献给罗杰·彭罗斯

目录

前言 11
致谢 19
第一章 欧式与非欧几何 23
1.1 欧式与双曲几何 23
1.2 球面几何 27
1.3 球面三角形的角盈 29
1.4 曲面的内蕴和外在几何 31
1.5 利用直性构造测地线 33
1.6 空间的本质 36
第二章 高斯曲率 39
2.1 引言 39
2.2 圆的周长和面积 41
2.3 局部高斯-博内定理 45
第三章 前言和第一幕练习 47
前言：牛顿终极等式 (≍) 47
欧式与非欧几何 48
高斯曲率 50
第四章 映射曲面：度量 53
4.1 引言 53
4.2 球面的射影地图 54
4.3 一般曲面的度量 56
4.4 度量曲率公式 59
4.5 共形映射 61
4.6 一些复分析 63
4.7 球面的共形球极地图 67
4.8 球极射影的公式 71
4.9 球极射影的保圆性质 73
第五章 伪球与双曲面 75
5.1 贝尔特拉米的洞见 75
5.2 轨迹与伪球 76
5.3 伪球的共形映射 78
5.4 贝尔特拉米-庞加莱半平面 80
5.5 用光学找到测地线 82
5.6 平行的角度 85
5.7 贝尔特拉米-庞加莱圆盘 87
第六章 等距变换与复数 90
6.1 介绍 90
6.2 默比乌斯变换 92
6.3 主要结果 97
6.4 爱因斯坦的时空几何 100
6.5 3 维双曲几何 105

第七章 第二幕练习.....110
映射曲面：度量.....110
伪球面与双曲平面.....114
等距同构与复数.....117
第八章 平面曲线的曲率.....121
8.1 引言.....121
8.2 曲率圆.....122
8.3 牛顿曲率公式.....124
8.4 曲率作为转弯的速率.....125
8.5 示例：牛顿的曳物线.....129
第九章 3 维空间中的曲线.....132
第十章 曲面的主曲率.....135
10.1 欧拉曲率公式.....135
10.2 欧拉曲率公式的证明.....137
10.3 旋转曲面.....138
第十一章 平面曲线的曲率.....142
11.1 测地曲率和法曲率.....142
11.2 Meusnier 定理.....144
11.3 测地线是“直的”.....145
11.4 测地曲率的内蕴测量.....146
11.5 测量测地曲率的一种简单外在方法.....147
11.6 测地线粘胶带构造的新解释.....148
11.7 旋转曲面上的测地线.....148
11.7.1 球面上的克莱罗定理.....148
11.7.2 开普勒第二定律.....150
11.7.3 牛顿对开普勒第二定律的几何演示.....152
11.7.4 克莱罗定理的动力学证明.....154
11.7.5 应用：双曲面中的测地线（重温）.....157
第十二章 曲面的外在曲率.....159
12.1 引言.....159
12.2 球面映射.....159
12.3 曲面的外在曲率.....160
12.4 哪些形状是可能的？.....164
第十三章 高斯绝妙定理.....168
13.1 引言.....168
13.2 高斯美丽定理（1816）.....168
13.3 高斯绝妙定理（1827）.....170
第十四章 尖刺的曲率.....174
14.1 引言.....174
14.2 圆锥尖刺的曲率.....174
14.3 多面体尖刺的内蕴和外在曲率.....177
14.4 多面体绝妙定理.....178
第十五章 形状算子.....181
15.1 方向导数.....181

15.2 形状算子 $\mathbf{S}$ 183
15.3 $\mathbf{S}$ 的几何效应 185
15.4 绕行：奇异值分解和转置的几何 186
15.5 $\mathbf{S}$ 的一般矩阵 190
15.6 $\mathbf{S}$ 的几何解释与 $[\mathbf{S}]$ 的简化 191
15.7 $[\mathbf{S}]$ 完全由三个曲率决定 193
15.8 渐近方向 195
15.9 经典术语与符号：三种基本形式 197
第十六章 全局高斯-博内定理简介 198
16.1 一些拓扑和结果的陈述 198
16.2 球面和环面的总曲率 201
16.2.1 球面的总曲率 201
16.2.2 环面的总曲率 203
16.3 通过厚煎饼看见 $\mathcal{KS}\mathbf{g}$ 204
16.4 通过百吉饼和桥看见 $\mathcal{KS}\mathbf{g}$ 205
16.5 球面映射的拓扑度 206
16.6 历史笔记 208
第十七章 全局高斯-博内定理的第一个（启发式）证明 209
17.1 平面环路的总曲率：霍普夫环量定理 209
17.2 变形圆的总曲率 212
17.3 霍普夫环量定理的启发式证明 213
17.4 变形球的总曲率 214
17.5 全局高斯-博内定理的启发式证明 215
第十八章 全局高斯-博内定理的第二个（角盈）证明 218
18.1 欧拉示性数 218
18.2 欧拉（经验）多面体公式 218
18.3 欧拉多面体公式的柯西证明 221
18.3.1 铺平多面体 221
18.3.2 多边形网的欧拉示性数 222
18.4 欧拉多面体公式的勒让德证明 224
18.5 向曲面添加环柄增加其亏格 226
18.6 全局高斯-博内定理的角盈证明 229
第十九章 全局高斯-博内定理的第三个（向量场）证明 231
19.1 引言 231
19.2 平面向量场 231
19.3 奇点的指数 232
19.4 原型奇点：复数的幂 234
19.5 曲面上的向量场 238
19.5.1 蜂蜜流向量场 238
19.5.2 蜂蜜流与地形图的关系 240
19.5.3 定义曲面上的指数 241
19.6 庞加莱-霍普夫定理 243
19.6.1 示例：拓扑球面 243
19.6.2 庞加莱-霍普夫定理的证明 245

19.6.3 应用：欧拉-惠里尔公式的证明.....246
19.6.4 庞加莱微分方程与霍普夫线场.....247
19.7 全局高斯-博内定理的向量场证明.....253
19.8 前路.....257
第二十章 第三幕练习.....258
平面曲线的曲率.....258
3 维空间中的曲线.....258
曲面的主曲率.....259
高斯绝妙定理.....260
形状算子.....261
全局高斯-博内定理简介.....264
GGB 的第一个（启发式）证明.....264
GGB 的第二个（角盈）证明.....265
GGB 的第三个（向量场）证明.....266
第二十一章 历史谜题.....268
第二十二章 外在构造.....270
22.1 边走边投影到曲面！.....270
22.2 测地线与平行移动.....272
22.3 土豆削皮器移动.....273
第二十三章 内蕴构造.....277
23.1 通过测地线的平行移动.....277
23.2 内蕴（又名“协变”）导数.....278
第二十四章 和乐.....282
24.1 示例：球面.....282
24.2 一般测地三角形的和乐.....283
24.3 和乐是可加的.....285
24.4 示例：双曲平面.....285
第二十五章 绝妙定理直观的几何证明.....289
25.1 引言.....289
25.2 定义的一些符号和提示.....290
25.3 迄今为止的故事.....290
25.4 球面投影保持平行移动.....292
25.5 美丽定理与绝妙定理的证明.....293
第二十六章 全局高斯-博内定理的第四个（和乐）证明.....295
26.1 引言.....295
26.2 沿开曲线的和乐？.....295
26.3 全局高斯-博内定理的霍普夫内蕴证明.....296
第二十七章 度量曲率公式的几何证明.....300
27.1 引言.....300
27.2 围绕环路之向量场的环量.....301
27.3 演练：平面中的和乐.....303
27.4 和乐作为地图中度量诱导向量场的环量.....305
27.5 度量曲率公式的几何证明.....307
第二十八章 曲率作为相邻测地线之间的力.....309

28.1 雅可比方程简介.....309
28.1.1 零曲率：平面.....310
28.1.2 正曲率：球面.....311
28.1.3 负曲率：伪球面.....313
28.2 雅克比方程的两个证明.....315
28.2.1 测地极坐标.....315
28.2.2 相对加速度=速度的和乐.....317
28.3 小测地圆的周长和面积.....319
第二十九章 黎曼曲率.....321
29.1 导言和总结.....321
29.2 n-流形中的角盈.....322
29.3 平行移动：三种构造.....324
29.3.1 常角度圆锥上的最近向量.....324
29.3.2 平行移动平面内的常角度.....325
29.3.3 希尔德梯子.....326
29.4 内蕴（又名“协变”）导数 $\nabla \mathbf{v}$.....327
29.5 黎曼曲率张量.....328
29.5.1 围绕小“平行四边形”的平行移动.....328
29.5.2 用向量交换子闭合“平行四边形”.....329
29.5.3 一般黎曼曲率公式.....330
29.5.4 黎曼曲率是张量.....333
29.5.5 黎曼张量的分量.....334
29.5.6 对于给定的 $\mathbf{wo}$，向量和乐仅取决于环路的平面及其面积.....336
29.5.7 黎曼张量的对称性.....337
29.5.8 截面曲率.....339
29.5.9 黎曼张量起源的历史笔记.....341
29.6 n-流形中的雅克比方程.....342
29.6.1 截面雅克比方程的几何证明.....342
29.6.2 截面雅克比方程的几何含义.....344
29.6.3 雅克比方程和截面雅克比方程的计算证明.....345
29.7 里奇张量.....346
29.7.1 被测地线束包围面积的加速度.....346
29.7.2 里奇张量的定义和几何意义.....349
29.8 尾声.....351
第三十章 爱因斯坦的弯曲时空.....352
30.1 引言：“我一生中最幸福的想法”.....352
30.2 引力潮汐力.....354
30.3 几何形式的牛顿万有引力定律.....357
30.4 时空度量.....359
30.5 时空图.....361
30.6 几何形式的爱因斯坦真空场方程.....362
30.7 史瓦西解和理论的第一次检验.....365
30.8 引力波.....368
30.9 几何形式的爱因斯坦场方程（含物质）.....371

30.10 引力坍塌到黑洞.....375
30.11 宇宙常数：“我一生中最大的错误。”.....378
30.12 结束.....380
第三十一章 第四幕练习.....381
外在构造.....381
内蕴构造.....382
和乐.....382
曲率作为相邻测地线之间的力.....383
黎曼曲率.....384
爱因斯坦弯曲时空.....387
第三十二章 1-形式.....391
32.1 引言.....391
32.2 1-形式的定义.....392
32.3 1-形式的示例.....394
32.3.1 引力功.....394
32.3.2 可视化引力功 1-形式.....395
32.3.3 地形图与梯度 1-形式.....396
32.3.4 行向量.....399
32.3.5 狄拉克的左矢.....399
32.4 基 1-形式.....400
32.5 1-形式的分量.....401
32.6 作为 1-形式的梯度：$\mathbf{df}$.....402
32.6.1 梯度作为向量的回顾：∇f.....402
32.6.2 作为 1-形式的梯度：$\mathbf{df}$.....403
32.6.3 笛卡尔 1-形式基：$\{\mathbf{dxj}\}$.....404
32.6.4 $\mathbf{df} = (\partial \mathbf{xf})\,\mathbf{dx} + (\partial \mathbf{yf})\,\mathbf{dy}$ 的 1-形式解释.....405
32.7 以几何方式加入 1-形式.....406
第三十三章 张量.....409
33.1 张量的定义：价.....409
33.2 示例：线性代数.....410
33.3 张量：从旧得新.....410
33.3.1 加法.....410
33.3.2 乘法：张量积.....411
33.4 分量.....411
33.5 度量张量与经典线元的关系.....413
33.6 示例：线性代数（再次）.....414
33.7 缩并.....415
33.8 使用度量张量更改价.....417
33.9 对称与反对称.....419
第三十四章 2-形式.....421
34.1 2-形式和 p 形式的定义.....421
34.2 示例：面积 2-形式.....422
34.3 两个 1-形式的楔积.....423
34.4 极坐标中的面积 2-形式.....426

34.5 基 2-形式和投影 427
34.6 将 2-形式与$\mathbb{R}\mathbf{3}$中的向量关联：通量 429
34.7 $\mathbb{R}\mathbf{3}$中向量与楔积的关系 431
34.8 法拉第和麦克斯韦电磁 2-形式....434
第三十五章 3-形式 440
35.1 3-形式需要 3 个维度 440
35.2 2-形式和 1-形式的楔积 440
35.3 体积 3-形式 442
35.4 球极坐标中的体积 3-形式....442
35.5 三个 1-形式和 p 个 1-形式的楔积 444
35.6 基 3-形式 445
35.7 $\mathbf{\Psi} \wedge \mathbf{\Psi} \neq \mathbf{0}$ 可能么？ 446
第三十六章 微分....447
36.1 1-形式的外导数 447
36.2 2-形式和 p-形式的外导数 449
36.3 形式的莱布尼茨规则....450
36.4 闭的和恰当的形式....451
36.4.1 基本结果：$\mathbf{d2} = \mathbf{0}$ 451
36.4.2 闭的和恰当的形式....452
36.4.3 复分析：柯西-黎曼方程....453
36.5 通过形式的向量分析....454
36.6 麦克斯韦方程....457
第三十七章 积分....461
37.1 1-形式的线积分 461
37.1.1 环量与功....461
37.1.2 路径无关⟺环路积分为零 462
37.1.3 恰当形式的积分：$\boldsymbol{\varphi} = \mathbf{df}$....463
37.2 作为积分的外导数....464
37.2.1 $\mathbf{d}$(1-形式)....464
37.2.2 $\mathbf{d}$(2-形式)....468
37.3 外微分的基本定理（广义斯托克斯定理） 470
37.3.1 外微分的基本定理....470
37.3.2 历史旁白....470
37.3.3 示例：面积....471
37.4 边界的边界是零！ 472
37.5 向量分析的经典积分定理....473
37.5.1 Φ=0-形式 473
37.5.2 Φ=1-形式 473
37.5.3 Φ=2-形式 475
37.6 外微分基本定理的证明....476
37.7 柯西定理....478
37.8 1-形式的庞加莱引理 479
37.9 德拉姆上同调入门....480
37.9.1 引言....480

37.9.2 一种特殊的二维涡旋向量场 481
37.9.3 涡旋 1-形式是闭的 482
37.9.4 涡旋 1-形式的几何意义 483
37.9.5 闭 1-形式环量的拓扑稳定性 484
37.9.6 第一德拉姆上同调群 486
37.9.7 $\mathbb{R}3$中的平方反比点源 488
37.9.8 第二德拉姆上同调群 490
37.9.9 环面的第一德拉姆上同调群 492
第三十八章 途经形式的微分几何 495
38.1 引言：嘉当的活动标架法 495
38.2 联络 1-形式 497
38.2.1 符号约定和两个定义 497
38.2.2 联络 1-形式 498
38.2.3 警告：前方有符号欺凌仪式！！ 500
38.3 姿态矩阵 501
38.3.1 途径姿态矩阵的联络形式 501
38.3.2 示例：圆柱形标架场 503
38.4 嘉当的两个结构方程 505
38.4.1 按照 **ej** 的对偶 **dxj** 得 **mi** 的对偶 **θi** 505
38.4.2 嘉当第一结构方程 506
38.4.3 嘉当第二结构方程 507
38.4.4 示例：球面标架场 508
38.5 曲面的六个基本形式方程 514
38.5.1 嘉当活动标架适用于曲面：形状算子与外在曲率 514
38.5.2 示例：球面 515
38.5.3 基分解的唯一性 516
38.5.4 曲面的六个基本形式方程 517
38.6 对称方程和 Peterson–Mainardi–Codazzi 方程的几何意义 519
38.7 高斯方程的几何形式 520
38.8 度量曲率公式和绝妙定理的证明 521
38.8.1 引理：**ω12** 的唯一性 521
38.8.2 度量曲率公式的证明 521
38.9 一种新的曲率公式 523
38.10 希尔伯特引理 524
38.11 利布曼刚性球定理 525
38.12 n-流形的曲率 2-形式 526
38.12.1 导言和总结 526
38.12.2 广义外导数 528
38.12.3 从曲率 2-形式中提取黎曼张量 530
38.12.4 重温比安奇恒等式 531
38.13 史瓦西黑洞的曲率 532
第三十九章 第五幕练习 538
1-形式 538
张量 539

2-形式 540
3-形式 540
微分 541
积分 545
途径形式的微分几何 546
延伸阅读 551
整体推荐 551
一般几何 552
拓扑 553
霍普夫的线场和物理学中的庞家莱-霍普夫定理 553
微分几何 555
黎曼曲率 556
爱因斯坦的弯曲时空（广义相对论） 557
形式（数学） 558
形式（物理学） 558
参考文献 560

前言

浮士德式的协议

代数学是魔鬼给数学家的协议。魔鬼说："我要给你这个强大的机器，它可以回答你想问的任何问题。你所要做的就是把你的灵魂给我：放弃几何学，你就会拥有这个奇妙的机器。"对我们灵魂的威胁就在于，当你进入代数计算时，你基本上就停止了思考：你停止了用几何思考，你停止了思考"一切的意义是什么"。

迈克尔·阿提亚爵士[1]

"微分几何"包含"几何"一词。

同义词？好吧，第一次打开指定教科书的本科生可能不会同意！相反，我们不幸的学生面对的是大量的公式，他们的证明包括冗长和不透明的计算。雪上加霜的是，这些计算通常很丑陋，涉及"指标的放荡"[2]（debauch of indices）——这是埃利·嘉当（我们戏剧的主要英雄之一）在 1928 年创造的短语。如果这学生足够诚实勇敢，教授可能被迫面对一个令人尴尬的直率问题："几何哪里去了？！"

实话说，现在大多数现代教科书实际上包含了许多图片，通常是计算机生成的曲线和曲面。但是，除了极少数例外，这些图片都是特定的、具体的例子，它们只是解释了定理，这些定理的证明完全由符号操纵。这些图片说明不了什么！

这本书有两个不同但同样雄心勃勃的目标，第一个目标是前四幕的主题——将"几何"重新纳入导论"微分几何"介绍。本书所包含的 235 张手绘图与仅由计算机生成的示例在本质上不同。它们是多年反复深刻思考的萃取成果——它们是对惊人几何事实的直观解释与视觉呈现。

我在《可视化复分析》[3]（Visual Complex Analysis，简称 VCA）前言中写的那些话今天仍然适用："就我所知，本书中包含的几何观点和提法，有很大一部分是全新的。我没有在正文中强调这点，是因为并不必要：学生们不需要知道这些，专家们不说也会知道。然而，如果一个想法显然不同寻常，我又知道它由别人发表过，那我肯定会毫不犹豫的给予点赞。此外，容我声明，课后练习看似原创，但实际并非如此。

就个人而言（但有一个严肃的数学观点），不辞辛苦的根源可以追溯到几十年前，我的青年时代。这个故事与两本书有关。

第一本书点燃了我对微分几何和爱因斯坦广义相对论的深刻迷恋。也许这次经历之所以

---

[1] 20 世纪的数学 (Shenitzer & Stillwell, 2002, p. 6)

[2] 完整的引文开始揭示嘉当的英雄形象："利玛窦和列维-奇维塔绝对微积分的效用必须通过避免过度形式化的计算来缓和，在这种计算中，指数的放荡掩盖了通常非常简单的几何现实。 我试图揭示的正是这个现实。"（摘自 Cartan 1928 的序言。）

[3] 鉴于我经常有机会回顾我的第一本书，Visual Complex Analysis (Needham 1997)，我将采用简洁的自负，将其简称为 VCA。

如此刻骨铭心，是因为这是我的初恋：那年我 19 岁。有一天，在我在牛津默顿学院学习物理的第一年结束时，我在布莱克威尔的商店里偶然发现了一本巨大的黑书。当时我并不知道，这本 1217 页的大部头已被相对论物理学家们委婉地称为“圣经”。可以毫不夸张地说这部非凡的作品改变了我的整个人生历程。 如果我没有读《引力论》（米斯纳、索恩和惠勒 1973 年），我就永远没有机会[4]在罗杰·彭罗斯手下学习（并与之成为终身朋友），而他从根本上改变了我对数学和物理的理解。

1982 年夏天，我对韦斯特福尔（1980 年）优秀的《牛顿传》中的数学闪光很感兴趣，因此我对牛顿（1687 年）的杰作，《自然哲学的数学原理》，通常简称为《原理》，进行了深入的研究。这是第二本从根本上改变了我生活的书。阿诺德[5]和钱德拉塞卡（V. I. Arnol'd and S. Chandrasekhar，1995 年）试图揭示牛顿在《原理》中阐述结果的非凡性质，而本书却源于对牛顿方法的迷恋。

正如我们在其他地方争辩的那样[6]，牛顿学者们费尽心思地驳斥了一个有害的神话[7]，即 1687 年《原理》的结果首先是牛顿用他最初的 1665 年版本的微积分导出的，直到后来才重新塑造成我们在完成的作品中发现的几何形式。

相反，现在人们明白，到了 1670 年代中期，特别是在研究阿波罗尼乌斯、帕普斯和惠更斯之后，成熟的牛顿对他年轻时最初发现微积分的形式不再抱有幻想——这与我们今天在大学里学习的莱布尼茨形式再次不同——而是接受了纯粹的几何方法。

因此，到了 1680 年代，牛顿对代数幂级数的迷恋让位于一种新的微积分形式——他称之为“流数的综合方法”[8]（synthetic method of fluxions）——在这种方法中，古人的几何学被改造并应用于无限缩小并最终消失的几何图形，从而重获生机。这就是我们在 1687 年他伟大的《原理》中发现的那种强大的非演算的微积分形式。

就像我在 VCA 中所做的那样，我现在希望在这本书中充分利用牛顿的方法。因此，让我立即详细说明它，而且比我在 VCA 中做的要详细得多，希望这第二本书能激励更多的数学家和物理学家采用牛顿的直觉（但严格[9]）方法，而不是我的第一本书。

如果两个量 A 和 B 依赖于一个小量 ϵ，而它们的比率在 ϵ 趋于零时趋于单位 1，那么我们将避免更繁琐的极限语言，遵循牛顿在《原理》的先例，简单地说，“A 最终等于 B。”此外，正如我们在早期作品（李约瑟 1993 年，2014 年）中所做的那样，我们将使用符号$\asymp$来表示最终相等[10]的概念。简言之，

---

[4] 多年后，我有幸与惠勒多次会面，并与他通信，所以我终于能够直接感谢他的万有引力对我生活的影响。

[5] 见 Arnol'd and Vasil'ev (1991); Arnol'd (1990).

[6] 参见 Needham (1993)、VCA 前言和 Needham (2014)。

[7] 可悲的是，这个神话起源于牛顿本人，当时他正与莱布尼茨就微积分的发现展开激烈的争斗。 参见 Arnol'd (1990)、Bloye 和 Huggett (2011)、de Gandt (1995)、Guicciardini (1999)、Newton (1687, p. 123) 和 Westfall (1980)。

[8] 参见 Guicciardini (2009, Ch. 9)。

[9] 易读的精美印刷品！

[10] 这一符号随后被诺贝尔奖得主物理学家 Subrahmanyan Chandrasekhar 采用（参见 Chandrasekhar 1995，第 44 页）。

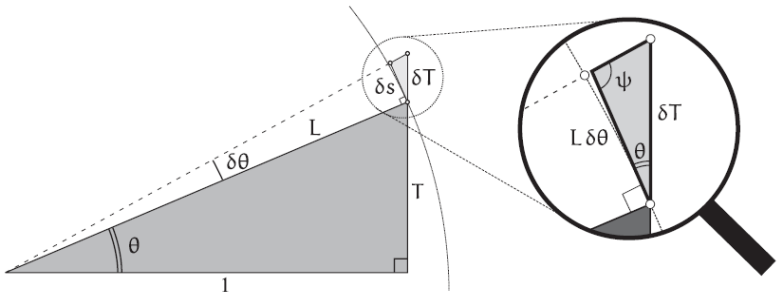
$$\text{“A 最终等于 B”} \iff A \asymp B \iff \lim_{\epsilon\to 0}\frac{A}{B} = 1.$$

根据极限定理【练习】，最终等式是等价关系，它还继承了一般等式的附加性质，例如 $X \asymp Y \,\&\, P \asymp Q \Rightarrow X \cdot P \asymp Y \cdot Q$, 和 $A \asymp B \cdot C \Leftrightarrow (A/B) \asymp C$.

在我们开始认真应用这一想法之前，我们还注意到，最终相等的适用范围可以自然地扩展到数字以外，例如，使人们能够说，两个三角形“最终相似”，这意味着它们的角度最终是相等的。

掌握了牛顿的方法后，我立即尝试用它来简化我的入门微积分教学，直到后来才意识到我如何将它应用于复分析（在 VCA 中），现在又应用于微分几何。虽然我可能会选择任何数量的简单、说明性的例子（更多信息见李约瑟 1993），但我将重复使用我在 VCA 前言中给出的具体例子，原因很简单：这一次，我将使用“≍”——符号来严格地表示论点，而在 VCA 中，我没有。事实上，这个例子可以被看作是将 VCA 的大多数“解释”转化为“证明”[11]的良方，仅仅是洒在必要的 ≍ 上。

让我们证明，如果 $T = \tan\theta$，那么 $\frac{dT}{d\theta} = 1 + T^2$。见下图。如果我们增加一个小的（最终趋于零的）量 $\delta\theta$，那么 T 将增加小三角形的垂直斜边 δT 的长度，其中，该三角形的其他两侧被构造成位于 $(\theta + \delta\theta)$ 和 $(\theta + \frac{\pi}{2})$ 方向上，如图所示。为了得到结果，我们首先观察到，在 $\delta\theta$ 趋于零的极限中，斜边 δT 的小三角形最终与斜边 L 的大三角形相似，因为 $\psi \asymp \frac{\pi}{2}$。接下来，正如我们在放大镜中看到的，小三角形中与 θ 相邻的边 δs 最终等于半径为 L 的圆的图示弧，因此 $\delta s \asymp L\,\delta\theta$。因此，



$$\frac{dT}{L\,d\theta} \asymp \frac{\delta T}{L\,\delta\theta} \asymp \frac{\delta T}{\delta s} \asymp \frac{L}{1} \quad\Longrightarrow\quad \frac{dT}{d\theta} = L^2 = 1 + T^2$$

就我所知，牛顿从来没有写下这个具体的例子，而是将他的几何推理风格[12]的启发性直

[11] 在撰写 VCA 时，我已经在使用 ≍ 符号（私下和印刷中），事后看来，我没有在该作品中使用它是一个错误； 这导致一些人认为 VCA 中提出的论证不如实际严格（并且仍然如此）。

[12] 牛顿方法的最佳大使将是你自己。 因此，我们建议你立即通过做第 24 页的练习 1、2、3 和 4 来尝试自己的牛顿推理。

接与三个多世纪后我们今天教给学生的无启发性的计算进行比较！正如牛顿本人所说[13]，几何方法是首选的，因为“所涉及的推理的清晰和简洁，以及所需的结论和插图的简单。”事实上，牛顿更进一步，认为只有综合方法才是“值得公开的”。

牛顿本人并没有用任何符号来表示他的“最终相等”概念，相反，他对古人几何方法的热爱溢出到模仿古人的表达方式上，使他每当在证明中援引这个概念时，写出了“最终有相等的比例”这个词。正如牛顿（1687 年，第 124 页）所解释的那样，《原理》“用古人的方式详细地写成文字”。即使牛顿声称两个比率最终是相等的，他也坚持用文字表达每个比率。因此，我本人在不首先将牛顿的每一段文字转录和总结为“现代”形式（事实上，这在 1687 年已经相当普遍）的情况下，完全无法跟随牛顿的推理。事实上，早在 1982 年，这就是我私下引进和使用 ≈ 符号的催化剂。

我认为，牛顿选择不引入“最终相等”的符号，对数学的发展来说是一个悲剧性的重大错误。当莱布尼茨的符号微积分席卷世界时，牛顿更具穿透力的几何方法被搁置一边。在这几个世纪里，只有少数人试图修复这种破坏，恢复牛顿的方法，最近最著名和最杰出的冠军是阿诺德[14]（V.I.Arnol'd，1937－2010 年）。

如果牛顿摆脱了这种古老的阐述方式的束缚，而是使用一些符号（任何符号！）来代替“最终相等”这个词，他在《原理》中的密集、长篇大论的证明可能会被简化为几行 ，他的思维方式今天可能仍然被广泛采用。 VCA 和本书都试图非常具体地展示牛顿几何方法在数学领域的持续现实意义和活力，这些数学领域的发现在他 1727 年去世后的未来一个世纪。

请允许我插入一些关于我使用“严格”和“证明”这两个词的细则。 是的，与我在 VCA 中的阐述相比，我在本书中明确使用牛顿终极等式代表了严格性的量子跃迁，但是会有一些数学家会反对（有理由！）即使这种严格性的增加也是不够的， 并且本书中的任何“证明”都不配这个标题，包括刚刚给出的那个：我实际上并没有证明三角形的边最终等于圆的弧。

我不能给出任何合乎逻辑的辩护，而只能重复我在二十多年前在 VCA 前言中所写的话：“毫无疑问，我的书在许多方面存在我尚不知晓的缺陷，但有一个“罪”是我故意犯下的，我不会为此悔改：许多论点并不严谨，至少就目前而言。如果一个人认为我们的数学理论只是精心设计的心理结构，摇摇晃晃地被高高举起，这就是一种严重的犯罪。然后，严谨就变成了一种令人伤脑筋的平衡行为，以防止整个结构在我们周围倒塌。但假设有人像我一样相信，我们的数学理论正试图捕捉非我们所创造的强大的柏拉图世界的各个方面。那么我会争辩说，如果它能让读者比其他方式更直接、更愉快地看到这个世界，则最初严谨性的缺失是一个很小的代价。”因此，为了先发制人地回应我的批评者，让我从一开始就承认，当我声称一个断言被“证明”时，它可以被解读为“排除合理的怀疑的证明”[15]（proved beyond a reasonable doubt）！

与严谨性无关的是一个可悲的事实，即在重新思考这么多经典数学时，我几乎肯定犯了

---

[13] 参见 Guicciardini (2009, p. 231)

[14] 例如，参见 Arnol’d (1990)。

[15] 读到这些文字后，一位强烈支持普林斯顿大学出版社的编辑委员会成员向我的编辑建议，在“Q.E.D.”的位置，我用字母“P.B.R.D.”结束我的每个证明！

错误：所有这些错误的责任在我，而且只有我一个人。 但请不要因为我的几何工具做工如此糟糕而责备我——我在执行符号计算时同样会犯错误！若有更正请发送 VDGF.correction@gmail.com 不胜感谢。

这本书阅读过程中无需瞻前顾后，只需随着五幕戏剧的正常展开即可完全理解。也就是说，我认为情节很重要，而且这本书的非正统结构和标题很合适，原因如下。首先，我试图以我自己所见的那样戏剧性地呈现这些想法，不仅在它们的历史发展[16]方面，而且（更重要的是）在这些想法本身的层叠串接、相互关联的流动及其对于其余的数学和物理惊人的影响方面。其次，更多是出于本能而非设计，五幕中每一幕的角色确实（或多或少）遵循莎士比亚戏剧的古典结构；特别是，预期的“高潮”确实是第三幕：“曲率”。事实上，在我开始写这本书的几年后，有一天我突然明白我一直在创作的作品是*一部五幕的数学戏剧*。就在那天我“更正”了作品的标题，并相应地把它原来的五个“部分”改为了“幕”：

- 第一幕：空间的本质
- 第二幕：度量
- 第三幕：曲率
- 第四幕：平行移动
- 第五幕：形式

前四幕实现了对微分几何进行独立的几何介绍的承诺。 第四幕是真正的数学强国，它最终使得为前三幕中的许多断言提供几何证明成为可能。

*主题*的几个方面与处理它们的几何方法一样非正统（在第一门课程中）。 在这里，我们将只描述三个最重要的例子。

首先，第三幕高潮中的高潮是*全局高斯-博内定理*——局部几何与全局拓扑之间的显著联系。 虽然包含这个主题是通常做法，但我们对它的处理却与众不同。 的确，我们用数学烟花的奢华展示来庆祝它的中心性和根本重要性：我们用*五章*来讲述它，提供*四个*截然不同的证明，每一个都为结果和微分几何本身的性质提供了新的启示。

其次，从二维表面到 n 维空间（称为“流形”）的过渡（通常在研究生院中）常常让学生感到困惑和恐惧。 第 29 章——本书第二长的一章——试图通过（最初）关注可以*可视化*的 3 维流形的曲率来弥合这一差距；然而，我们将讨论框架化以适用于*任何*数量的维度。我们使用这种方法对著名的*黎曼张量*（测量 n 维流形的曲率）进行了直观的、几何的、技术上完整的介绍。

第三，在致力于全面处理黎曼张量后，我们认为向读者隐瞒它在自然世界舞台上的最大胜利是*不道德的*。 因此，我们在第四幕结束时对爱因斯坦光荣的*广义相对论*进行了长时间的几何介绍，该理论将引力解释为物质和能量对 4 维时空的曲率。这是本书第三长的章节。它不仅处理（完整的几何细节）著名的*引力场方程*（爱因斯坦于 1915 年发现），而且还解释了一些关于它对黑洞、引力波和宇宙学的影响的最新和令人兴奋的发现！

[16] 就像我在 VCA 中所做的那样，我强烈推荐 Stillwell（2010 年）的杰作《数学及其历史》作为本书的配套，因为它对许多我们只能在这里触及的历史发展提供了深刻而深入的分析。

现在让我们转向第五幕，它在性质上与之前的四幕完全不同，因为它寻求完成工作的第二个目标，一个与第一个完全不同的目标，但同样雄心勃勃。

即使是最狂热的几何狂热者也必须承认，阿蒂亚的恶魔机器（在开场白中有描述）是一种必要的邪恶。 但如果我们必须计算，让我们至少优雅地计算！ 幸运的是，从 1900 年开始，埃利 • 嘉当（Élie Cartan）开发了一种强大而优雅的新计算方法，最初是为了研究李群，但后来为微分几何提供了一种新方法。

嘉当的发现被称为“外微积分”，它所研究、区分和整合的对象被称为“微分形式”，这里简称为“形式”。 我们最终将追随嘉当的领导，在第五幕的最后一章（本书最长的一章）中说明他的方法的力量和优雅，重新用符号方法证明前四幕中几何证明的结果。 但是形式将带领我们超越前四幕能够企及的高度：特别是，它们将提供一种非常有效的方法来计算 n 流形的黎曼张量，通过它的曲率 2-形式。

然而，首先，我们将充分发展嘉当的思想，提供完全独立于前四幕的形式的自成一体的介绍。 为了避免混淆，我们重复一遍，第五幕的前六章（共七章）没有提及微分几何！我们这样做是因为形式（Forms）在数学、物理学和其他学科的不同领域中找到了富有成效的应用。我们的目标是让尽可能多的读者容易理解形式,即使他们的主要兴趣不是微分几何。

为此，我们试图以比习惯更直观和几何的方式处理形式。 也就是说，读者不应该抱有任何幻想：第五幕的主要目的是在本科阶段构建“魔鬼机器”——一种非常强大的计算方法。

这些形式的巨大力量让人想起复数：一小滴水落入，一片汪洋倾泻而出——嘉当的形式所解释的远远超过了它的发现者所要求的，这是他发现柏拉图形式的明确迹象！

仅举一个例子，形式（Forms）以一种对本科生（只要他们被允许看到它）来说是一种启示的方式统一并阐明了所有向量微积分。 事实上，格林定理、高斯定理和斯托克斯定理只是关于形式的单个定理的不同表现形式，它比任何这些特殊情况都简单！ 尽管微分形式在数学和物理学中的重要性毋庸置疑，但大多数本科生会在从未见过它们的情况下离开大学，我一直认为这是一个丑闻。 甚至只有少数[17]珍贵的本科教科书（关于向量微积分或微分几何）提到了它们的存在，而这些课程也常被贬谪到研究生院。

这种可悲的事态现在已经进入了第二个世纪，我没有看到即将发生巨变的迹象。 作为回应，第五幕不是要诅咒黑暗，而是要点燃一支蜡烛[18]，努力让读者相信嘉当的形式（及其基础“张量”）既简单又美丽，而且它们（以及名字嘉当 Cartan！）应该成为本科课程的标准部分。 这是第五幕雄心勃勃的目标。在前四幕让读者沉浸在纯几何中之后，我们希望最后一幕的计算方面可以作为一个适当的宣泄结局！

在我们结束之前，让我们简单地列出一些家务细节：

---

[17] 请参阅本书末尾的进一步阅读。

[18] 我们当然不是第一个点燃这样的蜡烛。 事实上，就在我们的工作即将完成之际，Fortney (2018) 出版了一整本书，致力于实现同样的目标。 然而，Fortney 的著作不包括任何对微分几何的讨论，而且，Fortney 的书（461 页）比本书第五幕中包含的 100 页微分形式介绍要长得多。

- 首先，我没有尝试将这本书写成课堂教科书。 尽管我希望一些勇敢的人仍然会选择将它用于此目的——就像以前对 VCA 所做的那样——我的主要目标是尽可能诚实和清晰地向读者传达一个雄伟而强大的主题，不管那个读者是一个稚嫩的新手，还是一个老练的专家。

- 我的主题选择有时可能看起来不拘一格：例如，为什么没有关注极小曲面这个迷人而重要的主题？ 通常，在这种情况下，这是出于以下两个原因之一（或两者）：（1）我们的重点是内蕴几何，而不是外在[19]几何；(2) 已有关于该主题的优秀文献；在这种情况下，我试图在本书末尾的延伸阅读部分提供有用的指导。

- 方程式用（圆）括号编号，而图片用[方]括号编号。

- 粗斜体用于突出**新术语的定义**。

- 为便于翻阅本书时参考，值得注意的结果采用了框，而双重显着的事实采用了双重框。在整个工作中，只有少数结果如此基础，以至于它们被三重框起； 我们希望读者会喜欢找到它们，就像复活节彩蛋一样。

- 我试图让读者成为发展这些想法的积极参与者。 例如，随着争论的进行，我经常故意将一对合乎逻辑的踏脚石放在足够远的地方，以便您可能需要暂停并稍微伸展一下才能从一个过渡到另一个。 这些地方标有“[练习]”； 它们通常只需要一个简单的计算或片刻的反思。

- 最后，我们鼓励读者充分利用指数； 它的创造是爱的痛苦劳动！

我们以更广泛的工作哲学目标来结束这个前言，一个超越我们将寻求解释的特定数学的目标。

从数学青春期到成年期的通行证 [原文如此] 之一是能够区分真正的和虚假的奇迹。数学本身充满了前者，但后者的例子也比比皆是：“我不敢相信所有这些丑陋的项都约简了，给我留下了如此简单的答案！”； 或者，“我不敢相信这个复杂的表达有这么简单的意思！”

在这种情况下，与其庆幸自己，不如羞愧地低下头。 因为如果所有这些丑陋的项都约简了，它们一开始就不应该在那里！ 如果这个复杂的表达有一个非常简单的含义，它本来就不应该那么复杂！

就我自己而言，我很尴尬地承认，我的数学青春期一直持续到我 20 多岁，而我成为研究生后才开始长大，这要归功于彭罗斯和我的密友乔治·伯内特-斯图尔特（彭罗斯的一位顾问）的奇妙双重影响。

数学实在的柏拉图形式总是完美地美丽，也总是完美地简单；与之相反的短暂印象是我们自己不完美的表现。 我希望这本书可以帮助读者在面对这种完美时保持谦逊，就像多年前在牛津超现实的埃舍尔式尖顶之间，我的两个朋友第一次推动我走上这条路时一样。

---

[19] 1.4 节解释了“内蕴”和“外在”的含义。

T. N.

加利福尼亚米尔谷
2019 年 牛顿节

致谢

罗杰·彭罗斯改变了我对数学和物理学的理解。从我 20 岁读到他的第一篇论文开始，他的思想的完美、美丽和几乎是音乐的对位，让我产生了一种深刻的审美快感，我只能把它比作聆听巴赫的康塔塔 101 或贝多芬的大赋格开场旋律的体验。

从我还是他的学生开始，罗杰通过几何解开最深奥的奥秘的能力就（给我）留下了不可磨灭的印记——它给我灌输了一种终生的、不可动摇的信念，即几何证明必然永远存在。（后来我对牛顿《原理》的研究加深了我对几何方法普遍性的信念。）没有这种信念，这件作品就不可能存在，因为有时我需要花很多年的摸索才能发现某个数学现象的几何证明。

40 年来，能成为罗杰的朋友，我感到非常高兴和崇高的荣誉，我把这件不完美的作品奉献给罗杰，几乎无法偿还我欠他的知识债，但这是我可以做到最好的。

为了恰当地介绍下一个我要感谢的人，我不得不透露一个关于我自己的一些不堪的细节：1989 年我第一次从英国来到美国时，我每天抽两包烟！ 1995 年，我终于能够戒烟：这是我做过的最艰难的事情，如果不是尼古丁贴片的发明，我很可能会失败。

1997 年发表我的《可视化复分析》(VCA) 大约五年后，作为回应，我收到了杜克大学一位医学研究员的“粉丝来信”； 他打算来湾区看看，问能不能见面。我有些忐忑地答应了。我的来访者竟然是我的（救世主）尼古丁贴片的发明者杰德·罗斯（Jed Rose）教授！ 杰德从数学和物理出发，从未失去对这些学科的热爱，但他精明地算计到如果将精力投入到医学研究上，他会产生更大的影响。 我很庆幸他这么做了！

2011 年我开始创作 VDGF（《微分几何和形式：可视化方法》，这本书）后，杰德成为我最热情的支持者，他非常慷慨地使用他的医学发明资金买断了我的部分课程，从而极大地帮助了我对 VDGF 的研究。在写作这本书的九年时间中，每次杰德拜访我和我在加利福尼亚的家人时，我的精神都被杰德不屈不挠的乐观性格和他对我试图完成事情重要性的信念所鼓舞。 而且，随着手稿的缓慢演化，杰德提供了大量详细而有用的建议和更正。 由于他的有益观察，完成的工作明显更好。 因此，如你所见，杰德在三个线性独立的方向上帮助了我，我对他感激不尽。 而这一切似乎还不够，最初是纯粹的知识性关系，后来发展成为我们两个家庭之间非常温暖和亲密的友谊。

下一个关键人物我想感谢布朗大学杰出的几何学家托马斯·班绰夫（Thomas Banchoff）教授。在写这本书的过程中，我设法安排汤姆分两年来南佛罗里达大学做访问学者，每次一个学期。在这两个学期中，汤姆对我都非常慷慨，愿意阅读我不断发展的手稿，并给了我非常宝贵的反馈。每周他都会阅读最新一期的手稿，热火朝天，然后他和我会在每个星期五下午在他的办公室会面，仔细检查他的详细更正和建议（在页边空白处用红笔逐行书写）。虽然这本书只写了一半这种合作就不幸结束了，但我基本上采纳了他所有有益的建议和更正，我非常感谢他与我分享他深厚的几何智慧和专业知识。

衷心感谢刘伟[20]（Wei Liu）博士（一位专门从事光学研究的物理学家，我希望在某个美好的日子里最终能与他见面）。 2019 年，他写信给我表达了他对 VCA 的赞赏，并附上了他的[21]一篇研究论文，其中引用了我对庞家莱-霍普夫定理的处理。这篇论文让我大开眼界，让我看到物理学家如何继续精彩地利用似乎已从所有数学教科书中完全消失的霍普夫的美丽结果。该结果是第 19.6.4 节的主题，它是这样说的：庞家莱-霍普夫定理不仅适用于向量场，而且适用于霍普夫线场[22]，它极大地概括了向量场并且可以具有带分数指数的奇点。看看第 212 页 [19.14] 中显示的例子。应我的要求，刘博士随后进一步帮助我，指出物理学家对霍普夫的思想所做的许多其他应用。亲爱的读者，我在本书末尾的延伸阅读部分与你分享了他的善意指导。

除了上面的主要参与者，我还收到了来自远近的同事和朋友的各种指导、支持和建议。

我心爱的兄弟盖伊是一位主播，他对我的爱和信任常常被视为理所当然，但不应该如此！

斯坦利 • 内尔（Stanley Nel）和保罗 • 蔡茨（Paul Zeitz），我 30 年的朋友，一直比我自己更相信我，他们的鼓励对我来说意义重大，多年来我一直在努力创作这本书。 发现所需的几何洞察力，然后写和画这本书。

道格拉斯 • 霍夫施塔特（Douglas Hofstadter）——作为一名本科生，他的哥德尔、埃舍尔、巴赫让我（以及数以百万计的人）目瞪口呆——20 多年来，他一直对我的支持让我感到很荣幸。 首先，他在印刷品和采访中一再大力宣传 VCA。 其次，他阅读并提供了关于 Needham (2014) 的非常有价值的反馈，最终被纳入本书。

埃德 • 卡特穆尔（Ed Catmull）博士 — 与史蒂夫 • 乔布斯 (Steve Jobs) 的皮克斯的联合创始人，后来成为皮克斯和沃尔特 •迪斯尼动画工作室的总裁 — 早在 1999 年就给我写了一封关于 VCA 的非常讨人喜欢的电子邮件。起初，我确信我的大学之一 旧金山 (USF) 的数学伙伴在跟我开玩笑，但电子邮件是真实的。 Ed 邀请我参观 Pixar 校园（当时还在里士满角），带我参观了工作室，带我出去吃午饭，还给了我一份工作！ （我将留给读者为我拒绝那个提议的愚蠢行为分配一个数值。）虽然我与 Ed 的联系是零星的，但他一直是 VCA 的忠实支持者（在他所做的采访中称赞它）， 他还给我写了一封推荐信以获得资助，他一直非常支持我写作 VDGF 的努力。 我深深感谢 Ed 多年来对我的鼓励，感谢他在本书背面写的非常客气的话。

普林斯顿大学出版社最初与弗兰克摩根教授（我只知道他的名声）接洽，要求对 VDGF 的手稿进行匿名审查。 但是当他将他的评论提交给我的编辑时，他也直接以他自己的名义发给了我。 我非常感谢他选择这样做，因为现在我可以公开感谢他的具体建议和更正。 此外，我特别感谢他的报告极大地鼓舞了我当时的士气。 最后，我衷心感谢他愿意在本书的背面分享他对 VDGF 的非常慷慨的评价。

同样，我也感谢所有真正匿名的审稿人对我提出的建设性批评、建议和更正——我已尝试将他们的所有改进纳入其中，但很抱歉我不能点名感谢他们每个人。

---

[20] 国防科技大学前沿交叉学科学院，湖南长沙。

[21] Chen et al. (2020b).

[22] 这是现代术语； Hopf (1956) 最初称它们为线元场（fields of line elements）。

我感谢 M. C. Escher 公司允许复制圆极限 I（Circle Limit I）的两个修改：[5.11] 和 [5.12]，后者是由约翰 • 斯蒂尔威尔（John Stillwell）执行的显式数学变换，并在他的许可下使用。请注意，M. C. Escher 的圆极限 I 是 © 2020，M. C. Escher 公司，荷兰。 版权所有。www.mcescher.com。

最后，我非常感谢亨利 •塞格曼（Henry Segerman）教授向我提供了他的拓扑笑话[18.8]，第 191 页的图像，并允许我在这里复制它。

这是我的第二本书，也是我的最后一本书。 因此，我不仅要感谢所有直接帮助我写作上述 VDGF 的人，还要感谢那些在我生命早期影响和支持我的人。 其中一些人是如此天衣无缝地融入了我的存在结构，以至于他们变得隐形了，而且，可耻的是，我未能在 VCA 中适当地感谢他们； 现在是我纠正错误的最后机会。

其中首先是我最老的朋友安东尼 • 利维（Anthony Levy），我们在牛津大学默顿学院的本科时代。 令人费解的是，安东尼（或托尼，我当时认识他）早在有任何证据支持这种信念之前就相信了我，而且在几十年来数学自我怀疑期间，对我的持续相信不断地鼓舞着我。 而且，在纯粹的智力世界之外，安东尼睿智的建议和爱帮助我度过了我生命中最令人担忧的一些插曲。

同样从默顿本科时代开始，我将永远感谢我的两位物理导师迈克尔贝克博士（1930-2017）和迈克尔 G.鲍勒博士，他们不仅教了我很多物理知识，而且还 他们不遗余力地为我安排了更高级的广义相对论和自旋量的个人辅导，由默顿的两位杰出研究员布赖恩 • D • 布拉姆森博士和理查德 • 斯沃德博士授课。 特别是，布拉姆森博士对科学的热情极具感染力，是他让我第一次接触到彭罗斯的（具有启发性的！）工作，并促使我申请在彭罗斯的“相对论小组”攻读博士学位。 ”

进入我在彭罗斯（Penrose）学习的研究生时代，我想再次感谢我的朋友乔治 • 伯内特-斯图尔特（George Burnett-Stuart），他是彭罗斯（Penrose）的顾问。 乔治和我在大克拉伦登街合住了几年的小房子，因为我们进行了博士工作，在我们无休止地讨论音乐、物理和数学的过程中，乔治帮助我完善了我对数学的本质，以及什么构成了该主题中可接受的解释。 无论好坏，乔治对我今天的数学家负有重大责任。

再次向前迈进，对于我在旧金山大学的生活，我感谢刚刚退休的约翰 •斯蒂尔威尔（John Stillwell）和他出色的妻子伊莱恩（Elaine），他们作为同事和朋友度过了 20 年快乐而富有成效的岁月。 虽然我在写这本书的过程中多次动了他的脑筋，但我最大的亏欠是他的许多著作。 事实上，我们的关系始于我写给他的一封“粉丝信”，以回应他宏伟的巨著《数学及其历史》的第一版。 几年后，在担任科学学院副院长期间，我成功地将约翰引诱离开澳大利亚，在旧金山大学为他设立了教授职位。 VCA 和 VDGF 很大程度上归功于约翰（John）对整个数学领域的整体把握，以及他利用这种观点赋予数学思想意义的能力，所有这些都是他通过他的许多精彩书籍慷慨地与世界分享的。

1996 年，我在 VCA 的致谢中总结了这样的话：“最后，我感谢我最亲爱的妻子玛丽。在写这本书的过程中，她让我假装科学是生活中最重要的事情。 现在这本书已经结束了，

她每天都是我的证明，证明还有更重要的事情。” 二十多年后的今天，我对玛丽的爱只是越来越深，但我现在比以前多了两个日常证明！

1999 年，玛丽和我有幸生下了双胞胎：从那时起，信念和希望一直是我们骄傲和快乐的耀眼灯塔。 很抱歉，VDGF 像乌云一样笼罩了我女儿近一半的生活，它剥夺了我们在一起的时间。 然而，正是这三个灵魂的爱赋予了我生命的意义和目标，并在创作这部作品的漫长斗争中支撑着我。

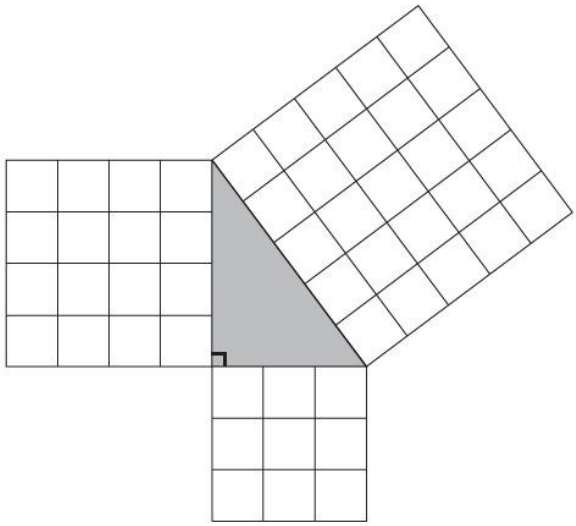
第一章 欧式与非欧几何

1.1 欧式与双曲几何

微分几何是将微积分应用于弯曲的空间几何。 但是要理解弯曲的空间，我们首先要尝试理解平坦的空间。

我们居住在一个充满弯曲物体的自然世界，如果一个孩子问我们“平坦”这个词的含义，我们最有可能回答的是没有弯曲：没有任何凹凸的光滑表面。然而，最早的数学家似乎已经被平面的独特简单性和统一性所吸引，作为回报，他们发现了关于在平面内构造的几何图形的惊人美丽事实。事后看来，其中一些事实可以被视为平面平坦度的特征。

毕达哥拉斯定理是最早和最深刻的此类事实之一。当然，古人一定是敬畏的，就像今天任何有悟性的人一样，一个看似纯粹的关于数字的事实， $3^2+4^2=5^2$ ，实际上具有几何意义，如[1.1][1]所示。



[1.1] 毕达哥拉斯定理

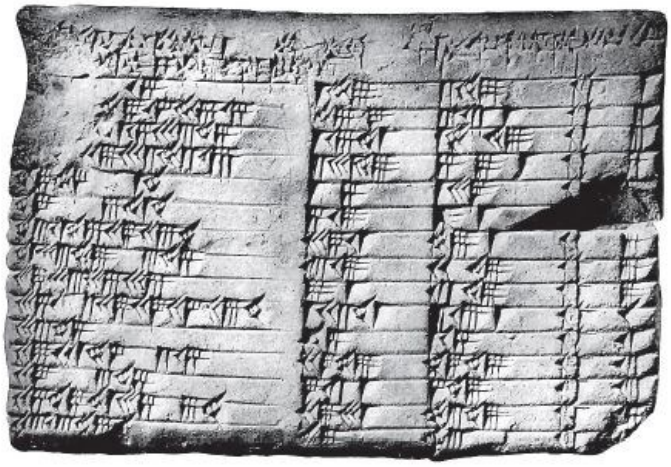
虽然毕达哥拉斯本人在公元前 500 年左右生活在希腊，但毕达哥拉斯定理在世界各地早已被发现。已知最早的记录在 [1.2] 中所示的巴比伦粘土碑（目录为“普林顿 322”）中，该碑出土于现在的伊拉克，可追溯到公元前 1800 年左右。

碑文中列出了毕达哥拉斯三元数组[2]（Pythagorean triples）：整数 (a,b,h)，使得 h 是具有直角边 a 和 b 的直角三角形的斜边，因此 $a^2+b^2=h^2$。有些古老三元组异乎寻常的

1 我们重复序言中所说的话：方程用括号（圆括号）标记——（…），而图形用方括号标记——[…]。

2 事实上，粘土碑只记录了三元组 (a,b,h) 的两个成员 (a,h)。

大，显而易见，它们并非被偶然发现，而是有一个生成解的数学过程。例如，碑文第四行记录了 $13500^2 + 12709^2 = 18541^2$。



[1.2] 普林顿 322：公元前 1800 年的毕达哥拉斯三元数组粘土碑

支撑这些古老结果的更深层次的知识尚不清楚[3]，但要找到“现代的”、按照逻辑的、演绎式的数学方法的第一个证据，我们必须跨越粘土碑之后的 1200 年。学者们认为，是米利都的泰勒斯（公元前 600 年左右）首先开创了从先前建立的结果中推导出新结果的想法，逻辑链始于少数明确表达的假设或公理。

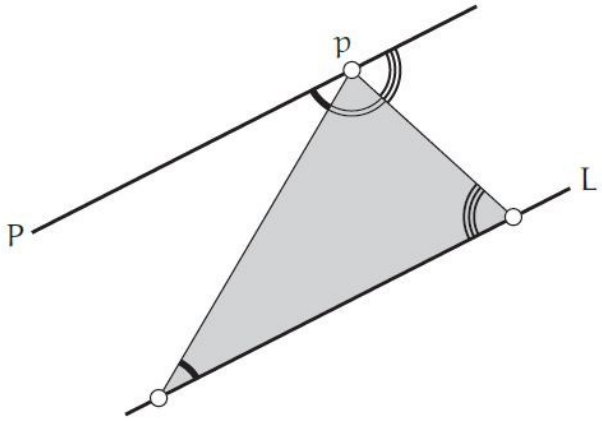
再次向前跨越泰勒斯之后的 300 年，我们在欧几里德的《几何原本》中找到了这种新方法最完美的范例之一，可以追溯到公元前 300 年。本书试图通过从五个简单的公理中推断出一切，从而为几何带来秩序、清晰和严谨，其中第五个也是最后一个公理涉及平行线。

定义两条不相交的线是平行的，欧几里得第五公理[4]在 [1.3] 中说明：

> **平行公理** 通过不在直线 L 上的任意点 p，有且只有一条平行于 L 的直线 P。

3 17 世纪，由于丢番图的功劳，费马和牛顿重建并推广了一种生成通解的几何方法。见练习 5。

4 欧几里得并没有以这种形式陈述他的公理，但它在逻辑上是等价的。



[1.3] 欧几里德平行公理：P 是通过 p 与 L 平行的唯一直线，三角形的内角和是 π。

但是，这个公理的性质比前四个公理更复杂，也不那么直接，数学家们开始了长期的斗争，试图摒弃这个假设，而是试图证明它一定是前四个公理的逻辑结果。

在接下来的 2000 年里，这种紧张局势一直没有得到解决。 随着几个世纪的流逝，人们进行了许多尝试来证明平行公理，这些努力的数量和强度在 18 世纪达到了高潮，但都失败了。

然而，在此过程中，公理的有用等价形式出现了。例如：存在不同大小的相似三角形（Wallis 在 1663 年；参见 Stillwell (2010)）。但是，第一个等价形式在欧几里德时代已经存在了，而且它是每个学童都教过的：三角形内角和等于两个直角（见[1.3]）。

对这些失败的解释是在 1830 年左右才出现的。尼古拉·罗巴切夫斯基（Nikolai Lobachevsky）和鲍耶·亚诺什（János Bolyai）完成了 4000 年前开始的旅程，独立宣布发现了一种全新的几何形式（现在被称为双曲几何），发生在一种新的平面（现在被称为双曲平面）中。在这个几何学中，前四个欧几里德公理成立，但平行公理不成立。相反，以下情况是正确的：

双曲公理 至少有两条通过 p 的平行线不与 L 相交。 (1.1)

这些先驱们探索了这个公理的逻辑推论，并通过纯粹的抽象推理，在一个与欧几里得的几何学截然不同的丰富新几何学中得出了许多令人着迷的结果。

在他们之前的许多其他人，也许最引人注目的是萨凯里（Saccheri，1733）和兰伯特（Lambert，1766），发现了（1.1）的一些结果，但他们探索的目的仅为了找到矛盾，他们相信这将最终证明欧几里德几何是唯一真正的几何。

当萨凯里出版《消除所有缺陷的欧几里德》（Euclid Freed of Every Flaw）时，他当然相信自己已经发现了一个明显的矛盾。但兰伯特是一个更令人困惑的案例，他可能是这个故事中的无名英雄。 他的结果如此深入地渗透到了这种新几何学中，以至于他有时似乎不可能不

相信他所做的事情的真实性。不管他的动机和信仰如何[5]，兰伯特（如[1.4]所示）无疑是第一个发现公理（1.1）下三角形内角和的非凡事实[6]的人，他的结果将成为第二幕后面的许多内容的核心。

尽管如此，罗巴切夫斯基和鲍耶是第一个认识到并完全接受他们发现了一个全新的、一致的、非欧几里德几何的想法的人，这一点他们当之无愧。但这种新几何的真正含义，以及它可能有用的地方，甚至他们也说不出来[7]。

[1.4] 约翰·海因里希·兰伯特（1728 年-1777 年）

令人惊讶的是，曲面的微分几何最终解决了这些问题。正如我们将解释的那样，1868 年，意大利数学家尤金尼奥·贝尔特拉米（Eugenio Beltrami）终于成功地给双曲几何一个具体的解释，将其建立在一个坚实而直观的基础上，并由此而发展壮大。可悲的是，罗巴切夫斯基和鲍耶都没能活着看到这一天：他们分别于 1856 年和 1860 年去世。

这种非欧几里德几何实际上已经在历史上数学的各个分支中均有所表现，但总是晦暗不明。亨利·庞加莱（Henri Poincaré，从 1882 年左右开始）不仅是第一个揭开它的伪装的人，而且还在复分析、微分方程、数论和拓扑等不同领域认识到并利用它的力量。它在 20 世纪和 21 世纪数学中的持续活力和中心地位体现在瑟斯顿关于 3 流形的工作、怀尔斯对费马大定理的证明和佩雷尔曼对庞加莱猜想的证明（作为瑟斯顿几何化猜想的一个特例），仅举

---

5 感谢罗杰·彭罗斯（Roger Penrose）让我看到兰伯特应该得到比通常更多的荣誉。 彭罗斯通过以下类比做到了这一点："我们是否应该将宇宙常数归功于爱因斯坦，即使他出于错误的原因引入它？ 我们应该责怪他后来收回它，称其为"我一生中最大的错误"吗？ 或者说广义相对论本身呢，随着时间的推移，爱因斯坦似乎越来越不相信它是正确的理论（需要被某种非奇异统一场论取代）？" 【私下交流。】

6 如果你等不及了，那就请看（1.8）。

7 事实上，罗巴切夫斯基确实使用这种几何来评估以前未知的积分，但是（至少事后看来）这个特殊的应用必须被视为相对次要的。

三个例子。

在第二幕中，我们将介绍贝尔特拉米的突破，以及双曲几何的性质，但现在我们希望探索一种不同的、更简单的非欧几里得几何（译者注：以下简称非欧几何），一种古人已经知道的非欧几何。

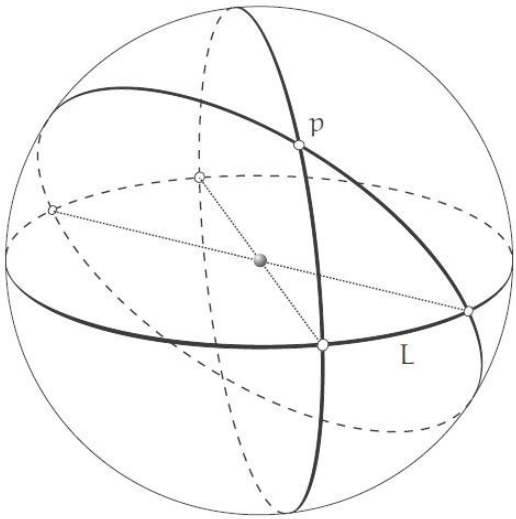
1.2 球面几何

要构造非欧几里得几何，我们必须否认唯一平行线的存在。双曲公理假设有两个或多个平行线，但还有另一种逻辑上的可能性——一条平行线都没有：

> **球面公理** 通过 p 没有平行于 L 的直线：每一条直线都与 L 相交。 (1.2)

因此，实际上有两种非欧几何[8]：球面几何和双曲几何。

顾名思义，球面几何可以在球体的表面上实现（若为单位球体，则可记为 $\mathbb{S}^2$）我们可以把它想象成地球表面。在这个球面上，连接曲面上两点的“直线”应该是什么？答案：它们之间最短的路线！具体说，如果你想从伦敦航行或飞行到纽约，那么最短的路线是什么？



[1.5] $\mathbb{S}^2$的大圆在一组对径点相交

古代水手们已经知道的答案是，最短的路线是一个大圆的弧线，例如赤道，通过用穿过球心的平面切割球面而获得。在[1.5]中，我们选择了 L 作为赤道，很明显，(1.2)是满足的：穿过 p 的每一条线在一对对径（即直径两端的）点与 L 相交。

在平面中，最短的路线也是最直的路线，事实上，在球面上也是如此：在稍后讨论的精确意义上，大圆轨迹在穿过球面时既不向右也不向左弯曲。

---

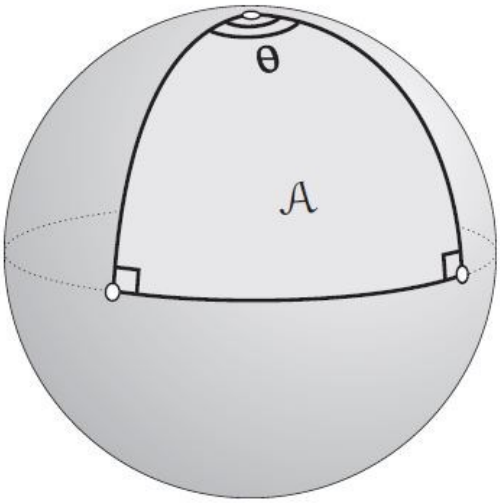
8 尽管如此，读者应该知道，在现代用法中，“非欧几何”通常与“双曲几何”同义。

在地球上构造大圆还有其他的方法，这些方法不需要考虑穿过球心的平面，而球心完全无法到达。例如，在一个地球仪上，你可以通过把一根绳子的一端按在伦敦，把绳子紧紧地绷直在曲面上，使得另一端就在纽约，来画出你的大圆。绷紧的绳子自动找到了最短、最直的路线——穿过两个城市的大圆被这两座城市分割成两段弧线，其中较短[9]的一条。

现在既然找到了直线的对应物，我们就可以在这个球面上“做几何”。例如，给定地球表面上的三个点，我们可以用大圆弧把它们连接起来，得到一个“三角形”。图[1.6]说明了一个顶点位于北极，另外两个位于赤道的情况。

但是，如果古代水手已经使用这种非欧几里得球面几何来导航海洋，天文学家已经使用它来绘制球形夜空，那么罗巴切夫斯基和博利亚的非欧几里得几何又有什么如此令人震惊和新奇的呢？

答案是，这个球面几何仅仅被认为是从球面所在的 3 维空间的欧几里得几何继承而来的。在那个时代，没有人把球面内二维几何当做欧几里德平面的替代方案。它不仅违反了欧几里德的第五公理，而且还违反了一个更基本的公理（欧几里德的第一公理），即两点之间仅有一条直线，因为当两点是对径点时这并不成立。



[1.6] 球面上的一个特别简单的“三角形”

另一方面，罗巴切夫斯基和鲍耶的双曲几何则是对欧式几何的更严重的挑战，它包含了熟知的无限长的直线，但标榜了多个平行线，可笑的角度和，以及许多其他看似荒谬的结果。然而，21 岁的鲍耶对他的发现充满信心和热情，他写信给他的父亲，“*从无到有，我创造了另一个全新的世界。*”

我们以一个悲剧的故事结束。鲍耶的父亲是高斯的朋友，他把亚诺什的成就寄给了他。这时，高斯本人已经在这方面取得了一些重要的发现，但一直保密。无论如何，亚诺什比高斯看得更远。世界上最著名的数学家高斯在公开场合说一句好话，就会让这位年轻的数学家有一个光明的未来。但是，先天和后天有时会合力将非凡的数学天赋注入一个被非常普通的

---

9 如果两点是对径点，比如北极和南极，那么这两条弧的长度相同。 此外，大圆本身不再是唯一的：每条子午线都是连接两极的大圆。

人类缺陷所破坏的容器中，而高斯对鲍耶奇妙发现的反应是极端卑鄙和自私自利。

首先，高斯让鲍耶焦灼地等待了六个月，然后他回答如下：

现在谈谈你儿子的工作。当我开始说我不能赞美它时，你可能会感到震惊，但我不能做任何其他事情，因为赞美它就是赞美我自己。这篇论文的全部内容，你儿子所走的道路，以及他所获得的结果，几乎在所有地方都与我自己的沉思一致，这些沉思部分占据了我过去 30 到 35 年的时间。

然而，高斯确实“感谢”鲍耶的儿子“为他省去了写下他已经知道了几十年的定理的麻烦[10]”。

亚诺什·鲍耶从未从高斯的外科手术式打击中恢复过来，他余生都放弃了数学[11]。

1.3 球面三角形的角盈

正如我们所说，平行公理等价于三角形内角之和为 π。因此，球面公理和双曲公理都必然导致内角和不等于 π 的几何。为了量化这种与欧几里德几何的偏离，我们引入了**角盈**（angular excess），定义为内角和超过 π 的量 $\mathcal{E}$：

$$\boxed{\mathcal{E} \equiv (\text{三角形内角和}) - \pi}$$

例如，对于[1.6]中所示的三角形， $\mathcal{E} = \left(\theta + \frac{\pi}{2} + \frac{\pi}{2}\right) - \pi = \theta$。

如果我们将三角形的角盈与其面积 $\mathcal{A}$ 进行比较，就会出现一个重要的见解。设球面的半径为 R。由于三角形占据了北半球的 $(\theta/2\pi)$，$\mathcal{A} = \left(\frac{\theta}{2\pi}\right) 2\pi R^2 = \theta R^2$。因此，

$$\boxed{\mathcal{E} = \frac{1}{R^2}\mathcal{A}} \tag{1.3}$$

1603 年，英国数学家托马斯·哈里奥特（Thomas Harriot，见[1.7]）做出了惊人的发现

---

10 高斯以前曾以完全相同的方式贬低阿贝尔对椭圆函数的发现。 见 Stillwell (2010, p. 236)。

11 如果这让你感到沮丧，请将想想莱昂哈德·欧拉（Leonhard Euler）令人振奋的事迹来平衡一下。欧拉像一座智慧的火山喷发出众多非常新颖的想法（其中一些我们稍后会遇到），他也是善良和慷慨精神的楷模。我们举一个相应的例子。 当尚不为人知的 19 岁的拉格朗日写信给他，提到微积分变分法中的重复发现时，欧拉回信说：“…我自己也导出了这些内容。 不过，我决定在你发布结果之前隐瞒这件事，因为我绝不想夺走你应得的任何荣耀。” 参见 Gindikin (2007, p. 216)。 顺带一提，欧拉还亲自出手挽救了兰伯特的事业！

[12]，即这种关系适用于球面上的任何三角形 Δ；见[1.8a]。哈里奥特的基本但巧妙的论证[13]如下。

延长 Δ 的大圆边将球面分成八个三角形，标记为Δ、Δ_α、Δ_β、Δ_γ的四个三角形各自与一个全等的对映三角形配对。这在 [1.8b] 中更清楚。由于球的面积是$4\pi R^2$，我们推得

$$\mathcal{A}(\Delta) + \mathcal{A}(\Delta_\alpha) + \mathcal{A}(\Delta_\beta) + \mathcal{A}(\Delta_\gamma) = 2\pi R^2 \tag{1.4}$$

另一方面，在 [1.8b] 很明显，Δ 和 Δ_α 一起形成了一个楔形，其面积是球面面积的一小部分 $(\alpha/2\pi)$：

$$\mathcal{A}(\Delta) + \mathcal{A}(\Delta_\alpha) = 2\alpha R^2$$

相似地，

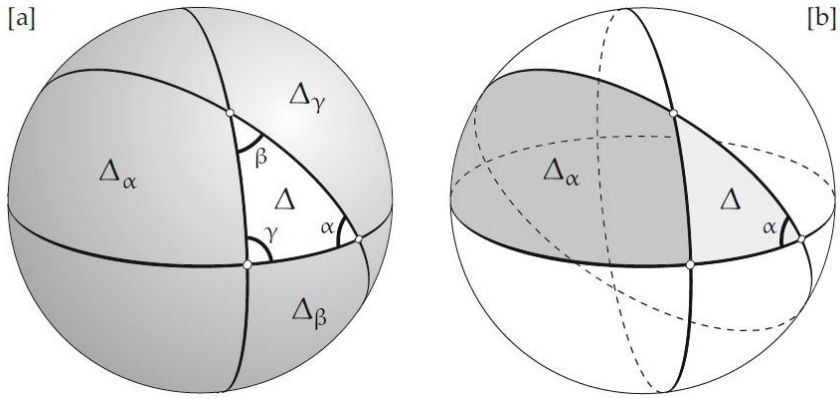
$$\mathcal{A}(\Delta) + \mathcal{A}(\Delta_\beta) = 2\beta R^2$$

$$\mathcal{A}(\Delta) + \mathcal{A}(\Delta_\gamma) = 2\gamma R^2$$

[1.7] 托马斯·哈里奥特（1560－1621）

12 这一发现通常归功于 Girard，他在大约 25 年后重新发现了它。

13 这个论证后来被欧拉在 1781 年重新发现。



[1.8] 哈里奥特定理（1603）：$\mathcal{E}(\Delta) = \mathcal{A}(\Delta)/\mathrm{R}^2$

将最后三个方程相加，我们发现

$$3\mathcal{A}(\Delta) + \mathcal{A}(\Delta_\alpha) + \mathcal{A}(\Delta_\beta) + \mathcal{A}(\Delta_\gamma) = 2(\alpha + \beta + \gamma)\mathrm{R}^2 \tag{1.5}$$

最后，从（1.5）中减去（1.4），我们得到

$$\mathcal{A}(\Delta) = \mathrm{R}^2(\alpha + \beta + \gamma - \pi) = \mathrm{R}^2\mathcal{E}(\Delta)$$

从而证明了（1.3）。

1.4 曲面的内蕴和外在几何

本书后面将深入探讨与这种“直线”的拉紧细绳构造相关的数学。现在我们只观察到该构造可以同样可以很好地应用于非球形曲面，例如[1.9]中所示的弯颈南瓜。

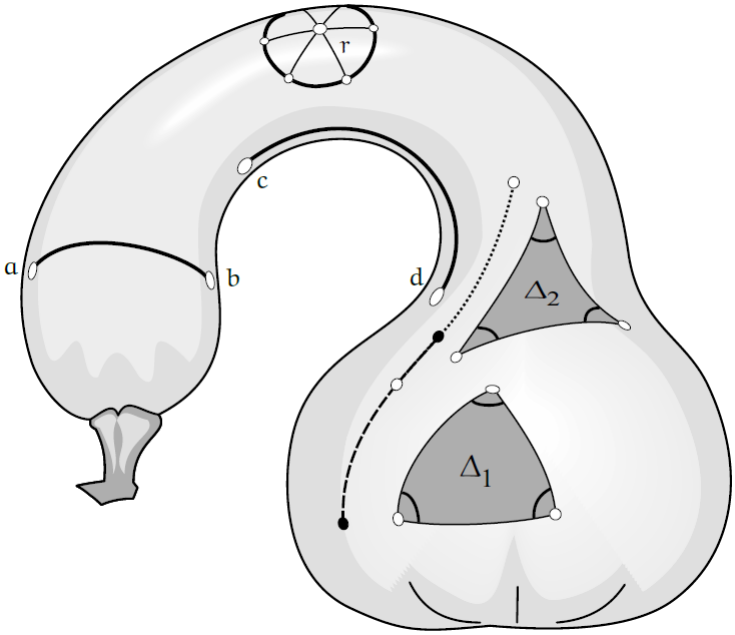
就像在球面上一样，我们在曲面上拉紧一根细绳，从而找到两点之间最短、最直的路线，例如 a 和 b。如果绳子可以轻松地在曲面上滑动，绳子中的张力确保产生的路径尽可能短。请注意，在 cd 的情况下，我们必须想象细绳在曲面的内部运动。

因此，为了以一致的方式处理所有可能的点对，最好将曲面想象为由两个薄薄的分离层组成，细绳夹在它们之间。另一方面，这只对思想实验有用，而不是实际实验。我们将很快通过提供一种在物理对象表面上构建这些最直曲线的实用方法来克服这个障碍，即使曲面片以错误的方式弯曲，使绳子紧紧地拉伸在物体的外部。

这些曲面上的最短路径相当于平面中的直线，它们将在本书中发挥至关重要的作用——它们被称为测地线（geodesics）。因此，用这个新词，我们可以说平面上的测地线是直线，球面上的测地线是大圆。

但即使在球面上，“测地线”的长度最小化定义也是临时的，因为我们看到非对径点由穿过它们的大圆的两条弧连接：短的（这是最短的路线）和长的。然而，长弧与短弧一样是

测地线。球面上还有一个额外的复杂性，即对径点由多个测地线连接，这种非唯一性也出现在更一般的曲面上。事实是，任何两个足够靠近的点都可以通过唯一的测地线段连接起来，该测地线段是它们之间的最短路径。



[1.9]弯颈南瓜曲面的**内蕴几何**：测地线相当于直线，由它们形成的三角形其角盈可能带任何符号，这取决于曲面的弯曲方式：$\mathcal{E}(\Delta_1) > 0$ 和 $\mathcal{E}(\Delta_2) < 0$。

正如平面中的线段可以通过重叠线段在任一方向无限延伸一样，测地线段也可以在曲面上延伸，而且这种延伸是独一无二的。例如，在[1.9]中，我们通过在白点之间放置重叠的虚线段来扩展连接黑点的虚线测地线段。

由于与测地线的长度最小化特征相关的微妙之处，不久我们将提供一种替代的、纯局部的测地线特征，基于它们的直线性。

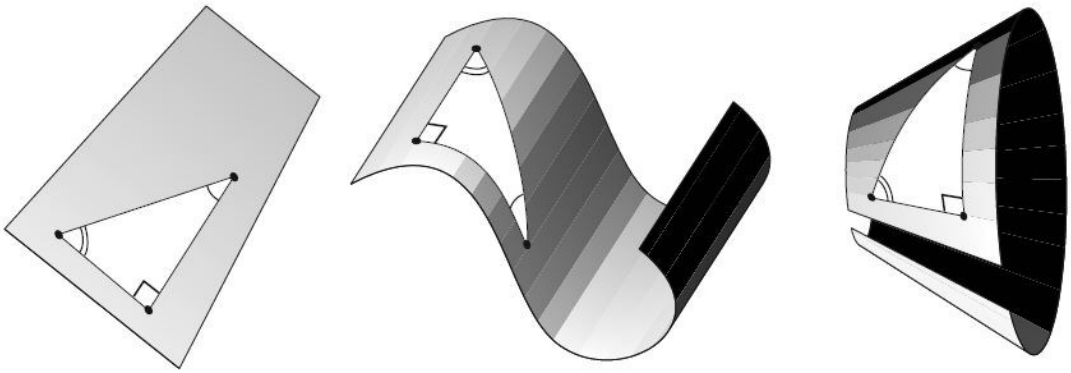
有了这些注意事项，现在很清楚我们应该如何定义曲面内的距离，例如 [1.9]：两个足够接近的点 a 和 b 之间的距离是连接它们的测地线段的长度。

图[1.9]显示了我们如何定义，例如，“半径 r 和中心 c 的圆”作为与 c 距离为 r 的点的轨迹。为了构建这个测地圆（geodesic circle），我们可以取一根长度为 r 的绳子，将一端固定在 c 处，然后（保持绳子拉紧）将另一端拖到曲面上。但正如三角形的内角和不再等于 π，圆的周长也不再等于 2πr。 事实上，你应该能够说服自己，对于图示的圆，周长小于 2πr。

给定曲面上的三个点，我们可以将它们与测地线连接起来形成一个测地三角形（geodesic triangle）；[1.9]显示了两个这样的三角形，Δ_1 和 Δ_2：

- 查看 Δ_1 中的角度，似乎很明显，它们的总和大于 π，所以$\mathcal{E}(\Delta_1) > 0$，就像球面几何中的三角形。
- 另一方面，同样清楚的是，Δ_2的内角和小于 π: $\mathcal{E}(\Delta_2) < 0$，而且（正如我们将解

释的）这种相反的行为实际上由双曲几何中的三角形表现出来。还要注意，如果我们在曲面的这个马鞍形部分构建一个圆，则周长现在大于 2πr。



[1.10]弯曲一张纸会改变外在几何，但不会改变内蕴几何

测地线的概念属于所谓的曲面内蕴几何（intrinsic geometry）——一种全新的几何学观点，由高斯（1827）引入。这意味着生活在曲面上的微小的、类似蚂蚁的、智能的（但二维的！）生物可以知道的几何形状。正如我们所讨论的，例如，这些生物可以将连接两个附近点的测地线“直线”定义为它们世界（表面）内连接这两个点的最短路径。 从那里他们可以继续定义三角形，依此类推。 以这种方式定义，只要曲面内的距离没有以任何方式拉伸或扭曲，当曲面在空间中弯曲（如一张纸可以）成完全不同的形状时，很明显内蕴几何形状不会改变。对于曲面上的蚂蚁状生物来说，这种变化是完全无法察觉的。

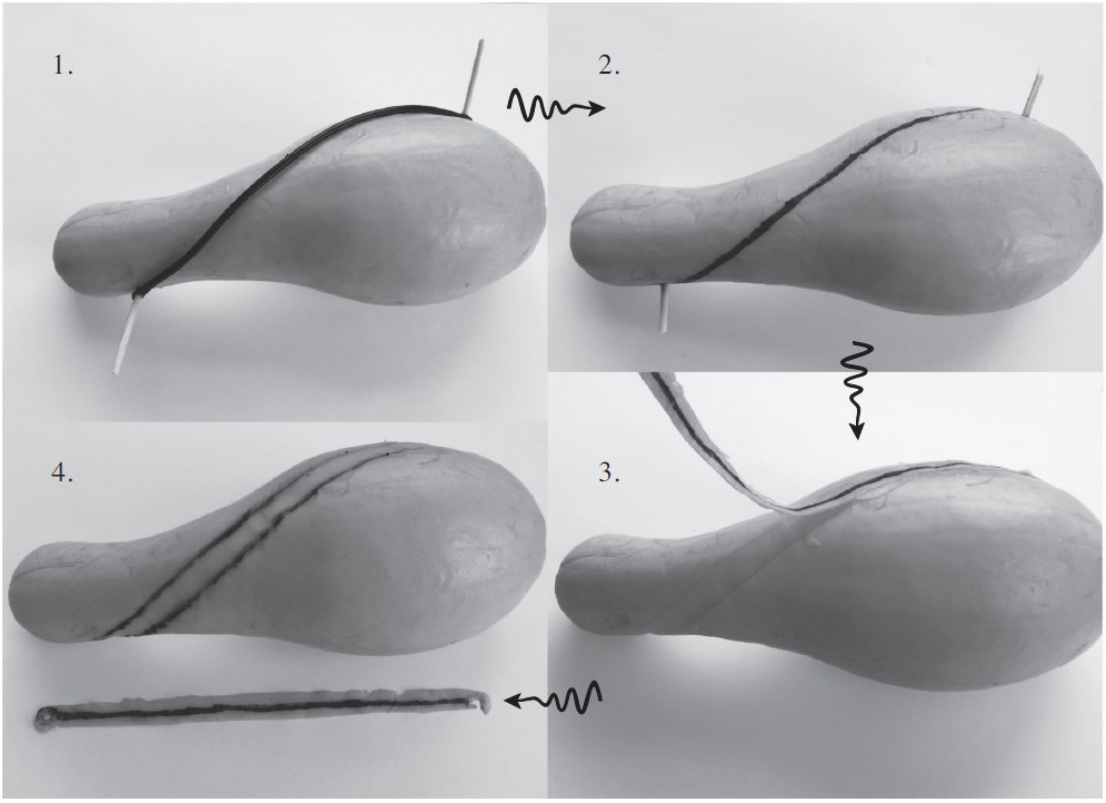
在这样的弯曲下，所谓的外在几何（extrinsic geometry，曲面置入空间中的方式）肯定会发生变化。见[1.10]。左边是一张平面纸，我们在上面画了一个三角形 Δ，角为(π/2)、(π/6)和(π/3)。当然$\mathcal{E}(\Delta) = 0$。显然，我们可以将这样一张平坦的纸弯曲成右侧的两个（外在）曲面中的任何一个[14]。然而，内蕴地看，这些曲面根本没有发生任何变化——它们都像煎饼一样平坦! 这些曲面上的图示三角形（Δ 由纸的无拉伸弯曲带入其中）与智能蚂蚁使用测地线构建的三角形相同，并且在这两种情况下$\mathcal{E} = 0$：这些曲面上的几何形状都是欧几里德的。

即使我们取一片内蕴弯曲的曲面，使得其中的三角形具有 $\mathcal{E} \neq 0$，它通常也可以稍微弯曲而不拉伸或撕裂，从而改变其外在几何形状而保持其内蕴几何不变。例如，将一个乒乓球切成两半，然后轻轻挤压其中一个半球的边缘，将圆形边缘扭曲成椭圆形（但不是一个位于一个平面上的椭圆形）。

1.5 利用直性构造测地线

我们已经提到，曲面上的测地线与平面中的直线至少有两个共同特征：（1）它们提供了相距不太远的两点之间的最短路径； （2）它们提供了这些点之间的“最直”路线。在本节中，我们试图阐明“直”（straightness）的含义，从而得出一种在物理曲面上构建测地线的非常简单实用的方法。

14 但请注意，要将其弯曲成最右边的形状，我们必须先修剪矩形的边。



[1.11]在水果或蔬菜的曲面上，剥下围绕测地线的窄条，然后将其平放在桌子上。你将在平面上获得一条直线！

大多数关于微分几何的课本很少关注这些实际问题，也许正是由于这个原因，我们将要描述的构造在文献中鲜为人知[15]。与之形成鲜明对比的是，在这本书中，我们敦促你通过一切可能的方式探索这些想法：理论思考、绘图、计算机实验，以及（尤其是！）实际曲面的物理实验。你当地的水果和蔬菜店可以为你的实验室提供许多有趣的形状，例如[1.11]中所示的黄色西葫芦。

我们现在可以使用这种蔬菜通过一个实验来揭示测地线隐藏的直性，我们希望你自己重复一遍：

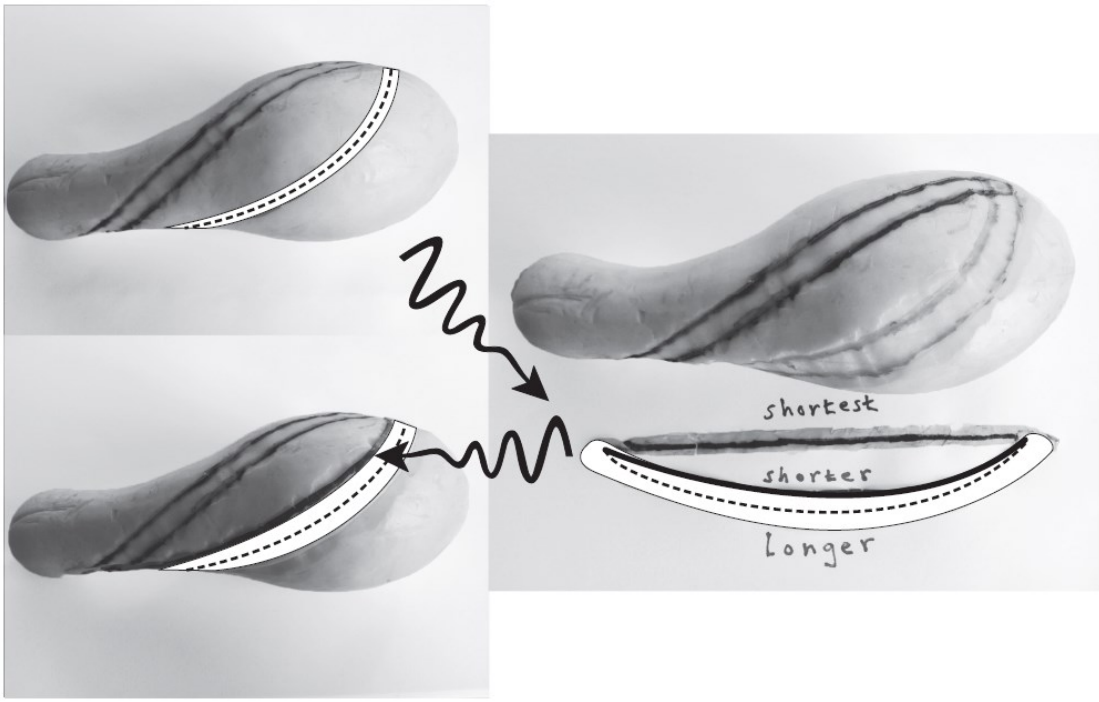
1. 在水果或蔬菜上，通过在其曲面上拉紧一根绳子来构建测地线。
2. 用笔描出绳子的路径，然后将绳子取出。
3. 在着墨路径的两侧（并靠近）做浅切口，然后用蔬菜削皮器或小刀去除两个切口之间的窄条果皮。
4. 将果皮条平放在桌子上，见证剥皮条内的测地线在平面上变成一条直线的神奇事实！

但为什么?!

