

Math textbook

相轉移 P (@phasetr)¹

2014-04-04

¹phasetr@gmail.com

Contents

I	Aim of this book	10
II	Mathematical/Physical English	12
1	Common phrases	14
2	Mathematical English	16
3	Physical English	18
III	Set theory	20
4	Introduction to cardinal number	21
IV	Linear Algebra	22
5	Talk: Linear algebra and Calculus: Introduction to university mathematics	23
5.1	Introduction	23
5.2	What is linear algebra?	23
5.2.1	Abstract definition of vectors	23
5.2.2	Examples of linear spaces	24
5.2.3	Examples of finite, but higher dimensional spaces than three	25
5.2.4	Function spaces: examples of infinite dimensional dimensional spaces	25
5.2.5	Linear maps, functionals	25
5.2.6	Eigenvalues, eigenvectors	26
5.2.7	Mathematical application of linear algebra	27
5.2.8	Physical application	27
5.2.9	Linear algebra and statistics	27
5.3	Integration	28
5.3.1	Let's define inner products!	28
	Inner products in higher dimensional spaces	28
5.3.2	Axiom for inner products	30
5.3.3	Examples of inner products	30
5.3.4	Physics for the space $L^2(\Omega)$	31
5.4	Differentiation	31

5.4.1	Taylor expansion	31
5.4.2	Unitary representation of a group	32
	Fourier expansion	32
5.4.3	Analitical mechanics and quantum mechanics	33
5.4.4	Physics and representation theory	33
5.5	Variational problem: mixture of differentiation and integration	33
5.5.1	Examples of variational problems	33
	shortest time problem	33
5.5.2	Mathematical formulatrion: functionals, linear operators	33
6	あとで大幅 rewrite: 学部 3 年のときに書いた「本」	35
6.1	はじめに	35
6.1.1	数学「を」学ぶ	35
6.1.2	数学「で」学ぶ	36
6.1.3	一年で物理の講義がほとんど無い理由	36
6.1.4	物理で使う数学	38
6.1.5	数学記号の記法	39
6.1.6	物理と近似: まだ書きかけ	40
6.2	調和振動子と線型代数	41
6.2.1	導入	41
6.2.2	記法の確認と線型写像	43
6.2.3	線型空間の導入	47
	物理法則とベクトル	47
	抽象ベクトル	48
	回転の表現	50
	Galilei の相対性原理	51
	抽象ベクトルの微分	54
	定理の証明	55
6.2.4	行列式	57
6.2.5	内積と外積	64
6.2.6	線型独立と基底, 次元	66
6.2.7	固有値, 固有ベクトル	67
6.2.8	対称行列の対角化とスペクトル分解	70
6.2.9	物理への適用	72
6.2.10	無限次元の線型代数学	73
	抽象ベクトル 2	73
	線型空間	73
	線型独立と次元, 基底 2	76
	内積	77
	直交多項式と基底	79
	微分方程式と Fourier 解析ことはじめ	80
7	Application to Google's pagerank	82
8	量子力学を理解するために斎藤正彦『線型代数入門』[184] を読む人が注 目すべき点	83

V	Calculus	85
9	Is 0 area for a point, line segment, line?	86
10	あとで大幅 rewrite: ポテンシャルの振る舞いと微分積分	87
10.1	導入	87
10.2	ε - δ 論法と収束	88
10.2.1	ε - δ 論法	88
10.2.2	Cauchy 列と実数の性質	90
10.2.3	数列, 級数の収束	92
10.2.4	開集合, 閉集合, 近傍	93
10.2.5	コンパクト性	94
10.2.6	連続と一様連続	94
10.2.7	連結と弧状連結	94
10.2.8	完備な空間	94
10.3	1 変数の微分と Taylor 展開	94
10.4	初等超越関数の Taylor 展開	97
10.5	Euler の公式	98
10.6	多変数の微分と Taylor 展開	99
10.7	微分公式	103
10.8	多変数の Taylor の定理と微分	105
10.8.1	Taylor 展開の応用-波動方程式の導出	107
10.9	最大最小, 極値問題	109
10.10	物理への適用	113
10.11	微分と積分の掛け橋-微分方程式	114
10.12	運動方程式と常微分方程式	115
10.12.1	常微分方程式	115
10.12.2	初期値問題の解の存在と一意性	116
	準備	116
	微分方程式の解の存在と一意性の証明 1	117
	微分方程式の解の存在と一意性の証明 2 の準備-局所 Lip- schitz 条件	121
	解の延長	124
	解の初期値および解に対する連続依存性	125
10.12.3	常微分方程式の解法の為の一般論	127
10.131	階線型方程式の解法	131
10.142	階線型方程式の解法	131
11	Maxwell 方程式とベクトル解析	135
11.1	導入	135
11.2	微分作用素ナブラ	135
11.3	諸公式の導出とテンソル	135
11.4	Green の定理	135
11.5	Stokes の定理	135
11.6	Gauss の定理	135
11.7	微分形式とベクトル解析	135

VI	Complex Analysis	136
12	Talk: Complex Analysis in One variable	137
12.1	Definition: complex differentiability	137
12.2	Cauchy-Riemann equation	137
12.3	Complex line integral	138
12.3.1	Appendix: Green's theorem	139
12.4	Morela's theorem	140
12.5	Cauchy's integral representation formula and Taylor series expansion	141
12.6	Simple application of fundamental theorems: convergence of holomorphic functions	142
12.7	Zero points for holomorphic functions and theorem of identity . .	143
12.8	Theorem of residues	143
12.8.1	Laurent expansion	143
12.8.2	Theorem of residue	145
12.8.3	Argument principle	145
12.9	Convergence theorems	146
12.10	Representation of holomorphic functions	147
12.11	Riemann's mapping theorem	148
12.12	Miscellaneous results	149
12.13	Appendix	149
12.13.1	Vector Analysis	149
12.13.2	Analytic continuation	149
13	Talk: Complex analysis for Ising model	150
14	Several complex variables	151
VII	Functional Analysis	152
15	Introduction to Lebesgue integral	153
15.1	Why Lebesgue integral?	153
15.1.1	Area	153
15.1.2	Function space: approximation of functions	153
15.1.3	Distance by supremum norm	154
15.1.4	Comparison of these two distances	154
15.1.5	Limiting procedure	155
15.2	Definition and Properties of measure	155
15.2.1	Definition	155
15.2.2	Why countable?	157
15.2.3	Definition of Lebesgue measure	157
15.2.4	Null sets	158
15.2.5	Almost everywhere	158
15.2.6	Examples of other measures or measure spaces	159
15.3	Properties of measure	159
15.4	Product measure and extension theorem	160
15.5	Measurable functions	161
15.6	Definition of Lebesgue integral and its properties	163

15.7 Famous theorems for exchanging orders of limits	166
15.8 Fubini's theorem	168
15.9 Function spaces L^p	168
15.10 Riesz-Markov-Kakutani's theorem	168
15.11 Fourier Analysis	168
15.11.1 Hilbert space	168
15.11.2 Fourier series	168
15.11.3 Fourier transform	168
15.11.4 Functions of positive type: Bochner's theorem	168
15.12 Miscellaneous results	168
16 Introduction to functional analysis	169
17 Introduction to operator theory	170
17.1 Functional calculus	170
17.2 Normed field	170
18 Introduction to Hilbert Space Theory And Operator Theory	171
18.1 Introduction	171
18.1.1 概要	171
18.1.2 線型代数からの代数的展開	171
18.1.3 標準的なコースで出てくる諸概念の重要性	171
18.1.4 線型代数と物理や工学への応用	172
18.1.5 線型代数と量子力学	172
18.1.6 線型代数と Hilbert 空間	172
18.1.7 微分方程式	172
18.1.8 微分方程式を解くときに出てくる直交多項式	173
Legendre 多項式, 球 Bessel 多項式	173
Fourier	173
18.1.9 変分	173
18.1.10 作用素のスペクトル	174
18.1.11 Taylor 展開と作用素論	174
18.1.12 2 回目の概論	175
18.1.13 3-4 回目の概論	175
18.1.14 参考文献の紹介	175
19 Fourier Analysis	176
20 重ね合わせの原理と Fourier 解析	177
20.1 Fourier 級数-関数の大域近似	177
20.2 Fourier 変換	177
20.3 Dirac の δ 関数	177
20.4 波動方程式の解	177
20.5 Green 関数	177
20.6 スペクトル分解	177
21 Talk: Inverse analysis for heat equation—mathematics, physics, engineering	178

22 Talk: Analysis for the Thomas-Fermi functional after Lieb-Loss	179
[130]	
22.1 Introduction	179
22.1.1 Comparison for the two problems	179
22.1.2 Variational problem and its mathematical background . .	179
22.1.3 Schrödinger equation	181
22.1.4 Connection to other mathematical branches	182
Mathematics for quantum mechanics	182
Geometric variational problem	184
Nonlinear analysis	184
Index theorem	184
Stochastic optimization	185
Other optimization problem	185
Singular analysis	185
22.2 Mathematical preliminaries	185
22.2.1 The important difference in integration and differentiation	185
22.2.2 Important spaces	186
L^p spaces	186
$L^p + L^q$ spaces	186
the L^1_{loc} space	187
Sobolev spaces H^k	187
22.2.3 Convolution	187
Definition	187
Convolutions of functions in dual L^p spaces are continuous	187
22.2.4 Various inequalities	187
Hölder's inequality	187
Young's inequality	188
Sobolev's inequality	188
22.2.5 Topologies	189
Strong topology and weak topology in Hilbert spaces . . .	189
Example of weakly convergent sequence	189
Why we use the pointwise convergence instead of weak convergence?	189
22.2.6 Theorems in Lebesgue integral and functional analysis . .	190
Dominated convergence	190
Differentiability of norms, Gateaux derivative	190
Uniform boundedness principle	190
Banach-Alaoglu theorem	190
Derivative of the absolute value	190
Convexity inequality for gradients	191
Fundamental theorem of variational method, du Bois-Reymond's lemma	191
22.2.7 Potential theory	191
Positivity properties of the Coulomb energy	191
22.3 Thomas-Fermi theory	192
22.3.1 Thomas-Fermi problem	192
Outline	192
Existence of an unconstrained Thomas-Fermi minimizer .	193
Thomas-Fermi equation	197
The Thomas-Fermi minimizer	199

23 Talk: Analysis for the Friedrichs model: perturbation theory for embedded eigenvalues	202
24 Mathematical foundations for quantum mechanics	203
25 Mathematical foundations for quantum statistical mechanics	204
26 Mathematical problems in quantum mechanics, quantum field theory, and quantum statistical mechanics	205
27 Fundamental problems of peruturbation theory in quantum mechanics as an approximation	207
28 Direct integral and its physics	208
 VIII Probability and Statistics	 210
29 Talk: Law of large number and central limit theorem in coin tossing	211
29.1 参考文献	211
29.2 数学的準備	211
29.2.1 加法定理・倍角の公式	211
29.2.2 指数関数の微分積分	211
29.2.3 Rademacher 関数	212
29.2.4 Rademacher 関数の積分公式	212
29.2.5 Rademacher 関数の積分公式 2	212
29.2.6 Lebesgue 積分の定理	213
29.3 導入	213
29.4 硬貨投げの数理と独立性	215
29.4.1 Viète の公式	215
29.4.2 Rademacher 関数	216
29.4.3 どういうことなの?	217
29.4.4 2 進展開の確率論ことはじめ	218
29.4.5 独立性ことはじめ	219
29.5 大数の法則	221
29.5.1 大数の弱法則	221
29.5.2 大数の強法則	222
29.6 中心極限定理	224
29.6.1 正規分布	224
29.6.2 中心極限定理の正確な言明: Rademacher 関数版	224
29.6.3 Markov の方法とその厳密化	224
 30 Talk: Gaussian superprocess and its application to Quantum Field Theory: Sasakure Seminar	 226
30.1 Introduction	226
30.1.1 Application to quantum field theory [29, 28]	227
30.1.2 Application to geometry [29] chap. 6	227
30.1.3 Application to analytic number theory [23]	228
30.2 Gaussian superprocess and its construction	228
30.2.1 Basics for probability	228

30.2.2	Gaussian superprocesses	229
	Fundamental structures	230
	Uniqueness	230
	Existence of a Gaussian superprocess for any separable H	231
30.3	Concrete realization of the set Q	232
30.4	Application to quantum field theory and relation to operator al- gebraic formulation	233
30.5	Appendix	233
30.5.1	Stone-Weierstrass theorem	233
30.5.2	Minlos' theorem	234
31	Talk: Lecture on ergodic theory	235
32	Statistics for idol management	236
IX	Algebra	237
33	対称性と群論	238
33.1	運動方程式を不変にする変換	238
33.2	エネルギーを不変にする変換	238
33.3	群	238
33.4	Lorentz 群と Maxwell 方程式	238
33.5	解析力学小論	238
X	Geometry	239
XI	Interesting counter examples	240
34	DVD: Playing with interesting counter examples	241
XII	Applied mathematics	242
35	Digital signal processing and algebraic geometry	243
36	Peano curve and its application to engineering	244
XIII	Quantum mechanics and Quantum field theory	245
XIV	Statistical Mechanics and condensed matter physics	246
37	LASER: song by electrones	247
38	Analysis for the Hubbard model	248

XV	Physical misc	249
39	Similarity for hydrodynamics and general relativity	250
XVI	Reference	251

Part I

Aim of this book

I would like to write this math textbook by **social coding** Github. It will be written in English and Japanese. I will also write numerical computation, animation, or simulation codes by some computer languages.

At first, the content is mainly from manuscripts of my seminars for students in department of mathematics or physics.

I start programming. In programmers' community, OSS softwares or social coding using Github is very popular. I want to introduce this to mathematics and physics, especially writing textbooks!

I want to gather our knowledge. There are many mathematics books written by physicists or engineers, and physics books written by engineers or mathematicians. Each is written in their own interest.

I hear that a mathematician (teacher) said, "professors of the department of engineering tell me I hope your lecture and the content will be engineering students friendly. But I do not understand their interest. I want to write books I can understand!"

OK, I will do it. I studied physics in my bachelor days and studied mathematics in my master days. So I will be "human sacrifice" for this difficult, but very interesting and exciting project.

There are many many good books, but they are systematic and "boring" sometimes. I want to fill the gaps: I want to write applicative contents in physics, engineering or mathematics itself, and connections among many branches.

Part II

**Mathematical/Physical
English**

数学や物理に関する英語表現や具体的な例文を少しずつ足していく。基本的な表現については体系的なまとめもほしいので, そういうのもやっていこう。

Chapter 1

Common phrases

- 以下では \mathcal{A} が明らかな場合, \mathbf{Ch} の代わりに単に $\mathbf{Ch}(\mathcal{A})$ と書く. [162]

In the sequel we will merely write \mathbf{Ch} for $\mathbf{Ch}(\mathcal{A})$ when \mathcal{A} is understood.

- 多くの初等的な量子力学の教科書を参照してほしい. これらのいくつかは数学者にも適した形で内容が説明されている. [131]

They can refer to a huge number of elementary quantum mechanics texts, some of which present the subject in a way that is congenial to mathematicians.

- 簡単にいうと.

Simply speaking.

- 記法を決める.

We fix notation.

- 教員は興味や時間の制約に応じて話題を取捨選択してもいい. [162]

The teacher may pick and choose topics according to interest and time constraints.

- 詳しくは A を参照すること.

See A for details.

- ここに書かれた内容は多かれ少なかれ自己充足的な複素代数幾何のコースを元になっている. [174]

The material presented here consists of a more or less self-contained advanced course in complex algebraic geometry.

- この章は以後の章で繰り返し使われる結果も含んでいる. [131]

The chapter also contains several results that will be used repeatedly in the chapters to follow.

- この節での主な論点は鎖複体がアーベル圏をなすことだ. [162]

The main point of this section will be that chain complexes form an abelian category.

- このノートは 1992 年秋の Leiden 大学での講義に基いている. [174]

These notes are based on courses given in the fall of 1992 at the University of Leiden.

- この本は量子力学のいろいろな分野の研究者に向けた内容になっている. [131]

The book is directed towards researchers on various aspects of quantum mechanics.

- この理論の目的は n を無限大にする極限をとったときの $T^n(x)$ の振る舞いを調べることにある. [185]

The aim of the theory is to describe the behavior of $T^n(x)$ as $n \rightarrow \infty$.

- これは大学院の 2-3 年生に向けた本だ. [162]

This book is aimed at a second- or third-year graduate student.

- これらの点は次の単純な問題を考えるとわかる. [162]

The following simple question illustrates these points.

- その間もホモロジー代数は発展を続けた. [162]

In the meantime, homological algebra continued to evolve.

- 同様にして.

In a similar vain.

- 読者は他のいくつかの話題についての紹介に出会おうだろう. [162]

The reader will find introductions to several other subjects.

- トポロジー, 正則局所環や半単純 Lie 環との歴史的な関係も説明する. [162]

The historical connections with topology, regular local rings, and semisimple Lie algebras are also described.

- はじめにアーベル圏とは何だったか思い出す必要がある. [162]

First we need to recall what an abelian category is.

- まとめると.

In summary.

- 最も単純に言えば, 力学系とは集合 X 上に定義された関数 T のことだ. [185]

At its simplest form, a dynamical system is a function T defined on a set X .

- 歴史的な経緯は追わない.

We will not necessarily follow the historical route.

Chapter 2

Mathematical English

- エルゴード理論の威力と有用性の多くは, 上で議論した抽象的な設定に関する次の確率的な解釈によっている. [185]

Much of the power and usefulness of ergodic theory is due to the following probabilistic interpretation of the abstract set up discussed above.

- 核がベクトル束になることはすぐわかる.

One can easily see that the kernel forms a vector bundle.

- 強混合性は測度論的同型の不変量であることはすぐにわかる.

It is easy to see that strong mixing is an invariant of measure theoretic isomorphism.

- 群 G は空間 S に作用する.

A group G acts on a space S .

- “どのようにして決定論的な系がランダムに振舞い得るのか” という問題の現代的な取り扱いはこのアイデアによっている. [185]

The modern treatment of the question “how come a deterministic system can behave randomly” is based on this idea.

- このとき, 写像 f は次の式で表される鎖写像になる. [162]

Then a map f is the chain map given by the formula.

- これが彼の結果を単純化しかつ改良する興味深い新たな不等式の導出に繋がった. [131]

This led to the invention of interesting new inequalities to simplify and improve his result.

- これが定理を証明するために必要な最小限の設定だ. [185]

This is the minimal setup needed to prove the theorem.

- これらの不等式は物質の安定性を理解する上で決定的な役割を果たす. [131]

These inequalities play a crucial role in our understanding of stability of matter.

- これを確認するため, \mathcal{A} を \mathcal{C} で最も小さいアーベル部分圏としよう.

To see this, let \mathcal{A} be the smallest abelian subcategory of \mathcal{C} .

- 指数 m が f_D と選んだ点 $p \in M$ に依存しないことはすぐにわかる.

It is easily verified that m does not depend on f_D and the chosen point $p \in M$.

- 射影曲線の種数を一般化した 2 つの数が高次元で重要な役割を果たす. [174]

Two numbers, generalizing the genus of a projective curve, play an important role in higher dimensions.

- 既に述べたように, 複素多様体 M 上, 同型を法とした正則な直線束の集合はテンソル積に関して群をなす. [174]

As already said before, the collection of holomorphic line bundles on a complex manifold M modulo isomorphism form a group under the tensor product.

- 剰余項.

The remainder term.

- 測度論での病的な現象を避けるため, (X, \mathcal{B}, μ) は常に標準確率空間の完備化と仮定する.

In order to avoid measure theoretic pathologies, we will always assume that (X, \mathcal{B}, μ) is the completion of a standard measure space.

- 微分積分学の基本定理.

Fundamental theorem of calculus.

- 複素代数幾何を研究する上で基本的な対象と写像を思い出そう. [174]

I recall the basic objects and maps one works with in complex algebraic geometry.

- 複素または射影多様体に関する一般論. [174]

Generalities on complex and projective manifolds.

- 部分積分.

Integration by parts.

- 便利な関数空間はいろいろあり, それらを状況に応じて使いわけていく.

There are many useful function spaces and we select a proper space as the situation demands.

- ホモロジー代数は代数的位相幾何, 群論, 可換環論, そして代数幾何などのいくつかの数学の分野で使われる道具だ. [162]

Homological algebra is a tool used in several branches of mathematics: algebraic topology, group theory, commutative ring theory, and algebraic geometry.

Chapter 3

Physical English

- Einstein の相対論的運動学がどんな効果をもたらすだろうか. [131]

What effect does Einstein's relativistic kinematics have?

- $Z\alpha$ が非常に大きい筈に 1 原子のときですら磁場が不安定性を引き起こす可能性があるとわかったとき, また新たな不等式が必要になった. [131]

Again new inequalities were needed when it was realized that magnetic fields could also cause instabilities, even for just one atom, if $Z\alpha$ is too large.

- 安定性に関する結論は精妙で普通の物理の学生が使える初等的な道具立てで簡単に導出できるわけでもないのに, 今日でさえほとんどの物理の教科書で議論されておらず, 問題が指摘されることすらない. [131]

Even today hardly any physics textbook discusses, or even raises this question, even though the basic conclusion of stability is subtle and not easily derived using the elementary means available to the usual physics student.

- 原子の存在と世界の安定性を理解するために解決しなければならない基本的な問題は次のように表現できる: 何故点粒子である電子が (ほぼ) 点粒子とみなせる原子核に落ちこんでいかないのか? [131]

The basic question that has to be resolved in order to understand the existence of atoms and stability of our world is: Why don't the point-like electrons fall into the (nearly) point-like nuclei?

- この奇妙で全く予期していなかった事実を理解するため, 適切な Schrödinger 方程式がゼロモードを持つことを認識する必要があった.

The understating of this strange, and totally unforeseen, fact requires the knowledge that the appropriate Schrödinger equation has '*zero-modes*'.

- この本では量子電気力学, Coulomb 多体系の安定性, 星の重力安定性, 平衡統計の基礎, そして熱力学的極限の存在を議論している. [131]

The topics covered in this book include quantum electrodynamics, stability of large Coulomb systems, gravitational stability of stars, basics of equilibrium statistical mechanics, and the existence of the thermodynamic limit.

- 古典力学から見たときのこの問題は 1915 年に Jeans によってうまくまとめられている. [131]

This problem of classical mechanics was nicely summarized by Jeans in 1915.

- 電気力は異なる符号の電荷を持つ荷電粒子間では引力になり, 同じ符号の電荷を持つ粒子間では斥力になる. [131]

The electric force is attractive between oppositely charged particles and repulsive between like-charged particles.

- 普通の物質の基本的な構成要素は電子と原子核だ. [131]

The basic constituents of ordinary matter are electrons and atomic nuclei.

- 物理学者が物質の世界を記述する基盤となる基礎理論は量子力学だ. [131]

The fundamental theory that underlies the physicist's description of the material world is quantum mechanics.

- 物理学者は量子力学を原子・分子レベルでの‘万物の理論’だと固く信じている. [131]

Physicists firmly believe that quantum mechanics is a ‘theory of everything’ at the level of atoms and molecules.

Part III

Set theory

Chapter 4

Introduction to cardinal number

動画から適当に切り出してくる.

Part IV

Linear Algebra

Chapter 5

Talk: Linear algebra and Calculus: Introduction to university mathematics

This note is based on a course for newcomers in the spring of 2014 at the Tokyo institute of Technology. See also [184, 31, 27, 26, 70, 159].

5.1 Introduction

We talk about linear algebra, especially it's relation to calculus, i.e., differentiation and integration.

In this talk we speak in Japanese but write in English on blackboard. It is because I want newcomers to get used to writing and reading in English.

5.2 What is linear algebra?

Main targets of linear algebra are vectors and matrices. We learn them in high school but linear algebra in university mathematics has somewhat different flavor from one in high school. Furthermore linear algebra is difficult to imagine where we use it compared to calculus. So we first consider its different points and its usage in mathematics and physics.

In high school we consider vectors as geometric objects. However we now consider them as algebraic objects, since they are ones in linear “algebra”! Our vectors are a generalized/abstract version of ones in high school. A vector needs not have direction nor length. Then, what properties should vectors have?

5.2.1 Abstract definition of vectors

In the following we sometimes use the terms linear spaces or vector spaces. These words have the same meaning.

Definition. 2.1.1. (Definition of linear/vector spaces) *Let \mathbb{R} be a field of real numbers: a real number is also called a scalar. A set L is called a **linear space***

if its elements have the following properties. Elements $x, y, z \in L$ are vectors and $a, b \in \mathbb{R}$ are real numbers in the following expressions.

1. Associativity:

$$(x + y) + z = x + (y + z), \quad \forall x, y, z \in L. \quad (5.2.1)$$

2. Commutativity of sum:

$$x + y = y + x, \quad \forall x, y \in L. \quad (5.2.2)$$

3. Existence of unit for sum: *there is an element $0 \in L$ such that*

$$x + 0 = x, \quad \forall x \in L. \quad (5.2.3)$$

4. Existence of an inverse element for sum: *for any $x \in L$ there is an element $y \in L$ such that*

$$x + y = 0 \quad (5.2.4)$$

for the above 0. In fact we can prove the uniqueness of an inverse element.

5. Relation of scalars (real numbers) and vectors:

$$a \cdot (x + y) = a \cdot x + a \cdot y, \quad (5.2.5)$$

$$(ab) \cdot x = a \cdot (bx), \quad (5.2.6)$$

$$(a + b) \cdot x = a \cdot x + b \cdot x, \quad (5.2.7)$$

6. Multiplication law for a scalar unit in \mathbb{R} :

$$1 \cdot x = x, \quad \forall x \in L. \quad (5.2.8)$$

Important points are

1. vectors are summable and sum is commutative operation, and
2. vectors and scalars satisfy some multiplication laws.

The above properties hold for geometric vectors in high school, of course. Our purpose is different from it: we consider vectors as algebraic objects and we derive many interesting properties using (mainly) algebraic thinking!

5.2.2 Examples of linear spaces

1. The set of real numbers \mathbb{R} itself with scalar \mathbb{R} .
2. The set of complex numbers \mathbb{C} itself with scalar \mathbb{R} .
3. The set of complex numbers \mathbb{C} itself with scalar \mathbb{C} .
4. A plane \mathbb{R}^2 with scalar \mathbb{R} .
5. A three dimensional space \mathbb{R}^3 with scalar \mathbb{R} .
6. Higher dimensional space \mathbb{R}^d ($d \geq 4$) with scalar \mathbb{R} .
7. Some subsets of (numerical) sequence, e.g., c_0 , c_∞ , ℓ^2 with scalar \mathbb{R} .
8. Some subsets of functions, e.g., $C^k(\Omega)$, $L^p(\Omega)$, $H^k(\Omega)$ with scalar \mathbb{R} .
9. A set of linear operators with scalar \mathbb{R} .

Today's main targets are introduction to the last three items and its relation to calculus.

5.2.3 Examples of finite, but higher dimensional spaces than three

We have many examples for higher dimensional objects in real world, e.g.,

1. (time evolution of) stock prices,
2. players' movement in soccer,
3. computerized control in robots.

For (time evolution of) stock prices, there are many listing companies and its time evolution is important in real world. This evolution is mathematically representable in higher dimensional space picture.

Other examples are also similar characterization. These are related to analytical mechanics in physics. Moreover analytical mechanics is closely related to geometry. See [69, 244, 243].

5.2.4 Function spaces: examples of infinite dimensional spaces

We have infinitely many infinite dimensional objects. Examples are weather maps and wind direction maps.

Take a point in a world map, and then there are infinitely many directions to blow wind. In classical point of view there are infinitely many space points, and hence there are infinitely many space points and directions of the wind are also infinitely many patterns.

What do we represent this infinitely many probability of the wind in real world? We use functions in several variables, $w(x, y, z, t) \in \mathbb{R}^3$. A value of a function $w(x, y, z, t)$ represents a direction of wind and its strength (length of a vector) at space-time point (x, y, z, t) . We always use this type of mathematics in physics.

Furthermore there are infinitely many types of wind distribution, i.e., we have infinitely many functions. For systematic thinking it is useful to think where functions live in. This is called a function space. There are many useful function spaces and we select a proper space as the situation demands.

In this way, in mathematics we will encounter various types of spaces other than a three dimensional, geometric space. We consider spaces where functions live and ones where spaces itself live.

5.2.5 Linear maps, functionals

We usually consider maps instead of functions in university mathematics. In fact a map is just a function whose domain and range are general sets. First we define a linear map and linear functional.

Definition. 2.5.1. Assume L_1 and L_2 are linear spaces. Then a function $F: L_1 \rightarrow L_2$ is called a **map or operator**. If F preserves linearity, i.e., F has a property

$$F(\alpha f + \beta g) = \alpha F(f) + \beta F(g), \quad \alpha, \beta \in \mathbb{R}, f, g \in L_1, \quad (5.2.9)$$

then an operator F is called a **linear operator**.

If L_2 is \mathbb{R} then F is usually called a **functional**. Furthermore F is called a **linear functional** if it is linear.

Here are some examples.

Example. 2.5.2. 1. Coordinate maps. Let $f = (f(1), f(2), \dots, f(d)) \in \mathbb{R}^d$. We write a vector $f = (f_1, f_2, \dots, f_d)$ as $f = (f(1), f(2), \dots, f(d))$ for later use. This is just a notational convention. Then we get functionals by

$$x_i: f \mapsto f(i). \quad (5.2.10)$$

This is a linear functional since this has a property

$$x_i(\alpha f + \beta g) = \alpha x_i(f) + \beta x_i(g). \quad (5.2.11)$$

2. Definite integrals: First we define a map I as

$$I: f \mapsto \int_{\mathbb{R}^d} f(x) dx \in \mathbb{R}. \quad (5.2.12)$$

This is a linear functional since this has a property

$$I(\alpha f + \beta g) = \alpha I(f) + \beta I(g). \quad (5.2.13)$$

3. An other type of a definite integral:

$$E: f \mapsto \int_{\mathbb{R}^d} \left(|\nabla f(x, t)|^2 + \left(\frac{\partial f(x, t)}{\partial t} \right)^2 \right) dx. \quad (5.2.14)$$

This is a nonlinear functional. In physics this E is called an energy functional.

4. Differential operators: Define an operator as

$$D: f \mapsto \frac{d}{dx} f. \quad (5.2.15)$$

This is a linear operator since it satisfies

$$D(\alpha f + \beta g) = \alpha Df + \beta Dg. \quad (5.2.16)$$

In this way we connects linear algebra with calculus. For analysis of nonlinear functionals we also need various linear spaces and some technique from linear algebra.

5.2.6 Eigenvalues, eigenvectors

These are not learned explicitly in high school. However they sometimes appears in entrance exams.

Definition. 2.6.1. Let A be a linear operator on a linear space L . A real number λ resp. a vector f are called an **eigenvalue** resp. **eigenvector** if they satisfy

$$Af = \lambda f. \quad (5.2.17)$$

Here are examples.

Example. 2.6.2. Let D^2 be a second order differential operator (this is linear) with respect to time and consider

$$mD^2x = -kx, \quad (5.2.18)$$

i.e.,

$$m \frac{d^2x(t)}{dt^2} = -kx(t). \quad (5.2.19)$$

This is an equation of motion for a spring in physics. A solution (eigenvector) is

$$x(t) = A \sin(\omega t + \theta), \quad \omega = \sqrt{\frac{k}{m}}. \quad (5.2.20)$$

We can write a solution using a celebrated Euler's formula:

$$x(t) = Ae^{i(\omega t + \theta)}. \quad (5.2.21)$$

Consideration of eigenvalues for a differential equation is somewhat difficult and we omit it. See, e.g., [46, 45] for details.

5.2.7 Mathematical application of linear algebra

There many branches related to linear algebra. Here are some examples. See also [175].

1. Theory of Lie group and its representation theory.
2. General algebra.
3. Algebraic geometry.
4. Analysis of differential equations.
5. Functional analysis.
6. Operator algebra.

5.2.8 Physical application

There many branches related to linear algebra in physics, too. See also [175].

For example, in quantum mechanics, one of the most fundamental physical theory, linearity is important and fundamental. We say “a superposition principle valids for wave functions,” and this “superposition” means linearity.

In high school we learn a superposition principle for wave. This holds because our wave equation is linear in high school.

5.2.9 Linear algebra and statistics

We use statistics in many branches, including humanities and sociology. E.g., natural language processing has many humanity and infomation theoretic elements. This area needs broad knowledge including probability and statistics. Interested readers should learn, e.g., principal component analysis.

5.3 Integration

5.3.1 Let's define inner products!

In this talk we mainly consider real linear spaces, e.g., \mathbb{R}^d . You may consider complex linear spaces if you know complex numbers.

We assume you agree with the existence of higher dimensional spaces. We want to define angles and length of vectors as in two or three dimensional ones. One reason to consider them is application to physics. There is a projection hypothesis in quantum mechanics: this projection comes from orthogonal projection, and it is just a shadow of objects in three dimensional objects when one shine a light from above.

At first we consider how to define angles between infinite dimensional vectors. Take a look at the following inner product formula in high school.

$$a \cdot b = |a| |b| \cos \theta. \quad (5.3.1)$$

We deform this:

$$\cos \theta = \frac{a \cdot b}{|a| |b|}. \quad (5.3.2)$$

We can derive an angle θ from the value of the cosine function. Lengths¹ of vectors can be computed by inner products, from $|a|^2 = a \cdot a$. So we can derive angles if we can define an inner product. Hence we have to define an inner product. Note that an inner product is not a divine concept.

At first we write an inner product in \mathbb{R}^3 . We write a vector $f = (f_1, f_2, f_3)$ as $f = (f(1), f(2), f(3))$ for later use. This is just a notational convention. We define an inner product for three dimensional vectors f and g as

$$f \cdot g := \langle f, g \rangle := \sum_{k=1}^3 f(k)g(k). \quad (5.3.3)$$

The first notation is one in high school. The second is usual one when considering a functional inner product. There are several notations for an inner product such as

$$(f, g), \quad (f|g), \quad \langle f|g \rangle. \quad (5.3.4)$$

The last one is the Dirac's bracket notation in quantum mechanics.

Inner products in higher dimensional spaces

We wrote an inner product as

$$\langle f, g \rangle = \sum_{k=1}^3 f(k)g(k). \quad (5.3.5)$$

¹In general we call it a norm of a vector. We can consider many norms for a vector since there are many senses of distances.

Since there is no reason to restrict our consideration to a three dimensional space we generalize a dimension three to general d .

$$\langle f, g \rangle = \sum_{k=1}^d f(k)g(k), \quad (5.3.6)$$

where we set

$$f = (f(1), f(2), \dots, f(d)), \quad g = (g(1), g(2), \dots, g(d)). \quad (5.3.7)$$

If you want to consider one in \mathbb{C}^d you should set

$$\langle f, g \rangle = \sum_{k=1}^d \overline{f(k)}g(k) \quad (5.3.8)$$

because it is desirable that a length (norm) of a vector is properly defined, i.e., $\|f\|^2 := \langle f, f \rangle \geq 0$.

We take a limit $d \rightarrow \infty$. We face a problem whether a series converges or not, but we can overcome this by considering converging ones.

$$\langle f, g \rangle = \sum_{k=1}^{\infty} f(k)g(k). \quad (5.3.9)$$

Remark. 3.1.1. *A infinite dimensional vector $f = (f(1), f(2), \dots)$ can be viewed as a sequence. In quantum mechanics we use sequences and matrices living in an infinite dimensional space when considering Heisenberg's matricial mechanics.*

Note that our temporal task is to make how to define an inner product for functions. Let's go back to a finite sum, think $1 = \Delta k$, and rewrite a sum as

$$\sum_{k=1}^d f(k)g(k)\Delta k. \quad (5.3.10)$$

Rewrite k as k/d and Δk as $1/d$. Then we get a sum

$$\sum_{k=1}^d f\left(\frac{k}{d}\right)g\left(\frac{k}{d}\right)\frac{1}{d}. \quad (5.3.11)$$

This is an expression, known as **区分布求積法** in high school².

Taking a limit $d \rightarrow \infty$ the above sum becomes an integral.

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx. \quad (5.3.12)$$

Now the interval is $[0, 1]$ here, but we can take any (measurable) subset of \mathbb{R} .

Hence we find that an inner product for functions can be defined using integral. In fact anything is good if it satisfies the axiom of an inner product.

A linear space with inner product is called pre-Hilbert space. A pre-Hilbert space is called a Hilbert space if it is complete for metric induced by inner product. A Hilbert space is famous as a space where wave functions in quantum mechanics live.

²This is a definition of the Riemann integral.

5.3.2 Axiom for inner products

Axiom. 5.3.1. Let L be a linear space and $\langle \cdot, \cdot \rangle: L \times L \rightarrow \mathbb{R}$ is a map of two variables. The pair $(L, \langle \cdot, \cdot \rangle)$ is an inner product space if the map $\langle \cdot, \cdot \rangle$ is an inner product, i.e., it satisfies the following properties.

1. Symmetry:

$$\langle f, g \rangle = \langle g, f \rangle. \quad (5.3.13)$$

2. Linearity in the second argument:

$$\langle f, \alpha g + \beta h \rangle = \alpha \langle f, g \rangle + \beta \langle f, h \rangle. \quad (5.3.14)$$

3. Positive definiteness:

$$\langle f, f \rangle \geq 0, \quad \langle f, f \rangle = 0 \iff f = 0. \quad (5.3.15)$$

5.3.3 Examples of inner products

There are many inner products in a function space. Let Ω be a (open) subset of \mathbb{R} and $h: \Omega \rightarrow \mathbb{R}_{\geq}$ be a function satisfying

$$h(x) \geq 0, \quad h \neq 0, \quad 0 < \int_{\Omega} h(x) dx < \infty. \quad (5.3.16)$$

Then the following is becomes an inner product:

$$\langle f, g \rangle_h := \int_{\Omega} f(x)g(x)h(x)dx. \quad (5.3.17)$$

We show some examples of inner products.

$$\langle f, g \rangle_1 := \int_{-1}^1 f(x)g(x)dx, \quad (5.3.18)$$

$$\langle f, g \rangle_2 := \int_0^{\infty} f(x)g(x)e^{-x}dx, \quad (5.3.19)$$

$$\langle f, g \rangle_3 := \int_{\mathbb{R}} f(x)g(x)e^{-x^2}dx. \quad (5.3.20)$$

The above inner products are related to the Legendre polynomials, Laguerre polynomials, Hermite polynomials.

Furthermore we take the following inner product.

$$\int_{-\pi}^{\pi} f(x)g(x)dx. \quad (5.3.21)$$

Then we get the following expressions.

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \sin nx \sin mx = \delta_{n,m}, \quad (5.3.22)$$

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \cos nx \cos mx = \delta_{n,m}, \quad (5.3.23)$$

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \cos nx \sin mx = 0. \quad (5.3.24)$$

This means that the functions $\{\cos nx\}$ and $\{\sin nx\}$ are orthogonal. This relates to the famous Forier series expansion. This is used for wave analysis in physics.

5.3.4 Physics for the space $L^2(\Omega)$

The symbol $L^2(\Omega)$ in the title means the space of square integrable functions in the sense of Lebesgue on Ω .

We introduce a physical meaning of the space $L^2(\Omega)$.³

The elements of the space $L^2(\Omega)$ have finite energy. Let us consider a wave equation.

$$\frac{\partial^2 u}{\partial t^2} = \Delta u. \quad (5.3.25)$$

The energy for its solution is written by

$$E = \int_{\mathbb{R}^d} \left((\nabla u)^2 + \left(\frac{\partial u}{\partial t} \right)^2 \right) dx. \quad (5.3.26)$$

We want to restrict solutions whose energy is finite since it is physically meaningless to consider infinite energy solutions. The elements of L^2 have finite energy, by definition, and hence it is the reason why the space L^2 is important in physics. To be precise our derived functions are in L^2 , so we have to consider a Sobolev space H^1 .

In quantum mechanics probabilistic interpretation forces us to consider functions in L^2 . Furthermore we have to impose more severe restriction to functions, i.e., the domain of Hamiltonians.

Linear operators have their domains and, for Hamiltonians, their domains are functions having finite energy.

5.4 Differentiation

5.4.1 Taylor expansion

We start from Taylor's theorem. Assume a function f is differentiable. Fundamental theorem of calculus and integration by parts lead

$$f(x) - f(x_0) = \int_{x_0}^x f'(y) dy = \int_{x_0}^x (x-y)^0 f'(y) dy \quad (5.4.1)$$

$$= [- (x-y) f'(y)]_{x_0}^x + \int_{x_0}^x (x-y) f^{(2)}(y) dy \quad (5.4.2)$$

$$= (x-x_0) f'(x_0) + \left[-\frac{(x-y)^2}{2} f^{(2)}(y) \right]_{x_0}^x + \int_{x_0}^x \frac{(x-y)^2}{2} f^{(3)}(y) dy. \quad (5.4.3)$$

$$= (x-x_0) f'(x_0) + \frac{(x-y)^2}{2} f^{(2)}(x_0) + \int_{x_0}^x \frac{(x-y)^2}{2} f^{(3)}(y) dy. \quad (5.4.4)$$

Further iteration leads the following Taylor's theorem.

$$f(x) = \sum_{k=0}^n \frac{(x-x_0)^k}{k!} (D^k f)(x_0) + \int_{x_0}^x \frac{(x-y)^n}{n!} f^{(n+1)}(y) dy, \quad (5.4.5)$$

³To be precise the proper space is not L^2 but H^1 .

where $D^k f$ is a shorthand notation for $d^k f/dx^k$, the last term of the RHS is called the remainder term and denoted by R_n .

Now assume the remainder term R_n converges to 0 as $n \rightarrow \infty$ and the above series also converges in some suitable sense. Let $x_0 = 0$. Then we get

$$f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} (D^k f)(0). \quad (5.4.6)$$

This series is called a **Taylor series** around 0 and this power series expansion is called a **Taylor expansion** around 0.

A Taylor expansion of an exponential function e^x around 0 is as follows.

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \quad (5.4.7)$$

We compare the above two expressions and set x to xD in the expansion of e^x . Then we have the following clear expression.

$$f(x) = (e^{xD} f)(0) \quad (5.4.8)$$

If we use quantum mechanical notation, $p = -iD$, the above expression becomes

$$f(x) = (e^{ixp} f)(0). \quad (5.4.9)$$

5.4.2 Unitary representation of a group

The above story closely related to a unitary representation theory of a group. Simply speaking a underlying space \mathbb{R} itself is a group and it acts on itself. This action lifts up to a function space living on a underlying space. See [119] for details.

One point. Physicists often define functions of operators by Taylor expansion. However it is not good and insufficient in view of mathematics. We can see it when we consider the action of e^{ixp} .

The operator e^{ixp} shifts an argument of a function x , and this shift is defined without differentiability of functions, i.e., we can define an action of e^{ixp} for more general functions. However if we define it by a Taylor expansion, the operator e^{ixp} is defined for only analytic functions.

Fourier expansion

A simple argument can deduce a Fourier expansion. We can formally consider for \mathbb{R} , but this is a little bit mathematically problematic. Instead we consider on an interval $[-\pi, \pi]$.

Consider a linear differential equation (eigenequation for self-adjoint operator p)

$$pf = kf, \quad k \in \mathbb{R}, \quad (5.4.10)$$

$$f(-\pi) = f(\pi). \quad (5.4.11)$$

Then the above solutions are

$$k \in \mathbb{Z}, \quad f_k(x) = e^{ikx}. \quad (5.4.12)$$

Euler's formula says

$$e^{ikx} = \cos kx + i \sin kx, \quad (5.4.13)$$

and this leads a Fourier expansion.

5.4.3 Analitical mechanics and quantum mechanics

In quantum mechanics a momentum becomes a differential operator and denotes $p = -iD$.

In analytical mechanics a momentum can be viewed as a generator of space shift. A differential operator $-id/dx$ is also a generator of space shift in view of representation theory. Hence these are common property, generator of space shift. This is an important **formal** connection to classical mechanics and quantum mechanics. This connection is also important in geometry.

5.4.4 Physics and representation theory

Representation theory is also important for theory of relativity and relativistic quantum field theory. See [83, 82] for details.

5.5 Variational problem: mixture of differentiation and integration

A variational problem is an infinite dimensional version of diffrential theory. This is important and interesting in both physics and mathematics.

It is a mathematically and physically famous problem that why soap bubble becomes round. This type of problems is called a geometric variational problem. See [157] for details.

5.5.1 Examples of variational problems

shortest time problem

Assume you are a lifesaver on beach and there is a person drowning. You have to save the person. How do you determine a path to the person in the shortest possible time? This is a famous variational problem.

You have to determine a path and it is a function mathematically. Hence you have to search a function in a function space. How to search it?

5.5.2 Mathematical formulatrion: functionals, linear operators

We formulate the above problem mathematically. In usual we make a some **functional** and search a function which optimize, i.e., minimize or maximize the value of it. In this procedure we differentiate a functional by functions, and this is a variational derivatice.

Recall an energy functional (5.2.14). This energy functional is a mixed type functional of differential and integral. We often face this type of functionals in theory of partial differential equations, in particular, quantum mechanics

and we need a minimum value of E and its minimizer. These are physically important quantities, ground state energy and ground state. Some types of quantum mechanical problems are formulated as a variational problem.

Chapter 6

あとで大幅 rewrite: 学部 3 年のときに書いた「本」

6.1 はじめに

6.1.1 数学「を」学ぶ

これは基本的に数学の本です。物理現象を解析するのに、ある数学の理論が有用なので、実際の解析を視野に入れつつ、そうした数学を展開していこう、という趣旨です。物理学にとって、数学とは驚異的に役に立つもので、場合によっては数学的な考察から新たな物理を生み出すことさえあるほどです。しかし、物理を学び、研究する際に何故数学を用いるのでしょうか？

物理現象の解析に数学を用いたのは、Galilei によると言われています。彼は「哲学は宇宙という大書物に書かれている」と考えていたようで、物理現象を数学的に定式化することに成功しました。実際にどんな現象が起こっているのかを確かめるために、(再現可能な) 実験を行って、位置の時間変化などを数値として表し、その上でその数値を比較する、という方法は、確かに優れたものだと思います。うまく数値を扱うために、数値の扱いに関してすでに整った体系を持っていた数学を用いよう、というのは自然な発想でしょう。

しかしこれから私たちが学ぶ数学には、単なる数値の扱いを超えた、かなり高級なものも含まれています。何故こうした数学を学ばなければならないのか考えてみると、そうすると現象の解析に便利だから、というだけです。他にもっと楽な方法はないのか、と思わないでもありません。そもそも、物理を学ぶのに何故数学が必要なのか、役に立つのか。非常に不思議です。Feynman は、物理学を日常言語で表現できないうちはまだまだ (人類の) 自然界への理解が足りないのだ、と考えていたようです。数学で表現された物理学の理論が (全てとはいわずとも) 日常言語に翻訳可能である、という事実も、良く考えてみると驚くべきことです。

現在の物理学を見渡す限り、数学はなくてはならない重要な道具であることは間違いないでしょう。しかしそれだけでしょうか。Bohr は前期量子論において、その理論構築の中で、実験結果とあわせるために、振動数の量子条件を数学的な関係式として導入したようですが、後になって、この物理的な意味が、物質波の理論とともに de Broglie によって与えられました。このように、数学的な考察がはじめにあり、その後に数学に物理的な生命を吹き込むという作業を行う場合があります。こうした事例は数学が、物理を数学的・抽象的な舞台に持ち込むことで、逆に単なる直感を超えた議論が展開できるようになり、物理的直観を伴った深い理

解をもたらす可能性を示唆します。

何はともあれ、以下、本書では数学に対して

何故だか分からないがとても役に立つ。
道具というよりもむしろ武器である。

というスタンスのもと、物理学を学ぶのに有用な数学の理論を展開していきます。とにかく実際に自然界で何が起きているか良く分からないので、そうした化け物を人間のつつましい能力で扱うための武器である数学は、自然界の事物の本質に従うべきもので、すっきりきれいにまとまるものばかりというわけにはいかないでしょう。いきおい泥臭い話にならざるを得ない部分があります。そして、頼みの武器がなまくらでは困ります。物理学を埋め込むのに、時として抽象性が高く、難しい数学を用いなければならないことがあります。これらは全て物理学のためです。頑張って食らいついてください。

また、単に理論を紹介するということなら、すでに世に良い本はたくさん出回っています。(筆者だけかもしれませんが、以前、特に数学の) 本を読む際に一番退屈だったのは、定義がごちゃごちゃと出てくることです。しかし、近い将来私達がそうした理論を作ろうというときがやってきます。このとき一番重要になるのは、先程「退屈」といった定義のところでしょう。何故かという、一般的に何かを調べようというとき、そもそも何を調べるか、どう調べるかということが一番の問題になります。調べたい事柄に対し、それを良く反映した物理量を導入(定義)していくことになるでしょう¹。

こうした点を踏まえ(筆者の勉強もかねて)、この本ではなるべくどのような定義を何故導入するかということに神経を使っていくことにします。これらは筆者が一番納得できると思った定義と論法であり、他の人から見れば、気に食わない定義・論法かもしれません。そういう場合、自分で積極的に納得のいく定義を考えてください。数学での例になりますが、L. Schwartz という人が超関数の理論を考え出しました²。しかし、日本人数学者の佐藤幹夫は、この定義がひどく気に食わなかったようで、今日「佐藤超関数」として知られる理論体系を組み上げ、世界的に有名になりました。納得のいかない定義を納得のいくものにすることで世界的な業績をあげてしまった、このような例があります。定義は納得するまで考えてください。定義こそが生命線です。

6.1.2 数学「で」学ぶ

作りかけ。

6.1.3 一年で物理の講義がほとんど無い理由

人によっては衝撃的なタイトルであるかもしれませんが、応物・物理に来たというのは、一年生では物理をほとんどやらない、というのですから。しかしこれにはもっともな理由があります。まず物理の代わりに何をやるのかということですが、数学をやります。先程述べたように、現在の枠組みの中では、物理学の理論を埋め込むべき数学を知らない、物理学が理解出来ないからです。

例をあげます。一年の授業で物理学 A という講義がありますが、これはいわゆる力学です。一般に空間の点はベクトルで表されますが、この点の動きを力学的に追跡するのに速度・加速度という概念を必要とします。数学的には、速度は位置(変位)を時間微分したもの、加速度はもう一回位置(変位)を時間微分した

¹後で具体例を加える。

²これは物理学において Dirac が導入したデルタ関数を数学的に正当化する理論であり、非常に重要な理論です。

ものです。運動方程式は加速度を含んだ式ですが、 $\mathbf{r} = (x, y, z)$ を位置ベクトル、 $\mathbf{f}(\mathbf{r}, t)$ を力のベクトルとすると

$$m \frac{d^2 \mathbf{r}}{dt^2} = \mathbf{f}(\mathbf{r}, t) \quad (6.1.1)$$

と書けます³。これは数学的には2階の常微分方程式となります。もちろん(物理学を学ぶのに必要になるレベルの)微分方程式は高校で学んでいないはずですが、

高校で力として重力、摩擦力、垂直抗力、電磁力など色々学んだと思いますが、その中に Lorentz 力というのがありました。これをベクトルで書くと

$$\mathbf{f}(\mathbf{r}, t) = q \{ \mathbf{E}(\mathbf{r}, t) + \mathbf{r} \times \mathbf{B}(\mathbf{r}, t) \} \quad (6.1.2)$$

となります。右辺にある \times はベクトルとベクトルの外積を表す記号として使われます。おそらく数学 A (線形代数) の講義で出て来る前に力学で出てくるでしょう。

また高校で位置エネルギーや電位というのを学んだでしょうが、これらはまとめてポテンシャルと呼ばれます。あまりいい加減なことをいうのも良くないのですが、ポテンシャルというのは空間微分すると力(の成分)が出てくるものだと思います。実際に重力ポテンシャルから重力を出してみます：

$$\mathbf{f}(\mathbf{r}, t) = - \frac{\partial (mgz)}{\partial z} = -mg. \quad (6.1.3)$$

ここでまた変な記号が出てきますが、一般式はポテンシャルを $U(\mathbf{r}, t)$ として

$$\mathbf{f}(\mathbf{r}, t) = - \text{grad } U(\mathbf{r}, t) \quad (6.1.4)$$

$$\text{grad} := e_1 \frac{\partial}{\partial x} + e_2 \frac{\partial}{\partial y} + e_3 \frac{\partial}{\partial z} \quad (6.1.5)$$

$$e_1 := e_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 := e_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 := e_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (6.1.6)$$

のようになります。ここで「 $\$:=\$$ 」は「右辺を左辺で定義する」という意味です。

ポテンシャルに限らず、物理に登場する関数は一般に空間と時間を変数に持ちますから、4変数関数です。そこで4変数関数の微分が出来るようになる必要があります。多変数の微積分は後期の数学 B で学ぶことになっていますが、遅すぎます。慣れれば別にどうということもないのですが、慣れるまでが大変なのです。そうかといって、詳しく数学を教えていたら物理の講義になりません。したがって数学の説明は必要最小限にして話がどんどん進みます。

以上、簡単に物理学 A のはじめの部分を書いてみました。この時点で「数学が便利」というのはよく分からないと思いますが、数学が必要ということは分かってもらえたと思います。そしてこんな講義についていけるのかと思う人が大半でしょう。

例えていえば、数学が分からないということは実験するのに実験装置の使い方が分からない、ひどい場合は説明書すら読めない、ということです。しかし実験装置は説明書を読むだけでなく、実際に使いながら操作法を学んでいくものです。数学書を(1人で)読み進めながら学んでいくことは非常に難しいことですが、物理がしたくて大学に来たのですから「何のために数学をするのか」というモチベーションが無いとどうしても途中で挫折してしまうでしょう。ここに生じる隙間を

³高校で学んだ記法 \vec{r} の代わりにこのように肉太の文字で書きます。普通の文字、 r (肉太のものは \mathbf{r}) と区別して下さい。実際に手書きするときは \mathbf{A} のように書きます。

埋めるには、実際にいろいろ物理で遊んでみて、どんな所でどのような数学を、どのように用いて、何をやるのかを自分で見てくることがよいでしょう。

本書ではその一例として、振動・波動現象の解析を実際に行ない、その中でどのような数学がどのように現れるかを示してみたいと思います。何故振動・波動を扱うのかというと、これが建築物の耐震性、共振の防止などの実用的な観点からも重要なだけでなく、物理でも振り子の連成振動がニュートリノに質量があることを証明したニュートリノ振動の力学モデルであること、電磁場が実は調和振動子の集合であることなど、初等的な所から最先端まで、物理学のいたるところに登場する重要なものだからです。そして、重要な数学もたくさん登場します。物理・数学共にそれ程詳しく論じるスペースはありませんが、この2年間の経験からこれだけあれば十分物理で遊べる、といえるぐらいの内容は盛り込んであります。

時にはかなり細かい注がついていることがあります。これらには(進んだ注)という、見なくても分からなくても良い注だという印をつけておきます。きちんと考えるとややこしいことがあるようだ、ということだけ把握してあれば十分です。

6.1.4 物理で使う数学

基本的に物理学で使う数学は、微分・積分と線形代数です。これらの計算と、多少の理論を(理解できずとも)知ってさえいれば、物理学の中で遊びまわれます。表題のとおり、各分野で使う数学を列挙してみます。大体の説明もつけます。

(一般、古典)力学では行列式の計算と対角化、各種計算の為に(多変数の)微分積分、現象を記述するための重要な道具である微分方程式が必要になります。

解析力学ではかなり込み入った偏微分の計算が出来なくてはなりません。2年の講義では電磁場の解析力学なども扱いますが、その為にベクトル解析が使えると便利です。

電磁気学では何よりもまずベクトル解析を使いこなせないといけません。これは電磁場の数学的取り扱いを非常に容易にしてくれます。数学的・形式的な面だけでなく、ベクトルによる物理法則の記述は相対論的共変性という観点からみて物理的に本質的な役割を果たします。また、Maxwell (マクスウェル) の方程式を実際に解くことがありますが、その際に Fourier (フーリエ) 級数・Fourier 変換が非常に便利な道具として活躍します。電磁気学から特殊相対性理論が生まれたわけですが、そこでは線形代数の理解が重要です。■重ね合わせ■の理解にも線形代数の理解が不可欠です。

熱力学では偏微分と凸関数の解析が出来ればどうにかになります。

回路理論は応物の皆さんは2年で必修になります。ここでは回路方程式をきちんと扱えるようにするために、複素数、(常)微分方程式、Fourier 級数、Fourier 変換、Laplace 変換が必要になります。

量子力学の基礎数理は線形代数です。透徹とした線形代数の世界を見せられることになります。また具体的な問題を解こうと思うと微分方程式 (Schrödinger 方程式) を解くことになりますが、そこでは特殊関数などを自由自在に使いこなせると便利です。これまでと異なり、もはや厳密解を求めることが出来ないようなケースを扱うことになりますが、そこで「摂動」という手法が出てきます。要は近似の度合いを上げたいということなのですが、そこで微分積分の計算力を問われます。

統計力学では: 書きかけ。

相対性理論の一番基本的なところは線形代数と Taylor 展開 (一次近似) さえ出来れば完璧に理解できます (100 年の記念で節目の年です。Einstein の論文を読みましょう)。もう少し細々としたところまで扱おうとするならば、テンソルや群論の力を借りて数学的な議論をなるべく簡単に済ませたいところです。

これだけでは実際にどういう数学をどこでどう使うのか、ということは良く分からないでしょうが、一年で学ぶ線形代数、微分積分、ベクトル解析を良く使うことは分かると思います。

見た方が早いです。実際に振動現象の解析に入ってみましょう。

6.1.5 数学記号の記法

本書では記述を簡便にするために、数学の記法をいくつか導入します。その中で基本的な記法と概念をいくつかここで紹介します。

作りかけ。

本書では振動・波動現象の解析を通じて、数学を学んでいくことにします。最終目標は、Fourier (フーリエ) 解析⁴ に慣れ親しみ、きちんと使えるようになることです。物理としては本末転倒ですが、本書では振動・波動を次の2つの(線形の微分)方程式に支配される現象であると定義します⁵:

$$\frac{d^2 u(t)}{dt^2} = -\omega^2 u(t) - \gamma \frac{du(t)}{dt}, \quad (6.1.7)$$

$$\frac{1}{v^2} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} = \frac{\partial^2 u(\mathbf{x}, t)}{\partial x^2} + \frac{\partial^2 u(\mathbf{x}, t)}{\partial y^2} + \frac{\partial^2 u(\mathbf{x}, t)}{\partial z^2}. \quad (6.1.8)$$

余計な項が入っていますが、上の式(で $\gamma = 0$ とした式)は、高校でも学んだ単振動の方程式です。下の式は、その名もずばり波動方程式です。

さて振動・波動を取り上げる理由ですが、これは2つあります。第一に、振動・波動は物理のなかでよく出てきます。単振動は一つの例ですが、これは電気回路の式とも数学的に等価です。2つのおもりをバネでつないだモデルが、ニュートリノに質量があることを証明したニュートリノ振動の力学モデルであること⁶、電磁場が空間の中を伝わる時波動として伝わっていくことなど、初歩から最先端まで、いたるところに出てきます。実用的なところでは、建築物の耐震性、共振の防止をモデル化して考えたとき、この方程式が出てきます。

もう一つは、色々な数学が出てくることです。これに付いては以下の本書の構成の中で説明します。

第2章は線形代数です。高校でいうと、行列とベクトルの理論です。物理としては、たくさんの質点をバネでつなげ、この振動を調べたいと思います。特に一般の n 個の質点があるとき(これは例えば、固体物理学で結晶の格子振動の模型として出てきます)。に、これをスマートに調べようと思うと、線形代数の力を借りるのが便利であることが分かります。また、あるポテンシャルが支配する質点系は、ポテンシャルの安定点近傍で微小振動することが分かっています。これを微分積分の章で証明するとき、ここで使った議論の助けを借ります。

他にも、線形代数は量子力学を学ぶときに決定的に重要な役割を果たします。ここで使われるのは、線形代数の抽象論(線形空間論)です。本書ではここに重点を置いて議論を展開します。

最後に、無限次元の線形代数を扱います。これは本書の目標、Fourier 解析の舞台です。ここで Fourier 解析の「こころ」を学ぶことにしましょう。本書のエッ

⁴簡単にいうと、三角関数で色々な関数を展開(近似)しよう、というものです。三角関数は周期を持っているので、特に周期関数の近似で威力を発揮します。

⁵誤解のないように言っておきますが、これらの式に従わない「波動」が存在します。非線形波動と呼ばれる現象群がその代表です。ソリトンだとか色々あるようです。

⁶2年のときに筑波の高エネルギー加速研に行ったことがあるのですが、その先生からこのお話を聞いたので間違いありません。「一つのモデルでたくさんの現象が説明できるのが物理の面白いところなんだ!」と力説していました。

センスがここに詰まっています、読むのも大変と思いますが、頑張ってついてきてください。

第3章は微分積分です。物理としては、ポテンシャルの安定点近傍での微小振動を証明します。まず復習をかねて1変数関数の微分を見直します。この1変数の微分をもとに、多変数関数の微分(偏微分)を考えます。次に積分ですが、物理でよく出てくる1変数の積分の計算結果と多変数の積分(重積分)の変数変換を中心に扱うことにします。

第4章は常微分方程式論(初期値問題)です。単振動の方程式は高校で学んだ、としてきましたが、これを考えなおします。物理としては、よく出てくる常微分方程式の解の性質を調べてみます。これらに関しては、式を見ただけで解の様子が把握できるようになってください。数学としては、常微分方程式の初期値問題の解の存在と一意性の議論が中心です。簡単な解法にも触れます。

第5章はベクトル解析です。物理としては、やはり電磁気学です。いくつかの実験式から Maxwell (マクスウェルの方程式) を導出し、これから電磁場が波として伝わることを見ます。

第6章は複素関数論です。第2章で Fourier 解析を扱うときに Euler (オイラー) の公式

$$e^{i\theta} = \cos \theta + i \sin \theta \quad (6.1.9)$$

を紹介します。これは指数関数の複素領域への拡張ですが、これが一意的であることを証明したいと思います。他にも留数解析などの重要な計算法を紹介します。物理として本質的なのはやはり量子力学なのでしょうが、本書の程度を超えます。この章は物理の色が少し薄いかもしれませんが。

第7章は偏微分方程式です。様々な物理法則は偏微分方程式で記述されます。学部レベルの物理で出てくるのは、ほぼ線形の偏微分方程式です。波動方程式はもちろんこの中の1つです。これらを解法を中心に説明していきます。ここで Fourier 解析をガンガン使います。

第8章は群論です。物理法則には共変性を要求するのが自然です。この共変性は群論という数学と深く関わっていることが知られています。これを電磁気学での Lorentz 変換と絡めて説明しようと思います。

6.1.6 物理と近似: まだ書きかけ

良い近似とは何か、みたいなことを書きたい。まだうまくまとまらない。節のタイトルも問題か?

本書では、数学的に厳密な解を(主に三角関数を使って)具体的に書き下せる物理現象しか扱いません。これは、物理現象を暴力的に単純化しているからです。例を挙げましょう。結晶を考えます。結晶は規則正しく格子状に原子が並んでいる、というイメージを持っていると思います。そして原子は、格子点のまわりで熱振動しているでしょう。温度が高くなると、原子の熱振動の振幅が大きくなり、最後には結合が切れて液体になると考えられます。ここから、結晶では原子はバネでつながっているというモデルが作れます。このとき、バネでつながれたのは、すぐ隣の原子どうしです。

ちょっと考えると、想像を絶する単純化をしていることが分かります。まず、原子を結び付けているのは Coulomb (クーロン) 力のはずですが、これは距離の2乗に反比例する力です。これをバネ(距離に比例)と近似しています。また、Coulomb 力は遠くの原子にもはたらきますが、この効果を切り捨てています。

しかし、実験と比較してみると、この乱暴なモデルでも定量的に精度良く説明できる現象があります。格子振動とその量子化、フォノンといった概念を登場さ

せ、物理を豊かにしてくれます。むしろ、Avogadro (アボガドロ) 数くらい原子が集まる (多体系) と、少数粒子の系とは全く違うふるまいを見せてくれることを教えてください。

物理では、近似はただの単純化ではありません。色々な意味で、1 番強烈な例は熱力学でしょう。例えば気体の熱力学を考えます。気体は Avogadro 数個の原子からなります。しかし、熱力学では Avogadro 数個の原子を直接扱う (ミクロな取り扱いをする) ことはありません。体積、温度、圧力といったマクロな量だけを用いて、この大自由度の系を特徴づけできないかと考えます。1 つ 1 つの原子を全く考えない、というある意味では暴力的な近似とも考えられます。

しかし、熱力学は定量的にも厳密な結果が出せる、物理の中でも 1, 2 を争うほど正確な理論です。それだけではなく、単純なミクロな系のあつまりではありえない不可逆過程や相転移といった現象や、いわゆる「熱」まで含めたエネルギー保存則の拡張など、新たな物理を見せ、説明してくれます。多体系での原子たちは、少数粒子のときとは全く別の論理に従っているようです。

6.2 調和振動子と線型代数

6.2.1 導入

高校で単振動を学んだことと思います。これは調和振動とも呼ばれます。以下では調和振動で統一します。調和振動の式 (運動方程式) を書くと ⁷,

$$m\ddot{x}(t) = -kx(t). \quad (6.2.1)$$

そしてこの方程式 (微分方程式) の解が

$$x(t) = A \sin(\omega t + \phi), \quad \omega := \sqrt{\frac{k}{m}} \quad (6.2.2)$$

であることは知っているとしします ⁸。これから両端が壁につながった 3 本のばね定数 k のばねの間に質量 m のおもりが 2 つつながれている系を調べます。

受験問題でもよく見かけるような状況です。これは固体の格子振動の模型などで重要です。2 つのおもりの平衡点からのずれを $x_1(t), x_2(t)$ として、おもりの運動方程式を書くと ⁹,

$$m\ddot{x}_1 = -kx_1 + k(x_2 - x_1) \quad (6.2.3)$$

$$m\ddot{x}_2 = -k(x_2 - x_1) - kx_2 \quad (6.2.4)$$

となります。あとできちんとやりますが、天下一に、

$$X_1 := \frac{1}{\sqrt{2}}(x_1 - x_2) \quad (6.2.5)$$

$$X_2 := \frac{1}{\sqrt{2}}(x_1 + x_2) \quad (6.2.6)$$

として ¹⁰ 上の式に代入して整理すると、

$$m\ddot{X}_1 = -3kX_1 \quad (6.2.7)$$

$$m\ddot{X}_2 = -kX_2 \quad (6.2.8)$$

⁷時間微分を $\dot{x}(t)$ のように上に点を打つことで表します。2 階の時間微分は $\ddot{x}(t)$ です。

⁸微分方程式の章でもう少し詳しく触れます。

⁹変数の時間 t を書くとは煩雑であり、また明らかなものなので省略しました。よくこうした省略をします。慣れてください。

¹⁰:= というのは右辺で左辺を定義するという意味でした。

となり実に綺麗に分解できてしまいました。こうすると高校でやってきたのと同じように

$$X_1 = A_1 \sin \left(\sqrt{\frac{3k}{m}} t + \phi_1 \right) \quad (6.2.9)$$

$$X_2 = A_2 \sin \left(\sqrt{\frac{k}{m}} t + \phi_2 \right) \quad (6.2.10)$$

となり, X_1, X_2 の定義式から x_1, x_2 が求まります。

今度は同じ状況ですが, バネが 4 つで 3 つの質点とそのバネの間につながれている系を考えます。先程と同様に質点の平衡点からのずれを x_i , $i = 1, 2, 3$ として運動方程式は

$$m\ddot{x}_1 = -kx_1 - k(x_1 - x_2) \quad (6.2.11)$$

$$m\ddot{x}_2 = +k(x_1 - x_2) - k(x_2 - x_3) \quad (6.2.12)$$

$$m\ddot{x}_3 = -kx_3 + k(x_2 - x_3) \quad (6.2.13)$$

となります。またもや天下りに

$$X_1 := \frac{1}{2} (\sqrt{2}x_1 + x_2 - x_3) \quad (6.2.14)$$

$$X_2 := \frac{1}{2} (\sqrt{2}x_2 + \sqrt{2}x_3) \quad (6.2.15)$$

$$X_3 := \frac{1}{2} (-\sqrt{2}x_1 + x_2 - x_3) \quad (6.2.16)$$

として上の式に代入して整理すると

$$m\ddot{X}_1 = -2X_1 \quad (6.2.17)$$

$$m\ddot{X}_2 = -(2 - \sqrt{2})X_2 \quad (6.2.18)$$

$$m\ddot{X}_3 = -(2 + \sqrt{2})X_3 \quad (6.2.19)$$

となり, これ以降は 2 質点の時と同じです。

これをさらに n 質点系へと一般化しましょう。何故一般化するのかということとは当然の疑問ですが, 上述の格子振動などでも一般には多数の質点の集まりですから, 一般の n での解析が必要です。また弾性体 (例えば弦) の振動を扱おうとすると, これは $n \rightarrow \infty$ とした極限を考えることになります。この一般化を見通しよく行なう為に私たちは線型代数学の力を借りることになります。2 質点ぐらいなら適当にやってもどうにかかなりそうですが, 3 質点ではもう既になかなか直観的にいけそうな気配がありません。一般の n 質点では適当にやってもどうにかするのはとても無理そうです。付記しておく, 2 質点系の場合の X は重心座標と相対座標になっています。つまりこれらは現象の見やすい座標系への座標系の変換であったとみなせます。しかし 3 質点系では一見して良く分かる, というわけにはいきません。こうした座標系の変換をどう見つけるかという考察をしたいというわけです¹¹

次章では一般のポテンシャルに対し, 安定点が存在すればその近傍ではそのポテンシャルに支配された系の運動が微小振動になることを証明します。そこへの接続に十分な範囲での議論をしなければなりません, それ以外にも重要なことは補足しながらいくことにします。

¹¹他にも線型代数は建築などで巨大行列の行列式の計算をすることがよくあるらしく, その為の行列式の計算の効率の良いプログラミングの作成などは今でも研究対象となっているようです。

6.2.2 記法の確認と線型写像

いくつかの一般的な記法をまとめておきます。今この場で覚えようとしても無理なので、使いながら覚えていって下さい。つまらないですが少し我慢して下さい。

数学的对象となるものの集まりを集合といいます。実数全体、自然数全体などはいずれも集合です。1つの集合 A があるとき、 A を構成する個々のものを A の元 (ゲンと読む) といいます。 x が A の元であることを記号 $x \in A$ あるいは $A \ni x$ で表します。集合 A と B があつたとし、 A の全ての元が B の元となつているとき、すなわち任意の A の元 a が $a \in B$ となるとき、 A は B の部分集合であると言い、 $A \subset B$ または $B \supset A$ と書きます。元が1つもない集合も特別な集合とみなし、これを空集合と呼び、 \emptyset と書きます。これは任意の集合の部分集合です。

集合 A が元 $x_1, x_2, \dots, x_n, \dots$ からなるとき、記号

$$A = \{x_1, x_2, \dots, x_n, \dots\} \quad (6.2.20)$$

$$= \{x; x_n, n \in N\} \quad (6.2.21)$$

で表すことがあります。すぐあとで述べますが N は自然数の集合です。一般にある条件 $C(x)$ を満たす元の集合を

$$\{x; C(x)\} \quad (6.2.22)$$

と書きます。具体的には

$$\{x; x \leq 3, x \in N\} \quad (6.2.23)$$

などです。

集合 A, B があり、写像¹² f が A を定義域とし B を値域とするとします。これを

$$f: A \rightarrow B \quad (6.2.24)$$

$$f: A \ni a \mapsto b = f(a) \in B \quad (6.2.25)$$

などと書きます。自然数¹³、整数、有理数、実数、複素数¹⁴ 全体の集合を各々 N, Z, Q, R, C と書きます¹⁵。また正の整数は $Z^+ := \{1, 2, \dots\}$ 、正の実数は $b_R^+ := [0, \infty)$ と書くことがあります。他にも n 次元空間を普通の数のように R の n 乗ということで R^n と書きます。あとで複素 n 次元空間 C^n も使います。さらに、 $n \times m$ 複素行列全体の集合を $M(n, m; C^n)$ と書きます。

ここからは行列の記法の確認です。特に行列式のところで証明を含めて大体的なことを3次正方行列について書きます¹⁶。2次だと単純すぎていろいろ簡単になってしまっていることがあり、4次以上は書くのが面倒だからです。証明については一般の n 次でも成立するようなものをつけます。

¹²数のことだと思って下さい。細かいことを言うと定義域、値域が一般の集合のとき主に写像といい、これら実数や複素数のときに主に関数というようです。

¹³ここでは自然数は0を入れます。

¹⁴自然数、整数、有理数、実数、複素数は各々英語で natural number, integral number (integer), rational number, real number, complex number といいます。 N, R, C はその頭文字です。

¹⁵それぞれ手書きするときは N, Z, Q, R, C と書きます。

¹⁶正確には n 次元でも同じように出来るけれども、一般的に行なうのが非常に面倒な場合に3次でやります。次元による面倒が全く無いような場合がありますから、そういう場合に n 次で書くことがあります。

まず A を 3 次正方行列としその (i, j) 成分を a_{ij} としましょう. このとき, 行列 A を次のように書きます.

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3) \quad (6.2.26)$$

ただし

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}. \quad (6.2.27)$$

また, 適当な場所もないようなので, この場で n 項単位ベクトルを導入しておきます. $n = 3$ として書くと

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{e}_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (6.2.28)$$

単位行列を I と書くことにしますが, ついでに Kronecker (クロネッカー) の δ と呼ばれるものを定義します. これは

$$\delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (6.2.29)$$

というもので, これを使うと単位行列は $I = (\delta_{ij})$ と書けます. 零行列は O と書きます.

単位行列を定義したので次は逆行列を定義しておきましょう. ある行列 A に対し X という行列が存在して $AX = I$ となったとしましょう. このとき X を右逆行列と呼びます. 左逆行列も同様に定義します. 有限次行列¹⁷の場合, 片方の逆行列が存在するともう片方の逆行列が存在して, この 2 つが一致します. 実際, X を A の右逆行列, Y を A の左逆行列とすると

$$X = IX = (YA)X = Y(AX) = YI = Y. \blacksquare \quad (6.2.30)$$

また逆行列が存在する行列のことを正則行列といいます.

次に高校で学ばなかったと思いますが, 転置行列 tA というものがあります. この定義は

$${}^tA := (a_{ji}) = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \quad (6.2.31)$$

です. 元の行列の行と列をひっくり返したものです.

書き忘れていましたが一般にベクトルは全て列ベクトルとします. 行ベクトルはあまり使わないのですが, スペースの関係もあって使いたいことがありますから, その為に今の転置の記法を流用します. またこれを用いて以下のように転置行列を書くことがあります.

$${}^t\mathbf{a}_1 := (a_{11} \ a_{21} \ a_{31}) \quad {}^t\mathbf{a}_2 := (a_{12} \ a_{22} \ a_{32}) \quad {}^t\mathbf{a}_3 := (a_{13} \ a_{23} \ a_{33}) \quad (6.2.32)$$

$${}^tA = \begin{pmatrix} {}^t\mathbf{a}_1 \\ {}^t\mathbf{a}_2 \\ {}^t\mathbf{a}_3 \end{pmatrix}. \quad (6.2.33)$$

¹⁷次元についてはあとでちゃんと述べます.

転置行列と元の行列が一致する行列, すなわち ${}^tA = A$ となる行列を対称行列といいます. 一番初めに見せた運動方程式を行列でまとめたときに出てくるのは対称行列です. 実際に書いてみると分かります.

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 - k(x_1 - x_2) \\ m\ddot{x}_2 &= +k(x_1 - x_2) - k(x_2 - x_3) \\ m\ddot{x}_3 &= -kx_3 + k(x_2 - x_3) \end{aligned}$$

$$\Leftrightarrow m \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{pmatrix} = -k \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

あとで少し使うので対称行列の性質を 1 つ証明しておきます. それは ${}^t(AB) = {}^tB {}^tA$ という性質です. 具体的に両辺を計算するとすぐ分かります. やってみましょう.

