vec{f}, \vec{g}) d(\mu \otimes \nu) = \int B\left(\vec{f}(x), \int \vec{g} d\nu\right) d\mu(x). \quad (\text{IV, 8; 43})$$

Так как \vec{f} μ -интегрируема, а $\vec{f} \rightarrow B\left(\vec{f}, \int \vec{g} d\nu\right)$ является линейным непрерывным отображением \vec{F} в \vec{H} , то та же теорема 45 дает

$$\int B\left(\vec{f}(x), \int \vec{g} d\nu\right) d\mu(x) = B\left(\int \vec{f} d\mu, \int \vec{g} d\nu\right), \quad (\text{IV, 8; 44})$$

чём и заканчивается доказательство теоремы.

Замечание. Если $\vec{F} = \vec{G} = \vec{H}$ — скалярные поля и B представляет собой обычное произведение, то конец доказательства упрощается. Достаточно доказать, что функция $(x, y) \rightarrow f(x)g(y)$ интегрируема, и тогда сразу можно будет написать равенства

$$\begin{aligned} \int \int f(x)g(y) d\mu(x) d\nu(y) &= \int d\mu(x) \int f(x)g(y) d\nu(y) = \\ &= \int f(x) d\mu(x) \int g(y) d\nu(y). \quad (\text{IV, 8; 44}_2) \end{aligned}$$

Следствие 1. Пусть A и B — части множеств X и Y соответственно. Если они соответственно μ -измеримы и ν -измеримы и имеют конечные или бесконечные меры, то множество $A \times B$ является $(\mu \otimes \nu)$ -измеримым, а его мера равна произведению мер

$$(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B) \leqslant +\infty, \quad (\text{IV, 8; 45})$$

где произведение считается равным нулю, если один из множителей равен нулю, даже если другой равен $+\infty$.

В частности, $(\mu \otimes \nu)(X \times Y) = \mu(X)\nu(Y)$.

Обратно, если множество $A \times B$ измеримо относительно $(\mu \otimes v)$ и ни одно из этих двух множеств не имеет нулевой меры, то A μ -измеримо, а B v -измеримо.

Следствие 2. Пусть μ и v — меры ≥ 0 на X и Y , и пусть p и q — комплексные функции, определенные соответственно на X и Y и такие, что функция p локально μ -интегрируема, а функция q локально v -интегрируема. Тогда функция $p \otimes q$: $(x, y) \rightarrow p(x)q(y)$ локально $(\mu \otimes v)$ -интегрируема и мера $(p \otimes q)(\mu \otimes v)$ совпадает с мерой $p\mu \otimes qv$.

Локальная $(\mu \otimes v)$ -интегрируемость вытекает из теоремы. Равенство произведения тензорных произведений $(p \otimes q)(\mu \otimes v)$ и тензорного произведения двух произведений $p\mu \otimes qv$ вытекает из того, что эти две меры принимают одно и то же значение на каждой функции φ из $\mathcal{C}(X \times Y)$ вида $u \otimes v$, где $u \in \mathcal{C}(X)$ и $v \in \mathcal{C}(Y)$, а именно значение $\left(\int p u d\mu \right) \left(\int q v d\nu \right)$ (см. примечание на стр. 661).

Обобщение на произвольные кратные интегралы

Пусть X , Y , Z — три локально компактных пространства, снабженных скалярными мерами λ , μ , v . Тогда с помощью теоремы, аналогичной теореме 73, можно сразу определить меру на $X \times Y \times Z$ как тензорное произведение мер $\lambda \otimes \mu \otimes v$. На этот раз для $u \in \mathcal{C}(X)$, $v \in \mathcal{C}(Y)$, $w \in \mathcal{C}(Z)$ следует положить

$$(\lambda \otimes \mu \otimes v)(u \otimes v \otimes w) = \lambda(u)\mu(v)v(w). \quad (\text{IV}, 8; 46)$$

Интегралы относительно $\lambda \otimes \mu \otimes v$ тройные. Точно так же определяются интегралы произвольной конечной кратности.

Можно также определить сначала меру $\lambda \otimes \mu$ на $X \times Y$, затем меру $(\lambda \otimes \mu) \otimes v$ на множестве $(X \times Y) \times Z$, совпадающем с множеством $X \times Y \times Z$. С помощью формул (IV, 8; 46) и (IV, 8; 4) легко убедиться, что обе меры $\lambda \otimes \mu \otimes v$ и $(\lambda \otimes \mu) \otimes v$ совпадают на $X \times Y \times Z$. Тем самым доказана ассоциативность тензорного произведения. Естественно, что метод вычислений, изложенный в теореме 75, сохраняет свою силу и при λ , μ , $v \geq 0$ имеет место теорема Фубини, обобщающая теоремы 77 и 78. Пусть, например, надо вычислить интеграл

$$\int \int \int \vec{f}(x, y, z) d\lambda(x) d\mu(y) dv(z), \quad (\text{IV}, 8; 47)$$

если заранее известно, что функция \vec{f} интегрируема относительно $\lambda \otimes \mu \otimes v$ или что она измерима и ≥ 0 .

Можно было бы, например, сначала вычислить двойной интеграл, а затем простой интеграл по формуле

$$\int d\mu(y) \int \int \vec{f}(x, y, z) d\lambda(x) dv(z), \quad (\text{IV}, 8; 48)$$

где двойной интеграл имеет смысл для μ -почти всех значений y .

Можно было бы сначала вычислить простой интеграл, а затем двойной по формуле

$$\int \int d\mu(y) dv(z) \int \vec{f}(x, y, z) d\lambda(x), \quad (\text{IV}, 8; 49)$$

где простой интеграл имеет смысл для $(\mu \otimes v)$ -почти всех значений (y, z) .

Можно было бы вычислить последовательно три простых интеграла по формуле

$$\int d\mu(y) \int d\lambda(x) \int \vec{f}(x, y, z) dv(z), \quad (\text{IV}, 8; 50)$$

где интеграл по $dv(z)$ имеет смысл для $(\lambda \otimes \mu)$ -почти всех значений (x, y) , следующий интеграл по $d\lambda(x)$ имеет смысл для μ -почти всех значений y .

Все три полученных интеграла совпадают.

Этот метод часто применяется для вычисления объемов.

Предположим, например, что все три пространства X , Y , Z представляют собой вещественную прямую \mathbb{R} и что $d\lambda = d\mu = dv$ являются мерами Лебега. Тогда мера-произведение $dx dy dz$ является мерой объемов в пространстве \mathbb{R}^3 .

Пусть теперь A — измеримая часть пространства \mathbb{R}^3 ; предположим, что нам надо вычислить ее объем, т. е. интеграл $\int_A dx \otimes dy \otimes dz$. Согласно формуле (IV, 8; 35), в рамках теоремы Фубини — Фату этот интеграл может быть вычислен, например, по формуле

$$\int S(z) dz \leq +\infty, \quad (\text{IV}, 8; 51)$$

где через $S(z)$ обозначена площадь сечения $A(z)$ множества A плоскостью, перпендикулярной оси z , т. е. двойной интеграл

$$\int \int_{A(z)} dx dy \leq +\infty. \quad (\text{IV}, 8; 52)$$

Эта площадь может быть также вычислена с помощью формулы

$$\int \int l(x, y) dx dy \leq +\infty, \quad (\text{IV}, 8; 53)$$

где через $l(x, y)$ обозначена «длина» сечения множества A прямой, параллельной оси z с координатами x, y , т. е. интеграл

$$\int\limits_{A(x, y)} dz \leqslant +\infty. \quad (\text{IV}, 8; 54)$$

Пусть рассматривается пространство \mathbb{R}^n , являющееся произведением n экземпляров пространства \mathbb{R} . Если через dx_1, dx_2, \dots, dx_n обозначены канонические меры (то есть меры Лебега) на пространствах-сомножителях¹⁾, то $dx_1 \otimes dx_2 \otimes \dots \otimes dx_n$ будет мерой, определенной на \mathbb{R}^n . Эту меру называют *канонической*, или *мерой Лебега*, на \mathbb{R}^n и обозначают через dx или $dx_1 dx_2 \dots dx_n$.

Широкая сходимость тензорных произведений

Теорема 80. *Если при n , стремящемся к бесконечности, последовательность мер μ_n на X широко сходится к мере μ , а последовательность мер v_n на Y широко сходится к мере v , то последовательность мер $\mu_n \otimes v_n$ при n , стремящемся к бесконечности, широко сходится к мере $\mu \otimes v$.*

Доказательство. Пусть φ — некоторый элемент пространства $\mathcal{C}(X \times Y)$. Воспользуемся методом вычислений, изложенным в теореме 75. Величины $\int \varphi(x, y) d\mu_n(y)$ сходятся к $\int \varphi(x, y) dv(y)$ для всех x , в силу предположения о широкой сходимости.

Кроме того, согласно следствию 3 теоремы 66, эта сходимость равномерна относительно точек x , принадлежащих некоторому компакту. Если через $H \times K$ обозначить произведение компактов, содержащее носитель функции φ , то все эти интегралы будут равны нулю для $x \notin H$. Иначе говоря, функции ψ_n , определенные в доказательстве теоремы 75, при n , стремящемся к бесконечности, сходятся к функции ψ равномерно на H с сохранением носителей в H . Теперь, учитывая следствие 2 теоремы 65, мы получаем, что $\mu_n(\psi_n)$ сходятся к $\mu(\psi)$, чем и заканчивается доказательство теоремы.

Следствие. *Существует такая последовательность P_l полиномов $\geqslant 0$ на пространстве \mathbb{R}^n , что меры $P_l dx$ (где $dx = dx_1 \otimes dx_2 \otimes \dots \otimes dx_n$) широко сходятся к δ при l , стремящемся к $+\infty$.*

¹⁾ Обозначение не корректно (см. примечание на стр. 443). Все пространства-сомножители одни и те же. Следовательно, это всюду одна и та же мера! Нужно было написать $1(x_1)dx_1, 1(x_2)dx_2, \dots, 1(x_n)dx_n$, что и в самом деле дает одну и ту же меру.

Доказательство. В теореме 68 мы видели, что существует такая последовательность полиномов $Q_l \geq 0$ на \mathbb{R} , для которой меры $Q_l(x)dx$ при l , стремящемся к бесконечности, широко сходятся к мере $\delta = d\delta(x)$. Теперь полиномы $P_l(x_1, x_2, \dots, x_n) = Q_l(x_1)Q_l(x_2) \dots Q_l(x_n)$ неотрицательны и, по теореме 80, мера $Q_l(x_1)dx_1 \otimes Q_l(x_2)dx_2 \otimes \dots \otimes Q_l(x_n)dx_n$ при l , стремящемся к бесконечности, стремится к мере $d\delta(x_1) \otimes \dots \otimes d\delta(x_2) \otimes \dots \otimes d\delta(x_n) = d\delta(x_1, x_2, \dots, x_n)$.

Теорема 80₂ (теорема Вейерштрасса о полиномиальной аппроксимации). *Пусть K — компакт пространства \mathbb{R}^n и \vec{f} — непрерывная функция на K со значениями в банаевом пространстве \vec{F} . Тогда, каким бы ни было число $\varepsilon > 0$, можно найти такой полином \vec{Q} , определенный на \mathbb{R}^n , со значениями в \vec{F} , что $\max_{x \in K} \|\vec{f}(x) - \vec{Q}(x)\| \leq \varepsilon$. Или иначе: существует такая последовательность полиномов \vec{Q}_k , которая сходится к функции \vec{f} равномерно на K при k , стремящемся к бесконечности.*

Доказательство. Формула $\vec{g} = \sum_{i=1}^N \vec{f}_N(a_i) a_i = \sum_{i=1}^N \vec{g}_i a_i$, $a_i \in C(\mathbb{R}^n)$, использованная в следствии 8 из теоремы 11 (о разложении единицы), дает сначала некоторую непрерывную разложимую функцию \vec{g} на \mathbb{R}^n с компактным носителем и такую, что $\|\vec{f} - \vec{g}\| \leq \varepsilon/2$ на K ¹).

Пусть теперь P_l — последовательность таких полиномов, что их меры, согласно предыдущему следствию, широко сходятся к δ . Положим

$$\begin{aligned} \vec{Q}_l(x) &= \int \int \dots \int_{\mathbb{R}^n} \vec{g}(x-t) P_l(t) dt = \int_{\mathbb{R}^n} P_l(x-t) \vec{g}(t) dt = \\ &= \int \int \dots \int_{\mathbb{R}^n} \vec{g}(x_1-t_1, x_2-t_2, \dots, x_n-t_n) \times \\ &\quad \times P_l(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n. \quad (\text{IV}, 8; 55) \end{aligned}$$

Все рассматриваемые величины имеют смысл, поскольку они представляют собой интегралы от непрерывных функций с компактным носителем.

¹) В условии следствия 8 функция \vec{f} предполагалась определенной на всем X (здесь \mathbb{R}^n) и имела компактный носитель. Но это не имело смысла, так как рассматриваются лишь значения \vec{f} на K .

Докажем, что последовательность величин \vec{Q}_l равномерно сходится к \vec{g} на K и что они представляют собой некоторые полиномы. Тогда среди них можно будет найти один такой полином \vec{Q}_l , что $\|\vec{Q}_l - \vec{g}\| \leq \varepsilon/2$ на K , и теорема будет доказанной.

Прежде всего, имеем: $P_l = \sum_p c_p x^p$ (в обозначениях, принятых на стр. 266, $p = (p_1, p_2, \dots, p_n) \in \mathbb{N}^n$ и $x^p = x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}$, для простоты мы убрали стрелки). Отсюда

$$(x-t)^p = \sum_{q \leq p} \binom{p}{q} x^q (-t)^{p-q},$$

$$P_l(x-t) = \sum_{\substack{p, q \\ q \leq p}} c_p \binom{p}{q} x^q (-t)^{p-q}, \quad (\text{IV, 8; 56})$$

$$\vec{Q}_l(x) = \sum_{\substack{p, q \\ q \leq p}} c_p \binom{p}{q} x^q \int_{\mathbb{R}^n} (-t)^{p-q} \vec{g}(t) dt = \sum_q d_q x^q,$$

а, следовательно, \vec{Q}_l является полиномом.

Докажем теперь равномерную сходимость. Так как \vec{g} разложима: $\vec{g} = \sum_{i=1}^N \vec{g}_i a_i$, то вопрос непосредственно сводится к доказательству того, что каждый интеграл $\int_{\mathbb{R}^n} a_i(x-t) P_l(t) dt$ сходится к $a_i(x) = \int_{\mathbb{R}^n} a_i(x-t) d\delta(t)$ равномерно относительно x в K .

Но $(x, t) \mapsto a_i(x-t)$ является непрерывной функцией на $K \times \mathbb{R}^n$. Когда x пробегает компакт K , частная функция $t \mapsto a_i(x-t)$ сохраняет носитель в некотором фиксированном компакте \mathbb{R}^n , а именно в векторной сумме $K - T_i$, где T_i является носителем a_i . Мы оказываемся в условиях применимости следствия 3 теоремы 66 с изменением ролей переменных x и t , чем и заканчивается доказательство теоремы.

Замечание. Существуют другие методы доказательства теоремы Вейерштрасса, но это доказательство позволяет получить некоторые гораздо более сильные результаты.

Предположим, что \vec{f} определена и непрерывна не только на K , но и на всем \mathbb{R}^n . Тогда для построения функции \vec{g} с компактным носителем, близкой к \vec{f} на K , имеется значительно более простой способ: возьмем $\vec{g} = \alpha \vec{f}$, где α — непрерывная функция, или даже класса C^∞ , с компактным носителем, равная 1 в некоторой окрестности компакта K (следствие 1 теоремы 11).

Тогда мы получим, что \vec{g} равна \vec{f} в некоторой окрестности K . Полиномы \vec{Q}_l сходятся к \vec{g} , а, следовательно, к \vec{f} равномерно на K . Правда, функция \vec{g} не разложима и непосредственно применить следствие 3 теоремы 66 невозможно. Легко убедиться, что доказательство этой теоремы проводилось лишь для скалярных функций в связи с наличием векторных мер. Однако она остается справедливой и для векторных функций в случае скалярных мер.

При этих условиях полиномы Q_l обладают гораздо более сильными свойствами сходимости. Предположим, что на скрестности компакта K функция \vec{f} принадлежит классу C^m . Тогда, если α принадлежит классу C^m и имеет носитель в этой окрестности, то \vec{g} принадлежит классу C^m во всем \mathbb{R}^n . Тогда для каждого $p \leq m$ производные $D^p \vec{Q}_l$ сходятся к $D^p \vec{f}$ равномерно на K . Другими словами, можно осуществить полиномиальную аппроксимацию одновременно и для \vec{f} , и для ее производных. В самом деле, мы имеем простейший случай, когда допустимо дифференцирование под знаком интеграла (см. далее следствие теоремы 115), так что

$$D^p \vec{Q}_l(x) = \int_{\mathbb{R}^n} D^p \vec{g}(x-t) P_l(t) dt, \quad (\text{IV}, 8; 57)$$

и мы пришли к исходной задаче, где вместо \vec{g} стоит $D^p \vec{g}$.

§ 9. ЧАСТНЫЕ СВОЙСТВА МЕР РАДОНА НА ВЕЩЕСТВЕННОЙ ПРЯМОЙ \mathbb{R}

Свойства, которые мы будем рассматривать в этом параграфе, используют тот факт, что вещественная прямая \mathbb{R} не только является локально компактным пространством, но и образует вполне упорядоченное множество. Если a и b — две точки пространства \mathbb{R} , то через $[a, b]$ мы будем обозначать интервал $a \leq x \leq b$ при $a < b$ и интервал $b < x \leq a$ при $b < a$. Точно такой же смысл будут иметь обозначения $]a, b]$, $[a, b]$, $]a, b[$. Следовательно, это будут скорее интервалы в смысле определения, данного на стр. 26, чем интервалы в смысле определения, данного на стр. 183.

Введение символа $\int_a^b d\mu$

Определение. Пусть $\overset{\rightarrow}{\mu}$ — мера Радона на интервале \mathbb{R}_1 (открытом, полуоткрытом или замкнутом) вещественной пря-

мой \mathbb{R} со значениями в банаховом пространстве \vec{E} . Предположим, что эта мера имеет в качестве базы вещественную меру $\geqslant 0$. Тогда ее можно, в частности, продолжить на характеристические функции компактных интервалов.

Обозначим, если a и b принадлежат \mathbb{R}_1 , через $\int_a^b d\vec{\mu}$ величину, определенную по формуле

$$\int_a^b d\vec{\mu} = \begin{cases} \int_{[a, b]} d\vec{\mu}, & \text{если } a < b, \\ \vec{0}, & \text{если } a = b, \\ - \int_{[b, a]} d\vec{\mu}, & \text{если } b < a. \end{cases} \quad (\text{IV, 9; 1})$$

В частности, если мера $d\vec{\mu}$ записана в виде $\vec{p}(x)dx$, где \vec{p} — локально интегрируема относительно меры dx , то по определению имеем

$$\int_a^b d\vec{\mu} = \begin{cases} \int_{[a, b]} \vec{p}(x) dx, & \text{если } a \leqslant b, \\ - \int_{[b, a]} \vec{p}(x) dx, & \text{если } a \geqslant b. \end{cases} \quad (\text{IV, 9; 2})$$

Теорема 81. Имеет место соотношение Шалля:

$$\int_a^b d\vec{\mu} + \int_b^c d\vec{\mu} + \int_c^a d\vec{\mu} = \vec{0}. \quad (\text{IV, 9; 3})$$

Доказательство очевидно. Достаточно рассмотреть все возможные случаи.

Неопределенные интегралы

Определение. Говорят, что функция M , определенная на \mathbb{R}_1 , со значениями в аффинном нормированном полном пространстве E , является *неопределенным интегралом от меры Радона* $\vec{\mu}$, заданной на \mathbb{R}_1 со значениями в \vec{E} и имеющей вещественную базу $\geqslant 0$, если, каковы бы ни были a и b из \mathbb{R}_1 , имеет

¹⁾ Будьте внимательны: мера $\vec{\mu}$, возможно, имеет точечные массы, и тогда небезразлично, что брать — полуинтервалы, включающие левый конец, или полуинтервалы, включающие правый конец.

место равенство

$$\overrightarrow{M(b) - M(a)} = \int_a^b d\vec{\mu}. \quad (\text{IV}, 9; 4)$$

Если $\vec{\mu}$ имеет вид $d\vec{\mu} = \vec{p}d\mu_0$, где μ_0 — некоторая вещественная мера $\geqslant 0$ и \vec{p} — локально μ_0 -интегрируемая функция, то говорят, что M есть *неопределенный интеграл от функции p по мере μ_0* .

Теорема 82. Пусть E — полное аффинное нормированное пространство, $\vec{\mu}$ — некоторая мера Радона на \mathbb{R}_1 со значениями в \vec{E} и базой $\geqslant 0$. Тогда мера $\vec{\mu}$ имеет бесчисленное множество неопределенных интегралов со значениями в \vec{E} . Если c — точка \mathbb{R}_1 , то функция $\vec{M}: x \rightarrow \int_c^x d\vec{\mu}$ является неопределенным интегралом со значениями в \vec{E} , а, добавляя к ней произвольные постоянные элементы из E , можно получить все неопределенные интегралы со значениями в E .

Доказательство. То, что функция M обладает указанным свойством, вытекает из соотношения Шаля. Кроме того, если две функции со значениями соответственно в \vec{E} и E являются неопределенными интегралами от меры $\vec{\mu}$, то из формулы (IV, 9; 4) следует, что их разность принимает одно и тоже значение в произвольных точках a и b пространства \mathbb{R}_1 , т. е. является постоянной на \mathbb{R}_1 . Обратно, сумма какого-либо неопределенного интеграла со значениями в \vec{E} и некоторой постоянной из E является неопределенным интегралом со значениями в E .

Пример. Рассмотрим функцию Хевисайда — вещественную функцию Y на \mathbb{R} , определенную по формуле

$$Y(x) = \begin{cases} 0 & \text{для } x \leqslant 0, \\ 1 & \text{для } x > 0. \end{cases} \quad (\text{IV}, 9; 5)$$

Эта функция является неопределенным интегралом от меры δ : $Y(x) = \int_{-\infty}^x d\delta$ или $\int_a^x d\delta$, где $a \leqslant 0$ — произвольное число.

Замечание. Можно также показать, что разность двух неопределенных интегралов со значениями в E является неко-

торой постоянной в \vec{E} и что сумма неопределенного интеграла со значениями в E и постоянной из \vec{E} является неопределенным интегралом в E .

Теорема 83. Каждая функция M со значениями в E , являющаяся неопределенным интегралом от меры Радона μ на \mathbb{R}_1 , со значениями в \vec{E} , непрерывна слева в любой точке \mathbb{R}_1 . Эта функция непрерывна справа в точке с прямой \mathbb{R}_1 тогда и только тогда, когда $\mu(\{c\}) = \vec{0}$. Во всяком случае у нее существует предел справа и ее скачек равен $\mu(c)$.

Доказательство. При $h \geq 0$ разность $\overrightarrow{M(c) - M(c-h)}$ является мерой $\mu([c-h, c])$ или интегралом относительно μ характеристической функции φ_h этого интервала. Пусть $d\mu = \vec{p}d\mu_0$, $\mu_0 \geq 0$. Тогда $\mu([c-h, c]) = \int \varphi_h \vec{p} d\mu_0$. Если $h > 0$ стремится к 0, то характеристическая функция φ_h сходится просто к функции 0, т. е. к характеристической функции пустого множества. Отсюда следует, что функция $\varphi_h \vec{p}$ просто сходится к $\vec{0}$. При $h \leq h_0$ она ограничена по норме функцией $\|\vec{p}\| \varphi_{h_0}$. Так как \vec{p} локально μ_0 -интегрируема, то эта функция ≥ 0 и μ_0 -интегрируема. Согласно теореме Лебега, распространенной на обобщенные пределы (см. замечание на стр. 541), предыдущие разности сходятся к 0 при h , стремящемся к 0, что доказывает непрерывность функции слева. Если мы теперь будем проводить рассуждения для разностей $\overrightarrow{M(c+h) - M(c)} = \int \Psi_h \vec{p} d\mu_0$, где Ψ_h — характеристическая функция интервала $[c, c+h]$, то увидим, что функции Ψ_h сходятся к характеристической функции Ψ_0 множества $\{c\}$, состоящего из одной точки c . Функции $\vec{p} \Psi_h$ будут мажорироваться некоторой интегрируемой фиксированной функцией ≥ 0 . Из теоремы Лебега следует, что эти разности при h , стремящемся к 0, сходятся к $\int \Psi_0 \vec{p} d\mu_0 = \mu(\{c\})$. Отсюда вытекает, что непрерывность справа имеет место тогда и только тогда, когда множество $\{c\}$ имеет нулевую меру относительно μ . Кроме того, функция M всегда имеет предел справа и $M(c+0) = M(c) + \mu(\{c\})$.

Замечание. Пусть a и c — точки \mathbb{R}_1 , $a < c$. Предыдущие рассуждения означают, что $\mu([a, c'])$ стремится к $\mu([a, c])$ при

c' , стремящемся к c по значениям $< c$, и $\mu([a, c'])$ стремится к $\mu([a, c]) = \mu([a, c]) + \mu(\{c\})$, когда c' стремится к c по значениям $> c$. Конечно, $\mu([a, c'])$ имеет пределами соответственно $\mu([a, c])$ и $\mu([a, c])$.

Следствие 1. Если функция \vec{p} определена на \mathbb{R}_1 со значениями в банаховом пространстве \vec{E} и локально интегрируема относительно меры Лебега dx , то каждый неопределенный интеграл $x \rightarrow \int_c^x \vec{p}(t) dt$ является непрерывной функцией x .

В самом деле, относительно меры $\vec{p} dx$ любая точка имеет нулевую меру.

Следствие 2. Неопределенный интеграл от меры на \mathbb{R}_1 является правильной функцией, а, значит, борелевской и, следовательно, измеримой по отношению к любой мере Радона на \mathbb{R}_1 .

Тот факт, что эта функция правильная, вытекает непосредственно из теоремы. Поскольку характеристическая функция интервала борелевская, то борелевской будет ступенчатая функция, а, значит, правильная функция, в чем можно убедиться переходом к пределу (теоремы 8 и 23).

Функции с ограниченной вариацией на прямой

Пусть M — отображение некоторой части \mathbb{R}_1 вещественной прямой \mathbb{R} в метрическое пространство E . Полной вариацией функции M называется точная верхняя грань $V(M) = V(\mathbb{R}_1, M)$ сумм

$$\sum_{\Delta} = \sum_{i=0}^{n-1} d(M(c_i), M(c_{i+1})) \quad (\text{IV}, 9; 6)$$

по всевозможным разбиениям Δ множества \mathbb{R}_1 , т. е. по всевозможным конечным возрастающим последовательностям точек c_0, c_1, \dots, c_n из \mathbb{R}_1 . Полная вариация — это некоторое неотрицательное число, конечное или равное $+\infty$. Если $\mathbb{R}_2 \subset \mathbb{R}_1$, то $V(\mathbb{R}_2, M) \leq V(\mathbb{R}_1, M)$. Если $V(\mathbb{R}_1, M)$ конечна, то говорят, что функция M имеет ограниченную вариацию на \mathbb{R}_1 (надо было бы говорить: имеет ограниченную полную вариацию). Чаще всего рассматриваются случаи, когда \mathbb{R}_1 — или интервал (ограниченный или нет, открытый, полуоткрытый или замкнутый), или множество \mathbb{N} целых чисел ≥ 0 (см. стр. 142), или множество \mathbb{Z} всех целых чисел. С этого момента мы будем предполагать, что \mathbb{R}_1 — интервал.

Функция, удовлетворяющая условию Липшица на ограниченном интервале \mathbb{R}_1 прямой \mathbb{R} , имеет ограниченную вариацию

В самом деле, из неравенства $d(M(x''), M(x')) \leq k|x'' - x'|$ следует, что полная вариация функции M мажорируется числом k , умноженным на длину интервала \mathbb{R}_1 . В частности, функция M на \mathbb{R}_1 со значениями в аффинном нормированном пространстве E , дифференцируемая и имеющая ограниченную производную, на интервале \mathbb{R}_1 конечной длины имеет ограниченную вариацию. Эта функция, в силу следствия 2 из теоремы 13 гл. III, удовлетворяет условию Липшица.

Конечно, если N принимает значения в аффинном нормированном пространстве E , а \vec{M} принимает значения в присоединенном векторном пространстве \vec{E} и k — некоторый скаляр, то

$$\begin{aligned} V(N + \vec{M}) &\leq V(N) + V(\vec{M}), \\ V(k\vec{M}) &= |k|V(\vec{M}). \end{aligned} \quad (\text{IV}, 9; 6_2)$$

Это говорит о том, что функции с ограниченной вариацией на \mathbb{R}_1 со значениями в \vec{E} образуют некоторое векторное пространство. Однако функция $\vec{M} \rightarrow V(\vec{M})$ не является нормой в этом пространстве. В самом деле, равенство $V(\vec{M}) = 0$ эквивалентно тому, что $\vec{M} = \text{const}$, но не обязательно $\vec{M} = \vec{0}$.

Заметим, что при замене последовательности c_i конечной последовательностью, содержащей данную, сумма (IV, 9; 6) не убывает. Следовательно, при вычислении точной верхней грани можно ограничиться рассмотрением таких разбиений Δ , у которых среди c_i фигурирует конечное число заранее заданных точек.

Теорема 83₂. *Если a, b, c — три заданные точки \mathbb{R}_1 , $a < b < c$, то полная вариация функции M в $[a, c]$ является суммой полных вариаций этой функции в $[a, b]$ и $[b, c]$:*

$$V([a, c]; M) = V([a, b]; M) + V([b, c]; M)^1.$$

В самом деле, левая часть может быть вычислена как точная верхняя грань сумм \sum_{Δ} , соответствующих конечным после-

¹⁾ Можно было бы ожидать, что справедливы соотношения вида

$$V([a, c]; M) = V([a, b]; M) + V([b, c]; M).$$

Однако, это, вообще говоря, не верно. Указанное соотношение справедливо, если M непрерывна слева в точке b , ибо тогда (лемма 2, стр. 689): $V([a, b]; M) = V([a, b[; M])$.

довательностям точек $c_0 = a, c_1, c_2, \dots, c_n = c$, среди которых находится точка b .

Теорема 84. *Каждая функция ограниченной вариации на интервале \mathbb{R}_1 прямой \mathbb{R} со значениями в метрическом пространстве E ограничена. Если E полно, то эта функция правильная, а следовательно, борелевская и, значит, она измерима относительно любой меры Радона на \mathbb{R}_1 .*

Если полная вариация функции M конечна, то для x и y из \mathbb{R}_1 она мажорирует число $d(M(x), M(y))$. Следовательно, функция M ограничена. Докажем теперь, что если пространство E полно, то эта функция правильная.

Пусть x_0, x_1, x_2, \dots — некоторая последовательность точек $>c$, стремящаяся к c при n , стремящаяся к бесконечности. Для упрощения доказательства сходимости последовательности $M(x_0), M(x_1), M(x_2), \dots$ к некоторому пределу при неограниченном возрастании n , переставим так члены этой последовательности, чтобы можно было считать последовательность x_n убывающей. Тогда независимо от n сумма $\sum_{i=0}^{n-1} d(M(x_i), M(x_{i+1}))$ ограничена полной вариацией M . Значит, $\sum_{i=0}^{\infty} d(M(x_i), M(x_{i+1})) < +\infty$. Таким образом, для $m \geq n+1$ при m и n , стремящихся к $+\infty$, сумма $\sum_{i=n}^{m-1} d(M(x_i), M(x_{i+1}))$ стремится к 0, и тем более стремится к нулю расстояние $d(M(x_n), M(x_m))$. Это означает, что последовательность $M(x_n)$ является последовательностью Коши. Поскольку E полно, эта последовательность сходится. Ее предел не зависит от выбора последовательности x_n . В самом деле, из двух последовательностей $(x'_n)_{n \in N}, (x''_n)_{n \in N}$ можно образовать смешанную последовательность $x'_0, x''_0, x'_1, x''_1, x'_2, x''_2, \dots$, на которой последовательность значений M должна иметь некоторый предел, а потому последовательности $M(x'_n)$ и $M(x''_n)$ также будут иметь тот же предел. Пусть $M(c+0)$ — общий предел этих последовательностей. Тогда $M(x)$ стремится к $M(c+0)$ при x , стремящемся к c по значениям $>c$. В противном случае существовало бы такое число $\varepsilon > 0$, что для каждого $n \geq 1$ можно было бы найти число $x_n, c < x_n \leq c + 1/n$, для которого выполняется неравенство $d(M(x_n), M(c+0)) > \varepsilon$. Но тогда последовательность x_n стремилась бы к точке c при n , стремящемся к бесконечности по значениям $>c$, а последовательность $M(x_n)$ не стремилась бы к $M(c+0)$, что противоречит ранее полученным результатам.

Точно так же можно показать, что $M(x)$ имеет предел $M(c=0)$, когда x стремится к c по значениям $\prec c$. Значит, M в точках \mathbb{R}_1 правильная.

Замечание. Обратное утверждение не верно. Правильная и даже непрерывная на интервале $[a, b]$ прямой \mathbb{R} функция с вещественными значениями не обязательно имеет ограниченную вариацию.

Рассмотрим, например, следующую вещественную функцию M на $[0, 1]$: в точках с абсциссой $1/n$, $n \geq 1$, $M(1/n) = (-1)^n/n$, в интервалах $[1/(n+1), 1/n]$ она линейна, а в точке $x = 0$ она равна 0. Функция M , очевидно, непрерывна, а ее полная вариация равна $(1 + 1/2) + (1/2 + 1/3) + (1/3 + 1/4) + \dots = +\infty$.

Теорема 84₂. Пусть M — отображение $[a, b]$ ($a < b$) в E ограниченной вариации. Обозначим через $V([a, x]; M)$ полную вариацию сужения функции M на интервал $[a, x]$. Функция $x \rightarrow V([a, x]; M)$ возрастает и непрерывна слева (соответственно справа) в точке c тогда и только тогда, когда M непрерывна слева (соответственно справа) в точке c .

Мы докажем две леммы, из которых вытекает эта теорема.

Лемма 1. Если M — отображение $[a, c[$ в метрическое пространство E , то при x , стремящемся к c по значениям $\prec c$, $V([a, x]; M) \leq +\infty$ стремится к $V([a, c[; M) \leq +\infty$.

Доказательство леммы 1. Рассматриваемая функция возрастающая, а, следовательно, при x , стремящемся к c по значениям $\prec c$, она имеет предел $\leq V([a, c[; M)$. Однако если V является произвольным числом $\prec V([a, c[; M)$, то по определению существует такое разбиение Δ интервала $[a, c[$: $c_0 = a \leq c_1 \leq c_2 \leq \dots \leq c_n < c$, что

$$\sum_{\Delta} \geq V. \quad (\text{IV}, 9; 7)$$

Тогда тем более для $c_n \leq x < c$

$$V([a, x]; M) \geq V([a, c_n]; M) \geq \sum_{\Delta} \geq V, \quad (\text{IV}, 9; 8)$$

что и требовалось доказать.

Лемма 2. Если M является отображением $[a, c]$ ($a < c$) в метрическое пространство E с ограниченной вариацией, то $V([a, c]; M) = V([a, c[; M)$ тогда и только тогда, когда M непрерывна слева в точке c ¹⁾.

¹⁾ Если M не является функцией с ограниченной вариацией на $[a, c]$, то она не имеет ограниченной вариации на $[a, c[$ и выполняется равенство $V([a, c]; M) = V([a, c[; M) = +\infty$ независимо от того, является ли функция M непрерывной слева в точке c или нет.

Доказательство леммы 2. Легко видеть, что $V([a, c]; M) \leq V([a, c]; M)$. Предположим сначала, что M непрерывна слева в точке c . Тогда при заданном $\varepsilon > 0$ существует такое $\eta > 0$, что из $c - x \leq \eta, x < c$, следует неравенство $d(M(c), M(x)) \leq \varepsilon/2$. С другой стороны, существует такое разбиение Δ : $c_0 = a, c_1, c_2, \dots, c_n = c$ отрезка $[a, c]$, что

$$\sum_{\Delta} \geq V([a, c]; M) - \frac{\varepsilon}{2}. \quad (\text{IV}, 9; 9)$$

Можно предполагать, кроме того, что $c_{n-1} \geq c - \eta$ (в противном случае можно пополнить Δ точкой $c - \eta$, отчего сумма \sum_{Δ} может лишь возрасти). Тогда

$$\begin{aligned} V([a, c]; M) &\geq V([a, c_{n-1}]; M) \geq \sum_{\Delta} - d(M(c), M(c_{n-1})) \geq \\ &\geq \left(V([a, c]; M) - \frac{\varepsilon}{2} \right) - \frac{\varepsilon}{2} = V([a, c]; M) - \varepsilon. \end{aligned} \quad (\text{IV}, 9; 10)$$

Так как ε произвольно, то $V([a, c]; M) = V([a, c]; M)$.

Обратно, предположим, что последнее равенство имеет место. Тогда существует такое разбиение Δ интервала $[a, c]$, при котором имеет место соотношение (IV, 9; 7), где $V = V([a, c], M) - \varepsilon$. Поэтому для $c_n \leq x < c$ будем иметь

$$V([a, c]; M) - \varepsilon \leq \sum_{\Delta} \leq \sum_{\Delta} + d(M(x), M(c)) \leq V([a, c]; M), \quad (\text{IV}, 9; 11)$$

откуда

$$d(M(x), M(c)) \leq \varepsilon \quad \text{для } c_n \leq x < c, \quad (\text{IV}, 9; 12)$$

а, значит, функция M непрерывна слева в точке c .

Доказательство теоремы. Мы только что доказали, что функция $x \rightarrow V([a, x]; M)$ непрерывна слева в точке c тогда и только тогда, когда функция M непрерывна слева в этой точке. Мы даже нашли в обоих случаях предел $V([a, x]; M)$, когда x стремится к c по значениям $< c$.

Если теперь x стремится к c по значениям $> c$, то $V([x, b]; M)$, как показывает рассуждение, подобное рассуждению, использованному в лемме 1, стремится к $V([c, b]; M)$. Рассуждая так же, как в лемме 2, можно убедиться, что последнее выражение равно $V([c, b]; M)$ тогда и только тогда, когда функция M непрерывна справа в точке c . Далее, $V([a, x]; M) = V([a, b]; M) - V([x, b]; M)$ стремится к $V([a, c]; M) = V([a, b]; M) - V([c, b]; M)$ тогда и только тогда, когда M непрерывна справа в точке c .

Если R_1 — замкнутый интервал $[a, b]$ и если M непрерывна слева¹⁾, то, вместо того чтобы выражать полную вариацию M

¹⁾ Или, естественно, непрерывна справа.

как точную верхнюю грань, ее можно выразить как предел сумм (IV, 9; 6), соответствующих произвольной последовательности $\Delta_0, \Delta_1, \Delta_2, \dots$ разбиений $[a, b]$ с наибольшей длиной интервалов разбиения, стремящейся к нулю, при n , стремящемся к бесконечности:

Теорема 85. Пусть M — отображение замкнутого интервала $[a, b]$ ($a < b$) прямой \mathbb{R} в метрическое пространство E , всюду непрерывное слева. Тогда, каким бы ни было число V_1 , строго меньшее полной вариации $V(M)$ функции M в $[a, b]$, найдется число $\eta > 0$, обладающее следующим свойством: какова бы ни была конечная возрастающая последовательность $c_0 = a \leq c_1 \leq c_2 \leq \dots \leq c_n = b$ точек $[a, b]$, такая, что длины всех интервалов $[c_i, c_{i+1}]$ не превосходят η , имеет место неравенство:

$$\sum_{i=0}^{n-1} d(M(c_i), M(c_{i+1})) \geq V_1. \quad (\text{IV, 9; 13})$$

Доказательство. Пусть V_2 — такое число, что $V_1 < V_2 < V(M)$. В силу определения полной вариации как точной верхней грани, существует такое разбиение Δ_0 : $d_0 = a, d_1, d_2, \dots, d_N = b$, что

$$\sum_{\Delta_0} = \sum_{j=0}^{N-1} d(M(d_j), M(d_{j+1})) \geq V_2. \quad (\text{IV, 9; 14})$$

Выберем теперь η настолько малым, чтобы:

1) $\eta < d_{j+1} - d_j$ для всех j . Тогда, если $c_{i+1} - c_i \leq \eta$, то интервал $[c_i, c_{i+1}]$ содержит не более одной точки d_j .

2) Из $d(x, d_j) \leq \eta$, $x < d_j$, следует, что $d(M(x), M(d_j)) \leq (V_2 - V_1)/2(N - 1)$ для всех $j = 1, 2, \dots, N - 1$.

Такой выбор η возможен, так как M непрерывна слева в каждой точке d_j .

Покажем, что такое число η решает поставленную задачу. Пусть Δ : $c_0 = a, c_1, \dots, c_n = b$ — разбиение $[a, b]$ на части, длины которых не превосходят η . Если к c_i добавить точки d_j , то мы получим новое разбиение Δ' , дающее сумму $\sum_{\Delta'}$, не меньшую, чем \sum_{Δ} , т. е. V_2 . Но тогда

$$\begin{aligned} \sum_{\Delta} = \sum_{\Delta'} - \sum_{i, j} (d(M(c_i), M(d_j)) + \\ + d(M(d_j), M(c_{i+1})) - d(M(c_i), M(c_{i+1}))), \end{aligned} \quad (\text{IV, 9; 15})$$

при этом последняя сумма распространяется на все системы (c_i, d_j, c_{i+1}) , такие, что $c_i < d_j < c_{i+1}$.

Имеет место оценка (неравенство треугольника):

$$d(M(d_j), M(c_{i+1})) - d(M(c_i), M(c_{i+1})) \leq d(M(c_i), M(d_j)), \quad (\text{IV, 9; 16})$$

откуда

$$\begin{aligned} d(M(c_i), M(d_j)) + d(M(d_j), M(c_{i+1})) - d(M(c_i), M(c_{i+1})) &\leqslant \\ &\leqslant 2d(M(c_i), M(d_j)) \leqslant 2 \frac{V_2 - V_1}{2(N-1)}. \quad (\text{IV}, 9; 17) \end{aligned}$$

Поэтому

$$\sum_{\Delta} \geqslant \sum_{\Delta'} - 2(N-1) \frac{V_2 - V_1}{2(N-1)} = \sum_{\Delta'} -(V_2 - V_1), \quad (\text{IV}, 9; 18)$$

и поскольку $\sum_{\Delta'} \geqslant V_2$, то $\sum_{\Delta} \geqslant V_1$.

Следующая теорема позволяет охарактеризовать функции ограниченной вариации со значениями в нормированном конечномерном аффинном пространстве:

Теорема 86. Пусть E — аффинное нормированное пространство размерности n над полем вещественных чисел с системой координат $0, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. Для того чтобы функция M на интервале \mathbb{R}_1 прямой \mathbb{R} со значениями в E имела ограниченную вариацию, необходимо и достаточно, чтобы каждая из ее составляющих имела ограниченную вариацию. Для того чтобы вещественная функция на \mathbb{R}_1 имела ограниченную вариацию, необходимо и достаточно, чтобы она была представима в виде разности двух возрастающих и ограниченных функций.

Доказательство. В силу теоремы 12 гл. II, при замене нормы в \tilde{E} эквивалентной нормой функция ограниченной вариации остается функцией ограниченной вариации¹⁾. Введение системы координат в E позволяет отождествить это пространство с пространством \mathbb{R}^n , и мы можем заменить норму в \tilde{E} эквивалентной нормой (теорема 13 гл. II): $\|(x_1, x_2, \dots, x_n)\| = \sqrt{\sum_{j=1}^n |x_j|^2}$ на \mathbb{R}^n . Если теперь функция M на \mathbb{R}_1 представима в виде $M(x) = 0 + \sum_{j=1}^n X_j(x) \vec{e}_j$, то для любого разбиения Δ отрезка \mathbb{R}_1 имеем:

$$\| \overrightarrow{M(c_{i+1}) - M(c_i)} \| = \sqrt{\sum_{j=1}^n |X_j(c_{i+1}) - X_j(c_i)|^2}. \quad (\text{IV}, 9; 19)$$

Следовательно, полная вариация функции M не меньше полной вариации каждой функции X_j и не больше суммы полных

¹⁾ Значение полной вариации, естественно, изменяется вместе с изменением нормы. Однако если она конечна в одной норме, то она остается конечной и в любой другой эквивалентной норме.

вариаций этих функций¹⁾). Это означает, что M имеет ограниченную вариацию в том и только в том случае, когда ограничены вариации функций X_j .

Пусть теперь M — вещественная функция на \mathbb{R}_1 . Если она моногонна, то все разности $M(c_{i+1}) - M(c_i)$, соответствующие некоторому разбиению Δ , имеют один и тот же знак и сумма (IV, 9; 6) при любом разбиении Δ равна $|M(c_n) - M(c_0)|$. Следовательно, если функция M монотонна и ограничена, то она имеет ограниченную вариацию. Поэтому разность двух возрастающих и ограниченных функций является функцией ограниченной вариации. Обратно, пусть M — вещественная функция ограниченной вариации на \mathbb{R}_1 . Выберем точку $c \in \mathbb{R}_1$. Для каждой точки x из \mathbb{R}_1 обозначим через $V(x)$ величину

$$V(x) = \begin{cases} \text{вариация } M \text{ на } [c, x], & \text{если } x > c, \\ 0, & \text{если } x = c, \\ -(\text{вариация } M \text{ на } [x, c]), & \text{если } x < c. \end{cases} \quad (\text{IV, 9; 20})$$

Очевидно, $V(x)$ является возрастающей и ограниченной функцией $x \in \mathbb{R}_1$. Если $y > x$, то разность $V(y) - V(x)$ будет полной вариацией функции M на интервале $[x, y]$, а, следовательно, она не меньше чем $M(y) - M(x)$. Поэтому

$$\begin{aligned} V(y) - V(x) &\geq M(y) - M(x), \\ \text{или} \quad V(y) - M(y) &\geq V(x) - M(x), \end{aligned} \quad (\text{IV, 9; 21})$$

т. е. разность $V - M$ является возрастающей и ограниченной функцией W , а, следовательно, $M = V - W$ будет разностью двух возрастающих и ограниченных функций.

Следствие. Пусть $|a, b|$ — некоторый интервал прямой \mathbb{R} . Если функция M на $|a, b|$ вещественна, ограничена и кусочно монотонна, то она имеет на $|a, b|$ ограниченную вариацию.

Говорят, что M кусочно монотонна, если интервал $|a, b|$ является объединением конечного числа таких интервалов $|a_0, a_1|, |a_1, a_2|, \dots, |a_{n-1}, a_n|$, $a_0 = a$, $a_n = b$, что в каждом из интервалов $|a_i, a_{i+1}|$ эта функция монотонна (все обычно используемые функции кусочно монотонны для ограниченного интервала $|a, b|$). Предполагая теперь, например, что $|a, b| =$

¹⁾ Если снова воспользоваться тем фактом, что сумма (IV, 9; 6) при замене последовательности c_i на содержащую ее последовательность может лишь возрасти, то легко убедиться в том, что вариация функции M в точности равна сумме вариаций функций X_j .

$=]a, b[$, можем записать:

$$\begin{aligned} V(|a, b|; M) &= V([a_0, a_1]; M) + V([a_1, a_2]; M) + \dots + V([a_{n-1}, a_n]; M) = \\ &= \sum_{i=0}^{n-1} |M(a_{i+1} - 0) - M(a_i + 0)| + \\ &+ \sum_{i=1}^{n-1} (|M(a_i) - M(a_i - 0)| + |M(a_i + 0) - M(a_i)|). \quad (\text{IV, 9; 22}) \end{aligned}$$

Функции ограниченной вариации и неопределенные интегралы

Функции ограниченной вариации были введены по той причине, что для них имеются простые соотношения, связанные с неопределенными интегралами от мер Радона. Будем говорить, что некоторая функция на \mathbb{R}_1 со значениями в \vec{E} имеет локально ограниченную вариацию, если ее сужение на каждый компактный интервал $[a, b]$ из \mathbb{R}_1 имеет ограниченную вариацию.