为了理解这一点，首先让我们明确一点，尽管条带可以在垂直于曲面的方向上自由弯曲

15 罕见的例外之一是 Henderson (1998)，我们强烈推荐给你； 有关详细信息，请参阅本书末尾的延伸阅读部分。

（即垂直于自身），但如果我们试图将其向侧面弯曲，与曲面相切，它就是刚性的。现在让我们用反证法来证明，想象一下，如果这种剥落的测地线平放在桌子上时没有产生一条直线会发生什么。进行这样的物理实验既是缺点也是优点，它们根本不允许我们构造一些不可能的东西，正如我们想要的数学矛盾证明所要求的那样。 尽管如此，让我们假设存在一条测地线路径，例如 [1.12] 左上角所示的虚线路径，当剥开并平放在桌子上（右侧）时，它不会变成一条直线。



[1.12] 假设所示的虚线路径是一条测地线，使得当将其平放在平面上时，围绕它的窄（白色）条不会变成直线。但在这种情况下，我们可以收缩平面中的虚线路径（朝向平面中最短的直线路径），从而生成实线路径。但是，如果我们然后将窄条贴回到曲面上，这条实线路径仍然比原来的虚线路径短，而虚线路径本应是曲面中最短的路径——矛盾！

这条虚线（非直线）平面曲线两端之间的最短路线是连接它们的直线。（如图所示，这是我们已经使用细绳找到的真正测地线的路径——但假装你现在不知道!）因此，我们可以通过将虚线稍微向这条直线最短路径变形来缩短虚线，产生沿剥离条带边缘的实线路径。因此，在将窄条重新粘回到曲面（左下角）后，实线曲线在曲面上提供了比虚线更短的路径，而虚线本应是最短的：矛盾！因此，我们证明了以前的断言：

如果从表面切出一条围绕测地线段 G 的窄条并平放在平面上，则 G 变为直线段。	(1.6)

我们现在非常接近所承诺的简单实用的测地线构造。 再次查看 [1.11] 的第 3 步，我们在此剥离了曲面条带。但是现在想象一下，我们正在将窄条粘回到曲面上。忽略我们是如何走到这一步的历史：我们现在在这个粘回过程中到底在做什么？我们拾起一条窄的直条（三维剥皮，但数学上理想化为二维条带），我们将它展开回到曲面上，进入我们从中切割它的浅通道。但这里有一个关键的观察角度：这个浅通道不需要存在——当我们展开它时，

曲面决定了窄条必须去哪里！

因此，作为（1.6）的一种时间反转，我们获得了一种在物理表面上构建测地线的非常简单实用的方法[16]：

要在曲面上构建测地线，从一点 p 沿方向 **v** 发出，将一段窄胶带的一端向下粘贴在 p 处，然后将其展开到表面上，从方向 **v** 开始。

(1.7)

（但是请注意，这并不提供将 p 连接到指定目标点 q 的测地线构造。）

如果这种构造看起来太简单而令人难以置信，请在你手头的任何曲面上进行尝试。你可以通过在胶带[17]上两点之间的曲面上拉紧一根绳子检查胶带是否确实描绘出测地线：绳子将遵循与胶带相同的路径。但请注意，作为承诺的好处，这种新的胶带构造适用于曲面的任何部分，即使曲面凹向你（此时拉紧细绳的构造会失效）。

当然，这一切都是数学理想化的具体体现。宽度非零的完全平坦的窄条带并不能[18]完美地贴合在真正弯曲的表面上，但它的中心线可以放在曲面上，而胶带的其余部分与曲面相切。

1.6 空间的本质

让我们回到发现非欧几里得几何的历史，首先看看这两种新几何与欧几里得几何的不同之处。

正如前面所说，欧几里得几何的特点是 $\mathcal{E}(\Delta)$ 为零。请注意，与平行公理的原始表述不同，该陈述可以通过实验进行检查：构造一个三角形，测量它的角度，看看它们加起来是否等于π。高斯可能是第一个想到物理空间可能不是欧几里得的人，他甚至尝试了上述实验，将三个山顶作为三角形的顶点，并使用光线作为三角形的边缘。

在他的设备允许的精度范围内，他发现$\mathcal{E} = 0$。非常正确，高斯并没有断定物理空间在结构上肯定是欧几里得的，而是说如果它不是欧几里得的，那么它与欧几里得几何的偏差非

---

16 这个重要的事实在文献中很难找到。在我们 30 多年前（重新）发现它之后，我们开始搜索，我们当时可以找到的最早提到基本想法的是在 Aleksandrov (1969, p. 99) 中，尽管形式不太实用：他想象着将一把灵活的金属尺压在曲面上。后来，基本想法也出现在 Koenderink (1990)、Casey (1996) 和 Henderson (1998) 中。然而，我们从那以后了解到，基本想法（尽管不是我们目前的实用形式）可以追溯到一个多世纪前的列维-奇维塔（Levi-Civita）！请参见第 236 页的脚注。

17 我们建议使用美纹纸胶带（又名画家胶带），因为它具有鲜艳的颜色，并且一旦创建了胶带，就可以轻松地反复撕下和重新粘贴。制作窄条（从通常很宽的一卷胶带）的一种简单方法是将一段胶带粘在厨房砧板上，然后用锋利的刀切成细长的段，制作出你喜欢的窄条。

18 这是我们稍后会遇到的一个基本定理的结果，称为绝妙定理（Theorema Egregium）。

常小。但他确实说（见罗森菲尔德 1988 年，第 215 页），他希望这种非欧几里得几何可以应用于现实世界。在第四幕中，我们将看到这是一个预言性的陈述。

尽管高斯曾向朋友们吹嘘说，他早在几十年前就预见到了罗巴切夫斯基和鲍耶的双曲几何，即使他在不知不觉中也被其一些核心成果所吸引。

1766 年（高斯出生前 11 年），兰伯特重新发现了哈里奥特在球面上的结果，然后在追求双曲公理（1.1）的类似结果方面开辟了全新的领域。首先，他发现双曲几何中的三角形（如果这种东西存在的话）的行为与球面几何中的三角形相反：

- 在球面几何中，三角形内角和大于π: $\mathcal{E} > 0$。
- 在双曲几何中，三角形内角和小于π: $\mathcal{E} < 0$。

因此，双曲三角形的行为类似于在鞍形表面上绘制的三角形，如[1.9]中的Δ_2。稍后我们将看到这并非偶然。

此外，兰伯特发现了一个关键事实，即 $\mathcal{E}(\Delta)$ 再次与 $\mathcal{A}(\Delta)$ 成正比：

在球面几何和双曲几何中都有，

$$\mathcal{E}(\Delta) = \mathcal{K}\mathcal{A}(\Delta) \tag{1.8}$$

其中$\mathcal{K}$在球面几何中是正常数，在双曲几何中是负常数。

与这个结果有关的几个有趣的观察结果：

- 尽管它们之间没有质的差异，但根据正常数 $\mathcal{K}$ 的值，仍然有无数种不同的球面几何。同样，$\mathcal{K}$ 的每个负值都会产生不同的双曲几何。

- 由于三角形的内角和不能为负，$\mathcal{E} \geq -\pi$。因此，在双曲几何（$\mathcal{K} < 0$）中，我们得到了一个奇怪而令人惊讶的结果，即没有三角形的面积可以大于$|\pi/\mathcal{K}|$。

- 从（1.8）我们推断出两个不同大小的三角形不可能有相同的角度。换句话说，在非欧几里得几何中，相似的三角形不存在！（这与瓦利斯 1663 年的发现一致，即相似三角形的存在等价于平行公理。）

- 与前一点密切相关的是，在非欧几里得几何中存在一个绝对长度单位。（高斯本人发现这个纯数学事实可能在物理世界中实现是一种令人兴奋的可能性。）例如，在球面几何中，我们可以将这个绝对长度单位定义为等边三角形的边，例如，内角和 1.01π；类似地，在双曲几何中，我们可以将其定义为内角和为 0.99π 的等边三角

形的边。

- 定义绝对长度单位的一种更自然的方法是根据常数 $\mathcal{K}$。由于角度的弧度度量定义为长度的比率，因此 $\mathcal{E}$ 是一个纯数。另一方面，面积 $\mathcal{A}$ 的单位为$(\text{长度})^2$。由此可见，$\mathcal{K}$ 必须具有$1/(\text{长度})^2$的单位，因此存在一个长度 R，使得 $\mathcal{K}$ 可以写成如下：在球面几何中 $\mathcal{K} = +(1/\mathrm{R}^2)$；在双曲几何中 $\mathcal{K} = -(1/\mathrm{R}^2)$。当然，在球面几何中，我们已经知道公式 $\mathcal{K} = +(1/\mathrm{R}^2)$ 中出现的长度 R 就是球面的半径。稍后我们将看到出现在公式 $\mathcal{K} = -(1/\mathrm{R}^2)$ 中的这个长度 R 可以在双曲几何中给出同样直观和具体的解释。

- 三角形越小，就越难将其与欧几里得三角形区分开来：只有当线性尺寸是 R 的很大一部分时，差异才会变得可辨别。 例如，与地球的半径相比，我们人类很小，所以如果我们发现自己在湖中央的一艘船上，它的表面似乎是一个欧几里得平面，而实际上它是一个球面的一部分。 这种对小图形的欧几里得错觉是高斯选择尽可能大的三角形来进行光线实验的原因，从而增加了他检测光线传播空间中可能存在的任何小曲率的机会。

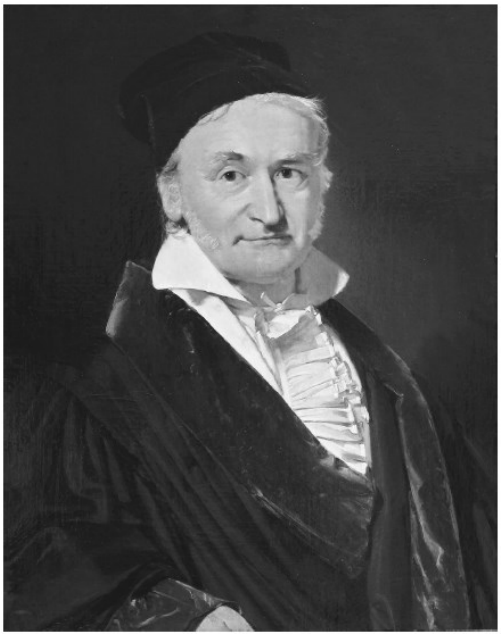
第二章 高斯曲率

2.1 引言

比例常数，

$$\mathcal{K} = +\frac{1}{\mathrm{R}^2}$$

通过哈里奥特的结果（1.3）进入球面几何，称为球面的**高斯曲率**[1]（Gaussian curvature）。半径 R 越小，球体表面的弯曲越紧密，高斯曲率 $\mathcal{K}$ 的值越大。



[2.1] 卡尔·弗里德里希·高斯（Carl Friedrich Gauss,1777-1855）

同样，在双曲几何中，负常数

$$\mathcal{K} = -\frac{1}{\mathrm{R}^2}$$

出现在（1.8）中的也称为高斯曲率，原因我们将稍后解释。

高斯（经过十年的不懈研究）在 1827 年出版的革命性的《关于曲面的一般研究》[2]中发表了这个内蕴[3]（intrinsic）概念 $\mathcal{K}$。

---

1 也称为高斯曲率、内蕴曲率、总曲率或就是普通的曲率。

2 高斯（1827）。

3 正如我们稍后将讨论的那样，奥林德·罗德里格斯（Olinde Rodrigues）早在 1815 年

正如我们现在所解释的，高斯引入了这个概念来测量一般的不规则曲面上每个点的曲率（如[1.9]中所示）。这个曲率的思想将主导未来的一切。根据哈里奥特和兰伯特的结果（1.8），

$$\mathcal{K} = \frac{\mathcal{E}(\Delta)}{\mathcal{A}(\Delta)} = \text{单位面积的角盈}$$

在球面几何和双曲几何中，这种解释适用于任何大小和任何位置的三角形 Δ。但在更一般的曲面上，例如在 [1.9] 中，这个定义是没有意义的，因为即使 $\mathcal{E}$ 的符号也在位于表面不同部分的三角形之间变化，例如 Δ_1 和 Δ_2。

[2.2] 一点处的高斯曲率 $\mathcal{K}(\mathrm{p})$ 是测地三角形收缩到该点时每单位面积的角盈。在此示例中，$\mathcal{K}(\mathrm{p}) > 0$ 且 $\mathcal{K}(\mathrm{q}) < 0$。

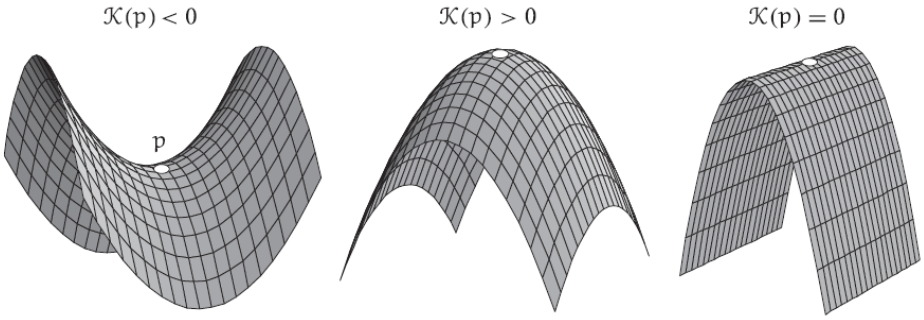
为了在这样的曲面上定义点 p 处的高斯曲率，我们现在想象一个包含 p 的小测地三角形Δ_{p}，然后让三角形向 p 收缩。

使用上一节中发现的测地线构造，[2.2]描绘了这样一个收缩三角形的序列，它们向充气游泳圈曲面（其数学名称是环面 torus）上的一个点收敛。我们现在将 p 处的高斯曲率 $\mathcal{K}(\mathrm{p})$ 定义为该三角形向 p 收缩时的极限：

$$\boxed{\mathcal{K}(\mathrm{p}) = \lim_{\Delta_{\mathrm{p}} \to \mathrm{p}} \frac{\mathcal{E}(\Delta)}{\mathcal{A}(\Delta)} = \mathrm{p}\ \text{点处每单位面积的角盈}} \tag{2.1}$$

就已经提出并发表了相同的概念，但是从外在的角度。 高斯不知道他已领先发表。

在这个阶段，这个极限的存在对你来说并不明显，与三角形的形状和它收缩的确切方式无关；稍后将证明这一点。随着剧情的展开，我们将发现其他几种解释高斯曲率并在具体案例中计算其值的方法[4]。



[2.3] 高斯曲率 $\mathcal{K}$ 是每单位面积的局部角盈：如果曲面看起来像马鞍，它的符号为负，如果它像一座小山，它的符号为正，如果它像一张卷曲的纸，它就会为零。

（2.1）中的定义超出了三角形。如果我们用一个小的 n-边形替换 Δ_p，那么（见例 10），

$$\mathcal{E}(\text{n 边形}) \equiv [\text{内角和}] - (n-2)\pi \tag{2.2}$$

并且（2.1）中将曲率解释为每单位面积的角盈适用而没有变化。

检查[2.2]中的充气游泳圈应该清楚地表明，在外半部的每个点 p 处 $\mathcal{K}(p) > 0$，其中p的直接邻域类似于一座小山，而内半部的每个点 q 处 $\mathcal{K}(q) < 0$，其中 q 的直接邻域类似于马鞍。图[2.3]总结了这一点。

2.2 圆的周长和面积

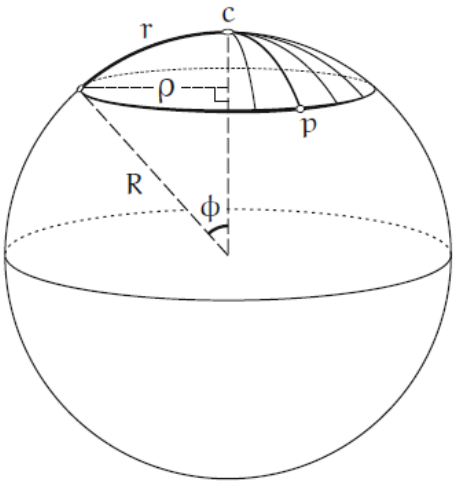
但为什么 $\mathcal{K}(p)$ 如此重要？是的，显然它在一定程度上控制了小三角形，但几何学不仅仅是三角形！答案是，虽然我们可能选择用小三角形来定义 $\mathcal{K}(p)$（暂时），但我们会逐渐发现曲率对曲面内几何的各个方面都有很强的控制力。我们现在只举两个例子。

在[1.9]中，我们指出了如何通过取固定长度 r 的测地线段 cp 的末端 p 并将其完全围绕 c 摆动来定义以 c 为中心的“半径为 r 的圆”。让我们计算在半径为 R 的球面上构造的这样一个圆的周长 $C(r)$。

参考[2.4]，我们看到

$$\rho = R\sin\phi \text{ 和 } \phi = \frac{r}{R} \quad \Longrightarrow \quad C(r) = 2\pi R\sin(r/R) \tag{2.3}$$

4 要使一个数学概念真正成为基本的，它必须位于不同数学分支的交叉点。因此，可以预期这些分支中的每一个都将提供一种看似不同但同样自然的方式来看待一个相同的概念。



[2.4] 半径为 R 的球面上半径为 r 的圆的周长为 $C(r)$，由 $C(r) = 2\pi R\sin(r/R)$ 给出。

正如曲率控制三角形内角和与欧几里得预测 π 的偏差一样，它也控制 $C(r)$ 与欧几里得预测 $2\pi r$ 的偏差。要看到这一点，请回忆正弦的幂级数：

$$\sin\phi = \phi - \frac{1}{3!}\phi^3 + \frac{1}{5!}\phi^5 - \cdots$$

因此，当 ϕ 趋于零时，

$$\phi - \sin\phi \asymp \frac{1}{6}\phi^3$$

（我们提醒读者，这里 $\asymp$ 表示牛顿的终极相等概念，正如在前言中介绍的那样。） 从（2.3）可以看出，当 r 缩小到零时，

$$2\pi r - C(r) = 2\pi R\left[\left(\frac{r}{R}\right) - \sin\left(\frac{r}{R}\right)\right] \asymp \frac{\pi r^3}{3R^2}$$

换句话说，$\mathbb{S}^2$ 的居民现在可以通过检查一个小圆的周长来确定他们世界的曲率，就像他们以前通过检查一个小三角形的角度一样容易：

$$\boxed{\mathcal{K} \asymp \frac{3}{\pi}\left[\frac{2\pi r - C(r)}{r^3}\right]} \tag{2.4}$$

值得注意的是，正如我们稍后将展示的那样，在第四幕中，这个公式继续正确地测量一般曲面上的高斯曲率！（请注意，分母中 r 的幂是可以预料的：我们知道 $\mathcal{K}$ 的单位是 $1/(\text{长度})^2$，而周长是长度，所以我们需要分母的单位是 $(\text{长度})^3$。）

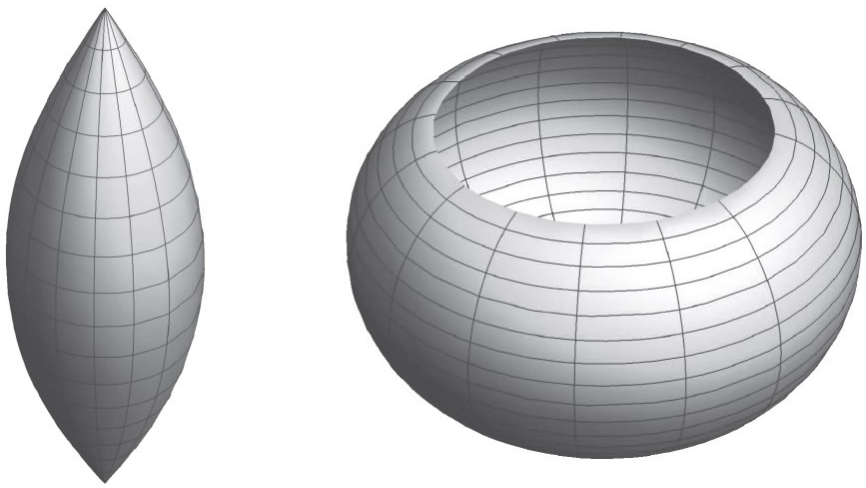
继续这个例子，让我们来检查由这个圆包围的极冠的面积$\mathcal{A}(\mathrm{r})$。同样，曲率决定了该面积如何偏离 $\pi\mathrm{r}^2$ 的欧几里得预测。在极冠公式的帮助下（见例 10，第 85 页），不难验证[练习]事实上，

$$\boxed{\mathcal{K} \asymp \frac{12}{\pi}\left[\frac{\pi\mathrm{r}^2 - \mathcal{A}(\mathrm{r})}{\mathrm{r}^4}\right]} \tag{2.5}$$

这个公式再次证明是通用的!（同样，与上述相同的推理解释了分母中的四次方。）

虽然我们还不能证明（2.4）和（2.5）的普遍性，但我们至少可以看到它们确实在可变曲面的每一点上产生了正确的符号，如[1.9]所示。因为如果这样一个曲面上的点的紧邻是正弯曲的，那么它在该点附近是山形的（正如[1.9]中包含Δ_1 的区域中的任何地方）。因此，以那里为中心的小圆的周长和面积确实会受到曲率的挤压，并且小于在平坦欧几里得平面中的情况。因此，上述两个公式都产生 $\mathcal{K} > 0$，因为它们应该如此。

另一方面，如果曲面在该点附近呈鞍形，则相反。回想一下我们在 [1.9] 中指出，在曲面的鞍形部分（Δ_2 所在的位置）绘制的圆将具有周长 $\mathrm{C}(\mathrm{r}) > 2\pi\mathrm{r}$。要把握这一点，站起来并伸出一只手臂，与你的身体成直角。如果你用脚后跟旋转，你的指尖会划出一个水平的圆圈。现在重复这个旋转，但这次在你转身时上下摆动你的手臂；显然，指尖比以前走得更远了。当我们在马鞍形曲面上画一个圆时，就会发生这种上下波动，因此前面的两个公式都产生 $\mathcal{K} < 0$，本该如此。



[2.5] 存在具有恒定正曲率的非球面旋转曲面，但它们必然具有尖刺或边缘。

我们说过曲率对几何体有“铁腕控制”，但这种控制到底有多绝对？例如，如果我们知道一块曲面具有恒定的正曲率$\mathcal{K} = (1/R^2)$，那么它是否真的是半径为 R 球面的一部分吗？好吧，拿一个乒乓球把它切成两半——现在稍微弯曲其中一个半球。显然我们得到了一个新的非球形曲面片，但由于我们没有拉伸曲面内的距离，测地线和角度没有改变，因此根据定义（2.3），曲率 $\mathcal{K}$ 没有改变。因此，我们当然可以获得至少恒定曲率的曲面片，这些曲面片外在几何不是球形，尽管它们都具有相同的内蕴几何。

图[2.5]表明，即使我们只关注旋转曲面，球面也不是唯一具有恒定正曲率的曲面。事实上，有一整套这样的曲面，球面代表两种图示类型之间的边界情况（见例 22）。尽管它们看起来不像球面，但生活在这些曲面上的聪明蚂蚁永远不会知道她不是生活在球面上。嗯，这几乎是真的：最终她可能会发现曲面不光滑的尖锐褶皱或尖刺，否则她可能会碰到曲面的边缘。1899 年，海因里希·利布曼（Heinrich Liebmann）证明[5]，如果一个具有恒定正曲率的曲面没有这些缺陷，那么它只能是一个球面。

忽略这些曲面的外在差异，两个曲面能否具有恒定的正曲率$\mathcal{K} = (1/\mathrm{R}^2)$ 却具有真正不同的内蕴几何形状？更明确地说，如果我们的智能蚂蚁突然从一个曲面转移到另一个曲面，她是否可以设计一个实验来发现她的世界发生了变化？ 1839 年，明丁（高斯的少数学生之一）发现了答案："不能！" 换言之，明丁发现[6]，如果两个曲面具有恒定的正曲率 $\mathcal{K} = (1/R^2)$，则它们的内蕴几何形状与半径为 R 的球面的几何形状局部相同。

我们已经讨论了[2.2]中游泳圈内半部具有负曲率这一事实，但它不是恒定的负曲率。事实上，如果 C 是环面和地面之间的接触圆，将内半部和外半部分开，那么很明显，当 q 接近 C 时，负曲率 $\mathcal{K}(\mathrm{q})$ 必然消失到零，如果 q 越过 C 到外半部，则切换为正曲率。（第二幕：Ex. 23 对此做了研究。）

[2.6] 基半径 R 的伪球具有恒定的负曲率 $\mathcal{K} = -(1/R^2)$。

事实上，确实存在具有恒定负曲率的曲面。尤金尼奥·贝尔特拉米（我们很快就会见到他）称所有这些曲面为伪球形的（pseudospherical），以纪念最简单的示例伪球面[7]（pseudosphere），如 [2.6] 所示。（这个曲面的精确构造将在第二幕中解释，它是由一条叫做曳物线（tractrix）的曲线产生的旋转曲面，该曲线由牛顿于 1676 年首次研究。）如果 R

5 证明必须等到第 38.11 节。

6 证明必须等到第四幕（练习 7，第 336 页）。

7 惠更斯于 1693 年首次研究了这个曲面，也称为 tractroid。 参见 Stillwell (2010, p. 345)。

是伪球圆底的半径，那么(我们稍后将证明）整个曲面上曲率取恒定负值 $\mathcal{K} = -(1/\mathrm{R}^2)$。

这个曲面的名称早已确定，无法更改（或许是种遗憾）。如你所见，它肯定不像球面那样封闭。事实上，我们将在本书后面证明不可能存在封闭的伪球面。此外，我们看到虽然伪球无限地向上延伸，但它的底部有一个圆形边缘。事实证明，在保持其恒定负曲率的同时，将伪球延伸到该边缘之外是不可能的。这绝非偶然：大卫•希尔伯特在 1901 年证明，如果一个具有恒定负曲率的曲面要嵌入到普通的欧几里得 3 维空间中，那么它必然有一条边，它不能超出该边。

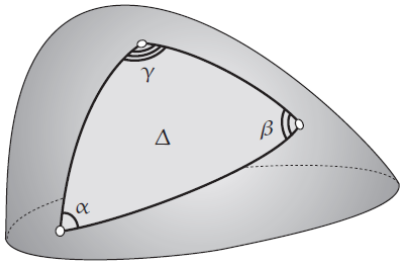
明丁的结果也适用于此：如果两个曲面片具有恒定的负曲率 $\mathcal{K} = -(1/\mathrm{R}^2)$，则它们的内蕴几何形状与半径为 R 的伪球面相同。

总而言之，如果一个曲面具有恒定的曲率 $\mathcal{K}$（正或负），则此单个数字完全决定了曲面的内蕴几何。

但是曲率变化的更典型的曲面呢？虽然曲率的控制仍然强大，但它不再是绝对的：两个曲面可能具有不同的内蕴几何，同时在对应点处仍具有相同的曲率。（具体例子见例 19。）

2.3 局部高斯-博内定理

回想一下哈里奥特 1603 年关于球面的结果（1.8）：*三角形的角盈等于曲率乘以面积*，我们可以将其视为三角形内的曲率总量。



[2.7] 一般曲面上的一般测地线三角形。

正如高斯在 1827 年的《关于曲面的一般研究》中最初表述的那样，*局部*[8]*高斯–博内定理*（Local Gauss–Bonnet Theorem）将该结果惊人地推广到可变曲率的一般曲面上的测地线[9]三角形 Δ，如 [2.7] 所示。它说这种三角形的角盈刚好是它内部的*总曲率*：

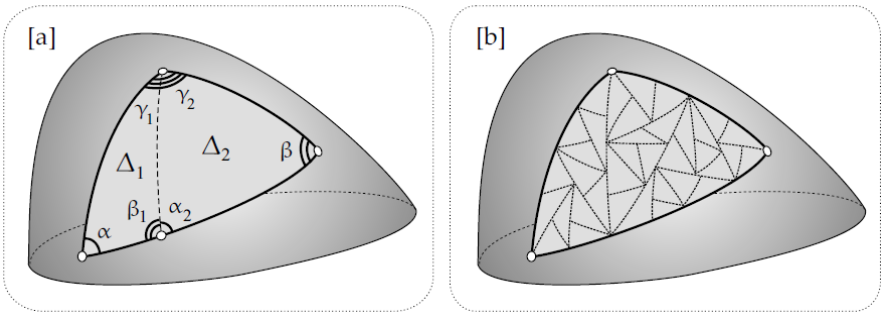
$$\mathcal{E}(\Delta) = \alpha + \beta + \gamma - \pi = \iint_{\Delta} \mathcal{K}\, d\mathcal{A}$$

8 这里的“局部”一词用于整个封闭曲面的“全局”版本区分开来。

9 1865 年，博内将该公式推广到非测地三角形，该定理因此得名。 直到第四幕（Ex. 6, p. 336）结束时才能证明该定理的最一般版本。

(2.6)

在球面的情况下，$\mathcal{K} = 1/R^2$，因此（2.6）得出哈里奥特公式（1.3）作为它的特例。



[2.8][a]角盈是可加的：$\mathcal{E}(\Delta) = \mathcal{E}(\Delta_1) + \mathcal{E}(\Delta_2)$。[b]如果我们继续细分三角形：$\mathcal{E}(\Delta) = \sum \mathcal{E}(\Delta_i)$，这种情况仍然成立。

要了解为什么会这样，首先回顾一下我们最初对曲率的定义（2.1）。当三角形 Δ_p 向 $\mathcal{S}$ 上的点 p 收缩时，

$$\mathcal{E}(\Delta_p) \asymp \mathcal{K}(p)\mathcal{A}(\Delta_p) \tag{2.7}$$

关键事实：角盈是可加的（additive）。

在[2.8a]中，从 Δ 的一个顶点到对边上的任意点画出了一个测地线段（虚线），从而将 Δ 分成两个测地线子三角形，Δ_1和 Δ_2。 观察到 $\beta_1 + \alpha_2 = \pi$，我们发现

$$\mathcal{E}(\Delta_1) + \mathcal{E}(\Delta_2) = [\alpha + \beta_1 + \gamma_1 - \pi] + [\alpha_2 + \beta + \gamma_2 - \pi] = \alpha + \beta + \gamma_1 + \gamma_2 - \pi$$

因此

$$\boxed{\mathcal{E}(\Delta) = \mathcal{E}(\Delta_1) + \mathcal{E}(\Delta_2)}$$

然后这些子三角形可以依次细分，依此类推，得出 [2.8b]，并且可加性质确保 $\mathcal{E}(\Delta) = \sum \mathcal{E}(\Delta_i)$。随着细分变得越来越细，曲率在每个 Δ_i 内变化越来越小，接近常数值 $\mathcal{K}_i$，并且在这个极限（2.7）中得到 $\mathcal{E}(\Delta) \asymp \sum \mathcal{K}_i \mathcal{A}_i$，因此我们得到局部高斯-博内定理（2.6）。

第三章 前言和第一幕练习

前言：牛顿终极等式 (≍)

1. （这是一个“终极相等”——见前言——如何变成相等的模型。）画一个边为 x 的立方体，因此体积为 $\mathcal{V} = x^3$。 在同一张图中，保持一个顶点固定，画一个稍大的立方体，边为$x + \delta x$。 如果 $\delta\mathcal{V}$ 是体积的增加，使用你的草图来推导当 δx 趋于零时，

$$\delta\mathcal{V} \asymp 3x^2\delta x \quad \Longrightarrow \quad \frac{d\mathcal{V}}{dx} \asymp \frac{\delta\mathcal{V}}{\delta x} \asymp 3x^2 \quad \Longrightarrow \quad \frac{d\mathcal{V}}{dx} \asymp 3x^2$$

但最终等式中的量与 δx 无关，因此它们相等：

$$(x^3)' = \frac{d\mathcal{V}}{dx} = 3x^2$$

2. （这个例子取自 Needham 1993）设 $c = \cos\theta$ 和 $s = \sin\theta$。 在$\mathbb{R}^2$的第一象限，在单位圆上画一点 $p = (c, s)$。 现在让 p 旋转一个小（最终趋于零）角度 $\delta\theta$。 以 p 为顶点，画出边为 δc 和 δs 的小三角形。 通过模仿序言中的牛顿几何论证，立即同时推断出

$$\frac{ds}{d\theta} = c \quad 和 \quad \frac{dc}{d\theta} = -s$$

3. （此示例取自 Needham 1993）令 L 为通过 $\mathbb{R}^2$ 第一象限中点 (a, b) 的一般直线，设 $\mathcal{A}$ 是由 x 轴、y 轴和 L 围成三角形的面积。
   (i) 用普通微积分方法求使 $\mathcal{A}$ 最小的 L，并证明 $\mathcal{A}_{\min} = 2ab$。
   (ii) 使用牛顿推理立即解决问题，无需计算！（提示：设 $\delta\mathcal{A}$ 是由 L 的小（最终趋于零）旋转 $\delta\theta$ 引起的面积变化。通过以两个三角形的形式画出 $\delta\mathcal{A}$，并观察每个三角形最终等于圆的一个扇区，基于$\delta\theta$ 写出 $\delta\mathcal{A}$ 的最终等式。现在设 $\delta\mathcal{A} = 0$。）

4. 以下问题取自 Arnol’d (1990, p. 28)，其中也包含答案。 计算极限

$$\lim_{x\to 0} \frac{\sin\tan x - \tan\sin x}{\sin^{-1}\tan^{-1}x - \tan^{-1}\sin^{-1}x}$$

   (i) 使用你能想到的任何传统方法。 （如果我们不祝愿你在本书中获得英国最好的运气，那就是我们的失职！ Arnol'd 指出，唯一能够快速解决这个问题的数学家是菲尔兹奖得主 Gerd Faltings。）
   (ii) 使用牛顿的几何推理方法。

欧式与非欧几何

5. 目前尚不清楚巴比伦人是怎么生成图 [1.2] 中所示的毕达哥拉斯三元组的，但 1500 年后，欧几里得（约公元前 300 年）是第一个陈述并证明此类三元组最通用公式的人。半个千年之后，丢番图（大约公元 250 年）率先使用几何构造系统地生成单位圆上的有理点，即具有有理坐标的点。然后可以使用这些有理点来生成毕达哥拉斯三元组，如下所示。

   (i) 令 L 为通过单位圆 C 上点 $(-1,0)$ 的直线 $y=m(x+1)$，令 (X,Y) 为 L 和 C 的第二个交点。 证明

$$X=\frac{1-m^2}{1+m^2},\qquad Y=\frac{2m}{1+m^2}$$

   (ii) 推导如果斜率 $m=(q/p)$ 是有理数，那么 X 和 Y 也是：

$$X=\frac{p^2-q^2}{p^2+q^2},\qquad Y=\frac{2pq}{p^2+q^2}$$

   (iii) 推导如果 (a,b,h) 是任意毕达哥拉斯三元组，则

$$\frac{a}{h}=\frac{p^2-q^2}{p^2+q^2},\qquad \frac{b}{h}=\frac{2pq}{p^2+q^2}$$

   其中 p 和 q 是整数。

   (iv) 推断出最一般的毕达哥拉斯三元组由以下公式给出，首先由欧几里德提出：

$$a=(p^2-q^2)r,\qquad b=2pqr,\qquad h=(p^2+q^2)r$$

   其中p,q,r是任意整数。

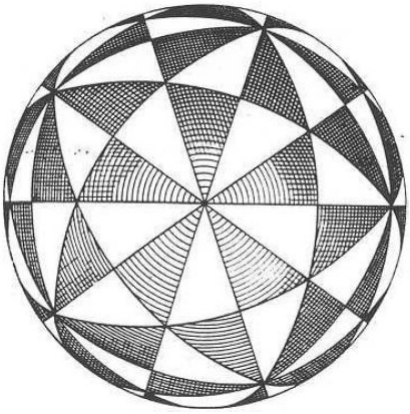
6. 使用（1.3）推导，随着球面上三角形的大小缩小到零，最终在球面的居民看来它是欧几里得的，即内角和等于 π。

7. 设 p 和 q 是球面上不同的非对径点，并考虑通过它们的唯一大圆 $\mathcal{C}$。$\mathcal{C}$ 被 p 和 q 分割成两段弧，设 m_1 和 m_2 是各自的中点。证明与 p 和 q 等距的点的轨迹是通过 m_1 和 m_2 正交切割 $\mathcal{C}$ 的大圆：这是 pq 的广义“垂直平分线”。（提示：虽然它在数学上无关紧要，但想象球面已经旋转使得 p 和 q 位于赤道上，这在心理上会有所帮助。）

8. 证明如果单位球面 $\mathbb{S}^2$上的三角形各边均小于 π，则该三角形包含在一个半球内。

（提示：虽然它在数学上无关紧要，但想象球面已经旋转使得一个顶点位于北极可能在心理上有所帮助。）

9. 欧几里得平面的特点之一是它具有正多边形镶嵌：它可以被正多边形平铺（完全填充，没有任何间隙）。 平面上正好有三种这样的正多边形镶嵌：等边三角形、正方形和正六边形。 球面也容许正多边形镶嵌。 想象一个内接在半径为 R 的球面中的线框二十面体。现在想象一个位于球心的点光源。球面上的阴影如 [3.1] 所示。在这里，二十面体的每个三角形面的像通过将其中心连接到其三条边的中点，被进一步划分为六个全等三角形。
   (i) 解释为什么二十面体边的阴影是球面的大圆弧，从而得到真正的球面三角形。
   (ii) 验证如果我们改为内接十二面体，将其五边形中心投影到球面上，然后将它们的中心连接到各边的中点，我们将获得相同的镶嵌。
   (iii) 球面（面积为 $4\pi R^2$）被分成多少个全等三角形？每个三角形的面积 $\mathcal{A}$ 是多少？
   (iv) 通过观察每个顶点有多少相似的角聚集在一起，推导三角形的角是$(\pi/2)$、$(\pi/3)$和$(\pi/5)$。 由此计算每个三角形的角盈 $\mathcal{E}$。
   (v) 验证前两个答案是否符合哈里奥特定理（1.3）。



[3.1] 球面的二十面体镶嵌。 在这里，二十面体的每个三角形面的像通过将其中心连接到其边缘的中点，被进一步划分为六个全等三角形。 （这幅可爱的图画取自 Fricke (1926)。）

10. 完成下述题目:
   (i) 证明在欧几里得几何中，四边形的内角和是 2π。
   (ii) 如果 Q 是半径为 R 球面上的测地四边形，则其角盈为

$$\mathcal{E}(Q) = \left(Q\text{ 的内角和}\right) - 2\pi$$

通过画出 Q 的对角线，从而将其分成两个测地三角形，推导出（1.3）可以推广为

$$\mathcal{E}(Q) = \frac{1}{R^2}\mathcal{A}(Q)$$

(iii) 证明 （2.2），从而将 (ii) 推广到球面上的测地 n 边形。

11. 制作窄条胶带，最好是彩色美纹纸胶带。 然后使用（1.7）在图 [11.7] 所示近似花瓶曲面上进行以下实验。如果你没有这样的花瓶，建议你借一个——这个实验太有趣了，不容错过！
    (i) 从最大半径 ρ_{max} 的水平圆上的一点开始，沿花瓶垂直向上发射测地线，形成旋转曲面的一条子午线（meridian）。
    (ii) 从同一点开始，在与子午线形成逐次增大角度 ψ 的方向上发射更多测地线。
    (iii) 请注意，在到达某个临界角 ψ_c 之后，测地线最初会沿着花瓶向上走，然后折返沿着花瓶向下走！
    (iv) 尽你所能，尝试找到临界测地线——将回头的和不回头的测地线分开。 通过在发射点将量角器压在曲面上，测量其角度 ψ_c。 注意：要找到临界测地线，你需要构建超长的测地线段，你可以通过扩展现有段，将新段与旧段重叠来完成，如 [1.9] 所示。
    (v) 设 ρ_{max} 为花瓶的最大半径（在发射点），设 ρ_{min} 为花瓶喉部的最小半径。 通过测量直径并除以 2，尽可能准确地找到这两个半径。 现在验证，在实验误差范围内，

$$\psi_c = \sin^{-1}[\rho_{min}/\rho_{max}]\,!$$

（这是克莱罗定理的物理实例化，我们将在稍后的 11.7.4 节证明。）

高斯曲率

12. **零曲率**。 制作窄条胶带，最好是彩色美纹纸胶带。 然后使用（1.7）进行以下实验。
    (i) 将一张纸卷成一个圆锥形，并用胶带将其粘住以免展开。 开始创建一个长的测地线段，起始于边缘的一个点，在将胶带粘到曲面之前尝试猜测它的形状。 从边缘上的同一点开始，重复此构造，向不同方向发射新的测地线。
    (ii) 接下来，构造一个测地三角形，并使用量角器验证 $\mathcal{E} = 0$（这适用于该曲面上的所有测地三角形，证明 $\mathcal{K} = 0$）。
    (iii) 最后，切开圆锥面（沿着母线）并再次将纸压平，观察胶带构造的欧几里得形式。

13. **正曲率**。 将一根牙签插入任何近似球形的水果（半径为 R）的北极，例如甜瓜。将一根绳子或牙线的一端系在该牙签上，然后将绳子在曲面上拉紧到距离赤道大约一半的位置。 在这个长度为 r 的测地线段的末端拿着一支笔，拖动笔（保持绳子拉紧）以创建一个长度为 C(r) 的纬度圆。将牙签均匀地插入这个圆圈中。现在将一根细绳或牙线绕在大约圆形的牙签环的外侧，轻轻拉紧，使其顺着水果表面的纬度圈。 用笔在圆圈的起点和终点仔细标记绳子。 打开它并沿着尺子拉伸以测量两个标记之间的长度 C(r)。

(i) 将（2.4）视为近似等式（而不是精确的最终等式），估计 $\mathcal{K}$。根据这些内蕴测量值，估计水果的外在半径 R。 将你的答案与 R 的直接测量值进行比较。

(ii) 从 (i) 继续，假设你对 r 和 C(r) 的测量是完美的。 使用 sin(r/R) 的（递减的和交替的）泰勒展开的第三项来表明以第 (i) 部分的方式应用（2.4）导致的 $\mathcal{K}$ 中百分比误差的上限——即， 不考虑最终相等所暗示的极限——由下式给出

$$\left|\frac{\Delta\mathcal{K}}{\mathcal{K}}\right| < 5\left[\frac{\mathrm{r}}{\mathrm{R}}\right]^2 \%$$

推导即使对于像你构建的那样大的圆，误差也不会大于约 3%！

(iii) 使用 (ii) 的结果推导出 R 中百分比误差上限的公式。

14. **负曲率**。制作窄条胶带，最好是彩色美纹纸胶带。 然后使用（1.7）进行以下实验。

(i) 按照第 53 页 [5.3] 附带的说明，用半径为 R 的圆盘构建你自己的伪球面——圆锥越多越好； 越大越好！

(ii) 从半径为 R 的圆底上的一点开始，向各个方向发射测地线，并在将胶带粘在曲面上之前尝试预测它们的路线。当一条胶带用完时，如图 [1.9] 所示，将新条与旧条重叠来继续延展测地线。唯一的例外是子午线测地线，它笔直地在曲面向上（沿着旋转面曳物线母线）。请注意，每个测地线最初都沿伪球面上行，然后转向折返并沿伪球面下行，最终返回基圆。

(iii) 构造一个直角测地线三角形 Δ，测量其角度，从而估计其角盈 $\mathcal{E}(\Delta)$。 估计（尽你所能）它的面积$\mathcal{A}(\Delta)$。 由此使用下列式子估计伪球面的（恒定）曲率 $\mathcal{K}$，

$$\mathcal{K} = \frac{\mathcal{E}(\Delta)}{\mathcal{A}(\Delta)}$$

(iv) 三角形越大，$\mathcal{E}(\Delta)$ 的值越大（即越负），使其测量更容易和更准确。 但代价是准确估计面积 $\mathcal{A}(\Delta)$ 变得更加困难。 要克服这个困难，请执行以下操作。 制作用于生成测地线的相同类型的窄胶带，使其都具有相同的（准确测量的）宽度 W，可能是 1/4 英寸。 现在用这些条带填充你的 Δ，当它们碰到边缘时将它们切断。 取下条带并将它们首尾相连地放在平坦的表面上，并测量总长度 L。然后 $\mathcal{A}(\Delta) \asymp \mathrm{LW}$。

(v) 对更多的三角形重复 (iii)，但不再将它们限制为直角，因为 (iv) 现在可以轻松测量任何形状三角形的 $\mathcal{A}(\Delta)$。 验证（在实验误差范围内）所有三角形得到相同的 $\mathcal{K}$ 值。

(vi) 假如

$$\mathcal{K} = -\frac{1}{\mathrm{R}^2}$$

估算 R，并将其与用于构建伪球的圆盘的实际半径进行比较。

第四章 映射曲面：度量

4.1 引言

球面的完美外在对称性的优势在于，它的内蕴几何形状同样是均匀的。相比之下，在图[2.5]中，是否可以在曲面上自由滑动和旋转，同时保持与曲面完全接触，这当然不清楚。但事实必然如此，因为上述讨论表明，曲面在空间中的实际形状不过是种干扰：这些曲面与球面无法内蕴地区分（至少在局部）。

从这个角度来看，最好有一个更抽象的模型来捕捉具有相同内蕴几何形状的所有可能曲面的本质。我们所说的“本质”是指任何两点之间距离的知识，因为这一点决定了内蕴的几何形状。事实上，这就是高斯对微分几何的根本见解，有一个规则来计算相邻点之间的无穷小距离就足够了。这个规则称为度量（metric）[1]。给定这个度量，就可以将曲线划分成无穷小段并求和（即积分），从而确定任何曲线的长度。因此，还可以将几何体的测地线确定为从一个点到另一个点的最短路径，同样还可以确定角度。

这引出了以下策略用于捕获任何曲面 $\mathcal{S}$（不一定是恒定曲率的）的本质。为了避免曲面在空间中形状的干扰，我们在一张平坦的纸上绘制了 $\mathcal{S}$ 的（制图）地图（map）。也就是说，在 $\mathcal{S}$ 上的点 $\hat{z}$ 和平面上的点 z 之间建立了一一对应的关系。当然，就球形地球和球形夜空而言，水手和天文学家数千年来一直在设计这样的地理和天体地图。

关于术语的重要说明：在大多数数学中，“地图”（map）是“映射”（mapping）或“函数”（function）的同义词。然而，当用作名词时，将始终使用当前制图意义上的“地图”。当指的是“映射”或“函数”时，应该说“映射”（mapping）或“函数”（function）！话虽如此，当涉及到动词时，我们将保留一些传统含义：例如，坐标函数将曲面上的这条曲线映射（map）到地图（map）上的那条曲线。

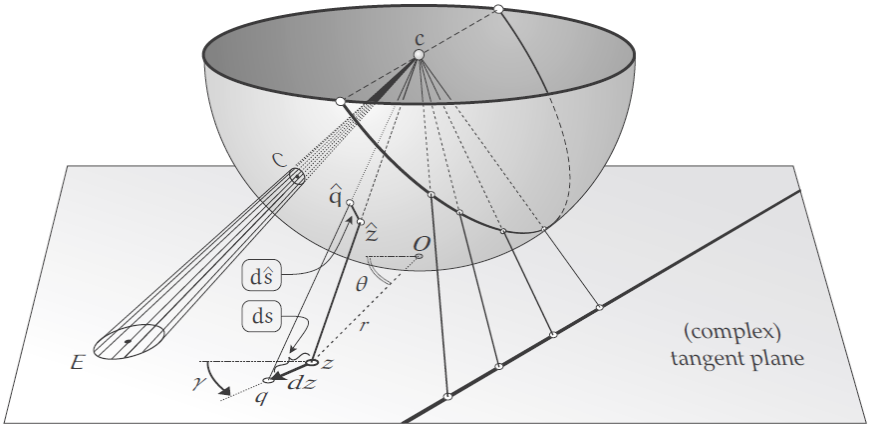
一般来说，如果不引入某种畸变，就无法创建这样的曲面地图：如果你剥掉橘子的一部分，然后尝试将橘子的那块果皮平压在桌子上，它就会撕裂。欧拉第一个证明了（1775 年）完美的地球制图在数学上是不可能的，其中地表上所有的“直线”（测地线）在地图上都变成了直线，并且所有陆地距离都用地图上的距离成比例表示。

前面的讨论解释了完美地图的不可能性。地球表面上的一个三角形的 $\mathcal{E} \neq 0$，如果这个三角形可以在不改变距离的情况下被压平，那么它在地图中的像将是一个 $\mathcal{E} = 0$ 的欧几里德三角形：矛盾。

我们最终会发现，非欧几何和复数之间存在着深刻而神秘的联系。因此，让我们从一开始就假设我们在上面绘制地图的那张平面纸实际上是复平面 $\mathbb{C}$。

---

1 另一个常见的名称，特别是在较旧的作品中，是第一基本形式。



图[4.1] 球面的中心射影将测地线映射到直线，将圆映射到椭圆。该度量告诉我们与地图距离 ds 和曲面距离 $d\hat{s}$ 相关的局部比例因子。

现在考虑在 $\mathcal{S}$ 上分隔两个相邻两点 $\hat{z}$ 和 $\hat{q}$ 的距离 $\delta\hat{s}$。然后点 $\hat{z}$ 和 $\hat{q}$ 将由复数 $z = re^{i\theta}$ 和 $q = z + \delta z$ 在这个复平面上表示，由（欧几里得）距离 $\delta s = |\delta z|$ 分隔。有关这样一张地图的具体示例（稍后将解释），请看[4.1]。一旦我们有了从地图中的表观间距 δs 计算 $\mathcal{S}$ 上实际间隔 $\delta\hat{s}$ 的规则，那么（原则上）我们就知道了关于 $\mathcal{S}$ 的内蕴几何的一切。

根据 δz 给出 $\delta\hat{s}$ 的规则称为度量（metric）。一般来说，$\delta\hat{s}$ 取决于 δz 的方向以及长度 δs：写成 $\delta z = e^{i\gamma}\delta s$，所以 $\delta\hat{s} \asymp \Lambda(z,\gamma)\delta s$。在此提醒读者，$\asymp$ 表示牛顿的终极相等（ultimate equality）概念，这在前言中介绍过。这种关系更传统地用无穷小符号表示为，

$$\boxed{d\hat{s} \asymp \Lambda(z,\gamma)ds} \tag{4.1}$$

根据这个公式，$\Lambda(z,\gamma)$ 是我们必须扩大地图中的表观间距 ds（位于 z 方向为 γ）以获得曲面 $\mathcal{S}$ 上的真实间距 $d\hat{s}$ 的量。

4.2 球面的射影地图

图[4.1]说明了（4.1）对于绘制南半球地图的一种特殊方法的含义，称为中心投影（central projection）。想象半球是一个玻璃碗，位于原点 0 的复平面上，并想象一个光源位于球体的中心，c。然后穿过半球上 $\hat{z}$ 的光线继续命中 $\mathbb{C}$ 中的点 z。由此产生的平面地图就是所谓南半球的射影地图（projective map），或射影模型（projective model）。

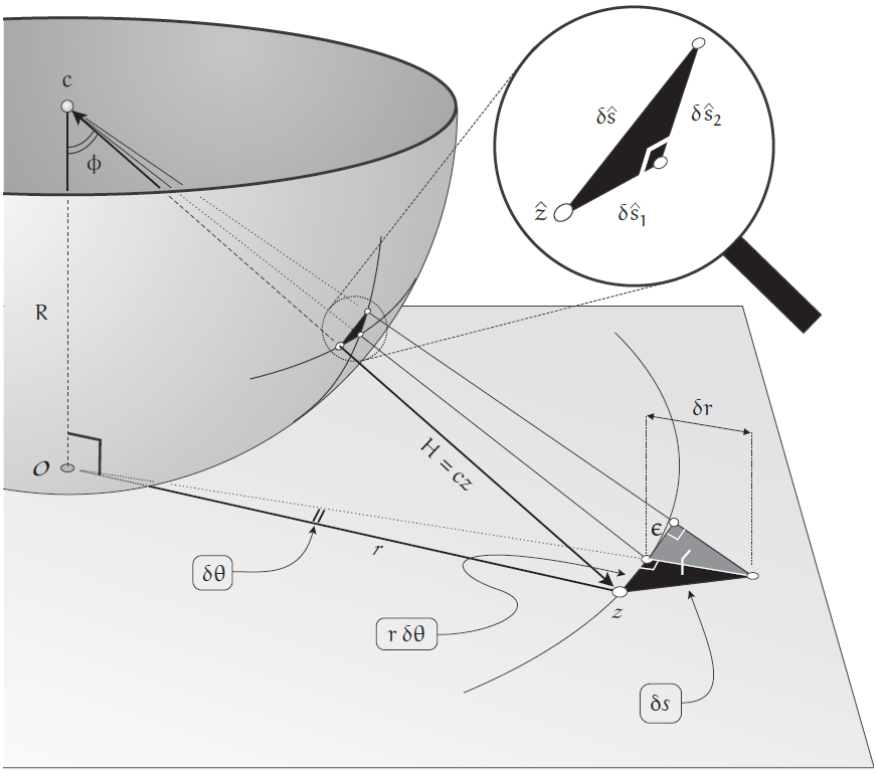
如果我们在半球上画一个圆 C，那么穿过它的光线会在空间中形成一个圆锥面，因此它们投影到 $\mathbb{C}$ 就形成了一个完美的椭圆 E。这是这种绘制地图的特殊方法的一个非常特殊和不寻常的特性。然而，事实证明[2]，如果 C 是一般曲面 $\mathcal{S}$ 上的一个内蕴定义的圆（例如在图[1.9]中），那么当它的半径收缩时，它在一般映射中的像 E 最终也将是一个椭圆。回到中心射影的完美椭圆，很明显，E 的主轴将是径向的，从碗与平面的接触点直接指向外。换句话说，如果你想象 dz 围绕 z 旋转，那么 $\Lambda(z,\gamma)$ 分别在 $\gamma = \theta$ 和 $\gamma = \theta + (\pi/2)$ 处达到其

2 因为如果映射是可微的，那么它的局部效应就是线性变换。

最小值和最大值。

选择哪种绘制曲面地图的方式取决于我们希望准确或忠实地表达哪些特征。例如，[4.1]说明射影地图忠实地表示直线：地图中的直线表示球面上的大圆测地线。但是，为保持直线的概念所付出的代价是角度没有被忠实地表示：两条曲线在球面上相交的角度（通常）不等于它们在地图上相交的角度。

尽管如此，实际上球面上有两个正交曲线族，它们映射到平面中的正交曲线：它们是经度和纬度的圆。纬度圆（即半球的水平横截面）映射到平面中以原点为中心的圆，经度（半）圆（半球通过其中心的垂直横截面）映射到经过原点的一条直线。正如所声称的，这些圆和直线确实以直角相交。我们现在使用这个事实来推导出射影地图中极坐标 (r,θ) 的球面度量公式。通过计算[练习] 不难做到这一点，但我们将改为使用在前言中引入的牛顿形式的几何推理，我们将在整本书中使用它。



图[4.2]平面上的小旋转 $\delta\theta$ 使 z 移动 $r\delta\theta$，并使 $\hat{z}$ 绕水平纬度圆移动 $\delta\hat{s}_1$。如果 z 然后径向向外移动 δr，那么 $\hat{z}$ 沿着经度的垂直圆向北移动 $\delta\hat{s}_2$。

见[4.2]。平面上的小旋转 $\delta\theta$ 使 z 绕半径 r 的圆旋转距离 $r\delta\theta$，并使 $\hat{z}$ 绕水平纬度圆旋转 $\delta\hat{s}_1$。如果 z 然后径向向外移动 δr，那么 $\hat{z}$ 沿着经度的垂直圆向北移动 $\delta\hat{s}_2$。根据毕达哥拉斯定理，$\delta\hat{s}^2 \asymp \delta\hat{s}_1{}^2 + \delta\hat{s}_2{}^2$，我们现在将分别找到这些项。（回顾一下前言，$\asymp$是牛顿“终极相等”概念的符号。）

令 $H = cz$ 表示$\mathbb{R}^3$中从球心 c 到复数 z 的距离，如 [4.2] 所示。由于 $\hat{z}$ 与 c 的距离为 R，并且由于旋转导致 $\hat{z}$ 和 z 与球心的距离成比例移动：

$$\frac{d\widehat{s_1}}{r\,d\theta} \asymp \frac{\delta\widehat{s_1}}{r\,\delta\theta} = \frac{R}{H}$$

接下来，想象 H 是一条连到 c 的固定长度细绳。我们在垂直于复平面的平面内将它在 z 处的自由端向上摆动一段距离 ε，那么 $\hat{z}$ 的向北运动 $\delta\widehat{s_2}$ 将与之前比例相同。此外，我们看到边为 ε 和 δr 的小垂直灰色三角形最终与边为 R 和 H 的大三角形 0cz 相似。于是，

$$\frac{\delta\widehat{s_2}}{\varepsilon} \asymp \frac{R}{H} \quad 与 \quad \frac{\varepsilon}{\delta r} \asymp \frac{R}{H} \quad \Longrightarrow \quad \frac{d\widehat{s_2}}{dr} = \frac{R^2}{H^2}$$

最后，再次由毕达哥拉斯定理，$H^2 = R^2 + r^2$。因此，回到行业标准的、基于无穷小的符号，根据射影地图中的坐标描述球面上的真实距离的度量由下式给出

$$\boxed{d\hat{s}^2 = \frac{1}{1+(r/R)^2}\left[\frac{dr^2}{1+(r/R)^2} + r^2 d\theta^2\right]} \tag{4.2}$$

在地理学中，纬度角 ϕ 通常在赤道处定义为零，但我们选择从极点测量它（如[4.2]和[2.4]所示）。回到[4.1]，我们现在可以通过查看小椭圆 E 的形状来量化地图诱导的畸变。随着 C 缩小，你现在应该能够确认[练习]

$$\left(\frac{E\ 长轴}{E\ 短轴}\right) \asymp \sec\phi \tag{4.3}$$

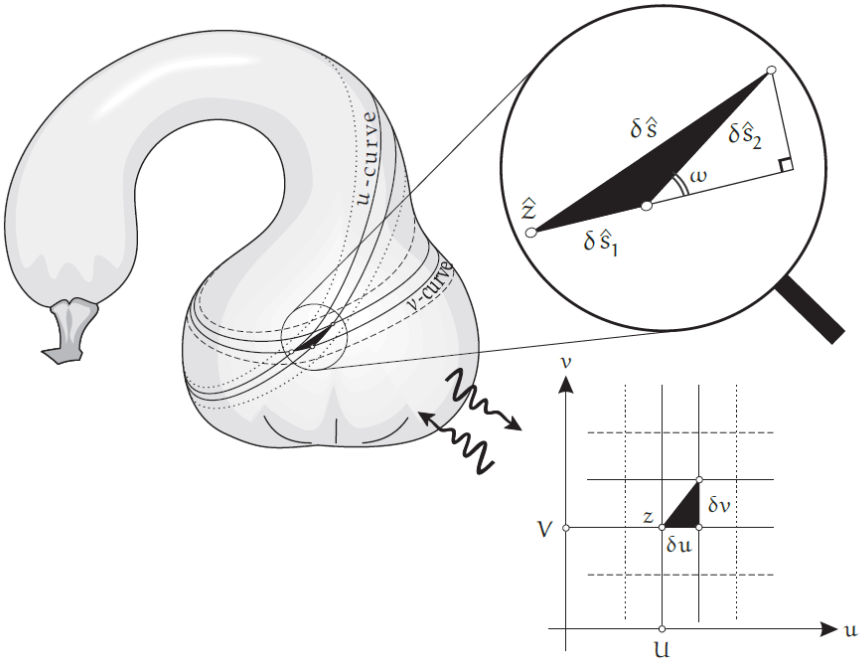
因此当 C 向北移向边缘 $(\phi = \frac{\pi}{2})$ 时，图形的畸变会无限增加。

4.3 一般曲面的度量

鉴于地图在导航中的重要性，几个世纪以来，数学家们探索了许多不同的绘制地球地图的方法，我们很快就会遇到一种特别重要的方法（其他地图绘制方法将在本幕结束时的练习中探讨）。现在我们只想指出，每个这样的地图都有一个与之相关的不同度量公式，尽管它们都描述了相同的内蕴几何。

例如，描述地球上一个地方位置的最常用方法是提供其经度 θ 和纬度 ϕ。假设使用这两个角度在平面上绘制一个非常直接的地图，θ沿着水平轴，ϕ沿着垂直轴。也就是说，如果一个特定的房子有经度 θ 和纬度 ϕ，那么在平面地图中，用笛卡尔坐标 (θ, ϕ) 的点表示它。在图[2.4]的帮助下，你现在应该很容易地发现[练习]，告诉我们与地图中相邻点相对应的球面上的真实距离的度量公式是

$$d\hat{s}^2 = R^2[\sin^2\phi\, d\theta^2 + d\phi^2] \tag{4.4}$$



图[4.3]在一般曲面上，绘制两个参数化曲线族："u-曲线"（u=const.）和"v-曲线"。然后由与那里相交的特定曲线对 (U,V) 指定一个点，从而在地图中产生一个点 $z=U+iV$。地图中的一个小水平移动 δu 最终会在曲面上上沿 v-曲线产生一个成比例移动 $\delta\hat{s}_1 \asymp A\delta u$。

这看起来肯定与以前的公式（4.2）大不相同，但我们知道它实际上描述了完全相同的内蕴几何。

即使简单地把曲面视为一个平面，也可能有无数种度量公式。例如，如果使用习惯的笛卡尔坐标，则 $d\hat{s}^2=dx^2+dy^2$，而如果使用极坐标，则$d\hat{s}^2=dr^2+r^2d\theta^2$。

现在让我们研究一般曲面的最一般度量公式的形式和含义。图[4.3]说明了如何绘制这样一个曲面 $\mathcal{S}$ 上面片的一般地图。对于这个面片中的每个点 $\hat{z}$，我们的第一个目标是分配一对坐标(u,v)，以便它可以表示为平面地图中的复数 $z=u+iv$。

开始时，我们很随意地简单地画出覆盖面片的曲线族，这样我们族中的一条（并且只有一条）曲线通过面片上的每个点 $\hat{z}$。 我们现在用唯一的数字 u 标记这些曲线中的每一条，并同意将这些曲线及其指定的 u 值称为 u-曲线。 标记可以非常随意地进行，除了我们将坚持编号平滑变化，即当我们在曲面上移动穿过 u-曲线时，u 值应该以确定的速率变化，即可微地——稍后再详细介绍。

为了完成坐标系，我们现在绘制了第二个曲线族，同样非常随意，除了它们应该穿过 u-曲线并且从不与它们重合。现在用 v 值标记这些新曲线，而且，v 值应该以确定的速率（以与以前相同的方式）可微分地变化；将这些曲线称为 v-曲线。因此，如图所示，点 $\hat{z}$ 可以由相交于此的唯一 u-曲线（例如$u=U$）和 v-曲线（例如$v=V$）标记。因此，在地图中，$\hat{z}$ 现在可以表示为$z=U+iV$。因此，如图所示，u-曲线在地图中由垂直线表示，而 v-曲线则映射到水平线。

现在我们至少在曲面的某个小块上有了坐标，任务是找到给出相邻点之间距离的度量公式。假设在地图中，我们沿着$\delta z = \delta u + i\delta v$远离 z 做一个小的移动。由于 u 值对 u-曲线的可微赋值，微小的变化 δu 产生了——我们现在得出可微（differentiable）的定义——曲面上（沿着 v-曲线）的小移动 $\delta\hat{s_1}$ 最终与 δu 成正比，我们将 A 定义为该点的比例常数：

$$\boxed{A \equiv \frac{\partial\hat{s_1}}{\partial u} \asymp \frac{\delta\hat{s_1}}{\delta u}}$$

这很重要，所以我们重申：A 是 (u, v) -地图水平方向上的局部比例因子，我们必须通过该因子在地图中拉伸一小段水平距离才能获得表面上的真实距离。

还有一种方法可以在不看地图的情况下将其可视化。如果在 [4.3] 中我们想象 u-曲线是以小的固定增量 $\varepsilon\,(u = U, u = U + \varepsilon, u = U + 2\varepsilon)$ 绘制的，那么 A 也可以可视化为与 u-曲线的拥挤或密度成反比。因为拥挤程度越大，曲面上给定运动 $\delta\hat{s_1}$ 产生的变化 δu 就越大。

以完全相同的方式，微小的变化 δv（同时保持 u 不变）在曲面上产生最终与其成正比的运动 $\delta\hat{s_2}$，使我们能够将 B 定义为地图垂直方向上的局部比例因子：

$$\boxed{B \equiv \frac{\partial\hat{s_2}}{\partial v} \asymp \frac{\delta\hat{s_2}}{\delta v}}$$

最后，如图所示，我们将 ω 定义为 u-曲线和 v-曲线之间的角度。显然这个角度不是一个常数：就像比例因子 A 和 B 一样，角度 ω 是位置的函数。

现在可以将毕达哥拉斯定理应用于[4.3]的放大镜中显示的直角三角形：

$$d\hat{s}^2 \asymp (\delta\hat{s_1} + \delta\hat{s_1}\cos\omega)^2 + (\delta\hat{s_1}\sin\omega)^2 \tag{4.5}$$
$$\asymp (A\delta u + B\delta v\cos\omega)^2 + (B\delta v\sin\omega)^2 \tag{4.6}$$

在简化并恢复到更标准的、基于无限小的符号之后，我们发现

一般曲面的一般度量公式为

$$d\hat{s}^2 = A^2\,du^2 + B^2\,dv^2 + 2F\,du\,dv \quad 其中 \quad F = AB\cos\omega \tag{4.7}$$

假设你最终会看其他书籍，我们应该立即发出符号警告。在他 1827 年的原作中，高斯

决定把度量写成[3]

$$d\hat{s}^2 = E\,du^2 + G\,dv^2 + 2F\,du\,dv$$

在随后的几个世纪里，几乎每一篇[4]关于微分几何的研究论文和教科书都奴性地延续了这种 E、F、G 符号。然而，我们已经看到，$\sqrt{E} = A$ 和 $\sqrt{G} = B$ 具有上面给出的简单几何解释，因此，在许多重要公式中表现出来的是它们（而不是 E 和 G）也就不足为奇了。其结果是，出现了不必要的平方根杂乱无章的文献。因此，我们将继续在整本书中使用符号 A 和 B（代替 $\sqrt{E}$ 和 $\sqrt{G}$），所以我们强调，当你看其他地方时，你必须用这个来翻译

$$\boxed{\text{符号词典：}\ E \equiv A^2,\quad G \equiv B^2,\quad F \equiv AB\cos\omega.} \tag{4.8}$$

接下来，通用度量公式可以简化如下。应该清楚的是，一旦我们绘制了 u-曲线族，我们就可以绘制正交轨迹族，然后选择这些曲线作为 v-曲线。由于该结构坚持 $\omega = (\pi/2)$，因此 $F = 0$，因此，

$$\boxed{\begin{array}{l}\text{局部地，我们总是可以在一般曲面上构造}\textbf{正交}\ (u,v)\\ \text{坐标系；然后度量采用以下形式，}\\[1em] \qquad\qquad d\hat{s}^2 = A^2\,du^2 + B^2\,dv^2\end{array}} \tag{4.9}$$

但是请注意，即使我们不坚持正交坐标，通常也不可能用单个(u,v)坐标系覆盖整个曲面。出现的问题是，两条 u-曲线（和/或 v-曲线）可能被迫相交，在这种情况下，必须为交叉点分配两个不同的 u 值。事实上，正如最终将在第 19 章中看到的那样，这些问题实际上在每个封闭的曲面上都是不可避免的，除了甜甜圈。

例如，在地球表面，假设选择 u-曲线是纬度圈（尽管不一定是 u=纬度）。那么，正交轨迹（即 v-曲线）必然是经度圈：在南北极相交的大圆。因此，必须为每个极点分配无限多个 v 值。

4.4 度量曲率公式

现在假设我们只得到了一个度量公式（4.9），没有任何关于它所描述的曲面 $\mathcal{S}$ 的直接几何知识，也没有坐标 u 和 v 本身的几何意义。这个曲面到底是什么样子的？就 $\mathcal{S}$ 的内蕴几何而言，这个公式告诉我们一切，但只是原则上。如何才能真正从公式中提取有用的信

---

3 事实上，高斯用 p 和 q 代替了 u 和 v，这后来成为标准符号。我们选择不听从历史上的 p 和 q。

4 只有少数例外，但它们的高血统可能为我们选择的符号提供了一些体面。 例如，Hopf (1956, p. 92) 偶尔会写出 $E = e^2$ 和 $G = g^2$，而 Blaschke (1929, p. 162) 偶尔会精确地使用我们的 A,B 表示法。

息？

如果我们知道每个点的曲率 $\mathcal{K}$，那么（通过[2.3]），我们将对 $\mathcal{S}$ 的性质和形状有一个清晰的理解。但是由于度量知道关于内蕴几何的一切，它必然（特别是）包含关于曲率的信息。因此，保证了 $\mathcal{K}$的公式的存在，并且由于度量公式的对称性，也很明显，在同时交换 $u \leftrightarrow v$ 和 $A \leftrightarrow B$ 的情况下，$\mathcal{K}$ 的公式必然是对称的。

尽管如此，$\mathcal{K}$的公式如此美丽[5]，仍然令人惊叹：

$$\boxed{\mathcal{K} = -\frac{1}{AB}\left(\partial_v\left[\frac{\partial_v A}{B}\right] + \partial_u\left[\frac{\partial_u B}{A}\right]\right)} \tag{4.10}$$

路还很长，但我们最终将以其应得的简单和优雅推导出这个公式，首先是在第四幕（第 27 章）中几何上，最后是在第五幕（第 38.8.2 节）中通过计算（使用嘉丹的微分形式）。不过，就目前而言，我们将其视为未来交付给我们的超先进技术，就像星际迷航 23 世纪的相位武器：我们可以在不开始了解（目前）它是如何工作[6]的情况下向目标扣动扳机。

例如，欧几里得度量 $d\hat{s}^2 = dr^2 + r^2 d\theta^2$具有$u = r$、$v = \theta$、$A = 1$和$B = r$，因此

$$\mathcal{K} = -\frac{1}{r}\left(\partial_\theta\left[\frac{\partial_\theta 1}{r}\right] + \partial_r\left[\frac{\partial_r r}{1}\right]\right) = -\frac{1}{r}(\partial_r 1) = 0$$

不出所料。

接下来，让我们瞄准球面度量（4.4）。在这种情况下，$u = \theta, v = \phi, A = R\sin\phi, B = R$，所以

$$\mathcal{K} = -\frac{1}{R^2\sin\phi}\left(\partial_\phi\left[\frac{\partial_\phi R\sin\phi}{R}\right] + \partial_\theta\left[\frac{\partial_\theta R}{R\sin\phi}\right]\right) = -\frac{\partial_\phi\cos\phi}{R^2\sin\phi} = +\frac{1}{R^2}$$

不出所料。

虽然计算更加冗长，仍鼓励你尝试将此公式应用于球面的射影度量公式（4.2）；这也应得到$\mathcal{K} = +(1/R^2)$。

在继续之前，注意到稍后需要的另一个结果。度量告诉我们如何将地图中的小距离转换为曲面上的距离。但应该如何转换面积呢？好吧，在[4.3]中，地图中的一个小矩形的面积为 $\delta u\delta v$，而在曲面上，相应的平行四边形的面积最终等于$(A\delta u)(B\delta v\sin\omega)$。因此，使用（4.7），

---

5 同样，大多数课本都保留了高斯的原始 E,G 符号，所以这个非凡公式的美在传统上被五个令人分心和不必要的平方根的外观所破坏！

6 这种方法并非没有风险：见证 20 世纪麦考伊博士在《永远边缘的城市》中被盗的移相器的悲惨使用！

曲面上无限小面元 $d\mathcal{A}$ 的公式为

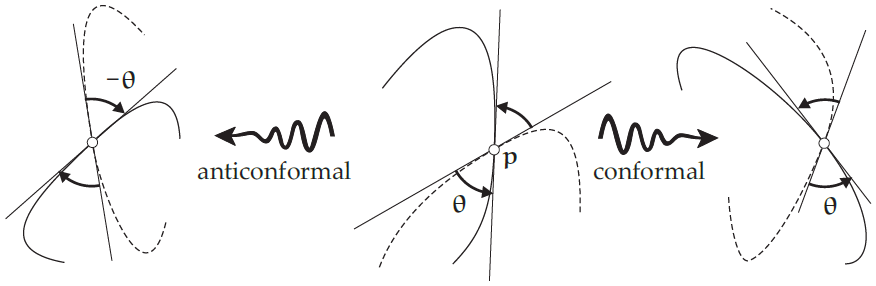
$$\boxed{d\mathcal{A} = \sqrt{(AB)^2 - F^2}\, du\, dv} \tag{4.11}$$

如果我们专门（就如通常所做）使用正交坐标系，使得$F = 0$，则这个公式简化为

$$\boxed{d\mathcal{A} = AB\, du\, dv} \tag{4.12}$$

4.5 共形映射

虽然球面的射影映射有保持直线的优点，但对于几乎所有目的而言，最好牺牲直线以保持角度。保持角度的大小和方向[7]的地图称为共形的（conformal）；如果保留了大小但反转了方向，则称为反共形的（anticonformal）。



[4.4]在中间，两条曲线在 p 处相交，它们之间的角度（从虚线到实线）定义为它们切线之间的角度 θ。共形地图（右）保持角度的大小和方向，而反共形地图（左）保持角度大小但反转了方向。

当我们谈到两条曲线之间的角度时，我们指的是它们的切线之间的角度。 见[4.4]。

根据度量公式（4.1），当且仅当扩展因子 Λ 不依赖于从 z 发出的无穷小向量 dz 的方向 γ 时，地图是共形的：

$$\boxed{\text{共形地图} \quad \Leftrightarrow \quad d\hat{s} = \Lambda(z)ds} \tag{4.13}$$

这种地图的最大优点是

> 曲面 $\mathcal{S}$ 上的一个无穷小形状在共形地图中由相似的形状表示，该形状与原始形状的不同之处仅在于大小：$\mathcal{S}$ 上形状的线性尺寸仅比地图中形状的线性尺寸大 Λ 倍。

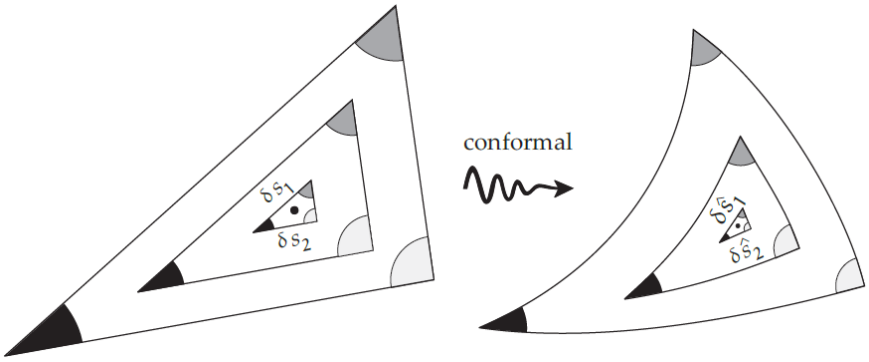
7 逆时针 (+) 或顺时针 (−)。

事实上，十八世纪的数学家用了*局部相似*（similar in the small）的表达方式，代替了现代术语*共形*（conformal）。

很明显，（4.13）意味着共形，并且要看到相反的情况是正确的（如所声称的），请考虑[4.5]。在左边，我们看到一个三角形逐渐缩小到曲面共形地图中的一个点。此时，曲面上相应的曲线三角形收缩为一个直线三角形，使用前言中引入的语言，最终是相似的：

$$\frac{\delta\widehat{s_1}}{\delta s_1} \asymp \frac{\delta\widehat{s_2}}{\delta s_2} \asymp \Lambda$$

对于某些 Λ，与三角形边 δs_1 和 δs_2 的方向无关。因此，我们看到共形意味着（4.13）。



图[4.5]当三角形在地图（左）中收缩时，它与曲面（右）上的共形像最终相似：$\frac{\delta\widehat{s_1}}{\delta s_1} \asymp \frac{\delta\widehat{s_2}}{\delta s_2}$。

在讨论一般度量公式（4.7）时，很明显，我们总是可以专门研究曲面上 u-曲线和 v-曲线的正交坐标系，对应于地图中的正交垂直线和水平线。但在这个阶段，这些方向上的扩展因子 A 和 B 是不同的，因此地图中的一个无限小圆被拉伸成曲面上的一个椭圆，角度通常会改变。