(証明) $A = (a_{ij})$ という書き方を使います. 記述の便宜を図るため, ${}^tA = (a_{ji}) = ({}^ta_{ij})$ という記法も一時的に使います.

$${}^t(AB) = {}^t \left(\sum_{k=1}^3 a_{ik} b_{kj} \right) = \left(\sum_{k=1}^3 a_{jk} b_{ki} \right) = \left(\sum_{k=1}^3 {}^tb_{ik} {}^ta_{kj} \right) = {}^tB {}^tA \quad (6.2.34)$$

あともう 1 つ転置について重要なものとして直交行列というものがあります. これは ${}^tUU = I$ となるような行列のことです. すなわち ${}^tU = U^{-1}$ ということです.

複素行列 A の共役をとる操作を普通の複素数のとき同様に \overline{A} と上にバーを書くことで表します. これにより, 随伴行列 A^* を

$$A^* := {}^t\overline{A} \quad (6.2.35)$$

と定義します. ここで先程の対称行列・直交行列に対応するものがありますから, それを定義しておきましょう. $A^* = A$ となる行列を Hermite (エルミート) 行列, $A^*A = I$ となる行列をユニタリ行列と呼びます.

また対角行列なるものがあります. 一般に a_{ii} の形の成分を対角成分といいます. 対角行列というのは対角成分しか持たない行列のことです. すなわち $A = (a_i \delta_{ij})$ という行列です. はじめに書いた運動方程式を X で書き直したものを行列で表現すると対角行列になります.

$$m \begin{pmatrix} \ddot{X}_1 \\ \ddot{X}_2 \\ \ddot{X}_3 \end{pmatrix} = -k \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 - \sqrt{2} & 0 \\ 0 & 0 & 2 + \sqrt{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

次に上三角行列を定義します. 上三角行列というのは行列の対角成分より下, a_{ij} の $j < i$ となるところが 0 となる行列です. 具体的にこう書くと分かるでしょう.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} \quad (6.2.36)$$

下三角行列も同様に定義します.

行列 A が正規行列であるということを次の式で定義します。これは対角化のところで用います。

$$AA^* = A^*A \quad (6.2.37)$$

上の条件は A と A^* が交換可能であるということであり、Hermite 行列やユニタリ行列はこれを満たします。行列が Hermite, ユニタリであるというのはそれぞれ $A^* = A$, $A^* = A^{-1}$ が成立するということでした。実数の範囲で考えれば、これは対称行列・直交行列のことです。

次は部分 (線型) 空間という概念です。 W が C^n の部分空間とすると、任意の W の元 x, y と任意の複素数 α, β に対し

$$\alpha x + \beta y \in W \quad (6.2.38)$$

が成立します。これだけ分かりづらいでしょうが、具体例としては R^3 内の平面や直線などです。

重要な注意ですが、部分集合とは違います。例えば平面内の半径 1 の円板は部分集合ですが、部分空間ではありません。 x を円板内の任意のベクトルとしましょう。 $\alpha \geq 1/|x|$ としてみましょう。この α に対し αx は円板内にいません。平面内の線分などでも駄目です。円板内の場合と同じようにある程度大きい値に対してはみだしてしまいます。ここから考えて、部分空間というのはかなり大きい集合であることが分かります。

また部分空間 W_1, W_2 が直交するというのは任意の $x_1 \in W_1, x_2 \in W_2$ が直交することを言います。さらに部分空間 W_1 の直交補空間を W_1^\perp ¹⁸ とすると、 W_1^\perp とは W_1 と直交する元全ての集合です。さらに W が W_1 と W_2 の直和であるとは、 $W = W_1 \cup W_2$ であってかつ $W_1 \cap W_2 = \{0\}$ となることをいいます。このとき $W = W_1 \oplus W_2 = W_1 \amalg W_2$ などと書きます。

以上でつまらない定義の羅列は一旦終わりです。ここから少し線型代数の説明に入ります。

線型代数で重要な概念の 1 つに 線型写像 があります。それでは線型写像の説明を、と行きたいところですがその前に関数と関数値の区別をはっきりさせておきます。

$$1 \longrightarrow 1, \quad 2 \longrightarrow 4, \quad 3 \longrightarrow 9 \quad (6.2.39)$$

という対応関係があったとしましょう。矢印の左にある数字が右側では 2 乗されていますこのときの対応を関数の形で書くと当然 $y = f(x) = x^2$ となります。関数値というのは x に対して対応させられている値 $y = f(x)$ のことです。それでは関数は、というと、この対応関係を与えるものであり、 f が関数です。また、上の式を実数 x に関数 f が作用して実数 x^2 になったと読んでもいいでしょう。このとき関数 f に作用という視点が与えられます。ここからこのような対応 f を作用素と呼ぶことがあります。「作用」を与える「素」だから「作用素」ですあるものが何らかの作用を受けると当然何か他のものになります。つまり作用は変換ということもできます。こうした言い換えは良く使いますから、慣れ親しんでおいて下さい。

現時点で関数と関数値の区別の重要性を理解するのはなかなか難しいと思います。3 年で学ぶ (人もいる) Lebesgue 積分論や関数解析学において関数を元とする集合 (関数空間) を考えねばならなくなります、そのときにこの区別をきち

¹⁸ 気軽に空間と書きましたが、これは本当に部分空間になっているでしょうか? $x, y \in W^\perp, z \in W$ としましょう。 $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z) = 0$ で、確かに部分空間になっています。

んとする必要がでてきます。要するに量子力学の数理です。あとで関数と関数値の区別の為の具体例を出します。

さて、ここからやっとな線型写像です。 f を R^n から R^m への写像 (関数) であるとしましょう。この f が線型写像であるというのは次の 2 つの性質が成立することを言います。

$$f(x+y) = f(x) + f(y) \quad (6.2.40)$$

$$f(\alpha x) = \alpha f(x), \alpha \in R \quad (6.2.41)$$

またはこれらを一本にまとめて

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y), \alpha, \beta \in R. \quad (6.2.42)$$

ここで $\alpha \in R$ というのは α が実数であることを表すのでした。まず f を R^1 から R^1 への写像、すなわち高校で学んできた普通の関数である場合でこの定義が何を言っているのかを見てみましょう。線型写像の名が示すとおり、このとき α を適当な実数とすると f は $f(x) = \alpha x$ という対応を与える一次関数 (直線!) です¹⁹。線型写像の「線型」たる所以が分かったかと思います。一般の n, m に対しては適当な行列 A を選ぶと $y = Ax$ と書けます。実数を 1 次正方行列とみなせば、自然な一般化になっていることが分かります。

すぐあとで使う一般の写像 (関数) に対する性質も述べておきます。 $f(x, y) = -f(y, x)$ という性質を満たす写像のことを交代性があるといいます。例えば $f(x, y) = x - y$ が一番分かりやすいでしょうか。高校で学んだ二次の行列の行列式にもこの交代性があります。

線型写像と行列には密接な関連があることさえ把握しておけば、今は十分です。細かいことは気にせず今は突っ走ってみましょう。

6.2.3 線型空間の導入

物理法則とベクトル

本章は主に行列に関することを述べるのですが、その前に線型代数で最も重要な概念である線型空間 (ベクトル空間) について議論することにします。現段階では振動・波動との直接の関係は少し薄れますが、これは量子力学の数理の基礎となる非常に重要なものであり、そこまで行かずとも物理学にとって基本的で重要な概念です。

私達が物理学を学ぶ、研究するというとき中心となるのは、やはり物理法則に関する理解・考察を深めることでしょう。しかし物理法則というものは一体何のことを指すのでしょうか? 抽象的な言い方になってしまいますが、これは人間が何らかの形で自然を理解できたとき、それを言葉や数式 (方程式) で表現したもので表現できるでしょう。今考えたいのは、「数式で表現」という部分です。物理学に限らず自然科学の理論は実験結果との整合性がなければなりません。また実験結果というとき、例えば同じ日の同じ時間に日本で彗星の運動を観察した結果とアメリカで彗星の運動を観察した結果や、適当に条件を合わせたときにある日に実験した結果とその 3 ヶ月後に実験した結果などが一致するようなものでないと困ります。ある理論が自然現象を記述できているというならば、この実験結果

¹⁹(進んだ注) f が連続であるとき、上の条件を満たす関数が本当に一次関数しかないことが実数の連続性を用いて証明されます。 $f(1) = 3$ などという適当な条件をつけると、定数 α までは完全に決定できます。証明は島内剛一、「数学の基礎」、日本評論社、p445- を参照して下さい。大略を述べると、自然数に対して成立することを述べたのち、整数に拡大して、さらに有理数に拡大し、 f の連続性から実数に拡大します。

が再現できなければなりません。つまり適当な条件が満たされている限り、任意の観測時間のずれと観測点のずれに対して、同じ実験結果を与えられるようなものでなくてはなりません。²⁰

これは物理法則と呼ばれるものの性質を規定する条件です。つまり物理法則は、現象を観測する時間と位置によらず成立するべきであるということです。物理法則は方程式で表現されることがありますが、当然この方程式自体も観測する位置と時間によらず成立するようなものでなければいけません。こうなると実際に我々が物理法則を方程式の形にすると、方程式に対して何らかの制限を加えないといけません。その制限とは何かを考えます。私達は物理現象を調べる際に直交座標・極座標など様々な座標系（基底）を用いますが、物理法則そのものの意味が座標系によって変わってしまっては困ります。ここで「意味が変わらない」ということをきちんと考える必要があります。例えば運動方程式は言葉で書けば「運動量の時間変化は加わる力に比例し、この力の方向に起こる」ということであり、数式で書けば $m\ddot{\mathbf{r}} = \mathbf{f}$ と定式化されます。具体的な座標系で加速度 $\ddot{\mathbf{r}}$ を考えましょう。これを直交座標で表現するか極座標で表現するかで

$$\ddot{\mathbf{r}} = (\ddot{x}, \ddot{y}, \ddot{z}) = \left(\frac{dv}{dt}, \frac{v^2}{\rho}, 0 \right) \quad (6.2.43)$$

のように加速度の「表現」の仕方が変わりますが、加速度であることそのものは変化しません。この例で重要なのは、運動方程式自体は具体的な座標表示によらずに書かれているということです。また一般に質量、加減度、力のような物理的に意味のある量を物理量と呼びますが、物理法則は何らかの形でこれら物理量の間の関係を述べたものとして定式化されるでしょう。上の例を見ればこれら物理量自体が具体的な座標によらずに書かれています。そうするとそもそも物理的に意味のある量（物理量）とは具体的な座標表示によらない意味を持つものである、ということまで要求しなければならないかもしれません。

以上のような考察を経て、物理法則に対して以下のような要請をします。

Request. 6.2.1. ある方程式がどの慣性系で見ても同じ形の方程式を満たすとき、その方程式は座標の変換に対し共変性を持つという。全ての物理法則は共変性を持つ。

この要請に沿えば物理法則が物理量の間の関係である以上、物理量も共変性を持たねばなりません。私達はこれら物理法則を実際に方程式にしなければならぬのですが、その際自動的に共変性が満たされているような上手い数学的な表現法があれば非常に便利です。答から言うならば、それはベクトル（一般にはテンソル）によって物理法則を表現すれば良いことが分かっています。

抽象ベクトル

ここでベクトルについて考察を深めます。上で「物理現象を調べる際に様々な座標系を用いる」と書きましたが、ベクトルによる物理法則の記述が座標系によらない意味を持つならば、ベクトルという概念自体座標系によらないものでなければいけません。つまりベクトルは座標系とは関係なく存在するものであるべきです。ここで座標系の入っていない、ベクトルのみが存在する空間を（抽象）ベクトル空間と呼び、これを V と書きましょう。これは法則が成立するべき空間（抽象ベクトル空間）と、私達が実際に現象の解析を行なう空間をはっきり区別した、ということです。ここまでは抽象ベクトルという言葉の意味（抽象ベクトルの定義）

²⁰ もっと上手い言い方を模索中。そしてこれは私自身指摘されて気づいたことですが、観測時間・観測点のずれに対して同じ結果を与えるということ、任意の時間と位置で成立するということは違います。

をはっきりさせていません。これから抽象ベクトルの持つべき性質を考え、矛盾のないような抽象ベクトルの理論を作ります。重要なのは抽象ベクトルとは何か、抽象ベクトルがどんな性質を持っていて欲しいのかは自分で考えて自分で決めるものであり、また決めてよいということです。

まず記述に必要な言葉を用意します。高校までは平面（空間）上の点を表すのに xy -平面というものを考えていました。これと対応する形で、現象を調べる際には適当に座標原点を取った上で x 軸や y 軸、 z 軸を設定していました。このとき、空間内のある点 r の座標が (x, y, z) であるというのは、

$$e_1 = e_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = e_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = e_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (6.2.44)$$

としたとき、

$$r = xe_1 + ye_2 + ze_3 \quad (6.2.45)$$

と表せる、ということです。このとき、 $\langle e_1, e_2, e_3 \rangle = \langle e_i \rangle_{i=1}^3$ を座標系（基底）、 x, y, z を（座標系に対する）座標（座標成分）といいます。このとき上の r が（抽象）ベクトル空間の元であるとする、これを適当な座標系で表現した右辺は普通の空間ベクトルです。空間ベクトルには自然に備わった（高校で学んだ）和（差）とスカラー倍（定数倍）という演算があります。右辺の空間ベクトルが和とスカラー倍のできるものである以上、左辺も和とスカラー倍ができないといけません。数学者が色々解析した結果、以下の 8 つの代数的な関係式が満たされていなければならないベクトルの和とスカラー倍も普通の空間ベクトルと同じように成立することが分かりました。これが抽象ベクトルの代数です。

1. $x + y = y + x$
2. $(x + y) + z = x + (y + z)$
3. 全ての x に対し、 $x + 0 = x$ を満たすベクトル 0 がただ 1 つ存在する。
4. 全ての x に対し、 $x + x' = 0$ を満たすベクトル x' がただ 1 つ存在する。
5. $1x = x$
6. $\alpha(\beta x) = (\alpha\beta)x$
7. $\alpha(x + y) = \alpha x + \alpha y$
8. $(\alpha + \beta)x = \alpha x + \beta x$

これで抽象ベクトルの代数的な性質を確立しました。その他に現象の解析に必要なのは、ベクトルの長さと 2 つのベクトルのなす角度です²¹。高校ではこれらが内積を用いて表現できることを学びました。ただし高校では $x = x_1e_1 + x_2e_2 + x_3e_3$ 、 $y = y_1e_1 + y_2e_2 + y_3e_3$ としたときの

$$x \cdot y = (x, y) = x_1y_1 + x_2y_2 + x_3y_3 \quad (6.2.46)$$

という形の内積しか考えていません。これを座標系（と原点）に依存しない形で書くことを考えなければいけません。このためには内積で表現すべき、長さと 2 つのベクトルのなす角度を基点にします。長さ・角度共にどの座標系を取るかに関わらず一定であってもらわないと困ります。特に長さは自分自身との内積 (x, x) で表現されますが、この値が必ず正になってもらわないと困ります。具体的な座標系を取ったときに内積に必要な演算が全て正当化されねばならないという要請をしたとき、内積の性質は次のように規定すればよいことが分かっています。

²¹例えば万有引力は 2 点の距離の 2 乗に反比例します。角度についてはフレミングの左手の法則など、各種「向き」の関わる法則を記述する際に重要です。

- 1) $(x, x) \geq 0$
- 2) $(x, x) = 0 \iff x = 0$
- 3) $(x, y) = (x, y)$
- 4) $(\alpha x_1 + \beta x_2, y) = \alpha(x_1, y) + \beta(x_2, y)$

ただしこれで本当に抽象ベクトルの長さと 2 つの抽象ベクトルのなす角度が座標系と原点によらず定まるのか、というのは当然の疑問です。そして単に変換というとき、数学的にはスカラー倍といった変換も含まれますが、この変換に対して 2 つのベクトルのなす角度が不変であっても長さは変わるというのは直観的に分かるでしょう。長さ・2 つのベクトルのなす角度を同時に不変にする変換それ自体も考察しないとはいけません。少し話が本筋からそれてしまいましたが、先にこれを考えます。まず変換というものは、何らかの作用を受けた結果であると考えることができます。作用は適当な写像 (関数) で表現できるということを前節で説明しました。長さ、角度を同時に不変にする変換は特殊な線型変換で表現できます。線型変換とは線型写像で表される変換です。また線型写像は行列で表現出来るということが証明できます。定理の証明前に予備考察を行ないます。まず長さを変えてはいけなくて、いわゆる拡大・縮小に対応する変換ではいけません。そのほか角度というところから回転という変換が思いつきます。実際この回転が長さと 2 つのベクトルのなす角度を変えない変換です。拡大・縮小のない回転で長さは当然不変であり、同時に同じだけ 2 つのベクトルを回転させれば角度も当然変換に対して不変になります。

回転の表現

そこでこれから回転を数理的に定式化します。下調べとして平面で考えましょう。角度 θ だけ回す回転を表す写像を R とし、適当に $x, y \in \mathbb{R}^2$ を考えます。標準基底に関して R を行列 \hat{R}_θ で表現すると

$$\hat{R}_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (6.2.47)$$

となります。これから

$$\hat{R}_\theta x = \begin{pmatrix} x_1 \cos \theta + x_2 \sin \theta \\ -x_1 \sin \theta + x_2 \cos \theta \end{pmatrix}, \quad \hat{R}_\theta y = \begin{pmatrix} y_1 \cos \theta + y_2 \sin \theta \\ -y_1 \sin \theta + y_2 \cos \theta \end{pmatrix} \quad (6.2.48)$$

$$\hat{R}_\theta(x + y) = \begin{pmatrix} (x_1 + y_1) \cos \theta + (x_2 + y_2) \sin \theta \\ -(x_1 + y_1) \sin \theta + (x_2 + y_2) \cos \theta \end{pmatrix} \quad (6.2.49)$$

と書けます。これから平面の場合には回転が行列で表示でき、したがって線型写像で表現できることが分かります。ここでは省略しますが、一般の次元でも同様に示せます。そこで問題は回転がどのような行列で表現できるか、というところ です。結果からいうと、これは直交行列です。これを証明します。証明自体は一般の (複素化した) ユニタリ行列に対して行ないます。ここで複素ベクトルを考えたときの内積は $(x, y) := \sum_{i=1}^3 x_i \bar{y}_i$ で定義します。

Theorem. 2.3.1. n 次正方行列 A に関する次の 4 条件は同値である。

- 1) A はユニタリ行列である。
- 2) 任意の n 項 (複素) 列ベクトル x に対して $|Ax| = |x|$ が成立する。
- 3) 任意の n 項 (複素) 列ベクトル x, y に対して $(Ax, Ax) = (x, y)$ が成立する。
- 4) $A = (a_1, \dots, a_n)$ とすれば、 $(a_i, a_j) = \delta_{ij}$ である。

$$(証明) 1) \Rightarrow 2) |Ax|^2 = (Ax, Ax) = {}^t x {}^t A \overline{A} x = {}^t x x = |x|^2.$$

1. \Rightarrow 3) 仮定より次の式が成立します.

$$|x + y|^2 = |x|^2 + (x, y) + \overline{(x, y)} + |y|^2 \quad (6.2.50)$$

$$|A(x + y)|^2 = |Ax|^2 + (Ax, Ay) + \overline{(Ax, Ay)} + |Ay|^2 \quad (6.2.51)$$

これと条件 2) から, $(x, y) + \overline{(x, y)} = (Ax, Ay) + \overline{(Ax, Ay)}$ が成立し, したがって (x, y) と (Ax, Ay) の実部が等しいことが分かります. 一方, x の代わりに ix を代入すれば, $i\{(x, y) - \overline{(x, y)}\} = i\{(Ax, Ay) - \overline{(Ax, Ay)}\}$ によって (x, y) と (Ax, Ay) は虚部も等しいことが分かります. したがって $(Ax, Ax) = (x, y)$.

1. \Rightarrow 1) 任意の x, y に対して

$$(x, (A^*A - I)y) = (x, A^*Ay) - (x, y) = (Ax, Ay) - (x, y) = 0. \quad (6.2.52)$$

したがって $A^*A - I = O$ となります.

1. \Leftrightarrow 4)

$A^*A = (c_{ij})$ とすれば $c_{ij} = \overline{(a_i, a_j)}$ であり, これから同値はすぐ分かります. ■

この定理によって長さ 2 のベクトルのなす角度を変えない変換, すなわち回転は直交行列で表現できることが分かりました. 上で述べたように行列の作用は線型写像に翻訳できます. 線型写像は行列と異なり, 座標系の取り方に依存しません. ここで一般に, 内積を変えない, つまり $(T_O(x), T_O(y)) = (x, y)$ となる線型変換 T_O を直交変換と呼びます.

Galilei の相対性原理

ここで物理法則の記述という問題に帰ります. 現象の解析を行なう際には適当な座標系 (と座標原点) を取ると書きました. さらにはじめにも例を挙げたように, 現実には時間原点の設定もしないといけません. 時間も含め座標系の選択にはかなりの自由度がありますから, これら異なる座標系の間の変換則が欲しいところです. 抽象的に物理法則, といっているのははじまらないので, ここでは力学の基本法則たる運動方程式とこれを不変にする変換を考えます. 力学は大概のことが運動方程式から導かれるので, 運動方程式の形不変性 (form invariance) が最も重要です.

考えるのが座標系と時間の変換ですから, まず Newton 力学の基礎にある時空間に関する概念を確認しなければなりません. 1 つ言葉を定義しておきます. 物理現象は常にある時にある場所で起こります. ある時刻に空間の 1 点で起こる出来事を事象と呼びます.

1) 任意の 2 個の事象の時間間隔はどの座標系で見ても同一である.

2) 任意の 2 個の同時刻の事象の間の距離はどの座標系で見ても同一である.

性質 1) は座標系 S と S' のそれぞれと共に動く時計で測定された任意の事象 A, B の時間間隔が $t_B - t_A = t'_B - t'_A$ であることを意味しています. これが常に成立する為には

$$\Delta t' = \Delta t \iff t' = t - \theta \quad (\theta = \text{const.}) \quad (6.2.53)$$

でなければなりません. ここから $\Delta t' = 0 \Leftrightarrow \Delta t = 0$ となり, 2 つの事象の同時性は座標系の選択に依存しません. 後でも使いますが, このとき微分作用素 d/dt は

$$\frac{d}{dt} = \frac{dt'}{dt} \frac{d}{dt'} = \frac{d}{dt'} \quad (6.2.54)$$

という変換を受けます.

同様に性質 2) は事象 A, B が同時刻であるとき, その間の距離が $|r_A - r_B| = |T_O(r_A) - T_O(r'_B)| (= |r'_A - r'_B|)$ であることを意味します. これが常に成り立つ為には, 上で示したように直交変換 T_O を用いて

$$\Delta r = T_O(\Delta r') (= \Delta r') \quad (6.2.55)$$

でなければなりません.

さて, Newton 力学の原理は運動方程式 (運動の第 2 法則)

$$m\ddot{r} = F \quad (6.2.56)$$

で与えられます. ここで $F = \sum f$ と表され, f は着目している物体に他から働く力で実体的な (物理的) な起源があり, F はその合力です. 特に他の諸物体から十分遠くに離れた物体 (孤立した物体) に対しては事実上 $F = 0$ としてよく, このとき

$$m\ddot{r} = 0 \implies \dot{r} = v_0 (\text{定ベクトル}) \quad (6.2.57)$$

となります. すなわち, 外から力の働いていない物体は同一の運動状態 (等速度運動) にあり続けます. 通常このことを慣性の法則と呼んでいます. そして慣性の法則が成立する座標系を慣性座標系, または簡単に慣性系と呼びます. ここで慣性の法則の独立した論理的な意味は

Request. 6.2.2. 宇宙には, 力を受けていない物体が等速度運動を続ける座標系 (慣性座標系) が少なくとも 1 つ存在する.

という慣性系の存在要請と考えるべきです. そうしないと, 慣性の法則が運動方程式の特別な場合として運動の第 2 法則に含まれてしまいます. 慣性の法則が運動の第 1 法則である理由です. 運動の第 3 法則は作用・反作用の法則です.

ここで時間と空間の性質として上記の 2 つの仮定を認めた上での慣性系どうしの座標変換を Galilei 変換 (Galilean transformation) といいます. Galilei 変換 $r \mapsto r', t \mapsto t'$ では $d/dt = d/dt'$ であることから

$$\ddot{r} = 0 \iff \frac{d^2}{dt'^2} T_O(r) = 0 \quad (6.2.58)$$

とならなければなりません. 座標系の変換で長さ, 2 つのベクトルのなす角度が変わってはならないので, \ddot{r} から \ddot{r}' への線型変換は直交変換になります. つまり直交変換 T_O を用いて

$$\ddot{r} = T_O(\ddot{r}') = \frac{d^2}{dt'^2} (T_O(r)) = \ddot{r}' \quad (6.2.59)$$

と書けます. 積分すれば

$$r = T_O(r') - ut - \alpha \quad (6.2.60)$$

となります。ここで \mathbf{u}, α は定ベクトルです。

特に $T_O = I$ (恒等写像), $\alpha = \mathbf{0}, \theta = 0$ のとき

$$\mathbf{r} \mapsto \mathbf{r}' = \mathbf{r} - \mathbf{u}t, \quad t \mapsto t' = t \quad (6.2.61)$$

となります。これはもとの座標系に対して速度 \mathbf{u} で動いている座標系への変換であり、これを Galilei ブーストといいます。Galilei ブーストに伴う速度の変換則は

$$\mathbf{v} = \dot{\mathbf{r}} \mapsto \mathbf{v}' = \dot{\mathbf{r}}' = \mathbf{v} - \mathbf{u} \quad (6.2.62)$$

で与えられます。この結果は当然、と思うかもしれませんが、Newton 力学での時空間に対する仮定と変換則に支えられて得られたものである、つまり相対論的な力学では破壊される変換則であることを注意しておきます。せっくなので相対論的な速度の変換則を紹介しておきます。

$$\mathbf{v}' = \frac{d\mathbf{r}'}{dt'} = \frac{\sqrt{1 - \frac{v^2}{c^2}} \mathbf{v}_\perp + \mathbf{v}_\parallel - \mathbf{u}}{1 - \frac{\mathbf{u} \cdot \mathbf{v}}{c^2}}. \quad (6.2.63)$$

ここで要請を 1 つおきましょう。

Request. 6.2.3. 全ての慣性系は対等であり、したがって力学法則はどの慣性系で見ても同じでなければならない、つまり Galilei 変換で形を変えてはならない。この性質を Galilei 共変性、そしてこの要請を Galilei の相対性原理という。

これが上で述べた物理法則の形不変性につながっています。

この要請を運動方程式に適用すれば、力は時間に t に陽によらずかつ力は変換則

$$\mathbf{F} = T_O(\mathbf{F}) \quad (6.2.64)$$

にしたがわなければなりません

最後に Galilei 変換 G を標準基底に関して行列表示してみましょう。まず 4 元ベクトル $\mathbf{R} := {}^t(x_0 \ x_1 \ x_2 \ x_3) = {}^t(ct \ x_1 \ x_2 \ x_3)$ と定義します。定数 c は速度の次元を持ってさえいけば良いのですが、一応光の速さとしておきます。ここで非斉次ベクトル $\tilde{\mathbf{R}}$ を $\tilde{\mathbf{R}} = {}^t(\mathbf{R} \ 1) = {}^t(x_0 \ x_1 \ x_2 \ x_3 \ 1)$ と定義します。この記法で Galilei 変換を書き直すと

$$G : \begin{cases} x_0 \mapsto x'_0 = x_0 - \delta \\ \mathbf{r} \mapsto \mathbf{r}' = O\mathbf{r} - \alpha\beta x_0 \end{cases} \quad (6.2.65)$$

となります。ただし $\beta := \mathbf{u}/c, \delta := c\theta$ です。さらに O は \mathbf{r} が適当な座標系で表現されているものとして、直交変換 T_O が直交行列 O で表現しておきました。Galilei 変換を表す行列を \hat{G} とすると、

$$G : \tilde{\mathbf{R}} \mapsto \tilde{\mathbf{R}}' = \hat{G}\tilde{\mathbf{R}} = \left(\begin{array}{c|ccc|c} 1 & 0 & 0 & 0 & -\delta \\ \hline -\beta_1 & & & & -\alpha_1 \\ -\beta_2 & & O & & -\alpha_2 \\ -\beta_3 & & & & -\alpha_3 \\ \hline 0 & 0 & 0 & 0 & -1 \end{array} \right) \tilde{\mathbf{R}} \quad (6.2.66)$$

となります。この計算をした後、 $\tilde{\mathbf{R}}'$ のはじめの 4 つの成分を取れば、めでたく座標変換が得られます。

抽象ベクトルの微分

さて、実はここまで曖昧にぼかしてきた部分があります。それは抽象ベクトルの微分とは何者かということです。まだきちんと議論していない偏微分を使ってしまいうことになってますが、これで例を出します。電場を $E(\mathbf{r}, t)$ とします。Maxwell 方程式から $\nabla \cdot E(\mathbf{r}, t) = \rho(\mathbf{r}, t)/\varepsilon_0$ となりますが、これは $\mathbf{r} = (x, y, z)$ としたとき

$$\frac{\partial E}{\partial x} + \frac{\partial E}{\partial y} + \frac{\partial E}{\partial z} = \frac{1}{\varepsilon_0} \rho(\mathbf{r}, t) \quad (6.2.67)$$

と書くことになります。 $\partial/\partial x$ というのは、他の変数 y, z, t を定数とみなして (固定して)、 x についてだけ微分を取るものです。例えば、 $f(x, y, z, t) = x^2 + xy^3 + \log yz + \sin zt$ としたとき

$$\frac{\partial f}{\partial x} = 2x + y^3, \quad \frac{\partial f}{\partial y} = 3xy^2 + \frac{1}{y}, \quad \frac{\partial f}{\partial z} = \frac{1}{z} + t \cos zt, \quad \frac{\partial f}{\partial t} = z \cos zt \quad (6.2.68)$$

などとなります。ここで具体的に座標系 (基底) を取る前の $\nabla \cdot E(\mathbf{r}, t) = \rho(\mathbf{r}, t)/\varepsilon_0$ はどう考えるべきでしょうか? $\partial/\partial x$ という表記を見ても分かるように、微分というものは本質的に座標の取り方に深く関わるものです。当然本節においては具体的な座標系 (基底) を選択する前の

$$\frac{d}{dt} \mathbf{r} \quad (6.2.69)$$

とは何者か、ということが問題です。さらに微分作用素それ自体も座標系の選択に深く依存しているので、微分作用素の表現自体もどうするか考えなければなりません。

この問題は、数学的には幾何学における多様体論で深く考察される問題です。現在、そもそも (偏) 微分自体が制御不可能なので、本小節では簡単に論じるとどめます。

解決策はとりあえず二つほど考えられます。つは、まず適当な座標系 (基底) を選択し、そのもとで微分を計算します。ここで全ての座標系 (基底) に対して成立するような、つまり座標変換に対して不変な性質を見出し、それを抽象ベクトルの微分の性質とする、という方法です。どういうことが説明しましょう。まず適当な座標系 (基底) を取ったとき、微分の値 $\partial f(x, y, z)/\partial x, \partial f(r, \theta, \varphi)/\partial r$ などには一般には変化します。例えば $f(x, y) := x + y = r \cos \theta + r \sin \theta = f(r, \theta)$ とします。

$$\frac{\partial f(x, y)}{\partial x} = 1, \quad \frac{\partial f(x, y)}{\partial y} = 1, \quad (6.2.70)$$

$$\frac{\partial f(r, \theta)}{\partial r} = \cos \theta + \sin \theta, \quad \frac{\partial f(r, \theta)}{\partial \theta} = -r \sin \theta + r \cos \theta \quad (6.2.71)$$

となります。そもそも微分する変数自体が変わっているのです、当たり前といえば当たり前で、微分する変数を変えずに $\partial f(r, \theta)/\partial x$ を考えれば、これは当然 $\partial f(x, y)/\partial x$ に一致します。ここで、このときに変わらない (であろう) 性質としては、ある座標系 (基底) を取ったときのある点において微分係数が 0 にならないとき、どんな座標系を取ろうともその同じ点においては微分係数が 0 にならない、ということです。もう少し何か書く。作りかけ。

2 つ目ははじめから座標系 (基底) の変換で不変な微分の形式を確立するというものです。これが多様体上で微分を行なう際に重要になる微分の形式で、(外) 微

分形式といいます。これは物理の方では解析力学 (の非常に高度な数学的定式化) で基本的かつ重要な役割を果たすほか、どうも素粒子理論の方でも良く使うようです。普通の微分ですらまともに論じていないので、ここであまり踏み込んだ議論はしません。ただこちらはベクトル解析のところでも多少触れるかもしれません。

ところで肝心の抽象ベクトルの微分はどうするのか、ということですが、現時点では、結局のところ、詳細な解析をする際には (都合のよい、上手い) 具体的な座標系 (基底) を選んで現象の解析を行なう、ということがありますので、第一の方法に従うことにしましょう作りかけ。

定理の証明

最後に積み残した「有限次元の間の線型写像は行列で表現出来る」ということを証明します。まず 1 つ言葉を定義します。ある (m, n) 型行列によって定まる線型写像とは (m, n) 型行列 A に対して

$$f_A(\mathbf{x}) = A\mathbf{x}, \quad \mathbf{x} \in \mathbb{C}^n \quad (6.2.72)$$

として定まる \mathbb{C}^n から \mathbb{C}^m への写像 f_A のことをいいます。まずは標準基底 $\langle \mathbf{e}_i \rangle_{i=1}^n, \langle \mathbf{e}'_i \rangle_{i=1}^m$ に関して定式化します。

Theorem. 2.3.2. \mathbb{C}^n から \mathbb{C}^m への写像が線型写像であるための必要十分条件は、写像がある (m, n) 型行列 A によって定まることである。またこのとき A はただ 1 つに定まる。

(証明) 十分性の方は

$$f_A(a\mathbf{x} + b\mathbf{y}) = A(a\mathbf{x} + b\mathbf{y}) = a(A\mathbf{x}) + b(A\mathbf{y}) = aT_A\mathbf{x} + bT_A\mathbf{y} \quad (6.2.73)$$

からすぐに分かります。

必要性を示しましょう。一般の n 次でも大した手間はないので、一般の状況で示します。 $f: \mathbb{C}^n \rightarrow \mathbb{C}^m$ が線型写像であるとしましょう。 \mathbb{C}^n の座標系を n 項単位ベクトルの組 $\langle \mathbf{e}_i \rangle_{i=1}^n$ (標準基底) に取ります。これらが f で写される先を \mathbf{a}_j とします。すなわち

$$f\mathbf{e}_j = \mathbf{a}_j, \quad (j = 1, 2, \dots, n) \quad (6.2.74)$$

とします。 $\mathbf{a}_j \in \mathbb{C}^m$, つまり \mathbf{a}_j は m 項列ベクトルです。ここで (m, n) 型行列 A を $A := (\mathbf{a}_1, \dots, \mathbf{a}_n)$ として定義し、 A によって定まる \mathbb{C}^n から \mathbb{C}^m への線型写像を T_A とすれば、

$$f_A(\mathbf{e}_j) = A\mathbf{e}_j = \mathbf{a}_j = f\mathbf{e}_j. \quad (6.2.75)$$

\mathbb{C}^n の任意のベクトル \mathbf{x} は、

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \quad (6.2.76)$$

と表せますから、

$$f_A(\mathbf{x}) = f_A\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) = \sum_{i=1}^n x_i f_A(\mathbf{e}_i) \quad (6.2.77)$$

$$= \sum_{i=1}^n x_i f(\mathbf{e}_i) = f\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) = f(\mathbf{x}). \quad (6.2.78)$$

したがって $f = f_A$ です.

次に一意性を示します. (m, n) 行列 B が存在して任意の $x \in \mathbb{R}^n$ に対し $f(x) = Bx$ となるとします. このとき

$$a_j = Ae_j = fe_j = Be_j = b_j (j = 1, \dots, n) \Rightarrow B = A. \blacksquare \quad (6.2.79)$$

上では標準基底で示しましたが, それ以外の適当な座標系を取ったときを考えます. まず予備定理を証明します.

Theorem. 2.3.3. $f: \mathbb{C}^n \rightarrow \mathbb{C}^m$ を線型写像, $\mathbb{C}^n, \mathbb{C}^m$ の新たな座標系をそれぞれ $\langle u_i \rangle_{i=1}^n, \langle v_j \rangle_{j=1}^m$ とする. さらに $\xi := {}^t(\xi_1 \cdots \xi_n)$, $\eta := {}^t(\eta_1 \cdots \eta_m)$ とする. このとき (m, n) 型行列 B がただ 1 つ存在し, 次の関係が成立する.

$$(f(u_1) \cdots f(u_n)) = (v_1 \cdots v_m)B, \quad \eta = B\xi. \quad (6.2.80)$$

(証明) $x = \xi_1 u_1 + \cdots + \xi_n u_n$ の f による像は $f(x) = \xi_1 f(u_1) + \cdots + \xi_n f(u_n)$ となります. そこで $f(u_j)$ の $\langle v_j \rangle_{j=1}^m$ に関する座標成分を

$$b_j = {}^t(b_{1j} \cdots b_{mj}) \iff f(u_j) = (v_1 \cdots v_m)b_j \quad (j = 1, \dots, n) \quad (6.2.81)$$

として $B := (u_1 \cdots u_n)$ とするとこれは (m, n) 行列で次が成立しています.

$$(f(u_1) \cdots f(u_n)) = (v_1 \cdots v_m)B. \quad (6.2.82)$$

そして

$$f(x) = \xi_1 f(u_1) + \cdots + \xi_n f(u_n) = (f(u_1) \cdots f(u_n))\xi \quad (6.2.83)$$

$$= (v_1 \cdots v_m)B\xi \quad (6.2.84)$$

一方

$$f(x) = \sum_{j=1}^m \eta_j v_j = (v_1 \cdots v_m)\eta \quad (6.2.85)$$

でありこれが上の式と等しく, $\langle v_j \rangle_{j=1}^m$ が座標系であって, $f(x)$ の座標成分は一意に決まっているので

$$\eta = B\xi. \quad (6.2.86)$$

これから B の一意性を証明します. (m, n) 型行列 C が存在し, $\eta = C\xi$ を満たすとします. このとき $x = u_j$ とすると $\xi = e_j (j = 1, \dots, n)$ となります. $f(u_j) = \eta_{1j} v_1 + \cdots + \eta_{mj} v_j$ とすると,

$$b_j = Be_j = {}^t(\eta_{1j} \cdots \eta_{mj}) = Ce_j = c_j \quad (j = 1, \dots, n) \quad (6.2.87)$$

$$\Rightarrow B = C. \quad \blacksquare \quad (6.2.88)$$

Remark. 2.3.4. 実は本定理で現れた B は座標系として $\langle u_i \rangle_{i=1}^n, \langle v_j \rangle_{j=1}^m$ を取ったときに線型写像 f を表現する行列になっています.

それでは本題に入りましょう.

Theorem. 2.3.5. \mathbb{C}^n から \mathbb{C}^m への写像 f が線型写像であるための必要十分条件は, 任意に座標系 (基底) を定めたとき写像がある (m, n) 型行列によって定まることである. またこのとき A は各座標系に対したただ 1 つに定まる.

(証明) まず $\langle e_i \rangle_{i=1}^n \rightarrow \langle u_i \rangle_{i=1}^n, \langle e'_j \rangle_{j=1}^m \rightarrow \langle v_j \rangle_{j=1}^m$ の座標系の変換をあらわす行列をそれぞれ P, Q とすると, 両方共に座標系であることから変換したきり逆に戻せないということはなく, きちんと逆変換が存在することから, 対応して P, Q には逆行列が存在します. これは

$$(u_1 \cdots u_n) = (e_1 \cdots e_n)P, \quad (v_1 \cdots v_m) = (e'_1 \cdots e'_m)Q \quad (6.2.89)$$

と書けます. さらに $X := {}^t(x_1 \cdots x_n), Y := {}^t(y_1 \cdots y_m)$ を標準基底での x, y の座標成分 (このようなベクトルを座標ベクトル) とします. すぐ上の注意から

$$Y = AX \quad (6.2.90)$$

となります. $x, f(x)$ を各座標系で表現すると

$$x = (e_1 \cdots e_n)X = (u_1 \cdots u_n)\xi \quad (6.2.91)$$

$$f(x) = (e'_1 \cdots e'_m)Y = (v_1 \cdots v_m)\eta \quad (6.2.92)$$

のようになります. 以上から

$$f(x) = f((u_1 \cdots u_n)\xi) = f((e_1 \cdots e_n)P\xi) = AP\xi, \quad (6.2.93)$$

$$f(x) = (v_1 \cdots v_m)\eta = (e'_1 \cdots e'_m)Q\eta = Q\eta \quad (6.2.94)$$

定理の前の注意から $\eta = B\xi$ であり, P^{-1} が存在することから

$$AX = QBP^{-1}X \quad (6.2.95)$$

となります. ここで $x = e_j$ とすると $X = e_j$ ですから,

$$A \text{ の第 } j \text{ 列} = Ae_j = (QBP^{-1})e_j = QBP^{-1} \text{ の第 } j \text{ 列} \quad (6.2.96)$$

となります. これから

$$A = QBP^{-1} \iff B = Q^{-1}AP \quad (6.2.97)$$

が成立します. これから任意の座標系で線型写像は行列で表現でき, その行列は標準基底で線型写像を表現した行列と座標系の変換 (基底変換の行列) で表されることが分かりました. ■

6.2.4 行列式

前節で述べたように線型変換は座標系 (基底) を選んだ上で一意的に行列で表すことが出来ます. 物理法則の形不変性 (form invariance) の観点からみても, ある (線型の) 座標変換に逆変換が存在するかどうか調べることは重要なことです. これは行列で線型変換を具体的に表して, その行列に逆行列があるか否かを調べれば良さそうです. というわけで, いくつかある逆行列の存在・非存在の判定法のうち, ここでは行列式を用いたものを紹介します.

高校でも 2 次正方行列の行列式を学びました. よく使う行列式の記法を確認しておきます. 行列 $A = (a_{ij}) = (a_1 \ a_2 \ a_3)$ に対し

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad |a_1 \ a_2 \ a_3|, \quad \det(a_1 \ a_2 \ a_3) \quad (6.2.98)$$

などと書きます。唐突ですが、2 次の行列式には以下のような性質がありました。暗算でも確かめられるでしょう。

$$\det(\mathbf{a}_1 \mathbf{a}_2) = -\det(\mathbf{a}_2 \mathbf{a}_1) \quad (6.2.99)$$

$$\det(\alpha \mathbf{a}_1 + \beta \mathbf{a}'_1 \mathbf{a}_2) = \alpha \det(\mathbf{a}_1 \mathbf{a}_2) + \beta \det(\mathbf{a}'_1 \mathbf{a}_2) \quad (6.2.100)$$

$$\det(\mathbf{a}_1 \alpha \mathbf{a}_2 + \beta \mathbf{a}'_2) = \alpha \det(\mathbf{a}_1 \mathbf{a}_2) + \beta \det(\mathbf{a}_1 \mathbf{a}'_2) \quad (6.2.101)$$

$$\det(\mathbf{e}_1 \mathbf{e}_2) = |E| = 1 \quad (6.2.102)$$

この性質を一般の n 次行列式にも要求します。3 次でいくつか具体的に書くと

$$\det(\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3) = -\det(\mathbf{a}_2 \mathbf{a}_1 \mathbf{a}_3) \quad (6.2.103)$$

$$\det(\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3) = -\det(\mathbf{a}_1 \mathbf{a}_3 \mathbf{a}_2) \quad (6.2.104)$$

$$\det(\alpha \mathbf{a}_1 + \beta \mathbf{a}'_1 \mathbf{a}_2 \mathbf{a}_3) = \alpha \det(\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3) + \beta \det(\mathbf{a}'_1 \mathbf{a}_2 \mathbf{a}_3) \quad (6.2.105)$$

$$\det(\mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3) = |E| = 1 \quad (6.2.106)$$

などです。前節で説明した言葉を使ってこの性質をまとめると以下ようになります。

Definition. 2.4.1. 行列式とは定義域が $M(n, n; C^n)$ で値域が C の、単位行列に対し 1 を対応させる、列に関する交代性を持った多重線型写像である。

ここで $M(n, m; C^n)$ とは $n \times m$ 型行列全体の集合でした。多重線型性というのは、行列式は多変数写像 (関数) ですが、この全ての変数 (全ての列) に対し線型性を要求するということです。この性質を用いて n 次行列式の持つ性質をいくつか調べて、具体的に計算できるようにします。

最終目標をはじめに示しておきます。

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \quad (6.2.107)$$

これが出来れば、2 次の行列式の計算は知っているわけですから、3 次の行列式が計算できるというわけです。4 次での対応する結果を書きおくのは無駄ではないでしょう。下に記しておきます。

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} \quad (6.2.108)$$

$$= (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{vmatrix} + (-1)^{2+1} a_{21} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \end{vmatrix} \quad (6.2.109)$$

$$+ (-1)^{3+1} a_{31} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \end{vmatrix} + (-1)^{4+1} a_{41} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{vmatrix} \quad (6.2.110)$$

まず次の定理を示しましょう。

Theorem. 2.4.2. 二つの列が一致すれば $\det A = 0$ となる。

(証明) これは交代性からほとんど明らかです. 確認してみましょう. 例えば $a_1 = a_2$ とすると交代性から

$$\det(a_1 \ a_1 \ a_3) = -\det(a_1 \ a_1 \ a_3) \implies \det(a_1 \ a_1 \ a_3) = 0 \quad (6.2.111)$$

他の列の場合でも同様に成立することが示せます. ■
次の定理にいきましょう.

Theorem. 2.4.3. ある列に他のある列の定数倍を加えて得られる行列の行列式は, もとの行列の行列式 $\det A$ に等しい.

意味が分からないかもしれないので, 何を示せばよいのかをまず確認しましょう. 式で書くと

$$\det(a_1 \ a_2 + ca_1 \ a_3) = \det(a_1 \ a_2 \ a_3). \quad (6.2.112)$$

といったことです.

(証明) 多重線型性から

$$\det(a_1 \ a_2 + ca_1 \ a_3) = \det(a_1 \ a_2 \ a_3) + c \det(a_1 \ a_1 \ a_3) \quad (6.2.113)$$

前定理から $\det(a_1 \ a_1 \ a_3) = 0$ なので

$$= \det(a_1 \ a_2 \ a_3) \quad (6.2.114)$$

先程同様, 他の列あるいは行の場合でも同じです. ■

次に示すのは

Theorem. 2.4.4. 転置行列と元の行列式の行列式の値は等しい.

つまり,

$$\det {}^t A = \det A \quad (6.2.115)$$

すなわち

$$\det \begin{pmatrix} {}^t a_1 \\ {}^t a_2 \\ {}^t a_3 \end{pmatrix} = \det(a_1 \ a_2 \ a_3) \quad (6.2.116)$$

です. これが示されると, 列ベクトルに対して示した結果が行ベクトルに対しても成立することになります. つまり

Theorem. 2.4.5. 二つの行が一致すれば $\det A = 0$ となる.

Theorem. 2.4.6. ある行に他のある行の定数倍を加えて得られる行列の行列式は, もとの行列の行列式 $\det A$ に等しい.

という定理を直接証明せずとも良いことが分かります. また定義としても

Definition. 2.4.7. 行列式は定義域が $M(n, m; C^n)$ で値域が C の, 単位行列に対し 1 を対応させる, 行に関する交代性を持った多重線型写像である.

を採用してよいことが分かります. 転置行列と元の行列式の行列式の値が等しいことは多重線型性を用いて証明します.
(定理の証明)

$$\mathbf{a}_i = \sum_{k=1}^3 a_{ki} \mathbf{e}_k, \quad {}^t \mathbf{a}_i = \sum_{k=1}^3 a_{ki} {}^t \mathbf{e}_k \quad (6.2.117)$$

と書けますから, これを用いて

$$\det \begin{pmatrix} {}^t \mathbf{a}_1 \\ {}^t \mathbf{a}_2 \\ {}^t \mathbf{a}_3 \end{pmatrix} = \det \begin{pmatrix} \sum_{i=1}^3 a_{i1} {}^t \mathbf{e}_i \\ \sum_{j=1}^3 a_{j2} {}^t \mathbf{e}_j \\ \sum_{k=1}^3 a_{k3} {}^t \mathbf{e}_k \end{pmatrix} \quad (6.2.118)$$

$$= \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 a_{i1} a_{j2} a_{k3} \det \begin{pmatrix} {}^t \mathbf{e}_i \\ {}^t \mathbf{e}_j \\ {}^t \mathbf{e}_k \end{pmatrix} \quad (\text{多重線型性}), \quad (6.2.119)$$

$$\det (\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3) = \det \left(\sum_{i=1}^3 a_{i1} \mathbf{e}_i \quad \sum_{j=1}^3 a_{j2} \mathbf{e}_j \quad \sum_{k=1}^3 a_{k3} \mathbf{e}_k \right) \quad (6.2.120)$$

$$= \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 a_{i1} a_{j2} a_{k3} \det (\mathbf{e}_i \ \mathbf{e}_j \ \mathbf{e}_k) \quad (\text{多重線型性}) \quad (6.2.121)$$

$$\iff 0 = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 a_{i1} a_{j2} a_{k3} \left\{ \det \begin{pmatrix} {}^t \mathbf{e}_i \\ {}^t \mathbf{e}_j \\ {}^t \mathbf{e}_k \end{pmatrix} - \det (\mathbf{e}_i \ \mathbf{e}_j \ \mathbf{e}_k) \right\} \quad (6.2.122)$$

示すべき式はこのようなになります. したがって焦点は

$$\det \begin{pmatrix} {}^t \mathbf{e}_i \\ {}^t \mathbf{e}_j \\ {}^t \mathbf{e}_k \end{pmatrix} = \det (\mathbf{e}_i \ \mathbf{e}_j \ \mathbf{e}_k) \quad (6.2.123)$$

が示せるかどうかです. 先程示した定理から i, j, k のうちのいずれか 2 つ以上が一致するときは両辺 0, つまりこのときはきちんと等号が成立します. そこで問題は i, j, k が異なる場合です.

実際に一致するのですが, 全部確認するのは面倒なのでいくつか具体的に見てみましょう.

$$\det \begin{pmatrix} {}^t \mathbf{e}_1 \\ {}^t \mathbf{e}_2 \\ {}^t \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \det (\mathbf{e}_1 \ \mathbf{e}_2 \ \mathbf{e}_3) \quad (6.2.124)$$

$$\det \begin{pmatrix} {}^t e_2 \\ {}^t e_3 \\ {}^t e_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad (6.2.125)$$

$$= - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{ (交代性を用いて 1 列と 2 列を交換)} \quad (6.2.126)$$

$$= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ (交代性を用いて 1 列と 3 列を交換)} \quad (6.2.127)$$

$$= \det(e_2 e_3 e_1) \quad (6.2.128)$$

したがって上の等号が確かに成立します. ■
さらに次に示すのは

Theorem. 2.4.8. (行列式の展開 1)

$$\begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \quad (6.2.129)$$

です. 3 次行列の行列式と 2 次行列の行列式を結ぶ式ですから, 重要そうな気配がするわけです.

(定理の証明) 前に示した定理と多重線型性から

$$\begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = \det(e_1 a_2 a_3) \quad (6.2.130)$$

$$= \det(e_1 - a_{12}e_1 + a_2 - a_{13}e_1 + a_3) \quad (6.2.131)$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = a_{22} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} + a_{32} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & 1 & a_{33} \end{vmatrix} \quad (6.2.132)$$

$$= a_{22} \left(a_{23} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{vmatrix} + a_{33} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \right) \quad (6.2.133)$$

$$+ a_{32} \left(a_{23} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} + a_{33} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{vmatrix} \right) \quad (6.2.134)$$

ここで前定理により第 1 項と第 4 項が 0 になります. 第 2 項と第 3 項には交代性を用いて

$$= a_{22}a_{33} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + a_{23}a_{32}(-1) \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \quad (6.2.135)$$

$$= a_{22}a_{33} - a_{23}a_{32} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \quad (6.2.136)$$

これで証明終了です. ■

準備が整ったので, はじめの懸案を片付けましょう. 示すべきは

Theorem. 2.4.9. (行列式の展開 2)

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \quad (6.2.137)$$

(証明) まず $\mathbf{a}_1 = a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + a_{31}\mathbf{e}_3$ を用いて多重線型性から

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad (6.2.138)$$

$$= a_{11} \begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + a_{21} \begin{vmatrix} 0 & a_{12} & a_{13} \\ 1 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 1 & a_{32} & a_{33} \end{vmatrix} \quad (6.2.139)$$

まず右辺第一項から考えましょう。つい先程示した定理から

$$\begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \quad (6.2.140)$$

となります。第 2 項は交代性を用いて

$$\begin{vmatrix} 0 & a_{12} & a_{13} \\ 1 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = - \begin{vmatrix} 1 & a_{22} & a_{23} \\ 0 & a_{12} & a_{13} \\ 0 & a_{32} & a_{33} \end{vmatrix} = - \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} \quad (6.2.141)$$

となります。第 3 項も同様です。これで証明終了です。■

手間暇はかかりますが、全く同じようにして一般の n 次の行列式をこのように「展開」することが出来ます。

行列式に関する定理をもう 2 つ紹介しておきます。

Theorem. 2.4.10. 2 つの行列の積の行列式はそれぞれの行列式の積に等しい。
すなわち $\det A \det B = \det(AB)$

(証明) これも多重線型性が命です。

$$\det(AB) = \det \left(\sum_{i=1}^3 b_{i1} \mathbf{a}_i \quad \sum_{j=1}^3 b_{j2} \mathbf{a}_j \quad \sum_{k=1}^3 b_{k3} \mathbf{a}_k \right) \quad (6.2.142)$$

$$= \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 b_{i1} b_{j2} b_{k3} \det(\mathbf{a}_i \quad \mathbf{a}_j \quad \mathbf{a}_k) \quad (6.2.143)$$

ここで交代性から i, j, k のうちで一致するものがあると 0 で、このうちのどれか

一つの添え字をひっくり返したのものにはマイナスがつきますから、

$$\det(AB) \quad (6.2.144)$$

$$= b_{11}b_{22}b_{33} \det(\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3) + b_{11}b_{32}b_{23} \det(\mathbf{a}_1 \ \mathbf{a}_3 \ \mathbf{a}_2) \quad (6.2.145)$$

$$+ b_{21}b_{12}b_{33} \det(\mathbf{a}_2 \ \mathbf{a}_1 \ \mathbf{a}_3) + b_{21}b_{32}b_{13} \det(\mathbf{a}_2 \ \mathbf{a}_3 \ \mathbf{a}_1) \quad (6.2.146)$$

$$+ b_{31}b_{22}b_{13} \det(\mathbf{a}_3 \ \mathbf{a}_2 \ \mathbf{a}_1) + b_{31}b_{12}b_{23} \det(\mathbf{a}_3 \ \mathbf{a}_1 \ \mathbf{a}_2) \quad (6.2.147)$$

$$= (b_{11}b_{22}b_{33} - b_{11}b_{32}b_{23} - b_{21}b_{12}b_{33} + b_{21}b_{32}b_{13} - b_{31}b_{22}b_{13} + b_{31}b_{12}b_{23}) \quad (6.2.148)$$

$$\times \det(\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3) \quad (6.2.149)$$

$$= \{b_{11}(b_{22}b_{33} - b_{32}b_{23}) - b_{12}(b_{21}b_{33} - b_{23}b_{31}) + b_{13}(b_{21}b_{32} - b_{22}b_{31})\} \quad (6.2.150)$$

$$\times \det(\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3) \quad (6.2.151)$$

$$= \left(b_{11} \begin{vmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{vmatrix} - b_{12} \begin{vmatrix} b_{12} & b_{13} \\ b_{32} & b_{33} \end{vmatrix} + b_{13} \begin{vmatrix} b_{12} & b_{13} \\ b_{22} & b_{23} \end{vmatrix} \right) \det(\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3) \quad (6.2.152)$$

$$= \det(\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3) \det(\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3) \quad (6.2.153)$$

最後のところで先程の行列式の展開を用いました。■

あとで用いる連立一次方程式に関する結果をここで示しておきます。

Theorem. 2.4.11. 連立一次方程式

$$c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n = 0 \quad (6.2.154)$$

$$c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n = 0 \quad (6.2.155)$$

$$\vdots \quad (6.2.156)$$

$$c_{n1}x_1 + c_{n2}x_2 + \cdots + c_{nn}x_n = 0 \quad (6.2.157)$$

が $x_i \neq 0$ となる i が存在するような解の組 $\mathbf{x} := (x_1 \ \cdots \ x_n)$ を持つ必要十分条件は係数 c_{ij} のなす正方行列 $C := (c_{ij})$ の行列式 $\det C$ が 0 になることである。

(証明) まず連立一次方程式は行列で $C\mathbf{x} = \mathbf{0}$ と書けることに注意します。

(必要性) 背理法で証明します。 $\det C = 0$ としましょう。すると逆行列が存在するので、これをかけると $\mathbf{x} = \mathbf{0}$ となり、仮定に反します。

(十分性) 一般には帰納法で証明できますが、帰納法で何をやっているのかを確認することにもなりますから、 $n = 1, 2, 3$ で具体的に調べてみます。 $n = 1$ のときは連立方程式は $c_{11}x_1 = 0$ で $\det C = c_{11} = 0$ なので、 $x_1 \neq 0$ となる解が存在することはすぐに分かります。

$n = 2$ のとき、2 行目に c_{11} をかけ、1 行目に c_{21} をかけて、2 行目から 1 行目を引くと

$$c_{11}x_1 + c_{12}x_2 = 0 \quad (6.2.158)$$

$$0x_1 + (c_{11}c_{22} - c_{12}c_{21})x_2 = 0 \quad (6.2.159)$$

$$(6.2.160)$$

となります。 x_2 の係数は $-\det C = 0$ なので、 $x_2 \neq 0$ としてもよいことが分かります。 $x_i \neq 0$ となる i が存在することさえ示せばよいので、これで十分です。

$n = 3$ のとき、もし全ての c_{ij} が 0 ならば、当然 $\det C = 0$ です。そして任意の \mathbf{x} が解となるので、証明終了です。次に $c_{ij} \neq 0$ となる i, j が存在するとします。例えばそれが c_{22} であったとしましょう。1 行目と 2 行目を入れ替え、1 列と

2 列目を入れ替えると, c_{22} が c_{11} の位置に来ます. こうして得られたものにあらためて番号を振りなおすと $c_{11} \neq 0$ となります. このようなことを考慮に入れると $c_{11} \neq 0$ と仮定してもよいことに注意しましょう. $n = 2$ のときと同じようにして,

$$c_{11}x_1 + c_{12}x_2 + c_{13}x_3 = 0 \quad (6.2.161)$$

$$0x_1 + (c_{11}c_{22} - c_{21}c_{12})x_2 + (c_{11}c_{23} - c_{21}c_{13})x_3 = 0 \quad (6.2.162)$$

$$0x_1 + (c_{11}c_{32} - c_{31}c_{12})x_2 + (c_{11}c_{33} - c_{31}c_{13})x_3 = 0 \quad (6.2.163)$$

となります. $n = 2$ のときと同様にして今度は 3 行目の x_2 の係数を 0 にします. このとき 3 行目の x_3 の係数を計算すると,

$$(c_{11}c_{22} - c_{21}c_{12})(c_{11}c_{33} - c_{31}c_{13}) - (c_{11}c_{32} - c_{31}c_{12})(c_{11}c_{23} - c_{21}c_{13}) \quad (6.2.164)$$

となります. ここで $\det C$ を計算してみます. C の第 i 行ベクトルを c_i として今まで証明してきた定理を色々使うと

$$0 = \det C = \begin{vmatrix} c_1 \\ c_2 \\ c_3 \end{vmatrix} = \begin{vmatrix} c_1 \\ -c_{21}c_1 + c_{11}c_2 \\ -c_{31}c_1 + c_{11}c_3 \end{vmatrix} \quad (6.2.165)$$

$$= \begin{vmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{11}c_{22} - c_{21}c_{12} & c_{11}c_{23} - c_{21}c_{13} \\ 0 & c_{11}c_{32} - c_{31}c_{12} & c_{11}c_{33} - c_{31}c_{13} \end{vmatrix} \quad (6.2.166)$$

$$= c_{11} \det \begin{vmatrix} c_{11}c_{22} - c_{21}c_{12} & c_{11}c_{23} - c_{21}c_{13} \\ c_{11}c_{32} - c_{31}c_{12} & c_{11}c_{33} - c_{31}c_{13} \end{vmatrix} \quad (6.2.167)$$

$c_{11} \neq 0$ なので $x_3 \neq 0$ と選べます.

一般の場合は帰納法で示せませんがここでは省略します. これで証明終了です.

ここ本節はじめに述べた, 「線型変換に逆変換があるかどうかを行列による表示だけから判定できるか?」という問題の答が述べられます. 線型変換 f に対する行列表示, は適当に座標系 (基底) を選んだときに一意的に定まるものです. 標準基底を取った時の線型写像を表す行列を A , 別の座標系 $\langle u_i \rangle_{i=1}^3$ を取った時の線型写像を表す行列を B とし, 座標系の変換を表す行列を P とします. このとき p 57, (6.2.97) から $A = PBP^{-1}$ となります. 今示した定理と $P^{-1}P = I$ から

$$\det A = \det(PBP^{-1}) = \det P \det B \det P^{-1} \quad (6.2.168)$$

$$= \det(PP^{-1}) \det B = \det I \det B = \det B \quad (6.2.169)$$

となります. 高校でも学んだとおり逆行列が存在することと行列式が 0 でないことは同値です. 上の式から適当な座標系で行列式が 0 でないならば, いかなる座標系で行列式を計算しようとも行列式の値は 0 ではありません. すなわち逆変換が存在するか否かは適当な座標系を取ったときの行列式を計算すれば十分であることが分かりました.

6.2.5 内積と外積

内積から書きます. ここでの内積は高校のときと同じです. $x = \sum_{i=1}^n x_i e_i \in C^n$, $y = \sum_{i=1}^n y_i e_i \in C^n$ としましょう. このとき x と y の内積 (x, y) を以下のように定義します.

Definition. 2.5.1. x と y の内積 (x, y) とは

$$(x, y) := \sum_{i=1}^n x_i \bar{y}_i \quad (6.2.170)$$

y_i についたバーは複素共役を表します。もちろん実数のときはあってもなくても同じです。また行列の積を用いると内積はこうも書けます。

$$(x, y) = {}^t x \bar{y} \quad (6.2.171)$$

内積を $\langle x | y \rangle$, $\langle x, y \rangle$ と書くこともあります。ややこしいですが、本冊子でも物理の記述をするときには従来どおり内積を $x \cdot y$ と書くことがあるでしょう。また定義から $(x, x) \geq 0$ ですが、これをベクトルの長さと呼び、

$$\sqrt{(x, x)} = |x| \quad (6.2.172)$$

と書くことも高校と同じです。

その他内積の値が一般に複素数になることが分かります。高校で $\cos \theta = (x, y)/(|x||y|)$ という 2 つのベクトルのなす角度が余弦の形で計算できる式を学んだはずですが、これがそのまま成立するとすれば、 \cos の値が複素数になります。複素関数論を学んだ後では別にどうということもないのですが、角度を定義するという観点からみると、 θ が複素数になってしまってまともな角度でなくなる（実数でなくなる）ので少し困ります。これの解決法としては、複素ベクトルの範囲で角度を考えると内積の値が $0 \Leftrightarrow$ 2 つのベクトルは直交というところにだけ着目して、後の角度に関する事項は一切考えない、というものがあります。実際問題として、直交性は非常に重要な役割を果たすことがよくありますが、それ以外の角度を論じることはほとんどありませんので、特に問題ありません。

次に外積です。これから書く外積は \mathbb{R}^3 のときにしか定義されないことに注意して下さい $x = {}^t(x_1, x_2, x_3)$, $y = {}^t(y_1, y_2, y_3)$ として外積を以下のように定義します。

Definition. 2.5.2. x と y の外積 $x \times y$

$$x \times y := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}. \quad (6.2.173)$$

簡単な記憶法を紹介します。このために行列式の後にしたのですが、行列式を用いた記憶法です。

$$x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \quad (6.2.174)$$

これを第 1 行に関して展開します。先程は列に関して展開しましたが、今度は行です。すると

$$x \times y = (x_2 y_3 - x_3 y_2) e_1 + (x_3 y_1 - x_1 y_3) e_2 + (x_1 y_2 - x_2 y_1) e_3 \quad (6.2.175)$$

となり、確かに上の定義と一致します。これを記憶法といって定義としなかった理由は成分にベクトルと実数が共存する行列式を定義していないからです。

6.2.6 線型独立と基底, 次元

とりあえず天下りですが, 線型結合と線型独立, 線型従属の定義を述べます. あまり変わらないので, 本節は一般の n で述べます. C^n のベクトル a_1, a_2, \dots, a_m に対して

$$c_1 a_1 + c_2 a_2 + \dots + c_m a_m \quad (6.2.176)$$

の形のベクトルを a_1, a_2, \dots, a_m の線型結合といいます. 高校では一次結合と呼んでいたかもしれませんが.

ベクトル a_1, a_2, \dots, a_m の間の関係

$$c_1 a_1 + c_2 a_2 + \dots + c_m a_m = 0 \quad (6.2.177)$$

を線型関係と呼びます. 全ての c_i が 0 の場合がありますから, 線型関係は必ず存在します. これを自明な線型関係といいます. 自明でない線型関係はあるとは限りませんが,

Definition. 2.6.1. 自明でない線型関係が存在するとき, a_1, a_2, \dots, a_m は線型従属である自明な線型関係が存在しないとき, a_1, a_2, \dots, a_m は線型独立である.

と定義します.

次に基底の定義を述べましょう.

Definition. 2.6.2. C^n の有限個のベクトル e_1, e_2, \dots, e_m が次の 2 条件を満たすとき, これらのベクトルが基底であるという.

1. C^n の任意のベクトルは e_1, e_2, \dots, e_m の線型結合として表される.
2. e_1, e_2, \dots, e_m は線型独立である.

定義から, 基底は何か特定の一組のものを指しているわけではありません. また基底はその順番も大事なので, 順番を込めたものとして基底を $\langle e_1, \dots, e_m \rangle = \langle e_i \rangle_{i=1}^m$ などとも書きます. とりあえずはこれを座標軸のようなものだと考えておいて下さい. 座標軸の選び方が色々あった方がいいので, それを数学的に定式化したものが基底です. これまで何度となく使ってきた (ものの n 次元版の) n 項単位ベクトルの組 $e_i, i = 1, 2, \dots, n$ (上のものとは別!) は標準基底と呼ばれます.

座標軸の選び方が色々あった方がいい, ということがよく分からないと思いますが, これは物理からの要請とも言えます. R^3 で考えてみましょう.

どうということかという, 例えば固体の結晶構造を解析したいとします. 手持ちの「キッテル固体物理学入門」によりますと, 三斜晶系, 単斜晶系, 斜方晶系, 正方晶系, 立方晶系, 菱面体晶系, 六方晶系というのがあるようで, さらにこれに格子の数がどうのこうのとあります. 今調べたい結晶の原子配列を見てみたら, お互いに 60 度の角度をなす, 長さの比が 2:1:1 のベクトルの先に原子がちょうど上手く乗っていたとしましょう. このときこのようなベクトルを取ってやれば, 現象の記述に便利そうです. この少し斜めに傾いた結晶の解析をするのに, 先程の「まっすぐ」な標準基底をとって調べるような人間がいたら, その人間の心は曲がっています.

そういうわけでこうした都合の良いベクトルを用いていきたいわけですが, これが標準基底と数学的に同じような資格をもっていて欲しいわけですが. 物理的には良くても数学的に問題があるようでは困ってしまいますから. ここで「標準基底と数学的に同じような資格」と書きましたが, これがまた曲者です. どういうことかという, 「標準基底の持っている性質」というのが何なのかよく分かっていないといけな訳ですが, これが良く分かりません. 困ってしまったわけで

すが、こういうとき数学でよく行なう手法というのがあります、それを紹介します。それというのは、標準基底がどんな性質を持っているかではなく、どんな性質を持っていて欲しいかを考えるということです。

R^3 の標準基底が持っていて欲しい性質を考えましょう。まず 1 つは標準基底で空間内の全ての点を表現出来ないと困ります。これは固体内の全ての原子の位置を書き表せない困る、ということです。ここから標準基底が持っていて欲しい性質を一つ要請しましょう。それはある程度たくさんのベクトルからなることです。例えば上の定義から基底は複数のベクトルからなるのが一般的ですが、仮に基底がただ 1 つのベクトル a_1 からなったとしましょう。そうすると a_1 では直線上の点しか記述できません。少なくとも基底がある程度たくさんのベクトルからなっていないと困るわけです。実際上で述べた標準基底はそうになっています。

そうなるのとどのくらい多ければいいのか、というのを考えたいわけです。あまり多くても邪魔なので、少なめにしたいところです。あまり少ないと先程の 1 つということになってしまうので、少し考えないといけません。これもまた、空間内の全ての点を表現出来ないと困るということが重要です。先程同様、基底が 2 つのベクトルからなるとしても高校で学んだとおり、良くて平面上の点しか記述出来ません。というわけで 3 本以上は欲しいことがわかります。詳しく述べるのはやめますが、3 次元空間内では 4 本以上あるとこれは上で定義した言葉を使うと、このベクトルの組は線型従属になります。線型従属だと、その中のベクトルのうちいくつかはその組の中のベクトルのいくつかで書くことが出来、空間の点を全て表現する、という観点からは別に無くても不自由しません。ということで、 \mathbb{R}^3 の基底は 3 つあればよい、ということです。実際標準基底はそうになっています。

今までのことをまとめると、基底は空間内の点を全て表現できること、数なるべく少ないことという性質を持っていて欲しいということになりました。基底の定義と比べると、「空間内の点を全て表現できること」が 1、 \equiv 数なるべく少ないこと」が 2 に対応することが分かります。

ちなみに 2 を考慮に入れると、定義で m と書いたところは実は n であったことが分かります。この数 n のことを次元といいます。ある空間の次元を \dim で表します。今の例で言えば、 $\dim R^3 = 3$ です。

基底に関して 1 つ重要なことを述べておきます。それは正規直交基底というものです。これは $\{e_i\}_{i=1}^m$ を基底としたとき、

$$(e_i, e_j) = \delta_{ij} \quad (6.2.178)$$

となるもののことです。ここで δ_{ij} は Kronecker の δ であり、 $i = j$ で 1、 $i \neq j$ で 0 になるというものでした。ここの言葉の意味を説明しておきましょう。正規というのは $\sqrt{(e_i, e_i)} = |e_i| = 1$ となるもの、つまりベクトルの長さが 1 であるということです。直交というのはもちろん $i \neq j$ のとき $(e_i, e_j) = 0$ となることです。例えば標準基底は正規直交基底です。

6.2.7 固有値, 固有ベクトル

まず固有値・固有ベクトルを定義しましょう。

Definition. 2.7.1. 行列 A に対し零ベクトルでないベクトル x と複素数 λ が以下の関係式を満たすとき、 λ を固有値、 x を固有ベクトルという。

$$Ax = \lambda x \quad (6.2.179)$$

要するに、行列の作用により方向を変えないベクトルを固有ベクトルといい、そのとき行列の作用は当然定数倍にしかなりえないわけですが、その倍数を固有値といいます。

また固有値を一般に複素数といいましたが, この理由を具体例で考えます.

$$R(\theta) := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (6.2.180)$$

この行列の固有値を考えてみましょう. 幾何学的に考えると, この行列をあるベクトルにける (作用させる) ことは平面内での回転を意味します. 回転させるので, 「方向を変えない」ということは $\theta = 2n\pi$, $n \in \mathbb{N}$ のときしかありえません. 実際この行列の固有値は $\cos \theta + i \sin \theta$ であって複素数です. 実行列だけを考えていても複素数の固有値が出てくるので, はじめから複素数で話を進めようということです.

また固有ベクトルを零ベクトルでないとした訳は, 当然ながら零ベクトルであれば固有値が何であれ上の式を満たしてしまうからです.

上の固有値・固有ベクトルの定義式では, 行列だけ分かっていて一般には固有値・固有ベクトルは未知です. これを求める方法を見つけないといけませんが, 一度に両方見つけようというのは欲張りです. 1 つずつ見つけることにしましょう. 上の式を以下のように変形します.

$$Ax = \lambda x = \lambda Ix \quad (I \text{ は単位行列}) \quad (6.2.181)$$

$$\iff (\lambda I - A)x = \mathbf{0} \quad (6.2.182)$$

固有ベクトルはゼロベクトルではないとしていますから, $\det(\lambda I - A) = 0$ が従います.

ここで行列式の計算をすると n 次行列のとき, 一般に行列式は λ の n 次方程式になります. 代数学の基本定理により, 重複を含めて n 個の根 (解) が必ず存在します. つまり固有値が存在します. ここで求めた固有値を定義式に代入して固有ベクトルを求めます. この手続きは 1 度やってみないと分からないと思いますので, 1 つ例をやってみましょう.

後で使うのにもちょうど良いので, 一番初めの運動方程式のところで出てきた行列を使いましょう.

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad (6.2.183)$$

つまり上の A の固有値・固有ベクトルを求めます. 上の教えにしたがってまず行列式を計算します. 行列式の具体的な計算練習も兼ねて, 少し丁寧にやりましょう.

$$0 = \det(\lambda I - A) = \begin{vmatrix} \lambda - 2 & 1 & 0 \\ 1 & \lambda - 2 & 1 \\ 0 & 1 & \lambda - 2 \end{vmatrix} \quad (6.2.184)$$

$$= (\lambda - 2) \begin{vmatrix} \lambda - 2 & 1 \\ 1 & \lambda - 2 \end{vmatrix} - 1 \times \begin{vmatrix} 1 & 1 \\ 0 & \lambda - 2 \end{vmatrix} \quad (6.2.185)$$

$$= (\lambda - 2) \{(\lambda - 2)^2 - 1\} - (\lambda - 2) \quad (6.2.186)$$

$$= (\lambda - 2) \{ \lambda - (2 + \sqrt{2}) \} \{ \lambda - (2 - \sqrt{2}) \} \quad (6.2.187)$$

したがって固有値は 2 , $2 + \sqrt{2}$, $2 - \sqrt{2}$ と分かりました.