Теорема 87. Пусть $\vec{\mu}$ — мера Радона на интервале \mathbb{R}_1 прямой \mathbb{R} со значениями в банаховом пространстве \vec{E} и базой $\lambda \geq 0$: $\vec{\mu} = \vec{p}\lambda$. Тогда каждый неопределенный интеграл M от меры $\vec{\mu}$ является функцией с локально ограниченной вариацией и эта функция имеет ограниченную вариацию на каждом таком интервале (открытом, полуоткрытом или замкнутом) $|a, b|$ из \mathbb{R}_1 , на котором \vec{p} λ -интегрируема. Для каждого интервала $|a, b|$ из \mathbb{R}_1 , $a < b$, имеет место равенство

$$V(|a, b|; M) = \int_{|a, b|} \|\vec{p}\| d\lambda \leqslant +\infty. \quad (\text{IV, 9; 23})$$

Все утверждения теоремы вытекают из (IV, 9; 23). Если $c_0, c_1, c_2, \dots, c_n$ — некоторое разбиение Δ интервала $|a, b|$, то

$$\sum_{\Delta} = \sum_{i=0}^{n-1} \|M(c_{i+1}) - M(c_i)\| = \sum_{i=0}^{n-1} \left\| \int_{c_i}^{c_{i+1}} \vec{p} d\lambda \right\| \leqslant \int_{|a, b|} \|\vec{p}\| d\lambda. \quad (\text{IV, 9; 24})$$

Переходя в суммах \sum_{Δ} к точной верхней грани, получаем неравенство

$$V(|a, b|; M) \leqslant \int_{|a, b|} \|\vec{p}\| d\lambda. \quad (\text{IV, 9; 24}_2)$$

Мы доказали соотношение (IV, 9; 23) со знаком \leqslant вместо знака $=$. Докажем сначала, что знак равенства имеет место, если p непрерывна на \mathbb{R}_1 и если $[a, b]$ является замкнутым, а, следовательно, компактным интервалом в \mathbb{R}_1 . В самом деле, положим $\Lambda = \int_a^b d\lambda$. Функция \vec{p} равномерно непрерывна на $[a, b]$.

Следовательно, для любого $\varepsilon > 0$ найдется такое $\eta > 0$, что из $|x' - x''| \leqslant \eta$ следует $\|\vec{p}(x') - \vec{p}(x'')\| \leqslant \varepsilon/2\Lambda$.

Пусть теперь Δ — некоторое разбиение интервала $[a, b]$: $c_0 = a$, $c_n = b$, максимальная из длин интервалов в котором $\leqslant \eta$. Тогда для $c_i \leqslant x \leqslant c_{i+1}$ имеем

$$\|\vec{p}(x)\| - \|\vec{p}(c_i)\| \leqslant \|\vec{p}(x) - \vec{p}(c_i)\| \leqslant \frac{\varepsilon}{2\Lambda}. \quad (\text{IV, 9; 25})$$

Теперь можно установить следующие соотношения:

$$\begin{aligned} \overrightarrow{M(c_{i+1}) - M(c_i)} &= \int_{c_i}^{c_{i+1}} \vec{p}(x) d\lambda(x) = \vec{p}(c_i) \int_{c_i}^{c_{i+1}} d\lambda + \\ &+ \int_{c_i}^{c_{i+1}} (\vec{p}(x) - \vec{p}(c_i)) d\lambda(x), \quad (\text{IV, 9; 26}) \end{aligned}$$

$$\|\overrightarrow{M(c_{i+1}) - M(c_i)}\| \geqslant \|\vec{p}(c_i)\| \int_{c_i}^{c_{i+1}} d\lambda - \frac{\varepsilon}{2\Lambda} \int_{c_i}^{c_{i+1}} d\lambda.$$

Однако

$$\begin{aligned} \|\vec{p}(c_i)\| \int_{c_i}^{c_{i+1}} d\lambda &\geqslant \int_{c_i}^{c_{i+1}} \|\vec{p}(x)\| d\lambda(x) - \int_{c_i}^{c_{i+1}} \|\vec{p}(x) - \vec{p}(c_i)\| d\lambda(x) \geqslant \\ &\geqslant \int_{c_i}^{c_{i+1}} \|\vec{p}\| d\lambda - \frac{\varepsilon}{2\Lambda} \int_{c_i}^{c_{i+1}} d\lambda. \quad (\text{IV, 9; 27}) \end{aligned}$$

Окончательно,

$$\|\overrightarrow{M(c_{i+1}) - M(c_i)}\| \geqslant \int_{c_i}^{c_{i+1}} \|\vec{p}\| d\lambda - \frac{\varepsilon}{\Lambda} \int_{c_i}^{c_{i+1}} d\lambda. \quad (\text{IV, 9; 28})$$

Отсюда получаем

$$V([a, b]; M) \geqslant \sum_{\Delta} \geqslant \int_a^b \|\vec{p}\| d\lambda - \frac{\varepsilon}{\Lambda} \int_a^b d\lambda \geqslant \int_a^b \|\vec{p}\| d\lambda - \varepsilon. \quad (\text{IV, 9; 29})$$

Так как ε произвольно, то $V([a, b]; M) \geq \int_a^b \|\vec{p}\| d\lambda$, что вместе с (IV, 9; 24₂) дает (IV, 9; 23). Если теперь \vec{p} — произвольная функция и $[a, b]$, как всегда, компактен, то функция $\vec{p} \lambda$ -интегрируема на $[a, b]$. Следовательно, согласно теореме 32, при заданном $\varepsilon > 0$ существует такая непрерывная на $[a, b]$ функция \vec{q} , что

$$\int_a^b \|\vec{p} - \vec{q}\| d\lambda \leq \frac{\varepsilon}{2}. \quad (\text{IV, 9; 30})$$

Пусть N — неопределенный интеграл от меры $\vec{q}\lambda$ с базой λ . Тогда

$$V([a, b]; M) \geq V([a, b]; N) \rightarrow V([a, b]; \overrightarrow{M - N}). \quad (\text{IV, 9; 31})$$

Так как функция \vec{q} непрерывна на $[a, b]$, то

$$V([a, b]; N) = \int_a^b \|\vec{q}\| d\lambda. \quad (\text{IV, 9; 32})$$

Согласно (IV, 9; 24₂), можно написать неравенство

$$V([a, b]; \overrightarrow{M - N}) \leq \int_a^b \|\vec{p} - \vec{q}\| d\lambda \leq \frac{\varepsilon}{2}. \quad (\text{IV, 9; 33})$$

Следовательно,

$$\begin{aligned} V([a, b]; M) &\geq \int_a^b \|\vec{q}\| d\lambda - \frac{\varepsilon}{2} \geq \\ &\geq \int_a^b \|\vec{p}\| d\lambda - \int_a^b \|\vec{p} - \vec{q}\| d\lambda - \frac{\varepsilon}{2} \geq \int_a^b \|\vec{p}\| d\lambda - \varepsilon. \end{aligned} \quad (\text{IV, 9; 34})$$

Отсюда, поскольку ε произвольно, вытекает (IV, 9; 23).

Если же функция \vec{p} произвольна, а отрезок $|a, b|$ не обязательно замкнут, то формула сохранится после перехода к пределу. Например, если $|a, b|$ является открытым интервалом, то согласно лемме 1 к теореме 84₂, $V(|a, b|; M)$ является пределом $V([a', b']; M)$, когда a' стремится к a по значениям $>a$, а b' стремится к b по значениям $<b$, и, согласно теореме 36 Фату, $\int_{|a, b|} \|\vec{p}\| d\lambda$ является пределом интеграла $\int_{[a', b']} \|\vec{p}\| d\lambda$.

Следствие. Если \vec{p} — некоторая функция на \mathbb{R}_1 со значениями в банаховом пространстве \vec{E} , локально dx -интегрируемая, то неопределенный интеграл $M: x \rightarrow \int\limits_c^x \vec{p}(t) dt$ есть непрерывная функция ограниченной вариации на каждом замкнутом интервале $[a, b]$ из \mathbb{R}_1 . Ее полная вариация на интервале $[a, b]$ ($a < b$) (открытом, полуоткрытом или замкнутом) из \mathbb{R}_1 определяется по формуле

$$V([a, b]; M) = \int\limits_{[a, b]} \|\vec{p}(t)\| dt. \quad (\text{IV}, 9; 35)$$

Если \vec{E} конечномерно, то имеет место обратное утверждение.

Теорема 88. Пусть M — функция, непрерывная слева и имеющая локально ограниченную вариацию на интервале \mathbb{R}_1 прямой \mathbb{R} , со значениями в конечномерном аффинном нормированном пространстве E . Тогда M является неопределенным интегралом от некоторой меры $\vec{\mu}$ на \mathbb{R}_1 со значениями в \vec{E} и с базой $\geqslant 0$. Эта мера определяется единственным образом, если \mathbb{R}_1 не имеет наибольшей точки или если предполагается, что $\vec{\mu}$ не содержит массы в этой точке. Кроме того, для каждого интервала $[a, b]$ из \mathbb{R}_1 , $a < b$, имеем:

$$\int\limits_a^b d[\vec{\mu}] = [\vec{\mu}]([a, b]) = V([a, b]; M) = V([a, b]; M), \quad (\text{IV}, 9; 36)$$

где $[\vec{\mu}]$ — наименьшая абсолютная мажоранта меры $\vec{\mu}$. Мера $\vec{\mu}$ имеет конечную норму на \mathbb{R}_1 тогда и только тогда, когда M имеет ограниченную вариацию на \mathbb{R}_1 .

Если L является вещественной возрастающей непрерывной слева функцией на \mathbb{R}_1 , то она будет неопределенным интегралом от некоторой меры $\lambda \geqslant 0$. Если для каждой пары точек x' , x'' из \mathbb{R}_1 , $x' \leqslant x''$, имеет место неравенство

$$\overrightarrow{|M(x'') - M(x')|} \leqslant L(x'') - L(x'), \quad (\text{IV}, 9; 37)$$

то, когда \mathbb{R}_1 не имеет наибольшей точки или когда $\vec{\mu}$ не имеет массы в наибольшей точке, мера $\vec{\mu}$ абсолютно мажорируется мерой $\lambda \geqslant 0$.

Легко видеть, почему следует предполагать функцию M непрерывной слева — этого требует теорема 83. Лемма 2 к теореме 84₂ обеспечивает равенство двух последних членов в соотношении (IV, 9; 36).

Обычно меру, для которой функция M является неопределенным интегралом, обозначают не через μ или $d\mu$, а через dM и говорят при этом, что эта мера является производной мерой для функции M .

Мы увидим, что в силу следствия 1 теоремы 89 для функции M , принадлежащей классу C^1 , эта мера равна $\vec{M}'(x)dx$.

Доказательство. Покажем прежде всего, что если M есть неопределенный интеграл от некоторой меры μ на \mathbb{R}_1 со значениями в \vec{E} и интервал \mathbb{R}_1 не имеет наибольшей точки, то мера μ единственна.

В самом деле, мы знаем выражение $\vec{\mu}([a, b])$ для полуоткрытых интервалов $[a, b[$ интервала \mathbb{R}_1 . Если $[a, b]$ — замкнутый интервал из \mathbb{R}_1 и b не является наибольшей точкой \mathbb{R}_1 , то $\vec{\mu}([a, b])$ есть предел значений $\vec{\mu}([a, b'])$ при b' , стремящемся к b по значениям $>b$ (замечание, следующее за теоремой 83), а, значит, меру $\vec{\mu}([a, b])$ также можно считать известной. Точно так же мера $\vec{\mu}([a, b[)$ может быть определена как предел чисел $\vec{\mu}([a', b[)$, когда a' стремится к a по значениям $>a$. Наконец, $\vec{\mu}([a, b])$ является пределом $\vec{\mu}([a', b])$ при a' , стремящемся к a по значениям $>a$. Если \mathbb{R}_1 не имеет наибольшей точки, то мера $\vec{\mu}(\varphi)$ будет известна для каждой ступенчатой функции φ с компактным носителем. Поскольку любая функция φ из $\mathcal{C}(\mathbb{R}_1)$ является равномерным пределом таких функций, имеющих носители в некотором фиксированном компакте из \mathbb{R}_1 , то $\vec{\mu}(\varphi)$ известна для каждой функции φ из $\mathcal{C}(\mathbb{R}_1)$, а, следовательно, μ определяется в рассматриваемом случае единственным образом. Проведенное рассуждение, очевидно, не годится для случая, когда \mathbb{R}_1 имеет наибольшую точку b_0 . В этом случае к μ можно добавить произвольную массу, расположенную в точке b_0 . Это не изменит неопределенные интегралы, в определение которых входят только открытые справа интервалы вида $[a, b[$ из \mathbb{R}_1 . Напротив, если μ определяет в точке b_0 некоторую массу $\vec{\mu}(\{b_0\}) \neq \vec{0}$, то ее можно удалить, т. е. из μ вычесть меру $\vec{\mu}(\{b_0\}) \delta_{(b_0)}$ и получить новую меру без массы в b_0 , для которой функция M является неопределенным интегралом. Такая мера без точки b_0 уже будет единственной.

Покажем теперь, что каждая функция M со значениями в конечномерном пространстве E , имеющая локально ограни-

ченную вариацию, есть неопределенный интеграл от некоторой меры.

В самом деле, пусть $0, (\vec{e}_i)_{i \in I}$ — некоторая система координат пространства E относительно поля \mathbb{R} . Тогда $M = 0 + \sum_{i \in I} M_i \vec{e}_i$, где каждая из функций M_i вещественна, непрерывна слева и имеет локально ограниченную вариацию. Если мы докажем, что она является неопределенным интегралом от некоторой вещественной меры μ_i , то M будет неопределенным интегралом от меры $\hat{\mu} = \sum_{i=1}^n \mu_i \vec{e}_i$. Поэтому нам достаточно ограничиться вещественными функциями.

Вещественная функция M , непрерывная слева и имеющая локально ограниченную вариацию, способом, изложенным в теореме 86, может быть выражена в виде разности двух возрастающих функций: $M = V - (V - M)$. Функции V и $V - M$ непрерывны слева (теорема 84₂). Если каждая из функций V и $V - M$ является неопределенным интегралом от некоторой меры ≥ 0 , то M будет неопределенным интегралом от некоторой вещественной меры. Таким образом, задача свелась к доказательству того, что всякая вещественная возрастающая непрерывная слева функция на \mathbb{R}_1 есть неопределенный интеграл от некоторой меры ≥ 0 .

Пусть M — такая функция. Если M непрерывна и строго возрастает, то утверждение очевидно, ибо тогда M является гомеоморфизмом $x \rightarrow y = M(x)$ интервала \mathbb{R}_1 на некоторый интервал \mathbb{R}_2 из \mathbb{R} (теорема 37 гл. II). Пусть M^{-1} — обратный гомеоморфизм. Обозначим через μ образ при отображении M^{-1} меры Лебега dy из \mathbb{R}_2 . Пусть теперь $[a, b]$ — некоторый интервал из \mathbb{R}_1 и $[\alpha, \beta] = [M(a), M(b)]$ — его образ в \mathbb{R}_2 при отображении M . Тогда, согласно следствию 1 из теоремы 60, получим, что $\mu([a, b]) = (M^{-1}(dy))([a, b]) = dy([M(a), M(b)]) = M(b) - M(a)$, а, следовательно, M действительно является неопределенным интегралом для меры μ .

Пусть теперь M — возрастающая функция (не обязательно строго), не обязательно непрерывная. Пусть a_i — ее точки разрыва (их счетное множество), и пусть c_i — ее скачки. Для каждого интервала $[a, b]$ из \mathbb{R}_1 имеем: $\sum_{a_i \in [a, b]} c_i \leq M(b) - M(a)$.

Рассмотрим меру Радона $\lambda = \sum_i c_i \delta_{(a_i)}$. Эта мера имеет неопределенный интеграл L (теорема 82). Интеграл L имеет точки разрыва a_i и скачки c_i в этих точках (теорема 83). Из предыдущего следует, что $M(b) - M(a) \geq \sum_{a_i \in [a, b]} c_i = L(b) - L(a)$,

поэтому $M - L$ также возрастающая функция, но теперь уже непрерывная. Далее, функция $M - L + \varepsilon h$ строго возрастает и непрерывна. Значит, она является неопределенным интегралом от некоторой меры Радона μ_e , а тогда M является неопределенным интегралом от меры $\mu = \mu_e + \lambda - \varepsilon dx$. Сразу же видно, что мера $\mu \geq 0$, но мера $\mu + \varepsilon dx \geq 0$. Поскольку μ не зависит от способа ее вычисления (единственность была доказана выше), т. е. от выбора ε , то $\mu \geq 0$.

Предположим теперь, что существует такая возрастающая вещественная функция L на \mathbb{R}_1 , для которой имеет место равенство (IV, 9; 37). Пусть μ и λ — меры, имеющие функции M и L своими неопределенными интегралами. На каждом полуоткрытом интервале $[a, b]$ выполняется неравенство $\|\mu([a, b])\| \leq \lambda([a, b])$. Если \mathbb{R}_1 не имеет наибольшей точки или если μ не имеет массы в наибольшей точке, то переходом к пределу, как это делалось ранее, можно убедиться, что последнее неравенство остается верным для интервалов других типов. Если $\varphi \geq 0$ — ступенчатая функция с компактным носителем на \mathbb{R}_1 , то $\|\mu(\varphi)\| \leq \lambda(\varphi)$. Переходом к пределу можно установить то же неравенство для любой функции $\varphi \geq 0$ из $\mathcal{C}(\mathbb{R}_1)$, а это означает, что μ абсолютно мажорируется мерой λ . Заметим, наконец, что если $\lambda \geq 0$ есть база для μ , $\mu = p\lambda$, то, согласно теореме 54, $|\mu|$ является мерой $\|p\|d\lambda$ и тогда соотношение (IV, 9; 36) эквивалентно соотношению (IV, 9; 23).

Мера μ имеет конечную норму на \mathbb{R}_1 тогда и только тогда, когда $|\mu|$ имеет конечную норму. [В любом случае $\|\mu\| \leq \||\mu|\|$. Мы видели, что в случае, когда размерность E не меньше 2, равенство, вообще говоря, места не имеет (см. теорему 15). Поэтому из неравенства $\|\mu\| < +\infty$ неравенство $\||\mu|\| < +\infty$ непосредственно не вытекает. Однако в случае, когда пространство E конечномерно, это будет именно так. Пусть $(e_i)_{i \in I}$ — некоторый базис в \mathbb{E} над полем \mathbb{R} , $\mu = \sum_{i \in I} \vec{e}_i \mu_i$. Тогда $\sum_{i \in I} \|\vec{e}_i\| |\mu_i|$ является абсолютной мажорантой меры μ (стр. 488). Если μ имеет конечную норму, то такими же будут меры μ_i , а, значит, и $|\mu_i|$, имеющие ту же норму (теорема 18). Если некоторая абсолютная мажоранта меры μ имеет конечную норму, то такой же будет наименьшая абсолютная мажоранта. То есть она имеет конечную норму тогда

и только тогда, когда $\int_{\mathbb{R}_1} d|\vec{\mu}| < +\infty$, или тогда и только тогда, когда M имеет ограниченную вариацию на всем \mathbb{R}_1 .

Следствие. Для того чтобы функция M на интервале \mathbb{R}_1 прямой \mathbb{R} со значениями в аффинном нормированном конечномерном пространстве E удовлетворяла условию Липшица с коэффициентом Липшица, равным k , необходимо и достаточно, чтобы она была неопределенным интегралом по отношению к мере Лебега от некоторой функции \vec{p} на \mathbb{R}_1 со значениями в \vec{E} с нормой, всюду мажорируемой числом k .

Доказательство. Если, для $x' \leqslant x''$, $M(x'') - M(x') = \int_{x'}^{x''} \vec{p}(x) dx$, $\|\vec{p}(x)\| \leqslant k$, то $\|\overrightarrow{M(x'') - M(x')} \| \leqslant k(x'' - x')$.

Обратно, если имеет место последнее неравенство, то имеет место неравенство (IV, 9; 37) при $L(x) = kx$. Следовательно, M является неопределенным интегралом некоторой меры $\vec{\mu}$, которую можно считать не имеющей массы в наибольшей точке интервала \mathbb{R}_1 (если такая точка существует). Очевидно, $\vec{\mu}$ абсолютно мажорируется мерой kdx . Поэтому из теоремы 53 Данфорда — Петтиса следует, что $d\vec{\mu} = \vec{p}dx$, где $\|\vec{p}\| \leqslant k$ на \mathbb{R}_1 .

Длина пути в метрическом пространстве

Согласно определению, данному на стр. 90, путем в метрическом пространстве E называется непрерывное отображение M отрезка $[a, b]$ прямой \mathbb{R} в метрическое пространство E .

Длиной этого пути называется полная вариация $V(M)$ функции M , $0 \leqslant V(M) \leqslant +\infty$. Путь имеет конечную длину тогда и только тогда, когда функция M имеет ограниченную вариацию. Если E является конечномерным аффинным пространством, снабженным некоторой системой координат, то путь будет иметь конечную длину тогда и только тогда, когда составляющие функции M имеют ограниченную вариацию.

Предположим, в частности, что E является аффинным нормированным пространством. Пусть $t \rightarrow A + t(B - A)$ ¹⁾ есть «прямолинейный» путь, концы которого A и B соответствуют значениям $t = 0$ и $t = 1$ соответственно. Для каждого разбиения

¹⁾ При изучении путей и их длин мы будем чаще пользоваться обозначением переменной через t или u вместо x . Именно так обычно поступают в аналитической геометрии при параметрическом задании кривых.

ния Δ интервала $[0, 1]$ имеем равенство $\|M(c_{i+1}) - M(c_i)\| = = (c_{i+1} - c_i) \overrightarrow{\|B - A\|}$, откуда $\sum_{\Delta} = \overrightarrow{\|B - A\|}$ и, значит, длина прямолинейного пути, как это и следовало ожидать, равна $\overrightarrow{\|B - A\|}$. Если теперь рассмотреть произвольный путь M и разбиение Δ : $c_0 = a, c_1, c_2, \dots, c_n = b$ интервала $[a, b]$, то величина \sum_{Δ} , определенная по формуле (IV, 9; 6), соответствующая этому пути, является длиной ломаной, вписанной в этот путь, с вершинами $M(c_0), M(c_1), \dots, M(c_n)$. Полная вариация или длина пути определяется как точная верхняя грань длин ломаных, вписанных в этот путь.

(Если E — произвольное метрическое пространство, то мы будем иногда пользоваться понятием точной верхней грани длин вписанных ломаных, когда, строго говоря, нельзя использовать понятие ломаной, а расстояние $d(M(C_i), M(C_{i+1}))$ можно определить, не проводя отрезка прямой, соединяющей эти точки.)

Рассмотрим два эквивалентных¹⁾ пути в E . Это означает, что мы имеем два отображения M и N двух отрезков $[a, b]$ и $[\alpha, \beta]$ прямой \mathbb{R} в E , и предполагаем, что существует такой гомеоморфизм ξ отрезка $[\alpha, \beta]$ на отрезок $[a, b]$, что $N = M \circ \xi$. В этом случае, очевидно, оба пути имеют одну и ту же длину.

Функция, которая каждому пути ставит в соответствие его длину, является функцией, определенной на всевозможных путях в E со значениями в $[0, +\infty]$ и обладающей двумя следующими очевидными свойствами:

1°) Каким бы ни был путь с началом в точке A и концом в точке B пространства E , его длина не меньше, чем расстояние $d(A, B)$, соответствующее простейшему разбиению $c_0 = a, c_1 = b$. В случае аффинного нормированного пространства это означает, что прямая является кратчайшим путем, соединяющим две точки²⁾.

2°) Если путь составлен последовательно из двух путей, то его длина равна сумме их длин³⁾.

¹⁾ В том же смысле, в каком эквивалентны параметрические многообразия; см. стр. 352.

²⁾ Однако она не обязательно будет единственной, обеспечивающей кратчайший путь. В некоторых нормированных аффинных пространствах можно взять систему из трех точек A, B, C , не расположенных на одной прямой и таких, что $\|C - A\| = \|B - A\| + \|C - B\|$.

³⁾ Следует обратить внимание на особый характер этой аддитивности. Если, например, при t , изменяющемся в $[a, b]$, $M(t)$ описывает сначала отрезок $[A, C]$, а затем отрезок $[C, A]$ так что $M(a) = M(b) = A$, то длина будет равна, конечно, $2\|C - A\|$, а не 0.

Уточним, о чём же здесь идет речь.

Пусть M_1 — некоторый путь, т. е. непрерывное отображение отрезка $[a, b]$, $a < b$, в E . Пусть M_2 — другой путь, т. е. отображение отрезка $[b, c]$, $b < c$, в E . Предположим еще, что $M_1(b) = M_2(b) = B$. Тогда M_1 и M_2 , очевидно, определят непрерывное отображение M отрезка $[a, c]$ в E :

$$M(t) = \begin{cases} M_1(t) & \text{для } a \leq t \leq b, \\ M_2(t) & \text{для } b \leq t \leq c, \end{cases} \quad (\text{IV}, 9; 37_2)$$

которое задает путь, называемый путем, составленным из первых двух путей. Утверждение 2°) тогда вытекает из (IV, 9; 6₃).

Функция длины является наименьшей функцией, определенной на путях из E , принимающей неотрицательные значения и обладающей двумя указанными свойствами.

В самом деле, если некоторая функция обладает двумя этими свойствами, то длина пути будет, очевидно, ≥ 0 длины любой из вписанных ломаных, а, значит, и их точной верхней грани, т. е. той длины пути, которую мы только что определили.

Из теоремы 85 следует, что длина пути может быть определена как предел длин последовательности вписанных ломаных, а не обязательно как их точная верхняя грань.

Пусть R_1 — некоторый не обязательно замкнутый интервал прямой R , а M — непрерывная функция на R_1 со значениями в метрическом пространстве E . Говорят, что эта функция определяет некоторый несобственный путь. Говорят, что этот несобственный путь *спрямляем*, если для каждого компактного интервала $[a, b]$ из R_1 сужение $M_{[a, b]}$ функции M на $[a, b]^1$ определяет путь конечной длины (сам несобственный путь может иметь бесконечную длину; это только означает, что функция M имеет локально ограниченную вариацию).

Пусть теперь c — некоторая точка R_1 . Каждой точке t из R_1 можно поставить в соответствие вещественное число $s(t)$, определенное следующим образом: если $t \geq c$, то $s(t)$ равно длине дуги $M_{[c, t]}$; если $t < c$, то $s(t)$ равно числу, противоположному длине дуги $M_{[t, c]}$. Таким образом на R_1 задается функция s : $t \rightarrow s(t)$ с вещественными значениями (уже определявшаяся в формуле (IV, 9; 20)), непрерывная (теорема 84₂),

¹⁾ На стр. 16 сужение функции f , заданной на E , на часть A этого множества мы обозначали через f_A . Иногда в этом случае используется также обозначение $f|A$. Тогда рассмотренная выше дуга является путем, определенным суженным отображением $M_{[a, b]}$, или $M|[a, b]$.

возрастающая и в общем случае строго возрастающая. Функция $s(t)$ называется *криволинейной абсциссой относительно* $t \in \mathbb{R}_1$, когда точка s берется за начало отсчета оси абсцисс¹⁾. Если функция s строго возрастает, то она определяет гомеоморфизм интервала \mathbb{R}_1 на некоторый интервал \mathbb{R}_2 вещественной прямой (теорема 37 гл. II). Так как при этом функция $s \rightarrow t(s)$ обратима, то можно взять s в качестве параметра и таким образом определить новую параметризацию несобственного пути, эквивалентную исходной параметризации. Различные параметризации, полученные при изменении s или при изменении знаков $+$ и $-$ в определении s , эквивалентны. Их называют *натуральными параметризациями* (они зависят, очевидно, не только от топологии, но и от метрики пространства E). Если функция s не является строго возрастающей, она не определяет эквивалентной параметризации пути. Однако если число s_0 задано и существует бесконечное множество значений t , для которых $s(t) = s_0$, то эти значения образуют некоторый интервал из \mathbb{R}_1 , на котором M постоянна, так что соответствующие значения $M(t)$ совпадают с одной и той же точкой P_0 из E . Следовательно, если положить $P_0 = M_0(s_0)$, то можно убедиться в том, что на некотором интервале \mathbb{R}_2 прямой \mathbb{R} существует такая функция M_0 , что $M(t) = M_0(s(t))$. Функция M_0 непрерывна и даже обладает свойством Липшица, ибо образ интервала $[s', s'']$ при отображении M_0 имеет длину, равную $|s'' - s'|$, а, следовательно,

$$\overrightarrow{|M_0(s'') - M_0(s')|} \leq |s'' - s'|. \quad (\text{IV}, 9; 37_3)$$

Эта функция не определяет эквивалентного пути, поскольку s не является инъекцией и потому не будет гомеоморфизмом \mathbb{R}_1 на \mathbb{R}_2 . Мы видим, что параметризация M_0 сохраняет «интересные» свойства параметризации M (и даже практически представляет больший интерес, чем параметризация M), и потому мы продолжаем называть ее натуральной параметризацией несобственного пути. (Допуская вольность речи, отмеченную на стр. 16 и состоящую в том, что даже после замены переменных для функции сохраняется прежнее обозначение, вместо $M_0(s)$ часто пишут $M(s)$, говоря, что это та же функция, выраженная через переменную s вместо переменной t .)

Возрастающая функция s на \mathbb{R}_1 является неопределенным интегралом от некоторой меры Радона ≥ 0 на \mathbb{R}_1 (теорема 88), по отношению к которой весь интервал $[t', t'']$ из $[a, b]$ имеет в качестве меры длину $|s'' - s'|$ пути $M_{[t', t'']}$. Обычно эту

¹⁾ Криволинейную абсциссу точки P нельзя, вообще говоря, обозначать через $P = M(t)$, поскольку одна и та же точка P из E может быть образом при отображении M нескольких точек t отрезка $[a, b]$ («кратные точки» пути).

меру на \mathbb{R}_1 обозначают через ds (когда переходят к натуральному параметрическому представлению, она становится настоящей мерой Лебега ds). Заметим, что ds является мерой на \mathbb{R}_1 , а не на образе пути. Эту меру называют *мерой длины* на пути.

Из самого определения s (формула (IV, 9; 20)) и теоремы 88 (формула (IV, 9; 36)) вытекает, что если E является аффинным нормированным конечномерным пространством, то ds есть наименьшая абсолютная мажоранта меры $d\vec{M}$. Из теоремы 54 Данфорда — Петтиса следует, что можно написать $d\vec{M} = \vec{p} ds$, $\|\vec{p}\| = 1$ ds -почти всюду.

Если E является локально компактным пространством, счетным в бесконечности, и если отображение M является ds -собственным, то существует образ предыдущей меры — мера Радона ≥ 0 на E , носитель которой содержится в замыкании образа пути. Зачастую, несмотря на путаницу, к которой могут привести неточные обозначения, вместо $M(ds)$ пишут ds .

Если \vec{f} является функцией, определенной на E , со значениями в банаховом пространстве \vec{F} , то можно говорить о *криволинейном интеграле* $\int \vec{f} ds$, полагая его по определению равным $\int_{\mathbb{R}} \vec{f}(M(t)) d(s(t))$. Согласно теореме 60, этот интеграл

является интегралом от функции \vec{f} по отношению к образу меры $M(ds)$, если он существует¹⁾.

Теорема 88₂. *Образ спрямляемого несобственного пути в пространстве \mathbb{R}^n имеет нулевую меру по отношению к мере $dx = dx_1 \otimes dx_2 \otimes \dots \otimes dx_n$.*

Доказательство. Поскольку путь спрямляем, то для доказательства этой теоремы можно заменить его путем $M_0 : s \rightarrow M_0(s)$, соответствующим натуральной параметризации, ибо образы путей M и M_0 совпадают.

Пусть $[a, b]$ — некоторый интервал из \mathbb{R}_2 . Пусть Δ — такое разбиение $[a, b] : c_0 = a, c_1, c_2, \dots, c_l = b$, что длина наибольшего из отрезков $\leq \eta$. Пусть c'_i — середина $[c_i, c_{i+1}]$. Поскольку длина любой дуги не меньше длины хорды, соединяющей ее

¹⁾ Обозначение $\int \vec{f} ds$ неполно, так как в него не входит M . Лучше будет писать $\int_{(M)} \vec{f} ds$ — интеграл относительно пути, определяемого отображением M .

концы, то расстояния от точек $M(t)$, $t \in [c_i, c_{i+1}]$, до $M(c'_i)$ мажорируются длинами дуг $M_{[c'_i, t]}$, т. е. числами $(c_{i+1} - c_i)/2$. Тогда образ пути $M_{[c_i, c_{i+1}]}$ содержитя в кубе с центром $M(c'_i)$ со сторонами длины $c_{i+1} - c_i$, параллельными осям. Объемная мера этого образа мажорируется, как нетрудно видеть, числом $(c_{i+1} - c_i)^n \leq \eta^{n-1}(c_{i+1} - c_i)$. Объемная же мера отрезка $[a, b]$ мажорируется величиной $\eta^{n-1}(b - a)$. Так как η произвольно, то объем этого образа равен нулю. Поскольку множество \mathbb{R}_2 является счетным объединением интервалов вида $[a, b]$, то его образ также имеет объем, равный нулю.

Замечание. Не надо думать, что этот результат будет верен в случае неспрямляемого пути. См. по этому поводу замечание к следствию 1 теоремы 102₂.

Неопределенный интеграл и первообразная

Теорема 89. Пусть \vec{f} — функция, определенная на интервале \mathbb{R}_1 вещественной прямой \mathbb{R} , со значениями в банаховом пространстве \vec{E} , локально интегрируемая по мере Лебега. Тогда неопределенный интеграл F от функции \vec{f} по dx со значениями в аффинном нормированном пространстве E с присоединенным векторным пространством \vec{E} имеет $\vec{f}(c)$ в качестве производной (соответственно производной слева или производной справа) в каждой точке c , в которой функция \vec{f} непрерывна (соответственно непрерывна слева или непрерывна справа).

Доказательство. Докажем, например, дифференцируемость F в случае, когда функция \vec{f} непрерывна.

Имеем

$$\frac{\overrightarrow{F(c+h) - F(c)}}{h} - \vec{f}(c) = \frac{1}{h} \int_c^{c+h} (\vec{f}(x) - \vec{f}(c)) dx, \quad (\text{IV}, 9; 38)$$

откуда вытекает оценка

$$\left\| \frac{\overrightarrow{F(c+h) - F(c)}}{h} - \vec{f}(c) \right\| \leq \sup_{|x-c| \leq h} \| \vec{f}(x) - \vec{f}(c) \| . \quad (\text{IV}, 9; 38_2)$$

В силу предположения о непрерывности, правая часть стремится к нулю при h , стремящемся к нулю, что и доказывает теорему.

Следствие 1. Пусть \vec{f} — непрерывная функция на \mathbb{R}_1 со значениями в \vec{E} . Тогда неопределенные интегралы от \vec{f} до dx со

значениями в E и первообразные функции \vec{f} со значениями в E , т. е. функции, имеющие \vec{f} в качестве производной, совпадают. Каждая функция на \mathbb{R}_1 со значениями в E класса C^1 является неопределенным интегралом от своей производной функции по отношению к мере Лебега dx .

Доказательство. Теорема говорит о том, что любой неопределенный интеграл является некоторой первообразной. Так как разность между двумя первообразными есть некоторая постоянная из \vec{E} (теорема 22 гл. III), то каждая первообразная отличается от неопределенного интеграла на постоянную, а, значит, также является неопределенным интегралом (теорема 82).

Замечания. 1°) Известно, что в основе интегрального исчисления лежит вычисление первообразных. Если F — некоторая первообразная функции \vec{f} , то

$$\overrightarrow{F(b) - F(a)} = \int_a^b \vec{f}(x) dx. \quad (\text{IV}, 9; 38_3)$$

2°) Пусть \vec{F} — функция двух переменных: $(x, y) \rightarrow \int_x^y \vec{f}(t) dt$, а функция \vec{f} всюду непрерывна. Тогда \vec{F} имеет, очевидно, полную производную, определяемую в дифференциальных обозначениях по формуле

$$d\vec{F} = (F'(x, y)) \cdot d(x, y) = \vec{f}(y) dy - \vec{f}(x) dx. \quad (\text{IV}, 9; 39)$$

В самом деле, согласно соотношению Шаля,

$$\vec{F}(x, y) = \int_a^y \vec{f}(t) dt - \int_a^x \vec{f}(t) dt, \quad (\text{IV}, 9; 40)$$

причем эти интегралы (один, являющийся лишь функцией x , другой — лишь функцией y) имеют в качестве производных функции $\vec{f}(x)$ и $\vec{f}(y)$.

3°) Дифференцируемая функция, производная которой не является всюду непрерывной функцией, не обязательно будет неопределенным интегралом от своей производной, которая, впрочем, не обязательно локально интегрируема по мере Лебега.

4°) Следствие связывает два довольно разных понятия. Первообразная связана с производной. Отыскание первообразной обобщается следующим образом. Пусть E и F — аффинные нормированные пространства и Ω — открытое множество из E . Пусть \vec{f} — некоторая непрерывная функция на Ω со значениями в $\mathcal{L}(E; F)$. Отыскивается функция на Ω со значениями в F , допускающая \vec{f} в качестве производной. Неопределенный интеграл связан с понятием меры Радона. Можно говорить о неопределенных интегралах от некоторой меры μ или от некоторой функции \vec{f} по отношению к мере $\lambda \geqslant 0$. Здесь сравниваются первообразные от \vec{f} и неопределенные интегралы от \vec{f} по отношению к мере dx .

Следствие 2. Пусть M — отображение класса C^1 интервала \mathbb{R}_1 прямой \mathbb{R} в аффинное нормированное пространство E . Это отображение определяет некоторый спрямляемый путь, и на каждом интервале $[a, b]$ из \mathbb{R}_1 , $a \leqslant b$, длина дуги $M_{[a, b]}$ вычисляется по формуле

$$s(b) - s(a) = \int_{[a, b]} \| \vec{M}'(t) \| dt. \quad (\text{IV}, 9; 40_2)$$

В самом деле, согласно следствию 1, функция M является неопределенным интегралом от $\vec{M}'(t) dt$. Следовательно, эта функция имеет локально ограниченную вариацию (теорема 87), а, значит, определяет некоторый спрямляемый путь. Согласно (IV, 9; 23) или (IV, 9; 35), длина пути $M_{[a, b]}$, или полная вариация M на $[a, b]$, равна (IV, 9; 40₂).

Теорема 89₂. Если функция M , определенная на \mathbb{R}_1 , со значениями в аффинном нормированном пространстве E принадлежит классу C^1 , то функция криволинейная абсцисса s также принадлежит классу C^1 . В каждой точке t_0 , где $\vec{M}'(t_0) \neq \vec{0}$, путь имеет касательную и единичный вектор этой касательной задается формулой $\overrightarrow{M'(t_0)} / \| \vec{M}'(t_0) \|$. В окрестности точки t_0 функция M_0 является функцией класса C^1 криволинейной абсциссы s , и имеет место равенство

$$\frac{\overrightarrow{M'(t_0)}}{\| \vec{M}'(t_0) \|} = \left(\frac{d\vec{M}_0}{ds} \right)_{s=s_0}, \quad (\text{IV}, 9; 40_3)$$

определяющее единичный вектор касательной.

Кроме того, если две точки t' и t'' стремятся к точке t_0 , то дуга, соединяющая точки $M(t')$ и $M(t'')$, имеет длину, эквивалентную длине стягивающей ее хорды.

Доказательство. Так как, согласно следствию 2 теоремы 89, s представляет собой неопределенный интеграл от не-прерывной функции $t \rightarrow \overrightarrow{\|M'(t)\|}$, то s является функцией t класса C^1 . Если теперь эта производная в точке t_0 отлична от нуля, то из теоремы о неявных функциях (теорема 29 гл. III) следует, что в некоторой окрестности точки t_0 можно выразить t как функцию s класса C^1 . Производную этой функции в точке t_0 можно вычислить по формуле

$$\left(\frac{dt}{ds} \right)_{s=s_0} = \frac{1}{\left(\frac{ds}{dt} \right)_{t=t_0}} = \frac{1}{\| \overrightarrow{M'(t_0)} \|}. \quad (\text{IV}, 9; 40_4)$$

Отсюда вытекает, что функция $M_0(s)$, которая может быть вычислена без каких либо предположений о ее дифференцируемости, дифференцируема по s в точках s , близких к $s(t_0)$, и ее производная задается формулой

$$\left(\frac{d\overrightarrow{M}_0}{ds} \right)_{s=s_0} = \frac{\overrightarrow{dM}}{dt} \cdot \frac{dt}{ds} = \frac{\overrightarrow{M'(t_0)}}{\| \overrightarrow{M'(t_0)} \|}. \quad (\text{IV}, 9; 40_5)$$

В условии теоремы говорилось о касательной к пути в точке.

Это вовсе не касательная к множеству образа пути в смысле определения, данного на стр. 218, ибо одна и та же точка, например $M(t_0)$, может быть также образом при отображении M бесконечного множества других точек $t \neq t_0$, соответствующих различным «ветвям кривой», проходящим через ту же точку, и рассмотрение только лишь производной M в точке t_0 не может дать никаких сведений относительно касательной в этой точке. Касательной к пути в точке t_0 мы называем предел, если он существует, прямой, соединяющей точки $M(t_0)$ и $M(t_0 + \Delta t)$ при $\Delta t \neq 0$, стремящемся к 0. Если существует производная $\neq 0$, то можно утверждать, что каждая касательная существует и что $\overrightarrow{M'(t_0)}$ является вектором этой касательной¹⁾.

Поскольку здесь $d\overrightarrow{M}_0/ds$ есть вектор с нормой 1, то он определяет единичный вектор касательной. Говорят также, что этот вектор является единичным вектором полукасательной, соответствующей «возрастающему направлению t », поскольку эта полукасательная является пределом полупрямой с началом в точке $M(t)$, проходящей через $M(t + \Delta t)$ при Δt , стремящемся к 0 по значениям > 0 .

¹⁾ Точно так же можно определить контингенцию в точке некоторого параметрического многообразия класса C^1 , не имеющую ничего общего с контингенцией образа параметрического многообразия.

Пусть теперь t' и t'' — два числа, стремящихся к t_0 . Соответствующие значения s' и s'' криволинейной абсциссы стремятся к s_0 . По формуле о конечных приращениях (следствие 1 теоремы 13 гл. III) можно записать:

$$\overrightarrow{M_0(s'') - M_0(s')} = \overrightarrow{M'_0(s_0)}(s'' - s') + \overrightarrow{\omega} |s'' - s'|, \quad (\text{IV}, 9; 40_6)$$

где

$$\|\overrightarrow{\omega}\| \leq \sup_{\theta \in [s', s'']} \|\overrightarrow{M'_0(\theta)} - \overrightarrow{M'_0(s_0)}\|.$$

Так как $\|\overrightarrow{M'(s_0)}\| = 1$, то мы получаем отсюда

$$\begin{aligned} \|\overrightarrow{M_0(s'') - M_0(s')} - |s'' - s'|\| &\leq \|\overrightarrow{M_0(s'') - M_0(s')} - \\ &- \overrightarrow{M'_0(s_0)}(s'' - s')\| \leq \|\overrightarrow{\omega}\| |s'' - s'|. \quad (\text{IV}, 9; 40_7) \end{aligned}$$

Непрерывность производной $\overrightarrow{dM_0/ds}$ обеспечивает стремление к нулю нормы $\|\overrightarrow{\omega}\|$ при s' и s'' , стремящихся к s_0 ; тем самым полностью доказана эквивалентность длины дуги $|s'' - s'|$ и длины хорды $\|\overrightarrow{M_0(s'') - M_0(s')}\|$.

Естественно, что теорема 89 дает достаточное условие для того, чтобы неопределенный интеграл F функции \tilde{f} был дифференцируемым. Он может иметь производные даже в тех точках, где \tilde{f} разрывна. Впрочем, иногда случается, что функция \tilde{f} , локально интегрируемая по мере dx , является разрывной во всех точках прямой, как в примере, построенном на стр. 532. Тем не менее, функция F для некоторых значений x необходимо имеет $\tilde{f}(x)$ в качестве производной, как это видно из следующей теоремы, доказанной Лебегом (мы примем ее без доказательства в силу его сложности).

Теорема 90. Если \tilde{f} является функцией, определенной на интервале \mathbb{R}_1 прямой \mathbb{R} , со значениями в банаевом пространстве \tilde{E} и локально интегрируемой по мере Лебега dx , то каждый неопределенный интеграл F от функции \tilde{f} относительно dx имеет $\tilde{f}(x)$ в качестве производной для dx -почти всех значений x .

Конечно большего, чем dx -почти всюду, ожидать не приходится, поскольку F не изменяется, если менять \tilde{f} на множестве меры нуль.

Следствие 1. Каждая функция F на \mathbb{R}_1 со значениями в аффинном нормированном пространстве E , имеющая всюду производную, ограниченную по норме числом k , удовлетворяет условию Липшица с коэффициентом Липшица k . Каждая функция F на \mathbb{R}_1 со значениями в конечномерном пространстве E , удовлетворяющая условию Липшица с коэффициентом k , имеет dx -почти всюду производную с нормой $\leq k$.

Первая часть является следствием 2 теоремы 13 гл. III, а вторая вытекает из теоремы 90 и следствия теоремы 88.

Замечание. Обе части следствия не будут в точности обратными друг к другу. Если известно, что некоторая функция F удовлетворяет условию Липшица, то отсюда еще не следует, что она имеет всюду производную (пример: $x \rightarrow |x|$ удовлетворяет условию Липшица с коэффициентом 1, но не имеет производной в начале координат). Если известно лишь, что некоторая функция F имеет dx -почти всюду производную с нормой $\leq k$, то она не обязательно будет удовлетворять условию Липшица и даже не обязательно будет непрерывной (функция Хевисайда (формула (IV, 9; 5)) имеет dx -почти всюду производную, равную нулю).

Следствие 2. Пусть $t \rightarrow M(t)$ — спрямляемый путь некоторого конечномерного аффинного нормированного пространства E , ds — мера Радона и криволинейная абсцисса s является неопределенным интегралом от ds . Тогда этот путь ds -почти всюду имеет касательную.

Здесь снова мы говорим «касательная к пути» вместо «касательная к образу пути».

Доказательство. Рассмотрим натуральное параметрическое представление кривой $s \rightarrow M_0(s)$. На стр. 705 мы видели, что $\overrightarrow{dM_0} = \overrightarrow{p} ds$, $\|\overrightarrow{p}\| = 1$ ds -почти всюду. Согласно теореме Лебега 90, M_0 имеет ds -почти всюду функцию \overrightarrow{p} в качестве производной. На стр. 709 мы видели, что в каждой точке, где существует производная $\neq 0$, имеется касательная к пути. Следовательно, путь ds -почти всюду имеет касательную с единичным вектором \overrightarrow{p} .

Последовательные первообразные непрерывной функции на прямой

Пусть, как и прежде, \mathbb{R}_1 — некоторый интервал прямой \mathbb{R} : Пусть \vec{f} — непрерывная функция на \mathbb{R} со значениями в

банаховом пространстве \vec{E} . Мы видели уже, как интегрирование позволяет находить первообразную функции \vec{f} .

Покажем теперь, как последовательные первообразные функции \vec{f} , требующие *a priori* нескольких последовательных интегрирований, могут быть вычислены с помощью одного интегрирования.

Теорема 91. *Если \vec{f} является непрерывным отображением некоторого интервала \mathbb{R}_1 прямой \mathbb{R} в банахово пространство \vec{E} , то существует, и при том единственная, m -я первообразная \vec{F}_m отображения \vec{f} , обращающаяся в нуль вместе со своими последовательными производными до $(m-1)$ -го порядка включительно в некоторой точке $x = c$ из \mathbb{R}_1 , и эта производная задается интегралом*

$$\vec{F}_m(x) = \int_c^x \frac{(x-\xi)^{m-1}}{(m-1)!} \vec{f}(\xi) d\xi^{(1)}. \quad (\text{IV}, 9; 41)$$

Доказательство. Единственность первообразной порядка m , удовлетворяющей заданным условиям, очевидна. В самом деле, согласно теореме 22 гл. III, разность между двумя первообразными порядка m является некоторым полиномом степени $\leq m-1$. Однако, если требовать, чтобы первообразные вместе с их производными до $(m-1)$ -го порядка обращались в нуль в точке c , то эти полиномы вместе с производными до $(m-1)$ -го порядка должны равняться нулю в точке c и, следовательно, в силу формулы Тейлора для полиномов, должны быть тождественно равны нулю. Нам остается лишь доказать, что формула (IV, 9; 41) действительно определяет некоторую первообразную порядка m , удовлетворяющую условиям теоремы.