现在考虑更专业化，即扩展因子在所有方向上都*相同*的情况，这样$A = B = \Lambda$，无限小圆被映射到无限小圆，*角度被保留*。在这种情况下，(u, v) 坐标称为*共形坐标*（conformal coordinates），或*等温坐标*（isothermal coordinates），（4.9）简化为欧几里得度量的简单倍数：

$$\boxed{d\hat{s}^2 = \Lambda^2\,[du^2 + dv^2]} \tag{4.14}$$

这是一个如此强大的限制，以至于人们可能会担心这样的地图甚至可能不存在。但正如高斯在 1822 年发现的那样，至少在局部，绘制这样一张一般曲面的地图*总是*可行的。值得注意的是，证明（见例 8）依赖于*复数*——事实上，复分析和共形地图之间的深层次联系是下一节的主题。

已经优雅的曲率公式（4.10）现在变得更加优雅[8]。回想一下，二阶拉普拉斯[9]微分算子∇^2定义为

$$\boxed{\nabla^2 \equiv \partial_u^2 + \partial_v^2} \tag{4.15}$$

既然现在$A = B = \Lambda$，很容易发现[练习]（4.10）简化为

$$\boxed{\mathcal{K} = -\frac{\nabla^2 \ln\Lambda}{\Lambda^2}} \tag{4.16}$$

4.6 一些复分析

即使把曲面$\mathcal{S}$变成平面（就可处理平面到自身的共形映射），这也是一个深入而丰富的研究领域。值得注意的是，这种共形映射与复数密不可分地交织在一起，我们将在本节中简要解释（在这一幕稍后将出现更多的具体例子）。为了全面探索复分析的壮观结果是如何从这个几何基础中产生的，我们将不谦虚地推荐我的第一本著作——《可视化复分析》（VCA）。

每个曲面 $\mathcal{S}$ 不只有一个共形映射，而是有无限多样的共形映射！我们首先注意到，产生具有度量（4.14）的共形 (u, v) 坐标系的特定 u-曲线和正交 v-曲线并不唯一：

$$d\hat{s} = \Lambda(u, v)ds$$

真正的魔法存在于共形映射 $F: \mathbb{C} \to \mathcal{S}$本身。给定这样的共形映射 F，可以简单地通过旋转、扩展和平移 $\mathbb{C}$ 中的 (u, v) 坐标网格，在$\mathcal{S}$上创建无限多种共形 $(\tilde{u}, \tilde{v})$ 坐标，（通过应用 F）在$\mathcal{S}$上产生了全新的$\tilde{u}$曲线和正交$\tilde{v}$曲线。而 S 上的这个新的正交$(\tilde{u}, \tilde{v})$坐标系和原始坐标系一样共形。

让我们引入一些符号来充分解释这一点。正如复分析中的习惯，将认为 $z = u + iv$居住在 $\mathbb{C}$ 的一个副本中，它的像$\tilde{z} = \tilde{u} + i\tilde{v}$在复变函数 $z \mapsto \tilde{z} = f(z)$ 作用下居住在$\mathbb{C}$的另一个副本中：

---

8 根据 Dombrowski (1979, p. 128)，这是高斯发现的第一个曲率公式，记录在他 1822 年 12 月 13 日的私人笔记中。 参见 Gauß (1973, p. 381)。 直到 1825 年，他才在更一般的（非共形的）正交坐标中找到了一个公式，但是，他再次保守了这个秘密。 最终，在 1827 年，他在非正交坐标系中找到了完全通用的公式，而最终的杰作（Gauss 1827）中也只有这个极其复杂丑陋的公式才出现。 我们注意到，高斯为他故意隐瞒自己的动机和发现之路的刻意和完全可悲的（在我们看来）习惯感到非常自豪，并宣称：“没有建筑师在完成建筑后将脚手架留在原处。”

9 虽然物理学家倾向于将此算子表示为 ∇^2，但数学家倾向于将其写为 Δ。 我们避免使用后一种符号，因为我们经常使用 Δ 来表示三角形。

$$z \mapsto \tilde{z} = \tilde{u} + i\tilde{v} = f(z) = f(u + iv)$$

在刚才讨论的例子中，有$f(z) = ae^{i\tau}z + w$，它通过（实）因子a进行扩张，通过角度τ旋转，然后通过（复）数w进行平移。

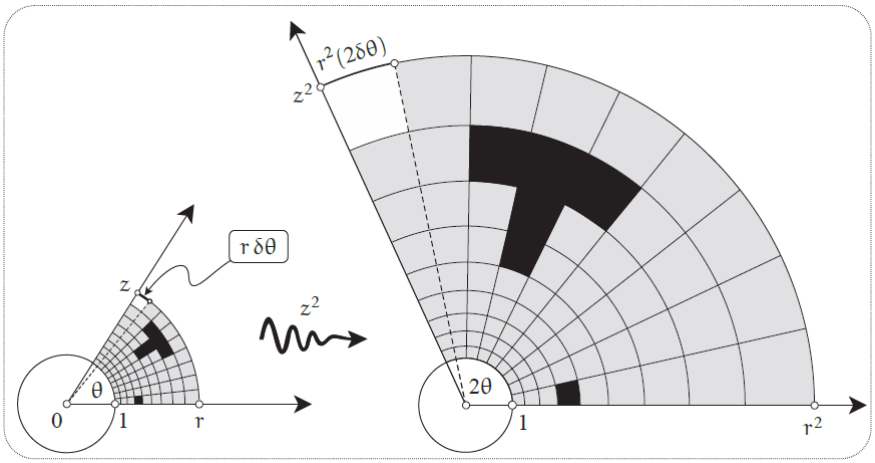
将映射 f 和映射 F 进行复合，就得到了从$\mathbb{C}$到$\mathcal{S}$的新的共形映射$\tilde{F} \equiv F \circ f$。如果 z 沿着一个小的复数δz移动，那么它在第一个映射$z \mapsto \tilde{z} = f(z)$下的像将沿着δz的“像”，从$\tilde{z}$出发的$\delta\tilde{z}$移动，很明显

$$\delta\tilde{z} = ae^{i\tau}\,\delta z \tag{4.17}$$

因此，长度$\delta s = |\delta z|$被因子 a 拉伸，使得$\delta\tilde{s} = |\delta\tilde{z}| = a\,\delta s$。接下来，在第二映射 F（直到曲面$\mathcal{S}$）下，$\delta\tilde{s}$的长度将由共形度量因子$\Lambda(\tilde{z})$拉伸。因此，在新的映射$\tilde{F}$下，δs的净放大率是由这两个膨胀系数的乘积给出的：

$$d\hat{s} = \tilde{\Lambda}(z)\,ds, \text{其中 } \tilde{\Lambda}(z) = a\,\Lambda(\tilde{z})$$

但即使是这种自由度也几乎没有划伤曲面，并解释为什么我们必须简单地引用复分析中的一个基本事实；有关详细信息，请参见 VCA。你研究过的每一个熟悉的、有用的实函数f(x)（如x^m、e^x和 sin x）都有一个唯一的到复数域的推广f(z)。和以前一样，把这想象成从一个$\mathbb{C}$到另一个$\mathbb{C}$的映射，复分析的奇迹是，所有这些自然发生的映射f(z)都是自动共形的！



图[4.6] 映射$z \mapsto f(z) = z^2$（z 的所有次幂均成立）是共形的，因此左侧“正方形”的精细网格映射到右侧的近似正方形，当正方形的大小缩小到零时这些像最终成为完美的正方形。

例如，图[4.6]说明了

$$z \mapsto \tilde{z} = f(z) = z^2 = \left[r\,e^{i\theta}\right]^2 = r^2 e^{i2\theta}$$

使每个点的长度平方并使角度加倍的效果。正如我们所看到的，左边的小“方块”被发送到右边的相似“方块”。当然，这两组“方块”只在它们缩小到零的极限情况时成为真正的正方形。同样，左边的黑色 T 形越小，右边像中的 T 形就越相似。

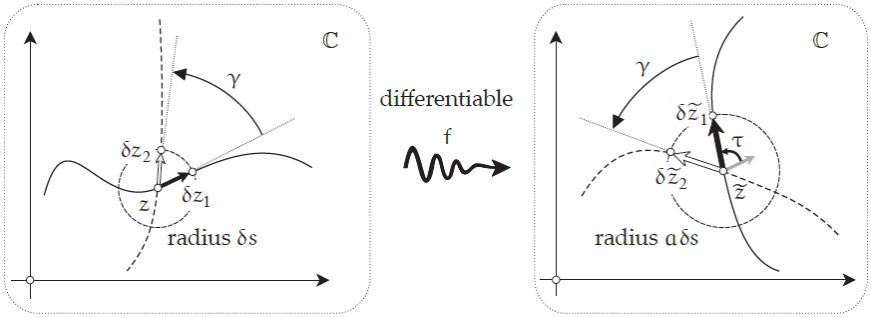
要想看看这有多了不起，假设你随机写下两个（实变量 u 和 v 的）实函数$\tilde{u}$和$\tilde{v}$，然后这两个函数“奉子成婚”，将它们合并成一个复合映射，$f(u,v) = \tilde{u} + \tilde{v}$。那么，f 基本上不可能是共形的，而且，正如我们现在将看到的，这也意味着$f'(z)$ 不可能存在！

让我们开始重做先前的分析，但现在用一般映射替换线性函数$f(z)$，使得导数$f'(z)$ 存在。这些函数被称为解析的（analytic）。正如我们刚才指出的，这种解析映射是非常罕见的，但它们基本上包括了数学和物理中自然产生的所有有用函数。

现在（4.17）的对应物是

$$\boxed{\delta\tilde{z} \asymp f'(z)\delta z = ae^{i\tau}\delta z} \tag{4.18}$$

主要的区别是，现在膨胀系数$a(z)$和旋转角度$\tau(z)$都取决于位置 z，而不是在整个$\mathbb{C}$中恒定。例如，在图[4.6]中，我们可以看到，与实轴相邻的黑色正方形比网格角中的白色正方形经历的膨胀要小，因此黑色正方形的$a(z)$比白色正方形的小。同样，很明显，黑色方块没有旋转，所以那里的$\tau = 0$，但白色方块显然需要旋转才能获得其图像。



图[4.7]**伸扭**（amplitwist） 可微复映射$f(z)$的局部效果是由复数$f'(z)$ = 伸扭 $f(z)$ = （伸缩）e^{i}（扭转） $= ae^{i\tau}$描述的伸扭（伸缩 amplify 和扭转 twist）。从 z 发出的任何一对微小复数$\delta z_{1,2}$被均匀伸扭，以产生从$\tilde{z} = f(z)$发出的一对$\delta\tilde{z}_{1,2} \asymp ae^{i\tau}\,\delta z_{1,2}$，因此，曲线之间的角 γ 被 保持：f 是保形的！

等式（4.18）说，从 z 发出的每个微小复箭头 δz 都经历相同的伸缩 a 和旋转 τ，以得到从 $\tilde{z} = f(z)$ 发出的像箭头 $\delta\tilde{z}$（见图[4.7]）。如图所示，从 z 发出的一对 δz 之间的角度将与 $\tilde{z} = f(z)$ 处的 δz 像对之间的角度相同，因此可微复映射是自动保形的！

我们看到导数 $f'(z)$ 描述了共形映射的局部行为，在 VCA 中，我们引入了（非标准）术语来几何地描述这一点。我们称局部膨胀系数 a 为伸缩（amplify）；局部旋转角度 τ 为扭转（twist）；我们称伸缩和扭转组合（将原始形状转换为其像所需）为伸扭（amplitwist）。总之，

$$\boxed{f'(z) = f\text{ 在 }z\text{ 处的伸扭} = (\text{伸缩})e^{i[\text{扭转}]} = ae^{i\tau}} \tag{4.19}$$

在我们转向空间中曲面的保形度量公式之前，让我们回到$f(z) = z^2$，并展示我们如何使用图[4.6]以几何方式推导其伸扭。

将注意力集中在图示的白色正方形上，其中一个角位于$z = re^{i\theta}$。由于在角度θ处通过 z 的径向边缘映射到在角度 2θ处通过z^2的径向边缘，因此该边缘经历了θ的旋转，因此

$$\tau = \text{扭转} = \theta$$

要找到放大系数a，请考虑正方形的高亮外边缘（最终等于圆形的弧度，通过 z 连接白点）。如果这在原点对角$\delta\theta$，那么它的长度最终是$r\,\delta\theta$。由于角度被映射翻倍，图像弧将对角$2\delta\theta$，并且由于它现在位于半径 r2 的圆上，因此该图像边缘的长度最终为$r^2(2\delta\theta)$。因此，

$$\left(\text{像的边长}\right) \asymp 2r\left(\text{原边长}\right) \Longrightarrow a = \text{伸缩} = 2r$$

我们的结论是，

$$(z^2)' = z^2\text{的伸扭} = \left(\text{伸缩}\right)e^{i[\text{扭转}]} = 2r\,e^{i\theta} = 2z$$

一个看起来与实际结果形式上相同的结果，$(x^2)' = 2x$，但现在这涵义更加丰富。

推广这个几何论证并得出$(z^m)' = mz^{m-1}$并非难事[练习]。其他重要映射的伸扭同样可以纯几何推导；有关详细信息，请参见 VCA。

现在让我们回到我们的主要兴趣：曲面上的共形坐标。我们现在可以用非常丰富的可微（即共形）映射类$f(z)$来取代简单的线性函数。再次从$\mathbb{C}$到$\mathcal{S}$定义$\tilde{F} \equiv F \circ f$，新的度量公式为

$$d\hat{s} = \tilde{\Lambda}(z)\,ds, \text{其中}\ \tilde{\Lambda}(z) = \left(\text{伸缩}\right)\Lambda(\tilde{z}) = |f'(z)|\Lambda(\tilde{z})$$

共形映射$F: \mathbb{C} \mapsto \mathcal{S}$的存在允许我们采取任何解析映射$f: \mathbb{C} \longmapsto \mathbb{C}$，将其传送到$\mathcal{S}$，使其成为$\mathcal{S}$到自身的共形映射。要看到这一点，请考虑

$$\hat{f} \equiv F \circ f \circ F^{-1} \tag{4.20}$$

作用于$\mathcal{S}$。首先F^{-1}共形映射$\mathcal{S}$到$\mathbb{C}$，然后 f 共形映射$\mathbb{C}$到$\mathbb{C}$，最后 F 共形映射$\mathbb{C}$回到$\mathcal{S}$。由于三个映射中的每一个都保持角度，它们的组合也保持角度，所以 $\hat{f}: \mathcal{S} \mapsto \mathcal{S}$确实是共形的。

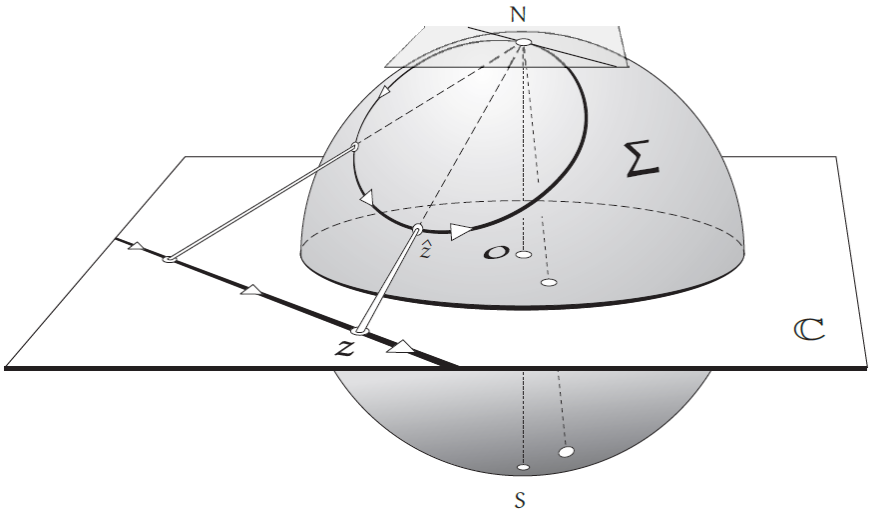
在下一节中，我们将遇见在$\mathcal{S} = \mathbb{S}^2$的情况下的极其重要的共形映射 F。稍后，我们将使用这个 F，通过（4.20），将$\mathbb{S}^2$的旋转表示为复函数（由（6.10）给出）。这些旋转不仅仅是共形的，它们也是球面的（保持方向）等距同构（isometries）。

4.7 球面的共形球极地图

喜帕恰斯（Hipparchus，约公元前 150 年）可能是使用图[4.8]所示方法构建球面共形[10]地图的第一人，该方法称为球极射影法（stereographic projection）。当然，到了公元 125 年，托勒密（他通常被认为是发现者）用这种方法来描绘天体在*天球*（celestial sphere）上的位置。

这种构造类似于中心射影，不同之处在于现在我们在北极 N 处想象一个点光源，并且我们投影到一个穿过赤道的平面上，而不是（投影到）接触南极[11]（的平面上）。一束通过 Σ 上点 $\hat{z}$ 的光线继续投影到 $\mathbb{C}$ 中的点 z，我们称之为 $\hat{z}$ 的*球极像*（stereographic image）。由于这给了我们 $\mathbb{C}$ 中的点和 Σ 上的点之间的一一对应关系，我们也可以说 $\hat{z}$ 是 z 的球极像。这样做不会引起混淆，从上下文就可以看清楚我们是把 $\mathbb{C}$ 映到 Σ，还是反之。

注意以下直接事实：(i) Σ 的南半球映射到圆 $|z| = R$ 的内域，特别是南极 S 映射到 $\mathbb{C}$ 的原点 0； (ii) 赤道上的每个点都映射到自身； (iii) 在 $\mathbb{C}$ 中，圆 $|z| = R$ 的外域映射到 Σ 的北半球，除了 N 不是平面中任何有限点的像。然而，很明显，随着 z 越来越远离原点（在任何方向上），$\hat{z}$ 越来越接近于 N。 在复分析，以及这一幕的后续篇幅，一般将 Σ 设定为*单位球面*，一旦它的点用复数进行球极标记，它就称为*黎曼球面*（Riemann sphere）。那么 N 点就是所谓*扩展复平面*（extended complex plane）上*无穷远点*（point at infinity）的具体表现。



图[4.8] **球极射影**：来自 N 的光穿过玻璃球面 Σ，投射到 $\mathbb{C}$ 上。通过 N 的圆的投影是 $\mathbb{C}$ 中与圆在 N 处的切线平行的直线。

图[4.8]说明了这样一个事实：

---

10 [4.8] 中构造的共形性并不意味着立即显而易见； 将在 [4.9] 中简要说明。

11 在一些较早的文献中，实际上的确取接触南极的平面； 这仅以恒定的比例因子 2 改变地图 [练习]。

平面中直线的球极像是Σ上的一个圆，该圆沿平行于原直线的方向通过 N。 (4.21)

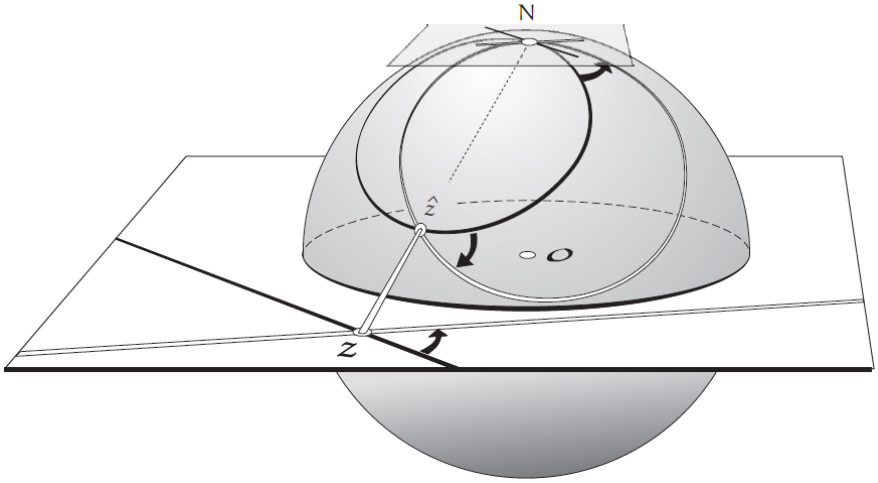
要看到这一点，首先观察当 z 沿着 [4.8] 中所示的线移动时，连接 N 到 z 的线扫过通过 N 的平面的一部分。因此 $\hat{z}$ 沿着该平面与 Σ 的交线运动，它就是过 N 的圆周。接下来，注意 Σ 在 N 点的切平面平行于ℂ。但是如果用第三个平面切开两个平行平面，会得到两条平行的交线。 因此，如所声称的，圆在 N 处的切线平行于原直线。

从这最后一个事实可以得出，球极射影保持角度不变。请看图[4.9]，其上画出两条相交于 z 的直线以及它们经过 N 的球极像圆周。

首先注意圆的交角在它们的两个交点 $\hat{z}$ 和 N 处的大小是相同的。这是因为该图形在通过球心和两个圆[12]心的平面上具有镜像对称性。因为两圆周在 N 点的切线平行于平面上的原直线，可知画在 z 和 $\hat{z}$ 处的两个角有相同的大小，但是要对球面上的角赋以方向后才能说球极射影是“共形”。

按我们的规定，图上所画的 z 点处的角（从由粗黑线到中空的双线的角）为正，即是说从平面上方看它是逆时针方向的。从我们画图 [4.9] 所用的透视来看，$\hat{z}$ 处的角是负的，即顺时针方向的。 然而，如果从球面内域来看这个角，它又是正的。于是，

如果我们依位于球内的观察者之所见来定义球面Σ上的角的方向，则球极射影是共形的。



图[4.9] **球极射影是共形的。** 当直线绕 z 旋转时，其圆周像的 N 处的切线也随之旋转，因此 z 和 N 处的角度相等。根据对称性，$\hat{z}$ 处的角等于 N 处的角，因此等于 z 处的角，证明球极射影是共形的。

12 如果你自己在橘子上画出任意两个相交的圆圈，这将变得一清二楚。

历史注释：值得注意的是，尽管球极射影自托勒密在公元 125 年左右首次投入实际应用以来就广为人知，但其美丽且极为重要的共形性在接下来的 1500 年后才被发现！这是由托马斯·哈里奥特[13]在 1590 年左右首次完成的——是的，就是同一位托马斯·哈里奥特，他在 1603 年发现了将角盈与球面面积联系起来的基本公式（1.3）！

根据共形性，度量采用（4.13）形式。 因此，球面上一个半径为$\delta\hat{s}$ 的很小的圆最终映射到平面上一个半径为δs 的圆，其中

$$\delta\hat{s} \asymp \Lambda\, \delta s$$

见图[4.10]现在的任务是找到Λ。

由于Λ与从 $\hat{z}$ 发出的半径$\delta\hat{z}$ 的方向无关，可以自由选择这个方向，以使几何分析尽可能简单。因此，如图所示，沿着纬度圆选择水平方向。

在球极射影下，这个纬度圆经过均匀扩展，产生一个以原点为中心的圆，通过 z，δs 指向它。 当 $\hat{z}$ 绕其纬度圆旋转时，z 也随之旋转，它们的运动与它们与 N 的距离成正比。因此，

$$\frac{\delta\hat{s}}{\delta s} = \frac{N\hat{z}}{Nz}$$

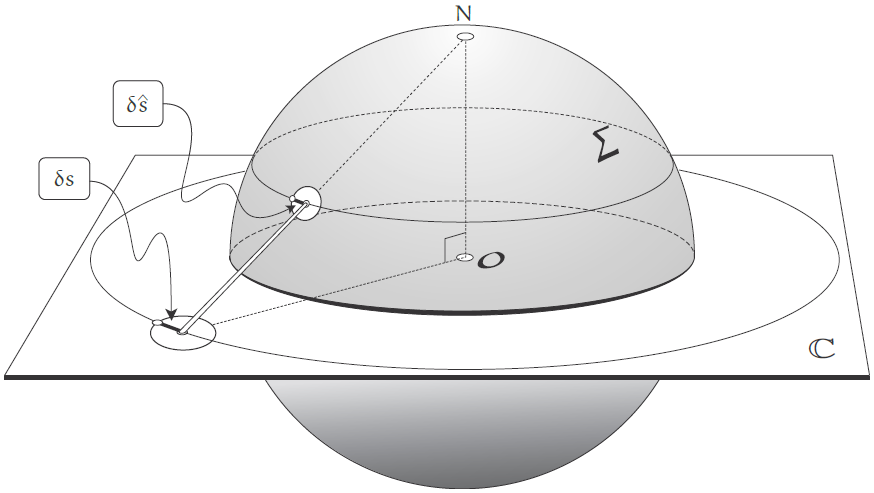
现在考虑图[4.11]，它显示了图[4.10] 通过 N、$\hat{z}$、z 截取的垂直截面。 三角形 $N\hat{z}S$ 相似于 $N0z$，所以

$$\frac{N\hat{z}}{2R} = \frac{R}{Nz}$$

将这与前面的方程结合起来，我们得出

$$\frac{\delta\hat{s}}{\delta s} = \frac{2R^2}{[Nz]^2}$$

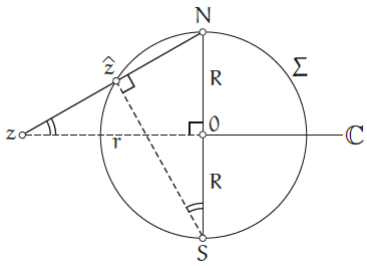
13 参见 Stillwell (2010, §16.2)。 有关哈里奥特生平的简短描述，请参见 Stillwell (2010, §17.7)。



图[4.10] 共形性意味着球体上一个半径为 $\delta\hat{s}$ 的非常小的圆最终映射到平面中一个半径为 δs 的圆。为了找到度量，我们必须找到它们的半径比，我们通过沿着图示的纬度圈选择 $\delta\hat{s}$ 来做到这一点。

最后，记 $r = |z|$，并将毕达哥拉斯定理应用于三角形 $N0z$，我们得到 $[Nz]^2 = R^2 + r^2$，因此球极地图的共形度量是

$$\boxed{d\hat{s} = \frac{2}{1+(r/R)^2}ds} \tag{4.22}$$



图[4.11] 三角形 $N\hat{z}S$ 相似于 $N0z$。

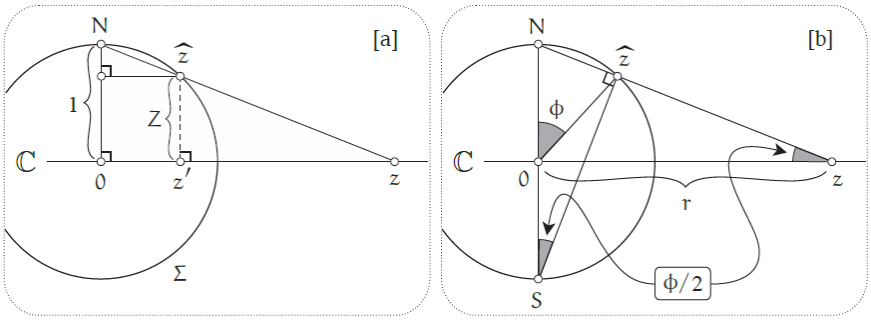
这是我们第一次（会有其他机会）尝试共形曲率公式（4.16）； 当然我们应该找到 $\mathcal{K} = (1/R^2)$。我们建议你通过两种方式来确认。首先，写出 $r^2 = x^2 + y^2$ 并使用拉普拉斯算子的原始笛卡尔形式 (4.15)。其次，使用在极坐标中拉普拉斯算子采用以下形式的事实

$$\boxed{\nabla^2 = \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2} \tag{4.23}$$

4.8 球极射影的公式

我们将在这一小节中导出把Σ上点$\hat{z}$的坐标与其在$\mathbb{C}$上的球极射影 z 的联系起来的显式的公式。为了简化问题，我们现在将限定在标准情况 $R = 1$。

首先，我们用 z 的笛卡尔坐标 $z = x + iy$ 来描述 z。类似地，设 (X, Y, Z) 是 Σ 上之 $\hat{z}$ 的笛卡尔坐标；这里取 X 轴和 Y 轴即为 $\mathbb{C}$ 上的 x 轴和 y 轴，所以正 Z 轴通过 N。为了能自如的应用这些坐标，先请检验一下以下的事实，看看你对上面讲的是否都清楚了：Σ 的方程是 $X^2 + Y^2 + Z^2 = 1$, N 的坐标是 $(0, 0, 1)$，类似地，$S = (0,0,-1), 1 = (1,0,0), i = (0,1,0)$。



图[4.12] [a]通过相似三角形，$|z|/|z'| = 1/(1 - Z)$。[b] $r = \cot(\phi/2)$。

令 $z \in \mathbb{C}$ 的球极像是 $\hat{z}$，现在来求用球极像 $\hat{z}$ 的坐标 (X, Y, Z) 来表示 $z = x + iy$ 的公式。由 $\hat{z}$ 向 $\mathbb{C}$ 作垂线，设其垂足是 $z' = X + iY$。很明显 z 与 z′ 在对原点方向相同，故

$$z = \frac{|z|}{|z'|} z'$$

现在看图[4.12a]，那里画的是 Σ 和 $\mathbb{C}$ 与通过 N 和 $\hat{z}$ 的垂直平面的截口；注意，这个截口平面也一定包含 z' 和 z。由图上所画的斜边为 $N\hat{z}$ 和 Nz 的两个直角三角形的相似性，立即可得 [练习]

$$\frac{|z|}{|z'|} = \frac{1}{1 - Z}$$

于是就有第一个球极射影公式：

$$x + iy = \frac{X + iY}{1 - Z} \tag{4.24}$$

为了后续使用，现在用这个公式反过来求把 $\hat{z}$ 的坐标 (X, Y, Z) 写成 z 的坐标 $z = x + iy$ 的公式。因为 [练习]

$$|z|^2 = \frac{1 + Z}{1 - Z}$$

可以得到 [练习]

$$X + iY = \frac{2z}{1+|z|^2} = \frac{2x+i2y}{1+x^2+y^2}, \qquad Z = \frac{|z|^2-1}{|z|^2+1} \tag{4.25}$$

用三个坐标 (X, Y, Z) 来描述 Σ 虽然时常是方便的，但肯定不自然，因为球面内在的是 2 维的。 如果我们代之以更自然的（2 维的）球极坐标 (ϕ, θ) 来描述，就会得到一个特别干净的球极射影公式。

先回忆一下，θ 量度绕 Z 轴的转角，而 $\theta = 0$ 是通过正 X 轴的垂直平面。 故对 $\mathbb{C}$ 中的点 z， θ 就是通常的由正实轴转到 z 的转角，角 ϕ 的定义[14]则如图 [4.12b] 所示，是点 N 和 $\hat{z}$ 在 Σ 之球心所张的角；例如赤道平面对应于 $\phi = (\pi/2)$。 我们规定，$0 \le \phi \le \pi$。

若 z 是坐标为 (ϕ, θ) 的点 $\hat{z}$ 的球极射影，显然有 $z = r\,e^{i\theta}$，所以现在只需求出作为 ϕ 的函数的 r 。由图 [4.12b] 很明显，三角形 $N\hat{z}S$ 和 $N0z$ 相似 [练习]，而且因为 $\angle NS\hat{z} = (\phi/2)$，立即有 $r = \cot(\phi/2)$ [练习]。所以，新的球极射影公式是

$$\boxed{z = \cot(\phi/2)\, e^{i\theta}} \tag{4.26}$$

我们还将在练习 33 中说明如何用这个公式来对球极射影给出另一种美丽的解释，那是应归功于罗杰 • 彭罗斯爵士的。

第一个应用是导出代表对径点的复数之间的关系。先回忆一下，位于球面上一直径两端的两点（例如南极和北极）互相称为对径点。现在证明

> 若 $\hat{p}$ 与 $\hat{q}$ 是 Σ 的对径点，则其球极射影 p 和 q 之间有关系式
>
> $$q = -(1/\bar{p})$$

(4.27)

注意，p 和 q 的上述关系其实是对称的（显然本应如此）：$p = -(1/\bar{q})$。为了证明（4.27），请自行验证一下，若 $\hat{p}$ 的坐标为 (ϕ, θ)，则 $\hat{q}$ 的坐标是 $(\pi - \phi, \pi + \theta)$。因此，

$$q = \cot\left[\frac{\pi}{2} - \frac{\phi}{2}\right] e^{i(\pi+\theta)} = -\frac{1}{\cot\left(\frac{\phi}{2}\right)} e^{i\theta} = -\frac{1}{\cot\left(\frac{\phi}{2}\right) e^{-i\theta}} = -\frac{1}{\bar{p}}$$

证毕。

有关（4.27）的初等几何证明，请参见练习 7。

---

14 这是美国的惯例； 在我的家乡英格兰，θ 和 ϕ 的作用与此处所述相反。 有关同一图表的一个可爱示例，但按照英国惯例进行标记，请参阅 Penrose 和 Rindler (1984, Vol. 1, p. 12)。

4.9 球极射影的保圆性质

本节致力于证明一个美丽、令人惊讶且至关重要的单一事实：

$$\boxed{\text{球极射影保持圆周！}} \tag{4.28}$$

也就是说，它不仅将无穷小圆发送到无穷小圆（所有共形地图都必然如此），而且它将球面上任何大小和位置的有限圆发送到平面中的完美圆，尽管球心没有映射到平面的中心。请注意，如果圆在 N 近处经过，那么它的像将非常大，并且在极限情况下它经过 N 时，这个非常大的圆变成了图[4.8]中所示的直线。

（4.28）有一个漂亮的、完全概念化的几何解释——见 VCA，第 142 页——但在我们目前无暇展开，只能求助于计算。

单位球面 Σ 上的每个圆都是Σ与距离 O 小于 1 的平面的交集：

$$\mathrm{lX+mY+nZ=k}, \text{其中 } \mathrm{l}^2+m^2+n^2 \geq \mathrm{k}^2$$

将 (4.25) 代入该平面的方程，我们发现 [练习] Σ上的交点球极射影到 ℂ 中的曲线，其方程为

$$2\mathrm{lx}+2\mathrm{my}+\mathrm{n}(x^2+y^2-1)=\mathrm{k}(x^2+y^2+1)$$

如果 k=n，那么[练习] Σ上的圆经过 N，它的图像是一条直线（本应如此！），方程为$\mathrm{lx}+\mathrm{my}=\mathrm{n}$。 如果 $\mathrm{k} \neq \mathrm{n}$ 那么我们可以凑成平方[练习] 将方程重写为

$$\left[x-\frac{l}{k-n}\right]^2+\left[y-\frac{m}{k-n}\right]^2=\frac{l^2+m^2+n^2-k^2}{(k-n)^2}$$

这确实是一个圆圈

$$\text{圆心}=\left(\frac{l}{k-n},\frac{m}{k-n}\right), \qquad \text{半径}=\frac{\sqrt{l^2+m^2+n^2-k^2}}{|k-n|}$$

（评论：这是一个很好的例子，展示了计算的诱人但具有破坏性的力量。我们调用了“魔鬼的机器”——见前言——仅仅几行就完成了它的工作，结果得到了证明。然而我们站在这里，不知道为什么结果是真的！）

根据保圆性质，很容易从图[1.5]推断出球面上的测地线（大圆）在地图中显示为在直径的两端与赤道圆相交的圆。

第五章 伪球与双曲面

5.1 贝尔特拉米的洞见

随着罗巴切夫斯基和鲍耶发现了双曲几何，对平行线的长期研究在 1830 年左右达到了高潮。在一个完全不同的平行线（双关语）上，随着高斯 1827 年的微分几何发现，对微分几何的研究几乎在同一时间达到了高潮。就像球面上最初平行[64]的线最终必然相交一样，这两条平行的思想之线也会以一种强大而富有成果的方式碰撞。

[5.1] 尤金尼奥·贝尔特拉米（1835–1900）。

1868 年，意大利几何学家尤金尼奥·贝尔特拉米（见[5.1]）认识到，这些看似不相关的思想领域的两个结果之间可能存在联系。一方面，在双曲几何中，他知道兰伯特的结果（1.8）——后来被高斯、罗巴切夫斯基和鲍耶重新发现，即三角形的角盈是其面积的固定负倍数。另一方面，他也知道局部高斯—博内定理。

贝尔特拉米的见解是，如果人们能找到一个常负曲率 $\mathcal{K} = -(1/\mathrm{R}^2)$ 的曲面，那么，根据（2.6），在其中构建的测地三角形将自动服从双曲几何的中心定律：

$$\mathcal{E}(\Delta) = -\frac{1}{\mathrm{R}^2}\mathcal{A}(\Delta)$$

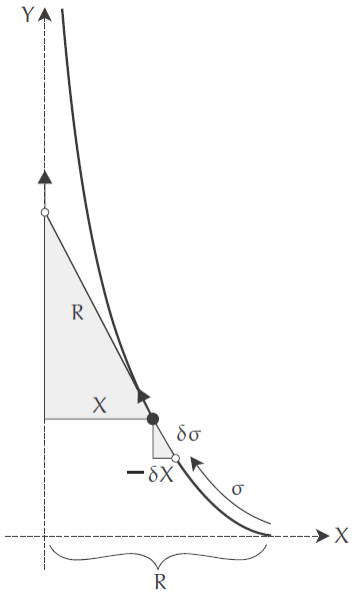
到目前为止，罗巴切夫斯基和鲍耶的奇异双曲几何已经默默无闻了近 40 年，被一些人诋毁，但被大多数人忽视。现在，贝尔特拉米终于有了一个想法，可以把它建立在一个安全和直观的基础上。也许双曲几何只是常负曲率曲面的内蕴几何！一场长达 2000 年的斗争即将结束。

64 想想赤道上的两个相邻点，以及穿过它们的经线。

5.2 轨迹与伪球

贝尔特拉米已经知道 [2.6] 中所示的伪球的确是一个常负曲率 $\mathcal{K} = -(1/\mathrm{R}^2)$ 的曲面，其中 R 是其圆形底部边缘的半径。（我们将在本节中证明这一事实。）因此，更具体地，他的见解是，这个曲面内的局部几何服从罗巴切夫斯基和鲍耶抽象非欧几何的定律。但这种抽象的双曲几何被理解为发生在一个无限大的双曲平面中（就像欧氏平面），它服从欧氏几何公理的前四个，但直线服从双曲公理（1.1），而不是欧氏几何的平行公理。

虽然伪球的常负曲率确保它忠实地体现了这个公理的局部后果，但伪球不会作为整个双曲面的模型，因为它以两种不可接受的方式偏离欧氏平面：（1）伪球类似于圆柱体，而不是平面，（2）线段不能无限地在两个方向上延伸：总会撞到边缘。（正如前面所提，希尔伯特在 1901 年发现，这样的边缘是所有常负曲率曲面的基本特征。它不是内蕴的，而是在试图迫使曲面可以装得进通常欧氏 3 维空间时产生的。）



[5.2]**曳物线**。重物连接到沿 X 轴布置的长度为 R 的细绳上。如果自由端沿 Y 轴向上移动，则重物将沿着所示的曲线拖动，轨迹称为曳物线。

贝尔特拉米认识到了这些障碍，在下一节中，我们将看到他是如何通过构造伪球的共形映射，一举克服了这两个障碍。不过，目前我们暂时转回伪球本身的构建。

尝试以下实验。拿一个小的重物，如镇纸，在上面绑上一段绳子。现在，将对象放置在桌面上，并通过沿桌子边缘移动细绳的自由端来拖动它。你将看到对象沿着 [5.2] 中的曲线移动，其中 Y 轴表示桌子边缘。这条曲线称为曳物线，Y 轴（曲线逐渐趋近）称为轴。1676 年，牛顿首次研究了曳物线。

如果细绳的长度为 R，则曳物线具有如下几何性质：从接触点到 Y 轴的切线段具有恒定长度 R。牛顿将其作为曳物线的定义。

回到图 [5.2]，令 σ 表示沿曳物线的弧长，$\sigma = 0$ 对应于我们要拖动对象的起始位置 $X = R$。就像物体即将通过 (X, Y) 一样，令 δX 表示物体沿着轨道移动距离 $\delta\sigma$ 时发生的 X 中非常小的变化。从图示三角形的最终相似性，我们推断出

$$\frac{-dX}{d\sigma} = \frac{X}{R}$$

并因此

$$\boxed{X = Re^{-\sigma/R}} \tag{5.1}$$

半径 R 的伪球[2.6]现在可以同时定义和构造为通过绕其轴线旋转曳物线获得的曲面。值得注意的是，这个曲面早在 1693 年（由克里斯蒂安·惠更斯）就被研究过，在它在接受双曲几何学方面的催化作用之前两个世纪，早在 1839 年，高斯的学生明丁就已经知道它曲率的恒定性。

[5.3]作者自制的伪球，由圆锥体建造，本身是由半径 R 的圆盘建造的。上半部已被移除，以使结构更容易看到。

为了证明曲率的恒定性，让我们找到伪球的度量。我们在曲面上选择非常自然的正交坐标曲线，如下所示。见[5.4a]（但请暂时忽略[5.4b]）。要确定伪球面上的一个点，我们只需要说（i）它位于哪条曳物线上，（ii）它沿着这条曳物线走了多远。为了回答（i），我们指定了围绕伪球轴线的角度 x，为了回答（ii），我们指定了从底部开始沿着轨道向上移动的距离σ。

因此，曲线$x = \text{const.}$ 是伪球的曳物线（这些显然是测地线[65]），曲线 $\sigma = \text{const.}$ 是伪球的圆形横截面（这些显然不是测地线）。由于这样一个圆的半径与[5.2]中的 X 坐标相同，因此从（5.1）可以看出

$$\boxed{\text{通过伪球上点 } (x, \sigma) \text{ 的圆} \sigma = \text{const. 的半径 } X \text{ 由 } X = R\,e^{-\sigma/R} \text{给出。}}$$

因此，增量dx会将点旋转圆弧长 Xdx（如图所示）。因此，度量为

65 通过对称性，任何旋转曲面的子午线都是测地线。

$$d\hat{s}^2 = X^2 dx^2 + d\sigma^2 = (Re^{-\sigma/R})^2 dx^2 + d\sigma^2 \tag{5.2}$$

最后，我们现在可以将曲率公式（4.10）应用于此度量，得到

$$\mathcal{K} = -\frac{1}{Re^{-\frac{\sigma}{R}}}\partial_\sigma\left[\frac{\partial_\sigma\left(Re^{-\frac{\sigma}{R}}\right)}{1}\right] = -\frac{1}{R^2}$$

从而证实了贝尔特拉米需要解释双曲几何的关键事实：

$$\boxed{\text{伪球具有恒定的负曲率 } \mathcal{K} = -(1/R^2)\text{，其中 R 是其底部半径。}} \tag{5.3}$$

这个命题在数学上和在历史上引起浓厚的兴趣，在这本书的过程中，我们将尝试尽可能直接地和几何地去理解它。事实上，在后面的几幕中，我们将提供两个几何证明：一个使用外在几何证明（第三幕），另一个使用内蕴几何证明（第四幕）。

在继续之前，我们想不出比自己做一个更好的方法来建立伪球几何感觉了！要查看构造背后的想法，请想象旋转[2.6]以创建伪球体。在这个过程中，无论拖动的重物的位置如何，旋转的细绳总是描出固定斜边长度 R 的圆锥体（与伪球相切）。

因此，用剪刀剪开尽可能多的纸，沿着三个边缘钉在一起，每个边缘钉几个钉子。找到可以放在在纸内的最大的碗或盘子，描出它的圆形边缘。现在沿着这个圆圈切割，产生相同的圆盘；重复此步骤，直到你至少有 20 张圆盘——越多越好！从第一个圆盘上切下一个小楔子，并将边缘用胶带粘在一起，形成一个非常浅的圆锥。取下一个圆盘，用稍微大一点的楔子[66]重复，以创建一个稍微高一点的圆锥体，但仍然具有相同的斜边长度。将此新圆锥体放在上一个圆锥体的顶部，重复，再重复，…… 最后，看看你自己的伪球吧！

5.3 伪球的共形映射

作为创建类似于欧氏平面的无限大双曲面地图的第一步，我们现在在 $\mathbb{C}$ 中构造伪球的共形映射。在我们的地图中，选择角度 x 作为水平轴，以便伪球的曳物线用垂直线表示。见[5.4b]。因此，伪球上具有坐标 (x, σ) 的点将在地图中由具有笛卡尔坐标 (x, y) 的点表示，我们将认为它是复数 $z = x + iy$。

如果我们的地图以最一般的方式构建，那么可以简单地选择 $y = y(x, \sigma)$ 作为 x 和σ的任意函数。但现在让我们坚持地图是共形的。因此，伪球上的无穷小三角形被映射到地图中

66 在实践中，你可能会发现简单地切割径向狭缝，然后将纸张重叠以创建圆锥面更容易。

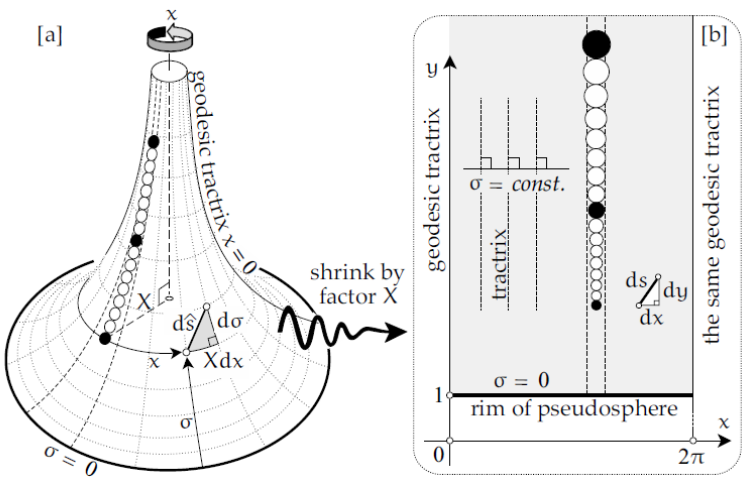
相似的无穷小[67]三角形，更普遍的是，伪球上的任何小形状在地图中看起来都是相同的（只是更大或更小）。在确定了这样的共形地图之后，结果发现（实际上）在选择 y 坐标方面没有自由。让我们看看为什么。首先，曳物线 $x = \text{const.}$ 正交于圆形横截面 $\sigma = \text{const.}$，所以它们在我们的共形地图中的像也是如此。因此，$\sigma = \text{const.}$ 的像必然用水平线$y = \text{const.}$表示。由此，我们推断 $y = y(\sigma)$ 必然仅是σ的函数。

其次，在伪球上，考虑连接点 (x, σ) 和 $(x + dx, \sigma)$ 的圆弧 $\sigma =$ 常量（半径 X）。根据 x 的定义，这些点对圆心的角为 dx，因此它们在伪球上的间隔为$X\,dx$（如图所示）。在地图中，这两个点具有相同的高度，并由距离 dx 分隔。因此，在从伪球传送到地图时，这个特定的线段被因子 X 缩小。

然而，由于映射是共形的，因此从 (x, σ) 发出的在任何方向上的无穷小线段必然乘以相同的因子$(1/X) = \frac{1}{R}\, e^{\sigma/R}$。换句话说，度量为

$$d\hat{s} = X\, ds$$

第三，考虑[5.4a]中所示的伪球上最上面的黑色圆盘。把这个圆盘想象成无穷小的，比如直径ε。在地图中，它将由另一个圆盘表示，其直径(ε/X)可以更生动地解释为在伪球轴上相同高度的观察者所看到的原始圆盘的角度宽度。现在假设我们反复将原始圆盘平移到伪球的边缘，每次将其移动一段距离 ε。图[5.4a]显示了由此产生的全等的依次相切的圆盘链。当圆盘沿伪球面向下移动时，它从轴退去，因此从轴看到来角度宽度减小。因此，地图中的图像盘在向下移动时似乎逐渐缩小，而连续的黑色盘之间的相等距离 ε 在地图中肯定不相等。



[5.4]伪球的共形映射。首先，选择围绕轴的角度作为 x 坐标。然后，由共形性导致无穷小三角形的相似性决定 y 坐标：$\frac{dy}{d\sigma} = \frac{1}{X}$。

在对地图的工作原理有了感觉之后，让我们实际计算与伪球上的点 (x, σ) 相对应的 y

67 在这里，我们将更喜欢直观的缩写“无穷小”，而不是拼写严格的“最终相等”。

坐标。根据上述观察（或直接从所示三角形相似的要求），可以得出

$$\frac{dy}{d\sigma} = \frac{1}{X} = \frac{1}{R} e^{\sigma/R} \quad \Longrightarrow \quad y = e^{\sigma/R} + \text{常数}$$

此常量的标准选择为 0，因此

$$y = e^{\sigma/R} = (R/X) \tag{5.4}$$

因此，整个伪球在地图中由位于线 $y = 1$ 上方的着色区域（其本身表示伪球的边缘）表示，与地图关联的度量为

$$\boxed{d\hat{s} = \frac{R\, ds}{y} = \frac{R\sqrt{dx^2 + dy^2}}{y}} \tag{5.5}$$

为了将来使用，还请注意，地图中具有边 dx 和 dy 的无限小矩形表示伪球上具有边 $(R\, dx/y)$和$(R\, dy/y)$的相似的无限小矩形。因此，地图中的表观面积 $dx\, dy$ 与伪球上的真实面积 $d\mathcal{A}$ 相关

$$\boxed{d\mathcal{A} = \frac{R^2\, dx\, dy}{y^2}} \tag{5.6}$$

（当然，这只是（4.12）的一个特例，$A = B = (R/y)$。）

5.4 贝尔特拉米-庞加莱半平面

我们现在有了一个圆柱状、有边的伪球的共形地图：$\{(x, y): 0 \le x < 2\pi, y \ge 1\}$。为了创建一个无限大双曲面的地图，贝尔特拉米知道他必须删除这两个形容词。请注意，虽然 R 的不同选择在数量上产生不同的几何形状，但它们在质量上都是相同的，因此做出特定的选择并没有坏处：

在基本上所有关于双曲几何的书籍和论文中，都做出了具体的选择 $R = 1$，这样 $\mathcal{K} = -1$。在本节中，我们也将遵循这一传统选择。

如果你希望从这个特定情况回到一般情况，你只需要在特殊情况（$R = 1$）公式中插入 R 的适当幂。例如，如果你处理的是面积，则必须乘以 R^2。

要删除“圆柱状”形容词，想象一下用标准的圆柱状油漆辊（单位半径）涂墙。旋转一圈后，你已经绘制了一条宽度为 2π 的墙条，并且辊子表面上的每个点都被映射到这个平墙条内的一个唯一点。要粉刷整面墙，你可以简单地继续滚动！现在想象一下，我们的油漆辊采取了伪球的形式。要使它匹配到平坦的墙面上，你必须首先根据度量（5.5）拉伸它的曲面，但是，就像以前一样，你可以继续滚动（比如说水平滚动）。例如，如果一个粒子沿着墙面上的水平线移动，则伪球上的对应粒子绕着水平圆$\sigma =$ 常量 转圈。“圆柱状”形容词已

经成功去除[68]，我们现在有了地图 $\{(x,y): \infty < x < \infty, y \geq 1\}$。

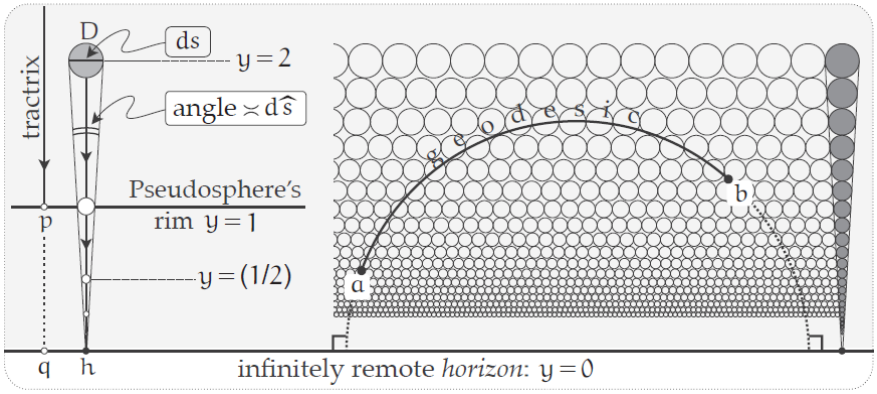
第二个问题，伪球的边缘，同样可以很容易通过共形映射解决。[5.5]的左侧是粒子在伪球面上沿着曳物线向下移动的图像。当然，在伪球面上，旅程在边缘$(\sigma = 0)$上的某个点$\hat{p}$被粗暴地中断，对应于线$y = 1$上的点 p。但在地图中，这个点 p 和其他点一样，绝对没有什么能阻止我们继续一直到$y = 0$上的点 q，真实距离 $d\hat{s}$ 继续由**标准双曲度量**给出，

$$\boxed{d\hat{s} = \frac{ds}{y}} \tag{5.7}$$

为什么停在 q？答案是，粒子甚至永远不会走那么远，因为 q 离 p 是无限远的!考虑[5.5]左侧显示的线$y = 2$上直径ds的小圆盘 D。它在伪球上的真实大小是$d\hat{s} = \frac{ds}{y}$，这最终等于它在线$y = 0$上正下方 h 处的图示角度。现在想象一下 D 以稳定的速度沿着伪球面移动。它在地图中的表观大小必须缩小，以便它在 h 处对置一个恒定的角度。在地图中，它碰到$y = 1$ ... 然后继续走！

假设从$y = 2$到$y = 1$需要一个时间单位，那么在下一个时间单位中，它将达到$y = (1/2)$，然后$y = (1/4)$，。..，因为这些点都被相同的双曲距离分隔：

$$\ln 2 = \int_1^2 \frac{dy}{y} = \int_{1/2}^1 \frac{dy}{y} = \int_{1/4}^{1/2} \frac{dy}{y} = \cdots$$



[5.5]圆盘D的双曲线直径是它在地平线上对着的（欧氏）角度。因此，当它向下移动时，它在地图中的像会缩小。右边的圆盘都是相同的大小，因此测地线 ab 穿过最小数量的圆盘。

因此，在地图内中看，运动减慢了，每个连续的时间单位只将到 $y = 0$ 的距离减半，因此 D 永远不会到达（这种现象的合适名称可能是“芝诺的复仇”!）。

最后，我们现在有了一个具体的模型

68 Stillwell (1996) 指出，这可能是数学中第一次出现拓扑学家现在所说的*万有覆盖*（universal cover）。

$$\boxed{\textbf{双曲平面}\ \mathbb{H}^2\text{：整个着色半平面}\ y>0\text{，度量}\ d\hat{s}=\frac{ds}{y}\text{。}} \tag{5.8}$$

实轴$y=0$上的点与普通点相距无限远，（严格来说）不被认为是双曲面的一部分。它们称为理想点，或无穷远点。无穷远点构成的直线 $y=0$ 称为地平线。

尽管贝尔特拉米在 1868 年发现了这张图（领先庞加莱 14 年），它现在普遍称为庞加莱半平面。但是，为了恢复一些历史的平衡，我们坚持将这张图称为贝尔特拉米–庞加莱半平面。

让我们尝试使这张地图的度量更加生动。在[5.5]的最右边是一串相同双曲大小 ε 的接触圆，如[5.4]所示。在左边，我们用这样的圆填充了双曲面的一部分，所有的圆都有相等的双曲直径 ε。因此，任何曲线的双曲线长度最终等于它截取的圆数乘以 ε。这清楚地表明，从 a 到 b 的最短路线是截取最小圆数的路线，因此具有所示的近似形状。

如果你遵循了我们之前的建议并制作了自己的伪球模型，你还可以通过在相似高度的两点之间拉伸一根细绳来检查测地线的形状。要研究无法拉伸细绳的区域的测地线，例如沿着曳物线本身，你可以使用窄胶带结构（1.7），它在任何地方都能工作。

我们的下一个任务是确认[5.5]中所示的可爱事实：这样的测地线的确切形式是一个完美的半圆，它以直角与地平线相交。唯一不是这种形式的测地线是垂直半线（曳物线），但即使是这些也可以视为半圆半径趋于无穷大的极限情况。

5.5 用光学找到测地线

为了解释双曲几何的贝尔特拉米-庞加莱半平面模型中测地线的半圆形形式，我们将借鉴物理学的思想，特别是光学[69]的思想。我们的灵感来自于 1662 年发现的一项名为**费马原理**[70]的定律：

$$\boxed{\text{光线传播的路径是需时最少的路径。}} \tag{5.9}$$

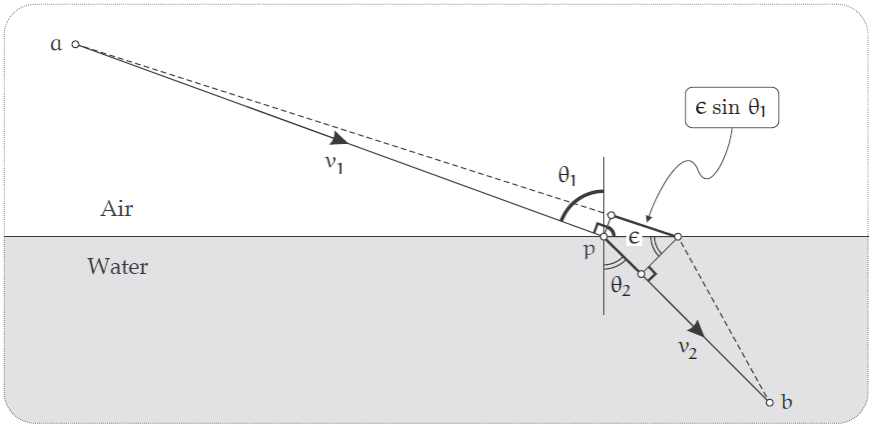
我们将使用牛顿推理[71]开始我们对物理学的短暂考察，看看费马原理如何允许我们在几

69 这种方法似乎并不广为人知。 感谢 Sergei Tabachnikov 向我指出它之前由 Gindikin 出版（2007 年，第 324 页）。 使用费马原理找到一条使某个量最小化路径的基本思想可以追溯到约翰·伯努利在 1697 年对最速降线问题的解。

70 首先，这就是以数论发现而闻名的皮埃尔·德·费马（1601-1665），包括费马大定理。其次，费曼发现这个原理有一个漂亮的量子力学解释，在费曼（1985）中可以找到一个精妙的解释。

71 我们很荣幸地意识到，在这里我们只是追溯了费曼所走的路：见费曼等（1963 年，

何上确定光从空气进入水（例如）时发生的突然弯曲（称为折射）。这就是为什么当你把勺子伸进一杯茶里时，你的勺子看起来会弯曲的原因。



[5.6]斯涅尔定律：为了使光沿着相邻的两条路线传播所需的时间相等，空中的额外时间必须由水中减少的时间精确补偿：$\sin\theta_1/v_1 = \sin\theta_2/v_2$。

在[5.6]中，一束光线从 a 点发出，以v_1的速度穿过空气，在 p 点处与垂直方向成θ_1角度入射入水。然后它被折射到θ_2方向，以较慢的速度v_2穿过水中，最终到达 b 点。早在公元 130 年，托勒密就进行了这样的实验，并编制了一个相当准确的角度θ_1到θ_2的对应表。但是，托勒密却没能得到这两个角度之间精确的数学关系，在接下来数个世纪里，科学家们依然没能得出。

最后，1621 年，荷兰数学家威里布里德·斯涅尔·范罗延（1591－1626 年）发现了正确的定律，现在被普遍称为**斯涅尔定律**[72]（Snell's Law）：

$$\boxed{\sin\theta_1 = n\ \sin\theta_2\text{，其中 } n = \text{const.}} \tag{5.10}$$

n 的值（称为折射率）取决于界面两侧的材料，但对于空气/水，$n \approx 1.33$。

至少在定性上，应该清楚的是，费马原理要求光弯曲。因为如果它从 a 到 b 直线飞行，那么它将浪费宝贵的时间在水中相对缓慢地飞行，而不是在空中快速飞行。从数量上讲，当时间的导数（相对于 p 的位置）等于零时，将发生最小化飞行时间所需的弯曲量。

从几何上讲，如果 p 的位置使时间最小化，那么 p 的无限小位移ε在时间中不应该产生

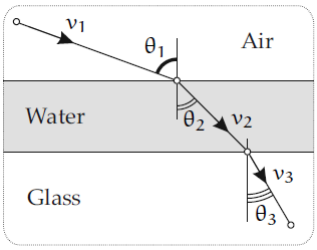
第 1 卷，第 26-3 期）。费马自己首先给出了一个解析证明，后来又给出了一个几何证明，但都没有现在的牛顿论证那么优雅。费马的两个证明都可以在 Mahoney (1994, pp. 399－401) 中找到。

72 像往常一样，历史远比名字所暗示的要复杂得多。我们之前遇到的同一位托马斯·哈里奥特（Thomas Harriot）也发现了“斯涅尔定律”，比斯涅尔早 20 年，但与他的大多数发现一样，哈里奥特是保密的，尽管他确实与开普勒沟通了这个结果。但即使是哈里奥特也被打败了——被打败了 600 年！伊斯兰数学家和物理学家伊本萨尔（Ibn Sahl）于公元 984 年发表了它，甚至用它来设计复杂的变形透镜。

变化（ε的一阶）。但是，正如我们在[5.6]中所看到的，这种位移导致光在空气中传播额外的距离 $\varepsilon \sin\theta_1$，使飞行时间最终增加 $(\varepsilon \sin\theta_1)/v1$。另一方面，根据同样的推理，在水中节省的时间等于$(\varepsilon \sin\theta_2)/v2$。要使净时间变化消失，这两个单独的时间变化必须相等。因此，通过约去ε，可得，

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2} \tag{5.11}$$

这不仅证明了斯涅尔定律（5.10），而且还作出了物理预测，并已经直接实验证实：折射率是两种材料中光速的比值，$n = (v_1/v_2)$。



[5.7] 斯涅尔定律应用于多层材料。

现在假设我们的水位于饮水杯的底部（见[5.7]）。当光线到达水底并进入玻璃底座时，它的速度是 v_3，它将如何弯曲？适用与以前相同的定律，

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2} = \frac{\sin\theta_3}{v_3}$$

因此，更普遍的是，如果我们有许多薄的水平材料条，光速在每个条内都是 vi 的，那么光穿过层的整个旅程都受定律的控制

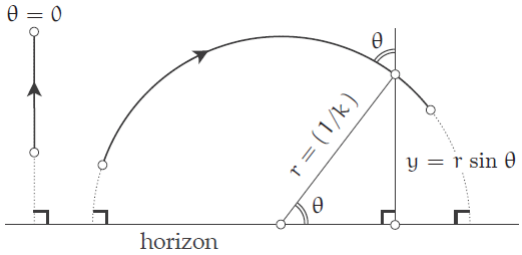
$$\frac{\sin\theta_i}{v_i} = \text{const.} = k$$

最后，让我们更进一步，想象光通过一个不均匀材料块，其密度在每个水平切片上都是相同的 $y =$ 常量，但它随着 y 的变化而连续变化，高度 y 的光速为v(y)，那里的光的角度为θ(y)。那么广义斯涅尔定律是

$$\boxed{\frac{\sin\theta(y)}{v(y)} = k} \tag{5.12}$$

你说，这一切都很有趣，但这与证明双曲平面中测地线是半圆形有什么关系？！好吧，假设伪球上的点 $\hat{a}$ 和 $\hat{b}$ 对应于 $\mathbb{H}^2$ 中的点 a 和 b（参见[5.5]）。想象一下，一个粒子沿着伪球表面从 $\hat{a}$ 到 $\hat{b}$ 的不同路线移动，但总是以相同的恒定速度移动，比如 1。当粒子沿着这些路径之一走时，它在双曲平面中的像将沿着从 a 到 b 的对应路径走，但不是以匀速的速

度走：正如我们在[5.5]中看到的，如果一个粒子以恒定的速度沿着伪球面向下移动，它在双曲平面中的像会减慢。



[5.8]双曲测地线必然满足广义斯涅尔定律$(\sin\theta/y) = k$，因此是与地平线正交的半圆和半线。

下一个关键的观察是，由于地图是共形的，减速只取决于粒子的位置，而不是它移动的方向。因为假设我们同时从$\hat{a}$向各个方向发射大量单位速度的粒子。在无限小的时间ε之后，它们将形成一个半径ε的圆，在伪球上以$\hat{a}$为中心，而在双曲平面（5.8）中，它们也将形成一个圆（以 a 为中心），但半径为εy，其中 y 是 a 的高度。换句话说，从 a 发出的粒子的速度是$v(y) = y$：我们越接近地平线，粒子移动得越慢。

当然，伪球面上从$\hat{a}$到$\hat{b}$每次旅行所需的时间将与双曲平面上沿 ab 移动所需时间相同。但在伪球上，测地线路线是最短的，也将是时间最少的路径，因此双曲平面中的测地线也是从 a 到 b 最快的路线：双曲平面中的测地运动自动服从费马原理，因此它的形状由广义斯涅尔定律决定！将 $v(y) = y$替换为（5.12），答案突然在[5.8]中变得清晰：

> 贝尔特拉米–庞加莱半平面模型 $\mathbb{H}^2$ 中的测地线满足 $(\sin\theta/y) = k$。如果$k \neq 0$，这是半径 $r = (1/k)$ 的半圆，以地平线为中心。如果$k = 0$，则它是垂直半线 $\theta = 0$。 (5.13)

稍后，在第 11.7.5 节中，我们将基于角动量对这个重要事实进行第二次物理解释！

5.6 平行的角度

现在我们掌握了双曲平面中测地线的形式，我们可以回到起点，在[5.9]中视觉上确认双曲公理（1.1）的真实性。确实有无限多的线（显示虚线）经过 p 的直线不与 L 相交。这样的直线称为超平行于 L。

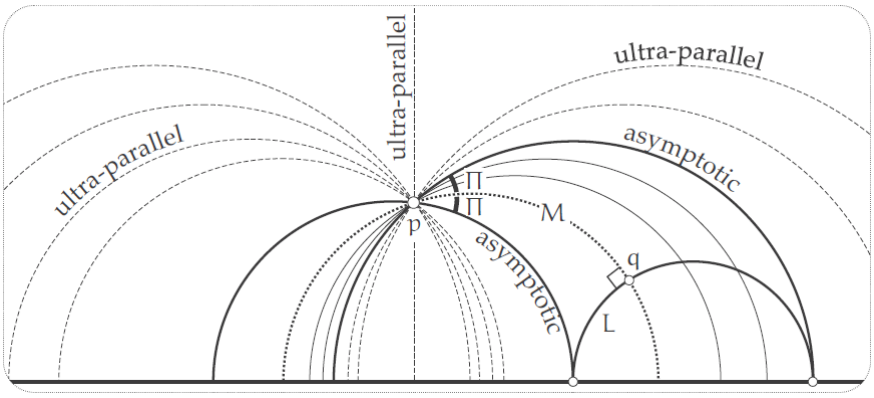
将超平行线与与 L 相交的线分开，我们看到，在双曲平面内的任何地方，正好有两条线不能与 L 相交，但它们确实在地平线上与 L 相交。这两条线称为渐近的[73]（asymptotic）。

就像欧氏几何一样，图中清楚地表明，有一条线 M（虚线）穿过经过 p 以直角（比如 q）切割 L。M 的存在使得可以以通常的方式定义点 p 与线 L 的距离，即作为 M 段 pq 的（双

73 另一个常用的名称是平行的。

曲线，$d\hat{s}$）长度 D。

事实上，如图所示，M 平分了渐近线在 p 处所张的角（尽管目前这还远不明显）。M 和任何一条渐近线之间的角称为平行角，通常用Π表示。当你围绕 p 旋转直线 M 时，它与 L 的交点向无穷大移动，Π告诉你在 M 开始完全错过 L 之前，可以旋转多少。



[5.9]双曲公理的显式呈现：有无穷多的线（超平行）（虚线显示）经过 p 且不与给定直线 L 相交。

正如罗巴切夫斯基和鲍耶都发现的那样，这个角度 Π 与 p 到 L 的双曲距离 D 之间存在不同寻常的联系：

$$\boxed{\tan(\Pi/2) = e^{-D}} \tag{5.14}$$

这通常称为**鲍耶–罗巴切夫斯基公式**。

因此，如果 p 接近 L，那么 $\Pi \approx (\pi/2)$，欧氏几何结果几乎成立：从 p 发出的大约一半的射线（半线）最终击中 L。当然，在欧氏几何中，无论 p 离 L 有多远，都有正好一半的射线击中 L。但在双曲平面中，我们看到——定性地在[5.9]和定量地在（5.14）中——当 p 从 L 远离时，从 p 发出的射线击中 L 的比例缩小到零！

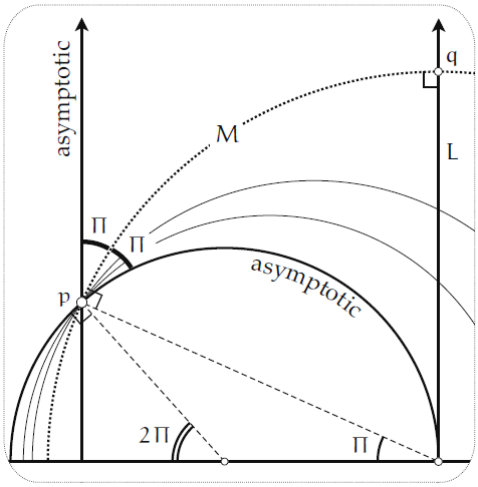
必须认识到，从伪球面或真正的双曲平面的微观居民的角度来看，没有办法区分测地线——每一条直线都是一模一样的。因此，从本质上讲，在我们的地图中由垂直半线表示的测地线与半圆表示的测地线完全无法区分。

但是，半圆在地平线上有两端，而垂直线似乎只有一端，这又如何解释呢？答案是，除了地平线上的点，在无穷远还有一个点，所有的垂直线都在那里交汇。根据（5.8），当我们沿着两条相邻的垂直线向上移动时，它们之间的距离消失为（1/y），它们会聚到无穷远的单点；这在伪球上尤其生动。

在强调了地图中线条的两种表现在数学上是相同的之后，我们现在做了一个转变，同时也强调了它们在心理上是不相同的。也就是说，站在这个非欧几里德世界之外，通过我们的地图往里看，如果我们看一个简单（但不那么典型）的情况，一条线是垂直的，而不是半圆形的，我们可能会发现一些数学关系成立。

赋予这个想法真正力量的是 H2 的刚性运动的存在，比如围绕一个点的旋转。这种保距运动称为等距运动，它们是下一章的主题。目前，我们注意到，等距的存在使任何涉及半圆线的双曲线图形都有可能刚性移动，使其在地图中显示为垂直线。

例如，回到[5.9]，我们可以围绕 p 旋转$\mathbb{H}^2$，直到 L 变为垂直，在这种情况下，得到[5.10]所示的更简单的形式。一个直接的收获是，现在我们可以看到 M 平分了 p 处的渐近线之间的角：只需验证所有标记的角度[练习]。但这意味着，在我们进行旋转之前，当 L 在[5.9]中处于一般位置时，M 一定的确已经将该角平分了。



[5.10] 与[5.9]相同的几何，但将双曲平面绕 p 旋转直到 L 显示为垂直。这张简化的图片可以用来导出鲍耶—罗巴切夫斯基公式，（5.14）。

同样，这张新图的更简单的几何形状也使确认鲍耶—罗巴切夫斯基公式的真实性变得容易，（5.14）；详情请见《可视化复分析》，第 306 页。

5.7 贝尔特拉米-庞加莱圆盘

具有度量 $d\hat{s} = ds/y$ 的贝尔特拉米-庞加莱上半平面只是描绘抽象双曲面 $\mathbb{H}^2$ 的一种方式。存在[74]着几种替代模式。虽然这些不同的模式从定义上讲，本质上都是相同的，但它们在心理上并不相同：在一个模式中很难看到一个特定的事实或公式，但在另一个模式中却很透明。因此，当试图掌握双曲几何的奇迹时，在模型之间切换的便利是一项宝贵的技能。

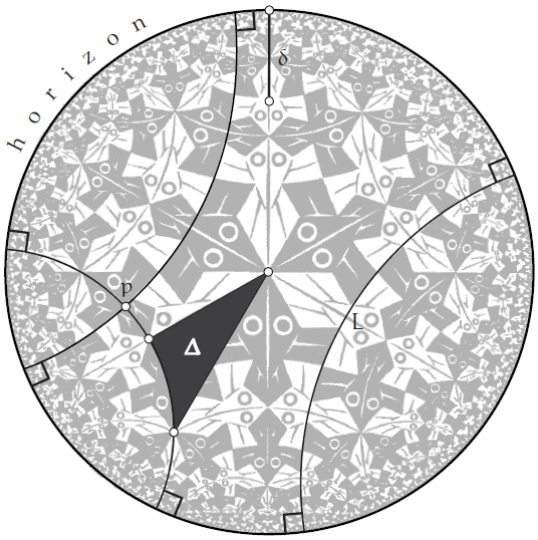
目前，我们只想说明一个特别有用和著名的模式。新模型绘制在单位光盘中；请参见[5.11]。和上半平面一样，这是一个共形模型，测地线再次表示为以直角与地平线相交的圆弧，但现在无限遥远的地平线是这个圆盘的边界，单位圆。如果 r 是距离圆盘中心的距离，则新的度量公式为（见练习 25）

$$\boxed{d\hat{s} = \frac{2}{1-r^2}ds} \tag{5.15}$$

74 有关完整说明，请参阅 VCA 或 Stillwell (2010)。

有关更多详细信息，我们请你访问 VCA 或斯蒂尔韦尔（2010 年）。你至少可以通过使用共形曲率公式（4.16）来验证[练习]此曲面具有恒定的负曲率 $\mathcal{K} = -1$ 来确认这确实是 $\mathbb{H}^2$。

贝尔特拉米首先发现了这个模型，在 1868 年与半平面模型同一篇论文中公布；见斯蒂尔韦尔（1996 年）。14 年后，庞加莱再次发现了这个模型，它再次被普遍称为“庞加莱圆盘”。然而，和以前一样，我们坚持将其称为贝尔特拉米-庞加莱光盘，以使两位都得到应得的荣誉；较少争议的是，它也称为共形圆盘模型。



[5.11] 贝尔特拉米-庞加莱双曲平面的圆盘模型。背景是埃舍尔的作品圆极限 I；叠加的是双曲直线，它们是与无穷远的边界圆（地平线）正交的直径和圆弧。显然，双曲公理（1.1）得到满足，$\mathcal{E}(\Delta) < 0$。

1958 年，著名的英国几何学家 H.S.M.考克斯特（1907－2003 年）向荷兰艺术家 M. C.埃舍尔（1898－1972 年）介绍了 $\mathbb{H}^2$ 的共形圆盘模型。这导致埃舍尔创作了他著名的圆极限木刻[75]系列，其中的第一幅在 [5.11] 中水印复现（故意做的若隐若现），但双曲直线与原作品显眼地叠加在一起。（这张图的想法直接来自于彭罗斯 (2005 年，图 2.12）。例如，我们看到，根据双曲公理（1.1），通过 p 确实有无限多的线不满足 L。圆的直径也是双曲线，因此图示的三角形 Δ 是一个真正的双曲三角形。请注意，$\mathcal{E}(\Delta) < 0$ 是显而易见的（而且很容易证明），因为它应该是。

当你盯着这个看的时候，试着把自己想象成其中的一条鱼。你的大小和形状和其他鱼完全一样，你可以永远在一条直线上游泳，而不会看到你周围的环境或你的同伴的任何变化。但是从外面看地图，距离的压缩让你看起来像是沿着一条圆弧形的路径转弯，而且在你游的同时身形也在缩小。事实上，如果 $\delta \equiv (1 - r)$ 是鱼与地平线的欧式距离，我们看靠近地图边缘的一点，我们发现（5.15）意味着 $\left(\text{鱼的表观大小}\right) \propto \delta$ [练习]。

75 可以在网上找到考克斯特（Coxeter）本人讨论这些埃舍尔（Escher）构造的的简短数学视频。

在练习 25 中，你将看到实际上有一个简单的共形转换，它将这个新的圆盘模型与我们以前的共形半平面模型联系起来，从而解释了新圆盘模型的共形性。计算机将这种转变应用于埃舍尔的原始艺术作品[5.11]，产生了埃舍尔本人从未创作过的一幅画，但他肯定会欣赏的。（本图片经许可转载自斯蒂尔韦尔(2005 年，第 195 页）。)

[5.12] 埃舍尔圆极限 I（约翰・斯蒂尔韦尔教授）从原来共形圆盘模型[5.11]转换为共形半平面模型。

让我们停下来，喘口气，回顾一下我们已经走了多远。我们开始这本书的故事[76]已经有了一个愉快的结局，某种意义上的结局。2000 年来，欧几里得几何一直被围绕平行公理的困惑和怀疑所困扰。现在，贝尔特拉米对双曲几何作为一种合法替代方案的具体辩护，终于从两千年的数学紧张中带来了清晰和宣泄。结束第二幕的好地方啊！

或者不是…

76 有关完整的传奇，请参见 Gray (1989)。

第六章 等距变换与复数

6.1 介绍

伟大的数学思想不仅平息过去的谜团，还揭示了新的谜团：光明尽头的隧道！我们现在正要一窥这些与其它数学和物理领域的奇怪新联系。

这些新的神秘联系中的第一个发生在三个常曲率几何（欧氏、球面和双曲）与复数之间。

我们已经要求你想象我们的曲面地图是在复平面中画出的。然而，精明的读者会注意到，到目前为止，我们只对复数的基本结构进行了少量使用。那是关于改变的。然而，首先，我们需要对等距变换的概念进行一些更普遍的审视。

等距变换（Isometry，简称等距）必然保持每个角度的大小，而那些也保持角度方向的（顺时针与逆时针）称为正向的（direct），而那些反转方向的称为反向的（opposite）。因此，正向等距是一种非常特殊的共形映射，而反向等距是一种非常特殊的反共形映射。例如，在平面中，旋转是正向等距，而关于直线反射是反向等距。

接下来，我们观察到，在组合的操作下，

> 给定曲面 $\mathcal{S}$ 的一组完备的等距集合（正向和反向）自动具有**群**结构，记为$\mathcal{G}(\mathcal{S})$。

为了确认这一点，设e =（不做任何事情），设 a、b、c 是 $\mathcal{S}$ 的任意三个等距。可见群公理[77]得到满足：

- 显然 e 保持距离，因此$e \in \mathcal{G}(\mathcal{S})$，并且由于$a \circ e = a = e \circ a$，可推断 e 是群的单位元。
- 若先应用 a，再应用 b（两者都保持距离），那么净变换也保持距离：$b \circ a \in \mathcal{G}(\mathcal{S})$。
- 由于变换 a 保持距离，它的逆变换也保持距离：$a^{-1} \in \mathcal{G}(\mathcal{S})$。
- 变换的组合（不一定是等距）满足结合律：$(a \circ b) \circ c = a \circ (b \circ c)$。

注意，（在组合下）正向和方向的等距在乘法下的行为类似于 $(+)$ 和 $(-)$：$(+)(+) = (+)$、$(+)(-) = (-)$ 和 $(-)(-) = (+)$。 它遵循

> 正向等距构成完整群 $\mathcal{G}(\mathcal{S})$ 的**子群** $\mathcal{G}_+(\mathcal{S})$。

另一方面，反向等距根本不构成群。 但它们确实属于完整群 $\mathcal{G}(\mathcal{S})$，那么它们与 $\mathcal{G}_+(\mathcal{S})$

[77] 这些定义了群的数学概念。 即使你以前没有遇到过这个概念，你仍然可以通过简单地接受以下公理作为定义来跟上。

有什么关系呢？

令 ξ 是一个特定的、固定的反向等距； 那么 ξ^{-1} 也是。 让 ζ 是一个一般的反向等距——想想这个在所有可能的反向等距上的变化。 然后 $\xi^{-1}\circ\zeta\in\mathcal{G}_+(\mathcal{S})\Longrightarrow\zeta\in\xi\circ\mathcal{G}_+(\mathcal{S})$。同理，$\zeta\in\mathcal{G}_+(\mathcal{S})\circ\xi$。 因此，

> 如果 ξ 是任意反向等距，则反向等距的完备集合是$\xi\circ\mathcal{G}_+(\mathcal{S})=\mathcal{G}_+(\mathcal{S})\circ\xi$，因此完整对称群是
>
> $$\mathcal{G}(\mathcal{S})=\mathcal{G}_+(\mathcal{S})\cup[\xi\circ\mathcal{G}_+(\mathcal{S})]=\mathcal{G}_+(\mathcal{S})\cup[\mathcal{G}_+(\mathcal{S})\circ\xi]$$

(6.1)

是否每个曲面 $\mathcal{S}$ 都拥有一个非平凡的等距群 $\mathcal{G}(\mathcal{S})$？ 不，因为等距变换还必然保持曲率。 假设一个等距变换将 p 处一个非常小的（最终趋于零的）三角形 Δ 搬运到 p' 处的全等三角形 Δ'。那么

$$\mathcal{K}(\mathrm{p})\asymp\frac{\mathcal{E}(\Delta)}{\mathcal{A}(\Delta)}=\frac{\mathcal{E}(\Delta')}{\mathcal{A}(\Delta')}\asymp\mathcal{K}(\mathrm{p}')$$

因此，像[1.9]中描述的南瓜这样的不规则曲面将没有任何等距[78]。

另一方面，对于等距的存在，曲率不必在 $\mathcal{S}$ 上保持恒定[79]。例如，考虑任何旋转曲面。就其结构而言，这个（通常）非常数曲率的曲面确实允许一个等距群。围绕曲面轴的旋转是正向等距，对穿过该轴平面的反射是反向等距；也可能存在其它等距。

$\mathcal{S}$ 的对称性越强，等距群就越大，最强的对称性发生在三种常曲率情况下：$\mathcal{K}=0$、$\mathcal{K}>0$ 和 $\mathcal{K}<0$。外在地看，具有这些几何形状的原型曲面是欧氏平面、球面和伪球面。然而，等距的概念属于内蕴几何；因此，例如，$\mathbb{H}^2$ 的贝尔特拉米-庞加莱半平面地图实际上比伪球更好地描述了双曲几何，并且正是这张地图（或共形盘模型）将与下面的内容相关。