固有ベクトルを求めましょう。分かりやすいように各固有値に対する固有ベクトルを u_2 , $u_{2+\sqrt{2}}$, $u_{2-\sqrt{2}}$ とします。式 (6.2.182) から

$$(2I - A)u_2 = 0 \quad (6.2.188)$$

$$\iff \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} u_{2,1} \\ u_{2,2} \\ u_{2,3} \end{pmatrix} = 0, \quad u_2 := \begin{pmatrix} u_{2,1} \\ u_{2,2} \\ u_{2,3} \end{pmatrix} \quad (6.2.189)$$

$$\implies u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad (6.2.190)$$

定義から分かるとおり, u が固有ベクトルだとすれば任意の実数 c に対して cu も固有ベクトルです。つまり, 固有ベクトルの大きさは特に指定されていません。それを良いことに固有ベクトルの大きさを 1 に取っておきました。後々 1 の方が都合がいいのです。ついでに述べておきますと $u_2 = 1/\sqrt{2} {}^t(-1 \ 0 \ 1)$ でも構いません。これについては趣味です。同様にして $u_{2+\sqrt{2}} = 1/2 {}^t(1 \ -\sqrt{2} \ 1)$, $u_{2-\sqrt{2}} = 1/2 {}^t(1 \ \sqrt{2} \ 1)$ となります。

固有値について注意をいくつかしておきます。1 つの固有値に対して固有ベクトルが 1 つであるとは限りません。具体的に見てみましょう。

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \quad (6.2.191)$$

上の行列の固有値は 1, 4 ですが, 固有値 1 の固有ベクトルは 2 つあります。選び方には任意性がありますが, 例えば ${}^t(1 \ -1 \ 0)$, ${}^t(1 \ 0 \ -1)$ です。さらに重要な注意として, 上の固有ベクトルは互いに直交する大きさ 1 のベクトルに選ぶことが出来ます。最後の以下の定理を証明しておきます。

Theorem. 2.7.2. 任意の *Hermite* 行列 H , ユニタリ行列 U の互いに異なる固有値に対する固有ベクトルは直交する。

(証明) 互いに異なる固有値を λ_1, λ_2 , 対応する固有ベクトルを x_1, x_2 とします。まず *Hermite* から示します。定義から $H^* = H$ であり, 内積の定義から A に対して $(Ax, y) = (x, A^*y)$ となることを思い出して下さい。まず *Hermite* 行列の固有値が (複素数ではなく) 実数になることを示します。

$$(Hx, x) = \lambda(x, x) \quad (6.2.192)$$

$$(Hx, x) = (x, Hx) = \bar{\lambda}(x, x) \quad (6.2.193)$$

$$\implies \lambda = \bar{\lambda} \quad (6.2.194)$$

これから

$$(Hx_1, x_2) = \lambda_1(x_1, x_2) \quad (6.2.195)$$

$$(Hx_1, x_2) = (x_1, Hx_2) = \lambda_2(x_1, x_2) \quad (6.2.196)$$

となりますから, 辺々引くと $(\lambda_1 - \lambda_2)(x_1, x_2) = 0$ であり, 固有値が異なるので, $(x_1, x_2) = 0$ となります。

ユニタリについてもほぼ同様に示せます。ユニタリの場合は $U^* = U^{-1}$ なので, $Ux_i = \lambda_i x_i$ から $U^*x_i = (1/\lambda_i)x_i$ となります。さらに

$$(x, x) = (U^*Ux, x) = (Ux, Ux) = (\lambda x, \lambda x) = \lambda \bar{\lambda}(x, x) \quad (6.2.197)$$

から $\lambda\bar{\lambda} = 1$ が分かります。これから

$$(Ux_1, x_2) = \lambda_1(x_1, x_2) \quad (6.2.198)$$

$$= (x_1, U^*x_2) = (x_1, \frac{1}{\lambda_2}x_2) = (x_1, \bar{\lambda}_2x_2) = \lambda_2(x_1, x_2) \quad (6.2.199)$$

となるので、先程と同様にして $(x_1, x_2) = 0$ です。■

6.2.8 対称行列の対角化とスペクトル分解

さて行列の話も大詰めです。ここでは対称行列が対角化可能であることを示します。対角化とは、ある行列 A に適当な行列 U をかけることで対角行列 Λ にしてしまうことです。これがあると何が嬉しいかということは次節で明らかになります。

さて、まず本節で示す定理の最も一般的な形を紹介します。行列 A が正規行列であるということは $AA^* = A^*A$ で定義されました。ここで A^* は $A^* = {}^t\bar{A}$ で定義される随伴行列と呼ばれるものでした。ここで示したい定理は

Theorem. 2.8.1. 正方行列 A に対し $U^{-1}AU$ が対角行列になるようなユニタリ行列 U が存在する為には A が正規行列であることが必要十分である。

というものです。しかし実用性の観点からは次の定理を示せば十分です。また実 Hermite 行列は直交行列であることを思い出して下さい。

Theorem. 2.8.2. Hermite 行列 H は適当なユニタリ行列によって対角化可能である。

(証明) 分かりやすくするために 3 次行列に対して証明します。全く同様にして一般の n 次でも出来ます。 H の固有値を h_1, h_2, h_3 , 対応する固有ベクトルをそれぞれ u_1, u_2, u_3 とします。固有ベクトルは規格化 (大きさが 1) しておきます。また $U = (u_1 \ u_2 \ u_3)$ としておきます。

一旦全ての固有値が異なるとして証明します。積の行列 U^*HU の (i, j) 成分を計算すると

$$U^*HU = \begin{pmatrix} u_1^* \\ u_2^* \\ u_3^* \end{pmatrix} (Hu_1 \ Hu_2 \ Hu_3) \quad (6.2.200)$$

$$= \begin{pmatrix} \overline{(u_1, Hu_1)} & \overline{(u_1, Hu_2)} & \overline{(u_1, Hu_3)} \\ \overline{(u_2, Hu_1)} & \overline{(u_2, Hu_2)} & \overline{(u_2, Hu_3)} \\ \overline{(u_3, Hu_1)} & \overline{(u_3, Hu_2)} & \overline{(u_3, Hu_3)} \end{pmatrix} \quad (6.2.201)$$

となります。つまり $U^*HU = \left(\overline{(u_i, Hu_j)} \right)$ です。ここで i, j を 1, 2, 3 のいずれかとして、

$$(u_i, Hu_j) = \overline{h_j}(u_i, u_j) = \overline{h_j}\delta_{ij} \quad (6.2.202)$$

となりますから、 U^*HU は対角行列です。

固有値に重複があったときは、重複した固有値に対する固有ベクトルが複数存在することになりますが、これらを全て互いに直交するように選びなおすことが出来ます。したがって上の議論がほぼそのまま成立します。これで定理が完全に証明されました。■

ここで前に正規行列を紹介したので、スペクトル分解を紹介したいと思います。まず射影 P を定義します。 W の直交補空間を W^\perp とすると、 C^n の任意のベクトル x は

$$x = x' + x'', \quad x' \in W, \quad x'' \in W^\perp \quad (6.2.203)$$

と一意的に分解されます。このとき x に対して上手く $Px = x'$ となる行列 P を部分空間 W への射影であるといいます。また P は

$$P^2 = P, \quad P^* = P \quad (6.2.204)$$

を満たします。

ここで W_1, W_2 を C^n の 2 つの部分空間とし、 P_1, P_2 をそれぞれ W_1, W_2 への射影としましょう。このとき部分空間 W_1, W_2 が直交する必要十分条件は $P_1 P_2 = P_2 P_1 = O$ となります。

実際、 W_1 と W_2 が直交すれば $W_1 \subset W_2^\perp$ から $P_1 P_2 x = 0$ です。逆に $P_1 P_2 = O$ ならば、任意の $x_1 \in W_1, x_2 \in W_2$ に対し、

$$(x_1, x_2) = (P_1 x_1, P_2 x_2) = (x_1, P_1 P_2 x_2) = 0 \quad (6.2.205)$$

となり、証明終了です。■

A を正規行列としましょう。 A の異なる固有値全部を $\beta_1, \beta_2, \dots, \beta_k$, 対応する i に対し $W_i := \{x; Ax = \beta_i x\}$ とします。このとき各 W_i は互いに直交し、 C^n はこれらの部分空間の直和になります。 P_i を各 W_i への射影とすると

$$P_1 + P_2 + \dots + P_k = I, \quad P_i P_j = O, \quad i \neq j \quad (6.2.206)$$

$$A = \beta_1 P_1 + \beta_2 P_2 + \dots + \beta_k P_k \quad (6.2.207)$$

が一意的に成り立ちます。これを正規行列 A のスペクトル分解と言います。

スペクトル分解の意味を少し説明しておきましょう。一番重要なのは固有ベクトルに対しては行列の積の作用が単なる定数倍になるということです。きちんと証明していませんが、ある行列に対してはその行列の固有ベクトルを基底として選ぶことができます。この行列を A , 固有ベクトルからなる基底を u_1, u_2, u_3 , 固有値はそれぞれ $\lambda_1, \lambda_2, \lambda_3$ とします。すなわち $Au_i = \lambda_i u_i$ です。この固有ベクトルで任意のベクトル x を $x = x_1 u_1 + x_2 u_2 + x_3 u_3$ と展開します。このとき行列の作用を考えると、

$$Ax = A(x_1 u_1 + x_2 u_2 + x_3 u_3) \quad (6.2.208)$$

$$= x_1 \lambda_1 u_1 + x_2 \lambda_2 u_2 + x_3 \lambda_3 u_3 \quad (6.2.209)$$

$$= x_1 (\lambda_1 I) u_1 + x_2 (\lambda_2 I) u_2 + x_3 (\lambda_3 I) u_3 \quad (6.2.210)$$

と書けます。ここで I は単位行列です。こうすると「行列をかける (作用させる)」という直観的に捉えにくいものが、ただの定数倍に還元され、非常に見やすくなります。あとでスペクトル分解と Fourier 解析との関係を少し紹介する予定なので、軽く紹介してみました。

6.2.9 物理への適用

準備が終わったので、早速これをはじめに言ってあった現象の解析に応用してみましょう。3 質点系での解析を行ないます。問題を再録しましょう。

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 - k(x_1 - x_2) \\ m\ddot{x}_2 &= +k(x_1 - x_2) - k(x_2 - x_3) \\ m\ddot{x}_3 &= -kx_3 + k(x_2 - x_3) \\ \iff m \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{pmatrix} &= -k \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned}$$

$$\iff m\ddot{\mathbf{x}} = -kA\mathbf{x}, \quad \mathbf{x} := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad A := \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

ここで A が対称行列 (正規行列) になっていますから、これが対角化できます。

対角化すれば一番初めに見せた X のようなものが上手に見つけられたことに相当し、首尾よく問題が解けたことになります。というわけで、6.2.8 節で示したように A を対角化するようなユニタリ行列を見つけられれば良い、ということになります。 A の固有値、固有ベクトルは 6.2.7 節で求めてしまいました。これにより A を対角化するには 6.2.8 節のはじめに示した定理を使えば十分です。つまり、固有ベクトルから作ったユニタリ行列 U を使えば良いということです。

6.2.7 節から固有値、固有ベクトルを再録してきましょう。

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \mathbf{u}_{2+\sqrt{2}} = \frac{1}{2} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \mathbf{u}_{2-\sqrt{2}} = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} \quad (6.2.211)$$

したがって $U = (\mathbf{u}_2 \ \mathbf{u}_{2+\sqrt{2}} \ \mathbf{u}_{2-\sqrt{2}})$ とすればこの U で $U^*AU = \Lambda$ として対角化できます。6.2.8 節のはじめに示した定理によれば、対角成分には固有値が並びます。このあたりが固有値を「固有」値と呼ぶ所以でしょう。

Λ を実際に書いてみると

$$\Lambda = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2+\sqrt{2} & 0 \\ 0 & 0 & 2-\sqrt{2} \end{pmatrix} \quad (6.2.212)$$

ここで $\mathbf{X} := U^*\mathbf{x} = {}^t(X_1 \ X_2 \ X_3)$ とします。上の行列に左から U^* をかけて

$$mU^*\ddot{\mathbf{x}} = -kU^*A\mathbf{x} = -kU^*AUU^*\mathbf{x} \quad (6.2.213)$$

$$\iff m\ddot{\mathbf{X}} = -k\Lambda\mathbf{X} \quad (6.2.214)$$

あとは各 X_i について解くだけです。もちろん、はじめに天下りに出した X_i がこれらです。これらはこのようにして導き出したものです。

あと詳しく詰めるのは各人で練習してみてください。弾性体への拡張 ($n \rightarrow \infty$ への拡張) も適当な本を見ると載っていますから、まずは自分でアタックしてみてください。本節もまだ残り一節ありますが、次の予告をしましょう。次章は一般のポテンシャルとその安定点近傍での振動です。数学的には微分積分学です。

6.2.10 無限次元の線型代数学

この節は後々への接続と量子力学のためのものです気楽に読んで下さい。

抽象ベクトル 2

6.2.3 節ではベクトルについて少し突っ込んだ議論をしました。ここではもう少し踏み込んでいきます。まず、 \mathbb{C}^n のベクトルを幾何ベクトルといいます。ということで他のベクトルがあるということを匂わすわけですが、その前に一般のベクトルが持つべき性質を考えて見ましょう。 x, y, z を任意の幾何ベクトル、 α, β を任意の複素数としましょう。このときこれらは以下の関係式を満たします。

1. $x + y = y + x$
2. $(x + y) + z = x + (y + z)$
3. 全ての x に対し、 $x + 0 = x$ を満たすベクトル 0 がただ 1 つ存在する。
4. 全ての x に対し、 $x + x' = 0$ を満たすベクトル x' がただ 1 つ存在する。
5. $1x = x$
6. $\alpha(\beta x) = (\alpha\beta)x$
7. $\alpha(x + y) = \alpha x + \alpha y$
8. $(\alpha + \beta)x = \alpha x + \beta x$

ここで一般のベクトルにも全く同じ性質を要求します。また時として、今まで持ってきたベクトルの幾何学的な描像は理解の妨げになります。例として普通の n 次元ベクトルでも幾何学的な描像が思い描きにくいようなものをあげましょう。

我々が八百屋だとします。仕入れなどの問題もありますから、野菜の売上というものは気にかかるものです。よく売れる品物や品種を調べたりしたいわけですが、このとき 3 品目を扱っているとするとこの 3 品目の個数を時間変化を調べてみればそれなりに売れ行きが分かるでしょう。確かにこれは 3 次元の空間ベクトルですが、これに空間的な描像を対応させる人はあまりいないでしょう。ここではそれをさらに先鋭化させていきます。

線型空間

ここで一般的なベクトルの定義（線型空間の定義）をしてもう一度はっきりさせておきましょう。

Definition. 2.10.1. ある集合 V が \mathbb{C} 上の (複素) 線型空間であるとは、 $f, g \in V$ に対して和 (加法) と呼ばれる演算があつて $f + g$ が定まり、かつ $f + g \in V$ となり、複素数 α と $f \in V$ に対してスカラー積と呼ばれる演算があつて αf が定まり、かつ $\alpha f \in V$ となつて、これらが次の演算規則を満たすことをいう

1. $f + g = g + f$
2. $(f + g) + h = f + (g + h)$
3. 全ての f に対し、 $f + \theta = f$ を満たすベクトル θ がただ 1 つ存在する。
4. 全ての f に対し、 $f + f' = \theta$ を満たすベクトル f' がただ 1 つ存在する。
5. $1f = f$
6. $\alpha(\beta f) = (\alpha\beta)f$
7. $\alpha(f + g) = \alpha f + \alpha g$
8. $(\alpha + \beta)f = \alpha f + \beta f$

このとき 3 で存在を要請した θ を零ベクトルと言います。零ベクトルは概念的には数の 0 とは異なりますが、全く無関係でもありません。実際、 $0f = \theta$ が成立します。

これを示すには 8 から

$$0f = (0 + 0)f = 0f + 0f \quad (6.2.215)$$

この両辺に 4 で存在を要請した $(0f)'$ を加えて 2 と 3 を用いると

$$\theta = 0f + \{0f + (0f)'\} = 0f + \theta = 0f \quad \blacksquare \quad (6.2.216)$$

4 で存在を要請した f' を $-f$ と書き, f の逆元と呼びます. ここでマイナス記号を使って混乱が無いのは $(-1)f = -f$ が成立するからです. 実際

$$\theta = 0f = (1 - 1)f = 1f + (-1)f \quad (6.2.217)$$

$$\implies (-1)f = -f \quad \blacksquare \quad (6.2.218)$$

この他, 今まで使った来たベクトルに対する算法は基本的に全て成立することが証明できます.

そろそろベクトルに対して今までの x と異なり, 太字ですら無い f をつかっている理由を説明しましょう. ここから出てくるベクトルは無限次元を射程に入れた議論をしています. 主に扱うものは数列, 関数 (写像), 線型作用素です. ここで数列は自然数を定義域とした複素数を値域とする関数 (写像) ですから, 関数の取り扱いを基本とします. 本冊子で扱う線型作用素はほぼ微分作用素, 積分作用素です. 高校までに学んできた微分・積分は「演算」という印象が強かったと思います. ここでは作用という性質を前面に出します.

まずは次の定理を証明します.

Theorem. 2.10.2. 関数空間 $\mathcal{F}(\mathbb{C}, \mathbb{C})$ は線型空間である.

(証明) ここではじめの方に述べた「関数」と「関数値」の区別が重要になります. 先程の定義によれば和とスカラー倍を定義しないといけません. 早速定義しましょう. f, g, h を \mathbb{C}^n から \mathbb{C}^n への関数, α, β, z を複素数とします. このとき和とスカラー倍は

$$f + g \longrightarrow (f + g)(z) := f(z) + g(z) \quad (6.2.219)$$

$$\alpha f \longrightarrow (\alpha f)(z) := \alpha f(z) \quad (6.2.220)$$

という風に, 関数の作用を用いて定義します. 前に関数の対応や作用という面を強調したのはこの為です. また零元 (零ベクトル) θ は, 恒等的に 0 を与える関数 (零関数) です. つまり任意の z に対し $\theta(z) = 0$. ある関数 f が恒等的に 0 を与える (つまり零関数) とき, $f \equiv \theta$ と 3 本線で書くことがあります.

また 2 つの関数 f, g が等しいことも作用で定義します. つまり任意の複素数に対し $f(z) = g(z)$ が成立するとき関数として (ベクトルとして) f と g が等しいと言います. ベクトルであることの証明ですが, 関数値は複素数なので, 上に述べた性質を満たすのは明らかでしょう. 一度きちんとやってみます. 1 は

$$(f + g)(z) = f(z) + g(z) \quad (6.2.221)$$

$$= g(z) + f(z) \quad (6.2.222)$$

$$= (g + f)(z) \quad (6.2.223)$$

つまり

$$f + g = g + f \quad (6.2.224)$$

2 は,

$$\{(f+g)+h\}(z) = (f+g)(z) + h(z) \quad (6.2.225)$$

$$= f(z) + g(z) + h(z) \quad (6.2.226)$$

$$= f(z) + (g+h)(z) \quad (6.2.227)$$

$$= \{f+(g+h)\}(z) \quad (6.2.228)$$

つまり

$$(f+g)+h = f+(g+h) \quad (6.2.229)$$

3 は,

$$(f+\theta)(z) = f(z) + \theta(z) \quad (6.2.230)$$

$$= f(z) \quad (6.2.231)$$

つまり

$$f+\theta = f \quad (6.2.232)$$

4 は,

$$\{f+(-f)\}(z) = f(z) + (-f)(z) \quad (6.2.233)$$

$$= f(z) - f(z) \quad (6.2.234)$$

$$= 0 = \theta(z) \quad (6.2.235)$$

つまり

$$f+(-f) = \theta \quad (6.2.236)$$

確かにこのとき f の逆元 $-f$ は一意的です. 5 は,

$$(1f)(z) = 1 \times f(z) \quad (6.2.237)$$

$$= f(z) \quad (6.2.238)$$

つまり

$$1f = f \quad (6.2.239)$$

6 は,

$$\{\alpha(\beta f)\}(z) = \alpha(\beta f)(z) \quad (6.2.240)$$

$$= \alpha\beta f(z) \quad (6.2.241)$$

$$= (\alpha\beta)f(z) \quad (6.2.242)$$

つまり

$$\alpha(\beta f) = (\alpha\beta)f \quad (6.2.243)$$

7 は,

$$\{(\alpha + \beta)f\}(z) = (\alpha + \beta)f(z) \quad (6.2.244)$$

$$= \alpha f(z) + \beta f(z) \quad (6.2.245)$$

$$= \{(\alpha f + \beta f)\}(z) \quad (6.2.246)$$

つまり

$$(\alpha + \beta)f = \alpha f + \beta f \quad (6.2.247)$$

8 は,

$$\{(\alpha + \beta)f\}(z) = (\alpha + \beta)f(z) \quad (6.2.248)$$

$$= \alpha f(z) + \beta f(z) \quad (6.2.249)$$

$$= (\alpha f + \beta f)(z) \quad (6.2.250)$$

つまり

$$(\alpha + \beta)f = \alpha f + \beta f \quad (6.2.251)$$

以上で関数がベクトルである証明は終わりました。■ これで関数の集合が線型空間であることが示されました。線型空間としての関数の集合を関数空間と呼ぶことがあります。

作用素についても同じように出来ます。作用素を A とします。作用素の場合、あくまで「作用」なので作用する空間を設定しないといけません。

例えば微分作用素のとき、後でも使いますが、 r 回微分可能でかつ各階の導関数が連続な、 $\Omega \in \mathbb{R}$ から \mathbb{R} への関数全体の集合を $C^r(\Omega, \mathbb{R})$ と書きますが、その上での作用素、といったような位置付けになります。今は作用素が作用する空間として適当な空間 V を取ります。このとき作用素を A, B 、作用する空間の元を $f \in V$ を取り、作用素の和とスカラー倍は以下のように定義します。

$$A + B \longrightarrow (A + B)f := Af + Bf \quad (6.2.252)$$

$$\alpha A \longrightarrow (\alpha A)f := \alpha Af \quad (6.2.253)$$

線型独立と次元、基底 2

退屈な話が続きました。そろそろもう少し身のある話に行きます。ここからは今までに有限次元 (\mathbb{C}^n) でしてきた話を無限次元を視野に入れて話しなおします。

まずは線型独立性です。

Definition. 2.10.3. $V = \{f_\lambda; \lambda \in \Lambda\}$ の元が線型独立であるというのは任意の有限個の V の元 $f_{\lambda_1}, \dots, f_{\lambda_n}$ が線型独立であることをいう。

直接に無限個の一次独立性を言わないのは、無限個の元の線型結合を考えると級数の収束の問題が入ってきて線型結合の係数に余計な制限が加わるからです。

次に次元です。

Definition. 2.10.4. ある正数 K が存在して V の任意の線型独立な元の組 $\{f_i\}_{i=1}^r$ に対し $r \leq K$ となるとき、 V は有限次元であるという。一方、どんなに大きい正数 K をとってもある線型独立な元の組 $\{f_i\}_{i=1}^r$ が存在して $r \geq K$ となるとき、 V は無限次元であるという。

基底はあまり変わりありません。

Definition. 2.10.5. V の部分集合 $E = \{f_\lambda; \lambda \in \Lambda\}$ が基底であるというのは次の 2 条件が成り立つときである。

1. V の任意の元は E の元の線型結合として表される。
2. E は線型独立である。

ここで

Theorem. 2.10.6. 関数空間は無限次元である。

を証明しましょう。関数空間としては、一旦実数から実数への関数の集合を取ります。この部分集合としては $f_n(x) = x^n$ となる関数の集合 $\mathcal{F} = \{f_n; n \in \mathbb{N}\}$ を取ります。 \mathcal{F} は無限個の元からなっています。したがってこの任意の有限個の元の線型結合がベクトルとしての零、つまり零関数 θ であることを示せば良いわけです。簡単のため、 $g = c_0 f_0 + c_1 f_1 + \cdots + c_n f_n$ を考えましょう。たまたま g は微分できるのでそれを使いましょう。 $g(x) = 0$ の両辺を k 階微分して 0 を代入すると $g^{(k)}(0) = n(n-1)\cdots(n-k+1)c_k = 0$ となりますから、任意の k について $c_k = 0$ です。したがって $g = \theta$ を示すことが出来ました。■

ついでに言うておくと、この f_n は適当な V の部分空間を取ると、その基底になることがあります。このあたりは後で Taylor 展開と絡めて説明する予定です。

内積

次に内積の話をしてします。無限次元での内積をどう定義するか、ということですが、まずは有限次元を思い出しましょう。 $x, y \in \mathbb{C}^n$ とします。このとき内積は

$$(x, y) = \sum_{i=1}^n x_i \bar{y}_i \quad (6.2.254)$$

で定義しました。ここで数列 $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ を考えます。これを縦に無限に長い列ベクトルのようにみなし、上の内積の定義の n を ∞ を無限大に飛ばしたものを考えると、内積のようなものが出来ました。

$$(a_n, b_n) := \sum_{n=1}^{\infty} a_n b_n \quad (6.2.255)$$

無限級数になってしまったので、収束の問題が真っ先に登場します。以後は和が上手く収束するようなものだけ考えましょう。これを l^2 空間といいます。 l^2 と略して言うこともあります。「空間」とつけてしまいましたが、これは問題があります。何故かと言うと、これは数列全体の部分集合にはなりませんが、部分空間になるか分かりません。きちんと部分空間になることを示しましょう。後できちんと書くか分かりませんが、

Theorem. 2.10.7. 上に有界な単調増加実数列は収束する。

Theorem. 2.10.8. 絶対収束する数列は収束する。

という 2 つの定理を拝借してきます。

部分空間の定義から和とスカラー倍がまた l^2 の元でないといけません。スカラー倍については良いでしょう。和を考えます。

$$|a_n + b_n|^2 \leq 2(|a_n|^2 + |b_n|^2) \quad (6.2.256)$$

という簡単な不等式を使います。両辺の和を取ると、まず右辺は収束します。左辺は正数の和ですから単調増加で、さらに収束級数の和で抑えられていますから上に有界です。上の定理によりこれは収束します。これで l^2 が線型空間であることが分かりました。■

次は l^2 の元の内積がきちんと収束することを示さないといけません。こちらでもまた初等的な不等式ですが、

$$|a_n b_n| \leq (|a_n| + |b_n|)^2 \leq 2(|a_n|^2 + |b_n|^2) \quad (6.2.257)$$

を使います。これもまた和を取ると右辺は収束します。先程同様これから左辺が収束しますが、これは絶対値の級数なので、絶対収束しています。したがって元の級数、すなわち内積も収束します。■

ちなみに「内積」とは言いますが、有限次元の内積を持つような性質を無限次元でも同様に持っていてももらいたいわけです。それでは有限次元の内積を持つ性質とは何でしょうか？ これは以下に述べるようなものです。それを定義として採用します。

Definition. 2.10.9. 内積とは線型空間 V の 2 元 f, g に対して複素数を対応させ、かつ以下の性質を満たす写像である。

$$(f, f) \geq 0 \quad (6.2.258)$$

$$(f, f) = 0 \iff f = \theta \quad (6.2.259)$$

$$(f, g) = \overline{(g, f)} \quad (6.2.260)$$

$$(\alpha f_1 + \beta f_2, g) = \alpha(f_1, g) + \beta(f_2, g) \quad (6.2.261)$$

上の 2 つの性質は正定値性と呼ばれます。これを使うと内積は正定値性を持つ双線型写像であると言えます。

また記法として主に無限次元のベクトルに対して、今後ベクトルの長さを

$$\|f\| := (f, f) \quad (6.2.262)$$

と書きます。こう書くときは長さではなくノルムと呼びます。

内積から以下のようにして角度（正確には余弦）が定義できました。

$$\cos \theta = \frac{(f, g)}{\|f\| \|g\|} \quad (6.2.263)$$

有限次元でこれが定義できるには $\cos \theta \leq 1$ から $|(f, g)| \leq \|f\| \|g\|$ を満たさないとはいけなかったわけですが、これは Cauchy-Schwarz の不等式で確かに成立していました。Cauchy-Schwarz の不等式を再録しましょう。

$$|(f, g)| \leq \|f\| \|g\| \quad (6.2.264)$$

無限次元と言わず、有限次元でも一般に内積が複素数なので、 $\cos \theta$ が複素数値になってしまいます。（複素）関数論を学んだあとなら別にあまり問題もないと思うのですが、それなりに困ったものではありません。しかも正規直交基底の定義に内積を使っているのに、内積は無いと困ります。ここで方向転換をしまして、基本的に内積で角度を測るときは直交しているかしていないか、つまり内積が 0 かどうかだけ見ているものだと思しましょう。無条件ではありませんが、無限次元でも Cauchy-Schwarz の不等式は成立します。

l^2 で考えましょう。テクニカルで腹立たしいですが、紙数の都合もあるのでこんなものを考えます。

$$\sum_{n=1}^{\infty} |ta_n + b_n|^2 = t^2 \sum_{n=1}^{\infty} |a_n|^2 + 2t \sum_{n=1}^{\infty} |a_n| |b_n| + \sum_{n=1}^{\infty} |b_n|^2 \quad (6.2.265)$$

これを t の 2 次式と考えましょう。 $\sum_{n=1}^{\infty} |a_n|^2 = 0$ のときは $(a_n)_{n \in \mathbb{N}} = \theta$ ですから、Cauchy-Schwarz の両辺は 0 で成立します。0 でないとき判別式が負とならねばならないわけですが、これがそのまま Cauchy-Schwarz の不等式です。■

というわけで内積が上手く定義できることがわかります。ここから正規直交基底なども定義できます。

次は内積を関数について定義してみましょう。手がかりはもちろん l^2 です。高校で区分求積法を学んだはずですが、それを使います。

$$\int_0^1 f(x) \overline{g(x)} dx = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f\left(\frac{k}{n}\right) \overline{g\left(\frac{k}{n}\right)} \quad (6.2.266)$$

積分区間が $(0, 1)$ である必要は無いのですが、ひとまず書いてみました。とりあえず無限級数さえ出てきてくれれば良いです。 a として $-\infty$ を b として ∞ を許して、関数の内積を

$$(f, g) := \int_a^b f(x) \overline{g(x)} dx \quad (6.2.267)$$

とします。こうするとこれは内積の性質を全て持ちます。当然これから直交などという事も色々できます。Lebesgue の意味で 2 乗絶対可積分な関数空間にこの内積を入れた空間を L^2 空間といいます。本冊子ではとても触れることは出来ませんが、これまでに述べた l^2 や L^2 空間は Hilbert (ヒルベルト) 空間と呼ばれるものの具体例です。これが量子力学の数理の主戦場です。可分な Hilbert 空間は全て同型であるという同型定理と呼ばれるものがありますが、 l^2 , L^2 空間はそれぞれ可分な Hilbert 空間です。量子力学は当初 Heisenberg (ハイゼンベルク) の行列力学と Schrödinger (シュレディンガー) の波動力学という形式があって、これが物理的な解釈なり何なり色々異なって大論争があったそうですが、物理的な結果は全く同じものを与えるので非常に不思議がられていたそうですが、これは数学的に行列力学は l^2 で考え、波動力学 L^2 で考えていたので、同型定理から数学的には同じ結果が出てもおかしくない、とかいう紆余曲折があったそうです。

直交多項式と基底

打って変わって基底のところでは「この f_n は適当な V の部分空間を取ると、その基底になることがあります」と書きました。そのような基底になる直交多項式としていくつか挙げておきましょう。

$$P_n(x) := \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (6.2.268)$$

$$H_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \quad (6.2.269)$$

$$L_n^\alpha(x) := \frac{d^n}{dx^n} (e^{-x} x^{n+\alpha}) \quad (6.2.270)$$

$$J_\nu(z) := \sum_{k=0}^{\infty} \frac{(-1)^k \left\{\frac{z}{2}\right\}^{\nu+2k}}{\Gamma(k+1)\Gamma(k+\nu+1)}, \quad \Gamma(z) := \int_0^{\infty} e^{-t} t^{z-1} dt \quad (6.2.271)$$

これらははじめから順に Legendre (ルジャンドル) 多項式, Hermite (エルミート) 多項式, Laguerre (ラゲール) 多項式, Bessel (ベッセル) 多項式, 最後のおまけは Γ 関数といいます. 本当はもう少し言わないといけないことがありますが, 紙数の都合もあり略します.

基底といえば, 我々は今振動・波動の数理を扱っているものであり, 振動・波動と言えば三角関数であり, 振動・波動と言えば重ね合わせの原理というのもあり, ということで基底として三角関数を選んで, それで波動を表す関数を三角関数の重ね合わせでかけると嬉しいな, と思ってみたりするわけですが, それは既に昔々(確かフランス人の) Fourier(フーリエ) さんが考えていまして, 関数を三角関数の重ね合わせで書く事を関数を Fourier 級数展開 (または Fourier 変換) するといいます. 紙数の余裕があればあとでやりますが, とりあえず展開式だけ書いておきましょう.

$$f(x) \sim \frac{1}{2}a_0 + \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx) \quad (6.2.272)$$

$$a_n := \int_{-\pi}^{\pi} f(t) \cos ntdt, \quad b_n := \int_{-\pi}^{\pi} f(t) \sin ntdt \quad (6.2.273)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\xi) e^{i\xi x} d\xi \quad (6.2.274)$$

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \quad (6.2.275)$$

微分方程式と Fourier 解析とはじめ

そろそろ筆者の知識と認識の限界に来ているのですが, もう少し書きましょう. 線型作用素についてです. t を時間として微分作用素を $D := \frac{d}{dt}$ とします. $D^2 = \frac{d^2}{dt^2}$ としましょう. 見やすくするため, $x(t) = f(t)$ と書きます. 運動方程式は以下のような方程式になります.

$$D^2 f = -\omega^2 f, \quad \omega := \sqrt{\frac{k}{m}} \quad (6.2.276)$$

これをもって何とするか, ですが, 前に $Ax = \lambda x$ というのをやりました. 固有値の話です. 全く同じ形をしています. すなわち運動方程式 (微分方程式) は無限次元空間での固有値問題を考えている, ということです. あとで証明しますが, ある条件下で微分方程式には解の存在と一意性が言えます. いつであっても, 適当な場所から適当な速度で放り出した物体が同一条件下では同じ運動をしてくれないと困るのですが, この定理は数学的にもきちんとそれが示された, という点で非常に有意義です.

上で述べたように解が存在してしかも一意的です. そんな訳で解を見つけてこれれば, それで全て片がつきます. 解自体は高校の段階で既に見つかっています. 三角関数でした. ここでそれを知らないとしてどうするかを考えてみましょう. 1 回微分して自分自身に戻る関数として, 指数関数 $g(x) = e^{kx}$ があります. これを運動方程式に代入してみましょう.

$$D^2 g = k^2 g = -\omega^2 g \quad (6.2.277)$$

となるわけです. となると, $k^2 = -\omega^2$ となる数を探してくれば答が見つかります. 2 乗して負になる数として虚数単位の i さんがいらっやいました. したがって $k = \pm i\omega$ です. ここで実は何処かの誰かがこの方程式の解として三角関数があ

る, ということを教えてくれたとしましょう. 実は 2 階の常微分方程式の解は 2 次元の線型空間になっていて, $\cos \omega t, \sin \omega t$ は基底になっています一意性から $e^{i\omega x}$ は三角関数の線型結合で書けないと困ります. しかも虚数が入っているので係数に虚数を含んだ線型結合になるでしょう. 結果としてはご存知の方もたくさんいることでしょう. Euler (オイラー) の公式という次式になります.

$$e^{ix} = \cos x + i \sin x \quad (6.2.278)$$

虚数乗とはこれいかに, といった問題も色々あって結構困るのですが, 関数論の解析接続というものがあって, これはきちんと数学的に正当化されます. 今のところはこれが虚数乗の定義だとでも思っておいて下さい. 微分積分の Taylor 展開のところでも再び登場させるつもりです.

あまりにもとんでもないところまで来てしまいましたが, これは本冊子の目標到達点です. そして非常に難しいので, 現段階で分からなくても全く問題ありません.

それではそろそろ微分積分に行きましょう.

Chapter 7

Application to Google's pagerank

動画の原稿から取ってくる.

Chapter 8

量子力学を理解するために斎藤 正彦『線型代数入門』[184]を読 む人が注目すべき点

Togetter からの転載.

[184] について見るべき所を挙げておきたい. まず p120 からが決定的に重要. 例 3 で関数空間 (多項式空間) に内積を叩き込んでいる. 量子力学で出てくる線型空間は原則関数空間なので, 正にここを使う. ヒルベルトとか鬱陶しいのはひとまずいい.

p122 例 5:ルジャンドル多項式. 多項式空間での基底としてのルジャンドルを出してきている. 関数のノルム (ベクトルの「長さ」) も出している. ルジャンドルは電磁気でも出てくる. マクスウェルは電磁場の「線型」方程式なので上手くはまってくれる.

p126 例 7, 有限フーリエ級数. フーリエは知っているとは仮定するが, 正にフーリエが線型代数の枠内で議論出来る (部分もある) 事を明示している. 三角関数が (適当な内積空間での) 「正規直交基底」になる. これは高校の教科書にも計算問題としては書いてある.

問題 7. これは量子統計で使う. (平衡) 状態はオブザーバブル A に対して適当な複素数を返す関数だが, これは必ず (密度) 行列を使ってトレースで書けることを言っている. ただしうるさいことを言うと原則有限次元でしか使えないので, 物理としては役に立たないと言い切ってもいいのだが.

p128 問題 8:ある種の変分とか何とか. 問題 10, 色々な内積とラゲール多項式. 他にもこの系で電磁気・量子力学で出てくる「 \sim 多項式」が再現される. 問題 12, 双対空間:初学者にとって直接どうというのが言いづらいが, 少し進んだ話をするとなんかそれに使うはず. 数学としては大事.

第 5 章:量子力学の魂. p137 例 8, 数列空間:数列も線型代数の枠内で議論できる (部分がある) という事.

p138 例 9. 線型常微分方程式. 先程から電磁気などそれっぽいのを挙げているが, 微分方程式論を応用先に持つことがここではっきりと分かる.

p144:スペクトル分解. 要するに対角化だが, スペクトル分解は無限次元化出来て量子力学で直接使える. 射影の値域がその固有値の固有空間で, 縮退度が固有空間の次元に対応したりとか何とか.

定理 2.10:行列の極分解. 量子力学的な解釈としては行列は大体複素数と思える. 複素数ならば極形式で表現出来るか考えたくなるのが人情らしいのでそれを

やってみた.

定理 2.11: 変分原理. 量子力学で基底エネルギーを評価するときに使う.

第 7 章. 行列の関数を定義するという荒業. ハミルトニアンを H とすると, その系の時間発展は $U_t = e^{itH}$ となる. ここで行列の指数関数が出てくる. あと量子光学でコヒーレンス扱うときもこの辺が出てくる.

量子力学ではより激しく微分作用素の指数関数も出てくる: 運動量作用素を $p = -id/dx$ として e^{itp} が出てくる. この場合, 数学的に正確な定義は面倒だが, とりあえず指数のテイラー展開に直接代入で「納得」されたい. ついでに言うとこれは本当にテイラー展開になる.

p217, ペロン・フロベニウスの定理. 正確に言うと別のバージョンだが, 固体物性, 磁性のモデルで使うハバードモデルの「研究」で本当に使う. 詳しくは田崎さんの本 [225] を参考にされたい. Google のページランクでも使う定理で応用は広い.

Part V

Calculus

Chapter 9

Is 0 area for a point, line segment, line?

動画から適当に切り出してくる.

Chapter 10

あとで大幅 rewrite: ポテンシャルの振る舞いと微分積分

10.1 導入

前章ではばねにつながれた質点の運動の解析を行ないました。本章では一般のポテンシャルにしたがう力学系 (多数の質点) の運動の解析を行ないます。

まずポテンシャルとは何かですけれども、系のエネルギーを書いたとき、運動エネルギー以外に $\frac{1}{2}kx^2$, $-G\frac{Mm}{r}$ などの項が現われることがあります。そのことです。正確には力 f に対し

$$\mathbf{f} = \frac{\partial U}{\partial x} \mathbf{e}_x + \frac{\partial U}{\partial y} \mathbf{e}_y + \frac{\partial U}{\partial z} \mathbf{e}_z \quad (10.1.1)$$

となる実数値関数 U が存在するとき、 f はポテンシャル U を持つといいます。一般にポテンシャルは U, V などと表します。ここでは一般のポテンシャルにしたがう力学系がつりあいの状態から少しずらすとその系は微小振動をすることを示します。

具体的な状況を考えてみましょう。先程のばねではつりあいの状態はばねの長さが 0 となるところです。これは振幅の小さいところでは振動します。これ以外には、例えば振り子の振動を考えましょう。あまり高くない適当な位置から振り子を静かに放すと、摩擦などが無いとすればずっと振動を続けるでしょう。鉛直方向から少しだけずらしたところでは、完全に単振動となることは高校で学んだと思います。他には放物線 $y = x^2$ に沿って質点を静かに放すと、これも摩擦が無いとすれば放物線に沿って転がりつづけるでしょう。ある程度の高さまで上ったらまた放物線に沿って戻ってくるということも問題ないでしょう。

上の例では適当に座標系を取ると、原点でつりあいの状態になります。これらは全て安定なつりあいになっています。安定なつりあいというのは、少しずらしたただけならばまたその元の状態に自然に戻ろうとするような状態のことです。不安定なつりあいを考えてみると良く分かるでしょう。 $y = -x^2$ の放物線に沿った運動では原点はつりあいの点ですが、少し動かすと放物線に沿って転がり落ちていき、原点に自然に戻ってくることはありません。これが不安定な釣り合いです。

少々唐突ですが、今の状況を見ると 1 変数では下に凸な曲線に従った運動は安定であるといえそうです。もちろん $y = x^2$ や (半) 円が基礎です。ただ、これを高次元に一般化しようと思うと大変です。大変そうな具体例が作れます。

例えば $z = f(x, y) = x^2 - y^2$ を考えます。これの微分をどう考えるか、という問題がまずありますが、それよりも比較的楽に図が描けるので、その図を元に考えます。はじめに言うておくと、原点で変な挙動を取ります。

原点の近くを x 軸に沿って進んでみましょう。つまり $z = f(x, 0) = x^2$ に沿って進みます。すると z は原点で極小になります。一方で y 軸に沿って進んでみましょう。つまり $z = f(0, y) = -y^2$ に沿って進みます。すると今度は z が原点で極大になります。ある方向から近づくと極大値を取り、ある方向から近づくと極小になるような点が存在します。また実はこの点で微分が 0 になりますから、微分が 0 になり、かつ極大とも極小ともいえないような変な点が存在することになります。ちなみにこの点を鞍点といいます。

次に $z = y^2$ を z 軸の周りに回転させたものを考え、これを $z = f(x, y)$ としましょう。これは簡単に想像できるでしょう。このとき、直観的に原点が安定点になることが分かります。先程との比較で言うと、こちらはどの方向から原点に近づいても $z = f(x, y)$ は極小になる事がわかります。

以上から、どの方向から原点に近づいても、というような条件を数学的にどう定式化していくかが鍵になりそうです。

そして私たちは前章で線型代数を学びました。出来ればそれを上手く使っていけると嬉しいわけです。その方法として線型化ということを考えます。高校で微分を用いた一次近似式を学んだと思います。接線、といえばもっとはっきりするでしょう。これは曲がったものでも、ある点の十分近くで考えればそれはまっすぐとみなせる、ということです。これは日常経験で常に感じていることです。

私たちは地球表面上に住んでいます。地球は楕円形です。要するに曲がっているわけですが、普段私たちは地面が曲がっているとは思わず、まっすぐになっていると感じています。これは曲がったものを十分小さいところで考えてまっすぐのものともみなす考えは、日常的な行為であり直観的にも正当なものであると考えられます。

私たちはこれを線型化と呼びましょう。そして上に述べた直観を数学の理論として取り込むべく、線型化法として微分を考えてみましょう。せっかくなので本章では微分だけでなく積分も説明します。こちらは微分方程式の実際の解法を考えるとから導入してみたいと思います。

10.2 $\varepsilon - \delta$ 論法と収束

10.2.1 $\varepsilon - \delta$ 論法

これから微分・積分を論じます。積分は微分の逆、という感じであまり極限の香りがしないと思っている方もいるかもしれませんが。しかし微分を考えれば分かるように、これから極限についての考察をすることになります。場合によっては非常に精密な取り扱いを要求されることがあります。その際、高校での極めていい加減な極限の定義では、ほとんど全く何も議論できません。そこでもっと議論しやすくするために極限の精密な定義を与えます。

$\lim_{n \rightarrow \infty} a_n = a$ というものを考えましょう。これは高校では「 n が限りなく大きくなる時 a_n は a に限りなく近づく」ということを意味している、という事でした。これを再考します。

今何か適当な実験、例えば重力加速度の測定実験をしてみましょう。まず考えるべきは実験の精度をどこまで高めたいか、という事です。高校でも実験をしてきたでしょうが、こうした実験では誤差がある程度大きくても許されるでしょう。一方で値の正確さを競うような研究の現場のレベルの実験においては可能な限り精度を高めねばならず、京大に国際重力基準点というのがあるそうですが、そ

こで行なっている実験では 9 桁の精度で重力加速度を測定しているようです。さらには実験を行なう際、高出力を出し過ぎると危険な装置もありますし、何処かの施設に行って時間制限のある中でやらねばならないこともあるでしょう。この際、これぐらいの結果を出したい（出力したい）がそのためにはどれぐらいの入力を行なえばよいかということを経験にきっちり決めておかないといけません。

ここで前章に述べた、「数列は関数の一種であり、関数とは対応である」ということを思い出しましょう。各種の入力（気温や湿度、装置の精度など）に対しある実験データが出てきますから、その出力の平均のようなものを対応させる関数（数列）を考えます。このとき、各種の入力が精度の順で適当に順序付けできるものとします。ここで上に述べたことを実行するには、出力に対しその出力を得られるような上手い入力を見つけることが重要になります。例えば理論で予測される値が a であったとして、 a の付近に適当な誤差で値が散らばるような実験が行ないたいものとします。

精度の程度（レベル）を自然数で n とし、それに対し実験値を与える対応（数列）を $(a_n)_{n \in \mathbb{N}}$ としましょう。理論値と実験で得るべき値との誤差を $\varepsilon > 0$ とします。すなわち $|a_n - a| < \varepsilon$ となってくれば良いわけです。今 a の値を実現するような n が存在したとし、それを N とします。 n の値については、ある精度の実験をすれば上の不等式を満たすのならば、それ以上の精度では必ず上の不等式を満たすでしょう。つまり適当な自然数 N が存在して $n \geq N$ ならば $|a_n - a| < \varepsilon$ となるはずです。

以上をまとめると、誤差範囲 $\varepsilon > 0$ を決めたのち、ある精度レベル N の実験で結果が上手くその誤差の範囲に収まるならば、その精度以上の実験 $n \geq N$ をすれば、実験結果 a_n と出したい値 a の差は $|a_n - a| < \varepsilon$ に収まってくれるだろう、ということです。ここで $\varepsilon \rightarrow 0$ をすると、誤差が「限りなく」小さくなるので、それに対応して入力も極めてシャープにせざるを得ず、 $N \rightarrow \infty$ となるでしょう。こうした現実を踏まえ、これ以降極限を次の定義で置き換えます。

Definition. 2.1.1. (ε - δ 論法 1) 実数列 $(a_n)_{n \in \mathbb{N}}$ が実数 a に収束するとは、任意の正数 ε に対して自然数 N が存在し、 $n \geq N$ を満たす全ての n に対し $|a_n - a| < \varepsilon$ が成立することをいう。このとき a は $(a_n)_{n \in \mathbb{N}}$ の極限であるといい、

$$a = \lim_{n \rightarrow \infty} a_n, \quad a_n \rightarrow a \quad (n \rightarrow \infty) \quad (10.2.1)$$

などとかく。

これを関数について言うとな下のようにになります。

Definition. 2.1.2. (ε - δ 論法 2) 区間 I 上定義された実数値関数 f が $x \rightarrow a$ で α に収束するとは、任意の正数 ε に対して実数 $\delta > 0$ が存在し、 $|x - a| < \delta$ を満たす全ての x に対し $|f(x) - \alpha| < \varepsilon$ が成立することをいう。このとき α は $f(x)$ の極限であるといい、

$$\alpha = \lim_{x \rightarrow a} f(x), \quad f(x) \rightarrow \alpha \quad (x \rightarrow a) \quad (10.2.2)$$

などと

最後に極限について 1 つ注意をしておきます。それは一般に極限をとる順序は交換出来ないということです。例えば二重数列 $a_{mn} = \frac{m}{m+n}$ というものを考えてみると

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} a_{mn} = \lim_{m \rightarrow \infty} 0 = 0, \quad \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} a_{mn} = \lim_{n \rightarrow \infty} 1 = 1 \quad (10.2.3)$$

$$\implies \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} a_{mn} \neq \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} a_{mn} \quad (10.2.4)$$

となります。物理にもよく出てくる極限の順序交換は微分と積分の順序交換、多変数の高階微分の順序交換です。物理をやっている場合、普通は性質の良いものを扱うからこうした順序交換が許される、という仮定の元でどんどん計算を進めていきますが、量子力学では、こうした演算の順序が交換出来ないということが本質的な条件として登場してきますので、一応注意を喚起しておきます。

10.2.2 Cauchy 列と実数の性質

高校までは基本的に具体的な数列を扱い (例えば $a_n = 1/n$), かつ極限も比較的楽に求まるようなものばかり扱ってきたと思います。しかしこれからはかなり一般的な数列を扱い、さらに具体的な数列であろうとも極限がまともに求まらないような数列まで扱うようになります。後で色々証明しますが、収束するだけことは分かっても収束する値が全く分からないことも良くあります。その際、収束先が分かっているのに上の極限の定義で極限値を表に出すのは気が引けるというものです。それで以下の Cauchy (コーシー) 列を導入します。

Definition. 2.2.1. 数列 $(a_n)_{n \in \mathbb{N}}$ が *Cauchy 列* (基本列) であるとは、任意の $\varepsilon > 0$ に対し自然数 N が存在して $n, m \geq N$ ならば $|a_n - a_m| < \varepsilon$ となることを言う。

すぐあとで使うこともあり、解析学で良く使う基本的な不等式である三角不等式を載せておきます。簡単に証明できるので証明は各自行なってみてください。

$$|a + b| \leq |a| + |b|, \quad ||a| - |b|| \leq |a - b| \quad (10.2.5)$$

ここで収束列が Cauchy 列であることを証明します。 $(a_n)_{n \in \mathbb{N}}$ が収束列で、 a に収束するものとする、収束の定義から任意の $\varepsilon > 0$ に対して自然数 N が存在して、 $n, m \geq N$ ならば $|a_n - a| < \varepsilon/2, |a_m - a| < \varepsilon/2$ が成立します。三角不等式を用いて

$$|a_m - a_n| \leq |a_m - a| + |a - a_n| < \varepsilon/2 + \varepsilon/2 = \varepsilon. \blacksquare \quad (10.2.6)$$

問題はこの逆、Cauchy 列は収束列であるが成立するかどうかです。Cauchy 列というのは、 n と m が十分大きければ a_n と a_m がいくらでも近い値を取るということなので、直観的にはこれが満たされそうな気がしますが、例えば有理数の集合内で考えたとき Cauchy 列は収束しないことがあります。例えば数列

$$a_1 = 1, \quad a_{n+1} = \frac{2a_n + 2}{a_n + 2} \quad (10.2.7)$$

を考えます。これが a に収束するとすれば、漸化式の両辺で $n \rightarrow \infty$ とすれば、 $a = \frac{2a+2}{a+2}$ から $a = \sqrt{2}$ を得ます。 $(a_n)_{n \in \mathbb{N}}$ は初項が 1 で有理数、漸化式から全ての n について a_n が有理数ですが、収束先の a は無理数です。

当然この原因は窮屈な有理数の中だけで考えている為に起こる現象です。したがって実数列ではこのようなことが起こらないという要請をおきましょう。

Axiom. 10.2.1. *Cauchy 列は収束列である。*

この他にももう 1 つの性質を要請します。

Axiom. 10.2.2. (*Archimedes の原理*) 任意の実数 $a, b > 0$ に対し、 $na > b$ となる自然数 n が存在する。

この Archimedes (アルキメデス) の原理は以下の式と同等です.

$$\lim_{n \rightarrow \infty} n = \infty, \quad \lim_{n \rightarrow \infty} 2^n = \infty. \quad (10.2.8)$$

これと同値な連続の公理を述べるのにいくつか新しい概念が必要なので, それを紹介します. 1 つは最大値を一般化した上限と呼ばれるものです. まず最大値をきちんと定義しておきます.

Definition. 2.2.2. ある集合 A の最大値が M であるというのは, $M \in A$ であって, 任意の $x \in A$ に対して $x \leq M$ となるものである. 最小値も同様に定義する.

重要なのは $M \in A$ というところです. 具体的に开区間 $I = (0, 1)$ を考えてみましょう. I の最大値とは何でしょうか? 当然真っ先に思いつくのは 1 という答ですが, $1 \notin I$ であり, 1 は最大値ではありません. しかし最大値のようなものを持ち出して議論をしたいときがあり, 1 に最大値のような資格を与えておくとう利なことが多々あります. これを定式化したものが上限です. その前に上界を定義しておきます.

Definition. 2.2.3. u が A の上界であるというのは, 任意の $a \in A$ に対して $a \leq u$ となる u のことである. A の上界全体の集合を $U(A)$ と書く. さらに下界も同様に定義する. こちらは $L(A)$ と書く.

ここで上界には必ず最小元が存在します. この最小元を上限と定義します.

Definition. 2.2.4. 上限とは上界の最小の元, $\min U(A)$ のことである. これを $\sup U(A)$ と書く. 下限も同様に定義する. こちらは $\inf L(A)$ と書く.

これを用いてもう 1 つ新しい概念を導入します.

Definition. 2.2.5. ある集合 A が上に有界であるというのは, 有限な上界が存在することである. すなわち $U(A) \neq \emptyset$ である. 下に有界も同様に定義する. 上下に有界な集合を単に有界であるという.

直観的には $\pm\infty$ を含まない, といったようなことです.

さて, やつとのことで連続の公理です. いくつかあるのでそれを列挙します.

Axiom. 10.2.3. (Weierstrass) 上に有界な空でない集合 $A \neq \emptyset$ は上限を持つ.

上に有界というのは A が一定以上の大きな値を元に持たないということです. から, 直観的にはこれも正しいと感じられるのではないのでしょうか.

Axiom. 10.2.4. 1) 上に有界な単調増加数列 $(a_n)_{n \in \mathbb{N}}$ は収束し, 極限は $\sup a_n$ に等しい.

2) 下に有界な単調減少数列 $(b_n)_{n \in \mathbb{N}}$ は収束し, 極限は $\inf b_n$ に等しい.

これは上に有界ならば, 数列の各項はある値以上にはなれません. さらに単調増加なので, 数列の各項はそのある値にいくらでも近くなれるでしょう. ただ近くなるだけでなく, 実際にその値に収束までしてしまうことを述べた (要請した) ものです.

Axiom. 10.2.5. (区間縮小法) $I_n := [a_n, b_n]$ で $(a_n)_{n \in \mathbb{N}}$ は単調増加数列, $(b_n)_{n \in \mathbb{N}}$ は単調減少数列であるとする. すなわち $a_1 \leq a_2 \leq \cdots a_n \leq b_n \leq \cdots b_2 \leq b_1$ であり, 全ての自然数 n に対して $I_n \supset I_{n+1}$ であるとする. このとき $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$ ならば, I_n の共通部分は 1 点 a のみからなる.

これはある区間 $I = [a, b]$ には隙間無く実数が詰まっていることを表しています。有理数と対比させると分かりやすいでしょう。

Axiom. 10.2.6. (*Bolzano-Weierstrass*) 有界数列は収束部分列を持つ。

数列 $(a_n)_{n \in \mathbb{N}}$ の部分列というのは $(a_n)_{n \in \mathbb{N}}$ の各項を a_1, a_5, a_7, \dots などのように適当に選んできたとき、上手い選び方をすると、もとの数列が何であろうとも有界でさえあれば、その部分列が上手く収束することを述べています。例えば $a_n = (-1)^n$ を考えると a_n 自体は収束しませんが、偶数番目、奇数番目だけ集めた数列 $(a_{2n})_{n \in \mathbb{N}}, (a_{2n+1})_{n \in \mathbb{N}}$ はそれぞれ $+1, -1$ に収束します。これらの公理は全て同値ですが、この同値性を証明するのは面倒なので略します。

10.2.3 数列, 級数の収束

本冊子では全ての定理を完全に証明しきることは出来ませんが、あとで登場する Taylor の定理ぐらいはきちんと証明したいので、そのための準備をします。ただひたすらに証明を行ないます。3 つ言葉を定義しておきます。各項が正、すなわち任意の $n \in \mathbb{N}$ に対して $a_n \geq 0$ となる級数 $\sum a_n$ を正項級数と呼びます。さらに数列 (一般に関数などでも良い) が上に有界であるとは定数 $M \in \mathbb{R}$ が存在し、任意の $n \in \mathbb{N}$ に対し $a_n \leq M$ となることをいいます。また x のべき x^n からなる級数 $\sum a_n x^n$ をべき級数といいます。

Theorem. 2.3.1. (級数に関する Cauchy の収束条件) 級数 $\sum a_n$ に対し次の 2 条件は互いに同値である。

- 1) $\sum a_n$ は収束する。すなわち部分和の数列 $(s_n)_{n \in \mathbb{N}}$ が収束する。
- 2) 任意の $\varepsilon > 0$ に対してある $n_0 \in \mathbb{N}$ が存在して、 $n > m \geq n_0$ となる全ての自然数 n, m に対して $|a_{m+1} + \dots + a_n| < \varepsilon$ となる。

(証明)

1. は $(s_n)_{n \in \mathbb{N}}$ が Cauchy 列であるということです。

Cauchy 列は収束することを公理として要請していますから、定理の statement が成立します。■

Theorem. 2.3.2. $\sum |a_n|$ が収束するとき、 $\sum a_n$ も収束する。

(証明) 三角不等式から $|a_{m+1} + \dots + a_n| \leq |a_{m+1}| + \dots + |a_n|$ ですから、 $\sum |a_n|$ が前定理の Cauchy の収束条件を満たすとき、 $\sum a_n$ も Cauchy の条件を満たします。■ またこの定理を元にして、級数 $\sum a_n$ が絶対収束するというのは各項を絶対値にした級数 $\sum |a_n|$ が収束することを言います。

Theorem. 2.3.3. 正項級数 $\sum a_n$ が収束するためには部分和の数列 $(s_n)_{n \in \mathbb{N}}$ が上に有界となることが必要十分である。

(証明) まず十分性から示します。部分和の数列とは $s_n = \sum_{k=0}^n a_k$ のことでした。任意の $n \in \mathbb{N}$ に対し、 $a_n \geq 0$ から部分和の数列 $(s_n)_{n \in \mathbb{N}}$ が上に有界ならば「上に有界な単調増加列は収束する」という公理から、この数列は収束します。

次に必要性を示しましょう。 $(s_n)_{n \in \mathbb{N}}$ が収束するならば、収束先を s とすると定義により任意の ε に対し $N \in \mathbb{N}$ が存在し、 $n \leq N$ ならば $|s_n - s| < \varepsilon$ となります。 ε を適当に固定します。色々な本を見ると大抵 1 にとってあるでしょう。別

に何でもよいのでここでは $\varepsilon = 10^{-2098}$ とでもしておきます. このときにも自然数 N_0 が存在して $n \leq N_0$ ならば $|s_n - s| < 10^{-2098}$ となります. これから

$$n \leq N_0 \implies s_n < s + 10^{-2098}. \quad (10.2.9)$$

ここで $M := \max\{|s_0|, |s_1|, \dots, |s_{N_0-1}|, s + 10^{-2098}\}$ としましょう. ここで $\max\{\dots\}$ とは $\{\dots\}$ の中の最大数を表すものとします. 第 n 項 s_n を考えたとき, $n \leq N_0$ ならば \max の定義から $s_n \leq M$ で, $n \geq N_0$ でも $s_n < s + 10^{-2098} \leq M$ であり, 確かに上に有界です. これは

Theorem. 2.3.4. 収束する数列は有界である

を示したことになります. ■

Theorem. 2.3.5. (比較定理) $\sum a_n, \sum c_n$ は正項級数で $\sum c_n$ は収束するものとする. このとき全ての n に対し $a_n \geq c_n$ ならば $\sum a_n$ は収束する.

(証明) このとき, 級数 $\sum c_n$ の部分和の数列は級数が収束することから上に有界です. 仮定から $\sum a_n$ の部分和の数列も上に有界となり, $\sum a_n$ が単調増加数列になることから「上に有界な単調増加数列は収束する」という公理を用いることでこの級数が収束することが分かります. ■

Theorem. 2.3.6. 正項級数 $\sum a_n$ に対し, $0 \leq k < 1$ となる実数 k が存在し, ある n_0 より大きな全ての $n \in \mathbb{N}$ に対し $a_{n+1}/a_n \leq k$ となるとき, $\sum a_n$ は収束する.

(証明) $a_{n+1} \leq ka_n$ から $a_n < k^{n-1}a_0$ となります. 両辺の和を取ると左辺は公比が 1 より小さいので収束します. よって上の比較定理から $\sum a_n$ は収束します. ■

Theorem. 2.3.7. (ratio test) 正項級数 $\sum a_n$ に対して

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = l \quad (10.2.10)$$

が存在するとき, l が 1 より小さいならば $\sum a_n$ は収束する.

(証明) このとき収束の定義から十分大きな $n \in \mathbb{N}$ を取ると

$$\frac{a_{n+1}}{a_n} < l + \varepsilon < k < 1 \quad (10.2.11)$$

をみたす実数 k が存在します. これから $\frac{a_{n+1}}{a_n} < k < 1$ となり, 前定理より級数 $\sum a_n$ は収束します. ■

10.2.4 開集合, 閉集合, 近傍

後の叙述の便宜の為にいくつかの概念を紹介します. まず集合族 (系) というものを定義します. 標語的には集合を元とする集合です. 具体的にいうと, $\mathcal{O} = \{A, B, \dots\}$ が集合族であるというのは, その元である A, B, \dots がまた集合であるということです. さらに \mathbb{R}^n の開区間とは, $(a_1, b_1) \times \dots \times (a_n, b_n)$ の形の集合とします. その上で開集合系を以下のように定義します.

Definition. 2.4.1. 開区間の適当な個数 (有限個でも無限個でも良い) の和集合で書ける集合を開集合という. 開集合からなる集合系を開集合系といい, \mathcal{O} と書く. 開集合系には全空間 \mathbb{R} と空集合 \emptyset も含む.

このように開集合とは開区間の一般化であると考えて構いません。
他にもいくつか言葉を定義します。

Definition. 2.4.2. \mathbb{R}^n の部分集合 A の 1 点 $a \in A$ の ε -近傍を,

$$U(a; \varepsilon) := \{x \in \mathbb{R}^n; |x - a| < \varepsilon\} \quad (10.2.12)$$

で定義する。これは中心を a とし, 半径 ε の円 (球) のことであり, さらに開集合である。

Definition. 2.4.3. \mathbb{R}^n の部分集合 A の 1 点 $a \in A$ が内点であるというのは, ある $\varepsilon > 0$ が存在して, a の ε -近傍全体が A に含まれることである。すなわち $U(a; \varepsilon) \subset A$ 。

近傍というとき普通 ε の値は小さいとして考えるのが普通です。したがって a の近傍とは文字通り a のすぐまわりの点の集合です。 a が A の内点であるというのは, a のある程度近くの点が全て A に属するということです。さらに定義から $a \in A$ が前提です。例えば $I = [0, 1)$ とします。 $0, 1$ の近くには I の点がありますが, 一方ですぐ近くに I に属さない点があります。つまり $0, 1$ は I の内点ではありません。 $-1, 9$ などそもそも A の点ですらありません。したがって内点ではありません。 I の内点とは開区間 $(0, 1)$ の点です。重要なことですが, 開集合の点は全て内点です。

閉集合書く。

まだ作りかけ。収束の話はなるべくきちんとする。

10.2.5 コンパクト性

作りかけ。

10.2.6 連続と一様連続

作りかけ。

10.2.7 連結と弧状連結

領域

10.2.8 完備な空間

10.3 1 変数の微分と Taylor 展開

まずは定義域が開区間 $I = (a, b) \in \mathbb{R}$ の 1 変数実数値関数 $f(x)$ の線型化を考えてみましょう。 $y = f(x)$ のグラフを考えます。簡単のため, はじめは座標を適当にとって原点の近くで考えることにします。直観的な描像から原点の近くで関数値の差が直線になっていてくれればよさそうです。 x を十分小さくとったとき, A をある定数とすると

$$f(x) - f(0) = A(x - 0) + O((x - 0)^2) \quad (10.3.1)$$

とすれば, 原点の十分近くでは直線になっているでしょう。ここで $O((x - 0)^2)$ は h がある程度大きくなってくると当然曲がりが大きくなるので, その補正です。曲がりなので $(x - 0)^2$ 以上の項で適当に近似できるものと考えます。

ここで原点の十分近くでの線型化ですから、定数 A は原点のみで決まって欲しいと思います。つまり x が十分小さいうちは A だけで $f(x)$ の挙動が完璧に把握できるものとしたいわけです。したがって A は原点のみの関数であることを要求します。また $O((x-0)^2)$ は 2 次以上の挙動をあらわすので、 $x \rightarrow 0$ で $O((x-0)^2)/(x-0) \rightarrow 0$ となることを要求します。

この事情を、原点を一般の点 x_0 にして次のように定義します。

Definition. 3.0.1. \mathbb{R} の開区間 $I = (a, b)$ で定義された実数値関数 f と $x \in I$ に対して

$$f(x) - f(x_0) = A(x_0)(x - x_0) + O((x - x_0)^2) \quad (10.3.2)$$

$$\lim_{x \rightarrow x_0} \frac{O((x - x_0)^2)}{x - x_0} = 0 \quad (10.3.3)$$

と書けるとき、 f は x_0 で線型化可能 (微分可能) であるといい、 A を f の x_0 における導値または微分係数という。このとき、

$$A = f'(x_0) = \frac{df}{dx} = \frac{df}{dx} \Big|_{x=x_0} (Df)(x_0) = (f(x))' = (f(x))' \Big|_{x=x_0} \quad (10.3.4)$$

などを書く。 I の各点で f が微分可能であるとき、 $I \rightarrow \mathbb{R}$ への関数 $f' : x \mapsto f'(x)$ が生じる。これを f の導関数という。

これが高校で学んできた微分可能性の定義と一致する事を確認しましょう。

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \left\{ A + \frac{O((x - x_0)^2)}{x - x_0} \right\} \quad (10.3.5)$$

$$= A \quad (10.3.6)$$

となり、確かに一致しました。

線型化はあくまで近似ですから近似の度合いを高め、さらに正確なものが得られるかどうかということを考えてみましょう。考えるのは $O((x-0)^2)$ を正確に評価することです。ここで高校で学んだ平均値の定理を考えましょう。

Theorem. 3.0.2. (平均値の定理) $a < b$ とする。 $f : I = [a, b] \rightarrow \mathbb{R}$ が I で連続、 $I^i = (a, b)$ で微分可能ならば定数 $\theta \in (0, 1)$ が存在して、

$$f(b) - f(a) = f'(b + \theta(b - a))(b - a) \quad (10.3.7)$$

この定理は証明なしに認めます。これで $b = x$, $a = x_0$ として

$$f(x) - f(x_0) = f'(x + \theta_1(x - x_0))(x - x_0) \quad (10.3.8)$$

が出来ます。

ここで f が I で 2 階微分可能であるとします。すると f' が線型化できて、

$$f'(x + \theta_1(x - x_0)) - f'(x_0) \quad (10.3.9)$$

$$= \frac{1}{2} D^2 f(x_0)(x + \theta_1(x - x_0) - x_0) + O((x - x_0)^2) \quad (10.3.10)$$

となります。 D^2 の前の $1/2$ は、上の式に代入した式を両辺 2 階微分して $x \rightarrow x_0$ としたときにきちんと両辺が一致するようにつけたものです。上の式に代入して

$$f(x) - f(x_0) \quad (10.3.11)$$

$$= \left\{ Df(x_0) + \frac{1}{2} D^2 f(x_0) \{x + \theta_1(x - x_0) - x_0\} + O((x - x_0)^2) \right\} (x - x_0) \quad (10.3.12)$$

$$= Df(x_0)(x - x_0) + \frac{1}{2} D^2 f(x_0)(x - x_0)^2 \quad (10.3.13)$$

$$+ \frac{1}{2} D^2 f(x_0) \theta_1 (x - x_0)^2 O((x - x_0)^2) \quad (10.3.14)$$

$$= Df(x_0)(x - x_0) + \frac{1}{2} D^2 f(x_0)(x - x_0)^2 + O((x - x_0)^3) \quad (10.3.15)$$

となります。ここで $O((x - x_0)^3)$ は 3 次以上の補正項です。これは $D^2 f(x_0)$ の分 f の情報が多くなっているため近似的度合いは高まっていることが予測されます。

これを繰り返していくことで以下の定理が得られます。

Theorem. 3.0.3. (*Taylor (テイラー) の定理*) $N \ni n \geq 1$ とする。区間 $[a, x] = I$ or $[x, a] = I$ で n 回微分可能な実数値関数 f に対し

$$f(x) = f(a) + \frac{Df(a)}{1!}(x - a) + \frac{D^2 f(a)}{2!}(x - a)^2 \quad (10.3.16)$$

$$+ \cdots + \frac{D^{n-1} f(a)}{(n-1)!}(x - a)^{n-1} + R_n(x) \quad (10.3.17)$$

によって R_n を定義するとき、定数 θ が存在して

$$R_n(x) = \frac{D^n f(x + \theta(x - a))}{n!}(x - a)^n \quad (10.3.18)$$

と書ける。これはまた積分を使って

$$R_n(x) = \int_a^x \frac{(x - t)^{n-1}}{(n-1)!} D^n f(t) dt \quad (10.3.19)$$

と書くこともできる。ここで R_n のことを n 次剰余項という。

上の式を Taylor の公式といいます。これを用いて関数がある条件を満たすとき x の多項式で関数の近似を行なうことが出来ます。 x の多項式ならば様子が簡単に掴めるので、これは非常に有用な定理です。さらに関数の性質が非常によいものならば、これはある区間で x のべき級数に展開できます。

これを定理として提出しましょう。

Theorem. 3.0.4. (*Taylor 展開*)

1. f が a を含む区間 I において C^∞ 級で、 I の各点 x で

$$\lim_{n \rightarrow \infty} R_n(x) = 0 \quad (10.3.20)$$

を満たすとき、 f は I 上で

$$f(x) = \sum_{n=0}^{\infty} \frac{D^n f(a)}{n!}(x - a)^n \quad (10.3.21)$$

の形の級数 (a を中心とするべき級数) で表される. これを f の a を中心とする *Taylor 展開* という.

2. 定数 $C \geq 0, M \geq 0$ が存在して, 全ての $n \in \mathbb{N}$ および $x \in I$ に対し

$$|Df^n(x)| \leq CM^n \quad (10.3.22)$$

を満たすとき, f は a を中心として *Taylor 展開* できる.

(証明) まずは 1 から. S_n を

$$S_n(x) := \sum_{k=0}^n \frac{D^k f(a)}{k!} \quad (10.3.23)$$

とします. 仮定から f は C^∞ 級であり, Taylor の定理から

$$f(x) - S_n(x) = R_{n+1}(x) \quad (10.3.24)$$

とかけます. ここで仮定から, R_{n+1} が任意の x に対し 0 に収束しますから,

$$\lim_{n \rightarrow \infty} \{f(x) - S_n(x)\} = 0 \quad (10.3.25)$$

となり, 証明終わりです.

次に 2 を示しましょう. こちらは少々長くなります. 仮定と剰余項の形から $c = x + \theta(x - a)$ として

$$|R_n(x)| = \frac{|D^n f(c)|}{n!} |x - a|^n \leq C \frac{M^n}{n!} |x - a|^n \quad (10.3.26)$$

となります. ここで前節で示した定理により級数 $\sum CM^n |x - a|^n / n!$ は収束します. 高校でも学んだと思いますが, 級数 $\sum a_n$ が収束するとき a_n は 0 に収束しますから, 1 の条件が満たされ f は Taylor 展開可能です. ■

次節ではいくつか良く使う Taylor 展開がありますからそれを示すことにします.

10.4 初等超越関数の Taylor 展開

何はともあれ初等超越関数とは何か, ですが $\cos x, \sin x, e^x, \log x$ のことです. まずはこれらの Taylor 展開を求めます. はじめに 0 のまわりの Taylor 展開の式を書いておきます. ほとんど 0 の場合しか使いませんので, これだけで十分です.

$$f(x) = \sum_{n=0}^{\infty} \frac{D^n f(0)}{n!} x^n \quad (10.4.1)$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \quad (10.4.2)$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad (10.4.3)$$

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n \quad (10.4.4)$$

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n \quad (10.4.5)$$

無限級数なので収束の話をついでにも考えないといけないのですが、まずは n 階微分を計算して上の計算が合っているかを確かめましょう。 $\cos x$ は $D \cos x = -\sin x$, $D^2 \cos x = -\cos x$, $D^3 \cos x = \sin x$, $D^4 \cos x = \cos x$ ですから、

$$D^n \cos x|_{x=0} = \begin{cases} (-1)^{n/2}, & (n \text{ は偶数}) \\ 0, & (n \text{ は奇数}) \end{cases} \quad (10.4.6)$$

となり、上の展開式は正しいです。 \sin はほとんど同じなので省略します。 e^x は $D^n e^x|_{x=0} = 1$ なのでやはり正しいです。 次に \log ですが、 $D \log(1+x) = \frac{1}{1+x} = (1+x)^{-1}$ であり、 $D^n \log(1+x)|_{x=0} = (-1)^{n-1}(n-1)!$ となり、やはり正しいです。

問題は収束です。 Taylor 展開の式を $f(x) = \sum a_n$, $a_n = D^n f(0)x^n/n!$ と書きましょう。

$\cos x, \sin x$ についてはまず x^{2n}, x^{2n+1} を一律に y^n として考えます。 前節に示した収束に関する定理を使います。 具体的には ratio test です。 まずは \cos 。

$$\frac{a_{n+1}}{a_n} = \frac{(-1)^{n+1}y^{n+1}}{(2n+2)!} \frac{(2n)!}{(-1)^n y^n} = \frac{-y}{2n+2} \rightarrow 0, n \rightarrow \infty \quad (10.4.7)$$

上の収束は y によらないので、任意の実数 y に対して上の Taylor 展開が成立します。 ここで $y = x^{2n}$ としたものが \cos です。 この展開はある点のそばでの線型化 (近似) から始まったものですが、その近似を高めていったところ任意の実数に対して成立してしまいました。 恐るべき結果です。 \sin もほとんど同じです。 ついでに e^x もほとんど同じです。

\log は何でもそう上手くいくものではないということを教えてくれるのでそれを見てみましょう。

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{(-1)^n x^{n+1}}{n+1} \right| \left| \frac{n}{(-1)^{n-1} x^n} \right| = \frac{nx}{(n+1)} \rightarrow x, n \rightarrow \infty \quad (10.4.8)$$

これは $n \rightarrow \infty$ が x に依存します。 ratio test によればこの極限値が 1 より小さいところならば級数が収束するので、対数の場合は Taylor 展開は 1 より小さいところでしか成立しません。 ただ $\log(1+x)$ の 0 まわりの展開が $|x| < 1$ でしか成立しないと言っているだけです。 1 のまわり、2 のまわりで展開すればまた少し事情が変わるはずですが。

10.5 Euler の公式

Euler (オイラー) の公式とは以下の想像を絶する式です。

$$e^{ix} = \cos x + i \sin x \quad (10.5.1)$$

ここで e の肩に乗っていて右辺の \sin の前に燦然と輝く i は虚数単位の i です。 この公式を「導出」しましょう。 前節の e^x の Taylor 展開の式で $x \rightarrow ix$ とす

ると

$$e^{ix} = \sum_{n=0}^{\infty} \frac{1}{n!} (ix)^n \quad (10.5.2)$$

$$= \sum_{n=0}^{\infty} \left\{ \frac{1}{(2n)!} (ix)^{2n} + \frac{1}{(2n+1)!} (ix)^{2n+1} \right\} \quad (10.5.3)$$

$$= \sum_{n=0}^{\infty} \left\{ \frac{1}{(2n)!} i^{2n} x^{2n} + i \frac{1}{(2n+1)!} i^{2n} x^{2n+1} \right\} \quad (10.5.4)$$

$$= \sum_{n=0}^{\infty} \frac{1}{(2n)!} (-1)^n x^{2n} + i \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} (-1)^n x^{2n+1} \quad (10.5.5)$$

ここで \cos, \sin の Taylor 展開と比べて上の式を得ます。

虚数に対して級数展開が成立するか、虚数乗とは何者か、といった大問題があるのですが、(複素) 関数論をきちんとやらないと決着が付きません。解析接続を持ち出すと、一般の複素数 z に対し $e^z = \cos z + i \sin z$ が成立しこれが複素数へのただ一つの拡張の仕方であることが分かります。余裕があれば後で触れるでしょう。今は有用な公式としてどんどん使っていくことにします。振動・波動との関連はとても分かりやすいと思います。我々は三角関数の新たな表現を手に入れたことになります。特に微分方程式を解くとき、これを導入すると計算が著しく簡単になります。

10.6 多変数の微分と Taylor 展開

変数が多いとややこしくなるので主に 2 変数 x, y (定義域は適当に $(a_1, b_1) \times (a_2, b_2)$ としておきます) の実数値関数に対して話を進めます。つまり $z = f(x, y)$ を考えていきます。1 変数の実数値関数のときの扱いと比べると、これが「平面」内の「曲線」を取り扱うことに相当したのと同様に今度は「空間」内の「曲面」を扱うことになり、ある程度は直観を働かせることも出来ます。

まず 2 変数の線型化 (線型近似) が何になるかを考えてみます。1 変数のときは高校で学んだとおり直線でした。直線とは 1 次元のものです。これから 2 変数での線型近似は 2 次元のもの、平面で行なうのが妥当であると考えられます。高校では空間内の平面をあまりやっていないものと思うので、少しこれを補足します。高校の教科書にも軽く載っていたと思いますが、ある平面は一直線上にない 3 点 A, B, C を適当に取ることで指定することが出来ます。3 点が一直線上にないので、例えば $\overrightarrow{AB}, \overrightarrow{AC}$ を考えれば、これらは線型独立であって、平面内の任意の点がこの 2 本のベクトルの線型結合で書けることになります。

つまり、1 変数の実数値関数のときは平面内の曲線の挙動を捉えるのに直線 (接線) による線型化を用いましたが、2 変数の実数値関数では空間内の曲面の挙動を捉えるのに平面 (接平面) による線型化を用い、さらに曲面上の各点での平面の挙動は 2 本の直線で捉えるということです。

それでは実際に線型化を開始します。適当に $x = {}^t(x, y)$ を取り、十分小さい h, k を取って $\mathbf{h} = {}^t(h, k)$ とし、 $\mathbf{r} + \mathbf{h}$ での様子を考えてみます。

$$f(x+h, y+k) - f(x, y) = ah + bk + O(h^2, hk, k^2) \quad (10.6.1)$$

ここで $O(h^2, hk, k^2)$ は 1 変数のとき同様曲がった部分の補正項ですが、2 変数なので 2 次の項のバリエーションが増えています。このままでは色々面倒なので、これ

を少し整理しておきます. 1 変数のときを参考にすると, O の項は微小増分 Δx が 0 に近づく時, $O(\Delta x)/\Delta x \rightarrow 0$ となれば良い, ということでした. ベクトルの割り算というものが正体不明である (定義していない) 以上, 直接 $O(h^2, hk, k^2)/h \rightarrow 0$ とすることは出来ません. そこで不等式 $\min\{h^2, k^2\} \leq |hk| \leq h^2 + k^2$ を参考に, 2 次の挙動は $(h^2, hk, k^2)/|h|^2 \rightarrow 0$ で上手く押さえられるものとします.