Теорема очевидна для $m = 1$. Предположим, что она верна относительно первообразной порядка $m-1$, и докажем ее справедливость относительно первообразных порядка $m \geq 2$. Иначе говоря, предположим, что функция

$$\vec{F}_{m-1}(x) = \int_0^x \frac{(x-\xi)^{m-2}}{(m-2)!} \vec{f}(\xi) d\xi \quad (\text{IV}, 9; 42)$$

есть первообразная $(m-1)$ -го порядка функции \vec{f} , обращающаяся в нуль вместе со своими производными порядков

¹⁾ Эта формула требует большого внимания; ξ является переменной интегрирования, а x , не являющаяся переменной интегрирования, фигурирует одновременно в \int и под знаком интеграла!

$\leq m - 2$ в точке c . Тогда существует первообразная \vec{F}_m порядка m от \vec{f} , обращающаяся в нуль вместе с производными порядков до $m - 1$ в точке c , и эта первообразная вычисляется по формуле

$$\vec{F}_m(x) = \int_c^x \vec{F}_{m-1}(t) dt = \int_c^x dt \int_c^t \frac{(t-\xi)^{m-2}}{(m-2)!} \vec{f}(\xi) d\xi. \quad (\text{IV}, 9; 43)$$

Предположим сначала, что $x > c$. Тогда два последовательных интегрирования можно заменить одним интегралом:

$$\int_A \int \frac{(t-\xi)^{m-2}}{(m-2)!} \vec{f}(\xi) dt d\xi, \quad (\text{IV}, 9; 44)$$

где через A обозначено множество точек плоскости \mathbb{R}^2 , определяемое неравенствами $c \leq \xi \leq t \leq x^1$, $x \in \mathbb{R}_1$. В самом деле, последний интеграл имеет смысл, поскольку речь идет о некоторой непрерывной функции, определенной на компакте, а тогда теорема Фубини показывает, что выражение (IV, 9; 43) равно (IV, 9; 44).

Полученный двойной интеграл можно вычислить, начиная интегрирование по t , а затем интегрируя по ξ , другими словами, изменяя порядок интегрирования. С помощью формулы (IV, 8; 35) получаем:

$$F_m(x) = \int_c^x \vec{f}(\xi) d\xi \int_\xi^x \frac{(t-\xi)^{m-2}}{(m-2)!} dt, \quad (\text{IV}, 9; 45)$$

где последний интеграл равен $(x-\xi)^{m-1}/(m-1)!$. Утверждение доказано, так как аналогичные рассуждения для $x < c$ приводят к той же формуле.

Следствие 1. Пусть \vec{f} — функция, удовлетворяющая условиям теоремы. Тогда на \mathbb{R}_1 существует, и притом единственная, первообразная F_m порядка m для функции \vec{f} со значениями в аффинном пространстве E с присоединенным векторным пространством \vec{E} , последовательные производные которой до $(m-1)$ -го порядка включительно принимают заданные значения $F_m(c), \vec{F}'_m(c), \dots, \vec{F}_m^{(m-1)}(c)$ в точке c . Эта первообразная

¹⁾ Можно было бы вместо \leq писать $<$, так как это изменяет подинтегральную функцию только на множестве меры нуль.

вычисляется по формуле

$$F_m(x) = F_m(c) + \frac{x-c}{1!} \vec{F}'(c) + \dots + \frac{(x-c)^{m-1}}{(m-1)!} \vec{F}_m^{(m-1)}(c) + \\ + \int_c^x \frac{(x-\xi)^{m-1}}{(m-1)!} \vec{f}(\xi) d\xi. \quad (\text{IV}, 9; 46)$$

Доказательство. Единственность очевидна. Если мы заметим, что функция F_m , определенная по формуле (IV, 9; 46), является суммой некоторой первообразной порядка m , производные которой до $(m+1)$ -го порядка в точке c обращаются в нуль, и некоторого полинома, задаваемого своим разложением Тейлора, производные которого до $(m+1)$ -го порядка в точке c принимают заданные значения, то результат становится очевидным.

Следствие 2 (новая формула Тейлора для функции одной вещественной переменной). *Пусть f — некоторая функция, определенная на интервале \mathbb{R}_1 прямой \mathbb{R} , со значениями в аффинном нормированном пространстве F , имеющая непрерывные производные порядков $\leq m+1$. Тогда имеет место формула*

$$f(x) = f(c) + \frac{x-c}{1!} \vec{f}'(c) + \dots + \frac{(x-c)^m}{m!} \vec{f}^{(m)}(c) + \\ + \int_c^x \frac{(x-\xi)^m}{m!} \vec{f}^{(m+1)}(\xi) d\xi, \quad (\text{IV}, 9; 47)$$

а также формула

$$f(x+h) = f(x) + \frac{h}{1!} \vec{f}'(x) + \dots + \frac{h^m}{m!} \vec{f}^{(m)}(x) + \\ + h^{m+1} \int_0^1 \frac{(1-t)^m}{m!} \vec{f}^{(m+1)}(x+th) dt \text{¹⁾.} \quad (\text{IV}, 9; 48)$$

Доказательство. Первая формула является на самом деле формулой (IV, 9; 46), в которой \vec{f} , \vec{F}_m и m заменены на $\vec{f}^{(m+1)}$, f и $m+1$. Формула применима, так как f является первообразной порядка $m+1$ непрерывной функции $\vec{f}^{(m+1)}$ и ее

¹⁾ Надо было бы предполагать пространство F полным, как это делалось в теореме с тем, чтобы быть уверенным в существовании интеграла. Однако здесь это не является необходимым, так как при заданной функции f этот интеграл существует.

производные порядка $0, 1, 2, \dots, m$ принимают значения $f(c), \vec{f}'(c), \vec{f}''(c), \dots, \vec{f}^{(m)}(c)$ в точке c .

Вторая формула получается заменой переменной или же применением первой формулы к интервалу $[0, 1]$ для функции $t \rightarrow f(x + t\vec{h})$.

Следствие 3 (новое общее выражение формулы Тейлора). Пусть f — отображение класса C^{m+1} открытого множества Ω аффинного нормированного пространства E в аффинное нормированное пространство F . Если весь отрезок $[x, x + \vec{h}]$ содержится в Ω , то имеет место формула:

$$\begin{aligned} f(x + \vec{h}) = & f(x) + \frac{f'(x)}{1!} \vec{h} + \dots + \frac{f^{(m)}(x)}{m!} \vec{h}^m + \\ & + \int_0^1 \frac{(1-t)^m}{m!} (f^{(m+1)}(x + t\vec{h}) \vec{h}^{m+1}) dt. \quad (\text{IV}, 9; 49) \end{aligned}$$

Доказательство. Для доказательства достаточно применить формулу Тейлора (IV, 9; 47) к интервалу $[0, 1]$ и функции $t \rightarrow f(x + t\vec{h})$. Согласно теореме о сложной функции, производная порядка k в точке t этой функции равна $f^{(k)}(x + t\vec{h}) \vec{h}^k$. Полученная формула Тейлора значительно важнее формулы, приведенной в гл. III. В самом деле, она дает явное выражение для остаточного члена. Однако она требует несколько больше, а именно — функция f должна иметь производную $(m+1)$ -го порядка, непрерывную в области Ω . При этом дополнительном предположении можно снова получить различные виды формулы Тейлора, приведенные в гл. III.

Если, например, f принимает вещественные значения, то можно воспользоваться теоремой о среднем в виде (IV, 1; 31). Тогда найдется такое число θ , $0 < \theta < 1$, при котором

$$\begin{aligned} \int_0^1 \frac{(1-t)^m}{m!} (f^{(m+1)}(x + t\vec{h}) \vec{h}^{m+1}) dt = \\ = (f^{(m+1)}(x + \theta\vec{h}) \vec{h}^{m+1}) \int_0^1 \frac{(1-t)^m}{m!} dt = \frac{f^{(m+1)}(x + \theta\vec{h}) \vec{h}^{m+1}}{(m+1)!}, \quad (\text{IV}, 9; 50) \end{aligned}$$

что снова приводит нас к формуле Тейлора (III, 7; 1₂). Если не делать никаких предположений относительно F , то согласно формуле (IV, 1; 24) имеет место следующая оценка остаточного

члена:

$$\left\| \int_0^1 \frac{(1-t)^m}{m!} (f^{(m+1)}(x + t\vec{h}) \vec{h}^{m+1}) dt \right\| \leqslant \|\vec{h}\|^{m+1} \|f^{(m+1)}\| \left\| \int_0^1 \frac{(1-t)^m}{m!} dt \right\| = \|f^{(m+1)}\| \frac{\|\vec{h}\|^{m+1}}{(m+1)!}, \quad (\text{IV, 9; 51})$$

и мы снова возвращаемся к формуле Тейлора (III, 7; 2).

Если L является $(m+1)$ -линейным непрерывным отображением \vec{E} в \vec{F} , то, учитывая равенство $\int_0^1 \frac{(1-t)^m}{m!} dt = \frac{1}{(m+1)!}$, можно записать:

$$\begin{aligned} \int_0^1 \frac{(1-t)^m}{m!} (f^{(m+1)}(x + t\vec{h}) \vec{h}^{m+1}) dt &= \\ &= \frac{L}{(m+1)!} \vec{h}^{m+1} + \int_0^1 \frac{(1-t)^m}{m!} ((f^{(m+1)}(x + t\vec{h}) - L) \vec{h}^{m+1}) dt, \end{aligned} \quad (\text{IV, 9; 52})$$

где последний интеграл мажорируется величиной

$$\begin{aligned} \sup_{0 \leqslant t \leqslant 1} \|f^{(m+1)}(x + t\vec{h}) - L\| \|\vec{h}\|^{m+1} \int_0^1 \frac{(1-t)^m}{m!} dt &= \\ &= \sup_{0 \leqslant t \leqslant 1} \|f^{(m+1)}(x + t\vec{h}) - L\| \frac{\|\vec{h}\|^{m+1}}{(m+1)!}; \end{aligned} \quad (\text{IV, 9; 53})$$

мы вернулись к формуле Тейлора в виде (III, 7; 9).

Окончательно, все найденные ранее выражения для формулы Тейлора являются следствиями нового выражения. Однако, напомним еще раз, что они были доказаны при *менее ограничительных* условиях и, как мы видели в главе о дифференциальном исчислении, эти менее ограничительные условия очень полезны в некоторых теоремах (см., например, доказательство теоремы 14 гл. III).

Формула интегрирования по частям

Теорема 92. 1°) Элементарный случай. Пусть \vec{u} и \vec{v} — функции класса C^1 , определенные на замкнутом интервале $[a, b]$ ¹⁾ прямой \mathbb{R} , со значениями в банаховых пространствах

¹⁾ В этой формуле a и b — произвольные точки \mathbb{R} : $a < b$, $a = b$ или $a > b$.

\vec{E} и \vec{F} соответственно. Пусть B — непрерывное билинейное отображение пространства $\vec{E} \times \vec{F}$ в банахово пространство \vec{G} . Тогда имеет место формула интегрирования по частям:

$$\int_a^b B(\vec{u}(x), \vec{v}'(x)) dx = \\ = B(\vec{u}(b), \vec{v}(b)) - B(\vec{u}(a), \vec{v}(a)) - \int_a^b B(\vec{u}'(x), \vec{v}(x)) dx. \quad (\text{IV}, 9; 54)$$

2°) Общий случай. Пусть $\vec{\mu}$ и \vec{v} — две меры Радона, определенные на замкнутом интервале $[a, b]$ прямой \mathbb{R} , со значениями в банаховых пространствах \vec{E} и \vec{F} соответственно, имеющие базы ≥ 0 . Пусть B — непрерывное билинейное отображение $\vec{E} \times \vec{F}$ в банахово пространство \vec{G} . Пусть \vec{M} и \vec{N} — неопределенные интегралы мер $\vec{\mu}$ и \vec{v} со значениями в \vec{E} и \vec{F} . Если $\vec{\mu}$ и \vec{v} не имеют общих точечных масс, то интегралы $\int_a^b B(\vec{M}, d\vec{v})$

и $\int_a^b B(d\vec{\mu}, \vec{N})$ имеют смысл и справедлива общая формула интегрирования по частям:

$$\int_a^b B(\vec{M}, d\vec{v}) = B(\vec{M}(b), \vec{N}(b)) - B(\vec{M}(a), \vec{N}(a)) - \int_a^b B(d\vec{\mu}, \vec{N}). \quad (\text{IV}, 9; 55)$$

Доказательство. Если учесть формулу дифференцирования билинейного непрерывного отображения (теорема 12 гл. III)

$$\vec{F}' = \vec{f} = B(\vec{u}', \vec{v}) + B(\vec{u}, \vec{v}'), \quad (\text{IV}, 9; 56)$$

то первая формула вытекает из формулы (IV, 9; 38₃), примененной к $\vec{F} = B(\vec{u}, \vec{v})$.

Легко видеть, почему формула общего случая является обобщением частной формулы интегрирования по частям. В самом деле, если положить $d\vec{\mu} = \vec{u}' dx$ и $d\vec{v} = \vec{v}' dx$, то $\vec{M} = \vec{u}$ и $\vec{N} = \vec{v}$ будут неопределенными интегралами от $\vec{\mu}$ и \vec{v} . Меры $\vec{\mu}$ и \vec{v} имеют в качестве базы меру $dx \geq 0$. Поскольку меры $\vec{\mu}$ и \vec{v} не имеют точечных масс, то они не имеют общей

точечной массы, а, следовательно, формула (IV, 9; 55) (с учетом определения (IV, 5; 13)) сводится к формуле (IV, 9; 54).

Докажем теперь общий случай. В силу соотношения $\int_a^b = - \int_b^a$ всегда можно считать $a < b$.

Пусть μ_0 и v_0 — такие вещественные меры ≥ 0 , для которых μ имеет в качестве базы меру μ_0 , а v имеет в качестве базы меру v_0 . Если положить $\lambda_0 = \mu_0 + v_0$, то мы увидим, что обе меры μ_0 и v_0 мажорируются мерой λ_0 , а, следовательно (теорема 53 Данфорда — Петтиса), имеют в качестве базы меру λ_0 . Поэтому $d\mu_0 = p_0 d\lambda_0$ и $dv_0 = q_0 d\lambda_0$, где p_0 и q_0 локально λ_0 -интегрируемы. Поскольку одновременно $d\mu = \vec{p} d\mu_0$, $dv = \vec{q} dv_0$, то окончательно получим (следствие теоремы 51):

$$\vec{d}\mu = (\vec{p}p_0) d\lambda_0 = \vec{f} d\lambda_0, \quad \vec{d}v = (\vec{q}q_0) d\lambda_0 = \vec{g} d\lambda_0. \quad (\text{IV, 9; 57})$$

Рассмотрим теперь двойной интеграл

$$\int_{[a, b]} \int_{[a, b]} B(\vec{f}(x), \vec{g}(x)) d\lambda_0(x) d\lambda_0(y). \quad (\text{IV, 9; 58})$$

Поскольку \vec{f} и \vec{g} являются $d\lambda_0$ -интегрируемыми на компакте $[a, b]$, этот интеграл, согласно теореме 79, существует. Та же самая теорема говорит, что этот интеграл равен

$$B \left(\int_a^b \vec{f} d\lambda_0, \int_a^b \vec{g} d\lambda_0 \right) = B(\vec{M}(b) - \vec{M}(a), \vec{N}(b) - \vec{N}(a)). \quad (\text{IV, 9; 59})$$

Представим теперь квадрат $[a, b] \times [a, b]$ в виде объединения трех непересекающихся множеств (и, конечно, λ_0 -измеримых как пересечение открытых и замкнутых множеств):

1°) множество $a \leq y < x < b$, отмеченное на рис. 19 цифрой 1;

2°) множество $a \leq x < y < b$, отмеченное на рис. 19 цифрой 2;

3°) диагональ квадрата, т. е. множество $a \leq x = y < b$, отмеченное на рис. 19 цифрой 3.

Рассмотрим интеграл, аналогичный (IV, 9; 59), на первом множестве. Согласно теореме Фубини — Лебега (теорема 77),

$$\int_{(1)} \int_{[a, b]} = \int_{[a, b]} d\lambda_0(x) \int_a^x B(\vec{f}(x), \vec{g}(y)) d\lambda_0(y). \quad (\text{IV, 9; 60})$$

Функция g является λ_0 -интегрируемой на интервале $[a, x]$, и, поскольку отображение $g \rightarrow B(\vec{f}(x), \vec{g})$ является непрерыв-

ным отображением \vec{F} в \vec{G} , можно записать

$$\int_a^x B(\vec{f}(x), \vec{g}(y)) d\lambda_0(y) = B\left(\vec{f}(x), \int_a^x \vec{g} d\lambda_0\right). \quad (\text{IV}, 9; 61)$$

Здесь интеграл $\int_a^x \vec{g} d\lambda_0$ равен $\vec{N}(x) - \vec{N}(a)$. Кроме того,

$\int_a^x \vec{g} d\lambda_0 = 0$ при $x=a$. Поэтому второй интеграл $\int_{[a, b]}$ можно

заменить на $\int_{[a, b]} = \int_a^b$. Отсюда мы получаем значение первого интеграла. С учетом (IV, 5; 13) получаем интеграл

$$\int \int_{(1)} = \int_a^b B(\vec{f}(x), \vec{N}(x) - \vec{N}(a)) d\lambda_0(x) = \int_a^b B(d\vec{\mu}(x), \vec{N}(x) - \vec{N}(a)), \quad (\text{IV}, 9; 62)$$

обязательно имеющий смысл.

Совершенно такие же рассуждения относительно интеграла по множеству 2 дадут следующий результат:

$$\int \int_{(2)} = \int_a^b B(\vec{M}(x) - \vec{M}(a), d\vec{v}(x)). \quad (\text{IV}, 9; 63)$$

Покажем теперь, что интеграл, рассматриваемый на множестве 3, т. е. на диагонали, равен нулю. Его можно вычислить по теореме Фубини. Если зафиксировать x , то надо будет

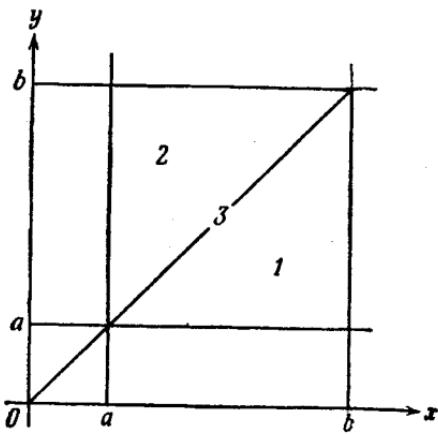


Рис. 19.

сначала на множестве $\{x\}$ вычислить $\int B(\vec{f}(x), \vec{g}(y)) d\lambda_0(y)$. Этот

интеграл всегда равен нулю, кроме того случая, когда его мера λ_0 имеет точечную массу в точке x . Обозначим через $\lambda_0(\{x\})$ эту точечную массу. Тогда предыдущий интеграл равен $B(\vec{f}(x), \vec{g}(x))\lambda_0(\{x\})$, и, следовательно, интеграл по диагонали будет определяться формулой

$$\int \int_{(3)} = \int_a^b B(\vec{f}(x), \vec{g}(x))\lambda_0(\{x\}) d\lambda_0(x). \quad (\text{IV}, 9; 64)$$

Заметим, что масса меры $\vec{\mu}$ в точке x , т. е. интеграл относительно $d\lambda_0$ от характеристической функции множества $\{x\}$, сводящегося к точке x , равен $\vec{f}(x)\lambda_0(\{x\})$. Точно так же масса меры \vec{v} в точке x равна $\vec{g}(x)\lambda_0(\{x\})$. Предположение, что $\vec{\mu}$ и \vec{v} не имеют общих точечных масс, равносильно утверждению, что для каждой точки x , в которой λ_0 имеет точечную массу, по крайней мере одна из функций \vec{f} или \vec{g} равна нулю. Но тогда функция $x \rightarrow B(\vec{f}(x), \vec{g}(x))\lambda_0(\{x\})$ тождественно равна нулю, а, следовательно, интеграл (IV, 9; 64) равен нулю.

Сравнивая (IV, 9; 59), (IV, 9; 62) и (IV, 9; 63), получаем:

$$\begin{aligned} B(\vec{M}(b) - \vec{M}(a), \vec{N}(b) - \vec{N}(a)) &= \\ &= \int_a^b B(d\vec{\mu}(x), \vec{N}(x) - \vec{N}(a)) + \int_a^b B(\vec{M}(x) - \vec{M}(a), d\vec{v}(x)). \end{aligned} \quad (\text{IV}, 9; 65)$$

Раскрывая и производя, как в (IV, 9; 61), перестановку интеграла и линейного непрерывного отображения, получаем

$$\begin{aligned} B(\vec{M}(b), \vec{N}(b)) - B(\vec{M}(a), \vec{N}(b)) - B(\vec{M}(b), \vec{N}(a)) + B(\vec{M}(a), \vec{N}(a)) &= \\ &= \int_a^b B(d\vec{\mu}, \vec{N}) - B\left(\int_a^b d\vec{\mu}, N(a)\right) + \int_a^b B(\vec{M}, d\vec{v}) - B\left(\vec{M}(a), \int_a^b d\vec{v}\right). \end{aligned} \quad (\text{IV}, 9; 66)$$

Поскольку $\int_a^b d\vec{\mu} = \vec{M}(b) - \vec{M}(a)$ и $\int_a^b d\vec{v} = \vec{N}(b) - \vec{N}(a)$, то после взаимного уничтожения равных членов в обеих частях равенства мы получим искомую формулу (IV, 9; 55).

Замена переменных при вычислении простых интегралов

Теорема 93. Пусть $t \rightarrow x = \xi(t)$ — некоторая функция класса C^1 , определенная на интервале $[\alpha, \beta]$ прямой \mathbb{R} , такая

что $\xi(\alpha) = a$ и $\xi(\beta) = b$ ¹⁾. Пусть \vec{f} — непрерывная функция, определенная на интервале $[a, b]$, со значениями в банаховом пространстве \vec{F} . Тогда справедлива формула «замены переменной»:

$$\int_a^b \vec{f}(x) dx = \int_a^\beta \vec{f}(\xi(t)) \xi'(t) dt. \quad (\text{IV}, 9; 67)$$

Доказательство. Рассмотрим функцию Φ , определенную равенством

$$\vec{\Phi}(u) = \int_a^u \vec{f}(x) dx. \quad (\text{IV}, 9; 68)$$

Так как \vec{f} предполагалась непрерывной, то функция $\vec{\Phi}$ принадлежит классу C^1 и является первообразной функции \vec{f} . Кроме того, $\vec{\Phi}(a) = \vec{0}$, а $\vec{\Phi}(b)$ совпадает с левой частью равенства (IV, 9; 67). Функция $s \rightarrow \vec{\Psi}(s) = \vec{\Phi}(\xi(s))$, согласно теореме о сложных функциях, принадлежит классу C^1 . Она удовлетворяет условию $\vec{\Psi}(\alpha) = \vec{\Phi}(a) = \vec{0}$, а ее производная вычисляется по формуле

$$\vec{\Psi}'(s) = \vec{\Phi}'(\xi(s)) \xi'(s) = \vec{f}(\xi(s)) \xi'(s). \quad (\text{IV}, 9; 69)$$

Отсюда следует, что для любого значения s можно записать

$$\vec{\Psi}(s) = \int_a^s \vec{f}(\xi(t)) \xi'(t) dt. \quad (\text{IV}, 9; 70)$$

При $s = \beta$ получаем $\vec{\Psi}(\beta) = \vec{\Phi}(b)$, левую часть равенства (IV, 9; 67), а так как правая часть равенства (IV, 9; 70) совпадает с правой частью (IV, 9; 67), то теорема доказана.

Применяемая здесь формула не обобщается на кратные интегралы, поскольку ее вывод опирается на интеграл \int_a^b , относящийся исключительно к прямой. Нам будет полезно воспользоваться другой формой интеграла, использующей символ $\int_{[a, b]}$.

¹⁾ Здесь α, β, a, b — произвольные числа на \mathbb{R} , $\alpha \leq \beta$, $a \leq b$.

Пусть I — интервал $[a, b]$ или $[b, a]$ в зависимости от того, какая из величин a или b больше, и пусть J — интервал $[\alpha, \beta]$ или $[\beta, \alpha]$.

Если к предположениям теоремы 93 мы добавим новое условие: ξ является монотонной функцией на J , то формулу (IV, 9; 67) можно будет записать в виде

$$\int_I \tilde{f}(x) dx = \int_J \tilde{f}(\xi(t)) |\xi'(t)| dt. \quad (\text{IV, 9; 71})$$

В самом деле, если ξ возрастает, то $|\xi'(t)|$ совпадает с $\xi'(t) \geq 0$. С другой стороны, интегралы \int_I и \int_J являются

интегралами вида $\varepsilon \int_a^b$ и $\varepsilon \int_a^b$, где $\varepsilon = \pm 1$. Так как функция ξ предполагается возрастающей, то значение ε будет одним и тем же в обоих случаях. Если же функция ξ будет убывающей, то функция $|\xi'(t)|$ равна $-\xi'(t)$. Значения, которые принимает ε в обоих предыдущих интегралах, противоположны друг другу.

Формулу (IV, 9; 71) мы докажем для несколько иного случая.

Теорема 94. Пусть \mathcal{O} и Ω — два открытых множества вещественной прямой \mathbb{R} . Пусть ξ — гомеоморфизм Ω на \mathcal{O} , принадлежащий вместе со своим обратным гомеоморфизмом классу C^1 . Пусть \tilde{f} — непрерывная функция на \mathcal{O} с компактным носителем, принимающая значения в банаевом пространстве \tilde{F} . При этих условиях имеет место формула

$$\int_{\mathcal{O}} \tilde{f}(x) dx = \int_{\Omega} \tilde{f}(\xi(t)) |\xi'(t)| dt. \quad (\text{IV, 9; 72})$$

Доказательство. Рассмотрим произвольную точку носителя \tilde{f} . Поскольку \mathcal{O} является окрестностью этой точки, существует не менее одного открытого интервала с центром в этой точке, замыкание которого целиком лежит в \mathcal{O} . Множество всевозможных таких открытых интервалов, образованных для всех точек носителя \tilde{f} , является покрытием этого носителя. Поскольку носитель компактен, то достаточно конечного числа интервалов, например $[a_i, b_i]$, $i \in I$, для покрытия этого носителя.

Пусть теперь γ_i — некоторое разбиение единицы, относящееся к покрытию интервалами $[a_i, b_i]$. Тогда

$$\int_{\mathcal{O}} \vec{f}(x) dx = \sum_{i \in I} \int_{[a_i, b_i]} \gamma_i(x) \vec{f}(x) dx. \quad (\text{IV}, 9; 73)$$

Функция ξ^{-1} должна быть гомеоморфизмом интервала $[a_i, b_i]$ на его образ. Значит, функция ξ^{-1} вместе со своей обратной функцией ξ строго монотонна в $[a_i, b_i]$ ¹⁾ (теорема 37 гл. II). Если мы положим $a_i = \xi(\alpha_i)$, $b_i = \xi(\beta_i)$, то в силу монотонности ξ можно применить формулу (IV, 9; 67) и записать равенство

$$\int_{[a_i, b_i]} \gamma_i(x) \vec{f}(x) dx = \int_{[\alpha_i, \beta_i]} \gamma_i(\xi(t)) \vec{f}(\xi(t)) |\xi'(t)| dt, \quad (\text{IV}, 9; 74)$$

из которого вытекает, что

$$\begin{aligned} \int_{\mathcal{O}} \vec{f}(x) dx &= \sum_{i \in I} \int_{[\alpha_i, \beta_i]} \gamma_i(\xi(t)) \vec{f}(\xi(t)) |\xi'(t)| dt = \\ &= \sum_{i \in I} \int_{\Omega} \gamma_i(\xi(t)) \vec{f}(\xi(t)) |\xi'(t)| dt = \int_{\Omega} \vec{f}(\xi(t)) |\xi'(t)| dt. \end{aligned} \quad (\text{IV}, 9; 75)$$

Теорема доказана.

Следствие 1. В условиях теоремы образ меры $|\xi'(t)| dt$ на множестве Ω при отображении ξ множества Ω в множество \mathcal{O} является мерой Лебега dx на множестве \mathcal{O} .

Для доказательства в качестве f следует взять какую-либо скалярную непрерывную функцию с компактным носителем на \mathcal{O} . При этом мы придем к определению образа некоторой меры.

Следствие 2. Если ξ удовлетворяет условиям теоремы и если \vec{f} является произвольной функцией на \mathcal{O} со значениями в банаховом пространстве \vec{F} , то \vec{f} будет интегрируемой относительно dx на \mathcal{O} тогда и только тогда, когда интегрируема на Ω .

¹⁾ В противном случае существовали бы такие три точки u, v, w этого интервала, что $u < v < w$ и $\xi^{-1}(u) \leqslant \xi^{-1}(v)$, $\xi^{-1}(w) \leqslant \xi^{-1}(v)$ или $\xi^{-1}(u) \geqslant \xi^{-1}(v)$, $\xi^{-1}(w) \geqslant \xi^{-1}(v)$. Рассмотрим, например, первое условие со знаком \leqslant . В действительности, в силу биективности ξ^{-1} , можно было бы взять даже знак $<$. Однако тогда в силу следствия теоремы 33 гл. II функция ξ^{-1} принимала бы не менее двух раз каждое значение, заключенное строго между $\xi^{-1}(v)$ и $\sup(\xi^{-1}(u), \xi^{-1}(w))$: один раз в интервале $[u, v]$ и один раз в интервале $[v, w]$, что невозможно, так как эта функция биективна.

относительно dt функция $t \rightarrow \vec{f}(\xi(t)) |\xi'(t)|$, и при этом справедливо равенство (IV, 9; 72).

В самом деле, по следствию 1, $d\vec{x}$ является образом $|\xi'(t)| dt$ при отображении ξ . При этом, в силу теоремы 60, функция \vec{f} интегрируема относительно $d\vec{x}$ тогда и только тогда, когда будет интегрируемой относительно $|\xi'(t)| dt$ функция $\vec{f}(\xi(t))$. В силу теоремы 51, последнее имеет место тогда и только тогда, когда функция $\vec{f}(\xi(t)) |\xi'(t)|$ интегрируема относительно dt . Соотношение (IV, 9; 72) следует из теорем 60 и 51.

Несобственные интегралы на прямой

Пусть $|a, b|$, $a < b$, — некоторый интервал прямой \mathbb{R} , не содержащий своего правого конца b . Пусть μ — мера Радона на $|a, b|$ со значениями в банаховом пространстве \vec{E} и с базой $\geqslant 0$. Несобственный интеграл $\int_{|a, \rightarrow b|} d\vec{\mu}$ называется сходящимся, если интеграл $\int_{|a, b|} d\vec{\mu}$ при каждом b' из $|a, b|$ имеет смысл и при b' , стремящемся к b по значениям из $|a, b|$, существует предел в \vec{E} . Вместо $\int_{|a, b|}$ мы пишем $\int_{|a, \rightarrow b|}$ (где $\rightarrow b$ означает, что в $|a, b|$ берутся значения b' , стремящиеся к b), поскольку это выражение в первом обозначении не имеет смысла как интеграл от некоторой интегрируемой функции. Выражение « $\int_{|a, \rightarrow b|}$ сходится» означает, что для каждого неопределенного интеграла M от $\vec{\mu}$, со значениями в аффинном нормированном пространстве E с присоединенным векторным пространством \vec{E} , $M(b')$ имеет предел при b' , стремящемся к b по значениям из $|a, b|$.

Пусть λ — некоторая мера $\geqslant 0$, а \vec{f} — локально λ -интегрируемая функция со значениями в банаховом пространстве \vec{F} . Говорят, что несобственный интеграл $\int_{\lambda} \vec{f} d\lambda$ от \vec{f} по λ сходится, если сходится $\int_{|a, \rightarrow b|} d\vec{\mu}$, $\vec{\mu} = \vec{f}\lambda$. Говорят, что $\int_{|a, \rightarrow b|} \vec{f} d\lambda$ абсолютно или нормально сходится, если сходится

интеграл $\int_{[a, \rightarrow b]} \|\vec{f}\| d\lambda$. Аналогичные определения даются для интегралов вида $\int_{[a \leftarrow, b]}$.

Можно рассматривать интегралы вида $\int_{[a \leftarrow, \rightarrow b]}$. Говорят, что такой интеграл сходится, если существует предел интеграла $\int_{[a', b']}$, когда a' и b' , оставаясь в $[a, b]$, стремятся независимо друг от друга соответственно к a и b . Если, например, $a < b$, то это означает, что, каково бы ни было число $\varepsilon > 0$, найдутся такие $a'_0 > a$ и $b'_0 < b$, что для всех $a < a' \leq a'_0$ и $b'_0 \leq b' < b$ имеет место неравенство $\left| \int_{[a \leftarrow, \rightarrow b]} - \int_{[a', b']} \right| \leq \varepsilon$.

Это также означает, что для произвольного c из $[a, b]$ оба интеграла $\int_{[c, \rightarrow b]}$ и $\int_{[a \leftarrow, c]}$ имеют смысл.

Нередко a и b равны $\pm \infty$. В частности, можно рассматривать $\int_{[-\infty \leftarrow, +\infty]}$. Вводят иногда также и обозначения вида \int_a^b , $\int_{b \leftarrow}^b$, $\int_{a \leftarrow}^b$.

Теорема 97. Пусть λ — мера Радона ≥ 0 на $[a, b]$ и \vec{f} — функция, определенная на $[a, b]$, со значениями в банаховом пространстве \vec{F} , λ -интегрируемая на каждом интервале $[a, b']$, где $b' \in [a, b]$. Для того чтобы интеграл $\int_{[a, \rightarrow b]} \vec{f} d\lambda$ был абсолютно сходящимся, необходимо и достаточно, чтобы при b' , изменяющемся в $[a, b]$, интеграл $\int_{[a, b']} \|\vec{f}\| d\lambda$ был ограничен.

Необходимо и достаточно также, чтобы функция \vec{f} была интегрируемой относительно меры λ на интервале $[a, b]$, и тогда интеграл $\int_{[a, \rightarrow b]} \vec{f} d\lambda$ сходится и равен $\int_{[a, b]} \vec{f} d\lambda$.

Как мы уже видели, эта теорема говорит о том, что понятие интегрируемости в смысле Лебега соответствует обобщенному понятию абсолютно сходящегося несобственного интеграла Римана, рассматриваемому в курсе математического анализа. Напротив, условно сходящиеся несобственные интегралы

Римана не являются интегралами от функций, интегрируемых в смысле Лебега. В этом случае запись $\int_{[a, b]} \vec{f} d\lambda$ не имеет смысла, и с точки зрения излагаемой теории абсолютно необходимо писать $\int_{[a, \rightarrow b]} \vec{f} d\lambda$). Эта теорема, с другой стороны, показывает, что на этой стадии понятие абсолютно сходящегося несобственного интеграла никакого интереса для нас не представляет, ибо оно не дает ничего нового по отношению к интегралу Лебега. Новыми являются лишь *условносходящиеся* несобственные интегралы, т. е. сходящиеся, но не абсолютно сходящиеся.

Доказательство. Обозначим через Φ характеристическую функцию интервала $[a, b]$.

1°) Поскольку $b' \rightarrow \int_{[a, b']} \|\vec{f}\| d\lambda$ является возрастающей функцией, эта функция имеет предел при b' , стремящемся к b , тогда и только тогда, когда она ограничена. Следовательно, $\int_{[a, \rightarrow b]} \|\vec{f}\| d\lambda$ сходится тогда и только тогда, когда при переменном b' ограничен интеграл $\int_{[a, b']} \|\vec{f}\| d\lambda$.

2°) Локально интегрируемая функция \vec{f} измерима (см. примечание на стр. 592). Следовательно, эта функция интегрируема тогда и только тогда, когда $\int_{[a, b]}^* \|\vec{f}\| d\lambda = \int_{[a, b]} \|\vec{f}\| \Phi_b d\lambda < +\infty$.

Однако, по теореме 36 Фату, это выражение является пределом от $\int_{[a, b']} \|\vec{f}\| \Phi_{b'} d\lambda = \int_{[a, b']}^* \|\vec{f}\| d\lambda = \int_{[a, b']} \|\vec{f}\| d\lambda$. Следовательно,

\vec{f} интегрируема тогда и только тогда, когда имеет смысл интеграл $\int_{[a, \rightarrow b]} \|\vec{f}\| d\lambda$.

¹⁾ Мы говорим «с точки зрения излагаемой теории необходимо», так как в большинстве случаев стрелка не ставится! Например, всегда пишут

$$\int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}, \text{ в то время как } \frac{\sin x}{x} \text{ не интегрируема по } dx \text{ на } [0, +\infty[.$$

Надо было бы писать $\int_{[0, \rightarrow +\infty[} \frac{\sin x}{x} dx$ или $\int_0^{+\infty} \frac{\sin x}{x} dx$.

3°) Если эти условия выполнены для фиксированной интегрируемой функции ≥ 0 , то при b' , стремящемся к b , $\vec{f}\Phi_{b'}$ будет просто сходится к $\vec{f}\Phi_b$ и имеет место неравенство $\|\vec{f}\Phi_{b'}\| \leq \|\vec{f}\|$. Из теоремы 35 Лебега следует, что при b' , стремящемся к b , интеграл $\int_{[a, b']} \vec{f} d\lambda = \int \vec{f}\Phi_{b'} d\lambda$ сходится к $\int_a^b \vec{f} d\lambda = \int_a^b \vec{f}\Phi_b d\lambda$, а, следовательно, $\int_a^b \vec{f} d\lambda$ сходится и равен $\int_{[a, b]} \vec{f} d\lambda$.

Следствие 1. На каждом ограниченном интервале $[a, b]$, $a < b$, содержащем начало координат, функция $1/|x|^\alpha$ интегрируема по мере Лебега тогда и только тогда, когда $\alpha < 1$. На интервале $[a, +\infty[$ или $]-\infty, a]$, не содержащем начала, это функция интегрируема тогда и только тогда, когда $\alpha > 1$.

Функция $1/|x|^\alpha$ непрерывна в $\mathbb{C}0$, а, следовательно, локально интегрируема в $]-\infty, 0[$ и $]0, +\infty[$. Пусть, например, $a \leq 0 \leq b$, $a < b$. Функция $\frac{\varepsilon}{1-a} \frac{1}{|x|^{\alpha-1}}$, где $\varepsilon = +1$ для $x > 0$ и $\varepsilon = -1$ для $x < 0$, является первообразной функции $1/|x|^\alpha$ в $\mathbb{C}0$ при $\alpha \neq 1$, а функция $\varepsilon \ln|x|$ является ее первообразной при $\alpha = 1$.

Для $0 < c < d < +\infty$ имеем:

$$\int_{[c, d]} \frac{dx}{|x|^\alpha} = \begin{cases} \frac{1}{1-\alpha} \left(\frac{1}{d^{\alpha-1}} - \frac{1}{c^{\alpha-1}} \right) & \text{при } \alpha \neq 1, \\ \ln d - \ln c & \text{для } \alpha = 1. \end{cases} \quad (\text{IV}, 9; 76)$$

Так как функция $1/x^\alpha \geq 0$, то она интегрируема в $]0, b[$ тогда и только тогда, когда сходится несобственный интеграл

$\int_{]0, b[}$, т. е. тогда и только тогда, когда его первообразные имеют конечный предел, если $x > 0$ стремится к 0. Для этого необходимо и достаточно, чтобы $\alpha < 1$, и тогда

$$\int_{]0, b[} \frac{dx}{x^\alpha} = \int_{]0, b[} \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} \cdot \frac{1}{b^{\alpha-1}}.$$

Тот же самый результат имеет место для интервала $[a, 0]$. Однако $1/|x|^\alpha$ интегрируема в интервале $[a, b]$ тогда и только тогда, когда она интегрируема в $[a, 0]$ и $[0, b]$, т. е. тогда и только тогда, когда $\alpha < 1$.

Пусть теперь $a > 0$. Функция $1/x^\alpha$ будет интегрируемой в $[a, +\infty[$ тогда и только тогда, когда сходится несобственный

интеграл $\int_{|a, \rightarrow +\infty|} dx/x^{\alpha}$, т. е. тогда и только тогда, когда $\alpha > 1$.

При этом

$$\int_{|a, +\infty|} \frac{dx}{x^{\alpha}} = \int_{|a, \rightarrow +\infty|} \frac{dx}{x^{\alpha}} = \frac{1}{\alpha - 1} \cdot \frac{1}{a^{\alpha-1}}.$$

Точно такое же рассуждение можно провести для интервала $]-\infty, a|$, $a < 0$.

Следствие 2. Обозначим через $1/x^{\alpha}$ при комплексном α и вещественном $x > 0$ число $e^{-\alpha \ln x}$. Функция $x \rightarrow 1/x^{\alpha}$ интегрируема по мере Лебега в $|0, a|$, $a > 0$, тогда и только тогда, когда $\operatorname{Re} \alpha < 1$. Эта функция интегрируема в $|a, +\infty|$ тогда и только тогда, когда $\operatorname{Re} \alpha > 1$. Кроме того, интегралы $\int_{|0 \leftarrow, a|} dx/x^{\alpha}$ и $\int_{|a, \rightarrow +\infty|} dx/x^{\alpha}$ сходятся лишь тогда, когда они сходятся абсолютно.

В самом деле, функция $1/x^{\alpha}$ непрерывна, а, значит, измерима. Она интегрируема тогда и только тогда, когда ее абсолютная величина $1/x^{\operatorname{Re} \alpha}$ интегрируема, и мы вернулись к следствию 1.

Для доказательства сходимости интегралов мы используем ту же самую первообразную $\frac{1}{1-\alpha} \cdot \frac{1}{x^{\alpha-1}}$ или $\ln x$. Опираясь на формулу (IV, 9; 76), можно заключить, что сходимость имеет место для $\int_{|0 \leftarrow, a|} \dots$ только при $\alpha < 1$ и для $\int_{|a, +\infty|} \dots$ только при $\alpha > 1$, т. е. тогда, когда имеет место абсолютная сходимость или интегрируемость.

Распространим теперь на несобственные интегралы относительно меры dx методы интегрирования по частям и замены переменной.

Теорема 98 (интегрирование по частям).

1°) Элементарный случай.

Пусть u и v — функции класса C^1 , определенные на интервале $[a, b]$ прямой \mathbb{R} , со значениями в банаховых пространствах \vec{E} и \vec{F} соответственно. Пусть B — непрерывное билинейное отображение $\vec{E} \times \vec{F}$ в банахово пространство \vec{G} . Если выражение $B(\vec{u}(x), \vec{v}(x))$ при $x < b$, стремящемся к b , имеет предел, то $\int_a^b B(\vec{u}, \vec{v}') dx$ существует тогда и только тогда, когда су-

ществует $\int_a^b B(\vec{u}', \vec{v}) dx$, и при этом имеет место формула

$$\int_a^b B(\vec{u}, \vec{v}') dx = \lim_{b' \rightarrow b} B(\vec{u}(b'), \vec{v}(b')) - B(\vec{u}(a), \vec{v}(a)) - \int_a^b B(\vec{u}', \vec{v}) dx. \quad (\text{IV, 9; 77})$$

2°) Общий случай¹⁾.

Пусть $\vec{\mu}$ и $\vec{\nu}$ — меры Радона на $[a, b]$ со значениями соответственно в банаховых пространствах \vec{E} и \vec{F} , имеющие базу ≥ 0 и не имеющие общей точечной массы. Пусть B — непрерывное билинейное отображение пространства $\vec{E} \times \vec{F}$ в банахово пространство \vec{G} .

Пусть \vec{M} и \vec{N} — неопределенные интегралы от $\vec{\mu}$ и $\vec{\nu}$ со значениями в \vec{E} и \vec{F} . Если при этом $B(\vec{M}(x), \vec{N}(x))$ имеет предел при x , стремящемся к b , то $\int_a^b B(\vec{M}, d\vec{\nu})$ существует тогда

и только тогда, когда существует интеграл $\int_a^b B(d\vec{\mu}, \vec{N})$, при этом

$$\int_a^b B(\vec{M}, d\vec{\nu}) =$$

$$= \lim_{b' \rightarrow b} B(\vec{M}(b'), \vec{N}(b')) - B(\vec{M}(a), \vec{N}(a)) - \int_a^b B(d\vec{\mu}, \vec{N}). \quad (\text{IV, 9; 78})$$

Доказательство получается сразу, если написать формулу для пары a, b' и устремить b' к b .

Следствие (критерий Абеля для полусходимости несобственных интегралов). Пусть \vec{u} и \vec{v} — функции, определенные на интервале $[a, b]$ прямой \mathbb{R} , $-\infty < a < b \leq +\infty$, со значениями в банаховых пространствах \vec{E} и \vec{F} . Пусть B — билинейное

¹⁾ В приложениях часто пользуются так называемой «второй формулой о среднем», которая является частным случаем рассматриваемой здесь формулы.

непрерывное отображение $\vec{E} \times \vec{F}$ в банахово пространство \vec{G} . Предположим, что функции u и v обладают следующими свойствами:

1°) функция u имеет ограниченную вариацию на $[a, b]$ и $\vec{u}(x)$ сходится к нулю, когда $x < b$ стремится к b ;

2°) функция v локально dx -интегрируема и имеет ограниченные неопределенные интегралы, т. е. величины $\vec{\sigma}_{c,d} = \int_{[c, d]} v(x) dx$ независимо от выбора c и d ограничены по норме.

Тогда, если пространство \vec{E} конечномерно¹⁾, имеет смысл интеграл $\int_{[a, b]} B(\vec{u}(x), \vec{v}(x)) dx$. Кроме того, если через $U(c)$ обозначить полную вариацию функции u на $]c, b[$ и через $V(c)$ — точную верхнюю грань величин $\|\vec{\sigma}_{c,d}\|$ для $d \in [c, b]$, то имеет место оценка

$$\left\| \int_{[c, b]} B(\vec{u}(x), \vec{v}(x)) dx \right\| \leq \|B\| U(c) V(c), \quad a \leq c < b. \quad (\text{IV}, 9; 79)$$

Для простоты мы проведем доказательство в случае, когда $\vec{E}, \vec{F}, \vec{G}$ совпадают с полем \mathbb{C} , а отображение B является умножением чисел. В этом случае рассматривается интеграл $\int_{[c, b]} uv dx$. По условию функция u имеет ограниченную вариацию, через $U(c)$ обозначается ее полная вариация в $]c, b[$ и считается, что $u(x)$ стремится к 0 при $x < b$, стремящемся к b . Предполагается, кроме того, что функция v локально dx -интегрируема, а неопределенные интегралы

$$V(c) = \sup_{c \leq d < b} \left| \int_{[c, d]} v(x) dx \right|$$

ограничены. Тогда интеграл $\int_{[c, b]} uv dx$ должен иметь смысл

¹⁾ Это ограничение необходимо, если с помощью теоремы 88 хотят получить, что u является неопределенным интегралом от некоторой меры μ с базой ≥ 0 . Однако, если этот факт известен заранее (например, если u принадлежит классу C^1 , — следствие 1 теоремы 89), то в таком ограничении необходимости нет.