我们将很快完善这个陈述，但简言之，但这里简要介绍这三种最强对称几何和复数之间的惊人联系——**主要结果：**

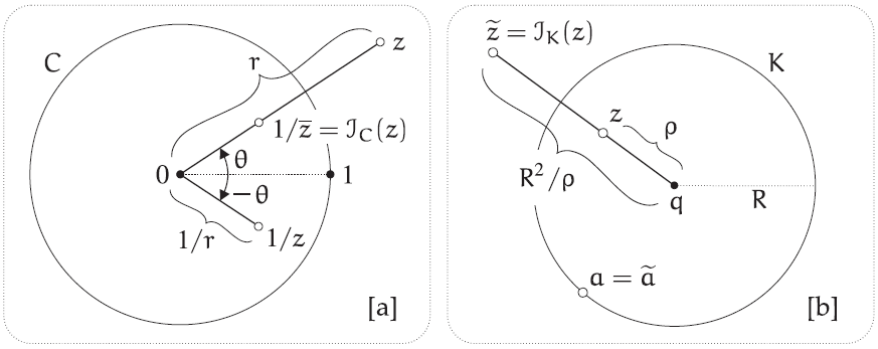
> 所有三个常曲率几何都有对称群 $\mathcal{G}_+(\mathcal{S})$（正向等距），它们是复平面上默比乌斯变换群的子群，$\mathrm{z}\mapsto\mathrm{M}(\mathrm{z})=\frac{\mathrm{az+b}}{\mathrm{cz+d}}$，其中 a、b、c、d 是复常数。

(6.2)

---

[78] 我还没有仔细考虑过如何量化排除等距存在可能性所需的不规则程度。 一个起点可能是观察如果 $\mathcal{K}(\mathrm{p})$ 的值在 p 处唯一获得，则 p 必然是任何假定等距的不动点——它无处可去！ 三个这样的点将导致三个不动点，我猜这将排除存在非平凡等距的可能性。

[79] 我们将关注无限的、连续的等距集合，但也可能有有限的等距集合，即使对于光滑的曲面也是如此。 例如，考虑一个游戏骰子的对称性（一个棱角圆润的立方体）。



[6.1] [a]**复反演**是几何反演和共轭的组合。[b]对一般圆的几何反演。

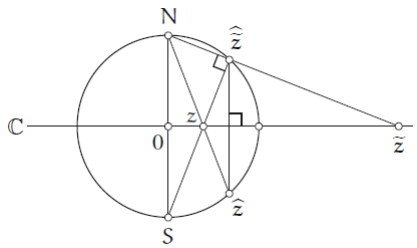
6.2 默比乌斯变换

仅上述结果就可以说服你，默比乌斯[80]变换在现代数学中非常重要（正如我们将看到的，在物理学中也是如此）。我们现在总结[81]一些关于这些变换我们需要知道的事实。

- **分解为简单的变换**。第一步我们把 $z \mapsto M(z) = \frac{az+b}{cz+d}$ 分解为以下的一串变换[练习]：

$$\left.\begin{array}{ll} \text{(i) } z \mapsto z + \dfrac{d}{c}, & \text{这是一个平移；} \\ \text{(ii) } z \mapsto \left(\dfrac{1}{z}\right), & \text{称为复反演；} \\ \text{(iii) } z \mapsto -\dfrac{(ad-bc)}{c^2} z, & \text{这是一个伸缩和一个旋转；} \\ \text{(iv) } z \mapsto z + \dfrac{a}{c}, & \text{这是另一个平移.} \end{array}\right\} \tag{6.3}$$

注意，若 $(ad - bc) = 0$，则 $M(z)$ 将整个复平面挤压到单个像点（a/c）；在这个例外情况下，$M(z)$ 无法撤消，称为奇异的。所以，在讨论默比乌斯变换时，我们将始终假设 $M(z)$ 是非奇异的，这意味着它是可逆的（即 $(ad - bc) \neq 0$）。



[6.2] 黎曼球的垂直横截面。黎曼球面对其赤道平面 ℂ 的反射发送 $\hat{z} \mapsto \hat{\tilde{z}}$ 和 $z \mapsto \tilde{z} = \mathcal{I}_C(z)$。

在上述四个变换中，只有第二个没有研究过。

---

[80] 也称为分式线性或双线性变换。

[81] 如需更深入的了解，请参阅 VCA 第 3 章和第 6 章。

- **对圆的反演**。这个映射 $z \mapsto (1/z)$ 是理解默比乌斯变换的关键。就像我们在 VCA 中所做的那样，我们将这种倒数映射称为*复反演*。在极坐标中，复反演下 $z = r\,e^{i\theta}$ 的像为 $1/(r\,e^{i\theta}) = (1/r)\,e^{-i\theta}$：新长度是原来的倒数，新角度是原来取负。见图[6.1a]。请注意单位圆 C 外的点如何映射到 C 内的点，反之亦然。图[6.1a]还说明了一种将复反演分解为两阶段过程的特别有效的方法：

  1. 把 $z = r\,e^{i\theta}$ 发送到与 z 方向相同但长度成为倒数的点，即点 $(1/r)\,e^{i\theta} = (1/\bar{z})$。
  2. 应用复共轭（即对实轴的反射），将 $(1/\bar{z})$ 发送到 $\overline{(1/\bar{z})} = (1/z)$。

  自己检查一下我们应用这两个映射的顺序是无关紧要的。（这并不典型：例如，在（6.3）中，顺序确实很重要。）

  虽然阶段（2）在几何上微不足道，但我们将看到阶段（1）中的映射充满了惊喜；它称为*几何反演*[82]，或简称为*反演*。显然，单位圆 C 在这种映射中扮演着特殊的角色：反演交换了 C 的内域和外域，而圆 C 上的点则不动（即映射到自身）。因此，我们将映射写为 $z \mapsto \mathcal{J}_C(z) = (1/\bar{z})$，我们称 $\mathcal{J}_C$ （比以前更精确一点）“对 C 的反演”。

  这种术语额外的精确性很重要，因为如图[6.1b]所示，可以很自然地把 $\mathcal{J}_C$ 推广为对任意圆 K（比如以 q 为中心、R 为半径）的反演。很明显，这个“对 K 的反演”（写成 $z \mapsto \tilde{z} = \mathcal{J}_K(z)$）应把 K 的内域和外域互换，而 K 上的点则不动。若 ρ 是从 q 到 z 的距离，则我们定义 $\tilde{z} = \mathcal{J}_K(z)$ 为这样的点，从 q 到它的方向与从 q 到 z 的方向相同，且它到 q 的距离为 (R^2/ρ)。如果我们将此图视为 [6.1a] 按因子 R 扩展而得，你自己检查一下这个定义是强加给我们的。

- **反演是黎曼球面的赤道反射**。如果我们将复数（用球极平面投影）从平面上吸到单位球面上，从而创建黎曼球面，那么反演的效果就会变得非常简单：

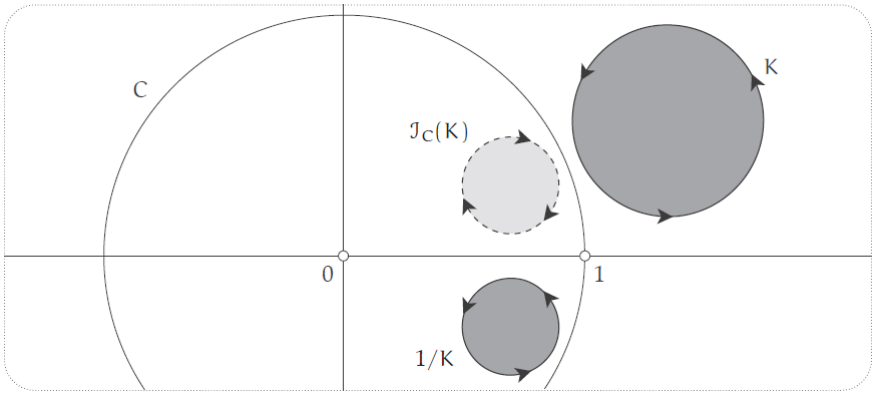
$$\boxed{\mathbb{C} \text{ 对单位圆的反演诱导了黎曼球面对其赤道平面 } \mathbb{C} \text{ 的反射。}} \tag{6.4}$$

要证实（6.4），请参见[6.2]，其中显示了黎曼球面的垂直横截面。点 z 被球极投影到 $\hat{z}$，穿过赤道平面反射到 $\hat{\tilde{z}}$，最后投影到$\tilde{z}$。我们看到三角形 $N0\tilde{z}$ 和 $z0N$ 相似[练习]，所以 $|\tilde{z}|/1 = 1/|z|$。因此，$\tilde{z} = \mathcal{J}_C(z)$，如声称的。

请注意，此图还显示，对单位圆的反演相当于先作南极的球极投影，再作是北极的球极投影（反之亦然）。

---

[82] 另一个常见且合适的名称是圆圈中的反射（原因将很快变得显而易见）。 在较早的作品中，它通常称为“倒数半径变换”。



[6.3]单位圆 C 中的复反演将定向圆 K 发送到 (1/K)，沿 K 行进方向左侧的阴影区域映射到沿 (1/K) 行进方向左侧的阴影区域。

- **反演保持圆**。根据（4.28）, $\mathbb{C}$ 上的圆 K 投影到 Σ 上的圆，球面关于其赤道平面作反共形反射（反演）将其映射到 Σ 上的另一个圆，它最终投影回 $\mathbb{C}$ 中的圆 $\mathcal{I}_C(K)$。见图 [6.3]。

  如果我们从一条直线开始，我们会在 Σ 得到一个通过 N 的圆（回忆 [4.21]），它通过 $\mathcal{S}$ 反射到一个圆，它投影回通过 0 的 $\mathbb{C}$ 中的圆。相反，由于反演互换点，通过 0 的圆 K 被发送到直线 $\mathcal{I}_C(K)$。见图 [6.4]。

我们可以将第二个结果视为第一个结果的极限情况，直线是圆的极限形式。事实上，在黎曼球面上，一条直线实际上是一个恰好穿过北极的圆。有了这种统一的语言，我们可以总结为，

$$\boxed{\text{反演是反共形的，将圆映射到圆。}} \tag{6.5}$$

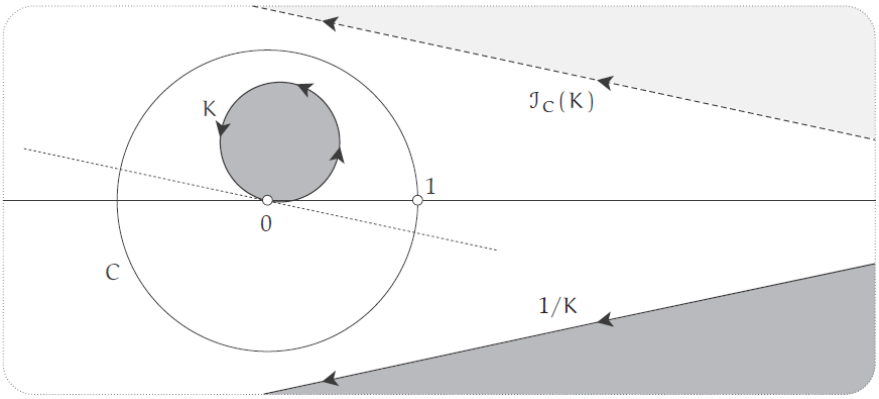
- **复反演是黎曼球面的旋转**。如果我们在这个反演之后执行一次共轭，$z \mapsto \bar{z}$，那么净效果是复反演。但是共轭对黎曼球面的影响是另一个反射，这次是对通过实轴的垂直平面上。你可以很容易地检查（或许在橘子的帮助下），这两次对穿过实轴彼此正交的两个平面反射的净效果是围绕该轴的旋转：

$$\boxed{\text{复反演 } z \mapsto (1/z) \text{ 是黎曼球面绕实轴旋转角度 } \pi\text{。因此它是共形的，并将圆映射到圆。}} \tag{6.6}$$

我们还可以用球极射影公式（4.26）给出第二个证明。首先请自行检查，如果 $\hat{z}$ 有坐标 (ϕ, θ)，然后围绕实轴旋转 π，将其带到坐标点 $(\pi - \phi, -\theta)$。所以这个旋转点对应于复数

$$\cot\left[\frac{\pi}{2} - \frac{\phi}{2}\right] e^{i(-\theta)} = \frac{1}{\cot(\phi/2)} e^{i(-\theta)} = \frac{1}{\cot(\phi/2)\, e^{i\theta}} = \frac{1}{z}$$

证毕。



[6.4] 关于单位圆 C 的复反演将定向圆 K 发送到 (1/K)，沿 K 行进方向左侧的阴影区域映射到沿 (1/K) 行进方向左侧的阴影区域。

图[6.3]说明了反演和复反演对定向圆 K 的影响。请注意，沿 K 行进方向左侧的深色阴影区域（K 的内域）是如何映射到沿 (1/K) 行进方向左侧的。如果 K 包含 0，则 (1/K) 的方向将被映射反转，但规则“左 ⟼ 左”仍然有效。请自行用直接地和使用（6.6）两种方法检验所有这些。

图[6.4]说明了在 K 通过 0（南极）情况下的相同现象，因此 (1/K) 是一条直线（通过北极的圆）。

请注意，这给出了半平面和圆盘内域之间的共形映射，实际上这正是从共形半平面模型 [5.12] 构建双曲平面的共形圆盘模型 [5.11] 的方式。（显式默比乌斯变换在练习 25 中给出，完整解释在 VCA 第 315 – 317 页中给出。）

- **角度和圆的保持**。如 [6.3] 和 [6.4] 所示，从（6.3）和（6.6）可以得出，使用新的广义的“圆”的方向，

默比乌斯变换是共形的，将每个定向圆 K 映射到定向圆 $\tilde{K}$，使得 K 左侧的区域映射到 $\tilde{K}$ 左侧的区域。	(6.7)

- **矩阵表示**。从几何学上讲，黎曼球面允许我们像其他任何点一样思考 ∞——它只是北极。从（6.6）中，我们看到复反演诱导了南北极交换的旋转，因此表述 $0 = 1/\infty$ 和 $\infty = 1/0$ 字面上是正确的。

  如果我们也能在代数层面上消除无穷大的例外角色，那就太好了。为此，我们采用了投影几何中的一个想法，将黎曼球面上的每个点描述为存在于 $\mathbb{C}^2$ 中的一对复数 $(\mathfrak{z}_1, \mathfrak{z}_2)$ 的比率：$z = (\mathfrak{z}_1/\mathfrak{z}_2)$。

  复数[83]的有序对 $[\mathfrak{z}_1, \mathfrak{z}_2]$ 称为 z 的射影坐标或齐次坐标。为了确保这个比率有良

[83] 遗憾的是，我们注意到，$[\mathfrak{z}_1, \mathfrak{z}_2]$ 也可以被视为 2-旋量（2-spinor）的坐标表示。 这个概念是罗

定义，我们要求 $[\mathfrak{z}_1, \mathfrak{z}_2] \neq [0,0]$。对于每个有序对 $[\mathfrak{z}_1$任意，$\mathfrak{z}_2 \neq 0]$，精确地对应一个普通的有限点 $z = (\mathfrak{z}_1/\mathfrak{z}_2)$，但是对于每个点 z，都对应着无限的射影坐标集，$[k\mathfrak{z}_1, k\mathfrak{z}_2] = k[\mathfrak{z}_1, \mathfrak{z}_2]$，其中 k 是任意的非零复数。例如，i 可以表示为 $[1+i, 1-i]$，$[-1, i]$，或 $[3+2i, 2-3i]$，这只是无限多可能中的三个例子。

另一方面，无穷远点现在可以表示为 $[\mathfrak{z}_1, 0]$ 形式的一个不起眼的复数对。

正如 $\mathbb{R}^2$ 的线性变换由实数 2×2 矩阵表示一样，$\mathbb{C}^2$ 的线性变换由复数 2×2 矩阵表示：

$$\begin{bmatrix} \mathfrak{z}_1 \\ \mathfrak{z}_2 \end{bmatrix} \longmapsto \begin{bmatrix} \mathfrak{m}_1 \\ \mathfrak{m}_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \mathfrak{z}_1 \\ \mathfrak{z}_2 \end{bmatrix} = \begin{bmatrix} a\,\mathfrak{z}_1 + b\,\mathfrak{z}_2 \\ c\,\mathfrak{z}_1 + d\,\mathfrak{z}_2 \end{bmatrix}$$

但若把 $[\mathfrak{z}_1, \mathfrak{z}_2]$ 和 $[\mathfrak{m}_1, \mathfrak{m}_2]$ 分别看做 $\mathbb{C}$ 中的点 $z = (\mathfrak{z}_1/\mathfrak{z}_2)$ 及其像点 $w = (\mathfrak{m}_1/\mathfrak{m}_2)$ 在 $\mathbb{C}^2$ 中的射影坐标，则上述 $\mathbb{C}^2$ 的线性变换将诱导出 $\mathbb{C}$ 的下述（非线性）变换：

$$z = \frac{\mathfrak{z}_1}{\mathfrak{z}_2} \mapsto w = \frac{\mathfrak{m}_1}{\mathfrak{m}_2} = \frac{a\mathfrak{z}_1 + b\mathfrak{z}_2}{c\mathfrak{z}_1 + d\mathfrak{z}_2} = \frac{a(\mathfrak{z}_1/\mathfrak{z}_2) + b}{c(\mathfrak{z}_1/\mathfrak{z}_2) + d} = \frac{az + b}{cz + d}$$

这正是最一般的默比乌斯变换！

因此，对于每个默比乌斯变换 $M(z)$，都有一个2×2矩阵 $[M]$，

$$M(z) = \frac{az + b}{cz + d} \longleftrightarrow [M] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

我们推导出，两个默比乌斯变换组合的矩阵是相应矩阵的乘积：

$$[M_2 \circ M_1] = [M_2][M_1] \tag{6.8}$$

同样，默比乌斯变换的逆可以通过对应矩阵的逆来求出：

$$[M^{-1}] = [M]^{-1} \tag{6.9}$$

容易验证[练习]，非奇异默比乌斯变换确实形成一个群（正如前面隐含的那样）。

因为默比乌斯变换的系数是非唯一的，相应的矩阵也不是唯一的：若 k 是任何非零常数，则矩阵 $k[M]$ 应与 $[M]$ 对应于同一个默比乌斯变换。然而，若令 $(ad - bc) = 1$ 而把 $[M]$ 规范化，则对应于同一个默比乌斯变换恰有两个可能的矩阵；若称其一为$[M]$，则另一个为$-[M]$；换句话说，矩阵可以确定到“相差一个符号”。

杰·彭罗斯爵士（Sir Roger Penrose）大量基础性、开创性工作的核心——技术细节参见彭罗斯和林德勒（Penrose and Rindler，1984）——与“自旋体”（spinorial objects）有关，旋转2π后不会恢复到原来的状态！ 有关此类看似不可能的物体存在的直观解释，请参见 Penrose (2005, §11.3)。

这个看似平凡不足道的事实，后来发现在数学和物理中却有深刻的含义；请参见 Penrose and Rindler [1984，第 1 章]和 Penrose[2005，第 11.3 节]。

6.3 主要结果

我们现在可以陈述并证明主要结果（6.2）的更详细版本：

1.

所有三个常曲率几何的对称群 $\mathcal{G}_+(\mathcal{S})$ 是默比乌斯变换群的子群。

1. 欧氏几何（$\mathcal{K}=0$）：

$$\mathrm{E}(z)=e^{i\theta}z+k$$

注意$z\mapsto\bar{z}$是反向等距，即对实轴的反射。因此（6.1）告诉我们，完整的等距群 $\mathcal{G}=\{\mathrm{E}(z)\}\cup\{\mathrm{E}(\bar{z})\}$。

2. 球极地图中的球面几何（$\mathcal{K}=+1$）：

$$\mathrm{S}(z)=\frac{az+b}{-\bar{b}z+\bar{a}},\quad \text{其中 } |a|^2+|b|^2=1 \tag{6.10}$$

注意$z\mapsto\bar{z}$是反向等距，即球面对通过实轴垂直平面的反射。因此（6.1）告诉我们，完整的等距群 $\mathcal{G}=\{\mathrm{S}(z)\}\cup\{\mathrm{S}(\bar{z})\}$。

3. 贝尔特拉米–庞加莱半平面地图中的双曲几何（$\mathcal{K}=-1$）：

$$\mathrm{H}(z)=\frac{az+b}{cz+d},\quad \text{其中 a、b、c、d 是实数，}(ad-bc)=1。 \tag{6.11}$$

注意$z\mapsto\bar{z}$是反向等距，即对虚轴的反射。因此，再次通过（6.1），完整的等距群 $\mathcal{G}=\{\mathrm{H}(z)\}\cup\{\mathrm{H}(-\bar{z})\}$。

矩阵结果（6.8）和（6.9）使验证这三个集合确实成群很简单[练习]。我们还注意到，在球面的情况下，矩阵是一种特殊的类型，在物理学中起着重要作用。它们称为酉矩阵，这意味着若先取共轭，然后转置（整个运算用*表示），即可得到矩阵的逆[练习]：

酉矩阵意思是 $[S][S]^* =$ 单位矩阵。

欧氏变换 $E(z)$ 是 θ 的旋转，然后是 k 的平移，很明显，这确实是平面最一般的刚性运动。因此，我们匆忙跳过这一点，并保留我们的精力来挑战证明 $\mathcal{K} = \pm 1$ 所声称的令人惊讶的结果。在这本作品中，我们不能停下来对这些可爱的结果进行几何正义； 相反，我们将向你推荐 VCA。 现在，让我们展示一下我们开发的度量机器的计算能力。

为此，我们需要暂停并介绍一些复分析。正如我们在第 4.6 节中解释的，复映射 $z \mapsto w = f(z)$ 的导数 $f'(z)$ 的定义与大学一年级微积分中的定义完全相同，所有通常的微分规则都适用而无需更改。因此，将商规则应用于规范化（$(ad - bc) = 1$）默比乌斯变换，我们发现

$$M(z) = \frac{az + b}{cz + d} \Longrightarrow M'(z) = \frac{1}{(cz + d)^2} \tag{6.12}$$

但是回想一下（4.19），复数给这个导数注入了“伸扭”的解释，它比普通微积分神奇地丰富。

我们现在可以返回到主要结果。正如欧拉是第一个证明（1775 年）球面的刚性运动只是旋转的人，而高斯（1819 年左右）首先认识到它们可以被表示为（6.10）形式的默比乌斯变换。

假设球面具有单位半径，我们可以基于复数 z 重写球极度量公式（4.22）：

$$d\hat{s} = \frac{2}{1 + |z|^2}|dz|$$

为了证明映射 $z \mapsto w = S(z) = \frac{az+b}{-\bar{b}z+\bar{a}}$ 是正向等距，因此我们必须证明

$$\frac{2}{1 + |w|^2}|dw| = \frac{2}{1 + |z|^2}|dz| \tag{6.13}$$

直接计算可得[练习]

$$1 + |w|^2 = \frac{1 + |z|^2}{\left|-\bar{b}z + \bar{a}\right|^2}$$

从（6.12）可以看出

$$\left|\frac{dw}{dz}\right| = |S'(z)| = \frac{1}{\left|-\bar{b}z + \bar{a}\right|^2} = \frac{1 + |w|^2}{1 + |z|^2}$$

从而证明（6.13），并与之一起证明主要结果的第2部分。（好吧，这并不能证明这些是唯一的正向等距。这将通过迟些的几何分析来解决。）

但是，任何凡人（甚至高斯）怎么能猜到这些默比乌斯转变的形式呢？！仅基于我们目前所知，就可得出一个简单的事实：*如果旋转把球面上的点* $\hat{z}$ *带到* $\widehat{M(z)}$，*那么它还必然把* $\hat{z}$ *的对径点带到* $\widehat{M(z)}$ *的对径点*。但我们知道，对径点的球极像与 49 页的（4.27）有关，因此

$$M\left(-\frac{1}{z}\right) = -\frac{1}{\overline{M(z)}}$$

因此，

$$\frac{a\left[-\frac{1}{z}\right]+b}{c\left[-\frac{1}{z}\right]+d} = -\frac{\overline{cz+d}}{\overline{az+b}} \Longrightarrow \frac{-b\bar{z}+a}{d\bar{z}-c} = \frac{\bar{c}\,\bar{z}+\bar{d}}{\bar{a}\,\bar{z}+\bar{b}}$$

如果 $c=-b$ 且 $d=a$ 显然是满足的，这正是（6.10）的形式！

在练习 27 中，我们给出了如何将旋转表示为默比乌斯变换的具体例子，我们还给出了一个建设性的证明，即每个旋转都必须是形式（6.10）的默比乌斯变换。

最后，考虑双曲平面。 度量公式（5.7）告诉我们必须证明 $z \mapsto w = H(z) = \frac{az+b}{cz+d}$ 满足

$$\frac{|dw|}{\operatorname{Im} w} = \frac{|dz|}{\operatorname{Im} z} \tag{6.14}$$

直接计算可得[练习，回忆 $\bar{a} = a, \ldots$]

$$\operatorname{Im} w = \frac{w-\bar{w}}{2i} = \frac{\operatorname{Im} z}{|cz+d|^2}$$

从（6.12）可知

$$\left|\frac{dw}{dz}\right| = |H'(z)| = \frac{1}{|cz+d|^2} = \frac{\operatorname{Im} w}{\operatorname{Im} z}$$

从而证明（6.14），以及主要结果的最后部分。

这些双曲等距不仅包含普通旋转和平移的类似物，而且还包括*第三种刚性运动*，称为*极限旋转*，它在普通欧氏几何中没有对应物。 它是 $\mathbb{H}^2$ 普通旋转的极限，因为旋转中心移到无穷远处，最终成为地平线上的一个点，$y=0$。 有关详细信息，请参阅 VCA，第 306-313 页。

6.4 爱因斯坦的时空几何

如果仅仅默比乌斯变换群的子群对几何就具有如此深刻的意义，那么很自然地要问，产生它们的完整群中是否存在更显著的力量。

的确，完整群至少在两个看似不相关的知识领域中发挥着重要作用：相对论和 3 维双曲几何。 事实上，这并非巧合。 这两个主题之间有着深刻的联系，但我们不能在这里探讨它，而是将你引向 Penrose 的精彩论述（2005，第 18.4 节）。

第一个角色可以毫不夸张地称为“根本的”：[84]

> 默比乌斯群描述了空间和时间的对称性，或者更准确地说，描述了爱因斯坦对两者的统一——时空。

显然，我们在这里详细探讨爱因斯坦的狭义相对论既不合适也不可行。[85] 然而，对于以前没有研究过这个理论的读者，我们将尝试说足够多的东西来理解与默比乌斯群的这种特殊联系。[86]

爱因斯坦理论的出发点是大自然的一个非凡而奇异的实验事实：

> 对于匀速相对运动的所有观察者来说，光速是相同的。

正如爱因斯坦在 1905 年首次认识到的那样，只有当这些观察者对空间和时间的测量存在分歧时，这才有可能！

为了量化这一点，让我们将时间 T 和事件 $\mathfrak{E}$ 的 3 维笛卡尔坐标 $(\mathrm{X},\mathrm{Y},\mathrm{Z})$ 组合成 4 维时空中的单个 4-向量 $(\mathrm{T},\mathrm{X},\mathrm{Y},\mathrm{Z})$。[87] 这些称为事件 $\mathfrak{E}$ 对第一个观察者的空间和时间坐标。

当然，第一个观察者向量的空间分量没有绝对意义：如果第二个观察者使用相同的原点但坐标轴相对于第一个旋转，那么第二个观察者会将不同的空间坐标 $(\tilde{\mathrm{X}},\tilde{\mathrm{Y}},\tilde{\mathrm{Z}})$ 归因于相同的事件 $\mathfrak{E}$。 然而，如果这两个观察者没有相对运动，他们仍然会在 $\tilde{\mathrm{X}}^2+\tilde{\mathrm{Y}}^2+\tilde{\mathrm{Z}}^2=\mathrm{X}^2+\mathrm{Y}^2+\mathrm{Z}^2$ 的值上达成一致，因为这代表到事件发生地距离的平方。

与此相反，我们习惯于认为时间分量 T 确实具有绝对意义。然而，爱因斯坦的理论—

---

[84] 我们不建议你偷看，尽管如此，这个结果的数学精确陈述将在 (6.20) 中给出。

[85] 对于该理论的精彩物理解释，我们推荐 Taylor 和 Wheeler (1992)。 有关该理论几何的更多信息，我们强烈推荐 Misner、Thorne 和 Wheeler (1973) 和 Penrose (2005)。

[86] 根据 Coxeter (1967, pp. 73 – 77) 的说法，H. Liebmann 在 1905 年几乎立即首次意识到了这种联系！

[87] 事实上，正是赫尔曼·闵可夫斯基在 1908 年用这些几何术语重铸了爱因斯坦 1905 年的理论，而起初爱因斯坦本人并不赞同！

—经过无数实验证实——告诉我们这是错误的。 如果两个（瞬间重合的）观察者处于相对运动中，他们会在事件发生的时间上存在分歧。 此外，他们将不再同意 $(X^2 + Y^2 + Z^2)$ 的值——这就是著名的洛伦兹收缩。

只有当两个观察者的相对速度是惊人的光速（大约每秒 186,000 英里）的很大一部分时，这种效应才可辨别。 例如，即使第二个观察者以步枪子弹的速度（每小时 2,000 英里）射离第一个观察者，那么即使在他们整个成年期间（比如 85 年），他们两个手表之间总累积差异仅相当于大约百分之一秒！ 正是我们相对于光速蜗牛般的生存状态这种意外数千年来掩盖了（并且每天都在继续掩盖）爱因斯坦发现的真相。

但是，如果第二个物体以相当一部分光速行进，那么时间和空间的扭曲就会非常剧烈，这种影响每天都在世界各地的粒子加速器中出现。 在这些奇怪的情况下，时空是否有某个方面具有绝对意义并且对此两个匀速相对运动的观察者必然达成一致？

爱因斯坦的惊人回答是“是的！”：时空确实拥有一个独立于所有观察者的绝对结构。 为此原因，爱因斯坦个人很不喜欢，多年来一直抗拒“相对论”这个名字——它应该是“绝对论”！

选一个方便的单位制使光速为 1，爱因斯坦[88]发现，两个观察者将对观察者和事件之间的时空间隔 λ 的值达成一致。 这个间隔[89] λ 是通过它的平方来定义的：

$$\boxed{\lambda^2 \equiv T^2 - (X^2 + Y^2 + Z^2) = \tilde{T}^2 - (\tilde{X}^2 + \tilde{Y}^2 + \tilde{Z}^2)} \tag{6.15}$$

正如闵可夫斯基所意识到的，这是适用于时空的距离概念的正确推广，并且时空的等距/对称性必然保留它。 但是，与普通距离完全不同的是，这里不同事件之间的间隔的平方可以为零甚至为负。

让我们在 $\lambda^2 > 0$ 的情况[90]下对 λ 进行更生动的解释。 正如已经说过的，我们假设两个观察者（彼此以匀速相对运动）在一个特定的地点和时间瞬间重合，定义了一个我们标记为 $\mathfrak{O}$ 的原点事件。（当然，要精确定义这个事件，我们必须想象我们的观察者是点状的，有一个特定的位置。）现在假设（偶然或有意）第一个观察者正好到达事件 $\mathfrak{E}$。 就她而言，她一直坐着不动，事件 $\mathfrak{O}$ 和 $\mathfrak{E}$ 都发生在她坐的地方 $(X = Y = Z = 0)$，因此 $\mathfrak{O}$ 和 $\mathfrak{E}$ 之间的时空间隔（所有观察者必然同意）就是 $\lambda = T$：

---

[88] 就像我们之前要求庞加莱稍微让开一点，为与 $\mathbb{H}^2$ 相关的贝尔特拉米腾出空间一样，现在我们必须要求爱因斯坦为庞加莱稍微让开一点：在 1905 年，庞加莱也已经发现了时空间隔的不变性。当然，亨德里克·洛伦兹应该被认为是狭义相对论的第三位父亲，但至少他在以他的名字命名的那组变换中是不朽的。

[89] 令人惊讶的是，这个间隔没有标准符号，但我们对 gimel（希伯来字母表的第三个字母）的使用似乎很合适：它（视觉上）类似于“间隔”的英文“I”，而一个内涵（文化上）是 gimel 是一座连接两点的桥梁。

[90] 如果 $\mathcal{J}^2 < 0$ 有不同但同样简单的解释； 参见 Taylor 和 Wheeler (1992, Ch. 1)。

$$\boxed{\text{如果匀速运动的观察者将手表从一个事件带到另一个事件，则两个事件之间的不变时空间隔 } \lambda \text{ 就是该手表上经过的时间。}} \tag{6.16}$$

接下来，假设在事件 $\mathfrak{O}$ 处发生了火花，因此光子（光粒子）从这次闪光向各个方向传播。 如果两个观察者都将注意力集中在一个光子上，他们都会同意沿着这个光子的时空轨迹的每个事件都有 $\lambda = 0$。 这种沿特定光线指向“长度”为零的 4-向量称为零向量。

洛伦兹变换 $\mathcal{L}$ 是时−空的一种线性变换（表示为 4×4 矩阵），它把一个观察者对某一事件的描述 $(\mathrm{T}, \mathrm{X}, \mathrm{Y}, \mathrm{Z})$ 变为另一观察者对同一事件的描述 $(\tilde{\mathrm{T}}, \tilde{\mathrm{X}}, \tilde{\mathrm{Y}}, \tilde{\mathrm{Z}})$ 。 换一个说法，$\mathcal{L}$ 就是保持量 λ^2 不变的线性变换，两个观察者所得到的此量之值必定相同。

现在设想从时−空的坐标原点发出一束闪光 —— 它的波在 3 维空间中就是一族以原定为中心的球面，其半径随光速增加。在时间 $\mathrm{T} = 1$ 之后，这些光子将形成一个单位半径的球面，我们现在将其标识为黎曼球面。 因此，现在这个球面上的点被时空坐标 $(1, \mathrm{X}, \mathrm{Y}, \mathrm{Z})$ 和复数同时标记，通过（4.25）球极射影赋值。

将投影坐标描述 $z = (\mathfrak{z}_1/\mathfrak{z}_2)$ 代入球极射影公式（4.25），我们得到，

$$\mathrm{X} = \frac{\mathfrak{z}_1\overline{\mathfrak{z}_2} + \mathfrak{z}_2\overline{\mathfrak{z}_1}}{|\mathfrak{z}_1|^2 + |\mathfrak{z}_2|^2}, \qquad \mathrm{Y} = \frac{\mathfrak{z}_1\overline{\mathfrak{z}_2} - \mathfrak{z}_2\overline{\mathfrak{z}_1}}{i(|\mathfrak{z}_1|^2 + |\mathfrak{z}_2|^2)}, \qquad \mathrm{Z} = \frac{|\mathfrak{z}_1|^2 - |\mathfrak{z}_2|^2}{|\mathfrak{z}_1|^2 + |\mathfrak{z}_2|^2}.$$

但是一条光线可以通过沿它的任何零向量来识别，因此我们现在不选择 $\mathrm{T} = 1$，而是通过将我们的零向量放大一个因子 $(|\mathfrak{z}_1|^2 + |\mathfrak{z}_2|^2)$ 来消除上述表达式中的分母（即通过选择 $\mathrm{T} = |\mathfrak{z}_1|^2 + |\mathfrak{z}_2|^2$）。 因此，新的零向量 $(\mathrm{T}, \mathrm{X}, \mathrm{Y}, \mathrm{Z})$（在与原始时空方向相同的时空方向上）由下式给出

$$\mathrm{T} = |\mathfrak{z}_1|^2 + |\mathfrak{z}_2|^2, \qquad \mathrm{X} = \mathfrak{z}_1\overline{\mathfrak{z}_2} + \mathfrak{z}_2\overline{\mathfrak{z}_1}$$

$$\mathrm{Y} = -i(\mathfrak{z}_1\overline{\mathfrak{z}_2} - \mathfrak{z}_2\overline{\mathfrak{z}_1}), \qquad \mathrm{Z} = |\mathfrak{z}_1|^2 - |\mathfrak{z}_2|^2$$

可以很容易地检验这些公式可以反转来得到

$$\begin{pmatrix} \mathrm{T} + \mathrm{Z} & \mathrm{X} + \mathrm{iY} \\ \mathrm{X} - \mathrm{iY} & \mathrm{T} - \mathrm{Z} \end{pmatrix} = 2\begin{pmatrix} \mathfrak{z}_1\overline{\mathfrak{z}_1} & \mathfrak{z}_1\overline{\mathfrak{z}_2} \\ \mathfrak{z}_2\overline{\mathfrak{z}_1} & \mathfrak{z}_2\overline{\mathfrak{z}_2} \end{pmatrix} = 2\begin{pmatrix} \mathfrak{z}_1 \\ \mathfrak{z}_2 \end{pmatrix}\begin{pmatrix} \overline{\mathfrak{z}_1} & \overline{\mathfrak{z}_2} \end{pmatrix}$$

或者，

$$\boxed{\begin{pmatrix} \mathrm{T} + \mathrm{Z} & \mathrm{X} + \mathrm{iY} \\ \mathrm{X} - \mathrm{iY} & \mathrm{T} - \mathrm{Z} \end{pmatrix} = 2\begin{pmatrix} \mathfrak{z}_1 \\ \mathfrak{z}_2 \end{pmatrix}\begin{pmatrix} \mathfrak{z}_1 \\ \mathfrak{z}_2 \end{pmatrix}^*} \tag{6.17}$$

其中“*”再次表示共轭转置。

注意，时空间隔现在可以巧妙地表示为该矩阵的行列式：

$$\boxed{\lambda^2 = T^2 - (X^2 + Y^2 + Z^2) = \det\begin{pmatrix} T+Z & X+iY \\ X-iY & T-Z \end{pmatrix}} \tag{6.18}$$

这使得很容易看到这个放大的时空向量仍然是空的 [练习]，因为它应该是。

现在已经使用球极射影来用复数标记光线，[91]通过（6.17），让我们找出 $\mathbb{C}$ 的默比乌斯[92]变换 $z \mapsto \tilde{z} = M(z)$ 对闪光的影响！或等价地，

$$\begin{bmatrix} \mathfrak{z}_1 \\ \mathfrak{z}_2 \end{bmatrix} \longmapsto \begin{bmatrix} \tilde{\mathfrak{z}}_1 \\ \tilde{\mathfrak{z}}_2 \end{bmatrix} = [M]\begin{bmatrix} \mathfrak{z}_1 \\ \mathfrak{z}_2 \end{bmatrix}$$

将其代入（6.17），我们得到构成闪光的零向量的以下线性变换：

$$\begin{pmatrix} T+Z & X+iY \\ X-iY & T-Z \end{pmatrix} \mapsto \begin{pmatrix} \tilde{T}+\tilde{Z} & \tilde{X}+i\tilde{Y} \\ \tilde{X}-i\tilde{Y} & \tilde{T}-\tilde{Z} \end{pmatrix} = [M]\begin{pmatrix} T+Z & X+iY \\ X-iY & T-Z \end{pmatrix}[M]^* \tag{6.19}$$

最后，想象一下这个线性变换应用于所有时空向量（不仅仅是空向量）。因为 $\det[M] = 1 = \det[M]$ ，从（6.18）可以看出，这个线性变换保持时空间隔：

$$\boxed{\tilde{\lambda}^2 = \det\left([M]\begin{pmatrix} T+Z & X+iY \\ X-iY & T-Z \end{pmatrix}[M]^*\right) = \lambda^2}$$

因此，

> $\mathbb{C}$ 的每一次默比乌斯变换都会产生一个独特的时空洛伦兹变换。 相反，可以证明（Penrose 和 Rindler，1984，第 1 章），对于每个洛伦兹变换，都对应一个唯一（最多相差一个符号）的默比乌斯变换。 (6.20)

即使在专业物理学家中，这个美丽的“奇迹”也没有应有的知名度。

事实证明，每一个默比乌斯或洛伦兹变换基本上都等价于四种原型中的一种。 基本思想是关注变换的所谓不动点，即在映射到自身的意义上保持固定的点：$M(z) = z$。

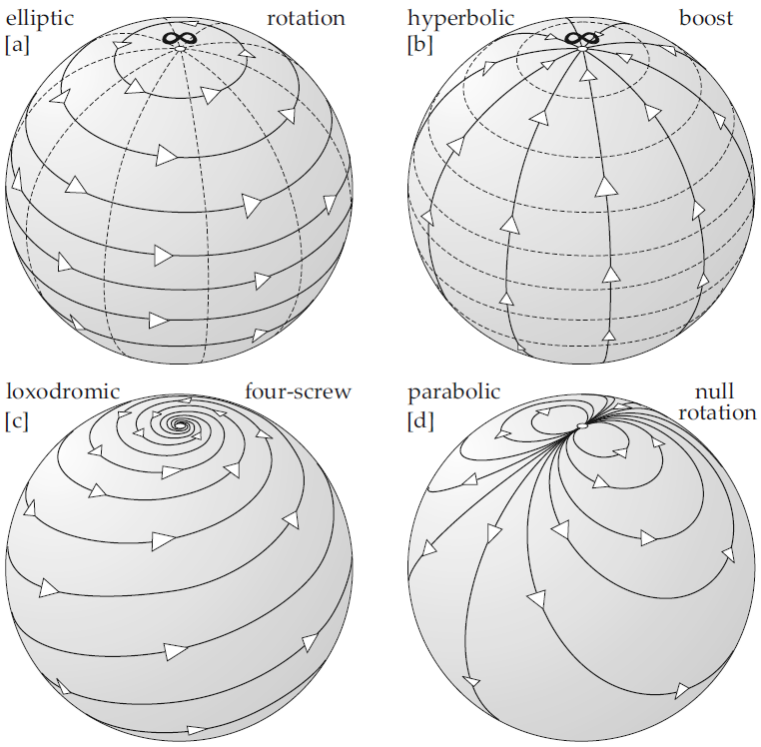
---

[91] 练习 33 解释了彭罗斯在时空中直接完成此任务的方法。

[92] 在相对论的语境中，默比乌斯变换称为自旋变换，对应的矩阵[M]称为自旋矩阵。 有关详细信息，请参见 Penrose 和 Rindler (1984, Ch. 1)。

显然这会导致二次方程有两个可能不同或重合的解 [练习]。 如果不动点不同，可以想象[93]它们是北极和南极； 这产生了 [6.5a-c] 中显示的三个原型。 如果两个不动点重合，可以想象它们都是北极； 这产生了 [6.5d] 中显示的第四个原型。

为了让这四种洛伦兹变换更加生动，想象自己在星际空间的飞船中，向四面八方眺望散落在天球上的无数星星。 这是一个想象中的单位球面，以你为中心，在它的表面上标记一个点，星光在它到达你眼睛的途中与球面相交的任何地方。 我们假设球面与你的宇宙飞船对齐，北极位于你的正前方。



[6.5] 默比乌斯和洛伦兹变换的分类。 四种类型的转换中的每一种都有两个名称，具体取决于它被视为作用于 $\mathbb{C}$（左侧的名称）还是作用于时空（右侧的名称）。

如果你现在发射你的横向推进器，让你的宇宙飞船围绕它的南北轴旋转，那么星场将围绕你旋转，如图 [6.5a] 所示。 不出所料，这种洛伦兹变换称为*旋转型*（rotation）。 在复平面中对应的默比乌斯变换也是一种旋转 $M(z) = e^{i\alpha}z$，称为*椭圆型*(elliptic)。

假设你不再旋转而是启动强大的主引擎。 几乎立即，这会将你的飞船以接近光速 v 直接冲向天球的北极。 如图 [6.5b] 中所示，你实际上会看到星星朝着你的行进方向拥挤——这与大多数科幻电影中的描述完全相反！ 这种类型的洛伦兹变换称为*加速型*(boost)。 在复平面中，相应的默比乌斯变换是由一个实因子 ρ 的简单伸缩 $M(z) = \rho z$，称为*双曲型*(hyperbolic)。 事实上（见例 34）膨胀系数 ρ 与航天器速度 v 的关系为

[93] 有关默比乌斯变换情况下这种简化的详细说明，请参见 VCA, Ch. 3； 对于洛伦兹变换，参见 Penrose 和 Rindler (1984, Ch. 1)。

$$\rho = \sqrt{\frac{1+v}{1-v}} \tag{6.21}$$

如果你同时启动主引擎和横向推进器，那么前两个效应结合起来会导致星场向你的运动方向螺旋，如图 [6.5c] 所示。 这种类型的洛伦兹变换称为*四螺杆型*(four-screw)，相应的默比乌斯变换 $M(z) = \rho e^{i\alpha} z$ 称为*斜驶型*(loxodromic)。

最后一种洛伦兹变换见图 [6.5d]。 它称为*零旋转型*(null rotation)，它的两个不动点在北极重合。 虽然很难对这种时空变换给出生动的物理描述，但相应的默比乌斯变换称为*抛物型*(parabolic)，是 $\mathbb{C}$ 的简单转换：$M(z) = z + \tau$。这将沿着平行于 τ 的直线移动每个复数。但是，正如我们从图 [4.8]中所看到的那样，这些直线的球极像是到通过北极的圆，在那里有一个平行于 τ 的公共切线，从而解释了图 [6.5d] 的形式。 请注意，随着接近北极，球面上的运动变得越来越小，从而使北极成为唯一的不动点。

6.5 3 维双曲几何

值得注意的是，罗巴切夫斯基和鲍耶都没有开始发展二维非欧几何。相反，从一开始，两人都独立地寻求了三维欧几里得[94]空间的双曲替代方案；双曲平面随后自然出现在这个双曲 3 空间中。

回想一下，我们在上半平面中创建了双曲面的共形映射，方法是将度量取为

$$d\hat{s} = \frac{ds}{\text{边界上方的欧氏高度}} = \frac{\sqrt{dX^2 + dY^2}}{y} \tag{6.22}$$

要将其推广到三维双曲空间 $\mathbb{H}^3$，请考虑欧几里德 3-空间 (X, Y, Z) 的水平 (X, Y) 平面上方的半空间 $Z > 0$。我们现在可以创建 $\mathbb{H}^3$的共形模型，再次将度量取为[95]

$$\boxed{d\hat{s} = \frac{ds}{\text{边界上方的欧氏高度}} = \frac{\sqrt{dX^2 + dY^2 + dZ^2}}{y}} \tag{6.23}$$

因此，(X, Y)平面 $(Z = 0)$上的点是无限远的，这个边界平面是二维地平线，我们现在将把它看作是复平面 $\mathbb{C}$，坐标为 $X + iY$。

从这种结构中，每个垂直半平面都是$\mathbb{H}^2$的副本。请参见图[6.6]并与图[5.5]进行比较。(通过延伸，想象这个空间充满了相等双曲线大小的小球面，随着地平线的接近，以明显的欧几

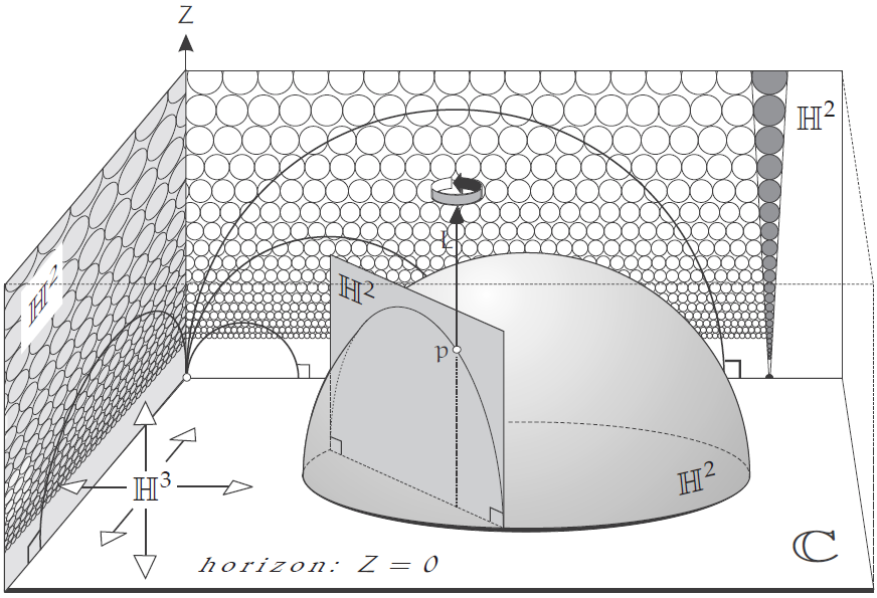
[94] 参见 Stillwell (2010, p. 365)。
[95] 这也是由贝尔特拉米在 1868 年发表的，连同对 n 维的推广。 见斯蒂尔韦尔 (1996)。

里德大小收缩。）

如果我们想象光在这样一个垂直平面内传播，对称性就规定它保持在该平面内，因此，我们之前对光所走路径的分析适用。因此，

> $\mathbb{H}^3$ 的测地线是垂直半直线和垂直于地平线 $\mathbb{C}$ 的半圆。

在$\mathbb{H}^2$中，垂直半直线是特殊的测地线，典型的是与地平线正交的半圆。同样，在 $\mathbb{H}^3$ 中，垂直半平面是特殊的双曲面；典型的平面实际上是正交于地平线 $\mathbb{C}$ 的半球，如图[6.6]中所示的半球。



[6.6] **3 维双曲空间** $\mathbb{H}^3$。 双曲平面表现为垂直的半平面和与 $\mathbb{C}$ 正交的半球。两个这样的平面的交集是一条双曲线：垂直的半直线或与 $\mathbb{C}$ 正交的半圆。

要开始看到这一点，请想象你住在 $\mathbb{H}^3$； “平面”这个词是什么意思？ 假设我递给你一个以 p 为中心的无穷小圆盘 D。 要将这个圆盘（唯一地）扩展到无限的“平面”H，你可能需要将其直径扩展为无限的双曲线，即与 $\mathbb{C}$ 正交的半圆。 如果圆盘是垂直的，这显然会产生包含 D 的垂直半平面。 如果圆盘不垂直，我们声称该构造将生成一个与 $\mathbb{C}$ 正交的半球 H，D 与在 p 处与 H 的切平面重合。 顺便说一下，我们可以很容易地将这个半球的中心构造为通过 p 与 D 的（欧氏）法线的 $\mathbb{C}$ 的交点。

在图 [6.6] 中，让我们将图示的半球作为这个 H。考虑图示的 H 与通过 p 的垂直平面的交点。 很明显，这是一条延伸直径为 D 的双曲线，因为我们已经构建了 H 使其切平面与 D 在 p 处重合。 现在让 L 是图中所示的通过 p 的垂直线。 如果我们绕 L 旋转垂直平面，这条双曲线相交会扫过所有的 H，所以这确实是 D 到平面的唯一延伸。 因此，

> $\mathbb{H}^3$ 的双曲平面是垂直的半平面和与地平线 $\mathbb{C}$ 正交的半球。

很明显，垂直半平面在两种意义上是双曲平面：它是 $\mathbb{H}^3$ 的居民所说的平面，其内部几何结构是 $\mathbb{H}^2$的几何结构。 半球面也一样吗？ 也就是说，如果我们使用环境空间 $\mathbb{H}^3$的双曲线度量 (6.23) 来测量这样一个半球内的距离，我们是否得到 $\mathbb{H}^2$？

答案是肯定的：每个半球本质上都是一个 $\mathbb{H}^2$。 事实上，如果你住在 $\mathbb{H}^3$，那么每个方向都像每个其他方向，每条线都像其他线，每个平面都像其他平面。 通过将几何反演推广到 3 维空间，可以给出一种自然的几何解释，通过表明存在将垂直双曲平面带入半球平面的运动，因此它们的内蕴几何必然相同。[96] 不过，现在我们将进行计算。

如果我们在半径为 R 的半球上使用球面极坐标（经度和纬度），则 $Z = R\cos\phi$，因此（4.4）意味着双曲度量（6.23）简化为

$$d\hat{s}^2 = \frac{\sin^2\phi d\theta^2 + d\phi^2}{\cos^2\phi} \tag{6.24}$$

要想看出这确实是 $\mathbb{H}^2$，我们可以引入新坐标将这个度量转换为标准形式（6.22）（见例 28），或者我们可以使用内蕴度量曲率公式（4.10）来确认[练习] $\mathcal{K} = -1$。

让我们描述几个简单的 $\mathbb{H}^3$ 等轴测图。 度量（6.23）显然没有被水平欧氏平移改变，

$$X + iY \longmapsto X + iY + (\text{复常量}), \qquad Z \longmapsto Z$$

同样，双曲距离不会被以原点为中心的空间欧氏扩张（膨胀）改变，

$$(X, Y, Z) \longmapsto k(X, Y, Z) = (kX, kY, kZ)$$

其中 $k > 0$ 是扩张因子，因为这将使 ds 和 Z 扩大相同的因子 k，而 $d\hat{s}$ 保持不变。

将此与平移相结合，我们看到以地平线为中心的任何膨胀也是等距。 这两种等距的存在再次证实了所有半球面都必须具有相同的内蕴几何，因为它们中的任何一个都可以刚体运动到另一个：平移第一个使其中心与第二个中心重合，然后扩展它直到半径相等。

任何等距都必须将平面带到平面，这意味着它们在 $\mathbb{C}$ 中的边界圆也必须携带到其他圆。事实证明（虽然我们不能在这里证明）这个圆守恒性质是如此特殊，仅凭它就足以完全确定[97]所涉及的复映射：它们只能是默比乌斯变换！

相反，$\mathbb{C}$ 的任何默比乌斯变换都会诱导所有 $\mathbb{H}^3$的唯一变换，因为 $\mathbb{H}^3$中的点 p 被唯一确定为三个平面的交点，可以将其编码为这些半球与 $\mathbb{C}$ 相交的三个圆。这些图像 三个圆然后确定三个新的半球，在 p 的图像中相遇。 此外，可以证明[98]这种诱导转换自动是 $\mathbb{H}^3$

---

[96] 有关详细信息，请参阅 VCA，第 133-136、322-327 页

[97] 卡拉西奥多里（Carathéodory，1937）以最一般的形式证明了这一非凡的事实。

[98] 要了解庞加莱是如何得到这一发现的，请参见 Stillwell (1996, pp. 113 – 122)。 有关结果背后的几何结构的更多信息，请参见 VCA，第 322-327 页。

的正向等距。总结一下庞加莱 (1883) 的这一奇妙发现，

双曲 3 空间的正向等距群 $\mathbb{H}^3$是 $\mathbb{C}$ 的默比乌斯变换群。

如果你在这个 $\mathbb{H}^3$ 世界长大，上过 $\mathbb{H}^3$ 的学校，你会知道三角形内角和小于两个直角，有无限多条直线平行于一条给定的直线等等。但是在大学里 你最终可能会学到一种理论几何，它违背了你所有的日常经验：无论三角形有多大，三角形内角和正好等于 π！ 为了直观地理解这种奇怪的现象，数学家试图构建一个曲面模型，在该曲面上这种“欧氏”几何真实成立。 而且，非同寻常的是，他们成功了！

从寓言回到现实，实际上是沃切特（Wachter，高斯的学生）于 1816 年首次发现了这样的曲面，称为双曲极限球面（英语：Horosphere），罗巴切夫斯基与鲍耶独立地重新发现了它，后来由罗巴切夫斯基命名。 就 [6.6] 而言，典型的极限球面是位于（接触）地平线的普通欧氏球面。 非常令人惊讶的是，如果我们使用双曲度量（6.23）来测量距离，则极限球面内的几何形状确实是平坦的欧氏平面！ 参见练习 30 和 32。

贝尔特拉米给出了一种更简单的方式来看待这一点。 使用允许我们交换垂直半平面和半球的相同几何反演的 3D 推广，他观察到典型的双曲极限球面可以被移动使其成为 $Z = \text{const.} = k$。在这个水平表面内， 双曲度量（6.23）随后减少到

$$d\hat{s} = \frac{\sqrt{dX^2 + dY^2}}{k}$$

这显然是欧氏的。

回想一下，希尔伯特证明了不能在欧几里得空间中构造一个完整的双曲平面。 因此，双曲空间中极限球面（完整欧几里得平面）的存在是我们对双曲几何优越性的第一次提示，欧几里得几何从属于它。 事实上，球面几何也包含在内。 因为事实证明，$\mathbb{H}^3$ 的居民所称的球面（即距中心双曲线距离为常数的点集）在我们的模型中实际上正是欧几里得球面，只是中心不同。 此外，这样的双曲球面具有常正曲率的内蕴几何：它如看起来一样是球形的！ 参见练习 31。

我们一再谈到生活在诸如 $\mathbb{H}^3$ 等非欧几里得世界中智慧生物的概念； 这并没有你想象的那么虚幻。 1915 年，爱因斯坦发现我们居住的实际时空并不平坦！ 事实上，能量和物质扭曲了时空结构，产生了一种复杂的曲率模式，有正有负，因地点、时间和方向而异。

然而，在正常情况下，曲率非常小，以至于三角形内角和与 π 相差一个完全无法察觉的小量，从而产生了我们的世界遵循欧几里得几何定律的错觉。 这种错觉是如此完美，如此引人注目，以至于它影响了世界 4000 年。

尽管如此，爱因斯坦在 1915 年发现的时空曲率的微妙模式是极其重要的，即使在日常生活中也是如此，它有一个名字…叫做重力！

第七章 第二幕练习

映射曲面：度量

1. $\mathcal{K}$ **的坐标独立性**。从平坦的欧几里德度量开始，

$$d\hat{s}^2 = dx^2 + dy^2$$

   (i) 现在通过编写 $x = f(u,v)$ 和 $y = g(u,v)$ 来更改坐标，其中 f 和 g 是你选择的显式简单函数，并在 (u,v) 坐标中找到新的度量公式。
   (ii) 应用曲率公式（4.10）来确认新度量仍然是平坦的，因为它应该是：$\mathcal{K} = 0$。
   (iii) 在 f 和 g 完全通用的情况下重做 (ii)。

2. **中心投影**。 考虑 [4.1] 中显示的球体的中心投影。
   (i) 证明球体上的一个无限小圆被扭曲成一个椭圆，其形状由（4.3）给出。
   (ii) 解释为什么这种扭曲是对称的：如果我们从平面中的一个无限小圆开始，球体上的图像是一个相同形状的椭圆。

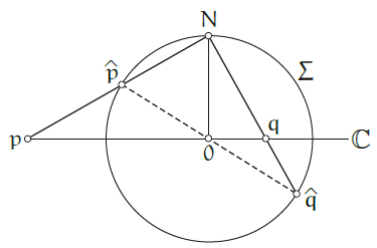
3. **中心投影**。 通过计算重新导出中心投影度量（4.2）。

4. **中心投影**。将曲率公式(4.10)应用于球体在中心投影下的度量(4.2)，并确认$K = +1/R^2$。

5. **n 边形的角盈**。 证明欧氏 n 边形的内角和为 $(n-2)\pi$，从而证明角盈定义为 $\mathcal{E} \equiv [\text{内角和}] - (n-2)\pi$。 现在将 Δp 取为一个包含曲面上的点 p 的小 n 边形，并通过将 p 连接到顶点将其划分为三角形。 推导出 (2.1) 不需改变即可适用。

6. **球极度量**。 假设在图[4.10]中我们选择正北的$\delta\hat{s}$（而不是图示的正西选择）。 证明这导致相同的度量（4.22），并以两种方式这样做：
   (i) 通过计算；
   (ii) 几何方法。

7. **对映点**。 下图显示了黎曼球 Σ 通过 $\hat{p}$ 及其对映点 $\hat{q}$ 的垂直截面。 证明三角形 p0N 和 N0q 是相似的。 推导出 (4.27)。



8. **共形（“等温”）坐标**。 给定一般曲面在一般 (u,v) 坐标中的度量，(4.7)，

$$d\hat{s}^2 = A^2\,du^2 + B^2\,dv^2 + 2F\,du\,dv$$

我们的目标是找到共形 (U,V) 坐标，如果 $Z = U + iV$，则

$$d\hat{s}^2 = \Lambda^2 (dU^2 + dV^2) = \Lambda^2 \, dZ \, d\bar{Z}$$

设 $M = \sqrt{A^2 B^2 - F^2}$ 是从 (u, v)-地图到曲面的面积放大系数，因此 (4.11) 变为 $d\mathcal{A} = M du \, dv$。

(i) 验证度量具有以下复因式分解：

$$d\hat{s}^2 = \left[A \, du + \frac{(F + iM)}{A} dv\right] \left[A \, du + \frac{(F - iM)}{A} dv\right]$$

(ii) 假设可以找到一个复积分因子Ω，使得

$$\Omega \left[A \, du + \frac{(F + iM)}{A} dv\right] = dU + i dV = dZ$$

推导在这种情况下 (U,V) 确实是曲面上的共形坐标：

$$d\hat{s}^2 = \Lambda^2 (dU^2 + dV^2) = \Lambda^2 \, dZ \, d\bar{Z}, \text{其中 } \Lambda = \frac{1}{|\Omega|}$$

(iii) 使用 $df = (\partial_u f) \, du + (\partial_v f) \, dv$ 推导出

$$\partial_u Z = \Omega A, \qquad \partial_v Z = \Omega \left[\frac{F + iM}{A}\right]$$

(iv) 推导出

$$[F + iM] \, \partial_u Z = A^2 \, \partial_v Z$$

(v) 将前面方程的两边同时乘以 $[F - iM]$ ，推导出

$$B^2 \partial_u Z = [F - iM] \partial_v Z$$

(vi) 通过实部和虚部各自相等，推导出 U 的变化率可以用 V 的变化率表示（反之亦然）如下：

$$\partial_u U = \frac{1}{M} [A^2 \, \partial_v V - F \, \partial_u V], \qquad \partial_v U = \frac{1}{M} [F \, \partial_v V - B^2 \, \partial_u V]$$

$$\partial_u V = \frac{1}{M} [F \, \partial_u U - F \, \partial_v U], \qquad \partial_v V = \frac{1}{M} [B^2 \, \partial_u U - F \, \partial_v U]$$

(vii) 假设 $\partial_u \partial_v \Phi - \partial_v \partial_u \Phi = 0$，推导出 $\Phi = U$ 和 $\Phi = V$ 都是贝尔特拉米-拉普拉斯方程的解：

$$\partial_v \left[\frac{A^2 \, \partial_v \Phi - F \, \partial_u \Phi}{M}\right] + \partial_u \left[\frac{B^2 \, \partial_u \Phi - F \, \partial_v \Phi}{M}\right] = 0$$

这个方程是椭圆型的，现在椭圆偏微分方程的一般理论保证这个方程确实有解，从而证实了共形坐标的存在！

(viii) 令 (x, y) 为第二对共形坐标，使得

$$d\hat{s}^2 = \lambda^2 (dx^2 + dy^2)$$

通过设$u = x$和$v = y$，推导出$A = B = \lambda$，$F = 0$，$M = \lambda^2$。 由 (vi) 推导出 $(x + iy)$ 平面和 $(U + iV)$ 平面之间的局部映射是一个伸扭（amplitwist），由著名的柯西–黎曼方程表征：

$$\partial_x U = \partial_y V, \qquad \partial_y U = -\partial_x V$$

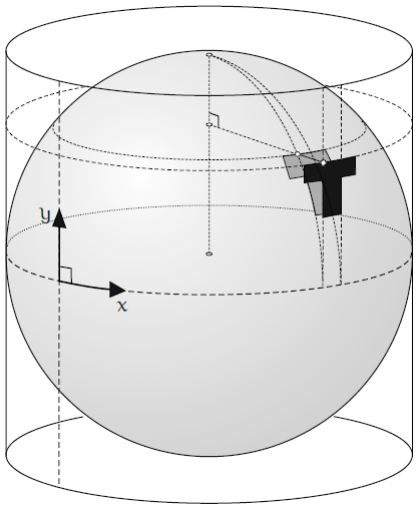
（有关这些现象的完整讨论，请参阅 VCA。）

(ix) 推导在这种情况下，贝尔特拉米-拉普拉斯方程简化为拉普拉斯方程：

$$\partial_x^2 \Phi + \partial_y^2 \Phi = 0$$

9. **共形曲率公式**。 将共形曲率公式（4.16）应用于立体投影下球体的度量（4.22），并证明 $K = +1/R^2$。

10. **阿基米德–兰伯特投影**。 在 (x, y) 平面中，考虑矩形 $\{0 \le x \le 2\pi R, -R \le y \le R\}$。 现在想象一下，首先将左右边缘用胶带粘在一起以创建一个圆柱体，然后将其滑到半径为 R 的球体上，使圆柱体沿其赤道接触球体（如下图所示）。 我们现在通过将球体的点水平投影到圆柱体上来绘制球体的地图，径向向外，垂直于圆柱体的轴，如图所示。 这是阿基米德–兰伯特投影，首先由阿基米德（公元前 250 年）进行研究，然后在大约 2000 年后的 1772 年由兰伯特重新发现并出版。兰伯特的开创性论文是第一个对制图学进行系统数学研究，基于制图保持什么属性。



(i) 证明（最好以几何方式）现在球面度量的形式是

$$d\hat{s}^2 = (R^2 - y^2)\, dx^2 + \frac{dy^2}{R^2 - y^2}$$

(ii) 应用曲率公式（4.10）证明 $\mathcal{K} = +1/R2$。

(iii) 使用（4.12）导出这种投影方法保持面积。例如，所示的两个 T 形具有相等的面积。正如阿基米德所意识到的那样，这意味着整个球面的面积必须等于或原始矩形：$(2\pi R)(2R) = 4\pi R^2$。阿基米德对此（以及卷上的相关发现）感到非常自豪，以至于他要求他的朋友们把这张特别的图表刻在他的坟墓上。大约一个半世纪后，公元前 75 年，西塞罗寻找并发现了这座坟墓，那里长满了灌木丛，但圆柱体和球体仍然可见。

(iv) 利用上一部分推导（不积分）极帽 $0 \le \phi \le \Phi$ 的面积$\mathcal{A}$为

$$\mathcal{A} = 2\pi R^2(1 - \cos\Phi)$$

11. **中心圆柱投影**。 重新考虑上一个问题图中的球体和圆柱体，但现在想象圆柱体向上和向下无限延伸。 我们再次将球体投影到圆柱体上，但这一次通过想象球体中心的点光源，将阴影投射到圆柱体上。 这称为中心圆柱投影。

(i) 绘制投影并推导出 $y = R\cot\phi$。

(ii) 使用第 (i) 部分或直接以几何形式显示度量现在采用以下形式

(iii) 应用曲率公式 (4.10) 来确认 $K = +1/R^2$。

12. **墨卡托投影**。 前两个练习中球体在圆柱体上的投影有两个共同点：1) 子午线 ($\theta = \text{const.}$) 映射到具有相同 θ 的圆柱体的垂直发生器； 2) 纬度圆 ($\phi = \text{const.}$) 映射到圆柱的水平圆形横截面。 1569 年，杰拉尔杜斯 • 墨卡托 (Gerardus Mercator) 发现了第三个投影，它具有这些特性，但与其他两个投影相比，它具有至关重要的优势，即共形。

(i) 从几何角度论证，如果 δbs 是沿球体纬度圆的小运动，那么上述两个所需属性意味着 $\delta\hat{s} \asymp \sin\phi\,\delta s$，其中 $\delta s = \delta x$ 是圆柱体上的水平运动。

(ii) 为了让我们的地图保形，我们必须坚持我们对所有方向的无穷小运动都有相同的比例因子。 特别地，假设$\delta\hat{s}$ 改为沿子午线选择，并让 $y = f(\phi)$ 是球体上 (θ, ϕ) 图像在圆柱体上的高度。 推导出

$$f'(\phi) = \frac{R}{\sin\phi}$$

(iii) 如果我们坚持将赤道地图上的点指向 x 轴，请验证（或推导，如果你还记得你的积分技术）墨卡托投影由下式给出

$$y = f(\phi) = R\ln\left[\frac{1+\cos\phi}{\sin\phi}\right]$$

（注意：在墨卡托时代，对数和微积分都不为人所知，那么他是怎么做到的？！请参阅 Osserman 在 Hayes 等人（2004，第 18 章）中的精彩章节。）

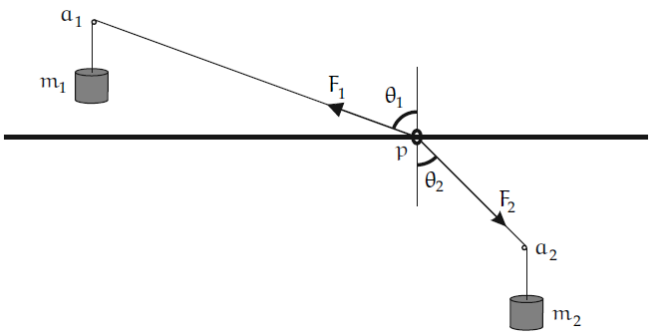
(iv) 如果你在固定的罗盘航向上驾驶飞机或航行船舶，你将沿着一条螺旋曲线（loxodrome）（在 [6.5c] 中说明），也称为恒向线。 如果你将圆柱体展开到桌面上以获得标准的平面墨卡托航海图，那么你的螺旋曲线路线如何在图表中显示？（提示：无需计算：只需借助共形性！）

13. **一张不可思议的好地图**。 在最近的几个练习中，我们见到了保持面积的球体地图和保

持角度的另一个球体地图。我们可以有一张同时保留两者的地图吗？ 证明对于任何曲面都不存在这样的地图，而不仅仅是球体。

伪球面与双曲平面

14. **波利亚对斯涅尔定律的机械证明**。 波利亚 (1954, pp. 149 – 152) 对斯涅尔定律提供了以下巧妙的机械解释，因此也是我们思考双曲平面测地线的另一种方式。 下图是[5.6]的修改形式，其中空气和水之间的边界被无摩擦的杆代替，沿着它滑动一个环，p。 这枚戒指上有两根固定长度的绳子。 如图所示，每根弦都经过 a_i 处的无摩擦栓钉，并连接到质量 m_i，该质量 m_i 悬挂在 a_i 下方距离 h_i 处。



(i) 通过考虑系统的势能，推导出当

$$m_1h_1 + m_2h_2 = \text{maximum}$$

(ii) 但是为了达到平衡，拉在环上的力 F_i 的水平分量必须抵消。 推导出

$$m_1 \sin\theta_1 = m_2 \sin\theta_2$$

(iii) 令 l_i 为从 a_i 到 p 的字符串长度，因此 $l_i + h_i = \text{const.}$. 现在让我们选择 $m_i = (1/v_i)$，其中 v_1 和 v_2 分别代表光在空气中和水中的速度， 就像最初的光学问题一样。 推导出 (i) 部分中的机械问题在数学上与寻找最短时间光路的原始问题在数学上相同：

$$\frac{l_1}{v_1} + \frac{l_2}{v_2} = \text{minimum}$$

(iv) 从（ii）部分力学问题的解中，推导出斯涅尔定律！

15. 如果 (x, y) 平面中的光速为 $v = 1/\sqrt{1-y}$，则证明测地线是抛物线。 描述这些抛物线。

16. **曳物线参数化**。 在 (X, Y, Z) 空间中，让我们切换到圆柱极坐标，即 (X, Y) 平面中的普通 (r, θ) 极坐标，由 Z 补充。因此，该空间中的欧几里得度量是

$$ds^2 = dr^2 + r^2 d\theta^2 + dZ^2$$

在 $\theta = 0$ 平面，即 (X, Z) 平面中，考虑一个粒子在时间 t 的位置为

$$X = r = \operatorname{sech} t\,, Z = t - \tanh t$$

(i) 证明这条曲线是曳物线：沿 Z 轴切线的距离是固定的。这个固定距离是多少？（注意：一些有趣的历史和美丽的几何图形是这种特殊表示的基础；参见 Stillwell (2010, pp. 341 – 342)。）

(ii) 推导出，如果我们绕 Z 轴旋转曲线，我们得到单位半径的伪球面。 证明表面的度量是

$$d\hat{s}^2 = \tanh^2 t\, dt^2 + \operatorname{sech}^2 t\, d\theta^2$$

(iii) 使用 (4.10) 确认 $\mathcal{K} = -1$。

(iv) 证明如果我们现在引入新坐标 $x = \theta$ 和 $y = \cosh t$，那么度量采用标准形式（5.7）。

17. **球面上的圆柱极面积**。 如果我们使用前面练习的圆柱极坐标，单位球体北半球的面积元素是，

$$d\mathcal{A} = \frac{r\, dr\, d\theta}{\sqrt{1 - r^2}}$$

(i) 通过计算证明这一点。

(ii) 使用牛顿终极等式从几何上证明这一点。

(iii) 将其积分以获得完整半球的面积。

18. **伪球面积有限！** 用（5.5）和（4.12）证明半径为 R 的无界伪球面有有限面积 $2\pi R^2$。这是惠更斯于 1693 年发现的。

19. **制作你自己的伪球！** 拿一叠十张纸并将它们装订在一起，沿三个边缘放置订书钉。 使用一对圆规（或倒盘或碗）绘制最大的圆圈，使其能够舒适地放入顶板。 在圆心刺穿所有十张纸。用重剪刀沿圆剪下，得到十个相同的圆盘，比如半径 R。重复整个过程，将圆盘数量加倍到 20。

(i) 从第一个圆盘上切出一个狭窄的扇形，并将边缘粘在一起形成一个浅锥。 对剩余的圆盘重复此过程，每次稳步增加扇形的角度，使锥体变得更尖、更高。 （你可能会发现在用胶带粘在一起之前切割径向狭缝并重叠纸张更容易。）确保在该过程结束时制作非常窄的锥体，仅使用四分之一圆盘或更少。

(ii) 按照你制作的顺序堆叠这些锥体。 解释如何制作半径为 R 的伪球体的一部分的模型。

(iii) 使用相同的想法创建一张类似圆盘的“双曲线纸”，就像你可以简单地从伪球中切出圆盘一样。将它压在伪球面上，并试验看是否可以在曲面上自由移动和旋转它。（有关详细说明，请参阅 Henderson (1998, p. 32)。）

20. **伪球的测地线**。 使用（1.7），在你在上一个练习中构建的玩具伪球体的表面上构建一段典型的测地线。 向两个方向延伸此段，一次一条胶带。 请注意测地线在再次向下螺

旋之前仅在伪球面上螺旋有限距离的令人惊讶的方式。

(i) 使用贝尔特拉米–庞家莱上半平面以数学方式验证曳物线母线是唯一无限向上延伸的测地线。

(ii) 设 L 是一个典型的测地线，设 α 是 L 和在 L 碰到边缘 $\sigma = 0$ 的点处的曳物线母线之间的角度。 证明 L 在伪球面上向上移动的最大距离$\sigma_{\max}$ 由 $\sigma_{\max} = |\ln \sin \alpha|$ 给出。

21. **共形曲率公式**。证明在具有度量 (4.14) 的共形映射的情况下，一般曲率公式 (4.10) 简化为 (4.16)。

22. **常曲率旋转曲面**。 想象一个粒子以单位速度在 (x, y) 平面中沿曲线运动，并令其在时间 t 的位置为 $[x(t), y(t)]$。 现在想象一下将这个平面绕 x 轴旋转角度 θ。 当θ从 0 到 2π 变化时，曲线扫过旋转表面。

(i) 解释为什么$\dot{x}^2 + \dot{y}^2 = 1$，其中点代表时间导数。

(ii) 从几何上证明曲面的度量是 $d\hat{s}^2 = dt^2 + y^2\, d\theta^2$。

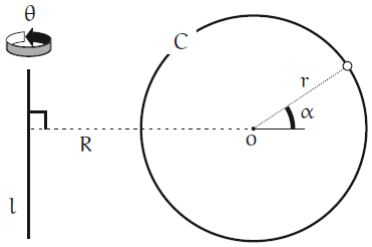
(iii) 从 (4.10) 推导出 $\mathcal{K} = -\ddot{y}/y$。

(iv) 求出三种常曲率情况下的通解 $y(t)$：$\mathcal{K} = 0$、$\mathcal{K} = +1$ 和 $\mathcal{K} = -1$。

(v) 在 (i) 部分的帮助下，计算三种情况下粒子的速度。

(vi) 在三种情况中的每一种情况下，绘制一些解曲线和产生的旋转曲面； 计算机可能会有所帮助。 特别是，在 $\mathcal{K} = +1$ 的情况下，确认确实获得了像 [2.5] 中所示的曲面。 同样，验证在 $\mathcal{K} = -1$ 的情况下，不仅可以获得伪球体本身，还可以获得看起来像两个伪球体在其狭窄的颈部粘在一起的表面。

23. **圆环的曲率**。 通过围绕 C 平面中距 o 距离为 $R(> r)$ 的直线 l 旋转圆心为 o 和半径为 r 的圆 C 来生成环面 T。 见下图。



(i) 如果 α 是绕圆的图示角度，而 θ 是图示绕对称轴 l 的旋转角，则从几何上证明 T 的度量是

$$d\hat{s}^2 = r^2 d\alpha^2 + (R + r\cos\alpha) d\theta^2$$

(ii) 使用（4.10）推导出环面的曲率是

$$\mathcal{K} = \frac{1}{r\,(r + R\sec\alpha)}$$

(iii) 画出 $\sec\alpha$ 的图，然后是 $(r + R\sec\alpha)$，最后是 $\mathcal{K}$ 本身。 从而确认 $\mathcal{K}$ 的这个

公式与图 [2.2] 的经验发现一致。

(iv) 你希望在 T 上的哪个位置找到 $\mathcal{K} = 0$？ 用公式确认这一点。

(v) 由公式可知，$\lim_{R\to\infty} \mathcal{K} = 0$。 用几何解释这一点。

(vi) 使用 (4.12) 求出 T 的总面积。

(vii) 检查你的面积公式是否与帕普斯质心定理的预测一致（如果你以前没有研究过这个定理，请先在书本或互联网上查找。）

(viii) 证明环面的总曲率为零：

$$\iiint_{T} \mathcal{K}\, d\mathcal{A} = 0$$

在第三幕中，我们将看到这绝非偶然！

等距同构与复数

24. **z^m 的伸扭**。 推广 [4.6] 中的几何论证以推导出 z^m 的伸扭与实数微积分的公式相同给出： $(z^m)' = mz^{m-1}$； 但现在我们可以明白为什么这是真的了！

25. **贝尔特拉米-庞加莱圆盘的度量**。 我们从 [6.3] 中知道，默比乌斯变换可以将半平面映射到圆盘，而且，正如我们稍后将能够解释的，结果证明，从 $\mathbb{H}^2$ 的上半平面模型 [5.12] 到 $\mathbb{H}^2$ 的圆盘模型 [5.11] 的特定默比乌斯变换是

$$z \mapsto w = D(z) = \frac{iz+1}{z+i}$$

.