さて, この線型化を 1 変数のときと同じように一次関数の形で書いてみましょう. $A := (a, b)$, $h := {}^t(h, k)$ とすれば,

$$f(x+h, y+k) - f(x, y) = A h + O(h^2, hk, k^2) \quad (10.6.2)$$

1 変数の場合と同じ形式で書いてみると, 微分係数が行ベクトル (1×2 行列) になる事がわかります. 直接に h を出してくると, 左辺が実数, 右辺がベクトルというすさまじい式になってしまいますから, 行列の形にすることが必要である, ともいえます. ここではあまり詳しくやりませんが, 多変数のベクトル値関数ではもっと露骨な形で行列になります. ここで一旦 2 変数 (多変数) の微分の定義をしておきましょう.

Definition. 6.0.5. \mathbb{R}^2 のある開区間 $I = (a_1, b_1) \times (a_2, b_2)$ で定義された実数値関数 f と $x \in I$ に対して

$$f(x+h) - f(x) = A(x)h + O(|h|) \quad (10.6.3)$$

$$\lim_{h \rightarrow 0} \frac{O(h^2, hk, k^2)}{|h|} = 0 \quad (10.6.4)$$

と書けるとき, f は x で微分可能であるといい, A を f の x における導値または微分係数という. このとき,

$$A = f'(x) = (Df)(x) = (f(x))' \quad (10.6.5)$$

など書く. I の各点で f が微分可能であるとき, $I \rightarrow \mathbb{R}$ への関数 $f' : x \mapsto f'(x)$ が生じる. これを f の導関数という.

ここから重要な役者である偏微分に登場してもらいます. 上では $h = {}^t(h, k)$ として h, k 両方が動いていたわけですが, 今度は k を k_0 に固定しておいて h だけに動いてもらいましょう. $k = k_0$ を固定すると $z = f(x, k_0) := g(x)$ は曲面を平面 $y = k_0$ で切った断面上を動くことになりますから, 幾何学的には g の接線を求める作業へ移ることに対応します. 本節はじめに述べたことと対応させると, 接平面を決める為の 2 本のベクトルの内の一本を決めることにあたります. 演算的な側面から言うと, 偏微分とはある着目している変数以外全て止めて (定数とみなし), 着目した変数に対する関数の挙動を調べるということです.

1 階の偏微分の定義にいきたいところなのですが, 後での都合もあるのでもう少し寄り道をします. 読者の皆さんがそうだと思いますが,

$$x = x e_x + y e_y, e_x = e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_y = e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (10.6.6)$$

として考えていると思います. e_i は前章で言うところの標準基底というものでした. 上では当然このつもりで説明をしてきました. これから述べる偏微分を標準基底を用いて表せば, $f(x + t e_1) = g(t)$ の t についての (1 変数の) 微分ということになります. この e_1 を一般に n としたものを n 方向への方向微分と言います. 何故こんなことを言い出すのか, 数学上の術学的趣味ではないか, という人もいるかと思いますが, 物理でも電磁気学などで「(曲面上での) 法線方向への微

分」というものを考えることがあります。それ以上に、我々の目的の 1 つに安定点の定式化がありますが、これはどの方向から近づいても微分が 0 になる点であると定式化出来そうである、という予備考察があります。方向微分とは、まさしくこのような方向性に沿うものとしての期待が持てます。

さて、ここで方向微分と特殊な方向微分としての偏微分を定義しましょう。

Definition. 6.0.6. a は \mathbb{R}^2 の開区間 $I = (a_1, b_1) \times (a_2, b_2)$ の 1 点で、 f は I で定義された実数値関数であるとする。今 \mathbb{R}^2 の任意の元 e に対して実変数 $t \in \mathbb{R}$ の関数 $g(t) := f(a + te)$ が $(-\varepsilon, \varepsilon), \varepsilon > 0$ で定義されているとする。この関数 g が $t = 0$ で微分可能ならば、関数 f は a において e 方向に微分可能であるといい、 $g'(0)$ を f の a における e 方向の導値 (または微分係数) といい、

$$g'(0) = (D_e f)(a) = \frac{\partial f}{\partial e}(a) \quad (10.6.7)$$

などと記す。 $(D_e f)(a)$ が I の各点 a で定義されるとき、 I 上の関数 $D_e f : a \mapsto (D_e f)(a)$ を、 f の e 方向の導関数という。

Definition. 6.0.7. $x = xe_x + ye_y = x_1e_1 + x_2e_2$ と表すこととし、 a は \mathbb{R}^2 の開区間 $I = (a_1, b_1) \times (a_2, b_2)$ の 1 点で、 f は I で定義された実数値関数であるとする。 f が $a \in \mathbb{R}^2$ で e_i 方向に微分可能であるとき、 f は第 i 座標 x_i について偏微分可能という。 $(D_{e_i} f)(a)$ を f の a における x_i についての偏導値 (または偏微分係数) といい、

$$(D_{e_i} f)(a) = \frac{\partial f}{\partial x_i}(a) = f_{x_i}(a) = (D_i f)(a) = \partial_i f(a) \quad (10.6.8)$$

などと記す。これを極限を用いて表せば、

$$D_1 f(x, y) = \frac{\partial f}{\partial x} = \lim_{h \rightarrow 0} \frac{f(r + he_1) - f(r)}{h} \quad (10.6.9)$$

などとなる。 f が I の各点で x_i について微分可能であるとき、 f の x_i についての偏導関数

$$\frac{\partial f}{\partial x_i} : a \mapsto \frac{\partial f}{\partial x_i}(a) \quad (10.6.10)$$

が生じる。 f から $\partial f / \partial x_i$ を得ることを f を x_i について偏微分するという。

さらに高階の偏導関数を先に定義してしまいましょう。

Definition. 6.0.8. x_i に関する偏導値 $(D_i f)(a)$ が I の全ての点 a で定義されるとき、 I 上で $D_i f : a \mapsto (D_i f)(a)$ が定義される。 D_i を f の第 i 成分 x_i に関する偏導関数という。もし $D_i f$ が第 j 座標に関して偏微分可能ならば 2 階偏導関数 $D_j(D_i f)$ が定義される。これを

$$D_j(D_i f)(x) = (D_{i,j} f)(x) = f_{x_i, x_j}(x) = \partial_j \partial_i f(x) = \begin{cases} \frac{\partial^2 f}{\partial x_j \partial x_i}(x) & , (i \neq j) \\ \frac{\partial^2 f}{\partial x_i^2}(x) & , (i = j) \end{cases} \quad (10.6.11)$$

などと記す。同様に高階の偏導関数も定義される。

Definition. 6.0.9. $k \in \mathbb{N}$ とする. $I = (a_1, b_1) \times (a_2, b_2)$ で定義された関数 f は, f の k 階までの全ての偏導関数 ($\sum_{l=1}^k n^l$ 個ある) が存在して I 上連続であるとき, I 上で C^k 級である, または k 回連続微分可能であるという. 全ての $k \in \mathbb{N}$ に対し C^k 級であれば, C^∞ 級または無限回微分可能という.

ここで本節はじめに述べた (多変数での) 微分係数と偏微分係数の関係を述べておきます.

$$f(x+h, y+k) - f(x, y) = ah + bk + O(h^2, hk, k^2) \quad (10.6.12)$$

において $k = 0$ とすると,

$$f(x+h, y) - f(x, y) = ah + O(h^2) \quad (10.6.13)$$

1 変数の微分の定義と上に示した偏微分の定義から, 辺々 h で割って $O(h^2)/h \rightarrow 0$ ($h \rightarrow 0$) から $a = D_1 f(x) = \partial f / \partial x$ となります. 同様にして $b = D_2(f(x)) = \partial f / \partial y$ です. これから微分係数 A は

$$A = (a \quad b) = \left(\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right) \quad (10.6.14)$$

と書けることになります.

前にも述べましたが, 偏微分の順序は一般に交換できません. つまり一般に $D_i D_j \neq D_j D_i$ です. しかし比較的緩い条件でこの交換が可能になります. 後でも使うので先んじてその証明を載せておきましょう.

Theorem. 6.0.10. $I = (a_1, b_1) \times (a_2, b_2)$ で定義された実数値関数 f に対し, 点 $c \in I$ で f_{xy}, f_{yx} が共に存在し c で連続ならば偏微分の順序が交換できる. すなわち

$$f_{xy}(c) = f_{yx}(c). \quad (10.6.15)$$

(証明) 定義によれば, f_{xy} は偏微分を 2 階続けて行なったものであり, 1 つの変数についての極限操作を 2 回繰り返すことで得られます. ところがこれを 2 次元空間 \mathbb{R}^2 における一つの極限で書くことが出来ます. 今 $l = {}^t(h, k) \in \mathbb{R}^2$ を小さく取り, $c+l \in I$ となるようにします. また $\varphi(x) := f(x, b+k) - f(x, b)$ とすれば, $\varphi'(x) = f_x(x, b+k) - f_x(x, b)$ となります. 今

$$\Delta(h, k) := f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b) \quad (10.6.16)$$

とします. 平均値の定理から $0 < \theta, \theta' < 1$ となる θ, θ' が存在して,

$$\Delta(h, k) = \varphi(a+h) - \varphi(a) = h\varphi'(a+\theta h) \quad (10.6.17)$$

$$= h\{f_x(a+\theta h, b+k) - f_x(a+\theta h, b)\} \quad (10.6.18)$$

$$= hk f_{xy}(a+\theta h, b+\theta' k) \quad (10.6.19)$$

となります. 今 $l \rightarrow 0$ とすれば,

$$\begin{pmatrix} a+\theta h \\ b+\theta' k \end{pmatrix} \rightarrow \begin{pmatrix} a \\ b \end{pmatrix} = c \quad (10.6.20)$$

であり, 仮定から f_{xy} が連続ですから, 上式から

$$f_{xy}(c) = \lim_{(h,k) \rightarrow 0, hk \neq 0} \frac{\Delta(h, k)}{hk} \quad (10.6.21)$$

を得る. 今定理の仮定が x, y について対称なので, 同じ議論を繰り返すことで

$$f_{yx}(c) = \lim_{(h,k) \rightarrow 0, hk \neq 0} \frac{\Delta(h,k)}{hk} \quad (10.6.22)$$

と出来ます. これで $f_{xy}(c) = f_{yx}(c)$ が証明されました. ■

10.7 微分公式

この節では合成関数の微分と関数の積の微分公式を扱います. まず合成関数の微分公式を示します.

Theorem. 7.0.11. f を $\Omega_1 \subset \mathbb{R}^n$ から \mathbb{R}^m への関数, g を $\Omega_2 \subset f(\Omega_1) \subset \mathbb{R}^m$ から \mathbb{R}^l への関数とし, f が点 $a \in \Omega_1$ で微分可能, g が点 $b = f(a) \in \Omega_2$ で微分可能とする. このとき合成関数 $\varphi := g \circ f$ は点 a で微分可能であり

$$D\varphi = D(g \circ f) = Dg \cdot Df, \quad (10.7.1)$$

$$\frac{\partial \varphi_r}{\partial x_j} = \sum_{l=1}^m \frac{\partial g_r}{\partial y_l} \frac{\partial f_l}{\partial x_j} \quad (1 \leq r \leq p, 1 \leq j \leq n). \quad (10.7.2)$$

(証明) 次元 n, m がほとんど表に出てきませんので, 一般の形で示していきます. f, g の微分可能性から,

$$f(a+h) = f(a) + Ah + O(|h|^2), \quad A := Df \quad (10.7.3)$$

$$g(b+k) = g(b) + Bk + O(|k|^2), \quad B := Dg \quad (10.7.4)$$

となります. $b = f(a)$ を g の式に代入して,

$$g(f(a+h)) = g(f(a) + Ah + O(|h|^2)) \quad (10.7.5)$$

$$= g(f(a)) + B(Ah + O(|h|^2)) + O(|Ah + O(|h|^2)|^2) \quad (10.7.6)$$

$$= (g \circ f)(a) + BAh + O(|h|^2) \quad (10.7.7)$$

と書けます. 合成 $h = g \circ f$ が点 a で微分可能ならば,

$$h(a+h) = h(a) + Ch + O(|h|^2) \quad (10.7.8)$$

なので, これを上との式と比べると合成の微分可能性と $C = BA = Dg \cdot Df$ が分かります, というのが直観的な議論です.

以下きちんと議論しなおします. f の変数を x で, g の変数を y で書くことにします. 各関数の微分可能性から, ε, δ を $h \rightarrow 0 \Rightarrow \varepsilon(h) \rightarrow 0, k \rightarrow 0 \Rightarrow \delta(k) \rightarrow 0$ となる関数として

$$f(x+h) - f(x) = Ah + |h|\varepsilon(h) \quad (10.7.9)$$

$$g(y+k) - g(y) = Bk + |k|\delta(k) \quad (10.7.10)$$

と書けます. ここで

$$k = k(h) = f(x+h) - f(x) \quad (10.7.11)$$

とすれば, f の連続性から $h \rightarrow 0 \Rightarrow k(h) \rightarrow 0$ となります. 合成関数 φ について考えると,

$$\varphi(x+h) - \varphi(x) = g(f(x+h)) - g(f(x)) \quad (10.7.12)$$

$$= g(y+k) - g(y) = Bk(h) + |k(h)|\delta(k(h)) \quad (10.7.13)$$

$$= B(Ah + |h|\varepsilon(h)) + |k(h)|\delta(k(h)) \quad (10.7.14)$$

$$= BAh + |h|\rho(h), \quad \rho(h) := B\varepsilon(h) + \frac{|k(h)|}{|h|}\delta(k(h)) \quad (10.7.15)$$

となります. ここで (m, n) 行列 $A = (a_{ij})$ に対して

$$|A| := \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}} \quad (10.7.16)$$

と書くことにします. A の第 i 行ベクトルを a_i とすれば,

$$|Ax|^2 = \sum_{i=1}^m |a_i x|^2 = \sum_{i=1}^m ({}^t a_i, x)^2 \quad (10.7.17)$$

$$\leq \sum_{i=1}^m |{}^t a_i|^2 |x|^2 \quad (\text{Cauchy-Schwartz の不等式}) \quad (10.7.18)$$

$$= |A|^2 |x|^2 \quad (10.7.19)$$

が一般に成立します. これと三角不等式を用いて

$$|k(h)| = |f(x+h) - f(x)| = |Ah + |h|\varepsilon(h)| \leq (|A| + |\varepsilon(h)|)|h| \quad (10.7.20)$$

となります. これから

$$\lim_{h \rightarrow 0} \rho(h) = 0 \quad (10.7.21)$$

となって証明終了です. ■

Theorem. 7.0.12. f, g が開集合 $\Omega \subset \mathbb{R}^n$ を定義域とする実数値関数とする. f, g が 1 点 a で微分可能なとき, 積 fg は a で微分可能であり, 導関数 $D(fg)$ は次のように書ける.

$$D(fg)(a) = f \cdot Dg + g \cdot Df \quad (10.7.22)$$

(証明) 以下の証明を見れば分かるように, 定義域の次元 n はまったく表に出てきません. そこで一般の n で証明を構成します. f, g が a で微分可能なので, ε, δ を先程と同様, $h \rightarrow 0 \Rightarrow \varepsilon(h), \delta(h) \rightarrow 0$ となるように取れば

$$f(a+h) = f(a) + A(a)h + |h|\varepsilon(h), \quad A := Df(a), \quad (10.7.23)$$

$$g(a+h) = g(a) + B(a)h + |h|\delta(h), \quad B := Dg(a) \quad (10.7.24)$$

と書けます.

積 fg について計算してみると

$$(fg)(\mathbf{a} + \mathbf{h}) = f(\mathbf{a} + \mathbf{h}) \cdot g(\mathbf{a} + \mathbf{h}) \quad (10.7.25)$$

$$= \{f(\mathbf{a}) + A(\mathbf{a})\mathbf{h} + |\mathbf{h}|\varepsilon(\mathbf{h})\} \cdot \{g(\mathbf{a}) + B(\mathbf{a})\mathbf{h} + |\mathbf{h}|\delta(\mathbf{h})\} \quad (10.7.26)$$

$$\begin{aligned} &= f(\mathbf{a}) \cdot g(\mathbf{a}) + f(\mathbf{a}) \cdot B(\mathbf{a})\mathbf{h} + g(\mathbf{a}) \cdot A(\mathbf{a})\mathbf{h} \\ &\quad + (g(\mathbf{a}) + B(\mathbf{a})\mathbf{h} + |\mathbf{h}|\delta(\mathbf{h})) |\mathbf{h}|\varepsilon(\mathbf{h}) \\ &\quad + (f(\mathbf{a}) + A(\mathbf{a})\mathbf{h} + |\mathbf{h}|\varepsilon(\mathbf{h})) |\mathbf{h}|\delta(\mathbf{h}) \end{aligned} \quad (10.7.27)$$

$$= (fg)(\mathbf{a}) + (fB + gA)(\mathbf{a}) \cdot \mathbf{h} + O(|\mathbf{h}|^2). \quad (10.7.28)$$

積 fg が微分可能ならば, 定義から

$$(fg)(\mathbf{a} + \mathbf{h}) = (fg)(\mathbf{a}) + C\mathbf{h} + O(|\mathbf{h}|^2) \quad (10.7.29)$$

と書けます. すぐ上の式と比べれば, 積の微分可能性と $C = fB + gA = f \cdot Dg + g \cdot Df$ となることが分かります. ■

10.8 多変数の Taylor の定理と微分

多変数の実数値関数に関しても Taylor の定理が成立します. 1 変数のときと同じようにして線型化を中心に組み立てれば良いのですが, 1 変数と同じなのでその詳細は控えます. まずは予備定理から証明します.

Theorem. 8.0.13. f を $I = \prod_{i=1}^n (a_i, b_i) \subset \mathbb{R}^n$ 上で C^k 級実数値関数とする. I の 2 点 $\mathbf{x}, \mathbf{x} + \mathbf{z}$ を結ぶ線分 $L: g(t) = \mathbf{x} + t\mathbf{z}, t \in [0, 1]$ が I に含まれているとき, 合成関数 $\varphi = f \circ g$ は $[0, 1]$ 上で C^k 級で, $1 \leq m \leq k$ のとき φ の m 階導関数は

$$\varphi^{(m)}(t) = \sum_{1 \leq i_1, \dots, i_m \leq n} \frac{\partial^m f}{\partial x_{i_1} \cdots \partial x_{i_m}}(\mathbf{x} + t\mathbf{z}) z_{i_1} \cdots z_{i_m} \quad (10.8.1)$$

で与えられる.

(証明) g は t の一次関数です. これは 2 階以降の導関数が 0 になってしまいますが, これを微分可能な関数と考えれば g は C^∞ の関数です. そこで $m = 1$ のときは単純な合成関数の微分から

$$\varphi'(t) = f'(\mathbf{x} + t\mathbf{z})g'(t) = f'(\mathbf{x} + t\mathbf{z})\mathbf{z} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x} + t\mathbf{z}) z_i \quad (10.8.2)$$

が成立します. そこで m についての帰納法で (10.8.1) が成立することを証明をします. 今 $m < k$ に対して (10.8.1) が成立すると仮定すれば, (10.8.1) を t について微分すれば, $m+1$ に対する (10.8.1) が得られます. 実際 f は C^k 級で $m+1 \leq k$ から, f の $m+1$ 階偏導値は, 前に示した定理から偏微分の順序によりません. したがって $m = 1$ の場合を用いて

$$\varphi^{(m)}(t) = \sum_{1 \leq i_1, \dots, i_{m+1} \leq n} \frac{\partial^{m+1} f}{\partial x_{i_1} \cdots \partial x_{i_{m+1}}}(\mathbf{x} + t\mathbf{z}) z_{i_1} \cdots z_{i_{m+1}} \quad (10.8.3)$$

が得られます. ■

ここで一旦便利な記法として m 次微分 $(d^m f)_x$ を導入しておきます. (10.8.1)
の右辺からそのまま

$$(d^m f)_x(z) := \varphi^{(m)}(t) = \sum_{1 \leq i_1, \dots, i_m \leq n} \frac{\partial^m f}{\partial x_{i_1} \cdots \partial x_{i_m}}(x + tz) z_{i_1} \cdots z_{i_m} \quad (10.8.4)$$

とします. これを f の x における m 次微分といいます. 特に $m = 1$ のときの $d^1 f$ を df と書きます.

それでは Taylor の定理です.

Theorem. 8.0.14. (多変数の Taylor の定理) f を $\prod_{i=1}^n (a_i, b_i)$ 上で C^k 級 ($k \geq 1$) の実数値関数とする. 2 点 $x, x + h$ を結ぶ線分 L が I に含まれるとき, $0 < \theta < 1$ なる実数 θ が存在して

$$f(x + h) = f(x) + \sum_{m=1}^{k-1} \frac{1}{m!} (d^m f)_x(h) + \frac{1}{k!} (d^k f)_{x+\theta h}(h) \quad (10.8.5)$$

が成立する.

(証明) $\varphi(t) := f(x + th)$ は $[0, 1]$ で C^k 級ですから, 1 変数関数についての Taylor の定理から, ある実数 $\theta \in (0, 1)$ が存在して

$$\varphi(1) = \sum_{m=0}^{k-1} \frac{1}{m!} \varphi^{(m)}(0) + \frac{1}{k!} \varphi^{(k)}(\theta) \quad (10.8.6)$$

となります. 前定理から $\varphi^{(m)}(0) = (d^m f)_x(z)$, $\varphi^{(k)}(\theta) = (d^k f)_{x+\theta h}(h)$ ですから定理が成立します. ■

ここで特に $m = 1$ とした 1 次微分 (普通単に微分という) の効能を見てみましょう. 我々の目的の 1 つは「安定点はあらゆる方向からの 1 階の微分が消える (微分係数が 0 になる) 点である」という予想をきちんと確認することです. 先程見たように, 偏微分は標準基底に沿った方向微分であり, 「あらゆる方向」というものをどう定式化するかには多少の困難があります. ここで, 標準基底を全て同じ角度だけ回転させたベクトルの組 $\{u_i\}$ はまた基底になります. この基底に沿った方向微分を考えれば, 新たな方向微分が得られます. これを繰り返せば, 原理的には「あらゆる方向からの微分」を考えることが出来ますが, 非常に面倒どころか変化の方向は無限にあるので, 人間の手では追いつきません. そこで線型代数で示した (空間内の) 任意の回転は直交行列で表現出来るという定理を使います. $x = {}^t(x_1, x_2, x_3)$ を回転させたベクトルを $y = {}^t(y_1, y_2, y_3)$ とすると $y = Ux$ ($x = {}^t U y$) となります. 成分で書くと, $y_k = U_{ki} x_i$ ($x_i = {}^t U_{ik} y_k$) です. さらに微分作用素の方も変換しないといけません. 高校でも学んできた 1 変数のときと同じで良いのですが, 多変数では $y_k = U_{k1} x_1 + U_{k2} x_2 + U_{k3} x_3$ となり, これらを各変数で微分しないといけないことに注意して下さい. これをふまえて計算すると以下のようになります.

$$\frac{\partial}{\partial x_i} = \sum_{k=1}^3 \frac{\partial y_k}{\partial x_i} \frac{\partial}{\partial y_k} = \sum_{k=1}^3 U_{ki} \frac{\partial}{\partial y_k}. \quad (10.8.7)$$

準備が整ったので微分を計算します. また, $f(x) = f(U^{-1}y) = g(y)$, z も変換し

て $w = Uz$ とすると

$$(df)_x(z) = \sum_{i=1}^3 \frac{\partial f}{\partial x_i}(x) z_i = \sum_{i=1}^3 \sum_{k=1}^3 U_{ki} \frac{\partial}{\partial y_k} g(y) z_i \quad (10.8.8)$$

$$= \sum_{i=1}^3 \sum_{k=1}^3 \frac{\partial}{\partial y_k} g(y) U_{ki} z_i = \sum_{i=1}^3 \sum_{k=1}^3 \frac{\partial}{\partial y_k} g(y) {}^t U_{ik} z_i \quad (10.8.9)$$

$$= \sum_{k=1}^3 \frac{\partial g}{\partial y_k}(y) w_k \quad (10.8.10)$$

となり、どのように回転させた基底に対する方向微分に対しても、微分は形を変えません。これは微分が基底の取り方に関係なく存在する量であることを示しています。つまり、方向に対する依存性がないということですから、これは「あらゆる方向への微分」に対応するものであるといえます。

10.8.1 Taylor 展開の応用-波動方程式の導出

流れとしてはこのまま本筋の極値問題へ行くべきですが、その前にもう 1 つの本筋である振動・波動論に対して非常に重要な波動方程式の導出を行ないます。

物理的な状況としては、両端が固定された (固定端の) 長さ l 、質量 m の弦を考えます。そして一方の端 (左端とする) を座標の原点とし、もう一方の端へと x 軸を取ります。時刻 t 、位置 x での弦の振幅を $u(x, t)$ で表し、また考察の便宜の為、弦を n 分割します (n 等分としても構いません)。.. n はあとで行なう近似が十分正確になる程度に大きく取り、最後には $n \rightarrow \infty$ とします。さらに作業上の都合から、等分した各点に各弦の部分の全質量が集まっているものとし、その各部分が適当に質量の無い (無視できる) 糸で結ばれているものとします..

このとき i 番目の質点の運動方程式を考えます。糸の張力は質点の左右で等しいとし、 T とします。また水平線と質点を結ぶ糸のなす角を左から θ_1, θ_2 とします。 μ を弦 (の質量) の線密度として、

$$(\mu \Delta x) \frac{\partial^2 u}{\partial t^2}(x_i, t) = T \sin \theta_2 - T \sin \theta_1 \quad (10.8.11)$$

ここで θ が十分小さく、 $\sin \theta \simeq \tan \theta$ が成立するように n を取ります。ここで \tan の幾何学的意味から、 $\tan \theta = \frac{\partial u}{\partial x}$ となります。上の式に代入して Taylor 展開を用いると

$$\mu \Delta x \frac{\partial^2 u}{\partial t^2}(x_i, t) = T \frac{\partial u}{\partial x}(x_i, t) - T \frac{\partial u}{\partial x}(x_i - \Delta x, t) \quad (10.8.12)$$

$$= T \left[\frac{\partial u}{\partial x}(x_i, t) - \left\{ \frac{\partial u}{\partial x}(x_i, t) - \Delta x \frac{\partial^2 u}{\partial x^2}(x_i, t) + O((\Delta x)^2) \right\} \right] \quad (10.8.13)$$

$$= T \frac{l}{n} \frac{\partial^2 u}{\partial x^2}(x_i, t) + O((\Delta x)^2) \quad (10.8.14)$$

$$\implies \mu \frac{\partial^2 u}{\partial t^2}(x_i, t) = T \frac{\partial^2 u}{\partial x^2}(x_i, t) + (\Delta x) \quad (10.8.15)$$

ここで $n \rightarrow \infty$ とし、 $x_i = x$ と書き直すと

$$\frac{1}{v^2} \frac{\partial^2 u}{\partial t^2}(x, t) = \frac{\partial^2 u}{\partial x^2}(x, t), \quad v := \sqrt{\frac{T}{\mu}} \quad (10.8.16)$$

上式を (1 次元の) 波動方程式といいます。あとで使うこともあるでしょうから, 3 次元版も付記しておきましょう. $\nabla := \partial_x e_x + \partial_y e_y + \partial_z e_z$ とし, $\nabla \cdot \nabla = \nabla^2 = \partial_x^2 + \partial_y^2 + \partial_z^2$, $\square := (1/v^2)\partial_t^2 - \nabla^2$ とすると,

$$0 = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = \square u \quad (10.8.17)$$

波動方程式が本当に波動を表しているかを簡単な例で確認してみましょう. 高校で単振動は $u(x, t) = \sin(\omega t - kx)$ で表されると学んだはず. ここで $\omega := 2\pi f$, $k := 2\pi f/\lambda$ です. ω を角振動数, λ を波数といいます. この 2 階の時間・空間微分を考えると $v := \sqrt{\omega/k}$ とすれば確かに波動方程式を満たします. また波動方程式は重ね合わせの原理も満たします. 2 つの波 u_1, u_2 があるとしましょう. この 2 つを重ね合わせたものもまた波である, というのは $u := u_1 + u_2$ が波動方程式の解となるということです. u_1, u_2 それぞれが波動方程式の解ですから, $\square u_i = 0 (i = 1, 2)$ となります. 微分作用素 \square は線型作用素ですから, $\square u = \square u_1 + \square u_2 = 0 + 0 = 0$ となり, 重ね合わせた u も確かに波動方程式を満たします.

少し脱線します. ここで Euler の公式 $e^{ix} = \cos x + i \sin x$ から $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$, $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$ です. 微分の手間を考えると, 数学的には \sin, \cos の代わりに e^{ix} を考えた方が楽ですからそうしてみます. この複素表示を用いると正弦波 (単振動) は $\psi_{k,\omega}(x, t) = e^{i(kx - \omega t)}$ と書けます. そして重ね合わせの原理から, 波の重ね合わせはまた波ですから, 波動方程式にしたがう波 u は正弦波の重ね合わせで記述できる (または非常に良く振る舞いが近似される) と仮定してみましょう. 重ね合わせの際には様々な波数・角振動数の正弦波を使うのがもっともらしいでしょう. つまり $\psi_{k,\omega}$ としては, 全実数を考えます. 全実数に対して重ね合わせるのだから, 積分を使えばよいでしょう. すなわち,

$$u(x, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(k, \omega) e^{i(kx - \omega t)} dk d\omega \quad (10.8.18)$$

と書けます. ここで $a(k, \omega)$ は重ね合わせる際の荷重です. ちなみにこれを u の Fourier 逆変換といいます. ついでにいうと, 上の表示では単発の (重ね合わせられていない) 正弦波そのものを表現できません. この不都合を解消する為に 1 つの「関数」を導入します. 線型代数で学んだ Kronecker の δ を真似た (と思われる) Dirac の δ 関数 $\delta(x)$ です. これは上の式の正当化をこめて, 次のように積分を用いて定義します.

$$\int_{-\infty}^{\infty} f(x') \delta(x - x') dx = f(x). \quad (10.8.19)$$

つまり, $\delta(x - x')$ にある関数 f をかけ全空間上を x' で積分すると, f の x での値がでてくるといいます. これを用いて正弦波が復元できるか確認してみましょう. $a(k, \omega) = \delta(k - k')\delta(\omega - \omega')$ として, Fourier 逆変換の定義式で積分変数を $k \rightarrow k', \omega \rightarrow \omega'$ と変えれば,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(k - k')\delta(\omega - \omega') e^{i(k'x - \omega't)} dk' d\omega' = e^{i(kx - \omega t)} \quad (10.8.20)$$

となり, 確かに単発の正弦波が復元できます. 少し話が飛びましたが, それでは Fourier 変換は? というのは当然の疑問でしょう. それは $a(k, \omega)$ のことです. これは u を用いて以下のように書けます.

$$a(k, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x, t) e^{-i(kx - \omega t)} dx dt \quad (10.8.21)$$

です。これを $\hat{u}(k, \omega), \mathcal{F}[u](k, \omega)$ などと書きます。Fourier 逆変換は $\mathcal{F}^{-1}[u](k, \omega)$ と書きます。

さらに少々線型代数を思い出してもらいましょう。行列 A に対して $Ax = \lambda x$ をみたす λ, x を固有値・固有ベクトルといいました。正確には変数分離という手法を用いた後に分かることですが、 $e^{i(kx - \omega t)}$ は線型作用素 \square の固有ベクトルになります。さらにスペクトル分解というのは、(正規) 行列を固有ベクトルと平行な成分への射影子と固有値の積を用いて分解するものでした。これを用いて微分作用素 (線型作用素!) \square のスペクトル分解を行なうと、

$$\square = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{1}{v^2} (-i\omega)^2 - (ik)^2 \right) e^{i(kx - \omega t)} dk d\omega \quad (10.8.22)$$

となります。ただし、これは作用素としての等式です。ここでは線型代数が無限次元に拡張されていく様を見ておいてください。

10.9 最大最小, 極値問題

本章のメインとなる節です。本節最後に安定点の定式化を行ないます。まず極値の正確な定義をしておきます。

Definition. 9.0.1. \mathbb{R}^n の部分集合 A で定義された実数値関数 f を考える。 $a \in A$ のある近傍 $U = U(a; \varepsilon)$ が A に含まれ (すなわち a が A の内点で), $f(a)$ が U における f の最大値である, すなわち任意の $x \in U$ に対して $f(x) \leq f(a)$ となるとき, f は a において極大であるという。極小についても同様。また f が a で極大であるだけでなく a 以外の任意の $x \in U$ に対し $f(a) > f(x)$ となるとき, f は a で狭義の極大であるという。狭義の極小についても同様。

f が a で極大または極小になるとき, f は a で極値を取るといい, a を f の極値点という。

1 変数のときの極値に関する結果は既知として, 多変数の極値に関して調べてみます。まず 1 変数のときとほとんど変わらない極値に関する必要条件が得られます。

Theorem. 9.0.2. \mathbb{R}^n の部分集合 D の内点 a で関数 $f: D \rightarrow \mathbb{R}$ が微分可能であるとすると, a で f が極値を取るならば次のことが成立する。

$$\frac{\partial f}{\partial x_i}(a) = 0 \quad (1 \leq i \leq n), \quad (10.9.1)$$

$$f'(a) = 0, \quad (df)_a = \mathbf{0}. \quad (10.9.2)$$

(証明) 本質が損なわれるわけでもなく, 理解がしやすくもなるので 2 変数関数 $f(x, y)$ で示します。 y を a の座標 a_2 に固定して得られる関数 $f_1(x) := f(x, a_2)$ は $x = a$ で極値を取るので, 1 変数の場合の対応する定理から

$$Df_1(x) = \frac{\partial f}{\partial x}(a) = 0 \quad (10.9.3)$$

が成立します。また, $(df)_a(x) = Df(a)x = 0$ が任意の $x \in \mathbb{R}$ に対して成立するので $df_a = 0$ です。■

本章はじめに述べたように $Df(a) = 0$ となる点を停留値といいます。上で示した極値に関する条件と安定点は極小点であることが必要という直観的考察から安定点では微分が消えます。また 1 変数のときの類推から, 2 次の導関数の挙

動を調べる必要があります。このとき座標系 (基底) の選択によらない 2 次微分 d^2f を考えるというのは以前書いたとおりです。ここで任意の $x \in \mathbb{R}^n$ に 2 次微分 d^2f を作用させると

$$(d^2f)_a(x) = \sum_{1 \leq i, j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) x_i x_j \quad (10.9.4)$$

となりますが、これは 2 次形式 (2 次同次式) と呼ばれるものです。そこで 2 次形式を正確に定義したのち、これを調べます。

Definition. 9.0.3. \mathbb{R}^n 上の 2 次形式 (quadratic form)

$$Q(x) = \sum_{i,j=1}^n b_{ij} x_i x_j, \quad b_{ij} = b_{ji} \in \mathbb{R} \quad (10.9.5)$$

に対し対称行列 $B = (b_{ij})$ をその係数行列という。 $Q(x) = (Bx, x) = {}^t x B x$ である。 Q に対し次のように定義する。

- 1) 全ての 0 でない $x \in \mathbb{R}^n$ に対し, $Q(x) > 0$ ならば, Q は正値であるという。
- 2) 全ての 0 でない $x \in \mathbb{R}^n$ に対し, $Q(x) < 0$ ならば, Q は負値であるという。
- 3) ある $x, y \in \mathbb{R}^n$ に対し, $Q(x) > 0 > Q(y)$ となるならば, Q は不定符号であるという。
- 4) $\det B \neq 0$ となるとき, Q は正則という。

Theorem. 9.0.4. 2 次形式 (10.9.5) に対して次の 5 つの条件は互いに同値である。

- 1) $x = 0$ で Q は狭義の最小値 0 を取る。
- 2) $x = 0$ で Q は狭義の極小値 0 を取る。
- 3) Q は正値である。
- 4) 係数行列 $B = (b_{ij})$ の固有値は全て正である。
- 5) 係数行列 B の全ての主小行列式 D_k は正である。すなわち

$$D_k = \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & b_{k2} & \cdots & b_{kk} \end{vmatrix} > 0 \quad (1 \leq k \leq n) \quad (10.9.6)$$

(証明) B が対称行列なので適当な直交行列 U を用いて Q を「対角化」できます。 $y := Ux$, B の固有値を λ_i とし, $UB^tU := \Lambda = (\lambda_i \delta_{ij})$ とすると,

$$Q(x) = {}^t x B x = {}^t (Ux) (UB^tU) (Ux) = {}^t y \Lambda y = \sum_{i=1}^n \lambda_i y_i^2 =: P(y). \quad (10.9.7)$$

これから P について考えていけば良いことになります。

- 1) \Rightarrow 2) 直観的には局所的な最小のことを極小というので、これは明らかでしょう。
- 2) \Rightarrow 3) 背理法で証明しましょう。 P が正値でないとします。このとき固有値の少なくとも 1 つ、例えば λ_1 が負になります。何故かという、正値の条件は「任意の y に対して $P(y) > 0$ となる」ことですが、この否定は「ある y が存在して $P(y) < 0$ となる」ことです。全ての λ_i が正であると、 y をどう取っても P は負になれません。

このとき $\varepsilon > 0$ を任意に取ります. さらに 0 の近傍 $U(0; \varepsilon)$ を考えます. そして $y_1 = (\varepsilon/2, 0, \dots, 0)$ とすると,

$$P(y_1) = \lambda_1 \left(\frac{\varepsilon}{2}\right)^2 < 0 \quad (10.9.8)$$

となり, $y = 0$ で狭義極小値 0 を取ることに反します.

3) \Rightarrow 4) こちらも背理法で示します. ある固有値, 例えば λ_1 が負であるとします. このとき例えば $y_0 = (1, 0, \dots, 0)$ を取ると $P(y_0) < 0$ となり, P , したがって Q が正値であることに反します.

4) \Leftrightarrow 5) 実際に主小行列式を計算します. このとき行列の積の行列式が行列式の積になること, すなわち $\det(AB) = \det A \cdot \det B$ となることを用いると,

$$D_n = \det B = \det B(\det U \det {}^t U) = \det(UB {}^t U) = \det \Lambda \quad (10.9.9)$$

$$= \begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \lambda_n \end{vmatrix} = (-1)^{2n} \lambda_n \begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \lambda_{n-1} \end{vmatrix} = \lambda_n D_{n-1} \quad (10.9.10)$$

これを繰り返していけば $D_n = \lambda_1 \cdots \lambda_n > 0$ となることが分かります.

4) \Rightarrow 1) 仮定から P の全ての固有値 λ_i は正です. 定義から任意の 0 でない $y \in \mathbb{R}^n$ に対して $P(y) > 0$ となります. したがって Q は $y = 0$ で狭義最小値 0 を取ります. ■

この定理を用いて, C^2 級関数 f の停留点 a が極値を取る為の 1 つの十分条件が与えられます.

Theorem. 9.0.5. \mathbb{R}^n の開集合 U で定義された C^2 級の実数値関数 f が, 1 点 $a \in U$ で $Df(a) = 0$ を満たすとき次のことが成立する.

- 1) \mathbb{R}^n 上の 2 次形式 $(d^2 f)_a$ が正値ならば, a は f の狭義の極小点である.
- 2) $(d^2 f)_a$ が負値ならば, a は f の狭義の極大点である.
- 3) $(d^2 f)_a$ が不定符号ならば, a は f の鞍点であり極値点ではない.

(証明) U が開集合ですから, a のある近傍で $U(a; \varepsilon) \in U$ となるものが存在します. そこで $|h| < \varepsilon$ となる任意の $h \in \mathbb{R}^n$ に対して Taylor の定理と仮定 $Df(a) = 0$ から

$$f(a+h) - f(a) = \frac{1}{2}(d^2 f)_{a+\theta h}(h), \quad \theta \in (0, 1). \quad (10.9.11)$$

1. の証明:

$(d^2 f)_a$ が正値ならば, 前定理からその係数行列の主小行列式 $D_k(a)$ に対して $D_k(a) > 0 (1 \leq k \leq n)$ が成立します. 今 f は C^2 級なので, 連続な関数 $D_{ij} f(x)$ を成分とする行列式 $D_k(x)$ は連続となります. そこで $\varepsilon > 0$ を十分小さく取ると

$$D_k(a + \theta h) > 0 (1 \leq k \leq n) \quad (10.9.12)$$

が成立します. そこで再び前定理を用いると, $|h| < \varepsilon, 0 < \theta < 1$ に対して 2 次形式 $(d^2 f)_{a+\theta h}$ は正値となります. そこで $|h| < \varepsilon$ となる任意の $h \neq 0$ に対して

$$f(a+h) - f(a) = \frac{1}{2}(d^2 f)_{a+\theta h}(h) > 0 \quad (10.9.13)$$

となります. これは a が f の狭義極小点であることを意味します.

1. の証明:

このとき $(d^2(-f))_a = -(d^2f)_a$ は正値ですから, 1) から $-f$ は a で狭義極小となります. したがって f は a で狭義極大となります.

1. の証明:

$(d^2f)_a$ が不定符号ならば,

$$(d^2f)_a(x) > 0 > (d^2f)_a(y) \quad (10.9.14)$$

となる $x, y \in \mathbb{R}^n$ が存在します. x, y を cx, cy で置き換えても上式が成立するので, $|x| < \varepsilon, |y| < \varepsilon$ と仮定できます. このとき実変数 $t \in (-1, 1)$ の 2 つの関数

$$g(t) = f(a + tx), \quad h(t) = f(a + ty) \quad (10.9.15)$$

を考えると, g, h は共に C^2 級で p 105, 定理 8.0.13 から,

$$D^k g(0) = (d^k f)_a(x), \quad D^k h(0) = (d^k f)_a(y) \quad (k = 1, 2) \quad (10.9.16)$$

となります. 今仮定から $(df)_a = 0$ ですから, $Dg(a) = Dh(a) = 0$ で, $D^2g(0) > 0 > D^2h(0)$ です. そこで $t = 0$ は g の狭義極小点で h の狭義極大点となります. すなわち a は f の鞍点で, f は x 方向には a で狭義極小となり, y 方向には a で狭義極大となります. したがって f は a で極大または極小になりません. ■

Remark. 9.0.6. この定理を固有値について述べなおすと, $(d^2f)_a$ の係数行列の全ての固有値が正のとき a で狭義極小を取り, 全ての固有値が負のとき a で狭義極大を取ることです.

またこれらを用いて安定点の定式化を行ないます. N 質点系がしたがうポテンシャルを $U(\mathbf{r}) := U(r_1, \dots, r_N)$ とし, U が適当に性質が良いものである (例えば C^∞ 級) とします. \mathbf{r}_0 がつりあいの位置でそこからのずれを $\mathbf{y} = \mathbf{r} - \mathbf{r}_0$ とします. ここで U を \mathbf{r}_0 のまわりで Taylor 展開すると

$$U(\mathbf{r}) = U(\mathbf{r}_0) + \sum_{i=1}^{3N} \frac{\partial U}{\partial x_i} \Big|_{\mathbf{r}_0} y_i + \frac{1}{2} \sum_{i,j=1}^{3N} \frac{\partial^2 U}{\partial x_i \partial x_j} \Big|_{\mathbf{r}_0} y_i y_j + O(|\mathbf{y}|^3) \quad (10.9.17)$$

$$= U(\mathbf{r}_0) + \frac{1}{2} \sum_{i,j=1}^{3N} \frac{\partial^2 U}{\partial x_i \partial x_j} \Big|_{\mathbf{r}_0} y_i y_j \quad (10.9.18)$$

となります. ここで $\mathbf{r}_1 = (x_1, x_2, x_3), \mathbf{r}_2 = (x_4, x_5, x_6), \dots, \mathbf{r}_N = (x_{3N-2}, x_{3N-1}, x_{3N})$ としました. 考えるのは平衡点まわりの微小振動なので, $O(|\mathbf{y}|^3) \approx 0$ とみなせます. 本章はじめの節での考察とエネルギーが低い方がより安定であるという一般的な原理から, 系が安定である為には

$$U(\mathbf{r}) > U(\mathbf{r}_0) \implies \sum_{i,j=1}^{3N} V_{ij} y_i y_j > 0 \quad (V_{ij} = V_{ji} = D_{ij} U(\mathbf{r}_0)) \quad (10.9.19)$$

でなければなりません. すなわち安定性の条件は $\hat{U} = (U_{ij})$ が正値行列であるということです. この条件を用いて本章の主題であるポテンシャルの安定点近傍で質点系は微小振動を行なうことを示します. これは一旦節を改めましょう.

10.10 物理への適用

本節では時間によらないポテンシャルの安定点近傍で質点系は微小振動を行なうことを示します。まず r_i の質点の質量を m_i とします。各座標成分に対する運動方程式は、

$$m_i \ddot{y}_i = - \sum_{j=1}^{3N} U_{ij} y_j \quad (10.10.1)$$

となります。この式を $\sqrt{m_i}$ で割って整理すると、

$$\sqrt{m_i} \ddot{y}_i = - \sum_{j=1}^{3N} \frac{1}{\sqrt{m_i m_j}} U_{ij} \sqrt{m_j} y_j \quad (10.10.2)$$

なります。ここでさらに $\tilde{y}_i := \sqrt{m_i} y_i$, $\tilde{U}_{ij} := U_{ij} / \sqrt{m_i m_j}$ と定義すると、

$$\ddot{\tilde{y}}_i = - \sum_{j=1}^{3N} \tilde{U}_{ij} \tilde{y}_j \iff \ddot{\tilde{\mathbf{y}}} = - \tilde{\hat{U}} \tilde{\mathbf{y}} \quad (10.10.3)$$

と書けます。今 $\tilde{\hat{U}}$ は対称行列なので、適当な直交行列 V で対角行列 $\Lambda = (\lambda_i \delta_{ij})$ に対角化可能です。また前節での注意 p 112, 注意 9.0.6 から正値行列 \hat{U} の固有値 λ_i は全て正であることから

$$\lambda_i = \omega_i^2, \quad \omega_i > 0 \quad (10.10.4)$$

としても良いことが分かります。ここで U が時間依存しないのでこれを対角化する V も時間依存しません。したがって $\dot{V} = 0$ となり、 $\Omega := (\omega_i^2 \delta_{ij})$ としてこれを用いると運動方程式は

$$\frac{d^2}{dt^2} V \mathbf{y} = -V \hat{U} V^{-1} V \mathbf{y} \iff \ddot{\mathbf{z}} = -\Omega \mathbf{z}, \quad \mathbf{z} := V \mathbf{y} \quad (10.10.5)$$

となります。 Ω が対角行列であることを用いて成分ごとに書けば

$$\ddot{z}_i = -\omega_i^2 z_i \quad (10.10.6)$$

となります。これは簡単に解けて

$$z_i = a_i e^{i\omega_i t} + b_i e^{-i\omega_i t} = A_i \sin(\omega_i t + \phi_i) \quad (10.10.7)$$

と書けます。これを逆に解いて

$$\mathbf{y} = \text{diag} \left(\frac{1}{\sqrt{m_i}} \right) V^{-1} \mathbf{z} \quad (10.10.8)$$

が得られます。ここで diag は括弧の中身を成分とする対角行列を表します。

以上から互いに相互作用を持つ N 個の質点に対する微小振動の方程式が N 個の互いに独立な調和振動子の集合として記述できたことになります。この振動を基準振動 (normal mode) といいます。基準振動によって系を記述する考え方は、多自由度振動系の基本であり、構造力学、分子モデルなど多くの応用例があるほか、場の量子論の基礎付けでも重要な役割を果たすようです。

これで本章の目的は果たしたのですが、以後のことも考えて積分についても議論することにします。その前に微分方程式と絡めて動機づけをしておきます。

10.11 微分と積分の掛け橋-微分方程式

運動方程式 $m\ddot{r} = f$ は 2 階の時間微分を含んでおり、さらにその微分されて埋まっている関数が求めるべきものです。こうしたものを微分方程式と呼びます。ここではこれを (近似的に) 解く方法を考えます。

本当に求めたいのは質点が力 f を受けながら描く運動の軌跡です。しかしいきなり 2 階微分を解きほぐすのは大変です。欲張らずにまずは速度 $\dot{r} = v$ を求めてみましょう。初期時刻を $t = t_i$ 、初期速度 $v(t_i) := v_0$ として、適当な時間 $t_f > 0$ での質点の速度を求めることにします。前の節で微分とは線型化であるということを見ました。非常に短い時間間隔 $\Delta t > 0$ を取れば、質点に働く力は Δt の間 $f(r_{t_i}, \dot{r}_{t_i})$ で一定であるとみなせるでしょう。ただし $r_{t_i} := r(t_i)$, $\dot{r}_{t_i} := \dot{r}(r(t_i), t_i)$ としました。すると線型化によって、運動方程式は

$$v(t_i + \Delta t) \simeq v_0 + \frac{\Delta t}{m} f(r_{t_i}, \dot{r}_{t_i}) \quad (10.11.1)$$

と書けます。速度のなす 3 次元空間で考えれば、当然これはその空間の中の線分となります。ここからさらに Δt 経ったときの速度 $v(t_0 + 2\Delta t)$ は、同様の線型化によって

$$v(t_i + 2\Delta t) \simeq v(t_i + \Delta t) + \frac{\Delta t}{m} f(r_{t_i + \Delta t}, \dot{r}_{t_i + \Delta t}) \quad (10.11.2)$$

$$\simeq v_0 + \frac{\Delta t}{m} f(r_{t_i}, \dot{r}_{t_i}) + \frac{\Delta t}{m} f(r_{t_i + \Delta t}, \dot{r}_{t_i + \Delta t}) \quad (10.11.3)$$

と書けます。先程と同じですが、こちらは折れ線になっています。これを n 回繰り返すと

$$v(t_0 + n\Delta t) \simeq v_0 + \sum_{k=0}^n f(r_{t_0 + k\Delta t}, \dot{r}_{t_0 + k\Delta t}) \quad (10.11.4)$$

となります。十分大きな N を取れば、 $t_i + N\Delta t \simeq t_f$ とできるでしょう。したがって

$$v(t_f) \simeq v_0 + \sum_{k=0}^N f(r_{t_0 + k\Delta t}, \dot{r}_{t_0 + k\Delta t}) =: v_p(t; \Delta t; N) \quad (10.11.5)$$

が得られます。作り方からこれは求めるべき速度の折れ線近似のグラフを与えることが分かります。

近似という程度が低いと思うかもしれませんが。しかし例えばボールを投げたとき、滞空時間が 10 秒であったとします。空気抵抗などが無いものとすれば、当然ボールは放物線を描きます。このとき Δt を $1/20000$ 秒、 $N = 2 \times 10^5$ として上の折れ線近似を頑張ってやれば、折れ線といえどもかなり正確に放物線様の軌道を描いてくれるでしょう。同様のプロセスを踏んで r の (近似) 解が求められます。つまり原理的にはこれできちんと近似解 $v_p(t; \Delta t; N)$ が求められます。また近似というならば精度を高めたいと思うのは当然の欲求でしょう。そのためには $\Delta t > 0$ をさらに小さく、したがって N もさらに大きくすれば良いと考えるのは普通でしょう。

このとき $\Delta t \rightarrow 0 (N \rightarrow \infty)$ という理想的な極限を考えます。そうすると各 Δt での線分が短くなっていき、次第に近似解の折れ線 $v_p(t; \Delta t; N)$ が真の解曲線 v

になっていくでしょう。このとき上の式の和の中では $\Delta t \rightarrow 0, N \rightarrow \infty$ という極限を同時に取ることになります。つまり

$$\mathbf{v}_p(t; \Delta t; N) = \mathbf{v}_0 + \sum_{k=0}^N \mathbf{f}(\mathbf{r}_{t_0+k\Delta t}, \dot{\mathbf{r}}_{t_0+k\Delta t}) \quad (10.11.6)$$

$$\xrightarrow{\Delta t \rightarrow 0, N \rightarrow \infty} \mathbf{v} = \mathbf{v}_0 + \lim_{\Delta t \rightarrow 0, N \rightarrow \infty} \sum_{k=0}^N \mathbf{f}(\mathbf{r}_{t_0+k\Delta t}, \dot{\mathbf{r}}_{t_0+k\Delta t}) \quad (10.11.7)$$

となります。右辺で現れる和の極限が上手く収束するときがあります。そこでこの和の極限を積分と名付けて以下のように書きます。

$$\int_{t_i}^{t_f} \mathbf{f}(\mathbf{r}(\tau), \dot{\mathbf{r}}(\tau)) d\tau \quad (10.11.8)$$

$$:= \lim_{\Delta t \rightarrow 0, N \rightarrow \infty} \sum_{k=0}^N \mathbf{f}(\mathbf{r}_{t_0+k\Delta t}, \dot{\mathbf{r}}_{t_0+k\Delta t}). \quad (10.11.9)$$

実際には収束の議論をするのは一苦労です。そのあたりを以下で簡単に見る他、実際に必要となる計算結果、変数変換公式などを導出していきます。

10.12 運動方程式と常微分方程式

10.12.1 常微分方程式

何度か述べたように、力学の基本法則である運動方程式は数学的には(常)微分方程式というものになります。運動方程式を考えれば分かるように、求める解(関数)の具体的な形は分からなくとも、その満たすべき関係式が分かっているときがあります。その関係式から関数の具体的な形を求めることを、一般に(微分)方程式を解くといいます。解く際に許される演算は

- 1)有限回の微分,
- 2)有限回の与えられた関数との合成,

のみです。波動方程式を考えれば分かるように、物理で現れる変数は一般に時間と空間の4変数があります。多変数の微分方程式を偏微分方程式といいます。ここでは時間のみを変数とする常微分方程式のみを考えます。

まず1つ注意しておきます。力として $\mathbf{f} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ という Lorentz 力を取ります。関数の引数をきちんと示して、質量 m 、電荷 q を持つ質点の運動方程式を書くと

$$m\ddot{\mathbf{r}}(t) = q\{\mathbf{E}(\mathbf{r}(t), t) + \dot{\mathbf{r}}(t) \times \mathbf{B}(\mathbf{r}(t), t)\} \quad (10.12.1)$$

という式になります。求めるべき \mathbf{r} が両辺に入っている上、時間 t の変化に対して一般にはひどく複雑な挙動を取るでしょう。しかし積分の導入のところで示したように、 \mathbf{E} や \mathbf{B} が時間的にあまりにも複雑すぎる変化をしなければ、非常に短い時間に区切って線型化して、その折れ線近似で近似解を構成できるでしょう。ここから時間分割を小さくしていくことで厳密な解が「求められる」でしょう。

「求められる」と書きましたが、これは具体的に $f = A \sin \omega t$ などの数式で書けるとは限りません。むしろ、このようにきちんと数式で解が表現出来ないほうが普通です。しかし何といたっても上で考えたようにすれば、運動方程式には解

が存在するだけでなく、実験的な状況を考えれば同じ初期条件を与えれば同じ結果が出る、という意味で解が一意的であろうと推測されます。この解の存在と一意性は数学的に比較的緩い条件下で厳密に示すことができます。この章ではこの解の存在と一意性を示したのち、具体的ないくつかの微分方程式の解法を述べることにします。

10.12.2 初期値問題の解の存在と一意性

準備

1 番初めに私達に課せられた課題がかなり重いものであることを改めて確認します。前章で N 質点系の安定点近傍での微小振動を示しました。これから私達は N 質点系でも成立するような形で定理を示さなければなりません。さらに前節の Lorentz 力のところで見たように、力としても求めるべき系の軌道 \mathbf{r} や速度 $\dot{\mathbf{r}}$ に依存します。 k 番目の質点の質量を m_k 、位置を \mathbf{r}_k 、加わる力を \mathbf{f}_k として、扱うべき微分方程式は

$$m_1 \ddot{\mathbf{r}}_1 = \mathbf{f}_1(\mathbf{r}_1, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N, t) \quad (10.12.2)$$

$$\vdots \quad (10.12.3)$$

$$m_N \ddot{\mathbf{r}}_N = \mathbf{f}_N(\mathbf{r}_1, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N, t) \quad (10.12.4)$$

$$\text{初期条件: } \mathbf{r}_k(t_0) = \mathbf{r}_{k0}, \quad \dot{\mathbf{r}}_k(t_0) = \mathbf{v}_k(t_0) = \mathbf{v}_{k0} \quad (k = 1, 2, \dots, N) \quad (10.12.5)$$

となります。ここで力は加速度には依存しないものとししました。ここでやや作為的ですが、以下のように考えます。まず $\mathbf{r}_i = {}^t(x_i, y_i, z_i) := {}^t(x_{6i-5}, x_{6i-3}, x_{6i-1})$ とします。そして次のように定義します。

$$\mathbf{r}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \dot{\mathbf{r}}_1 = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} := \begin{pmatrix} x_{3N+1} \\ x_{3N+2} \\ x_{3N+3} \end{pmatrix}, \quad (10.12.6)$$

$$\vdots \quad (10.12.7)$$

$$\mathbf{r}_N = \begin{pmatrix} x_{3N-2} \\ x_{3N-1} \\ x_{3N} \end{pmatrix}, \quad \dot{\mathbf{r}}_N = \begin{pmatrix} \dot{x}_{3N-2} \\ \dot{x}_{3N-1} \\ \dot{x}_{3N} \end{pmatrix} := \begin{pmatrix} x_{6N-2} \\ x_{6N-1} \\ x_{6N} \end{pmatrix}, \quad (10.12.8)$$

$$\mathbf{R} := \begin{pmatrix} x_1 \\ \vdots \\ x_{3N} \\ \dot{x}_1 \\ \vdots \\ \dot{x}_{3N} \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_{3N} \\ x_{3N+1} \\ \vdots \\ x_{6N} \end{pmatrix}, \quad \mathbf{f}_i := \begin{pmatrix} f_{3i-2}(\mathbf{R}, t) \\ f_{3i-1}(\mathbf{R}, t) \\ f_{3i}(\mathbf{R}, t) \end{pmatrix}, \quad (10.12.9)$$

$$\Rightarrow \dot{\mathbf{R}} = \begin{pmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_{3N} \\ \dot{x}_{3N+1} \\ \vdots \\ \dot{x}_{6N} \end{pmatrix} = \begin{pmatrix} x_{3N+1} \\ \vdots \\ x_{6N} \\ \ddot{x}_1 \\ \vdots \\ \ddot{x}_{3N} \end{pmatrix} = \begin{pmatrix} x_{3N+1} \\ \vdots \\ x_{6N} \\ f_1/m_1 \\ \vdots \\ f_{3N}/m_{3N} \end{pmatrix} := \mathbf{F}(\mathbf{R}, t) \quad (10.12.10)$$

$$\text{初期条件: } \mathbf{R}(t_0) = \mathbf{R}_0 \quad (10.12.11)$$

今までは多数の式を統制する必要がありましたが、こうすることで一本の式を制御すればよく、見かけ上 1 変数のように扱えるようになり、数学的な見通しが非常に良くなります。

正確には $D^n x = f(x, Dx, \dots, D^{n-1}x, t)$ という形、つまり最高階の微分について 1 価に解けているものです。}

先程 F があまり複雑な変化をしなければ解の存在と一意性が示せると書きました。どこまでの変化を許容できるか、ということで 1 つ定義をします。

Definition. 12.2.1. (*Lipschitz*(リプシッツ) 条件) 関数 F は $D := \mathbb{R}^n \times I$ を定義域とし、値域が \mathbb{R}^m であるとする。 F が *Lipschitz* 条件を満たすというのは、ある正数 L が存在して D の任意の 2 点 $(\mathbb{R}_1, t), (\mathbb{R}_2, t)$ に対して以下の式を満たすことをいう。

$$|F(\mathbb{R}_1, t) - F(\mathbb{R}_2, t)| \leq L|\mathbb{R}_1 - \mathbb{R}_2|. \quad (10.12.12)$$

このとき $|\mathbb{R}_1 - \mathbb{R}_2| < \varepsilon =: \delta/L$ とすれば $|F(\mathbb{R}_1, t) - F(\mathbb{R}_2, t)| < \varepsilon$, つまり F は R について連続となります。ただし x^2 などはこの Lipschitz 条件を満たしませんから、Lipschitz 条件はただの連続性よりは強い条件です。

微分方程式の解の存在と一意性の証明 1

以下では有界閉領域

$$D := \{(\mathbf{R}, t); |\mathbf{R} - \mathbf{R}_0| \leq b, |t - t_0| < a \ (0 < a, b < \infty)\} \quad (10.12.13)$$

上で考えます。連続関数は有界閉集合上で最大値、最小値を取るの

$$\max_{(\mathbf{R}, t) \in D} |F(\mathbf{R}, t)| = M < \infty \quad (10.12.14)$$

となります。

さて、定理の statement を述べましょう。

Theorem. 12.2.2. (微分方程式の解の存在と一意性 1) $F(\mathbf{R}, t)$ を $D = \mathbb{R}^N \times I$ において *Lipschitz* 条件を満たす 連続関数とする。このとき $t_0 \in I$ ならば、任意の初期値 $\mathbf{R}_0 \in \mathbb{R}^N$ に対して初期値問題

$$\frac{d}{dt}\mathbb{R}(t) = F(\mathbf{R}, t), \quad \mathbb{R}(t_0) = \mathbf{R}_0 \quad (10.12.15)$$

を

$$|t - t_0| \leq h, \quad h := \min\left(a, \frac{b}{M}\right) \quad (10.12.16)$$

で満たす一意的な解 R が存在する。

(証明) 第 1 段: “折れ線” \mathbb{R}_ε の作り方。

有界閉領域 D において連続なので、Heine の定理から F は D において一様連続です。すなわち任意の $\varepsilon > 0$ に対して $\kappa = \kappa(\varepsilon)$ を定めて

$$(\mathbb{R}_1, t_1) \in D, (\mathbb{R}_2, t_2) \in D \text{ かつ } |\mathbf{R}_1 - \mathbf{R}_2| \leq \kappa, |t_1 - t_2| \leq \kappa \quad (10.12.17)$$

$$\implies |F(\mathbf{R}_1, t_1) - F(\mathbf{R}_2, t_2)| \leq \varepsilon. \quad (10.12.18)$$

とできます. さらに

$$\delta := \min \left(\kappa, \frac{\kappa}{M}, \frac{b}{M} \right) > 0 \quad (10.12.19)$$

を取ります. $0 < |t - t_0| \leq h$ となる任意の t に対して 0 , または正整数 $n = n(\delta, t) = n(\delta(\varepsilon), t) = n_\varepsilon(t)$ を

$$n\delta < |t - t_0| \leq (n+1)\delta \quad (10.12.20)$$

となるように定めて, Cauchy の折れ線 $\mathbb{R}_\varepsilon(t)$ を

$$\begin{cases} \mathbb{R}_\varepsilon(t) = \mathbb{R}_n + (t - t_n)\mathbf{F}(\mathbf{R}_n, t_n), \\ \text{ただし} \\ t_j := t_0 + j \cdot \delta \frac{t - t_0}{|t - t_0|}, \\ \mathbf{R}_j := \mathbf{R}_0 + \delta \frac{t - t_0}{|t - t_0|} \mathbf{F}(\mathbb{R}_{j-1}, t_{j-1}) (j = 1, 2, \dots, n) \\ \text{また } \mathbf{R}_\varepsilon(t_0) = \mathbf{R}_0 \end{cases} \quad (10.12.21)$$

によって与えます. この折れ線が作成可能なことをいうためには, 点 (\mathbf{R}_j, t_j) が全て D に属することをいかなければなりません. そうでないと $\mathbf{F}(\mathbf{R}_j, t_j)$ が定義されないからです. まず (10.12.16), (10.12.20) によって

$$|t_j - t_0| \leq j\delta \leq n\delta < |t - t_0| \leq h \leq a \quad (j = 1, 2, \dots, n). \quad (10.12.22)$$

となります. 同じく $\{(10.12.19), (10.12.20)\}$ によって, $|\mathbf{R}_1 - \mathbf{R}_0| \leq \delta |\mathbf{F}(\mathbf{R}_0, t_0)| \leq \delta M \leq b$ となります. したがって $\mathbb{R}_2, \mathbb{R}_3, \dots, \mathbb{R}_{j-1}$ が全て $|\mathbb{R}_k - \mathbb{R}_0| \leq b$ ($k = 1, 2, \dots, j-1$) を満たすとする, $(\mathbf{R}_k, t_k) \in D$, したがって $\mathbf{F}(\mathbf{R}_k, t_k)$ ($k = 1, 2, \dots, j-1$) が定義されて, しかも

$$|\mathbf{R}_j - \mathbf{R}_0| \leq \sum_{k=1}^j |\mathbf{R}_k - \mathbf{R}_{k-1}| = \sum_{k=1}^j \delta |\mathbf{F}(\mathbf{R}_{k-1}, t_{k-1})| \quad (10.12.23)$$

$$\leq j\delta M \leq n\delta M < |t - t_0| M \leq hM \leq b \quad (10.12.24)$$

を満たします. こうして j に関する帰納法から点 (\mathbf{R}_j, t_j) ($j = 1, 2, \dots, n$) が全て D に属することが分かりました.

第2段:折れ線 $\mathbf{R}_\varepsilon(t)$ が“近似微分方程式”の解であることの証明.

折れ線 $\mathbb{R}_\varepsilon(t)$ はその角である有限個の点 x_1, x_2, \dots, x_n では

$$\frac{d\mathbf{R}_\varepsilon(t)}{dt} = \mathbf{F}(\mathbf{R}_n, t_n) = \mathbf{F}(\mathbf{R}_{n_\varepsilon(t)}, t_{n_\varepsilon(t)}) \quad (10.12.25)$$

を満たします. 一方 $\{(10.12.19), (10.12.20), (10.12.21)\}$ から

$$|t - t_{n_\varepsilon(t)}| \leq \delta \leq \kappa, \quad (10.12.26)$$

$$|\mathbf{R}_\varepsilon(t) - \mathbf{R}_{n_\varepsilon(t)}| \leq \delta |\mathbf{F}(\mathbf{R}_{n_\varepsilon(t)}, t_{n_\varepsilon(t)})| \leq \delta M \leq \kappa \quad (10.12.27)$$

となります. さらに $\{(10.12.18)\}$ によって $|\mathbf{F}(\mathbf{R}_\varepsilon(t), t) - \mathbf{F}(\mathbf{R}_{n_\varepsilon(t)}, t_{n_\varepsilon(t)})| \leq \varepsilon$ となります. したがって折れ線 $\mathbf{R}_\varepsilon(t)$ は有限個の角以外では

$$\left| \frac{d\mathbf{R}_\varepsilon(t)}{dt} - \mathbf{F}(\mathbf{R}_\varepsilon(t), t) \right| \leq \varepsilon \quad (10.12.28)$$

を満たします.

第 3 段: 2 つの折れ線 $R_\varepsilon(t), R_{\varepsilon'}(t)$ の比較.

$\varepsilon' > 0$ に対する折れ線 $R_{\varepsilon'}(t)$ も有限個の角以外では

$$\left| \frac{dR_{\varepsilon'}(t)}{dt} - F(R_{\varepsilon'}(t), t) \right| \leq \varepsilon' \quad (10.12.29)$$

を満たします. したがって Lipschitz 条件から有限個の角以外では

$$\left| \frac{dR_\varepsilon(t)}{dt} - \frac{dR_{\varepsilon'}(t)}{dt} \right| \leq L|R_\varepsilon(t) - R_{\varepsilon'}(t)| + \varepsilon + \varepsilon' \quad (10.12.30)$$

を満たします. そして $R_\varepsilon(t) - R_{\varepsilon'}(t)$ は連続関数, その上 $R_\varepsilon(t_0) - R_{\varepsilon'}(t_0) = 0$ ですから,

$$R_\varepsilon(t) - R_{\varepsilon'}(t) = \int_{x_0}^x \left\{ \frac{dR_\varepsilon(s)}{ds} - \frac{dR_{\varepsilon'}(s)}{ds} \right\} ds \quad (10.12.31)$$

であり, $U(t) := |R_\varepsilon(t) - R_{\varepsilon'}(t)|$ とすると, (10.12.30) によって

$$U(t) \leq \left| \int_{t_0}^t LU(s) ds \right| + (\varepsilon + \varepsilon') \left| \int_{x_0}^x ds \right| \quad (10.12.32)$$

となります. したがって $\max_{|t-t_0| \leq h} U(x) =: M'$ とすると (10.12.32) から

$$U(t) \leq |t - t_0| (LM' + \varepsilon + \varepsilon') \quad (10.12.33)$$

となるので, これを (10.12.32) の右辺に代入して

$$U(t) \leq \frac{|t - t_0|^2}{2!} L(LM' + \varepsilon + \varepsilon') + (\varepsilon + \varepsilon')|t - t_0|. \quad (10.12.34)$$

これをまた (10.12.32) の右辺に代入して

$$U(t) \leq \frac{|t - t_0|}{3!} L^2(LM' + \varepsilon + \varepsilon') + (\varepsilon + \varepsilon')L \frac{|t - t_0|}{2!} + (\varepsilon + \varepsilon')|t - t_0| \quad (10.12.35)$$

となります. 以下同様に繰り返し, さらに e^x の Taylor 展開も用いて

$$U(t) \leq \frac{|t - t_0|^n}{n!} L^{n-1}(LM' + \varepsilon + \varepsilon') + (\varepsilon + \varepsilon') \sum_{m=1}^{n-1} \frac{|t - t_0|^m L^{m-1}}{m!} \quad (10.12.36)$$

$$\leq \frac{|t - t_0|^n}{n!} L^{n-1}(LM' + \varepsilon + \varepsilon') + (\varepsilon + \varepsilon') \frac{1}{L} \left(e^{L|t-t_0|} - 1 \right) \quad (10.12.37)$$

となりますが, ここで $n \rightarrow \infty$ として

$$|R_\varepsilon(t) - R_{\varepsilon'}(t)| \leq \frac{\varepsilon + \varepsilon'}{L} \left(e^{L|t-t_0|} - 1 \right) \quad (10.12.38)$$

が得られます. 次に行く前に上の極限の成立を示します. 証明すべきは $a > 0$ に対して $a^n/n! \rightarrow 0 (n \rightarrow \infty)$ です. まずは $2a < k$ となる数 k を 1 つ取ります. $n > k$ に対して

$$\frac{a}{n} < \frac{k}{2n} < \frac{1}{2} \quad (10.12.39)$$

ですから,

$$\frac{a^n}{n!} = \frac{a}{1} \frac{a}{2} \cdots \frac{a}{k-1} \frac{a}{k} \frac{a}{k+1} \cdots \frac{a}{n} < \frac{a^k}{k!} \cdot \left(\frac{1}{2}\right)^{n-k} = \frac{a^k}{k!} \frac{1}{2^k} \left(\frac{1}{2}\right)^n \quad (10.12.40)$$

となります. $(1/2)^n \rightarrow 0 (n \rightarrow \infty)$ から $a^n/n! \rightarrow 0$ が分かります.

第 4 段: 解の存在.

$\varepsilon > 0$ を与えた場合に少なくとも 1 つ折れ線 $R_\varepsilon(t)$ が定まりました. 少なくとも 1 つ, というのは例えば (10.12.18) における $\kappa = \kappa(\varepsilon)$ の定め方が一意的でないことによります. しかし少なくとも 1 つ定まった $R_\varepsilon(t)$ は (10.12.38) を満足するので, $\varepsilon = 1/k (k \in \mathbb{Z}^+)$ とすれば, $|t - t_0| \leq h$ において一様に

$$R(t) := \lim_{n \rightarrow \infty} R_{1/n}(t) \quad (10.12.41)$$

が存在します. $a_n(t) := R_{1/n}(t)$ とすれば全ての n について a_n は連続関数となります. (10.12.38) を a_n で書き直すと

$$|a_n(t) - a_m(t)| \leq \frac{m+n}{Lmn} (e^{L|t-t_0|} - 1) \quad (10.12.42)$$

であり, これから $(a_n)_{n \in \mathbb{N}}$ が Cauchy 列であることが分かります. 一様ノルムに対して連続関数の空間は Banach 空間になるので収束先が連続関数として確かに存在します.

議論を進めましょう. $R(t)$ は各 $R_{1/n}(t)$ と共に初期条件 $R(t_0) = R_0$ を満たし, かつ連続です. $R_\varepsilon(t_0) = R_0$ となる連続関数 $R_\varepsilon(t)$ を代入し, (10.12.28) から

$$\left| R_\varepsilon(t) - R_0 - \int_{t_0}^t F(R_\varepsilon(s), s) ds \right| \leq \varepsilon |t - t_0| \leq \varepsilon h \quad (10.12.43)$$

が得られます. $\varepsilon = 1/n$ として $n \rightarrow \infty$ とすると (10.12.41) から一様収束で

$$R(t) - R_0 = \int_{t_0}^t F(R(s), s) ds \quad (10.12.44)$$

となります. 右辺括弧内が連続関数であり, 連続関数の合成はまた連続関数になるので, R は微分可能でかつ

$$\frac{dR(t)}{dt} = F(R(t), t) \quad (10.12.45)$$

であり, こうして $R(t_0) = R_0$ となる解が少なくとも 1 つ $|t - t_0| \leq h$ において存在することがわかりました.

第 5 段: 解の一意性.