и должна иметь место следующая оценка:

$$\left| \int_{[c, \rightarrow b]} uv \, dx \right| \leq U(c) V(c). \quad (\text{IV}, 9; 80)$$

Как мы сейчас увидим, все утверждения опираются на интегрирование по частям. В самом деле, поскольку функция u , которую мы будем обозначать через M , имеет ограниченную вариацию, то $dM = \mu$ есть мера Радона на $[a, b]$ конечной нормы (теорема 88). Точнее: $U(c) = \int_{[c, b]} d|\mu|$. Рассмотрим затем меру $vdx = v$ и через N обозначим какой-нибудь неопределенный интеграл от v . Положим, например, $N(x) = \int_c^x v(t) dt$ так, чтобы $N(c) = 0$. Функция N непрерывна, ограничена и $V(c) = \sup_{c \leq x < b} |N(x)|$. Нам надо вычислить интеграл $\int_{[c, \rightarrow b]} M(x) dN(x)$. Меры μ и v не имеют общей точечной массы, поскольку ее нет у меры v . Поэтому можно применить теорему 98: рассматриваемый интеграл имеет смысл тогда и только тогда, когда имеет смысл интеграл $\int_{[a, \rightarrow b]} N(x) dM(x)$ и первый интеграл равен

$$(\lim_{x \rightarrow b} M(x) N(x)) - M(a) N(a) = \int_{[a, \rightarrow b]} N(x) dM(x). \quad (\text{IV}, 9; 81)$$

Далее, при x , стремящемся к b , функция $M(x)$ стремится к 0. Так как функция $N(x)$ по предположению ограничена, то функция $M(x)N(x)$ также стремится к нулю. Кроме того, интеграл в правой части абсолютно сходится, так как $\mu = dM$ имеет конечную норму и N ограничена. Это означает, что рассматриваемый интеграл также сходится (но не обязательно абсолютно). Поэтому, заменяя a на c и учитывая, что $N(c) = 0$, получаем:

$$\int_{[c, \rightarrow b]} uv \, dx = - \int_{[c, b]} N(x) d\mu(x),$$

а, следовательно,

$$\left| \int_{[c, \rightarrow b]} uv \, dx \right| \leq \int_{[c, b]} |N(x)| d|\mu|(x) \leq V(c) U(c), \quad (\text{IV}, 9; 81_2)$$

чем и доказывается неравенство (IV, 9; 80).

В частном случае, когда $u = M$ принадлежит классу C^1 , формула (IV, 9; 81) записывается в виде:

$$\int_{[a, \rightarrow b]} uv \, dx = \int_{[a, \rightarrow b]} M(x) v(x) \, dx = \\ = (\lim_{x \rightarrow b} M(x) N(x)) - M(a) N(a) - \int_{[a, b]} M'(x) v(x) \, dx$$

и

$$\int_{[c, \rightarrow b]} uv \, dx = - \int_{[c, b]} M' v x \, dx,$$

а, следовательно,

$$\left| \int_{[c, \rightarrow b]} uv \, dx \right| \leq V(c) \int_{[c, b]} |M'(x)| \, dx = U(c) V(c). \quad (\text{IV, 9; 81}_3)$$

Замечание. Теперь становится ясным то, что не было столь очевидным в случае теоремы Абеля для рядов: *суть теоремы Абеля составляет формула интегрирования по частям*. Теорема Абеля о рядах является частным случаем теоремы Абеля относительно интегралов. В самом деле, пусть в условиях теоремы 63 гл. II заданы последовательности \vec{u}_n и \vec{v}_n . Обозначим через \vec{u} и \vec{v} функции, определенные на полуправой $[0, +\infty]$ по формулам:

$$\begin{aligned} \vec{u}(x) &= \vec{u}_n & \text{для } n \leq x < n+1, \\ \vec{v}(x) &= \vec{v}_n & \text{для } n \leq x < n+1. \end{aligned} \quad (\text{IV, 9; 82})$$

Мы видим, что последовательность \vec{u}_n имеет ограниченную вариацию тогда и только тогда, когда имеет ограниченную вариацию на \mathbb{R} функция \vec{u} . Мы видим также, что частные суммы $\sigma_{m,n}$ для последовательности \vec{v}_n ограничены тогда и только тогда, когда ограничены интегралы $\int_{[c, d]} \vec{v}(x) \, dx$. Но тогда из только что доказанной теоремы вытекает теорема 63 гл. II и оценка (II, 14; 31) при $c = m + 1$.

Примеры применения критерия Абеля

Пусть u — некоторая функция, определенная на полуправой $[a, +\infty]$, с комплексными значениями, имеющая ограниченную вариацию и стремящаяся к нулю на бесконечности.

Тогда для вещественного $\lambda \neq 0$ сходится интеграл

$$\int_{[a, \rightarrow +\infty]} e^{i\lambda x} u(x) dx, \quad (\text{IV}, 9; 83)$$

и имеет место неравенство

$$\left\| \int_{[c, \rightarrow +\infty]} e^{i\lambda x} u(x) dx \right\| \leq U(c) \frac{2}{|\lambda|}. \quad (\text{IV}, 9; 84)$$

В самом деле, если мы положим $v(x) = e^{i\lambda x}$, то увидим, что интеграл

$$\sigma_{c, d} = \int_{[c, d]} v(x) dx = \frac{e^{i\lambda d} - e^{i\lambda c}}{i\lambda}$$

допускает оценку: $|\sigma_{c, d}| \leq \frac{2}{|\lambda|}$ ¹⁾.

Естественно, результат не верен при $\lambda = 0$.

Точно так же можно рассмотреть аналогичные интегралы

$$\int_{[a, \rightarrow +\infty]} \cos \lambda x u(x) dx, \quad \int_{[a, \rightarrow +\infty]} \sin \lambda x u(x) dx, \quad (\text{IV}, 9; 85)$$

для которых проводится то же самое рассуждение. Последний интеграл, кроме того, имеет смысл при $\lambda = 0$, так как он тогда равен нулю²⁾.

1) Проведем в этом примере интегрирование по частям, лежащее в основе доказательства теоремы Абеля. Имеем:

$$\int_{[c, d]} e^{i\lambda x} u(x) dx = \left[\frac{e^{i\lambda x} u(x)}{i\lambda} \right]_c^d - \int_{[c, d]} \frac{e^{i\lambda x}}{i\lambda} du(x).$$

Предел существует при d , стремящемся к бесконечности.

2) Если функция u вещественна, ≥ 0 , убывает и стремится на бесконечности к 0, то теорема о знакопеременных рядах снова обеспечивает сходимость. В самом деле, интеграл

$$\int_{\left[\frac{n\pi}{\lambda}, \frac{(n+1)\pi}{\lambda} \right]} u(x) \sin \lambda x dx$$

имеет знак $(-1)^n$, при возрастании n убывает по абсолютной величине (поскольку $u(x + \pi/\lambda) \leq u(x)$) и стремится к 0 при n , стремящемся к бесконечности (этот интеграл мажорируется по модулю величиной $\frac{\pi}{\lambda} u\left(\frac{n\pi}{\lambda}\right)$). Значит, он является общим членом сходящегося ряда. Этим доказывается, что последовательность интегралов $\int_{[a, \frac{n\pi}{\lambda}]} \dots$ при $n \rightarrow \infty$ имеет некоторый предел.

$$\left| a, \frac{n\pi}{\lambda} \right|$$

Во всех трех рассматриваемых случаях, если функция u не интегрируема на $[a, +\infty]$, мы будем иметь дело с несобственными полусходящимися интегралами.

Особый интерес представляют интегралы

$$\int_{[0, \rightarrow +\infty]} \frac{e^{ix}}{x^\alpha} dx, \quad \int_{[0, \rightarrow +\infty]} \frac{\cos \lambda x}{x^\alpha} dx, \quad \int_{[0, \rightarrow +\infty]} \frac{\sin \lambda x}{x^\alpha} dx. \quad (\text{IV}, 9; 86)$$

Функция u , поскольку она имеет особенность в начале $x=0$, не имеет ограниченной вариации. Поэтому эти интегралы условиям теоремы Абеля не удовлетворяют. Однако для обоснования сходимости достаточно их разбить на сумму двух интегралов \int_0^1 и $\int_1^{+\infty}$.

Если считать для первых двух интегралов $\alpha < 1$, а для третьего $\alpha < 2$, то интегралы \int_0^1 будут иметь смысл, поскольку в этом случае мы имеем дело с функциями, интегрируемыми по Лебегу. Остается, следовательно, изучить интеграл $\int_1^{+\infty}$. Что касается этого интеграла, то мы находимся в условиях применимости теоремы Абеля, поскольку функция $1/x^\alpha$, будучи убывающей и ограниченной, имеет ограниченную вариацию и при $\alpha > 0$ стремится к нулю на бесконечности.

Таким образом, рассматриваемые интегралы сходятся: два первых при $0 < \alpha < 1$, а третий при $0 < \alpha < 2$. Заметим, что здесь функция u на интервале $[0, +\infty]$ принадлежит классу C^∞ и что мы рассматриваем «наиболее» классический случай интегрирования по частям! Эти интегралы играют очень важную роль в анализе и физике. Позже мы научимся их вычислять, основываясь на новой функции, так называемой Г-функции. Приведем для примера два из этих интегралов:

$$\int_{[0, \rightarrow +\infty]} \frac{\sin \lambda x}{x} dx = \frac{\pi}{2} \quad \text{для } \lambda > 0, \quad (\text{IV}, 9; 87)$$

$$\int_{[0, \rightarrow +\infty]} \frac{\cos \lambda x}{\sqrt{x}} dx = \int_{[0, \rightarrow +\infty]} \frac{\sin \lambda x}{\sqrt{x}} dx = \sqrt{\frac{\pi}{2\lambda}} \quad \text{при } \lambda > 0. \quad (\text{IV}, 9; 87_2)$$

Однако при $\frac{n\pi}{\lambda} \leq b' \leq \frac{(n+1)\pi}{\lambda}$ интеграл $\int_{[a, b']} \frac{\sin \lambda x}{x^\alpha} dx$ заключен между интегралами $\int_{[a, \frac{n\pi}{\lambda}]} \frac{\sin \lambda x}{x^\alpha} dx$ и $\int_{[a, \frac{(n+1)\pi}{\lambda}]} \frac{\sin \lambda x}{x^\alpha} dx$. Значит, интеграл сходится.

Теорема 99 (о замене переменной). Пусть $[\alpha, \beta]$ и $[a, b]$ — два ограниченных или нет интервала прямой \mathbb{R} . Пусть ξ — такое отображение класса C^1 интервала $[\alpha, \beta]$ на интервал $[a, b]$, что $\xi(\alpha) = a$ и $\lim_{\beta' \rightarrow \beta} \xi(\beta') = b$ ¹⁾. При этих условиях интеграл

$\int_a^b \tilde{f}(x) dx$, где \tilde{f} — локально dx -интегрируемая функция, определенная на $[a, b]$, со значениями в банаховом пространстве \tilde{F} , сходится тогда и только тогда, когда сходится интеграл

$\int_a^b \tilde{f}(\xi(t)) \xi'(t) dt$; при этом имеет место равенство

$$\int_a^b \tilde{f}(x) dx = \int_a^{\beta} \tilde{f}(\xi(t)) \xi'(t) dt. \quad (\text{IV}, 9; 88)$$

Для доказательства достаточно записать формулу для интервала $[\alpha, \beta']$ (следствие 2 теоремы 94) и устремить β' к $\beta^2)$.

Пример. Интегралы Френеля. Так называются интегралы

$$\int_{[0, \rightarrow +\infty]} \cos \lambda x^2 dx \quad \text{и} \quad \int_{[0, \rightarrow +\infty]} \sin \lambda x^2 dx. \quad (\text{IV}, 9; 89)$$

Так как подинтегральные функции при x , стремящемся к бесконечности, не сходятся даже к 0, то сходимость этих интегралов вовсе не очевидна. Однако изучение кривых, изображающих эти функции³⁾, показывает, что они сильно колеблются и что было бы естественным ожидать сходимости этих интегралов. В самом деле, достаточно здесь произвести замену переменной $x = \sqrt{t}$ для того, чтобы прийти к рассмотренным ранее интегралам (IV, 9; 87₂). Исходя из этих формул, справедливость которых будет доказана позже, приходим к следующим результатам:

$$\int_{[0, \rightarrow +\infty]} \cos \lambda x^2 dx = \int_{[0, \rightarrow +\infty]} \sin \lambda x^2 dx = \frac{1}{2} \sqrt{\frac{\pi}{2\lambda}}. \quad (\text{IV}, 9; 90)$$

¹⁾ Относительное расположение α и β , так же как a и b , произвольно.

²⁾ При монотонной функции ξ отсюда легко получаются формулы, соответствующие (IV, 9; 71).

³⁾ Нулями функции $\sin \lambda x^2$ являются точки $\lambda x^2 = n\pi$, или $x = \sqrt{\frac{n\pi}{\lambda}}$, $n = 0, 1, 2, \dots$. Расстояние между двумя последовательными нулями $\sqrt{\frac{\pi}{\lambda}} (Vn+1 - Vn) \sim \sqrt{\frac{\pi}{\lambda}} \frac{1}{2Vn}$ при n , стремящемся к бесконечности, стремится к нулю. Сходимость этих интегралов (см. примечание на стр. 733) вытекает сразу же из теоремы о знакопеременных рядах.

Главное значение в смысле Коши

Пусть задан некоторый интервал $|a, b|$ прямой \mathbb{R} и точка c из $]a, b[$. Пусть \vec{f} — функция, определенная на $|a, b| - \{c\}$, со значениями в банаховом пространстве \vec{F} , интегрируемая по мере Лебега на $|a, c - \delta|$ и на $|c + \delta', b|$ при любых δ и $\delta' > 0$, но не обязательно интегрируемая на $|a, b|$. Когда говорят, что $\int_{|a, b|} \vec{f} dx$ сходится как несобственный интеграл, то

этим хотят сказать, что оба интеграла $\int_{|a, \rightarrow c|} \vec{f} dx$ и $\int_{|c \leftarrow, b|} \vec{f} dx$

сходятся. Это означает, что интеграл $\int_{|a, c - \delta|} + \int_{|c + \delta', b|} \vec{f} dx$ сходится

к некоторому пределу при δ и δ' , стремящихся независимо друг от друга к нулю.

Может случиться, что этот предел не существует, однако, существует предел при δ , стремящемся к нулю, когда $\delta' = \delta$. В этом случае говорят, что *интеграл от \vec{f} сходится в смысле главного значения Коши*, и пишут

$$\text{в. п. } \int_{|a, b|} \vec{f}(x) dx = \lim_{\delta \rightarrow 0} \left(\int_{|a, c - \delta|} + \int_{|c + \delta, b|} \right).$$

Теорема 100. Для того чтобы существовал предел в. п. $\int_{|a, b|} \vec{f}(x) dx$, необходимо и достаточно, чтобы в окрестности точки $x=c$ функция \vec{f} являлась суммой антисимметричной функции \vec{f}_1 ($\vec{f}_1(c+u) = -\vec{f}_1(c-u)$) и такой симметричной функции \vec{f}_2 ($\vec{f}_2(c+u) = \vec{f}_2(c-u)$), для которой сходится интеграл

$\int_{|c \leftarrow, b|} \vec{f}_2(x) dx$.

При $c=0$ симметрия и антисимметрия называются четностью и нечетностью. Заменяя переменную $x=c+u$, можно всегда перейти к последнему случаю, когда $c=0$. Каждая функция \vec{f} единственным образом может быть представлена в окрестности точки $x=0$ в виде суммы четной и нечетной функций:

$$\vec{f}(x) = \frac{\vec{f}(x) + \vec{f}(-x)}{2} + \frac{\vec{f}(x) - \vec{f}(-x)}{2} = \vec{f}_1(x) + \vec{f}_2(x). \quad (\text{IV}, 9; 91)$$

Поскольку \vec{f}_1 нечетна, то ее интеграл в симметричном интервале равен нулю. Следовательно, если α выбрать так, чтобы $a \leq -\alpha < 0 < \alpha \leq b$, то мы получим:

$$\int_{[-\alpha, \delta]} \vec{f}_1(x) dx + \int_{[\delta, +\alpha]} \vec{f}_1(x) dx = 0. \quad (\text{IV}, 9; 92)$$

Сходимость в смысле v. p. в $|a, b|$, или, что то же самое, в интервале $[-\alpha, \alpha]$, одна и та же как для \vec{f} , так и для \vec{f}_2 . Так как функция \vec{f}_2 четна, то

$$\int_{[-\alpha, -\delta]} \vec{f}_2(x) dx + \int_{[\delta, +\alpha]} \vec{f}_2(x) dx = 2 \int_{[\delta, a]} \vec{f}_2(x) dx, \quad (\text{IV}, 9; 93)$$

так что интеграл от \vec{f} существует в смысле v. p. тогда и только тогда, когда сходится интеграл $\int_{[0, \alpha]} \vec{f}_2(x) dx$, чем и заканчивается доказательство теоремы.

Пример. Предположим, что в окрестности точки $x = c$ функция \vec{f} может быть представлена в виде $\vec{f}(x) = \frac{\vec{c}_{-1}}{x - c} + \vec{c}_0 + \vec{c}_1(x - c) + \dots$. Так как функция $1/(x - c)$ антисимметрична, то интеграл в смысле v. p. существует.

Впрочем, если мы будем вычислять $\int_{[c+\delta, b]} \vec{f}(x) dx$, то его можно будет записать в виде

$$\vec{c}_{-1}(\ln(b - c) - \ln \delta) - \int_{[c+\delta, b]} \left(\vec{f}(x) - \frac{\vec{c}_{-1}}{x - c} \right) dx.$$

Второе выражение при δ , стремящемся к нулю, имеет предел (и даже $\vec{f}(x) - \frac{\vec{c}_{-1}}{x - c}$ интегрируема в $|a, b|$). Первое же выражение из-за члена $-\ln \delta$ предела не имеет. Если теперь вычислить $\int_{[a, c-\delta]} \vec{f}(x) dx$, то мы получим

$$\vec{c}_{-1}(\ln \delta - \ln(c - a)) + \int_{[a, c-\delta]} \left(\vec{f}(x) - \frac{\vec{c}_{-1}}{x - c} \right) dx.$$

При этом второе слагаемое имеет предел при δ , стремящемся к нулю, а первое слагаемое предела не имеет из-за наличия

члена $\ln \delta$. Если образовать сумму $\int_{\pm c_1}^{a, c-\delta} + \int_{c+\delta, b}$, то члены $\pm c_1 \ln \delta$ взаимно уничтожаются и сумма при δ , стремящемся к 0, будет иметь предел. Поэтому мы получим:

$$\text{v. p. } \int_{[a, b]} \vec{f}(x) dx = c_1 \ln \left(\frac{b-c}{c-a} \right) + \int_{[a, b]} \left(\vec{f}(x) - \frac{\vec{c}_1}{x-c_1} \right) dx. \quad (\text{IV}, 9; 94)$$

Точно так же для функции \vec{f} , интегрируемой на каждом конечном интервале, существует понятие главного значения интеграла в смысле Коши для интервала $]-\infty, +\infty[$:

$$\text{v. p. } \int_{]-\infty, +\infty[} \vec{f}(x) dx = \lim_{B \rightarrow \infty} \int_{[-B, +B]}.$$

Для того чтобы существовало такое главное значение, необходимо и достаточно, чтобы \vec{f} можно было представить в виде $f = \vec{f}_1 + \vec{f}_2$, где \vec{f}_1 — четная, а \vec{f}_2 — нечетная функции, и чтобы интеграл $\int_{]-\infty, +\infty[} \vec{f}_2(x) dx$ был сходящимся.

Легко подсчитать, что

$$\text{v. p. } \int_{[a, b]} \frac{dx}{x} = \ln \left| \frac{b}{a} \right| \quad (\text{при } a \neq 0, b \neq 0), \quad (\text{IV}, 9; 95)$$

$$\text{v. p. } \int_{]-\infty, +\infty[} \frac{dx}{x} = 0 \quad (\text{функция нечетна}). \quad (\text{IV}, 9; 96)$$

Замечание. Аналогично вводится понятие главного значения в смысле Коши для рядов, множеством индексов которых является множество \mathbb{Z} целых чисел любого знака. Если $\sum_{n=-\infty}^{+\infty} \vec{u}_n$ — такой ряд, то, поскольку $\mathbb{Z} \neq \mathbb{N}$, понятие сходимости для него не определено. Существует лишь понятие безусловной сходимости (эквивалентное абсолютной сходимости, если \vec{u}_n принимают значения в конечномерном пространстве (см. замечание 2°) на стр. 133). Однако обычно говорят, что ряд $\sum_{n=-\infty}^{+\infty} \vec{u}_n$ сходится, если сходится каждый из двух рядов $\sum_{n=0}^{\infty} \vec{u}_n$ и $\sum_{n=0}^{\infty} \vec{u}_{-n}$. Это утверждение означает, что сумма $\sum_{n=-N_1}^{N_2} \vec{u}_n$ имеет предел, когда N_1 и N_2 независимо друг от друга стремятся к $+\infty$.

Однако можно было бы считать, что $\sum_{n=-\infty}^{+\infty} \vec{u}_n$ сходится в смысле главного значения Коши, если сходится сумма $\sum_{n=-N}^{+N} \vec{u}_n$ при N , стремящемся к бесконечности. Если, например, $\vec{u}_n = -\vec{u}_{-n}$, то такая сходимость всегда будет иметь место, а сумма будет равна нулю.

Главное значение в смысле Коши можно определить для кратных интегралов. Если \vec{f} является некоторой функцией на \mathbb{R}^n , суммируемой на каждом множестве $|x| \geq \delta > 0$, то говорят, что интеграл от этой функции сходится в смысле главного значения Коши, если имеет предел при δ , стремящемся к 0, интеграл $\int \dots \int_{|x| \geq \delta} \vec{f}(x) dx$. При $n = 1$ мы снова возвращаемся к главному значению на \mathbb{R} .

Приведем пример, который играет особо важную роль в квантовой механике и в котором используется понятие главного значения Коши.

Теорема 101. Пусть $|a, b|$ — такой интервал прямой \mathbb{R} , что $a < 0 < b$. Пусть \vec{f} — интегрируемая функция, определенная на $|a, b|$, со значениями в банаховом пространстве \vec{F} . Если \vec{f} дифференцируема в начале координат или, более общо, если величина $[\vec{f}(x) - \vec{f}(0)]/|x|^\alpha$, $\alpha > 0$, $x \neq 0$, остается ограниченной, когда x стремится к нулю, то интеграл $\int_{|a, b|} (\vec{f}(x)/x) dx$ существует в смысле главного значения Коши в начале координат. Кроме того, интеграл

$$\int_{|a, b|} \frac{\vec{f}(x)}{x + iy} dx \quad (IV, 9; 97)$$

сходится при y , стремящемся к 0, по значениям знака ϵ ($\epsilon = \pm 1$) к

$$\text{v. p. } \int_{|a, b|} \frac{\vec{f}(x)}{x} dx - \epsilon i \pi \vec{f}(0). \quad (IV, 9; 98)$$

Естественно было ожидать появления члена $\text{v. p. } \int$, а не второго члена $-\epsilon i \pi \vec{f}(0)$, хотя существенным здесь является именно наличие 2-го члена.

Доказательство. Функция $\vec{f}(x)/(x+iy)$ на $|a, -A|$ и $|+A, +b|$, $A > 0$, измерима и мажорируема по норме величиной $\|\vec{f}(x)\|/A$, а, значит, интегрируема. В силу теоремы 35 Лебега, при y , стремящемся к 0, интеграл $\int_{|A, +b|} \frac{\vec{f}(x)}{x+iy} dx$ сходится к $\int_{|A, +b|} \frac{\vec{f}(x)}{x} dx$. Остается показать, что в. р. $\int_{|-A, +A|} \frac{\vec{f}(x)}{x} dx$ существует и что $\int_{|-A, +A|} \frac{\vec{f}(x)}{x+iy} dx$ сходится к в. р. $\int_{|-A, +A|} \frac{\vec{f}(x)}{x} dx - \varepsilon i\pi \vec{f}(0)$.

Запишем очевидное разложение:

$$\int_{|-A, +A|} \frac{\vec{f}(x)}{x+iy} dx = \vec{f}(0) \int_{|-A, +A|} \frac{dx}{x+iy} + \int_{|-A, +A|} \frac{\vec{f}(x) - \vec{f}(0)}{x+iy} dx. \quad (\text{IV}, 9; 99)$$

Изучим сначала его первый член. Для определенности положим $\varepsilon = +1$. Рассмотрим в открытом дополнении к полуправой $\{z = x + iy; y = 0, x \leq 0\}$ комплексной плоскости функцию $\ln z$, определенную на стр. 164. Положим $z = re^{i\varphi}$, где $-\pi < \varphi < \pi$, и $\ln z = \ln r + i\varphi$. Таким образом определенная логарифмическая функция принадлежит классу C^1 , а ее производная равна $1/z$. Следовательно, функция $x \rightarrow \ln(x+iy)$ дифференцируема на прямой при $y \neq 0$ и ее производной является функция $x \rightarrow 1/(x+iy)$. Отсюда следует, что первый интеграл в (IV, 9; 99) равен

$$\vec{f}(0) [\ln(x+iy)]_{x=-A}^{x=+A}. \quad (\text{IV}, 9; 100)$$

Так как $|x+iA| = |x-iA|$, то логарифм модуля исчезает, и выражение (IV, 9; 100) будет равно $\vec{f}(0)i(\varphi_2 - \varphi_1)$, где через φ_2 обозначен аргумент точки $A+iy$, а через φ_1 — аргумент точки $-A+iy$. Если устремить y к 0, то эти аргументы будут стремиться соответственно к 0 и π и, следовательно, первый интеграл будет сходиться к $-i\pi\vec{f}(0)$.

Рассмотрим теперь второй интеграл. При y , стремящемся к 0, функция $x \rightarrow \frac{\vec{f}(x) - \vec{f}(0)}{x+iy}$ просто сходится к функции $x \rightarrow \frac{\vec{f}(x) - \vec{f}(0)}{x}$ всюду, кроме начала координат, где, впрочем, последняя функция не определена. Поэтому можно сказать, что

рассматриваемая функция просто сходится почти всюду к этой функции. Кроме того, она мажорируется по норме функцией $x \rightarrow \frac{\|\vec{f}(x) - \vec{f}(0)\|}{|x|}$, которая является неотрицательной и интегрируемой. Последнее вытекает из того, что функция $x \rightarrow \frac{\|\vec{f}(x) - \vec{f}(0)\|}{|x|}$ мажорируется функцией $\frac{\|\vec{f}(x) - \vec{f}(0)\|}{|x|^a} \cdot \frac{1}{|x|^{1-a}}$,

а последняя не превосходит постоянной, умноженной на $1/|x|^{1-a}$ при $1-a < 1$. Из теоремы Лебега следует, что переход к пределу для второго интеграла оправдан и что при y , стремящемся к 0, он сходится к интегралу.

$$\int_{[-A, +A]} \frac{\vec{f}(x) - \vec{f}(0)}{x} dx. \quad (\text{IV}, 9; 101)$$

Поскольку подинтегральная функция интегрируема, ее интеграл совпадает с ее интегралом в смысле главного значения Коши, а это в силу нечетности $1/x$ дает

$$\text{v. p. } \int_{[-A, +A]} \frac{\vec{f}(x)}{x} dx = \vec{f}(0) \text{ v. p. } \int_{[-A, +A]} \frac{dx}{x} = \text{v. p. } \int_{[-A, +A]} \frac{\vec{f}(x)}{x} dx. \quad (\text{IV}, 9; 102)$$

Складывая теперь оба предела, найденные для двух интегралов в (IV, 9; 99), получаем требуемый результат.

Если y стремится к 0 по значениям < 0 , то φ_2 также стремится к 0, φ_1 стремится к $-\pi$, а величина (IV, 9; 100) стремится к $+i\pi\vec{f}(0)$.

Замечание. Если изложенные результаты применить к скалярной непрерывной функции f с компактным носителем на вещественной прямой, то в этом случае хотелось бы сказать, что мера $dx/(x+iy)$, зависящая от параметра y , при y , стремящемся к 0 со знаком ϵ , широко сходится к мере v. p. $(dx/x) + \epsilon i \delta$. Однако этого делать нельзя, поскольку функция $1/x$ не является локально интегрируемой относительно dx , а следовательно, v. p. (dx/x) не является мерой Радона. Впрочем, предыдущая сходимость имеет место не для всех непрерывных функций f с компактным носителем, а лишь для дифференцируемых функций или функций, удовлетворяющих некоторым дополнительным условиям в начале координат.

В теории распределений можно увидеть, как интерпретируется предыдущий переход к пределу.

§ 10. КРАТНЫЕ ИНТЕГРАЛЫ НА \mathbb{R}^n . ДЛИНЫ, ПЛОЩАДИ И ОБЪЕМЫ В КОНЕЧНОМЕРНОМ АФФИННОМ ЕВКЛИДОВОМ ПРОСТРАНСТВЕ. ЗАМЕНА ПЕРЕМЕННЫХ В КРАТНЫХ ИНТЕГРАЛАХ НА \mathbb{R}^n

Формула (IV, 9; 72) может быть обобщена на кратные интегралы на \mathbb{R}^n относительно меры $dx = dx_1 dx_2 \dots dx_n$.

Теорема 102. Пусть \mathcal{O} и Ω — два открытых множества из \mathbb{R}^n и ξ — гомеоморфизм Ω на \mathcal{O} , принадлежащий вместе со своим обратным гомеоморфизмом классу C^1 . Пусть \vec{f} — непрерывная на \mathcal{O} функция с компактным носителем и значениями в банаховом пространстве \vec{F} . Тогда имеет место формула

$$\int \int \dots \int_{\mathcal{O}} \vec{f}(x) dx = \int \int \dots \int_{\Omega} \vec{f}(\xi(t)) |\det(\xi'(t))| dt, \quad (\text{IV, 10; 1})$$

в которой через dx обозначено $dx_1 dx_2 \dots dx_n$, а через dt обозначено $dt_1 dt_2 \dots dt_n$; что же касается $\det \xi'(t)$, то это — якобиан отображения ξ в точке t .

Доказательство. При $n = 1$ якобиан есть не что иное, как производная, а рассматриваемая теорема переходит в уже доказанную теорему 94. Поэтому дальнейшее доказательство можно будет провести по индукции.

Предположим, что теорема верна для кратного интеграла порядка $n - 1$; докажем ее справедливость для кратного интеграла порядка $n \geq 2$.

Доказательство в одном частном случае. Функция ξ определена системой n функций $\xi_1, \xi_2, \dots, \xi_n$ от n переменных t_1, t_2, \dots, t_n . В рассматриваемом частном случае мы будем проводить доказательство, предполагая, что функция ξ_i равна $\xi_i(t) = t_j$. По теореме Фубини можно записать

$$\begin{aligned} \int \int \dots \int_{\mathcal{O}} \vec{f}(x) dx &= \int da \int \dots \int_{\mathcal{O}(a)} \vec{f}(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) \times \\ &\quad \times dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n, \quad (\text{IV, 10; 2}) \end{aligned}$$

где $\mathcal{O}(a)$ является множеством точек $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ из \mathbb{R}^{n-1} , таких, что $(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) \in \mathcal{O}$. Множество $\mathcal{O}(a)$ является некоторым открытым множеством пространства \mathbb{R}^{n-1} (сечение множества \mathcal{O} гиперплоскостью $x_i = a$, совпадающей с \mathbb{R}^{n-1}). Интегрируемая функция имеет компактный носитель на $\mathcal{O}(a)$, поскольку этот носитель лежит в компакте, являющимся пересечением гиперплоскости с компактным носителем \vec{f} в \mathcal{O} .

По предположению индукции, мы можем с помощью замены переменных преобразовать $(n - 1)$ -кратный интеграл на $\mathcal{O}(a)$. Отображение

$$(t_1, \dots, t_{j-1}, t_{j+1}, \dots, t_n) \rightarrow (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n),$$

где $x_k = \xi_k(t_1, \dots, t_{j-1}, a, t_{j+1}, \dots, t_n)$, $k = 1, \dots, i - 1, i + 1, \dots, n$, является, вместе со своим обратным отображением, гомеоморфизмом, принадлежащим классу C^1 , множества $\Omega(a)$ на $\mathcal{O}(a)$, где $\Omega(a)$ — открытое множество точек $(t_1, \dots, t_{j-1}, t_{j+1}, \dots, t_n)$ из \mathbb{R}^{n-1} , таких, что $(t_1, \dots, t_{j-1}, a, t_{j+1}, \dots, t_n) \in \Omega$. При этом

$$\begin{aligned} & \int \dots \int \tilde{f}(x) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n = \\ & = \int \dots \int \tilde{f}(\xi(t)) \left| \frac{D(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)}{D(t_1, \dots, t_{j-1}, t_{j+1}, \dots, t_n)} \right| \times \\ & \quad \times dt_1 \dots dt_{j-1} dt_{j+1} \dots dt_n. \end{aligned} \quad (\text{IV, 10; 3})$$

Остается подставить это выражение в (IV, 10; 2). Однако, так как $x_i = t_j$, то

$$\frac{D(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)}{D(t_1, \dots, t_{j-1}, t_{j+1}, \dots, t_n)} = (-1)^{i+j} \frac{D(x_1, x_2, \dots, x_n)}{D(t_1, t_2, \dots, t_n)}, \quad (\text{IV, 10; 4})$$

а, значит, модули обоих определителей одни и те же.

Окончательно по формуле Фубини получаем

$$\begin{aligned} & \int \int \dots \int \tilde{f}(x) dx = \\ & = \int da \int \dots \int \tilde{f}(\xi(t_1, \dots, t_{j-1}, a, t_{j+1}, \dots, t_n)) \times \\ & \quad \times |\det \xi'(t_1, \dots, t_{j-1}, a, t_{j+1}, \dots, t_n)| dt_1 \dots dt_{j-1} dt_{j+1} \dots dt_n = \\ & = \int \int \dots \int \tilde{f}(\xi(t)) |\det \xi'(t)| dt, \end{aligned} \quad (\text{IV, 10; 5})$$

чем и заканчивается доказательство формулы (IV, 10; 1) в этом частном случае.

Доказательство общего случая. Пусть τ — некоторая точка Ω . Поскольку ξ и ξ^{-1} принадлежат классу C^1 , то в этой точке $\det \xi'(t)$ отличен от нуля, а значит, существуют хотя бы один индекс i и один индекс j , такие, что частная производная $\frac{\partial \xi_i}{\partial t_j}(\tau) \neq 0$.

Рассмотрим теперь отображение θ_τ , определенное по формулам

$$\begin{aligned} y_l &= t_l, \quad l \neq j, \\ y_j &= \xi_i(t_1, t_2, \dots, t_n). \end{aligned} \quad (\text{IV}, 10; 6)$$

Якобианом этого преобразования в точке τ является не равная нулю частная производная $\partial \xi_i / \partial t_j$. Согласно теореме о неявных функциях, существует такое открытое множество Ω_τ , содержащее τ , что отображение θ_τ является гомеоморфизмом (принадлежащим вместе с обратным гомеоморфизмом классу C^1) этого множества Ω_τ на открытое множество \mathcal{U}_τ из \mathbb{R}^n . Уравнения (IV, 10; 6) при заданном $y = (y_1, y_2, \dots, y_n)$ из \mathcal{U}_τ могут быть разрешены: $t_l = y_l$, $l \neq j$, и $t_j = \eta_j(y_1, y_2, \dots, y_n)$, где η_j принадлежит классу C^1 . Пусть \mathcal{O}_τ — образ Ω_τ при отображении ξ . При переменном τ открытые множества \mathcal{O}_τ образуют покрытие \mathcal{O} и, в частности, покрытие компактного носителя K функции \vec{f} . Но тогда обязательно среди них найдется конечное число множеств, покрывающих K . Обозначим эти множества через \mathcal{O}_{τ_i} , $i \in I$.

Рассмотрим теперь разложение единицы γ_i относительно этого конечного покрытия K . Можно написать:

$$\int \int \dots \int_{\mathcal{O}} \vec{f}(x) dx = \sum_{i \in I} \int \int \dots \int_{\mathcal{O}_{\tau_i}} \gamma_i(x) \vec{f}(x) dx. \quad (\text{IV}, 10; 7)$$

Теперь отображение ξ множества Ω_{τ_i} в \mathcal{O}_{τ_i} может быть разложено на два C^1 -диффеоморфизма, а именно: $\xi = (\xi \circ \theta_{\tau_i}^{-1}) \circ \theta_{\tau_i}$. Первый диффеоморфизм θ_{τ_i} определен формулой (IV, 10; 6), а второй $\zeta_i = \xi \circ \theta_{\tau_i}^{-1}$ может быть записан в виде $(y_1, y_2, \dots, y_n) \rightarrow (x_1, x_2, \dots, x_n)$, где

$$x_k = \xi_k(y_1, \dots, y_{j-1}, \eta_j(y_1, y_2, \dots, y_n), y_{j+1}, \dots, y_n), \quad k \neq i, \quad (\text{IV}, 10; 8)$$

$$x_i = y_j.$$

Теперь можно воспользоваться уже изученным случаем, ибо гомеоморфизм $\xi \circ \theta_{\tau_i}^{-1}$ имеет такой же вид, как и в первом случае. Поэтому можно записать формулу

$$\begin{aligned} \int \int \dots \int_{\mathcal{O}_{\tau_i}} \gamma_i(x) \vec{f}(x) dx &= \\ &= \int \int \dots \int_{\mathcal{U}_{\tau_i}} \gamma_i(\zeta_i(y)) \vec{f}(\zeta_i(y)) |\det \zeta'_i(y)| dy. \quad (\text{IV}, 10; 9) \end{aligned}$$

Гомеоморфизм θ_{τ_i} множества Ω_{τ_i} на \mathcal{U}_{τ_i} имеет такой же, и даже более простой, вид, поскольку здесь имеется лишь замена относительно одной из переменных. Поэтому

$$\begin{aligned} \int \int \dots \int_{\sigma_{\tau_i}} \gamma_i(x) \vec{f}(x) dx &= \\ &= \int \int \dots \int_{\Omega_{\tau_i}} \gamma_i(\xi(t)) \vec{f}(\xi(t)) |\det \zeta'_i(\theta_{\tau_i}(t))| |\det \theta'_{\tau_i}(t)| dt. \end{aligned} \quad (\text{IV}, 10; 10)$$

Из формулы умножения якобианов (следствие 3 теоремы 11 гл. II) теперь следует, что

$$|\det \zeta'_i(\theta_{\tau_i}(t))| \cdot |\det \theta'_{\tau_i}(t)| = |\det \xi'(t)|.$$

Поэтому

$$\begin{aligned} \int \int \dots \int_{\sigma_{\tau_i}} \gamma_i(x) \vec{f}(x) dx &= \\ &= \int \int \dots \int_{\Omega_{\tau_i}} \gamma_i(\xi(t)) \vec{f}(\xi(t)) |\det \xi'(t)| dt. \quad (\text{IV}, 10; 11) \end{aligned}$$

Суммируя по различным значениям i , получаем (IV, 10; 1), и теорема полностью доказана.

З а м е ч а н и е. Предположение биективности ξ , очевидно, необходимо. Так как отображение $t \rightarrow x = t^2$ не является гомеоморфизмом $]-1, +1[$ на $[0, 1[$, то в случае одной переменной

$$\int_{]-1, +1[} 2|t| dt = 2 \neq 1 = \int_{[0, 1[} dx. \quad (\text{IV}, 10; 12)$$

См. по этому поводу формулу (IV, 10; 23₂). Напротив, в весьма частном случае одной переменной, рассмотренном в теореме 93, не было необходимости предполагать, что ξ является гомеоморфизмом, и, положив $x = t^2$, мы получим формулу

$$\int_{-1}^{+1} 2t dt = \int_{+1}^{+1} dx = 0. \quad (\text{IV}, 10; 13)$$

Следствие 1. В условиях теоремы образ меры $|\det \xi'(t)| dt$ при отображении ξ на Ω является мерой $dx = 1(x) dx$ на \mathcal{O} ¹⁾.

¹⁾ Может показаться странным применение различных букв x и t , тем более что \mathcal{O} и Ω лежат в одном и том же пространстве \mathbb{R}^n . Однако из-за формулы (IV, 10; 1) такая запись удобна, поскольку ξ входит в формулу $t \rightarrow x = \xi(t)$. Позже, в следствии 4 и после следствия 5, мы будем применять одну и ту же букву. См. примечание на стр. 443.

Доказательство. В самом деле, если формулу (IV, 10; 1) применить к скалярной непрерывной функции f с компактным носителем, то мы придем к определению образа меры.

Замечание. Поскольку ξ является гомеоморфизмом, то, естественно, образ dx при отображении ξ^{-1} равен $|\det \xi'(t)| dt$. Образ dt при отображении ξ равен $|\det(\xi^{-1})'(x)| dx$.

В соответствии с обозначениями, введенными на стр. 627, можно будет писать $dx = d(\xi(t)) = |\det \xi'(t)| dt$.

Следствие 2. Пусть \tilde{f} — некоторая функция, определенная на \mathcal{O} , со значениями в банаховом пространстве \tilde{F} . Для того чтобы \tilde{f} была интегрируемой по мере dx , необходимо и достаточно, чтобы функция $t \rightarrow \tilde{f}(\xi(t)) |\det \xi'(t)|$ была интегрируемой относительно меры dt , и тогда будет иметь место формула (IV, 10; 1)¹.

Это утверждение является значительным обобщением теоремы, в которой функция \tilde{f} предполагалась непрерывной и имеющей компактный носитель.

Доказательство. Следствие 1 позволяет применить теорему 60: $\tilde{f}(x)$ является dx -интегрируемой тогда и только тогда, когда $\tilde{f}(\xi(t))$ будет $(|\det \xi'(t)| dt)$ -интегрируемой. Согласно теореме 51, это будет тогда и только тогда, когда dt -интегрируема функция $\tilde{f}(\xi(t)) |\det \xi'(t)|$.

Следствие 3. Пусть A — часть Ω , измеримая по мере Лебега в \mathbb{R}^n . Тогда ее образ $\xi(A) \subset \mathcal{O}$ измерим по мере Лебега и имеет меру, равную $\int\limits_{\xi(A)} dx = \int\limits_A |\det(\xi'(t))| dt$.

Для доказательства достаточно применить формулу (IV, 10; 1) к характеристической функции множества $\xi(A)$.

Следствие 4. Мера Лебега $dx = dx_1 dx_2 \dots dx_n$ на векторном пространстве \mathbb{R}^n инвариантна относительно любого сдвига \mathbb{R}^n . Мера подмножества A из \mathbb{R}^n равна мере его сдвигов. Другими словами, какими бы ни были элемент \vec{a} из \mathbb{R}^n и сдвиг $\tau_{\vec{a}}$: $x \rightarrow x + \vec{a}$, имеем $\tau_{\vec{a}}(dx) = dx$.

¹) На этот раз мы берем меры dx и dt , т. е. одну и ту же меру в \mathbb{R}^n .

Доказательство. Якобиан сдвига $\xi = \tau_{\vec{a}}$, где $\xi(t) = t + \vec{a}$, равен 1 в любой точке пространства \mathbb{R}^n .

Позже мы увидим, что мера dx , с точностью до множителя пропорциональности, является единственной мерой, инвариантной относительно сдвигов: любая комплексная мера Радона на \mathbb{R}^n , инвариантная относительно сдвигов, имеет вид $c dx$, где c — некоторая комплексная постоянная. Более общо, любая мера Радона на \mathbb{R}^n со значениями в нормированном векторном пространстве \vec{E} , инвариантная относительно всех сдвигов, имеет вид $\vec{c} \vec{dx}$, где \vec{c} — фиксированный вектор из \vec{E} .

Следствие 5. Пусть ξ — линейное отображение пространства \mathbb{R}^n в себя. Если A — множество из \mathbb{R}^n , измеримое по мере Лебега в \mathbb{R}^n , то его образ $\xi(A)$ также измерим по мере Лебега и, кроме того, (мера $\xi(A)$) = $|\Delta|$ (мера A), где Δ — определитель линейного отображения ξ .

Доказательство. Так как якобиан линейного отображения является постоянной величиной, равной Δ , то достаточно применить следствие 3; в нем предполагается, что ξ является гомеоморфизмом, а, значит, $\Delta \neq 0$. Однако результат остается верным, если $\Delta = 0$, поскольку здесь мера $\xi(\bar{A})$ равна нулю, ибо $\xi(\bar{A})$ содержится в векторном подпространстве пространства \mathbb{R}^n размерности $< n$ (см., например, следствие 1 теоремы 102₂).

Здесь мы получаем новое важное истолкование модуля определителя линейного отображения: $|\Delta| = \frac{\text{мера } \xi(A)}{\text{мера } A}$ для каждого измеримого по мере Лебега множества A . Пользуясь языком следствия 1, мы можем сказать, что образ меры $|\Delta| dx$ при отображении ξ является мерой dx или образ меры dx при линейном отображении ξ с отличным от нуля определителем $|\Delta|$ является мерой $\frac{1}{|\Delta|} dx$. (Если A является измеримым по мере Лебега множеством \mathbb{R}^n , то его мера относительно dx должна равняться мере его образа $\xi(A)$ по мере образа $\xi(dx)$. Это действительно так, поскольку $\xi(dx) = \frac{1}{|\Delta|} dx$ и поскольку мера $\xi(A)$ по dx равна определителю $|\Delta|$, умноженному на меру A .)

Если ξ является гомотетией с отношением k , то мера $\xi(A)$ по мере Лебега равна $|\Delta| = |k|^n$, умноженному на меру A :

$$\text{мера } kA = |k|^n \times \text{мера } A.$$

Следствие 5₂. *Мера Лебега n -параллелепипеда¹⁾ в \mathbb{R}^n с вершиной a , определенного векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$, равна модулю определителя этих векторов.*

Пусть X_{ij} , $j = 1, 2, \dots, n$, — координаты векторов \vec{X}_i . Определителем этих векторов является, по определению, определитель, составленный из чисел X_{ij} .

Рассмотрим теперь линейное отображение ξ , преобразующее i -й вектор \vec{e}_i канонического базиса \mathbb{R}^n в вектор \vec{X}_i . Так как векторы \vec{e}_i определяют n -параллелепипед объема 1 (этот параллелепипед является произведением $[0, 1] \times [0, 1] \times \dots \times [0, 1]$, а, следовательно, его мера равна n -й степени меры Лебега отрезка $[0, 1]$ прямой \mathbb{R} , т. е. равна 1), то из следствия 5 вытекает, что искомый объем равен $|\det \xi|$ — модулю определителя n векторов \vec{X}_i .

Следствие 6. *Пусть a — некоторая точка Ω , A — часть Ω , измеримая по мере Лебега. Тогда отношение мер Лебега множеств $\xi(A)$ и A сходится к $|\det \xi'(a)|$, когда множество A равномерно стягивается к точке a .*

Переменное множество A мы называем *равномерно стягивающимся к a* , если точная верхняя грань расстояния от a до точек множества A стремится к 0.

Доказательство. Запишем разность $\varepsilon = \frac{\text{мера } \xi(A)}{\text{мера } A} - |\det \xi'(a)|$ в виде

$$\begin{aligned} \varepsilon &= \frac{\iiint \dots \int \det(\xi'(t)) dt}{\iint \dots \int_A dt} - |\det \xi'(a)| = \\ &= \frac{\iint \dots \int_A (|\det(\xi'(t))| - |\det(\xi'(a))|) dt}{\iint \dots \int_A dt}; \quad (\text{IV}, 10; 14) \end{aligned}$$

¹⁾ В аффинном пространстве E над полем вещественных чисел p -параллелепипедом с вершиной a , определенным векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_p$, называется множество элементов из E вида $a + t_1 \vec{X}_1 + t_2 \vec{X}_2 + \dots + t_p \vec{X}_p$, где $t_i \in [0, 1]$, $i = 1, 2, \dots, p$. Векторы эти не обязательно должны быть независимыми. Число p может быть больше размерности E . В настоящем следствии задано n векторов в \mathbb{R}^n . Если они зависимы, то следствие применимо, и мера n -параллелепипеда будет равной нулю.

отсюда следует, что эта разность допускает оценку:

$$|\varepsilon| \leq \sup_{t \in A} ||\det(\xi'(t))| - |\det(\xi'(a))||. \quad (\text{IV}, 10; 15)$$

Из полученного неравенства вытекает утверждение леммы, если учесть, что по условию ξ принадлежит классу C^1 и что, следовательно, функция $t \rightarrow |\det \xi'(t)|$ непрерывна.