为了导出圆盘模型的度量 (5.15)，考虑从 z（在半平面中）出发的无穷小向量 dz 被放大为从 w（在圆盘中）出发的无穷小向量 $dw = D'(z)\, dz$。 根据定义，dw 的双曲长度 $d\hat{s}$ 是 dz 的双曲长度。 证明下式来验证 (5.15)

$$\frac{2\,|dw|}{1-|w|^2} = \frac{dz}{\operatorname{Im} z} = d\hat{s}$$

26. **贝尔特拉米-庞加莱圆盘的曲率**。 使用共形曲率公式 (4.16) 验证具有度量 (5.15) 的贝尔特拉米-庞加莱圆盘 [5.11] 具有恒定的负曲率 $\mathcal{K} = -1$。

27. **黎曼球面的默比乌斯旋转**。 我们的目的是给出一个建设性的、半几何的证明，证明黎曼球的每次旋转都是类型（6.10）的默比乌斯变换。 我们的方法在本质上与威尔逊先前发表的方法一致（2008 年，第 42-44 页）。 （注：矩阵的对应群写成SU(2)：“S”代表“特殊”，意思是归一化，有单位行列式；“U”代表“酉”；“2”代表“ 2 × 2” 矩阵。）在下面，让 R_Z^θ 表示黎曼球面绕 (X, Y, Z) 空间的正 Z 轴顺时针旋转角度 θ，因此 $R_Z^\theta(z) = e^{i\theta} z$。

(i) 在 SU(2) 中找到一个表示 R_Z^θ 的矩阵 $[R_Z^\theta]$。

(ii) 解释为什么结果（6.6）可以写成

$$[\mathrm{R}_X^{\pi}] = \pm\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \in \mathrm{SU}(2)$$

(iii) 参考式（4.24），用草图说明对应于 $\mathrm{R}_X^{(\pi/2)}$ 的复映射为

$$\mathrm{z} = \frac{\mathrm{X} + i\mathrm{Y}}{1 - \mathrm{Z}} \longmapsto \mathrm{R}_X^{\left(\frac{\pi}{2}\right)}(z) = \frac{\mathrm{X} - i\mathrm{Z}}{1 - \mathrm{Y}}$$

(iv) 为了确认 $\mathrm{R}_X^{(\pi/2)}$ 确实是一个默比乌斯变换，让我们暂时假设它是，并让我们尝试猜测它的形式。使用草图确认 $-i \mapsto 0$，因此分子必然与 $(\mathrm{z} + \mathrm{i})$ 成正比。同样，通过找到旋转到北极的点，推导出分母。这决定了 $\mathrm{R}_X^{(\pi/2)}$ 的形式，直至乘以因子；通过注意 0 7！ 一世。 结论是如果 $\mathrm{R}_X^{(\pi/2)}$ 是一个默比乌斯变换，它只能是

$$\mathrm{R}_X^{(\pi/2)} = \frac{\mathrm{z} + i}{i\mathrm{z} + 1}$$

(v) 通过证明它满足第 (iii) 部分中的方程，证明这实际上是 $\mathrm{R}_X^{(\pi/2)}$。（提示：顶部和底部乘以 $(\mathrm{X} - \mathrm{iZ})$，但只乘以分母。最后，在所得分母中，使用 $\mathrm{X}^2 + \mathrm{Y}^2 + \mathrm{Z}^2 = 1$ 的事实。）

(vi) 推导出 $\mathrm{R}_X^{(\pi/2)}$ 可以在 $\mathrm{SU}(2)$ 中表示为

(vii) 从几何上解释为什么 $\left[\mathrm{R}_X^{(\pi/2)}\right]^2 = [\mathrm{R}_X^{\pi}]$ ，并通过直接计算证实这一点。

(viii) 从几何上解释为什么对于任何角度 α，

$$[\mathrm{R}_Y^{\alpha}] = \left[\mathrm{R}_X^{-(\pi/2)}\right][\mathrm{R}_Z^{\alpha}]\left[\mathrm{R}_X^{(\pi/2)}\right]$$

并推导出 $[\mathrm{R}_Y^{\alpha}] \subset \mathrm{SU}(2)$。

(ix) 最后，考虑绕任意轴 A 的一般旋转。 通过将A旋转到(X, Y)平面，然后将其旋转到Y轴，推导出这也可以表示为SU(2)中的矩阵，从而得出证明。

28. $\mathbb{H}^3$ 中 $\mathbb{H}^2$ 的度量。让我们探索表示 $\mathbb{H}^3$ 中典型双曲平面（即以地平线为中心的半球）的度量 (6.24) 的不同方式。

(i) 证明如果我们定义 $\mathrm{u} \equiv \tan\phi$，则度量 (6.24) 变为

$$\mathrm{d}\hat{s}^2 = u^2 \mathrm{d}\theta^2 + \frac{\mathrm{d}u}{1 + u^2}$$

并使用（4.10）来确认 $\mathcal{K} = -1$。

(ii) 证明如果我们通过 $u \equiv 1/\sinh\xi$ 定义一个新变量 ξ，则前一部分的度量变为共形：

$$\mathrm{d}\hat{s}^2 = \frac{\mathrm{d}\theta^2 + \mathrm{d}\xi^2}{\sinh^2\xi}$$

并使用（4.16）来确认 $\mathcal{K}=-1$。

(iii) 最后，找到 (θ,ξ) 平面到 (x,y) 平面的共形映射，使得第 (ii) 部分中的共形双曲度量采用标准形式 (5.7)：

$$d\hat{s}^2=\frac{dx^2+dy^2}{y^2}$$

（提示：设$dy/y=d\xi/\sinh\xi$。）

29. **H3 中半球形 H2 的曲率**。 使用曲率公式（4.10）验证具有度量（6.24）的 $\mathbb{H}^3$ 的水平中心半球确实是具有恒定负曲率 $\mathcal{K}=-1$ 的$\mathbb{H}^2$。

30. **地层：度量和曲率**。 正如所讨论的，$\mathbb{H}^3$ 的典型地平层表现为一个停在 C 上的球体。使用 (6.24) 的推导作为你的灵感，证明这样一个地平层的度量是

$$d\hat{s}=\frac{\sin^2\phi\ d\theta^2+d\phi^2}{(1+\cos\phi)^2}$$

应用曲率公式 (4.10) 来证明地层本质上是一个平坦的欧几里得平面：$\mathcal{K}=0$。

31. $\mathbb{H}^3$ 中的球体。 考虑一个半径为 R 的欧几里得球体，其中心位于 $\mathbb{H}^3$ 地平线上方的高度 kR。 以 (6.24) 的推导为灵感，证明这样一个球体的度量是

$$d\hat{s}=\frac{\sin^2\phi\ d\theta^2+d\phi^2}{(k+\cos\phi)^2}$$

应用曲率公式（4.10）来证实文中的说法，如果这个球体完全位于地平线上方（即 $k>1$），那么这个表面本质上是真正的球面，具有恒定的正曲率，$\mathcal{K}=k^2-1$。 （注意：前两个练习的结果分别作为特殊情况 $k=0$ 和 $k=1$。）

32. **地平层度量**。 让我们以另一种方式推导出习题 30 的结果。 定义坐标 r 使得

$$dr=\frac{d\phi}{1+\cos\phi}$$

证明地层的度量（见例 30）可以写成 $d\hat{s}^2=dr^2+r^2\,d\theta^2$，这就是极坐标中欧氏平面的度量。

33. **Penrose 用复数直接标记光线**。 在文本中，我们使用黎曼球体通过立体投影来标记具有复数的光线。 Roger Penrose 爵士（参见 Penrose 和 Rindler (1984, Vol. 1, p. 13)）发现了一种从光线直接传递到相关复数的显着替代方法。 令 p 点垂直于（水平）复平面原点上方一个单位。 现在想象一下，同时 p 发出闪光，$\mathbb{C}$ 开始以光速 (= 1) 向 p 直

线向上（沿方向 $\phi = 0$）行进。 （你可以想象整个 $\mathbb{C}$ 闪烁，产生一个平面波。）将 p 在 (θ, ϕ) 方向上发射的光子 F 的速度分解为垂直和平行于 $\mathbb{C}$ 的分量。因此找到 F 的时间 击中 C。推导出 F 在点 $z = \cot(\phi/2)\, e^{i\theta}$ 击中 $\mathbb{C}$。 因此，彭罗斯的构造等价于球极射影！

34. **爱因斯坦的像差公式**。 回想一下狭义相对论，沿 Z 轴的“加速”（boost） v 产生洛伦兹变换公式，

$$\tilde{T} = \frac{T + vZ}{\sqrt{1 - v^2}}, \qquad \tilde{X} = X, \qquad \tilde{Y} = Y, \qquad Z = \frac{Z + vT}{\sqrt{1 - v^2}}$$

(i) 证明这个变换可以改写为

$$\tilde{T} + \tilde{Z} = \rho(T + Z), \qquad \tilde{X} = X, \qquad \tilde{Y} = Y, \qquad \tilde{T} - \tilde{Z} = (1/\rho)(T - Z)$$

其中

$$\rho = \sqrt{\frac{1 + v}{1 - v}}$$

(ii) 从（6.19）推导出这种加速可以用自旋变换表示，

$$\begin{bmatrix} \tilde{\delta}_1 \\ \tilde{\delta}_2 \end{bmatrix} = \begin{bmatrix} \sqrt{\rho} & 0 \\ 0 & 1/\sqrt{\rho} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}$$

(iii) 从而确认在[6.5b]和（6.21）中提出的主张。

(iv) 如果 (θ, ϕ) 是在你的飞船引擎点火之前恒星的视向，推论在你在方向 $\phi = 0$ 上加速到速度 v 之后，恒星的新方向 $(\theta, e\phi)$ 由下式给出

$$\cot\frac{\tilde{\phi}}{2} = \rho \cot\frac{\phi}{2}$$

这是标准像差公式的一种更优雅和令人难忘的形式，它是爱因斯坦在 1905 年发现的。

(v) 推导出星域似乎朝着北极聚集在一起，即朝着你的行进方向。

(vi) 当你的飞船速度接近光速时，星域会发生什么变化？（!）

第八章 平面曲线的曲率

8.1 引言

曲率的研究并不是从曲面开始的。相反，顾名思义，曲率（curvature）源于对曲线（curves）的研究，特别是平面中的曲线。

我们当然有一个直观的想法，即平面曲线在一个部分更直，在另一个部分更弯曲，但是我们如何用数学精确量化这种弯曲程度？

在第一幕中，我们专注于内蕴（高斯）曲率的概念。 在第一幕结束时，我们为这种对内蕴（相对于外在）几何的痴迷提供了另一个令人信服的理由：它是唯一一种对我们作为时空居民有意义的时空几何，而时空的内蕴曲率具有根本的重要性——这就是重力。

但是，虽然给定二维曲面片只能以非常有限的方式变形，同时保留其内蕴几何（即，不拉伸其中的距离），但对于一维曲线则不是这样：我们可以弯曲它（而无需拉伸它）使其跟任何一条相同长度的曲线完全一样。 因此，一维内蕴曲率的概念根本不存在！ 换句话说，如果你是一个生活在曲线中的非常短的一维生物，只能在你沿着曲线移动时测量沿曲线来回的距离，那么你曲线的形状对你来说不仅是不可知的，而且毫无意义。

因此，从一开始，很明显，我们所能期望的最好的是曲率的外在定义：曲线如何位于平面内。 最大的惊喜——“奇迹”会是一个更好的词——是这种外在的一维曲率概念最终与二维曲面的内蕴曲率有直接关系。

正如我们所说，曲线中的一维生物无法通过其中的测量来感知其世界的曲率。 但是，假设我们放松了对这种存在的可测量和可知的概念。 假设我们承认物理概念，例如质量、速度和力，而不是只允许测量沿曲线的长度。 现在，让我们继续将自己限制在平面曲线上。将此平面曲线想象为位于外层空间的无摩擦线，没有重力（或其他）作用力，假设我们沿线以单位速度发射单位质量无摩擦的珠子。[99]

由于没有外力，珠子将在沿着金属丝的整个行程中继续以单位速度行进。但是牛顿第一运动定律告诉我们，如果没有力作用在珠子上，它只会沿直线运动。这里发生的情况是，当珠子试图沿直线飞行时，金属丝以与运动方向成直角的方式压在它上面，旋转速度向量，使其继续沿着金属丝指向，但没有改变速度的大小，也就是速率。

一段电线弯曲得越紧，速度矢量在通过它时必须转动得越快。牛顿第二运动定律告诉我们，速度向量的这种变化率（加速度）实际上等于（对于单位质量）弯曲线施加在珠子上的力 F。 金属丝的弯曲越尖锐，金属丝必须施加的力就越大，以使珠子的轨迹弯曲。

为了让这个想法更加深入人心，请回想一下在游乐园或游乐场沿着（过山车）弯曲轨道

[99] 当然，虽然这本质上仍然涉及二维平面中的一维曲线，但我们现在将所有这些视为嵌入在我们的物理 3 维世界中。 事实上，我们至少需要三个维度才能让珠子环绕并附着在金属丝上。

高（恒定）速行驶的感觉。 你实际上已经成为一个一维存在，被沿着运动的路径运输。（我们很快就会讨论通过 3 维空间扭曲的曲线，但是，现在，假设你正在沿着一条水平（因此是平面）延伸的轨道行驶。）即使你闭上眼睛，你在绕过一个弯道时仍然可以感受到力量，座舱的侧面压在你的身体上：转得越急，力就越大。

牛顿实际上是第一个引入曲率 κ 的纯几何定义的人，我们马上就会遇见它，并且由于他的定义和他的运动定律，他确实能够推断出

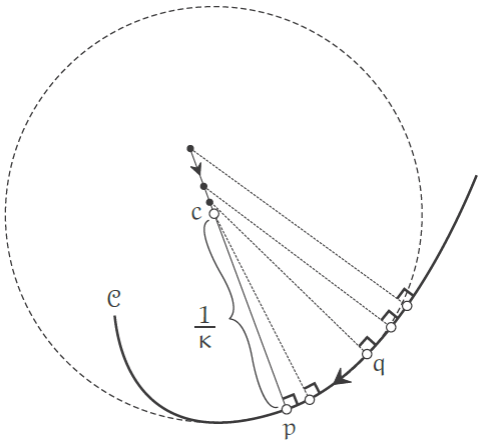
> 如果单位质量的珠子以单位速度沿平面曲线形式的金属丝发射，则该曲线的曲率 κ 等于垂直于曲线的，由金属丝施加在珠子上的力 F。 (8.1)

几何学与物理学之间的这种联系（轨道的几何弯曲与将物体保持在该轨道上所需的力之间）成为牛顿巨著（可以说是历史上最重要的科学著作，即 1687 年伟大的《原理》）的关键。

在其中，牛顿使用无穷小几何——前言中描述的“终极等式”——来解释天体的运行，包括行星围绕太阳的椭圆轨道作为主要焦点。只是现在行星没有被拧在电线上，它们的轨道被无形的引力之手从直线运动中偏转。这个引力之手从太阳延伸到太空，强度与距离的平方成正比递减：此即牛顿著名的*万有引力的平方反比定律*。

8.2 曲率圆

当 21 岁的牛顿开始研究他所谓的平面曲线的“弯曲性”时（就在他 1664 年圣诞节生日[100]前不久），他首次将曲率的概念引入了数学中，从而为二十多年后的《原理》奠定了基础。

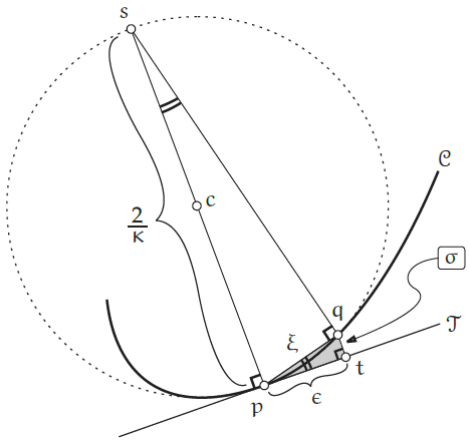


[8.1] p 处的曲率圆是最接近那里曲线的圆。 曲率 κ 定义为其半径的倒数。

[100] 在我的家里，我们（或者至少我）把这个事件（圣诞节）称为“牛顿节”!

牛顿将曲线 $\mathcal{C}$ 上点 p 处的曲率圆定义为最接近 p 附近曲线的圆（正如切线是p 处最接近 p 附近曲线的直线）（参见 [8.1]）。

他构造了这个近似圆的中心 c（曲率中心）作为 p 处的法线与相邻点 q 处的法线的交点，在$q \to p$过程中的极限位置。 pc 称为曲率半径，牛顿最初将 $\kappa \equiv (1/pc)$ 称为“弯曲度”，后来重新命名为曲率（关于符号的重要说明：这里的“pc”表示点 p 和 c 之间的距离）。



[8.2] 曲线从其切线的偏离 σ 最初随距离 ϵ 二次增加（如抛物线），比例常数是曲率的二分之一。

接下来，跟随牛顿的脚步，我们通过查看曲线偏离其自身切线的速度来测量 κ（参见 [8.2]）。 根据 p 处的曲率 κ 的定义，直径的长度为 $ps = (2/\kappa)$。 现在让 q 是 $\mathcal{C}$ 上靠近 p 的一个点（其中 $\xi = pq$），并从 q 到 p 处的切线 $\mathcal{T}$ 放下长度为 $qt = \sigma$ 的垂线，最后令 $\epsilon = pt$。

由于 $\mathcal{T}$ 与 $\mathcal{C}$ 相切， $\lim_{\epsilon \to 0}(\sigma/\epsilon) = 0$，因此

$$\frac{\xi^2}{\epsilon^2} = \frac{\epsilon^2 + \sigma^2}{\epsilon^2} = 1 + \left[\frac{\sigma}{\epsilon}\right]^2 \asymp 1 \Longrightarrow \xi \asymp \epsilon$$

关于符号的重要说明：在此以及在整本书中，我们将使用 $\asymp$ 符号，它代表了在前言中介绍和定义的牛顿终极相等概念。

此外，阴影三角形 ptq 最终相似于三角形 sqp，所以

$$\frac{\xi}{[2/\kappa]} \asymp \frac{\sigma}{\xi}$$

这实际上是牛顿《原理》第一卷（Newton 1687 p. 439）的引理 II，来自（另见 Brackenridge and Nauenberg (2002, p. 112)）。 根据需要求出的是 κ 还是 σ，我们可以结合前两个结果来推导出

$$\boxed{\kappa \asymp \frac{2\sigma}{\epsilon^2} \quad 或 \quad \sigma \asymp \frac{1}{2}\kappa\epsilon^2} \tag{8.2}$$

在曲率上附加一个符号很方便：如果 $\mathcal{C}$ 向上凹，κ 为正，如果向下凹，则为负。例如，紧接着，笛卡尔方程 $y = ax^2$ 的抛物线在原点处具有曲率 $\kappa = 2a$。同样，$\cos x$ 的泰勒展开告诉我们图 $y = \cos x$ 在点 $(0, 1)$ 处具有 $\kappa = -1$，因此曲率圆是单位圆，以原点为中心 [练习]。

8.3 牛顿曲率公式

假设曲线 $\mathcal{C}$ 是图 $y = f(x)$。如果 x 轴平行于切线 $\mathcal{T}$ 绘制，则泰勒定理意味着 [练习] $\sigma \asymp (1/2)f''(x_p)\epsilon^2$，其中 x_p 表示 p 的 x 坐标。因此（8.2）意味着

$$\boxed{\kappa = f''(x_p)}$$

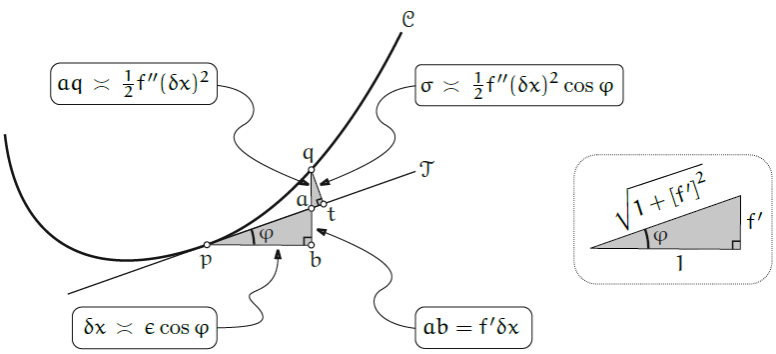
注意，此公式自动符合我们关于 κ 符号的约定。

更一般地，假设现在沿任意方向绘制 x 轴，使得 $\mathcal{T}$ 与其成角度 φ 倾斜，因此 $f' = \tan\varphi$。在这个一般情况下，牛顿发现上述公式的正确推广是，

$$\boxed{\kappa = \frac{f''}{\{1 + [f']^2\}^{3/2}}} \tag{8.3}$$

这里可以理解为导数是在 $x = x_p$ 处计算的。

现在已经接触了几个使用无穷小几何的例子，你应该不难遵循牛顿对其公式的原始证明，该证明在 Knoebel (2007, pp. 182 – 185) 中被复制。然而，在这里，我们将提供一个不同的几何论证，试图将（8.3）的复杂的新分母解释为旋转 x 轴的简单结果（见图 [8.3]）。



[8.3] 简单公式 $\kappa = f''$（在此情况下 $\varphi = 0$）在一般情况下必须乘以$\cos^3\varphi$；如右侧方框所示，$\cos\varphi = 1/\sqrt{1 + [f']^2}$。

和以前一样，这里 $\epsilon \equiv pt$ 和 $\sigma \equiv qt$，我们现在假设 x 轴水平绘制，而不是沿着切线 $\mathcal{T}$。泰勒定理现在说如果 $\delta x \equiv x_q - x_p$ 那么 q 的高度高于 $\mathcal{T}$ 的部分由 $aq \asymp (1/2)f''(\delta x)^2$

给出。 因为两个阴影三角形明显相似，所以当 aq 投影到垂直于 $\mathcal{T}$ 的方向时，我们拾取一个 $\cos\varphi$ 因子，得到 $\sigma \asymp (1/2)f''(\delta x)^2\cos\varphi$。

接下来，$pa \asymp pt = \epsilon$，因此 $\delta x \asymp \epsilon\cos\varphi$。 因此，在从切线投影到水平 x 轴时，我们获得了两个额外的 $\cos\varphi$ 因子，总共产生三个：

$$\sigma \asymp \frac{1}{2}f''\epsilon^2\cos^3\varphi$$

最后，[8.3]右边的方框子图证明了

$$\cos\varphi = \frac{1}{\sqrt{1+[f']^2}}$$

所以牛顿的公式（8.3）现在直接从（8.2）得到。

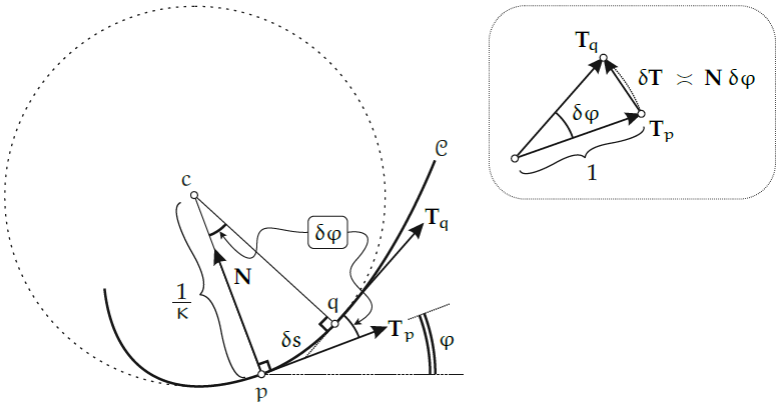
8.4 曲率作为转弯的速率

1761 年，也就是在牛顿引入曲率概念大约一个世纪之后，凯斯特纳（Kaestner）发表了一个简单的替代解释，最终证明该解释比牛顿的更适合泛化。

> 曲率是切线相对于弧长的转动率。换句话说，如果 φ 是切线的角度，s 是弧长，那么 $\kappa = (d\varphi/ds)$。 (8.4)

注意，这种通过在一小段曲线附近进行的局部测量来定义曲率的方法具有直接优势，即：我们不再需要绘制法线并沿着它们延伸到远处，直到它们在曲率中心相交。

为了证实（8.4）等价于牛顿的定义，只要验证半径为 ρ 且曲率 $\kappa = (1/\rho)$ 的正圆就足够了，因为这个圆可以被认为是一般情况的曲率圆。 由于在圆的情况下转动率显然是均匀的，考虑一整圈：切线在穿过 $2\pi\rho$ 的整个圆周时旋转 2π，因此转动率是 $(2\pi)/(2\pi\rho) = \kappa$，正如将要展示的那样。



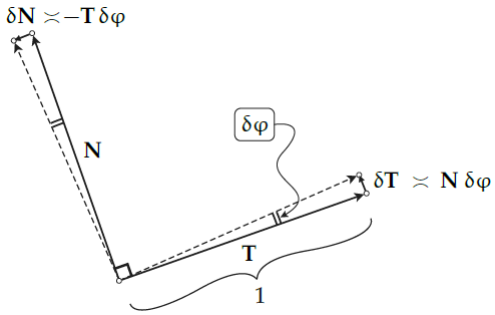
[8.4] 曲率是切线的转动率。 如果单位切线 $\mathbf{T}$ 沿曲线 $\mathcal{C}$ 每移动距离 δs 时转过角度 $\delta\varphi$，则 $\kappa \asymp (\delta\varphi/\delta s)$。右边的框中，向量上，单位切线的变化率为 $(d\mathbf{T}/ds) \asymp (\delta\mathbf{T}/\delta s) \asymp (\mathbf{N}\delta\varphi/\delta s) \asymp \kappa\mathbf{N}$。

图 [8.4] 反驳了这种等价性，也让我们更进了一步。 这里 $\mathbf{T}$ 表示 $\mathcal{C}$ 的单位切向量，N 表示单位法线，指向曲率中心 c。当我们从 p 移动一小段距离 δs 到 q 时，$\mathbf{T}$ 旋转了 $\delta\varphi$。由于 δs 最终等于半径为 $(1/\kappa)$ 的曲率圆的相应段（根据牛顿最初的定义），$\delta s \asymp (1/\kappa)\delta\varphi$。因此 $\kappa = (d\varphi/ds)$，再次证明了（8.4）。

但是现在，此外，图[8.4] 右侧的子图帮助我们计算从 p到 q 时向量 $\mathbf{T}$ 本身的变化 $\delta\mathbf{T} \equiv (\mathbf{T}_q - \mathbf{T}_p)$。 通过画出从同一点发出的两个单位切线，我们可以将 $\delta\mathbf{T}$ 可视化为连接它们尖端的图示向量，最终将指向 $\mathbf{N}$。但该向量的长度最终等于单位圆的弧（虚线所示）。综合上述内容，可得 $\delta\mathbf{T} \asymp \mathbf{N}\delta\varphi$。 因此，

$$\boxed{\frac{\delta\mathbf{T}}{\delta s} \asymp \mathbf{N}\frac{\delta\varphi}{\delta s} \quad \Longrightarrow \quad \frac{\delta\mathbf{T}}{\delta s} = \kappa\mathbf{N}} \tag{8.5}$$

我们现在可以回到曲率的开篇讨论，即单位质量的珠子沿曲线 $\mathcal{C}$ 以单位速度行进。因为它以单位速度行进，珠子的速度向量 $\mathbf{v}$ 只是 $\mathbf{T}$，而且 $\delta s = \delta t$. 因此在（8.5）中 $(d\mathbf{T}/ds) = (d\mathbf{v}/dt)$ 实际上是珠子的加速度，而后者又是钢丝施加的力。 我们因此确认了最初的声明（8.1）。



[8.5] 当 $\mathbf{T}$ 和 $\mathbf{N}$ 一起旋转时，它们的尖端分别沿着 $\mathbf{N}$ 和 $-\mathbf{T}$ 做等量移动。

当然，我们也可以将曲率视为法线 $\mathbf{N}$ 的转动率（而不是切线）。 确实图 [8.5]简单而形象地表明，当 $\mathbf{T}$ 的尖端向 $\mathbf{N}$ 的方向旋转时， $\mathbf{N}$ 的尖端向$-\mathbf{T}$ 的方向做等量旋转，所以

$$\frac{d\mathbf{N}}{ds} = -\kappa\mathbf{T}$$

事实上，当我们将注意力从曲线转向曲面时，使用法线代替切线将被证明是至关重要的，因为在后一种情况下，没有“切线”这样的东西，但仍然存在一个唯一的法向量，垂直于曲面的切平面。这个法向量在曲面上一个点的附近如何变化确实会告诉我们该点的高斯曲率。

（顺便说一句，我们再次注意到几何理解和盲目计算之间的区别，这在上面的例子中非常简单。让 φ 表示切线与水平 x 轴的夹角，因此

$$\mathbf{T} = \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}, \qquad \mathbf{N} = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}$$

用撇号表示关于弧长的微分，则$\kappa = \varphi'$。 计算立即可得 $\mathbf{T}' = \varphi'\,\mathbf{N}$ 和 $\mathbf{N}' = -\varphi'\,\mathbf{T}$。）

如果珠子以任意（但恒定）的速度 v 行进并且具有任意质量 m，该怎么办？ 在那种情况下，根据定义，$\delta s \asymp v\,\delta t$，并且 $\mathbf{v} = v\mathbf{T}$。 因此，（8.5）变成了以下著名结果的广义版本，即将物体保持在圆形轨道上所需的力（向心力）：

$$\mathbf{F} = m\,\frac{d\mathbf{v}}{dt} = \kappa m v^2 \mathbf{N}$$

牛顿早在 1665 年就第一次看到了这一点，[101]当时他努力理解使月球保持在轨道上的力。

这种对曲率（作为转动率）的新解释使我们能够处理不能表示为函数图的曲线，因此超出了牛顿原始公式（8.3）的范围。 令 $\mathcal{C}$ 为在时间 t 的位置为 $(x[t], y[t])$ 的粒子的轨道，因此速度为

$$\mathbf{v} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}, \quad \text{因此} \quad \tan\varphi = \frac{\dot{y}}{\dot{x}}$$

其中，在牛顿之后，点[102]表示相对于时间的微分。 请注意，我们不再假设速度是恒定的：$\mathbf{v}$ 的长度可能变化。

通过对前面方程 $\tan\varphi = \dot{y}/\dot{x}$ 的两边进行微分，利用链式法则，不难得到一个更通用的公式（同样是牛顿发现的）[练习]：

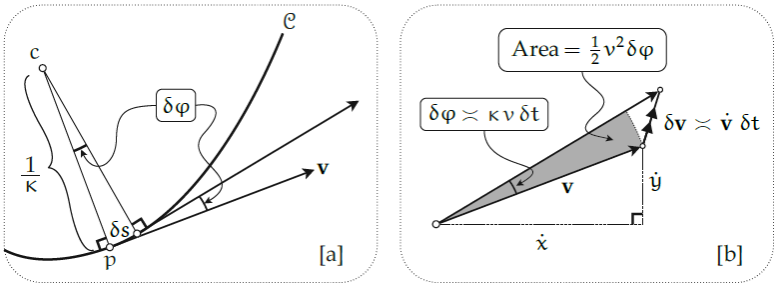
$$\boxed{\kappa = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{[\dot{x}^2 + \dot{y}^2]^{3/2}}} \tag{8.6}$$

然而，让我们试着用更直接的几何术语来理解这个公式。 参见图 [8.6a]，它显示了 p 处的速度 $\mathbf{v}$ 和一瞬间 δt 后的速度，之后粒子沿着 $\mathcal{C}$ 行进了 $\delta s \asymp v\,\delta t$。在图 [8.6b] 中，两个速度向量都从同一点画出，因此连接它们尖端的向量是速度的变化，$\delta\mathbf{v} \asymp \dot{\mathbf{v}}\,\delta t$。

---

[101] 参见 Westfall (1980, pp. 148 – 150)。

[102] 牛顿完美的极简符号有时会令人不悦，因为圆点很难看到——在本书中，我们试图通过简单地扩大圆点来解决这个问题！



[8.6] [a] 在 p 处的速度 v 和一个时刻 δt 之后，在沿 $\mathcal{C}$ 行进 $\delta s \asymp v\,\delta t$ 之后。 [b] 两个速度向量都是从同一点发出的，所以连接它们尖端的向量是 $\delta\mathbf{v} \asymp \dot{\mathbf{v}}\,\delta t$。 以两种不同的方式观察三角形的面积，得出κ 的公式（8.6）。

因此，半径 $v = \sqrt{\dot{x}^2 + \dot{y}^2}$，对角 δφ 的圆的阴影扇形面积由下式给出

$$\left(\text{阴影扇形面积}\right) = \frac{1}{2}v^2\delta\varphi \asymp \frac{1}{2}v^2\kappa\,\delta s \asymp \frac{1}{2}v^3\kappa\,\delta t$$

但是这个面积最终等于带有边向量的图示三角形的面积

$$\mathbf{v} = \begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix}, \qquad \delta\mathbf{v} = \begin{bmatrix}\ddot{x}\\ \ddot{y}\end{bmatrix}\delta t$$

并且通过基本结果（其本身可以通过几何证明[练习]），该区域是具有这两列的矩阵的行列式的一半：

$$\left(\text{以速度向量为界的三角形面积}\right) \asymp \frac{1}{2}(\dot{x}\ddot{y} - \dot{y}\ddot{x})\delta t$$

因此，

$$\frac{1}{2}v^3\kappa\,\delta t \asymp \frac{1}{2}(\dot{x}\ddot{y} - \dot{y}\ddot{x})\delta t$$

从而证明（8.6）。

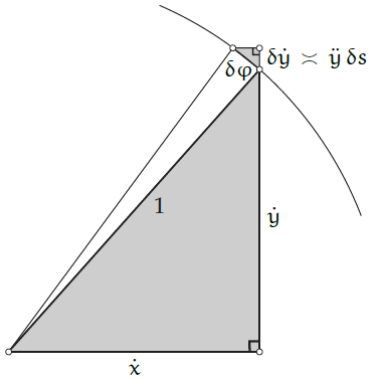
注意，如果粒子的水平速度固定在 $\dot{x} = 1$，因此其轨道为 (t, y[t])，那么我们将牛顿原始公式（8.3）恢复为（8.6）的特例。 事实上，这正是牛顿本人进行简化的方式； 再次参见 Knoebel (2007, pp. 182 – 185)。

最后，考虑另一个特别重要的特殊情况，其中粒子以单位速度运动，使得

$$|\mathbf{v}| = \sqrt{\dot{x}^2 + \dot{y}^2} = 1$$

并且 $\delta s = \delta t$。 因此，与一般图 [8.6b] 的变化是 δv 现在与 v 正交，与单位圆相切。如图 [8.7] 中所示，在时间 δt 期间，v 的尖端现在画出单位圆的弧 δφ，其高度增加了

$\delta\dot{y} \asymp \ddot{y}\,\delta t = \ddot{y}\,\delta s$。



[8.7] 如果粒子以单位速度运动，那么在 $\delta t = \delta s$ 时刻，**v** 的尖端跟踪单位圆的弧 $\delta\varphi$，其高度增加 $\delta\dot{y} \asymp \ddot{y}\,\delta s$。 两个阴影三角形的最终相似性可推出 (8.7)。

从两个阴影三角形的最终相似性，我们推断出

$$\frac{\delta\varphi}{\ddot{y}\,\delta s} \asymp \frac{1}{\dot{x}}$$

因此，回想 $\kappa \asymp (\delta\varphi/\delta s)$，（8.6） 简化为一个极其简单的公式，该公式（相当奇怪地）在标准课本中不容易找到：对于单位速度轨道，

$$\boxed{\kappa = \frac{\ddot{y}}{\dot{x}}} \tag{8.7}$$

同样，从相同的三角形，$\kappa = -\ddot{x}/\dot{y}$。 （请参阅例 1 以通过计算获得较少启发性的证明。）

8.5 示例：牛顿的曳物线

伪球面的恒定负高斯曲率实际上可以追溯到产生它的牛顿曳物线的曲率。

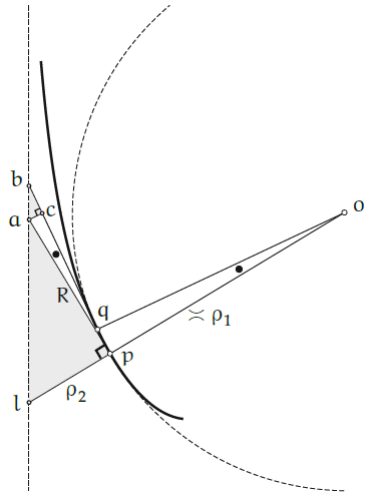
给定曳物线的参数表示（参见第 88 页的例 16），它的曲率可以通过基于 (8.6) 的常规计算找到[练习]。 然而，对问题的几何分析更加优雅，并以一种对研究伪球本身更有用的形式给出答案。 与其使用我们迄今为止开发的任何通用公式，不如提出一个针对该特定曲线量身定制的论证[103]。

令

[103] 先前发表于 VCA, p. 295.

$$\rho_1 = \text{生成曳物线的曲率半径}$$
$$\rho_2 = \text{从曳物线到其轴的法线段 pl}$$

如图 [8.8]所示。（稍后我们将解释 ρ_2 也是曲率半径，因此两个距离使用相同的希腊字母。）



[8.8] 曳物线有 $\kappa = (\rho_2/R^2)$

根据定义，该图中的曳物线具有恒定长度为 R 的切线。在相邻的点 p 和 q，图 [8.8] 说明了两个这样的切线，pa 和 qb，包含角•。因此相应的法线 po 和 qo 包含相同的角度•。注意，ac 已垂直于 qb 画出。

现在让我们看看当 q 与 p 合并时会发生什么，p 本身保持不变。在这个极限过程中，o 是曲率圆的圆心，pq 是这个圆的圆弧，ac 是一个以 p 为中心的半径为 R 的圆的圆弧。因此，

$$\rho_1 \asymp op, \qquad \frac{pq}{op} \asymp \bullet \asymp \frac{\mathrm{ac}}{\mathrm{R}}$$

所以

$$\frac{\mathrm{ac}}{\mathrm{pq}} \asymp \frac{\mathrm{R}}{\rho_1}$$

接下来我们求助于曳物线的定义属性$\mathrm{pa} = \mathrm{R} = \mathrm{qb}$ 即可导出

$$\mathrm{bc} \asymp \mathrm{pq}$$

最后，利用三角形 abc 最终与三角形 lap 相似的事实，我们推导出

$$\frac{R}{\rho_1} \asymp \frac{ac}{pq} \asymp \frac{ac}{bc} \asymp \frac{\rho_2}{R}$$

因此，

$$\boxed{\kappa = \frac{1}{\rho_1} = \frac{\rho_2}{R^2}} \tag{8.8}$$

第九章 3 维空间中的曲线

前面对平面曲线的分析只需稍作修改即可用于从平面逃逸并在 3 维空间中扭曲的曲线。

基本的洞察是，即使对于这样一个扭曲的 3 维轨道，每个无穷小的线段仍然位于一个平面（因此我们之前的分析依然适用），但是这个瞬时运动平面（称为*密切平面*，英文 osculating plane）在空间不再固定不动，而是在粒子沿曲线行进时围绕粒子的运动方向旋转。密切平面的旋转速率称为*挠率*（torsion），用 τ 表示。

对数学事实的直接物理体验是无可替代的：不要仅仅试图想象下面的实验——请去做！

请自行构造图 [9.1] 中所示的装置（如下所示）。 剪下一小段窄吸管，然后用胶带将其粘在一块硬纸板上（或其他任何平坦而坚硬的东西）上。 在硬纸板上画一个向量 **T** 延伸吸管段，并画一个等长的向量 **N** 垂直于吸管，如图所示。 最后构造第三个等长的向量并将其附加为平面的法向量（在图[9.1] 中标记为 **B**）。

[9.1] **Frenet 标架**。 将纸板压在弯曲的金属丝曲线上揭示出密切平面，**N** 指向该平面内的曲率中心。 当我们沿着曲线移动标架时，**挠率** τ 是平面法向量 **B**（称为**副法线**）绕 **T** 旋转的速率。

取一根坚固的电线（可能是金属衣架，切开），将其弯曲成任何你喜欢的非平面曲线。现在将电线穿过吸管部分，然后将纸板压在曲线上，使其成为密切平面。 一旦以这种方式贴合曲线，正交向量集 $(\mathbf{T},\mathbf{N},\mathbf{B})$ 称为曲线的 Frenet 标架。

将电线固定在空间中（最好在朋友的帮助下），沿着曲线滑动 Frenet 标架，同时尽可能保持纸板与弯曲的电线接触，从而确保它继续代表密切平面。 通过这种方式，你将体验必须旋转纸板以使其与弯曲的电线保持接触的速度 τ。

或者，如果你只对曲率 κ 如何沿曲线变化感兴趣，请用一只手固定纸板，用那只手的拇指将线压在平面上，以确保它是密切平面。 然后，你可以用另一只手逐渐将电线拉过吸管，同时用拇指将其压在纸板上，并观察你手中固定的纸板平面上曲线的曲率变化。

回到我们之前对平面曲线的讨论，回想一下曲线由 **T** 和 **N** 所张的平面，并进一步回想 **N** 的方向可以从 **T** 的变化率推导出，通过（8.5）：$\mathbf{T}' = \kappa\mathbf{N}$。 （回想一下，这里的

导数是关于弧长的，只有当粒子以单位速度运动时，它才与时间导数相同。）

在扭曲的 3 维曲线的当前情况下，我们可以将其扭转并将 **N** 定义为切线转向的方向：

$$\boxed{\mathbf{N} \equiv \frac{\mathbf{T}'}{|\mathbf{T}'|}}$$

这个法线 **N** 被称为*主法线*（principal normal），以区别于垂直于 **T** 的平面上的无数其他“法线”； 它的区别在于它位于密切平面内，并且它直接指向曲率中心（而不是远离它）。因此，比以前更明确一点的是，粒子在其中瞬间移动的密切平面是由 **T** 和 **T'** 所张的平面。

有了这些约定，单位速度粒子的加速度总是指向曲率中心，其大小就是曲率。 因此，可以将该加速度重新命名为该单位速度粒子的*曲率向量* $\boldsymbol{\kappa}$ ：

$$\boxed{\boldsymbol{\kappa} \equiv \kappa\,\mathbf{N}} \tag{9.1}$$

如图 [9.1] 所示，密切平面的空间方向方便地编码在其单位法向量中，表示为 **B**，并称为曲线的*副法线*（binormal）：

$$\boxed{\mathbf{B} \equiv \frac{\mathbf{T} \times \mathbf{T}'}{|\mathbf{T}'|} = \mathbf{T} \times \mathbf{N}}$$

正如我们所说，挠率是密切平面围绕运动方向 **T** 的旋转速率。等效地，它是 **B** 围绕 **T** 的转动速率。

我们注意到以下简单但基本的事实：

> *当单位向量开始旋转时，它的尖端在单位球面上（在该点单位球面的切平面内）移动，因此在垂直于向量本身的方向上移动。* (9.2)

当 **B** 旋转时，它的尖端因此必须开始在密切纸板 $(\mathbf{T},\mathbf{N})$ 平面的平行平面中移动。 但是请注意，这里 **B** 不能向 **T** 方向倾斜，因此它的变化率纯粹是在 **N** 方向； 确保你可以在几何上看到它（也许在你的玩具 Frenet 框架的帮助下）。这当然也可以通过计算来显示：参见练习 2。

因此

$$\boxed{\mathbf{B}' = -\tau\mathbf{N}}$$

注释：τ 定义中的负号在我们看来并无可取之处，但绝大多数作者都包含，在此我们谨慎起见仍旧保留！

至于曲率，上述分析只需很小的变化即可适用。 和以前一样，**T** 在密切平面内以 κ 给定的速率自旋，而 **N** 随之自旋，如图 [8.5] 中所示。 但是现在 **N** 不仅仅在密切平面内旋转，因为它也必须与 **B** 保持正交，因此它以与 **B** 围绕 **T** 旋转相同的速率旋转出密切平面。换句话说，应用图 [8.5] 所示几何思想两次，根据（9.2），**N** 的总变化率由下式给出

$$\mathbf{N}' = -\kappa\mathbf{T} + \tau\mathbf{B}$$

因此当我们沿着曲线移动时，整个 Frenet 标架的旋转可以总结为以下矩阵方程，称为弗勒内-塞雷（Frenet-Serret）方程：[104]

$$\boxed{\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix}' = \begin{bmatrix}0 & \kappa & 0\\ -\kappa & 0 & \tau\\ 0 & -\tau & 0\end{bmatrix}\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix} = [\Omega]\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix}} \tag{9.3}$$

让我们牢牢把握 $[\Omega]$ 结构背后的几何：消失的主对角线只是（9.2）的代数表现； 同样，反对称性 $([\Omega]^{\mathrm{T}} = -[\Omega])$ 只是图 [8.5] 的代数表现。

矩阵 $[\Omega]$ 本身告诉我们 $(\mathbf{T}, \mathbf{N}, \mathbf{B})$ 框架从一个时刻到下一个时刻如何旋转。 如果我们观察标架沿曲线移动一小段时间 δt，则

$$\delta t\text{ 后的新标架} \asymp [I + [\Omega]\,\delta t]\,[\text{原标架}]$$

有关标架旋转的更多信息，请参见练习 3。

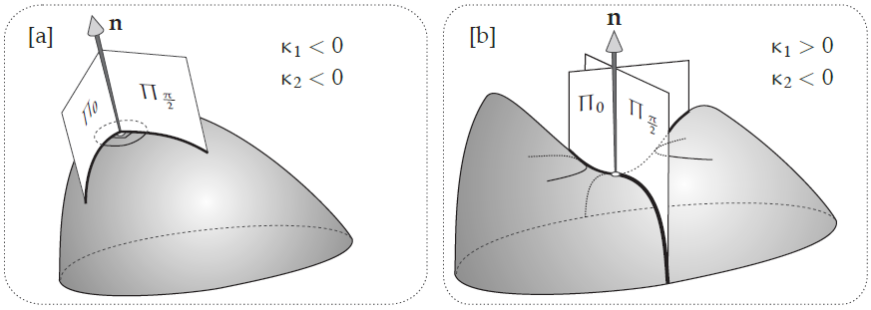
[104] 1847 年由 Frenet 和 1851 年由 Serret 独立发现。

第十章 曲面的主曲率

10.1 欧拉曲率公式

在第一幕中，我们迫不及待且不合时宜地开始讨论了高斯 1827 年革命性的曲面内蕴几何，以及相关的内蕴曲率 $\mathcal{K}$ 等概念。但是，从历史上看，牛顿对平面曲线外在曲率的研究的自然进展是对曲面的外在曲率的研究：即它们是如何在周围空间内弯曲的？

欧拉在 1760 年取得了第一个根本性的突破。我们将首先介绍他的发现，然后证明它。在曲面 $\mathcal{S}$ 上的一点 p 处，设想有平面 Π_θ 围绕 p 处的曲面法线 $\mathbf{n}_\mathrm{p}$ 旋转，考虑曲面 $\mathcal{S}$ 与平面 Π_θ 相交得到的平面曲线 $\mathcal{C}_\theta$。参见图 [10.1]，其中说明了该平面 Π_θ 在两种不同类型的曲面上的两个正交位置。 这里 θ 表示 Π_θ 从任意（至少现在）初始方向开始的旋转角度。 随着 Π_θ 旋转，相交曲线 $\mathcal{C}_\theta$ 的形状发生变化，因此其在 p 处的曲率 κ(θ)（通常）也会发生变化。



[10.1] [a] 如果 κ_1 和 κ_2 具有相同的符号，则曲面局部类似于小山，平行于（并接近）切平面的切片会产生椭圆（或与曲面不相交）。 [b] 如果 κ_1 和 κ_2 具有相反的符号，则曲面局部类似于马鞍。 切平面上方的平行切片产生双曲线的两个分支（如图所示）； 切平面下方的切片产生正交双曲线的两个分支（未显示）。

在继续之前，我们应该解释 κ(θ) 附有一个符号，根据这个约定：从 p 到 $\mathcal{C}_\theta$ 的曲率中心 c 的向量被定义为 $\frac{1}{\kappa(\theta)}\mathbf{n}$。 因此，如果 c 位于 $+\mathbf{n}$ 的方向，则 κ(θ) 为正，而如果它位于 $-\mathbf{n}$ 的方向，则 κ(θ) 为负。 当然，$\mathbf{n}$ 实际上有两个相反（同样有效）的选择。 反转 $\mathbf{n}$ 的选择会反转 κ 的符号。注意，因此 $\mathcal{C}_\theta$ 的主法线（如上一节中所定义）$\mathbf{N} = [\kappa(\theta)$ 的符号$]\,\mathbf{n}$。

随着 θ 的变化，让 κ_1 和 κ_2 表示 κ(θ) 的最大值和最小值。 欧拉[105]优雅而重要的发现是，这些曲率的极值（所谓的**主曲率**）总是会出现在垂直的方向上，也就是所谓的**主方向**。此外，如果我们选择 $\theta = 0$ 以与曲率 κ_1 的方向重合，他发现

[105] 更熟悉的形式 (10.1) 是大约 50 年后由 Charles Dupin 在 1813 年从欧拉的结果推导出来的。参见 Knoebel (2007, p. 188)。

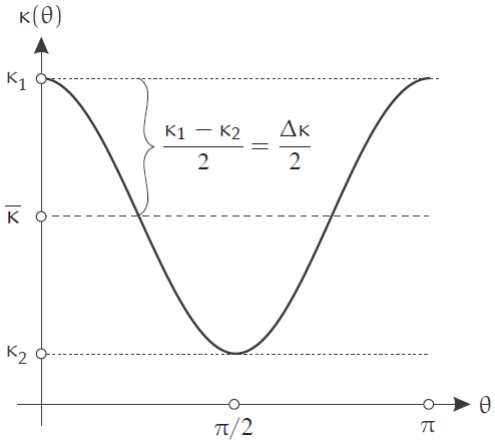
$$\boxed{\text{欧拉曲率公式：}\kappa(\theta) = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta} \tag{10.1}$$

虽然这是在大多数现代教科书中该公式的标准形式，但有一种更好的书写方式，这实际上是欧拉本人首次发表它的形式。

代入 $\cos^2\theta = (1+\cos 2\theta)/2$ 和 $\sin^2\theta = (1-\cos 2\theta)/2$，我们得到

$$\boxed{\kappa(\theta) = \bar{\kappa} + \frac{\Delta\kappa}{2}\cos 2\theta} \tag{10.2}$$

其中 $\bar{\kappa} \equiv [\frac{\kappa_1+\kappa_2}{2}]$ 是**平均曲率**，值围绕该曲率振荡，$(\Delta\kappa/2) \equiv (\kappa_1 - \kappa_2)/2$ 是振荡幅度。参见图 [10.2]。 注意，可以直接从这个公式推导出 κ_1 和 κ_2 的极值性质，及主方向的正交性。 此外，这种形式还体现了将法平面旋转 π 使其返回到同一位置的几何事实，因此曲率的变化具有周期 π。



[10.2] 欧拉曲率公式告诉我们，当我们改变法向截面 $\mathcal{C}_\theta$ 的角度 θ 时，其曲率 κ(θ) 呈正弦振荡，在相互垂直的方向上达到其最大值和最小值。

如图 [10.1a] 所示，如果 κ_1 和 κ_2 具有相同的符号，则 κ(θ) 始终具有相同的符号；即，曲面局部类似于小山。 但是如果κ_1和κ_2符号相反，则 κ(θ) 改变符号； 即，$\mathcal{C}_\theta$ 从切平面的一侧翻转到另一侧，并且如 [10.1b] 所示，曲面局部类似于马鞍。

当然，$\kappa_1 = \kappa_2$ 是可能的，这种情况下在p 点有 $\kappa(\theta) = \text{const}$; 这时，p 称为**脐点**，曲面局部类似于球体。 一般来说，可以证明这种脐点只能出现在曲面的孤立位置。 当然，半径为 R 的球体是一个强烈的例外：每个点都是一个脐点，$\kappa(\theta) = 1/R$。 但可以证明，球体是唯一具有这种性质的曲面。

在脐点具有 $\kappa_1 = 0 = \kappa_2$ 的特殊情况下，该点周围的曲面形状可能要复杂得多。 这将在第 12.4 节中讨论和说明。

10.2 欧拉曲率公式的证明

我们现在将提供欧拉曲率公式的几何证明[106]。 选择 p 作为笛卡尔 $(\mathrm{x},\mathrm{y},\mathrm{z})$ 坐标的原点，并让 x 和 y 轴位于 p 处的切线平面 $\mathrm{T_p}$ 中。 然后局部曲面可以由形式为 $\mathrm{z}=\mathrm{f}(\mathrm{x},\mathrm{y})$ 的方程表示，使得在原点 $\mathrm{f}(0,0)=0$ 和 $\partial_{\mathrm{x}}\mathrm{f}=0=\partial_{\mathrm{y}}\mathrm{f}$。 将 $\mathrm{f}(\mathrm{x},\mathrm{y})$ 展开成泰勒级数，我们推导出当x 和 y 趋于零时

$$\mathrm{z} \asymp \mathrm{a}x^2+\mathrm{b}y^2+\mathrm{c}xy \tag{10.3}$$

用平行于 $\mathrm{T_p}$ 且非常接近 $\mathrm{T_p}$ 的平面 $\mathrm{z}=\mathrm{const.}=\mathrm{k}$ 切开曲面，因此产生相交曲线，其方程（当 k 趋于零时）是二次方程，$\mathrm{a}x^2+\mathrm{b}y^2+\mathrm{c}xy \asymp \mathrm{k}$，因此它们（最终）是圆锥曲线。

图 [10.1] 说明了这样一个事实，如果 κ_1 和 κ_2 具有相同的符号，则这些圆锥曲线是椭圆，如果 κ_1 和 κ_2 具有相反的符号，则它们是双曲线。 在这两种情况下，圆锥曲线都有两个垂直的对称轴，它们与切片平面的高度 k 无关。这是从方程的齐次二次性质得出的。例如，将切片的高度增加四倍只会使圆锥的大小增加一倍，而不会改变其形状：$\mathrm{k}\to 4\mathrm{k}$ 和 $(\mathrm{x},\mathrm{y})\to(2\mathrm{x},2\mathrm{y})$ 产生与以前相同的方程。

1813 年，夏尔• 迪潘（Charles Dupin）使用这个圆锥来量化 $\mathcal{S}$ 在 p 处的弯曲量和类型，并以他的名义将其称为**迪潘标形**（Dupin indicatrix）。 根据 p 处 迪潘标形的类型，点 p 本身被称为椭圆型、双曲线型或抛物线型。

至关重要的是，圆锥截面的对称性意味着曲面本身在两个垂直方向上具有局部镜像对称性。[107] 我们现在可以推导出欧拉曲率公式并推导出这两个垂直的对称平面实际上是产生最大和最小曲率的平面； 即，这些局部镜像对称方向与主方向相同。 总结一下我们将要证明的内容，

> 曲面上一般点的无穷小邻域在两个垂直平面（均包含曲面法线）中具有镜像对称性，并且这些平面与切面相交的垂直方向是最大和最小曲率的主方向。 (10.4)

细化我们的坐标系，我们现在将 x 轴和 y 轴与这些对称方向对齐。由于方程（10.3）现在在反射 $\mathrm{x}\mapsto-\mathrm{x}$ 和 $\mathrm{y}\mapsto-\mathrm{y}$ 下不变，因此 $\mathrm{c}=0$，因此曲面的局部方程变为

$$\mathrm{z} \asymp \mathrm{a}x^2+\mathrm{b}y^2 \tag{10.5}$$

为了找到系数 a 和 b 的几何意义，我们现在参考图 [8.2] 并将其视为描绘 Π_θ 与 $\mathcal{S}$ 的交点：曲线 $\mathcal{C}$ 现在是 $\mathcal{C}_\theta$，切线 $\mathcal{T}$ 现在是切平面的交点 $\mathrm{T_p}$ 与 Π_θ，曲线与其切线的偏

---

[106] Aleksandrov (1969) 的可爱文章中的证明是类似的，但它需要两次计算，这里用几何代替。 使用相对于切平面的二次近似的现在标准思想比欧拉的原始方法更简单。 它是由 Jean-Baptiste Meusnier 于 1776 年发现的。参见 Knoebel (2007, p. 194)。

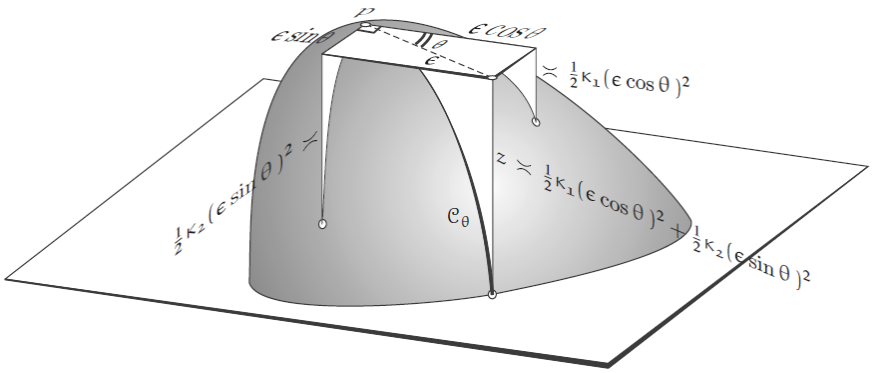
[107] 至少在一般情况下，其中 $\kappa_1\neq\kappa_2$： 正如我们将在第 12.4 节中看到的，在脐点的情况下，对称性可能要复杂得多。

差 σ 现在只是曲线在切平面上方的高度 z。

设 $\theta = 0$ 对应于 x 轴，设 $\kappa_1 = \kappa(0)$ 为 $\mathcal{C}_0 = (\mathcal{S}$ 与 xz 平面的交点$)$ 的曲率，方程为 $z = ax^2$。 那么结果（8.2）表明$a = \frac{1}{2}\kappa_1$。 以完全相同的方式，定义$\kappa_2 = \kappa(\frac{\pi}{2})$ 为与 yz 平面相交曲线的曲率，我们发现$b = \frac{1}{2}\kappa_2$。 因此（10.5）可以更几何地表示为

$$\boxed{z \asymp \frac{1}{2}\kappa_1 x^2 + \frac{1}{2}\kappa_2 y^2} \tag{10.6}$$

现在考虑图 [10.3]，它描绘了一般角度 θ 的曲线 $\mathcal{C}_\theta$。（该图假设（以及其他后续图）高斯曲率是正的，但伴随的推理同样适用于负曲面。）如果我们在 T_p 内沿方向 θ 移动距离 ϵ 则我们到达 图示点 $x = \epsilon\cos\theta, y = \epsilon\sin\theta$。 因此将（10.6）插入（8.2）产生



[10.3] 主曲率告诉我们，当我们开始沿两个垂直的主方向中的每一个行进时，曲面从切平面下降的速度（在此图中为向下）。 当我们在切平面内沿大致方向 θ 移动距离 ϵ 时，曲面下降的距离 z 只是由于这两个分量中的每一个单独造成的下降的总和（根据（10.6））。

证明欧拉曲率公式，从而确定与局部镜像对称的正交方向相关的曲率 κ_1 和 κ_2 的极值性质。

10.3 旋转曲面

如果我们在平面曲线 $\mathcal{C}$ 的平面内绕直线 L 旋转 $\mathcal{C}$，则我们得到一个旋转曲面 $\mathcal{S}$，其主方向很容易识别。 图 [10.4] 在特定情况下说明了这一点，其中 $\mathcal{C}$ 是曳物线，L 是其轴，因此 $\mathcal{S}$ 是伪球体。

很明显，$\mathcal{S}$ 在通过其轴 L 的每个平面上都具有镜像对称性，因此这样一个平面与 $\mathcal{S}$（它只是 $\mathcal{C}$ 的副本）的相交会产生 $\mathcal{S}$ 内的一个主要方向。

在 $\mathcal{S}$ 上的点 p 让我们沿着 $\mathcal{C}$ 的副本选择这个方向以对应于 $\theta = 0$，以便使用先前的符号 $\mathcal{C} = \mathcal{C}_0$。 因此，第一主曲率 $\kappa_1 = (1/\rho_1)$ 只是（最多符号）原始平面曲线 $\mathcal{C}$ 的曲率。

因此，p 处的第二个主方向必须是 $\mathcal{S}$ 内垂直于通过 L 的该平面的方向。 与该第二个主方向相关的曲率半径 ρ_2 实际上只是距离 pl，在曲面法线 **n** 的方向上， 从 p 到 L 轴上的点 l：

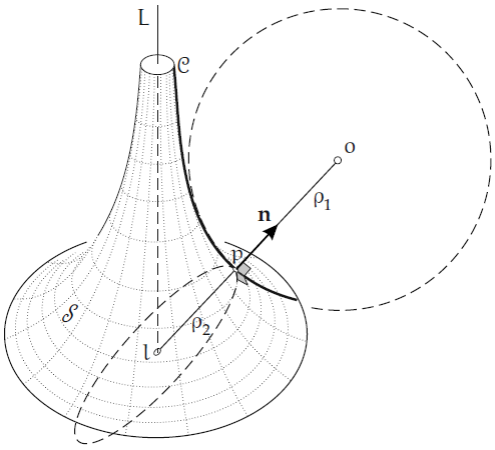
$$\rho_1 \asymp \text{生成曲线}\,\mathcal{C}\,\text{的曲率半径 op} \tag{10.7}$$
$$\rho_2 \asymp \text{沿法线}\,\mathbf{n}\,\text{到轴 L 的距离 pl} \tag{10.8}$$

为了证实这一点，让我们使用与之前相同的符号，并将 $\mathcal{S}$ 与该垂直平面的交点表示为 $\mathcal{C}_{(\pi/2)}$。 如果我们围绕 L 旋转 lp，那么它会扫出一个圆锥体，并且 p 在 $\mathcal{S}$ 内沿着这个圆锥体的圆形边缘移动，它最初与 $\mathcal{C}_{(\pi/2)}$ 重合。

现在回想一下牛顿的原始方法，将曲率中心定位为曲线相邻法线的交点。 在 $\mathcal{C}_{(\pi/2)}$ 的当前情况下，相邻法线是该锥体的母线，在 l 处相遇，从而证明 $\kappa_2 = \pm(1/\rho_2)$，符号取决于 **n** 的选择。

返回到图 [10.4]，根据法向量 **n** 的图示选择，凭借（8.8）我们发现对于

$$\text{伪球：}\ \kappa_1 = +\frac{1}{\rho_1}, \qquad \kappa_2 = -\frac{1}{\rho_2} = \frac{\rho_1}{R^2} \tag{10.9}$$

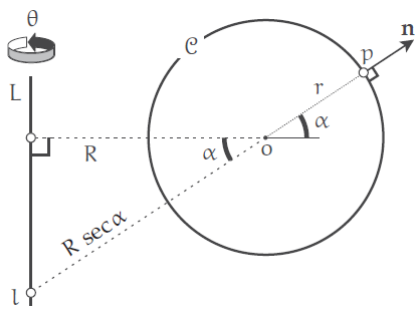


[10.4] 旋转曲面的主曲率半径 ρ_1 和 ρ_2，此处用伪球体表示。

（当然，一般来说，反转 **n** 的方向会反转两个主曲率的符号。）注意，我们现在已经兑现了我们的承诺并解释了我们在图 [8.8] 中看到的符号 ρ_2。

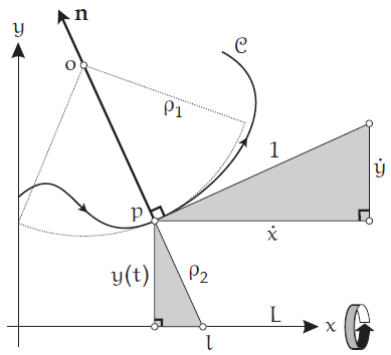
对于我们的第二个示例，考虑通过在距 o 的距离为 R 处绕直线 L 旋转半径为 r 和中心为 o 的圆 $\mathcal{C}$ 获得的环面（甜甜圈）（见图[10.5]）。我们假设$\mathcal{C}$被一个绕 o 旋转的粒子 p 跟踪，半径 op 与水平面成 α 角。 从图中，通过图示的 **n** 选择，我们看到对于这个

$$\text{环面：}\ \kappa_1 = -\frac{1}{r}, \qquad \kappa_2 = -\frac{1}{r + R\sec\alpha} \tag{10.10}$$



[10.5] 对于通过围绕 L 旋转圆 $\mathcal{C}$ 获得的环面，主曲率半径是 $\rho_1 = -r$ 和 $\rho_2 = -(r + R\sec\alpha)$。

抛开具体的例子，现在假设 $\mathcal{C}$ 是一条一般曲线，但由在 (x, y) 平面中运动的粒子以单位速度追踪，其在时间 t 的位置为 $x = x(t)$ 和 $y = y(t)$（见图 [10.6]）。



[10.6] 由 $\mathcal{C}$ 绕 L（x 轴）旋转产生的旋转曲面的主曲率半径 ρ_1 和 ρ_2。

让我们将 L 作为水平 x 轴，并围绕该轴旋转 $\mathcal{C}$ 以生成曲面 $\mathcal{S}$。让我们选择指向运动方向左侧的 **n**，如图所示。 使用（10.7）和（8.7），我们推导出跟踪在的曲线的第一主曲率

$$\boxed{\text{单位速度：}\ \kappa_1 = \ddot{y}/\dot{x}} \tag{10.11}$$

正如我们之前提到的（8.7），结果（10.11）同样可以表示为 $\kappa_1 = -\ddot{x}/\dot{y}$。 无论我们在曲线上的哪个位置，请自行检查这些公式为 κ_1 得出了正确的符号[练习]。

最后，根据（10.8），$\rho_2 = pl$。 通过利用图 [10.6] 中两个阴影三角形的相似性，我们发现 $(y/\rho_2) = (\dot{x}/1)$，因此我们推导出轨迹的第二个主曲率

$$\boxed{\text{单位速度：}\ \kappa_2 = -\dot{x}/y} \tag{10.12}$$

无论我们在曲线上的哪个位置，请自行检查该公式为 κ_2 得出了正确的符号[练习]。

最后，注意到上述分析收获的重要一课是：

让曲线 $\mathcal{C}$ 绕直线 L 旋转以生成旋转曲面 $\mathcal{S}$。然后 $\mathcal{C}$ 中向 L 凹的部分生成 $\mathcal{S}$ 上具有正曲率的部分，而 $\mathcal{C}$ 中向远离 L 方向凹的部分生成 $\mathcal{S}$ 上具有负曲率的部分。 $\mathcal{C}$ 的拐点生成 $\mathcal{S}$ 上曲率消失的圆； 这些圆将相反曲率的区域分开。 (10.13)

第十一章 平面曲线的曲率

11.1 测地曲率和法曲率

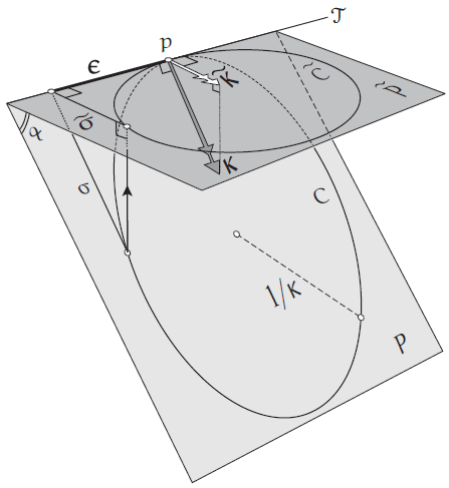
对于我们（地球表面的居民）来说，大圆不仅在“给出两点之间最短路径”这点上类似于一条直线，它看起来也是直的（没有明显的曲率）。 如果你沿一条“直线”穿过一个看似平坦的沙漠，你实际上是在沿着这样一个大圆走，它在 3 维空间中的曲率是1/(地球的半径)。 我们该如何调和对同一条曲线的这两种相互矛盾的观点呢？

简而言之，答案是 3 维空间中一般曲面 $\mathcal{S}$ 上的一般曲线之完整曲率可以分解为两个分量：一个在曲面内（对其居民可见），另一个垂直于曲面（对其居民不可见）。 曲面内的可见分量称为**测地曲率**，记为 κ_g； 垂直于曲面的不可见分量称为**法曲率**，记为 κ_n。

如果密切平面垂直于曲面，即包含曲面法线 $\mathbf{n}$，则所有曲率都是“法向曲率”（$\kappa_n = \kappa$），并且在曲面内没有一个可见的“测地曲率”（$\kappa_g = 0$）。 我们在地球表面的大圆就是这种情况。

现在假设你正站在一个看似平坦的沙漠中间，你在脚下的沙地上划了一个半径为 r 的小圆。 因此，你将获得一条看似平坦的曲线，曲率 $\kappa = 1/r$ 非常大。 但当然，沙漠是地球曲面的一部分，从这个角度来看，你的曲线的特别之处在于它的密切平面几乎与你所在位置的地球切平面重合。 在这种情况下，我们几乎与前一种情况相反：现在几乎所有的曲率都是测地曲率，但实际上仍然存在与以前同样大小的不可见法曲率（我们将在稍后解释）。

事实上，一般情况只是前面描述的两种极端的适当混合。



[11.1] 当平面 P 中的圆 C 正交投影到与 P 成角度 α 的平面 $\tilde{P}$ 上时，到公切线 $\mathcal{T}$ 的距离收缩 $\cos\alpha$，p 处的曲率向量大小也同步收缩： $\tilde{\kappa} = \kappa\cos\alpha$。

这里的关键点实际上与曲面本身无关，而是与平面曲线在另一个平面上投影时的曲率如

何变化有关。 参见图 [11.1]，其中画出了平面 P 内半径为 (1/κ) 的圆 C，p 处的切线为 $\mathcal{T}$。该图还画出了 C 在第二个经过 $\mathcal{T}$ 的平面 $\tilde{P}$ 上的正交投影 $\tilde{C}$，两个平面之间的夹角为 α。 你可能知道 $\tilde{C}$ 实际上是一个椭圆[练习]，原始圆 C 已在垂直于 $\mathcal{T}$ 的方向上被压缩。因此很明显 $\tilde{C}$ 在 p 处的曲率 $\tilde{\kappa}$ 小于原始曲率 κ。

更准确地说，该图画出在此投影下，到公切线 $\mathcal{T}$ 的距离压缩到原来的 $\cos\alpha$，因此 $\tilde{\sigma} = \sigma\cos\alpha$。 因此，借助于（8.2）可知，

$$\tilde{\kappa} \asymp \frac{2\tilde{\sigma}}{\epsilon^2} = \frac{2\sigma\cos\alpha}{\epsilon^2} \asymp \kappa\cos\alpha$$

通过将 C 视为一般曲线 $\mathcal{C}$ 的曲率圆，我们看到该公式也适用于 $\mathcal{C}$：

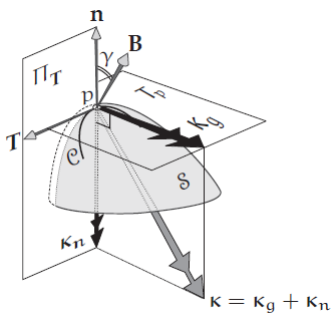
> 如果平面曲线 $\mathcal{C}$（在 p 处有切线 $\mathcal{T}$）垂直投影到经过 $\mathcal{T}$ 并与第一个平面成角度 α 的第二个平面上，则在 p 处投影曲线 $\tilde{\mathcal{C}}$ 有 $\tilde{\kappa} = \kappa\cos\alpha$。 (11.1)

这个结果也可以从运动的角度来理解（可能更直观）。 正如我们已经讨论过的，如果一个粒子以单位速度沿 $\mathcal{C}$ 运动，那么可以将粒子的加速度重命名为曲率向量 $\boldsymbol{\kappa} = \kappa\mathbf{N}$，这个加速度向量指向曲率中心且它的大小等于曲率。

接下来注意到投影的加速度等于加速度的投影 [练习]。 但是注意，粒子的投影通常不会以恒定速度沿投影轨道 $\tilde{\mathcal{C}}$ 运动，于是其加速度向量将具有与 $\tilde{\mathcal{C}}$ 相切的分量，即与 $\tilde{\kappa}$ 相反，不会与 $\tilde{\mathcal{C}}$ 正交（记住，$\tilde{\kappa}$ 是粒子以恒定单位速度穿过 $\tilde{\mathcal{C}}$ 时的加速度）。这种正交性的分解在几何上也说得通：在图 [11.1] 中，C 的半径（通常）投影到 $\tilde{P}$ 中的一条直线不再与像椭圆 $\tilde{\mathcal{C}}$ 正交。 也就是说，p 是一个例外，在粒子和它的投影沿 $\mathcal{T}$ 以单位速度一起移动并经过 p 的瞬间，

$$\tilde{\kappa} = (\kappa \text{ 在 } \tilde{P} \text{ 上的投影})$$

（如图 [11.1]所示）。 通过使这些向量的长度相等，立即得到（11.1）。



[11.2] 净加速度 κ 可以分解为与曲面相切的测地曲率向量 κ_g 和垂直于曲面的法曲率向量 κ_n：$\kappa = \kappa_g + \kappa_n$。

现在让我们回到最初的问题，其中 $\mathcal{C}$ 是曲面 $\mathcal{S}$ 上的一般曲线（参见图 [11.2]）。 和以前一样，令 $\mathbf{T}$ 表示在 p 处与 $\mathcal{C}$ 相切的单位长度切线，T_p 为 $\mathcal{S}$ 在 p 处的切平面，引入 Π_T 作为 $\mathbf{n}$ 和 $\mathbf{T}$ 所张成的法平面。设倾斜的密切平面与切平面 T_p 的夹角为 γ； 等效地，γ 也是 $\mathcal{C}$ 的副法线 $\mathbf{B}$ 和 $\mathcal{S}$ 的曲面法线 $\mathbf{n}$ 之间的夹角。因此，κ_g 和 κ_n 分别是 $\mathcal{C}$ 在 T_p 和 Π_T 上的投影的曲率。 那么（11.1）意味着

$$\boxed{\kappa_g = \kappa\cos\gamma, \qquad \kappa_n = \kappa\sin\gamma} \tag{11.2}$$

这也可以从加速度的角度来理解。 想象粒子以单位速度经过 $\mathcal{C}$ 在 T_p 和 Π_T 上的投影。 它们各自的加速度将是*测地曲率向量* κ_g 和*法曲率向量* κ_n，指向切平面和法平面中各自的曲率中心，大小分别等于 κ_g 和 κ_n。 如图 [11.2] 中所示，加速度可以分解为这两个正交分量：

$$\boxed{\kappa = \kappa_g + \kappa_n} \tag{11.3}$$

从上式立即可得（11.2）。

11.2 Meusnier 定理

考虑曲面 $\mathcal{S}$ 上沿特定方向 $\mathbf{T}$ 通过 p 的所有曲线（例如 $\mathcal{C}$）的族。 $\mathcal{S}$ 的居民可以自由地画出这种类型的曲线，这些曲线是紧紧弯曲的、略微弯曲的或根本不弯曲的：κ_g 可以自由和任意地指定。 然而，法曲率 κ_n *并非如此*：曲面本身的弯曲迫使其内的所有曲线沿 $\mathbf{n}$ 方向弯曲。

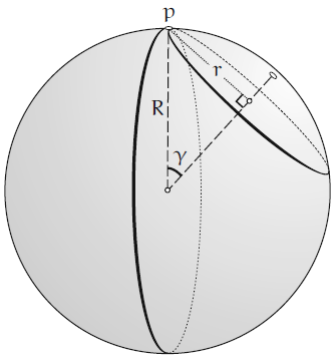
事实上，正如 Meusnier 在 1779 年发现的那样，曲面迫使所有这些曲线弯曲*相同*的量：κ_n 与特定曲线 $\mathcal{C}$ 无关，因此，特别是，κ_n 必须等于该方向上法向截面的曲率。 更直观地说，这表示所有这些曲线局部在 $\Pi_{\mathbf{T}}$ 上具有*相同的投影*：在 p 附近，这些投影都看起来像法线截面，而法线截面又看起来像 $\Pi_{\mathbf{T}}$ 内以 $(1/\kappa_n)\mathbf{n}$ 为中心，半径为 $(1/\kappa_n)$ 的圆。

然而即使如此，很明显，这种与曲线无关的法曲率通常必须取决于曲线族通过 p 时的方向 $\mathbf{T}$。 因此，我们可以将族的共同法向曲率写为该方向的函数 $\kappa_n(\mathbf{T})$。（例如，如果 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 是主方向，则 $\kappa_n(e_{1,2}) = \kappa_{1,2}$ 是主曲率。）将这个（至今还）未经证实的主张与（11.2）结合起来，我们可以说

> Meusnier 定理。 *在同一方向 $\mathbf{T}$ 上通过曲面上点 p 的所有曲线都具有与该方向上的法线截面相同的法向曲率 $\kappa_n(\mathbf{T})$。 如果一条这样的曲线在 p 处的密切平面与 p 处的切平面所成角度为 γ，且其在此处的曲率为 κ_γ，则 $\kappa_\gamma\sin\gamma = \kappa_n(\mathbf{T})$ 与 γ 无关。* (11.4)

在给出一般论证之前，考虑半径为 R 的球体，因为在这种特殊情况下，定理的真实性很容易形象化（见图 [11.3]）。 以 p 为北极，法向截面是半径为 R 的大圆（子午线），因此曲率 $\kappa_n = (1/R)$。 与 p 处的水平切面成角度 γ 且经过北极的平面与球面的交线是一个

半径为 $r = R\sin\gamma$ 的圆，因此 $\kappa_\gamma = 1/r = 1/(R\sin\gamma)$。



[11.3] 将半径为 R 的球面用经过北极 p 与水平面成角度 γ 的平面切割，得到半径为 $r = R\sin\gamma$ 的圆。

回到在沙子里画一个小圆的介绍性例子，想象 γ 趋于零，在这种情况下，圆的半径缩小到零，它的曲率向量的长度趋于无穷大。 但同时这个曲率向量趋向于与曲面法线正交，因此它越来越少地投射到该方向上。 这两种效应完全相互抵消，因此曲率向量在曲面法线上的投影具有恒定的大小：

$$\kappa_\gamma \sin\gamma = \frac{1}{R} = \kappa_n$$

证实了定理的这个例子。

对于一般情况，想象粒子以单位速度沿 $\mathcal{C}$ 运动，令 $\mathbf{T}$ 表示它的速度向量（沿着整个轨道，不仅在 p 处）。 然后由（11.3），

$$\dot{\mathbf{T}} = \kappa = \kappa_g + \kappa_n \Longrightarrow \kappa_n(\mathbf{T}) = \dot{\mathbf{T}} \cdot \mathbf{n}$$

出于与图 [8.5] 中相同的本质原因，$\dot{\mathbf{T}} \cdot \mathbf{n} = -\mathbf{T} \cdot \dot{\mathbf{n}}$，所以

$$\kappa_n(\mathbf{T}) = -\mathbf{T} \cdot \dot{\mathbf{n}} \tag{11.5}$$

$\dot{\mathbf{n}}$ 是 $\mathbf{T}$ 方向曲面法线的变化率，与 $\mathcal{C}$ 无关，从而证明定理。

11.3 测地线是“直的”

我们已经通过曲面内的长度最小的特性定义了测地线。 但是欧氏平面（直线）的测地线也可以通过它们的直性（straightness）来识别。 同样，在球面上，我们刚刚看到测地线（大圆）不仅给出最短路线，而且它们也是“直的”，从某种意义上说，它们的曲率对地表的居民来说是不可见的：$\kappa_g = 0$ 。事实上，长度最小化和内蕴直性之间的这种联系是普遍的：

测地线对于地表的居民来说看起来是直线：它们内蕴上是“直的”，因为它们的测地曲率为零：即在测地线的每个点上 $\kappa_g = 0$。

为了理解这一点，想象一根吉他弦在其直线平衡位置 L 被一个空间位置为 p 的拨片拨动。 当拨片将琴弦拉离 L 时，琴弦在包含 p 和 L 的平面 Π 中形成一个三角形。一旦松开，由于琴弦被迫缩短其长度而产生的合力作用在 Π 内，由此产生的运动返回 L 都发生在这个平面内。 接下来，假设不是将弦从 L 拉到 Π 中的锐角三角形，而是将其拉成位于 Π 中的平缓凸曲线的形式，那么所产生的力和运动显然仍会在平面 Π 中。

现在让我们回到我们的曲面并将测地线想象成一根吉他弦在曲面上（无摩擦地）拉紧以连接两个固定点。 绳子处于静止状态：它在曲面上处于平衡状态，已经滑过曲面，收缩到尽可能短。 现在将注意力集中在测地线的一个非常短的段 apb 上。 该线段几乎位于单个平面 Γ 中，即通过 a、p 和 b 的平面。 当 a 和 b 接近 p 时，Γ 的极限位置是 p 处的密切平面，我们将其表示为 Π_p。 根据上述推理，作用在 apb 上的净长度收缩力 $\mathbf{F}_p$ 最终将位于 Π_p 之内。

如果 $\mathbf{F}_p$ 具有与曲面相切的任何分量，则弦将在此方向上自由移动，从而减少其长度。但由于绳子已经尽可能短，这不可能发生！ 因此，在曲面内没有分量，力 $\mathbf{F}_p$（位于 Π_p 内）必须垂直于曲面，沿着法线 $\mathbf{n}_p$。 因此，我们已经根据曲面的外在几何得到测地线的一个重要的表征[108]：

在测地线的每一点 p 处，密切平面 Π_p 包含曲面法线 $\mathbf{n}_p$，因此测地曲率消失：$\kappa_g = 0$。 (11.6)

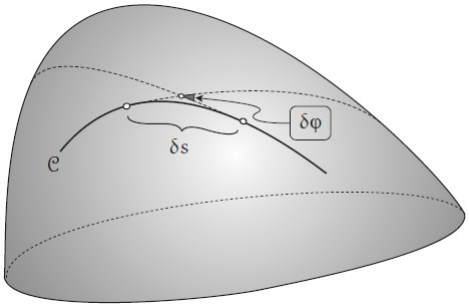
内蕴和外在几何似乎属于完全不同的世界，但我们在这里看到两者奇怪地纠缠在一起。稍后我们将见证这两个世界之间更深、更神秘的联系。

11.4 测地曲率的内蕴测量

我们刚刚看到消失的测地曲率表征测地线，我们也可把测地线当做测量非测地曲线（其中$\kappa_g \neq 0$）测地曲率的内蕴工具，。

在（8.4）中，我们看到欧氏平面中曲线的曲率可以看作是其切线相对于沿曲线距离的旋转率。 嗯，这个原始结构（如图 [8.4] 中所示）对地表的居民也非常有意义。

[108] 这种表征首先由约翰·伯努利在 1697 年发现，然后他把它教给了他的学生欧拉。



[11.4] 曲面的居民可以通过在相距 δs 的相邻点处构造曲线的切线测地线，然后找到这些切线相交的角度 $\delta\varphi$，从而测量曲线 $\mathcal{C}$ 的测地曲率 κ_g。 随着两点趋近并合二为一，$\kappa_g \asymp (\delta\varphi/\delta s)$。

从他们的角度来看，构造根本没有变化（见图[11.4]）。 他们在相距 δs 的相邻点处画出曲线的切线（虚线测地线），找到这些切线相交的角度 $\delta\varphi$，然后计算曲率（在点合并的范围内）为 $\kappa_g \asymp (\delta\varphi/\delta s)$.