背理法で証明します. $R(t)$ の他に $P(t_0) = R_0$ となる解 P が $|t - t_0| \leq h' \leq h$ において存在したとしましょう. P を誤差 $\varepsilon' = 0$ の近似解と考えると

$$\left| \frac{dP(t)}{dt} - F(P(t), t) \right| = 0 \quad (10.12.46)$$

と書けます. (10.12.38) を得たときと同様にして

$$|R_\varepsilon(t) - P(t)| \leq \frac{\varepsilon}{L} (e^{L|t-t_0|} - 1) \quad (10.12.47)$$

と書けます. したがって $\varepsilon = 1/n$ の n を ∞ にすると $R = P$ となり, 証明終了です. ■

微分方程式の解の存在と一意性の証明 2 の準備-局所 Lipschitz 条件

まず次の定理の証明からはじめます。

Theorem. 12.2.3. $F(\mathbf{R}, t) : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^l$ が変数 \mathbf{R} に対して Lipschitz 条件を満たすならば, K を Lipschitz 定数, C をある定数として

$$|F(\mathbf{R}, t)| \leq L(|\mathbf{R}| + C). \quad (10.12.48)$$

(証明) 適当に $\mathbf{R}_0 \in \mathbb{R}^m$ を取ってきます。このとき三角不等式を用いて

$$|F(\mathbf{R}, t)| \leq |F(\mathbf{R}, t) - F(\mathbf{R}_0, t)| + |F(\mathbf{R}_0, t)| \quad (10.12.49)$$

$$\leq L|\mathbf{R} - \mathbf{R}_0| + |F(\mathbf{R}_0, t)| \quad (F \text{ は Lipschitz 条件を満たす}) \quad (10.12.50)$$

$$\leq L(|\mathbf{R}| + |\mathbf{R}_0|) + |F(\mathbf{R}_0, t)| \leq L(|\mathbf{R}| + C). \blacksquare \quad (10.12.51)$$

この定理によって, Lipschitz 条件を満たす関数は高々一次の増大度であるということが分かります。 x^2 が Lipschitz 条件を満たさないと前に書きましたが, それはこのことです。今しがた示した定理は非常に適用範囲が狭いということです。つまり Lipschitz 条件から非常に嬉しい結果が出ましたが, 物理への適用という点から見て, こんなものはハリボテのおもちゃです。もっと緩い条件で定理が成立してくれないと物理としてはとても困ります。多少制限が強くなりますが, 幸いにもこれよりも緩い条件で同じような結果が出せます。その為にまず定義を 1 つ出します。

Definition. 12.2.4. (局所 Lipschitz 条件) $G : \Omega \rightarrow \mathbb{R}^n$ が Ω で \mathbf{R} に対して局所 Lipschitz 条件を満たすというのは, 任意の $K \subset \Omega$ となる有界閉集合 K に対して定数 $L_K > 0$ が存在して, 任意の $(\mathbf{R}_1, t), (\mathbf{R}_2, t)$ に対して以下の式が成立することをいう。

$$|G(\mathbf{R}_1, t) - G(\mathbf{R}_2, t)| \leq L_K |\mathbf{R}_1 - \mathbf{R}_2|. \quad (10.12.52)$$

ある集合 Ω に K が有界閉集合として含まれるということを $K \Subset \Omega$ と書くことにします。すぐあとで示しますが, 局所 Lipschitz 条件を満たす為の 1 つの十分条件として G が C^1 級である, というものがあります。(普通の) Lipschitz 条件を満たしても微分不可能な関数が存在するので, C^1 級というのはかなり厳しい条件ではないか, という人もいるかもしれませんが, 当面物理では不連続性の激しいものは扱いません。むしろ C^1 級というのは緩すぎるくらい緩い条件です。

一意性存在定理 2 の証明に入る前に予備定理を 2 つ証明しておきます。

Theorem. 12.2.5. $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ が C^1 級ならば, これは局所 Lipschitz 条件を満たす。

(証明) 任意に $K \Subset \mathbb{R}^m$ を取ります。ただしこのとき K は凸集合であるとし, 凸集合というのは, 任意の $x, y \in K$ に対して x と y を結ぶ線分も K に含まれる集合です。 $F = (F_1, \dots, F_n)$, $\mathbf{R} = (x_1, \dots, x_m)$ とします。 C^1 級の仮定から F の全ての偏導関数は連続です。 K は有界閉集合 (コンパクト集合) であり, 有界閉集合上の連続関数は必ず最大・最小値を取ります。そこで以下のように定義します。

$$L_K^k := \max_{\mathbf{X} \in K, 1 \leq j \leq m} \left| \frac{\partial}{\partial x_j} F_k(\mathbf{X}) \right| < +\infty \quad (\text{有限値}). \quad (10.12.53)$$

三角不等式と平均値の定理から,

$$|\mathbf{F}(\mathbf{X}) - \mathbf{F}| \leq \sum_{i=1}^n |F_i(\mathbf{X}) - F_i(\mathbf{Y})| \text{ (三角不等式)} \quad (10.12.54)$$

$$\sum_{i=1}^n \left| \sum_{j=1}^m \frac{\partial}{\partial x_j} F_i(\mathbf{X} + \theta(\mathbf{Y} - \mathbf{X}))(Y_j - X_j) \right|, \theta \in (0, 1) \text{ (平均値の定理)} \quad (10.12.55)$$

$$\leq \sum_{i=1}^n \left| \sum_{j=1}^m L_K^i |Y_j - X_j| \right| \times 1 \quad (10.12.56)$$

$$\leq \sum_{i=1}^n L_K^i \sqrt{\sum_{j=1}^m (Y_j - X_j)^2} \times \sqrt{\sum_{j=1}^m 1^2} \text{ (Cauchy-Schwartz の不等式)} \quad (10.12.57)$$

$$\leq \sqrt{m} \sum_{i=1}^n L_K^i |\mathbf{X} - \mathbf{Y}| \quad (10.12.58)$$

$$\leq L_K |\mathbf{X} - \mathbf{Y}| \quad \left(L_K := \sqrt{m} \sum_{i=1}^n L_K^i \right) \quad (10.12.59)$$

$$\iff |\mathbf{F}(\mathbf{X}) - \mathbf{F}| \leq L_K |\mathbf{X} - \mathbf{Y}| \quad (10.12.60)$$

が得られます. 最後の式は局所 Lipschitz 性を示す式です. ■

次の予備定理の証明に入る前にカットオフ関数 (cut off function) として任意の正の実数 r に対して $\chi_r: \mathbb{R}^n \rightarrow \mathbb{R}^n$ (カイと読む) というものを次のように定義します.

$$\chi_r(\mathbf{X}) = \begin{cases} \mathbf{X} & |\mathbf{X}| < r \\ \frac{\mathbf{X}}{|\mathbf{X}|} r & |\mathbf{X}| \geq r \end{cases} \quad (10.12.61)$$

したがって任意の $\mathbf{X} \in \mathbb{R}^n$ に対して $|\chi_r(\mathbf{X})| \leq r$ となります. これから示すのは, この χ が Lipschitz 定数 2 の Lipschitz 連続関数であるということです.

Theorem. 12.2.6.

$$|\chi_r(\mathbf{X}) - \chi_r(\mathbf{Y})| \leq 2|\mathbf{X} - \mathbf{Y}|. \quad (10.12.62)$$

(証明) 3 通りに場合を分ければ十分です. 解析らしく, 延々不等式と格闘します.

1. $|\mathbf{X}|, |\mathbf{Y}| \leq r$ のとき.

このときカットオフ関数は恒等写像となるので, 何もする必要はありません.

1. $|\mathbf{X}| \geq r, |\mathbf{Y}| \leq r$ のとき.

まず三角不等式と $|\mathbf{Y}| \leq r$ から

$$|\mathbf{X}| - r \leq |\mathbf{X}| - |\mathbf{Y}| \leq |\mathbf{X} - \mathbf{Y}| \quad (10.12.63)$$

となります。愚直に三角不等式を用いて

$$|\chi_r(\mathbf{X}) - \chi_r(\mathbf{Y})| = \left| \frac{\mathbf{X}}{|\mathbf{X}|}r - \mathbf{Y} \right| \quad (10.12.64)$$

$$\leq \left| \frac{\mathbf{X}}{|\mathbf{X}|}r - \frac{\mathbf{Y}}{|\mathbf{X}|}r \right| + \left| \frac{\mathbf{Y}}{|\mathbf{X}|}r - \mathbf{Y} \right| = \frac{r}{|\mathbf{X}|}|\mathbf{X} - \mathbf{Y}| + \frac{|\mathbf{Y}|}{|\mathbf{X}|}|\mathbf{X} - r| \quad (10.12.65)$$

$$\leq \left(\frac{r}{|\mathbf{X}|} + \frac{|\mathbf{Y}|}{|\mathbf{X}|} \right) |\mathbf{X} - \mathbf{Y}| \leq 2|\mathbf{X} - \mathbf{Y}|. \quad (10.12.66)$$

1. の証明: $|\mathbf{X}|, |\mathbf{Y}| \geq r$ のとき.

$$|\chi_r(\mathbf{X}) - \chi_r(\mathbf{Y})| = \left| \frac{\mathbf{X}}{|\mathbf{X}|}r - \frac{\mathbf{Y}}{|\mathbf{Y}|}r \right| \quad (10.12.67)$$

$$\leq \left| \frac{\mathbf{X}}{|\mathbf{X}|}r - \frac{\mathbf{Y}}{|\mathbf{X}|}r \right| + \left| \frac{\mathbf{Y}}{|\mathbf{X}|}r - \frac{\mathbf{Y}}{|\mathbf{Y}|}r \right| \quad (10.12.68)$$

$$= \frac{r}{|\mathbf{X}|}|\mathbf{X} - \mathbf{Y}| + r \left| \frac{|\mathbf{X}| - |\mathbf{Y}|}{|\mathbf{X}| \cdot |\mathbf{Y}|} \right| |\mathbf{Y}| \quad (10.12.69)$$

$$= \frac{r}{|\mathbf{X}|}|\mathbf{X} - \mathbf{Y}| + \frac{r}{|\mathbf{X}|}|\mathbf{X} - \mathbf{Y}| \quad (10.12.70)$$

$$\leq \frac{2r}{|\mathbf{X}|}|\mathbf{X} - \mathbf{Y}| \leq 2|\mathbf{X} - \mathbf{Y}|. \blacksquare \quad (10.12.71)$$

それでは目標の定理の証明に入ります。はじめに色々と定義が書いてありますが、あまり気にしないで構いません。また、定義域を狭めた所で成立する解のことを局所解といいます。

Theorem. 12.2.7. (局所解の存在と一意性) $I := [t_0, T)$, $B(\mathbf{R}_0; R) := \{\mathbf{R} \in \mathbb{R}^m; |\mathbf{R} - \mathbf{R}_0| < R\}$ ($0 < R \leq \infty$), $\Omega := B(\mathbf{R}_0; R) \times I = \Omega(\mathbf{R}_0, T, R)$ とする。 \mathbf{F} が Ω で局所 Lipschitz 条件を満たすならば、任意の $\mathbf{Y}_0 \in B(\mathbf{X}_0; R)$ に対して $T_0 = T_0(\mathbf{Y}_0) \in (t_0, T)$ が定まり、次の微分方程式を一意的に満たす解 \mathbb{R} が存在する。

$$\begin{cases} \frac{d}{dt}\mathbb{R}(t) = \mathbf{F}(\mathbf{R}(t), t), & t_0 < t \leq T_0 \\ \mathbb{R}(t_0) = \mathbf{R}_0. \end{cases} \quad (10.12.72)$$

(証明) 記号が煩雑になりますので、読むときはゆっくり読んで下さい。 $|\mathbf{R}_0| =: R_0 (< R)$ とします。まずは適当なコンパクト集合 (有界閉集合) を作りましょう。上の不等式から R_1, T_1 が存在して

$$R_0 < R_1 < R, \quad t_0 < T_1 < T \quad (10.12.73)$$

と書けます。これを用いて

$$K_1 := \{\mathbf{X} \in \mathbb{R}^m; |\mathbf{X} - \mathbf{R}_0| \leq R_1\} \times [t_0, T_1] \quad (10.12.74)$$

とすると、これは Ω のコンパクト集合 (有界閉集合) です。カットオフ関数 χ_{R_1} を用いて

$$\tilde{\mathbf{F}}(\mathbf{R}, t) := \mathbf{F}(\chi_{R_1}(\mathbf{R}), t) \quad (10.12.75)$$

として、次の微分方程式を考えます。

$$(\tilde{E}) \begin{cases} \frac{d}{dt} \mathbf{R}(t) = \tilde{\mathbf{F}}(\mathbf{R}, t) (= \mathbf{F}(\chi_{R_1}(\mathbf{R}), t)) & (t_0 < t < T_1) \\ \mathbf{R}(t_0) = \mathbf{R}_0 \end{cases} \quad (10.12.76)$$

ここで $\tilde{\mathbf{F}}$ が K_1 で Lipschitz 条件を満たすことを示します。カットオフ関数の定義から

$$|\chi_{R_1}(\mathbf{X})| \leq R_1, \quad |\chi_{R_1}(\mathbf{Y})| \leq R_1 \quad (10.12.77)$$

となります。これによって

$$(\chi_{R_1}(\mathbf{X}), t), (\chi_{R_1}(\mathbf{Y}), t) \in K_1 \quad (10.12.78)$$

が分かります。 \mathbf{F} の局所 Lipschitz 性から任意の $(\mathbf{X}, t), (\mathbf{Y}, t) \in K_1$ に対して

$$\left| \tilde{\mathbf{F}}(\mathbf{X}, t) - \tilde{\mathbf{F}}(\mathbf{Y}, t) \right| = \left| \mathbf{F}(\chi_{R_1}(\mathbf{X}), t) - \mathbf{F}(\chi_{R_1}(\mathbf{Y}), t) \right| \quad (10.12.79)$$

$$\leq L_{K_1} |\chi_{R_1}(\mathbf{X}) - \chi_{R_1}(\mathbf{Y})| \leq 2L_{K_1} |\mathbf{X} - \mathbf{Y}| \quad (10.12.80)$$

となります。最後のところで前の定理を使いました。微分方程式の解の一意的存在定理 12.2.2 と、 $\tilde{\mathbf{F}}$ の K_1 における Lipschitz 性から (\tilde{E}) の解 $\tilde{\mathbf{R}}$ が一意的に存在することが分かりました。

最後に、この $\tilde{\mathbf{R}}$ がある時刻のあいだで (定義域を小さく取った所で) 元の方程式を満たすことを示さなければなりません。まず $\tilde{\mathbf{R}}$ が $[t_0, T_1]$ で連続であること、はじめの不等式 (10.12.73) から $\tilde{\mathbf{R}}(t_0) = \mathbf{R}_0, |\mathbf{R}_0| = R_0 < R_1$ であることから適当な時刻 $t_1 \in (t_0, T_1]$ が存在して、任意の $t \in [t_0, t_1]$ に対して $|\tilde{\mathbf{R}}| \leq R_1$ と出来ることを証明しましょう。これが示されれば、時刻を $[t_0, t_1]$ を制限した $\tilde{\mathbf{R}}$ が元の方程式を満たすことが分かります。それでは証明。

$\tilde{\mathbf{R}}$ の連続性から $\varepsilon = R_1 - R_0$ とすると、これに対応して $\delta > 0$ が存在して任意の $t \in [t_0, t_0 + \delta]$ に対して $|\tilde{\mathbf{R}}(t) - \mathbf{R}_0| < \varepsilon$ となります。三角不等式から

$$\left| \tilde{\mathbf{R}}(t) \right| \leq \left| \tilde{\mathbf{R}}(t) - \mathbf{R}_0 \right| + \left| \mathbf{R}_0 \right| < \varepsilon + R_0 < R_1 \quad (10.12.81)$$

となり、 $t_1 := t_0 + \delta$ とすればめでたく証明完了です。

$t \in [t_0, t_1]$ としましょう。このとき $|\tilde{\mathbf{R}}| < R_1$ となるので、 $\chi_{R_1}(\tilde{\mathbf{R}}) = \tilde{\mathbf{R}}$ となります。したがって

$$\tilde{\mathbf{F}}(\tilde{\mathbf{R}}(t), t) = \mathbf{F}(\chi_{R_1}(\tilde{\mathbf{R}}(t)), t) = \mathbf{F}(\tilde{\mathbf{R}}(t), t) \quad (10.12.82)$$

が分かるので、 $\tilde{\mathbf{R}}$ が元の方程式

$$\begin{cases} \frac{d}{dt} \mathbf{R}(t) = \mathbf{F}(\mathbf{R}(t), t), & t_0 < t \leq T_0 \\ \mathbf{R}(t_0) = \mathbf{R}_0. \end{cases} \quad (10.12.83)$$

を満たす、つまり定義域を制限した所で成立する局所的な解であることが分かりました。■

解の延長

前小節で局所解の一意的な存在を証明できましたが、上の証明では定義域 (時刻) を狭めた所まででしか一意的な存在を保証できません。しかし場合によってはその定義域を拡張しなければ (もっと先の時刻まで解の存在を保証しなければ) ならないときがあるかもしれません。つまり解の延長ができるか否か、という問題にめくりをつけなければなりません。まず解の延長を正確に定義します。

Definition. 12.2.8. (解の延長) $R = R_1(t)$ は t_1 を右端とする閉区間 $[t_0, t_1]$ において, $R = R_2(t)$ は t_1 を左端とする左に閉じた区間 $[t_1, t_2)$ において $D_t R(t) = F(R(t), t)$ の解であって, かつ $R_1(t_1) = R_2(t_1)$ であるとする. このとき解 R_1 は t_1 を越えて右に向かって R_2 に延長できるという. 同様に t_1 を越えて左に延長できるということも定義できる.

解の最大存在時刻 T_m を次の式で定義します.

$$T_m := \sup\{T_0; \text{微分方程式が } [t_0, T_0] \text{ で解を持つ}\} \quad (10.12.84)$$

Theorem. 12.2.9. $I := [t_0, T)$, $B(R_0; R) := \{R \in \mathbb{R}^m; |R - R_0| < R\} (0 < R \leq \infty)$, $\Omega := B(R_0; R) \times I = \Omega(R_0, T, R)$ とする. 局所解の一意的存在定理 12.2.7 によって保証される解の延長可能性は以下の 2 通りである.

- 1) $T_m < T, \lim_{t \rightarrow T_m} |R(t) - R_0| = R$,
- 2) $T_m = T$.

(証明)

1. の証明:

背理法で証明します. この解の上の点列 $(R(t_n), t_n), t_0 \leq t_n \leq T_m$ が $n \rightarrow \infty$ のとき Ω の境界からの距離 η が正であるような Ω の内点 (R', T) に収束するものとします. n を十分大きく取れば, 点 $(R(t_n), t_n)$ と Ω との距離が $\eta/2$ であるようにできます. このとき t_n で $R(t_n)$ となるような $DR = F(R(t), t)$ の解が $|t - t_n| \leq h' := \min(\eta/2, \eta/2M)$ において一意的に存在します. この t の領域 $|t - t_n| \leq h'$ は $\lim_{n \rightarrow \infty} t_n = T_m$ によって, n が十分大きいときには点 T_m を内点とします. これは $R(t_0) = R_0$ がとなるもとの解 R が T_m を超えてなお右に延長できることを示し, T_m の定義に反します.

1. の証明:

先程の解 R は $|t - t_n| \leq h'$ において $DR(t) = F(R(t), t)$ の解として連続であり, しかも $|t - t_n| \leq h'$ が点 T_m を内部に含むので, 仮定 $\lim_{n \rightarrow \infty} (R(t_n), t_n) = (R', T_m)$ から

$$R(T_m) = R' \quad (10.12.85)$$

となり, 証明終了です. ■

解の初期値および解に対する連続依存性

これまでは決まった初期値に対する解の存在と一意性について調べてきました. 実験的な状況を考えれば, カオス的な振る舞いでもない限り, 初期値が少し異なる場合その解も少しだけ「異なる」ということが実現されていないと困ります. これをある範囲内で示すことができます. ここではもう少し一般化して示しますが, その証明に入る前に予備定理を 1 つ証明します.

Theorem. 12.2.10. (Gronwall) $I = [a, b]$, $x_0 \in I$ とする. $g, h: I \rightarrow \mathbb{R}$ とし, g については任意の $x \in I$ に対して $g(x) \geq 0$ となる連続関数とする. このとき a を定数として

$$h(x) \leq a + \int_{x_0}^x h(s) g(s) ds \quad (10.12.86)$$

ならば

$$h(x) \leq a \int_{x_0}^x e^{\int_{x_0}^x g(s) ds} ds. \quad (10.12.87)$$

(証明)

$$H(x) := a + \int_{x_0}^x h(s) g(s) ds \quad (10.12.88)$$

とすると条件の不等式を用いて

$$DH(x) = h(x)g(x) \leq H(x)g(x) \iff DH(x) - H(x)g(x) \leq 0 \quad (10.12.89)$$

となります. さらに

$$D \left(H(x) e^{-\int_{x_0}^x g(s) ds} \right) = e^{-\int_{x_0}^x g(s) ds} \{ DH(x) - g(x)H(x) \} \leq 0 \quad (10.12.90)$$

となることから,

$$H(x) e^{-\int_{x_0}^x g(s) ds} \quad (10.12.91)$$

は単調減少になります. したがって

$$H(x) e^{-\int_{x_0}^x g(s) ds} \leq H(x_0) e^{-\int_{x_0}^{x_0} g(s) ds} = H(x_0) = a \quad (10.12.92)$$

であり, これから示すべき不等式が導かれます. ■

Theorem. 12.2.11. (解の初期値および解に対する連続依存性) $F^i (i = 1, 2)$ を *Lipschitz* 条件を満たす関数として

$$(E)_i = \begin{cases} \frac{d}{dt} \mathbf{R}^i(t) = \mathbf{F}^i(\mathbf{R}^i(t), t) \\ \mathbf{R}^i(t_0) = \mathbf{R}_0^i \end{cases} \quad (i = 1, 2) \quad (10.12.93)$$

とする. このとき任意の $t \in I$ に対して

$$|\mathbf{R}^1(t) - \mathbf{R}^2(t)| \leq \left(|\mathbf{R}_0^1 - \mathbf{R}_0^2| + \frac{1}{L} \sup_{t \in I, \mathbf{R} \in \mathbb{R}^m} |\mathbf{F}^1(\mathbf{R}, t) - \mathbf{F}^2(\mathbf{R}, t)| \right) e^{L|t-t_0|} \quad (10.12.94)$$

(証明) 各微分方程式の差を取ると

$$D \{ \mathbf{R}^1(t) - \mathbf{R}^2(t) \} = \mathbf{F}^1(\mathbf{R}^1(t), t) - \mathbf{F}^2(\mathbf{R}^2(t), t) \quad (10.12.95)$$

となります. 辺々積分して整理すると,

$$\mathbf{R}^1(t) - \mathbf{R}^2(t) = \mathbf{R}_0^1 - \mathbf{R}_0^2 + \int_{t_0}^t \{ \mathbf{F}^1(\mathbf{R}^1(s), s) - \mathbf{F}^2(\mathbf{R}^2(s), s) \} ds \quad (10.12.96)$$

すぐ後で出て来るので先んじて β の定義をしておきます.

$$\beta := \sup_{t \in I, \mathbf{R} \in \mathbb{R}^m} |\mathbf{F}(\mathbf{R}, t) - \mathbf{F}(\mathbf{R}, t)|. \quad (10.12.97)$$

先程の積分した式の両辺の絶対値を取り, 三角不等式と Lipschitz 条件を用いた式

$$|F^1(R^1(s), s) - F^2(R^2(s), s)| \quad (10.12.98)$$

$$\leq |F^1(R^1(s), s) - F^1(R^2(s), s)| + |F^2(R^1(s), s) - F^2(R^2(s), s)| \quad (10.12.99)$$

$$\leq L|R^1(s) - R^2(s)| + \beta \quad (10.12.100)$$

を用いて, $h(x) := |R^1(x) - R^2(x)|$, $\alpha := |R_0^1 - R_0^2|$ とすれば

$$h(x) \leq \alpha + \int_{t_0}^x \{Lh(s) + \beta\} ds = \alpha + \int_{t_0}^x L \left\{ h(s) + \frac{\beta}{L} \right\} ds \quad (10.12.101)$$

となります. $H(x) := h(x) + \beta/L$ とすれば,

$$H(x) \leq \left(\alpha + \frac{\beta}{L} \right) + \int_{x_0}^x LH(s) ds \quad (10.12.102)$$

であり, Gronwall の不等式を用いると

$$H(x) \leq \left(\alpha + \frac{\beta}{L} \right) e^{L|x-t_0|} \quad (10.12.103)$$

となります. $h(x) \leq H(x)$ から

$$|R^1(t) - R^2(t)| \leq \left(\alpha + \frac{\beta}{L} \right) e^{L|t-t_0|}. \blacksquare \quad (10.12.104)$$

この定理において $F^1 = F^2$, $R_0^1 = R_0^2$ とすれば, 全ての $t \in I$ に対して

$$|R^1(t) - R^2(t)| \leq 0 \implies R^1(t) = R^2(t) \quad (10.12.105)$$

となります. つまり解の一意性の別証明が得られました.

10.12.3 常微分方程式の解法の為の一般論

具体的な微分方程式として 1 階と 2 階の常微分方程式をあとで扱います. 以下で主に使うものは単独の (連立でない) 方程式であり, 方程式系に対する扱いは少々煩雑にもなるので, 単独の方程式に対して議論を展開します. とりあえずこんなものか, と思って眺めておいて下さい. まずは定義から.

Definition. 12.3.1. n 階線型常微分方程式

$$\Lambda(D)x(t) := D^n x(t) + p_{n-1}(t)D^{n-1}x(t) + \cdots + p_1(t)Dx(t) + p_0(t)x(t) = q(t) \quad (10.12.106)$$

は $q = \theta$ (関数として 0, つまり恒等的に 0 を返す関数) のとき斉次方程式といい, $q \neq \theta$ のとき非斉次方程式という.

次に斉次方程式について重ね合わせの原理 (重畳原理) を証明します.

Theorem. 12.3.2. (重ね合わせの原理) 全ての係数関数 p_i を連続関数とする n 階の斉次の線型常微分方程式

$$\begin{aligned} \Lambda(D)x(t) &:= D^n x(t) + p_{n-1}(t)D^{n-1}x(t) + \cdots \\ &+ p_1(t)Dx(t) + p_0(t)x(t) = 0 \end{aligned} \quad (10.12.107)$$

の解 x_1, \dots, x_m の線型結合

$$x(t) = \sum_{i=1}^m c_i x_i(t) \quad (10.12.108)$$

はまた (10.12.107) の解である.

(証明) すぐに終わります. x_1, \dots, x_m がそれぞれ (10.12.107) の解なので, $\Lambda(D)x_i = 0$ となります. 微分作用素の線型性を用いて

$$\Lambda(D)x = \Lambda(D) \left(\sum_{i=1}^m c_i x_i \right) = \sum_{i=1}^m c_i \Lambda(D)x_i = \sum_{i=1}^m c_i 0 = 0. \blacksquare \quad (10.12.109)$$

次に解の線型独立性を定義します. これは微分方程式の解のなす集合を考える際に重要なものです.

Definition. 12.3.3. (10.12.108) の形に表せない解が存在するとき, 素の解は x_1, \dots, x_m に対して線型独立であるという. 一般に m 個の解の中のどの 1 つも残りの $(m-1)$ 個に対して線型独立であるときに, これら m 個の解は互いに線型独立であるという.

次に線型独立性の判定法を紹介します.

Theorem. 12.3.4. n 個の解 $x_1(t), \dots, x_n(t)$ が線型独立であるための必要十分条件は Wronski 行列式 (Wronkian)(ロンスキアン)

$$W(x_1(t), x_2(t), \dots, x_n(t)) := \begin{vmatrix} x_1(t) & x_2(t) & \dots & x_n(t) \\ Dx_1(t) & Dx_2(t) & \dots & Dx_n(t) \\ \dots & \dots & \dots & \dots \\ D^{n-1}x_1(t) & D^{n-1}x_2(t) & \dots & D^{n-1}x_n(t) \end{vmatrix} \quad (10.12.110)$$

が定義域 I 内で決して 0 にならないことである.

(証明) (必要性) まず $W(x_1(t), x_2(t), \dots, x_n(t)) \neq 0$ となる $t_1 \in I$ が少なくとも一つ存在することを示しましょう. 背理法を使います. もしそのような t_1 が存在しないとすると, 定理 (2.4.11) から $c_i (i = 1, \dots, n)$ に関する連立一次方程式

$$c_1 x_1(t) + c_2 x_2(t) + \dots + c_n x_n(t) = 0 \quad (10.12.111)$$

$$c_1 Dx_1(t) + c_2 Dx_2(t) + \dots + c_n Dx_n(t) = 0 \quad (10.12.112)$$

\vdots

$$c_1 D^{n-1}x_1(t) + c_2 D^{n-1}x_2(t) + \dots + c_n D^{n-1}x_n(t) = 0 \quad (10.12.113)$$

に対して $c_i \neq 0$ となる i が存在する解の組 $c := {}^t(c_1 \ c_2 \ \dots \ c_n)$ が存在します. $x(t) = c_1 x_1(t) + \dots + c_n x_n(t)$ を考えると, これは $D^i x(t) = 0 (i = 0, 1, \dots, n-1)$ を満たします. $D^{n-1}x(t) = 0$ を積分すると $x(t) = d_0 + d_1 t + \dots + d_{n-2} t^{n-2}$ となりますが, $d_i = 0 (i = 0, 1, \dots, n-2)$ となり, $x = \theta$; 恒等的に 0 を返す関数 θ となり, $x_i (i = 1, 2, \dots, n)$ の一次独立性に反します.

次に

$$DW(x_1(t), x_2(t), \dots, x_n(t)) \quad (10.12.114)$$

$$= \begin{vmatrix} x_1(t) & x_2(t) & \dots & x_n(t) \\ Dx_1(t) & Dx_2(t) & \dots & Dx_n(t) \\ \dots & \dots & \dots & \dots \\ D^n x_1(t) & D^n x_2(t) & \dots & D^n x_n(t) \end{vmatrix} \quad (10.12.115)$$

$$(10.12.116)$$

となります。置換による行列式の定義を採用したほうが楽なので、一般の場合の証明は見送りますが、そうかといって読者に任せる、というのも気が引けるので $n = 2$ で確認してみます。

$$DW(x_1, x_2) = D(x_1 Dx_2 - x_2 Dx_1) \quad (10.12.117)$$

$$= Dx_1 Dx_2 + x_1 D^2 x_2 - Dx_2 Dx_1 - x_2 D^2 x_1 \quad (10.12.118)$$

$$= x_1 D^2 x_2 - x_2 D^2 x_1 = \begin{vmatrix} x_1 & x_2 \\ D^2 x_1 & D^2 x_2 \end{vmatrix}. \quad (10.12.119)$$

話を元に戻しましょう。DW の第 n 行は各 x_i が (10.12.107) の解なので、

$$x_i(t) = - \sum_{k=1}^n p_k(t) D^k x_i(t) \quad (10.12.120)$$

と書けることを用いて、行列式の列に関する交代性から

$$DW(x_1(t), x_2(t), \dots, x_n(t)) = -p_{n-1}(t) W(x_1(t), x_2(t), \dots, x_n(t)) \quad (10.12.121)$$

となります。 $n = 2$ で確認してみましょう。

$$DW(x_1, x_2) \quad (10.12.122)$$

$$= x_1 D^2 x_2 - x_2 D^2 x_1 = -x_1(p_1 Dx_2 + p_0 x_2) + x_2(p_1 Dx_1 + p_0 x_1) \quad (10.12.123)$$

$$= -p_1(x_1 Dx_2 - x_2 p_1 Dx_1) = -p_1 W(x_1, x_2). \quad (10.12.124)$$

これは W に関する 1 階の線型の斉次微分方程式となります。 p_1 は連続であるとしているので、ある 1 点 t_0 で $W(t_0) = 0$ となるならば W は恒等的に 0 になります。しかし $t = t_1$ で $W(t_1) \neq 0$ なので、 W は定義域内で 0 となることはありません。

(十分性) 背理法で証明します。定義域 I 内で $W(x_1(t), \dots, x_n(t)) \neq 0$ かつ y_1, \dots, y_n が線型独立でないと仮定します。線型独立ではないので、少なくとも 1 つは零でない定数 c_i を用いて

$$c_1 x_1(t) + \dots + c_n x_n(t) \equiv 0 \quad (10.12.125)$$

となります。これを微分して

$$c_1 D^i x_1(t) + \dots + c_n D^i x_n(t) \equiv 0 (i = 0, 1, \dots, n-1) \quad (10.12.126)$$

となりますが、 $\hat{W}(t) := (D^{i-1} x_j(t))$ とすると $\hat{W}(t)c = 0$ と行列でまとめて書けます。Wronskian W が 0 でないので、 \hat{W} には逆が存在して、それをかければ $c = 0$ となりますが、これは不合理です。■

ここで n 階の斉次の線型常微分方程式において $n+1$ 個の解を持ってくるとこれらは線型独立ではないことが分かります. 見やすいので $n=1$ で書いてみます.

$$Dx(t) = p(t)x(t) \quad (10.12.127)$$

の 2 つの解を x_1, x_2 としましょう. 前定理によって Wronskian を計算すると

$$W(x_1, x_2) = \begin{vmatrix} x_1 & x_2 \\ Dx_1 & Dx_2 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 \\ px_1 & px_2 \end{vmatrix} = p \begin{vmatrix} x_1 & x_2 \\ x_1 & x_2 \end{vmatrix} = 0 \quad (10.12.128)$$

つまり (互いに線型独立な) 解を持つとしたら, それは n 個まで, ということが分かります. さらに n 個の解 $x_i (i=1, \dots, n)$ を $D^{i-1}x_i(t_1) = 1, D^jx_i(t_1) = 0 (j \neq i-1)$ となる初期条件で求めれば, Wronskian が 1 になるので, x_i は線型独立です. したがって n 階の斉次線型常微分方程式の解の集合は n 次の線型空間になります. そこで次の定義をします.

Definition. 12.3.5. (10.12.107) の n 個の線型独立な解 x_1, x_2, \dots, x_n を (10.12.107) の基本解系という.

Theorem. 12.3.6. (10.12.107) の任意の解 x は (10.12.107) の任意の基本解 x_1, x_2, \dots, x_n の線型結合として一意に表される.

(証明) $W(x_1, \dots, x_n) \neq 0$ から定数 c_i を

$$D^k x(t_1) = \sum_{i=1}^n c_i D^k x_i(t_1) \quad (k=0, 1, \dots, n-1) \quad (10.12.129)$$

となるように定めることが出来ます. x_1, x_2, \dots, x_n の線型独立性から表現 $x(t) = \sum c_i x_i(t)$ の一意性が分かります. ■

最後に非斉次方程式について議論します. まず斉次方程式

$$\begin{aligned} \Lambda(D)x(t) &:= D^n x(t) + p_{n-1}(t)D^{n-1}x(t) + \dots \\ &\quad + p_1(t)Dx(t) + p_0(t)x(t) = 0 \end{aligned} \quad (10.12.130)$$

の解を斉次解, 非斉次方程式

$$D^n x(t) + p_{n-1}(t)D^{n-1}x(t) + \dots + p_1(t)Dx(t) + p_0(t)x(t) = q(t) \quad (10.12.131)$$

の解を特解 (特殊解) と呼びます. このもとで

Theorem. 12.3.7. (10.12.131) の任意の解はその任意の 1 つの特解と斉次方程式 (10.12.130) の一般解の和として表される.

(証明) (10.12.131) の 2 つの解を x_1, x_2 としましょう. ここで $z := x_1 - x_2$ とすると

$$\Lambda z = \Lambda x_1 - \Lambda x_2 = q - q = 0 \quad (10.12.132)$$

となります. これは z が斉次方程式 (10.12.130) の解であることを表します. $x_1 = z + x_2$ と書けますが, これが証明すべきことでした. ■

10.13 1 階線型方程式の解法

物理的に重要な微分方程式の結果をいくつか紹介しました。ここからは良く出てくる具体的な微分方程式の解法を紹介します。解法を紹介する前に 1 つ確認しておきます。解の存在と一意性をきちんと証明した以上、どんな手を使ってでも解を 1 つ見つけられればそれで終わりです。どれほど納得のいかない変な導出法であったとしても、です。逆にいうならば、変な導出法だが解が 1 つしかないのだから見つけた者勝ちだと強弁をはるためにこれまで色々な飛び道具を用意してきた、ともいえます。

さて、1 階の方程式で重要なのは電気回路についての方程式

$$\Lambda(D)j(t) := L \frac{dj(t)}{dt} + Rj(t) = v(t) \quad (10.13.1)$$

でしょう。ここで L はコイルのインダクタンス、 j は電流、 R は抵抗、 v は電源電圧です。

斉次解、特解をそれぞれ j_1, j_2 とします。つまり

$$\Lambda(D)j_1(t) = 0 \quad (10.13.2)$$

$$\Lambda(D)j_2(t) = v(t). \quad (10.13.3)$$

j_1 から求めます。今解の一意性存在定理での F としたところが $-kj_1$ となっています。これは高々一次の増大度の C^1 級関数なので Lipschitz 条件を満たし、解が一意的に存在します。これは $Dj_1 = -kj_1, k = R/L$ となる関数を求めれば良いのですが、 e^{-kt} はこれを満たします。したがって

$$j_1(t) = ce^{-\frac{R}{L}t}. \quad (10.13.4)$$

c_1 は初期条件から決める定数です。

次に特解を求めます。これには定数変化法と呼ばれる手法を用います。これは先程の ce^{-kt} で、 $c = c(t)$ としてみよう、というものです。これを非斉次方程式に代入してみます。

$$v(t) = \Lambda(D)(c(t)e^{-kt}) = LD(c(t)e^{-kt}) + Rc(t)e^{-kt} \quad (10.13.5)$$

$$= L(Dc(t) - kc(t))e^{-kt} + Rc(t)e^{-kt} = Le^{-kt}Dc(t) \quad (10.13.6)$$

となり、これから

$$c(t) = \frac{1}{L} \int v(s)e^{ks} ds \quad (10.13.7)$$

と求められます。一般解は

$$j(t) = ce^{\frac{R}{L}t} + \frac{1}{L} \int v(s)e^{\frac{R}{L}(t-s)} ds \quad (10.13.8)$$

となります。

10.14 2 階線型方程式の解法

2 階の線型方程式は振動の方程式です。摩擦などの抵抗が速度に比例するものとしましょう。比例定数を Γ 、 F を外力として運動方程式は

$$mD^2x(t) + \Gamma Dx(t) + kx(t) = F(t) \quad (10.14.1)$$

となります。ここで $f := F/m$, $2\gamma := \Gamma/m$, $k/m := \omega^2$, $\omega_0 = 0$ とすると

$$D^2x(t) + 2\gamma x(t) + \omega^2 x(t) = f(t) \quad (10.14.2)$$

となります。実際に解く際にはこれで考えると楽なことはおいおい分かるでしょう。各種考察の前に抵抗, コンデンサ, コイルが全て入った回路の方程式を書いておきます。高校では $i = \Delta q / \Delta t$ と学んだと思いますが, この $\Delta \rightarrow 0$ とした極限で $i = dq/dt = Dq$ となります。

$$LDj(t) + Rj(t) + \frac{1}{C} \int j(s)ds = v(t) \quad (10.14.3)$$

これは一般に微分積分方程式と呼ばれるものですが, v を例えば C^1 級であるとしておいて両辺を微分すると 2 階の常微分方程式になります。

それでは実際に (10.14.2) を解いていきましょう。まず解の存在と一意性を確かめます。 $y_0(t) := x(t)$, $y_1(t) := Dx(t)$, $\mathbf{Y}(t) := {}^t(y_0(t) \ y_1(t))$ としましょう。このとき

$$D\mathbf{Y}(t) = \begin{pmatrix} Dy_0(t) \\ Dy_1(t) \end{pmatrix} = \begin{pmatrix} y_1(t) \\ -2\gamma y_1(t) - \omega^2 y_0(t) \end{pmatrix} =: \mathbf{F}(\mathbf{Y}(t), t) \quad (10.14.4)$$

となります。示すべきは \mathbf{F} の Lipschitz 性です。2 行目から 3 行目への不等式は $(a+b)^2 \leq 2(a^2+b^2)$ を使っています。

$$|\mathbf{F}(\mathbf{Y}, t) - \mathbf{F}(\mathbf{Z}, t)| \quad (10.14.5)$$

$$= \sqrt{(y_0 - z_0)^2 + \{-2\gamma(y_1 - z_1) - \omega^2(y_0 - z_1)\}^2} \quad (10.14.6)$$

$$\leq \sqrt{(y_0 - z_0)^2 + 2\left[\{-2\gamma(y_1 - z_1)\}^2 + \{-\omega^2(y_0 - z_1)\}^2\right]} \quad (10.14.7)$$

$$= \sqrt{(1 + 2\omega^4)(y_0 - z_0)^2 + 8\gamma^2(y_1 - z_1)^2} \quad (10.14.8)$$

$$\leq \max\left(\sqrt{1 + 2\omega^4}, \sqrt{8\gamma^2}\right) \sqrt{(y_0 - z_0)^2 + (y_1 - z_1)^2} \quad (10.14.9)$$

$$= \max(1 + 2\omega^4, 8\gamma^2) |\mathbf{Y} - \mathbf{Z}|. \blacksquare \quad (10.14.10)$$

次に斉次解を求めます。 $\gamma = 0$, つまり摩擦がない時の解は高校で学んだものと同じです。以下に示す方法を使えば各自で容易に導出できるはずなので, ここでは省略して, $\gamma \neq 0$ の場合を扱っていきます。かなりの荒技ですが, ここで微分作用素 $\Lambda(D)$ を因数分解します。まず x が $\Lambda(D)$ の固有関数 (ベクトル) であるとします。このとき D はただ数をかけるだけの掛け算作用素に変わるので, 固有関数を考える限りは普通の数と同じように扱ってよいでしょう。これは (10.14.2) の係数が実数であって, 関数ではないからできることです。 $d := \gamma^2 - \omega^2$ として $D^2 + 2\gamma D + \omega^2$ を因数分解すると,

$$\left\{D - (-\gamma + \sqrt{d})\right\} \left\{D - (-\gamma - \sqrt{d})\right\} \quad (10.14.11)$$

となります。したがって (10.14.2) は

$$0 = \left\{D - (-\gamma + \sqrt{d})\right\} \left\{D - (-\gamma - \sqrt{d})\right\} x(t) \quad (10.14.12)$$

$$= \left\{D - (-\gamma - \sqrt{d})\right\} \left\{D - (-\gamma + \sqrt{d})\right\} x(t) \quad (10.14.13)$$

となります。 $\{D - (-\gamma - \sqrt{d})\}x(t) = 0, \{D - (-\gamma + \sqrt{d})\}x(t)$ の解をそれぞれ x_1, x_2 とすると、これらは自動的に (10.14.2) を満たします。前節での計算結果から c_i を (初期条件により定まる) 定数として

$$x_1(t) = c_1 e^{(-\gamma - \sqrt{d})t}, \quad x_2(t) = c_2 e^{(-\gamma + \sqrt{d})t} \quad (10.14.14)$$

となります。またこの Wronskian を計算すると

$$W(x_1, x_2) = \begin{vmatrix} x_1 & x_2 \\ Dx_1 & Dx_2 \end{vmatrix} = c_1 c_2 e^{-\gamma t} \{(-\gamma + \sqrt{d}) - (-\gamma - \sqrt{d})\} \quad (10.14.15)$$

$$= 2c_1 c_2 \sqrt{d} e^{-\gamma t} \neq 0 (d \neq 0) \quad (10.14.16)$$

となり、 $d \neq 0$ のとき x_1, x_2 は線型独立です。 $d = 0$ のときはどうするかということが問題として浮上してきました。このとき (10.14.2) は

$$(D + \gamma)^2 x(t) = 0 \quad (10.14.17)$$

となります。つまり $x_1(t) := c_1 e^{-\gamma t}$ とすれば、これは上の方程式の解になっています。もう 1 つ線型独立な解を見つける必要があります。Laplace 変換を使うと何も考えずにすんなり出てきますが、今はひとまず天下りに与えます。 $x_2(t) := c_2 t e^{-\gamma t}$ が求める答です。実際

$$(D + \gamma)^2 x_2 = c_2 (D^2 + 2\gamma D + \gamma^2) t e^{-\gamma t} \quad (10.14.18)$$

$$= c_2 e^{-\gamma t} \{(\gamma^2 - 2\gamma) + 2\gamma(1 - \gamma) + \gamma^2\} = 0 \quad (10.14.19)$$

となります。Wronskian も

$$W(x_1, x_2) = \begin{vmatrix} x_1 & x_2 \\ Dx_1 & Dx_2 \end{vmatrix} = \begin{vmatrix} c_1 e^{-\gamma t} & c_2 t e^{-\gamma t} \\ -c_1 \gamma e^{-\gamma t} & c_2 (1 - \gamma t) e^{-\gamma t} \end{vmatrix} \quad (10.14.20)$$

$$= c_1 c_2 e^{-\gamma t} (2 - \gamma t) \neq 0 (t \neq 2/\gamma) \quad (10.14.21)$$

となります。したがっていずれにせよ齊次方程式の任意の解は x_1, x_2 の線型結合で書くことができることが分かります。ここでずっと気になっていた読者もいることと思いますが、 d はいわゆる判別式であって、 d の正負が次の問題です。

数学は一旦忘れて物理を考えてみましょう。いま (10.14.2) がばねの振動を記述しているものとします。ある程度ばねを伸ばして (縮ませて) から手を放したとしましょう。摩擦力が十分小さいとき、はじめは大きく振動していても時間が経つごとに振幅が小さくなっていくはずですが、つまり減衰振動になるはずですが、逆に摩擦力が非常に大きいとき、ほとんど全く振動せずに振動が止まってしまうでしょう。

この状況が $d = \gamma^2 - \omega^2$ の正負と直接に関係します。 $d < 0$ のときを考えます。つまり $\gamma < \omega$ であり、上でいうところの「摩擦係数が小さいとき」です。このとき $\sqrt{d} = i\sqrt{-d} = i\sqrt{\omega^2 - \gamma^2}$ であり、

$$x_1(t) = c_1 e^{-\gamma t} e^{-i\sqrt{\omega^2 - \gamma^2}t}, \quad x_2(t) = c_2 e^{-\gamma t} e^{+i\sqrt{\omega^2 - \gamma^2}t} \quad (10.14.22)$$

です。時間と共に振幅 $c e^{-\gamma t}$ は減衰していき、Euler の公式 $e^{it} = \cos t + i \sin t$ からきちんと振動項も入っていて、減衰振動を記述していることが分かります。

次に $d > 0$ を考えましょう。このときは何もせずに

$$x_1(t) = c_1 e^{(-\gamma - \sqrt{d})t}, \quad x_2(t) = c_2 e^{(-\gamma + \sqrt{d})t} \quad (10.14.23)$$

として良いことになります。ここで場合によっては x_2 で $-\gamma + \sqrt{d} > 0$ となることがあり、これは時間と共に振幅が無限に大きくなっていく、という非物理的 (非現実的) な解になってしまいます。しかしこれはもともと Hooke の法則が振動の小さいところでしか成り立たない近似的な法則であり、その近似の限界を超えてしまったところから法則の成立を仮定したために起こったことです。実際には非線型項 x^2, x^3, \dots などが出てきて、振幅の発散は起こりません。 x_1 (と d が十分小さいときの x_2) は過減衰と呼ばれます。

最後に $d = 0$ ですが、これは振動と過減衰を分ける臨界 (現象) であり、精査すると色々面白いようですが、踏み込むのはやめます。

非斉次方程式ですが、この解について一般論を述べるのはひどく大変です。ここでは回路理論で最も自然だと思われる $f(t) = f \sin \omega_0 t$ について述べます。これについても解の一意性存在と物理的な直観を頼りに議論しましょう。今考えているのは外力による振動を加えた時のばねの伸び縮みです。普通に考えて、ばねの振動と外力による振動は同じような振動になるでしょう。つまり、解を $x_0(t) = A \sin \omega_0 t$ と仮定して良いでしょう。しかしこれだと $Dx(t) = A\omega_0 \cos \omega_0 t$ となり、計算がやや面倒になります。そこで複素数の世界に潜って、 $f(t) = f e^{i\omega_0 t}$, $x_0(t) = A e^{i\omega_0 t}$ とします。すると微分の計算が簡単になる他、 $\operatorname{Re} x_0(t)$, $\operatorname{Im} x_0(t)$ とすることで、それぞれ外力が \cos, \sin で与えられる場合に一度に対応できます。実際に計算してみましょう。

$$f e^{i\omega_0 t} = (D^2 + 2\gamma D + \omega^2) A e^{i\omega_0 t} = A(-\omega_0^2 + 2i\gamma\omega_0 + \omega^2) e^{i\omega_0 t} \quad (10.14.24)$$

$$\implies A = \frac{f}{-\omega_0^2 + 2i\gamma\omega_0 + \omega^2} \quad (10.14.25)$$

したがって非斉次解は

$$x_0(t) = \frac{f}{-\omega_0^2 + 2i\gamma\omega_0 + \omega^2} e^{i\omega_0 t} \quad (10.14.26)$$

となります。

Chapter 11

Maxwell 方程式とベクトル解析

- 11.1 導入
- 11.2 微分作用素ナブラ
- 11.3 諸公式の導出とテンソル
- 11.4 Green の定理
- 11.5 Stokes の定理
- 11.6 Gauss の定理
- 11.7 微分形式とベクトル解析

Part VI

Complex Analysis

Chapter 12

Talk: Complex Analysis in One variable

See, e.g., [207] for consulation. In the following we assume $U \subset \mathbb{C}$ is a domain without further mention.

12.1 Definition: complex differentiability

Definition. 1.0.8. Let $f: U \rightarrow \mathbb{C}$ be a function. A function f is called **(complex-) differentiable** at $a \in U$ if there exists a constant $A \in \mathbb{C}$ such that

$$f(z) = f(a) + A(z - a) + o(z - a). \quad (12.1.1)$$

A constant A is written by $f'(a)$. A function f is **holomorphic** if f is differentiable at all points in U .

Note that we do not assume continuity of a derivative for holomorphy.

Remark. 1.0.9. (What is different from differentiability for functions in two real variables?) Consider the following function $f(x, y)$ with two real variables,

$$f(x, y) = \begin{cases} 0, & x = y = 0, \\ \frac{xy}{x^2 + y^2}, & \text{otherwise.} \end{cases} \quad (12.1.2)$$

This function f is C^∞ for x or y but f is not **continuous** in two variables. A deep result, **Hartogs' theorem**, exists for the theory in several complex variables. See, e.g., [158].

12.2 Cauchy-Riemann equation

Theorem. 2.0.10. (The Cauchy-Riemann equation.) Let $f: U \rightarrow \mathbb{C}$ be holomorphic and denote $f = u + iv$, $u = u(x, y)$ and $v = v(x, y)$. Then we have $u_x = v_y$ and $u_y = -v_x$.

Proof. Calculate as below:

$$f'(0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x} = u_x + iv_x \quad (12.2.1)$$

$$= \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{iy} = \frac{1}{i}(u_y + iv_y). \quad (12.2.2)$$

□

12.3 Complex line integral

First we review a line integral in \mathbb{R}^2 .

Definition. 3.0.11. (Line integral in \mathbb{R}^2) Let $C: I = [a, b] \rightarrow \mathbb{R}^2$ be a piecewise C^1 curve and $\text{Im } C \subset U$ and $F: U \rightarrow \mathbb{R}^2$ be a continuous vector field. Then we define a **line integral** over C of F as

$$\int_C F \cdot dC = \int_{\text{Im } C} F(C) \cdot dC = \int_a^b F(C(t)) \cdot C'(t) dt. \quad (12.3.1)$$

Definition. 3.0.12. (Complex line integral) Assume $f: U \rightarrow \mathbb{C}$ is continuous and $C \subset U$ is a piecewise C^1 curve. Then we define a **complex line integral** as

$$\int_C f(z) dz := \int_C (u dx - v dy) + i \int_C (u dy + v dx), \quad (12.3.2)$$

where we denote $f = u + iv$. Furthermore we define another type of integral.

$$\int_C f(z) |dz| := \int_a^b f(z(t)) |z'(t)| dt. \quad (12.3.3)$$

We define two more concepts for assertion of the Cauchy's integral theorem.

Definition. 3.0.13. (Homolog 0 curve.) Assume $C \subset U$ is a piecewise C^1 closed curve, i.e., $C: [0, 1] \rightarrow U$ satisfied $C(0) = C(1)$, and $a \in U$. Then a number $n(C, a)$, called a **winding number**, is defined by

$$n(C, a) := \frac{1}{2\pi i} \int_C \frac{1}{z - a} dz. \quad (12.3.4)$$

Furthermore a curve C is called **homolog 0** for U if $n(C, a) = 0$ for any $a \notin U$. This is denoted by $C \sim 0$ in U .

Definition. 3.0.14. (Simply connectedness.) Assume $U \subset \mathbb{C}$. A domain U is called **simply connected** if each piecewise C^1 closed curve C is homolog 0 in U .

Here we can state two forms of the Cauchy integral theorem.

Theorem. 3.0.15. (Cauchy's integral theorem 1: for any U but confined C .) Assume a function f on U is holomorphic and a curve $C \subset U$ is a homolog 0 curve. Then we have $\int_C f(z) dz = 0$.

Theorem. 3.0.16. (Cauchy's integral theorem 2: for confined U but any C .) Assume a function f on U is holomorphic, a region U is simply connected, and a curve $C \subset U$ is piecewise C^1 . Then we have $\int_C f(z)dz = 0$.

Remark. 3.0.17. The assumption of piecewise smoothness for curves in the Cauchy integral theorems is needed for application, e.g., the residue theorem.

Proof. We prove the theorem in a easy form: we use the famous Green theorem in the vector calculus [207]. Assume a function f is in C^1 and we denote a region surrounded by C as B . Then we have, by Green' theorem,

$$\int_C f = \int_C (udx - vdy) + i \int_C (udy + vdx) \quad (12.3.5)$$

$$= - \int_B (u_y + v_x) dxdy + i \int_B (u_x - v_y) dxdy \quad (12.3.6)$$

$$= 0, \quad (12.3.7)$$

where we use the Cauchy-Riemann equation in the last equal sign. \square

This theorem is wonderful in, at least, two points.

1. The value of f is determined by line integrals. We can take curves arbitrarily, so holomorphic functions cannot determine their values freely. See also Cauchy's integral representation formula.
2. It defines an another function by line integrals, $F_{z_0} := \int_C f$, where C is a curve which connects two points z_0 and z in U . See Morela's theorem.

12.3.1 Appendix: Green's theorem

We introduce a simple version. Consider C^1 functions φ and ψ defined on a compact interval $I = [a, b]$ satisfying

$$\varphi(x) \leq \psi(x), \quad x \in I. \quad (12.3.8)$$

The region U is defined by

$$U = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \quad \varphi(x) \leq y \leq \psi(x)\}. \quad (12.3.9)$$

By similar functions f, g the region U is also described by

$$U = \{(x, y) \in \mathbb{R}^2 : \alpha \leq y \leq \beta, \quad f(y) \leq x \leq g(y)\}. \quad (12.3.10)$$

Theorem. 3.1.1. (Green's theorem.) Assume a C^1 vector field $F = (u, v)$ is defined in U . Then it holds that

$$\int_U (v_x - u_y) dxdy = \int_{\partial U} udx + vdy. \quad (12.3.11)$$

Proof. The fundamental theorem in calculus leads

$$- \int_U u_y dxdy = - \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} u_y dy \right) dx \quad (12.3.12)$$

$$= \int_a^b \{u(x, \varphi(x)) - u(x, \psi(x))\} dx = \int_{C_1+C_2} udx = \int_{\partial U} udx, \quad (12.3.13)$$

where the curves C_1 and C_2 are assumed to be properly defined. Similar computation for v shows

$$\int_U v_x dx dy = \int_\alpha^\beta \left(\int_{f(y)}^{g(y)} v_x dx \right) dy \quad (12.3.14)$$

$$= \int_\alpha^\beta \{v(g(y), y) - v(f(y), y)\} dy = \int_{\partial U} v dy. \quad (12.3.15)$$

Summing up the above two expressions leads the result. \square

12.4 Morela's theorem

Morela's theorem is a converse of Cauchy's integral theorem: i.e., differentiability and integrability coincide for holomorphic functions. For proof we need some preparations.

Definition. 4.0.2. A function $F: U \rightarrow \mathbb{C}$ is a primitive function of f in U is $F' = f$.

Theorem. 4.0.3. Assume a function $f: U \rightarrow \mathbb{C}$ is continuous and $\int_C f(z) dz = 0$ for any piecewise C^1 closed curves. Then there exists a function F on U such that $F' = f$.

Proof. Fix a point $z_0 \in U$ and let $z \in U$ be arbitrary. Since U is connected we have two piecewise C^1 curve C_j which do not intersect. A curve $C := C_1 + (-C_2)$ is a closed piecewise C^1 curve. Then, by assumption,

$$0 = \int_C f(z) dz = \int_{C_1} f(z) dz - \int_{C_2} f(z) dz. \quad (12.4.1)$$

Hence we can define a function F as

$$F(z) := \int_{z_0}^z f(z) dz := \int_{C_1} f(z) dz = \int_{C_2} f(z) dz. \quad (12.4.2)$$

Next we would like to show $F' = f$. Since f is continuous, for any $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(z+h) - f(z)| < \epsilon$ if $|h| < \delta$. Then we obtain

$$\left| \frac{F(z+h) - F(z)}{h} - f(z) \right| = \left| \frac{1}{h} \int_z^{z+h} (f(w) - f(z)) dw \right| \quad (12.4.3)$$

$$\leq \frac{1}{|h|} \epsilon |h| = \epsilon. \quad (12.4.4)$$

\square

Then we obtain

Theorem. 4.0.4. (Morela's theorem.) Assume a function $f: U \rightarrow \mathbb{C}$ is continuous and $\int_C f = 0$ for any piecewise C^1 closed curve $C \subset U$. Then f is holomorphic in U .

Proof. By the above theorem we have a function F such that $F' = f$. The following Cauchy's integral representation formula and Taylor series expansion theorem show that F and f are holomorphic. \square

12.5 Cauchy's integral representation formula and Taylor series expansion

Theorem. 5.0.5. *Let $f: U \rightarrow \mathbb{C}$ be holomorphic and $C \subset U$ be a homolog 0 curve. Then we have a following representation:*

$$n(C, a)f(a) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z-a} dz. \quad (12.5.1)$$

Proof. Let $a \in U \setminus C$ and put

$$\varphi(z) = \begin{cases} \frac{f(z)-f(a)}{z-a}, & z \neq a, \\ f'(a), & z = a. \end{cases} \quad (12.5.2)$$

Since φ is holomorphic in U^1 we get, by Cauchy's integral theorem,

$$0 = \frac{1}{2\pi i} \int_C \varphi(z) dz = \frac{1}{2\pi i} \int_C \frac{f(z)}{z-a} dz - \frac{f(a)}{2\pi i} \int_C \frac{1}{z-a} dz. \quad (12.5.3)$$

□

Theorem. 5.0.6. (Taylor series expansion.) *Assume a function $f: U \rightarrow \mathbb{C}$ is holomorphic. Then a function f is complex analytic and has a following expansion:*

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n, \quad c_n := \frac{1}{2\pi i} \int_C \frac{f(z)}{(z-a)^{n+1}} dz. \quad (12.5.4)$$

Proof. For sufficiently small $|z-a|$, we have the following expansion:

$$\frac{1}{w-z} = \frac{1}{w-a-(z-a)} = \frac{1}{w-a} \frac{1}{1-\frac{z-a}{w-a}} \quad (12.5.5)$$

$$= \frac{1}{w-a} \sum_{n=0}^{\infty} \left(\frac{z-a}{w-a} \right)^n. \quad (12.5.6)$$

Due to Cauchy's integral representation formula and exchange of limits we have

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n, \quad c_n := \frac{1}{2\pi i} \int_C \frac{f(z)}{(z-a)^{n+1}} dz. \quad (12.5.7)$$

□

A function $f: \mathbb{C} \rightarrow \mathbb{C}$ is called **entire** if it is holomorphic in all \mathbb{C} .

Theorem. 5.0.7. (Liouville's theorem) *If an entire function f is bounded f is constant.*

¹There are some gaps. See the proof of Theorem 2.4, Chapter 9 in [207].

12.6 Simple application of fundamental theorems: convergence of holomorphic functions

Definition. 6.0.8. (Uniform convergence in the wider sense.) Let $f_n, f: U \rightarrow \mathbb{C}$ be functions. A sequence of functions (f_n) converges to f uniformly in the wider sense if, for any compact subset $K \subset U$ such that $\lim_{n \rightarrow \infty} \|f_n - f\|_K = 0$, where

$$\|f\| := \sup_{z \in K} |f(z)|. \quad (12.6.1)$$

We denote this convergence as $f_n \xrightarrow{uw} f$.

Theorem. 6.0.9. (Weierstrass. Theorem 3.5 in [207].) Assume functions $f_n: U \rightarrow \mathbb{C}$ are holomorphic, $f: U \rightarrow \mathbb{C}$ is defined by a uniform limit in the wider sense, i.e., $f_n \xrightarrow{uw} f$.

1. For any piecewise C^1 curves we have the following equality,

$$\int_C f(z) dz = \lim_{n \rightarrow \infty} \int_C f_n(z) dz. \quad (12.6.2)$$

2. The limit function f is holomorphic.

3. We have

$$f'(z) = \lim_{n \rightarrow \infty} f'_n(z). \quad (12.6.3)$$

Proof. 1)-2): Due to Cauchy's integral theorem, for any homolog 0 curve C , we obtain $\int_C f_n = 0$. Since C is compact and a sequence (f_n) converges to f , a sequence $(\|f - f_n\|_C)_{n \in \mathbb{N}}$ converges to 0 and hence is bounded. Then it holds that

$$\left| \int_C f \right| = \left| \int_C f - \int_C f_n \right| \leq \int_C |f - f_n| |dz| \quad (12.6.4)$$

$$= \|f - f_n\|_C \int_C |dz| \rightarrow 0. \quad (12.6.5)$$

3): By Cauchy' integral representation theorem it holds that, for $a \in U \setminus C$,

$$f'(a) = \frac{-1}{2\pi i} \int_C \frac{f(z)}{(z-a)^2} dz, \quad (12.6.6)$$

$$f'_n(a) = \frac{-1}{2\pi i} \int_C \frac{f_n(z)}{(z-a)^2} dz. \quad (12.6.7)$$

The remaining argument is the same as above. \square

Remark. 6.0.10. (Points for the proof.) Assume functions $f_n: K \rightarrow \mathbb{C}$ are continuous on a compact set $K \subset \mathbb{C}$ and that a sequence (f_n) converges to a function f on K .

1. A new sequence $(\|f_n - f\|)_n$ converges to 0 and hence is bounded.

2. It holds that $\int_K f_n \rightarrow \int_K f$.

12.7 Zero points for holomorphic functions and theorem of identity

Theorem. 7.0.11. (Zero points for holomorphic functions are isolated.) *Let a function $f: U \rightarrow \mathbb{C}$ be holomorphic. Then each zero point of f is isolated.*

Proof. Suppose $z_0 \in U$ is a zero point of f , i.e., $f(z_0) = 0$, and its degree is k -th. Let a point z be sufficiently near z_0 . Then, by Cauchy's integral representation formula, we get

$$f(z) = \sum_{n=k}^{\infty} c_n(z - z_0)^n = (z - z_0)^k g(z), \quad g(z) := \sum_{n=k}^{\infty} c_n(z - z_0)^{n-k}. \quad (12.7.1)$$

Since a function g is continuous there exists a neighborhood $B \subset U$ of z_0 such that $g(w) \neq 0$ for all $w \in B$. Hence $f(z) \neq 0$ for all $z \in B \setminus z_0$. \square

Theorem. 7.0.12. (Theorem of identity) *Let functions $f, g: U \rightarrow \mathbb{C}$ be holomorphic and $C \subset U$ have a limit point in U . If $f = g$ on C then $f = g$ in U .*

Proof. Let $h = f - g$. Then $h = 0$ on C . Since zero points are isolated for holomorphic functions in one variable it holds that $h = 0$. \square

Theorem. 7.0.13. (Theorem of identity in several complex variables.) *Let $U \subset \mathbb{C}^n$ be an open set. Assume $f: U \rightarrow \mathbb{C}$ is holomorphic. If $f|_V = 0$ for an open subset $V \subset U$ then $f = 0$.*

Example. 7.0.14. (Counterexamples for Theorem of identity in several complex variables.)

1. (Using the proof of isolated zero points.) *Consider a holomorphic functions f in one variables and set $F(z_1, z_2) = z_1 f(z_2)$. Then $F = 0$ on a complex line $z_1 = 0$ but this is not a zero function.*
2. (Using algebraic geometry.) *Assume a polynomial f is not zero and nonconstant. Consider an algebraic variety $V(f) = \{f = 0\}$. Then f is 0 on a closed set $V(f)$ and $f \neq 0$.*

12.8 Theorem of residues

12.8.1 Laurent expansion

Let $a \in U$ and $R > 0$. We call the following $B_0(a, R)$ an **exclusive neighborhood** of a :

$$B_0(a, R) := \{z \in \mathbb{C} : 0 < |z - a| < R\}. \quad (12.8.1)$$

Definition. 8.1.1. (Singular points.) *Let $a \in U$ and $R > 0$. If a function $f: B_0(a, R) \rightarrow \mathbb{C}$ is holomorphic and not holomorphic in $B(a, R)$, then a point a is called an **isolated singular point**. I.e., a function f is not defined at a point a , or f is defined and not differentiable at a .*

Theorem. 8.1.2. (Laurent expansion.) *If a function f is holomorphic in $B_0(a, R)$ for a point $a \in U$, then f is represented as the following **Laurent series** in $B = B_0(a, R)$.*

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z-a)^n, \quad (12.8.2)$$

where coefficients c_n is defined by

$$c_n := \frac{1}{2\pi i} \int_C \frac{f(z)}{(z-a)^{n+1}} dz \quad (\forall n \in \mathbb{Z}). \quad (12.8.3)$$

Here the above C is a circle with a center $a \in U$ and a radius $r \in (0, R)$.

Proof. Let $0 < R_2 < R_1 < R$, define an annulus $A(a; R_1, R_2) := \{z \in \mathbb{C} : R_2 < |z-a| < R_1\}$, and take a point $z \in A(a; R_1, R_2)$. Take a positive number $\rho > 0$ such that $\overline{B(z, \rho)}$ is included by $A(a; R_1, R_2)$. Let C_i be circles with positive direction for $B(a, R_i)$ and Γ is a boundary of $A(a; R_1, R_2)$. Then we have

$$\int_{\Gamma} \frac{f(w)}{w-z} dw = \int_{C_1} \frac{f(w)}{w-z} dw - \int_{C_2} \frac{f(w)}{w-z} dw. \quad (12.8.4)$$

Since $|(z-a)/(w-a)| < 1$ when $w \in C_1$ we obtain

$$\frac{1}{w-z} = \sum_{n=0}^{\infty} \frac{(z-a)^n}{(w-a)^{n+1}}. \quad (12.8.5)$$

This leads

$$\frac{1}{2\pi i} \int_{C_1} \frac{f(w)}{w-z} dw = \sum_{n=0}^{\infty} c_n (z-a)^n, \quad c_n = \frac{1}{2\pi i} \int_{C_1} \frac{f(w)}{(w-a)^{n+1}} dw. \quad (12.8.6)$$

Furthermore, due to holomorphy of f in $B(z, \rho)$, Cauchy's integral representation formula leads

$$f(z) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(w)}{w-z} dw \quad (12.8.7)$$

Since $|(w-a)/(z-a)| < 1$ when $w \in C_2$ we obtain

$$\frac{1}{w-z} = - \sum_{n=1}^{\infty} \frac{(w-a)^{n-1}}{(z-a)^n}. \quad (12.8.8)$$

This leads

$$-\frac{1}{2\pi i} \int_{C_2} \frac{f(w)}{w-z} dw = \sum_{n=1}^{\infty} c_{-n} (z-a)^{-n}, \quad c_{-n} = \frac{1}{2\pi i} \int_{C_2} \frac{f(w)}{(w-a)^{-n+1}} dw. \quad (12.8.9)$$

Taking limits for R_1 and R_2 and arguments for changes of paths end the proof. \square

Definition. 8.1.3. Assume $a \in U$ is an (isolated) singular point of a function f and its Laurent expansion is

$$f(z) = \sum_{n=-\infty}^{\infty} c_n(z-a)^n. \quad (12.8.10)$$

We classify a singular point a as follows.

1. A singular point a is called a **removable singular point** if $c_n = 0$ for all $n \leq -1$.
2. A singular point a is called a **pole of n -th order** if there exists a negative integer $n \leq -1$ such that $c_n \neq 0$ and $c_m = 0$ for all $m < n$.
3. A singular point a is called a **essential singular point** if there exists infinitely many negative integer n such that $c_n \neq 0$.

We have several analyses and characterizations for singular points but we omit them.

12.8.2 Theorem of residue

Definition. 8.2.1. Let $a \in U$ and $R > 0$ and a function $f: B_0(a, R) \rightarrow \mathbb{C}$ is holomorphic. The Laurent expansion of f around a is

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z-a)^n. \quad (12.8.11)$$

Then the number a_{-1} is called the **residue of f at a** and denoted by

$$a_{-1} = \text{Res}(f, a). \quad (12.8.12)$$

Theorem. 8.2.2. (Theorem of residue.) Assume a function f is holomorphic in $U = B_0(a, R)$ for $a \in U$ and $R > 0$. For any piecewise C^1 curves C in U it holds that

$$\frac{1}{2\pi i} \int_C f(z) dz = n(C, a) \text{Res}(f, a). \quad (12.8.13)$$

Proof. Trivial. □

12.8.3 Argument principle

Proposition. 8.3.1. Assume a function f is holomorphic in $B_0(a, R)$ for $a \in U$ and $R > 0$, and that a is a n -th zero point or pole. Then $a \in U$ is a pole of first order for a function $g = f'/f$ and its residue at a is given by

$$\text{Res}\left(\frac{f'}{f}, a\right) = \begin{cases} n, & a \text{ is a zero point,} \\ -n, & a \text{ is a pole.} \end{cases} \quad (12.8.14)$$

Proof. We show this the case a is a zero point. Then we have an expression

$$f(z) = (z-a)^n f_1(z) \quad (12.8.15)$$

for some holomorphic function f_1 with $f(a) \neq 0$. Hence we obtain

$$\frac{f'(z)}{f(z)} = \frac{n}{z-a} + \frac{f_1'(z)}{f_1(z)} \quad (12.8.16)$$

Since the second term in the RHS is holomorphic we obtain the result. \square

This result is generalized to the famous Argument principle and Louche's theorem.

Definition. 8.3.2. (Riemann sphere.) *One point compactification space $\overline{\mathbb{C}}$ of the complex line \mathbb{C} is called the **Riemann sphere**.*

Definition. 8.3.3. *Let $U \in \overline{\mathbb{C}}$ a domain. A function $f: U \rightarrow \mathbb{C}$ is holomorphic without several poles is called **meromorphic** in U .*

Theorem. 8.3.4. (Argument principle.) *Assume a function $f: U \rightarrow \mathbb{C}$ is meromorphic in U , the set of zero points of f is $\{a_1, \dots, a_n\}$ with multiplicity, and the set of poles is $\{b_1, \dots, b_p\}$ with multiplicity. If C is a homolog 0 curve in U and walk through no zero points and no poles of f then it holds that*

$$\frac{1}{2\pi i} \int_C \frac{f'(z)}{f(z)} dz = \sum_{j=1}^n n(C, a_j) - \sum_{k=1}^p n(C, b_k). \quad (12.8.17)$$

Proof. Use the theorem of residue. \square

Theorem. 8.3.5. (Louche's theorem.) *Assume a piecewise C^1 closed curve C surround a region U and that functions f and g are holomorphic in $U \cup C$. If $|f(z) - g(z)| < |f(z)|$ holds over C then f and g have the same number of zero points in U with multiplicity.*

12.9 Convergence theorems

Theorem. 9.0.6. (Hurwitz, [2] P191.) *Assume functions $f_n: U \rightarrow \mathbb{C}$ are holomorphic and vanish nowhere and a function f is defined by a uniform limit of a sequence (f_n) in the wider sence. Then a function $f = 0$ or f vanishes nowhere.*

Proof. Assume $f \neq 0$. Zero points of f , if they exist, are isolated. For any $z \in U$ there exists a small radius $r > 0$ such that $B(z, r) \subset U$ and $f(w) \neq 0$ for all $w \in B(z, r) \setminus \{z\}$. In particular $|f(z)|$ attains its minimum on $C := \partial B(z, r)$ and it is positive. Since sequences (f_n) and (f'_n) converge to f respectively f' we obtain

$$\lim_{n \rightarrow \infty} \frac{1}{2\pi i} \int_C \frac{f'_n(z)}{f_n(z)} dz = \frac{1}{2\pi i} \int_C \frac{f'(z)}{f(z)} dz \quad (12.9.1)$$

By Argument principle the LHS of the above is 0 for all n and hence the LHS is 0. Thus $f(z) \neq 0$. Since a point z is taken arbitrarily this ends the proof. \square

Definition. 9.0.7. *Denote the abelian algebra of continuous functions on U as*

$$C(U) := \{f: U \rightarrow \mathbb{C} : f \text{ is continuous.}\} \quad (12.9.2)$$

*A set $H \subset C(U)$ is **normal family** if, for any subset $(f_n)_n \subset H$, there exists a subsequence $(f_{n_k})_k$ and a (continuous) function f such that $f_{n_k} \xrightarrow{uw} f$.*

Theorem. 9.0.8. (Ascoli-Arzelà.) *Let $U \subset \mathbb{R}^n$ and $H \subset C(U)$. A subset H is a normal family if and only if the following two properties hold.*

1. *H is an equicontinuous family,*
2. *H is uniformly-bounded on any compact subset $K \subset U$.*

We denote an algebra of holomorphic functions on U as $\mathcal{O}(U)$.

Theorem. 9.0.9. (Montel.) *Let $H \subset \mathcal{O}(U)$ be a subset. Then the followings are equivalent.*

1. *H is a normal family.*
2. *H is uniformly-bounded on any compact subset $K \subset U$.*

Theorem. 9.0.10. (Vitali. See [207] P.374.) *Let $E = z_{n \in \mathbb{N}} \subset U$, $z_n \rightarrow z \in U$, and a family $(f_n) \subset \mathcal{O}(U)$ be uniformly-bounded in U . If sequences $(f_n(z))_n$ converge for all $z \in E$ then there exists a holomorphic function $f \in \mathcal{O}(U)$ such that $f_n \xrightarrow{\text{uw}} f$.*

12.10 Representation of holomorphic functions

First we explain the setup of Mittag-Leffler's theorem. Assume an arbitrary sequence $(a_n)_{n \in \mathbb{N}}$ satisfying

$$|a_0| \leq |a_1| \leq \cdots \leq |a_n| \leq |a_{n+1}| \leq \cdots, \quad \lim_{n \rightarrow \infty} |a_n| = \infty, \quad (12.10.1)$$

$$a_n \neq a_m \quad (n \neq m), \quad (12.10.2)$$

and that a sequence of polynomials $(P_n)_{n \in \mathbb{N}}$ in $(z - a_n)^{-1}$ have forms

$$P_n(z) = \sum_{k=1}^{m(n)} \frac{c_{n,k}}{(z - a_n)^k}. \quad (12.10.3)$$

Since P_n is holomorphic in $|z| < |a_n|$ we have a Taylor expansion

$$P_n(z) = \sum_{k=0}^{\infty} b_{n,k} z^k, \quad |z| < |a_n|, \quad n \geq 1. \quad (12.10.4)$$

Choosing natural numbers $k(n)$ for each $n \in \mathbb{N}$ properly and set functions $(\varphi_n)_n$ as

$$\varphi_n(z) = \sum_{k=0}^{k(n)} b_{n,k} z^k. \quad (12.10.5)$$

Put $\varphi_0 = 0$.