Замечание. Доказанное следствие дает важную интерпретацию для $|\det(\xi'(a))|$. Интуитивно это можно представлять следующим образом. Пусть $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ — канонический базис \mathbb{R}^n . Его образ при производном отображении $\xi'(a)$ образован векторами $\frac{\partial \vec{\xi}}{\partial t_1}(a), \frac{\partial \vec{\xi}}{\partial t_2}(a), \dots, \frac{\partial \vec{\xi}}{\partial t_n}(a)$. Если в качестве множества A взять малый n -параллелепипед с вершиной в a , определенной векторами $k\vec{e}_i$ при малом k , то его образ будет «приближенно» представлять собой n -параллелепипед с вершиной $\xi(a)$, определяемый векторами $k \frac{\partial \vec{\xi}}{\partial t_i}(a)$. Отношение объемов этих n -параллелепипедов равно модулю определителя векторов $\frac{\partial \vec{\xi}}{\partial t_i}(a)$, т. е. $|\det(\xi'(a))|$.

Однако здесь мы очень далеки от строгого доказательства следствия 6.

Пример замены переменных. Вычисление интеграла в полярных координатах. Займемся сначала полярными координатами на плоскости. Здесь мы имеем дело с отображением $P: (r, \varphi) \rightarrow (x, y)$ плоскости \mathbb{R}^2 в \mathbb{R}^2 , изученным в гл. III на стр. 242,

$$x = r \cos \varphi, \quad y = r \sin \varphi. \quad (\text{IV}, 10; 16)$$

Якобиан отображения P равен r . Если мы перейдем к сужению отображения на открытое множество Ω , определяемое неравенствами $r > 0$ и $0 < \varphi < 2\pi$, то P будет гомеоморфизмом, принадлежащим вместе со своим обратным гомеоморфизмом классу C^1 и отображающим Ω на свой образ \mathcal{O} , являющийся в \mathbb{R}^2 дополнением к множеству $y = 0, x \geq 0$. Пусть теперь \tilde{f} — некоторая функция $(x, y) \rightarrow \tilde{f}(x, y)$, определенная на \mathbb{R}^2 , со значениями в банаевом пространстве \tilde{F} . Она будет интегрируемой на \mathbb{R}^2 тогда и только тогда, когда она интегрируема на \mathcal{O} ; при этом ее интегралы на \mathbb{R}^2 и \mathcal{O} будут одними и теми же, поскольку дополнение к \mathcal{O} имеет нулевую меру Лебега в \mathbb{R}^2 .

Согласно следствию 2, эта функция интегрируема на \mathcal{O} тогда и только тогда, когда будет интегрируемой на Ω функция $(r, \varphi) \rightarrow \tilde{f}(r \cos \varphi, r \sin \varphi)r$, причем справедлива формула

$$\int \int_{\mathbb{R}^2} \tilde{f}(x, y) dx dy = \int \int_{\Omega} \tilde{f}(r \cos \varphi, r \sin \varphi) r dr d\theta. \quad (\text{IV}, 10; 17)$$

В этой формуле интеграл по Ω можно заменить на интеграл по области $\bar{\Omega}$, т. е. по множеству, определенному расширенными неравенствами $r \geq 0$, $0 \leq \varphi \leq 2\pi$, ибо по мере $dr d\varphi$ мера множества $\bar{\Omega} - \Omega = \dot{\Omega}$, определяемого соотношениями $\{r = 0, 0 \leq \varphi \leq 2\pi \text{ или } r > 0, \varphi = 0, \text{ или } r > 0, \varphi = 2\pi\}$, равна нулю.

Часто бывает выгодно для вычисления якобиана применить геометрический метод. Выполним эти вычисления на примере полярных координат в трехмерном пространстве \mathbb{R}^3 . У нас имеется отображение P пространства \mathbb{R}^3 в \mathbb{R}^3 , определяемое по формулам $m(r, \theta, \varphi) \rightarrow M(x, y, z)$, где

$$\begin{aligned} x &= r \sin \theta \cos \varphi, \\ y &= r \sin \theta \sin \varphi, \\ z &= r \cos \theta. \end{aligned} \quad (\text{IV}, 10; 13)$$

Если ограничиться открытым множеством

$$\Omega = \{r, \theta, \varphi; r > 0, 0 < \theta < \pi, 0 < \varphi < 2\pi\},$$

то P будет представлять собой C^1 -диффеоморфизм множества Ω на его образ — открытое множество \mathcal{O} из \mathbb{R}^3 , являющееся дополнением к полуплоскости $\{x, y, z; y = 0, x \geq 0\}$.

Как и в предыдущем случае, $C\mathcal{O}$ и $\dot{\Omega} = \bar{\Omega} - \Omega$ имеют нулевую меру.

Частная производная $\partial \vec{P} / \partial r$ в точке $m = (r, \theta, \varphi)$ является единичным вектором \vec{i} , направленным по радиус-вектору точки $M = (x, y, z) = P(m)$. Частная производная $\partial \vec{P} / \partial \theta$ является вектором \vec{rj} , где \vec{j} — единичный вектор касательной к меридиану в точке M в направлении возрастания значений угла θ . Частная производная $\partial \vec{P} / \partial \varphi$ является вектором $r \sin \theta \vec{k}$, где \vec{k} — единичный вектор, направленный по параллели, проходящей через M , в направлении возрастания значений φ . Поэтому в дифференциальных обозначениях можно написать формулу

$$d\vec{M} = P'(m) d\vec{m} = \vec{i} dr + \vec{rj} d\theta + r \sin \theta \vec{k} d\varphi. \quad (\text{IV}, 10; 19)$$

Якобианом отображения P в точке $m = (r, \theta, \varphi)$ является определитель системы трех векторов $\frac{\partial \vec{P}}{\partial r}, \frac{\partial \vec{P}}{\partial \theta}, \frac{\partial \vec{P}}{\partial \varphi}$ по отношению к векторам базиса \mathbb{R}^3 . Этот определитель равен произведению $r^2 \sin \theta$ на определитель системы векторов $\vec{i}, \vec{j}, \vec{k}$. Последний же, как определитель системы трех единичных взаимно перпендикулярных векторов, равен $\pm 1^1$). Поэтому имеем

$$\det P'(r, \theta, \varphi) = \pm r^2 \sin \theta \neq 0.$$

Формула замены переменных в полярных координатах в трехмерном пространстве имеет, таким образом, следующий вид:

$$\begin{aligned} \iiint_{\mathbb{R}^3} \vec{f}(x, y, z) dx dy dz &= \\ &= \iint_{\bar{\Omega}} \vec{f}(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta dr d\theta d\varphi. \end{aligned} \quad (\text{IV, 10; 20})$$

Обобщение теоремы 102 и следствий 1, 2, 3, 6. Теорема 102 и следствия 1, 2, 3 остаются справедливыми, если ξ является гомеоморфизмом, принадлежащим классу C^1 . Обратный гомеоморфизм ξ^{-1} при этом не обязан принадлежать классу C^1 , т. е. $\det \xi'(t)$ для некоторых значений $t \in \Omega$ может обращаться в нуль. Следствие 6 остается справедливым, если ξ является некоторым отображением класса C^1 и не обязательно представляет собой гомеоморфизм. Эти факты мы примем без доказательства. Из них теперь можно получить следующую теорему:

Теорема 102₂. Пусть ξ — некоторое отображение класса C^1 открытого множества Ω из \mathbb{R}^n в \mathbb{R}^n . Обозначим через A множество точек Ω , в которых якобиан отображения ξ равен нулю. Тогда его образ $\xi(A)$ относительно меры Лебега имеет меру нуль в \mathbb{R}^n .

Доказательство. Ограничимся для простоты случаем, когда $\Omega = \mathbb{R}^n$. Меру Лебега некоторого множества будем обозначать через mes . Пусть $a \in A$. Согласно следствию 6, распространенному, как только что было сказано, на произвольное отображение ξ класса C^1 , существует при заданном $\varepsilon > 0$

¹⁾ Это означает, что параллелепипед, определяемый векторами $\vec{i}, \vec{j}, \vec{k}$, имеет объем, равный 1. Конечно, этот факт «известен», но не забывайте, что до построения теории интегрирования теория мер объемов не была строго обоснована! Это будет сделано в этом параграфе.

такой открытый шар B_a с центром в a , что для любой части B шара B_a , измеримой по мере Лебега в \mathbb{R}^n , имеет место неравенство

$$\text{mes } \xi(B) \leq \varepsilon \text{ mes } B. \quad (\text{IV}, 10; 21)$$

Пусть K — компакт множества A . Тогда B_a , $a \in K$, образуют открытое покрытие K . Для него существует некоторое конечное подпокрытие, например B_1, B_2, \dots, B_N . Положим $B'_1 = K \cap B_1, B'_2 = K \cap B_2 \cap CB'_1, B'_3 = K \cap B_3 \cap CB'_1 \cap CB'_2$ и т. д. Тогда множества B'_1, B'_2, \dots, B'_N , являющиеся пересечениями измеримых множеств, будут измеримыми. Кроме того, они не пересекаются и образуют покрытие K . Но тогда $\xi(B'_1), \xi(B'_2), \dots, \xi(B'_N)$ образуют покрытие $\xi(K)$.

Поэтому имеем

$$\text{mes } \xi(K) \leq \sum_{i=1}^N \text{mes } \xi(B'_i) \leq \varepsilon \sum_{i=1}^N \text{mes } (B'_i) = \varepsilon \text{ mes } K. \quad (\text{IV}, 10; 22)$$

Так как ε произвольно, то это означает, что $\text{mes } \xi(K) = 0$.

Заметим теперь, что множество A замкнуто, поскольку оно является прообразом множества $\{0\}$ при непрерывном отображении $t \rightarrow \det \xi(t)$ пространства \mathbb{R}^n в \mathbb{R} . Если через K_m обозначить пересечение A с шаром $\|\vec{x}\| \leq m$ из \mathbb{R}^n , то K_m будет компактом A , а само A — объединением K_m ¹⁾. Каждое множество $\xi(K_m)$ имеет, как мы только что видели, нулевую меру. Множество $\xi(A)$, представляющее собой объединение счетного числа множеств $\xi(K_m)$ нулевой меры, имеет меру, равную нулю.

Следствие 1. Пусть V — многообразие класса C^1 , счетное в бесконечности. Пусть H — отображение класса C^1 множества V в \mathbb{R}^n . Если размерность $V \leq n - 1$, то его образ $H(V)$ имеет нулевую меру Лебега в \mathbb{R}^n . В частности, если множество V является многообразием класса C^1 в \mathbb{R}^n размерности $\leq n - 1$, то оно имеет меру, равную нулю.

Доказательство. Пусть l — размерность многообразия V , $l \leq n - 1$. Пусть $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ — карта многообразия V . Тогда сложное отображение $H \circ \Phi$ является отображением \mathcal{O} в \mathbb{R}^n , принадлежащим классу C^1 .

Если положить

$$\xi(t, u) = (H \circ \Phi)(t) \quad \text{для } t \in \mathcal{O}, u \in \mathbb{R}^{n-l}; \quad (\text{IV}, 10; 23)$$

¹⁾ В этом месте использовано упрощенное предположение $\Omega = \mathbb{R}^n$. Оно позволяет непосредственно показать, что A является счетным объединением компактов. Однако утверждение справедливо и в общем случае.

то рассматриваемое отображение продолжается до отображения класса C^1 открытого множества $\mathcal{O} \times \mathbb{R}^{n-l}$ из \mathbb{R}^n в пространство \mathbb{R}^n .

Так как $d\xi/di \equiv 0$, то якобиан этого отображения будет равен нулю и, следовательно, согласно теореме, образ $H \circ \Phi(\mathcal{O}) = \xi(\mathcal{O} \times \mathbb{R}^{n-l})$ имеет меру Лебега, равную нулю. Это означает, что образ при отображении H образа $\Phi(\mathcal{O})$ карты Φ имеет нулевую меру. Пусть теперь K является некоторым компактом V . Так как он может быть покрыт конечным числом образов карт, то его образ при отображении H заведомо имеет нулевую меру. Поскольку V , по предположению, является объединением счетного множества компактов, то теперь ясно, что образ V при отображении H также имеет меру, равную нулю.

Замечание. Не следует думать, что этот результат остается верным для многообразия класса C^0 , т. е. топологических не дифференцируемых многообразий. Если, например, в плоскости мы рассмотрим *кривую Жордана*, т. е. множество, гомеоморфное отрезку прямой $[a, b]$, и если гомеоморфизм H отрезка $[a, b]$ на эту дугу только непрерывен, но не дифференцируем, то мера этой кривой Жордана по мере $dx \otimes dy$ не обязательно равна нулю. Непрерывное отображение H отрезка \mathbb{R} в \mathbb{R}^n является путем в \mathbb{R}^n в смысле определения, данного на стр. 90. Путь же, если он только непрерывен; т. е. не дифференцируем, может иметь образ, целиком заполняющий куб в \mathbb{R}^n , с мерой >0 .

Следствие 2. Пусть ξ — отображение класса C^1 открытого множества Ω из \mathbb{R}^n в открытое множество \mathcal{O} из \mathbb{R}^n . Обозначим для каждой точки $x \in \mathcal{O}$ через $v(x)$ число точек прообраза $\xi^{-1}(\{x\})$. Отображение ξ является собственным относительно меры $|\det \xi'(t)| dt$ множества Ω тогда и только тогда, когда v локально dx -интегрируема в \mathcal{O} ; в этом случае образ меры $|\det \xi'(t)| dt$ множества Ω при отображении ξ является мерой $v(x) dx$ множества \mathcal{O} . При этом, если \tilde{f} — функция, определенная на \mathcal{O} , со значениями в банаховом пространстве \tilde{F} , то функция $v(x)\tilde{f}(x)$ является dx -интегрируемой тогда и только тогда, когда dt -интегрируема функция $\tilde{f}(\xi(t)) |\det \xi'(t)|$, причем имеет место равенство

$$\int \int \dots \int_{\mathcal{O}} \tilde{f}(x) v(x) dx = \int \int \dots \int_{\Omega} \tilde{f}(\xi(t)) |\det \xi'(t)| dt. \quad (\text{IV}, 10; 23_2)$$

Полагая здесь $v = 1$, мы снова получим следствия 1 и 2 теоремы 102.

Измерение объемов в аффинных евклидовых конечномерных пространствах

Пусть E — аффинное евклидово пространство размерности N . Возможность измерения объемов в таком пространстве кажется интуитивно очевидной; остается обосновать ее со всей строгостью.

Пусть $0, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_N$ — ортонормированная система координат в E . С ее помощью можно отождествить E с \mathbb{R}^N и таким образом определить меру как тензорное произведение $dx = dx_1 \otimes dx_2 \otimes \dots \otimes dx_N$.

Более строго, следует говорить, что выбор системы координат определяет линейную биекцию $P: (x_1, x_2, \dots, x_N) \rightarrow a + x_1\vec{e}_1 + \dots + x_N\vec{e}_N$ пространства \mathbb{R}^N на E и что при этих условиях можно определить образ меры $P(dx)$ при отображении P меры Лебега $dx = dx_1 dx_2 \dots dx_n$ пространства \mathbb{R}^n . Допускаемая нами вольность речи заключается в том, что вместо $P(dx)$ мы пишем просто dx .

Теорема 103 (основная теорема измерения объемов). *Пусть E — аффинное евклидово пространство размерности N . Мера, определенная в ортонормированной системе координат по формуле $dx = dx_1 dx_2 \dots dx_N$, не зависит от выбранной системы координат.*

Доказательство. Пусть $a, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_N$ — некоторая ортонормированная система координат, а $b, \vec{f}_1, \vec{f}_2, \dots, \vec{f}_N$ — другая ортонормированная система. Координаты x и y одной и той же точки в этих системах координат связаны формулой замены переменных:

$$y_i = a_i + \sum_{j=1}^N m_{ij} x_j, \quad (IV, 10; 24)$$

$$a_i = (\overrightarrow{a - b} | \vec{f}_i), \quad m_{ij} = (\vec{e}_j | \vec{f}_i).$$

Согласно следствию 1 теоремы 102 о замене переменных в кратных интегралах имеем: $dy = |\Delta| dx^1$, где Δ — определитель матрицы \mathcal{M} , составленной из чисел m_{ij} . Однако определитель ортогональной матрицы равен ± 1 , так как она удовлетворяет условию $\mathcal{M}^{-1}\mathcal{M} = I$, где \mathcal{M} — транспонированная,

¹⁾ Здесь допускается отмеченная вольность речи. На самом деле две системы координат определяют две биекции P и Q пространства \mathbb{R}^N на E и $P = Q \circ \mathcal{M}$, где \mathcal{M} является отображением \mathbb{R}^N на \mathbb{R}^N , определенным по формуле (IV, 10; 24). Теперь $P(dx) = Q(\mathcal{M}(dx)) = Q(|\Delta^{-1}| dy)$, или $|\Delta| P(dx) = Q(dy)$, что мы записали в виде $|\Delta| dx = dy$.

а I — единичная матрицы, и поэтому $\Delta^2 = 1$ ¹⁾. Абсолютная величина этого определителя равна 1. Отсюда следует, что $dx = dy$, и теорема доказана.

Итак, в аффинном евклидовом пространстве существует вполне определенная мера Радона $\geqslant 0$, которую мы будем обозначать через dx . Эта мера инвариантна относительно сдвигов и, более общо, относительно движений пространства, т. е. относительно аффинных обратимых преобразований, сохраняющих скалярные произведения. Так как эта мера в ортонормированной системе координат равна $dx_1 dx_2 \dots dx_N$, то к ней применимы следствия 5 и 5₂ теоремы 102.

Пусть теперь A — некоторое измеримое множество из E . Его объемом в евклидовом пространстве называется его мера относительно dx . Если A является прямоугольным параллелепипедом со сторонами a_1, a_2, \dots, a_N , то его объем, очевидно, равен произведению $a_1 a_2 \dots a_N$, если выбрать ортонормированную систему координат, базисные векторы которой параллельны его сторонам.

Теорема 103₂. *Если $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_N$ являются системой N векторов с общим началом, то объем N -параллелепипеда, образованного этими векторами, вычисляется по формуле:*

$$V^2 = \begin{vmatrix} (\vec{X}_1 | \vec{X}_1) & (\vec{X}_1 | \vec{X}_2) & \dots & (\vec{X}_1 | \vec{X}_N) \\ (\vec{X}_2 | \vec{X}_1) & (\vec{X}_2 | \vec{X}_2) & \dots & (\vec{X}_2 | \vec{X}_N) \\ \dots & \dots & \dots & \dots \\ (\vec{X}_N | \vec{X}_1) & (\vec{X}_N | \vec{X}_2) & \dots & (\vec{X}_N | \vec{X}_N) \end{vmatrix}. \quad (\text{IV}, 10; 25)$$

Доказательство. Выберем ортонормированную систему координат. Мы знаем, что V^2 является квадратом определителя системы N векторов (следствие 5₂ теоремы 102). Поэтому, для того чтобы получить предыдущую формулу, достаточно вычислить по правилам умножения определителей произведение определителя этой системы на себя.

Эта формула, очевидно, удобнее, чем формула, использующая определитель n векторов относительно некоторой системы координат, поскольку она включает лишь скалярные произведения и потому не зависит от какой-либо системы координат.

¹⁾ Невольно чувствуешь большой соблазн доказать это свойство определителя ортогональных матриц, опираясь на геометрическое понятие объема! Однако понятие объема не дано свыше и должно быть введено со всей строгостью с помощью той же теории интегрирования! Поэтому надо иметь алгебраическое доказательство свойства $\Delta = \pm 1$.

Измерение длин в аффинном евклидовом пространстве

Аффинное евклидово пространство нормировано, а, значит, в нем можно измерять длины кривых (см. стр. 701). Если $t \rightarrow M(t)$ при t , пробегающем некоторый интервал \mathbb{R}_1 прямой \mathbb{R} , является параметрически заданной кривой, принадлежащей классу C^1 , то длина этой кривой равна интегралу $\int_{\mathbb{R}_1} \| \overrightarrow{M'(t)} \| dt$ (формула (IV, 9; 40₂)), равному $\int_{\mathbb{R}_1} (\overrightarrow{M'(t)} | \overrightarrow{M'(t)})^{1/2} dt$. Если x_i — составляющие M в ортонормированной системе координат, то это дает $\int_{\mathbb{R}_1} \sqrt{\sum_{i=1}^N x_i'^2(t)} dt$.

Измерение n -мерных площадей в линейном многообразии размерности n аффинного евклидова конечномерного пространства

В аффинном евклидовом конечномерном пространстве мы смогли определить, с одной стороны, длины дуг кривых, а с другой, — меры объемов. Однако нам бы хотелось определить все промежуточные величины — площади поверхностей и, более общо, n -мерные площади n -мерных многообразий при $0 < n \leq N$.

Займемся для начала случаем, когда F является линейным n -мерным многообразием аффинного пространства E . В этом случае само F является некоторым аффинным евклидовым пространством, и, следовательно, в F существуют меры объемов. Если в F выбрать ортонормированную систему координат и отождествить его с \mathbb{R}^n , то эта мера равна $dx_1 dx_2 \dots dx_n$. Однако, поскольку F является пространством n измерений, то вместо слова «объем» мы будем употреблять слово « n -мерная площадь», или « n -площадь». Если F_1 и F_2 — два линейных многообразия одной и той же размерности n , то каждое аффинное отображение F_1 на F_2 , сохраняющее скалярные произведения, переводит ортонормированную систему координат пространства F_1 в ортонормированную систему координат пространства F_2 , а, следовательно, сохраняет n -площади. Каждое аффинное отображение, являющееся подобием с отношением k , т. е. отображением, умножающим длины на k и скалярные произведения на k^2 , имеет якобиан относительно ортонормированных систем координат F_1 и F_2 , равный $\pm k^n$, а, следовательно, умножает n -площади на $|k|^n$. Если $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$ — векторы из F с общим началом, то n -мерная площадь определяемого ими n -парал-

лелепипеда задается формулой:

$$S^2 = \begin{vmatrix} (\vec{X}_1 | \vec{X}_1) & (\vec{X}_1 | \vec{X}_2) & \dots & (\vec{X}_1 | \vec{X}_n) \\ (\vec{X}_2 | \vec{X}_1) & (\vec{X}_2 | \vec{X}_2) & \dots & (\vec{X}_2 | \vec{X}_n) \\ \dots & \dots & \dots & \dots \\ (\vec{X}_n | \vec{X}_1) & (\vec{X}_n | \vec{X}_2) & \dots & (\vec{X}_n | \vec{X}_n) \end{vmatrix}_{i, j} = \det(\vec{X}_i | \vec{X}_j). \quad (\text{IV}, 10; 26)$$

Упражнение. Иногда полезно предыдущую формулу записать в иной форме. Предположим, например, что речь идет о двух векторах \vec{X} и \vec{Y} . Тогда она запишется в виде

$$S^2 = (\vec{X} | \vec{X})(\vec{Y} | \vec{Y}) - (\vec{X} | \vec{Y})^2. \quad (\text{IV}, 10; 27)$$

В силу тождества Лагранжа ее можно переписать так:

$$\sum_{i, j} (X_i Y_j - X_j Y_i)^2, \quad (\text{IV}, 10; 28)$$

где X_i (соответственно Y_i) — координаты векторов \vec{X} (соответственно \vec{Y}) относительно некоторой ортонормированной системы координат. Таким образом, S^2 является суммой квадратов площадей параллелограммов — ортогональных проекций исходного параллелограмма на плоскости, порожденные всевозможными парами векторов базиса. Эта формула обобщается на n -мерные площади при произвольном n .

Теорема 104. n -мерная площадь n -параллелепипеда равна произведению $(n-1)$ -мерной площади какого-либо из его оснований на длину соответствующей высоты.

Доказательство. n -параллелепипед определяется своей вершиной и n векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$. Предположим сначала, что эти n векторов линейно зависимы. Тогда n -мерная площадь параллелепипеда равна нулю. Однако в этом случае или же $n-1$ векторов $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$ зависимы, и тогда $(n-1)$ -мерная площадь основания равна нулю, или же они независимы, и тогда вектор \vec{X}_n лежит в векторном подпространстве \vec{F}_n , порожденном векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, а, значит, в нуль обращается высота. В обоих случаях утверждение теоремы справедливо.

Предположим теперь, что исходные n векторов линейно независимы. Рассмотрим основание параллелепипеда, определенное его вершиной и векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$. Выберем такой ортонормированный базис в \vec{F} , в котором первые $n-1$

элементов лежат в векторном подпространстве \vec{F}_{n-1} , порожденном векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, а n -й вектор базиса является единичным вектором e_n , ортогональным к пространству \vec{F}_{n-1} . Далее, n -мерная площадь n -параллелепипеда является модулем определителя, составленного из n векторов относительно этого базиса. Можно написать, что $\vec{X}_n = \vec{\xi}_n + \vec{Y}_n$, где ортогональная проекция $\vec{\xi}_n$ вектора \vec{X}_n на \vec{F}_{n-1} является линейной комбинацией векторов $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, а длина вектора $\vec{Y}_n = (\vec{X}_n | e_n) e_n$ равна $|(\vec{X}_n | e_n)|$. Вектор \vec{Y}_n является высотой параллелепипеда, относящейся к выбранному основанию. В этом случае определитель векторов $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$ является также определителем векторов $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}, \vec{Y}_n$. Однако его последняя строка состоит из элементов 0, 0, ..., 0, $(\vec{X}_n | e_n)$. Поэтому определитель равен произведению определителя, составленного из векторов $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, модуль которого является $(n-1)$ -мерной площадью основания, на число $(\vec{X}_n | e_n)$, модуль которого дает длину высоты \vec{Y}_n , чем и заканчивается доказательство теоремы.

Следствие 1. *n -мерная площадь n -параллелепипеда, определяемого n векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$, не превосходит произведения длин этих векторов. Она всегда строго меньше этого произведения, кроме того случая, когда все векторы попарно ортогональны.*

Доказательство. Утверждение очевидно для $n=1$. Предполагая, что оно верно для $n-1$ векторов, докажем его справедливость для n векторов.

n -мерная площадь n -параллелепипеда равна произведению $(n-1)$ -мерной площади основания, порожденного векторами $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, на длину $\|\vec{Y}_n\|$ высоты \vec{Y}_n . По предположению индукции, первый сомножитель не превосходит произведения $\|\vec{X}_1\| \|\vec{X}_2\| \dots \|\vec{X}_{n-1}\|$ и строго меньше его, если векторы $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$ не являются попарно ортогональными. Второй сомножитель не превосходит $\|\vec{X}_n\|$ и строго меньше его, если \vec{X}_n не ортогонален векторам $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, чем и заканчивается доказательство следствия.

Следствие 2. *Пусть $(X_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ — квадратная матрица, составленная из вещественных чисел и состоящая из*

n строк и *n* столбцов. Тогда определитель Δ этой матрицы допускает следующую оценку:

$$|\Delta| \leq \sqrt{X_{11}^2 + \dots + X_{1n}^2} \sqrt{X_{21}^2 + \dots + X_{2n}^2} \dots \sqrt{X_{n1}^2 + \dots + X_{nn}^2}. \quad (\text{IV}, 10; 28_2)$$

Для доказательства достаточно аналитически записать предыдущий результат.

Замечание. В случае комплексных чисел X_{ij} это неравенство сохранится, если заменить X_{ij}^2 на $|X_{ij}|^2$, однако его доказательство будет значительно сложнее. Неравенство (IV, 10; 28₂) было доказано Адамаром и нашло значительное приложение в теории интегральных уравнений Фредгольма.

Теорема 105 (об ортогональной проекции площадей гиперплоскостей). *Если F и G — гиперплоскости евклидова аффинного пространства E размерности N , то $(N-1)$ -мерная площадь ортогональной проекции на G измеримого множества A гиперплоскости F равна произведению площади A на косинуса острого угла, образованного гиперплоскостями F и G .*

Доказательство. Можно считать, что F и G не параллельны друг другу, так как в противном случае теорема очевидна. Пусть H — пересечение F и G . Размерность этого пересечения равна $N-2$. Пусть \vec{f} и \vec{g} — единичные векторы пространств \vec{F} и \vec{G} , перпендикулярные \vec{H} и выбранные так, что они составляют острый угол θ , равный по определению углу между гиперплоскостями F и G (этот угол является также углом между нормалями к \vec{F} и \vec{G}). Таким образом, $(\vec{f}|\vec{g}) = \cos \theta \geq 0$. Выберем в H ортонормированную систему координат $0, \vec{h}_1, \vec{h}_2, \dots, \vec{h}_{N-2}$. Тогда $0, \vec{h}_1, \vec{h}_2, \dots, \vec{h}_{N-2}, \vec{f}$ образуют ортонормированную систему координат в F , а $0, \vec{h}_1, \vec{h}_2, \dots, \vec{h}_{N-2}, \vec{g}$ — ортонормированную систему координат в G . Введение указанных систем координат отождествляет эти пространства с \mathbb{R}^{N-1} .

Проекция F на G является аффинным отображением, оставляющим неизменной точку 0 , а, следовательно, определяется некоторой матрицей относительно этих систем координат. Проекции векторов $\vec{h}_1, \vec{h}_2, \dots, \vec{h}_{N-2}$ совпадают с самими векторами, а проекция вектора \vec{f} равна $\vec{g} \cos \theta$. Следовательно, матрица этого отображения является диагональной с элементами $1, 1, \dots, \cos \theta$. Определитель Δ этой матрицы равен $\cos \theta$, и потому результат вытекает из следствия 5 теоремы 102.

***n*-мерная площадь *n*-мерного параметрического многообразия**

Попытаемся теперь определить *n*-мерную площадь *n*-мерного многообразия класса C^1 аффинного евклидова пространства E . Мы не будем здесь строить общую теорию, как это мы делали для длин и для объемов. Мы рассмотрим только случай многообразия класса C^1 .

1°) Пусть $\Psi: \mathcal{O} \rightarrow \Psi(\mathcal{O})$ — отображение класса C^1 некоторого открытого множества \mathcal{O} из \mathbb{R}^n в пространство E .

Производное отображение $\Psi'(u)$ в точке u из \mathbb{R}^n переводит систему *n* векторов базиса \mathbb{R}^n в систему *n* векторов $\frac{\partial \vec{\Psi}}{\partial u_i}(u)$ из E , $i = 1, 2, \dots, n$. Она определяет некоторый параллелепипед, *n*-мерная площадь которого $D(u)$ определяется по формуле

$$D(u) = \sqrt{\det_{i,j} \left(\frac{\partial \vec{\Psi}}{\partial u_i}(u) \mid \frac{\partial \vec{\Psi}}{\partial u_j}(u) \right)}. \quad (\text{IV}, 10; 29)$$

По определению, *n*-мерная площадь *n*-мерного параметрического многообразия класса C^1 , определяемого отображением Ψ , задается интегралом

$$S = \int \dots \int_{\mathcal{O}} D(u) du = \int \dots \int_{\mathcal{O}} D(u_1, u_2, \dots, u_n) du_1 du_2 \dots du_n. \quad (\text{IV}, 10; 30)$$

Причины, по которым мы дали такое определение, ясны. Разобъем множество \mathcal{O} на малые параллелепипеды со сторонами, параллельными осям. Образ при отображении Ψ параллелепипеда с вершиной u и сторонами, параллельными осям, длины du_1, du_2, \dots, du_n является «приближенно» параллелепипедом с вершиной $\Psi(u)$, определяемым векторами $\frac{\partial \vec{\Psi}}{\partial u_1}(u) du_1, \frac{\partial \vec{\Psi}}{\partial u_2}(u) du_2, \dots, \frac{\partial \vec{\Psi}}{\partial u_n}(u) du_n$, площадь которого равна $D(u) \times du_1 du_2 \dots du_n$. Искомая площадь является «суммой этих элементарных площадей», откуда и получается интеграл (IV, 10; 30).

Теорема 106. *n*-мерная площадь параметрического многообразия Ψ , определяемого по формуле (IV, 10; 30), не изменится, если данное многообразие заменить эквивалентным многообразием. Другими словами, если Ψ_1 и Ψ_2 — отображения класса C^1 двух открытых множеств \mathcal{O}_1 и \mathcal{O}_2 из \mathbb{R}^n в E и если существует C^1 -диффеоморфизм H множества \mathcal{O}_1 на \mathcal{O}_2 , такой,

что $\Psi_1 = \Psi_2 \circ H$, то имеет место равенство

$$\int_{\sigma_1} \dots \int D_1(u) du = \int_{\sigma_2} \dots \int D_2(u) du. \quad (\text{IV}, 10; 31)$$

Доказательство. Пусть $\alpha_1 \in \mathcal{O}_1$, $\alpha_2 = H(\alpha_1) \in \mathcal{O}_2$ и $x = \Psi_1(\alpha_1) = \Psi_2(\alpha_2)$. Подпространство, порожденное n векторами $\frac{\partial \vec{\Psi}_1}{\partial u_i}(\alpha_1)$, имеет размерность $\leq n$ и равную n , если эти векторы независимы. Обозначим через F n -мерное подпространство, содержащее рассматриваемое подпространство. Поскольку линейное отображение $\Psi'_1(\alpha_1)$ является композицией $\Psi'_2(\alpha_2) \circ H'(\alpha_1)$, в которой $H'(\alpha_1)$ есть биекция, то это подпространство совпадает с подпространством, порожденным векторами $\frac{\partial \vec{\Psi}_2}{\partial u_i}(\alpha_2)$ (следствие 4 теоремы 11 гл. III). Выбирая в этом векторном пространстве F произвольный ортонормированный базис, мы можем отождествить его с пространством \mathbb{R}^n .

Если теперь мы положим $A_1 = D_1(\alpha_1)$ и $A_2 = D_2(\alpha_2)$, то множитель A_1 , определяющий площадь параллелепипеда, порожденного векторами $\frac{\partial \vec{\Psi}_1}{\partial u_i}(\alpha_1)$, равен модулю определителя линейного отображения $\Psi'_1(\alpha_1)$ пространства \mathbb{R}^n в \mathbb{R}^n , а множитель A_2 равен модулю определителя линейного отображения $\Psi'_2(\alpha_2)$ из \mathbb{R}^n в \mathbb{R}^n .

Соотношение $\Psi'_1(\alpha_1) = \Psi'_2(\alpha_2) \circ H'(\alpha_1)$ означает, что между определителями имеет место равенство

$$D_1(\alpha_1) = |\det H'(\alpha_1)| D_2(\alpha_2) = |\det H'(\alpha_1)| D_2(H(\alpha_1)). \quad (\text{IV}, 10; 32)$$

Так как это равенство справедливо для любого α_1 из \mathcal{O}_1 , то из формулы (IV, 10; 1) замены переменных для кратных интегралов следует, что имеет место равенство

$$\begin{aligned} \int_{\sigma_2} \dots \int D_2(\alpha_2) d\alpha_2 &= \int_{\sigma_1} \dots \int D_2(H(\alpha_1)) |\det H'(\alpha_1)| d\alpha_1 = \\ &= \int_{\sigma_1} \dots \int D_1(\alpha_1) d\alpha_1 \end{aligned} \quad (\text{IV}, 10; 33)$$

и, следовательно, оба интеграла, определяющие n -мерные площади для Ψ_1 и Ψ_2 , равны между собой.

Пример. Поверхности в аффинном евклидовом трехмерном пространстве. В этом случае обычно через u , v принято

обозначать координаты множества параметров $\mathcal{O} \subset \mathbb{R}^2$, через x, y, z — составляющие отображения Ψ по осям ортонормированной системы координат в $E = \mathbb{R}^3$, через A, B, C — якобианы:

$$\begin{aligned} A &= \frac{\partial y}{\partial u} \frac{\partial z}{\partial v} - \frac{\partial z}{\partial u} \frac{\partial y}{\partial v}, \\ B &= \frac{\partial z}{\partial u} \frac{\partial x}{\partial v} - \frac{\partial x}{\partial u} \frac{\partial z}{\partial v}, \\ C &= \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial y}{\partial u} \frac{\partial x}{\partial v}, \end{aligned} \quad (\text{IV, 10; 34})$$

через E, F, G — величины:

$$\begin{aligned} E &= \left\| \frac{\overrightarrow{\partial\Psi}}{\partial u} \right\|^2 = \left(\frac{\partial x}{\partial u} \right)^2 + \left(\frac{\partial y}{\partial u} \right)^2 + \left(\frac{\partial z}{\partial u} \right)^2, \\ F &= \left(\frac{\overrightarrow{\partial\Psi}}{\partial u} \mid \frac{\overrightarrow{\partial\Psi}}{\partial v} \right) = \frac{\partial x}{\partial u} \frac{\partial x}{\partial v} + \frac{\partial y}{\partial u} \frac{\partial y}{\partial v} + \frac{\partial z}{\partial u} \frac{\partial z}{\partial v}, \\ G &= \left\| \frac{\overrightarrow{\partial\Psi}}{\partial v} \right\|^2 = \left(\frac{\partial x}{\partial v} \right)^2 + \left(\frac{\partial y}{\partial v} \right)^2 + \left(\frac{\partial z}{\partial v} \right)^2 \end{aligned} \quad (\text{IV, 10; 35})$$

и через H — то, что мы ранее обозначали через D . Согласно (IV, 10; 27), имеем

$$H = \sqrt{EG - F^2}, \quad (\text{IV, 10; 36})$$

а согласно (IV, 10; 28),

$$H = \sqrt{A^2 + B^2 + C^2}. \quad (\text{IV, 10; 37})$$

2º Вернемся теперь к параметрическому многообразию V , определенному отображением Ψ .

Выражение $D(u)du$ определяет некоторую меру Радона dS на \mathcal{O} . Если A является множеством из \mathcal{O} , то меру этого множества относительно меры Радона можно рассматривать как n -мерную площадь сужения Ψ на это множество. Если Ψ является собственным относительно этой меры dS , то образ при отображении Ψ этой меры Радона есть некоторая мера $\Psi(dS) \geq 0$ в аффинном пространстве E . Обычно через dS мы будем обозначать меру на \mathcal{O} или ее образ. Эта мера служит для определения поверхностных интегралов. Если f является функцией, определенной на E , со значениями в банаховом пространстве F ,

то интегралом $\int \dots \int \vec{f} dS$ ¹⁾, по определению, называют

$$\int \dots \int \vec{f} dS = \int \dots \int (\vec{f} \circ \Psi) dS = \int \dots \int \overrightarrow{f(\Psi(u))} D(u) du.$$
(IV, 10; 38)

Это — интеграл от \vec{f} относительно образа меры $\Psi(dS)$, если последний существует. Впрочем, если мы вернемся к ситуации теоремы 106, то увидим, что если через dS_1 и dS_2 обозначить меры, определенные на \mathcal{O}_1 и \mathcal{O}_2 с помощью предыдущего процесса относительно эквивалентных многообразий Ψ_1 и Ψ_2 , то dS_2 является образом $H(dS_1)$ меры dS_1 при гомеоморфизме H . Об этом, действительно, говорит следствие 1 теоремы 102. Так как

$$\Psi_1(dS_1) = (\Psi_2 \circ H)(dS_1) = \Psi_2(H(dS_1)) = \Psi_2(dS_2),$$

то $\Psi_1(dS_1)$ и $\Psi_2(dS_2)$ представляют собой одну и ту же меру на E (если они существуют).

3°) Теперь мы можем дать корректное определение n -мерной площади *произвольного параметрического многообразия размерности n класса C^1* в аффинном пространстве E размерности N .

Пусть V — многообразие (абстрактное или погруженное в аффинное пространство) размерности n класса C^1 . Пусть H — некоторое отображение V в E класса C^1 , определяющее тем самым некоторое n -мерное многообразие, параметрическое или особое, класса C^1 .

Пусть $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ — некоторая карта многообразия V . Тогда $\Psi = H \circ \Phi$ является отображением \mathcal{O} в E класса C^1 и, значит, можно определить n -мерную площадь dS_Φ меры $\geqslant 0$ в \mathcal{O} . Ее образ при *гомеоморфизме* Φ является мерой $\Phi(dS_\Phi)$ на открытом множестве $\Phi(\mathcal{O})$ многообразия V . Если теперь рассмотреть две карты этого вида, например Φ_1 и Φ_2 , и если открытые множества $\Phi_1(\mathcal{O}_1)$ и $\Phi_2(\mathcal{O}_2)$ имеют непустое пересечение Ω , то можно через Ω_1 и Ω_2 обозначить прообразы Ω при отображениях Φ_1 и Φ_2 . Сужения Φ_1 и Φ_2 на Ω_1 и Ω_2 определяют теперь эквивалентные параметрические многообразия, поскольку

¹⁾ Как и для криволинейного интеграла (см. примечание на стр. 705), надо было написать $\int \dots \int \vec{f} dS$ — интеграл относительно параметрического

многообразия (Ψ). С другой стороны, символ кратного интеграла $\int \dots \int$ совершенно не обоснован, поскольку нет и речи об интегралах относительно тензорного произведения мер. Это, к сожалению, всего лишь некоторый прием, позволяющий указать, что речь идет об интеграле на многообразии размерности > 1 .

$H_{1,2} = \Phi_2^{-1} \circ \Phi_1$ является гомеоморфизмом класса C^1 (следствие 1 теоремы 33 гл. III) и поскольку $\Phi_1 = \Phi_2 \circ H_{1,2}$. Сделанное выше замечание показывает теперь, что мера dS_{Φ_2} на \mathcal{O}_2 является образом меры dS_{Φ_1} на \mathcal{O}_1 при отображении $H_{1,2}$ и что образы этих мер, т. е. меры $\Phi_1(dS_{\Phi_1})$ и $\Phi_2(dS_{\Phi_2})$, в Ω совпадают.

Рассмотрим теперь систему всех открытых множеств $\Phi(\mathcal{O})$, являющихся образами всех карт Φ на V . В каждом из этих открытых множеств мы имеем некоторую вещественную меру $\geq 0 \Phi(dS_{\Phi})$, а в пересечении Ω двух из этих открытых множеств $\Phi_1(\mathcal{O}_1)$ и $\Phi_2(\mathcal{O}_2)$ меры $\Phi_1(dS_{\Phi_1})$ и $\Phi_2(dS_{\Phi_2})$, определенные соответствующими картами Φ_1 и Φ_2 , совпадают. Значит, мы находимся в условиях теоремы о кусочной склейке мер (теорема 17) и можно утверждать, что на V существует некоторая вполне определенная мера Радона, которая на каждом образе карты $\Phi(\mathcal{O})$ совпадает с n -мерной площадью $\Phi(dS_{\Phi})$, определенной этой картой. Именно эту меру Радона на V называют мерой n -мерных площадей на особом многообразии и обозначают через dS . Если A — часть V , то ее называют n -мерной площадью A (или сужения H в A), т. е., по определению, это мера A относительно dS . Если норма dS конечна, то она, т. е. интеграл $\int_V \dots \int dS$, является n -мерной площадью пар-

метрического многообразия. Эта площадь не изменяется, если параметрическое многообразие будет заменено эквивалентным параметрическим многообразием. Если отображение H является dS -собственным, то существует на E образ меры $H(dS)$, часто обозначаемый через dS , носитель которого лежит в замыкании $\overline{H(V)}$ образа $H(V)$. Если \vec{f} является некоторой функцией на E со значениями в банаховом пространстве \vec{F} , то можно говорить об интеграле $\int_V \dots \int \vec{f} dS^1$, равном по определению $\int_V \dots \int (\vec{f} \circ H) dS$. Это — интеграл от функции \vec{f} относительно $H(dS)$; если только образ меры существует.

Предыдущие результаты можно подытожить и сформулировать в виде следующей теоремы:

Теорема 107. Пусть $H: V \rightarrow E$ — отображение класса C^1 многообразия V класса C^1 размерности n в аффинное евклидово

¹⁾ Или лучше интеграл $\int_V \dots \int \vec{f} dS$ относительно параметрического многообразия, определенного отображением H многообразия V в E . Заметим, что мы ввели три рода мер ≥ 0 : меру dS_{Φ} на \mathcal{O} , меру $\Phi(dS_{\Phi})$ на $\Phi(\mathcal{O})$, склейку мер $\Phi(dS_{\Phi})$, дающую меру dS на V , и, наконец, меру $H(dS)$ на E .

пространство E размерности N , определяющее параметрическое многообразие в E . Для каждой карты $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ многообразия V через dS_Φ обозначим меру $D(u)du \geq 0$ на \mathcal{O} , определенную по формуле (IV, 10; 29). Тогда на V существует, и при том единственная, мера $dS \geq 0$, совпадающая с $\Phi(dS_\Phi)$ на каждом множестве $\Phi(\mathcal{O})$. Эту меру называют n -мерной площадью на параметрическом многообразии. Если A является (dS -измеримой) частью многообразия V , то интеграл $\int_A \dots \int dS$ дает площадь A . В частности, $\int_V dS = \|dS\|$ является площадью параметрического многообразия.

Эти результаты, естественно, применимы, если V — многообразие, лежащее в E , а H — тождественное отображение.

Замечания. 1°) Если V является открытым множеством \mathcal{O} из \mathbb{R}^n , то в качестве карты Φ можно взять тождественное отображение, и тогда сразу же получаем $dS = dS_\Phi = D(u)du$. Впрочем, этот случай был уже нами разобран перед теоремой 106 (тогда V и H обозначались через \mathcal{O} и Ψ).

2°) Пусть при $n = 1$ многообразие V является открытым интервалом \mathbb{R}_1 вещественной прямой \mathbb{R} и H представляет собой отображение, обозначаемое через M , интервала \mathbb{R}_1 в E . Тогда мы приходим к случаю, рассмотренному выше в замечании 1°). Здесь $D(t)$ (для переменной из \mathbb{R}_1 мы пишем t вместо u) является длиной $\overrightarrow{M'(t)}$ и $ds = D(t)dt = \|\overrightarrow{M'(t)}\|dt$. Мы снова возвращаемся к длине (формула (IV, 9; 40₂)). Однако длина кривой может быть определена в произвольном нормированном аффинном пространстве и даже в произвольном метрическом пространстве (стр. 701), в то время как n -мерные площади при $n > 1$ могут быть определены только в аффинных евклидовых пространствах.

3°) Если V при $n = N$ является открытым множеством Ω из E , а H — тождественным преобразованием, то мера N -мерных площадей будет просто мерой объемов dx . В самом деле, теперь мы находимся в условиях замечания 1°), и, положив $\Phi = H =$ — тождественное преобразование, мы получим $D(u) = 1$.

4°) Если Ω — некоторое открытое множество из \mathbb{R}^N , а ξ — отображение Ω в E класса C^1 , то оно определяет некоторое параметрическое многообразие в E класса C^1 . Точно так же, как в п. 3°), N -мерную площадь некоторой части A множества Ω можно называть ее объемом. Однако, поскольку отображение ξ не обязательно инъективно, эта величина не обязательно будет объемом образа $\xi(A)$. Если в E выбрать ортонормированную систему координат и отождествить это пространство с \mathbb{R}^N , то,

согласно (IV, 10; 30), этот объем всегда будет равен $\int \int \dots \int_A |\det \xi'(t)| dt$. В следствии 3 теоремы 102 мы видели,

что это выражение равно объему $\xi(A)$, если только ξ является некоторым гомеоморфизмом. В общем случае, если для каждой точки $x \in E$ через $v(x)$ обозначить число тех точек Ω , образ которых при отображении ξ равен x , $0 \leq v(x) \leq +\infty$, то, согласно (IV, 10; 23₂), будет иметь место равенство

$$\int \int \dots \int_A |\det \xi'(t)| dt = \int \int \dots \int_{\xi(A)} v(x) dx. \quad (\text{IV, 10; 38}_2)$$

Мы получили «объем $\xi(A)$, в котором каждая точка учитывается столько раз, сколько она покрывается при рассматриваемом отображении».

5°) Предположим теперь, что $n = 0$. Многообразие нулевой размерности в E является множеством изолированных точек E . В этом случае 0-мерная площадь такого многообразия определяется как число (конечное или бесконечное) этих точек; 0-мерная площадь точки равна 1.

Следствие 1. При преобразовании подобия с отношением k n -мерные площади умножаются на $|k|^n$.