但是他们所谓的曲线曲率，我们认为只是曲率的一部分，即测地曲率 κ_g； 法曲率 κ_n 对他们来说既不可见也不可知。 他们的内蕴视角和我们的外在视角（俯视曲面）之间的唯一另一个区别是，他们所谓的“直线”，在我们看来是曲面内的测地线。

当然，如果 $\mathcal{C}$ 本身是测地线，那么两条切线都与 $\mathcal{C}$ 重合，因此 $\delta\varphi = 0$，因此 $\kappa_g = 0$，正该如此。

11.5 测量测地曲率的一种简单外在方法

在第一幕中，我们讨论了这样一个事实：如果我们从曲面上剥离以测地线 G 为中心的窄条，那么当该条放在平面上时，它会变成一条直线 $\tilde{G}$（参见图 [1.11]）。因此 G 在曲面上的内蕴直性 $(\kappa_g = 0)$ 表现为平面中的普通直性：扁平条带 $\tilde{G}$ 的曲率 $\tilde{\kappa}$ 消失了。

如果我们同样从曲面剥去一条以非测地曲线 $\mathcal{C}$ 为中心的窄带，其 $\kappa_g \neq 0$，然后将其展平到平面上，我们得到一个平面曲线 $\tilde{\mathcal{C}}$，其普通曲率 $\tilde{\kappa} \neq 0$（请参阅图 [1.12]）。

那么，曲面上特定点 p 处的测地曲率 $\kappa_g(p)$ 与平面曲线 $\tilde{\mathcal{C}}$ 中相应点 $\tilde{p}$ 处的扁平条带的曲率 $\tilde{\kappa}(\tilde{p})$ 有何关系呢？

令 $\kappa_g(p)$ 表示曲面上曲线 $\mathcal{C}$ 在 p 处的测地曲率，令 $\tilde{\kappa}(\tilde{p})$ 表示当以 $\mathcal{C}$ 为中心的窄带被剥离曲面，压平在平面上得到平面曲线 $\tilde{\mathcal{C}}$ 时对应点处的曲率。

则

$$\kappa_g(p) = \tilde{\kappa}(\tilde{p})$$

要看到这一点，请用图 [11.4] 的内蕴构造并想象从曲面以 $\mathcal{C}$ 以及（以角度 $\delta\varphi$ 相交

的切线测地线的）两个小虚线段为中心剥离窄条。当连接剥皮条的小三角形在平面上平放时，我们又回到了原来的构造图 [8.4]来测量曲率。摊平的测地线在平面上变成了直线，由于这些线仍然与 $\tilde{\mathcal{C}}$ 相切，所以它们是切线。但是 δs 和 $\delta\varphi$ 在展平过程中都没有改变，因此 $\kappa_g(p) \asymp (\delta s/\delta\varphi) \asymp \tilde{\kappa}(\tilde{p})$，如前所述。

11.6 测地线粘胶带构造的新解释

我们最初使用测地线的长度最小化特性来解释构造（1.7），其中测地线是通过从我们的选择任意点开始，将一条窄胶带向下卷到曲面上，并沿任意方向前进来构造的。 但是，正如我们一直在讨论的，测地线也可以用它们的*直性*来表征：即测地曲率为零。 我们现在使用这个属性来为我们的测地线构造提供一个新的、第二个解释。

考虑一条窄而直的胶带 $\mathcal{L}$，平放在平面上，中心线为 L。线 L 在 $\mathbb{R}^3$ 中是外在的直线；即$\kappa = 0$。 它在 $\mathcal{L}$ 上内蕴地也是直的：$\kappa_g = 0$。 现在让我们拿起 $\mathcal{L}$ 作等距弯曲并在空间中将其扭曲成我们喜欢的任何形状。 无论 $\mathcal{L}$ 现在采用什么新形式，中心线 L 内蕴的仍是直线 $\kappa_g = 0$。 因此曲率向量沿 L 与条带垂直：

$$\kappa = \kappa_g + \kappa_n = \kappa_n$$

现在将窄胶带向下卷到光滑的曲面 $\mathcal{S}$ 上。 正如我们在第 14 页上所解释的，只有中心线 L 才能真正与 $\mathcal{S}$ 接触，但是在每个接触点 p 处，胶带的切平面与曲面的切平面重合：$T_p(\mathcal{L}) = T_p(\mathcal{S})$。

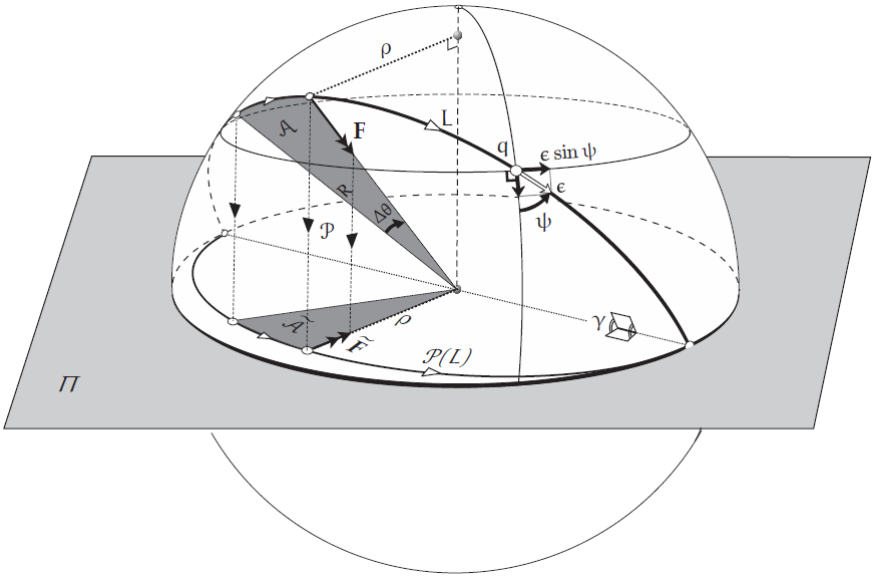
但是我们刚刚确定 κ 沿 L 一直垂直于胶带 $\mathcal{L}$，但这意味着它也垂直于曲面 $\mathcal{S}$。因此，将其视为 $\mathcal{S}$ 内的曲线，L 的测地曲率消失，它确实是 $\mathcal{S}$ 的测地线。

11.7 旋转曲面上的测地线

11.7.1 球面上的克莱罗定理

除了最简单的曲面之外，很难在所有曲面上明确地找到测地线。 然而，存在一类通用的曲面，我们可以为测地线的路径给出明确的几何公式（*克莱罗定理*）：旋转曲面。

在这些曲面中，球面是最简单的曲面之一。 古人已经知道它的测地线——大圆。我们现在重新审视这些大圆，并揭示它们拥有的隐藏特性，然后我们将能够将其推广到所有旋转曲面。



[11.5] 当粒子 q 以单位速度沿测地线 L 行进时，半径以恒定速率扫过区域，因此它在 Π 上的投影也是如此。 因此，$\rho \sin \psi$ 沿 L 是常数，这是克莱罗定理的一个特例。

考虑球面 [11.5]，它被认为是通过围绕垂直 z 轴的半圆旋转而产生的旋转曲面。 当这个半圆（或实际上是一般的母曲线）绕轴旋转时，它会创建所谓的曲面子午线。 等效地，这些是曲面与通过其对称轴的平面的相交曲线，在一般旋转曲面上也是如此。 在球面上，这些子午线是穿过两极的大圆，也许更广为人知的名称是经度圆。 球面上的这些子午线是测地线这点并非偶然；正如我们将很快讨论的，一般旋转曲面上的子午线也必然是测地线。

如果我们观察粒子沿其大圆轨道 L 以单位速度移动 Δt 时间，那么半径将在 L 平面内旋转 $\Delta\theta = (\Delta t/R)$。 因此半径扫出面积$\mathcal{A} = \frac{1}{2} R^2 \Delta\theta = \frac{1}{2} R \Delta t$，因此面积被以恒定速率扫出：在每个相等的时间间隔 Δt 内将扫出面积 $\mathcal{A}$ 相同。 从历史上看，这种情况总结为相等的时间间隔扫出相等的面积。

现在让我们将该轨道垂直向下投影到赤道平面 Π 上。 这个投影 $\mathcal{P}$ 是一个线性变换，因此所有的面积都被同一个因子压缩，这个因子是变换的决定因素。 事实上，从几何学很容易看出，如果 L 的平面与 Π 成角 γ，则 $\det \mathcal{P} = \cos\gamma$ [练习]。

也容易看出，圆轨道 L 的投影是 Π 中的椭圆轨道 $\mathcal{P}(L)$。 因此，当粒子绕这个椭圆运动时，Π 中的半径也在相等的时间内扫过相等的面积。 更详细地说，

$$\frac{d\tilde{\mathcal{A}}}{dt} = \cos\gamma \frac{d\mathcal{A}}{dt} = \frac{1}{2} R \cos\gamma$$

由于 $\mathcal{P}(L)$ 是一个椭圆，因此只有当投射的粒子远离时移动得更慢，而当它靠近时移动得更快，这种稳定的面积生成才有可能。通过注意以下几点可以很容易地确认和量化这种速度的变化 [练习]：

如果 $\mathbf{q}(t)$ 是空间中的任何轨迹，则其在平面 Π 上的正交投影 $\mathcal{P}(\mathbf{q})$的速度是速度的投影：$\frac{d}{dt}\mathcal{P}(\mathbf{q}) = \mathcal{P}(\frac{d\mathbf{q}}{dt})$。 (11.7)

切平面 T_q（未画出）由经度和纬度的正交圆的切线所张。 如图所示，q 处的测地线的方向可以用它与通过 q 的子午线所成的角度 ψ 来描述。 如果粒子 q 移动一小段时间 ϵ 它将沿大圆轨道行进一小段距离 ϵ。因此，沿纬度圆运动的水平分量是 $\epsilon \sin\psi$，如图所示。因此，位于 T_q 内的测地线的单位速度矢量可以分解为沿纬度圆的水平分量 sinψ 和沿经度圆（子午线）的分量 cosψ。

沿子午线的分量投影到 Π 中的径向运动，不产生任何区域。面积生成完全是由于水平分量 sinψ，它投影到 Π 中垂直于半径 ρ 的相等分量； 请注意，ρ 是原始粒子与对称轴的距离，如图所示。

由 于Π 中面积的产生率为 $\frac{1}{2}\rho \sin\psi$，我们推导出该量沿原始大圆轨道是恒定的，由下式给出，

$$\boxed{\frac{1}{2}\rho \sin\psi = \frac{1}{2} R\cos\gamma = \text{const.}}$$

虽然这个公式可以直接从球面的几何形状中得到证明，但上述论证的优点是我们很快就能将它推广以得到，

> **克莱罗定理。** 设 $\mathcal{S}$ 是通过围绕轴 $\mathcal{L}$ 旋转曲线 $\mathcal{C}$ 产生的旋转曲面。 如果 ρ 是轴 $\mathcal{L}$ 到测地线 g 上点一点 q 的距离，而 ψ 是经过 q 的子午线 $\mathcal{C}_q$ 与 g 的方向之间的夹角，则在 q 沿测地线 g 运动时 $\rho \sin\psi$ 保持不变。 反过 (11.8)

（回想一下，平行线是 $\mathcal{S}$ 上的水平圆，$\mathcal{S}$ 与垂直于 $\mathcal{L}$ 的平面的交集。）

历史记录：亚历克西斯 • 克劳德 • 克莱罗 (Alexis Claude Clairaut，1713-1765 年) 是法国数学家、天文学家和地球物理学家，他帮助扩展了牛顿在《原理》中的成果。 1752 年，他发表了一个准确的、可用的、近似的解决太阳、地球和月球三体问题的方法，欧拉宣称它是“. . . 数学史上最重要、最深刻的发现。” (Hankins, 1970, p. 35) 上述一般定理的命名源于克莱罗 1733 年对旋转二次曲面的研究。

11.7.2 开普勒第二定律

为了理解克莱罗定理的一般版本，让我们提问，使 $\mathcal{P}(q)$ 保持在其轨道 $\mathcal{P}(L)$ 上的力的大小和方向是什么？ 为了回答这个问题，我们注意到对（11.7）的简单推广（从速度到加

速度）：

> 如果 $\mathbf{q}(t)$ 是空间中的任何轨迹，则正交投影 $\mathcal{P}(q)$ 在平面 Π 上的加速度是加速度的投影： $\frac{d^2}{dt^2}\mathcal{P}(\mathbf{q}) = \mathcal{P}(\frac{d^2\mathbf{q}}{dt^2})$ 。 (11.9)

为简单起见，假设在图 [11.5] 中绕球面运行的粒子具有单位质量，因此力等于加速度。保持 q 在轨道 L 上的力 $\mathbf{F}$ 沿着球面法线指向直指球体中心 O，且具有恒定的大小 $(v^2/R) = (1/R)$（因为粒子具有单位速度）。 因此，将其投影保持在平面 Π 椭圆轨道上的力 $\tilde{\mathbf{F}}$ 也指向 O，且凭借所示的相似三角形可知，

$$\frac{|\tilde{\mathbf{F}}|}{|\mathbf{F}|} = \frac{\rho}{R} \quad \Longrightarrow \quad |\tilde{\mathbf{F}}| = (1/R^2)\,\rho$$

指向单个点 O 的力场称为中心力场。我们刚刚确定了中心力场，其中指向中心的力的大小与粒子与 O 的距离成正比，导致以 O 为中心的椭圆轨道在相等的时间内扫过相等的面积。 这首先由牛顿在《原理》（命题 10）中（使用完全不同的论证）证明。

实际上，可以按如下方式创建该力场和轨道。想象 Π 是一块无摩擦的冰，在 O 处有一个小洞。拿一个小冰球，在上面系上一根长度为 l 的弹性绳。将绳子穿过孔，将冰球放在孔的顶部。取松紧带的另一端，在孔下方距离 l 处垂下，并将其固定在固定点上。坐在 O 处，因为绳子在其自然长度 l 处松弛，冰球不会受到绳子的拉力。但是，如果我们将冰球从 O 移开距离 ρ，从而将绳子拉伸 ρ，那么胡克定律告诉我们，拉伸的弹性会以与ρ成正比的力将冰球拉回 O。如果我们现在以任意速度向任意方向发射冰球，它确实会沿着以 O 为中心的椭圆轨道滑行，在相同的时间内扫过相同的面积。在家里创建这个实验的近似（无冰）并不难，我们鼓励你这样做。

轨道是一个以 O 为中心的椭圆这一事实与力随距离线性变化的事实特别相关。 但是，正如牛顿是第一个认识到并证明的那样，相等的面积在相等的时间内被扫过的事实是所有中心力场的一个异乎寻常的普遍性质！ 我们将暂时描述牛顿的美丽证明，但首先让我们停下来了解这个结果在《原理》中的关键作用。

通过分析第谷·布拉赫（Tycho Brahe，1546-1601 年）多年来对行星进行的艰苦准确的观测，约翰内斯·开普勒（Johannes Kepler，1571-1630 年）能够在他面前的海量数据中辨别出数学模式。 这些经验数学事实现在被称为开普勒的行星运动三定律。 开普勒在 1609 年宣布了前两条定律，在 1618 年宣布了第三条（经过巨大的斗争）：

开普勒定律

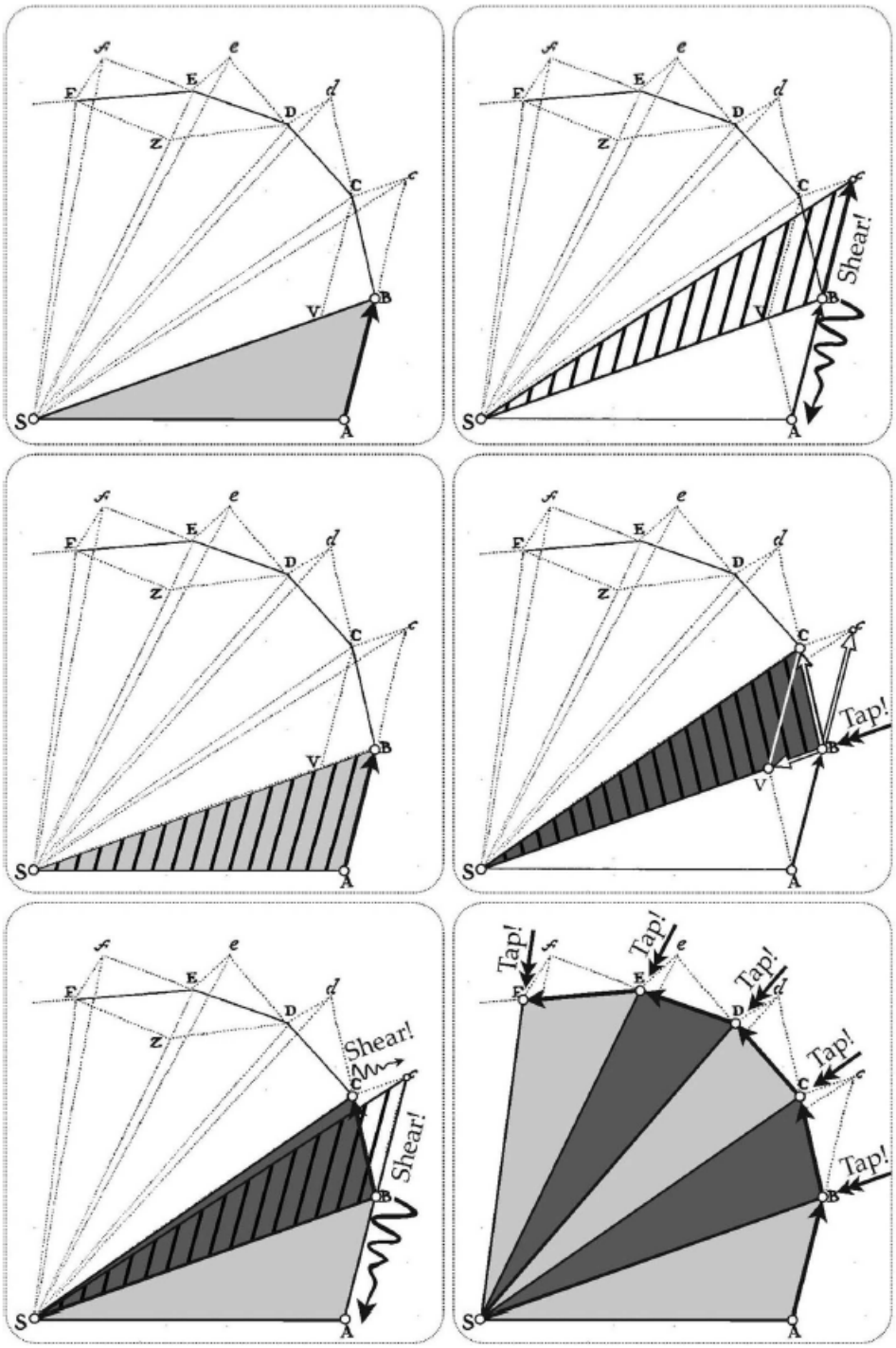
> （I） 行星的轨道是一个椭圆，太阳位于两个焦点之一。
> （II） 连接太阳和行星的线段在相等的时间内扫过相等的面积。
> （III）行星公转周期的平方与其轨道半长轴的立方成正比。

在接下来的 70 年里，这些定律仍然是个谜。 最终，在 1687 年，牛顿成功地在数学上将开普勒定律解释为他万有引力平方反比定律的逻辑结果——这是对他思想的惊人证明。但是，为了让牛顿能够仅使用几何学（见序言）来实现他的动力学分析，他必须能够将时间表示为几何量。

因此，开普勒第二定律，或者更确切地说是牛顿将其推广到任意中心力场，对整个事业来说绝对是至关重要的——它代表了《原理》的命题 1。 如果没有这个绝对基本的事实：面积就是时钟，牛顿在《原理》中的大量几何图和相关证明根本不可能实现。

11.7.3 牛顿对开普勒第二定律的几何演示

图 [11.6] 包含牛顿自己的原理图 1 的六份副本，经过修饰以讲述牛顿论证的故事（如连环画中），确定在指向固定点 S 的任何中心力场中， 轨道 ABCDEF 在相等的时间内扫过相等的面积。



[11.6] 开普勒第二定律。 牛顿自己的原理图 1 的六份副本已被修饰，以讲述牛顿论证的故事（如连环画中），确定在指向固定点 S 的任何中心力场中，轨道 ABCDEF 在相等的时间扫过相等的面积。

在没有任何力的情况下，第一张图画出了一个粒子从 A 到 B 以直线、匀速运动，正如牛顿第一运动定律所规定的那样； 在这段时间内，从固定点 S 出发的半径扫出阴影区域 SAB。 在下一个相等的时间增量中，粒子将继续在从 B 到 c 的相等距离上，仍然在平面 SAB 中，扫过交叉影线区域 SBc。 但是这两个区域面积相等，因为图示的沿 BA 的剪切使 SBc 与 SAB 重合。 因此，在没有力的情况下，相等的时间内扫出相等的面积。

现在假设，在粒子到达 B 的那一刻，它接收指向 S 的尖锐敲击。如果粒子最初在 B 处静止，那么这次敲击会导致它从 B 行进到 v 的时间与它之前从 B 行进到 c 的时间相同。 因此，粒子的实际运动是这两个运动的总和，从 B 到 C，仍在平面 SAB 中，扫过暗阴影交叉区域 SBC。 而这个面积再次等于原始面积 SAB，因为图示的平行于 SB 的剪切将 SBC 带到 SBc，然后（和以前一样）第二个剪切使 SBc 与 SAB 重合。

假设我们在粒子到达 C 时给它第二次指向 S 的轻敲[109]，然后它将从 C 行进到 D，并且，根据与之前相同的推理，面积 SCD 将再次等于面积 SAB。 以这种方式继续，以相等的时间间隔（在 D、E、F、...）向 S 敲击粒子，多边形轨道 ABCDEF 在相等的时间内扫过相等的区域。

牛顿总结道，

现在让三角形的数量增加，宽度无限减少，它们的最终周长ADF将是一条曲线； 因此，使物体从曲线切线处不断拉回的向心力将不间断地起作用，而所描述的任何面积都将与时间间隔成正比，证明完毕（Q.E.D.）。

1964 年，理查德·费曼 (Richard Feynman) 对他的加州理工学院 (Caltech) 班级说[110]：

你刚刚看到的演示是牛顿《数学原理》中的一个精确副本，而你可能从中获得或未获得的独创性和喜悦是在时间之初就已经存在。

最后，观察反过来也是正确的； 这是《原理》的命题 2：

每个在平面中某条曲线上运动的物体，画出物体到某点的半径，若此半径绕该点扫出的面积与时间间隔成正比，则该物体受到指向该点向心力的作用。

11.7.4 克莱罗定理的动力学证明

我们现在距离在一般旋转曲面上对克莱罗定理的令人满意的证明仅一步之遥，如 [11.7] 中的花瓶所示。

当点 q 在这个曲面上沿着测地线 g 移动时，轨道的加速度（根据定义）总是沿着曲面法线 $\mathbf{n}$ 定向，因此它的方向与对称轴 $\mathcal{L}$ 相交。 但是，由于（11.9），g 在 Π 上的投影是一个轨道 $\mathcal{P}(\mathrm{g})$，其加速度指向 O，因此根据开普勒第二定律的牛顿广义版本，$\mathcal{P}(\mathrm{q})$ 在 Π 中以恒定速率扫过面积。 但是，在短时间 $\delta t = \epsilon$ 内，粒子 q 沿着曲面上的测地线移动了距离 ϵ，它的投影在平面 Π 上扫出了一个面积 $\delta\mathcal{A}$，该面积最终等于以 ρ 为底，$\epsilon \sin\psi$ 为高的白色三角形的面积。 因此 $\delta\mathcal{A} \asymp \frac{1}{2}\rho\,\epsilon\sin\psi$，因此 $\frac{\mathrm{d}\mathcal{A}}{\mathrm{dt}} = \frac{1}{2}\rho\sin\psi$ 是常量。 这样

[109] 请注意，牛顿图似乎认为每个轻敲的大小相等，但如果轻敲不相等，他的论证同样有效。

[110] 参见他所谓的“失传的演讲”（Goodstein and Goodstein, 1996, p. 156），其录音可在互联网上找到。

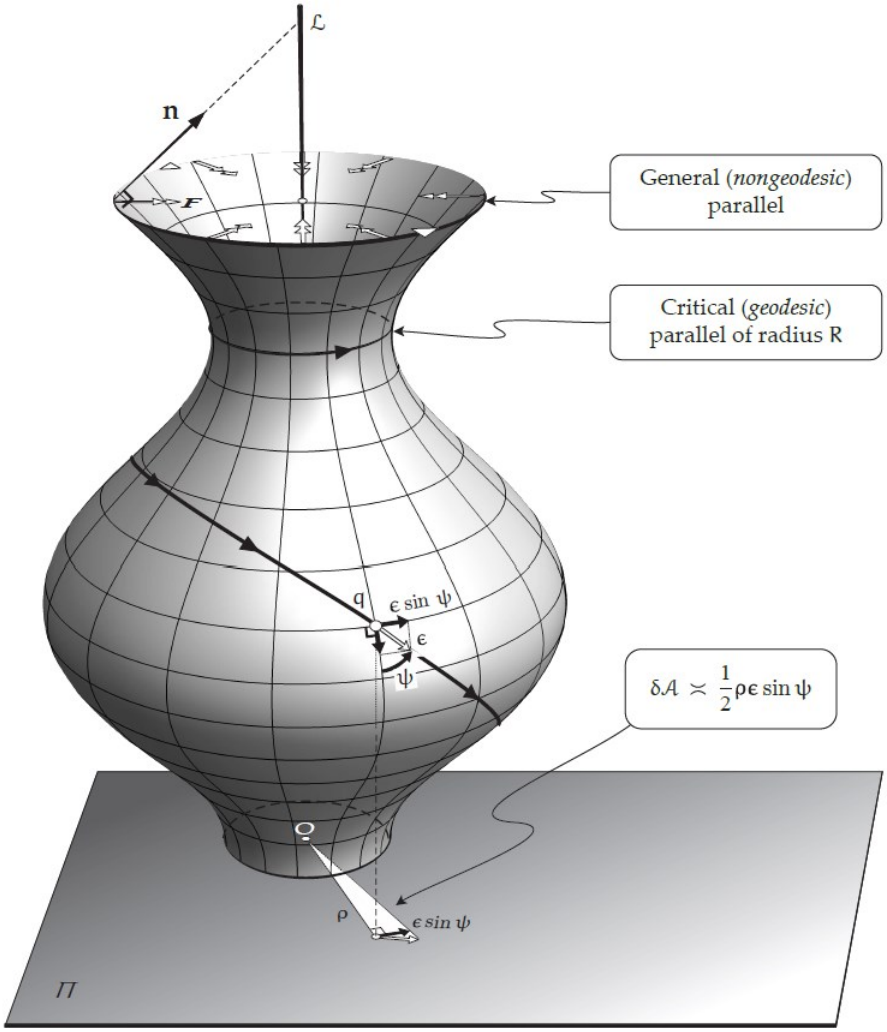
就完成了对克莱罗定理的第一部分的解释。

请注意，子午线 $(\psi = 0)$ 是特殊的测地线，因为它们在 Π 中的投影径向移入和移出，不产生任何面积。

反过来，假设沿 $\mathcal{S}$ 上的曲线 g 运动时 $\rho \sin\psi$ 是常量。那么 $\mathcal{P}(q)$ 在 Π 中以恒定速率扫过面积。 因此，根据《原理》命题 2（以上），可以得出 $\mathcal{P}(q)$ 的加速度总是指向 O。由此可以得出，g 本身的加速度总是指向 $\mathcal{L}$ 轴； 等效地，它位于经过 q 和 $\mathcal{L}$ 且包含 **n** 的垂直平面内； 等效地，它在 q 处垂直于通过 q 的平行线。

接下来，假设 g 中包含 q 的小段不是平行线。那么 **v** 不与平行线相切 $(\psi \neq \frac{\pi}{2})$。因此，*加速度垂直于 $\mathcal{S}$ 中的两个不同方向*：平行方向和 g 的方向 **v**。 因此加速度是沿 **n** 方向的，所以 g 是测地线，正如将要显示的。

但是假设包含 q 的 g 的线段是水平的圆形平行线 $\wp$ 的一部分，例如图 [11.7] 中花瓶的顶部边缘。请注意，$\rho \sin\psi$ 在 $\wp$ 上确实是常数，因为 ρ 和 ψ 分别是常数，$\psi = (\pi/2)$。但是现在，如图所示，加速度是水平的并且指向平行圆的中心，并且这个水平方向通常不垂直于曲面，因此$\wp$（通常）不是测地线。



[11.7] **克莱罗定理**。 在短时间 $\delta t = \epsilon$ 内，粒子 q 沿着曲面上的测地线移动了距离 ϵ，并且它在 Π 上的投影扫出了一个面积 δA，该面积最终等于底为 ρ 和高为 $\epsilon\sin\psi$ 的白色三角形的面积。 因此

$\frac{d\mathcal{A}}{dt} = \frac{1}{2}\rho\sin\psi$，根据牛顿对开普勒第二定律的概括，这个量是常数。

然而，在沿 $\wp$ 的法线 **n** 是水平的例外情况下，$\wp$ 是测地线（图 [11.7] 中给出了一个例子）。 这种平行 $\wp$ 被称为*临界的*，可以用各种方式来表征。 例如，如果轴 $\mathcal{L}$ 包含 $\wp$ 的垂直圆柱面沿 $\wp$ 接触 $\mathcal{S}$，则 $\wp$ 是测地线。 大多数教科书通过假设 $\mathcal{S}$ 的生成曲线 $\mathcal{C}$ 可以用图形 $\rho = \rho(z)$ 来描述这种情况，其中 z 是沿 $\mathcal{L}$ 的垂直距离。那么平行的 $\wp$ 是测地线当且仅当它是“临界的”，即 $\rho'(z) = 0$； 换句话说，轮廓曲线与 $\mathcal{L}$ 的距离具有最大值、最小值或拐点。

对于学过一些物理学的读者，我们注意到量 $\rho\sin\psi = \Omega$ 是（单位质量）粒子绕轴 $\mathcal{L}$ 转动的*角动量*。Ω 保持恒定的事实可以从物理上[111]理解：这是将 q 保持在 $\mathcal{S}$ 上的测地轨道上的力经过 $\mathcal{L}$ 的结果，因此没有关于 $\mathcal{L}$ 的*矩*。与旋转的滑冰运动员在收紧手臂时旋转得更快的原理完全相同，因此，为了保持其角动量守恒，在曲面上运行的粒子必须在靠近 $\mathcal{L}$ 时更快地绕 $\mathcal{L}$ 旋转。 但是由于曲面上的粒子沿测地线具有恒定的线速度，它只能通过将其速度指向水平线来增加其绕轴的角速度。

[111] Pressley (2010, p. 230) 也提供了这种物理解释。

注意，测地线 g 的角动量 Ω 实际上告诉我们可达的最接近于 $\mathcal{L}$ 的 g：

$$\rho = \frac{\Omega}{\sin\psi} \geq \Omega = \rho_{\min}$$

例如，假设 g 从图 [11.7] 中花瓶喉部的临界纬线（半径 R）以下出发，然后向上爬向它（临界纬线）。 如果角动量太大（$\Omega = \rho_{\min} > R$），那么 g 永远无法到达花瓶的喉部； 相反，它会弹回沿花瓶曲面向下。 图示的测地线就是这种情况，当它接近狭窄的底部时，它同样会被迫再次弹回。 为了更好地感受这一切，请尝试使用真正的花瓶，使用窄胶带构建测地线（1.7）（参见练习 11）。

11.7.5 应用：双曲面中的测地线（重温）

现在让我们将克莱罗定理应用于伪球面，以便对双曲平面的测地线有一个全新的认识。

由于伪球面越往上走就变得任意狭窄，因此从上面的讨论可以得出结论，如果一个粒子在伪球面上行进并具有丝毫角动量，它最终必然折返并下行。 因此，子午线、曳物线母线（角动量为零）是唯一可以无限向上延续的测地线。

回想一下伪球面的贝尔特拉米-庞加莱上半平面图的构造（我们在图 [5.4] 中以几何方式推导出来的）。该图中称为 X 的现在称为 ρ。 因此地图中的高度 y，由（5.4）给出，是

$$y = \frac{1}{\rho}$$

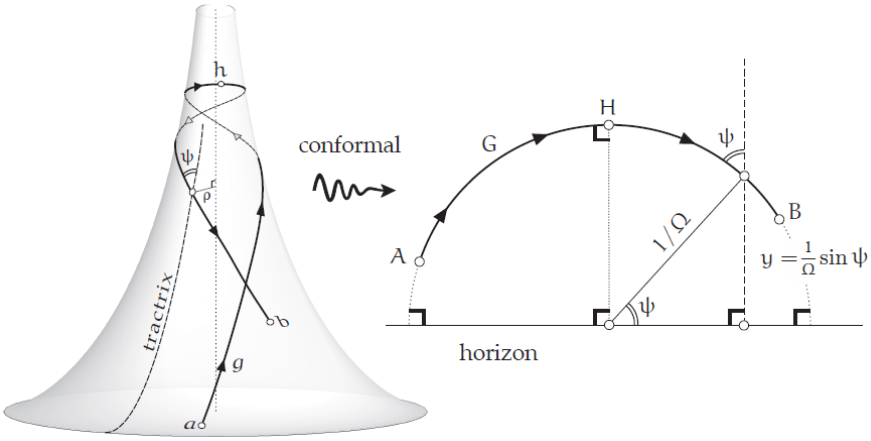
其中我们选择了伪球面的半径为 $R = 1$，以得到双曲平面的标准模型，其中

$$\mathcal{K} = -(1/R^2) = -1$$

图 [11.8] 画出了伪球面上的测地线 g（从 a 开始到 b 结束）及其在地图中的图像（从 A 开始到 B 结束）。 回想一下，我们之前证明（使用光学）该像是一个与地平线 $y = 0$ 以直角相交的半圆。 我们现在可以使用克莱罗定理对这一事实进行新的证明，并在此过程中对表示 g 的半圆的大小给出新的解释。

回想一下，这张地图在构造上是共形的。这意味着 g 和伪球面上曳物线母线（子午线）之间的角度 ψ 保持不变：地图中 g 的像与曳物线的像（垂直半线）成相同的角度 ψ。 如果沿 g 运动的单位质量粒子具有角动量 Ω，则克莱罗定理得出，

$$\rho \sin\psi = \Omega \Longrightarrow y = \frac{1}{\Omega}\sin\psi$$



[11.8] 单位质量的粒子以单位速度沿测地线 g 在半径 $R = 1$ 的伪球面上运动。 根据克莱罗定理，它的角动量 $\Omega = \rho \sin\psi$ 是常数。 但由于贝尔特拉米-庞加莱地图是共形的，角度 ψ 保持不变，因此地图中的像 G 有方程$y = \frac{1}{\Omega}\sin\psi$，即半径为 $(1/\Omega)$ 的半圆与地平线 $y = 0$ 相交于直角。

因此，使用与图 [5.8] 中完全相同的推理方法，我们得出结论，

> 如果一个单位质量的粒子在伪球面上沿测地线 g 以单位速度运动，并且它绕对称轴的角动量为 Ω，那么它在贝尔特拉米–庞加莱上半平面中的图像沿半圆运动，与 $y = 0$ 相交于直角。 此外，该半圆的半径为 $(1/\Omega)$。换句话说，半圆的（欧氏）曲率是粒子的角动量。

最后，令 h 是粒子在角动量 Ω 迫使它折返并沿伪球面向下运动之前可以到达的最高点。 很明显，h 被映射到 g 的像 G 上的最高点 H，h 和 H 都对应$\psi = (\pi/2)$。 从（5.1）可知，沿着曳物线母线（未画出）从边缘笔直上行到 h 的曲线段的弧长 σ_{max} 由 G 的半径的对数给出：

$$\sigma_{max} = \ln\frac{1}{\Omega}$$

第十二章 曲面的外在曲率

12.1 引言

我们已经通过欧拉曲率公式 (10.1) 详细了解了两个主曲率（连同它们相关的主方向）如何表征曲面的外在几何形状。 但是，是否有一个数字（没有任何关联的方向）可以表征曲面某一点的整体外在几何，就像高斯曲率 $\mathcal{K}$ 表征内蕴几何一样？

为了表征整体外在几何，我们大概必须取主曲率的某种平均值。 两种最明显的方法是取算术平均值 $\frac{\kappa_1+\kappa_2}{2}$ 或几何平均值 $\sqrt{\kappa_1\kappa_2}$。 这两个平均值在几何上都是自然的，而且非常重要。

算术平均值通常用字母 H（或$\bar{\kappa}$）表示，简称为平均曲率：

$$\boxed{H = \bar{\kappa} \equiv \frac{\kappa_1 + \kappa_2}{2}} \tag{12.1}$$

再看看 H 在图 [10.2] 中对欧拉曲率公式图的重要性：它是曲率正弦振荡的中心线。这个平均曲率 H 是理解所谓的极小曲面形状的基础，极小曲面包括跨越复杂弯曲线框的所有可能的肥皂膜的形状。 根据定义，这些极小曲面必然在每个点都满足 $H = 0$，因此 $\kappa_2 = -\kappa_1$，并且该曲面是马鞍形的。 正如序言中所指出的，我们不会在本书中探索这些迷人的曲面； 相反，我们推荐您阅读延伸阅读部分，其中推荐了几部优秀作品。

$\sqrt{\kappa_1\kappa_2}$ 的几何意义，或者更确切地说，它的平方 $\kappa_1\kappa_2$ 的几何意义比 H 的几何意义更重要，但我们会故意让你悬念再久一点，以了解它的意义可能是什么。 原谅我们，但即使是伟大的高斯也宣布它是所有数学中最伟大的妙语之一，所以几页鼓点是正常的。 （如果你不堪等待，请直接跳到答案：第 142 页的 (13.3)。）

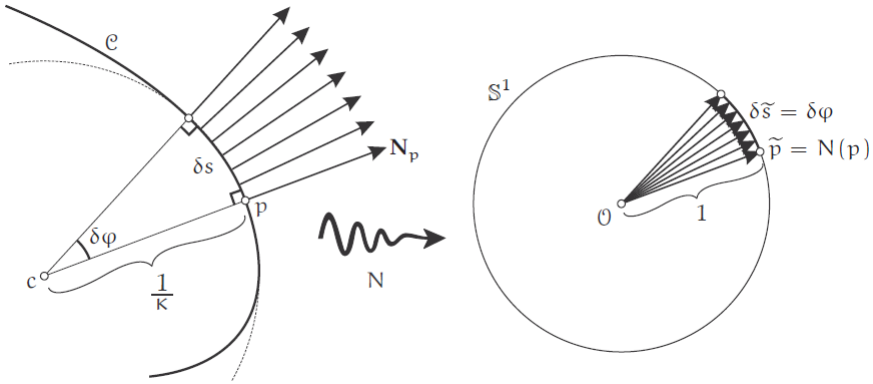
12.2 球面映射

众所周知，平面曲线 $\mathcal{C}$ 的曲率 κ 可以定义为其切线的转动速率。 等价地，κ 也可以被视为法线的转动速率。 后一种解释将允许我们将曲率的外在定义从曲线推广到曲面，因为后者也容许唯一的 $(\pm)$ 法向量。

然而，首先，如果我们将平面曲线法向量的转动速率重新解释为发生在曲线的一小段上的法向量方向的扩展，例如长度 δs，这将是有帮助的。

为了量化这一点，想象一下将这些单位法向量中的每一个并移植到从一个公共点 $\mathcal{O}$ 发

出。参见 [12.1]。这样，法线 $\mathbf{N}$ 可以看作是从平面曲线 $\mathcal{C}$ 上点 p 到位于 $\mathrm{N_p}$ 尖端的单位圆 $\mathbb{S}^1$（以$\mathcal{O}$为中心）上的点 $\tilde{\mathrm{p}} = \mathrm{N(p)}$ 的映射 N。



[12.1] 映射 N 被定义为将 $\mathcal{C}$ 上的点 p 发送到单位圆上与 $\mathrm{N_p}$ 方向相同的点。 然后 $\kappa \asymp (\delta\varphi/\delta s)$ 可以重新解释为 N 的局部长度放大系数，测量法线的扩展率。

在 $\mathcal{C}$ 的小段 δs 上，法向量的方向分布在角度 $\delta\varphi$ 上，因此这些法向量的尖端在单位圆上填充了一个长度为 $\delta\tilde{s} = \delta\varphi$ 的弧。现在可以通过考虑法向映射下弧长的局部放大来量化法向量的局部扩展：

$$\kappa = \mathrm{N}\ \text{映射的局部长度放大系数} \asymp \frac{\delta\tilde{s}}{\delta s} \tag{12.2}$$

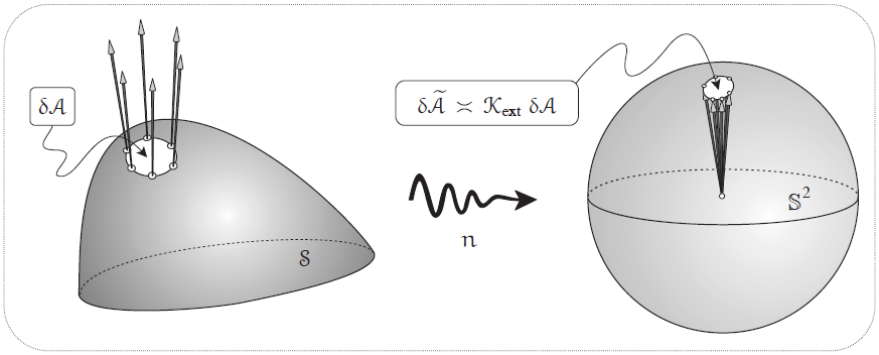
这提出了一种将构造推广到曲面 $\mathcal{S}$ 的方法。考虑曲面 $\mathcal{S}$ 上面积为 $\delta\mathcal{A}$ 并包含点 p 的一小片，那里的法向量为 $\mathbf{n}_\mathrm{p}$（见[12.2]）。 类比[12.1]，我们引入球面映射——最常称为高斯映射，或法线映射——从曲面到单位球面（$\mathrm{n}: \mathcal{S} \to \mathbb{S}^2$），将点 p 发送到与 $\mathbf{n}_\mathrm{p}$ 相同方向的单位球面上点 $\tilde{\mathrm{p}} = \mathrm{n(p)}$ 处。

术语的历史说明：在基本上所有其他教科书中，“球面映射”被称为“高斯映射”，但这在历史上是不准确的。 是的，高斯确实在 1827 年发表了它（并在几年前私下使用了它），但奥林德·罗德里格斯（Olinde Rodrigues，1795-1851）（法国银行家和业余数学家）于 1815 年首次发表了这个概念，并将其用于深入研究曲面的曲率。 认识到这一事实，马塞尔·伯杰（Marcel Berger，二十世纪后期最重要的几何权威之一）将其称为罗德里格斯-高斯映射，在我们看来它达到了适当的平衡。话虽如此，由于其清晰和简洁性，我们通常更喜欢术语“球面映射”[1]。

12.3 曲面的外在曲率

因此，我们得到了一种全新的曲面曲率外在测量法线向量的扩展，我们将暂时将其表示为 $\mathcal{K}_\mathrm{ext}$。

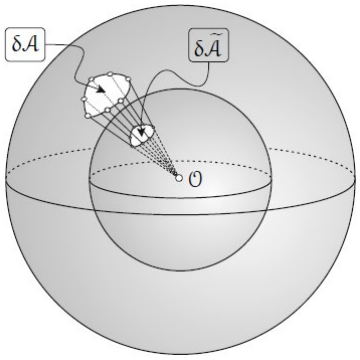
[1] 选择“球面映射”还有另一个强有力理由：它是希尔伯特（Hilbert，1952）和霍普夫（Hopf，1956）等传奇人物首选的名称。 它同样被 I. M. Singer、V. A. Toponogov 和其他现代著名的微分几何学家所采用。



[12.2] 外在曲率 $\mathcal{K}_{\text{ext}}$ 是球面映射的局部面积放大系数：$\mathcal{K}_{\text{ext}} \asymp \frac{\delta\tilde{\mathcal{A}}}{\delta\mathcal{A}}$。

在 [12.2] 中，我们将小形状缩小到 p 并定义

$$\boxed{\mathcal{K}_{\text{ext}} = \text{球面映射面积的局部放大系数} \asymp \frac{\delta\tilde{\mathcal{A}}}{\delta\mathcal{A}}} \tag{12.3}$$



[12.3] 如果 $\mathcal{S}$ 是一个以 $\mathcal{O}$ 为中心、半径为 R 的球面，并且假设 n 下的像 $\mathbb{S}^2$ 与其同心，则 n 只是 $\mathcal{O}$ 的径向投影。因此 $\mathcal{S}$ 内的距离被压缩 (1/R) 且 $\mathcal{S}$ 上的面积被 $(1/R)^2$ 压缩。

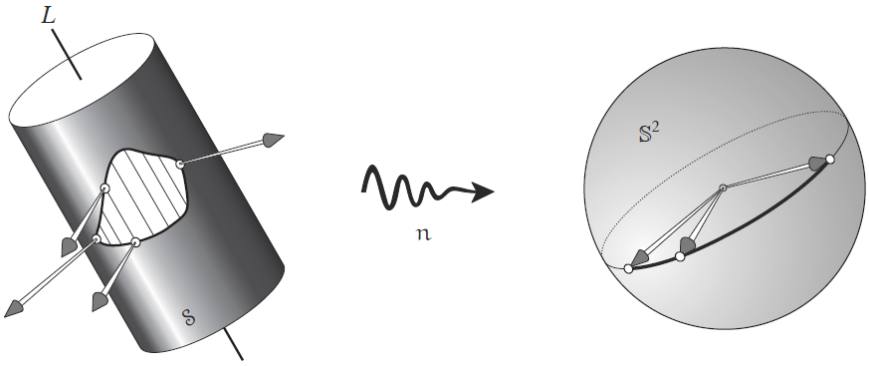
例如，假设$\mathcal{S}$ 是一个以 $\mathcal{O}$ 为中心的半径为 R 的球面，其球面映射下的像 $\mathbb{S}^2$ 具有相同的中心。 那么 n 只是从 $\mathcal{O}$ 的径向投影（如图 [12.3] 中所示）。 明显线性尺寸缩小到 (1/R)，而面积缩小到 $(1/R)^2$，所以

$$\boxed{\text{半径为 R 的球面具有外在曲率 } \mathcal{K}_{\text{ext}} = (1/R)^2\text{。}} \tag{12.4}$$

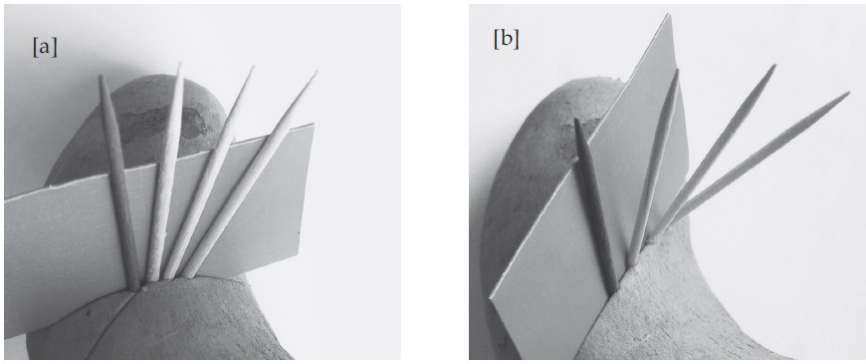
作为第二个例子，令 $\mathcal{S}$ 是一个半径为 R 的圆柱面，轴为 L（如图 [12.4] 中所示）。 圆柱面同一母线上的所有点都具有相同的法向量，因此球面映射下的像也相同。 由于 S 的所有法线都垂直于 L，因此它们在 n 下的像位于 S2 的大圆上，该大圆位于与 L 垂直的平面上。因此，$\mathcal{S}$ 上的任何区域都被 n 压成这个大圆的弧，具有零的面积。 注意，相同的母线的破碎也发生在圆锥面上，因此

$$\boxed{\text{圆柱面和圆锥面都有 } \mathcal{K}_{\text{ext}} = 0\text{。}} \tag{12.5}$$

与（12.2）相反，（12.3）中的放大系数在一般情况下是唯一定义的并不是很明显。 然而，稍后我们将能够证明它确实是：在给定点的所有无穷小区域都经过相同的放大，而不管它们的形状如何。 现在，让我们假设这一点并寻找一个特定的形状，它将揭示 $\mathcal{K}_{\text{ext}}$ 的明确公式。



[12.4] 球面映射 n 将圆柱体的每个母线（平行于其轴 L）压缩到垂直于 L 的大圆上的一个点，因此 $\mathcal{K}_{\text{ext}} = 0$。



[12.5] [a] 如果我们在一个主方向上移动，那么法向量会在那个方向上倾斜并最初停留在法线平面内； [b] 如果我们在一个一般的方向上移动，那么法向量会立即超出法线平面。

为了帮助我们找到面积放大系数在几何上不言自明的特殊形状，需要两个引理，第一个是这样的：

$$\boxed{\text{当 p 开始沿主方向移动时，}\mathbf{n}_{\text{p}}\text{ 保持在该方向的法线平面 }\Pi\text{ 内。}} \tag{12.6}$$

更准确地说，我们的意思是，如果 ζ 表示 $\mathbf{n}$ 和 Π 之间的夹角，使得 $\zeta(\text{p}) = 0$，则 $\dot{\zeta}(\text{p}) = 0$。 这是从 $\mathcal{S}$ 在 Π 中的局部镜像对称性得出的，因为如果 $\mathbf{n}$ 立即从 Π 旋转到一侧或另一侧（$\dot{\zeta} > 0$ vs. $\dot{\zeta} < 0$），这将违反对称性（10.4）。

图 [12.5a]在山药上说明了（12.6）。 在同一根山药上，在同一点上，图 [12.5b] 说明了这样一个事实，如果我们改为沿一般方向移动，那么 $\dot{\zeta}(\mathrm{p}) \neq 0$ 并且 $\mathbf{n}$ 立即旋转出 Π。 我们强烈鼓励您使用任何现成的水果或蔬菜进行自己的此类实验。

如果我们将图 [12.2] 的左侧视为 $\mathcal{S}$ 的横截面，绘制在平面 Π 中，将右侧视为 $\mathbb{S}^2$ 在平行于 Π 到 O 的平面中的横截面，则立即可得如下的第二个引理：

$$\boxed{\text{如果 p 在 } \mathcal{S} \text{ 的第 i 个主方向上沿小向量 } \mathbf{v}_\mathrm{i} \text{ 移动，则 } \mathrm{n}(\mathrm{p}) \text{ 最终在 } \mathbb{S}^2 \text{ 上沿 } -\kappa_\mathrm{i}\mathbf{v}_\mathrm{i} \text{ 移动。}} \tag{12.7}$$

请注意，此公式中的负号是由于我们的约定：在图 [12.5a] 中，$\mathrm{n}(\mathrm{p})$ 与 $\mathbf{v}_\mathrm{i}$ 同向移动，是 $\mathbf{v}_\mathrm{i}$ 的正倍数，但（根据我们的约定）这里 $\kappa_\mathrm{i} < 0$，因为 法向截面弯曲远离我们选择的 $\mathbf{n}$。

因此，我们被引导考虑在一个边与主方向对齐的小矩形的球面图下的命运。 设这些边的长度分别为沿第一和第二主方向的 ϵ_1 和 ϵ_2。 凭借（12.7），球面图最终将这个矩形发送到 $\mathbb{S}^2$ 上的另一个矩形，平行于原始矩形，但边最终被拉伸到 $\kappa_1\epsilon_1$ 和 $\kappa_2\epsilon_2$。 因此，

$$\boxed{\delta_c\tilde{\mathcal{A}} \asymp (\kappa_1\epsilon_1)(\kappa_2\epsilon_2) \asymp (\kappa_1\kappa_2)\,\delta_c\mathcal{A}}$$

因此（12.3）得到

$$\boxed{\mathcal{K}_\mathrm{ext} = \kappa_1\kappa_2} \tag{12.8}$$

例如，在半径为 R 的球体上，$\kappa_1 = \kappa_2 = (1/\mathrm{R})$，得出 $\mathcal{K}_\mathrm{ext} = (1/\mathrm{R}^2)$，与（12.4）一致。 请注意，如果一般曲面的主曲率具有相同的符号，那么我们对其几何平均值的解释如下：曲面的外在曲率与半径为 $1/\sqrt{\kappa_1\kappa_2}$ 的球面的曲率相同。

作为另一个例子，半径为 R 的圆柱体具有 $\kappa_1 = (1/\mathrm{R})$ 和 $\kappa_2 = 0$，产生 $\mathcal{K}_\mathrm{ext} = 0$，与（12.5）一致。 锥体也有一个消失的主曲率，因此在这种情况下 $\mathcal{K}_\mathrm{ext} = 0$。

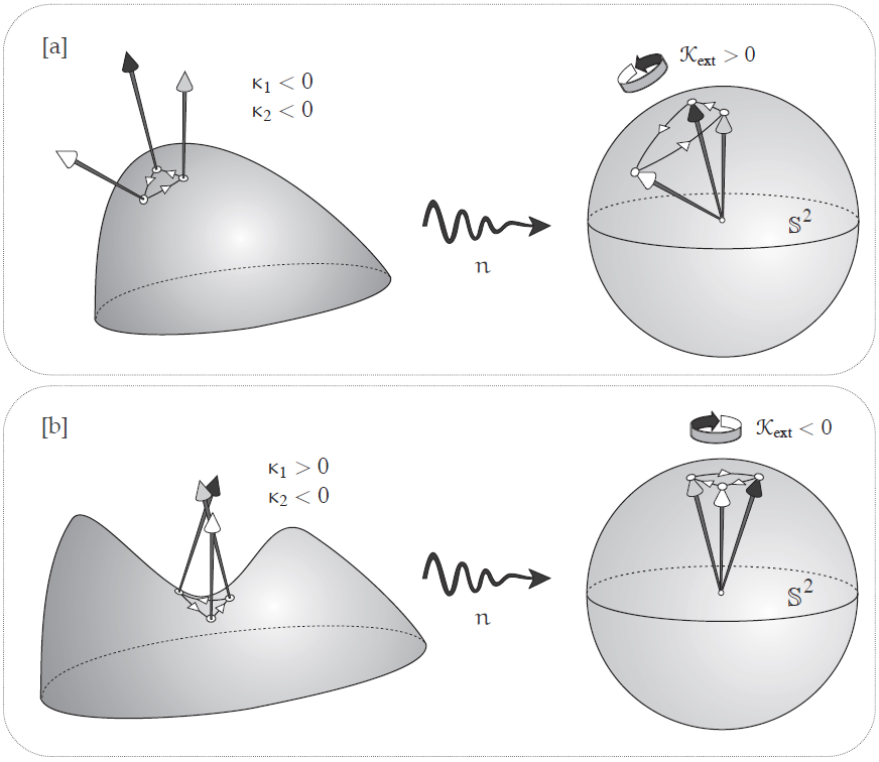
历史注解：曲率的外在定义 (12.3) 和显式公式 (12.8) 几乎普遍归因于高斯 (1827)。然而，与球面映射（又名罗德里格斯-高斯地图）本身一样，这两个见解实际上都是由罗德里格斯于 1815 年首次发表的，比高斯早 12 年。看起来高斯（以及大多数 20 世纪的数学家！）根本不知道罗德里格斯的发现。 有关这段历史的更多信息，请参见 Kolmogorov 和 Yushkevich (1996, p. 6) 和 Knoebel (2007, p. 118)。

公式（12.8）给 $\mathcal{K}_\mathrm{ext}$ 附加了一个符号：如果 p 是椭圆的（即κ_1和κ_2具有相同的符号）那么$\mathcal{K}_\mathrm{ext} > 0$； 如果它是双曲线的（即，$\kappa_1$ 和 κ_2 具有相反的符号），则 $\mathcal{K}_\mathrm{ext} < 0$； 如果它是抛物线的（即 κ_i 之一消失），则 $\mathcal{K}_\mathrm{ext} = 0$。

为了根据其原始定义（12.3）理解 $\mathcal{K}_\mathrm{ext}$ 的这个符号，我们取 $\delta_c\mathcal{A}$ 为正，并使用球面图

的简单几何特性将符号附加到 $\delta\tilde{\mathcal{A}}$，如下所示。

想象一下，当 $\mathbf{n}$ 指向您的眼睛时，您会看到 $\mathcal{S}$ 上逆时针追踪的 $\delta\mathcal{A}$ 边界，如 [12.6] 中所示。 这个想法是根据球面图是保留还是反转边界的方向，在 $\mathbb{S}^2$ 上的区域 $\delta\tilde{\mathcal{A}}$ 上附加一个符号。 也就是说，如果 $\delta\tilde{\mathcal{A}}$ 的边界以与原始相同的逆时针方向（从球体外部看）追踪，则我们认为 $\delta\tilde{\mathcal{A}}$ 为正，如果以相反的顺时针方向追踪，则将其视为负方向。



[12.6] $\mathcal{K}_{\text{ext}}$的符号取决于球面映射 n 是保持方向（[a]：$\mathcal{K}_{\text{ext}} > 0$）还是反转方向（[b]：$\mathcal{K}_{\text{ext}} < 0$）。

12.4 哪些形状是可能的？

让我们系统地分析一下任意曲面可能具有哪些形状，至少是局部的。

在一个通用点的附近，我们已经看到曲面由（10.6）中的二次方程描述：

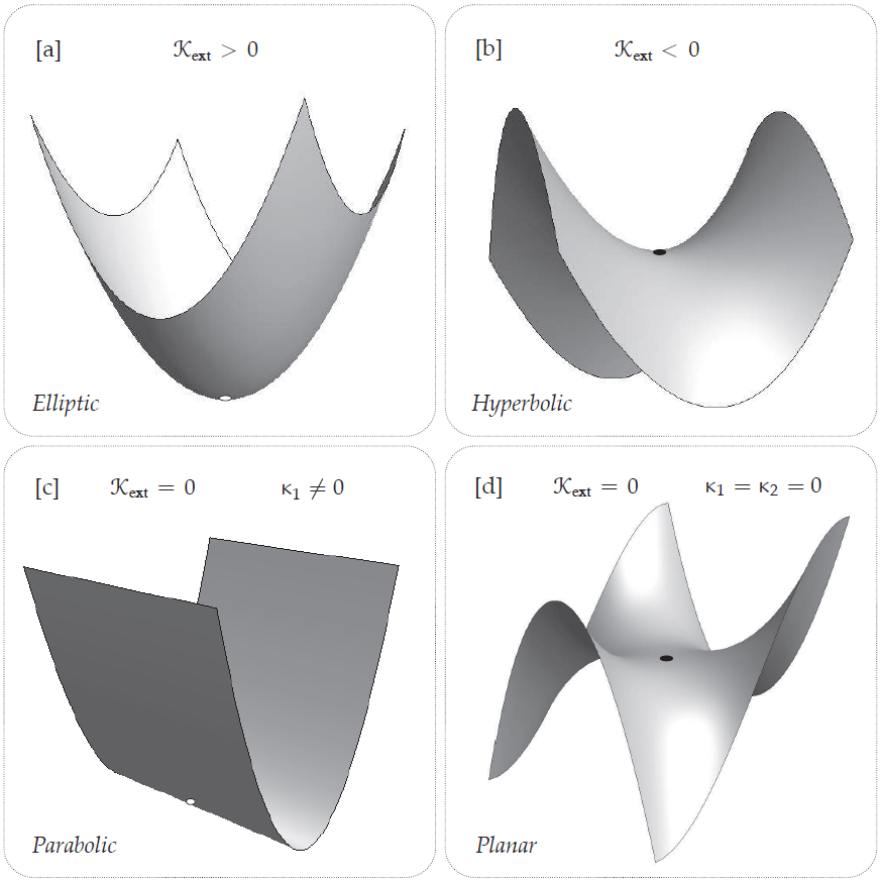
$$z \simeq \frac{1}{2}\kappa_1 x^2 + \frac{1}{2}\kappa_2 y^2$$

一般来说，主曲率都不消失，所以 $\mathcal{K}_{\text{ext}} = \kappa_1\kappa_2 \neq 0$。 回想一下，如果 $\mathcal{K}_{\text{ext}} > 0$ 点称为椭圆形，如果 $\mathcal{K}_{\text{ext}} < 0$ 则称为双曲线点。 这种通用点附近的曲面形状完全由关于 $\mathcal{K}_{\text{ext}}$ 符号的这种区别决定：

(12.8)

如果 $\mathcal{K}_{\text{ext}} > 0$，则曲面局部为碗状（如图［12.7a］中所示）。
如果 $\mathcal{K}_{\text{ext}} < 0$，则曲面局部为鞍形（如图［12.7b］中所示）。

其余情况 $\mathcal{K}=0$ 可以以两种不同的方式出现：一个主曲率消失，或者两者都消失。 在前一种情况下，该点称为抛物线，在后一种情况下称为平面，原因很快就会清楚。



[12.7]曲面的局部形状是： [a] 如果 $\mathcal{K}_{\text{ext}} > 0$ 则为碗形； [b] 如果 $\mathcal{K}_{\text{ext}} < 0$ 则为马鞍形； [c] 如果 $\mathcal{K}_{\text{ext}} = 0$ 则为槽形（并且只有 κ_1 和 κ_2 之一消失）。 如果两个主曲率都消失，则曲面可以是任意复杂的； [d] 仅说明了一种可能性，称为猴鞍。

在前一种抛物线情况下，假设 $\kappa_1 \neq 0$ 且$\kappa_2 = 0$。 然后曲面由 $z \asymp \frac{1}{2}\kappa_1 x^2$ 局部给出，它是一个沿 y 方向运行的槽（具有抛物线截面），如图 [12.7c] 中所示。

甜甜圈的曲面提供了迄今为止考虑的所有三种情况的示例。 甜甜圈外半部的每个点都有正曲率，看起来像图 [12.7a]，甜甜圈内半部的每个点都有负曲率，看起来像图 [12.7b]。如果我们想象把甜甜圈放在盘子上，与盘子接触的圆周将两个相反曲率的半部分分开，并且由抛物线点组成，其中 $\mathcal{K}_{\text{ext}} = 0$。 果然，这个圆周围的曲面带确实是一个槽，就像[12.7c]。有关关注此圆形槽的图表，请参见图 [13.3a]，第 141 页。

在最后的“平面”情况下，再次有 $\mathcal{K} = 0$，但现在两个主曲率都消失了。 由于 $\kappa_1 = \kappa_2 = 0$，我们看到 z 最终必须等于 x 和 y 中的三次（或更高次幂）齐次多项式。 将这样的点称为平面是有道理的，因为法线截面的曲率在各个方向上都消失了，就像一个平面，并且该曲面离开其切平面的速度非常缓慢，以至于它在局部看起来确实很像平面。 但当然，当我们移得更远，或者只是更精确地观察时，我们可以看到曲面远离切平面弯曲。 正如我们现在将要解释的，靠近这样一个平面点的曲面形状可以任意复杂。

为了让我们为曲面在 $\mathcal{K}_{\text{ext}} = 0$ 平面点附近可能表现出的更复杂的行为做好准备，让我们首先重新审视常规的 $\mathcal{K}_{\text{ext}} < 0$ 马鞍面。 为简单起见，让我们假设 $\kappa_1 = -\kappa_2 = 2$。 现在让我们将切平面视为复平面，以便其中的每个点都可以通过复数 $x + \mathrm{i}y = \mathrm{r}\,\mathrm{e}^{\mathrm{i}\theta}$ 来描述（使用笛卡尔坐标和极坐标）。 那么，

$$\boxed{\text{马鞍高度} = (x^2 - y^2) = \mathrm{Re}(x + \mathrm{i}y)^2 = \mathrm{Re}\left[\mathrm{r}^2 e^{\mathrm{i}2\theta}\right] = \mathrm{r}^2 \cos 2\theta}$$

高度公式的这种新形式使得当我们绕原点旋转一圈时，曲面会在高度上发生两次完整的振荡。 如果我们想象马背上的真实马鞍，那么马一侧马鞍的“谷”对应于这两个完整的高度摆动之一，它容纳骑手的一条腿，而第二个完整的摆动为另一条腿给出另一边的谷。 在骑手的正前方和后方，马鞍在我们称之为“山丘”的地方升起。

现在让我们制作一个适合有两条腿和一条尾巴的猴子的马鞍； 注意，该曲面的官方名称确实是猴鞍！ 为了创建猴子需要的三个等距的山谷，我们只需要对我们的复数取立方：

$$\boxed{\text{猴鞍高度} = \mathrm{Re}(x + \mathrm{i}y)^3 = \mathrm{Re}\left[\mathrm{r}^3 e^{\mathrm{i}3\theta}\right] = \mathrm{r}^3 \cos 3\theta}$$

这个猴鞍在图 [12.7d] 中有说明。

注意，三个山谷现在等距间隔 $(2\pi/3)$，并且在山谷之间的三个山丘同样等距：每个山丘都直接与山谷相对。 如您所见，这意味着法线截面必然在原点有一个拐点，这意味着（无需计算）曲率必须在那里消失。 由于法向曲率在各个方向都消失了，因此猴鞍确实在其中心有一个平面点。

现在很容易看出，存在着无数这样的鞍座，其复杂性不断增加，其中猴鞍可以被描述为 3 鞍座。 例如，如果我们想创建一个猫鞍——我们必须承认这不是标准术语！——我们只需要用五次方替换猴鞍的三次方，来创建 5 鞍座，高度由 $\mathrm{r}^5 \cos 5\theta$给出。

注意，从视觉上可以清楚地看出，除了平面点本身（其中 $\mathcal{K}_{\text{ext}} = 0$）之外，所有这些鞍座都具有严格的负曲率。 事实上，可以证明（见例 20）一般 n 鞍座的曲率 $\mathcal{K}_{\text{ext}}$ 围绕原点对称：它仅取决于 r，因此在每个 $\mathrm{r} = \text{constant}$ 圆上都有一个恒定的负值。

我们还没有穷尽围绕平面点的所有可能的形状。 虽然所有的高阶鞍座都有负曲率围绕着消失曲率的平面点，但也有可能使平面点被正曲率的海洋包围。 例如，考虑几乎平底的碗，方程 $\mathrm{z} = \mathrm{r}^4 = (\mathrm{x}^2 + \mathrm{y}^2)^2$ [练习]。

第十三章 高斯绝妙定理

13.1 引言

1827 年，高斯宣布了绝妙定理（Theorema Egregium）——拉丁语为“绝妙的定理”。 这一年虽然见证了贝多芬的逝世，但定理的出现意味着它也见证了现代微分几何的诞生。

这一结果催生了数学和物理学的根本性进步，其中一些我们已经涉及到，而另一些必须等待未来的章节。 它为贝尔特拉米在 1868 年接受双曲几何的关键步骤铺平了道路，将其解释为恒定负高斯曲率的鞍形曲面的内蕴几何。 而黎曼将高斯的内蕴曲率出色地推广到更高维的流形，反过来又使爱因斯坦在 1915 年能够为他极其美丽和极其准确的广义相对论赋予精确的数学形式，其中引力被理解为物质和能量在空间和时间的内蕴几何上施加的曲率。

但是为了正确地理解绝妙定理本身，我们将首先描述它的起源。 Dombrowski (1979) 的迷人而富有洞察力的工作使用高斯的私人笔记本、给朋友的信件和官方出版物，仔细拼凑出高斯对微分几何的发展见解的年表，特别是对这个定理的见解。 正如 Dombrowski 所解释的那样，高斯的史诗般的探索始于 1816 年，当时他对曲面的曲率做出了非局域性的发现，同样意义深远且出乎意料。

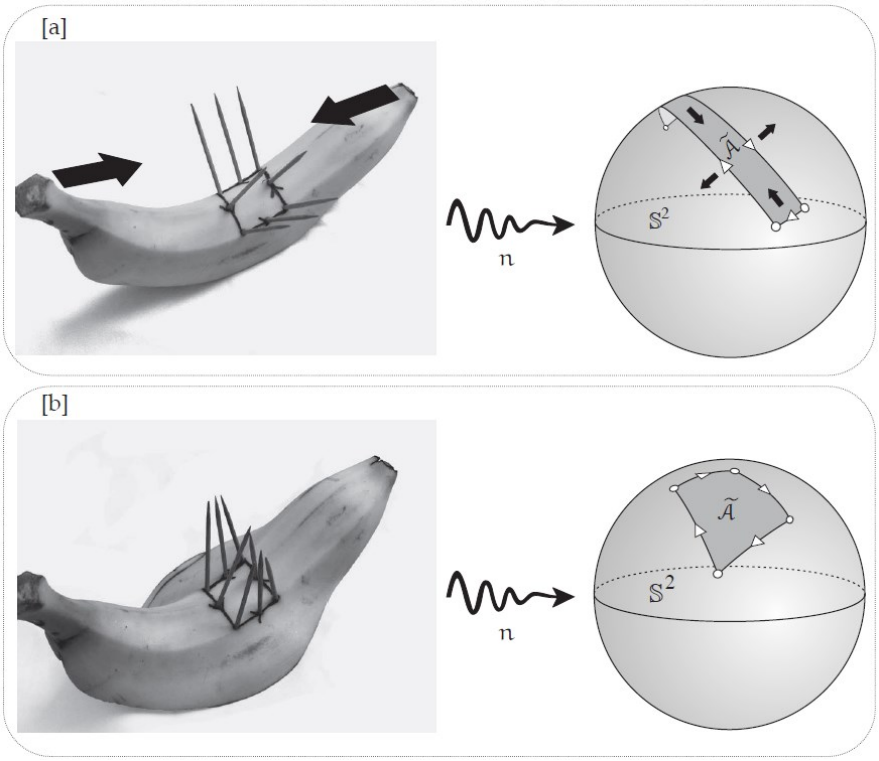
13.2 高斯美丽定理（1816）

高斯不喜欢滔滔不绝。 但如果他对别人的赞美是吝啬的，那么至少他不是伪君子。 他的一系列重大发现一直隐藏在他的私人笔记本中，直到他去世，因为他认为它们不够完善，不值得发表。 事实上，他的拉丁格言是 pauca，sed matura，“很少，但成熟了”。 其中一项未发表的发现是 1816 年的非本地结果。

令人苦恼的是，即使在他的私人笔记本[1]上，高斯也没有留下任何引导他猜测到结果的痕迹，也没有留下他如何证明的痕迹，但他的确以一反常态的热情为它命名：

> **美丽定理。** 如果固定图形的曲面在空间中呈现不同的形状，则图形的球面映射像的面积总是相同的。 (13.1)

[1] (Gauß, 1973, p. 372)



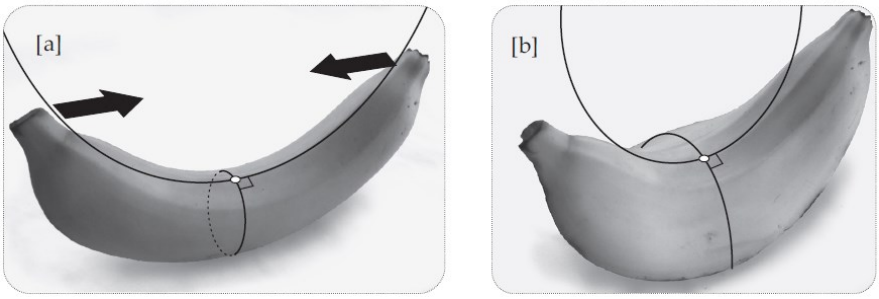
[13.1] 高斯**美丽定理**： [a] 在球面映射 n 作用下，香蕉上的方形环映射到 $\mathbb{S}^2$ 上面积为 $\tilde{\mathcal{A}}$ 的狭窄四边形，由于负曲率，以与原始相反的方向遍历。 通过将两端拉向彼此来弯曲香蕉皮会扭曲 $\mathbb{S}^2$ 上的像，产生 [b]：$\mathbb{S}^2$ 上的像现在是一个方形四边形，但面积 $\tilde{\mathcal{A}}$ 与以前完全相同！

高斯在这里所说的“球面像”是指球面映射下 $\mathbb{S}^2$ 上的像。 图[13.1]用香蕉皮说明了定理的含义； 我们强烈建议你亲自尝试这个（或类似的）实验。 首先，我们从正曲率一侧剥下一条果皮，并去除香蕉本身，保留果皮曲面的其余部分完好无损。 在果皮的负弯曲部分，我们画了一个方形的逆时针环，在这个环上我们竖立了法线（牙签）。

在球面映射下，这个环被映射到面积 $\widetilde{\mathcal{A}}$ 的 $\mathbb{S}^2$ 上的一个狭窄的四边形，以与原始相反的方向遍历，如图 [13.1a] 所示。 确保用你的眼睛跟随香蕉上的环，确认法线的尖端确实在 $\mathbb{S}^2$ 上追踪了这样一个反转的像。

如果我们现在将香蕉的两端拉向彼此，香蕉皮会经历等距变换，并且沿着（香蕉皮）曲面片边界的法线变形以在 $\mathbb{S}^2$ 上产生更方形的像。 根据美丽定理，这个新的球面像与以前的面积完全相同！

这样的实验为潜在的数学真理（13.1）提供了令人兴奋和有形的表现，但它们什么也解释不了。 在第四幕中，我们将介绍**平行移动**的概念，在它的帮助下，我们将能够为美丽定理提供一个优雅的概念解释； 不过现在，我们将简单地假设它成立，从而推导出绝妙定理作为其结果。



[13.2] 高斯**绝妙定理**： [a] 香蕉上图示点的两个曲率圆的曲率半径为 ρ_1 和 ρ_2，那里的外在曲率是 $\mathcal{K}_{\text{ext}} = 1/\rho_1\rho_2$。 通过相互拉动两端来弯曲香蕉皮产生 [b]：一个圆扩大而另一个收缩，但它们的半径乘积保持不变：$\mathcal{K}_{\text{ext}}$ 是不变的。

13.3 高斯绝妙定理（1827）

高斯在接下来的十年里没有发表关于这个主题的一个词，但在私下里，他多次回来研究微分几何和他的美丽定理，尤其是在 1822 年和 1825 年，甚至写作然后放弃了完整的手稿[2]。 最后，在 1827 年，他感到满意，将结果发表为他的 Disquisitiones Generales Circa Superficies Curvas（《曲面的一般研究》[3]）的核心，他让被压抑的兴奋占了上风。 他私下曾自称“美丽”的东西，现在向世界宣布为“绝妙”——绝妙定理（Theorema Egregium）。

但此时高斯几乎完全覆盖了他从 1816 年最初的非局部发现开始的踪迹，并且他向世界展示的结果形式纯粹是*局部*的。 要了解他是如何得到结果的这个新的局部版本的，取美丽定理中出现的图形，并简单地将其缩小到点 p。

令 $\delta\mathcal{A}$ 为$\mathcal{S}$ 上 p 周围的原始面积，$\delta\tilde{\mathcal{A}}$ 为 $\mathbb{S}^2$ 上n(p) 周围的像面积。 自然地，$\delta\mathcal{A}$ 在 $\mathcal{S}$ 的等距下是不变的，并且凭借美丽定理 （13.1）$\delta\tilde{\mathcal{A}}$ 也是不变的。 “绝妙定理”紧接而来：

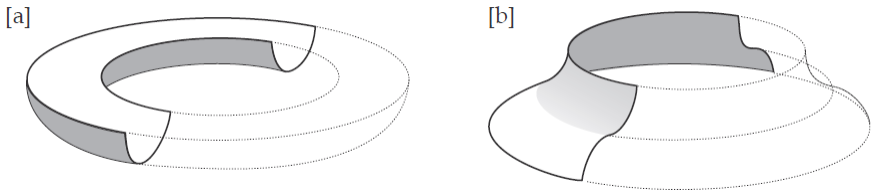
> **绝妙定理**。 外在曲率 $\mathcal{K}_{\text{ext}} \asymp (\delta\tilde{\mathcal{A}}/\delta\mathcal{A})$ 在 $\mathcal{S}$ 的等距下是不变的，因此属于 $\mathcal{S}$ 的内蕴几何。更明确地说，虽然主曲率分别取决于曲面在空间中的形状，但它们的乘积没有 ：$\kappa_1\kappa_2$ 在等距下是不变的。 (13.2)

图 [13.2] 用另一个香蕉皮说明了这一点，我们再次敦促你亲自尝试这个实验。 将香蕉的两端拉向彼此，观察在特定点 p 发生的情况：曲率半径 ρ_1 缩小（对于沿香蕉长度延伸的主方向），而横截面曲率半径 ρ_2 扩大。 但根据 (13.2) 乘积 $\rho_1\rho_2$ 在整个弯曲过程中保持完全恒定。

---

[2] 1825 年流产手稿的英文译本附在（Gauss 1827）中。

[3] (Dombrowski 1979) 和 (Gauss 1827) 都有英文翻译。



[13.3] 当 [a] 中圆形槽的外半部分在它所在的平面上反射时，我们得到等距曲面 [b]。 然而，[a] 不能在不拉伸曲面的情况下连续变形为 [b]。 尽管如此，绝妙定理向我们保证两个曲面在对应点具有相等的曲率。

为了通过实验证实这一点，取两段相当硬的金属丝，在两个主方向的特定点弯曲它们，使其与自然状态下的香蕉曲面（在去除里面的水果之前）贴合，从而获得两小块那里的主曲率圆。 现在将这两个小圆片平放在桌子上，用尺子估计它们的半径 ρ_1 和 ρ_2，从而估计它们的乘积 $\rho_1\rho_2$。 现在让一位朋友弯曲香蕉皮并保持稳定，同时你再次将金属丝在与以前相同的点，在新的主方向上贴合到新曲面上。[4] 最后，在实验误差范围内，确认 $\rho_1\rho_2$ 的新值与之前相同。

“弯曲”这个词意味着连续变形，但这实际上并不是定理所要求的：确实存在不能通过渐进的、连续的等距变形来实现的等距映射，但凭借定理它们仍然保持曲率不变。[5]

为了说明这一点，我们将描述 Aleksandrov (1969, p. 101) 的一个例子。 图 [13.3] 是 Aleksandrov 原图的手绘副本。图 [13.3a] 描绘了一个圆形槽，我们想象它停在一个平面上，沿圆 C 与之接触。为了获得 [13.3b] 中的曲面，沿 C 将曲面切成两半，然后将外半部分对平面反射。 很明显，这个新曲面与原始曲面等距，但它也直观地清楚（并且可以证明）原始曲面是刚性的，如果在此过程中不拉伸曲面，就不能弯曲成新形状。 注意，如果我们改为对直槽（半圆柱体）进行类似的变换，则可以通过连续变形获得新曲面，而无需拉伸。

关于物理与数学的“曲面”弯曲多说一句。 一个物理曲面，无论多么薄，实际上都不能在没有任何拉伸的情况下弯曲。 例如，拿一张纸，把它卷成一个圆柱面，用胶带把边缘粘在一起。 纸张的两侧开始时长度相同，但圆柱面的外圆周比内圆周略大：外侧必然被拉伸，从而在材料内部产生张力。 正是由于这个原因，当你撕下胶带时，纸张会自发地弹回原来的平面形状。 只有在厚度消失的数学极限内，曲面才能弯曲而没有任何拉伸。

高斯的结果（13.2）确实很了不起，但他更进一步。 由于结果表明 $\mathcal{K}_{\text{ext}}$ 实际上是曲率的内蕴量度，因此很自然地会问它与高斯曲率 $\mathcal{K}$ 的原始内蕴定义 （2.1）有什么关系，即每单位面积的角盈 ：$\mathcal{K} \asymp \mathcal{E}(\Delta)/\mathcal{A}(\Delta)$。 高斯的回答非常异乎寻常：

[4] 我们特意选择了一个点和一种形变，使得主方向没有改变。 然而，一般来说，当发生弯曲时，主方向将在曲面内旋转。

[5] 尚不清楚高斯本人是否意识到这种区别，但他确实用等距而不是弯曲来表达绝妙定理。 只是后来的作者解释了高斯的说法，即 $\mathcal{K}_{\text{ext}}$“在弯曲下是不变的”。

外在定义的曲率 $\mathcal{K}_{\text{ext}} = \kappa_1\kappa_2$ 和内蕴定义的高斯曲率 $\mathcal{K}$ 数值上相等：

$$\mathcal{K}_{\text{ext}} = \mathcal{K} \tag{13.3}$$

根据这个结果，我们可以并且将会忽略这两种曲率度量之间的区别，使我们能够简单地谈论曲面的曲率 $\mathcal{K}$。

与美丽定理本身一样，这个奇妙结果的最简单和最普遍的证明必须等待第四幕中平行移动概念的引入。 然而，在下一章中，我们将能够通过关于多面体的更有限的论证来为结果提供一些合理性。

同时，让我们确认（13.3）对于一些我们已经计算了外在主曲率和内蕴高斯曲率的特定曲面的有效性。

- 圆柱面和圆锥面。 内蕴平坦的平面可以卷成圆柱或圆锥，因此这两个曲面的高斯曲率消失了。 将此事实与 （12.5）结合，

$$\mathcal{K}_{\text{ext}} = 0 = \mathcal{K}$$

- 球面。 结合结果（12.4）和（1.3），

$$\mathcal{K}_{\text{ext}} = (1/\mathrm{R}^2) = \mathcal{K}$$

- 伪球面。 结合结果 （10.9） 和（5.3），

$$\mathcal{K}_{\text{ext}} = -(1/\mathrm{R}^2) = \mathcal{K}$$

- 环面。 将结果 （10.10）和练习 23 结合起来，

$$\mathcal{K}_{\text{ext}} = \frac{1}{\mathrm{r}(\mathrm{r} + \mathrm{R}\sec\alpha)} = \mathcal{K}$$

这五个曲面在历史和数学上都非常重要，以至于我们赋予它们单独处理的尊严，但实际上，（13.3）对所有五个曲面的真实性用下述事实可以一举斩获：

- 一般旋转曲面。 结合结果（10.11）、（10.12）和练习 22，其中粒子以单位速度沿母线运动，

$$\mathcal{K}_{\text{ext}} = -\ddot{y}/y = \mathcal{K} \tag{13.4}$$

第十四章 尖刺的曲率

14.1 引言

到目前为止，我们一直认为我们的曲面非常光滑，每个点都有一个明确定义的切平面和相关的法向量。 但是考虑图 [14.1]，它显示了榴莲[1]的尖刺曲面，这里显示的是它的一个内部黄色果实暴露在外。

[14.1] 我们如何测量榴莲尖刺曲面的曲率？

不管我们是从外在的还是从内蕴的角度来接近这个曲面，我们到底应该如何定义榴莲的许多尖刺之一的尖端的曲率？！

一旦我们回答了这个问题，我们将能够对绝妙定理有一些新的了解。

14.2 圆锥尖刺的曲率

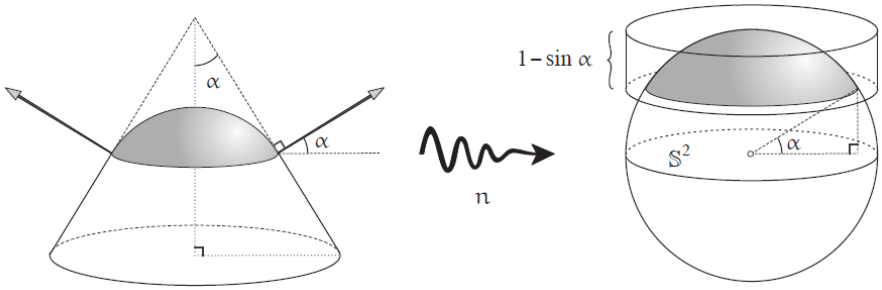
假设我们的尖刺采取锥尖的形式。 正如我们在上一节中讨论的那样，圆柱和圆锥的曲率都消失了，至少在圆锥尖端以外的每一点都消失了。 虽然圆柱体消失的内蕴曲率很容易接受，但我们有充分的理由拒绝圆锥体的“平坦性”。

想象自己是圆锥曲面的二维居民，站在顶点处。 您构造一个以顶点为中心 r 为半径的圆并测量其周长 $C(r)$。 您很快就会发现 $C(r)$ 小于欧氏预测的 $2\pi r$。 这种差异清楚地表明曲率的存在，并且这种曲率必然位于尖端，而不是其他地方，因为一般点的曲率已被证明消失了。

为了更详细地讨论这个问题，假设从对称轴到圆锥曲面的内角是 α（见图 [14.2]）。 直观地说，任何合理的曲率度量都应该随着 α 变小和尖刺变得更尖锐而增加； 相反，当 α 接近 $\pi/2$ 时，它应该减小到零，因为在那个界限，圆锥变平成欧氏平面。