Theorem. 10.0.11. (Mittag-Leffler. Theorem 10.1 in [207].) *We assume the above setup. Then the series*

$$g(z) := \sum_{n=0}^{\infty} (P_n(z) - \varphi_n(z)) \quad (12.10.6)$$

converges uniformly in the wider sense on $\mathbb{C} \setminus \{a_n : n \in \mathbb{N}\}$ for properly chosen $(k(n))$. Furthermore it satisfies the following properties.

1. The function g is meromorphic on \mathbb{C} .
2. The set $A = \{a_n : n \in \mathbb{N}\}$ is the set of poles of g in \mathbb{C} .
3. The principal parts of g at each a_n is P_n .

Proof. Since a series P_n converges in $B(0, |a_n|/2)$, there exists a natural number $k(n)$ such that, letting $\varphi_n(z) = \sum_{k=0}^{k(n)} b_{n,k} z^k$, it follows that

$$|P_n(z) - \varphi_n(z)| < \frac{1}{2^n} \quad \left(z \in B(0, \frac{|a_n|}{2}) \right). \quad (12.10.7)$$

For any $R > 0$ there exists $N \in \mathbb{N}$ such that $|a_n| > 2R$ for all $n > N$. Since functions (P_n) is holomorphic in $B(0, R)$ for $n > N$ it holds that

$$h(z) = \sum_{n=N+1}^{\infty} (P_n(z) - \varphi_n(z)) \quad (12.10.8)$$

uniformly converges in $B(0, R)$ and hence the above function h is holomorphic in $B(0, R)$. Here the below function

$$g(z) = \sum_{n=0}^N (P_n(z) - \varphi_n(z)) + h(z) \quad (12.10.9)$$

is meromorphic in $B(0, R)$, poles are $\{a_n : 0 \leq n \leq N\}$, and the principal part of the function g are P_n at a_n .

Since $R > 0$ is arbitrary the function g is meromorphic in \mathbb{C} , and the set of poles is $\{a_n : n \in \mathbb{N}\}$, and the principal part of g is P_n at a_n . \square

Remark. 10.0.12. *Mittag-Leffler's theorem in a several complex variables version is the famous Cousin first problem.*

Theorem. 10.0.13. (Weierstrass. Theorem 5.5.4. in [161]) *Let $Z \subset U$ be a discrete set. Set numbers $\nu_\zeta \in \mathbb{Z} \setminus \{0\}$ for each $\zeta \in Z$. Then there exists a meromorphic function f on U such that, for each neighborhood of ζ , the function f can be represented as*

$$f(z) = (z - \zeta)^{\nu_\zeta} h(z), \quad h(\zeta) \neq 0, \quad (12.10.10)$$

by a holomorphic function h . The function f is holomorphic in $U \setminus Z$ and has no zero points.

Remark. 10.0.14. *Weierstrass' theorem in a several complex variables version is the famous Cousin second problem.*

12.11 Riemann's mapping theorem

Definition. 11.0.15. *Two regions U_1 and U_2 are analytically homeomorphic if there exists a biholomorphic function $f: U_1 \rightarrow U_2$.*

Theorem. 11.0.16. (Theorem 13.1 in [207].) *The following three simply connected regions are not analytically homeomorphic each other.*

1. The Riemann sphere, $\overline{\mathbb{C}}$.

2. The complex line, \mathbb{C} .
3. The unit disc, $D = B(0, 1)$.

Proof. Since the Riemann sphere is compact the image $f(\overline{\mathbb{C}})$ by a continuous function is compact. On the other hand \mathbb{C} and D is not compact and hence $f(\overline{\mathbb{C}})$ is not analytically homeomorphic to \mathbb{C} or D . If a homeomorphic function f on \mathbb{C} satisfies $f(\mathbb{C}) \subset D$ then, since f is a bounded entire function, Liouville's theorem asserts f is a constant function, i.e., $f(\mathbb{C}) \neq D$. \square

Theorem. 11.0.17. (Riemann's mapping theorem. Theorem 13.2 in [207].) *If U is a simply connected region in \mathbb{C} and $U \neq \mathbb{C}$, then there exists a holomorphic and univalent function f on U such that it maps U onto the unit disc D . Furthermore we set an additional condition*

$$f(a) = 0, \quad f'(a) > 0 \quad (12.11.1)$$

for an arbitrary point $a \in U$. Under the above additional condition a function f is unique and it is a analytically holomorphic map from U to D .

12.12 Miscellaneous results

We give examples of unexplained but important results.

1. Omitted proofs.
2. Theorem of mean: Corollary 2.4 in [207].
3. Maximum principle: Theorem 3.9 in [207].
4. Inverse mapping theorem.
5. Relations between holomorphic functions and harmonic functions.
6. Calculation of definite integrals using theorem of residue.
7. Interrelation for Fourier transform.
8. Theory of elliptic functions.
9. Theory of automorphic forms.
10. Algebraic function theory.
11. Theory of Riemann surface.

12.13 Appendix

12.13.1 Vector Analysis

12.13.2 Analytic continuation

Chapter 13

Talk: Complex analysis for Ising model

Chapter 14

Several complex variables

Cousin's problems.

Part VII

Functional Analysis

Chapter 15

Introduction to Lebesgue integral

15.1 Why Lebesgue integral?

15.1.1 Area

Theory of Lebesgue integral is one for integral, of course. However integral is closely connected with a theory of area. Consider finitely many disjoint plane sets (A_n) . We denote an area of a (plane) set A as $m(A)$. Then an area of their union $\bigcup A_n$ is $\sum_n m(A_n)$.

Next consider a infinitely many disjoint plane sets (A_n) and an area of their union. In this case we may face a severe problem if we treat “infinity” too roughly. Suppose sets $A_y = [0, 1] \times \{y\}$ for $y \in [0, 1]$ and their union $\bigcup_{y \in [0, 1]} A_y$. An area $m(A_y)$ of each A_y is 0 (later we will prove), but their union is $\bigcup_y A_y = [0, 1] \times [0, 1]$ An area is $m(\bigcup_y A_y) = 1$. Thus it holds that

$$m\left(\bigcup_y A_y\right) \neq \sum_y m(A_y). \quad (15.1.1)$$

Clearly uncountable operation leads contradiction.

One way to overcome this difficulty may be just ignoring of a union of infinitely many sets. However, if we would like to define areas for general sets by fundamental sets such as rectangles, we must consider a union of infinitely many rectangles. It suffice to consider an area of circle.

How do we define an area for a (disjoint) union of infinitely many sets?

15.1.2 Function space: approximation of functions

In application it is important to consider approximation of functions, e.g., application to analysis of differential equations. Then we would like to set distances for functions since, in computer simulation, we cannot use an exact solution but an approximated solution. However what is the distance between functions?

15.1.3 Distance by supremum norm

We can use a distance defined by supremum norm, e.g.,

$$d_\infty(f, g) := \|f - g\| := \sup_{x \in \Omega} |f(x) - g(x)|. \quad (15.1.2)$$

This distance detects the maximum pointwise distance between functions. This seems intuitive and natural, and is useful, in fact.

Distance by L^2 norm

However here is another natural distance for functions, i.e., L^2 distance. We are going to “derive” this distance. Consider vectors $f, g \in \mathbb{C}^n$. Then we can define their inner product and distance.

$$\langle f, g \rangle := \sum_{x=1}^n \overline{f(x)} g(x), \quad (15.1.3)$$

$$d_2(f, g) := \|f - g\|_2 := \sqrt{\langle f - g, f - g \rangle} \quad (15.1.4)$$

If the sum converges we can take a limit $n \rightarrow \infty$.

$$\langle f, g \rangle := \sum_{x=1}^{\infty} \overline{f(x)} g(x). \quad (15.1.5)$$

A distance d_2 is defined similarly. Now we sum up values in increments of 1. If this increment tends to 0 we obtain an integral:

$$\langle f, g \rangle := \int_0^\infty \overline{f(x)} g(x) dx. \quad (15.1.6)$$

Here an interval can be arbitrary. This distance is called a L^2 distance.

This distance is also intuitive, natural and useful. Furthermore this is a natural interpretation in physics, i.e., energy of function. Consider a rope and a system of N subropes. Then the energy of N subropes is

$$E = \sum_{k=1}^N \frac{1}{2} m v_k^2 + \sum_{k=1}^N k (x_{k-1} - x_k)^2. \quad (15.1.7)$$

A division number N is virtual, so we would like to take a limit $N \rightarrow \infty$.

$$E = \int_0^1 \left(\frac{1}{2} m \left(\frac{\partial u}{\partial t} \right)^2 + \frac{1}{2} k \left(\frac{\partial u}{\partial x} \right)^2 \right) dx. \quad (15.1.8)$$

This is an (almost) L^2 .

15.1.4 Comparison of these two distances

There are some difference in these two distances. We want to see these “sense of distances” are totally different.

First consider two functions $f, g: I = (0, N) \rightarrow \mathbb{R}$ such as

$$f(x) = 1, \quad g(x) = 1 + \varepsilon, \quad (15.1.9)$$

where $\varepsilon > 0$. Then we have

$$\|f - g\|_{\infty} = \varepsilon, \quad (15.1.10)$$

$$\|f - g\|_2 = \varepsilon\sqrt{N}. \quad (15.1.11)$$

Next consider two functions $h_n, h: [0, 1] \rightarrow \mathbb{R}$ such as

$$h_n(x) = \begin{cases} n+1, & x \in [0, \frac{1}{n^4}), \\ 1, & x \in [\frac{1}{n^4}, 1], \end{cases} \quad (15.1.12)$$

$$h(x) = 1. \quad (15.1.13)$$

Then we obtain

$$\|h_n - h\|_{\infty} = n, \quad (15.1.14)$$

$$\|h_n - h\|_2 = \frac{1}{n}. \quad (15.1.15)$$

Hence these two distances have totally different metric structures.

There is one more problem: a sequence of functions (h_n) seems to converge to h , but true convergence destination is a function h_{∞} such that

$$h_{\infty} = \begin{cases} \infty, & x = 0, \\ 1, & \text{otherwise.} \end{cases} \quad (15.1.16)$$

Furthermore a function h_{∞} is not Riemann integrable. Since a line has an area 0 (later we will prove) an area of the graph of h_{∞} must be finite, or 1. Then how do we define an integral for functions which may have an infinite value?

15.1.5 Limiting procedure

In analysis our main task is to control various limits. In “functional” analysis we must control limits for functions. Here limits, i.e., topologies are defined by integrals. We must search a manageable integral theory. We require the following points.

1. We can identify the above two functions h and h_{∞} .
2. We can manage limits and integral easily.
3. Our limit theorems are simple, easy to memorize, and analogous to the famous theorems in the theory of real numbers.

15.2 Definition and Properties of measure

15.2.1 Definition

Definition. 2.1.1. (Finitely additive family.) *Suppose X is a set and \mathcal{F} is a family of subsets X . Then the family \mathcal{F} is **finitely additive** if it satisfies the following three conditions.*

1. $\emptyset \in \mathcal{F}$.
2. A set A^c is in \mathcal{F} if $A \in \mathcal{F}$.

3. $A \cup B \in \mathcal{F}$ if $A, B \in \mathcal{F}$.

Proposition. 2.1.2. For a finitely additive family (X, \mathcal{F}) it holds that

1. $X \in \mathcal{F}$,
2. $\bigcup_{k=1}^n A_k, \bigcap_{k=1}^n A_k \in \mathcal{F}$ if $A_k \in \mathcal{F}$ for $k = 1, \dots, n$.

Definition. 2.1.3. (Finitely additive measure.) Suppose X is a set and \mathcal{F} is a finitely additive family of subsets X . Then a function $m: \mathcal{F} \rightarrow \mathbb{R}$ is called a **finitely additive measure (on \mathcal{F})** if it satisfies the following three conditions.

1. For any $A \in \mathcal{F}$ we obtain $0 \leq m(A) \leq \infty$. In particular $m(\emptyset) = 0$.
2. If $A, B \in \mathcal{F}$ and $A \cap B = \emptyset$, then $m(A + B) = m(A) + m(B)$.
3. $A \cup B \in \mathcal{F}$ if $A, B \in \mathcal{F}$.

Proposition. 2.1.4. Suppose (X, \mathcal{F}) is a finitely additive family and m is a finitely additive measure. Then it follows that

1. (**finite additivity**) if $A_1, \dots, A_n \in \mathcal{F}$ and they are pairwise disjoint, then

$$m\left(\bigcup_{j=1}^n A_j\right) = \sum_{j=1}^n m(A_j). \quad (15.2.1)$$

2. (**monotonicity**) if $A, B \in \mathcal{F}$ and $A \subset B$, then $m(A) \leq m(B)$. In particular, if $m(B) < \infty$ then

$$m(A \setminus B) = m(A) - m(B). \quad (15.2.2)$$

3. (**finite subadditivity**) If $A_1, \dots, A_n \in \mathcal{F}$ then

$$m\left(\bigcup_{j=1}^n A_j\right) \leq \sum_{j=1}^n m(A_j). \quad (15.2.3)$$

Definition. 2.1.5. (σ -additive family.) Suppose X is a set and \mathcal{B} is a family of subsets of X . The family \mathcal{B} is called a **completely additive (countably additive, or σ -additive) family, or σ -algebra** if the following three conditions hold.

1. $\emptyset \in \mathcal{B}$.
2. A set E^c is in \mathcal{B} if $E \in \mathcal{B}$.
3. $\bigcup_{n=1}^{\infty} A_n \in \mathcal{B}$ if $E_n \in \mathcal{B}$ for all n .

Definition. 2.1.6. (Measure.) Suppose X is a set and \mathcal{B} is a σ -additive family of subsets X . Then a function $\mu: \mathcal{B} \rightarrow \mathbb{R}$ is called a **measure (on \mathcal{B})** if it satisfies the following three conditions.

1. (**non-negativity**) $0 \leq \mu(A) \leq \infty$, $\mu(\emptyset) = 0$.
2. (**countable additivity**) If $A_n \in \mathcal{B}$ ($n = 0, 1, 2, \dots$) and (A_n) is pairwise disjoint, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n). \quad (15.2.4)$$

We call a pair (X, \mathcal{B}, μ) a **measure space**.

15.2.2 Why countable?

Assume a Lebesgue measure (ordinary area) μ is uncountable additive. Suppose sets $A_y = [0, 1] \times \{y\}$ for $y \in [0, 1]$ and their union $\bigcup_{y \in [0, 1]} A_y = [0, 1] \times [0, 1]$. An area $\mu(A_y)$ of each A_y is 0, but their union is $[0, 1] \times [0, 1]$. An area is $\mu(\bigcup_y A_y) = 1$. Thus it holds that

$$\mu\left(\bigcup_y A_y\right) \neq \sum_y \mu(A_y). \quad (15.2.5)$$

Clearly uncountable operation leads contradiction.

Proposition. 2.2.1. *An area of a set $A = [0, 1] \times \{0\}$ is 0.*

Proof. Let $\varepsilon > 0$ and $b = [0, 1] \times [\varepsilon/2, \varepsilon/2]$. Then $A \subset B$ and

$$0 \leq \mu(A) \leq \mu(B) \leq \varepsilon. \quad (15.2.6)$$

□

15.2.3 Definition of Lebesgue measure

Definition. 2.3.1. (Carathéodory outer measure.) *If there is a set function $\Gamma: 2^X \rightarrow \mathbb{R}$ on X and Γ satisfies the following three conditions*

1. **(non-negativity)** $0 \leq \Gamma(A) \leq \infty$, $\Gamma(\emptyset) = 0$,
2. **(monotonicity)** if $A \subset B$ then $\Gamma(A) \leq \Gamma(B)$,
3. **(subadditivity)** $\Gamma(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \Gamma(A_n)$,

then a set function Γ is called Carathéodory's outer measure or, simply, outer measure.

Theorem. 2.3.2. ([90] Theorem. 5.1) *Let \mathcal{F} be a finitely additive family of X and m be a finitely additive measure on \mathcal{F} .*

1. *If we define a set function Γ as*

$$\Gamma(A) := \inf \sum_{n=1}^{\infty} m(E_n), \quad (15.2.7)$$

where inf is taken for all countable covering $A \subset \bigcup E_n$ ($E_n \in \mathcal{F}$), then Γ is an outer measure.

2. *If m is countably additive on \mathcal{F} then $\Gamma(E) = m(E)$ for $E \in \mathcal{F}$.*

Let $-\infty \leq a_k < b_k \leq \infty$ and $R = \prod_{k=1}^n (a_k, b_k)$. We call R an open hyperrectangle and a set defined by a finite union of (open) hyperrectangles as hyperrectangle mass. Let \mathcal{R}^n be a family of (open) hyperrectangles in \mathbb{R}^n and \mathcal{F}^n a family of hyperrectangle masses. For $R = \prod_{k=1}^n (a_k, b_k) \in \mathcal{R}^n$ we define a (finitely) additive measure on \mathcal{R}^n as

$$m(R) := \prod_{k=1}^n (b_k - a_k). \quad (15.2.8)$$

Then we define the **Lebesgue outer measure** μ^* on \mathcal{F}^n as an outer measure from m .

Suppose Γ is an outer measure on X . If a set $E \subset X$ satisfies the condition

$$\Gamma(A) = \Gamma(A \cap E) + \Gamma(A \cap E^c), \quad \forall A \subset X, \quad (15.2.9)$$

then E is called **Γ -measurable**.

Theorem. 2.3.3. ([90] Theorem. 6.1) *If Γ is an outer measure on X then the family of Γ -measurable sets \mathcal{M}_Γ is a σ -algebra and Γ is a measure on \mathcal{M}_Γ .*

Consider the Lebesgue outer measure μ^* . A set belonging to \mathcal{M}_{μ^*} is called a **Lebesgue measurable set**. We denote μ for a measure μ^* on \mathcal{M}_{μ^*} and call the **Lebesgue measure**.

Theorem. 2.3.4. ([90] Theorem. 6.3) *Let \mathcal{A}_0 be an arbitrary family of subsets of a set X . Then there is a smallest σ -algebra \mathcal{B}_0 contains \mathcal{A}_0 . We denote this \mathcal{B}_0 as $\mathcal{B}(\mathcal{A}_0)$.*

Proof. There is a σ -algebra containing \mathcal{A}_0 , e.g., the power set 2^X . If we consider all σ -algebra $\{\mathcal{B}\}$ containing \mathcal{A}_0 and set $\mathcal{B}_0 := \bigcap_{\{\mathcal{B}\}} \mathcal{B}$, then this \mathcal{B}_0 is a σ -algebra and minimal. \square

Let $\mathcal{O}(\mathbb{R}^n)$ be a set of open sets in \mathbb{R}^n . Then a σ -algebra $\mathcal{B}^n = \mathcal{B}(\mathcal{O}(\mathbb{R}^n))$ is called the **Borel family** and its elements **Borel sets**. In the following our main concern is a measure space $(\mathbb{R}^n, \mathcal{B}^n, \mu)$, where μ is the Lebesgue measure. However main arguments holds for general measure spaces.

15.2.4 Null sets

Definition. 2.4.1. *Suppose (X, \mathcal{B}, μ) is a measure space and $A \in \mathcal{B}$. If $\mu(A) = 0$ then A is called a **null set** (for μ).*

For the Lebesgue measure μ a point $a \in \mathbb{R}^n$ is a null set. A countable set $A = \{a_n\}_n \in \mathbb{R}^n$ is also null since

$$\mu(A) = \sum_{n=1}^{\infty} \mu(\{a_n\}) = 0. \quad (15.2.10)$$

Note that there is an uncountable Lebesgue null set in \mathbb{R} , e.g., the **Cantor set**. For details see P.41 in [90].

15.2.5 Almost everywhere

Suppose (X, \mathcal{B}, μ) is a measure space. If there is a proposition P related to points in a set $E \in \mathcal{B}$ and it holds on $E \setminus E_0$, where $E_0 \subset E$ is null, then we say that the proposition P holds almost everywhere on E for μ .

For example, assume $X = E = \mathbb{R}$, μ is the Lebesgue measure, and $E_0 = \mathbb{Q}$. Note that \mathbb{Q} is null for μ since it is countable. Then a function

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & \text{otherwise,} \end{cases} \quad (15.2.11)$$

is 0 almost everywhere on \mathbb{R} and we denote

$$f(x) = 0 \quad \text{a.e.} \quad (15.2.12)$$

15.2.6 Examples of other measures or measure spaces

Example. 2.6.1. Let \mathcal{B}^n be the Borel σ -algebra and μ is the Lebesgue measure. We define a new measure ν as

$$\nu(A) := \int_A f(x) d\mu(x) = \int_A f(x) dx, \quad A \in \mathcal{B}, \quad (15.2.13)$$

where a function $f: \mathbb{R}^n \rightarrow \mathbb{R}_{\geq}$ is integrable.

Example. 2.6.2. Let $X = \mathbb{N}$, $\mathcal{B} = 2^{\mathbb{N}}$, and μ be a counting measure, i.e., for $A = \{n_k\}_{k=1}^m$,

$$\mu(A) := \sum_{k=1}^m 1 = |A|. \quad (15.2.14)$$

Then, for a function $f: \mathbb{N} \rightarrow \mathbb{C}$, an integral is defined by

$$\int_{\mathbb{N}} f(x) d\mu(x) = \sum_{n=1}^{\infty} f(n), \quad (15.2.15)$$

i.e., an integral is a series in this case. Hence a series theory must be contained in Lebesgue integration theory.

15.3 Properties of measure

Theorem. 3.0.3. ([90] Theorem 6.2.) Suppose (X, \mathcal{B}, μ) is a measure space and $(A_n)_n \subset \mathcal{B}$. If a sequence of sets (A_n) is monotone increasing or monotone decreasing and $\mu(A_1) < \infty$, then

$$\mu\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n). \quad (15.3.1)$$

In general

1.

$$\mu\left(\liminf_{n \rightarrow \infty} A_n\right) \leq \liminf_{n \rightarrow \infty} \mu(A_n). \quad (15.3.2)$$

2. If $\mu(\bigcup A_n) < \infty$ then

$$\limsup_{n \rightarrow \infty} \mu(A_n) \leq \mu\left(\limsup_{n \rightarrow \infty} A_n\right). \quad (15.3.3)$$

3. If $\mu(\bigcup A_n) < \infty$ and $\lim_{n \rightarrow \infty} A_n$ exists then we obtain

$$\mu\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n). \quad (15.3.4)$$

Proof. If (A_n) is increasing we have $A_n = \bigcup_{k=1}^n (A_k \setminus A_{k-1})$, where $A_0 := \emptyset$. Since $\lim_{n \rightarrow \infty} A_n = \sum_{k=1}^{\infty} (A_k \setminus A_{k-1})$ additivity shows that

$$\mu\left(\lim_{n \rightarrow \infty} A_n\right) = \sum_{k=1}^{\infty} \mu(A_k \setminus A_{k-1}) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mu(A_k \setminus A_{k-1}) = \lim_{n \rightarrow \infty} \mu(A_n). \quad (15.3.5)$$

If (A_n) is decreasing then a sequence $(A_1 - A_n)_n$ is increasing and $\lim_{n \rightarrow \infty} (A_1 - A_n) = A_1 - \lim_{n \rightarrow \infty} A_n$. Hence the above and $\mu(A_1) < \infty$ shows that

$$\begin{aligned} \mu(A_1) - \mu\left(\lim_{n \rightarrow \infty} A_n\right) &= \mu\left(A_1 \setminus \lim_{n \rightarrow \infty} A_n\right) \\ &= \lim_{n \rightarrow \infty} \mu(A_1 \setminus A_n) = \lim_{n \rightarrow \infty} (\mu(A_1) - \mu(A_n)) = \mu(A_1) - \lim_{n \rightarrow \infty} \mu(A_n). \end{aligned} \quad (15.3.6)$$

(15.3.7)

For a general case set $B_n = \bigcap_{k=n}^{\infty} A_k$. Then $(B_n)_n$ is monotone increasing and $\liminf_{n \rightarrow \infty} B_n = \liminf_{n \rightarrow \infty} A_n$. Hence it holds that

$$\mu\left(\liminf_{n \rightarrow \infty} A_n\right) = \mu\left(\liminf_{n \rightarrow \infty} B_n\right) \leq \liminf_{n \rightarrow \infty} \mu(A_n). \quad (15.3.8)$$

The remaining parts are derived from similar arguments. \square

15.4 Product measure and extension theorem

Definition. 4.0.4. Suppose $(X, \mathcal{B}_X, \mu_X)$ is a measure space. A measure space X is **σ -finite** if there is a monotonically increasing sequence of sets (X_k) belonging to \mathcal{B}_X such that $\mu_X(X_k) < \infty$ ($k \in \mathbb{N}$) and $\lim_{n \rightarrow \infty} X_n = X$.

Suppose $(X, \mathcal{B}_X, \mu_X)$ and $(Y, \mathcal{B}_Y, \mu_Y)$ are σ -finite measure spaces. We want to define a **product measure space** on $Z := X \times Y$.

We call a set $C \in Z$ a **rectangle** if it has the form

$$C = A \times B, \quad A \in \mathcal{B}_X, \quad B \in \mathcal{B}_Y. \quad (15.4.1)$$

The family \mathcal{F} of rectangles in C becomes a finitely additive family. Then we call $\mathcal{B}_Z := \mathcal{B}(\mathcal{F})$ a **product Borel family** or **product σ -algebra**, sometimes denoted by $\mathcal{B}_Z = \mathcal{B}_X \times \mathcal{B}_Y$.

Next we define a measure μ on \mathcal{B}_Z . For a rectangle $C = A \times B$, $\mu_X(A)$, $\mu_Y(B) < \infty$ we define

$$\mu(C) := \mu_X(A)\mu_Y(B). \quad (15.4.2)$$

We call μ a **product measure**.

Theorem. 4.0.5. (Theorem 9.4 [90].) *The above μ is unique.*

For the proof we need the following

Theorem. 4.0.6. (E. Hopf's extension theorem, Theorem 9.1 [90].) *Suppose X is a set, \mathcal{F} is a finitely additive family, $\mathcal{B} = \mathcal{B}(\mathcal{F})$, and m is a finitely additive measure on \mathcal{F} . A finitely additive measure m can be extended to a measure μ on \mathcal{B} if and only if m is countably additive on \mathcal{F} . If, furthermore, X is σ -finite then the extension is unique.*

15.5 Measurable functions

Suppose X is a set and \mathcal{B} is a σ -algebra on X . Fix $E \in \mathcal{B}$ and set a function $f: E \rightarrow \bar{\mathbb{R}}$, where $\bar{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$. Define the following sets:

$$E(f > a) := \{x \in E : f(x) > a\}, \quad (15.5.1)$$

$$E(f \leq a) := \{x \in E : f(x) \leq a\}, \quad (15.5.2)$$

$$E(f = a) := \{x \in E : f(x) = a\}, \quad (15.5.3)$$

$$E(a < f \leq b) := \{x \in E : a < f(x) \leq b\}, \quad \text{etc.} \quad (15.5.4)$$

$$(15.5.5)$$

A function f on E is **\mathcal{B} -measurable** or **measurable** if

$$E(f > a) \in \mathcal{B} \quad \forall a \in \mathbb{R}. \quad (15.5.6)$$

A function $f: E \rightarrow \bar{\mathbb{R}}$ is called a **simple function** if, for a finite pairwise disjoint decomposition $E = \bigsqcup_{k=1}^n E_k$, f has a form

$$f(x) = \sum_{j=1}^n \alpha_j \chi_{E_j}(x), \quad (15.5.7)$$

where $\alpha_j \in \mathbb{R}$, $\alpha_j \neq \alpha_k$, and χ_E is a defining function of E .

Theorem. 5.0.7. *If a function $f: E \rightarrow \mathbb{R}_{\leq}$ is measurable, then there is a non-negative and monotonically increasing sequence of simple functions (f_n) such that $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ for all $x \in E$.*

Proof. Define (f_n) as

$$f_n(x) = \begin{cases} \frac{k-1}{2^n}, & x \in E \left(\frac{k-1}{2^n} \leq f < \frac{k}{2^n} \right) \quad (k = 1, 2, \dots, 2^n n), \\ n, & x \in E(f \geq n). \end{cases} \quad (15.5.8)$$

□

Theorem. 5.0.8. *If f and g are measurable then sets $E(f > g)$, $E(f \leq g)$, and $E(f = g)$ are all measurable.*

Proof. The claims holds, e.g.,

$$E(f > g) = \bigcup_{q \in \mathbb{Q}} (E(f > q) \cap E(q > g)) \in \mathcal{B}. \quad (15.5.9)$$

□

Theorem. 5.0.9. *If f is measurable then $|f|^\alpha$ is also measurable for any $\alpha \in \mathbb{R}$. For the case $\alpha < 0$ we define $|f(x)|^\alpha = \infty$ for the point x where $f(x) = 0$.*

Proof. We consider only the case $\alpha > 0$ here. Let $a \in \mathbb{R}$. If $a > 0$ then we have

$$E(|f|^\alpha > a) = E(|f| > a^{1/\alpha}) = E(f > a^{1/\alpha}) \cup E(f < -a^{1/\alpha}) \in \mathcal{B}. \quad (15.5.10)$$

If $a \leq 0$ then it holds that

$$E(|f|^\alpha > a) = E \in \mathcal{B}. \quad (15.5.11)$$

□

Theorem. 5.0.10. *If f and g are measurable then, for any $\alpha, \beta \in \mathbb{R}$, a function $\alpha f + \beta g$ is measurable.*

Proof. It suffices to show the case $\alpha > 0$. If $\beta = 0$ then, for any $a \in \mathbb{R}$,

$$E(\alpha f > a) = E\left(f > \frac{a}{\alpha}\right) \in \mathcal{B}. \quad (15.5.12)$$

If $\beta \neq 0$ then, for any $a \in \mathbb{R}$,

$$E(\alpha f + \beta > a) = E\left(f > \frac{a - \beta}{\alpha}\right) \in \mathcal{B}. \quad (15.5.13)$$

Thus a function $\alpha f + \beta$ is measurable. Hence it follows that, for $\beta \geq 0$

$$E(\alpha f + \beta g > a) = E\left(g \geq -\frac{\alpha}{\beta}f + \frac{a}{\beta}\right) \in \mathcal{B}. \quad (15.5.14)$$

□

Theorem. 5.0.11. *If f and g are measurable then the product fg is also measurable.*

Proof. This is due to the fact

$$fg = \frac{1}{4}((f+g)^2 - (f-g)^2). \quad (15.5.15)$$

□

Theorem. 5.0.12. *If all f_n are measurable then the following functions are also measurable.*

$$g := \sup_{n \in \mathbb{N}} f_n, \quad h := \inf_{n \in \mathbb{N}} f_n, \quad \limsup_{n \rightarrow \infty} f_n, \quad \liminf_{n \rightarrow \infty} f_n. \quad (15.5.16)$$

Proof. The function g is measurable since, for any $a \in \mathbb{R}$,

$$E(g > a) = \bigcup_{n=1}^{\infty} E(f_n > a) \in \mathcal{B}. \quad (15.5.17)$$

The function h is measurable since $h = -\sup_{n \in \mathbb{N}}(-f_n)$. Setting $g_n = \sup_{k \geq n} f_k$, g_n is measurable and $\inf_{n \in \mathbb{N}} g_n = \limsup_{n \rightarrow \infty} f_n$ is also measurable since g_n is measurable. □

Sum up:

Theorem. 5.0.13. *The set of measurable functions is a linear space and closed in pointwise limits.*

15.6 Definition of Lebesgue integral and its properties

Fix a measure space (X, \mathcal{B}, μ) and $E \in \mathcal{B}$.

Definition. 6.0.14. (Definition of integral for non-negative simple functions.)
If a simple function $f: E \rightarrow \mathbb{R}_{\leq}$ defined by

$$f = \sum_{j=0}^n \alpha_j \chi_{E_j}, \quad E = \bigsqcup_{j=0}^n E_j, \quad \alpha_0 = 0 < \alpha_j, \quad (15.6.1)$$

then we define and denote an integral of f as

$$\int_E f(x) d\mu(x) := \sum_{j=1}^n \alpha_j \mu(E_j). \quad (15.6.2)$$

Theorem. 6.0.15. The above definition is well-defined.

Proof. A simple function can be described by

$$f = \sum_{k=0}^m \beta_k \chi_{F_k}, \quad E = \bigsqcup_{k=0}^m F_k, \quad \beta_0 = 0 < \beta_k. \quad (15.6.3)$$

If there exists $j, k \geq 1$ such that $\mu(E_j \cap F_k) = \infty$ then

$$\sum_{j=1}^n \alpha_j \mu(E_j) = \sum_{k=1}^m \beta_k \mu(F_k) = \infty. \quad (15.6.4)$$

□

Lemma. 6.0.16. (Linearity of integrals.)

1. If functions f and g are non-negative and simple then

$$\int_E (f + g) d\mu = \int_E f d\mu + \int_E g d\mu. \quad (15.6.5)$$

2. If a function f is non-negative and simple and if $A \cup B \subset E$ then it holds that

$$\int_{A+B} f d\mu = \int_A f d\mu + \int_B f d\mu. \quad (15.6.6)$$

Proof. Proof for 1). By assumption it follows that

$$f = \sum_{j=0}^n \alpha_j \chi_{E_j}, \quad E = \bigsqcup_{j=0}^n E_j, \quad \alpha_0 = 0 < \alpha_j, \quad (15.6.7)$$

$$g = \sum_{k=0}^m \beta_k \chi_{F_k}, \quad E = \bigsqcup_{k=0}^m F_k, \quad \beta_0 = 0 < \beta_k. \quad (15.6.8)$$

A function $(f + g)$ is also simple and described by

$$f + g = \sum_{j,k} (\alpha_j + \beta_k) \chi_{E_j \cap F_k}. \quad (15.6.9)$$

If there exists a number $j \geq 1$ such that $\mu(E_j) = \infty$ then the both hand sides are ∞ and the claim is shown. If $\mu(E_j \cap F_k) < \infty$ for any j, k then we have

$$\int_E (f + g) d\mu = \sum_{j,k} (\alpha_j + \beta_k) \mu(F_j \cap G_k) \quad (15.6.10)$$

$$= \sum_{j=1}^n \alpha_j \sum_{k=0}^m \mu(F_j \cap G_k) + \sum_{k=1}^m \beta_k \sum_{j=0}^n \mu(F_j \cap G_k) \quad (15.6.11)$$

$$= \sum_{j=1}^n \alpha_j \mu(F_j) + \sum_{k=1}^m \beta_k \mu(G_k) \quad (15.6.12)$$

$$= \int_E f d\mu + \int_E g d\mu. \quad (15.6.13)$$

Proof for 2).

$$\int_{A+B} f d\mu = \sum_{j=1}^n \alpha_j \mu(E_j \cap (A \sqcup B)) \quad (15.6.14)$$

$$= \sum_{j=1}^n \alpha_j \mu(E_j \cap A) + \sum_{j=1}^n \alpha_j \mu(E_j \cap B) = \int_A f d\mu + \int_B f d\mu. \quad (15.6.15)$$

□

Lemma. 6.0.17. (Lemma 12.2 in [90].) *Suppose functions (f_n) and g are non-negative and simple, g is finite, a sequence (f_n) is monotonically increasing, and $\lim f_n(x) \geq g(x)$ for any $x \in E$ then it holds that*

$$\lim_{n \rightarrow \infty} \int_E f_n d\mu \geq \int_E g d\mu. \quad (15.6.16)$$

Proof. We may assume g is positive on E . Since g can be described by

$$g = \sum_{j=m} \alpha_j \chi_{E_j}, \quad 0 \leq \alpha_j < \infty, \quad (15.6.17)$$

putting $\alpha = \min \{\alpha_j\}$ and $\beta = \max \{\alpha_j\}$ it holds that

$$0 < \alpha \leq g \leq \beta < \infty. \quad (15.6.18)$$

For $\varepsilon \in (0, \alpha)$ a function $g - \varepsilon > 0$ is simple. If we set $F_n := E(f_n > g - \varepsilon)$, then, since (f_n) is monotonically increasing and $\lim_{n \rightarrow \infty} f_n \geq g$, it follows that (F_n) is monotonically increasing and

$$\lim_{n \rightarrow \infty} F_n = E, \quad \lim_{n \rightarrow \infty} \mu(F_n) = \mu(E). \quad (15.6.19)$$

If $\mu(E) < \infty$, since $\lim \mu(E \setminus F_n) = \mu(E) - \lim \mu(F_n) = 0$, for any $\varepsilon > 0$ there is n_0 such that

$$\mu(E \setminus F_n) < \varepsilon, \quad n \geq n_0. \quad (15.6.20)$$

Since it holds that

$$\int_E f_n d\mu \geq \int_{F_n} f_n d\mu \geq \int_{F_n} (g - \varepsilon) d\mu = \int_{F_n} g d\mu - \varepsilon \mu(F_n) \quad (15.6.21)$$

$$\geq \int_E g d\mu - \int_{E \setminus F_n} g d\mu \varepsilon \mu(E) \quad (15.6.22)$$

$$\geq \int_E g d\mu - \beta \mu(E \setminus F_n) + \varepsilon \mu(E) \quad (15.6.23)$$

$$> \int_E g d\mu - \varepsilon (\beta + \mu(E)), \quad (15.6.24)$$

we obtain

$$\lim_{n \rightarrow \infty} \int_E f_n d\mu \geq \int_E g d\mu - \varepsilon (\beta + \mu(E)). \quad (15.6.25)$$

For the case $\mu(E) = \infty$, we use a similar reasoning. \square

Lemma. 6.0.18. (Lemma 12.3 in [90].) *Suppose sequences of nonnegative simple functions (f_n) and (g_n) are monotonically increasing in n and that $\lim_{n \rightarrow \infty} f_n = \lim_{n \rightarrow \infty} g_n$, then it follows that*

$$\lim_{n \rightarrow \infty} \int_E f_n d\mu = \lim_{n \rightarrow \infty} \int_E g_n d\mu. \quad (15.6.26)$$

Proof. By assumption we have

$$\lim_{n \rightarrow \infty} \int_E f_n d\mu \geq \int_E g_m d\mu. \quad (15.6.27)$$

Clearly the converge inequality holds. \square

Definition. 6.0.19. *There is a monotonically increasing sequence of simple functions (f_n) such that $\lim_{n \rightarrow \infty} f_n = f$. Then We define an integral for non-negative measurable function f by*

$$\int_E f d\mu := \lim_{n \rightarrow \infty} \int_E f_n d\mu. \quad (15.6.28)$$

This definition makes sense since the above limit does not depend on a choice of a sequence.

For a general real-valued function f , we decompose f as

$$f_+(x) := \max\{f(x), 0\}, \quad f_-(x) := \max\{-f(x), 0\}. \quad (15.6.29)$$

*A function f is **integralable** if $\int_E f_+ < \infty$ or $\int_E f_- < \infty$, and we define*

$$\int_E f d\mu := \int_E f_+ d\mu - \int_E f_- d\mu. \quad (15.6.30)$$

If f is complex-valued we decompose f as $f = \Re f + i \Im f$ and define an integral

$$\int_E f d\mu := \int_E \Re f d\mu + i \int_E \Im f d\mu. \quad (15.6.31)$$

Theorem. 6.0.20. (Theorem 12.1 in [90].) *If a function f is integralable on E then*

$$\left| \int_E f d\mu \right| \leq \int_E |f| d\mu. \quad (15.6.32)$$

Theorem. 6.0.21. (Theorem 12.2 in [90].)

1. *If $\mu(E) = 0$ then*

$$\int_E f d\mu = 0. \quad (15.6.33)$$

2. *If f is integralable on E then*

$$\mu(E(f = \infty)) = \mu(E(f = -\infty)) = 0. \quad (15.6.34)$$

Proof. We show only 2). Set $E_0 := E(f = \infty)$. Then, for $\alpha > 0$, we obtain

$$\int_E f_+ d\mu = \int_{E_0} f_+ d\mu + \int_{E \setminus E_0} f_+ d\mu \geq \int_{E_0} \alpha d\mu = \alpha \mu(E_0). \quad (15.6.35)$$

Since α is arbitrary $\mu(E_0) = 0$. □

Theorem. 6.0.22. (Theorem 12.3 in [90].)

1. (Linearity of integral) *For integralable functions f, g and any complex number α, β it holds that*

$$\int_E (\alpha f + \beta g) d\mu = \alpha \int_E f d\mu + \beta \int_E g d\mu. \quad (15.6.36)$$

2. *If f is integralable and $A \sqcup B \subset E$ then we have*

$$\int_{A \sqcup B} f d\mu = \int_A f d\mu + \int_B f d\mu. \quad (15.6.37)$$

15.7 Famous theorems for exchanging orders of limits

We fix a measure space (X, \mathcal{B}, μ) .

Theorem. 7.0.23. (Theorem. 13.1 in [90].) *If functions (f_n) is nonnegative and $f = \sum_{n=1}^{\infty} f_n$ then*

$$\int_E f d\mu = \sum_{n=1}^{\infty} \int_E f_n d\mu. \quad (15.7.1)$$

Theorem. 7.0.24. (Monotone convergence theorem, Beppo-Levi, Theorem. 13.2 in [90].) *If a sequence (f_n) is nonnegative and monotonically increasing and that $\lim_{n \rightarrow \infty} f_n = f$ it holds that*

$$\int_E f d\mu = \lim_{n \rightarrow \infty} \int_E f_n d\mu. \quad (15.7.2)$$

Proof. We may assume functions (f_n) is finite on E . Letting

$$g_1 = f_1, \quad g_n = f_n - f_{n-1} \quad (n \geq 2), \quad (15.7.3)$$

we obtain

$$f_n = \sum_{k=1}^n g_k, \quad f = \sum_{k=1}^{\infty} g_k, \quad g_k \geq 0. \quad (15.7.4)$$

Then it follows that

$$\int_E f d\mu = \sum_{k=1}^{\infty} \int_E g_k d\mu = \lim_{n \rightarrow \infty} \sum_{k=1}^n \int_E g_k d\mu = \lim_{n \rightarrow \infty} \int_E f_n d\mu. \quad (15.7.5)$$

□

Theorem. 7.0.25. (Absolute continuity, Theorem 13.4 in [90].) *Assume a function f is integrable on X . Then, for any $\varepsilon > 0$, there is a $\delta > 0$ such that*

$$\mu(E) < \delta \implies \left| \int_E f d\mu \right| < \varepsilon. \quad (15.7.6)$$

Theorem. 7.0.26. (Fatou's lemma, Theorem 13.5 in [90].) *If functions (f_n) is nonnegative then it holds that*

$$\int_E \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_E f_n d\mu. \quad (15.7.7)$$

Theorem. 7.0.27. (Lebesgue's dominant convergence theorem, Theorem 13.6 in [90].) *If there is a function φ on E and that*

$$|f_n| \leq \varphi \quad (n \in \mathbb{N}), \quad (15.7.8)$$

then it follows that

$$\int_E \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_E f_n d\mu, \quad (15.7.9)$$

$$\limsup_{n \rightarrow \infty} \int_E f_n d\mu \leq \int_E \limsup_{n \rightarrow \infty} f_n d\mu. \quad (15.7.10)$$

In particular, if there is $f = \lim_{n \rightarrow \infty} f_n$ then we have

$$\lim_{n \rightarrow \infty} \int_E f_n d\mu = \int_E \lim_{n \rightarrow \infty} f_n d\mu. \quad (15.7.11)$$

Theorem. 7.0.28. (Theorem 13.7 in [90].) *If a sequence of functions (f_n) satisfies*

$$\sum_{n=1}^{\infty} \int_E |f_n| d\mu < \infty, \quad (15.7.12)$$

then there is an a.e. finite function $\sum_{n=1}^{\infty} f_n$ such that

$$\sum_{n=1}^{\infty} \int_E f_n d\mu = \int_E \sum_{n=1}^{\infty} f_n d\mu. \quad (15.7.13)$$

Theorem. 7.0.29. (Theorem 14.2 in [90].) *If a function $f(x, t) = f: X \times (a, b) \rightarrow \mathbb{C}$ is integralable on X for x and differentiable for t and that there is a integralable function φ such that $\left| \frac{\partial f}{\partial t} \right| \leq \varphi$, then it holds that*

$$\frac{d}{dt} \int_X f(x, t) d\mu(x) = \int_X \frac{\partial f}{\partial t}(x, t) d\mu(x). \quad (15.7.14)$$

15.8 Fubini's theorem

15.9 Function spaces L^p

Riemann-Lebesgue. Def of Sobolev using Fourier transform. P83, Theorem 12.6. P84, Theorem 12.7, P85, Theorem 12.8.

15.10 Riesz-Markov-Kakutani's theorem

15.11 Fourier Analysis

15.11.1 Hilbert space

15.11.2 Fourier series

15.11.3 Fourier transform

15.11.4 Functions of positive type: Bochner's theorem

15.12 Miscellaneous results

1. Omitted proofs.
2. Examples and construction of measures. ([90], Section 4-6, Theorem 4.2.)
3. Details for outer measure. ([90], Section 5.)
4. Important (topological) properties of the Lebesgue measure. ([90], Section 7.)
5. Completion of a measure. ([90], Section 8.)
6. Borel measurability and Lebesgue measurability. ([90], Section 11.)
7. Radon-Nikodym's theorem. ([90], Section 18.)
8. Probability theory. ([238, 72])

Chapter 16

Introduction to functional analysis

Chapter 17

Introduction to operator theory

17.1 Functional calculus

17.2 Normed field

Chapter 18

Introduction to Hilbert Space Theory And Operator Theory

18.1 Introduction

[8, 27, 26, 30, 46, 45, 81, 101, ?, 124, 157, 152, 160, 167, 114, 222]

18.1.1 概要

この間 Twitter で新入生が線型代数何ぞ的なこと言っていたのもあるので、新入生向けに線型代数の世界を見せたい。私が話せるのは解析学周辺しかないが、ないよりはましだろう。Hilbert 空間と線型作用素を基本に話す。線型代数へのモチベーションを高めることを目的に概論的な話で 4-5 回に分けて解説する。主な参考書は [81, 27, 26] とする。

やる予定の内容を書いておきたい。基本的には抽象論をやる。作用素論方面の話に行ったらスペクトル定理まで進む。作用素環でも大事になる。非可換幾何への展開でまた L^2 など具体的な話との関係が返ってくる。あと、 L^2 のような話は具体的な話はもちろん大事だが、これはイントロで少し触れるだけにする。

まずイントロでする予定の話。まず超大雑把に言って教養でやる線型代数らしい線型代数と、微分方程式方面と関わる方面の話と、関数解析または作用素論的な抽象論みたいな感じの話がある的な話をする。

18.1.2 線型代数からの代数的展開

加群への展開とか、Lie 群への展開とか数学として取り逃すところは色々出てくるが、この辺は私の数学力的に手に負えないところが出てくるので色々ある、とだけ言って逃げる。ただ表現論と Fourier と、みたいなところと量子力学とかは少し触れたい。

18.1.3 標準的なコースで出てくる諸概念の重要性

あと標準的なコースの重要性はきちんと言わないといけない。行列式と固有値、固有ベクトルあたりは何をネタにしよう。物理の各所で出てくるが、固体物理とい

うか連成振動とかその辺か。あと統計学での主成分分析とかそういう話か。この辺、具体例を仕入れる必要がある。

固有値、固有ベクトルは量子力学とかその他物理でも色々展開があるという話はしよう。物理の話ばかりしているのもどうかという気はするが、応用はそれしか知らない無学な市民だった。

Google のページランクみたいな話もしよう。確率との関係とかエルゴードとか言っておくと響く向きには響くだろう。これ、数値計算とも関係するかなりクールな話なので盛り込みたい。

とりあえず有限次元はこんなものか。無限次元というか微分積分への接続として平衡点近傍の話をもってくる方がいいか。あとは微分作用素と積分作用素の線型性は必ず触れる。

18.1.4 線型代数と物理や工学への応用

有限次元と言ったところで専門に近い所で見ても色々あるし困る。とりあえずハバードだとか、直接的に研究に結び付くくらいやばい、という話はしよう

あとは数値計算でも使う的な話は入れよう。微積分との絡みで平衡点近傍の安定性とかそんな話もしよう。

脱線するが、平衡点近傍の話、多分力学系とかそういうところでも使う。あまりきちんと勉強していないが、山本義隆の解析力学にも解説あるし、ゆきみさんいわく常微分方程式と解析力学にも解説あるらしい。

これは適当な線型化から系の性質を調べるとかいう話で、微分積分や力学とも深い関係がある。機械工学とかその辺でも確か出てくるはずとかそんな話をしたい。

考えてみれば、Hubbard や Google のページランクについては動画を作ったのだった。

18.1.5 線型代数と量子力学

Hilbert 空間の抽象論と作用素論的な展開と量子力学との関係的なアレはあとで詳しくやるから、軽くこなす。まずは有限次元の方が。

18.1.6 線型代数と Hilbert 空間

線型代数は (数学内部または少なくとも物理と物理に近い工学で) 役に立つという話はされるだろうが、あまり具体的な話はされない (時間がない) だろうから、その辺の話から入る。新入生向けなので、まず Hilbert 空間は何ぞというところを話す。高校でもやった三角関数の積分が実は Hilbert 空間で意味を持つというところ、微分積分と線型代数の交点というか親玉みたいな話としての関数解析で大事な空間という話をする。また、物理でそれなりに色々な数学が出てくるが、線型代数という視点でクリアで統一的な理解ができるから大事という話をする。微分作用素、積分作用素の線型性とかも話す必要がある。

18.1.7 微分方程式

物理または工学上大事な数学的道具立てとして大事な微分方程式があるが、初等的な方程式なら具体的に解ける。「線型の微分方程式」という中で既に線型性が出ていたので、そういうところで解析と線型代数の関わりみたいな話がしたい。これを解く中で現われる直交多項式の話の「直交」も線型代数由来の話で、これが

Hilbert 空間の話という感じで、[27] の 1 章の演習問題にいくつか書いてあるので、一応参考文献として挙げておこう。

今回、個別の話をやっている余裕はなからうが Legendre やら Bessel やら、量子力学とか電磁気周りでの微分方程式を解くときにも出てくるという話も盛り込みたいこれは個別の関数の相手もそれはそれで大事なのだが、理屈としては線型空間論で一括処理できるのだ、という認識を持つことで数学的、精神的な負担を減らすことを目的に、必ず触れるようにしたい。

あとアレだ、モノによっては多重極展開とか応用上の意味があったりもするから、単なる数学ではない部分もある的なアレ。変分とか無限次元の微分とかいう話はすると楽しいかよくわからないが、ネタとして書いておこう。

18.1.8 微分方程式を解くときに出てくる直交多項式

第 1 回のイントロでは特に特殊関数周りの具体例を色々あげておきたいと思っている。全体的な話として、まだ買っていないのだが [8] がかなり気に入っている。

とりあえず触れようと思っているのは、Legendre 多項式、Legendre 陪関数、Hermite 多項式、Laguerre 多項式、Fourier 級数のあたりだ。ちなみに今はじめて知ったのだが、Chebyshev 多項式は [222] によると、計算機の中での応用があるらしい。

Legendre 多項式、球 Bessel 多項式

Legendre や球 Bessel については [144] が参考になるかと思う。自分が知っている話、ということで物理への応用について話す予定で、正にそういう話だ。Laguerre は例えば [101] を検討している。上記多項式もそうだが、Hermite についても手元にある本含め、まだ資料をあさっている。

今すぐ参考文献を知りたい向きは、基本的には偏微分方程式を解くところで使うので、その辺で探すといい。「物理数学 Legendre 多項式」などで探せば色々出てくる。

Fourier

Fourier は熱方程式、波動方程式、電磁気学あたりで探すといいだろう、数学の本ではあるが、逆問題を通じた応用的な色彩が強い本として、波動方程式への応用については [160] を、熱方程式への応用については [114] を参考にすると楽しいだろう。

物理への応用に関してよい参考書は今探しているところだ。波動の本でもいいが、電磁気（電磁波）からの話が個人的に気に入っているというか感覚が掴みやすかったんで、その辺で探すといい。もちろん、自分の専門に近いところ、自分にとって分かりやすいところで探すのが一番いい。いいのがあったら教えてほしい。

微分作用素の固有値展開からの Fourier という線もある。Fourier は高校でやった三角の積分が直交関係を表す。

18.1.9 変分

多項式から話題を変えるが、例えば変分というのがある。変分原理として物理の各所で現われるが、量子力学で基底エネルギーを出すのに使うこともある。実係数の微分方程式への数学的応用ということでは [46] が定評がある。もちろんかっちりとした数学の本だ。Hilbert 空間を中心に議論されている。最近演習問題も追加された英語版 [45] も出版されたので、買うならそちらを買った方がいいかもし

れない。東大の微分方程式系の研究室での学部 4 年のセミナーでも使われることがあるようなので、そのくらいきちんとした本だ。

また、何度も紹介しているが、解析力学というか幾何学での変分ということでは [157] が比較的分かりやすく、しかも面白い。読んだことはないのだが、物理での変分原理については [167] もある。

18.1.10 作用素のスペクトル

これまでの微分方程式の話とは大分変わるが、作用素論につなげるので、量子力学とスペクトルの話もしたいと思っている。これについては [81] はもちろんのこと、数理物理としては新井先生の本 [27, 30] がいい。

量子力学での変分に関する数学的に精密な話も書いてある。他には、作用素の関数やユニタリ表現に関する話も大事だ。作用素の関数については先日ワヘイハイオフで詳しい話を聞かせる、という要望を受けたので、別途早めにまとめようと思っている。

18.1.11 Taylor 展開と作用素論

また、Taylor 展開と作用素論ということで e^{ipx} の話もしよう。簡単に説明しておくところな感じ。 $f(x)$ を原点周りで Taylor 展開するとこうなる：

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{d}{dx} \right)^n f(0). \quad (18.1.1)$$

どうでもいいが、量子力学っぽく $p = -id/dx$ と書こう：

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} (ip)^n f(0). \quad (18.1.2)$$

ここで指数関数の Taylor 展開は

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n \quad (18.1.3)$$

となる。ここで Taylor 展開の $\sum_{n=0}^{\infty} \frac{1}{n!} (ip)^n$ は x に ip を代入したものと同じ形をしていることに注意して次のように書き換えてみる。

$$f(x) = (e^{ipx} f)(0). \quad (18.1.4)$$

指数関数に微分作用素を叩き込むという荒技を披露したが、作用素論を使ってこれが正当化できます、みたいなことも言いたい。また、作用素の指数関数 e^{ipx} は Taylor 展開で定義してしまうと解析関数に対してしか定義できないが、 x だけずらす作用素と思えば一般の関数に対して定義できる。ここでユニタリ作用素とかそういう話になる。あと x だけずらす作用素 e^{ipx} の無限小生成子としての運動量という所から、解析力学と量子力学の関係がどうの、みたいな話もちろっと触れたい。

以上大体イントロで話す予定のこと。

18.1.12 2 回目の概論

2 回目から実際にもう少し踏み込んだ話をしていく。まずは Hilbert 空間自体の話をする。[?] には Banach 空間の話もあるが、時間的に多分カットだろう。演習問題になっている定理にも少し触れたい。完備性の話などもあるので、証明もポイントをおさえて触れていきたい。

18.1.13 3-4 回目の概論

引き続き 2, 3 章を力づくでやっていく。非有界作用素はゴツ過ぎて触れられないが、スペクトル定理は紹介する。スペクトル定理は無限次元版の対角化だ。スペクトル測度や解析関数カルキュラスとか出てきてやばいのだが、むしろ色々な数学との関係話を話す機会として採り上げたい。Stone の定理と量子力学の話とかも一応入れる予定。

18.1.14 参考文献の紹介

参考文献をまとめておこう。[81, 27] の他、1 つの展開としての作用素環方面、特に (非可換) 幾何方面ということで、数学会で PDF が公開されている [152] も紹介しておく。触れる予定はないが、微分方程式関係と共に関数解析をやろうという感じの本も紹介だけはしておこう。[160, 114, 46, 45] だ。[45] は [46] のあと、演習問題を追加して増補されたバージョンだ。英語に慣れることも含め、こちらを読んだ方が楽しいだろう。こういう具体的な方から学ぶのが好きな人は頑張ってアタックしてみてほしい。また、こちらに興味があるという人は声をかけてほしい。トークしろと言われると困る部分はあるが、一緒に勉強しようというなら時間さえ合えば付き合いたい。そしてプロデュースしたい。

Chapter 19

Fourier Analysis

動画から適当に切り出してくる.

Chapter 20

重ね合わせの原理と Fourier 解析

20.1 Fourier 級数–関数の大域近似

20.2 Fourier 変換

20.3 Dirac の δ 関数

20.4 波動方程式の解

20.5 Green 関数

20.6 スペクトル分解

Chapter 21

Talk: Inverse analysis for heat equation—mathematics, physics, engineering

DVD の原稿をもう少し整理して張る.

Chapter 22

Talk: Analysis for the Thomas-Fermi functional after Lieb-Loss [130]

22.1 Introduction

22.1.1 Comparison for the two problems

Here we talk about two problems, i.e.,

1. theory for the Schrödinger equation,
2. theory for Thomas-Fermi problem.

The former is connected to

- Lebesgue integral and various inequalities,
- linear algebra,
- functional analysis,
- the theory of distribution,
- partial differential equations,

and so on. On the other hand the latter is connected to

- Lebesgue integral and various inequalities,
- functional analysis.

In the seminar on Feb. 14th, 2014 we talk about the latter because the first talking plan is to explain the latter and the latter is a little bit easier because the latter needs less knowledge.

22.1.2 Variational problem and its mathematical background

We give a lecture on a book [130] and a paper [133].

Here we give two examples of solving optimization (variational) problems. In general we consider optimization for a **functional** if we say ‘variational’.

The first comes from quantum mechanics and is the problem of determining the energy of an atom—primarily the lowest one. The second is a classical type minimization problem—the Thomas-Fermi problem—that arises in chemistry. In these two cases the difficult part is showing the existence of a minimizer, and hence of a solution to a partial differential equation. Needless to say, the following considerations (known as the **direct method in the calculus of variations**) for establishing a solution to a differential equation are not limited to these elementary examples, but should be viewed as a general strategy to attack optimization problems.

Historically, and even today in many places, it is customary to dispense with the question of existence as a mere subtlety. By simply assuming that a minimizer or maximizer exists, however, and then trying to derive properties for it, one can be led to severe inconsistencies—as the following amusing example taken from [253] and attributed to Perron shows:

Let N be the largest natural number. Since $N^2 \geq N$ and N is the largest natural number, $N^2 = N$ and hence $N = 1$.

What this example tells us is that even if the ‘variational equation’, here $N^2 = N$, can be solved explicitly, the resulting solution need not have anything to do with the problem we started out to solve.

Let us continue this overview with some general remarks about minimization of functions. A general theorem in analysis says that a bounded continuous real function f defined on a bounded and closed set K in \mathbb{R}^n attains its minimum value. To prove this, pick a sequence of points x_j such that

$$f(x_j) \rightarrow \lambda := \inf_{x \in K} f(x) \quad \text{as } j \rightarrow \infty. \quad (22.1.1)$$

Since K is compact there exists a subsequence, again denoted by x_j , and a point $x \in K$ such that $x_j \rightarrow x$ as $j \rightarrow \infty$. Hence, since f is continuous,

$$\lambda = \lim_{j \rightarrow \infty} f(x_j) = f(x), \quad (22.1.2)$$

and the minimum value is attained at x .

Instead of \mathbb{R}^n , consider now $L^2(\Omega, d\mu)$ and let $\mathcal{F}(\psi)$ be some (nonlinear) functional defined on this space. In many examples $\mathcal{F}(\psi)$ is strongly continuous, i.e., $\mathcal{F}(\psi_j) \rightarrow \mathcal{F}(\psi)$ as $j \rightarrow \infty$ whenever $\|\psi_j - \psi\|_2 \rightarrow 0$ as $j \rightarrow \infty$. Suppose we wish to show that the infimum of $\mathcal{F}(\psi)$ is attained on $K := \{\psi \in L^2(\Omega, d\mu) : \|\psi\|_2 \leq 1\}$. This set is certainly closed and bounded, but for a bounded sequence $\psi_j \in K$ there need not be a strongly convergent subsequence (see Sect. 2.9 in [130]).

The idea now is to relax the strength of convergence (topology). Indeed, if we use the notion of weak convergence instead of strong convergence, then, by Theorem 2.18 in [130], every sequence in K has a weakly convergent subsequence: see also Section 22.2.5. In this way, the set of convergent sequences has been enlarged—but a new problem arises. The functional $\mathcal{F}(\psi)$ need not be weakly continuous—and it rarely is. Thus, to summarize, the more sequences exist that have convergent subsequences the less likely it is that $\mathcal{F}(\psi)$ is continuous on these sequences. The way out of this apparent dilemma is that in many examples the functional turns out to be **weakly lower semicontinuous**, i.e.,

$$\liminf_{j \rightarrow \infty} \mathcal{F}(\psi_j) \geq \mathcal{F}(\psi) \quad \text{if } \psi_j \xrightarrow{w} \psi. \quad (22.1.3)$$

Thus, if ψ_j is a minimizing sequence, i.e., if

$$\mathcal{F}(\psi_j) \rightarrow \inf \{ \mathcal{F}(\psi) : \psi \in K \} = \lambda, \quad (22.1.4)$$

then there exists a subsequence ψ_j such that $\psi_j \xrightarrow{w} \psi$, and hence

$$\lambda = \lim_{j \rightarrow \infty} \mathcal{F}(\psi_j) \geq \mathcal{F}(\psi) \geq \lambda. \quad (22.1.5)$$

Therefore, $\mathcal{F}(\psi) = \lambda$, and the goal is achieved!

22.1.3 Schrödinger equation

The **time independent Schrödinger equation** for a particle in \mathbb{R}^n , interacting with a force field $F(x) = -\nabla V(x)$, is

$$-\Delta\psi + V\psi = E\psi. \quad (22.1.6)$$

The function $V: \mathbb{R}^n \rightarrow \mathbb{R}$ is called a potential¹. The ‘wave function’ ψ is a complex-valued function in $L^2(\mathbb{R}^n)$ subject to the **normalization condition**

$$\|\psi\|_2 = 1. \quad (22.1.7)$$

An L^2 solution to (22.1.6) may or may not exist for any E ; often does not. The special real numbers E for which such solutions exist are called **eigenvalues** and the solution, ψ , is called an **eigenfunction**. These have, of course, its origin in linear algebra: $H := -\Delta + V$ is a linear operator, and E ’s are its eigenvalues and ψ is an eigenfunction belonging to E .

Associated with (22.1.6) is a variational problem. Suppose ψ is some nice function (i.e., in $\mathcal{D}(\mathbb{R}^n) := C_c^\infty(\mathbb{R}^n)$) and multiplied $\bar{\psi}$ to (22.1.6):

$$-\bar{\psi}\Delta\psi + V|\psi|^2 = E|\psi|^2. \quad (22.1.8)$$

By integration by parts (Stokes’ theorem), we get

$$\int |\nabla\psi|^2 + \int V|\psi|^2 = E\|\psi\|_2^2. \quad (22.1.9)$$

Consider the following functional defined for a suitable class of functions in $L^2(\mathbb{R}^n)$ (to be specified later):

$$\mathcal{E}(\psi) = T_\psi + V_\psi, \quad (22.1.10)$$

with

$$T_\psi = \int_{\mathbb{R}^n} |\nabla\psi(x)|^2 dx, \quad (22.1.11)$$

$$V_\psi = \int_{\mathbb{R}^n} V(x) |\psi(x)|^2 dx. \quad (22.1.12)$$

Physically T_ψ is called the **kinetic energy** of ψ , V_ψ is its **potential energy** and $\mathcal{E}(\psi)$ is the **total energy** of ψ .

¹ただし三角の微分があってそこは高校 3 年理系相当だから世間的にはあまり馴染みもないだろうとも思っている。

The variational problem we shall consider is to minimize $\mathcal{E}(\psi)$ subject to the constant $\|\psi\|_2 = 1$. This relates to the fundamental physical inspection: *nature favors the states having lower energy*. Note that we change the problem to easier one, i.e., differential equation to integro-differential equation. Integral equation is easier because we can use many inequalities. Furthermore integral does conserve the order of functions, i.e.,

$$f \leq g \implies \int f \leq \int g. \quad (22.1.13)$$

This property is essential in the following analysis.

As we shall show below, a minimizing function ψ_0 , if one exists, will satisfy equation (22.1.6) with $E = E_0$, where

$$E_0 := \inf \left\{ \mathcal{E}(\psi) : \int |\psi|^2 = 1 \right\}. \quad (22.1.14)$$

Such a function ψ_0 will be called a **ground state**. The number E_0 is called the **ground state energy**.

Thus the variational problem determines not only ψ_0 but also a corresponding eigenvalue E_0 , which is the smallest eigenvalue of (22.1.6).

Our route to finding a solution to (22.1.6) takes us to the main problem.

1. Show, under suitable assumptions on V , that a minimizer exists.
 1. Note that there are examples where a minimizer does **not** exist, e.g., take $V = 0$.
2. Show that the minimizers satisfy (22.1.6) in the sense of distributions.
3. Show that, under suitable additional assumptions on V , the distributional solutions of (22.1.6) are sufficiently regular to yield classical solutions.
4. Show that (non-) uniqueness of the minimizer.
 1. Here uniqueness means that it is unique apart from an overall phase, i.e., $\psi_0 \rightarrow e^{i\theta} \psi_0$.

Considering distributional solutions is usual procedure in mathematics. If we want to solve algebraic equations we first consider solutions in \mathbb{C} for existence. Then we consider properties of solutions. Considering distributional solutions amounts to considering in \mathbb{C} .

22.1.4 Connection to other mathematical branches

Mathematics for quantum mechanics

- Time dependent Schrödinger equation
A time dependent Schrödinger equation is

$$i \frac{\partial}{\partial t} \phi = H \psi, \quad (22.1.15)$$

where H is a Hamiltonian, i.e., selfadjoint operator. There are many PDE analysis including scattering theory, etc.

- Operator theory
In operator theoretic language, our main task in this seminar is a proof that

1. the **Hamiltonian** $H = -\Delta + V$ is bounded below,
2. the infimum of the spectrum $\sigma(H)$ of H is in the point spectrum (the set of eigenvalues).

Note that, in operator theory, the ground state does not appear explicitly. Remember the theory of eigenvalues and eigenvectors in linear algebra: eigenvectors are defined after the definition of eigenvalues.

Main concern of operator theory is, of course, the properties of operators, not the properties of functions. Hence operator theoretic point of view is different from real analytic one.

This originates in the physics of quantum mechanics: in quantum mechanics the spectrum of an operator is the set of measured values, i.e., the spectrum of an operator has direct physical meaning. Since my main concern is physics my main mathematical tool is operator theory or operator algebras.

Today's analysis is one for an operator by real analytic methods.

- Probability

We can use probability theory for operator analysis. Usually we can only see properties before or after action of an operator. However, using probability, we can see some property 'in course of action'. For example, by heat semigroup technique and Feynman-Kac formula, we get a formula for ground state energy, **functional integral representation**,

$$E_0 = - \lim_{t \rightarrow \infty} \frac{1}{t} \log \int_W \overline{\psi(\phi(0))} \psi(\phi(t)) e^{-\int_0^t V(\phi(s)) ds} d\mu_0, \quad (22.1.16)$$

where

$$W := \mathbb{R}^n \times M, \quad M := (\mathbb{R}^n)^{[0, \infty)}, \quad (22.1.17)$$

$$\phi(t): \mathbb{R}^n \times M \rightarrow \mathbb{R}^n; \quad \phi(t)(x, \omega) := x + \omega(t), \quad (x, \omega) \in W \quad \forall t \geq 0. \quad (22.1.18)$$

Or we can express the heat semigroup action by path integral. For the Nelson model we obtain

$$\langle F, e^{-tH_N} G \rangle = \int_{\mathbb{R}^d} dx E^x \left[e^{-\int_0^t V(B_s) ds} \langle I_0 \overline{F(B_0)}, e^{\phi_E(\int_0^t \delta_s \otimes \varphi(\cdot - B_s) ds)} \rangle I_t G(B_t) \right] \quad (22.1.19)$$

See [137], [33], [38] for details.

- Algebraic analysis

I do not know this area well. A Keyword is the exact WKB analysis. See, for example, [110].

- Nonstandard analysis

I do not know also this area well. See, for example, [150]. This book explains, e.g., Feynman's path integral. Nonstandard analysis relates to the domain of mathematical logic: quantum physics connects mathematical logic to other mathematical branches!

- Number theory: noncommutative harmonic oscillator

I do not know also this area well. See, for example, [173, 237].

We define a non-commutative harmonic oscillator. Consider a Hilbert space $L^2\{\mathbb{R}\} \otimes \mathbb{C}$ and a differential operator of matrix type $Q = Q_{\alpha, \beta}$:

$$Q := \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \left(-\frac{1}{2} \frac{d^2}{dx^2} + \frac{1}{2} x^2 \right) + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \left(x \frac{d}{dx} + \frac{1}{2} \right), \quad (22.1.20)$$

where $\alpha, \beta > 0$, $\alpha\beta > 1$. Then the operator Q is of strictly positive operator and has only discrete spectrum. For investigating the spectral structure we define the following Dirichlet series (spectral zeta).

$$\zeta_Q(s) := \sum_{n=1}^{\infty} \lambda_n^{-s}. \quad (22.1.21)$$

This series absolutely converges in $\Re(s) > 1$ and is analytically continued to all \mathbb{C} as a meromorphic function. In particular, if $\alpha = \beta = \sqrt{2}$ then we have

$$\zeta_Q(s) = 2(2^s - 1)\zeta(s), \quad (22.1.22)$$

i.e., this spectral zeta is Riemann's zeta.

This relates to many other mathematical branches, e.g., a differential equation of Picard-Fuchs type for a universal family of an elliptic curve, theory of automorphic forms and functions, Heun's differential equation and its connection problem. Furthermore Taniguchi shows some interesting property using stochastic analysis. For details see [173].

- In connection to topology: Aharonov-Bohm effect

We may consider Schrödinger operators on topologically defected regions, especially, non-simply connected regions. Topological properties on the domain give some effects on the properties of the operator. In this respect Aharonov-Bohm effect is very famous and important in physics or gauge theory. See, e.g., chapter 3 in [30] and its references.

Geometric variational problem

A variational method is a basic tool in many physical branches. Analytical mechanics and its relation to (differential) geometry is famous. Variational argument for energy of a curve introduces connection, covariant derivative, parallel displacement, geodesic, curvature and so on. We can generalize the theory of energy for curves to the one for energy for maps between (Riemann/Kähler) manifolds, i.e., theory of harmonic maps. Geometric variational problem also relates to Morse theory. See [157]. See also the isoperimetric problem.

Nonlinear analysis

In the analysis for the Thomas-Fermi theory we face nonlinear analysis. In physics there are many nonlinear partial differential equations and we have many analysis for these equations. In geometry we have many interesting nonlinear functionals and nonlinear partial differential equations, such as the Yamabe equation or the Monge-Ampère equation. See [37].

Index theorem

In Aharonov-Bohm effect the topology of spaces affects operators on it. E.g., if a Riemannian manifold is compact the spectrum of Laplacian becomes discrete².

Conversely properties of some operators determines some interesting properties in topology. A Laplacian on a Riemannian manifold, especially its spectrum,

²数学ではよくある.

has many topological information. This type of problems offers the region, spectral geometry [96].

In this course one of the most interesting thing is the celebrated index theorem by Atiyah and Singer. The Atiyah–Singer index theorem, proved by Michael Atiyah and Isadore Singer (1963), states that for an elliptic differential operator on a compact manifold, the analytical index (related to the dimension of the space of solutions) is equal to the topological index (defined in terms of some topological data). It includes many other theorems, such as the Riemann–Roch theorem, as special cases, and has applications in theoretical physics.

Path integral (functional integral) has a deep connection to this problem. We have a (formal?) proof for index theorem using path integral.

Stochastic optimization

Stochastic optimization is the process of maximizing or minimizing the value of a mathematical or statistical function when one or more of the input parameters is subject to randomness. The word stochastic means involving chance or probability. This is applied to, e.g., machine learning. This is used to improve a user’s dictionary in converting “kanji-kana”.

Other optimization problem

In real world we encounter many optimization problems. In general engineering solves optimization problems in real constraints.

Singular analysis

The Coulomb potential has a singular point at $x = 0$ and a solution has a singularity at $x = 0$. That is, singularity of a potential affects a singularity of a solution.

Both physics and mathematics have many singular problems, such as resolution of singularity in algebraic geometry, the residue theorem in complex analysis, existence of black holes in general relativity, definition of phase transition, and so on.

22.2 Mathematical preliminaries

In this section we assume $\Omega \subset \mathbb{R}^n$ is an open set.

22.2.1 The important difference in integration and differentiation

In analysis integration is very important because it saves an order of functions: we define $f \leq g$ by

$$f(x) \leq g(x) \quad \forall x. \quad (22.2.1)$$

Then it follows that

$$f \leq g \Rightarrow \int f \leq \int g. \quad (22.2.2)$$

Furthermore, if f and g is continuous, $f \leq g$, and there is x such that $f(x) < g(x)$, then we get

$$\int f < \int g. \quad (22.2.3)$$

This is called the strong monotonicity of integration.

Can you make counterexamples for differentiation?

22.2.2 Important spaces

In this seminar we mainly consider L^p spaces and Sobolev spaces H^k on \mathbb{R}^n .

L^p spaces

Here we consider only Lebesgue or Borel measurable functions. For details see [90] or [130].

For $p \in [1, \infty)$, L^p spaces is defined by

$$L^p(\Omega) := \left\{ f : \int_{\Omega} |f(x)|^p dx < \infty \right\}. \quad (22.2.4)$$

For $p = \infty$

$$L^p = \{f : \text{ess.sup } |f| < \infty\}. \quad (22.2.5)$$

We introduce the topology on L^p by the norms:

$$\|f\|_p = \begin{cases} \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p}, & 1 \leq p < \infty, \\ \text{ess.sup } |f|, & p = \infty. \end{cases} \quad (22.2.6)$$

Then, by the Minkowski inequality, L^p spaces are linear for $p \in [1, \infty)$. L^∞ is also linear. In fact L^p spaces are complete for $p \in [1, \infty]$. For details see [90] or [130].

Moreover, for $p = 2$, we can introduce the inner product:

$$\langle f, g \rangle := \int_{\Omega} \overline{f(x)} g(x) dx. \quad (22.2.7)$$

For this inner product the space L^2 becomes a Hilbert space.

$L^p + L^q$ spaces

Let $p, q \in [1, \infty]$. Then a function f is in $L^p + L^q$ if f can be decomposed as

$$f = f_1 + f_2, \quad f_1 \in L^p, \quad f_2 \in L^q. \quad (22.2.8)$$

This decomposition is simple but ingenious.

Example. 2.2.1. Let $V = -\frac{1}{|x|}$ in \mathbb{R}^3 and

$$V = V_1 + V_2, \quad V_1 := -\chi_{|x| \leq 1} V, \quad V_2 := -\chi_{|x| > 1} V, \quad . \quad (22.2.9)$$

Then we have

$$V_1 \in L^{5/2}, \quad V_2 \in L^4 \cap L^\infty. \quad (22.2.10)$$

the L^1_{loc} space

The space $L^1_{\text{loc}}(\Omega)$ is defined by

$$L^1_{\text{loc}}(\Omega) := \{f : \forall A \subset \Omega, A \text{ is compact, } \|\chi_A f\|_1 < \infty\}, \quad (22.2.11)$$

where a function χ_A is an indicator function for $A \subset \Omega$. This space is useful when considering the theory of distributions.