Пусть V — некоторое многообразие класса C^1 размерности n и $H: V \rightarrow E$ — параметрическое многообразие пространства E класса C^1 . Если отображение ξ пространства E является подобием с отношением k , то $\xi \circ H$ будет новым параметрическим многообразием, имеющим V в качестве многообразия параметров. Поскольку ξ изменяет n -мерные площади n -параллелипедов в $|k|^n$ раз, то непосредственно видно, что мера $D(u)du$ относительно $\xi \circ H$ равна мере относительно H , умноженной на $|k|^n$, а, следовательно, мера dS на V относительно $\xi \circ H$ равна мере относительно H , умноженной на $|k|^n$, откуда и следует сформулированное утверждение.

Следствие 2. Пусть V — истинное многообразие размерности n аффинного евклидова пространства E класса C^1 . Если V замкнуто, то n -мерная площадь любой ограниченной dS -измеримой части A многообразия V конечна. Если многообразие V замкнуто и ограничено, то его площадь конечна.

В самом деле, замыкание \bar{A} подмножества A относительно V замкнуто в многообразии V . Так как V замкнуто в E , то \bar{A} замкнуто в E и, поскольку это множество по условию еще и ограничено, то оно компактно (теорема 23 гл. II). Следовательно, мера действительно конечна.

Следствие 3. Пусть $H: V \rightarrow E$ — параметрическое многообразие размерности n аффинного евклидова пространства E класса C^1 . Если V счетно в бесконечности и W является некоторым подмногообразием V размерности $< n$ класса C^1 , то его n -мерная площадь равна нулю.

Это следствие является обобщением следствия 1 теоремы 102₂.

Доказательство. В самом деле, пусть $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ — некоторая карта многообразия V . Поскольку Φ является C^1 -диффеоморфизмом, то прообраз при отображении Φ многообразия-пересечения W и $\Phi(\mathcal{O})$ является многообразием из \mathcal{O} размерности $< n$ класса C^1 (следствие 2 теоремы 32 гл. III). Тогда этот прообраз имеет нулевую меру относительно du (следствие теоремы 102₂), а, значит, и относительно $D(u)du = dS_\Phi$. Поскольку Φ является некоторым гомеоморфизмом, то $W \cap \Phi(\mathcal{O})$ имеет нулевую n -мерную площадь в V . Так как каждый компакт K из V покрывается конечным числом образов таких карт, как $\Phi(\mathcal{O})$, то $W \cap K$ также имеет нулевую площадь. Наконец, поскольку V является счетным множеством компактов, то W является счетным объединением множеств нулевой n -мерной площади, а, значит, также имеет n -мерную площадь, равную нулю.

Пример 1. Площадь сфер радиуса R в аффинном евклидовом трехмерном пространстве. Выбор некоторой ортонормированной системы координат отождествляет E с \mathbb{R}^3 . Если в \mathbb{R}^2 взять открытое множество $0 < \theta < \pi$, $0 < \varphi < 2\pi$ и положить

$$\begin{aligned} x &= R \sin \theta \cos \varphi, \\ y &= R \sin \theta \sin \varphi, \\ z &= R \cos \theta, \end{aligned} \tag{IV, 10; 39}$$

то полярные координаты (IV, 10; 18) определят некоторую карту $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ является открытой частью сферы, дополнительной к полумериану $y = 0$, $x \geq 0$. Поскольку этот полумериан является объединением некоторой полуокружности без концов, являющейся многообразием размерности 1 класса C^1 , и двух полюсов — многообразий размерности 0 класса C^1 , то он имеет нулевую площадь. Следовательно, площадь сферы равна площади $\Phi(\mathcal{O})$. Если A является каким-либо подмножеством сферы, то его площадь равна площади его пересечения с $\Phi(\mathcal{O})$. Следовательно, знание только лишь карты Φ совершенно достаточно для изучения сферических площадей; $\Phi(\mathcal{O})$ dS -почти всюду совпадает с многообразием. Именно так чаще всего бывает на практике. Согласно вычислениям, проведенным на стр. 750, мы знаем векторы $d\Phi/d\theta$ и $d\Phi/d\varphi$. Эти векторы

взаимно ортогональны и имеют длину R и $R \sin \theta$, а, следовательно, площадь образованного ими параллелограмма равна $D(\theta, \varphi) = R^2 \sin \theta$. Значит, площадь части A сферы будет вычисляться по формуле

$$S(A) = \iint_{\Phi^{-1}(A)} R^2 \sin \theta \, d\theta \, d\varphi. \quad (\text{IV}, 10; 40)$$

В частности, площадь сферы равна

$$S = \iint_{\begin{array}{c} 0 < \theta < \pi \\ 0 < \varphi < 2\pi \end{array}} R^2 \sin \theta \, d\theta \, d\varphi = R^2 \left(\int_{[0, \pi]} \sin \theta \, d\theta \right) \left(\int_{[0, 2\pi]} d\varphi \right) = 4\pi R^2. \quad (\text{IV}, 10; 41)$$

Позже мы вычислим площадь сферы в евклидовом пространстве произвольной размерности.

Пример 2. Вычисление телесных углов в аффинном евклидовом пространстве. Пусть $H: V \rightarrow E$ — некоторая параметрическая гиперповерхность, т. е. $(N - 1)$ -мерное параметрическое многообразие класса C^1 в аффинном пространстве E размерности N . Пусть O — некоторая точка пространства E . Поставим задачу определения «телесного угла», под которым заданная гиперповерхность видна из точки O ¹⁾. Мы определим абсолютный телесный угол, а не алгебраический угол, так что он всегда будет неотрицательным числом. Мы должны будем всегда предполагать, что точка O не принадлежит образу многообразия V . (В противном случае через V_0 можно обозначать дополнение в V к прообразу $H^{-1}(\{O\})$. Так как прообраз $H^{-1}(\{O\})$ замкнут, то множество V_0 открыто и, следовательно, является некоторым многообразием, а сужение H на V_0 является некоторым параметрическим многообразием, образ которого не содержит точки O . Телесный угол относительно V , по определению, будет телесным углом относительно V_0 .) Пусть M — произвольная точка множества $E - O$. Полупрямая OM пересекает единичную сферу с центром O в некоторой точке m . Тем самым определяется некоторое отображение $\xi: M \rightarrow m = \xi(M)$ множества $E - O$ в E , принадлежащее, очевидно, классу C^1 , образ которого является единичной сферой с центром O . Отображение $\xi \circ H$ определяет новое параметрическое многообразие класса C^1 , где V есть, как всегда, многообразие параметров. Образ этого параметрического многообразия содержится в единичной сфере с центром O . Тогда, по определе-

¹⁾ Понятие определено весьма интуитивно. При этом результат (IV, 10; 48) просматривается почти сразу же. Мы его установим со всей строгостью в качестве типичного примера. На практике обычно знают, что хотят получить, и к цели идут быстрее!

нию, телесный угол, под которым из точки O видно параметрическое многообразие, определяемое отображением H , является $(N - 1)$ -мерной площадью параметрического многообразия, определяемого композицией $\xi \circ H: V \rightarrow E$.

Мера площадей, связанная с параметрическим многообразием $\xi \circ H$, является некоторой мерой $d\omega \geq 0$ на V , которую называют мерой телесных углов. Можно говорить также о мере относительно ω части A множества V . Полный телесный угол параметрического многообразия равен $\|d\omega\| = \int_V \dots \int d\omega \leqslant +\infty$. Очевидно, что если V является просто

некоторым открытым множеством единичной сферы и H является тождественным отображением, то этот телесный угол будет площадью самого множества V . В частности, если V представляет собой всю единичную сферу или какую-либо сферу с центром O , то в качестве телесного угла получают площадь единичной сферы в E . Телесный угол инвариантен относительно гомотетии с центром в точке O .

Найдем производное отображение отображения ξ в точке $M_0 \in E$. В пространстве E мы можем произвольно выбрать ортонормированную систему координат. Выберем ее так, чтобы $N - 1$ первых координат точки M_0 были равны нулю, а N -я координата равнялась расстоянию r_0 от точки M_0 до точки O . Формулы, определяющие преобразование $\xi: x \rightarrow y$, тогда будут иметь вид:

$$y_i = \frac{x_i}{r} = \frac{x_i}{\sqrt{x_1^2 + x_2^2 + \dots + x_N^2}}, \quad i = 1, 2, \dots, N. \quad (\text{IV}, 10; 42)$$

Отсюда дифференцированием получаем

$$dy_i = \frac{dx_i}{r} - \frac{x_i}{r^2} dr, \quad \text{где} \quad dr = \sum_{j=1}^N \frac{x_j}{r} dx_j. \quad (\text{IV}, 10; 43)$$

Ввиду особого расположения точки M_0 относительно выбранной системы координат это дает $dr = dx_N$, откуда

$$dy_i = \frac{dx_i}{r_0} \quad \text{для } i = 1, 2, \dots, N - 1, \quad dy_N = 0. \quad (\text{IV}, 10; 44)$$

Отсюда вытекает, что искомое производное отображение $\xi'(M_0)$ сводится к ортогональному проектированию каждого вектора \vec{X} на гиперплоскость, перпендикулярную радиус-вектору OM_0 в точке M_0 и последующей гомотетии с отношением $1/r_0$.

Пусть теперь $\Phi: \mathcal{O} \rightarrow \Phi(\mathcal{O})$ — некоторая карта V , $u_0 \in \mathcal{O}$ и $M_0 = H(\Phi(u_0))$. Тогда множитель $D(u_0)$, относящийся

к $(N - 1)$ -мерным площадям на V , является $(N - 1)$ -мерной площадью в E параллелепипеда, построенного на системе векторов $\vec{X}_i = H'(\Phi(u_0)) \frac{\partial \vec{\Phi}}{\partial u_i}(u_0)$. С другой стороны, для телесного угла мы имеем множитель $\Delta(u_0)$, который должен быть $(N - 1)$ -мерной площадью параллелепипеда, построенного на системе векторов \vec{Y}_i , полученных из предыдущих векторов при отображении $\xi'(M_0)$. Согласно теореме 105 о проекциях площадей и следствию 1 теоремы 107, $\Delta(u_0) = D(u_0) \cos \theta_0 / r_0^{N-1}$, где $\theta_0 = \theta(u_0)$ — острый угол, образованный OM_0 с нормалью к многообразию V в точке M_0 .

Окончательно, на открытом множестве \mathcal{O} из \mathbb{R}^n имеет место формула

$$d\omega_{\Phi} = \Delta(u) du = \frac{\cos \theta(u)}{(r(u))^{N-1}} D(u) du. \quad (\text{IV}, 10; 45)$$

Образ этой меры при отображении Φ дает меру углов в открытом множестве $\Phi(\mathcal{O})$ многообразия V . Это можно записать так:

$$\Phi(d\omega_{\Phi}) = \frac{\cos \theta}{r^{N-1}} \Phi(dS_{\Phi}), \quad (\text{IV}, 10; 46)$$

где для точки $v \in V$ величины θ и r относятся к точке $H(v) \in E$. Легко находится та же самая мера для любого открытого множества $\Phi(\mathcal{O})$ из V :

$$d\omega = \frac{\cos \theta}{r^{N-1}} dS = \frac{\cos \theta(v)}{(r(v))^{N-1}} dS(v). \quad (\text{IV}, 10; 47)$$

Телесный угол dS -измеримого подмножества A из V или всего множества V находится, следовательно, по формуле

$$\Omega = \int \dots \int_A d\omega \quad \text{или} \quad \int \dots \int_V \frac{\cos \theta}{r^{N-1}} dS, \quad (\text{IV}, 10; 48)$$

как этого и следовало ожидать из геометрических соображений.

Вычисление объемов с помощью поверхностных интегралов

Пусть Φ есть C^1 -диффеоморфизм открытого множества $\mathcal{O} \subset \mathbb{R}^n$ на открытое множество Ω аффинного евклидова пространства E .

Если \vec{f} — функция, определенная на Ω , со значениями в банаховом пространстве \vec{F} , то интеграл $\int \int \dots \int_{\Omega} \vec{f}(x) dx$, где

dx — мера, связанная с аффинным евклидовым пространством, может быть, согласно следствию 2 теоремы 102 о замене переменных в кратных интегралах, вычислен по формуле (IV, 9; 76):

$$\int \int \dots \int_{\Omega} \vec{f}(x) dx = \int \int \dots \int_{\sigma} \vec{f}(\Phi(u)) |\det \Phi'(u)| du, \quad (\text{IV, 10; 49})$$

в которой выписанный определитель является якобианом преобразования относительно произвольно выбранной в E ортонормированной системы координат. Этот определитель дает объем параллелепипеда, определяемого векторами $\vec{X}_i = \frac{\partial \vec{\Phi}}{\partial u_i}(u)$, $i = 1, 2, \dots, N$.

Положим $u = (v, w)$, где $v = (u_1, \dots, u_{N-1})$ и $w = u_N$. Зададим w . Тогда сужение отображения Φ на множество значений u , соответствующее этому значению w , является гиперповерхностью Σ_w в Ω класса C^1 . Рассмотрим векторы \vec{X}_i . Первые $N - 1$ из них будут векторами, касательными к гиперповерхности в рассматриваемой точке, а N -й вектор \vec{X}_N будет расположен вне касательной гиперплоскости.

Обозначим через v единичный вектор нормали к гиперповерхности Σ_w в рассматриваемой точке $\Phi(u)$. Объем параллелепипеда, определяемого векторами \vec{X}_i , является произведением $(N - 1)$ -мерной площади параллелепипеда, порожденного первыми $N - 1$ векторами, на длину ортогональной проекции последнего вектора на нормаль: $|(\vec{X}_N | v)|$. Однако указанная $(N - 1)$ -мерная площадь определяется множителем $D(v, w)^1$, позволяющим вычислить площади на поверхности Σ_w . Поэтому мы имеем право записать:

$$\int \int \dots \int_{\Omega} \vec{f}(x) dx = \int \int \dots \int_{\sigma} \vec{f}(\Phi(u)) \left| D(v, w) \left(\frac{\partial \vec{\Phi}}{\partial w}(u) | v \right) \right| du. \quad (\text{IV, 10; 50})$$

Если известно, что функция \vec{f} интегрируема, или если известно, что она вещественна, $\geqslant 0$ и измерима, то можно

¹) Эта функция обязательно зависит от v и w . Для фиксированного w она зависит лишь от v . Следовательно, $D(v, w)dv = dS_w$ является мерой площадей относительно параметрического многообразия Σ_w : $v \rightarrow \Phi(v, w)$. Ее образ при гомеоморфизме Φ является мерой площадей на самом многообразии Σ_w . Хотя эта мера на Σ_w изменяется вместе с w , она всегда допускает представление через dS . (Обыходная математика была бы слишком сложной, если бы мы не пользовались упрощенным языком!)

применить теорему Фубини и результат записать в виде

$$\int d\omega \int \dots \int \vec{f}(\Phi(u)) \left| D(v, \omega) \left(\frac{\partial \vec{\Phi}}{\partial \omega}(u) | \vec{v} \right) \right| dv, \quad (\text{IV}, 10; 51)$$

или

$$\int d\omega \int \dots \int \vec{f}(x) \left| \left(\frac{\partial \vec{\Phi}}{\partial \omega} | \vec{v} \right) \right| dS. \quad (\text{IV}, 10; 52)$$

Таким образом этот N -кратный интеграл может быть вычислен последовательно как некоторый поверхностный интеграл по гиперповерхности Σ_ω , зависящей от параметра ω , а затем как простой интеграл относительно ω .

Рассмотрим теперь частный случай, представляющий самостоятельный интерес. Пусть h — некоторая вещественная функция класса C^1 , определенная на открытом множестве Ω аффинного пространства E , производное отображение которой $\overleftarrow{h'(a)}$ отлично от нуля в каждой точке $a \in \Omega$. Зафиксируем некоторую точку a . Поскольку производное отображение $\overleftarrow{h'(a)}$ отлично от нуля, существует по крайней мере одна частная производная h относительно некоторой ортонормированной системы координат в E , не равная нулю. Пусть, для определенности, это $\frac{\partial h}{\partial x_N}(a)$. Тогда, согласно теореме о неявных функциях (теорема 28 гл. III), существует такая окрестность Ω_a точки a в Ω , в которой уравнение $w = h(x_1, x_2, \dots, x_N)$ может быть разрешено относительно x_N в виде $x_N = g(x_1, \dots, x_{N-1}, w)$, где g — функция класса C^1 . Кроме того, согласно правилу (III, 8; 23), та же самая теорема дает $\frac{\partial g}{\partial w} = \frac{1}{\partial h / \partial x_N}$. Если теперь перейти к сужению на открытое множество Ω_a и воспользоваться формулами

$$x_i = u_i = v_i, \quad i = 1, 2, \dots, N-1, \quad (\text{IV}, 10; 53)$$

$$x_N = g(u_1, \dots, u_{N-1}, u_N) = g(v_1, \dots, v_{N-1}, w),$$

то мы окажемся в описанных выше условиях. Мы определили C^1 -диффеоморфизм Φ открытого множества $\mathcal{O}_a \subset \mathbb{R}^N$ на Ω_a так, что подпространства $w = \text{const}$ стали гиперповерхностями Σ_w с уравнением $h(x) = w$. Поскольку $x_i, i \leq N-1$, не зависят от w , то вектор $\frac{\partial \vec{\Phi}}{\partial \omega}(u)$ является вектором, параллельным оси x_N , с составляющей $\frac{\partial g}{\partial w}(u) \left(= \frac{1}{\frac{\partial h}{\partial x_N}(x)} \right)$ по этой оси. Если учесть, что, согласно следствию 5 теоремы 33₄ гл. III,

единичный вектор нормали $\vec{v} = \pm \overrightarrow{\text{grad}} h / \|\overrightarrow{\text{grad}} h\|$, то проекция вектора \vec{v} на ось x_N равна $\pm \frac{\partial h}{\partial x_N}(x) / \|\overrightarrow{\text{grad}} h(x)\|$.

Отсюда следует формула

$$\left| \left(\frac{\partial \Phi}{\partial \omega} \Big| \vec{v} \right) \right| = \frac{1}{\left| \frac{\partial h}{\partial x_N} \right|} \cdot \frac{\left| \frac{\partial h}{\partial x_N} \right|}{\|\overrightarrow{\text{grad}} h\|} = \frac{1}{\|\overrightarrow{\text{grad}} h\|}. \quad (\text{IV}, 10; 54)$$

Если \vec{f} является некоторой интегрируемой на Ω_a функцией со значениями в банаховом пространстве \vec{F} , то можно вычислить ее объемный интеграл с помощью поверхностного по следующей формуле:

$$\iint_{\Omega_a} \dots \int \vec{f}(x) dx = \int d\omega \int_{\Sigma_w} \dots \int \frac{\vec{f}(x)}{\|\overrightarrow{\text{grad}} h(x)\|} dS. \quad (\text{IV}, 10; 55)$$

Эти вычисления, естественно, могут быть выполнены только для открытого множества Ω_a , являющегося окрестностью точки a . Однако окончательный результат не зависит от специального выбора координат относительно точки a . Формула (IV, 10; 55) представляет самостоятельный интерес и не связана с выбором системы координат. Если теперь покрыть компакт K из Ω конечным числом таких открытых множеств Ω_a , то мы получим, что этот результат верен для каждой функции с компактным носителем в Ω и со значениями в \vec{F} . Пользуясь переходом к пределу, для произвольной интегрируемой на Ω функции со значениями в \vec{F} можно получить формулу

$$\iint_{\Omega} \dots \int \vec{f}(x) dx = \int d\omega \int_{\Sigma_w} \dots \int \frac{\vec{f}(x)}{\|\overrightarrow{\text{grad}} h(x)\|} dS. \quad (\text{IV}, 10; 56)$$

Таким образом, справедлива

Теорема 108. Пусть Ω — открытое множество аффинного евклидова пространства E , и пусть h — вещественная функция класса C^1 , определенная на Ω , производная которой в каждой точке $\neq 0$. Если через Σ_w обозначить гиперповерхность, определенную уравнением $h(x) = w$, и если \vec{f} является функцией, определенной на Ω , со значениями в банаховом пространстве \vec{F} , интегрируемой по мере объемов dx в E , то для почти всех значений w функция $\vec{f} / \|\overrightarrow{\text{grad}} h\|$ интегрируема на Σ_w относительно поверхности меры dS , интеграл $\int_{\Sigma_w} \dots \int \frac{\vec{f}}{\|\overrightarrow{\text{grad}} h\|} dS$,

определенный для почти всех значений w , является функцией w со значениями в \vec{F} , интегрируемой по мере dw , и имеет место формула (IV, 10; 56). Если \vec{f} есть dx -измеримая, вещественная, не обязательно интегрируемая функция ≥ 0 , то обе части равенства (IV, 10; 56) равны и одновременно являются конечными или бесконечными.

Замечание. Геометрически предыдущий результат довольно ясен. Элемент объема dx может быть представлен как произведение поверхности элемента dS поверхности Σ_w на расстояние вдоль нормали dv , разделяющее поверхность Σ_w от поверхности Σ_{w+dw} . Вектор $\overrightarrow{\text{grad}} h$ является вектором нормали к поверхности Σ_w длины $\left| \frac{dw}{dv} \right|$. Отсюда следует, что $dv = \frac{dw}{\|\overrightarrow{\text{grad}} h\|}$ и, следовательно, $dx = \frac{dS dw}{\|\overrightarrow{\text{grad}} h\|}$. Это рассуждение лишено строгости; четкие рассуждения были проведены ранее.

Следствие 1. Пусть E — евклидово аффинное пространство, и пусть r — функция «расстояние от точки O в E », т. е. $r(x) = \|x - O\|$. Пусть \vec{f} — функция, определенная на E , со значениями в банаховом пространстве \vec{F} , интегрируемая по объемной мере dx . Тогда почти для всех значений r функция \vec{f} интегрируема относительно поверхностной меры на сфере Σ_r с центром O радиуса r .

Интеграл $\int_{\Sigma_r} \dots \int \vec{f}(x) dS$ от этой функции, определенный для почти всех значений r , интегрируем по мере dr , и имеет место равенство

$$\int \int \dots \int \vec{f}(x) dx = \int_{[0, +\infty[} dr \int_{\Sigma_r} \dots \int \vec{f}(x) dS. \quad (\text{IV}, 10; 57)$$

Если f есть dx -измеримая, вещественная, не обязательно интегрируемая функция ≥ 0 , то обе части равенства (IV, 10; 57) равны и одновременно принимают конечное или бесконечное значение.

Для доказательства достаточно применить теорему к дополнению Ω к точке O в E относительно функции $h = r$. Градиент этой функции в каждой точке M является единичным вектором, лежащим на OM , и, следовательно, длина градиента тождественно равна 1, откуда следует требуемый результат.

Следствие 2. Пусть \vec{f} — некоторая функция, определенная на полуправой \mathbb{R}_+ , т. е. на множестве чисел ≥ 0 прямой \mathbb{R} . Пусть r — скалярная функция, определенная на аффинном евклидовом пространстве E размерности N соотношением $x \rightarrow r(x) = \|\overrightarrow{x-O}\|$, где $\|\overrightarrow{x-O}\|$ — расстояние от x до точки $O \in E$. Тогда функция $\vec{f} \circ r: x \rightarrow \vec{f}(r(x))$, определенная на E со значениями в \vec{F} , зависит только от расстояния переменной x до точки O (что коротко записывается в виде $\vec{f}(r)$). Функция $\vec{f} \circ r$ измерима на E по мере объемов тогда и только тогда, когда функция \vec{f} измерима на \mathbb{R}_+ по мере Лебега. Функция $\vec{f} \circ r$ интегрируема на E по мере объемов тогда и только тогда, когда функция $t \rightarrow S_N t^{N-1} \vec{f}(t)$ интегрируема на \mathbb{R}_+ по мере Лебега, и имеет место равенство

$$\iint_E \dots \int \vec{f}(x) dx = S_N \int_{\mathbb{R}_+} \vec{f}(t) t^{N-1} dt, \quad (\text{IV}, 10; 58)$$

где S_N — площадь сферы радиуса 1 в E . Наконец, образ меры объемов dx при отображении $r: x \rightarrow r(x)$ из E в \mathbb{R}_+ является мерой $S_N t^{N-1} dt$ на \mathbb{R}_+ .

Вообще говоря, r используется как переменная интегрирования в правой части (IV, 10; 58): $S_N \int_{[0, +\infty)} \vec{f}(r) r^{N-1} dr$.

Доказательство. Если \vec{f} — непрерывная скалярная функция с компактным носителем на \mathbb{R}_+ , то к ней можно применить следствие 1. Поскольку на сфере Σ_r радиуса r функция f постоянна и равна $f(r)$, то величина $\int_{\Sigma_r} \dots \int f(r) dS$ равна

произведению $f(r)$ на площадь сферы радиуса r . С помощью гомотетии (следствие 1 теоремы 107) получаем, что эта площадь равна произведению r^{N-1} на площадь S_N сферы радиуса 1, откуда и следует равенство (IV, 10; 58).

Эта формула означает, что образ меры dx при отображении $x \rightarrow r(x)$ является мерой $S_N t^{N-1} dt$ на \mathbb{R}_+ . Но тогда из теоремы 60 следует, что если \vec{f} является некоторой функцией на \mathbb{R}_+ , то функция $\vec{f}(r) dx$ -интегрируема тогда и только тогда, когда $\vec{f}(S_N t^{N-1} dt)$ -интегрируема на \mathbb{R}_+ , а из теоремы 51 следует, что \vec{f} интегрируема относительно $(S_N t^{N-1} dt)$ тогда и только тогда, когда dt -интегрируема функция $S_N \vec{f}(t) t^{N-1}$.

Это следствие означает также, что вычисление объемного интеграла от функции, зависящей лишь от одной переменной r , можно свести к вычислению обычного интеграла. Достаточно лишь знать площадь S_N единичной сферы в \mathbb{R}^N , а она хорошо известна для $N = 1, 2, 3$:

$$S_1 = 2^1), \quad S_2 = 2\pi, \quad S_3 = 4\pi.$$

Обратно, если на \mathbb{R}_+ существует функция \tilde{f} , для которой известны значения $\int \int \dots \int_E^{\mathbb{R}_+} \tilde{f}(r) dr$ и $\int_{\mathbb{R}_+} \tilde{f}(t) t^{N-1} dt$, то они дают значение S_N (если интегралы не равны нулю). Это как раз то, что мы сделаем позже, положив $\tilde{f}(t) = e^{-t^2}$. Так можно найти (сделайте это в качестве упражнения), что

$$S_4 = 2\pi^2, \quad S_5 = \frac{8}{3}\pi^2, \quad S_6 = \pi^3, \dots .$$

Площадь сферы радиуса R равна $S_N R^{N-1}$, что дает

$$2, \quad 2\pi R, \quad 4\pi R^2, \quad 2\pi^2 R^3, \quad \frac{8}{3}\pi^2 R^4, \quad \pi^3 R^5, \dots .$$

Следствие 3. В аффинном евклидовом пространстве размерности N объем шара радиуса R равен $S_N \frac{R^N}{N}$ (первообразная площади сферы $S_N R^{N-1}$, обращающаяся в нуль при $R = 0$).

В самом деле, этот объем вычисляется по формуле (IV, 10; 58), ибо он является интегралом от характеристической функции шара — функции, зависящей только от r . Здесь $f(r) = 0$ для $r > R$ и $= 1$ для $r \leq R$. Мы приходим к значению $S_N \int_{[0, R]} r^{N-1} dr = S_N \frac{R^N}{N}$.

Таким путем для $N = 1, 2, 3, 4, 5, 6$ получаем:

$$2R, \quad \pi R^2, \quad \frac{4}{3}\pi R^3, \quad \frac{1}{2}\pi^2 R^4, \quad \frac{8}{15}\pi^2 R^5, \quad \frac{1}{6}\pi^3 R^6.$$

¹⁾ S_1 является 0-мерной площадью. Сфера в \mathbb{R} с центром в 0 радиуса 1 сводится к множеству $\{-1, +1\}$, состоящему из двух элементов, площадь которого равна 2 (замечание 5°) после теоремы 107). Впрочем, если некоторая функция на прямой \mathbb{R} зависит только от расстояния до начала координат, то это означает, что она четна, и в этом случае формула (IV, 10; 58)

примет вид $\int_{]-\infty, +\infty[} \tilde{f}(x) dx = S_1 \int_{[0, +\infty[} \tilde{f}(x) dx$, откуда $S_1 = 2$.

Следствие 4. Пусть r — расстояние от точки O в N -мерном аффинном евклидовом пространстве. Интеграл $\int \int \dots \int_{r \leq 1} \frac{dx}{r^\alpha}$ (соответственно $\int \int \dots \int_{r \geq 1} \frac{dx}{r^\alpha}$) конечен тогда и только тогда, когда $\alpha < N$ (соответственно $\alpha > N$).

Это следствие является обобщением следствия 1 теоремы 97, соответствующим $N = 1$.

Доказательство. По следствию 2, предыдущие интегралы конечны тогда и только тогда, когда конечны соответствующие простые интегралы $\int \frac{r^{N-1}}{r^\alpha} dr = \int \frac{1}{r^{\alpha-N+1}} dr$. Это так в том случае, когда $\alpha - N + 1 < 1$ (соответственно > 1) или $\alpha < N$ (соответственно $> N$).

§ 11. ФУНКЦИИ, ПРЕДСТАВИМЫЕ РЯДАМИ ИЛИ ИНТЕГРАЛАМИ

Функции, представимые рядами

Пусть $\vec{f}_0, \vec{f}_1, \vec{f}_2, \dots$ — последовательность отображений множества E в множество F . Предположим, что эта последовательность просто сходится при n , стремящемся к бесконечности, к некоторому отображению f из E в F . Функция f называется *пределом последовательности функций*.

Поставим задачу изучения различных свойств функции f в зависимости от выбора E и F и свойств функций f_n : непрерывность f для топологических пространств E и F ; интегрируемость f , если $E = X$ представляет собой локально компактное пространство с мерой Радона $\mu \geq 0$, а \vec{F} — некоторое банахово пространство; дифференцируемость f в случае аффинных нормированных пространств E и F . Вообще говоря, вместо того чтобы иметь дело с последовательностями, в анализе чаще всего рассматривают ряды. В этом случае E является некоторым множеством, \vec{F} — пространством Банаха, а функцию \vec{S} определяют по формуле

$$\vec{S}(x) = \sum_{n=0}^{\infty} \vec{u}_n(x), \quad x \in E. \quad (\text{IV}, 11; 1)$$

Предположим, что с помощью некоторых критериев можно доказать простую сходимость этого ряда; рассмотрим задачу изучения непрерывности, интегрируемости и дифференцируемости суммы \vec{S} , исходя из аналогичных свойств функций \vec{u}_n .

Совершенно очевидно, что в этом случае подход, использующий ряды, не отличается от подхода, использующего последовательности, ибо \vec{S} является пределом последовательности функций $\vec{S}_n = \sum_{l=0}^n \vec{u}_l$, и, обратно, если функция \vec{f} есть предел некоторой последовательности функций \vec{f}_n , то она будет суммой ряда $\vec{f}_0 + (\vec{f}_1 - \vec{f}_0) + (\vec{f}_2 - \vec{f}_1) + \dots$. С теоретической точки зрения, в том случае, когда F — банахово пространство, оба подхода равносильны. Представление функций в виде пределов последовательностей является более общим подходом, поскольку оно возможно в произвольных топологических пространствах. На практике для представления функций чаще всего используют ряды.

Если F — поле комплексных чисел, то может случиться, что нам придется иметь дело с функциями, представленными в виде бесконечного произведения, а именно $\Pi(x) = \prod_{n=0}^{\infty} u_n(x)$.

В этом случае предполагается, что можно доказать простую сходимость этого произведения, и ставится задача изучения непрерывности, интегрируемости или дифференцируемости функции Π , исходя из аналогичных свойств членов произведения.

Большая часть критериев уже рассматривалась в предыдущих главах. В таких случаях мы будем довольствоваться лишь их напоминанием, а здесь приведем только новые критерии.

Непрерывность суммы ряда

В нашем распоряжении имеется теорема 65 гл. II, которую мы сейчас лишь напомним:

Пусть f_n — последовательность отображений топологического пространства E в метрическое пространство F , сходящаяся к некоторому пределу f локально равномерно на E . Если при этом функции f_n непрерывны, то непрерывной будет и функция f .

Естественно, эта теорема, сформулированная для предела некоторой последовательности, также верна для суммы ряда или для предела бесконечного произведения.

Интегрируемость суммы ряда относительно некоторой меры ≥ 0

Здесь мы располагаем теоремами 34—36, их следствиями и теоремой 37 предыдущей главы. Приведем только следующее их уточнение:

Теорема 110. Пусть \vec{f}_n — последовательность μ -интегрируемых на X функций со значениями в некотором банаевом пространстве \vec{F} . Предположим, что \vec{f}_n просто сходятся μ -почти всюду к некоторому пределу \vec{f} и по норме мажорируются одной и той же фиксированной μ -интегрируемой функцией $g \geq 0$. Тогда, какова бы ни была μ -измеримая часть Y множества X , интегралы $\int_Y \vec{f}_n$ сходятся к $\int_Y \vec{f}$ равномерно относительно части Y , пробегающей множество μ -измеримых частей множества X .

Доказательство. В теореме 35 Лебега мы видели, что $\int \|\vec{f} - \vec{f}_n\| d\mu$ сходятся к 0, т. е. при заданном $\varepsilon > 0$ существует такое целое число p , что для всех $n \geq p$ имеет место неравенство $\int \|\vec{f} - \vec{f}_n\| d\mu \leq \varepsilon$.

Для каждой измеримой части Y множества X мы имеем одно и то же неравенство $\int_Y \|\vec{f} - \vec{f}_n\| d\mu \leq \varepsilon$, откуда заведомо следует неравенство $\left\| \int_Y \vec{f}_n d\mu - \int_Y \vec{f} d\mu \right\| \leq \varepsilon$, полностью доказывающее теорему.

Следствие. Пусть \vec{f}_n — последовательность функций, определенных на замкнутом интервале $[a, b]$ прямой \mathbb{R} , со значениями в банаевом пространстве \vec{F} , dx -интегрируемых и равномерно сходящихся при n , стремящимся к бесконечности, к некоторой предельной функции \vec{f} . Тогда $\int_a^b \vec{f}_n(x) dx$ сходятся к $\int_a^b \vec{f}(x) dx$ равномерно относительно α и β из $[a, b]$.

Дифференцируемость суммы ряда

Что касается дифференцируемости, то у нас до сих пор не было никаких теорем.

Пусть f_n — последовательность отображений аффинного нормированного пространства E в аффинное нормированное пространство F .

Предположим, что все эти отображения дифференцируемы. Выясним, можно ли из достаточно сильной сходимости f_n к f , например равномерной, получить дифференцируемость пре-

дельной функции и сходимость f'_n к f' ? Легко видеть, что такое утверждение не верно, поскольку из оценок, налагаемых на функции, получить оценки на производные невозможно. Рассмотрим, например, последовательность комплексных функций одной вещественной переменной, определенных по формулам

$$f_n(x) = \frac{e^{inx}}{n}, \quad n \geq 1. \quad (\text{IV, 11; 2})$$

Эта последовательность равномерно сходится к 0 на всей вещественной прямой \mathbb{R} при n , стремящемся к бесконечности. Однако, ее производные определяются по формуле

$$f'_n(x) = ine^{inx}, \quad (\text{IV, 11; 3})$$

и модуль производной f'_n стремится к бесконечности при n , стремящемся к бесконечности, для всех значений x .

Точно так же можно рассмотреть ряд Вейерштрасса:

$$S(x) = \sum_{n=0}^{\infty} \frac{\sin 3^n x}{2^n}. \quad (\text{IV, 11; 4})$$

Этот ряд, составленный из вещественных функций вещественной переменной, равномерно сходится на прямой \mathbb{R} , поскольку его общий член мажорируется по модулю числом $1/2^n$. Значит, этот ряд представляет собой некоторую непрерывную функцию.

Ряд же, составленный из производных

$$\sum_{n=0}^{\infty} \left(\frac{3}{2}\right)^n \cos 3^n x, \quad (\text{IV, 11; 5})$$

не сходится. Более того, Вейерштрасс показал, что функция S не имеет производной ни при каком значении x — это был первый пример функций такого рода.

Таким образом, налагаемые условия должны быть связаны не только со сходимостью функций, но также и со сходимостью их производных. Мы сейчас убедимся, что если сделать предположение о равномерной сходимости производных и сходимости функций всего лишь в одной точке, то из него будет следовать равномерная сходимость самих функций и будет обоснован переход к пределу для производной.

Теорема 111. Пусть f_n — дифференцируемые (соответственно принадлежащие классу C^1) отображения открытого множества Ω аффинного нормированного пространства E в аф-

финное нормированное пространство F . Пусть производные f'_n (функции на Ω со значениями в $\mathcal{L}(\vec{E}; \vec{F})$) сходятся локально равномерно к некоторому пределу g , а $f_n(a)$ сходятся к пределу $f(a)$ хотя бы в одной точке a множества Ω . Если множество Ω связно, а пространство F полно, то f_n сходятся локально равномерно в Ω к некоторому пределу f . Функция f дифференцируема (соответственно принадлежит классу C^1), и ее производная равна g .

Если же множество Ω не связно, а пространство F не полно, но f'_n сходятся локально равномерно к некоторому пределу g и f_n сходятся просто к некоторому пределу f , то сходимость f_n к f будет локально равномерной, а f будет дифференцируемой (соответственно принадлежащей классу C^1), при этом $f' = g$.

Доказательство. Поскольку сходимость f'_n локально равномерна, то существует такая окрестность точки a , в которой f'_n равномерно сходятся к функции g . Обозначим через $B = B(a; \rho)$ шар с центром a радиуса $\rho > 0$, содержащийся в этой окрестности. Если x принадлежит этому шару, то мы можем применить формулу конечных приращений (теорема 13 гл. III) к функции $\vec{f}_{mn} = \overrightarrow{f_m} - \overrightarrow{f_n}$ и написать:

$$\|\overrightarrow{f_{mn}}(x) - \overrightarrow{f_{mn}}(a)\| \leq \rho \sup_{\|\xi - a\| \leq \rho} \|f'_{mn}(\xi)\|. \quad (\text{IV, 11; 6})$$

Функции $f'_{mn} = f'_m - f'_n$ сходятся равномерно к 0 в шаре B при m и n , стремящихся к бесконечности. Следовательно, левая часть сходится равномерно к 0 в этом шаре. Поскольку, по предположению, $\overrightarrow{f_{mn}}(a) = \overrightarrow{f_m}(a) - \overrightarrow{f_n}(a)$ сходятся к 0, то отсюда следует, что $\overrightarrow{f_{mn}} = \overrightarrow{f_m} - \overrightarrow{f_n}$ сходятся равномерно к 0 в этом шаре при m и n , стремящихся к бесконечности.

Функции f_n не обязательно являются ограниченными в B . Но если мы выберем целое число n_0 так, чтобы при $n \geq n_0$ имело место неравенство $\sup_{\|\xi - a\| \leq \rho} \|\overrightarrow{f_{n_0}}(\xi) - \overrightarrow{f_n}(\xi)\| \leq 1$, то все

$\overrightarrow{f_n} - \overrightarrow{f_{n_0}}$ будут ограниченными по норме числом 1 в шаре B и функции $(\overrightarrow{f_m} - \overrightarrow{f_{n_0}}) - (\overrightarrow{f_n} - \overrightarrow{f_{n_0}})$, $m \geq n_0$, $n \geq n_0$, будут равномерно сходиться к 0 в шаре B при m и n , стремящихся к бесконечности. Это говорит о том, что функции $\overrightarrow{f_n} - \overrightarrow{f_{n_0}}$, $n \geq n_0$, образуют в метрическом пространстве $(\vec{F}^B)_{cb}$ последовательность Коши. Если \vec{F} полно, то это метрическое пространство также

полно (следствие 2 теоремы 65 гл. II), и тогда разность $\overrightarrow{f_n} - \overrightarrow{f}_{n_0}$, а вместе с ней и \overrightarrow{f}_n сходятся равномерно к некоторому пределу в B при n , стремящемся к бесконечности. Если F не полно, но если заранее известно, что функции f_n сходятся просто к некоторому пределу f , то отсюда следует, что сходимость равномерна в B . В самом деле, каково бы ни было $\varepsilon > 0$, существует такое целое число p , что из неравенств $m \geq p, n \geq p$ для всех x из B следует: $\|f_m(x) - f_n(x)\| \leq \varepsilon$. Переходя к пределу, при m , стремящемся к бесконечности (при фиксированном x), получаем отсюда, что при $n \geq p$ для каждого $x \in B$ имеет место неравенство $\|f(x) - f_n(x)\| \leq \varepsilon$, доказывающее наше утверждение.

Если заранее известно, что f_n сходятся просто к некоторому пределу f , то только что проведенное рассуждение сохранится для произвольной точки Ω , а это означает, что сходимость f_n к f локально равномерна. Если же известно лишь, что $f_n(a)$ сходятся к $f(a)$, F полно, а Ω связно, то необходимо провести новые рассуждения.

Обозначим через \mathcal{E} множество таких точек x из Ω , для которых последовательность $f_n(x)$ сходится к некоторому пределу при n , стремящемся к бесконечности. Это множество открыто. В самом деле, если точка b принадлежит \mathcal{E} , то, проведя для точки b то же самое рассуждение, что и для точки a , мы увидим, что существует содержащийся в Ω шар с центром в точке b , в каждой точке которого функции f_n сходятся к некоторому пределу. Следовательно, множество \mathcal{E} открыто. Покажем теперь, что оно замкнуто в Ω . Для этого предположим, что $b_0, b_1, \dots, b_j, \dots$ является последовательностью точек из Ω , сходящейся при j , стремящемся к бесконечности, к некоторой предельной точке $b \in \Omega$, и докажем, что если все b_j принадлежат \mathcal{E} , то и b принадлежит \mathcal{E} .

По предположению, существует некоторая окрестность точки b , в которой последовательность f'_n сходится равномерно к g . Следовательно, существует некоторый шар $B(b; r')$ с центром в точке b , полностью лежащий в этой окрестности. Поскольку при j , стремящемся к бесконечности, b_j сходятся к b , существует такое целое j_0 , что при $j \geq j_0$ выполняется неравенство $\|b_j - b\| \leq r'/3$. Но тогда шар с центром в b_j радиуса $\frac{2}{3}r'$ содержитя в $B(b; r')$, и, следовательно, в этом шаре f'_n равномерно сходятся к g , а $f_n(b_j)$ сходятся к $f(b_j)$. К точке b , можно теперь применить то же рассуждение, которое было только что проведено относительно точки a . Мы увидим, что

в каждой точке x шара $B(b_j; 2\rho/3)$ функции $f_n(x)$ сходятся к некоторому пределу. Точка b составляет часть этого шара, а, следовательно, она также принадлежит множеству \mathcal{E} , которое тем самым оказывается замкнутым. Так как множество Ω предполагалось связным, то непустое множество \mathcal{E} (содержащее точку a) может быть одновременно и открытым, и замкнутым только в том случае, когда оно совпадает с самим множеством Ω . Мы доказали, что f_n сходятся к некоторому пределу f всюду в Ω , а в этом случае, как мы видели выше, сходимость заведомо локально равномерна.

Нам надо теперь доказать, что f дифференцируема и ее производная равна g . Поскольку a уже не играет особой роли, то достаточно показать, что производная функции f в точке a равна $g(a)$. Обозначим через \vec{k}_n функции, определенные на $B(a; \rho)$, со значениями в \vec{F} по формуле:

$$\vec{k}_n(x) = \begin{cases} \frac{\overrightarrow{f_n(x) - f_n(a)} - \vec{f}'_n(a)(\overrightarrow{x - a})}{\|x - a\|} & \text{для } x \neq a, \\ \vec{0} & \text{для } x = a. \end{cases} \quad (\text{IV, 11; 7})$$

Эти функции просто сходятся в B к функции \vec{k} , определенной формулой:

$$\vec{k}(x) = \begin{cases} \frac{\overrightarrow{f(x) - f(a)} - g(a)(\overrightarrow{x - a})}{\|x - a\|} & \text{для } x \neq a, \\ \vec{0} & \text{для } x = a. \end{cases} \quad (\text{IV, 11; 8})$$

Из формулы конечных приращений, взятой в виде, указанном в следствии 1 теоремы 13, следует неравенство:

$$\|\vec{k}_m(x) - \vec{k}_n(x)\| \leq \sup_{\|\xi - a\| \leq \rho} \|(\vec{f}'_m(\xi) - \vec{f}'_n(\xi)) - (\vec{f}'_m(a) - \vec{f}'_n(a))\|. \quad (\text{IV, 11; 9})$$

Отсюда вытекает, что разность $\vec{k}_m - \vec{k}_n$ сходится равномерно к $\vec{0}$ в B при m и n , стремящихся к бесконечности. Но тогда проведенные ранее рассуждения относительно \vec{f}_n показывают, что \vec{k}_n сходится равномерно к \vec{k} в B при n , стремящемся к бесконечности. Согласно определению производной $\vec{f}'_n(a)$, каждая функция \vec{k}_n непрерывна в точке $x = a$. По теореме 65 гл. II равномерный предел \vec{k} также непрерывен в точке $x = a$. Это означает, что функция f дифференцируема в точке a и $f'(a) = g(a)$.

Наконец, если функции f_n принадлежат классу C^1 , то \vec{f}'_n непрерывны, и, поскольку они сходятся локально равномерно

к g , то, согласно теореме 109, функция g непрерывна, а, значит, функция f принадлежит классу C^1 , и теорема полностью доказана.

Замечания. 1°) Поскольку в теореме речь идет лишь о дифференциальном исчислении, то ее можно было привести в гл. III. Однако мы решили собрать в этом параграфе ряд результатов одного характера.

2°) Конечно, первая часть теоремы будет неверной, если не считать пространство F полным, а множество Ω — связным.

Приведем контрпример для случая, когда множество Ω не связно. Пусть Ω — дополнение к началу координат на вещественной прямой \mathbb{R} , а f_n — функция, равная нулю для $x < 0$ и равная постоянной n для $x > 0$. Тогда производные f'_n тождественно равны нулю, и, следовательно, последовательность равномерно сходится к нулю при n , стремящемся к бесконечности. Функции f_n сходятся к некоторому пределу в любой точке $x < 0$. Однако они не сходятся ни к какому пределу на всей вещественной оси \mathbb{R} .

3°) Существенным во всех этих вопросах является равномерная локальная сходимость. Если даже предполагать, что сходимость f'_n к g равномерна в Ω , то можно утверждать лишь, что f_n сходится к f локально равномерно.

Рассмотрим, к примеру, на вещественной прямой функции $f_n(x) = x/n$. Они просто сходятся к 0 при n , стремящемся к бесконечности. Их производные $f'_n(x) = 1/n$ сходятся к 0 равномерно на \mathbb{R} . Отсюда следует (и это было ясно a priori), что функции f_n сходятся к 0 равномерно на любом компакте (или локально равномерно). Однако они не сходятся к 0 равномерно на \mathbb{R} , поскольку $\|f_n\| = +\infty$.

Следствие 1. Пусть f_n — отображение открытого множества Ω аффинного нормированного пространства E в аффинное нормированное пространство F , принадлежащее классу C^m . Если для каждого $k \leq m$ функции $f_n^{(k)}$, определенные на Ω , со значениями в $\mathcal{L}_k(E^k; F)$ локально равномерно сходятся к некоторому пределу f_k и $f_0 = f$, то f принадлежит классу C^m и $f_k = f^{(k)}$ для $k \leq m$.

Доказательство очевидно.

Следствие 2. Если Ω является открытым множеством аффинного нормированного пространства E , а F — аффинным банаховым пространством, то пространство $(F^\Omega)_{cb; m}$ полно.

Напомним, что $f \in (F^\Omega)_{cb; m}$, если f является непрерывным ограниченным отображением Ω в F и если для каждого $k \leq m$ производная этого отображения порядка k является непрерыв-

ным ограниченным отображением Ω в $\mathcal{L}_k(\vec{E}^k; \vec{F})$. В этом случае $\|f\|_m = \sum_{0 \leq k \leq m} \|f^{(k)}\|_0$ (см. стр. 267).