[1] 这种东南亚美食被称为“水果之王”是理所当然的：它的味道（和气味！）和它看起来一样奇怪。

与数学锥体不同，物理锥体的尖端不会非常尖锐，而是稍微钝和圆。如果我们现在想象这个圆的尖端是一个极冠的形式，我们可能会获得一些数学见解。 平滑地贴合到锥体上的球体，从某种意义上说，沿着锥体和极冠之间的连接处有一个明确定义的切平面（参见[14.2]）。



[14.2] 球面映射将圆锥的钝尖发送到 $\mathbb{S}^2$ 上类似的极冠，其面积等于所示高度圆柱面的面积 $(1-\sin\alpha)$。

如图所示，球面映射只是将这个钝尖扩展成单位球面上类似的极冠。 现在（暂时）将原始数学圆锥锥尖的曲率定义为物理圆锥的钝尖内的总曲率似乎是合理的。 因为很明显，这个定义不依赖于钝尖的大小：随着我们缩小它的半径，让它越来越锋利，[2]单位球面上的球面像的大小不会改变。 此外，即使没有计算，这种曲率度量对 α 的依赖性似乎也符合我们之前描述的直觉。

因此我们做出这样的定义：

$$\boxed{\mathcal{K}(\text{尖峰}) \equiv \text{钝尖的}\,\mathcal{K}\text{的总和} = \text{（极冠）球面像的面积}}$$

也可以直接从原始（未钝化的）锥尖传递到相同的球面像。 想象一个水平面位于圆锥体的顶端，并想象一个单位法向量附着在平面上，笔直向上。 将这个单位法线作为我们球面映射的 n，平面的这个水平位置由 n 映射到图 [14.2] 右侧的北极。 平面可以在任何方向自由摆动，直到它碰到锥体的一侧并与曲面相切，此时 n 位于极冠的边界上。 因此，当位于尖端的平面呈现所有可能的位置时，n 像以前一样填充相同的极冠。我们称其为广义球面映射。

不管我们如何得出这个球面像，现在让我们找到它的面积公式。 根据阿基米德的结果（见例 10），极冠的面积等于其图示的投影到圆柱体上的面积，圆柱体沿其赤道接触单位球体。 因此，由于圆柱体的周长为 2π，而它的这一部分的高度为 $(1-\sin\alpha)$，则

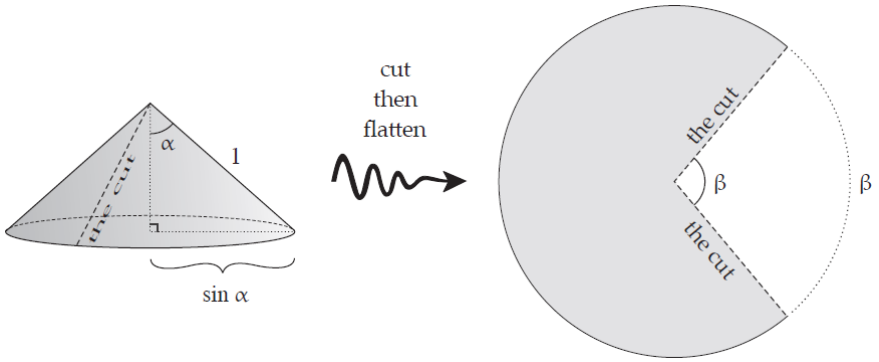
$$\boxed{\mathcal{K}(\text{尖峰}) = \text{球面像的面积} = 2\pi(1-\sin\alpha)} \tag{14.2}$$

[2] 当钝头宽 10 厘米时，你会毫不犹豫地将手掌向下按到钝头上，但当它宽 0.01 厘米时，你则不会如此！

请注意，这个曲率公式确实在 $\alpha = 0$ 时达到了最大值，并且在 $\alpha = (\pi/2)$ 时它也像预期的那样消失。

有另一种方式来看待这个结果，这种方式将很快提供多面体尖刺曲率的自然概括（见图[14.3]）。 假设我们沿着锥体的母线切开一个单位，从尖端（顶点）开始，然后围绕圆锥水平切开。 如果我们将其展平，我们将获得单位圆的一个扇形区域，以切割的两侧为边界，现在以角度 β 分开。 显然，β 也是原始尖刺曲率的度量。 事实上，我们现在将证明它是完全相同的曲率度量。

在图 [14.3] 的左边，我们看到底边的周长是 $2\pi \sin\alpha$。 但是当圆锥被切割并压平以产生右侧单位圆的扇区时，该长度不会改变。



[14.3] 当圆锥沿母线切开一个单位并铺平时，底周长 $2\pi \sin\alpha$ 变为弧 $(2\pi - \beta)$。 因此，$\mathcal{K}(\text{尖刺}) = \beta$。

因此，

$$2\pi \sin\alpha = 2\pi - \beta$$

因此（14.1）可以重新表示为

$$\boxed{\mathcal{K}(\text{尖峰}) = \beta = \text{尖峰铺平后的分割角}} \tag{14.2}$$

$\mathcal{K}_{\text{spike}}$ 的这两个公式 （14.1） 和 （14.2）都使用了圆锥居民不知道的外在思想：第一个涉及曲面法线的 3 维概念，而第二个涉及将曲面铺平在三维空间内。 尽管如此，内蕴地重新解释第二个公式实际上是一件简单的事情。

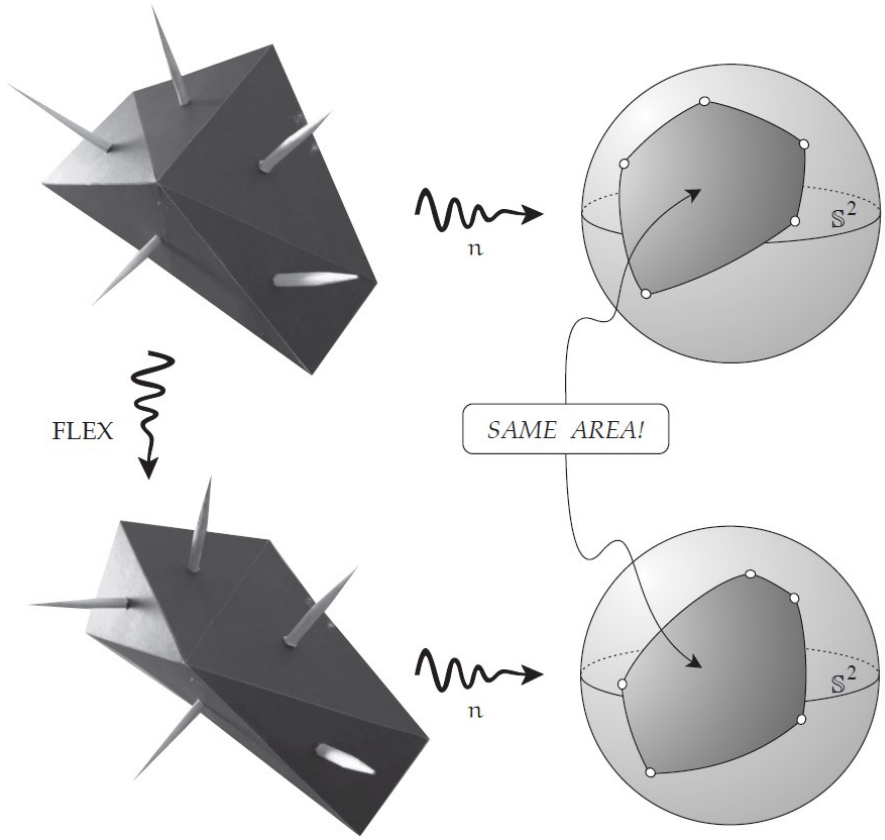
回到本节开头的讨论，以顶点为中心的半径为 r 的圆的周长 $C(r)$，通过[练习]以下 β 的内蕴公式与曲率相关：

$$\boxed{\mathcal{K}(\text{尖峰}) = \frac{2\pi r - C(r)}{r} = 2\pi - C(1)} \tag{14.3}$$

14.3 多面体尖刺的内蕴和外在曲率

在圆锥的情况下，我们刚刚看到有一种自然的方法来测量尖刺的内蕴和外在曲率，这两种测量结果是一致的。 正如我们现在所解释的，多面体尖刺的顶点 v 也是如此，我们将假设 m 个平坦的多边形面 $f_1, f_2, \cdots, f_m$ 聚在一起，f_j 的外法线为 $\mathbf{n}_j$。

$\mathcal{K}(\text{尖刺})$ 的后两个解释（14.2）和（14.3）的推广直截了当，所以我们从那里开始。 与圆锥一样，假设我们沿着其中一条边切开，从 v 开始，到任意点 p 结束，然后在所有 f_j 上围绕 v 切出一个横向回路，最后回到 p，从而将尖刺从多面体切下来。



[14.4] **多面体绝妙定理**。 广义球面映射将多面体尖刺的顶点发送到 $\mathbb{S}^2$ 上的测地多边形 P_m，随着多面体尖刺的折曲，这个球面多边形改变形状，但其面积不变！ 将此与 [13.1] 中弯曲香蕉皮的图片进行比较。

想象一下多面体由坚硬的纸板多边形组成，沿着它们的边缘用胶带粘在一起，这样每个边缘都像一个铰链（图[14.4] 中显示了这样的模型）。然后，由于它已沿 vp 切割，分离的尖刺现在可以展平到平面上，切割的 vp 的两侧最终以角度 β 分开。 如果 θ_j 是在 v 处相交的 f_j 边缘之间的角度，并且 $\Theta \equiv \sum \theta_j$ 是在 v 处相交的角度之和，那么显然 $\beta = 2\pi - \Theta$。 最后，这可以内蕴地重新表示为在原始多面体表面内绘制的以 v 为中心的单位圆的周长 $C(1)$：

$$\boxed{\mathcal{K}_{\text{int}}(v) = \text{尖峰铺平后的分割角} = \beta = 2\pi - \Theta = 2\pi - C(1)}$$

(14.4)

曲率的外在定义也可以通过广义球面映射进行推广。 正如我们对圆锥所做的那样，想象一个平面 Π（法向 n）与尖刺接触。 当 Π 摇摆并占据所有可能的位置时，一直与尖刺保持接触，n 会在 $\mathbb{S}^2$ 上填充哪个像区域？

Π 的运动界限由面决定：我们可以摇动 Π 直到它碰到面 f_j 并与面 f_j 重合，此时 n 与 $\mathbf{n}_j$ 重合，即 f_j 的法线。 因此，Π 运动的极限由 $\mathbb{S}^2$ 上的 m 个点 $\mathbf{n}_j$ 确定。 接下来，观察

> 当平面 Π 在顶点 v 上方晃动时，$\mathbb{S}^2$ 上的球面像 n 填充 $\mathbb{S}^2$ 上的以 n_j 为顶点的 m 边形 P_m 内部。

要看到这一点，我们只需要想象 Π 最初与 f_j 重合，然后在 f_j 和 f_{j+1} 相交的边 e 上滚动。 这样做时，n 必须在垂直于 e 的平面内摆动，因此 $\mathbb{S}^2$ 上的像 n 沿着将 n_j 连接到 n_{j+1} 的大圆的弧旅行。因为 Π 在所有可能的位置上晃动，连接连续 n_j 的测地弧形成了 n 覆盖的区域的边界。

因此，我们得出了 v 的外在曲率的自然概念：

$$\boxed{\mathcal{K}_{\text{ext}}(v) = \text{连接多面体法线的 } \mathbb{S}^2 \text{ 上 m 边形的面积}} \tag{14.5}$$

14.4 多面体绝妙定理

正如锥体所发生的那样，多面体尖刺的内在和外在曲率测量结果是一致的，我们可以合理地称之为多面体绝妙定理的结果：

$$\boxed{\mathcal{K}_{\text{ext}}(\mathrm{v}) = \mathcal{K}_{\text{int}}(\mathrm{v})} \tag{14.6}$$

图 [14.4] 说明了该定理的作用：随着尖刺弯曲，法线都向不同方向旋转，球面图像 P_m 发生变化，但根据定理，它的面积实际上是尖刺固有的，因此不能改变！ 将此与图 [13.1] 中弯曲香蕉皮的图片进行比较。

要理解这一点，请考虑图 [14.5]。 简而言之，解释是球面多边形的面积仅取决于它的角度，而该图表明，这些反过来又由多面体面的内角决定。

在左侧，多面体的两条边以 θ 角相交，与这些边垂直的平面以 $\tilde{\theta}$ 角相交。 从标记的直角可以得出 $\theta + \tilde{\theta} = \pi$。 因此，在广义球面图下，球面多边形的相应边以角度 $\tilde{\theta} = \pi - \theta$

相交。

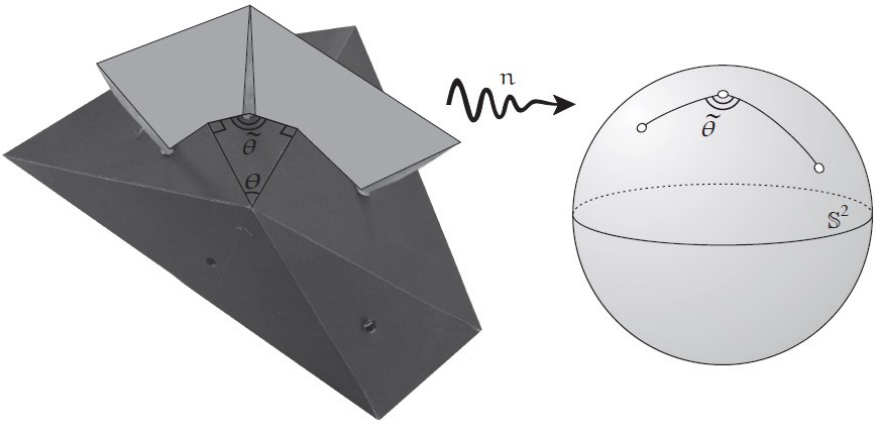
我们现在可以详细验证结果。 回想一下，在练习 5 中，我们将 Harriot 定理从三角形推广到多边形，表明球形 m 边形 $\mathrm{P_m}$ 的面积 $\mathcal{A}(\mathrm{P_m})$ 再次等于其角盈：

$$\boxed{\mathcal{A}(\mathrm{P_m}) = \mathcal{E}(\mathrm{P_m}) \equiv \left[\text{内角和}\right] - (\mathrm{m}-2)\pi}$$

所以，

$$\mathcal{K}_{\mathrm{ext}}(v) = \mathcal{A}(\mathrm{P_m}) = 2\pi - \sum_{i=1}^{m}\left[\pi - \tilde{\theta}_{\mathrm{i}}\right] = 2\pi - \Theta = \mathcal{K}_{\mathrm{int}}(v)$$

于是证明了多面体绝妙定理。



[14.5] 在左侧，多面体的两条边以 θ 角相交，垂直于这些边的平面以 $\tilde{\theta}$ 角相交。 从标记的直角可以得出 $\theta + \tilde{\theta} = \pi$。 因此，在广义球面映射下，右侧球面多边形的相应边以角度 $\tilde{\theta} = \pi - \theta$ 相交。

历史注释：上述优雅的洞察通常归功于希尔伯特，他出现在[3]永恒的杰作《直观几何》（Geometry and the Imagination，希尔伯特，1952 年，第 195 页）中，德国原版于 1932 年出现。但事实上[4]，伟大的英国物理学家詹姆斯• 克拉克•麦克斯韦 1854 年（James Clerk Maxwell）也进行了相同的观察，比希尔伯特早 78 年。 麦克斯韦 1854 年写给威廉汤姆森的原始信可以在麦克斯韦（2002 年，第 1 卷，第 243 页）中找到，他在 1856 年的最终论文出现在麦克斯韦（2003 年，第 1 卷，第 4 节）中。

最后，我们注意到这些概念可以扩展到非凸多面体和具有负曲率的顶点。 有关优美且完全一般的调研，请参见 Banchoff (1970)。

---

[3] 然而，希尔伯特并没有使用我们的晃动平面。 相反，他对从 $\mathrm{n_j}$ 到 $\mathrm{P_m}$ 的过渡的理由仅在于：“为了将其与曲面的球面表示联系起来，我们…用大圆弧将点 $[\mathrm{n_j}]$ 连接起来…[以创建 $\mathrm{P_m}$]”

[4] 我在翻阅我的麦克斯韦文集（麦克斯韦 2003 年）的副本以寻找一些完全不相关的东西时，偶然发现了这个鲜为人知的事实。

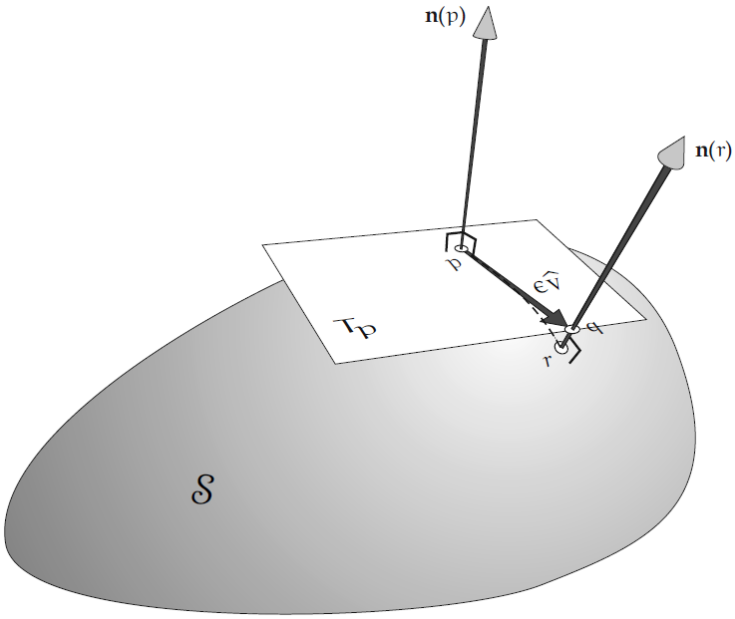
第十五章 形状算子

15.1 方向导数

从外在看，高斯曲率 $\mathcal{K}$ 测量点 p 附近曲面法线 $\mathbf{n}$ 的变化，但仅以模糊的平均方式测量。 我们现在不是在包含 p 的曲面 $\mathcal{S}$ 的一小块区域上查看 $\mathbf{n}$ 的方向的整体展开，而是转向一种更精确的量化 $\mathbf{n}$ 变化的方法，通过查看当在 $\mathcal{S}$ 内在特定方向上远离 p 时它变化的速度。

令 $\hat{\mathbf{v}}$ 是从 p 发出并位于$\mathcal{S}$ 在 p 处的切平面 $\mathrm{T_p}$ 内的单位向量（见图 [15.1]）。 我们想定义 $\mathbf{n}$ 在 $\hat{\mathbf{v}}$ 方向上的变化率。 如果我们在这个方向上远离 p 移动一小段距离 ϵ ，那么我们会到达一个点 q，因此其位置向量 $\mathbf{q} = \mathbf{p} + \epsilon\hat{\mathbf{v}}$。 然后很容易尝试将 $\mathbf{n}$ 的变化率定义为

$$\lim_{\epsilon\to 0}\frac{\mathbf{n}(\mathrm{q}) - \mathbf{n}(\mathrm{p})}{\epsilon}$$



[15.1] 为了找到 $\mathbf{n}$ 在 $\hat{\mathbf{v}}$ 方向上的导数 $\nabla_{\hat{\mathbf{v}}}\mathbf{n}$，我们考虑在曲面内沿该方向移动距离 ϵ 导致的变化 $\delta\mathbf{n} \equiv \mathbf{n}(\mathrm{r}) - \mathbf{n}(\mathrm{p})$。 然后 $\nabla_{\hat{\mathbf{v}}}\mathbf{n} \asymp \frac{\delta \mathrm{n}}{\epsilon}$ 。

但这不行。 因为 q 实际上是在 $\mathrm{T_p}$ 内，而不是 $\mathcal{S}$ 上，因此 $\mathrm{n(q)}$ 甚至没有定义。

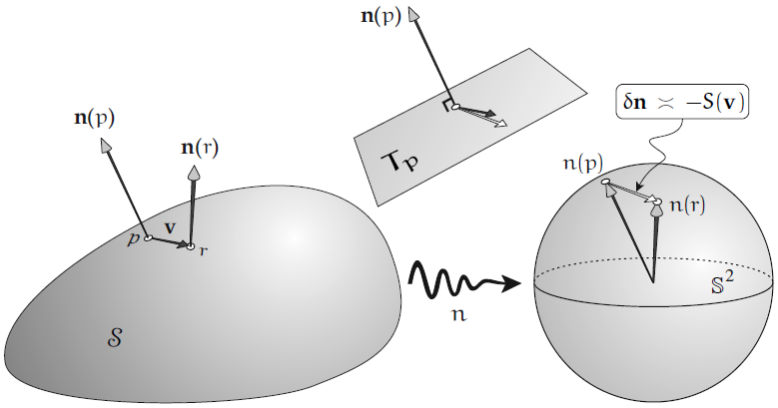
尽管如此，似乎很清楚，我们正在摸索的概念确实是明确定义的：我们只需要在曲面内移动距离 ϵ ，而不是切平面，但我们必须在所需的方向 $\hat{\mathbf{v}}$ 上移动。例如实现这个目的的一种方法是从 q 到 $\mathcal{S}$ 做垂线，在 r 处相交。 似乎很清楚（并且可以证明）从 p 到 r 的测地线段的长度 pr 最终等于 ϵ：

$$\mathrm{pr} \asymp \epsilon$$

回到我们对变化率的失败定义，我们现在可以定义

$$\boxed{\delta\mathbf{n} \equiv \mathbf{n}(\mathrm{r}) - \mathbf{n}(\mathrm{p})}$$

是在曲面内沿 $\hat{\mathbf{v}}$ 方向移动距离 ϵ （最终）导致的 $\mathbf{n}$ 变化。 由于 $\mathbf{n}$ 具有单位长度，因此当我们从 p 移动到 r 时，它只会稍微旋转。 因此，如果我们将 $\mathbf{n}(\mathrm{p})$ 和 $\mathbf{n}(\mathrm{r})$ 画到共同的原点（如图 [15.2] 中所示），那么其尖端的运动最终与 $\mathbf{n}$ 正交，因此 $\delta\mathrm{n}$ 最终位于切平面 $\mathrm{T_p}$ 内。



[15.2] 应用于 $\mathrm{T_p}$ 中从 p 发出的短切向量 $\mathbf{v}$ 的**形状算子**告诉我们从 $\mathbf{v}$ 的尾部到尖端的法线变化有多少。这种变化 $\delta\mathbf{n} \asymp -\mathrm{S}(\mathbf{v})$ 最终位于 $\mathbb{S}^2$ 在 $\mathrm{n(p)}$ 处的平行切平面。 叠加两个切平面，S 可以被视为 $\mathrm{T_p}$ 到自身的线性变换，这里画成与 $\mathcal{S}$ 分离，漂浮在两个曲面之间。

最后，$\mathbf{n}$ 沿 $\hat{\mathbf{v}}$ 的**方向导数** $\nabla_{\hat{\mathbf{v}}}$ 可以定义为

$$\boxed{\nabla_{\hat{\mathbf{v}}}\mathbf{n} \equiv \lim_{\epsilon\to 0}\frac{\delta\mathbf{n}}{\epsilon}} \tag{15.1}$$

正如我们所解释的，这个导数存在于 $\mathrm{T_p}$ 中。

这是每单位长度的方向导数。 相反，如果我们想要沿长度为 v 的一般切向量 $\mathbf{v} = v\,\hat{\mathbf{v}}$ 的导数，那么我们必须简单地按这个长度放大：

$$\nabla_{\mathbf{v}}\mathbf{n} = \nabla_{v\hat{\mathbf{v}}}\mathbf{n} = v\nabla_{\hat{\mathbf{v}}}\mathbf{n} \tag{15.2}$$

如果我们将 $\mathbf{v}$ 视为在曲面上运动的粒子通过 p 时的速度，就更直观。 那么单位向量 $\hat{\mathbf{v}}$ 对应于以单位速度运动的粒子。 因此，当 $\nabla_{\hat{\mathbf{v}}}$ 应用于沿粒子轨迹（不一定是 $\mathbf{n}$）定义的任何量时，它会给出该量随时间的变化率。 因此，如果粒子沿着相同的轨迹以三倍的速度行进，将 $\hat{\mathbf{v}}$ 更改为 $3\hat{\mathbf{v}}$，那么数量将沿着轨迹以三倍的速度变化。 这是看待（15.2）的一

种方式。 注意，方向导数的概念可以应用于定义在曲面上或仅轨迹沿线上的任何量（向量或标量）。

有另一种方式来看待方向导数的定义，首先，我们需要将“终极相等”的概念扩展到向量。 因此，假设两个向量 $\mathbf{a}$ 和 $\mathbf{b}$ 依赖于一个趋于零的小量 ϵ。 $\mathbf{a} \asymp \mathbf{b}$ 的一个明显定义是 $\mathbf{a}$ 的每个分量最终都等于 $\mathbf{b}$ 的相应分量。 但是如果一个或多个分量同时为零，这个定义是有问题的 [为什么？]。 因此，我们采用以下更令人满意的几何定义，即两个向量的大小和方向最终相等：

> 如果两个向量 $\mathbf{a}$ 和 $\mathbf{b}$ 依赖于一个趋于零的小量 ϵ，我们将它们定义为最终相等，写成 $a \asymp b$，当且仅当 $|a| \asymp |b|$ 并且它们的方向之间的角度以零为极限。

有了这个，我们现在可以将（15.1）改写为

$$\nabla_{\hat{\mathbf{v}}}\mathbf{n} \asymp \frac{\delta\mathbf{n}}{\epsilon} \quad \Leftrightarrow \quad \delta\mathbf{n} \asymp \epsilon\nabla_{\hat{\mathbf{v}}}\mathbf{n} = \nabla_{\epsilon\hat{\mathbf{v}}}\mathbf{n}$$

那么，这里是由此产生的观点，我们将在本书的其余部分反复使用：

> 当 ϵ 趋于零，由此 $\mathbf{v} = \epsilon\,\hat{\mathbf{v}}$ 趋于零时，$\nabla_{\mathbf{v}}\mathbf{n}$ 最终等于 $\mathbf{n}$ 从 $\mathbf{v}$ 的尾部到尖端的变化 $\delta\mathbf{n}$：$\nabla_{\mathbf{v}}\mathbf{n} \asymp \delta\mathbf{n}$。 (15.3)

当然，如前所述，这句话带有一丝诗意，因为 $\mathbf{v}$ 的尖端实际上并不位于 $\mathcal{S}$ 内。但是随着 ϵ 趋于零，所需的诗意也是如此！ 因为，在此极限过程中，图 [15.1] 中 q 和 r 之间的差别很快变得无关紧要。

15.2 形状算子 S

与点 p 处的曲面相关联的*形状算子*[122]（Shape Operator） S 告诉我们当我们沿着任意切向量 $\mathbf{v}$ 远离 p 时，法向量如何变化（它所指的方向，以及多快），该向量不再一定很短，而是具有任意长度 v。更准确地说，它被简单地定义为 $\mathbf{n}$ 沿 $\mathbf{v}$ 的方向导数的负数：

$$\boxed{S(\mathbf{v}) \equiv -\nabla_{\mathbf{v}}\mathbf{n}} \tag{15.4}$$

在定义中加入符号是标准做法； 正如我们将很快看到的，它是由我们早先对法向截面

[122] 还有另外两个数学对象可以编码与形状运算符完全相同的信息：第二基本形式和 Weingarten 映射。我们不会在本书中使用它们。

曲率符号的定义所引发的。

曲面 $\mathcal{S}$ 和球面 $\mathbb{S}^2$ 分别在 p 和 n(p) 处具有相同的法向量 $\mathbf{n}(p)$。因此，它们在这两个点的切平面 T_p 和 $T_{n(p)}$ 是平行的，我们可以想象移动它们以使其重合。因此，从 n(p) 发出的 $\mathbb{S}^2$ 的任何切向量都可以被画成居住在 T_p 内，从 p 发出。在图 [15.2]中，我们画出了漂浮在两个曲面之间的这两个重合平面。

如果我们设想非常接近 p 的点 r 被球面映射带到 $\mathbb{S}^2$ 上接近 n(p) 的点 n(r)，那么形状算子的含义现在变得特别清楚（见图 [15.2]）。如果 $\mathbf{v}$ 是连接 p 到 r 的非常短的向量，而 δn 是连接 n(p) 到 n(r) 的相应向量，那么，根据 (15.3)，

$$\boxed{S(\mathbf{v}) \asymp -\delta\mathbf{n}}$$

如果我们把 p 看作这两个重合平面的原点，那么 S 实际上是 T_p 到自身的线性变换。也就是说，如果 $\mathbf{v}$ 和 $\mathbf{w}$ 是 p 处的任意切向量，而 a 和 b 是任意常数，则

$$\begin{aligned} S(a\mathbf{v}+b\mathbf{w}) &= -\nabla_{a\mathbf{v}+b\mathbf{w}}\mathbf{n} \\ &= -\nabla_{a\mathbf{v}}\mathbf{n} - -\nabla_{b\mathbf{w}}\mathbf{n} \\ &= -a\nabla_{\mathbf{v}}\mathbf{n} - -b\nabla_{\mathbf{w}}\mathbf{n} \\ &= aS(\mathbf{v}) + bS(\mathbf{w}) \end{aligned}$$

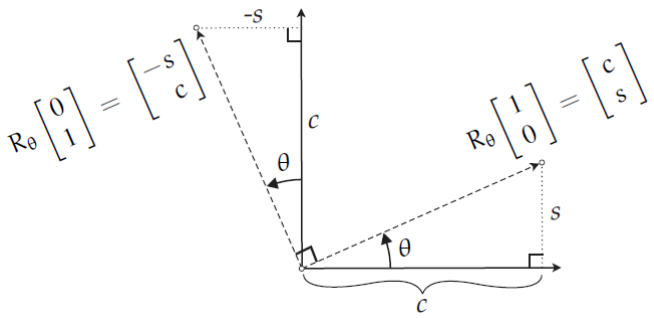
回忆一下线性代数，这意味着 S 可以表示为一个长方形的数列 [S]，称为 S 的矩阵。一般来说，矩阵的第 j 列是第 j 个基向量的像，或者更确切地说是该向量的数值分量。在我们的二维情况下，这意味着 [S] 是一个 2×2 的正方形。

在本章的后面，我们将研究矩阵 [S]，但现在我们将通过一个我们暂时需要的例子来说明“矩阵化”的一般概念，即矩阵 $[R_\theta]$ 表示表示平面绕原点旋转角度 θ 的旋转 R_θ。

图 [15.3] 说明了 R_θ 对基向量的影响，从中我们立即推导出

$$[R_\theta] = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \tag{15.5}$$

其中 $c \equiv \cos\theta$ 且 $s \equiv \sin\theta$。



[15.3] 矩阵 $[R_\theta]$ 的第一/第二列是第一/第二基向量旋转 R_θ 后的像。

15.3 S 的几何效应

设 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 分别是沿（正交的）第一和第二主方向的单位向量。 当参考这个特殊的正交基 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 时，S 对 T_p 中向量的几何影响是什么？ 我们之前的结果 (12.7) 提供了简洁的答案：

$$\boxed{\begin{array}{c}\text{主方向是形状算子 S 的}\textbf{特征向量}\text{，主曲率是对应的}\textbf{特征值}\text{：}\\ S(\mathbf{e}_i) = \kappa_i \mathbf{e}_i\end{array}} \tag{15.6}$$

注意：正是为了抵消（12.7）中的负号，在 S 的定义（15.4）中引入了一个负号。

因此，知道 S 对主方向的影响是通过它们各自的主曲率拉伸它们，线性现在表明 S 对一般切向量的影响是在这两个垂直方向上通过这两个因子拉伸它。

如果我们选择我们的基向量为 $\{\mathbf{e}_1, \mathbf{e}_2\}$，那么 S 的矩阵采用特别简单的对角线形式，

$$\boxed{[S] = \begin{bmatrix}\kappa_1 & 0\\ 0 & \kappa_2\end{bmatrix}} \tag{15.7}$$

回想一下，线性变换以相同的因子扩展每个形状的面积。 这个通用的面积扩展因子是其矩阵的**行列式**。 这在当前情况下尤为明显，其中与主轴对齐的单位正方形被拉伸为边长为 κ_1 和 κ_2 的长方形，因此单位面积的原正方形扩展为面积 $|[S]| = \kappa_1\kappa_2$， 所以这是所有面积的扩展因子。 我们再次恢复我们熟悉的高斯曲率外在版本的表达式：

$$\boxed{\mathcal{K}_{\text{ext}} = \text{形状算子的面积扩展因子} = |[S]| = \kappa_1\kappa_2} \tag{15.8}$$

回想一下，当我们第一次推导出这个结果时（见 (12.8)），我们是通过考虑一个特定形状的面积扩展来得出的，我们只是声称结果与这个选择无关； 对形状算子线性的认识现在证明了这一说法。 此外，鉴于其几何意义，该面积扩展因子（即行列式）在所有坐标系中必然具有相同的值。

注意，这个矩阵 $[S]$ 是对称的（也称为自伴随矩阵），这意味着它在其主对角线上具有镜像对称性，从左上角到右下角。 由于这条对角线上的反射是通过交换行和列来实现的，这会创建（根据定义）转置 $[S]^T$ ，这种对称性可以写成，

$$[S] = [S]^T \tag{15.9}$$

矩阵 [S] 的对称性 （15.9）并非偶然，而是反映了（原谅双关语！）线性变换 S 本身的潜在对称性，即使基不与主方向对齐，它也会继续存在，在这种情况下 [S] 将不再是对角的（矩阵）。

15.4 绕行：奇异值分解和转置的几何

在这个可选的小节中，我们首先通过提出以下问题来试图阐明形状算子的对称性（15.9）：如果一般线性变换 M 用矩阵 [M] 表示，那么转置矩阵 $[M]^T$表示什么变换 M^T？ 为了回答这个问题，我们将首先对所有线性代数中最重要的结果之一，即所谓的**奇异值分解**[123]（Singular Value Decomposition，简称 SVD）进行几何解释和推导。

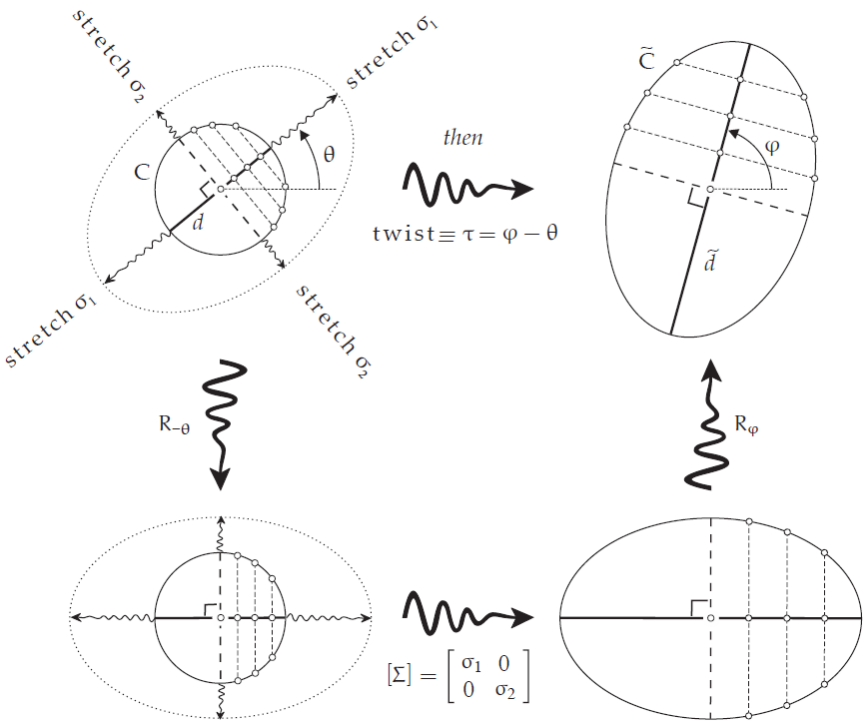
我们现在将展示的几何解释和证明似乎适用于线性代数的入门课程。 然而，令人惊讶的是，我们未能在任何标准的介绍性教科书中找到这些想法； 事实上，我们怀疑许多在职数学家也可能不熟悉它们。 不过，如果你急于继续学习微分几何，请随时跳到下一部分。

我们首先回顾一个熟悉的事实，即如果平面的线性变换具有两个实特征值，则其几何效果很容易可视化，如下所示：通过这些因子沿两个（通常是非正交的）特征向量方向拉伸。但是如果变换没有实的特征值呢？ 这种变换的一个简单示例（即不保留任何方向的转换）是旋转，但至少这也很容易被可视化。 但是，我们如何使不保留任何方向的一般线性变换具有几何意义？

SVD 为这个问题提供了一个非常简单而生动的答案，它适用于所有线性变换； 因此，即使是我们认为已经理解的变换（在两个非正交特征向量方向上伸展的变换）也因此获得了新的意义。

奇异值分解 (SVD)。平面的每一次线性变换都相当于在两个正交方向上拉伸（通常通过不同的因子 σ_1 和 σ_2，称为**奇异值**的），然后旋转角度 τ，我们称为**扭转**。 (15.10)

[123]正如我们已经注意到并试图纠正的那样，贝尔特拉米的名字目前并未附在他的双曲平面模型上。雪上加霜的是，普通的“大街上的数学家”不知道正是同一个欧金尼奥·贝尔特拉米（Eugenio Beltrami）也第一个发现了 SVD ！ 见斯图尔特 (1993)。



[15.4] 奇异值分解 (SVD)：该图的上半部分从左到右说明了以原点为中心的圆的一般线性变换的效果。 保留中点使我们能够证明 M 等价于两个垂直拉伸（通过“奇异值”σ_1 和 σ_2），然后是旋转，我们称之为扭转 $\equiv \tau$； 这里$\tau = \varphi - \theta$。 图的下半部分表明这个 SVD 等价于能够写出 $M = R_{\varphi} \circ \Sigma \circ R_{-\theta}$，其中 Σ 水平延伸 σ_1，垂直延伸 σ_2。

要理解这一点，请考虑图 [15.4] 的上半部分，它显示了一般线性变换对以原点为中心的圆 C 的影响（从左到右）。由于 C 的笛卡尔方程是二次的，由变换引起的坐标线性变化将导致像曲线的另一个二次方程。 因此像曲线 $\tilde{C}$ 是一个圆锥截面，并且由于 C 的有限点没有被发送到无穷远，这个圆锥曲线必然是一个椭圆（如 [15.4] 的右上角所示）。

我们刚刚使用了线性性质的代数陈述； 接下来我们使用基本的几何事实是，如果我们将两个向量相加然后映射结果，或者如果我们先映射向量然后将其相加，它们没有区别。 说服自己承认这两个简单的结果：

- 平行线映射到平行线。
- 线段的中点映射到像线段的中点。

我们现在将这些事实应用[124]于 $\tilde{C}$。

由于 C 的所有直径都被 C 的中心平分，因此 $\tilde{C}$ 上弦的像必须全部通过一个公共的平分点。 因此圆 C 的中心映射到椭圆 $\tilde{C}$ 的中心。

粗线画出圆的特定直径 d 使得其像映射到椭圆长轴 $\tilde{d}$（也是粗线画出）。 现在考虑垂直于 d 的 C 的弦（以虚线画出）。 由于这些都被 d 一分为二，因此它们的像必然是 $\tilde{C}$ 的一组平行的弦，使得 $\tilde{d}$ 是它们的公共平分线。 因此，这组平行的弦必然垂直于 $\tilde{d}$。 （通

[124] 以下论点以前出现在 VCA，第 208 页。

过在一般方向上画出一组 $\tilde{C}$ 的平行弦来说服自己。）

现在很明显，线性变换是在 d 方向上的拉伸，另一个垂直于它的拉伸，最后是扭转，从而证实了 SVD 的存在。 注意，我们同样可以在扭转之后而不是之前进行正交拉伸。 但在那种情况下，拉伸将在新的、扭转的方向上，沿着它，并与之正交。

图的下半部分演示了 SVD 相当于可以写

$$\boxed{M = R_{\varphi} \circ \Sigma \circ R_{-\theta}} \tag{15.11}$$

其中 Σ 水平拉伸 σ_1，垂直拉伸 σ_2。

请注意，此结果在计数自由度的水平上也是有意义的。 正如矩阵有四个独立的数字条目一样，我们变换的规格也需要四位几何信息：d 的方向、该方向的拉伸因子、垂直方向拉伸因子和扭转。

如果我们要求将两个拉伸因子设置为相等，则会出现一个非常重要的特殊情况：$\sigma_1 = \sigma_2 = a$，比方说，在这种情况下，一个圆的像是另一个 a 倍大的圆。 这显然将自由度的数量从四个减少到三个。 然而，由于我们现在在所有方向上产生相等的膨胀，为 d 选择的方向变得无关紧要，因此我们只剩下两个真正的自由度：膨胀因子 a 和扭转 τ。

在 VCA 中，我们试图从几何上利用一个事实，即复变量的可微函数 f(z) 的局部效应正是刚刚描述的类型的映射。 在这种情况下，a(z) 和 $\tau(z)$ 通常都随位置变化：它们描述了 z 的无穷小邻域所经历的膨胀和旋转。

更明确地说，回顾 （4.18），从 z 发出的每个微小的复箭头 δz 都被“伸扭”为从 $\tilde{z} = f(z)$ 发出的像箭头 $\delta\tilde{z}$，其中 $\delta\tilde{z} \asymp [ae^{i\tau}]\,\delta z$。在 VCA 中，我们将 a(z) 称为 f(z) 的拉伸，我们将局部拉伸和扭转的组合效应称为 f(z) 的伸扭，编码为复数：

$$\boxed{f'(z) = f(z)\text{的伸扭} = \left(\text{拉伸}\right)e^{i(\text{扭转})} = ae^{i\tau}}$$

现在让我们继续我们对 M^T 含义的探索。 为了促进这个探索，首先考虑原始线性变换 M 的逆（见 [15.4] 的顶部）。 这个逆 M^{-1} 在图 [15.5a] 中从右到左绘制：首先它通过扭转 $-\tau$ 消除 M 的扭转 τ，然后它通过 $1/\sigma_1$ 和 $1/\sigma_2$ 在相同的两个正交 M 最初做拉伸的方向。 颠倒图 [15.4] 下半部分的箭头，我们也可以将其表示为

$$M^{-1} = R_{\theta} \circ \Sigma^{-1} \circ R_{-\varphi}$$

最后我们可以揭示 M^T 的几何意义。我们假设读者熟悉矩阵转置运算的代数性质，我们将使用它们来揭示底层几何。 首先请注意，对旋转矩阵（15.5）进行转置会产生相反的旋

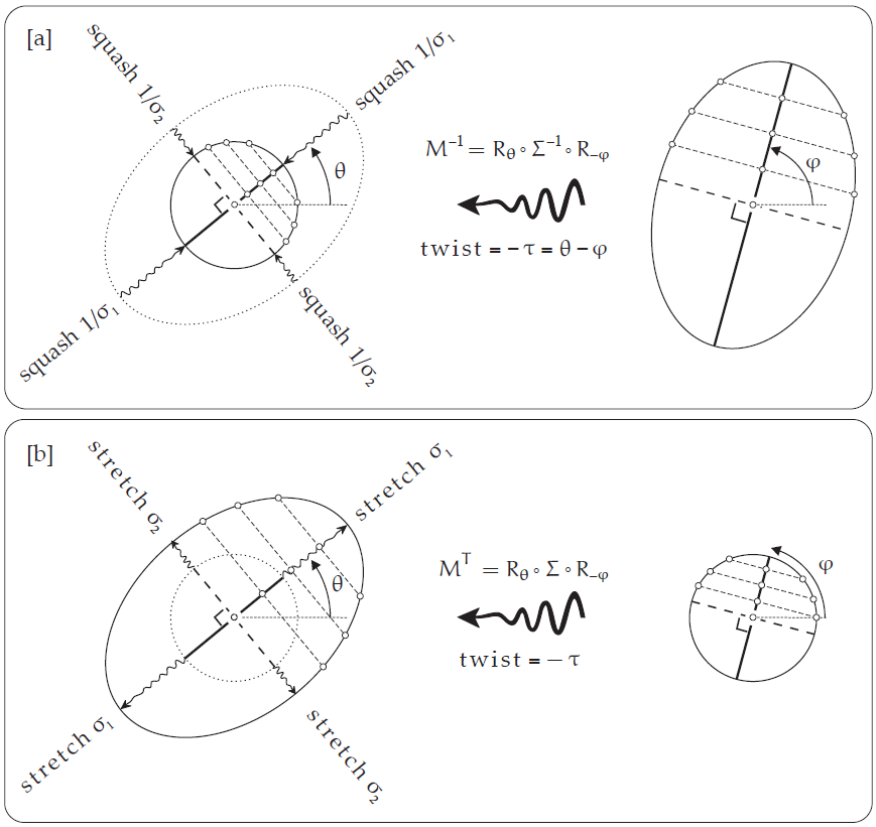
转：$[R_\theta]^T = [R_{-\theta}]$。 另外，回想一下 $([P][Q])^T = [Q]^T [P]^T$ 。 因此（15.11）得出，

图 [15.5b] 说明了这个意思：

$$\boxed{\text{线性变换 } M^T \text{ 对 } M \text{ 进行了相反的扭转，然后是与 } M \text{ 相同的两个正交拉伸。}} \tag{15.12}$$

掌握了 M^T 的几何意义，对称的几何意义也一下子明白了：

$$\boxed{\text{线性变换 } M \text{ 是对称的（即 } MT = M\text{）当且仅当扭转消失。 但是如果}\tau = 0\text{，那么正交的展开方向就变成了特征向量，所以我们也得出结论，当且仅当它有正交的特征向量时是不对称的。}} \tag{15.13}$$



[15.5] [a] [15.4] 顶部原始线性变换的逆 M^{-1}（从右到左）：它通过首先做相反的扭曲 $-\tau$，然后按 $1/\sigma_1$ 和 $1/\sigma_2$ 在与 M 最初进行拉伸相同的两个正交方向上压缩来取消 M。 [b] 转置 MT 也从扭曲$-\tau$ 开始，但随后它进行了与 M 相同的两个正交拉伸。

如果 $\tau = \varphi - \theta = 0$，则图 [15.4] 下半部分的两个旋转相等且相反，因此（15.11）特化：

$$\boxed{\text{如果 } M^T = M \text{ 则 } M^{-1} = R_\theta \circ \Sigma \circ R_{-\theta}} \tag{15.14}$$

这有时称为**谱定理**。

相反，我们可以将（15.12）作为我们对 M^T 的定义，然后转置的所有代数性质将成为易于可视化的几何定理。 例如，检查 $R_\theta^T = R_{-\theta}$，这意味着 $[R_\theta]T = [R_{-\theta}]$。 同样，检查 $(P \circ Q)^T = Q^T \circ P^T$，这意味着 $([P][Q])^T = [Q]^T [P]^T$ 。

这是传统上通过计算证明的结果的一个新示例，现在通过几何图形透明呈现：M^T 以与 M 相同的系数扩展面积：

$$\boxed{\left|M^T\right| = \sigma_1\sigma_2 = |M|}$$

这是在每个标准课本中都用代数证明的另一个结果，而现在几何让我们理解。 如果我们先做 M（图 [15.4] 的顶部）然后做 M^T（图[15.5b]），那么相反的扭转取消，我们可以看到

$$\boxed{\text{变换 } (M^T \circ M) \text{ 是对称的，具有正交特征向量（沿 } d \text{ 的和垂直于它的），特征值为 } \sigma_1^2 \text{ 和 } \sigma_2^2\text{。}} \tag{15.15}$$

最后，我们提到另一种常用的对称表现形式，请自行在代数和几何上验证这一点：

$$\boxed{\begin{array}{c}\text{线性变换 } M \text{ 是对称的（即，} M^T = M\text{） 当且仅当对于所有 } \mathbf{a} \text{ 和 } \mathbf{b} \\ \mathbf{a} \cdot M(\mathbf{b}) = \mathbf{b} \cdot M(\mathbf{a})\end{array}} \tag{15.16}$$

有了转置的几何意义，练习 12 提供了另外八个我们可以称之为“可视化线性代数”的例子。这些例子强调了我们的总体理念：直接的几何推理经常允许我们完全绕过符号操作，以获得对数学现实的直观、视觉的把握。

15.5 S 的一般矩阵

形状算子 S 是一个几何实体，独立于任何特定的基向量选择。 但是表示 S 的矩阵 $[S]$ 确实取决于这个选择。

假设我们选择一组任意的正交基 $\{\mathbf{E}_1, \mathbf{E}_2\}$，而不先找到主方向和曲率； 矩阵看起来如何？ 虽然我们可能还不知道主方向，但它们肯定存在，所以让我们假设 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 是通过将 $\{\mathbf{E}_1, \mathbf{E}_2\}$ 旋转某个未知角度 θ 获得的。 图 [15.6] 在 $\kappa_1 > \kappa_2 > 0$ 的情况下说明了这一点。在左边，我们看到 S 对单位圆的影响，它被带入一个椭圆，长半轴 κ_1 与 $\mathbf{e}_1$ 对齐，短半轴 κ_2 与 $\mathbf{e}_2$ 对齐。

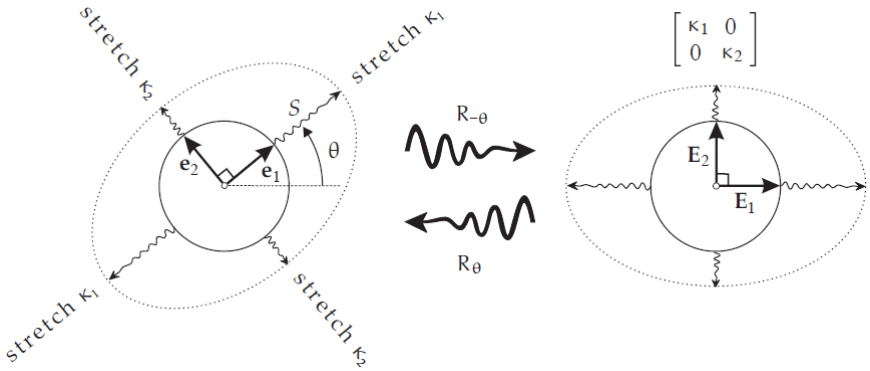
正如我们在右侧看到的，就 $\{\mathbf{E}_1, \mathbf{E}_2\}$ 而言，S 的效果等效于以下三个连续变换：

$$S = (\text{旋转} - \theta)\ \text{然后}\ (\text{沿}\ \mathbf{E}_{1,2}\ \text{扩展}\ \kappa_{1,2})\ \text{然后}\ (\text{旋转}\ \theta)$$

如果你阅读了上一节，那么这作为（15.14）的结果你应该已经很熟悉了。

回忆一下线性代数，两个矩阵的乘积 $[B][A]$ 被（或至少应该是！）定义为复合线性变换的矩阵：$[B][A] \equiv [B \circ A]$。 由 （15.5）可知

$$\begin{aligned}[S] &= [R_\theta]\begin{bmatrix}\kappa_1 & 0\\ 0 & \kappa_2\end{bmatrix}[R_{-\theta}] \\ &= \begin{bmatrix}c & -s\\ s & c\end{bmatrix}\begin{bmatrix}\kappa_1 & 0\\ 0 & \kappa_2\end{bmatrix}\begin{bmatrix}c & s\\ -s & c\end{bmatrix} \\ &= \begin{bmatrix}\kappa_1 c^2 + \kappa_2 s^2 & (\kappa_1 - \kappa_2)sc\\ (\kappa_1 - \kappa_2)sc & \kappa_1 s^2 + \kappa_2 c^2\end{bmatrix}\end{aligned} \tag{15.17}$$



[15.6] 在左侧，形状算子 S 以等于主曲率的因子拉伸正交主方向，将单位圆变形为与主方向对齐的椭圆。 S 的这种效果相当于首先旋转 $-\theta$，产生右边的图形，然后水平拉伸 κ_1，垂直拉伸 κ_2，然后旋转 θ。

这四个条目中的至少一个应该看起来很熟悉。 令 $\kappa(\mathbf{E}_1)$ 表示沿 $\mathbf{E}_1$ 方向截取的 S 法向截面的曲率，其与第一主方向的夹角为 $-\theta$。 然后欧拉公式（10.1）告诉我们左上角的条目是 $\kappa_1 c^2 + \kappa_2 s^2 = \kappa(\mathbf{E}_1)$。

但是为什么会出现这种简化呢？

15.6 S 的几何解释与 [S] 的简化

答案揭示了形状运算符的含义：

> 如果 $\hat{\mathbf{v}}$ 是任意单位切向量，则该方向法向截面的曲率 $\kappa(\hat{\mathbf{v}})$ 由 $\kappa(\hat{\mathbf{v}}) = \hat{\mathbf{v}} \cdot S(\hat{\mathbf{v}})$ 给出。 因此，
>
> $\kappa(\hat{\mathbf{v}}) = S(\hat{\mathbf{v}})$ 在 $\hat{\mathbf{v}}$ 方向上的投影。

(15.18)

（注意：根据 Meusnier 定理（11.4），$\kappa(\hat{\mathbf{v}})$ 不仅是法向截面的曲率，它也等于之前引入的曲面上在方向 $\hat{\mathbf{v}}$ 通过该点的任何曲线的法向曲率 $\kappa_n(\hat{\mathbf{v}})$。）

这种解释在图 [15.7] 中有说明。 请特别注意，这种构造正确地得出 $\kappa(\mathbf{e}_1) = \kappa_1$ 和 $\kappa(\mathbf{e}_2) = \kappa_2$。

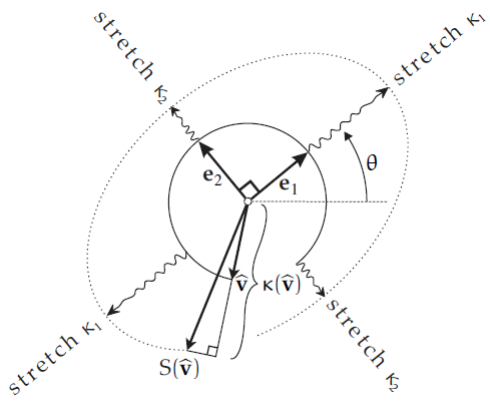
对此的基本几何解释与图 [8.5] 中所示的相同。事实上，我们之前证明了形式（11.5）中的结果，但我们在这里重复证明以说明我们的新符号。 将 $\hat{\mathbf{v}}$ 视为粒子以单位速度沿法向截面运动的速度，使得 $\nabla_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = \kappa(\hat{\mathbf{v}})\,\mathbf{n}$ 是其朝向曲率中心的加速度，我们发现

$$0 = \nabla_{\hat{\mathbf{v}}}(\hat{\mathbf{v}} \cdot \mathbf{n}) = \hat{\mathbf{v}} \cdot \nabla_{\hat{\mathbf{v}}}\mathbf{n} + \mathbf{n} \cdot \nabla_{\hat{\mathbf{v}}}\hat{\mathbf{v}} = -\hat{\mathbf{v}} \cdot S(\hat{\mathbf{v}}) + \kappa(\hat{\mathbf{v}})$$

从上式立刻可以得出 （15.18）。

如果我们采取欧拉公式（10.1），我们可以给出另一种推导。 目前我们使用 θ 来表示从 $\mathbf{E}_1$ 到 $\mathbf{e}_1$ 的角度，因此，为了避免与我们之前对欧拉公式的讨论混淆，让我们改用 α 来表示一般 $\hat{\mathbf{v}}$ 与 $\mathbf{e}_1$ 所成的角度，这样欧拉公式现在就成为

$$\kappa(\hat{\mathbf{v}}) = \kappa_1 \cos^2 \alpha + \kappa_2 \sin^2 \alpha$$



[15.7] 沿一般单位切向量 $\hat{\mathbf{v}}$ 方向截取的曲面法向截面的曲率由 $\kappa(\hat{\mathbf{v}}) = S(\hat{\mathbf{v}})$ 在 $\hat{\mathbf{v}}$ 方向上的投影 给出。

由于（15.18）是一个纯粹的几何陈述，正如在[15.7]中所见证的，它足以证明任何特定基的真实性。 在主基 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 中，一个简单的坐标计算（由欧拉公式）证实了结果：

$$\hat{\mathbf{v}} \cdot S(\hat{\mathbf{v}}) = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} \cdot S \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} \cdot \begin{bmatrix} \kappa_1 \cos\alpha \\ \kappa_2 \sin\alpha \end{bmatrix} = \kappa(\hat{\mathbf{v}})$$

有了结果（15.18），我们现在可以回到（15.17）的解释和简化。 回想一下，即使 S 是任意线性变换，其矩阵的第 j 列将是第 j 个基向量的像，因此 $[S]$ 的第一列是向量 $S(\mathbf{E}_1)$，因此其沿 $\{\mathbf{E}_1, \mathbf{E}_2\}$（即其分量）的投影是该向量与这些基向量的点积。 同样，对于 $[S]$ 的第二列也是如此，所以

$$[S] = \begin{bmatrix} \mathbf{E}_1 \cdot S(\mathbf{E}_1) & \mathbf{E}_1 \cdot S(\mathbf{E}_2) \\ \mathbf{E}_2 \cdot S(\mathbf{E}_1) & \mathbf{E}_2 \cdot S(\mathbf{E}_2) \end{bmatrix} \tag{15.19}$$

因此（15.18）解释了[125]为什么（15.17）的左上角条目结果是$\kappa(\mathbf{E}_1)$。 现在看看另一个对角线条目。无论我们使用欧拉公式[练习]还是参考（15.18）和（15.19），我们看到（15.17）的右下方条目是 $\kappa_1 s^2 + \kappa_2 c^2 = \kappa(\mathbf{E}_2)$。如果您阅读了前面（可选）小节，还请注意（15.16）表明 $[S]$ 的对称性现在遵循 S 的对称性，因为 $\mathbf{E}_1 \cdot S(\mathbf{E}_2) = \mathbf{E}_2 \cdot S(\mathbf{E}_1)$。

在我们进行更多计算之前，让我们停下来更明确地讨论（15.19）中条目的几何意义。正如我们所说，第一列是 $S(\mathbf{E}_1) = -\nabla_{\mathbf{E}_1}\mathbf{n}$，它测量法线在我们沿 $\mathbf{E}_1$ 方向移动时如何转动。设 Π_1 为该方向的法线平面，即由 $\mathbf{E}_1$ 和 $\mathbf{n}$ 张成的平面。有关具体示例，请参阅图 [12.5b]。然后 $S(\mathbf{E}_1)$ 的第一个分量，即 $\mathbf{E}_1 \cdot S(\mathbf{E}_1)$，测量当我们开始沿 $\mathbf{E}_1$ 移动时 $\mathbf{n}$ 在 Π_1 内指向 $\mathbf{E}_1$的速度；这是由法向截面的曲率决定的，实际上它是曲率 $\kappa(\mathbf{E}_1)$。$S(\mathbf{E}_1)$ 的第二个分量，即 $\mathbf{E}_2 \cdot S(\mathbf{E}_1)$，测量 $\mathbf{n}$ 垂直于运动方向旋转的速度，即它从 Π_1 *旋转出来*的速度。

当然，第二列 $S(\mathbf{E}_2)$ 的含义完全类似。 第一个分量 $\mathbf{E}_1 \cdot S(\mathbf{E}_2)$ 测量当我们开始沿 $\mathbf{E}_2$ 移动时 $\mathbf{n}$ 从 Π_2 旋转的速度。 第二个分量 $\mathbf{E}_2 \cdot S(\mathbf{E}_2)$，测量 $\mathbf{n}$ 在 Π_2 内旋转的速度，这就是 $\kappa(\mathbf{E}_2)$。

对于图 [12.5b] 中所示的典型方向，Π 内与 κ 相关的倾斜和 Π 外的旋转都清晰可见。然而，如果我们在一个主要方向上移动，那么所有的倾斜都在运动方向上，在 Π 内，并且没有从 Π 开始的旋转，这就是我们在图 [12.5a] 中看到的。

让我们更仔细地看一下非对角线项，它衡量 $\mathbf{n}$ 摆出 $\Pi_{1,2}$ 的速度有多快。 如果我们再次让 $\Delta\kappa = (\kappa 1 - \kappa 2)$，如 （10.2）中所示，那么表示一般基 $\{\mathbf{E}_1, \mathbf{E}_2\}$ 中的形状算子的矩阵由下式给出

$$\boxed{[S] = \begin{bmatrix} \kappa(\mathbf{E}_1) & \frac{\Delta\kappa}{2}\sin 2\theta \\ \frac{\Delta\kappa}{2}\sin 2\theta & \kappa(\mathbf{E}_2) \end{bmatrix}} \tag{15.20}$$

注意，如（15.9）所要求，这个通用矩阵确实是对称的：$[S]^T = [S]$。 还要注意，如果基与主基重合，使得 $\theta = 0$，那么（15.20）就应该简化为对角线形式（15.7）。如果 $\theta = (\pi/2)$，一般矩阵也简化为这种对角线形式，因为在这种情况下，基再次与主基对齐，只是现在 $\{\mathbf{E}_1, \mathbf{E}_2\} = \{-\mathbf{e}_2, \mathbf{e}_1\}$。 另一方面，当 $\theta = (\pi/4)$ 时，非对角线项最大（$\mathbf{n}$ 以最快的速度从 $\Pi_{1,2}$ 摆出），在这种情况下，$\Pi_{1,2}$ 平分主方向。

15.7 [S] 完全由三个曲率决定

接下来，观察到两个相等的非对角线项只是欧拉公式 (10.2) 中的振荡项，相移 $\pm(\pi/4)$。

[125] 反过来，我们又得到了欧拉公式的另一个推导。

事实上，这个公式告诉我们法向截面在 $\mathbf{E}_1+\mathbf{E}_2$ 方向上的曲率，平分基向量之间的角度，由下式给出

$$\kappa(\mathbf{E}_1+\mathbf{E}_2)=\kappa\left(\frac{\pi}{4}-\theta\right)=\bar{\kappa}+\frac{\Delta\kappa}{2}\sin 2\theta \tag{15.21}$$

当然，在正交方向上同样好使：

$$\kappa(\mathbf{E}_1-\mathbf{E}_2)=\kappa\left(\theta-\frac{\pi}{4}\right)=\bar{\kappa}-\frac{\Delta\kappa}{2}\sin 2\theta \tag{15.22}$$

但是，选择（有点随意）使用第一个方向，（15.20）可以写成

$$[S]=\begin{bmatrix}\kappa(\mathbf{E}_1) & \kappa(\mathbf{E}_1+\mathbf{E}_2)-\bar{\kappa}\\ \kappa(\mathbf{E}_1+\mathbf{E}_2)-\bar{\kappa} & \kappa(\mathbf{E}_2)\end{bmatrix} \tag{15.23}$$

乍一看，我们似乎仍然需要知道 $\kappa_{1,2}$（或至少它们的总和）才能计算 κ。 但实际上我们不需要，我们将看到形状算子矩阵可以纯粹用三个方向上的法线截面的曲率来表示：$\mathbf{E}_1$、$\mathbf{E}_2$ 和 $(\mathbf{E}_1+\mathbf{E}_2)$。

回想一下，矩阵的**迹**是其对角元素之和，因此对于原始对角矩阵（15.7），迹是 $\mathrm{Tr}\,[S]=\kappa_1+\kappa_2=2\bar{\kappa}$。 但是我们从线性代数知道，如果[A]和[B]是平面任意两个线性变换的矩阵，则$\mathrm{Tr}\,[A][B]=\mathrm{Tr}\,[B][A]$。 因此，

$$\mathrm{Tr}[R_\theta][A][R_{-\theta}]=\mathrm{Tr}[A][R_{-\theta}][R_\theta]=\mathrm{Tr}[A]$$

换句话说，

任何线性变换的迹在基向量的旋转下保持不变。 (15.24)

（这实际上有一个几何原因；参见 Arnol’d (1973, §16.3)。）

因此，回到手头的情况，其中 $[A]=[S]$，我们推导出即使在一般基的情况下，使用矩阵(15.23)，

$$\kappa(\mathbf{E}_1)+\kappa(\mathbf{E}_2)=\mathrm{Tr}\,[S]=\kappa_1+\kappa_2=2\bar{\kappa}$$

事实上，我们甚至不需要求助于线性代数的一般定理来看到这一点，因为在我们的例子中，我们已经明确计算了 $[S]$，并且我们的断言的真实性立即从 （15.17）得出。 另一个具体例子通过将（15.21）和（15.22）相加得出。

换句话说，我们有一个本身就很有趣的结果：

任意两个垂直方向上的曲率之和等于主曲率之和。

因此，正如所声称的，矩阵（15.23）确实仅取决于三个方向的曲率，对于 $\bar{\kappa} = \frac{1}{2}\,[\kappa(\mathbf{E}_1) + \kappa(\mathbf{E}_2)]$，无需知道主曲率。 因此，正如所声称的，形状算子的一般矩阵（15.23）可以根据这三个曲率明确写出：

$$[\mathrm{S}] = \begin{bmatrix} \kappa(\mathbf{E}_1) & \kappa(\mathbf{E}_1 + \mathbf{E}_2) - \dfrac{1}{2}\,[\kappa(\mathbf{E}_1) + \kappa(\mathbf{E}_2)] \\ \kappa(\mathbf{E}_1 + \mathbf{E}_2) - \dfrac{1}{2}\,[\kappa(\mathbf{E}_1) + \kappa(\mathbf{E}_2)] & \kappa(\mathbf{E}_2) \end{bmatrix}$$

最后，回忆一下 （15.8），高斯曲率的外在版本 $\mathcal{K}_{\mathrm{ext}}$ 是 $[\mathrm{S}]$ 矩阵的任何这些形式的行列式。 例如，（15.23）得出

$$\mathcal{K}_{\mathrm{ext}} = |[\mathrm{S}]| = \kappa(\mathbf{E}_1)\kappa(\mathbf{E}_2) - [\kappa(\mathbf{E}_1 + \mathbf{E}_2) - \bar{\kappa}]^2$$

15.8 渐近方向

回想一下，**迪潘标形**（Dupin indicatrix）$\mathcal{D}$ 来自曲面与平行于切平面 $\mathrm{T_p}$ 且距离为 ϵ 的平面 $\mathrm{T_p}(\epsilon)$ 的交集，因此 $\mathrm{T_p}(0) = \mathrm{T_p}$。 当 ϵ 从 0 增加并且 $\mathrm{T_p}(\epsilon)$ 刚开始沿法线方向移开时，$\mathcal{D}$ 是与曲面相交的新生圆锥截线。

若$\mathcal{K}(\mathrm{p}) > 0$则 $\mathcal{D}$ 为椭圆型，p 称为椭圆点； 如果 $\mathcal{K}(\mathrm{p}) = 0$ 则 $\mathcal{D}$ 是抛物型，p 称为抛物点； 如果 $\mathcal{K}(\mathrm{p}) < 0$，则 $\mathcal{D}$ 是双曲型，而 p 称为双曲点。

在这里，我们将重点关注双曲线（负曲率）情况，并将研究双曲线 $\mathcal{D}$ 的渐近线的方向，称为*渐近方向*。 正如我们现在将要解释的，渐近方向与形状算子 S 具有简单的几何关系。

如前一章所述，如果 (x,y) 轴与主方向对齐，则共轭双曲线对 $\mathcal{D}$ 的方程为

$$\kappa_1 x^2 + \kappa_2 y^2 = \pm 1$$

符号由 ϵ 的符号决定，即我们是否向上或向下移动切平面。 对称轴（在椭圆和双曲线情况下）与主轴/坐标轴重合，在双曲线情况下，这些对称轴将渐近方向之间的角度一分为二。

让 $\kappa_1 = \mathrm{p}^2$ 和 $\kappa_2 = -\mathrm{q}^2$（p 和 q 为正），让我们为 $\mathcal{D}$ 选择上述等式中的加号。然后，

$$\mathcal{D}: \mathrm{p}^2 x^2 - \mathrm{q}^2 y^2 = 1 \quad \text{且} \quad [\mathrm{S}] = \begin{bmatrix} \mathrm{p}^2 & 0 \\ 0 & -\mathrm{q}^2 \end{bmatrix}$$

因此渐近线具有方程 $\mathrm{y} = \pm\frac{\mathrm{p}}{\mathrm{q}}\mathrm{x}$（参见图 [15.8]），其中我们添加了单位圆 $\mathcal{C}$ 及其椭圆图像 $\mathrm{S}(\mathcal{C})$，以帮助可视化映射 S 的效果，该映射 S 水平和垂直拉伸，并在 x 轴上翻转。因此，渐近方向上的向量及其在 S 下的图像由下式给出，

$$\mathbf{a}_\pm \propto \begin{bmatrix} q \\ \pm p \end{bmatrix} \quad \Longrightarrow \quad \mathrm{S}(\mathbf{a}_\pm) \propto \begin{bmatrix} p \\ \mp q \end{bmatrix}$$

如图 [15.8] 中所示，这意味着尽管 S 将主方向映射到同一方向，

> 形状算子将渐近方向映射到正交方向：
>
> $$\mathbf{a}_\pm \cdot \mathrm{S}(\mathbf{a}_\pm) = 0$$

(15.25)

事实上，正如我们在 [15.8] 中看到的，S 将两个渐近方向转向相反的方向，$\mathbf{a}_+$ 为 $-\frac{\pi}{2}$，$\mathbf{a}_-$ 为 $+\frac{\pi}{2}$。

但为什么会这样呢？ 我们已经将 $\mathcal{D}$ 画为曲面与 $\mathrm{T_p}(\epsilon)$ 的放大交叉曲线，就像 ϵ 从 0 增加一样。但是，最初，当 $\epsilon = 0$ 时，这个交叉双曲线退化为渐近线本身。 由于这条相交曲线（方向为 $\mathbf{a}_\pm$）位于切平面内，它具有消失的法向曲率，因此根据形状算子的几何解释（15.18）得出

$$\kappa(\hat{\mathbf{a}}_\pm) = \hat{\mathbf{a}}_\pm \cdot \mathrm{S}(\hat{\mathbf{a}}_\pm) = 0$$

从而解释结果。

鉴于与法曲率的这种联系，渐近方向的定义已被概括为法向曲率消失的任何方向。 因此，抛物线点也有一个“渐近方向”[练习：它是什么？] 尽管抛物线 $\mathcal{D}$ 没有任何渐近线。

如果我们选择 $\mathbf{E}_1$ 作为渐近方向之一，那么 Π_1 是这个方向的法线平面，那么当我们沿着 $\mathbf{E}_1$ 在曲面内移动时，法线 n 必须围绕 $\mathbf{E}_1$ 旋转，直接从 Π_1 摆动出来，它的 头部沿与曲面相切并垂直于 $\mathbf{E}_1$ 的方向移动，即沿 $\pm\mathbf{E}_2$ 方向移动。 因此 $\mathrm{S}(\mathbf{E}_1) = \pm\tau\,\mathbf{E}_2$，其中 τ 是 $\mathbf{n}$ 围绕 $\mathbf{E}_1$ 的自旋速率，即第 106 页介绍的所谓的挠率[126]。从（15.19）得出

> 在双曲线点，曲面的曲率可以用渐近曲线的挠率 τ 表示为
>
> $$\mathcal{K}_{\mathrm{ext}} = |[\mathrm{S}]| = -\tau^2$$

(15.26)

根据斯托克 (1969, p. 101) 的说法，这个结果归功于贝尔特拉米 (Beltrami) 和恩内珀 (Enneper)。

---

[126] 我们希望这里不会因为字母 τ 的双重职责而引起混淆：我们最近还用它来表示 SVD 分解的（不相关的）“扭转”。

15.9 经典术语与符号：三种基本形式

在查阅有关微分几何的较旧的经典著作时，你会遇到本书中未使用的术语和符号。 特别是，你肯定会遇到所谓的三种基本形式，分别表示为 I、II 和 III。

所有三种基本形式都是成对的切向量 $\mathbf{u}$ 和 $\mathbf{v}$ 的对称函数。

第一基本形式： $\mathrm{I}(\mathbf{u},\mathbf{v}) \equiv \mathbf{u}\cdot\mathbf{v}$
第二基本形式： $\mathrm{II}(\mathbf{u},\mathbf{v}) \equiv \mathrm{S}(\mathbf{u})\cdot\mathbf{v}$
第三基本形式： $\mathrm{III}(\mathbf{u},\mathbf{v}) \equiv \mathrm{S}(\mathbf{u})\cdot\mathrm{S}(\mathbf{v})$

形状算子本身没有出现在经典文献[127]中； 它的数学内容由第二基本形式表示。 例如，由（15.18）给出的 $\hat{\mathbf{v}}$ 方向的法线截面的曲率，在过去应该写成，$\kappa(\mathrm{bv}) = \mathrm{II}(\hat{\mathbf{v}},\hat{\mathbf{v}})$。

警告：虽然在第五幕之前我们不会介绍和定义**微分形式**（简称形式），但我们应该立即提醒读者，这三种经典的“形式”根本不是形式！ 虽然我们当然不会责怪继续使用古典语言的现代作者，但我们对真正形式的依赖促使我们将古典语言变成了死语言！

[127]形状运算符首先受到 Barrett O'Neill 在他开创性的导论性课本（O'Neill 2006）中的大力支持并被普遍使用，该课本的第一版于 1966 年出版。

第十六章 全局高斯-博内定理简介

16.1 一些拓扑和结果的陈述

全局高斯–博内定理被广泛认为是所有数学中最美丽的结果之一。 此外，它也是基础性的，催生了更强大的概括的发现，其高潮可能是（目前）1963 年发现的**阿蒂亚–辛格**（Atiyah-Singer）**指标定理**。反过来，这个定理在数学和理论物理的其他领域引起了翻天覆地的变化。这些后续的发展远远超出了当前工作的范围，可悲的是超出了当前作者的能力，但 GGB 原始形式的陈述（我们将在下文中缩写）出奇地简单且易于理解。 在我们可以陈述结果之前，只需要少许预备知识。

首先，曲面上区域 P 的**总曲率** $\mathcal{K}(\mathrm{P})$ 定义（很自然）为，

$$\mathcal{K}(\mathrm{P}) \equiv \iint_{\mathrm{P}} \mathcal{K}\, \mathrm{d}\mathcal{A}$$

（注意：即使在现代，这个概念偶尔仍会以其古老的拉丁名称 Curvatura Integra 来指代。）例如，如果 P 是一张以任意简单曲线为边界的平面纸，则 $\mathcal{K}(\mathrm{P}) = 0$. 如果我们现在将纸弯曲成不同的形状 $\tilde{\mathrm{P}}$（不拉伸），比如说，变成圆柱或圆锥的一部分，那么根据绝妙定理，仍然有 $\mathcal{K}(\tilde{\mathrm{P}}) = 0$。

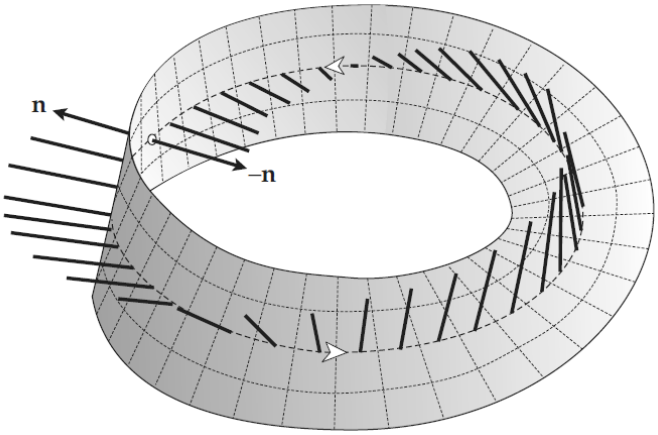
但现在假设 P 由高度可拉伸的材料制成，例如橡胶。 如果我们在球面曲面上拉伸 P，则 $\mathcal{K}(\tilde{\mathrm{P}}) > 0$。 事实上，通过选择球面尽可能小，我们可以让它的曲率随心所欲地变大，从而使 $\mathcal{K}(\tilde{\mathrm{P}})$ 变得随心所欲。 如果我们改为将 R 拉伸为伪球的一部分，则 $\mathcal{K}(\tilde{\mathrm{P}}) < 0$，并且我们可以将其设为负值。

如此连续、一对一的拉伸，$\mathrm{P} \mapsto \tilde{\mathrm{P}}$ 不保留长度甚至角度，称为**拓扑映射**或**拓扑变换**，或**同胚**。 正如古希腊几何寻找在严格的、保持距离的映射下不变的图形的特性一样，19 世纪出现了一个新的数学领域，称为**拓扑学**，寻找在拓扑变换下不变的特性。

显然，曲率的概念不属于拓扑：在点 p 附近拉伸曲面会改变 $\mathcal{K}(\mathrm{P})$ 的值。 事实上，我们刚刚看到 $\mathcal{K}(\tilde{\mathrm{P}})$ 可以取任何我们喜欢的值，无论是正值还是负值。 同样，更原始的长度和角度概念也不属于拓扑。 因此，乍一看，拓扑学似乎是一个相当琐碎或贫瘠的研究领域：任何有趣或微妙的东西怎么能在拓扑映射的任意复杂和极端扭曲中幸存下来？！

值得注意的是，事实并非如此。 在其主要父母（黎曼和庞加莱）的热情拥抱下，拓扑学迅速成长为强大而仁慈的九头蛇，解释和统一了不同和遥远的思想领域中的现象。

为了引入我们的第一个拓扑不变量，我们将注意力限制在同样可定向的闭合曲面上。任何这样的曲面都可以被描绘成 $\mathbb{R}^3$ 中实体对象的边界。 这样的曲面是自动可定向的，这意味着人们可以始终如一地决定 $\mathbf{n}$ 的两个相反选择中的哪一个是曲面的法线——让它从实体对象向外指向空旷的空间。



[16.1] **莫比乌斯带**是不可定向的：沿着中心线的完整回路携带法线 $\mathbf{n}$，返回到起点时变成 $-\mathbf{n}$。

但不是每个曲面都是可定向的吗？ 不！ 正如默比乌斯（Möbius）和 雷思丁（Listing）在 1858 年独立发现的那样，取一条纸带，将其扭转半圈，然后将纸带的末端粘合在一起形成一个环，从而产生只有一面的曲面！ 如图 [16.1] 中所示，从 $\mathbf{n}$ 的任意初始选择开始，然后沿着中心线的完整回路连续携带 $\mathbf{n}$，我们将其返回到其起点为 $-\mathbf{n}$。 因此，这个所谓的**莫比乌斯带**(Möbius band[128])是不可定向的。

回到普通的、可定向的、封闭的曲面，拓扑上区分一个和另一个的基本特征是它包含的孔的数量。 这称为曲面的**亏格** g，它的值在图 [16.2] 中对一些曲面进行了说明。 每对具有相同属的曲面在**拓扑上是等价的**，或**同胚的**，这意味着一个曲面可以通过拓扑映射（又名**同胚**）变成另一个。 但是不同亏格的两个曲面不能拓扑等价，因为 g 在拓扑映射下是不变的。

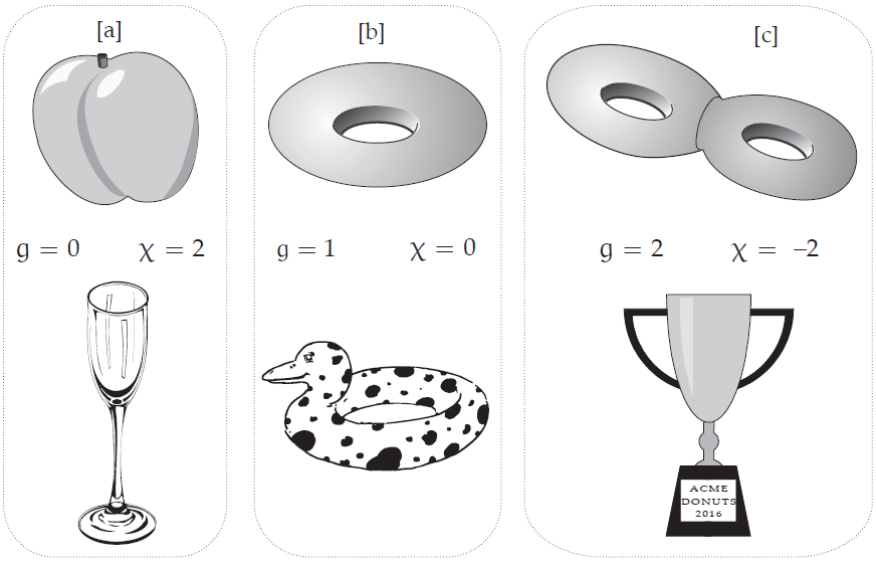
亏格由黎曼（Riemann）在 1851 年[129]更精确地定义为沿着曲面上闭合的、不相交的环路进行切割的最大数量，使得不将曲面分成两个不连续的部分。 例如，沿任何闭环切割球面会将其一分为二，因此亏格为零。 在环面上，我们可以只沿一个环路切割，而无需将环面分成两部分：例如，沿环绕对称轴的赤道切割，或沿穿过洞的圆（其平面包含对称轴）切割 ）。 但是如果我们现在沿着任何避开第一个的环切割所得曲面，将把它分成两个，所以环面的亏格是 1。 尝试（在你的脑海中）在一个有两个孔的甜甜圈上进行环路切割（如图 [16.2c] 中所示）并验证 $g = 2$。

正如默比乌斯在 1863 年发现的那样，每个封闭的、可定向的曲面在拓扑上都相当于一个 g 孔环面。 我们将在没有证据的情况下接受这些视觉上看似合理的陈述，但寻求精确定义和证明的读者应该参考我们在本书末尾的延伸阅读中推荐的优秀拓扑课本。