Sobolev spaces H^k

A Sobolev space $H^1(\mathbb{R}^n)$ is defined by

$$H^1(\mathbb{R}^n) := \left\{ \psi \in L^2(\mathbb{R}^n) : \int_{\mathbb{R}^n} (|\nabla \psi(x)|^2 + |\psi(x)|^2) dx < \infty \right\}, \quad (22.2.12)$$

where $\nabla \psi$ is in the distributional sense. We have another definition of a Sobolev space $H^1(\mathbb{R}^n)$ using Fourier transform,

$$H^1(\mathbb{R}^n) := \left\{ \psi \in L^2(\mathbb{R}^n) : \int_{\mathbb{R}^n} \left(|k \hat{\psi}(k)|^2 + |\hat{\psi}(k)|^2 \right) dk < \infty \right\}. \quad (22.2.13)$$

22.2.3 Convolution

Definition

Here we consider $\Omega = \mathbb{R}^d$. For suitable functions f and g we can define a some type of product, called the convolution, by

$$(f * g)(x) := \int_{\mathbb{R}^d} f(y)g(x-y)dy. \quad (22.2.14)$$

For this product the space $L^1(\mathbb{R}^d)$ becomes a Banach algebra. There exists a useful inequality, Young's inequality: see Theorem 2.4.2.

Convolutions of functions in dual L^p spaces are continuous

Lemma. 2.3.1. (Lemma 2.20. [130]) *Let $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$ with $p, q > 1$ dual indices, i.e.,*

$$\frac{1}{p} + \frac{1}{q} = 1. \quad (22.2.15)$$

*Then the convolution $f * g$ is a continuous function on \mathbb{R}^n that tends to zero at infinity in the strong sense that, for any $\varepsilon > 0$ there is \mathcal{R}_ε such that*

$$\sup_{|x| > \mathcal{R}_\varepsilon} |(f * g)(x)| < \varepsilon. \quad (22.2.16)$$

22.2.4 Various inequalities

Hölder's inequality

Theorem. 2.4.1. (Theorem 2.3 [130].) *Let p and q be dual indices, i.e., $1/p + 1/q = 1$ with $1 \leq p \leq \infty$. Let $\Omega \subset \mathbb{R}^n$ be open, $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$. Then $fg \in L^1(\Omega)$ and*

$$\left| \int_{\Omega} fg dx \right| \leq \|f\|_p \|g\|_q. \quad (22.2.17)$$

This means that

$$L^p(\Omega) \subset L^q(\Omega)', \quad p \in [1, \infty], \quad (22.2.18)$$

where, for a linear space L , L' is a dual space. In fact we can prove

$$L^q(\Omega) = (L^p(\Omega))', \quad L^p(\Omega) = (L^q(\Omega))' = (L^p(\Omega))'', \quad p \in (1, \infty), \quad (22.2.19)$$

i.e., L^p spaces are **reflexive** for $p \in (1, \infty)$.

It is well-known L^∞ is not reflexive. See, e.g., [46, 45].

Young's inequality

Theorem. 2.4.2. (Theorem 4.2 [130].) *Let $\Omega = \mathbb{R}^n$.*

1. *Assume that*

$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \quad 1 \leq p, q, r \leq \infty. \quad (22.2.20)$$

Then, for any $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, it holds that

$$\|f * g\|_r \leq \|f\|_p \|g\|_q. \quad (22.2.21)$$

2. *Assume that*

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2, \quad 1 \leq p, q, r \leq \infty. \quad (22.2.22)$$

Then, for any $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, and $h \in L^r(\mathbb{R}^n)$ it holds that

$$\left| \int_{\mathbb{R}^n} f(x) (g * h)(x) dx \right| \leq \|f\|_p \|g\|_q \|h\|_r. \quad (22.2.23)$$

Sobolev's inequality

A function $f: \mathbb{R}^n \rightarrow \mathbb{C}$ is in $D^1(\mathbb{R}^n)$ if it is in $L^1_{\text{loc}}(\mathbb{R}^n)$, if its distributional derivative $\nabla f \in L^2(\mathbb{R}^n)$, and if f vanishes at infinity, i.e.,

$$|\{x : |f(x)| > a\}| < \infty, \quad \forall a > 0. \quad (22.2.24)$$

Theorem. 2.4.3. (Sobolev's inequality for $n \geq 3$, Section 8.3 in [130].) *For $n \geq 3$ let $f \in D^1(\mathbb{R}^n)$. Then $f \in L^q(\mathbb{R}^n)$ with $q = 2n/(n-2)$ and the following inequality holds:*

$$S_n \|f\|_q^2 \leq \|\nabla f\|_2^2, \quad (22.2.25)$$

where

$$S_n = \frac{n(n-2)}{4} |\mathbb{S}^2|^{2/n} = \frac{n(n-2)}{4} 2^{2/n} \pi^{1+1/n} \Gamma\left(\frac{n+1}{2}\right)^{-2/n}. \quad (22.2.26)$$

22.2.5 Topologies

Strong topology and weak topology in Hilbert spaces

Here we explain only Hilbert spaces since it suffices in this seminar.

Suppose \mathcal{H} is a (separable) Hilbert space, i.e., \mathcal{H} has a complete inner product space. The space \mathcal{H} has a natural topology, called the **strong topology**, defined by its norm. We may consider other topology for \mathcal{H} , called the **weak topology**. It is defined by the convergence for sequences. The strong topology is defined by convergence in norm, i.e.,

$$\text{s-lim}_{j \rightarrow \infty} \psi_j = \psi \iff \|\psi_j - \psi\| \rightarrow 0 \text{ as } j \rightarrow \infty. \quad (22.2.27)$$

The weak topology is defined by some convergence using inner product, i.e.,

$$\text{w-lim}_{j \rightarrow \infty} \psi_j = \psi \iff \langle \psi_j - \psi, \phi \rangle \rightarrow 0 \text{ as } j \rightarrow \infty \quad \forall \phi \in \mathcal{H}. \quad (22.2.28)$$

Proposition. 2.5.1. *If $\psi_j \xrightarrow{s} \psi$ then $\psi_j \xrightarrow{w} \psi$.*

Proof. By Cauchy-Schwarz. □

Remark. 2.5.2. *Suppose a functional $F: \mathcal{H} \rightarrow \mathbb{C}$ is strongly continuous. However F is not necessarily weakly continuous.*

Example of weakly convergent sequence

Before introducing concrete examples, we state a general example.

Theorem. 2.5.3. (Inequality of Bessel, Sec. 2.21 [130].) *If ψ_j is orthonormal system, then we have, for any $\phi \in \mathcal{H}$,*

$$\sum_{j=1}^{\infty} |\langle \psi_j, \phi \rangle|^2 \leq \|\phi\|^2. \quad (22.2.29)$$

Suppose a sequence (ψ_j) is orthonormal. Then the Bessel inequality states that (ψ_j) converges weakly to 0. However (ψ_j) does not converges strongly because $\|\psi_j - \psi_k\| = \sqrt{2}$ if $j \neq k$.

A concrete example is as follows. A function ψ_j is defined by $\psi_j = \chi_{A_j}$, where $A_j = (j-1, j]$ and χ_A is a indicator function for a set A . Then a sequence (ψ_j) is orthonormal and $\psi_j \xrightarrow{w} 0$. This (ψ_j) also converges to 0 pointwise but (ψ_j) does not converge strongly. The weak topology fills in this discrepancy.

Why we use the pointwise convergence instead of weak convergence?

A function space has many interesting and important convergence, e.g., pointwise convergence and uniform convergence. The weak topology is natural if we consider its linear space structure. A (normed) linear space always has its dual space³ and the weak topology uses a dual space in its definition. Furthermore the importance of the weak topology is also due to the theorem 2.6.3.

³これがまた分かりづらい。29.4.4 節でこれをすっきりさせる。

22.2.6 Theorems in Lebesgue integral and functional analysis

Dominated convergence

Theorem. 2.6.1. (Theorem 1.8. [130]) *Assume functions (f_j) and f are measurable and $f_j \rightarrow f$ pointwise. If a function $g \in L^1(\Omega)$ exists and $|f_j| \leq g$ then $|f| \leq g$ and*

$$\lim_{j \rightarrow \infty} \int_{\Omega} f_j(x) dx = \int_{\Omega} f(x) dx. \quad (22.2.30)$$

Differentiability of norms, Gateaux derivative

Theorem. 2.6.2. (Theorem 2.6. [130]) *Suppose f and g are functions in $L^p(\Omega)$ with $p \in (1, \infty)$. The function defined on \mathbb{R} by*

$$N(t) := \int_{\Omega} |f(x) + tg(x)|^p dx \quad (22.2.31)$$

is differentiable and its derivative at $t = 0$ is given by

$$\frac{d}{dt} N(t) \Big|_{t=0} = p \int_{\Omega} |f(x)|^{p-2} \Re \left(\overline{f(x)} g(x) \right) dx. \quad (22.2.32)$$

Uniform boundedness principle

Theorem. 2.6.3. (Theorem 2.12. [130]) *Let $f_j \in L^p(\Omega)$ with the property*

$$\{L(f_j) : j \in \mathbb{N}\} < \infty \quad \forall L \in L^p(\Omega)^*. \quad (22.2.33)$$

Then the sequence of norms $(\|f_j\|_p)$ is bounded.

Banach-Alaoglu theorem

Theorem. 2.6.4. (Theorem 2.18. [130])

1. *Strongly Bounded and strongly closed sets of a Hilbert space \mathcal{H} are weakly compact.*
2. *Strongly Bounded and strongly closed sets of each $L^p(\Omega)$ space are weakly compact.*

Remark. 2.6.5. *This theorem implies that a **strong** bounded sequence in $L^p(\Omega)$ has a **weakly** convergent subsequence! Note also that this is the extremely important property for the analysis on \mathbb{R} .*

Derivative of the absolute value

Theorem. 2.6.6. (Theorem 6.17. [130]) *Let f be in $W^{1,p}(\Omega)$. Then the function $|f|$, defined by $|f|(x) := |f(x)|$, is in $W^{1,p}(\Omega)$ with $\nabla |f|$ being the function*

$$(\nabla |f|)(x) = \begin{cases} \frac{1}{|f|(x)} (R(x) \nabla R(x) + I(x) \nabla I(x)), & f(x) \neq 0, \\ 0 & f(x) = 0, \end{cases} \quad (22.2.34)$$

here $R(x)$ and $I(x)$ denote the real and imaginary parts of f . In particular, if f is real-valued,

$$(\nabla |f|)(x) = \begin{cases} \nabla f(x) & f(x) > 0, \\ -\nabla f(x) & f(x) < 0, \\ 0 & f(x) = 0. \end{cases} \quad (22.2.35)$$

Thus

1. $|\nabla |f|| \leq |\nabla f|$ a.e. if f is complex-valued,
2. $|\nabla |f|| = |\nabla f|$ a.e. if f is real-valued.

Convexity inequality for gradients

Theorem. 2.6.7. (Theorem 7.8. [130]) Let f, g be real-valued functions in $H^1(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n} \left| \nabla \sqrt{f^2 + g^2} \right|^2(x) dx \leq \int_{\mathbb{R}^n} \left(|\nabla f(x)|^2 + |\nabla g(x)|^2 \right) dx. \quad (22.2.36)$$

If, moreover, $g(x) > 0$, then equality holds if and only if there exists a constant c such that

$$f(x) = cg(x) \quad \text{a.e.} \quad (22.2.37)$$

Remark. 2.6.8. For complex-valued F the above inequality becomes

$$\int_{\mathbb{R}^n} |\nabla |F|(x)|^2 dx \leq \int_{\mathbb{R}^n} |\nabla F(x)|^2 dx. \quad (22.2.38)$$

Fundamental theorem of variational method, du Bois-Reymond's lemma

Theorem. 2.6.9. Let $f \in L^1_{\text{loc}}$. If, for any $\phi \in C_c^\infty(\Omega)$, it holds that

$$\int_{\Omega} f(x)\phi(x)dx = 0, \quad (22.2.39)$$

then we must have $f = 0$.

22.2.7 Potential theory

We use this theory in Section 22.3.1. See the chapter 9 in [130] for details.

Positivity properties of the Coulomb energy

We define a sesquilinear form D as

$$D(f, g) := \frac{1}{2} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{f(x)g(x)}{|x - y|} dx dy. \quad (22.2.40)$$

Then we have the

Theorem. 2.7.1. (Theorem 9.8. [130]) *If $f: \mathbb{R}^n \rightarrow \mathbb{C}$ satisfies $D(|f|, |f|) < \infty$, then*

$$D(f, f) \geq 0. \quad (22.2.41)$$

If g satisfies the same property as f , then we have Schwartz' inequality for D , i.e.,

$$|D(f, g)|^2 \leq D(f, f)D(g, g). \quad (22.2.42)$$

Furthermore $f \mapsto D(f, f)$ is strictly convex.

In [130] more properties for the Coulomb potential are explained.

22.3 Thomas-Fermi theory

22.3.1 Thomas-Fermi problem

Outline

We turn now to our second example of a variational problem—the Thomas-Fermi (TF) problem. See [133] for details and further developments. It goes back to the idea of L. H. Thomas and E. Fermi in 1926 that a large atom, with many electrons, can be approximately modeled by a simple **nonlinear** problem for a ‘charge density’ $\rho(x)$. We do not derive the energy functional. See a textbook.

Here we state the mathematical problem. The potential function $Z/|x|$ that appears in the following can easily be replaced by

$$V(x) := \sum_{j=1}^K \frac{Z_j}{|x - R_j|} \quad (22.3.1)$$

with $Z_j > 0$ and $R_j \in \mathbb{R}^3$, but we refrain from doing so in the interest of simplicity.

In [130] the authors leave many steps as an exercise for the reader. If you want details you should read the original papers, e.g., [133].

TF theory is defined by an energy function \mathcal{E} on a certain class of **nonnegative** functions ρ on \mathbb{R}^3 :

$$\mathcal{E}(\rho) := \frac{3}{5} \int_{\mathbb{R}^3} \rho(x)^{5/3} dx - \int_{\mathbb{R}^3} \frac{Z}{|x|} \rho(x) dx + D(\rho, \rho), \quad (22.3.2)$$

where $Z > 0$ is a fixed parameter (the charge of the atom's nucleus) and

$$D(\rho, \rho) := \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \rho(x) \rho(y) \frac{1}{|x - y|} dx dy = \frac{1}{2} \int_{\mathbb{R}^3} \rho(x) \left(\rho * \frac{1}{|x|} \right) (x) dx \quad (22.3.3)$$

is the Coulomb energy of a charge density⁴. The class of admissible functions is

$$\mathcal{C} = \left\{ \rho : \rho \geq 0, \rho \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3) \right\}. \quad (22.3.4)$$

⁴指数関数 2^n が前にかかっているのに発散せずに x という有限確定値に収束するのはちょっとした驚きだ。

Our problem is to minimize $\mathcal{E}(\rho)$ under the condition $\int \rho = N$, where N is any fixed positive number (identified as the ‘number’ of electrons in the atom). The case $N = Z$ is special and is called the **neutral case**. We define two subsets of \mathcal{C} :

$$\mathcal{C}_N := \mathcal{C} \cap \left\{ \rho : \int_{\mathbb{R}^3} \rho = N \right\} \subset \mathcal{C}_{\leq N} := \mathcal{C} \cap \left\{ \rho : \int_{\mathbb{R}^3} \rho \leq N \right\}. \quad (22.3.5)$$

Corresponding to these two sets are two energies: The ‘constrained’ energy

$$E(N) := \inf \{ \mathcal{E}(\rho) : \rho \in \mathcal{C}_N \}, \quad (22.3.6)$$

and the ‘unconstrained’ energy

$$E_{\leq}(N) := \inf \{ \mathcal{E}(\rho) : \rho \in \mathcal{C}_{\leq N} \}. \quad (22.3.7)$$

Obviously it holds that

$$E_{\leq}(N) \leq E(N). \quad (22.3.8)$$

The reason for introducing the unconstrained problem will become clear later. A minimizer will not exist for the constrained problem when $N > Z$: atoms cannot be negatively charged in TF theory. But a minimizer will always exist for the unconstrained problem. It is often advantageous, in variational problems, to relax a problem in order to get at a minimizer; in fact, we already used this device in the study of the Schrödinger equation⁵. When a minimizer for the constrained problem does exist it will later be seen to be the ρ that is a minimizer for the unconstrained problem.

Existence of an unconstrained Thomas-Fermi minimizer

The main theorem in this section is as follows.

Theorem. 3.1.1. (Theorem 11.12. [130]) *For each $N > 0$ there is a unique minimizing $\rho_{N,\infty}$ for the unconstrained TF problem (22.3.7), i.e.,*

$$\mathcal{E}(\rho_{N,\infty}) = E_{\leq}(N). \quad (22.3.9)$$

The constrained energy $E(N)$ and the unconstrained energy $E_{\leq}(N)$ are equal. Moreover $E(N)$ is a convex and nonincreasing function of N .

The proof is divided into several steps.

Lemma. 3.1.2. (Theorem II.2. (c) [133]) *The expression (22.3.2) is well-defined and finite when $\rho \in \mathcal{C}$. The functional $\mathcal{E}(\rho)$ is bounded from below on each $\mathcal{C}_{\leq N}$.*

Proof. Write

$$V(x) = -\frac{1}{|x|} = V_1 + V_2, \quad V_1 \in L^{5/2}, \quad V_2 \in L^\infty, \quad (22.3.10)$$

⁵正確にはその内の一番簡単な例.

e.g.,

$$V_1 = \chi_{|x| \leq 1} V, \quad V_2 = \chi_{|x| > 1} V. \quad (22.3.11)$$

Then, by Hölder's inequality we obtain

$$\int |\rho V| dx \leq \|V_1\|_{5/2} \|\rho\|_{5/3} + \|V_2\|_{\infty} \|\rho\|_1, \quad (22.3.12)$$

and Young's inequality shows that

$$D(\rho, \rho) = \frac{1}{2} \int \left(\rho * \frac{1}{|x|} \right) \rho dx \leq \frac{1}{2} \|\rho\|_1 \left(\|V_1\|_{5/2} \|\rho\|_{5/3} + \|V_2\|_{\infty} \|\rho\|_1 \right), \quad (22.3.13)$$

positive definiteness of $1/|x|$ (Theorem 2.7.1) implies

$$D(\rho, \rho) \geq 0, \quad (22.3.14)$$

and, of course,

$$\int \rho^{5/3} dx = \|\rho\|_{5/3}^{5/3}. \quad (22.3.15)$$

Summing up we get

$$E_{\leq}(N) \geq \frac{3}{5} \|\rho\|_{5/3}^{5/3} - c_1 \|\rho\|_{5/3} - c_2 \quad \text{on each } \mathcal{C}_{\leq N}, \quad (22.3.16)$$

where the number c_1 is N independent ($\|V_1\|_{5/2}$) and c_2 is N dependent ($ZN \|V_2\|_{\infty}$). Since $x^{5/3} - cx$ is bounded from below on $[0, \infty)$, the desired properties hold. \square

Let (ρ_j) be a minimizing sequence for $\mathcal{E}(\rho)$.

Lemma. 3.1.3. (Theorem II.2. (d) [133]) *The sequence of numbers $(\|\rho_j\|_{5/3})$ is also bounded.*

Proof. The inequality (22.3.16) and the fact that $\{x : 0 \leq x, x^{5/3} - cx \leq d\}$ is bounded for each fixed c, d . \square

Therefore, by passing to a subsequence we can assume that $\rho_j \xrightarrow{w} \rho_{N,\infty}$ in $L^{5/3}(\mathbb{R}^3)$ for some $\rho_{N,\infty} \in L^{5/3}(\mathbb{R}^n)$. Here we use the Banach-Alaoglu theorem, Theorem 2.6.4.

Lemma. 3.1.4. (Theorem II.13. [133]) *The functional $\mathcal{E}(\rho)$ is lower semicontinuous on each $\mathcal{C}_{\leq N}$ in the weak $L^{5/3}(\mathbb{R}^3)$ topology, i.e., if $\rho_j \xrightarrow{w} \rho_{N,\infty}$ in weak $L^{5/3}$ with $\sup_j \|\rho_j\|_1 < \infty$, $\rho \in L^1$, and $\rho_j, \rho_{N,\infty} \geq 0$, then*

$$\mathcal{E}(\rho_{N,\infty}) \leq \liminf \mathcal{E}(\rho_j). \quad (22.3.17)$$

Moreover, if $\mathcal{E}(\rho_{N,\infty}) = \lim \mathcal{E}(\rho_j)$, then $\|\rho_j - \rho_{N,\infty}\|_{5/3} \rightarrow 0$ and each term in $\mathcal{E}(\rho_j)$ converges to the corresponding term in $\mathcal{E}(\rho_{N,\infty})$.

Proof. By passing to a subsequence, we can suppose that $\lim \mathcal{E}(\rho_j)$ exists and we may as well suppose the limit is finite since (22.3.17) is trivial otherwise. Then by Lemma 3.1.3 it follows that $\sup_j \|\rho_j\|_{5/3} < \infty$.

First claim for (22.3.17):

$$\|\rho_{N,\infty}\|_{5/3} \leq \liminf \|\rho_j\|_{5/3}. \quad (22.3.18)$$

This is due to the lower semicontinuity of norms. For explicit proof, noting that $\rho_{N,\infty}^{2/3} \in L^{5/2}$, the definition of weak convergence and Hölder's inequality leads

$$\int \rho_{N,\infty}^{5/3} dx = \lim_{j \rightarrow \infty} \int \rho_j \rho_{N,\infty}^{2/3} dx \leq \|\rho_{N,\infty}\|_{5/3}^{2/3} \liminf \|\rho_j\|_{5/3}. \quad (22.3.19)$$

Second claim for (22.3.17):

$$\lim \int \rho_j V = \int \rho_{N,\infty} V. \quad (22.3.20)$$

Write $V = V_1 + V_2$ with $V_1 \in L^{5/2}$ and $V_2 \in L^p$ ($5/2 < p < \infty$). Hölder's inequality shows that

$$\int V_1 \rho_j \rightarrow \int V_1 \rho_{N,\infty}. \quad (22.3.21)$$

The following Lemma 3.1.5 implies $\rho_j \xrightarrow{w} \rho_{N,\infty}$ in weak L^q ($1 < q \leq 5/3$) sense. Hence it holds that

$$\int V_2 \rho_j \rightarrow \int V_2 \rho_{N,\infty}. \quad (22.3.22)$$

Third claim for (22.3.17):

$$D(\rho_{N,\infty}, \rho_{N,\infty}) \leq \liminf D(\rho_j, \rho_j). \quad (22.3.23)$$

We use positive definiteness of $|x|^{-1}$, Theorem 2.7.1 and the resulting Schwartz inequality. Since $\rho \in L^1$ and $|x|^{-1} \in L^{5/2} + L^4$, Young's inequality shows that $\rho * |x|^{-1} \in L^{5/2} + L^4$. Since the Lemma 3.1.5 leads that $\rho_j \xrightarrow{w} \rho_{N,\infty}$ both in $L^{5/3}$ and $L^{4/3}$, we obtain

$$D(\rho_{N,\infty}, \rho_j) = \int \left(\rho_{N,\infty} * \frac{1}{|x|} \right) \rho_j dx \quad (22.3.24)$$

$$\rightarrow \int \left(\rho_{N,\infty} * \frac{1}{|x|} \right) \rho_{N,\infty} dx = D(\rho_{N,\infty}, \rho_{N,\infty}). \quad (22.3.25)$$

Therefore

$$D(\rho_{N,\infty}, \rho_{N,\infty}) = \lim_{j \rightarrow \infty} D(\rho_{N,\infty}, \rho_j) \quad (22.3.26)$$

$$\leq \liminf \left(D(\rho_{N,\infty}, \rho_{N,\infty})^{1/2} D(\rho_j, \rho_j)^{1/2} \right), \quad (22.3.27)$$

where we use Schwartz' inequality in the final step.

The above three estimates imply that if $\mathcal{E}(\rho_{N,\infty}) = \lim \mathcal{E}(\rho_j)$, then each term of $\mathcal{E}(\rho_{N,\infty})$ converges and in particular $\lim \|\rho_j\|_{5/3} = \|\rho_{N,\infty}\|_{5/3}$. Since $L^{5/3}$ is uniformly convex, convergence of norms and weak convergence implies norm convergence. \square

Here we show the Lemma used in the above Lemma.

Lemma. 3.1.5. *If a sequence of functions $(f_j) \subset L^1 \cap L^{5/3}$ weakly converges to f in $L^{5/3}$ and is bounded in $L^{5/3}$, then there exists a subsequence (f_{j_k}) such that*

$$f_{j_k} \xrightarrow{w} f \quad \text{in } L^q, \quad (22.3.28)$$

where $q \in (1, 5/3]$.

Proof. We assume $q \in (1, 5/3)$. Set a number α as follows:

$$\frac{1}{q} = \alpha + (1 - \alpha)\frac{3}{5}, \quad \text{i.e.,} \quad \alpha = \frac{5}{2} \left(\frac{1}{q} - \frac{3}{5} \right) \in (0, 1). \quad (22.3.29)$$

Note that

$$q\alpha < 1, \quad q(1 - \alpha)\frac{3}{5} < 1, \quad (22.3.30)$$

and that

$$|f_j|^q = |f_j|^{q\alpha} |f_j|^{q(1-\alpha)}, \quad |f_j|^{q\alpha} \in L^{\frac{1}{q\alpha}}, \quad |f_j|^{q(1-\alpha)} \in L^{\frac{5}{3(1-\alpha)q}}. \quad (22.3.31)$$

Then Hölder's inequality implies that

$$\int |f_j|^q \leq \int |f_j|^{q\alpha} |f_j|^{q(1-\alpha)} \quad (22.3.32)$$

$$\leq \left(\int |f_j| \right)^{q\alpha} \left(\int |f_j|^{5/3} \right)^{\frac{3}{5}(1-\alpha)q} = \|f_j\|_1^{q\alpha} \|f_j\|_{5/3}^{(1-\alpha)q}. \quad (22.3.33)$$

Hence we obtain

$$\|f_j\|_q \leq \|f_j\|_1^\alpha \|f_j\|_{5/3}^{1-\alpha} \leq \left(\sup \|f_j\|_1 \right)^\alpha \left(\sup \|f_j\|_{5/3} \right)^{1-\alpha}. \quad (22.3.34)$$

The Banach-Alaoglu theorem 2.6.4 implies our assertion. \square

We return to the main path. From The Lemma 3.1.4, the function $\rho_{N,\infty}$ is a minimizer because we can deduce now

$$E_{\leq}(N) = \lim_{j \rightarrow \infty} \mathcal{E}(\rho_j) \geq \mathcal{E}(\rho_{N,\infty}) \geq E_{\leq}(N). \quad (22.3.35)$$

Next we show the

Lemma. 3.1.6. *The above $\rho_{N,\infty}$ is unique.*

To do this we need the following

Lemma. 3.1.7. (Theorem II.6. [133].) *The functional $\mathcal{E}(\rho)$ is a strictly convex functional of ρ on the convex set $\mathcal{C}_{\leq N}$. Furthermore $E_{\leq}(N)$ is nonincreasing.*

Proof. Write

$$\mathcal{E}(\rho) = K(\rho) - A(\rho) + R(\rho) \quad (22.3.36)$$

corresponding to the three terms in the definition (22.3.16) of $\mathcal{E}(\rho)$. The functional $A(\rho)$ is linear and $K(\rho)$ is strictly convex since $f(x) = x^{5/3}$ is strictly convex on $[0, \infty)$. Finally $R(\rho)$ is strictly convex since $|x|^{-1}$ is strictly positive definite by Theorem 2.7.1. \square

Proof. We start the proof for Lemma 2.3.1. Assume there were two different minimizers ρ^1 and ρ^2 in $\mathcal{C}_{\leq N}$. Then $\rho = (\rho^1 + \rho^2)/2$, which is also in $\mathcal{C}_{\leq N}$, has strictly lower energy than $E_{\leq}(N)$, which is a contradiction. This reasoning also shows that $E_{\leq}(N)$ is a convex function. It is a simple consequence of its definition that $E_{\leq}(N)$ is nonincreasing. This is the end of the proof for Lemma 3.1.6. \square

By definition we have

$$E_{\leq}(N) \leq E(N). \quad (22.3.37)$$

Lemma. 3.1.8. *It holds that*

$$E_{\leq}(N) = E(N), \quad (22.3.38)$$

i.e., we can show that

$$E(N) \leq E_{\leq}(N). \quad (22.3.39)$$

Proof. We can suppose that $\int \rho_{N,\infty} = M < N$, for otherwise the desired conclusion is immediate. Take any nonnegative function $g \in L^{5/3} \cap L^1$ with

$$\int g = N - M \quad (22.3.40)$$

and consider, for each $\lambda > 0$, the function

$$\rho_{\lambda}(x) := \rho_{N,\infty}(x) + \lambda^3 g(\lambda x). \quad (22.3.41)$$

As $\lambda \rightarrow 0$, $\rho_{\lambda} \xrightarrow{s} \rho_{N,\infty}$ in every L^p with $p \in (1, 5/3]$. Therefore we have $\mathcal{E}(\rho_{\lambda}) \rightarrow \mathcal{E}(\rho_{N,\infty})$. On the other hand it holds that $\mathcal{E}(\rho_{\lambda}) \geq E(N)$, and hence $E(N) \leq E_{\leq}(N)$. Here we use the fact that our domain is the whole of \mathbb{R}^3 . \square

Note that we travel around (infinitely) many L^p spaces and use many useful and important properties for them!

Thomas-Fermi equation

Theorem. 3.1.9. (Theorem 11.13 [130], Theorem II.10. [133].) *The minimizer of the unconstrained problem, $\rho_{N,\infty}$, is not the zero function and it satisfies the following equation, in which $\mu \geq 0$ is some constant that depends on N :*

$$\rho_{N,\infty}(x)^{2/3} = \frac{Z}{|x|} - \left(\frac{1}{|x|} * \rho_{N,\infty} \right)(x) - \mu \quad \text{if } \rho_{N,\infty}(x) > 0, \quad (22.3.42)$$

$$0 \geq \frac{Z}{|x|} - \left(\frac{1}{|x|} * \rho_{N,\infty} \right)(x) - \mu \quad \text{if } \rho_{N,\infty}(x) = 0. \quad (22.3.43)$$

Remark. 3.1.10. *An equivalent way to write the above is*

$$\rho_{N,\infty}(x)^{2/3} = \left[\frac{Z}{|x|} - \left(\frac{1}{|x|} * \rho_{N,\infty} \right)(x) - \mu \right]_+. \quad (22.3.44)$$

Proof. Clearly $E_{\leq}(N)$ is strictly negative because we can easily construct some small ρ for which $\mathcal{E}(\rho) < 0$. This implies that $\rho_{N,\infty} \neq 0$ as a function.

For any real-valued function $g \in L^{5/3} \cap L^1$ and all $t \in [0, 1]$ consider the family of functions

$$\rho_t(x) := \rho_{N,\infty}(x) + t \left(g(x) - \frac{\int g}{\int \rho_{N,\infty}} \rho_{N,\infty}(x) \right), \quad (22.3.45)$$

which are defined since $\rho_{N,\infty} \neq 0$. Clearly $\int \rho_t = \int \rho_{N,\infty}$ and it is easy to check that $\rho_t \geq 0$ for all $t \in [0, 1]$ provided that g satisfies the two conditions:

$$g(x) \geq -\frac{\rho_{N,\infty}}{2}, \quad -\int g \geq -\frac{1}{2} \int \rho_{N,\infty}. \quad (22.3.46)$$

Define the function

$$F(t) := \mathcal{E}(\rho_t), \quad (22.3.47)$$

which certainly has the property that

$$E_{\leq}(N) \leq F(t), \quad \forall t \in [0, 1]. \quad (22.3.48)$$

We check the differentiability of F :

1. the $\int \rho^{5/3}$ term is differentiable by Theorem 2.6.2 (Gateaux derivative),
2. the second and third term in the expression (22.3.2) are trivially differentiable since they are polynomials in t .

Since $F(0) = E_{\leq}(N)$ and F must be nondecreasing at $t = 0$, it follows that

$$F'(0) \geq 0. \quad (22.3.49)$$

Define the function

$$W(x) := \rho_{N,\infty}^{2/3}(x) - \frac{Z}{|x|} + \left(\frac{1}{|x|} * \rho_{N,\infty} \right)(x) \quad (22.3.50)$$

and set

$$\mu := -\frac{\int \rho_{N,\infty}(x) W(x) dx}{\int \rho_{N,\infty}(x) dx}. \quad (22.3.51)$$

Since

$$\frac{d}{dt} \int \rho_t(x)^{5/3} dx \Big|_{t=0} = \frac{5}{3} \int \rho_{N,\infty}(x)^{2/3} \left(g(x) - \frac{\int g}{\int \rho_{N,\infty}} \rho_{N,\infty}(x) \right) dx, \quad (22.3.52)$$

$$\frac{d}{dt} \int \frac{1}{|x|} \rho_t(x) dx \Big|_{t=0} = \int \frac{1}{|x|} \left(g(x) - \frac{\int g}{\int \rho_{N,\infty}} \rho_{N,\infty}(x) \right) dx, \quad (22.3.53)$$

$$\frac{d}{dt} D(\rho_t, \rho_t) \Big|_{t=0} = \int \left(\rho_{N,\infty} * \frac{1}{|x|} \right)(x) \left(g(x) - \frac{\int g}{\int \rho_{N,\infty}} \rho_{N,\infty}(x) \right) dx, \quad (22.3.54)$$

it holds that

$$0 \leq \frac{d}{dt} F(t) \Big|_{t=0} = \int W(x) \left(g(x) - \frac{\int g}{\int \rho_{N,\infty}} \rho_{N,\infty}(x) \right) dx \quad (22.3.55)$$

$$= \int W(x) g(x) dx + \int g(x) dx \int \frac{-W(x) \rho_{N,\infty}(x)}{\int \rho_{N,\infty}} dx \quad (22.3.56)$$

$$= \int g(x) (W(x) + \mu) dx \quad (22.3.57)$$

for all functions g with the properties stated above.

In particular (22.3.57) holds for all nonnegative functions g with

$$\int g \leq \frac{1}{2} \int \rho_{N,\infty}, \quad (22.3.58)$$

and hence (22.3.57) holds for all nonnegative functions in $L^{5/3} \cap L^1$. From this it follows that $W(x) + \mu \geq 0$ a.e., which yields (22.3.43). From (22.3.51) we see that $-\mu$ is the average of W with respect to the measure $\rho_{N,\infty}(x)dx$, and hence the condition $W(x) + \mu \geq 0$ forces us to conclude that

$$W(x) + \mu = 0 \quad \forall x \in \{x : \rho_{N,\infty}(x) > 0\}; \quad (22.3.59)$$

this proves (22.3.42).

The last task is to prove that $\mu \geq 0$. If $\mu < 0$ then (22.3.42) implies that for $|x| > -\mu/Z$, $\rho_{N,\infty}(x)^{2/3}$ equals an L^6 -function plus a constant function, i.e., $-\mu$. If $\rho_{N,\infty}$ had this property it could not be in L^1 . \square

The Thomas-Fermi equation (22.3.44) reveals many interesting properties of $\rho_{N,\infty}$ and we refer the reader to [133] for this theory.

The Thomas-Fermi minimizer

Theorem. 3.1.11. (Theorem 11.14. [130]) *As before the function $\rho_{N,\infty}$ be the minimizer for the unconstrained problem. Then it holds that*

$$\int_{\mathbb{R}^3} \rho_{N,\infty}(x) dx = N \quad \text{if } 0 < N \leq Z, \quad (22.3.60)$$

$$\rho_{N,\infty} = \rho_{Z,\infty} \quad \text{if } N \geq Z. \quad (22.3.61)$$

In particular the former implies that $\rho_{N,\infty}$ is the minimizer for the constrained problem when $N \leq Z$. If $N > Z$, there is no minimizer for the constrained problem.

The number μ in (22.3.44) is 0 if and only if $N \geq Z$ and in this case $\rho_{N,\infty}(x) = \rho_{Z,\infty} > 0$ for all $x \in \mathbb{R}^3$.

*The **Thomas-Fermi potential** Φ_N defined by*

$$\Phi_N(x) := \frac{Z}{|x|} - \left(\frac{1}{|x|} * \rho_{N,\infty} \right)(x) \quad (22.3.62)$$

satisfies $\Phi_N(x) > 0$ for all $x \in \mathbb{R}^3$. Hence, when $\mu = 0$, corresponding to $N = Z$, the TF equation becomes

$$\rho_{Z,\infty}(x)^{2/3} = \Phi_Z(x). \quad (22.3.63)$$

Proof. We shall start by proving that there is a minimizer for the constrained problem if and only if $\int \rho_{N,\infty} = N$, in which case the minimizer is then obviously $\rho_{N,\infty}$. If $\int \rho_{N,\infty} = N$, then $\rho_{N,\infty}$ is a minimizer for $E(N)$. If the $E(N)$ problem has a minimizer (call it ρ^N), then $\int \rho^N = N$ and, by the monotonicity statement in Lemma 3.1.6, ρ^N is a minimizer for the unconstrained problem. Since this minimizer is unique $\rho^N = \rho_{N,\infty}$.

Now suppose that there is some $M > 0$ for which $M > \int \rho_{M,\infty} =: N_c$ (we shall soon see that $N_c = Z$). By uniqueness, we have that $E(M) = E(N_c)$. Then two statements are true:

1. $\int \rho_{N,\infty} = N_c$ and $\rho_{N,\infty} = \rho_{N_c,\infty}$ for all $N \geq N_c$, and
2. $\int \rho_{N,\infty} = N$ for all $N \leq N_c$.

To prove the former suppose that $N \geq N_c$. We shall show that $E(N) = E(N_c)$ (recall that $E(N) = E_{\leq}(N)$), and hence that $\rho_{N,\infty} = \rho_{N_c,\infty}$ by uniqueness. Clearly $E(N) \leq E(N_c)$. If $E(N) < E(N_c)$ and if $N < M$, we have a contradiction with the monotonicity of the function E . If $E(N) < E(N_c)$ and if $N > M$, we have a contradiction with the convexity of the function E . Thus $E(N) = E(N_c)$ and the former statement is proved. The latter statement follows from the former, for suppose that $\int \rho_{N,\infty} =: P < N$. Then the conclusion of the former holds with N_c replaced by P and M replaced by N . Thus, by the former, $\int \rho_{Q,\infty} = P$ for all $Q \geq P$. By choosing

$$Q = N_c \geq N \geq P, \quad (22.3.64)$$

we find that $N_c = \int \rho_{N_c} = P$, which is a contradiction.

We have to show that $N_c = Z$ and this will be done in conjunction with showing the nonnegativity of the TF potential.

Let

$$A = \{x \in \mathbb{R}^3 : \Phi_N(x) < 0\}. \quad (22.3.65)$$

By Lemma 3.1.6, Φ_N is continuous away from $x = 0$ and vanishes uniformly as $|x| \rightarrow \infty$. Hence A is an open set. In some small neighborhood of $x = 0$, $\Phi_N(x)$ is clearly positive, so $0 \notin A$. From the TF equation with $\mu \geq 0$, we see that $\rho_{N,\infty}(x) = 0$ for $x \in A$. But

$$\Delta \Phi_N = 4\pi \rho_{N,\infty} = 0 \quad \text{in } A, \quad (22.3.66)$$

and Theorem 9.3 in [130] tells us that Φ_N is harmonic in A . Since Φ_N is continuous, Φ_N vanishes on the boundary of A . Since Φ_N also vanishes uniformly at ∞ , the strong maximum principle, Theorem 9.4 in [130], states that $\Phi_N = 0$ for $x \in A$. Thus A is empty, as claimed. We omit the proof that Φ_N is strictly positive as an exercise.

Let $N > Z$ and consider the unconstrained optimizer $\rho_{N,\infty}$. We claim that $\int \rho_{N,\infty} \leq Z$. By the fact that $\rho_{N,\infty}$ is a radial function we get from equation 9.7 (5) [130] (Newton's theorem), that

$$\left(\frac{1}{|x|} * \rho_{N,\infty} \right)(x) = \frac{1}{|x|} \int_{|y| \leq |x|} \rho_{N,\infty}(y) dy + \int_{|y| > |x|} \frac{1}{|y|} \rho_{N,\infty}(y) dy. \quad (22.3.67)$$

From this and the definition of Φ_N it follows easily that $\lim_{|x| \rightarrow \infty} |x| \Phi_N(x) = Z - \int \rho_{N,\infty}$. Hence $\int \rho_{N,\infty} \leq Z$, for otherwise it would contradict the positivity

of Φ_N . Thus, for $N > Z$ the constrained TF problem does not have a minimizer and we conclude that $N_c \leq Z$.

Because $E(N_c)$ is the absolute minimum of $\mathcal{E}(\rho)$ on \mathcal{C} , and because $\rho_{N_c, \infty}$ is the absolute minimizer, a proof analogous to that of Theorem 3.1.1 (indeed, an even simpler proof), shows that this $\rho_{N_c, \infty}$ satisfies the TF equation with $\mu = 0$. Since Φ_N is nonnegative, this is the equation (22.3.63) with $\rho_{Z, \infty}$ replaced by $\rho_{N_c, \infty}$. We have seen that $\Phi_N(x)$ behaves like $(Z - N_c)/|x|$ for large $|x|$. If $N_c < Z$, then, from (22.3.63), $\rho_{N_c, \infty} \notin L^{5/3}$, which is a contradiction. \square

Chapter 23

Talk: Analysis for the Friedrichs model: perturbation theory for embedded eigenvalues

手書きのメモしかないので, そのうち \TeX におこす.

Chapter 24

Mathematical foundations for quantum mechanics

動画から適当に切り出してくる.

Chapter 25

Mathematical foundations for quantum statistical mechanics

動画から適当に切り出してくる.

Chapter 26

Mathematical problems in quantum mechanics, quantum field theory, and quantum statistical mechanics

Togetter (Twitter) からの転載.
元の質問は次の通り.

摂動論を勉強した頃、摂動を加える前後で固有関数の様相が大きく変わらないのかが疑問でモヤモヤしていたのですが、これも難しい問題だったのですね。別の話題かつ大雑把な質問でアレなのですが、いま「この辺」の他にはどんなトピックがホットなのですか。

本当に最近の動向は良く分かりませんし、私の専門は場の理論・量子統計なので量子力学方面は微妙ですが、何にせよこのあたりで知っているのをいくつか。まず観測・量子情報関係が色々あります。早稲田だと中里・安倍研が正にここでしょう。数学というか数理物理だと理科大の大矢先生や電通大の長谷川さんあたりがやっています。具体的な問題はあまり詳しくないですがこの辺も作用素不等式の研究といった数学色が強い話題も色々あるので。

新井先生の『量子現象の数理』[30]にも書いてありますが、アハラノフ-ボーム効果もまだ死んでいません。比較的最近、数学会の解析学賞の話題になる程度には色々面倒な話があります。3次元でのアハラノフ-ボームは物理的には本道ですが、数学的には多分未開です。

物理でも当然大事な話ですが、散乱理論もまだまだ数学的には未開の部分がたくさんあります。筑波の磯崎先生の数学会賞にもなっている話は量子力学ですが、相対論的場の理論でも Dybalski さんなどの話 [59] があります。こちらは量子力学以上に手付かず。

またもや新井先生の本に書いてありますが、超対称性量子力学もまだまだ発展途上でしょう。特にこれは幾何との絡みがあるようなので、研究者がたくさんいるはず。あとアンダーソン局在も数学の人が熱心に取り組んでいる話題です。

「この辺の話題」といった丁度その部分ですが、場の理論だと、発散の困難関係を処理しないと何も前に進まない状況で、いまだにそこではまっているところです。同じく摂動の問題で、励起状態は準安定になってほしいというのが以前言った所ですが、発散の困難の場合は基底状態の摂動でいろいろあります。

基底状態は基底状態のまま、基底エネルギーは固有値でいてほしいのですが、これが固有値でなくなる（固有状態がどこかに行ってしまう）というのが発散の困難です。ちょうど量子力学の δ 関数のガウシアンによる近似で、 $t > 0$ のときはまともな関数なのに極限を取ると超関数になってしまうのと同じ問題です。

その前の基底状態・平衡状態の存在証明（発散の困難の回避）で手間取っているためほとんど進んでいませんが、物性理論への応用が私のやりたい部分です。電子-フォノン相互作用系での強磁性やフォノンの BEC など、相転移関係で基本的な部分をはっきりさせたいというところ。

励起状態関係の話やレーザーの話と数学的に近い部分・重なる部分があるのですが、平衡への回帰という非平衡統計の問題も議論されています。このあたりは早稲田の田崎先生が C^* を使う手法で研究されていましたが、数年前に亡くなってしまいました。

場の理論というより統計ですが、数学として扱っているモデル・手法がかなり重なる、または全く同じだったりするので、研究者も多少重なりがあります。かなり数学よりになりますが、流体力学極限という確率論の話も使いつつ非平衡関係のことをしている Spohn という人もいます。

量子力学方面に戻りますが、物理でも最近話題になっている（らしい）circuit QED も数学ベースでの研究があります。廣川先生や廣島先生がやっています。これは一次元調和振動子の摂動みたいなものなのでモデル自体は非常にシンプルですが、最近ホットな話題と聞いています。あまり物理っぽい話を紹介できなくて申し訳ないのですが、学部にも毛が生えた程度、または本当に学部の物理水準でも既に数学としては色々問題があります。解決済みですが、量子統計だと平衡状態が物理で書かれる $\text{Tr } e^{-\beta H}$ という形では存在しないという問題もあったのでした。

Chapter 27

Fundamental problems of perturbation theory in quantum mechanics as an approximation

Twitter での次の発言に対する応答をまとめた。

ちなみに物理していると「収束するかの難しい問題はおいといて」的な文言と共に作用素の級数が出てくることがありますね。

やや別件ですが、作用素級数とか収束以前に摂動が原理的にまずいということがあります。 $H = -\Delta + \alpha V$ で α が小さいときに摂動展開する、という状況でエネルギー（固有値）の振る舞いが大きな問題になります。 $V_1 = x^2$, $V_2 = -r^{-1}$ としましょう。もちろん調和振動子と水素原子に対応するハミルトニアンです。固有値（スペクトル）、固有関数ともに厳密解が分かっているわけですが、いくら α が小さかろうと、測定値であるエネルギーに対応する固有値の振る舞いがこの 2 つで決定的に違います。もっというなら $\alpha = 0$ のときとも全く違います。調和振動子などは連続スペクトルが全て消えてしまうという凄まじさ。

基底エネルギーのずれくらいならそこそこ値が合うはずですが、元々のスペクトル自体が激変しているので、あまり適当なことをすると問題になります。応用上も大事なレーザーでも問題があります。ラムシフトが起きますが、物理的に言って励起状態は準安定状態に遷移して欲しいわけで、数学的には「摂動前の固有値に対応するものは固有値でなくなる」という現象が起きて欲しい。ただし摂動論だと「摂動前の固有値は摂動では少しずれた値の固有値になる」と思って計算するので、ここでも原理的な問題が起きます。学部 3 年の量子力学でやったと思いますが、実際にはここで手で虚軸方向へのずれを入れることで回避します。この間「数学的に細かいことは気にするな。もはや数学徒としては研究レベルの問題だから」と書きましたが、今本当にこの辺の数理解物理の研究が進行中です。

Chapter 28

Direct integral and its physics

これは Togetter に【微分ヤクザ P による直積分講座】というのでまとまっている。いったん、やりとりをそのまま書いておく。

#0 は素イデアル 0t_uda 2012-01-08 00:27:35

整域でない環に思いを馳せた。 alg_d 2012-01-08 00:35:50

@alg_d 非可換環 phasetr 2012-01-08 14:13:42

@phase_tr 両側イデアルでない右イデアル alg_d 2012-01-08 14:22:52

@alg_d 左閉イデアルなら大量生産できます phasetr 2012-01-08 14:27:33

@phase_tr 「大量」とは具体的にはどのくらいでしょうか。 alg_d 2012-01-08 14:28:34

@alg_d 非可算無限程度なら。作用素環で大事な GNS 構成定理というのがありますが、これは作用素環上の有界線型汎関数の kernel みたいなものからヒルベルト空間などを作る定理です。この kernel みたいなものが大雑把に線型汎関数の分だけあります。 phasetr 2012-01-08 14:30:46

@phase_tr そんなにたくさん作って、一体何を企んでるんですか。 alg_d 2012-01-08 14:32:27

@alg_d 最近ヒッグスなどで自発的対称性の破れというのが取り沙汰されていますが、ここでは非可算無限個 (例えば $U(1)$ くらい) の線型汎関数を重ね合わせて線型汎関数を 1 つ作ったりします。物理の実用上どうしても必要です。 phasetr 2012-01-08 14:37:14

@phase_tr 自発的対称性の破れてそういうの使うんですか。よく分かりませんが凄いです。 alg_d 2012-01-08 14:40:16

@alg_d 直和の代わりに直積分とか使いますが、気持ちの上では大したことはしません。めんどいのはテクニカルな部分だけだと見ています。 phasetr 2012-01-08 14:42:28

@phase_tr 直積分 wwwwww フヒヒ wwwwww alg_d 2012-01-08 14:43:40

直積分ってなんぞ www 恐ろしい語感に衝撃を受けた. alg_d 2012-01-08 14:45:17

【 Direct integral 】<http://t.co/sTQT3B6T> alg_d 2012-01-08 14:45:38
インテグラルの上に \oplus が付いていて笑劇を受けている. alg_d 2012-01-08 14:50:04

@alg_d von Neumann が導入したというように書いてありますが, 実際に von Neumann 環の分類では環を直積分分解して簡単な所に還元します. 特に中心が自明な環 (歴史的に因子, factor と呼ぶ) に落とし込むので factor の解析が大事です. 物理でもこれが出ます. phasetr 2012-01-08 14:55:04

ちなみに今やっている所でも定ファイバー直積分を日常的に使っている. 非相対論的 QED (量子電気力学の方だ) で使うヒルベルト空間自体が直積分で書かれている (書ける) ので, 大体全部直積分になる. phasetr 2012-01-08 14:58:56

直積分とは何か知らない僕に花束を life_wont_wait 2012-01-08 15:01:42

@life_wont_wait 強磁性を考えます. 理論的には磁化の方向はどこを向いていても構わないので $SO(3)$ だけの自由度があります. $SO(3)$ の分だけ基底状態がありますが, これの重ね合わせもまた基底状態になります. 重ね合わせを数学的に書くと直積分になるというだけです. phasetr 2012-01-08 15:08:54

@phase_tr 無限次元空間の基底の線型結合とかですか? life_wont_wait 2012-01-08 15:12:41

@life_wont_wait 規格化したいので正確には凸結合ですし, ベクトルだけでなく, 汎関数, 空間, 演算子全て直積分で書くことになりましたが大体そんな感じです

phasetr 2012-01-08 15:16:57 @phase_tr 空間も...了解ありがとうございます. やはりルベグ積分必須な感じですか?

life_wont_wait 2012-01-08 15:20:27 @life_wont_wait 数学的にきちんとやりたいならもちろん必要ですが, そんなものはどうでもいいでしょう. 対称性の群が連続群だから連続にして, それで積分になったというだけです. うるさいことを言う人には「数学やりたければ数学科に行け」と言っておいてください. phasetr 2012-01-08 15:23:00

「数学やりたければ数学科に行け」これは名言だ faogr 2012-01-08 15:23:37

Part VIII

Probability and Statistics

Chapter 29

Talk: Law of large number and central limit theorem in coin tossing

29.1 参考文献

話は [97] を元に組んでいる. 確率論それ自体については [72] が参考になる.

29.2 数学的準備

29.2.1 加法定理・倍角の公式

きちんとした証明は面倒なので, Euler の公式 $e^{ix} = \cos x + i \sin x$ を使う簡単な導出だけしておこう.

$$e^{2ix} = \cos 2x + i \sin 2x \quad (29.2.1)$$

$$= (e^{ix})^2 = (\cos^2 x - \sin^2 x) + i(2 \cos x \sin x). \quad (29.2.2)$$

あとは実部と虚部を比較する.

29.2.2 指数関数の微分積分

まず微分.

$$\frac{d}{dx} e^{ax} = a e^{ax}. \quad (29.2.3)$$

これで (不) 定積分もできる.

$$\int_0^1 e^{ax} dx = \frac{1}{a} (e - 1). \quad (29.2.4)$$

あとで使う計算もやっておこう.

$$\int_0^1 e^{ix(1-2t)} dt = \int_{-1}^1 e^{ixs} \frac{ds}{2} = \frac{\sin x}{x}. \quad (29.2.5)$$

29.2.3 Rademacher 関数

定義は式 (29.4.12) だ. 周期 1 の周期関数として考えると関数 $r_k(t)$ は

$$r_k(t) = r_1(2^{k-1}t) \quad (29.2.6)$$

と書ける. この周期性は実際の計算のときに役に立つ.

29.2.4 Rademacher 関数の積分公式

Rademacher 関数 (29.4.12) に関する積分公式を導出しておく.

Theorem. 2.4.1.

$$\int_0^1 e^{iy r_k(t)} dt = \cos y. \quad (29.2.7)$$

Proof. (29.2.6) を使って計算していく.

$$\int_0^1 e^{iy r_k(t)} dt = \int_0^1 e^{iy r_k(2^{k-1}t)} dt \quad (29.2.8)$$

$$= \int_0^{2^{k-1}} e^{iy r_1(s)} \frac{ds}{2^{k-1}} \quad (29.2.9)$$

$$= \sum_{l=1}^{2^{k-1}} \int_{l-1}^l e^{iy r_1(s)} \frac{ds}{2^{k-1}} \quad (29.2.10)$$

$$= \sum_{l=1}^{2^{k-1}} \frac{1}{2^{k-1}} \int_0^1 e^{iy r_1(s)} ds \quad (29.2.11)$$

$$= \int_0^1 e^{iy r_1(s)} ds \quad (29.2.12)$$

$$= \int_0^{1/2} e^{iy} ds + \int_{1/2}^1 e^{-iy} ds \quad (29.2.13)$$

$$= \frac{1}{2} (e^{iy} + e^{-iy}) = \cos y. \quad (29.2.14)$$

□

29.2.5 Rademacher 関数の積分公式 2

あとで使うので, [97] P.11 の演習問題 3. を解く.

Proposition. 2.5.1. $s \geq 1, k_1 < k_2 < \dots < k_s$ に対して

$$\int_0^1 \prod_{l=1}^s r_{k_l}(t) dt = 0. \quad (29.2.15)$$

が成り立つ.

Hilbert 空間論 (関数の直交性) も想起させるなかなか含蓄の深い関係式だ. もちろん, 正確には確率論でより深く大事な概念, 独立性を表わしている.

Proof. $s = 1$ のときは自明だ. $s = 2$ のときは式 (29.2.6) を使うと

$$\int_0^1 r_{k_1}(t) r_{k_2}(t) dt = \int_0^1 r_1(2^{k_1-1}t) r_1(2^{k_2-1}t) dt \quad (29.2.16)$$

となる. 正と負が適当に互い違いになるだけで同じ回数出てくるので, 効果が打ち消しあって積分は 0 になる. 一般の s でも同じことを議論すればいい. \square

29.2.6 Lebesgue 積分の定理

Theorem. 2.6.1. 関数列 (f_n) の各関数は非負で可積分だとする.

$$\sum_{n=1}^{\infty} \int_0^1 f_n(t) dt < \infty, \quad (29.2.17)$$

つまり上記の積分の和が収束するなら関数項級数

$$\sum_{n=1}^{\infty} f_n \quad (29.2.18)$$

はほとんどいたる所で収束する.

29.3 導入

今回のひなしろ数学会では確率論, 特に大数の法則と中心極限定理について議論しよう. まず大数の法則と中心極限定理の説明が必要だ.

大数の法則の説明を Wikipedia から引いてみよう.

ある試行において事象が起きる確率 (数学的確率, 理論的確率などともいう) が p であり, その試行は, 繰り返し行ったとしてもある回の試行が他の回の試行に影響を及ぼすことがない (独立試行) ものとする. このような前提条件の下で, その事象が起きる比率が試行回数を増やすにつれて近づく値 (統計的確率あるいは経験的確率) は p である. つまり, 各回の試行において各事象の起こる確率というものが, 試行回数を重ねることで, 各事象の出現回数によって捉えられるというのが大数の法則の主張するところである.

「確率」という言葉を 2 つの意味で使っているのでかなり面倒な話をしているが, ひとまずあとできちんと書く数学的な内容からいうと次のような感じになる.

1. (変な細工がない) 硬貨を投げるとき, 表と裏が出るのは平均的には $1/2$ ずつになる: 表が出るのを $+1$, 裏が出るのを -1 と考えると, 平均 (期待値) が 0 になる.
2. (変な細工がない) 骰子を投げるとき, 出る目の平均 (期待値) は $(1 + 2 + 3 + 4 + 5 + 6) / 6 = 3.5$ になる.

色々微妙なところはあるが, これらを一般的にきちんと証明しようとして頑張った数学的成果を大数の法則と呼んでいる.

応用上も色々な話がある. まずは平均にどんな情報・意味があるかを考えてみよう. よくニュースなり何なりで平均の数値が出てくる. しかし, 必ずしも平均に意味があるとは限らない. 次のような状況を考えよう.

3 年 1 組の生徒、40 人に数学のテストを受けさせたところ、簡単だが授業をきちんと聞いていないと出来ない問題を出題した。きちんと勉強している生徒 20 人は 100 点を取れたが、クラスのもう半分の 20 名は 0 点だった。

このとき平均点はもちろん 50 点だが、この 50 という数値に何の意味があるだろうか。

次のような状況を考えてもいい。

ある国の国民の 1 年の所得を調べた。富が偏在していて、人口の 1 # は各人 100 億円の所得があるものの、他の 99 # の所得は 1 万円以下だった。

こんな歪な状況で平均に何の意味があるだろうか。意味が全くないとは言わない。ただ、平均だけを見ていても仕方がないことはいくらでもあるということだ。

話はそれだが、そういうときに真っ先に見るべきは実際のデータとそのグラフだ。グラフを見るだけでも色々なことが分かる。これは統計学、特に記述統計学の基本中の基本で、色々なところでグラフが出てくる理由でもある。数値だけでなくグラフを見ることが大事なのだ。

ちなみに、グラフを徹底利用して統計学の応用に決定的な影響を与えた人物として Florence Nightingale がいる。看護師としてばかり有名なのがあるが、実際には医療統計学と統計学の普及に大きな影響を与えた不朽の業績がある。

それはそれとして、平均が持つ意味、平均からのずれ、そしてそれを表す量も調べなければいけないことが分かってもらえたとして、これらの問題について考えていこう。先程「平均には意味があるかは微妙なところだ」と書いた。しかし、当然のことだが、平均がきちんと意味を持つときもある。そういう場合に「どのくらい頑張ると平均に辿り着けるか」という問題が出てくる。例えば物理の実験では（理論から予測される）適当な真の値がある。特に理論の検証というタイプの実験・測定では当然その値を得ようと実験をする。しかし誤差など色々な問題があるため本来の値からは少しずつずれてふらついた値が実験値として出てくる。このふらつきをおさえた値を実験値から計算するにはどのくらい実験すればいいか、といった問題を是非解きたい。1-2 回やっただけではふらつきが多過ぎたときにその影響をおさえきれないし、実験の試料が高かったり、施設利用費が高いとそう何度も実験できないという現実的な問題もある。この辺をどうにかしたいということだ。これが次の中心極限定理の内容にあたる。

難しすぎるので今日は話さないし勉強不足もあってそもそも話せないが、分散または偏差という大事な量があってそれについて大偏差原理という命題がある。そちらはそちらで応用上大事だ。統計力学とも深い関わりがあるようで勉強したい。

大偏差原理はそれとして、今度は中心極限定理だ。ひとまず Wikipedia から説明を引いてみよう。

大数の法則によると、ある母集団から無作為抽出された標本平均はサンプルのサイズを大きくすると真の平均に近づく。これに対し中心極限定理は標本平均と真の平均との誤差を論ずるものである。多くの場合、母集団の分布がどんな分布であっても、その誤差はサンプルのサイズを大きくしたとき近似的に正規分布に従う。

要は何回くらい実験すればまともな値と思えるようになるか、そのレートを教えてくれるのが中心極限定理だ。正規分布などは一旦無視する。

今日はこうした応用上の目的（確率論というより統計学？）を背景にした確率論の数学を議論する。高校の確率論とはかなり毛色が違うし、大学の普通のコースともかなり違う。参考文献 [97] を元に、ちょっと変わった確率論入門コースを走ってみる。メインテーマはひたすらにコイン投げだ。それでは始めよう。

29.4 硬貨投げの数理と独立性

29.4.1 Viète の公式

まずは Viète の公式を出してみよう. 次の式を Viète の公式という.

$$\frac{2}{\pi} = \prod_{k=1}^{\infty} \cos \frac{\pi}{2^{k+1}} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2+\sqrt{2}}}{2} \cdot \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \cdots \quad (29.4.1)$$

どうやって発見したのか, どうやって導いたのか・証明を思い付いたのかが謎だし, 比較的簡単に計算できる $\cos \pi/2^k$ が出てくるのも不思議で, さらには $\cos \pi/2^k$ をきちんと頑張って計算してみたというところも尊い, 色々な感慨をかきたてるとても不思議な式だ.

まず計算自体は簡単な, しかし不思議な次の式から始める. 29.2.1 で導出した倍角の公式を繰り返し使うのだ.

$$\sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2} \quad (29.4.2)$$

$$= 2^2 \sin \frac{x}{4} \cos \frac{x}{4} \cos \frac{x}{2} \quad (29.4.3)$$

$$= 2^3 \sin \frac{x}{8} \cos \frac{x}{8} \cos \frac{x}{4} \cos \frac{x}{2} \quad (29.4.4)$$

$$= 2^n \sin \frac{x}{2^n} \prod_{k=1}^n \cos \frac{x}{2^k}. \quad (29.4.5)$$

割り算がしたいのでいったん $x \neq 0$ (正確には $x \neq 2n\pi, n \in \mathbb{Z}$) とする. ここで \sin の原点での微分係数から

$$1 = \lim_{n \rightarrow \infty} \frac{\sin \frac{x}{2^n}}{\frac{x}{2^n}} = \frac{1}{x} \lim_{n \rightarrow \infty} 2^n \sin \frac{x}{2^n} \quad (29.4.6)$$

となるから

$$\lim_{n \rightarrow \infty} 2^n \sin \frac{x}{2^n} = x \quad (29.4.7)$$

が導ける. 最終結論のこの式自体は $x = 0$ でも成り立つことに注意しよう²⁹.

再び $x \neq 0$ を仮定しよう¹. 式 (29.4.5) と (29.4.7) から

$$\frac{\sin x}{x} = \prod_{k=1}^{\infty} \cos \frac{x}{2^k} \quad (29.4.8)$$

が出る. ここで何をどう思ったか $x = \frac{\pi}{2}$ を代入してみると次のようになる.

$$\frac{2}{\pi} = \prod_{k=1}^{\infty} \cos \frac{\pi}{2^{k+1}} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2+\sqrt{2}}}{2} \cdot \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \cdots \quad (29.4.9)$$

これが示したかった Viète の公式だ.

¹ハイパー どうでもいいが「再びマタタビ楕円型」という天啓を得たことを記録しておく.

29.4.2 Rademacher 関数

これまではまさに高校の数学だった²⁶. 本番はここからだ. 式 (29.4.8) を別の観点から眺めてみようという体で, 唐突に次の無茶苦茶な理論展開を見せる.

実数 $t \in [0, 1]$ を取り, これを 2 進展開しよう. $\varepsilon_k \in \{0, 1\}$ とすると実数 $t \in [0, 1]$ は

$$t = \frac{\varepsilon_1}{2} + \frac{\varepsilon_2}{2^2} + \cdots \quad (29.4.10)$$

と書ける. もちろん上の式で 2 が 10 のときがいつもの 10 進法での展開だ. 展開を一意に決めるため $0.1111\cdots = 1$ のようないつもの処方箋を出しておく.

各桁の ε_i は t の関数だから 2 進展開をきちんと書くと

$$t = \frac{\varepsilon_1(t)}{2} + \frac{\varepsilon_2(t)}{2^2} + \cdots = \sum_{k=1}^{\infty} \frac{\varepsilon_k(t)}{2^k} \quad (29.4.11)$$

となる.

関数 $\varepsilon_k(t)$ の図を描いてみよう. 見ているだけでも特徴的で結構楽しい. 数学的には次の Rademacher 関数が便利なのでこれを使っていく. 今後の主役でもある.

$$r_k(t) := 1 - 2\varepsilon_k(t), \quad k \in \mathbb{Z}_+. \quad (29.4.12)$$

これを使って 2 進展開を書き直そう.

$$1 - 2t = \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k}. \quad (29.4.13)$$

29.2.4 節の結果と式 (29.4.8) を使うと

$$\frac{\sin x}{x} = \int_0^1 e^{ix(1-2t)} dt = \int_0^1 \exp\left(ix \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k}\right) dt \quad (29.4.14)$$

$$= \prod_{k=1}^{\infty} \cos \frac{x}{2^k} = \prod_{k=1}^{\infty} \int_0^1 \exp\left(ix \frac{r_k(t)}{2^k}\right) dt \quad (29.4.15)$$

が分かる. つまり

$$\int_0^1 \exp\left(ix \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k}\right) dt = \prod_{k=1}^{\infty} \int_0^1 \exp\left(ix \frac{r_k(t)}{2^k}\right) dt. \quad (29.4.16)$$

一般に極限と極限の順序交換ができないことには注意しておこう². そう思うと上式は積分と無限積が交換できることを言っていることになる! 印象的にいうなら「積の積分が積分の積になる」のだ.

²次の例をよく使っている.

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{n}{n+m} = 1, \quad (29.4.17)$$

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{n}{n+m} = 0. \quad (29.4.18)$$

29.4.3 どういうことなの？

式 (29.4.15) は偶然だろうか。本質が見えるようになるまでは偶然というほかない。ここで式 (29.4.8) を別の方法で導いてみよう。

唐突だが次の関数を考える²⁷。いったん少し面倒な方法で取り組んでみよう。29.4.4 節で整理するので少しのあいだ我慢してほしい。

$$\sum_{k=1}^n c_k r_k(t). \quad (29.4.19)$$

これは区間

$$\left(\frac{s}{2^n}, \frac{s+1}{2^n} \right), \quad s = 0, 1, \dots, 2^n - 1 \quad (29.4.20)$$

上でそれぞれ一定の値を取る階段関数だ。その一定の値は

$$\pm c_1 \pm c_2 \pm \dots \pm c_n \quad (29.4.21)$$

になる。したがって次の式が成り立つ。

$$\int_0^1 \exp \left[i \sum_{k=1}^n c_k r_k(t) \right] dt = \frac{1}{2^n} \sum_{\pm} \exp \left[i \sum_{k=1}^n \pm c_k \right]. \quad (29.4.22)$$

ここで右辺の和は全ての ± 1 の組み合わせで取る²⁸。したがって

$$\frac{1}{2^n} \sum_{\pm} \exp \left[i \sum_{k=1}^n \pm c_k \right] = \prod_{k=1}^n \frac{e^{ic_k} + e^{-ic_k}}{2} = \prod_{k=1}^n \cos c_k \quad (29.4.23)$$

だから、29.2.4 の結果を使って

$$\int_0^1 \exp \left[i \sum_{k=1}^n c_k r_k(t) \right] dt = \prod_{k=1}^n \cos c_k = \prod_{k=1}^n \int_0^1 e^{ic_k r_k(t)} dt \quad (29.4.24)$$

が分かる。 $c_k = x/2^k$ とすると、

$$\int_0^1 \exp \left[ix \sum_{k=1}^n \frac{r_k(t)}{2^k} \right] dt = \prod_{k=1}^n \cos \frac{x}{2^k} \quad (29.4.25)$$

となる。(29.4.13) の収束

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{r_k(t)}{2^k} = 1 - 2t \quad (29.4.26)$$

は区間 $(0, 1)$ で一様³ だから、積分と極限の順序の交換ができて次の式が導ける。

$$\frac{\sin x}{x} = \int_0^1 e^{ix(1-2t)} dt = \lim_{n \rightarrow \infty} \int_0^1 \exp \left[ix \sum_{k=1}^n \frac{r_k(t)}{2^k} \right] dt = \lim_{n \rightarrow \infty} \prod_{k=1}^n \cos \frac{x}{2^k} = \prod_{k=1}^{\infty} \cos \frac{x}{2^k}. \quad (29.4.27)$$

これで (29.4.8) の別証明ができた。

29.4.1 節での導出よりは長いし初等的でもないが、2 進展開との関係を示唆する不思議な証明といえる。これは一体何だろうか。ここからは 2 進展開と上の証明について深く考えていこう。

³大学の数学科でしか使わない概念なので気にしないでいい。あとで自分が困らないように数学的にきちんと書いておいただけだから。

29.4.4 2 進展開の確率論ことはじめ

次の条件を満たす実数 t はいくつあるだろうか. 正確にいうなら次の条件を満たす t の集合は何だろうか:

$$r_1(t) = 1, \quad r_2(t) = -1, \quad r_3(t) = -1. \quad (29.4.28)$$

グラフを書くとこの集合は端点を除いて区間 $(3/8, 4/8)$ になる. 区間の長さ (今後測度と呼ぶ) は当然 $1/8$ で, さらに

$$\frac{1}{8} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \quad (29.4.29)$$

とも書ける. これを測度 (区間の長さ) の記号を使って面倒くさく書くところなる.

$$\mu(r_1(t) = 1, r_2(t) = -1, r_3(t) = -1) = \mu(r_1(t) = 1) \cdot \mu(r_2(t) = -1) \cdot \mu(r_3(t) = -1). \quad (29.4.30)$$

ここで $\mu(A)$ は集合 (区間) A の長さを表わす.

これは次のように一般化できる: $\delta_k = \pm 1$ とすると,

$$\mu(r_1(t) = \delta_1, \dots, r_n(t) = \delta_n) = \prod_{k=1}^n \mu(r_k(t) = \delta_k). \quad (29.4.31)$$

何も言わなければわざわざ

$$\left(\frac{1}{2}\right)^n = \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} \quad (29.4.32)$$

を難しく書いていただけで何がしたいのか全く分からないし, 馬鹿ではないかと思うだろう. しかしこれが関数 r_k , ひいては 2 進展開に深く関わっていることが分かる.

ここまで来ると式 (29.4.24) は次のように機械的に計算できる.

$$\int_0^1 \exp \left[i \sum_{k=1}^n c_k r_k(t) \right] dt = \sum_{\delta_1, \dots, \delta_n} \exp \left[i \sum_{k=1}^n c_k \delta_k \right] \mu(r_1(t) = \delta_1, \dots, r_n(t) = \delta_n) \quad (29.4.33)$$

$$= \sum_{\delta_1, \dots, \delta_n} \prod_{k=1}^n e^{i c_k \delta_k} \prod_{k=1}^n \mu(r_k(t) = \delta_k) \quad (29.4.34)$$

$$= \sum_{\delta_1, \dots, \delta_n} \prod_{k=1}^n e^{i c_k \delta_k} \mu(r_k(t) = \delta_k) \quad (29.4.35)$$

$$= \prod_{k=1}^n \sum_{\delta_k = \pm 1} e^{i c_k \delta_k} \mu(r_k(t) = \delta_k) \quad (29.4.36)$$

$$= \prod_{k=1}^n \int_0^1 e^{i c_k r_k(t)} dt. \quad (29.4.37)$$

前の議論だと組み合わせ的に面倒な話が必要で, 少なくとも私には難しい⁴.

⁴ センターレベルの組み合わせ・確率論ですら私には難しい.

29.4.5 独立性とはじめ

硬貨投げを考えると、普通は次の条件を仮定している。

1. 硬貨投げは「公平」である：表か裏かが出る確率に偏りはなく、両方とも確率 $1/2$ である。
2. 硬貨投げの結果は他の回の結果とは関係ない：それぞれの事象は独立である。

独立を数学的に書くと次のようになる⁵。

事象 A_1, \dots, A_n が独立であることの定義は次式が成り立つことと定める。

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2) \cdots P(A_n). \quad (29.4.38)$$

ここで $P(A)$ は事象 A が起きる確率を表わす。

表と裏が適当な回数ずつ出るときの確率は明らかに

$$\frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} = \frac{1}{2^n} \quad (29.4.39)$$

になる。これまで計算してきた r_k に関する計算結果と同じなので、計算の便法として高校以来やってきた組み合わせ的な計算を r_k に関する計算で置き換えることができるかもしれない。そこで次の「辞書」を作って翻訳できるか試してみよう。

表	+1
裏	-1
k 回目の結果	$r_k(t)$
事象の確率	事象と対応する t の集合の測度

辞書の使い方を見るために次の問題を考えてみる。

公平な硬貨投げを n 回繰り返すときにちょうど l 回表が出る確率を求めよ。

辞書を使うと次のように翻訳できる。

r_1 から r_n のうち、ちょうど l 個が $+1$ になる t の集合の測度を求めよ。

こうすると、面倒な組み合わせの議論が積分（測度）の問題として考えられるようになる。組み合わせは色々チェックが大変なので、個人的には精神的な負担が減り苦しみから解放された気になる⁶。

実際に確率を計算してみよう。組み合わせ的に考えたときの結果

$$\frac{1}{2^n} \binom{n}{l} \quad (29.4.40)$$

が再現できるだろうか。もちろんできなければ困る。

まず求める値が

$$\mu \left(\sum_{k=1}^n r_k(t) = 2l - n \right) \quad (29.4.41)$$

⁵正確にはこう定義する。

⁶センターの組み合わせや確率論すらろくに解けない方の市民だった。

と書けることに注意しよう. 整数 m に対して

$$\frac{1}{2\pi} \int_0^{2\pi} e^{imx} dx = \begin{cases} 1, & m = 0, \\ 0, & m \neq 0 \end{cases} \quad (29.4.42)$$

となるから

$$\phi(t) := \frac{1}{2\pi} \int_0^{2\pi} \exp \left[ix \left(\sum_{k=1}^n r_k(t) - (2l - n) \right) \right] dx = \begin{cases} 1, & \sum_{k=1}^n r_k(t) = 2l - n, \\ 0, & \sum_{k=1}^n r_k(t) \neq 2l - n. \end{cases} \quad (29.4.43)$$

したがって

$$\mu \left(\sum_{k=1}^n r_k(t) = 2l - n \right) = \int_0^1 \phi(t) dt \quad (29.4.44)$$

$$= \int_0^1 \frac{1}{2\pi} \int_0^{2\pi} \exp \left[ix \left(\sum_{k=1}^n r_k(t) - (2l - n) \right) \right] dx dt \quad (29.4.45)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \left(\int_0^1 \left[ix \sum_{k=1}^n r_k(t) \right] dt \right) dx \quad (29.4.46)$$

式 (29.4.24) で $c_k = x$ とすれば

$$\mu \left(\sum_{k=1}^n r_k(t) = 2l - n \right) = \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \cos^n x dx \quad (29.4.47)$$

となる. ここで

$$\cos^n x = \left(\frac{e^{ix} + e^{-ix}}{2} \right)^n = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e^{i(n-2k)x} \quad (29.4.48)$$

に注意すると,

$$\frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \cos^n x dx = \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e^{i(n-2k)x} dx \quad (29.4.49)$$

$$= \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} \frac{1}{2\pi} \int_0^{2\pi} e^{2i(k-l)x} dx \quad (29.4.50)$$

$$= \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} \delta_{k,l} \quad (29.4.51)$$

となる.