Доказательство. Пусть f_n — последовательность Коши, принадлежащая $(F^\Omega)_{cb; m}$. Согласно определению норм, производные функции $f_n^{(k)}$, $0 \leq k \leq m$, образуют последовательность Коши в $\mathcal{L}_k(\vec{E}^k; \vec{F})_{cb; 0}^\Omega$. Так как пространство \vec{F} полно, то полно пространство $\mathcal{L}_k(\vec{E}^k; \vec{F})$ (теорема 53 гл. II), а вместе с ним полно и пространство $(\mathcal{L}_k(\vec{E}^k; \vec{F}))_{cb; 0}^\Omega$ (следствие 2 теоремы 65 гл. II). Следовательно, функции $f_n^{(k)}$ сходятся равномерно на Ω к непрерывному ограниченному пределу f_k на Ω .

Согласно следствию 1, функция f принадлежит классу C^m и $f^{(k)} = f_k$, т. е. f лежит в $(F^\Omega)_{cb; m}$. Поскольку каждая последовательность $f_n^{(k)}$ равномерно сходится к $f_k = f^{(k)}$, то f_n сходится к f в пространстве $(F^\Omega)_{cb; m}$, которое тем самым оказывается полным.

Примеры. Предыдущая теорема была сформулирована для последовательностей, но на практике она чаще применяется к рядам.

Пусть задан ряд (IV, 11; 1), где функции u_n определены на открытом множестве Ω аффинного нормированного пространства E и имеют значения в банаховом пространстве \vec{F} . Вообще говоря, заранее известно, что этот ряд просто сходится (и даже локально равномерно) в Ω . Функции u_n непрерывны, а это значит, что сумма ряда \vec{S} также непрерывна. Если u_n дифференцируемы, то возникает вопрос, дифференцируема ли функция \vec{S} . Для его решения применяется следующий метод: *почленно дифференцируют предыдущий ряд и полагают a priori*

$$S'(x) = \sum_{n=0}^{\infty} u'_n(x). \quad (\text{IV, 11; 13})$$

Если ряд из производных $\sum_{n=0}^{\infty} u'_n(x)$ локально равномерно сходится в Ω , то \vec{S} дифференцируема и формула (IV, 11; 13) справедлива.

В самом деле, в этом случае функции $\vec{S}_m = \sum_{n=0}^m \vec{u}_n$ сходятся к \vec{S} , и в то же самое время функции $S'_m = \sum_{n=0}^m u'_n$ локально

равномерно сходятся к $\sum_{n=0}^{\infty} u'_n$. Из теоремы следует, что \vec{S} дифференцируема и ее производная равна предыдущему пределу.

Рассмотрим для примера ряд Фурье, в котором \vec{a}_n и $\vec{S}(x)$ являются векторами некоторого банахового пространства \vec{F} :

$$\vec{S}(x) = \sum_{n=-\infty}^{+\infty} \vec{a}_n e^{inx}, \text{ где } x \text{ вещественно; (IV, 11; 14)}$$

предположим, что \vec{a}_n удовлетворяют неравенству

$$\|\vec{a}_n\| \leq \frac{A}{|n|^{p+2}}, \quad (\text{IV, 11; 15})$$

где p — целое число ≥ 0 и A — постоянная > 0 .

Функция \vec{S} , представимая этим рядом, принадлежит по крайней мере классу C^p . В самом деле, если мы продифференцируем ряд почленно k раз, $k \leq p$, то получим следующие формулы:

$$\vec{S}^{(k)}(x) = \sum_{n=-\infty}^{+\infty} (in)^k \vec{a}_n e^{inx}. \quad (\text{IV, 11; 16})$$

Так как $\|(in)^k \vec{a}_n\| \leq A/|n|^{p+2}$, то каждый из этих рядов равномерно и даже нормально сходится, и, следовательно, необходимый результат вытекает непосредственно из теоремы.

Если же лучшей оценки, чем (IV, 11; 15), нет, то, вообще говоря, мы не можем произвести еще одно дифференцирование.

З а м е ч а н и я. 1°) Тот факт, что теорема не применима, еще не означает, что предел f функций f_n или сумма \vec{S} ряда $\sum_{n=0}^{\infty} \vec{u}_n$ не может иметь производной. Все, что можно сказать о производной \vec{S}' функции \vec{S} в случае, когда ряд из производных расходится, — это то, что производную \vec{S}' , если она существует, нельзя представить в виде ряда $\sum_{n=0}^{\infty} \vec{u}'_n$.

Рассмотрим, например, тригонометрический ряд

$$\sum_{n=1}^{\infty} \frac{\sin 2\pi n x}{\pi n}. \quad (\text{IV, 11; 17})$$

Согласно теореме Абеля (теорема 63 гл. II), он сходится всюду и сходится равномерно в каждом интервале $[\delta, 1 - \delta]$, $\delta > 0$. В самом деле, сумма

$$|\sin(2\pi x) + \sin(2\pi 2x) + \dots + \sin(2\pi nx)|$$

мажорируется величиной $2/|e^{2inx} - 1|$ и, следовательно, числом $2/|e^{2i\pi\delta} - 1|$ при $\delta \leq x \leq 1 - \delta$, а последовательность $1/n!$ убывает и стремится к 0 при n , стремящемся к бесконечности. Согласно (II, 14; 31), остаток при этом допускает оценку:

$$|R_n| \leq \frac{2}{\pi(n+1)|e^{2i\pi\delta} - 1|},$$

доказывающую наше утверждение. Однако сходимость, очевидно, не будет равномерной на $[0, 1]$. Следовательно, имеет место локально равномерная сходимость на $]0, 1[$, а ряд представляет собой непрерывную функцию на этом открытом интервале. Почленное дифференцирование дает всюду расходящийся ряд

$$2 \sum_{n=1}^{\infty} \cos 2\pi n x. \quad (\text{IV, 11; 17}_2)$$

Отсюда вовсе не следует, что функция S не дифференцируема. Позже мы увидим, что S определяется по формуле:

$$S(x) = \begin{cases} 0 & \text{для } x = 0 \text{ и } x = 1, \\ \frac{1}{2} - x & \text{для } 0 < x < 1. \end{cases} \quad (\text{IV, 11; 18})$$

Эта функция разрывна при $x = 0$ и $x = 1$. Из-за этого разрыва, несмотря на сходимость на $[0, 1]$, равномерная сходимость ряда имеет место только в интервале $[\delta, 1 - \delta]$, $\delta > 0$.

Во всем интервале $]0, 1[$ эта функция дифференцируема и даже принадлежит классу C^∞ . Ее производную, равную -1 , в виде ряда (IV, 11; 17) представить нельзя, поскольку он расходится. Из этого примера видно, что метод, с помощью которого Вейерштрасс доказал, что ряд (IV, 11; 4) представляет функцию, не имеющую ни в одной точке производной, не является очевидным. *Из того факта, что почленное дифференцирование ряда приводит к расходящемуся ряду, вовсе не следует, что рассматриваемая функция не дифференцируема. Всего навсего он говорит лишь о том, что такая возможность не исключена!*

Дифференцируемость бесконечного произведения

Предположим, что некоторая функция определена в виде бесконечного произведения функций с вещественными или комплексными значениями:

$$\Pi(x) = \prod_{n=0}^{\infty} u_n(x). \quad (\text{IV, 11; 19})$$

В этом случае можно, опираясь на сходимость этого произведения и сходимость ряда логарифмических производных

$$\sum_{n=0}^{\infty} \frac{u_n'(x)}{u_n(x)}, \quad (\text{IV}, 11; 20)$$

выяснить, будет ли функция Π дифференцируема и будет ли ее логарифмическая производная определяться по формуле

$$\frac{\Pi'}{\Pi} = \sum_{n=0}^{\infty} \frac{u_n'}{u_n}. \quad (\text{IV}, 11; 21)$$

Следует заметить, что доказательство этого утверждения нельзя свести непосредственно к рассмотрению функций $\ln u_n$, ибо речь здесь идет о функциях с комплексными значениями, не обязательно лежащими в области $\{z = re^{i\varphi}, r > 0, -\pi < \varphi < \pi\}$ комплексной плоскости, а, значит, нет уверенности в том, что вычисление логарифмов законно. Однако, если f является дифференцируемой функцией, определенной в открытом множестве аффинного нормированного пространства E и принимающей всюду комплексные значения $\neq 0$, то логарифмическую производную f'/f можно вычислять без предварительного перехода к логарифмам¹⁾.

Теперь можно сделать следующие замечания.

1°) Если f и g — функции с комплексными значениями, дифференцируемые и $\neq 0$ на Ω , то логарифмическая производная произведения (соответственно дроби) является суммой (соответственно разностью) логарифмических производных:

$$\frac{(fg)'}{fg} = \frac{f'}{f} + \frac{g'}{g}, \quad \frac{(f/g)'}{f/g} = \frac{f'}{f} - \frac{g'}{g}. \quad (\text{IV}, 11; 22)$$

В самом деле, речь идет об уже известных формулах

$$(fg)' = f'g + fg', \quad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}. \quad (\text{IV}, 11; 23)$$

2°) Если f — дифференцируемая на Ω функция с комплексными значениями, то по теореме о сложной функции логарифмическая производная функции e^f равна $(e^f f')/e^f = f'$.

3°) Если множество Ω связно, а логарифмическая производная функции f равна нулю в Ω , то эта функция постоянна. В самом деле, производная f' этой функции равна нулю.

¹⁾ $f(x)$ является элементом поля K ($= \mathbb{R}$ или \mathbb{C}), а $f'(x)$ является линейным непрерывным отображением \vec{E} в K . Поэтому $f'(x)/f(x)$ также является линейным непрерывным отображением \vec{E} в K (элементом пространства, сопряженного к пространству \vec{E}).

Отсюда следует, что если f и g — две функции, имеющие одну и ту же логарифмическую производную, то их отношение постоянно, ибо логарифмическая производная отношения f/g равна нулю.

4°) Для комплексных функций f_n на топологическом пространстве Ω (см. стр. 168) мы отмечали, что *равномерная сходимость* f_n к f и *равномерная сходимость* f_n/f к 1 (если f *всюду* $\neq 0$) не эквивалентны. Однако, если функции f_n *непрерывны*, то *соответствующие локально равномерные сходимости* эквивалентны (см. стр. 168 гл. II).

Точно так же, если функции f_n (соответственно g_n) сходятся равномерно к f (соответственно к g) при n , стремящемся к бесконечности, то отсюда еще не следует, что $f_n g_n$ сходятся равномерно к $f g$ ¹⁾. Однако это утверждение справедливо для *равномерной сходимости*, если f_n и g_n *непрерывны*.

В самом деле, пусть $a \in \Omega$. Тогда найдется такая окрестность \mathcal{U}' точки a , в которой f_n (соответственно g_n) равномерно сходятся к f (соответственно к g). Обозначим через \mathcal{U}'' такую окрестность точки a , в которой функции f и g , непрерывные как локально равномерные пределы непрерывных функций, ограничены по модулю. Тогда имеем:

$$|f_n g_n - fg| \leq |f_n||g_n - g| + |f_n - f||g|. \quad (\text{IV}, 11; 23_2)$$

Покажем, что на множестве $\mathcal{U} = \mathcal{U}' \cap \mathcal{U}''$ правая часть равномерно сходится к 0. Второе слагаемое, очевидно, обладает этим свойством. Что же касается первого слагаемого, то заметим, что существует такое целое число n_0 , что для $n \geq n_0$ выполняется неравенство $|f_n - f| \leq 1$. Так как функция $|f|$ ограничена в \mathcal{U}'' , то при $n \geq n_0$ функция $|f_n|$ также будет ограниченной в \mathcal{U}'' , а, значит, первое слагаемое также стремится равномерно к 0 в \mathcal{U} .

5°) Если Ω является интервалом \mathbb{R} , а f — комплексной функцией $\neq 0$ на Ω , принадлежащей классу C^1 , то имеет место формула:

$$f(x) = f(a) e^{a \int_a^x (f'(\xi)/f(\xi)) d\xi}. \quad (\text{IV}, 11; 24)$$

В самом деле, левая и правая части, согласно п. 2°), имеют едину и ту же логарифмическую производную f'/f ; следовательно, они, согласно п. 3°), пропорциональны, а так как обе части совпадают при $x = a$, то они равны между собой.

¹⁾ Произведение, как отображение $\mathbb{C} \times \mathbb{C}$ в \mathbb{C} , не является равномерно непрерывным.

В случае, когда $|f'/f| \leq M$, можно указать следующие оценки:

$$\left| \frac{f(x)}{f(a)} \right| \leq e^{M|x-a|} \quad (\text{IV, 11; 25})$$

и

$$\left| \frac{f(x)}{f(a)} - 1 \right| \leq e^{M|x-a|} - 1. \quad (\text{IV, 11; 26})$$

Если теперь f является комплексной функцией класса C^1 , заданной на открытом множестве Ω аффинного пространства E , всюду $\neq 0$, и если ее логарифмическая производная по норме мажорируется числом M , то, применяя неравенства (IV, 11; 25) и

(IV, 11; 26) к функции $t \rightarrow f(a + t \overrightarrow{x-a})$, определенной на $[0, 1]$, получим такие же неравенства, где вместо $|x-a|$ будет стоять $\|\overrightarrow{x-a}\|$.

Имеет место следующий результат:

Теорема 112. Пусть f_n — последовательность комплексных функций класса C^1 , определенных на открытом множестве Ω аффинного нормированного пространства E и всюду $\neq 0$. Если их логарифмические производные f'_n/f_n локально равномерно сходятся к пределу g и если значения функций $f_n(a)$ сходятся к некоторому пределу $f(a) \neq 0$ ¹) хотя бы в одной точке $a \in \Omega$, а множество Ω связно, то f_n сходятся к некоторой предельной функции f локально равномерно в Ω . Функция f нигде не обращается в нуль, она принадлежит классу C^1 и имеет функцию g в качестве логарифмической производной. Если Ω не связно, но заранее известно, что f_n просто сходятся всюду в Ω к некоторой не равной нулю функции, то сформулированные выводы полностью сохраняются.

Доказательство. Положим $f_{mn} = f_m/f_n$. Пусть $B = B(a; \rho)$ — некоторый шар с центром в a , в котором функция f'_n/f_n равномерно сходится к g . Тогда при m и n , стремящихся к бесконечности, f'_{mn}/f_{mn} равномерно сходится к 0 в B . Из (IV, 11; 26) вытекает, что $f_{mn}(x)/f_{mn}(a)$ равномерно сходится к 1. Поскольку $f_{mn}(a)$ сходится к 1, то $f_{mn} = f_m/f_n$ при m и n , стремящихся к бесконечности, равномерно сходятся к 1.

В частности, существует такое n_0 , что для всех $n \geq n_0$ имеем

$$\left| \frac{f_n}{f_{n_0}} - 1 \right| \leq \frac{1}{2}, \text{ а, значит, для } n \geq n_0 \text{ и } x \in B \text{ выполняется нे-}$$

¹⁾ Если $f(a) = 0$ и Ω связно, то без труда доказывается, что функции f_n сходятся локально равномерно к нулю.

равенство $\frac{1}{2}|f_{n_0}| \leq |f_n| \leq \frac{3}{2}|f_{n_0}|$. Теперь для всех $x \in B$, $m \geq n_0$, $n \geq n_0$ будет иметь место неравенство

$$|f_m(x) - f_n(x)| \leq |f_n(x)| \left| \frac{f_m(x)}{f_n(x)} - 1 \right| \leq \frac{3}{2} |f_{n_0}(x)| \left| \frac{f_m(x)}{f_n(x)} - 1 \right|,$$

а, следовательно, $|f_m(x) - f_n(x)|$ сходится к нулю при m и n , стремящихся к бесконечности. Поскольку поле комплексных чисел \mathbb{C} полно, отсюда следует, что $f_n(x)$ для всех $x \in B$ имеют предел $f(x)$. Кроме того, из $|f_n(x)| \geq |f_{n_0}(x)|/2$ следует, что $|f(x)| \geq |f_{n_0}(x)|/2 > 0$, т. е. что f не обращается в нуль в B . Наконец, если по заданному $\varepsilon > 0$ выбрать p так, чтобы при $m \geq p$ и $n \geq p$ для всех $x \in B$ имело место неравенство $|f_m(x)/f_n(x) - 1| \leq \varepsilon$, то при $n \geq p$ для всех $x \in B$ будет справедливым неравенство $|f(x)/f_n(x) - 1| \leq \varepsilon$, а это означает, что f/f_n равномерно сходится к 1 в B .

Теперь, используя метод теоремы 111, можно убедиться, что если Ω связно (или если известно заранее, что f_n сходятся просто к некоторой функции f , которая всюду $\neq 0$), то f_n сходятся просто к некоторому пределу f всюду $\neq 0$ и f/f_n сходятся локально равномерно к 1, или f_n сходятся локально равномерно к f согласно результатам п. 4°) на стр. 789. Отношение f'_n/f_n тогда локально равномерно сходится к g , а f_n к f и поэтому $f'_n = (f'_n/f_n)f_n$ сходятся к gf , поскольку речь идет о непрерывных функциях (см. замечание 4°) на стр. 789). Теорема 111 теперь утверждает, что функция f дифференцируема и ее производная равна gf , т. е. логарифмическая производная равна g , чем и заканчивается доказательство теоремы.

Следствие. Если комплексные функции u_n на Ω принадлежат классу C^1 и всюду $\neq 0$ и если ряд $\sum_{n=0}^{\infty} u'_n/u_n$ локально равномерно в Ω сходится к некоторой предельной функции g и, кроме того, произведение $\prod_{n=0}^{\infty} u_n$ сходится хотя бы в одной точке a из Ω , то когда Ω связно, это произведение сходится на всем Ω локально равномерно, а определяемая им функция Π принадлежит классу C^1 и имеет логарифмическую производную, равную g . Если множество Ω не связно, но заранее известно, что произведение $\prod_{n=0}^{\infty} u_n(x)$ сходится для любого $x \in \Omega$, то указанные выводы полностью сохраняются.

Функции, представимые интегралами

Рассмотрим интеграл

$$\vec{f}(x) = \int_T \vec{h}(x, t) d\mu(t), \quad (\text{IV}, 11; 26_2)$$

где μ — мера Радона ≥ 0 на локально компактном пространстве T , счетном в бесконечности, и \vec{h} — функция на $X \times T$ со значениями в банаховом пространстве \vec{F} .

Предположим, что для каждого $x \in X$ частная функция $\vec{h}_x: t \rightarrow \vec{h}(x, t)$ является μ -интегрируемой. В дальнейшем это условие будет считаться выполненным, и мы не будем больше об этом напоминать. Предыдущий интеграл в этом случае определяет некоторый вектор $\vec{f}(x)$ из \vec{F} и, следовательно, определяет некоторую функцию $\vec{f}: x \rightarrow \vec{f}(x)$. Рассмотрим задачу изучения непрерывности, интегрируемости и дифференцируемости этой функции, исходя из аналогичных свойств функции \vec{h} .

Непрерывность функции, представимой интегралом

Наиболее важный критерий дает теорема 35 Лебега, которую мы сформулируем следующим образом.

Теорема 114. Пусть пространство X метризуемо. Если для μ -почти всех значений t функция \vec{h} раздельно непрерывна по x в точке a и существует такая окрестность \mathcal{U} точки a , для которой на $\mathcal{U} \times T$ имеет место оценка $\|\vec{h}(x, t)\| \leq k(t)$, где k есть μ -интегрируемая функция ≥ 0 на T , то \vec{f} непрерывна в точке a пространства X .

В самом деле, если x устремить к a , то, начиная с некоторого момента, x будет находиться в \mathcal{U} , и тогда, в силу теоремы 35 Лебега и следующего за ней замечания, $\vec{f}(x)$ будет стремится к $\vec{f}(a)$.

Надо быть внимательнее и не смешивать различные предположения.

Для μ -почти всех значений t мы считаем, что функция \vec{h} непрерывна по x в точке a . Это естественно, поскольку изучается непрерывность функции \vec{f} по x в точке a . С другой стороны, предполагается, что $\|\vec{h}\|$ мажорируется μ -интегрируемой функцией k переменной t , что также естественно, поскольку

производится интегрирование по μ в пространстве T переменной t .

Следствие. Пусть \vec{h} — непрерывная функция на $X \times T$ со значениями в \vec{F} и для каждой точки $a \in X$ существует такая ее окрестность \mathcal{U} и такой компакт K из T , что для $x \in \mathcal{U}$ носитель частной функции \vec{h}_x : $t \rightarrow \vec{h}(x, t)$ лежит в K . Тогда \vec{f} является непрерывной функцией из X в \vec{F} .

Доказательство. При x , стремящемся к a , \vec{h}_x сходится к \vec{h}_a в $(\vec{F}^K)_{cb}$ (теорема 66). Следовательно, существует такая окрестность $\mathcal{U}_0 \subset \mathcal{U}$ точки a , что для $x \in \mathcal{U}_0$ норма $\|\vec{h}_x\|$ ограничена. Обозначим через M ее границу. Тогда можно применить предыдущую теорему с $k(t) = M\chi_K(t)$, где χ_K — характеристическая функция компакта K .

Непосредственно можно получить соотношение:

$$\begin{aligned} \|\vec{f}(x) - \vec{f}(a)\| &\leq \mu(K) \sup_{t \in K} \|\vec{h}(x, t) - \vec{h}(a, t)\| = \\ &= \mu(K) \|\vec{h}_x - \vec{h}_a\|, \quad (\text{IV, 11; 26}_3) \end{aligned}$$

где норма берется в $(\vec{F}^K)_{cb}$ и результат получается из теоремы 66 без ссылки на теорему Лебега (это позволяет не предполагать пространство X метризуемым).

Интегрируемость функции, представимой интегралом

Речь идет о следующей задаче. Предполагается, что пространство X локально компактно и снабжено мерой Радона $\lambda \geq 0$. Надо выяснить, будет ли функция \vec{f} λ -интегрируемой и можно ли будет вычислять интеграл от этой функции, интегрируя сначала \vec{h} по λ при фиксированном t , а затем интегрируя результат по t и по мере μ , т. е. имеем ли мы право писать:

$$\int_X d\lambda(x) \int_T \vec{h}(x, t) d\mu(t) = \int_T d\mu(t) \int_X \vec{h}(x, t) d\lambda(x)? \quad (\text{IV, 11; 27})$$

Основные критерии справедливости последнего равенства дают теоремы 77 и 78 Фубини и следствие теоремы 78. Формула (IV, 11; 27) справедлива в том случае, когда функция \vec{h} интегрируема на $X \times T$ по мере тензорного произведения $\lambda \otimes \mu$, или если она измерима и ≥ 0 , или если она измерима и одна из величин, аналогичных (IV, 11; 27), относящихся к $\|\vec{h}\|$, конечна.

Дифференцируемость функции, представимой интегралом

Будем считать, что $X = \Omega$ является открытым множеством аффинного нормированного пространства E .

То, что нам приходилось видеть в рядах, наводит на мысль, что \vec{f} может быть дифференцируемой только тогда, когда интеграл, относящийся к частной производной \vec{h} по x , сам обладает подходящими свойствами.

В противоположность тому, что было сделано в теореме 111, мы не будем стараться свести дело к общему случаю, в котором f предполагается существующей только в точке a области Ω .

Теорема 115. Предположим, что для μ -почти каждого значения t частная функция $x \rightarrow \vec{h}(x, t)$ дифференцируема (соответственно принадлежит классу C^1) на Ω . Пусть ее производная функция $x \rightarrow \frac{\partial \vec{h}}{\partial x}(x, t)$ является отображением Ω в $\mathcal{L}(\vec{E}; \vec{F})$.

Предположим, что в каждой точке a функция $t \rightarrow \frac{\partial \vec{h}}{\partial x}(a, t)$ (определенная μ -почти всюду на T со значениями в $\mathcal{L}(\vec{E}, \vec{F})$) μ -измерима¹⁾ и что существует в Ω такая окрестность U точки a , что на $U \times T$ имеет место неравенство $\left\| \frac{\partial \vec{h}}{\partial x}(x, t) \right\| \leq k(t)$, где k — неотрицательная и μ -интегрируемая функция на T .

При этих условиях функция \vec{f} , отображающая Ω в \vec{F} , дифференцируема (соответственно принадлежит классу C^1), и для каждой точки $a \in \Omega$ ее производная вычисляется по формуле

$$\vec{f}'(a) = \int_T \frac{\partial \vec{h}}{\partial x}(a, t) d\mu(t) \in \mathcal{L}(\vec{E}; \vec{F}). \quad (\text{IV, 11; 28})$$

¹⁾ Эта измеримость будет автоматически выполняться, если пространство E конечномерно. Предположим для простоты, что $E = \mathbb{K}$ есть поле скаляров. Тогда для μ -почти всех значений t имеет место равенство

$$\frac{\partial \vec{h}}{\partial x}(x, t) = \lim_{\substack{\xi \rightarrow 0 \\ \xi \neq 0}} \frac{\vec{h}(x + \xi, t) - \vec{h}(x, t)}{\xi} \in \vec{F}.$$

Функции $t \rightarrow \vec{h}(x + \xi, t)$ и $t \rightarrow \vec{h}(x, t)$ μ -измеримы (и даже μ -интегрируемы). Следовательно, для фиксированного x функция $t \rightarrow \frac{\partial \vec{h}}{\partial x}(x, t)$ является μ -почти всюду пределом некоторой последовательности μ -измеримых функций, а, следовательно, μ -измерима.

Напомним, кроме того, что все встречающиеся на практике функции измеримы!

Прежде чем проводить доказательство этой теоремы, сделаем несколько замечаний по поводу ее условий.

Подинтегральная функция при заданной точке a определена μ -почти всюду на T , измерима и мажорируема некоторой интегрируемой функцией $k \geq 0$, а, значит, интегрируема (согласно теореме 50 гл. II, пространство $\mathcal{L}(\vec{E}; \vec{F})$ полно). Таким образом, условия теоремы обеспечивают существование правой части равенства (IV, 11; 28).

Добавим, что если для μ -почти всех значений t частная функция \vec{h}_t не только дифференцируема по x , но и принадлежит классу C^1 в Ω , то условие мажорируемости производной функцией k влечет за собой непрерывность производной \vec{f} , т. е. принадлежность \vec{f} классу C^1 на Ω . Однако это условие необходимо даже для доказательства дифференцируемости \vec{f} .

Доказательство. Надо доказать только тот факт, что функция \vec{f} имеет производную, определяемую формулой (IV, 11; 28). Все сводится к тому, чтобы доказать, что при \vec{X} , стремящемся к $\vec{0}$, величина

$$\frac{1}{\|\vec{X}\|} \left\| \int_T \left(\vec{h}(a + \vec{X}, t) - \vec{h}(a, t) - \frac{\partial \vec{h}}{\partial x}(a, t) \vec{X} \right) d\mu(t) \right\| \quad (\text{IV, 11; 29})$$

стремится к 0. Эта величина не превосходит

$$\int_T \left(\frac{1}{\|\vec{X}\|} \left\| \vec{h}(a + \vec{X}, t) - \vec{h}(a, t) - \frac{\partial \vec{h}}{\partial x}(a, t) \vec{X} \right\| \right) d\mu(t). \quad (\text{IV, 11; 30})$$

В силу частной дифференцируемости \vec{h} , подинтегральная функция t сходится просто μ -почти всюду к $\vec{0}$, когда \vec{X} стремится к $\vec{0}$. Если мы сможем промажорировать ее некоторой неотрицательной μ -интегрируемой функцией t , то результат будет получен в силу теоремы 35 Лебега. Далее, если через ρ обозначить такое число > 0 , при котором шар $B(a; \rho)$ будет содержаться в \mathcal{V} , то для μ -почти всех значений t можно применить формулу конечных приращений в виде следствия 1 теоремы 13 гл. III, лишь бы только $\|\vec{X}\| \leq \rho$, а это даст

$$\begin{aligned} \frac{1}{\|\vec{X}\|} \left\| \vec{h}(a + \vec{X}, t) - \vec{h}(a, t) - \frac{\partial \vec{h}}{\partial x}(a, t) \vec{X} \right\| &\leq \\ &\leq \sup_{\xi \in [a, a + \vec{X}]} \left\| \frac{\partial \vec{h}}{\partial x}(\xi, t) - \frac{\partial \vec{h}}{\partial x}(a, t) \right\| \leq 2k(t) \end{aligned} \quad (\text{IV, 11; 31})$$

(где k по условию ≥ 0 и μ -интегрируема), откуда следует доказательство теоремы.

Практически поступают следующим образом. Если функция \vec{f} определена интегралом (IV, 11; 26₂), то для того, чтобы убедиться в ее дифференцируемости, формально дифференцируют под знаком интеграла и пишут равенство (IV, 11; 28). Формула будет обоснованной, если выполнены условия теоремы (которые в случае, когда h_1 — почти всюду на Ω принадлежит классу C^1 , позволяют утверждать, что функция, определяемая правой частью равенства (IV, 11; 28), непрерывна).

Укажем простое, но весьма важное следствие.

Следствие. Пусть \vec{f} — функция, определенная по формуле (IV, 11; 26₂) при следующих условиях:

- a) \vec{h} является непрерывным отображением $\Omega \times T$ в \vec{F} ;
- b) \vec{h} имеет всюду частную производную по x и $\partial \vec{h} / \partial x$ является непрерывным отображением $\Omega \times T$ в $\mathcal{L}(\vec{E}; \vec{F})$;
- c) для каждой точки $a \in \Omega$ существует в Ω такая окрестность \mathcal{U} этой точки и такой компакт K из T , что для всех x из \mathcal{U} непрерывная на T частная функция \vec{h}_x со значениями в \vec{F} имеет носитель в K .

Тогда функция \vec{f} принадлежит классу C^1 в Ω и ее производная задается формулой (IV, 11; 28).

Доказательство аналогично доказательству следствия теоремы 114. Следует только убедиться, что из этих условий вытекают условия теоремы 115.

Замечание относительно исследования последовательных первообразных функций, непрерывной на интервале прямой \mathbb{R} . Вернемся к ситуации, которая рассматривалась в теореме 91.

Функция \vec{F}_m , определенная по формуле (IV, 9; 41), является первообразной функции \vec{f} порядка m , обращающейся в нуль в точке c вместе со всеми своими производными до порядка $m - 1$ включительно. Это было доказано с помощью преобразования двух последовательных интегралов к одному интегралу. Теперь мы можем дать другое доказательство этого утверждения путем непосредственной его проверки.

Дифференцируема ли функция \vec{F}_m ? Непосредственно этого не видно. Эта функция зависит от x , с одной стороны, потому, что x является верхним пределом интегрирования, а с другой, — потому, что x находится под знаком интеграла.

Функцию $\vec{F}_m(x)$ можно рассматривать как значение при $y = x$ функции двух переменных

$$\vec{G}_m(x, y) = \int_c^y \vec{f}(\xi) \frac{(x - \xi)^{m-1}}{(m-1)!} d\xi. \quad (\text{IV}, 11; 32)$$

Покажем, что эта функция имеет полную производную. Для этого достаточно доказать, что она имеет непрерывные частные производные по x и y , а затем применить теорему 15 гл. III (см. далее формулу (IV, 11; 36)). Прежде всего, согласно теореме 89,

$$\frac{\partial \vec{G}_m}{\partial y}(x, y) = \vec{f}(y) \frac{(x - y)^{m-1}}{(m-1)!}, \quad (\text{IV}, 11; 33)$$

и полученная функция непрерывна по (x, y) .

Для фиксированного y к функции $(x, \xi) \rightarrow \vec{f}(\xi) \frac{(x - \xi)^{m-1}}{(m-1)!}$ применима теорема 115 (здесь вместо x и t у нас написаны x и ξ , вместо $d\mu(t)$ написано $d\xi$, а отрезок $[c, y]$ взят в качестве компакта $K = T$). Поэтому функция \vec{G}_m имеет частную производную по x , определяемую по формуле

$$\frac{\partial \vec{G}_m}{\partial x}(x, y) = \int_c^y \vec{f}(\xi) \frac{(x - \xi)^{m-2}}{(m-2)!} d\xi \quad \text{для } m \geq 2. \quad (\text{IV}, 11; 34)$$

Покажем, что $\partial \vec{G}_m / \partial x$ также непрерывна по (x, y) ¹⁾. Зафиксируем (x_0, y_0) и предположим, например, что $y_0 \geq c$. Тогда можно написать:

$$\frac{\partial \vec{G}_m}{\partial x}(x, y) = \int_{R_1} \Phi_y(\xi) \vec{f}(\xi) \frac{(x - \xi)^{m-2}}{(m-2)!} d\xi, \quad (\text{IV}, 11; 35)$$

где Φ_y является характеристической функцией интервала $[c, y]$. Для фиксированного ξ при (x, y) , стремящемся к (x_0, y_0) , $\Phi_y(\xi) \vec{f}(\xi) \frac{(x - \xi)^{m-2}}{(m-2)!}$ сходится к $\Phi_{y_0}(\xi) \vec{f}(\xi) \frac{(x_0 - \xi)^{m-2}}{(m-2)!}$ всюду, кроме, быть может, $\xi = y_0$ (поскольку тогда $\Phi_y(\xi)$ не сходится к $\Phi_{y_0}(\xi)$, если y стремится к y_0 по значениям $< y_0$). Следовательно, для $d\xi$ -почти всех значений ξ (всех, кроме одного: y_0) частная функция $(x, y) \rightarrow \Phi_y(\xi) \vec{f}(\xi) \frac{(x - \xi)^{m-2}}{(m-2)!}$ непрерывна в точке (x_0, y_0) . Если x и y будут пробегать компактные окре-

¹⁾ В действительности, если учсть замечание 1°), следующее за теоремой 15 гл. III, проводить доказательство этого факта не нужно.

стности точек x_0, y_0 в \mathbb{R}_1 , то подинтегральная функция будет ограниченной, иметь носитель в фиксированном компакте, а, следовательно, из теоремы 114 будет следовать, что $\partial \vec{G}_m / \partial x$ непрерывна в точке (x_0, y_0) .

Таким образом, \vec{G}_m принадлежит классу C^1 на $\mathbb{R}_1 \times \mathbb{R}_1$. Но тогда по теореме о сложной функции (следствие 5 теоремы 11) функция \vec{F}_m принадлежит классу C^1 и ее производная вычисляется по формуле

$$\frac{d\vec{F}_m}{dx}(x) = \frac{\partial \vec{G}_m}{\partial x}(x, x) + \frac{\partial \vec{G}_m}{\partial y}(x, x). \quad (\text{IV}, 11; 36)$$

Если учесть, что $(x - y)^{m-2}$ при $m \geq 2$ обращается в нуль при $y = x$, то для $m \geq 2$ получаем:

$$\frac{d\vec{F}_m}{dx}(x) = \frac{\partial \vec{G}_m}{\partial x}(x, x) = \int_c^x \vec{f}(\xi) \frac{(x - \xi)^{m-2}}{(m-2)!} d\xi. \quad (\text{IV}, 11; 37)$$

Таким же образом можно поступать и далее, последовательно дойти до производной $(m-1)$ -го порядка включительно и получить формулу

$$\frac{d^k}{dx^k} \vec{F}_m(x) = \int_c^x \vec{f}(\xi) \frac{(x - \xi)^{m-k-1}}{(m-k-1)!} d\xi = \vec{F}_{m-k}(x), \quad k \leq m-1, \quad (\text{IV}, 11; 38)$$

из которой следует, что производные в точке c функции \vec{F}_m до $(m-1)$ -го порядка обращаются в нуль.

Вычислим теперь производную порядка m от функции \vec{F}_m .

К функции \vec{F}_1 можно применить ту же формулу, но на этот раз мы получим:

$$\begin{aligned} \vec{F}_1(x) &= \int_c^x \vec{f}(\xi) d\xi, \\ \frac{d\vec{F}_1}{dx}(x) &= \vec{f}(x). \end{aligned} \quad (\text{IV}, 11; 39)$$

Таким образом, нами доказано, что производная m -го порядка функции \vec{F}_m равна \vec{f} , и, следовательно, функция \vec{F}_m , определенная по формуле (IV, 9; 41), является первообразной порядка m для функции \vec{f} , обращающейся вместе со своими производными до $(m-1)$ -го порядка включительно в нуль в точке c , что является новым доказательством теоремы 91.

Случай несобственных сходящихся интегралов

Пусть \tilde{h} — функция, определенная на произведении $X \times \mathbb{R}_1$, со значениями в банаховом пространстве \tilde{F} . Здесь \mathbb{R}_1 — некоторый интервал $[a, b]$ прямой \mathbb{R} . Рассмотрим интеграл

$$\tilde{f}(x) = \int_{[a, b]} \tilde{h}(x, t) d\mu(t). \quad (\text{IV}, 11; 40)$$

Будем предполагать, что μ является некоторой мерой Радона на \mathbb{R}_1 , и что для каждого $x \in X$ несобственный интеграл, определяющий функцию $\tilde{f}(x)$ (см. определение на стр. 724), является сходящимся. Говорят, что несобственный интеграл равномерно сходится при x , пробегающем некоторую часть A множества X , если собственный интеграл $\int_{[a, b']}$ сходится

к несобственному интегралу $\int_{[a, b']}$ (при b' , стремящемся к b

так, что $b' < b$) равномерно относительно $x \in A$. Это еще означает, что «остаток» $\int_{[b', b]}$ сходится к 0 при b' , стремя-

щемся к b равномерно относительно $x \in A$. Можно сказать также, что при любом $\varepsilon > 0$ можно найти такое число $b_0 < b$ (не зависящее от x), что для любого $x \in A$ из $b_0 \leq b' < b$ следует неравенство $\left\| \int_{[b', b]} \right\| \leq \varepsilon$.

Необходимые свойства функции f могут быть получены в два приема. Сначала доказывают, что собственный интеграл $\int_{[a, b']}$ обладает требуемыми свойствами, а затем устремляют b' к b .

Не ставя целью рассмотреть возможно более общие случаи, мы приведем следующие теоремы:

Теорема 116. Пусть \tilde{f} — функция, определенная на X по формуле (IV, 11; 40), в которой несобственный интеграл предполагается сходящимся для каждого $x \in X$. Предположим, что

1°) X является метризуемым пространством и для каждого $b' < b$ собственный интеграл

$$\tilde{f}_{b'}(x) = \int_{[a, b']} \tilde{h}(x, t) d\mu(t) \quad (\text{IV}, 11; 41)$$

определяет на X некоторую функцию $\tilde{f}_{b'}$ со значениями в \tilde{F} ,

непрерывную в точке $x = x_0$ пространства X (например, по теореме 114).

2°) Существует такая окрестность \mathcal{U} точки x_0 , что несобственный интеграл равномерно сходится относительно x из \mathcal{U} .

В этом случае функция \vec{f} непрерывна в точке x_0 пространства X .

Теорема очевидна. По предположению, функция $\vec{f}_{b'}$ непрерывна на \mathcal{U} в точке x_0 и функция $\vec{f}_{b'}$ сходится к $\vec{f}_b = \vec{f}$ равномерно на \mathcal{U} . Остается лишь применить теорему 65 гл. II.

Теорема 117. Пусть задан несобственный интеграл (IV, 11; 40), в котором $X = \Omega$ есть открытое множество аффинного нормированного пространства E . Будем предполагать, что

1°) Для каждого $b' < b$ собственный интеграл (IV, 11; 41) представляет собой некоторую функцию $\vec{f}_{b'}$ на Ω , удовлетворяющую условиям теоремы 115, из которой следует, что $\vec{f}_{b'}$ дифференцируема (соответственно принадлежит классу C^1) на Ω , а ее производная определяется по формуле

$$\vec{f}'_{b'}(x) = \int_{|a, b'|} \frac{\partial h}{\partial x}(x, t) d\mu(t). \quad (\text{IV, 11; 42})$$

2°) Несобственный интеграл

$$g(x) = \int_{|a, \rightarrow b|} \frac{\partial h}{\partial x}(x, t) d\mu(t) \quad (\text{IV, 11; 43})$$

сходится для каждого $x \in \Omega$ локально равномерно на Ω .

3°) Множество Ω связно, F полно, а несобственный интеграл (IV, 11; 40) сходится хотя бы в одной точке $x = a$ области Ω или же Ω и F произвольны, а этот несобственный интеграл сходится для каждого $x \in \Omega$.

Тогда несобственный интеграл (IV, 11; 40) сходится для каждого x из Ω ; эта сходимость локально равномерна на Ω , функция \vec{f} , определенная по формуле (IV, 11; 40), дифференцируема (соответственно принадлежит классу C^1) на Ω , а ее производной является функция g .

В самом деле, когда b' стремится к b , функции $\vec{f}_{b'}$ удовлетворяют условиям теоремы 111¹).

¹⁾ Здесь речь идет не о последовательности \vec{f}_n , а о семействе $\vec{f}_{b'}$, зависящем от вещественного параметра b' и некотором пределе при b' , стремящемся к b . Легко проверяется, что для этих функций все сказанное остается в силе.

Пример 1. Рассмотрим интеграл (в котором $x \in \mathbb{R}$):

$$f(x) = \int_{\mathbb{R}} \frac{e^{itx}}{1+t^2} dt. \quad (\text{IV, 11; 44})$$

Функция $t \rightarrow 1/(1+t^2)$ интегрируема на \mathbb{R} . Поэтому, применяя теорему 114 с $\mathcal{Y} = \mathbb{R}$ и $k(t) = 1/(1+t^2)$, получаем, что функция f непрерывна на \mathbb{R} .

Дифференцируя по x под знаком интеграла, получаем интеграл

$$\int_{\mathbb{R}} \frac{ite^{itx}}{1+t^2} dt, \quad (\text{IV, 11; 45})$$

не имеющий смысла, поскольку функция $t/(1+t^2)$ не интегрируема. Поэтому доказать дифференцируемость f , рассматривая лишь собственные интегралы (теорема 115), невозможно.

Рассмотрим теперь несобственный интеграл

$$g(x) = \int_{|-\infty \leftarrow, \rightarrow +\infty|} \frac{ite^{itx}}{1+t^2} dt. \quad (\text{IV, 11; 46})$$

Применяя теорему Абеля, можно убедиться, что для $x \neq 0$ он является условно сходящимся. В самом деле, функция $t \rightarrow t/(1+t^2)$ (производная которой равна $(1-t^2)/[(1+t^2)^2]$) монотонна в каждом из интервалов $]-\infty, -1]$, $[-1, +1]$ и $[+1, +\infty[$. Так как эта функция непрерывна и при t , стремящемся к бесконечности, стремится к 0, то она ограничена и, согласно следствию теоремы 86, имеет ограниченную вариацию. Функция $t \rightarrow e^{itx}$ имеет при $x \neq 0$ «ограниченные неопределенные интегралы» (оценка (IV, 9; 107)):

$$|\sigma_{c, d}| = \left| \int_c^d e^{itx} dt \right| \leqslant \frac{2}{|x|}. \quad (\text{IV, 11; 47})$$

Поэтому из критерия Абеля (следствие теоремы 98) следует условная сходимость (IV, 11; 46) для $x \neq 0$. Эта условная сходимость локально равномерна на открытом множестве $\Omega = C_0$ прямой \mathbb{R} . В самом деле, для каждого заданного числа $\delta > 0$ при $c \geqslant 1$ и $|x| \geqslant \delta$ имеет место оценка (в обозначениях теоремы Абеля):

$$\left| \int_{|-\infty \leftarrow, -c]} \right| \text{ и } \left| \int_{|+c, \rightarrow +\infty|} \right| \leqslant U(c)V(c) = \frac{c}{1+c^2} \cdot \frac{2}{\delta} \quad (\text{IV, 11; 47}_2)$$

¹⁾ Для $c \geqslant 1$ функция $t/(1+t^2)$ монотонна в $[c, +\infty[$, и ее полная вариация в точности равна $c/(1+c^2)$.

— величина, стремящаяся к 0 при c , стремящемся к бесконечности, и фиксированном δ .

При фиксированном c собственный интеграл

$$f_c(x) = \int_{[-c, +c]} \frac{e^{itx}}{1+t^2} dt, \quad (\text{IV, 11; 48})$$

в силу теоремы 115 или даже ее следствия, принадлежит на Ω (или даже на \mathbb{R}) классу C^1 . Поэтому мы можем устремить c к $+\infty$ и применить теорему 117. Функция f оказывается дифференцируемой в $\Omega = \mathbb{C} \setminus \{0\}$, а ее производная определяется несобственным интегралом g из (IV, 11; 46).

Если теперь мы формально вычислим вторую производную

$$\int \frac{-t^2 e^{itx}}{1+t^2} dt, \quad (\text{IV, 11; 49})$$

то получим выражение, не имеющее смысла ни при каком способе рассуждений.

Позже мы увидим, что функция f задается формулой

$$f(x) = \pi e^{-|x|}. \quad (\text{IV, 11; 50})$$

Эта функция непрерывна на \mathbb{R} , принадлежит классу C^1 на дополнении к началу координат и не имеет производной в начале координат. Однако эта функция принадлежит классу C^∞ , т. е. имеет последовательные производные всех порядков на дополнении к началу координат. Здесь, как и в замечании 1^o) на стр. 787, предыдущие теоремы позволяют предвидеть, когда функция, представимая интегралом, дифференцируема; невыполнение условий теорем еще не означает, что она не будет дифференцируемой.

В качестве второго примера мы докажем очень важную в теории рядов и интегралов Фурье формулу.

Пример 2.

Теорема 118. Имеет место формула

$$\int_{[0, \rightarrow +\infty]} \frac{\sin \lambda t}{t} dt = \frac{\pi}{2}, \quad \lambda > 0. \quad (\text{IV, 11; 51})$$

Эта формула уже была приведена в (IV, 9; 87). Полагая $\lambda t = t'$, перейдем к случаю $\lambda = 1$. Для доказательства рассмотрим функцию J , определенную следующим образом:

$$J(x) = \int_{[0, \rightarrow +\infty]} e^{-xt} \frac{\sin t}{t} dt. \quad (\text{IV, 11; 52})$$

При $x > 0$ под интегралом в правой части стоит интегрируемая функция. Поэтому этот интеграл собственный (теорема 97).

При $x = 0$ это не так. Интеграл будет несобственным сходящимся по теореме Абеля.

Докажем, прежде всего, что интеграл правой части является несобственным интегралом, равномерно сходящимся при $x \geq 0$. В самом деле, если к $\int_{[c, \rightarrow +\infty[} c > 0$ применить теорему Абеля, где

$$u(t) = \frac{e^{-xt}}{t}, \quad v(t) = \sin t, \quad (\text{IV, 11; 53})$$

то u будет убывающей функцией, стремящейся к 0 при t , стремящемся к бесконечности, т. е. функцией ограниченной вариации, а v будет функцией с ограниченными неопределенными интегралами: $\left| \int_{[c, d]} \sin t dt \right| \leq 2$. Поэтому непосредственно применима оценка (IV, 9; 79):

$$\left| \int_{[c, \rightarrow +\infty[} e^{-xt} \frac{\sin t}{t} dt \right| \leq \frac{2e^{-xc}}{c} \leq \frac{2}{c} \quad (\text{IV, 11; 54})$$

— величина, стремящаяся к 0 равномерно относительно $x \geq 0$ при c , стремящемся к бесконечности. Поскольку при конечном фиксированном c , согласно следствию теоремы 114, $J(x, c) = \int_{[0, c]} e^{-xt} \frac{\sin t}{t} dt$ является непрерывной функцией x , то функция J , по теореме 116, непрерывна на полупрямой \mathbb{R}_+ , т. е. на множестве вещественных чисел ≥ 0 . Выясним теперь, будет ли эта функция дифференцируемой? Продифференцируем ее формально под знаком интеграла:

$$J'(x) = - \int_{[0, \rightarrow +\infty[} e^{-xt} \sin t dt. \quad (\text{IV, 11; 55})$$

Интеграл, стоящий в правой части, является интегралом от некоторой интегрируемой при $x > 0$ функции, тогда как при $x = 0$ он является несобственным расходящимся интегралом. Поэтому нет более необходимости пользоваться здесь теоремой Абеля, поскольку нет возможности каким-либо образом получить информацию о производной в начале координат. Однако можно просто воспользоваться теоремой 115. Функция $x \rightarrow e^{-xt} \frac{\sin t}{t}$ для всех значений t принадлежит классу C^1 на $\Omega =]0, +\infty[$, имеет производную, равную $-e^{-xt} \sin t$, и при $x \geq \delta > 0$ выполняется неравенство

$$|e^{-xt} \sin t| \leq e^{-\delta t}, \quad (\text{IV, 11; 56})$$

где справа стоит интегрируемая функция ≥ 0 .

Из теоремы 115 следует, что функция J на $\Omega =]0, +\infty[$ принадлежит классу C^1 и ее производная определяется формулой (IV, 11; 55). Но тогда интеграл, стоящий в правой части, может быть вычислен элементарно. В самом деле, производя дважды интегрирование по частям (теорема 98), его можно записать в виде

$$\begin{aligned} J'(x) &= - \int_{]0, \rightarrow +\infty[} \frac{e^{-xt}}{x} \cos t dt = \\ &= \left[-\frac{e^{-xt}}{x^2} \cos t \right]_{]0, \rightarrow +\infty[}^{+\infty} + \int_{]0, \rightarrow +\infty[} \frac{e^{-xt}}{x^2} \sin t dt = -\frac{1}{x^2} - \frac{J'(x)}{x^2}, \quad (\text{IV, 11; 57}) \end{aligned}$$

откуда

$$J'(x) = -\frac{1}{1+x^2}, \quad x > 0. \quad (\text{IV, 11; 58})$$

Заметим, что $J'(x)$ имеет предел, когда x стремится к 0. Из теоремы 14 гл. III следует, что функция J имеет при $x = 0$ производную и эта производная равна -1 . Однако мы видели, что это невозможно было ни предвидеть, ни представить $J'(0) = -1$ с помощью не имеющего никакого смысла интеграла

$$-\int_{]0, \rightarrow +\infty[} \sin t dt.$$

Из предыдущего следует, что функцию J можно получить, отыскивая первообразную правой части равенства (IV, 11; 58), что дает

$$J(x) = -\arctg x + c. \quad (\text{IV, 11; 59})$$

Постоянная c легко вычисляется. Действительно, при x , стремящемся к $+\infty$, функция $t \rightarrow (e^{-xt} \sin t)/t$ просто сходится к 0, и, поскольку она при $x \geqslant 1$ мажорируется неотрицательной интегрируемой функцией $t \rightarrow (e^{-t} |\sin t|)/t$, то из теоремы Лебега следует, что $J(x)$ стремится к 0 при x , стремящемся к $+\infty$.

Постоянная правой части равенства (IV, 11; 59) тогда равна $\pi/2$. Следовательно, в силу непрерывности функции J в начале координат, $J(0) = \pi/2$, чем и заканчивается доказательство теоремы. В теории эйлеровых функций мы познакомимся с одним обобщением этой формулы.

Применение к делимости дифференцируемых функций

Теорема 119. Пусть \vec{U} — функция, определенная на открытом множестве Ω пространства \mathbb{R}^{N-1}), со значениями в базахом пространстве \vec{F} . Предположим, что пересечение Ω с ги-

¹⁾ Можно было бы всюду заменить \mathbb{R} на \mathbb{C} .

перплоскостью $x_N = 0$ состоит из нулей функции \vec{U} порядка, не меньшего k , т. е. функция \vec{U} на нем равна нулю вместе со своими частными производными до порядка $k - 1$ включительно.

Если функция \vec{U} принадлежит классу C^m с $m \geq k$, то найдется единственная функция \vec{V} , определенная на Ω , со значениями в \vec{F} класса C^{m-k} , такая, что будет иметь место тождество

$$\vec{U}(x) = x_N^k \vec{V}(x). \quad (\text{IV}, 11; 60)$$

В некотором смысле можно сказать, что тот факт, что функция \vec{U} имеет на гиперплоскости $x_N = 0$ нуль порядка k , влечет за собой «делимость \vec{U} на x_N^k ». Следует заметить, что отношение не принадлежит тому же классу дифференцируемости, что и функция \vec{U} .

Доказательство. 1°) Если такая функция \vec{V} существует, то она определяется единственным образом, ибо в каждой точке, где $x_N \neq 0$, она задается отношением $\vec{V}_0(x) = \vec{U}(x)/x_N^k$. Так как по условию $m \geq k$, т. е. $m - k \geq 0$, то эта функция непрерывна и с помощью предельного перехода может быть определена в точке, где $x_N = 0$).

2°) Докажем существование функции \vec{V} . Для простоты предположим, что область Ω выпукла и можно применить формулу Тейлора²⁾. Для каждой точки $x \in \Omega$ с координатами x_1, x_2, \dots, x_N к функции одной переменной $u - \vec{U}(x_1, \dots, x_{N-1}, u)$ можно применить формулу Тейлора до порядка $k - 1$ относительно точки 0 и приращения x_N . Поскольку производные до порядка $k - 1$ включительно на H равны нулю, то эта формула сводится к остаточному члену, и мы можем написать (формула (IV, 9; 49)), что

$$\vec{U}(x_1, \dots, x_N) = x_N^k \int_0^1 \frac{\partial^k \vec{U}}{\partial x_N^k}(x_1, \dots, x_{N-1}, tx_N) \frac{(1-t)^{k-1}}{(k-1)!} dt. \quad (\text{IV}, 11; 61)$$

¹⁾ Это пункт 1°) доказательства теоремы 45 гл. II.

²⁾ Существование \vec{V} будет очевидным, если $\vec{V}_0(x) = \vec{U}(x)/x_N^k$, определенная для $x_N \neq 0$, имеет предел при x , стремящемся к такой точке a , что $a_N = 0$, и если функция, равная $\vec{V}_0(x)$ для $x_N \neq 0$ и равная предыдущему пределу при $x_N = 0$, принадлежит классу C^{m-k} . Таким образом, для каждой точки a , такой, что $a_N = 0$, можно ограничиться рассмотрением того, что происходит в некоторой окрестности точки a . Это позволяет ограничиться открытой выпуклой окрестностью.

Мы видим, что в качестве \vec{V} можно взять функцию, определяемую интегралом

$$\begin{aligned}\vec{V}(x_1, x_2, \dots, x_N) &= \\ &= \int_0^1 \frac{\partial^k \vec{U}}{\partial x_N^k}(x_1, \dots, x_{N-1}, tx_N) \frac{(1-t)^{k-1}}{(k-1)!} dt. \quad (\text{IV, 11; 62})\end{aligned}$$

Эта функция принадлежит классу C^{m-k} в Ω . В самом деле, поскольку функция \vec{U} , по предположению, принадлежит классу C^m , то ее производная $\partial^k \vec{U} / \partial x_N^k$ принадлежит классу C^{m-k} , и тогда для вычисления производной порядка $\leq m - k$ функции \vec{V} мы можем воспользоваться дифференцированием под знаком интеграла. Теперь мы находимся в условиях применимости следствия теоремы 115, что и доказывает теорему.

З а м е ч а н и я. 1°) Тот факт, что \vec{F} является пространством Банаха, т. е. полно, существен. Так как $\vec{V}(x) = \vec{U}(x)/x_N^k$, то для точек x , не лежащих в гиперплоскости $x_N = 0$, $\vec{V}(x)$ является элементом \vec{F} . Для точек x , находящихся в этой гиперплоскости, интеграл (IV, 11; 62) в неполном пространстве может не иметь смысла. Поэтому предел $\vec{V}(x)$ при x , стремящемся к некоторой точке гиперплоскости $x_N = 0$, существовать не обязан.

2°) Можно было бы надеяться, что функция \vec{V} принадлежит классу $C^{m'}$, где $m' > m - k$. Однако это не верно, и наилучшим является число $m - k$.

Рассмотрим частный случай $N = 1$ ($\mathbb{R}^N = \mathbb{R}$), $\vec{F} = \mathbb{R}$ и функцию U , определенную формулой

$$U(x) = |x|^{m+1/2}. \quad (\text{IV, 11; 63})$$

Она, очевидно, принадлежит классу C^m , но не принадлежит классу C^{m+1} . В начале координат она имеет нуль порядка $m + 1$ и, следовательно, в частности, любого порядка $k \leq m$. Отношение V здесь запишется в виде

$$V(x) = |x|^{m-k+1/2}. \quad (\text{IV, 11; 64})$$

Это отношение принадлежит классу C^{m-k} и не принадлежит классу C^{m-k+1} .

Следствие. Пусть \vec{U} — некоторая функция класса C^m на открытом множестве Ω аффинного конечномерного пространства

Е со значениями в банаховом пространстве \vec{F} . Пусть Σ — гиперповерхность класса C^m , содержащаяся в Ω , а $f(x) = 0$ — нормальное уравнение этой гиперповерхности¹⁾. Если эта гиперповерхность состоит из нулей функции \vec{U} порядка, не меньшего k , то существует функция \vec{V} класса C^{m-k} , определенная единственным образом на Ω со значениями в \vec{F} и такая, что

$$\vec{U}(x) = (f(x))^k \vec{V}(x). \quad (\text{IV}, 11; 65)$$

Доказательство. Здесь, как и при доказательстве теоремы (см. замечание 2°) на стр. 805), достаточно для каждой точки a гиперповерхности дать доказательство, относящееся к некоторой окрестности этой точки. Выберем некоторую систему координат в E . Поскольку $f(x) = 0$ является нормальным уравнением поверхности Σ , то в точке $a \in \Sigma$ хотя бы одна из частных производных f относительно x_1, x_2, \dots, x_N отлична от 0. Предположим, например, что это $\frac{\partial f}{\partial x_N}(a)$.

Рассмотрим теперь функции, определенные равенствами

$$\begin{aligned} u_1 &= x_1, & u_2 &= x_2, \dots, & u_{N-1} &= x_{N-1}, \\ u_N &= f(x_1, x_2, \dots, x_N). \end{aligned} \quad (\text{IV}, 11; 66)$$

Эти функции определяют некоторое отображение Φ множества Ω в \mathbb{R}^N класса C^m . Якобиан этого отображения $\neq 0$ в точке a , поскольку он равен $\frac{\partial f}{\partial x_N}(a)$. Согласно теореме о неявных функциях (следствие 2 теоремы 31 гл. III), существует такая открытая окрестность ω точки a в Ω , что сужение Φ на ω является C^m -диффеоморфизмом ω на некоторое открытое множество ω_1 из \mathbb{R}^n . Обозначим через Φ^{-1} его обратный диффеоморфизм. С помощью Φ можно перенести каждую фигуру из ω на ω_1 .

Пересечение Σ с ω перейдет на гиперповерхность $\Sigma_1 = \Phi(\Sigma)$, являющуюся пересечением гиперплоскости $u_N = 0$ с множеством ω_1 .

Диффеоморфизм Φ преобразует функцию \vec{U} на ω со значениями в \vec{F} в функцию \vec{U}_1 на ω_1 со значениями в \vec{F} , равную $\vec{U}_1 = \vec{U} \circ \Phi^{-1}$.

Легко проверить, что гиперплоскость $u_N = 0$ состоит из нулей порядка k функции \vec{U}_1 ²⁾.

¹⁾ Смотри следствие теоремы 33₂ гл. III.

²⁾ Так как $\vec{U}_1 = \vec{U} \circ \Phi^{-1}$, то производные \vec{U}_1 определяются по теореме о сложных функциях. Вычисление их громоздко, но *индукцией по k* легко доказать, что все они равны нулю до порядка $k - 1$ включительно.

Из теоремы следует, что в ω_1 существует такая функция \vec{V}_1 класса C^{m-k} , определяемая единственным образом, для которой имеет место тождество $\vec{U}_1 = u_N^k \vec{V}_1$. В этом случае \vec{V}_1 является результатом преобразования Φ^{-1} функции \vec{V} на ω , определенной соотношением $\vec{V} = \vec{V}_1 \circ \Phi$. Функция \vec{V} на ω принадлежит классу C^{m-k} , и на ω имеет место тождество $\vec{U} = f^k \vec{V}$, чем и заканчивается доказательство теоремы.

Следствие 2. Пусть Σ — гиперповерхность класса C^∞ открытого множества Ω конечномерного аффинного пространства E с нормальным уравнением $f(x) = 0$. Все другие ее нормальные уравнения $g(x) = 0$ можно получить, полагая $g = fV$, где V — некоторая скалярная функция класса C^∞ на Ω всюду $\neq 0$.

Доказательство. Прежде всего, если $g = fV$, а функция V не имеет нулей и принадлежит классу C^∞ , то функция g также принадлежит классу C^∞ и $g(x) = 0$ определяет гиперповерхность Σ . Кроме того, для точки $a \in \Sigma$

$$\overleftarrow{g'(a)} = \overleftarrow{f'(a)V(a)} + f(a)\overleftarrow{V'(a)} = \overleftarrow{f'(a)V(a)} \neq 0,$$

а, следовательно, $g(x) = 0$ есть нормальное уравнение гиперповерхности Σ .

Обратно, пусть $g(x) = 0$ — нормальное уравнение гиперповерхности Σ . Тогда g принадлежит классу C^∞ и обращается в нуль на Σ , т. е., согласно следствию 1, $g = fV$, где V принадлежит классу C^∞ . Можно также сказать, что $g(x) = 0$ является нормальным уравнением Σ , а f принадлежит классу C^∞ и обращается в нуль на Σ , т. е. $f = gW$, где W принадлежит классу C^∞ . Поэтому $g = fV = gWV$, откуда следует, что $WV = 1$ там, где g не обращается в нуль, т. е. вне Σ . Поскольку W и V непрерывны, то функция WV также непрерывна, а, следовательно, $WV \equiv 1$ всюду. Это означает, что $V \neq 0$ всюду, и следствие доказано.

Замечания. 1°) Интересно отметить, что интегральное исчисление оказывается полезным для изучения таких чисто дифференциальных свойств.

2°) Исходя из следствия 2, можно доказать (но это очень сложно), что каждая замкнутая гиперповерхность Σ класса C^∞ аффинного конечномерного пространства может быть полностью определена одним нормальным уравнением $f(x) = 0$.

Для простоты следующую теорему мы приведем лишь в случае $k = 1$.

Теорема 120. Пусть Ω — открытое множество пространства \mathbb{R}^N и \vec{U} — некоторая функция на Ω класса C^m со значениями в банаховом пространстве \vec{F} . Предположим, что функция \vec{U} обращается в нуль на пересечении Ω с векторным n -мерным подпространством, определяемым уравнением $x_{n+1} = \dots = x_N = 0$. Тогда можно найти (вообще говоря, бесконечным числом способов) такие функции $V_{n+1}, V_{n+2}, \dots, V_N$ класса C^{m-1} на Ω со значениями в \vec{F} , что

$$\vec{U}(x) = x_{n+1}\vec{V}_{n+1}(x) + \dots + x_N\vec{V}_N(x). \quad (\text{IV, 11; 67})$$

Доказательство аналогично доказательству теоремы 119. Формула конечных приращений (формула Тейлора для $k=1=0$) применяется здесь к функции $N-n$ переменных: $(u_{n+1}, u_{n+2}, \dots, u_N) \rightarrow \vec{U}(x_1, x_2, \dots, x_n, u_{n+1}, \dots, u_N)$. При этом получается, что

$$\begin{aligned} U(x_1, x_2, \dots, x_N) &= \\ &= \int_0^1 (U'(x_1, x_2, \dots, x_n, tx_{n+1}, \dots, tx_N) \cdot (0, \dots, 0, x_{n+1}, \dots, x_N)) dt = \\ &= \sum_{i=n+1}^N x_i \int_0^1 \frac{\partial \vec{U}}{\partial x_i}(x_1, x_2, \dots, x_n, tx_{n+1}, \dots, tx_N) dt. \end{aligned} \quad (\text{IV, 11; 68})$$

Поэтому для $i=n+1, n+2, \dots, N$ можно положить

$$\vec{V}_i(x) = \int_0^1 \frac{\partial \vec{U}}{\partial x_i}(x_1, x_2, \dots, x_n, tx_{n+1}, \dots, tx_N) dt, \quad (\text{IV, 11; 69})$$

чём и заканчивается доказательство теоремы.

Замечание. Здесь при $n \leq N-2$ единственного решения не будет. К функции \vec{V}_{N-1} можно добавить функцию x_N , а к функции \vec{V}_N функцию x_{N-1} , и это приведёт лишь к добавлению и вычитанию $x_{N-1}x_N$ из \vec{U} .

Следствие. Пусть V — многообразие класса C^m открытого множества Ω конечномерного аффинного пространства E , определенное системой нормальных уравнений $f_1(x) = 0, f_2(x) = 0, \dots, f_l(x) = 0$. Пусть \vec{U} — функция класса C^m на Ω со

значениями в банаховом пространстве \vec{F} , равная нулю на V . Тогда существует система функций $\vec{V}_1, \vec{V}_2, \dots, \vec{V}_l$ (вообще говоря, их бесконечно много) класса C^{m-1} на Ω со значениями в \vec{F} , такая, что

$$\vec{U} = f_1 \vec{V}_1 + f_2 \vec{V}_2 + \dots + f_l \vec{V}_l. \quad (\text{IV}, 11; 70)$$

Доказательство аналогично доказательству следствия 1 теоремы 119 с некоторыми усложнениями, требующими применения разложения единицы. На нем мы останавливаться не будем.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Абсолютно мажорируемая векторная мера** 488
— сходящееся бесконечное произведение 165
— сходящийся несобственный интеграл Римана 725
— ряд 129, 160
Абстрактное многообразие размерности n класса C^m 334
Аксиома отделимости Хаусдорфа 44
Алгебра 127
σ-алгебра частей множества 505
Антисимметричность 21
Аппроксимирующая последовательность 424, 524, 563
Ассоциативность тензорного произведения 677
Атлас многообразия 326
Атомическая мера 441
Аффинная контингенция 219
— система координат 176
Аффинное многообразие 176
— нормированное пространство 179
— отображение 178
— пространство 174
— касательное в точке к многообразию 338
Аффинно линейная функция 151

Банахова алгебра 127
Банахово пространство 117
Барицентр 184
Безусловно сходящийся ряд 131
Бесконечно дифференцируемое многообразие 319
Бесконечное произведение 162
Биективное отображение 13
Биекция 13
Билинейное отображение 121
Борелевская σ -алгебра 505
Борелевское отображение 517
— подмножество 505
Брахистохрона 369, 384

Вектор касательный в точке к множеству 218
Векторная контингенция 219
— мера 450
Векторное подпространство, касательное к многообразию в точке 220
— поле 222
— пространство, касательное к абстрактному многообразию в точке 338
— — — многообразию в точке 338
Верхний интеграл 520
— — Римана 421, 635, 653
— предел последовательности 79
Верхняя огибающая 25
Вещественная мера Радона 476
Вещественное многообразие 333
Взвешенное геометрическое среднее 360
Внешность множества 48
Внешняя мера 490, 494
Внутренность множества 47
Внутренняя мера 492
Возрастающее отображение 25
Вполне нулевая мера 586
— отрицательная мера 586
— положительная мера 586
— упорядоченное множество 23
Выпуклая функция 204
Выпуклое множество (часть) 43
— — в аффинном пространстве 184

Галилеева система координат 195
Гармоническая функция 415
Геодезическая кривая 389, 398
Гиперплоскость 177
Гиперповерхность 327
Главное значение в смысле Коши 736
Гомеоморфизм 54
Гомеоморфные пространства 55
Градиент функции 221
Граница 48

- Двойная последовательность 61
 Двойной интеграл 657
 Делимость дифференцируемых функций 804
 Дизьюнктные меры 588
 C^m -диффеоморфизм 336
 Дифференциал 213, 217
 Дифференцируемая гиперповерхность 292
 — функция 227
 Дифференцируемое многообразие 218, 319
 Дифференцируемость бесконечного произведения 787
 — неявной функции 298
 — суммы ряда 779
 — функции, представимой интегралом 794
 Длина пути 701
 Достаточные условия экстремума 287
 Дуга 90
- Евклидово аффинное пространство 185
 — векторное пространство 185
 — скалярное произведение 184
 Единичный шар 42
 Единственность тензорного произведения мер 658
 Естественная топология 64
- Зависимые функции 350
 Задача Дирихле 415
 — Плато 414
 Заключение 36
 Закон инерции 194
 Замена переменных 16, 394, 720, 742
 Замкнутая кривая 353
 — часть 45
 Замкнутый интервал 26, 46
 — параллелепипед 46
 — шар 41, 46
 Замыкание 49
- Измельчение разбиения 418
 Измеримое множество 500, 501
 — отображение 510, 537
 Изолированная точка 47
 Изотропные векторы 192
 Изотропный конус 192
 Индуцированная метрика 50
 — топология 59
 Интеграл Лебега от функции с векторными значениями 524
 — от векторной этажной функции по мере μ 520
- — функции, определенной почти всююде 533
 — — — по мере μ 443
 — Римана 426
 — — на прямой 416
 — — ступенчатой функции с компактным носителем 419
 — сходящийся в смысле главного значения Коши 736
 — Френеля 735
 Интегрируемая функция с векторными значениями 524
 μ -интегрируемая функция с векторными значениями 524
 Интегрируемая функция со значениями в пространстве Банаха 423
 Интегрируемость в смысле Лебега 725
 — суммы ряда 778
 — уравнений Эйлера 383
 — функции по нескольким мерам 612
 — — представимой интегралом 793
 Инъективное отображение 13
 Инъекция 13
 Истинное параметрическое представление 321
 Исходное множество отображения 11
- Каноническая инъекция 12
 — мера прямой R 448
 — сюръекция 19
 Кардинальное число 28
 Карта 326
 — множества 326
 — покрывающая точку 326
 Касательная к пути 709
 Касательный вектор к многообразию в точке 344
 Квантор 37
 Класс эквивалентности 18
 Колебание функции в точке 196, 638
 — — на окрестности точки 638
 Компактная часть 68
 Компактное топологическое пространство 69
 Комплексная мера Радона 476
 Комплексное многообразие 333
 Комплексно сопряженная мера Радона 477
 Композиция отображений 15
 Конечное множество отображения 11
 Континуум-гипотеза 35
 Кратный интеграл 677
 — — на R^n 742
 Кривая 327
 — Жордана 753
 — класса C^m 353

- Криволинейная абсцисса 704
 Криволинейный интеграл 705
 Критерий Абеля 732
 — равномерной сходимости 161
 — условной сходимости 142
 Круг 41
- Лемма Хаара 379
 Линейное многообразие, касательное к многообразию в точке 220
 — L -непрерывное отображение 579
 — отображение 110
 Линейно связное пространство 92
 Локальная карта класса C^m 326
 — равномерная сходимость бесконечного произведения 168
 — сходимость последовательности мер Радона по норме 631
 Локально компактное пространство 74
 — линейно связное пространство 95
 — равномерно сходящаяся последовательность функций 153
 — связное пространство 94
 Локальный гомеоморфизм 312
- Мажоранта 23
 Мажорируемая часть 23
 Максимум 23, 25
 Матрица Якоби 211
 Мера Дирака 441
 — — в точке 440
 — допускающая борелевское продолжение 583
 — Лебега 448
 — множества 501
 — плотности ρ по отношению к вещественной мере 594
 — Радона 440
 — на вещественной прямой 682
 — — локально компактном пространстве 447
 — — — равная нулю 465
 — с базой 594
 — μ -сингулярная 589
 — сосредоточения на части пространства 589
 n -мерная площадь n -мерного параметрического многообразия 760, 763
 Метризуемое топологическое пространство 59
 Метрическое пространство 40
 — свойство 58
 Минимум 24
 Миноранта 23
 Минорируемая часть 23
- Многообразие, дифференцируемое m раз 319
 — принадлежащее классу C^m 320
 — размерности n , принадлежащее классу C^m 320
 — m раз непрерывно дифференцируемое 319
 Множество 9
 — линейных отображений 15
 — непрерывных отображений 15
 — нулевой меры 506
 — отображений 15
 Множитель Лагранжа 353
 Монотонная функция 202
 Мощность 27
 — континуума 33
- Натуральная параметризация 704
 Независимые функции 350
 Необходимые условия экстремума 285, 287
 Неограниченно удаляющаяся последовательность точек 619
 Неопределенный интеграл, 694, 706
 — — от меры Радона 683
 — — — функции по мере 684
 Неподвижная точка отображения 107
 Непрерывная функция многих переменных 65
 — — с компактным носителем 463
 Непрерывно дифференцируемая функция 227
 — дифференцируемое многообразие 319
 Непрерывное в бесконечности отображение 618
 — отображение 52
 — слева отображение 195
 — справа отображение 195
 Непрерывность суммы ряда 778
 — функции, представимой интегралом 792
 Неравенство выпуклости 562
 — Гельдера 364, 567
 — Коши — Буняковского — Шварца 43
 — Коши — Шварца 366
 — Минковского 361, 562
 — счетной выпуклости 546, 562, 565
 Несобственный интеграл 724
 — путь 703
 Нечетность 736
 Неявная функция 293
 Нижний предел последовательности 79
 Нижняя огибающая 25
 Норма 41

- билинейного отображения 124
- линейного отображения 113
- Нормальная система уравнений 331
- Нормальное пространство 462
- Нормально сходящийся ряд 129, 160
- Норма меры 449
 - порядка p 359, 561
- Нормированная алгебра 127
- Нормированное векторное пространство 42
- Носитель 417
 - меры 663
 - Радона 466
- Обобщенное евклидово пространство 192
 - эрмитово пространство 192
- Образ карты 326
- меры при отображении 617, 618, 621
- множества при отображении 13
- отображения 113
- параметрического многообразия 352
- Обратное отображение 139
- Общая, или аффинная, система координат 176
 - теорема Лебега 539
- Объединение 10
- Ограниченнная часть 23
- Окрестность точки 46
- Окружность 41
- Ортонормированный базис 190
- Основная теорема измерения объемов 754
- Особое многообразие размерности n класса C^m 352
- Отделимое пространство 58
- Открытое множество 44
 - отображение 311
- Открытый интервал 26
 - параллелепипед 45
 - прямоугольник 64
 - шар 41, 45
- Относительно компактная часть 68
- Относительный максимум 285
 - минимум 285
 - экстремум 393
- Отношение порядка 21
 - эквивалентности 17
- Отображение 11
 - μ -измеримое 510
 - класса C^k 336
 - непрерывное в точке 52
 - μ -собственное 623
- Отрезок 43
- Отрицательная мера Радона 478
 - часть элемента 481
- Паракомпактное пространство 453
- Параллельный перенос меры Радона 632
- Параметрическое многообразие размерности n класса C^m 352
- Параметрические многообразия, C^m -эквивалентные 353
- Первообразная 706
- Пересечение множества частей 10
- Перестановка 13
 - членов ряда 130
- Σ и \int для рядов с положительными членами 547
- Плотная часть 50
- Плотность по отношению к мере Лебега 449
- Площадь n -мерная 756
- Поверхность 327
- Подмножество (или часть) 9
- Под покрытие 67
- Покрытие 67
- Полная вариация последовательности 142
 - функции 686
 - дифференцируемость 253, 266
 - масса 474
 - производная 213
 - решетка 482
- Полное метрическое пространство 101
- Полный дифференциал 213
- Положительная мера Радона 478
 - часть элемента 481
- Полулинейная форма 187
- Полулинейное отображение 187
- Полунорма 564
- Полуоткрытый интервал 26
- Полусходящийся несобственный интеграл Римана 725
- Полупоралинейная форма 187
- Полупорядоченное векторное пространство 480
- Полуполненная прямая 26
- Последовательность 15
 - имеющая ограниченную вариацию 142
 - Коши 101
 - мер Радона, широко сходящаяся к предельной мере 632
 - L -сходящаяся 578
- Постоянное отображение 13
- Правило нахождения относительного максимума или минимума 289
- Правильная функция 197
- Предел последовательности 60
 - функций 777
- Преобразование 13
- Принцип кусочной склейки мер 475

- Продолжение меры 473
 — с векторными значениями 610
 — положительной меры 489
 — равномерно непрерывных отображений 104
 Проекция 12
 Произведение меры на функцию 590, 592
 — множеств 11
 — нормированных векторных пространств 118
 — топологий 64
 — числовых рядов 137
 Производная 199
 — билинейного непрерывного отображения 225
 — матрица 211
 — мера 698
 — отношения 238
 — отображения 198
 — — слева 198
 — — справа 198
 — порядка m 199
 — произведения 236, 268
 Производное отображение 213
 Производный вектор 215
 Прообраз множества при отображении 13
 — функции при замене переменных 16
 Простая сходимость последовательности функций 149
 Пространство n раз дифференцируемых функций 267
 Псевдоизмеримое множество 500, 501
 — подмножество 496
 Пустой носитель меры Радона 466
 Путь 90
- Равномерно дифференцируемая функция** 253
 — непрерывное отображение 86
 — стягивающееся множество 748
 — сходящаяся последовательность 150, 152
 — сходящийся ряд 160
 Разбиение 418
 Раздельно непрерывная функция 67
 Разложение единицы 452
 — подчиненное покрытию 452
 Разложимая функция 463, 654
 Рассеянная мера 443
 Расстояние 40, 82
 Расходящееся бесконечное произведение 163
 Регулярное пространство 59
- Решетка 481
 Рефлексивность 17, 21
- Свойство Больцано — Вейерштрасса**
 76
 — выполняющееся почти всюду 508
 — — μ -почти всюду 508
 — Гейне — Бореля 68
 — Лузина 536
 Связанные точки 93
 Связная компонента 93
 Связное пространство 88
 Сдвиг 150
 Сепарабельное пространство, 50, 510
 Сжатие 107
 Симметричное билинейное отображение 259
 — π -линейное отображение 262
 Симметричность 17
 Скачок функции в точке 196
 Собственное отображение 618
 Совершенное множество 507
 Соотношение Шаля 175, 683
 Сопряженное пространство 115
 Спрямляемый несобственный путь 703
 Среднее значение функции на интервале 438
 — порядка p 359
 Стационарная точка 287
 Строгий относительный максимум 285
 — — минимум 285
 Строго возрастающее отображение 25
 — убывающее отображение 25
 Ступенчатая функция 418
 Сумма безусловно сходящегося ряда 131
 — Коши — Римана 436
 — ряда 128
 Суммирование по блокам безусловно сходящегося ряда 134
 Существенная верхняя грань 574
 Существенно ограниченная функция 574
 Существование тензорного произведения мер 661
 Сфера 41
 Сходимость мер Радона по норме 631
 Сходящаяся двойная последовательность 61
 — последовательность 60
 — — мер Радона 631
 Сходящееся бесконечное произведение 162
 Сходящийся несобственный интеграл 724
 — ряд 128
 Счетное множество 27, 31

- Сюръективное отображение 13
 Сюръекция 13
- Тензорное произведение мер 658
 Теорема Абеля 143, 732
 — Банаха — Штейнгауза 641
 — Бернштейна 27
 — Бэра 641
 — Вейерштрасса о полиномиальной аппроксимации 680
 — Даламбера 84
 — Данфорда — Петтиса 594
 — Егорова 534
 — Лебега — Никодима 594
 — — о сходимости 534
 — об интегрировании по частям 728
 — — ортогональной проекции площадей гиперплоскостей 759
 — о замене переменной 735
 — — неподвижной точке 108
 — — неявной функции 305
 — — перестановке интеграла и линейного непрерывного отображения 430
 — — переходе через таможню 92
 — — плотности 659
 — — постоянном ранге 345
 — — применении билинейного непрерывного отображения 567
 — — — линейного непрерывного отображения 566
 — — сложной функции 230
 — — среднем 433
 — Рисса 106
 — Ролля 200
 — Фату 542
 — Фишера — Рисса 573
 — Фубини — Лебега 664
 — — Фату 670
- Тождественное отображение 12
 Топологическая группа 66
 Топологическое векторное пространство 66
 — множество 320
 — подпространство 59
 — произведение 63
 — пространство 58
 — свойство 58
 Топология пополненной прямой 60
 Точка прикосновения 49
 — разрыва первого рода 196
 — сгущения 75
 Точная верхняя грань 24
 — нижняя грань 24
 Транзитивность 17, 21
 — образов мер 624
- Трансфинитное число 28
 Трансцендентное число 33
- Убывающее отображение 25
 Узкая сходимость последовательности мер конечной нормы 652
 — — — Радона 654
 Узко сходящаяся последовательность мер 653
 Универсальный атлас 326
 Уравнение колебания струны 239
 — Эйлера 378
 Уравнения Гамильтона 408
 Условие 36
 — Гельдера 87
 — Липшица 87
 Условия трансверсальности 400
 Условно сходящееся бесконечное произведение 165
 — сходящийся ряд 129
 Условный максимум 353
 — минимум 353
- Факторгруппа 19
 Фактормножество 19
 Факторпространство 20
 Формула интегрирования по частям 716
 — конечных приращений 200, 247
 — Лейбница 268
 — Тейлора 201, 272, 276, 279, 292, 714, 715
 Фундаментальная последовательность 100
 — система окрестностей 47
 Функции, μ -эквивалентные в смысле Лебега 572
 Функциональное пространство 146
 Функция 11
 — Гамильтона 407
 — длины 702
 — имеющая интегрируемую p -ю степень 563
 — — локально ограниченную вариацию 694
 — интегрируемая по Риману 423
 — μ -интегрируемая 610
 — — по Риману 634
 — Лагранжа 407
 — полуунпредымянная сверху 85
 — — снизу 85
 — представимая интегралом 777, 792
 — — рядом 777
 — принадлежащая классу C^m 199
 — m раз непрерывно дифференцируемая 199, 266

- Римана 168
- с ограниченной вариацией 686
- μ -существенно ограниченная 574
- удовлетворяющая условию Гельдера 87
- — — Липшица 87
- характеристическая 30
- Хевисайда 684

- Частная дифференцируемость 253, 266
- производная вдоль вектора 209
- — отображения 224
- сумма с индексом n 128
- теорема Лебега 538
- Часть (подмножество) множества 9
- Четность 736

Шар 41

- Широкая сходимость последовательности мер к мере Дирака 646

- — — Радона 630, 632, 654
- — тензорных произведений 679

- Эквивалентные метрики 56
- нормы 56
- пути 702
- функции 572
- Экстремаль 378
- Экстремум 285
- Эрмитова положительно определенная матрица 192
- Эрмитово аффинное пространство 187
- векторное пространство 187
- скалярное произведение 187
- Этажная функция 513
- μ -этажная функция 513

Ядро отображения 113

- Якобиан 212

ОГЛАВЛЕНИЕ

Предисловие редактора перевода	5
Предисловие	7
Г л а в а I. Т е о р и я м н о ж е с т в	9
§ 1. Множества. Элементарные операции	9
Части множества	9
Отношение включения. Дополнение	10
Объединение. Пересечение	10
Произведение множеств	11
§ 2. Отображения. Функции	11
Примеры отображений	12
Инъекции. Сюръекции. Биекции	13
Образ и прообраз подмножества	13
Множество отображений. Семейства. Последовательности	15
Композиция отображений	15
Замена переменных и замена функций	16
§ 3. Отношения эквивалентности. Фактормножество	17
Классы эквивалентности. Разбиения	18
Фактормножество	19
Факторгруппа по инвариантной подгруппе	19
Факторпространство векторного пространства по векторному подпространству	20
§ 4. Отношения порядка	21
Примеры отношений порядка	22
Мажорируемые части. Мажоранты. Максимум. Точная верхняя грань	23
Возрастающие функции	25
Положительная прямая	26
§ 5. Мощности. Счетные множества	27
Мощности. Кардинальные числа	27
Счетные множества	31
Мощность континуума	33
Трансцендентные числа	33
Континуум-гипотеза	35
§ 6. Некоторые основные понятия логики	36
Г л а в а II. Т о п о л о г и я	40
§ 1. Метрические пространства. Элементарные примеры	40
Сфера. Шары	41
Нормированные векторные пространства	41
§ 2. Открытые и замкнутые части. Окрестности. Внутренность. Граница. Замыкание. Плотные подмножества	44
Открытые части	44
Замкнутые части	45
Окрестности	46

Внутренность	47
Внешность	48
Граница	48
Замыкание	49
Плотные подмножества	50
Подпространства. Индуцированная метрика	50
§ 3. Непрерывные функции. Гомеоморфизмы	52
Гомеоморфизмы	54
§ 4. Метрические пространства и топологические пространства	55
Топология пополненной прямой $\bar{\mathbb{R}}$	60
§ 5. Последовательности. Пределы. Сходимости	60
§ 6. Топологическое произведение	63
Сходящиеся последовательности в произведении	64
Непрерывные функции многих переменных	65
Топологические группы. Топологические векторные пространства	66
Раздельная непрерывность функций двух переменных	67
§ 7. Компактные пространства. Элементарные свойства	67
Локально компактные пространства	74
Точка сгущения последовательности	75
Верхний и нижний пределы вещественной последовательности	79
§ 8. Свойства непрерывных функций на компактных пространствах	79
Равномерная непрерывность	86
§ 9. Связные пространства	88
Линейно связные пространства	90
§ 10. Дополнение по общей топологии связных пространств	92
Некоторые применения понятия связности. Критерии негомеоморфности	97
Существование и непрерывность обратной функции для строго монотонной непрерывной функции	98
Применение: метрики, определяющие топологию в $\bar{\mathbb{R}}$	99
§ 11. Полные метрические пространства	100
Продолжение равномерно непрерывных отображений	104
Частные свойства конечномерных топологических векторных пространств	106
§ 12. Теорема о неподвижной точке	107
§ 13. Элементарная теория нормированных векторных пространств и пространств Банаха	110
Ядро и образ непрерывного линейного отображения	113
Произведения нормированных векторных пространств	118
Билинейные непрерывные отображения произведения нормированных векторных пространств в нормированное векторное пространство	121
Мультилинейные непрерывные отображения	126
Алгебры. Нормированные алгебры	127
§ 14. Ряды в нормированных векторных пространствах	128
Перестановка членов ряда	130
Суммирование по блокам безусловно сходящегося ряда	134
Действие линейного непрерывного отображения на ряд	136
Произведение двух числовых рядов. Применение билинейного непрерывного отображения к двум рядам	137
Обратимые отображения в банаховых пространствах	139
Критерий условной сходимости	142
§ 15. Наиболее употребительные примеры функциональных пространств. Сходимость простая и равномерная	146
Функциональные пространства	146
Простая сходимость последовательности функций	149
Равномерная сходимость последовательности функций	150

Другие применения выражения «равномерная сходимость»	152
Пространства, порожденные структурами пространств E и F	154
Непрерывность локально равномерного предела последовательности непрерывных функций	155
Некоторые контрпримеры	157
Ряды функций со значениями в нормированном векторном пространстве	159
§ 16. Бесконечные произведения вещественных или комплексных чисел и функций	162
Бесконечные произведения и логарифмические ряды	164
Бесконечные произведения вещественных или комплексных функций	167
Применение к функции ζ Римана	168
Глава III. Дифференциальное исчисление	174
§ 1. Аффинные пространства	174
Аффинные многообразия	176
Линейные отображения. Аффинные отображения	178
Аффинные нормированные пространства	179
Выпуклые множества в аффинных пространствах	183
Евклидовы векторные и евклидовы аффинные пространства	184
Эрмитовы векторные и эрмитовы аффинные пространства	187
Изоморфизм (или полуизоморфизм) конечномерного евклидова (или эрмитова) пространства и его сопряженного пространства	189
Ортонормированные базисы	190
Обобщенные евклидовы или эрмитовы пространства	192
§ 2. Вещественные функции вещественной переменной. Непрерывность справа и слева	195
Разрывы первого рода. Правильные функции	196
Производная вещественной функции вещественной переменной	198
Монотонные функции	202
Дифференцируемые функции и теоремы о промежуточных значениях	204
Выпуклые функции	204
§ 3. Производная отображения одного аффинного пространства в другое. Производный вектор функции скалярной переменной	208
Общий случай. Частная производная вдоль вектора	209
Матрица Якоби. Якобиан	211
Недостатки понятия производной вдоль вектора	212
Полная производная, или производное отображение	213
Понятие дифференциала	217
Геометрическая интерпретация производного отображения: дифференцируемое многообразие и линейное касательное многообразие	218
Градиент вещественной функции в евклидовом пространстве	221
Случай, когда F является произведением аффинных пространств	223
Случай, когда E является произведением аффинных пространств. Частные производные отображения	224
Производная билинейного непрерывного отображения	225
Дифференцируемые функции. Непрерывно дифференцируемые функции	227
Примеры непрерывно дифференцируемых функций	228
Пространства дифференцируемых функций	229
§ 4. Теорема о сложной функции	230
Примеры вычисления обычных производных	235
§ 5. Формула конечных приращений	247
Полная дифференцируемость и частная дифференцируемость	253

§ 6. Производные высших порядков	257
Последовательные производные	261
Случай произведения пространств. Полная и частная дифференцируемости	266
Пространства n раз дифференцируемых функций	267
Производная произведения (формула Лейбница)	268
§ 7. Формула Тейлора. Максимум и минимум	272
Применение формулы Тейлора для вычисления производных	276
Формула Тейлора относительно некоторой системы координат	279
Применение к изучению максимумов и минимумов. Определения	285
Необходимые условия экстремума	285
Нахождение необходимых и достаточных условий экстремума функции	287
Частный случай вещественной функции f двух вещественных переменных x, y	290
Применение формулы Тейлора к изучению расположения гиперповерхности по отношению к касательной гиперплоскости	292
§ 8. Теорема о неявной функции. Постановка задачи	293
Существование неявной функции	294
Дифференцируемость неявной функции	298
Дифференцируемость функции $u \rightarrow u^{-1}$ на $\mathcal{L}(\vec{F}, \vec{G})$	300
Частный случай, когда $E = F = G = \mathbb{K}$ — скалярное поле	306
Случай, когда E, F, G конечномерны	308
Обратная функция как неявная функция	309
Вычисление производных высших порядков неявной функции	314
Техника замены переменных и замены функций	318
§ 9. Дифференцируемые многообразия	319
Определение многообразия при помощи его параметрического представления	321
Определение многообразия с помощью неявных уравнений	331
Вещественные и комплексные многообразия	333
Абстрактные многообразия	334
Векторное пространство, касательное в точке к многообразию аффинного пространства E размерности N	338
Векторное пространство, касательное к абстрактному многообразию в точке	343
Теорема о постоянном ранге	345
Зависимые и независимые функции	350
Особые, или параметрические, многообразия	352
§ 10. Условные максимумы и минимумы	353
Практический способ вычисления условного максимума или минимума	356
Применение теории условных максимумов. Неравенства Гельдера и Минковского	358
§ 11. Вариационное исчисление	369
Постановка задачи	369
Дифференцируемость J	372
Необходимые условия экстремума	378
Лемма Хаара	379
Простые случаи интегрируемости уравнений Эйлера	383
Уравнение геодезических на поверхности	389
Относительный экстремум	393
Замена переменных	394
Приложение к задаче о геодезических	396
Переменные концы. Условие трансверсальности	400
Применение к геодезическим кривым	405

Канонические уравнения Гамильтона	406
Применения к механике	409
Вариационное исчисление для кратных интегралов	410
Г л а в а IV. Интегральное исчисление	416
§ 1. Интеграл Римана на прямой	416
Ступенчатые функции	418
Верхний интеграл Римана от ограниченной функции $f > 0$ с компактным носителем	421
Интегрируемые функции со значениями в пространстве Банаха	423
Интеграл от интегрируемой функции	425
Примеры интегрируемых по Риману функций	433
Вычисление интеграла функции с помощью сумм Коши — Римана	435
Среднее значение функции на интервале	438
§ 2. Меры Радона на локально компактном пространстве	439
Мера Радона на компактном пространстве	439
Примеры мер Радона	440
Меры на локально компактном пространстве	446
Примеры мер Радона	448
Применения к механике и физике	450
Векторные меры	450
Разложение единицы	452
Носитель меры Радона	465
Продолжение меры на непрерывные функции φ с некомпактным носителем	473
Принцип кусочной склейки мер	475
Комплексные и вещественные меры	476
Вещественные положительные меры	478
Решетки	481
§ 3. Продолжение положительной меры. Теория Лебега	489
Внешние меры открытых множеств	490
Внутренняя мера компакта	492
Измеримые множества. Мера множеств	494
Множества нулевой меры	506
Свойства, выполняющиеся почти всюду	508
μ -измеримые функции со значениями в метризуемом сепарабельном пространстве	510
μ -этажные функции	513
Борелевские функции	517
Интеграл от векторной этажной функции	520
Верхний интеграл от вещественной неотрицательной функции	520
Интегрируемость функций с векторными значениями	524
Интеграл Лебега от функции с векторными значениями	524
Интегрируемость и интегралы от функций, определенных почти всюду	533
§ 4. Теорема Лебега о сходимости. Пространство L^1	534
Примеры применений теоремы Лебега	540
Характеристика интегрируемых функций. Интегрируемость и измеримость	551
Теория интегрирования, основанная на свойствах непрерывных и полуунепрерывных снизу функций	555
Пространства $\mathcal{L}^p(X, \mu; F)$	561
Пространства $\mathcal{L}^p(X, \mu; \vec{F})$. Теорема Фишера — Рисса	572
Пространства $\mathcal{L}^\infty(\vec{F})$ и $L^\infty(\vec{F})$	574
Продолжение мер, не обладающих свойством неотрицательности	578

§ 5. Умножение меры на функцию	590
Произведение векторной меры на непрерывную скалярную функцию	590
Элементарные свойства	591
Случай когда μ — вещественная мера ≥ 0	591
Мера с базой μ . Мера с базой ≥ 0	594
Применение к продолжению меры с векторными значениями	610
Применение к интегрируемости функций по нескольким мерам	612
Сопряженность пространств L^p и $L^{p'}$	614
§ 6. Образ меры при отображении	617
Случай, когда H является гомеоморфизмом X на Y	626
Обобщение теоремы 59 на случай, когда μ не ≥ 0	627
Различные примеры образов мер	628
§ 7. Широкая сходимость мер Радона	630
Сходимость по норме. Локальная сходимость по норме	630
Широкая сходимость	632
Функции, μ -интегрируемые по Риману	634
Широкая сходимость и равномерная сходимость	641
Компактные подмножества пространства $C_k(X)$	645
Широкая сходимость последовательности мер к мере Дирака	646
Узкая сходимость последовательности мер конечной нормы	652
Сходимость широкая и сходимость узкая	654
§ 8. Тензорное произведение мер. Кратные интегралы	657
Постановка задачи	657
Существование и единственность тензорного произведения	658
Примеры тензорных произведений	662
Элементарные свойства	663
Носитель меры $\mu \otimes \nu$	663
Вычисление двойного интеграла путем двух последовательных простых интегрирований	664
Случай, когда интегрируемая функция является произведением функции от x и функции от y	673
Окончание доказательства прямого утверждения теоремы	675
Обобщение на произвольные кратные интегралы	677
Широкая сходимость тензорных произведений	679
§ 9. Частные свойства мер Радона на вещественной прямой R	682
Введение символа $\int\limits_a^b d\mu$	682
Неопределенные интегралы	683
Функции с ограниченной вариацией на прямой	686
Функция, удовлетворяющая условию Липшица на ограниченном интервале R_1 прямой R , имеет ограниченную вариацию	687
Функции ограниченной вариации и неопределенные интегралы	694
Длина пути в метрическом пространстве	701
Неопределенный интеграл и первообразная	706
Последовательные первообразные непрерывной функции на прямой	711
Формула интегрирования по частям	716
Замена переменных при вычислении простых интегралов	720
Несобственные интегралы на прямой	724
Примеры применения критерия Абеля	732
Главное значение в смысле Коши	736
§ 10. Кратные интегралы на R^n . Длины, площади и объемы в конечномерном аффинном евклидовом пространстве. Замена переменных в кратных интегралах на R^n	742

Измерение объемов в аффинных евклидовых конечномерных про-	
странствах	754
Измерение длин в аффинном евклидовом пространстве	756
Измерение n -мерных площадей в линейном многообразии раз-	
мерности n аффинного евклидова конечномерного пространства	756
n -мерная площадь n -мерного параметрического многообразия	760
Вычисление объемов с помощью поверхностных интегралов	770
§ 11. Функции, представимые рядами или интегралами	777
Функции, представимые рядами	777
Непрерывность суммы ряда	778
Интегрируемость суммы ряда относительно некоторой меры ≥ 0	778
Дифференцируемость суммы ряда	779
Дифференцируемость бесконечного произведения	787
Функции, представимые интегралами	792
Непрерывность функции, представимой интегралом	792
Интегрируемость функции, представимой интегралом	793
Дифференцируемость функции, представимой интегралом	794
Случай несобственных сходящихся интегралов	799
Применение к делимости дифференцируемых функций	804
Предметный указатель	811