まとめると次のようになることが分かった.

$$\frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \cos^n x dx = \frac{1}{2^n} \binom{n}{l} \quad (29.4.52)$$

2 項係数が指数関数の 2 項展開から出てくるところが証明のポイントで、組み合わせ的な議論がこれと積分計算に落ちていることが見て取れる。計算ではさっと流してしまっているのが、Rademacher 関数の独立性、(29.4.31) が本質的に効いていることを注意し直しておきたい。

29.5 大数の法則

前節では Viète の公式の初等的な証明からはじめ、硬貨投げにどう落とし込むか、さらには確率論、特に独立性がどう効いてくるのかを見た。

この節では主題の 1 つ、大数の法則について考えてみよう。弱法則と強法則の 2 種類がある。ここで「強弱」は論理的な強弱を表わしている。

29.5.1 大数の弱法則

また硬貨投げを考えよう。「公平」な硬貨を 1 枚投げたとき、表が出れば 1 円もらえ、裏が出れば 1 円払う (−1 円もらう) ことにする⁷。 n 回硬貨を投げたあと、いくらくらい儲かっているか (損しているか) を予測したい。このとき、硬貨投げという運の要素が絡んでいるため、正確な値ももちろん出せず、どのくらいの確率でどのくらい儲かるか、ということを実測することしかできないことに改めて注意しよう。

29.4.5 節の辞書を使うと儲け自体は次のように書ける。

$$\sum_{k=1}^n r_k(t). \quad (29.5.1)$$

問題は n 回硬貨投げをしたあとに想定額 A_n を越えられるかどうかだろう。要はいくら儲けたいという目標があって、それに到達できるかだ。辞書を使うと儲けの見込み額は次の値 (集合の測度、区間の長さ) を求める問題に帰着する：

$$\mu \left(\sum_{k=1}^n r_k(t) > A_n \right). \quad (29.5.2)$$

ここからは数学的な結果の解釈が絡んできて難しい。次のように考えて議論を組み立てよう。

A_n が十分に大きいとき、(29.5.2) は小さいはずだ。

もう少し噛み砕くと「公平な硬貨投げで大儲けすることは難しいだろう」という予想を立てたと言ってもいい。 A_n の取り方も結構面倒だが、 $\varepsilon > 0$ として $A_n = \varepsilon n$ と取ることにする。その上で次の式を証明しよう。

$$\lim_{n \rightarrow \infty} \mu \left(\left| \sum_{k=1}^n r_k(t) \right| > \varepsilon n \right) = 0. \quad (29.5.3)$$

いきなり絶対値が入ったが、あまり気にしないでほしい。「公平な硬貨投げなので、大きな得がないなら大きな損もないはずだ」(ローリスクローリターン) という気持を表わしている⁸。

⁷Rademacher 関数を適当に変換すれば他の金額設定もできないことはないはずだが、独立性などがうまくいくのが確認していない。

⁸これが成り立つような状況しか考えに入れていないといってもいい。実際に「大きな得はないが大損の可能性はある」状況はあるから。

式 (29.4.52) を使うことはできる. 実際次のように書ける.

$$\mu \left(\left| \sum_{k=1}^n r_k(t) \right| > \varepsilon n \right) = \sum_{|2l-n| > \varepsilon n} \mu \left(\sum_{k=1}^n r_k(t) = 2l - n \right) = \sum_{|2l-n| > \varepsilon n} \frac{1}{2^n} \binom{n}{l}. \quad (29.5.4)$$

このルートだと問題は次の式を示すことに帰着する.

$$\lim_{n \rightarrow \infty} \sum_{|2l-n| > \varepsilon n} \frac{1}{2^n} \binom{n}{l} = 0. \quad (29.5.5)$$

Stirling の公式を使って示すこともできるが面倒だ. Chebyshev による簡明かつ重要な結果を使った証明だけ紹介しておこう. 次のように変形していく.

$$\int_0^1 \left(\sum_{k=1}^n r_k(t) \right)^2 dt \geq \int_{\left| \sum_{k=1}^n r_k(t) \right| > \varepsilon n} \left(\sum_{k=1}^n r_k(t) \right)^2 dt > \varepsilon^2 n^2 \mu \left(\left| \sum_{k=1}^n r_k(t) \right| > \varepsilon n \right). \quad (29.5.6)$$

29.2.5 節を参考に計算すると

$$\int_0^1 \left(\sum_{k=1}^n r_k(t) \right)^2 dt = n \quad (29.5.7)$$

となるから,

$$\mu \left(\left| \sum_{k=1}^n r_k(t) \right| > \varepsilon n \right) < \frac{1}{\varepsilon^2 n} \quad (29.5.8)$$

が分かる. これで簡単に (29.5.3) が証明できた.

Chebyshev の不等式だけでなく, Rademacher 関数の独立性, つまり 29.2.6 の結果も本質的に効いている. 実は式 (29.5.3) が大数の弱法則だ³⁰. 弱法則ではまだまだほしい内容に足りない. 「儲けが際限なく大きくなる確率が 0 に収束する」くらいの微妙なことしか言っていないからだ. 平均がばっちり 0 に行くことを証明したい. これは次節で議論する大数の強法則として数学的に定式化されている.

29.5.2 大数の強法則

Borel は次の定理を示した.

Theorem. 5.2.1. (Borel) ほとんど全ての $t \in [0, 1]$ に対して

$$\lim_{n \rightarrow \infty} \frac{r_1(t) + r_2(t) + \cdots + r_n(t)}{n} = 0 \quad (29.5.9)$$

が成り立つ.

Rademacher 関数 $r_k(t)$ での「硬貨投げ」では期待値 (平均) 0 であることに注意してほしい. これはまさに, 算術平均がその「期待値」に収束することを言っている.

それと, 定理の言明の中で \sum を使っていないのは「何となく全部和で書いた方が気分が出る感じがする」という程度の理由で特に意味はない. また「ほとんど全て」という魔術的な言葉は測度論の専門用語だ. かなり難しいが測度論・確率論には本質的な言葉でもある. ここでは無視しよう: いったん全部の t で成り立つと思ってもらって構わない.

Proof. 証明は読むだけなら簡単だ. 自分でそんなものをすんなり思いつくか? と言われたらつらいところはある.

$$f_n(t) = \left(\frac{1}{n} \sum_{k=1}^n r_k(t) \right)^4 \quad (29.5.10)$$

として $\int_0^1 f_n(t) dt$ を考える. 4 乗をかつぎ出すというのが楽に進める上でとても大事なポイントになっている.

和を展開して 29.2.5 節の結果を使うと

$$\int_0^1 f_n(t) dt = \frac{n + \frac{4!}{2!2!} \binom{n}{2}}{n^4} \quad (29.5.11)$$

となる. 分子は全ての r_k が偶数個だけしか表われないときだけ抜き出した項の数だ. 29.2.5 節の結果から奇数次が出てくる場合の積分は消えてしまう.

分子は高々 2 次で分母が 4 次だから $\sum_{n=1}^{\infty} \int_0^1 f_n < \infty$ になる. したがって定理 2.6.1 から

$$\sum_{n=1}^{\infty} f_n \quad (29.5.12)$$

はほとんどいたる所収束する. 収束する級数を構成する数列は極限で 0 になるという一般的な定理があるから,

$$\lim_{n \rightarrow \infty} \left(\frac{r_1(t) + r_2(t) + \cdots + r_n(t)}{n} \right)^4 = 0 \quad (29.5.13)$$

になる. つまり

$$\lim_{n \rightarrow \infty} \frac{r_1(t) + r_2(t) + \cdots + r_n(t)}{n} = 0 \quad (29.5.14)$$

となる.

□

ここで

$$r_k(t) = 1 - 2\varepsilon_k(t) \quad (29.5.15)$$

だったから実は

$$\lim_{n \rightarrow \infty} \frac{\varepsilon_1(t) + \varepsilon_2(t) + \cdots + \varepsilon_n(t)}{n} = \frac{1}{2} \quad (29.5.16)$$

でもある. これは次のようにまとめられる.

Theorem. 5.2.2. ほとんど全ての実数 t に対してその 2 進展開に 0 と 1 は漸近的に同じ個数だけ表われる.

確率論的というと次のようになる.

繰り返し公平に硬貨を投げる. 各回の試行が独立ならば確率 1 (ほとんどいたる所の t にあたる表現) で表・裏が出る頻度 (相対度数) は極限で $1/2$ になる.

これが「(試行が独立なら) 表と裏が出る確率は同じ」という言明の数学的精密化であって, 29.4.5 節の辞書の正当性も主張している.

この話は正規数という解析数論の話題にも繋がっていく. 魔界の住人のような結果もある. 興味がある向きは [97] 2.2 節を読み, さらに解析数論にもアタックしてみてほしい.

29.6 中心極限定理

29.6.1 正規分布

まずは定義を書く. 平均 μ , 分散 σ^2 の正規分布というのは次の確率密度関数を持つ分布のことだ.

$$N(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right]. \quad (29.6.1)$$

適当な仮定, 特に独立性がある場合, 試行の確率分布が試行を無限回繰り返す極限で正規分布に漸近していくという言明が中心極限定理なのだった.

29.6.2 中心極限定理の正確な言明: Rademacher 関数版

まず中心極限定理 (の Rademacher 関数版) を数学的に正確に書いておこう.

Theorem. 6.2.1. Rademacher 関数に関して次の式が成り立つ.

$$\lim_{n \rightarrow \infty} \mu \left(\omega_1 < \frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} < \omega_2 \right) = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \quad (29.6.2)$$

上式右辺が正に平均 0, 分散 1 の正規分布だ.

これは Stirling の公式と組み合わせ的な議論からも導ける. しかしもっと確率論的にクリアな形で導出したい.

29.6.3 Markov の方法とその厳密化

Proof. 実数 $\omega_1 < \omega_2$ を固定しておき,

$$g(x) = \xi_A(x), \quad A = \{x : \omega_1 < x < \omega_2\} \quad (29.6.3)$$

とする. さらに補助関数を定義しよう. 式で書くと面倒だが図を描くとすぐ分かる. $0 < \varepsilon, 2\varepsilon < \omega_2 - \omega_1$ として

$$g_\varepsilon^+(x) = \begin{cases} \frac{x - (\omega_1 - \varepsilon)}{\varepsilon}, & x \in [\omega_1 - \varepsilon, \omega_1], \\ 1, & x \in [\omega_1, \omega_2], \\ -\frac{x - (\omega_2 + \varepsilon)}{\varepsilon}, & x \in [\omega_2, \omega_2 + \varepsilon], \\ 0, & \text{otherwise,} \end{cases} \quad (29.6.4)$$

$$g_\varepsilon^-(x) = \begin{cases} \frac{x - \omega_1}{\varepsilon}, & x \in [\omega_1, \omega_1 + \varepsilon], \\ 1, & x \in [\omega_1 + \varepsilon, \omega_2 - \varepsilon], \\ -\frac{x - \omega_2}{\varepsilon}, & x \in [\omega_2 - \varepsilon, \omega_2], \\ 0 & \text{otherwise.} \end{cases} \quad (29.6.5)$$

こうすると

$$g_\varepsilon^-(x) \leq g(x) \leq g_\varepsilon^+(x) \quad (29.6.6)$$

であり,

$$\int_0^1 g_\varepsilon^- \left(\frac{\sum_{k=1}^n r_k(t)}{\sqrt{n}} \right) dt \leq \mu \left(\omega_1 < \frac{\sum_{k=1}^n r_k(t)}{\sqrt{n}} < \omega_2 \right) \leq \int_0^1 g_\varepsilon^+ \left(\frac{\sum_{k=1}^n r_k(t)}{\sqrt{n}} \right) dt. \quad (29.6.7)$$

ところで ξ の関数 (g_ε^+ , g_ε^- の逆 Fourier 変換)

$$G_\varepsilon^-(\xi) := \int_{\mathbb{R}} g_\varepsilon^-(y) e^{iy\xi} dy, \quad G_\varepsilon^+(\xi) := \int_{\mathbb{R}} g_\varepsilon^+(y) e^{iy\xi} dy, \quad (29.6.8)$$

は絶対可積分になる. Fubini の定理が使えるから,

$$\lim_{n \rightarrow \infty} \int_0^1 g_\varepsilon^- \left(\frac{\sum_{k=1}^n r_k(t)}{\sqrt{n}} \right) dt = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\xi^2/2} \int_{\mathbb{R}} g_\varepsilon^-(y) e^{i\xi y} dy d\xi \quad (29.6.9)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g_\varepsilon^-(y) e^{-y^2/2} dy. \quad (29.6.10)$$

関数 g_ε^+ に対しても同じことが言えるので

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g_\varepsilon^-(y) e^{-y^2/2} dy \leq \liminf_{n \rightarrow \infty} \mu \left(\omega_1 < \frac{\sum_{k=1}^n r_k(t)}{\sqrt{n}} < \omega_2 \right) \quad (29.6.11)$$

$$\leq \limsup_{n \rightarrow \infty} \mu \left(\omega_1 < \frac{\sum_{k=1}^n r_k(t)}{\sqrt{n}} < \omega_2 \right) \quad (29.6.12)$$

$$\leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g_\varepsilon^+(y) e^{-y^2/2} dy \quad (29.6.13)$$

が任意の $\varepsilon > 0$ に対して成立する. したがって

$$\lim_{n \rightarrow \infty} \mu \left(\omega_1 < \frac{\sum_{k=1}^n r_k(t)}{\sqrt{n}} < \omega_2 \right) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(y) e^{-y^2/2} dy \quad (29.6.14)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy \quad (29.6.15)$$

が成り立つ. □

この証明を精査すると確率論的には分布関数という大事な概念に導かれる. その他のトピックへの展開も面白い. 興味がある向きは是非 [97] の 3 章を読んでほしい.

Chapter 30

Talk: Gaussian superprocess and its application to Quantum Field Theory: Sasakure Seminar

30.1 Introduction

Today we consider Gaussian superprocesses and its application to quantum field theory. We talk about construction of a probability space on the space of real tempered distributions, $\mathcal{S}'_{\text{real}}(\mathbb{R}^3)$ using a Gaussian superprocess. We are going to construct it from the Sobolev space $H^{-1/2}(\mathbb{R}^3)$.

Almost all of you may not know how large the space $L^2(\mathcal{S}'_{\text{real}}(\mathbb{R}^3))$. So we state a theorem on it. It relates to quantum field theory.

We set some notations. The function $\omega: \mathbb{R}^3 \rightarrow \mathbb{R}_{\geq}$ is $\omega(k) := \sqrt{k^2 + m^2}$, $m > 0$, and the real Hilbert space $H^{-1/2}(\mathbb{R}^3)$ is defined by

$$H^{-1/2}(\mathbb{R}^3) := \left\{ f \in \mathcal{S}'_{\text{real}}(\mathbb{R}^3) : \omega^{-1/2} \hat{f} \in L^2(\mathbb{R}^3) \right\}. \quad (30.1.1)$$

For any (separable) Hilbert space \mathcal{K} we set a new Hilbert space, called a (boson) Fock space [29, 28], by

$$\mathcal{F}_b(\mathcal{K}) := \bigoplus_{n=0}^{\infty} \bigotimes_s^n \mathcal{K}, \quad (30.1.2)$$

where \oplus means a direct sum, \otimes_s means a symmetric tensor product, and $\otimes_s^0 \mathcal{K} := \mathbb{C}$.

Then we can state a

Theorem. 1.0.1. ([25], Theorem 4.4.1) *We have a unitary equivalence $\mathcal{F}_b(L^2(\mathbb{R}^3)) = L^2(\mathcal{S}'_{\text{real}}(\mathbb{R}^3), d\mu)$, where $d\mu$ is a probability measure.*

Theorem. 1.0.2. *A L^2 space $L^2(\mathcal{S}'_{\text{real}}(\mathbb{R}^3), d\mu)$ can be constructed from a Gaussian superprocess indexed by $H^{-1/2}(\mathbb{R}^3)$.*

Note that Prof. Obata in University of Tohoku uses the above Fock space in his study of quantum probability [1].

30.1.1 Application to quantum field theory [29, 28]

The above Fock space \mathcal{F}_b is used in quantum field theory. We use the space $L^2(\mathbb{R}^3)$ for one particle **quantum mechanics**, use the space $L^2(\mathbb{R}^{3*2})$ for two particle quantum mechanics, and use the space $L^2(\mathbb{R}^{3n})$ for n particle quantum mechanics.

Since particles are created or annihilated in quantum field theory, i.e., a particle number is not a constant in quantum field theory, we must use the space for the space with variable particle number. For this, our space must include all subspaces for n particle spaces, and hence our particle space is the (boson) Fock space.

One more feeling for the need of tempered distributions

We use a Brownian motion for quantum mechanics, i.e., integration theory on the set of continuous functions $C(\mathbb{R}^3)$. Roughly speaking the set of tempered distributions $\mathcal{S}'_{\text{real}}$ is a dual space for a space $C(\mathbb{R}^3)$, i.e., $\mathcal{S}'_{\text{real}}$ is on $C(\mathbb{R}^3)$. Quantum field theory is one on quantum mechanics. So our space must be constructed as integration theory on the space over $C(\mathbb{R}^3)$, the space $\mathcal{S}'_{\text{real}}$.

Usage of Lévy process

We use Lévy processes for probabilistic representations of quantum mechanical systems. See [137] for details.

30.1.2 Application to geometry [29] chap. 6

Let \mathcal{H} and \mathcal{K} be separable Hilbert spaces and define a **Fermion Fock space** by

$$\mathcal{F}_f(\mathcal{K}) := \bigoplus_{n=0}^{\infty} \bigotimes_{\text{as}}^n \mathcal{K}, \quad (30.1.3)$$

where \bigotimes_{as}^n means an antisymmetric tensor product. We define the Hilbert space $\mathcal{F}(\mathcal{H}, \mathcal{K})$ by

$$\mathcal{F} := \mathcal{F}(\mathcal{H}, \mathcal{K}) := \mathcal{F}_b(\mathcal{H}) \bigotimes \mathcal{F}_f(\mathcal{K}). \quad (30.1.4)$$

This space relates to supersymmetric quantum field theory.

We can define a coboundary operator, a cochain complex, and p -th Laplace-Beltrami operator on \mathcal{F} . We can define p -th cohomology groups and prove a de Rham-Hodge-Kodaira decomposition. Furthermore we can define an infinite dimensional Dirac operator.

For finite dimensional Dirac operators celebrated index theory has developed. Its relation to supersymmetric quantum mechanics has been also studied. There exists some studies for infinite dimensional Dirac operators and its functional integral (path integral) representations for their indices. This suggests a geometric study, in particular, topological invariants for infinite dimensional spaces.

30.1.3 Application to analytic number theory [23]

We can derive the Dirichlet series and other arithmetic functions including the famous Riemann ζ , using quantum statistical partition functions. This relates to the Hilbert-Polya conjecture. Supersymmetric quantum theory and duality are encoded into many relations of functional equations of arithmetic functions.

30.2 Gaussian superprocess and its construction

30.2.1 Basics for probability

First we introduce some probabilistic notions.

Let (M, Σ, μ) be a probability space, i.e., Σ is a σ -algebra on a space M and μ is a probability measure on Σ . A real-valued measurable function on a measurable space (M, Σ) is called a **random variable**. For a random variable f we define a **expectation** for μ by

$$\mathbb{E}[f] := \int_M f d\mu. \quad (30.2.1)$$

A set of random variables of finite number $\{f_1, \dots, f_n\}$ is called **Gaussian** if its characteristic function has the following form:

$$C_{f_1, \dots, f_n}(t_1, \dots, t_n) := \mathbb{E} \left[\exp \left[i \sum_{i=1}^n t_i f_i \right] \right] = \exp \left[-\frac{1}{2} \sum_{i,j} a_{i,j} t_i t_j \right], \quad t_i \in \mathbb{R}, \quad (30.2.2)$$

where $A := \{a_{i,j}\}$ is a real positive definite matrix.

In the following we denote Gaussian random variables as GRV. One of GRVs' fundamental property is

Proposition. 2.1.1. *For a GRV $\{f_1, \dots, f_m\}$ we have*

$$\mathbb{E}[f_1 \cdots f_{2n}] = \sum \mathbb{E}[f_{i_1} f_{j_1}] \cdots \mathbb{E}[f_{i_n} f_{j_n}], \quad (30.2.3)$$

$$\mathbb{E}[f_1 \cdots f_{2n-1}] = 0. \quad (30.2.4)$$

Remark. 2.1.2. *A quasi-free state amounts to a representation of GRV in operator algebra.*

Wick product

Let f_1, \dots, f_n be mutually different random variables. We define their **Wick products** $: f_1^{n_1} \cdots f_k^{n_k} :$ recursively as follows:

$$: f_1^0 \cdots f_k^0 : = 1, \quad (30.2.5)$$

$$\frac{\partial}{\partial f_i} : f_1^{n_1} \cdots f_k^{n_k} : = \begin{cases} n_i : f_1^{n_1} \cdots f_i^{n_i-1} \cdots f_k^{n_k} : & (n_i \geq 1), \\ 0 & (n_i = 0), \end{cases} \quad (30.2.6)$$

$$\mathbb{E}[: f_1^{n_1} \cdots f_k^{n_k} :] = 0 \quad (n_1 + \cdots + n_k \neq 0). \quad (30.2.7)$$

We define an operator $\partial/\partial f_i$ is defined by

$$\frac{\partial}{\partial f_i} f_1^{n_1} \cdots f_k^{n_k} = \begin{cases} n_i f_1^{n_1} \cdots f_i^{n_i-1} \cdots f_k^{n_k} & (n_i \geq 1), \\ 0 & (n_i = 0), \end{cases} \quad (30.2.8)$$

and extend it by linearity for $\{f_i\}$ polynomials.

Proposition. 2.1.3. *For a GRV f we get*

$$: f^n : = \sum_{k=0}^{[n/2]} \frac{n!}{k!(n-2k)!} f^{n-2k} \left(-\frac{1}{2} E[f^2] \right)^k. \quad (30.2.9)$$

Proposition. 2.1.4. *Let f be a GRV. Then, for any $\alpha \in \mathbb{C}$,*

$$: e^{\alpha f} : := \sum_{n=0}^{\infty} \frac{: f^n :}{n!} \alpha^n \quad (30.2.10)$$

is defined by a.e. absolute convergence and it holds that

$$: e^{\alpha f} : = e^{-\frac{\alpha^2}{2} E[f^2]} e^{\alpha f}. \quad (30.2.11)$$

The RHS of the above converges in $L^p(M, d\mu)$ for all $p \in [1, \infty)$.

30.2.2 Gaussian superprocesses

Definition. 2.2.1. *Suppose (M, Σ, μ) is a probability space, H is a real Hilbert space, and $\phi: H \ni f \rightarrow \phi(f)$ is a map from H to random variables on M . The map ϕ is a **Gaussian superprocess with index H** if ϕ satisfies the following four conditions:*

1. (linearity) we have, for all $f, g \in H$ and $\alpha, \beta \in \mathbb{R}$,

$$\phi(\alpha f + \beta g) = \alpha \phi(f) + \beta \phi(g), \quad \text{a.e.} \quad (30.2.12)$$

2. $\{\phi(f) : f \in H\}$ is full, i.e., Σ is the smallest σ -algebra where enables $\{\phi(f)\}$ measurable.

3. each $\phi(f), f \in H$ is a GRV, i.e.,

$$E \left[e^{it\phi(f)} \right] = \exp \left[-\frac{1}{2} t^2 E [\phi(f)^2] \right], \quad t \in \mathbb{R}. \quad (30.2.13)$$

4. We have

$$E [\phi(f)\phi(g)] = \frac{1}{2} \langle \phi(f), \phi(g) \rangle, \quad f, g \in H, \quad (30.2.14)$$

where $\langle f, g \rangle$ is an inner product for H .

A probability space which realizes a Gaussian superprocess with index H clearly depends on the space H . Hence, in general, we denote a measure space Q_H , a probability measure μ_H , and a superprocess $\phi_H(f)$. If there is no confusion we simply denote them Q , μ , and $\phi(f)$.

In the following we consider its fundamental structures, uniqueness, existence.

Fundamental structures

Lemma. 2.2.2. Suppose $\{f_\alpha\}_{\alpha \in I}$ is a set of random variables on a probability space (M, Σ, μ) with index I and set

$$\mathcal{A} := \{F(f_{\alpha_1}, \dots, f_{\alpha_n}) : F \in \mathcal{S}(\mathbb{R}^n), \alpha_i \in I, n \geq 1\}. \quad (30.2.15)$$

A set $\{f_\alpha\}$ is full if and only if \mathcal{A} is dense in $L^2(M, d\mu)$.

Proof. We show only that fullness implies density. Denote $H = L^2(M, d\mu)$. We consider the following action of $F \in \mathcal{A}$ over $\Psi \in H$, i.e.,

$$(F\Psi)(m) := F(m)\Psi(m), \quad m \in M. \quad (30.2.16)$$

The set \mathcal{A} becomes a non-unital abelian $*$ -subalgebra of the operator algebra of all bounded operators on H . Then the strong operator topological closure \mathcal{A}_S is an abelian von Neumann algebra. In fact $\mathcal{A}_S = L^\infty(M)$. (You may need a proof: Since L^∞ is a von Neumann algebra, this is made of projections. Projections in it is defining functions for measurable sets of M . Thus we must approximate them by rapidly decreasing functions.)

Next we define $\tilde{\mathcal{A}}_S = \mathcal{A}_S 1$, where $1 \in H$ is a constant function. Clearly $\tilde{\mathcal{A}}_S \subset \bar{\mathcal{A}}$ and $\tilde{\mathcal{A}}_S = L^\infty(M)$. This proves $\bar{\mathcal{A}} = H$. \square

Theorem. 2.2.3. Let ϕ be a Gaussian superprocess indexed by a real Hilbert space H . We define closed subspaces $\Gamma_n(H)$, $n \geq 0$, of $L^2(Q, d\mu)$ as follows:

$$\Gamma_0(H) := \mathbb{C}, \quad (30.2.17)$$

$$\Gamma_n(H) := \overline{\{:\phi(f_1) \cdots \phi(f_n): : f_j \in H, j = 1, \dots, n\}}. \quad (30.2.18)$$

Then we have

$$L^2(Q, d\mu) = \bigoplus_{n=0}^{\infty} \Gamma_n(H). \quad (30.2.19)$$

In the proof we use the following interesting orthogonal properties.

Lemma. 2.2.4. Suppose $\{f_1, \dots, f_n, g_1, \dots, g_m\}$ is a GRV. Then we obtain

$$\mathbb{E}[:f_1 \cdots f_n : : g_1 \cdots g_m :] = \begin{cases} \sum_{\sigma \in \mathfrak{S}_n} \mathbb{E}[f_1 g_{\sigma(1)}] \cdots \mathbb{E}[f_n g_{\sigma(n)}] & (n = m), \\ 0 & (n \neq m). \end{cases} \quad (30.2.20)$$

Remark. 2.2.5. The above spaces $\Gamma_n(H)$ are ‘ n -particle subspaces’.

Uniqueness

Theorem. 2.2.6. Suppose ϕ_1 and ϕ_2 are Gaussian superprocesses for H and their probability spaces are $(Q_i, d\mu_i)$, $i = 1, 2$. Then there exists a unitary operator $U: L^2(Q_1) \rightarrow L^2(Q_2)$ satisfies

$$U1 = 1, \quad (30.2.21)$$

$$U\phi_1(f)U^{-1} = \phi_2(f), \quad f \in H. \quad (30.2.22)$$

Existence of a Gaussian superprocess for any separable H

Proof. Let H be a separable real Hilbert space and $\{e_n\}_{n=1}^\infty$ is a CONS. Set $\dot{\mathbb{R}}$ be the one-point compactification

$$\dot{\mathbb{R}} := \mathbb{R} \cup \{\infty\} \quad (30.2.23)$$

and $\mathbb{R}^{(n)}$ is a copy of $\dot{\mathbb{R}}$. Then

$$Q := \prod_{n=1}^{\infty} \mathbb{R}^{(n)} \quad (30.2.24)$$

is a compact Hausdorff space by Tychonoff's theorem. We denote $C(Q)$ is an abelian C^* -algebra of continuous functions on Q and $C_{\text{fin}}(Q)$ is a vector subspace of $C(Q)$ whose functions depends on finite variables. We set a functional l for $F \in C_{\text{fin}}(Q)$ by

$$l(F) := \frac{1}{\pi^{n/2}} \int_{\mathbb{R}^n} F(x_1, \dots, x_n) \exp \left[- \sum_{i=1}^n x_i^2 \right] dx_1 \cdots dx_n. \quad (30.2.25)$$

The functional l is a continuous linear functional on $C_{\text{fin}}(Q)$. By the way $C_{\text{fin}}(Q)$ is dense in $C(Q)$ thanks to celebrating Stone-Weierstrass' theorem. Hence the functional l can be uniquely extended to a functional on $C(Q)$; we denote it l too. Since $F \geq 0$ leads $l(F) \geq 0$ the functional l is positive. Riesz-Markov-Kakutani theorem shows the unique existence of the probability measure μ on Q such that

$$l(F) = \int_Q F d\mu. \quad (30.2.26)$$

We define a real valued function $\phi(e_n)$ on Q , for each e_n , by

$$\phi(e_n)(q) := x_n, \quad q = \{x_k\}_{k=1}^\infty \in Q. \quad (30.2.27)$$

Functions $\{\phi(e_n)\}$ are random variables on Q and clearly $\phi(e_n) \in L^2(Q, d\mu)$. Hence $\{\phi(e_n)\}_{n=1}^\infty$ is full. Expand $f \in H$ by $\{e_n\}$:

$$f = \sum_{n=1}^{\infty} \alpha_n e_n. \quad (30.2.28)$$

Set functions

$$\phi_n(f) := \sum_{k=1}^n \alpha_k \phi(e_k). \quad (30.2.29)$$

The sequence $\{\phi_n(f)\}$ converges in $L^2(Q)$ because, for $n > m$, we have

$$\|\phi_n(f) - \phi_m(f)\|^2 = \frac{1}{2} \sum_{k=m+1}^n \alpha_k^2 \rightarrow 0. \quad (30.2.30)$$

Denote the above limit function $\phi(f)$.

We can show

$$\mathbb{E}[\phi(f)\phi(g)] = \frac{1}{2} \langle f, g \rangle_H. \quad (30.2.31)$$

Since $\phi(f)$ is a GRV and the correspondence $f \rightarrow \phi(f)$ is linear, the constructed $\phi(\cdot)$ is a Gaussian superprocess indexed by H . \square

30.3 Concrete realization of the set Q

Finally we consider a useful realization of Q because we would like to describe Q concretely for concrete H . The reason why we want to use $\mathcal{S}'_{\text{real}}$ is a representation theory: a Minkowski space \mathcal{M}^4 and the Lorentz group actions are extended to a function space on it, and their actions are extended to the space of distributions. Their actions are closely related to relativistic causality, light cone analysis, and so on.

Let $\mathcal{S}_{\text{real}}(\mathbb{R}^\nu)$ be the subset of real valued rapid decreasing functions $\mathcal{S}(\mathbb{R}^\nu)$. For a Borel set B of \mathbb{R}^n and functions $f_1, \dots, f_n \in \mathcal{S}_{\text{real}}(\mathbb{R}^\nu)$, a subset in $\mathcal{S}'_{\text{real}}(\mathbb{R}^\nu)$

$$\{T \in \mathcal{S}'_{\text{real}}(\mathbb{R}^\nu) : (T(f_1), \dots, T(f_n)) \in B\} \quad (30.3.1)$$

is called a **cylinder set**. In the following the σ -algebra of $\mathcal{S}'_{\text{real}}(\mathbb{R}^\nu)$ is the minimal σ -algebra generated by the above cylinder sets.

Theorem. 3.0.7. (Minlos, Theorem 3.2.12 in [25], Theorem D.20 in [33]) Let C be a functional on $\mathcal{S}_{\text{real}}(\mathbb{R}^\nu)$. Then the existence of a probability measure on $\mathcal{S}'_{\text{real}}(\mathbb{R}^\nu)$ satisfying

$$C(f) = \int_{\mathcal{S}'_{\text{real}}(\mathbb{R}^\nu)} e^{iT(f)} d\mu(T), \quad f \in \mathcal{S}_{\text{real}}(\mathbb{R}^\nu) \quad (30.3.2)$$

is equivalent to the following three conditions:

1. $C(0) = 1$;
2. a map $f \rightarrow C(f)$ is continuous in the strong topology of $\mathcal{S}_{\text{real}}(\mathbb{R}^\nu)$;
3. C is of positive type, i.e., for each $f_i \in \mathcal{S}_{\text{real}}(\mathbb{R}^\nu)$ and $z_i \in \mathbb{C}$, $n = 1, 2, \dots$, we have

$$\sum_{i,j=1}^n \bar{z}_i z_j C(f_i - f_j) \geq 0. \quad (30.3.3)$$

Remark. 3.0.8. This theorem is an infinite dimensional version of finite dimensional Bochner's theorem.

Theorem. 3.0.9. Suppose a real Hilbert space H satisfies the following two conditions:

1. an inclusion $\mathcal{S}_{\text{real}}(\mathbb{R}^\nu) \subset H$ is dense;
2. embedding is continuous.

Then we can take $\mathcal{S}'_{\text{real}}(\mathbb{R}^\nu)$ as a measure space Q realizing a Gaussian superprocess $\phi(\cdot)$ indexed by H and $\phi(f)$ is given by

$$\phi(f)(T) = T(f), \quad T \in \mathcal{S}'_{\text{real}}. \quad (30.3.4)$$

Remark. 3.0.10. Real α -th Sobolev spaces $H^\alpha(\mathbb{R}^\nu)$ are examples of Hilbert spaces satisfying the condition in the above theorem.

Proof. A characteristic function of a Gaussian superprocess indexed by H is given by

$$C(f) := \mathbb{E} \left[e^{i\phi(f)} \right] = \exp \left[-\frac{1}{4} \|f\|^2 \right]. \quad (30.3.5)$$

Since $\mathcal{S}_{\text{real}}(\mathbb{R}^\nu) \subset H$, $C(\cdot)$ is a functional on $\mathcal{S}_{\text{real}}(\mathbb{R}^\nu)$ and $C(0) = 1$. By the assumption 2, a correspondence $f \rightarrow C(f)$ is continuous in the strong topology of $\mathcal{S}_{\text{real}}(\mathbb{R}^\nu)$. Clearly $C(f)$ is of positive type. Hence, by Minlos' theorem, there exists a probability measure μ on $\mathcal{S}'_{\text{real}}(\mathbb{R}^\nu)$ such that

$$C(f) = \int_{\mathcal{S}'_{\text{real}}(\mathbb{R}^\nu)} e^{iT(f)} d\mu(T), \quad f \in \mathcal{S}_{\text{real}}(\mathbb{R}^\nu). \quad (30.3.6)$$

Define $\phi(f)(T) := T(f)$, $T \in \mathcal{S}'_{\text{real}}$, $f \in \mathcal{S}_{\text{real}}$. Then $\phi(f)$ becomes a GRV on $(\mathcal{S}'_{\text{real}}(\mathbb{R}^\nu), d\mu)$ and satisfies

$$\mathbb{E}[\phi(f)\phi(g)] = \frac{1}{2}\langle f, g \rangle, \quad f, g \in \mathcal{S}_{\text{real}}(\mathbb{R}^\nu). \quad (30.3.7)$$

The assumption 1 shows that we can define $\phi(f)$ for $f \in H$ by

$$\phi(f) := \lim_{n \rightarrow \infty} \phi(f_n), \quad \|f_n - f\|_H \rightarrow 0. \quad (30.3.8)$$

Since $|e^{ix} - e^{iy}| \leq |x - y|$, we obtain

$$\lim_{n \rightarrow \infty} \mathbb{E}[e^{i\phi(f_n)}] = \mathbb{E}[e^{i\phi(f)}], \quad \lim_{n \rightarrow \infty} \exp\left[-\frac{1}{4}\|f_n\|^2\right] = \exp\left[-\frac{1}{4}\|f\|^2\right]. \quad (30.3.9)$$

Hence $\phi(f)$ is a GRV. \square

30.4 Application to quantum field theory and relation to operator algebraic formulation

If having time.

30.5 Appendix

30.5.1 Stone-Weierstrass theorem

We consider the algebra $C_0(X, \mathbb{C})$ of complex-valued continuous functions on a locally compact Hausdorff space X with the topology of uniform convergence. This is a C^* -algebra with the $*$ -operation given by pointwise complex conjugation.

Theorem. 5.1.1. *Let X be a locally compact Hausdorff space and let A be a subset of $C_0(X, \mathbb{C})$ which satisfies the following two conditions:*

1. *for any pair $x, y \in X$, $x \neq y$ there exists a function $p \in A$ which satisfies $p(x) \neq p(y)$;*
2. *for any point $x \in X$ there exists a function $p \in A$ which satisfies $p(x) \neq 0$.*

Then the complex unital $$ -algebra generated by A is dense in $C_0(X, \mathbb{C})$.*

30.5.2 Minlos' theorem

First we review Bochner' theorem. A function $\phi: \mathbb{R}^n \rightarrow \mathbb{C}$ is of positive type if, for any $x_1, \dots, x_n \in \mathbb{R}^n$ and any $z_1, \dots, z_n \in \mathbb{C}$,

$$\sum_{i,j=1}^n \bar{z}_i z_j \phi(x_i - x_j) \geq 0. \quad (30.5.1)$$

Theorem. 5.2.1. *(Bochner, [90], Section 30) A function $\phi: \mathbb{R}^n \rightarrow \mathbb{C}$ is of positive type with $\phi(0) = 1$ if and only if ϕ has the form*

$$\phi(x) = \int_{\mathbb{R}^n} e^{2\pi i \langle x, k \rangle} d\mu(k), \quad (30.5.2)$$

where μ is a Borel probability measure on \mathbb{R}^n .

Minlos' theorem is an infinite dimensional version of Bochner. One of its proof uses Bochner directly: we use a finite dimensional net and apply it for each finite dimensional subspaces. Then a net limit functional is the desired one.

There are two proofs in [33]. One is abstract and uses **nuclear extension**. The other is concrete and uses a **quantum mechanical harmonic oscillator**

Chapter 31

Talk: Lecture on ergodic theory

This note is based on a course for third-year Bachelor students in the May of 2014. See [185, 97] for details.

Chapter 32

Statistics for idol management

動画から適当に切り出してくる.

Part IX

Algebra

Chapter 33

対称性と群論

33.1 運動方程式を不変にする変換

33.2 エネルギーを不変にする変換

33.3 群

33.4 Lorentz 群と Maxwell 方程式

33.5 解析力学小論

Part X

Geometry

Part XI

Interesting counter examples

Chapter 34

DVD: Playing with interesting counter examples

DVD の原稿をもう少し整理して張る.

Part XII

Applied mathematics

Chapter 35

Digital signal processing and algebraic geometry

動画から適当に切り出してくる.

Chapter 36

Peano curve and its application to engineering

動画から適当に切り出してくる.

Part XIII

Quantum mechanics and Quantum field theory

Part XIV

Statistical Mechanics and condensed matter physics

Chapter 37

LASER: song by electrones

動画から適当に切り出してくる.

Chapter 38

Analysis for the Hubbard model

動画から適当に切り出してくる.

Part XV

Physical misc

Chapter 39

Similarity for hydrodynamics and general relativity

動画から適当に切り出してくる.

Part XVI

Reference

Index

- ある (m, n) 型行列によって定まる線型写像, 55
Archimedes の原理, 90
安定性の条件, 112
安定なつりあい, 87
鞍点, 88
 ε -近傍, 94
 ε - δ 論法, 89
上三角行列, 45
上に有界, 91
運動の第 1 法則, 52
運動の第 3 法則, 52
運動の第 2 法則, 52
運動方程式, 52
 n 項単位ベクトル, 44
Hermite 行列, 45
Euler の公式, 40, 98, 133
開集合, 93
回転, 50
解の最大存在時刻, 125
下界, 91
角振動数, 108
掛け算作用素, 132
下限, 91
重ね合わせの原理, 108, 127
形不変性, 51
関数, 46
関数空間, 74
関数値, 46
慣性系, 52
慣性座標系, 52
慣性の法則, 52
外積, 65
Galilei 共変性, 53
Galilei の相対性原理, 53
Galilei ブースト, 53
Galilei 変換, 52
基準振動, 113
基底, 66, 77
(常微分方程式の) 基本解系, 130
狭義の極大, 109
共変性, 48
極限, 89
極小, 109
局所解, 123
局所 Lipschitz 条件, 121
極大, 109
極値, 109
極値点, 109
逆行列, 44
行列式, 58
空集合, 43
Kronecker の δ , 44
元, 43
減衰振動, 133
交代性, 47
固有値, 67
固有値問題, 80
固有ベクトル, 67
Cauchy の折れ線, 118
Cauchy 列, 90
最小値, 91
最大値, 91
作用素, 46
作用・反作用の法則, 52
座標ベクトル, 57
自然数, 43
下三角行列, 45
下に有界, 91
写像, 43
集合, 43
集合族, 93
収束, 89
主小行列式, 110
 C^k 級, 102
次元, 67, 76
事象, 51
実数, 43
上界, 91
上限, 91
数列, 89
スペクトル分解, 71, 109

随伴行列, 45
 正規行列, 46
 正規直交基底, 67
 正項級数, 92
 斉次解, 130
 斉次方程式, 127
 整数, 43
 正則, 44, 110
 正則行列, 44
 正值, 110
 正定値性, 78
 正の実数, 43
 正の整数, 43
 積分, 115
 線型化, 88, 94, 99
 線型関係, 66
 線型空間, 73
 線型結合, 66
 線型写像, 47
 線型従属, 66
 線型独立, 66, 76, 128
 線型変換, 50
 零行列, 44
 零ベクトル, 73
 対角化, 70
 対角行列, 45
 対角成分, 45
 対称行列, 45
 多重線型性, 58
 多様体, 54
 単位行列, 44
 重畳原理, 127
 調和振動子, 113
 直和, 46
 直交行列, 45, 50
 直交変換, 51
 直交補空間, 46
 定数変化法, 131
 Taylor 展開, 97
 Taylor の公式, 96
 Taylor の定理, 96, 106
 停留値, 109
 転置行列, 44
 δ 関数, 108
 特殊解, 130
 凸集合, 121
 導関数, 95, 100
 導値, 95, 100
 内積, 65, 78
 内点, 94
 2 次形式, 110
 ノルム, 78
 波数, 108
 波動方程式, 108
 非斉次ベクトル, 53
 非斉次方程式, 127
 標準基底, 55, 66
 微分, 106
 微分可能, 95, 100
 微分が消える, 109
 微分形式, 55
 微分係数, 95, 100
 微分積分方程式, 132
 微分方程式, 114, 115
 微分方程式を解く, 115
 複素数, 43
 負値, 110
 不定符号, 110
 Fourier, 80
 Fourier 級数, 80
 Fourier 逆変換, 108
 Fourier 変換, 80, 108
 物理法則, 47
 物理量, 48
 部分空間, 46
 部分集合, 43
 変換, 46
 偏導関数, 101
 偏導値, 101
 偏微分可能, 101
 偏微分係数, 101
 偏微分方程式, 115
 べき級数, 92
 方向微分, 100
 ポテンシャル, 87
 (微分方程式の解が) 右に延長できる, 125
 有界, 91
 有理数, 43
 ユニタリ行列, 45
 (時間に) 陽によらない, 53
 4 元ベクトル, 53
 Lipschitz 条件, 117
 連続の公理, 91
 ロンスキアン, 128

Bibliography

- [1] 類似明出伊, 伸明尾畑. 『量子確率論の基礎』. 牧野書店, 9 2003.
- [2] L. V. Ahlfors, 乾吉笠原. 『複素解析』. 現代数学社, 3 1982.
- [3] M. Aigner and G. Ziegler. 『天書の証明』. 丸善出版, 9 2012.
- [4] 康夫秋月. 『近代数学の展望』. 筑摩書房, 12 2009.
- [5] J. Almira and E. Linares. An application of the gelfand-mazur theorem: the fundamental theorem of algebra revisited. *Divulgaciones Matemáticas*, Vol. 13, No. 2, pp. 123–125, 2005.
- [6] Allen Altman and Steven Kleiman. *A Term of Commutative Algebra*. Worldwide center of mathematics, 4 2013.
- [7] 哲哉安藤. 『不等式 21 世紀の代数的不等式論』. 数学書房, 9 2012.
- [8] 和彦青本. 『直交多項式入門』. 数学書房, 4 2013.0.
- [9] A. Arai. Essential spectrum of a self-adjoint operator on an abstract hilbert space of fock type and applications to quantum field hamiltonians. *J. Math. Anal. Appl.*, pp. 189–216, 2000.
- [10] A. Arai. Ground state of the massless nelson model without infrared cutoff in a non-fock representation. *Rev. Math. Phys.*, Vol. 13, pp. 1075–1094, 2001.
- [11] A. Arai. Ground state of the massless nelson model without infrared cutoff in a non-fock representation. *Rev. Math. Phys.*, pp. 1075–1094, 2001.
- [12] A. Arai. *Mathematical Principles of Quantum Statistical Mechanics (in Japanese)*. Kyoritsu-syupan, 2008.
- [13] A. Arai and M. Hirokawa. On the existence and uniqueness of ground states of generalized spin-boson model. *J. Funct. Anal.*, Vol. 151, pp. 455–503, 1997.
- [14] Asao Arai. On a model of a harmonic oscillator coupled to a quantized, massless scalar field i. *J. Math. Phys.*, Vol. 22, pp. 2539–2548, 1981.
- [15] Asao Arai. On a model of a harmonic oscillator coupled to a quantized, massless scalar field ii. *J. Math. Phys.*, Vol. 22, pp. 2549–2552, 1981.

- [16] Asao Arai. Rigorous theory of spectra and radiation for a model in quantum electrodynamics. *J. Math. Phys.*, Vol. 24, pp. 1896–1910, 1983.
- [17] Asao Arai. An asymptotic analysis and its applicaiton to the nonrelativistic limit of the pauli-fierz and a spin-boson model. *J. Math. Phys.*, Vol. 31, pp. 2653–2663, 1990.
- [18] Asao Arai. Perturbation of embedded eigenvalues: a general class of exactly soluble models in fock spaces. *Hokkaido Math. Jour.*, Vol. 31, pp. 1–34, 1990.
- [19] Asao Arai. Spectral analysis of a quantum harmonic oscillator coupled to infinitely many scalar bosons. *J. Math. Anal. Appl.*, Vol. 31, pp. 270–288, 1990.
- [20] Asao Arai. Non-invertible bogoliubov transformations and instability of embedded eigenvalues. *J. Math. Phys.*, Vol. 32, pp. 1838–1846, 1991.
- [21] Asao Arai. Essential spectrum of a self-adjoint operator on an abstract hilbert space of fock type and applications to quantum field hamiltonians. *J. Math. Anal. Appl.*, Vol. 246, pp. 189–216, 2000.
- [22] Asao Arai. *Fock spaces and quantum fields (in Japanese)*. Nihon-hyoronsha, 2000.
- [23] Asao Arai. *Infinite-Dimensional Analysis and Analytic Number Theory*, Vol. 63. 2000.
- [24] Asao Arai. Infinite dimensional analysis and analytic number theory. *Acta Applicandae Mathematica*, Vol. 63, pp. 41–78, 9 2000.
- [25] 朝雄新井, 洋江沢. 『場の量子論と統計力学』. 日本評論社, 6 1988.
- [26] 朝雄新井, 洋江沢. 『量子力学の数学的構造 II』. 朝倉書店, 7 1999.
- [27] 朝雄新井, 洋江沢. 『量子力学の数学的構造 I』. 朝倉書店, 7 1999.
- [28] 朝雄新井. 『フォック空間と量子場 下』. 数理物理シリーズ. 日本評論社, 8 2000.
- [29] 朝雄新井. 『フォック空間と量子場 上』. 数理物理シリーズ. 日本評論社, 8 2000.
- [30] 朝雄新井. 『量子現象の数理』. 朝倉物理学体系. 朝倉書店, 2 2006.
- [31] 朝雄新井. 『物理の中の対称性—現代数理物理学の観点から』. 日本評論社, 1 2008.
- [32] 朝雄新井. 『量子統計力学の数理』. 共立出版, 7 2008.
- [33] 朝雄新井. 『量子数理物理学における汎関数積分法』. 共立出版, 8 2010.
- [34] 晋司新井. 『もし性暴力の被害にあったら』, 6 1999.
- [35] H. Araki, R. Haag, D. Kastler, and M. Takesaki. Extension of kms states and chemical potential. *Comm. Math. Phys.*, Vol. 53, No. 2, pp. 97–134, 1977.

- [36] M. Atiyah and I. Macdonald. *Introduction to Commutative Algebra*. Westview Press, 2 1994.
- [37] T. Aubin. *Nonlinear Analysis on Manifolds. Monge-Ampère Equations*. Springer New York, 12 1982.
- [38] Simon B. *Functional Integration And Quantum Physics*. Chelsea Pub Co., 12 2004.
- [39] J. Baez, I. E. Segal, and Z. Zhou. *Introduction to Algebraic and Constructive Quantum Field Theory*. Princeton Univ. Pr., 1992.
- [40] H. Baumgartel. *Operator Algebraic Methods in Quantum Field Theory*. Vch Pub, 10 1995.
- [41] B. Blackadar. *K-Theory for Operator Algebras*. Cambridge University Press, 9 1998.
- [42] J. Bost and A. Connes. Hecke algebras, type iii factors and phase transitions with spontaneous symmetry breaking in number theory. *Selecta Math*, Vol. 1, No. 3, pp. 411–457, 1995.
- [43] O. Bratteli and D. Robinson. *Operator Algebras and Quantum Statistical Mechanics*, Vol. 1 of *Theoretical and Mathematical Physics*. Springer Berlin Heidelberg, 11 2010.
- [44] O. Bratteli and D. Robinson. *Operator Algebras and Quantum Statistical Mechanics*, Vol. 2 of *Theoretical and Mathematical Physics*. Springer Berlin Heidelberg, 7 2013.
- [45] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 11 2011.
- [46] H. Brezis, 宏藤田, 芳雄小西. 『関数解析 その理論と応用に向けて』. 産業図書, 10 1988.
- [47] D. Buchholz and H. Grundling. *Quantum Systems and Resolvent Algebras*. 6 2013.
- [48] D. Buchholz and S. Summers. Scattering in relativistic quantum field theory: Fundamental concepts and tools. *arxiv:0509047*, pp. 1–14, 2005.
- [49] 忠躬遅塚. 『フランス革命 歴史における劇薬』. 岩波書店, 12 1997.
- [50] M. Damak and V. Georgescu. On the spectral analysis of many-body systems. *arxiv:0911.5126*, pp. 1–52, 2009.
- [51] 悦朗伊達, 哲二三輪, 道夫神保. 『ソリトンの数理』. 岩波書店, 2 2007.
- [52] Florence N. David. 『確率論の歴史—遊びから科学へ』. 海鳴社, 1975.
- [53] J. Dereziński. Introduction to representations of the canonical commutation and anticommutation relations. *arxiv:0511030*, pp. 1–79, 2005.
- [54] J. Dereziński and V. Jakšić. Spectral theory of pauli-fierz operators. *J. Func. Anal.*, pp. 243–327, 2001.

- [55] J. Dereziński and V. Jakšić. Return to equilibrium for pauli-fierz systems. *Ann. Henri Poincaré*, pp. 739–793, 2003.
- [56] J. Dereziński, V. Jakšić, and A. Pillet. Perturbation theory of w^* -dynamics, liouvilleans and kms-states. *Rev. Math. Phys*, Vol. 15, pp. 447–489, 2003.
- [57] W. Dunham, 重雄一樂, 敏實川. 『微積分名作ギャラリー ニュートンからルベークまで』. 日本評論社, 11 2009.
- [58] Chet Van Duzer. *Sea Monsters on Medieval and Renaissance Maps*. British Library Board, 6 2013.
- [59] W. Dybalski. Spectral theory of automorphism groups and particle structures in quantum field theory. *arxiv:0901.3127v1*, 2009.
- [60] W. Dybalski and A. Pizzo. Coulomb scattering in the massless nelson model i. foundations of two-electron scattering. *arxiv:1302.5001*, pp. 1–39, 2 2013.
- [61] W. Dybalski and A. Pizzo. Coulomb scattering in the massless nelson model ii. regularity of ground states. *arxiv:1302.5012*, pp. 1–45, 2 2013.
- [62] Lieb H. E. Thomas-fermi and related theories of atoms and molecules. *Rev. Modern Phys.*, Vol. 311, pp. 603–641, 1981.
- [63] 洋江沢. 『理科が危ない 明日のために』. 新曜社, 6 2001.
- [64] 洋江沢. 『だれが原子をみたか』. 岩波書店, 1 2013.
- [65] M. Fannes, B. Nachtergaele, and A. Verbeure. The equilibrium states of the spin-boson model. *Commun. Math. Phys.*, Vol. 114, pp. 537–548, 1988.
- [66] J. Fröhlich, M. Merkli, and I. M. Sigal. Ionization of atoms in a thermal field. *arxiv:0410012*, pp. 1–53, 8 2007.
- [67] Kazuhiro Fujiwara and Fumiharu Kato. *Foundations of Rigid Geometry*. 8 2013.
- [68] 賢治深谷. 『数学者の視点』. 岩波書店, 1 1996.
- [69] 賢治深谷. 『解析力学と微分形式』. 岩波書店, 4 2004.
- [70] 賢治深谷. 『電磁場とベクトル解析』. 岩波書店, 1 2004.
- [71] 直久舟木, 耕平内山. 『ミクロからマクロへ 1 界面モデルの数理』. シュプリンガー・フェアラーク東京, 4 2002.
- [72] 直久舟木. 『確率論』. 朝倉書店, 11 2004.
- [73] 直久舟木. 『書評 伊藤清の数学』. 2011.
- [74] 直久舟木, 耕平内山. 『ミクロからマクロへ 2 格子気体の流体力学極限』. 丸善出版, 8 2012.
- [75] 幹雄古田. 『大学院で幾何の勉強を目指す学部生の方たちへ』. 4 2007.

- [76] B. Gelbaum and J. Olmsted. *Counterexamples in Analysis*. Dover, 6 2003.
- [77] Sheldon Goldstein, Joel L. Lebowitz, Christian Mastrodonato, Roderich Tumulka, and Nino Zanghi. Canonical typicality. *arxiv:0511091v2*, pp. 1–6, 2005.
- [78] Sheldon Goldstein, Joel L. Lebowitz, Christian Mastrodonato, Roderich Tumulka, and Nino Zanghi. On the approach to thermal equilibrium of macroscopic quantum systems. *arxiv:0911.1724v1*, pp. 1–19, 2009.
- [79] Daniel M. Gusfield and Robert W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. The MIT Press, 1 2003.
- [80] D. Hasler and I. Herbst. *Ground States in the Spin Boson Model*. arXiv, 8 2010.
- [81] 文雄日合, 研二郎柳. 『ヒルベルト空間と線型作用素』. 牧野書店, 7 1995.
- [82] 武平井. 『線形代数と群の表現 II』. 朝倉書店, 11 2001.
- [83] 武平井. 『線形代数と群の表現 I』. 朝倉書店, 11 2001.
- [84] M. Hirokawa. On the ground state energy of the spin-boson model without infrared cuto and the superradiant ground state of the wigner-weisskopf model. *RIMS*, Vol. 1156, pp. 85–105, 2000.
- [85] M. Hirokawa, F. Hiroshima, and J. Lörinczi. *Spin-Boson Model through a Poisson-Driven Stochastic Process*. arXiv, 9 2012.
- [86] 重雄一樂. 『集合と位相 そのまま使える答えの書き方』. 講談社サイエンティフィク, 4 2001.
- [87] 康隆伊原. 『志学数学 -研究の諸段階 発表の工夫』. 丸善出版, 7 2012.
- [88] 茂飯高, 幸彦浪川, 健爾上野. 『デカルトの精神と代数幾何』. 日本評論社, 10 1993.
- [89] 誠一郎井上, 拓郎土江, 将太井上. 『パーフェクト JavaScript』. 技術評論社, 9 2011.
- [90] 清三伊藤. 『ルベーグ積分入門』. 裳華房, 4 1963.
- [91] 秀一伊藤. 『常微分方程式と解析力学』. 共立出版, 1 1998.
- [92] 恭雄岩永, 眞久佐藤. 『環と加群のホモロジー代数的理論』. 日本評論社, 10 2002.
- [93] V. Jakšić and A. Pillet. On a model for quantum friction, i. fermi’s golden rule and dynamics at zero temperature. *Ann. Inst. Henri. Poincaré*, pp. 47–68, 1995.
- [94] V. Jakšić and A. Pillet. On a model for quantum friction, ii. fermi’s golden rule and dynamics at positive temperature. *Comm. Math. Phys.*, pp. 619–644, 1996.

- [95] V. Jakšić and A. Pillet. On a model for quantum friction iii. ergodic properties of the spin-boson system. *Comm. Math. Phys.*, pp. 627–651, 1996.
- [96] Mark Kac. Can one hear the shape of a drum? *American Mathematical Monthly*, Vol. 73, pp. 1–23, 1966.
- [97] M. Kac, 陽一郎高橋, 眞澄中嶋. 『Kac 統計的独立性』. 数学書房, 4 2011.
- [98] R. Kadison and J. Ringrose. *Fundamentals of the Theory of Operator Algebras*. American Mathematical Society, 7 1997.
- [99] R. Kadison and J. Ringrose. *Fundamentals of the Theory of Operator Algebras: Advanced Theory*. American Mathematical Society, 7 1997.
- [100] 謙二郎上條. 『ロケットターボポンプの研究・開発 35 年間の思い出』. 東北大学出版会, 3 2013.
- [101] 健一金谷. 『これなら分かる応用数学教室 最小二乗法からウェーブレットまで』. 共立出版, 6 2003.
- [102] 正樹柏原. 『代数解析概論』. 岩波書店, 3 2008.
- [103] 正樹柏原, 隆裕河合, 達雄木村. 『代数解析学の基礎』. 紀伊国屋, 11 2008.
- [104] 和也加藤. 『 p 進 hodge 理論とゼータの値』.
- [105] 利行桂. 『代数幾何入門』. 共立出版, 10 1998. ろくに読んでいないが, 代数幾何の符号理論への応用も書いてある. 符号理論への応用が面白そう.
- [106] 利行桂. 『代数学 I 群と環』. 東京大学出版会, 3 2004.
- [107] 利行桂. 『代数学 III 体とガロア理論』. 東京大学出版会, 9 2005.
- [108] 利行桂. 『代数学 II 環上の加群』. 東京大学出版会, 3 2007.
- [109] Victor J. Katz. 『カツ 数学の歴史』. 共立出版, 7 2005.
- [110] 隆裕河合, 義次竹井. 『特異摂動の代数解析学』. 岩波書店, 5 2008.
- [111] みゆき川村. 『多面体の折紙 正多面体・準正多面体およびその双対』. 日本評論社, 11 1995.
- [112] 愛川添. 『白と黒のとびら オートマトンと形式言語をめぐる冒険』. 東京大学出版会, 4 2013.
- [113] S. Khaleelulla. *Counterexamples in Topological Vector Spaces*. Springer Berlin Heidelberg, 7 1982.
- [114] 成煥金, 昌宏山本. 『熱方程式で学ぶ逆問題 Fourier 解析 関数解析から数値解析まで』. サイエンス社, 3 2008.
- [115] 達雄木村. 『佐藤幹夫の数学』. 日本評論社, 8 2007.
- [116] Donald E. Knuth. *Stable Marriage and Its Relation to Other Combinatorial Problems: An Introduction to the Mathematical Analysis of Algorithms*. Amer. Mathematical Society, 10 1996.

- [117] 昭七小林. 『複素幾何』. 岩波書店, 9 2005. 幾何の人に聞いたら, 「話題は豊富だが読みづらい. ただ, このくらいのページ数でこれだけ色々書いてある本もなかなかないから, そういう意味では役に立つ. 」とのことだった.
- [118] 昭七小林. 『顔をなくした数学者—数学つれづれ』. 岩波書店, 7 2013.
- [119] 俊行小林, 利雄大島. 『リー群と表現論』. 岩波書店, 4 2005.
- [120] 天文台国立. 『理科年表 平成 25 年版 机上版』. 丸善出版, 11 2013.
- [121] 天文台国立. 『理科年表 平成 25 年版』. 丸善出版, 11 2013.
- [122] A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis*. Dover, 6 1975.
- [123] 彦三郎小松. 『佐藤超函数論入門』. 数理解析研究所講究録, pp. 1–174, 10 1973.
- [124] T. Koornwinder. Orthogonal polynomials, a short introduction. *arXiv:1303.2825*, 3 2013.
- [125] ことりん. 『偏微分方程式のお話 解の存在について』. 関西すうがく徒のつどい, pp. 1–25, 3 2012.
- [126] 幹雄久保, 知己松井, 明久田村. 『応用数理計画ハンドブック』. 朝倉書店, 5 2002.
- [127] 治樹倉澤. 『量子力学』. 10 2013.
- [128] Clement Lam. The search for a finite projective plane of order 10. pp. 305–318.
- [129] Amy Langville, Carl Meyer, 和生岩野, 利明黒川, 洋黒川. 『Google PageRank の数理 最強検索エンジンのランキング手法を求めて』. 共立出版, 10 2009.
- [130] E. H. Lieb and M. Loss. *Analysis*. Amer. Math. Soc., 4 2001.
- [131] E. H. Lieb and R. Seiringer. *The Stability of Matter in Quantum Mechanics*. Cambridge University Press, 11 2009.
- [132] E. H. Lieb, R. Seiringer, J. Solovej, and J. Yngvason. *The Mathematics of the Bose Gas and its Condensation (Oberwolfach Seminars)*. Birkhaeuser Basel, 7 2005.
- [133] E. H. Lieb and B. Simon. Thomas-fermi theory of atoms, molecules and solids. *Adv. in Math.*, Vol. 23, pp. 22–116, 1977.
- [134] E. H. Lieb, W. Thirring, and F. Dyson. *The Stability of Matter: From Atoms to Stars Selecta of Elliott H. Lieb*. Springer, 12 2005.
- [135] E. H. Lieb and J. Yngvason. The entropy concept for non-equilibrium states. *arxiv:1305.3912*, pp. 1–23, 2013.
- [136] J. Littlewood. 『不等式』. シュプリンガーフェアラーク東京, 8 2012.

- [137] J. Lórinzi, F. Hiroshima, and V. Betz. *Feynman-Kac-Type Theorems and Gibbs Measures on Path Space: With Applications to Rigorous Quantum Field Theory*. Walter De Gruyter, 6 2011.
- [138] 幸夫松本. 『多様体の基礎』. 岩波書店, 9 1988. 議論がぐどくらい丁寧. もう少し程よく省略してもいいくらいだが, 初学者または非専門家にはむしろこのくらいがいいのだろうとも思う. 実多様体がメインで複素多様体はほとんどない. このくらいの丁寧さを『Morse 理論』にも期待したかった.
- [139] 英之松村. 『可換環論』. 共立出版, 9 2000.
- [140] 和夫松坂. 『集合・位相入門』. 岩波書店, 6 1968.
- [141] R. Melrose. 『境界つき多様体のディラック作用素』. 丸善出版, 7 2012.
- [142] John Milnor. *Morse Theory*.
- [143] John Willard Milnor. *Topology from the Differentiable Viewpoint*. Princeton Univ Pr, 11 1997.
- [144] 典克三尾. 『電磁気学第 2 参考資料』.
- [145] Charles W. Misner, Kip S. Thorne, and John A. Wheeler. *Gravitation*. W H Freeman and Co, 9 1973.
- [146] 穀森. 『数学史入門』. 筑摩書房.
- [147] 光生森本. 『復刊 佐藤超函数入門』. 共立出版, 9 2000.
- [148] 亮介長岡. 『改訂版 数学の歴史』. 放送大学教育振興会, 4 2003.
- [149] 啓中嶋. 『非線形問題と複素幾何学』. 10 2008.
- [150] 徹中村. 『超準解析と物理学』. 日本評論社, 6 1998.
- [151] 美知代中根. 『 ϵ - δ 論法とその形成』. 共立出版, 7 2010.
- [152] 利一夏目, 仁志森吉. 『作用素環と幾何学』. 数学メモワール, 6 2001.
- [153] J. v. Neumann, 徹広重, 健井上, 敏彦恒藤. 『量子力学の数学的基礎』. みすず書房, 11 1957.
- [154] 日本数学協会. 『数学文化 特集 高木貞治没後 50 周年』. 14. 日本数学協会, 7 2010.
- [155] 数学会日本. 『原田耕一郎氏 有限単純群の研究』. 日本数学会, 2000.
- [156] Weaver Nik. *Forcing for Mathematicians*. World Scientific Pub Co Inc, 6 2014.
- [157] 青季西川. 『幾何学的変分問題』. 岩波書店, 4 2006.
- [158] 利雄西野. 『多変数函数論』. 東京大学出版会, 11 1996.
- [159] 敏雄丹羽. 『微分方程式と力学系の理論入門-非線形現象の解析にむけて』. 遊星社, 10 2004.

- [160] 宣好登坂, 和栄大西, 昌宏山本. 『逆問題の数理解法 偏微分方程式の逆解析』. 東京大学出版会, 12 1999.
- [161] 潤次郎野口. 『多変数解析関数論 学部生へおくる岡の連接定理』. 朝倉書店, 4 2013.
- [162] Charles A. Weibel. *An Introduction to Homological Algebra*. Cambridge University Press, 10 1995.
- [163] Michael OConnor. An introduction to smooth infinitesimal analysis. <http://www.math.cornell.edu/~oconnor/sia.pdf>, pp. 1–13, 5 2008.
- [164] 芳子緒方, 登高小澤. 『東大数理ビデオアーカイブス』, 12 2009. 小澤先生のアロハ装の真相についての当人の発言動画:3:40 くらいからご自身で理由を語っている. 東大数理ビデオアーカイブス (<http://www.ms.u-tokyo.ac.jp/video/>) 自体面白いので, 興味がある向きは見てみよう.
- [165] 洋子小川. 『博士の愛した数式』. 新潮社, 11 2005.
- [166] 泉小嶋, 和弥岡村. 『無限量子系の物理と数里』. サイエンス社, 4 2013.
- [167] 誠岡崎. 『べんりな変分原理』. 共立出版, 5 1993.
- [168] 博司大栗. 『大栗先生の超弦理論入門』. 講談社, 8 2013.
- [169] 泰生大野. 『書評 佐藤幹夫の数学』. pp. 1–4, 2009.
- [170] 浩一太田. 『電磁気学 II』. 丸善出版, 10 2000.
- [171] 浩一太田. 『電磁気学 I』. 丸善出版, 10 2000.
- [172] 浩一太田. 『マクスウェル理論の基礎 相対論と電磁気学』. 東京大学出版会, 7 2002.
- [173] Alberto Parmeggiani. *Spectral Theory of Non-Commutative Harmonic Oscillators: An Introduction*. Springer Berlin Heidelberg, 4 2010.
- [174] Chris Peters. An introduction to complex algebraic geometry with emphasis on the theory of surfaces. pp. 1–129, 2004.
- [175] phasetr. *Math textbook*. 2014.
- [176] Bott Raoul and Tu W. Loring. *Differential Forms in Algebraic Topology*. Springer, 5 1995.
- [177] M. Reed and B. Simon. *Functional Analysis*. Academic Press, 4 1981.
- [178] Alvin E. Roth and Marilda A. Oliveira Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, 6 1992.
- [179] るびきち. 『Emacs テクニックバイブル 作業効率をカイゼンする 200 の技』. 技術評論社, 8 2010.
- [180] るびきち. 『Emacs Lisp テクニックバイブル』. 技術評論社, 11 2011.
- [181] D. Ruelle. *Statistical Mechanics: Rigorous Results*. World Scientific Pub Co Inc, 4 1999.

- [182] D. Ruelle. *Thermodynamic Formalism: The Mathematical Structure of Equilibrium Statistical Mechanics*. Cambridge University Press, 2nd edition, 11 2004.
- [183] S2S. 『0 は自然数 faq.』 . p. 1, 2012.
- [184] 正彦齋藤. 『線型代数入門』 . 東京大学出版会, 3 1966.
- [185] Omri Sarig. *Lecture Notes on Ergodic Theory*. 12 2009.
- [186] Mikio Sato. Theory of hyperfunctions i. *Journal of the Faculty of Science, University of Tokyo. Sect. 1, Mathematics, astronomy, physics, chemistry,*, pp. 139–193, 8 1959.
- [187] Mikio Sato. Theory of hyperfunctions ii. *Journal of the Faculty of Science, University of Tokyo. Sect. 1, Mathematics, astronomy, physics, chemistry,*, pp. 387–437, 8 1960.
- [188] 文隆佐藤, 英雄小玉. 『一般相対性理論』 . 岩波書店, 6 2000.
- [189] 健太郎佐藤. 『炭素文明論 「元素の王者」が歴史を動かす』 . 新潮社, 7 2013.
- [190] Pierre Schapira. General topology. pp. 1–78, 2010.
- [191] Y. Sekine. Magnetism and infrared divergence in a hubbard-phonon interacting system. *arxiv:10082056*, pp. 1–9, 8 2010.
- [192] Y. Sekine. Phonon bose-einstein condensation in a hubbard-phonon interacting system with infrared divergence. *arxiv:13085589*, pp. 1–14, 8 2013.
- [193] 士郎瀬山. 『はじめての現代数学』 . 早川書房, 3 2009.
- [194] 士郎瀬山. 『数学と算数の遠近法 方眼紙を見れば線形代数がわかる』 . 早川書房, 11 2010.
- [195] 浩二志賀. 『複素数 30 講』 . 朝倉書店, 4 1989.
- [196] 浩二志賀. 『無限からの光芒 ポーランド学派の数学者たち』 . 日本評論社, 4 1988.
- [197] 明清水. 『量子論の基礎 その本質のやさしい理解のために』 . サイエンス社, 4 2004.
- [198] 明清水. 『熱力学の基礎』 . 東京大学出版会, 3 2007.
- [199] B. Simon. *The Statistical Mechanics of Lattice Gases*. Princeton Univ Press, 9 1993.
- [200] J. Solovej. Review of stability of matter in quantum mechanics, by elliot h. lieb and robert seiringer. *arXiv*, pp. 1–7, 11 2011.
- [201] Jan Philip Solovej. Book review: Stability of matter in quantum mechanics, by elliot h. lieb and robert seiringer. *arXiv*, pp. 1–7, 2011.
- [202] 相転移 P. 『よくわからない数学 色々な反例で遊ぼう』 , 10 2013.

- [203] Jordan Stoyanov. *Counterexamples in Probability: Third Edition*. Dover Publications, 9 2013.
- [204] R. Streater and A. Wightman. *PCT, Spin and Statistics, and All That*. Princeton Univ. Pr., 12 2000.
- [205] Sho Sugiura and Akira Shimizu. Canonical thermal pure quantum state. *arXiv:1302.3138*, pp. 1–5, 2 2013.
- [206] 光夫杉浦. 『解析入門 I』. 東京大学出版会, 3 1980.
- [207] 光夫杉浦. 『解析入門 II』. 東京大学出版会, 4 1985.
- [208] 利一砂田. 『ダイヤモンドはなぜ美しい?』. 丸善出版, 7 2012.
- [209] 重信砂川. 『理論電磁気学』. 紀伊國屋書店, 9 1999.
- [210] 編集部数学のたのしみ. 『数学まなびはじめ 第 1 集』. 日本評論社, 1 2006.
- [211] 編集部数学のたのしみ. 『数学まなびはじめ 第 2 集』. 日本評論社, 1 2006.
- [212] 編集部数学セミナー. 『数学の 50 年』. 2 2013.
- [213] 編集部数学セミナー. 『数学セミナー 2013 年 2 月号 小林昭七』. 日本評論社, 1 2013.
- [214] 編集部数学セミナー. 『数学セミナー 2014 年 01 月号 グラフ理論の新展開』. 日本評論社, 12 2013.
- [215] 編集部数学セミナー. 『数学セミナー 2014 年 02 月号 多様体が分かりたい』. 日本評論社, 1 2014.
- [216] George G. Szpiro. 『ケプラー予想』. 新潮社, 4 2005.
- [217] 康高橋. 『物性研究者のための場の量子論 1』. 培風館, 10 1974.
- [218] 康高橋. 『物性研究者のための場の量子論 2』. 培風館, 4 1976.
- [219] 陽一郎高橋. 『力学と微分方程式』. 岩波書店, 2 2004.
- [220] M. Takesaki. Non-commutative integration. *arxiv:1208.5197*, Vol. <http://arxiv.org/abs/1208.5197>, pp. 1–14, 2012.
- [221] 正道竹崎. 『作用素環への入口』. 日本数学会, 1 2003.
- [222] 宏志田中. 『物理数学 i』. 2013.
- [223] H. Tasaki. From nagaoka’s ferromagnetism to flat-band ferromagnetism and beyond -an introduction to ferromagnetism in the hubbard model. *Progr. Theor. Phys.*, pp. 489–548, 1998.
- [224] Hal Tasaki. The approach to thermal equilibrium and thermodynamic normality — an observation based on the works by goldstein, lebowitz, mastrodonato, tumulka, and zanghi in 2009, and by von neumann in 1929. *arxiv:1003.5424v4*, 2010.
- [225] 晴明田崎. 『数学：物理を学び楽しむために』. 2013.

- [226] 晴明田崎. 『熱力学—現代的な視点から』. 培風館, 4 2000.
- [227] 晴明田崎. 『統計力学 II』. 培風館, 12 2008.
- [228] 晴明田崎. 『統計力学 I』. 培風館, 12 2008.
- [229] 愛護協会多様体. 『多様体愛護協会』.
- [230] 俊坪井. 『ベクトル解析と幾何学』. 朝倉書店, 5 2002.
- [231] 俊坪井. 『幾何学 I 多様体入門』. 東京大学出版会, 4 2005.
- [232] 俊坪井. 『幾何学 III 微分形式』. 東京大学出版会, 5 2008.
- [233] 善文土基. 『 x の x 乗の話』. 日本評論社, 7 2002.
- [234] 元辻. 『複素多様体論講義』. 10 2012.
- [235] Tomoaki Uda. 『0 は自然数 にしたい理由.』 . p. 1, 3 2012.
- [236] 東介ト部. 一次元代数的特異点とディンキン図形. 遊星社, 10 2007.
- [237] Masato Wakayama. 『非可換調和振動子のスペクトルゼータ関数』. 数理解析研究所講究録, Vol. 1479, pp. 26–39, 2006.
- [238] D. Williams, 次郎赤堀, 啓介原, 俊雄山田. 『マルチンゲールによる確率論』. 培風館, 2 2004.
- [239] G. Woeginger. When cauchy and holder met minkowski: A tour through well-known inequalities. Vol. 82, No. 3, pp. 1–6, 6 2009.
- [240] 昌宏山本. 『逆問題入門』. 1 2002.
- [241] 哲朗山本. 『行列解析の基礎』. サイエンス社, 12 2010.
- [242] 哲朗山本. 『行列解析ノート』. サイエンス社, 1 2013.
- [243] 義隆山本, 孔一中村. 『解析力学 II』. 朝倉書店, 9 1998.
- [244] 義隆山本, 孔一中村. 『解析力学 I』. 朝倉書店, 9 1998.
- [245] 恭彦山内, 光夫杉浦. 『連続群論入門』. 培風館, 8 2010.
- [246] 邦夫保江. 『ヒルベルト空間論』. 日本評論社, 5 2000.
- [247] 邦夫保江. 『量子の道草 方程式のある風景』. 日本評論社, 1 2009.
- [248] 武吉田. 『素数夜曲 女王陛下の LISP』. 東海大学出版会, 6 2012.
- [249] 輝義吉田. 『保型表現と galois 表現 –初学者のために–』 . pp. 1–21, 2010.
- [250] 輝義吉田. 『ガロア理論の基本定理』 . pp. 1–5, 2011.
- [251] 輝義吉田. 『現代数学のめざすもの–ガロア理論 整数論とその周辺から–』 . pp. 1–12, 12 2011.
- [252] K. Yosida. *Functional Analysis*. Springer Berlin Heidelberg, 8 1996.
- [253] L. C. Young. *Lectures on the calculus of variations and optimal control theory*. Saunders, Philadelphia, 5 1969.

- [254] Yogui. 『初めての Ruby』. オライリージャパン, 6 2008.
- [255] レヴェンソントマス, のぶ子寺西. 『ニュートンと贋金づくり 天才科学者が追った世紀の大犯罪』. 白揚社, 11 2012.
- [256] 亮介長岡. 『東大の数学入試問題を楽しむ 数学のクラシック鑑賞』. 日本評論社, 2 2013.