---

128 通常也称为莫比乌斯带（Möbius strip）。

129 参见 Stillwell (1995)，第 58 页。



[16.2] 每对曲面在拓扑上是相同的，通过洞的数量亏格 g 区别于其他对。 还显示了相应的欧拉示性数$\mathcal{X} = 2 - 2g$。

我们现在可以宣布惊人的结果：

全局高斯-博内定理（GGB）。 封闭的可定向曲面 $\mathcal{S}_g$ 的总曲率仅取决于其拓扑亏格 g，并由下式给出

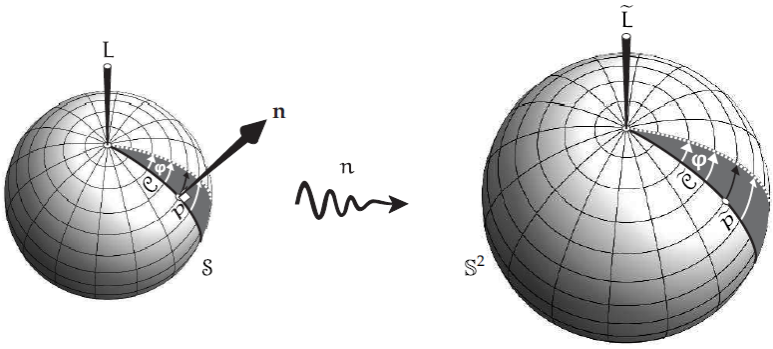
$$\boxed{\mathcal{K}(\mathcal{S}_g) = 4\pi(1-g) = 2\pi\mathcal{X}(\mathcal{S}_g)} \tag{16.1}$$

此处的量

$$\boxed{\mathcal{X}(\mathcal{S}_g) \equiv 2 - 2g} \tag{16.2}$$

（目前）仅作为标记不同曲面亏格的替代方法（参见图 [16.2] 示例）。 这个量 $\mathcal{X}(\mathcal{S}_g)$ 称为曲面的**欧拉示性数**； 它自然出现在许多拓扑结果中，并且实际上具有其自身的含义，这将在第 18.1 节中解释。

暂停片刻，让结果的惊喜和美丽沉浸一下。 如果我们取任何给定的 $\mathcal{S}_g$（例如 $g = 1$ 的简单甜甜圈），由橡胶或橡皮泥制成，然后拉伸、扭曲它， 挤压它，并以我们喜欢的任何方式变形，曲面某一点曲率的每一次增加都必然立即被曲面其他地方的曲率完全相反的减少所抵消。 在拓扑甜甜圈（又名**环面**）的情况下，在我们的整个操作过程中，总曲率保持为零。



[16.3] 当半圆 $\mathcal{C}$ 绕 L 旋转角度 ϕ 时，它的球面像 $\tilde{\mathcal{C}}$ 以相同的速率在 $\mathbb{S}^2$ 上旋转，产生相等 ϕ 的月牙面。 随着 ϕ 增加到 2π，$\tilde{\mathcal{C}}$ 扫过 $\mathbb{S}^2$ 的整个曲面，因此 $\mathcal{K}(\mathcal{S}) = 4\pi$。

16.2 球面和环面的总曲率

16.2.1球面的总曲率

如果曲面 $\mathcal{S}$ 在拓扑上等效于球面（即，$g = 0$），则 GGB 预测其应具有总曲率 4π。 在几何球面 $\mathcal{S}$ 的情况下，这很容易通过多种方式验证。我们给出了三个证明，最后一个证明可以推广到球面之外。

首先，半径为 R 的球面的每个法向截面本身就是一个半径为 R 的圆，即曲率 $(1/R)$。所以 $\mathcal{K} = \kappa_1\kappa_2 = (1/R^2)$，总曲率

$$\mathcal{K}(\mathcal{S}) = \iint_{\mathcal{S}} \mathcal{K}\, d\mathcal{A} = \iint_{\mathcal{S}} \frac{1}{R^2}\, d\mathcal{A} = \frac{1}{R^2} 4\pi R^2 = 4\pi$$

第二个更有启发性的解释是基于图 [12.2]：如果 P 是 $\mathcal{S}$ 上的一个区域，并且 $\tilde{P} = n(P)$ 是 P 的球面像，那么

$$\boxed{\mathcal{K}(P) = \left[\mathbb{S}^2\text{上 P 的球面像面积}\right] = \mathcal{A}(\tilde{P})} \tag{16.3}$$

现在取与 $\mathcal{S}$ 具有相同中心单位球面 $\mathbb{S}^2$，则球面映射变为径向投影（如图 [12.3]所示）。这清楚地表明 $\tilde{\mathcal{S}} = n(\mathcal{S}) = \mathbb{S}^2$，所以（16.3）意味着 $\mathcal{K}(\mathcal{S}) = \mathcal{A}(\mathbb{S}^2) = 4\pi$。

第三，最后，我们将球面 $\mathcal{S}$ 描绘为旋转曲面，通过将半径为 R 的半圆 $\mathcal{C}$ 通过其两端围绕直径 L 旋转获得，如图 [16.3] 中所述。

尽管我们可以通过将 $\mathcal{C}$ 旋转整整 2π 来生成整个球面，但该图描绘了在 $\mathcal{C}$ 旋转了角度 ϕ 之后生成所谓的**月牙面**的过程。

关键观察是这样的，它适用于由任意平面曲线 $\mathcal{C}$ 围绕其平面内的任意直线旋转而成的任意旋转曲面：

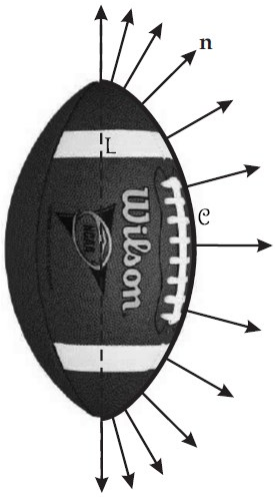
> 如果平面曲线 $\mathcal{C}$ 及其单位法向量 **n** **一起**绕 $\mathcal{C}$ 所在平面中的任意一条直线旋转，则旋转后的 **n** 是 $\mathcal{C}$ 生成的旋转曲面的法线。

从图中的特殊情况很容易理解。 典型点 p 处曲面的切平面由以下两个方向张成：(1) 旋转后 $\mathcal{C}$ 的方向； (2) p 旋转时移动的方向（即垂直于 $\mathcal{C}$ 所在平面）。 但是 **n** 最初垂直于这两个方向，因此当 $\mathcal{C}$ 和 **n** 一起旋转时，它仍然垂直于 (1) 和 (2)。 由于 **n** 垂直于张成 $\mathcal{S}$ 的切平面的两个方向，因此它是 $\mathcal{S}$ 的法线。

凭借（16.3），这立即意味着以下内容：

> 设 $\mathcal{C}$ 为平面曲线，L 为该平面内的一条直线，设 $\tilde{\mathrm{L}}$ 为通过 $\mathbb{S}^2$ 中心且与 L 平行的直线。那么，当 $\mathcal{C}$ 绕 L 旋转时，其球面像 $\tilde{\mathcal{C}} = \mathrm{n}(\mathcal{C})$ 绕 $\tilde{\mathrm{L}}$ 以相同的速率旋转，并且由 $\mathcal{C}$ 扫出的曲面的总曲率等于在 $\mathbb{S}^2$ 上由 $\tilde{\mathcal{C}}$ 扫出的总（有符号）面积。 (16.4)

特别地，如图所示，这个月牙面的球面像只是在 $\mathbb{S}^2$ 上等角的月牙面，如果 $\mathcal{C}$ 扫过所有的 $\mathcal{S}$，那么它的球面像扫过所有的 $\mathbb{S}^2$，再次确认 $\mathcal{K}(\mathcal{S}) = 4\pi$。



[16.4] 美式足球的球面像是整个 $\mathbb{S}^2$，所以根据GGB，$\mathcal{K}(\text{足球}) = 4\pi$。

最后一种观点的优势在于，我们现在对 GGB 本身有了第一次真正的了解。 考虑图 [16.4] 所示的美式足球，它是由曲线 $\mathcal{C}$ 围绕直线 L 旋转产生的。这里我们以相等的角度增量画出了法线，这清楚地表明了曲率的变化：在两极附近最大，并且 至少在赤道附近。 尽管如此，由于足球在拓扑上是球形的，所以总曲率应该是 4π，我们现在看到正是如此。 因

为 $\mathcal{C}$ 的球面像和之前一样是半圆 $\tilde{\mathcal{C}}$（连接 $\mathbb{S}^2$ 的两极），因此当 $\mathcal{C}$ 旋转产生足球时，$\tilde{\mathcal{C}}$ 旋转扫过所有 $\mathbb{S}^2$，就像以前一样，所以足球的总曲率确实是 4π！

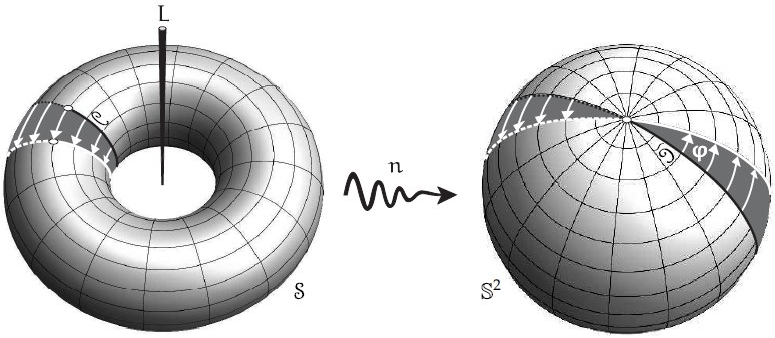
显然，关键是足球的球面像覆盖了整个 $\mathbb{S}^2$。 在下一章中，我们将详细阐述这个想法，以产生我们对 GGB 的第一个（启发式）证明。

16.2.2环面的总曲率

正如我们在（10.13）中所观察到的，如果 $\mathcal{C}$ 的一部分向远离 L 的方向凹，那么当它围绕 L 旋转时，它会生成具有负曲率的曲面部分。 在这种情况下，应从总曲率中减去 $\tilde{\mathcal{C}}$ 在 $\mathbb{S}^2$ 上产生的面积。 我们现在用环面来说明这一点，环面当然具有正曲率和负曲率的区域。

根据 GGB，甜甜圈（圆环）应该具有消失的总曲率，在第 89 页的练习 23 ，你实际上通过蛮力验证了这一点（在整个曲面上对曲率公式进行积分）。 我们现在可以给出一个真正的解释，事实上建立一个（曲面上）更强的、局部的（相对于全局的）结果。

假设我们只想吃甜甜圈的一部分：我们可以切出一个楔子——图[16.5] 中的深色阴影区域——就像切一块蛋糕一样，通过在对称轴上切两片。 我们现在将（16.4）应用到图 [16.5] 以证明甜甜圈的这个楔形必须具有消失的总曲率。 全局结果随之而来的是贪婪：我们切出越来越大的甜甜圈楔子，直到我们全部吃掉！



[16.5] 当圆 $\mathcal{C}$ 绕 L 旋转角度 ϕ 时，它的（大圆）球面像 $\tilde{\mathcal{C}}$ 在 $\mathbb{S}^2$ 上以相同的速率旋转，产生两个角度为 ϕ 的弯。 正如我们所见，球面映射保留了甜甜圈外半部分的方向，并在内半部分反转了方向。 甜甜圈的楔形的总曲率是两个大小相等符号相反月牙面的面积总和，因此它消失了。 因此，甜甜圈的总曲率也消失了。

当圆 $\mathcal{C}$ 绕 L 旋转角度 ϕ 时，它的（大圆）球面像 $\tilde{\mathcal{C}}$ 在 $\mathbb{S}^2$ 上以相同的速率旋转，产生两个角度为 ϕ 的弯。 正如我们所见，球形贴图保留了甜甜圈外半部分的方向，并在内半部分反转了方向。 为了突出甜甜圈负弯曲内半部分的方向反转，我们将标签 $\mathcal{C}$ 放在那里，因此它被映射到一个向后的 $\mathcal{C}$。

甜甜圈的楔形的总曲率是两个月牙面大小相等符号相反面积的总和，因此它消失了。让

ϕ 增加到2π，我们看到甜甜圈外半部分的像正向完全覆盖了 $\mathbb{S}^2$ 一次，而内半部分的像负向覆盖了 $\mathbb{S}^2$ 一次，因此总数为 $\mathcal{K}(\text{整个甜甜圈}) = 4\pi + (-4\pi) = 0$。

看看别的东西。 圆环顶部的整个圆周（将相反曲率的外半部分和内半部分分开）映射到一个点，即 $\mathbb{S}^2$ 的北极。 同样，如果将甜甜圈放在盘子上，它会停留在最低的圆上，同样也映射到南极。 $\mathbb{S}^2$ 的这两点称为**分支点**，它们将覆盖 $\mathbb{S}^2$ 的两个独立的层连接在一起。

我们下一个任务是找到可视化有多个洞曲面总曲率的方法。

16.3 通过厚煎饼看见 $\mathcal{K}(\mathcal{S}_g)$

想象一下，将非常稠密的煎饼面糊倒入煎锅中，制成一个又大又厚的煎饼。 在它开始煮熟并凝固之前，我们迅速拿起一个圆柱形饼干刀，从煎饼内部取出 g 个圆柱形圆盘，留下 g 个洞。面糊开始渗回孔中——从而产生像甜甜圈内半部分的形状——然后开始烹饪和硬化，做出一个亏格为 g ，形状如图 [16.6] 的厚煎饼。

球面映射将整个煎饼的平坦底部发送到 $\mathbb{S}^2$ 的南极，同样将煎饼的平坦顶部发送到北极。 正如预期，这些在球面上没有产生任何面积，因此对总曲率没有贡献。

[16.6] 这里的 $\mathcal{S}_g$ 是有 g 个洞的厚煎饼。 顶部和底部是平的，因此没有曲率，但外缘的 $\mathcal{K} = +4\pi$，g 个洞的每个边缘处 $\mathcal{K} = -4\pi$。 因此总曲率是 $4\pi(1-g)$。

另一方面，煎饼的外边缘就像一个圆环的外半部分，因此它的球面图覆盖了整个球面一次，对总曲率正向贡献了 4π。 但是煎饼上 g 个洞的边缘处就像一个环面的内半部分，所以每个都负向覆盖整个球面一次，对总曲率有 -4π 的贡献。 因此（16.3）让我们看到[130]

$$\mathcal{K}(\mathcal{S}_g) = 4\pi + (-4\pi)g = 4\pi(1-g) = 2\pi\mathcal{X}(\mathcal{S}_g)$$

[130] 虽然我从未见过这个想法被写下来，但我知道（通过私下谈话）我的朋友汤姆班乔夫教授早在我之前就想到了同样的想法，一直到基于煎饼的思考！

16.4 通过百吉饼和桥看见 $\mathcal{K}(\mathcal{S}_g)$

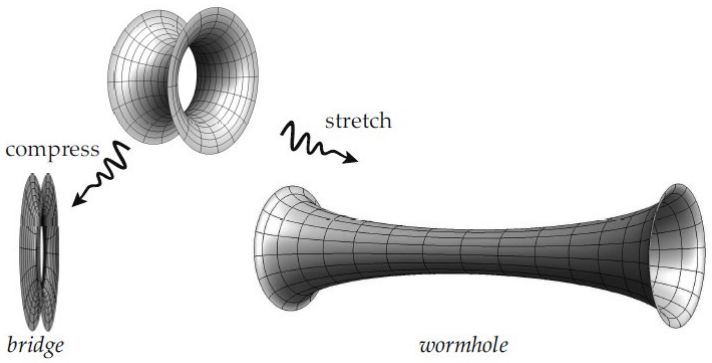
我们现在暂时回到第一人称来讲一个个人但相关的轶事。

一天，我在旧金山大学校园附近的一家百吉饼店等着上菜时，发现陈列柜里只剩下六个普通的百吉饼，仍然在烤盘里整齐地排成一排——正是我希望购买的数量！ 当我继续等待时，我带着两个念头打发时间：(1) 热切的（虽然是无神论的）“祈祷”：我前面的三个人不喜欢普通的百吉饼，以及（2）很高兴确定：即使没有我想要的 6 孔物体的几何学的任何详细知识，GGB 也绝对保证它的总曲率正好是 -20π。

我的祈祷 (1) 应验，服务员轻轻地从一排上撕下六个百吉饼，依次将它们放入一个棕色纸袋中，然后递给我。 当我离开商店时，让我震惊的是，虽然我包里的每个百吉饼看起来都没有被改变过，但它们最初的总曲率 -20π 已经蒸发到零了！ 邪恶的服务员看似不留痕迹，把 GGB 答应给我的负曲率都给抢走了！

显然，唯一的几何变化发生在百吉饼之间非常小的连接处（或者我们称之为**桥梁**），在那里它们被撕裂了，所以那里一定是戏剧性的 -20π 曲率抢劫的罪魁祸首。由于 6 个百吉饼由 5 个桥连在一起，所以每一个在被撕开时都可能增加了 4π。 在那种情况下，每个桥必然最初存储了 -4π 的曲率。

我们现在将证实这个理论，从而解释其中的奥秘，让我们以一种新的方式看见 $\mathcal{K}(\mathcal{S}_g)$（顺便说一句，免除服务员的渎职行为！）。



[16.7] 顶部的曲面是环面的内半部分，$\mathcal{K} = -4\pi$。 如果曲面被压缩以创建百吉饼**桥**（左）或拉伸以创建**虫洞**（右），则该总曲率不会改变。

考虑图 [16.7]。 顶部是一个曲面，看起来像我们厚薄煎饼上的一个洞的边缘，但现在是翻过来的； 或者，想想环面的内半部分。 我们已经讨论过，它的球面像负向覆盖了 $\mathbb{S}^2$ 一次，所以它的总曲率是 $\mathcal{K} = -4\pi$。

现在假设我们大幅度地水平压缩这个曲面以得到左侧的曲面，类似于烤盘中两个连接的百吉饼之间的桥梁。 如果我们跟随曲面经历上述压缩而跟踪球面像的演变，我们会看到随着桥梁狭窄喉部周围的区域收缩，其球面像实际上会扩大。 尽管如此，当球面像的某些部

分膨胀而其他部分收缩时，压缩后的桥面作为一个整体，与未压缩的原始球面像具有相同的球面像，因此其总曲率仍然是 $\mathcal{K} = -4\pi$，恰如预期。

为了将来使用，我们还注意到，同样，如果我们拉伸曲面以在图 [16.7] 的右侧产生**虫洞**（我们将称之为），则总曲率不会改变； 它也有 $\mathcal{K} = -4\pi$。

正如图 [16.8] 所表明的那样，每个桥具有 $\mathcal{K} = -4\pi$ 的结果是

$$\mathcal{K}(\mathcal{S}_g) = g \cdot \mathcal{K}(\text{百吉饼}) + (g-1) \cdot \mathcal{K}(\text{桥}) = g \cdot 0 + (g-1) \cdot (-4\pi) = 4\pi(1-g)$$

符合 GGB。

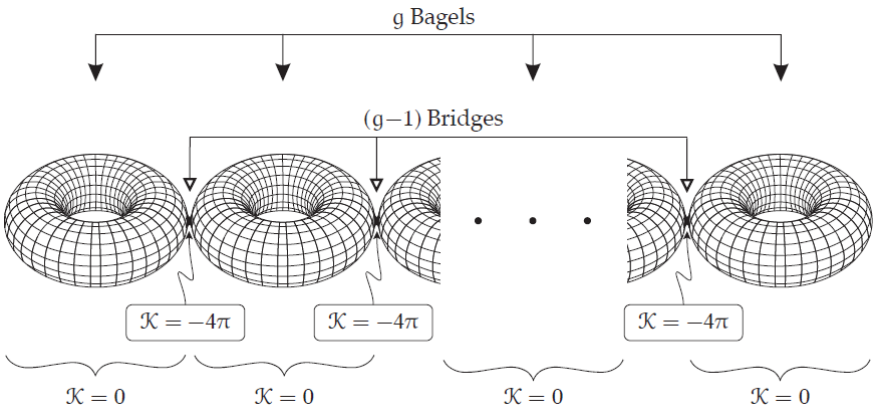
谜题：将三个百吉饼放在一个大三角形的顶点上，然后用三个虫洞沿着三角形的边缘将它们连接在一起，形成一个单一的封闭曲面 $\mathcal{S}$，其中 $\mathcal{K}(S) = 3\mathcal{K}_{\text{虫洞}} = -12\pi$。 这不违反GGB吗？！ （更多例子可以在例 22 中找到。）

16.5 球面映射的拓扑度

到目前为止，我们希望我们的例子已经清楚地表明，如果我们能理解这一点，我们就会理解 GGB：

> 无论曲面 $\mathcal{S}_g$ 采用何种形式，$n(\mathcal{S}_g)$ 将始终覆盖球面的几乎每一点 $(1-g)$ 次，前提是我们以代数方式计算层数，同时考虑正向和负向。

(16.5)



[16.8] 这里 $\mathcal{S}_g$ 采用 g 个百吉饼的形式，仍然通过 $(g-1)$ 桥在烤盘中连接在一起。 每个百吉饼的总曲率为零，而每个桥的 $K = -4\pi$。 因此总曲率是 $(-4\pi)(g-1) = 2\pi\mathcal{X}$。

为了澄清这个说法，本节将引入**拓扑度**[131]（topological degree）的概念，它对 $\mathbb{S}^2$ 的每个点被覆盖的次数进行精确的代数计数。我们说“几乎每一点”的原因是为了允许分支点，正如在环面的情况下所提到的。

我们首先注意到（16.5）并不意味着只有 $|1-g|$ 总共覆盖 $\mathbb{S}^2$ 的床单。实际上，在 [16.6] 中的厚煎饼的情况下，$\mathbb{S}^2$ 的覆盖 $n(\mathcal{S}_g)$ 采取了一个正向覆盖和 g 个负向覆盖的形式：总共 $(g+1)$ 个覆盖。换句话说，对于 $\mathbb{S}^2$ 上的任何给定点 $\tilde{p}=n(p)$，在煎饼上有 $(g+1)$ 个位置 p_i，其中法线指向 $\tilde{p}$ 的方向，即 $n(p_i)=\tilde{p}$，对于 $i=1,2,\ldots,(g+1)$。

令 $\mathcal{P}(\tilde{p})$ 表示曲面 $\mathcal{S}_g$ 上曲率 $\mathcal{K}(p_i)$ 为正的点 p_i 的数量，因此 n 是方向保持的，并且球面上的图像 $\tilde{p}$ 的邻域被正覆盖。同样，让 $N(s)$ 表示 $\mathcal{K}(p_i)$ 为负的点数，因此 n 是方向反转，并且 $\tilde{p}$ 被负覆盖。例如，在上面的厚煎饼[16.6]的情况下，$P(s)=+1$，与 $\tilde{p}$无关；并且 $N(s)=g$，也与 $\tilde{p}$ 无关。（在这里以及接下来的内容中，我们排除了 $K=0$ 的点，因为这些点对覆盖 $\mathbb{S}^2$ 没有贡献。）

我们现在可以如下定义球面映射的**拓扑度**（或者，更常见的是，只是“度”）：

> 给定一个封闭的、有向的亏格为 g 的曲面 $\mathcal{S}_g$ 和 $\mathbb{S}^2$ 上的一个点 $\tilde{p}$，球面映射的**拓扑度**——写成 $\deg[n(\mathcal{S}_g),\tilde{p}]$——是 $n(\mathcal{S}_g)$ 覆盖 $\tilde{p}$ 次数的代数计数，考虑到方向：
>
> $$\deg[n(\mathcal{S}_g),\tilde{p}]\equiv\mathcal{P}(\tilde{p})-\mathcal{N}(\tilde{p})$$

(16.6)

在 $\mathcal{S}_g$ 是厚煎饼的情况下，$\mathcal{P}$ 和 $\mathcal{N}$ 与 $\tilde{p}$ 无关，因此度数也是如此：

$$\deg[n(g\text{ 孔厚煎饼})]=\mathcal{P}-\mathcal{N}=1-g$$

在 $\mathcal{S}_g$ 由桥接百吉饼组成的情况下（图 [16.8]），$\mathbb{S}^2$ 被 g 个百吉饼的像覆盖了 $2g$ 次，并且桥的像覆盖了 $(g-1)$ 次。因此，$\mathbb{S}^2$ 的每个点都被 $(3g-1)$ 层覆盖。同样在这个例子中，覆盖的数量与 $\tilde{p}$ 无关，因此这些覆盖的代数计数也是如此：

$$\deg\big[n\big(g\text{ 个桥接的百吉饼}\big)\big]=\mathcal{P}-\mathcal{N}=g-[g+(g-1)]=1-g$$

就像以前一样。

理解 GGB 的关键（至少从目前的角度来看）是要能够看到这种反复出现的结果并非巧合，而是该度在本质上确实是拓扑的——每个 $\mathcal{S}_g$ 必须满足相同的等式：

---

[131] 也称为布劳威尔（Brouwer）度，以荷兰拓扑先驱布劳威尔（1881-1966）命名，他是第一个系统地使用这一概念的人。

$$\boxed{\deg[\mathrm{n}(\mathcal{S}_{\mathrm{g}})] = \mathcal{P} - \mathcal{N} = (1-\mathrm{g}) = \frac{1}{2}\mathcal{X}(\mathcal{S}_{\mathrm{g}})} \tag{16.7}$$

简而言之，如果我们能证明（16.7），那么我们就证明了 GGB（16.1）。

16.6 历史笔记

虽然高斯（Gauss）和博内（Bonnet）确实为 GGB 铺平了道路，但他们都没有意识到这个非凡的结果，更不用说宣布了！

但名字已经卡住了。 即使那些熟知这个名字不准确历史的人现在也不敢碰它：似乎我们都同意一个名字的含义比这个名字本身在历史上准确更重要。

正如我们所讨论的，高斯在 1827 年宣布他发现了局部的高斯-博内定理（2.6），指出如果 Δ 是一般曲面上的测地三角形，则其角盈等于其总曲率：

$$\mathcal{E}(\Delta) = \mathcal{K}(\Delta)$$

事实上，**局部高斯–博内定理**是指高斯的原始结果（博内在 1848 年）对三角形的边不再需要是测地线的情况的推广。这在上述等式的右侧添加了一个表示边的总测地曲率的项。[132]然而，这些先生们都没有提及封闭曲面。

看起来，发现 GGB 的荣誉实际上分两个不同的步骤属于利奥波德 • 克罗内克（Leopold Kronecker）和沃尔特 • 戴克[133]（Walther Dyck）。首先，在 1869 年，克罗内克引入了度的概念——后来被布劳威尔（Brouwer ）澄清和利用——并证明了 $\mathcal{K}(\mathcal{S}_{\mathrm{g}}) = 4\pi\deg(\mathrm{n})$。 其次，1888 年戴克证明了 $\deg(\mathrm{n}) = \frac{1}{2}\mathcal{X}$，从而完成了 GGB 现代形式（16.1）的证明。

---

[132] 这将在第四幕结束时得到证明：练习 6，第 336 页。

[133] 参见赫希 (1976)。 有时，如在 Berger（2010 年，第 380 页）中，GGB 被归功于 Werner Boy。然而，Boy (1903) 明确地将 GGB 归功于 Kronecker 和 Dyck，而 Boy 本人将 GGB 推广到不可定向的曲面。

第十七章 全局高斯-博内定理的第一个（启发式）证明

17.1 平面环路的总曲率：霍普夫环量定理

为有助于理解球面映射（即（16.7））度数的拓扑性质，以及 GGB（即（16.1）），我们将降低一个维度。 也就是说，取代 $\mathbb{R}^3$ 中二维曲面 $\mathcal{S}$ 到二维球面 $\mathbb{S}^2$ 的球面/法线贴图 n，我们将考虑 $\mathbb{R}^2$ 中一维曲线 $\mathcal{C}$ 到一维圆周 $\mathbb{S}^1$ 的法向映射 N。

由于 GGB 涉及闭合曲面，因此相应地我们将注意力限制在 $\mathcal{C}$ 是闭合曲线的情况。 此外，让我们从最基本的情况开始，其中 $\mathcal{C}$ 不仅是封闭的，而且是**简单的**——即，没有任何自交。 让我们称这种简单的闭合曲线为**环路**（loop）。

尽管我们最终将关注处处具有明确定义的切线和法线的平滑曲线，但让我们从三角形 Δ 开始，其内角为 θ_i，外角为 φ_i，因此 $\theta_i + \varphi_i = \pi$ （见[17.1a]）。

如果我们想象一个粒子逆时针绕 Δ 运动一次，那么很明显它的速度向量 **v** 环绕了一整圈，这确实等价于 Δ 内角和为 π 的基本事实：

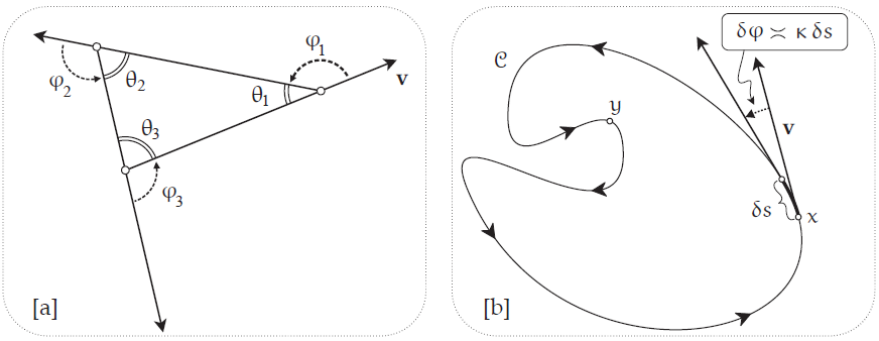
$$\text{净旋转} = \varphi_1 + \varphi_2 + \varphi_3 = 2\pi \iff \theta_1 + \theta_2 + \theta_3 = \pi$$

注意，净旋转表述在某种程度上是两者中更简单和更基本的，因为如果我们推广到 n 边形，它保持不变，而内角方面的等价表述现在取决于 n：

$$\text{净旋转} = \sum_{i=1}^{n} \varphi_i = 2\pi \iff \sum_{i=1}^{n} \theta_i = (n-2)\pi$$

当并非所有旋转都在同一方向时，净旋转的概念变得很重要。 螺母沿螺栓的运动生动地说明了这一点。 假设你观察螺母的初始位置，然后闭上眼睛，同时你的朋友以复杂的正向和反向旋转组合旋转螺母，沿着螺栓来回移动螺母。 当你睁开眼睛时，你不知道你的朋友到底做了什么旋转，但你知道净旋转是什么：它可以仅通过螺母从其起始位置移动了多远来衡量。

图 [17.1b] 考虑了速度 **v** 的这种净旋转，但现在是光滑环路 $\mathcal{C}$，而不是多边形。 而 **v** 的方向在 Δ 的顶点处执行了 φ_i 的突然跳跃，它的方向现在沿着 $\mathcal{C}$ 平滑地变化，对于图示的曲线，它有时正向（逆时针）转，如在 x 处，有时向负转（顺时针），如在 y 处。 尽管如此，直观上似乎很清楚，当粒子沿着 $\mathcal{C}$ 行进时，它的速度 **v** 沿途来回摇摆，但在 $\mathcal{C}$ 的完整正轨道之后，净效果是 **v** 执行了完整的正转。 同样，在谈到 **v** 的这种净旋转时，我们的意思是允许负向旋转抵消正向旋转。



[17.1] 霍普夫的环量定理：当一个粒子跟踪一个简单的环路时，它的速度执行一个正向旋转。 在[a]中，$\varphi_1 + \varphi_2 + \varphi_3 = 2\pi$； 在 [b] 中，$\oint_{\mathcal{C}} d\varphi = \mathbf{v}$ 的净旋转 $= 2\pi$。

[17.2] 海因茨·霍普夫 (1894 – 1971)。 照片由恩斯特·阿曼 (Ernst Ammann) 拍摄，CC BY-SA 4.0

环绕一个环路轨道运动一周会导致速度向量 $\mathbf{v}$ 旋转一整圈这一事实称为霍普夫环量定理（Hopf Umlaufsatz，来自德语，“Umlauf”（环量）和“Satz”（定理））。 虽然你可能不会怀疑其对图 [17.1b] 中简单曲线的真实性，但对图 [17.3] 中的“简单”曲线真的如此明显吗？！此外，结果不适用于与自身相交的曲线。 例如，8 字形曲线的净旋转是多少[练习]？

虽然这个结果在某种意义上自古以来就为人所知，但沃森（Watson 1917）似乎是第一个清楚地阐明它的人，而霍普夫（1935）是第一个给出纯拓扑证明的人。 练习 23 中描述了霍普夫巧妙的几何思想。 然而，在这里，我们希望给出一个不同的启发式证明，作为 GGB 相应证明的精确模型。

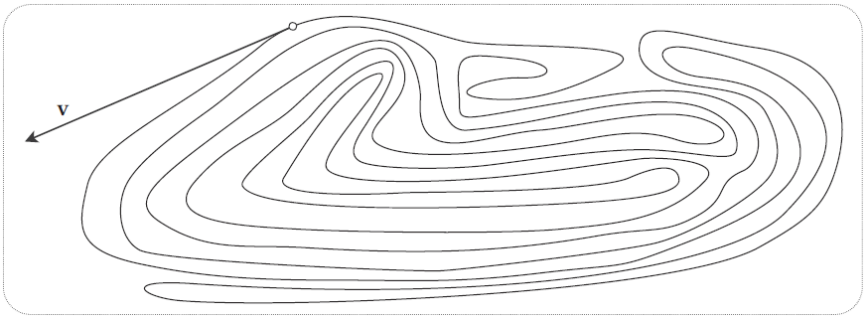
首先与之前对 GGB 的讨论建立明确联系。首先，回想曲率 κ 只是 $\mathbf{v}$ 在单位弧长 s（或时间，如果粒子以单位速度行进）的旋转速率。 因此，环量定理可以重新表述为环路总曲率的拓扑不变性：

$$\boxed{\oint_{\mathcal{C}} \kappa\, ds = \oint_{\mathcal{C}} d\varphi = \mathbf{v} \text{ 的净旋转} = 2\pi} \tag{17.1}$$

与 $\mathcal{C}$ 的形状无关，这一点与 GGB 惊人的相似！

为了将其与球面/法线映射相关联，令 $\mathbf{N}$ 为 $\mathcal{C}$ 的单位法线，并且，如 [12.1]，将其视为从平面曲线 $\mathcal{C}$ 的点 p 到单位圆 $\mathbb{S}^1$ 上位于 $\mathbf{N}_\mathrm{p}$ 尖端的点 $\tilde{\mathrm{p}} = \mathrm{N}(\mathrm{p})$ 的映射 N。

虽然霍普夫用切线旋转来说明他的成果，但我们同样可以说是 $\mathbf{N}$ 在绕 $\mathcal{C}$ 轨道旋转一次后进行了一次净旋转。 等价地，我们可以说，在代数上，$\mathrm{N}(\mathcal{C})$ 覆盖 $\mathbb{S}^1$ 一次。



[17.3] ($\mathbf{v}$ 的净旋转) $= 2\pi$ 真的那么明显吗？

在 $\mathcal{C}$ 的小段 δs 上，法向量的方向分布在角度 $\delta\varphi$ 上，因此这些法向量的尖端填充了 $\mathbb{S}^1$ 上长度为 $\delta\tilde{s} = \delta\varphi$ 的弧。 因此，正如我们在第 131 页首次讨论的那样，

$$\kappa = \mathrm{N}\text{ 映射的局部长度放大系数} \asymp \frac{\delta\tilde{s}}{\delta s}$$

因此，

$$\boxed{\oint_{\mathcal{C}} \kappa\, ds = 2\pi[\mathrm{N}(\mathcal{C})\text{覆盖}\mathbb{S}^1\text{的次数}]} \tag{17.2}$$

与我们讨论 GGB 一样，我们接下来介绍 N 的度数，以阐明如何代数计算 $\mathrm{N}(\mathcal{C})$ 覆盖 $\mathbb{S}^1$ 的次数。

即使 $\mathcal{C}$ 不是一个简单的环路，而是允许自相交，我们也可以像以前一样定义球面/法线映射 N 的度数。 如果 $\kappa(\mathrm{p}) > 0$ 那么当粒子在 C 上穿过 p 时，$\mathrm{N}(\mathrm{p})$ 绕 $\mathbb{S}^1$ 正向旋转（逆时针）。同样，如果 $\kappa(\mathrm{p}) < 0$ 那么 $\mathrm{N}(\mathrm{p})$ 将绕 $\mathbb{S}^1$ 反向旋转（顺时针）。 于是，

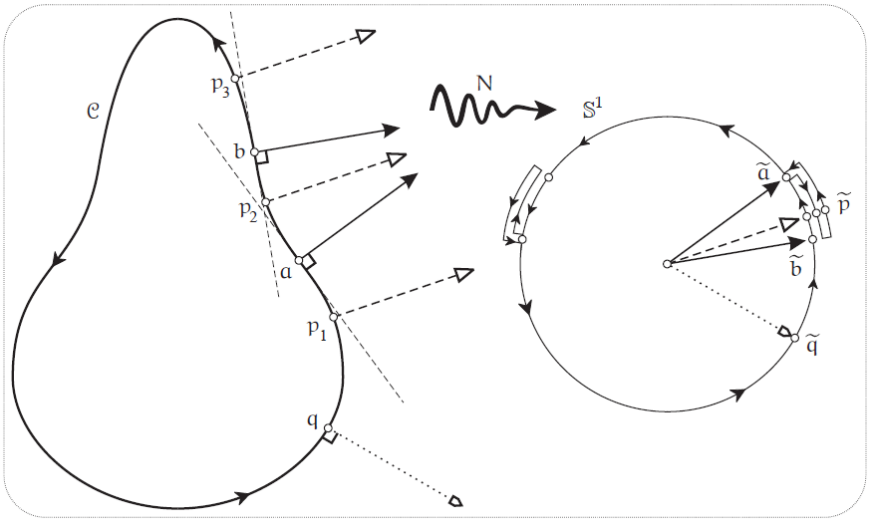
给定逆时针追踪的闭合曲线 $\mathcal{C}$ 和 $\mathbb{S}^1$ 上的点 $\tilde{\mathrm{p}}$，球面/法线映射 N 的度数是 $\mathrm{N}(\mathcal{C})$ 覆盖 $\tilde{\mathrm{p}}$ 的次数的代数计数，考虑到方向：

$$\deg[\mathrm{N}(\mathcal{C}), \tilde{\mathrm{p}}] \equiv \mathcal{P}(\tilde{\mathrm{p}}) - \mathcal{N}(\tilde{\mathrm{p}}) \tag{17.3}$$

其中 $\mathcal{P}(\tilde{\mathrm{p}})$ 是 $\kappa > 0$ 的 $\tilde{\mathrm{p}}$ 原像数，$\mathcal{N}(\tilde{\mathrm{p}})$ 是 $\kappa < 0$ 的 $\tilde{\mathrm{p}}$ 原像数。

有了这个更精确的定义，关键事实是度数与 $\tilde{p}$ 的选择无关，因此（17.2）采用以下形式，

$$\boxed{\oint_{\mathcal{C}} \kappa \, ds = 2\pi \deg\left[N(\mathcal{C})\right]} \tag{17.4}$$



[17.4] N 的**度数**是 $N(\mathcal{C})$ 覆盖 $\mathbb{S}^1$ 的次数的代数计数。κ 的符号在拐点 a 和 b 处发生变化，导致 $\mathbb{S}^1$ 上的图像在 $\tilde{a}$ 和 $\tilde{b}$ 处反转方向。这反过来可以被认为是轨道向自身折叠。 由于 $\tilde{p}$ 被遍历了 3 次，两次是正向运动，一次是负向运动，因此 $\tilde{p}$ 的净覆盖数为 $2-1=1$。

霍普夫环量定理指出 $\deg[N(\text{简单环路})]=+1$，在这种情况下，（17.4）简化为（17.1）。

17.2 变形圆的总曲率

让我们用一个具体的例子来阐明这些想法。 如果 $\mathcal{C}$ 是一个圆，那么当 x 绕 $\mathcal{C}$ 运行一次时，它的像 $\tilde{x}=N(x)$ 以匹配的角速度绕 $\mathbb{S}^1$ 运行一次。 显然，$\deg[N(\mathcal{C})]=+1$。现在假设我们逐渐对称地使圆变形，使得 $\mathcal{C}$ 具有 [17.4] 左侧所示的形状，类似于梨的横截面。

对于 $\mathbb{S}^1$ 上的图示点 $\tilde{q}$，$\mathcal{C}$ 上正好有一个原像 q。但对于点 $\tilde{p}$，则有三个原像 p_1、p_2、p_3。

定位 $\tilde{p}$ 的这些原像的一个有用的心理工具是首先观察 $\tilde{p}$ 处的 $\mathbb{S}^1$ 的切线（未显示）必须平行于 $\tilde{p}$ 的每个原像处的 $\mathcal{C}$ 的切线。 现在想象在 $\tilde{p}$ 处取这条切线并让它平行于自身向 $\mathcal{C}$ 移动，最终扫过所有 $\mathcal{C}$。注意每次移动线接触 $\mathcal{C}$：这些包括 $\tilde{p}$ 的所有原像，但它们也包括 $\tilde{p}$ 的原像对映点 $-\tilde{p}$。[17.4] 中 $-\tilde{p}$ 的原像在哪里[练习]？

注意 $\mathcal{C}$ 的右手边，我们观察到正好有两个拐点 a 和 b，区别在于 $\mathcal{C}$ 在这些点处（且仅在这些点处）与切线相交。

要看到拐点在球面/法线映射 N 中起着至关重要的作用，请想象 x 从 q 开始并沿 $\mathcal{C}$ 的右侧向上移动。然后 $\tilde{x} = N(x)$ 从 $\tilde{q}$ 到 $\tilde{b}$ 和 $\tilde{p}$ 沿 $\mathbb{S}^1$ 向前移动直到它碰到 $\tilde{a}$。但是在 $\tilde{a}$ 它会反弹并向后移动，在碰到 $\tilde{b}$ 之前第二次通过 $\tilde{p}$。现在它再次反弹并继续沿 $\mathbb{S}^1$ 向前运动，第三次通过 $\tilde{p}$。接下来它第二次到达 $\tilde{a}$ [练习：此时 x 在哪里？] 而这次它直接穿过 $\tilde{a}$ 并继续前进。我们强烈建议你在 x 完成 $\mathcal{C}$ 的完整轨道时在你的脑海中跟随 x，当 x 穿过 $\mathcal{C}$ 的左侧时，$\tilde{x}$ 执行另一个来回运动。

现在在这个运动的可视化中出现了一个关键的精神飞跃：把 $\tilde{x}$ 的运动想象成一个珠子沿着一条连续的、不间断的线运行的轨道。因此，如图所示，当 $\tilde{x}$ 第一次到达 $\tilde{a}$ 并开始在 $\mathbb{S}^1$ 上向后移动时，这只能是因为线在自身上折回。同样，当珠子下一次到达 $\tilde{b}$ 并再次反向时，恢复其沿 $\mathbb{S}^1$ 的向前运动，这只能是因为线已第二次向后折叠。因此，如图所示，随着 x 通过 p_1，然后是 p_2，最后是 p_3，$\tilde{x}$ 通过 $\tilde{p}$ 三次，首先是向前，然后是沿着折叠线向后，然后再次沿着两次折叠线向前。

实际上，$\tilde{a}$ 和 $\tilde{b}$ 之间的所有三个折叠的线段都在 $\mathbb{S}^1$ 上彼此叠在一起，但是我们将它们稍微抬起以显示褶皱和线上对应于 $\mathbb{S}^1$ 的三个不同位置 $\mathbb{S}^1$ 的单点 $\tilde{p}$。

q 处的正曲率 κ 通过 $\tilde{q}$ 引起正运动。因此，由于 q 是 $\tilde{q}$ 的唯一原像，$P(\tilde{q}) = 1$ 且 $N(\tilde{q}) = 0$，因此 $\deg[N(\mathcal{C}), \tilde{q}] \equiv P(\tilde{q}) - N(\tilde{q}) = 1$，如预期。

同样，p_1 和 p_3 处的正曲率 κ 通过 $\tilde{p}$ 引起正运动，而 p_2 处的负值 κ 通过 $\tilde{p}$ 引起负运动。因此，$\mathcal{P}(\tilde{p}) = 2$ 且 $\mathcal{N}(\tilde{p}) = 1$，因此 $\deg[N(\mathcal{C}), \tilde{p}] \equiv \mathcal{P}(\tilde{p}) - \mathcal{N}(\tilde{p}) = 1$，与之前一样。

17.3 霍普夫环量定理的启发式证明

现在让我们将我们从这个例子中学到的知识应用到一般情况下。想象一下从一个圆开始，让它不断变形，演变成最一般的简单闭合曲线。在这里，我们将曲线的形式视为时间的函数 $\mathcal{C}(t)$，例如，时间 t 从 0 到 1，$\mathcal{C}(0)$ 是初始圆，$\mathcal{C}(1)$ 是最终圆曲线（参见图 [17.4] 中的 $\mathcal{C}$）。

但是我们限制这种演变以排除自相交，并且使得 κ 在中间曲线上的任何地方都是明确定义的：如果我们希望我们的映射 N 在每个 $\mathcal{C}(t)$ 上是连续的，这是必要的。例如，如果出现尖角，则 N 将在那里出现跳跃不连续性。

随着 $\mathcal{C}(t)$ 的不断演变，$\tilde{x}$ 的轨道也在不断演变，具体化和可视化为一条可拉伸、牢不可破、可折叠的绳子，缠绕在 $\mathbb{S}^1$ 周围。每当 $\mathcal{C}(t)$ 中出现拐点，表示 κ 的符号发生变化时，图像串就会形成折叠。

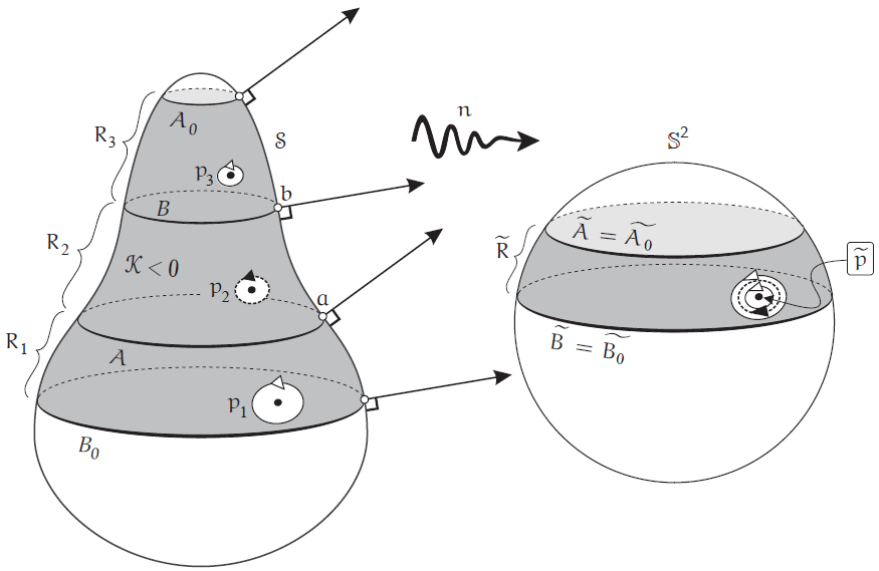
如果 $\tilde{x}$ 远离这种折叠，则 $\mathcal{P}$ 和 $\mathcal{N}$ 分别保持不变，因此它们的差异 $\deg[N[\mathcal{C}(t)], \tilde{x}]$ 也是如此。然而，如果 $\tilde{x}$ 穿过一个折叠点，我们要么获得两层相反方向的新层，要么失去两层相反方向。但是，这意味着 $\mathcal{P}$ 和 $\mathcal{N}$ 都增加 1 或它们都减少 1。在任何一种情况下，

关键的观察是，

> $\mathcal{C}$ 上 κ 符号的变化导致 $N(\mathcal{C})$ 在 $\mathbb{S}^1$ 上折叠，但这种折叠对覆盖 $\mathbb{S}^1$ 的层数（按代数计算）没有影响。

也就是说，$\deg[N\{\mathcal{C}(t)\}] = \mathcal{P}(\tilde{x}) - \mathcal{N}(\tilde{x})$ 在我们穿过折叠时保持不变。因此 $\deg[N\{\mathcal{C}(t)\}]$ 是曲线整体的一个明确定义的属性，与 $\tilde{x}$ 无关。

$\deg[N\{\mathcal{C}(t)\}]$ 的值必然随时间不断变化，并且保持整数：因此它根本不会改变。但是最初 $\deg[N\{\mathcal{C}(0)\}] = 1$。因此，由于度数的最终值与其初始值相同，因此 $\deg[N\{\mathcal{C}(1)\}] = 1$，如霍普夫所断言。



[17.5] n 的次数是 n(S) 覆盖 $\mathbb{S}^2$ 的次数的代数计数。 当我们穿过 A 和 B 时，$\mathcal{K}$ 的符号发生变化，导致 $\mathbb{S}^2$ 上的图像在我们穿过 $\tilde{A}$ 和 $\tilde{B}$ 时反转方向。 由于 $\tilde{p}$ 被覆盖了 3 次，正方向覆盖了两次，负方向覆盖了一次，因此 $\tilde{p}$ 的净覆盖数为 $2 - 1 = 1$。

17.4 变形球的总曲率

在图 [17.4] 中，我们把 $\mathcal{C}$ 比作梨的横截面。现在我们将注意力转向梨本身的曲面！

也就是说，让我们取 $\mathcal{C}$ 的右半部分，并围绕通过其两端的垂直轴将其旋转 2π，以生成 [17.5] 中所示的梨形旋转曲面 $\mathcal{S}$。正如视觉上明显的（以及之前在（10.13）中提到的），$\mathcal{C}$ 上 $\kappa > 0$ 的曲线段生成 $\mathcal{S}$ 的椭圆型区域，$\mathcal{K} > 0$；$\mathcal{C}$ 上 $\kappa < 0$ 的曲线段生成 $\mathcal{S}$ 的双曲型区域，$\mathcal{K} < 0$。拐点 a 和 b 旋转以生成两个圆 A 和 B（均为抛物型；即 $\mathcal{K} = 0$），将梨曲面的正弯曲部分和负弯曲部分分开。

但该图也展示了另外两个圆 A_0 和 B_0，它们在球面映射下具有与 A 和 B 相同的图像：$\tilde{A} = \widetilde{A_0}$ 和 $\tilde{B} = \widetilde{B_0}$。这四个圆圈将 $\mathcal{S}$ 分为顶部区域、底部区域和中间的三个区域，我们

用灰色阴影标记并标记为 R_1、R_2、R_3。 所有这三个区域 R_i 都映射到 $\mathbb{S}^2$ 上的相同区域 $\tilde{R}$。

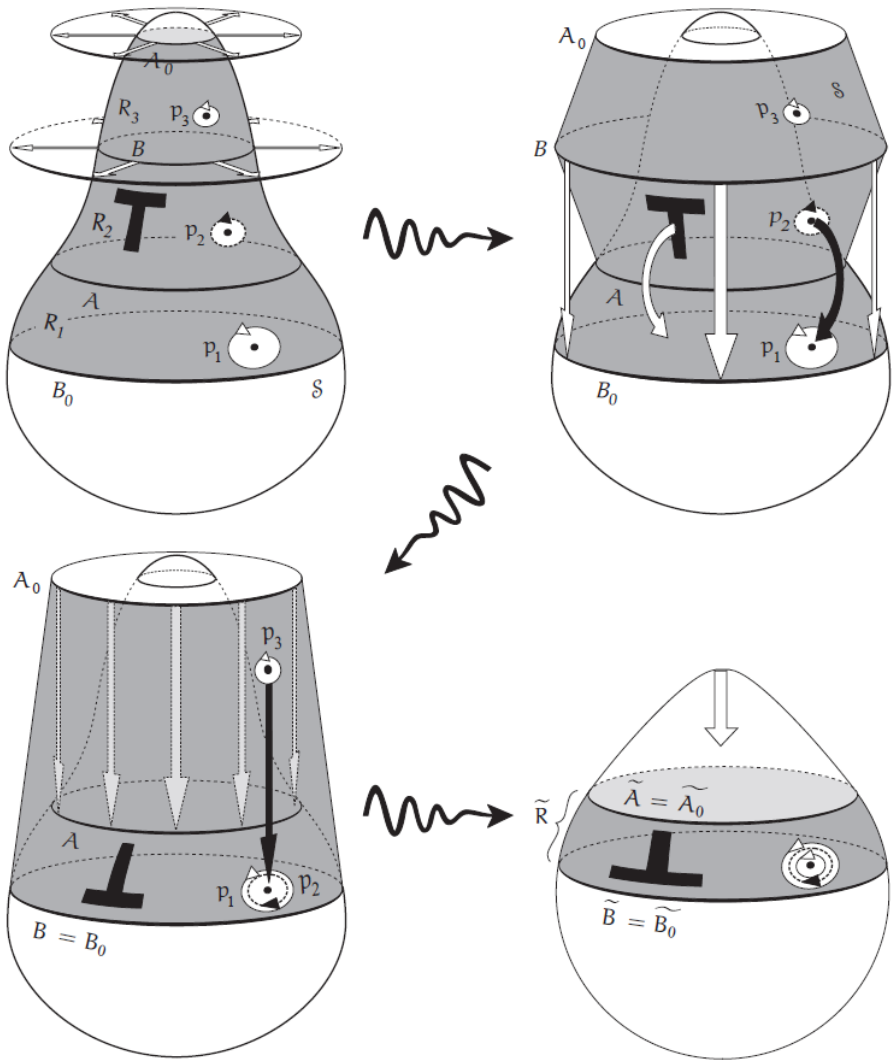
由于 R_1 和 R_3 上的 $\mathcal{K} > 0$，球面映射在这些区域中是保持方向的，因此这些区域中的一个点被映射到 $\mathbb{S}^2$ 上一个被正向覆盖的点。 另一方面，由于 R_2 上的 $\mathcal{K} < 0$，球面映射在该区域是方向反转的，因此该区域中的一个点被映射到 $\mathbb{S}^2$ 上被负向覆盖的点。

特别地，该图说明了 $\tilde{R}$ 内点 $\tilde{p}$ 的三个原像 p_1、p_2、p_3。 因此，由于 $\tilde{p}$ 被覆盖 3 次，正方向覆盖两次，负方向覆盖一次，$\tilde{p}$ 的净覆盖数为 $2 - 1 = 1$。

让我们不要忘记 GGB 在曲率方面的原始公式，以及我们对为什么闭合曲面的总曲率是拓扑不变之理解的求索。 在我们变形的球体 $\mathcal{S}$ 的情况下，位于梨形无阴影顶部的总曲率是 $\mathbb{S}^2$ 北极冠的面积； 同样，$\mathcal{S}$ 的无阴影底部的总曲率同样是较大的无阴影南部区域的面积。 三个阴影区域中的每一个都包含相同的曲率量，即 $\tilde{R}$ 的面积。 但由于 R_2 上的 $\mathcal{K} < 0$，n 在那里方向相反，因此它的总曲率是 $\tilde{R}$ 面积的负数。 因此，对所有 $\mathcal{S}$ 的曲率进行积分会产生 $\mathbb{S}^2$ 的整个面积，一次：即 4π。

17.5 全局高斯-博内定理的启发式证明

在低维情况下，通过将曲线的球面像视为覆盖 $\mathbb{S}^1$ 的可拉伸、可折叠线，我们获得了对度数直观理解。 在目前的情况下，将 $n(\mathcal{S})$ 想象为覆盖 $\mathbb{S}^2$ 的高度可拉伸、牢不可破、可折叠的膜同样非常有帮助——也许想想非常薄的披萨面皮。 但是要明白，这个特殊的数学面皮不仅可以按我们需要的任何方式拉伸，而且还可以根据需要轻松收缩。



[17.6] $\mathcal{K}$ 符号变化的拓扑结果是球面像的折叠。当我们穿过相反曲率区域之间的边界时，球面像要么增加要么失去两层相反方向，所以$\mathbb{S}^2$的覆盖次数（即度）的代数计数保持不变，因此总曲率是拓扑不变量。

在图[17.5]的左边，把梨 $\mathcal{S}$ 的整个曲面想象成覆盖着一层薄薄的披萨面皮，松散地粘在曲面上。 现在让我们不要担心球面映射的几何，而是只关注它如何将 $\mathcal{S}$ 上的灰色区域 $R_1 \cup R_2 \cup R_3$ 映射到 $\mathbb{S}^2$ 上的灰色区域 $\widetilde{R}$ 的拓扑（见图[17.6]）。

该图中的箭头表示要执行的拓扑变换的顺序。 首先，假设我们在 A 和 B_0 周围系两根绳子固定，然后径向向外拉伸 B 直到与 B_0 大小相同，同样向外径向拉伸 A_0 直到与 A 大小相同，拉伸 R_2 和 R_3 进入锥体的两个图示部分。

接下来，将新的 A_0 留在原处，让我们垂直向下移动新的 B 并将其粘贴到 B_0 上，在此过程中将 R_2 折叠到 R_1 的顶部。 请注意，至关重要的是，这种折叠如何导致 p_2 周围的正向环绕被逆转。 R_2 中大写字母“T”的命运强调了这一点。

现在垂直向下移动新的 A_0 并将其粘贴到 A 上，在此过程中将 R_3 折叠到新的 R_2（和 R_1）的顶部。

至此我们已经正确地体现了球面映射的拓扑结构，我们可以想象做一个最终的几何看家，将梨的顶部向下推使其成为球形（和单位半径），并拉伸三个折叠的灰色中的每一个层以获得正确的球形贴图的几何形状，因此特别是 $\widetilde{p_1}$、$\widetilde{p_2}$ 和 $\widetilde{p_3}$ 最终都将彼此重叠，但它们的方向已通过折叠在连续层之间反转。

我们如何到达这个最终状态的细节无关紧要； 选择我们特定的中间转换只是为了使这个网络转换更容易想象和绘制。此外，这个例子涉及了一个小技巧：如果 $g \neq 0$ 那么 $n(\mathcal{S}_g)$ 必须有多层，我们无法想象从覆盖 $\mathcal{S}_g$ 的单层面团开始并操纵它来获得 $n(S)$。

在这一点上，我们可以尝试直接（没有中间步骤）想象球面图对演化曲面 $\mathcal{S}$ 的影响。如果我们从球面 $\mathcal{S}_0$ 开始，逐渐将其变形为图 [17.5] 中的梨形曲面，那么我们可以想象数学面团 $n(\mathcal{S}_0)$ 在相应的进化后覆盖 $\mathbb{S}^2$。 由于在抛物线的变形 $\mathcal{S}_0$ 上出现负曲率，因此球面图像在这些抛物线的图像处向后流动，在那里产生折叠。 但是当我们在 $\mathbb{S}^2$ 上的 $n(\mathcal{S}_0)$ 中越过这样的折叠时，我们总是会获得或失去两个相反方向的新层，因此覆盖数量的代数和没有改变。

如果我们改为从某个 $\mathcal{S}_g$（其中 $g \neq 0$）开始并允许它连续变化，则同样的推理也适用。例如，假设最初 $\mathcal{S}_g$ 采用图 [16.8] 的 g 桥接百吉饼的形式。 现在假设我们用拇指向内按压其中一个百吉饼的外（正弯曲）曲面，形成一个以点 p 为中心的凹陷。当我们这样做时，$n(\mathcal{S}_g)$ 将在 $\tilde{p}$ 附近经历相应的演变，位于积极覆盖 $\mathbb{S}^2$ 的 g 片之一上。 在总共 $(3g-1)$ 张纸中，那张纸将经历上一段中描述的演变。 这个包含 $\tilde{p}$ 的单片将自身折叠，在以 $\tilde{p}$ 为中心的拓扑环中创建两个相反方向的新层，但是（在代数上）这个折叠区域将继续覆盖 $\mathbb{S}^2$ 一次，因此 $\mathbb{S}^2$ 的净覆盖数将保持为 $(1-g)$，即使在 $\tilde{p}$ 附近。

我们对 GGB 的第一个（启发式）证明到此结束。

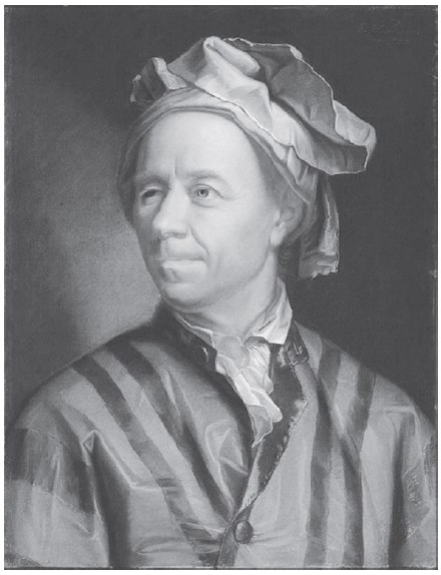
第十八章 全局高斯-博内定理的第二个（角盈）证明

18.1 欧拉示性数

到目前为止，我们已经介绍了欧拉示性数 $\mathcal{X}$，但仅仅作为标记给定亏格为 g 、封闭的、可定向曲面 $\mathcal{S}_{\mathrm{g}}$ 的一种方便的替代方法。 然而，$\mathcal{X}$ 实际上有自己的定义和含义，它适用于比封闭的、可定向曲面更广泛的一类对象。特别是，一旦我们解释了 $\mathcal{X}$ 的这个更深层次的含义，我们就可以将它应用到 $\mathcal{S}_{\mathrm{g}}$，于是 $\mathcal{X}(\mathcal{S}_{\mathrm{g}}) = 2 - 2\mathrm{g}$ 就成为一个重要的定理（不是定义）。

18.2 欧拉（经验）多面体公式

欧拉示性数的故事始于 1750 年 11 月 14 日，不是光滑的曲面，而是多面体。 那天，欧拉——如图[18.1]——写了一封信给克里斯蒂安• 哥德巴赫[1]，概述了一个非凡的经验发现，欧拉在两年后终于成功地证明了这一点。[2]



[18.1] 莱昂哈德•欧拉 (1707 – 1783)

这一发现建立在一个对现代人来说似乎非常明显的认识之上：欧拉是第一个清楚地认识

[1] 哥德巴赫最著名的是他在 1742 年写给欧拉的一封信。在信中，他推测每个大于 2 的偶数都是两个素数之和； 这现在被称为哥德巴赫猜想。 虽然它被认为是真实的，并且（截至 2017 年）已被计算机验证多达 4, 000, 000, 000, 000, 000, 000，但在近 300 年后仍未得到证明。

[2] 有关这段历史和相关数学思想的精湛、数学上准确又引人入胜的描述，请参阅 Richeson (2008)。另请参阅本书末尾的延伸阅读。

到多面体**顶点**、**边**和**面**的存在性的人。 正如他在信中所写，

因此，在任何类型的物体中都需要考虑三种边界，即 1）点，2）线和 3）面，这些名称专门用于此目的。

在欧拉之前，数学家转而关注在顶点相交的面之间包含的立体角的大小，而不是顶点本身。 如果我们逐渐改变多面体的尺寸，立体角将不断变化，但顶点数的离散计数 V 将保持不变。 这个离散计数 V 是欧拉现在抓住的。

同样，在欧拉之前，没有人考虑过欧拉所描述的“两个面沿着它们的边连接在一起的接合点，由于缺乏公认的术语，我称之为边缘”。 然后欧拉继续计算这些边的数量 E 和面的数量 F。 这种从观察长度和角度不断变化的多面体转变为观察离散拓扑特征的转变，这又是欧拉天才的一个例子。

在实现了这一微妙而深刻的飞跃之后，欧拉很快就发现了一个非凡的经验模式。 图 [18.2] 说明了五种柏拉图立体，每一个旁边都列出了 V、F、(V + F) 和 E 的值。现在很明显，并且因为欧拉是第一个[3]宣布的，(V + F) 条目比 E 条目多 2。在所有情况下，正如欧拉写给哥德巴赫的那样，我们有

欧拉的原始公式：
$$\boxed{\mathrm{V}+\mathrm{F}=\mathrm{E}+2}$$

虽然我们仅仅验证了这五个柏拉图多面体，欧拉测试了更多非柏拉图的例子，在 1750 年给哥德巴赫的信中，他表示相信这是关于多面体的普遍真理，尽管他承认他不知道如何以证明当时。 结果现在被称为**欧拉多面体公式**。 数学家们只是逐渐发现了其普遍性的精确极限：事实上，该公式仅适用于拓扑球面的多面体。

然而，上述公式并不是结果的现代形式。 如果我们将 E 移到等式的另一侧，我们得到的结果在代数上是非常等价的。 然而，这个看似微不足道的步骤代表了一个非常重要的概念进步（由庞家莱于 1895 年首次采用），因为方程的左侧现在是一个表征多面体 $\mathcal{P}$ 的单个整数，这是我们对其的新定义

欧拉示性数：
$$\boxed{\mathcal{X}(\mathcal{P})\equiv\mathrm{V}+\mathrm{F}-\mathrm{E}} \tag{18.1}$$

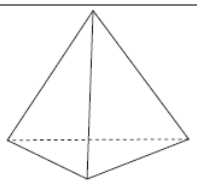
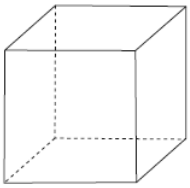
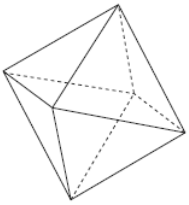
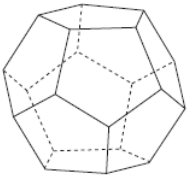
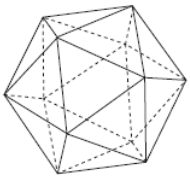
有了这个概念，这里是欧拉多面体公式的现代形式：

$$\boxed{\mathcal{X}\big(\text{拓扑球形多面体}\big)=2} \tag{18.2}$$

[3] 1860 年，在欧拉宣布他的奇妙发现一个多世纪后，一份失传已久的手稿在写成两个世纪后奇迹般地浮出水面。 它揭示了笛卡尔早在 1630 年，比欧拉早了一个多世纪，就已经做出了与欧拉基本相同的发现（但形式不同）！ Stillwell (2010, p. 469) 和 Richeson (2008, Ch. 9) 都很好地讲述了这个引人入胜的故事。

这与我们之前对 $\chi(Sg) = 2 - 2g$ 的定义一致，其中 $g = 0$，但关键的区别在于，现在我们将能够证明 (18.2) 来自 (18.1)。 事实上，我们将提出两种截然不同的证明。

欧拉本人最终确实成功地给出了他自己公式的第一个证明，并且在某些方面他的证明比我们将要提出的两个证明在拓扑上更自然，但是为了使他的论点完全有说服力，必然克服一些微妙的障碍。 （详见 Richeson 2008，第 7 章。）

Platonic Solid	Name	V	F	V + F	E
	Tetrahedron	4	4	8	6
	Cube	8	6	14	12
	Octahedron	6	8	14	12
	Dodecahedron	20	12	32	30
	Icosahedron	12	20	32	30

[18.2] 欧拉多面体公式，对所有五种柏拉图多面体都进行了经验验证：$V + F = E + 2$。

虽然我们还没有提供 (18.2) 的任何证明，但我们在本节结束时指出它有许多后果，其中最引人注目的一个涉及 [18.2] 中所示的五个柏拉图多面体。 首先，回想一下欧几里得的《几何原本》[4]提供了几何证明（例 24 中概述），这五个多面体是现存仅有的“正多面体”。正多面体是一种所有面都是全等正多边形的多面体，并且相同数量的这些正多边形在每个顶点处相交。

现在假设我们完全放宽几何约束，只坚持每个“面”——我们现在想象是弯曲的、可弯

[4]尽管欧几里德负责在他的《几何原本》中发布证明，但证明本身被认为来自雅典的泰阿泰德（柏拉图的朋友），可追溯到公元前 400 年左右。

曲的和可拉伸的——具有相同数量的波浪边，并且这些不规则、弯曲的边的数量相等面在每个顶点相遇。 练习 25 使用 (18.2) 来证明一个显着的事实，即只有五种拓扑可能性仍然是正确的，每一种在拓扑上都与古代的五种柏拉图多面体中的一种没有区别。 因此，五个柏拉图多面体令人着迷的几何美和规律性原来是一个非凡的、长达 2000 年的红鲱鱼！

18.3 欧拉多面体公式的柯西证明

18.3.1铺平多面体

从古希腊时代一直到 18 世纪，多面体都被视为固体； 事实上，我们仍然在谈论柏拉图多面体（Platonic solids，solids 也可翻译成“固体”）。 1813 年，柯西似乎是第一个实现概念飞跃的人，将多面体从它所包围的实体上剥离，将其视为一个中空的曲面。

他证明欧拉多面体公式（18.2）的下一个关键步骤是将这个空心曲面展平到平面上。 柯西对于如何实现这一点有些含糊不清，而瑞奇逊（Richeson, 2008, Ch. 12) 在阐明柯西的方法和其他数学家随后的澄清方面做了令人钦佩的工作。

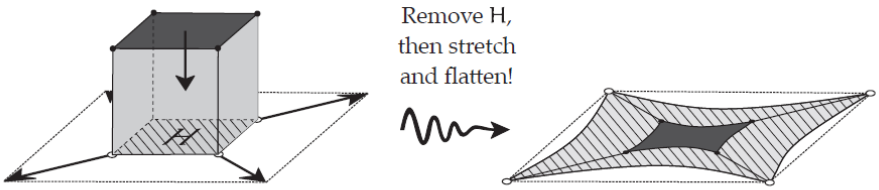
不过有一件事是肯定的：柯西的思想仍然牢牢植根于几何学中，他的多面体必然有直边和平面，而且它们的平铺版本也必然有直边。 此外，他的证明要求多面体是凸的：这个严格的几何要求意味着如果你的眼睛在多面体内部的任何地方，你可以看到整个曲面——在凸多面体内部无处可藏！ 但凸性显然不是拓扑条件，因此它不是欧拉多面体公式有效性的必要条件。 因此，我们选择回避所有这些历史文物，取而代之的是提供一个更现代的、纯粹的柯西论证的拓扑版本，但仍然忠实于他辉煌的原始洞察力。

为此，想象一下多面体的面是由可弯曲、可拉伸的橡胶布制成的，并将顶点和边描绘成只不过是在这个封闭的拓扑球形多面体气球上用钢笔绘制的点和连接曲线。 即使我们从一个经典的直线多面体开始，我们现在可能会不断地变形它（不切割或连接它的任何部分），并且产生的曲面将具有与原来画上相同数量的顶点、边和面。

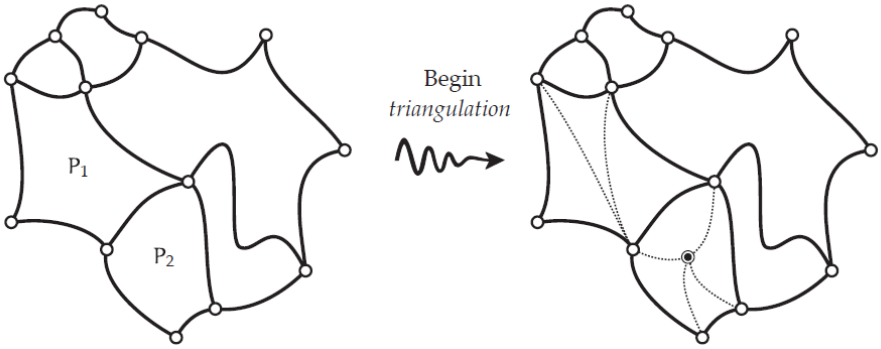
为了展平多面体，再次想象我们从具有平面的经典刚性曲面开始。 现在切出并丢弃其中一个面 H，并将多面体放在平坦的表面上，H 面朝下。图 [18.3] 的左侧说明了对立方体如何做到这一点。 接下来，返回将多面体描绘成橡胶布（现在带有孔 H），顶点是绘制在其曲面上的点。 取包围孔 H 的顶点，在平面内径向向外拉出每个顶点，使孔的边界越来越大，其余的多面体被拉下，最终在扩大的孔内拉平 H（见 [18.3] 的右手边）。

在以这种方式展平多面体后，我们可以进一步变形边缘（同时保持它们在平面上），以便它们呈现我们喜欢的任何形状，并且我们这样做是拓扑不受惩罚，因为这样的变形不会改变 V，E 或 F. 我们将生成的图形称为**多边形网**[5]。

[5]更传统的名称是网或图，但这些术语的使用只允许“多边形”有两条边，而我们将坚持我们的多边形至少有三个边。



[18.3] *To flatten the cube, we remove its bottom face H, then stretch out the resulting hole (hatched) until we have pulled the remaining faces down into the plane.*



[18.4] 多边形网的欧拉示性数不会被三角剖分改变。

18.3.2多边形网的欧拉示性数

图 [18.4] 展示了一个更典型的多边形网示例，它是通过展平非柏拉图多面体而获得的。

正如柯西推理的那样，在将多面体展平的过程中，我们移除了一个面，并且没有改变顶点或边的数量：

$$\mathrm{V} \rightsquigarrow \mathrm{V}, \qquad \mathrm{E} \rightsquigarrow \mathrm{E}, \qquad \mathrm{F} \rightsquigarrow \mathrm{F}-1 \quad \Longrightarrow \quad \mathcal{X} \rightsquigarrow \mathcal{X}-1$$

简而言之，展平多角体使其欧拉示性数减少了 1。

因此，要证明欧拉多面体公式（18.2），只需证明这个优雅的结果：

$$\boxed{\mathcal{X}(\text{多边形网}) = 1} \tag{18.3}$$

你可能需要检查一下我们画出的特定示例是否如此。

为了证明这总是正确的，我们首先观察我们可以将每个 n 边形（$n > 3$）分解成三角形，并且至关重要的是，这个**三角剖分**过程不会改变 $\mathcal{X}$ 的值。

该图说明了两种剖分方法。 在多边形 P_1（我们假设它是一个 n 边形）中，我们从一个顶点到所有其他顶点画出 $(n-3)$ 条曲线“对角线”，将这个面分成 $(n-2)$ 个三角形，

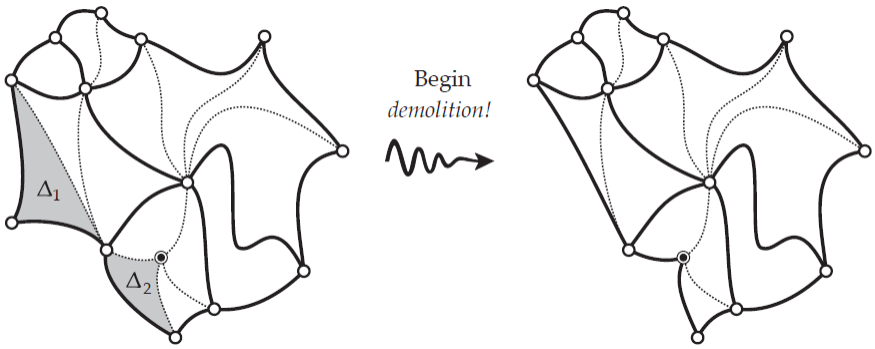
净增加（$n - 3$）个面。

因此，

$$\text{如果 } P_1\text{：} V \rightsquigarrow V, \qquad E \rightsquigarrow E + (n-3), \qquad F \rightsquigarrow F + (n-3) \quad \Longrightarrow \quad \mathcal{X} \rightsquigarrow \mathcal{X}$$

在 P_2（我们假设它是一个 m 边形）中，我们改为在内部某处添加一个顶点，然后将其连接到 m 个顶点，创建 m 个新边。 这将这个面分成 m 个三角形，净增加 $(m - 1)$ 个面。 因此，

$$\text{如果 } P_2\text{：} V \rightsquigarrow V + 1, \qquad E \rightsquigarrow E + m, \qquad F \rightsquigarrow F + (m-1) \quad \Longrightarrow \quad \mathcal{X} \rightsquigarrow \mathcal{X}$$



[18.5] 三角剖分的欧拉示性数不因它的拆除而改变。

[18.5]的左边显示了一种可能的三角剖分方式，我们刚刚证明了它的欧拉示性数必然与原始多边形网的欧拉示性数相同。

费尽心思构造这个三角剖分之后，（18.3）证明的最后一步是拆除它！ 也就是说，我们将一个一个地逐渐擦除每个三角形，直到只剩下一个三角形。 为了减少需要考虑的情况，我们将同意只蚕食边缘周围的三角形； 我们不会在网的中间挖洞，因为这可能会将网分成两个不同的、互不相连的孤岛。

有了这个协议，只需要考虑两种情况：边界三角形与网络内部共享一条（虚线）边（例如，Δ_1），或者它共享两条（虚线）边（例如，Δ_2） .

首先，假设我们

$$\text{擦除 } \Delta_1\text{：} V \rightsquigarrow V - 1, \qquad E \rightsquigarrow E - 2, \qquad F \rightsquigarrow F - 1 \quad \Longrightarrow \quad \mathcal{X} \rightsquigarrow \mathcal{X}$$

其次，假设我们

$$\text{擦除 } \Delta_2\text{：} V \rightsquigarrow V, \qquad E \rightsquigarrow E - 1, \qquad F \rightsquigarrow F - 1 \quad \Longrightarrow \quad \mathcal{X} \rightsquigarrow \mathcal{X}$$

最终，只有一个三角形 Δ 将保持站立，其欧拉示性数将为

$$\mathcal{X}(\Delta) = \mathrm{V} - \mathrm{E} + \mathrm{F} = 3 - 3 + 1 = 1$$

由于 $\mathcal{X}$ 在三角剖分过程[18.4]下是不变的，在拆除过程[18.5]下，它的初值必然等于它的终值 1，所以我们证明了（18.3）。 至此，我们也完成了柯西对欧拉多面体公式的证明。

18.4 欧拉多面体公式的勒让德证明

回想哈里奥特（Harriot）的美丽结果 1.3），将球体上测地三角形的角盈 $\mathcal{E}$ 与其面积 $\mathcal{A}$ 相关联。 正如你在练习 5 中证明的那样，这可以推广到测地线 n 边形。 欧氏 n 边形的内角和为 $(\mathrm{n} - 2)\pi$，因此曲面上测地 n 边形的角盈 $\mathcal{E}$ 为

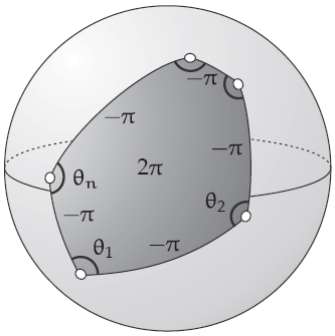
$$\mathcal{E}(\mathrm{n}\ 边形) = [内角和] - (\mathrm{n} - 2)\pi \tag{18.4}$$

从已被证明的 $\mathcal{E}$ 的可加性，哈里奥特的结果然后得到，

$$\frac{1}{\mathrm{R}^2}\mathcal{A}(\mathrm{n}\ 边形) = \mathcal{E}(\mathrm{n}\ 边形) = [内角和] - \mathrm{n}\pi + 2\pi \tag{18.5}$$

其中 R 是球面的半径。

图 [18.6] 提供了该公式右侧的角度过剩表达式的有价值的可视化[6]。 在 n 边形内，(A) 标记内角 $\theta_1, \theta_2, \cdots\cdots, \theta_n$; (B) 在每条边旁边写上“$-\pi$”； (C) 在中间写上“$2\pi$”。 那么 $\mathcal{E}$ 就是图中所有内容的总和：$(\mathrm{A}) + (\mathrm{B}) + (\mathrm{C})$。



[18.6] 测地 n 边形的 $\mathcal{E}$ 值可以可视化为图片中所有内容的总和。

勒让德在 1794 年提出了一个巧妙的证明，第一步是将多面体投影到球体上。 他进行这种投影的具体方式（在例 26 中描述）要求他假设多面体是凸的，正如柯西在 20 年后也会做的那样。 但是，正如我们在柯西证明的上下文中所指出的，凸性实际上不能与拓扑结果相关。 因此，我们将再一次避开这一历史产物，并提出一个更明显的拓扑版本的论证——一个不取决于凸性，但忠实于勒让德的基本洞察力。

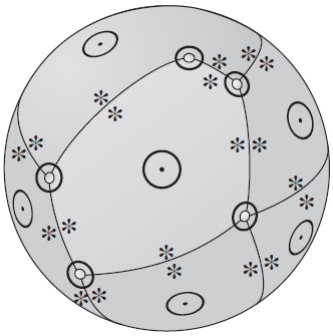
[6] 我们独立地想到了这个想法，但后来发现它已由 Richeson (2008, p. 95) 发表。

和以前一样，想象我们的多面体 P 是一个弯曲的、拓扑球形的橡胶膜，在其曲面上简单地绘制了点（顶点）和连接曲线（边）。 不像我们在柯西的证明中那样折叠多面体，这次让我们像气球一样充气！ 因此，我们得到了一个多边形网，覆盖了最终球形气球的曲面。请注意，由此产生的边缘不是测地线，但是，正如我们现在解释的那样，我们可以轻松地使它们如此而无需改变欧拉示性数。

想象这个球体是刚性的，在我们刚刚构建的每个顶点处都有小钉子。 接下来，将当前的非测地线边想象成由拉伸的弹性绳制成，暂时固定并连接到它们末端的钉子（顶点）。 进一步想象这些弹性绳和球体之间没有摩擦：它们可以在曲面上自由滑动而没有阻力。 现在让我们将钉子敲平在曲面上，这样任何没有连接到钉子的绳子都可以扫过它的位置而不会被卡住。

最后，放开弹性绳！ 它们将自动收缩以在钉子之间创建最短（= 测地线 = 大圆）路线。 这样我们就得到了一个完全覆盖球体的测地多边形网，它的欧拉示性数必然与原始值 $\mathcal{X}(\mathcal{P})$ 相同。

现在我们来到勒让德巧妙想法的第二部分：对构成这个测地球面覆盖网的所有测地多边形 P_j 应用公式（18.5）两边求和（[18.7] 给出了一个例子）。 让我们先看看右边的总和。可视化 [18.6] 的功效现在很明显，因为它使我们能够看见整个网络的总和值，如下所示。



[18.7] $\sum_j \mathcal{E}(P_j)$ 是这张图片中所有东西的总和。 这里，$* = -\pi$，且 $\odot = 2\pi$。

为了避免混乱，[18.7] 使用了两个视觉缩写：$* = -\pi$ 和 $\odot = 2\pi$。 将注意力集中在任何一个顶点上：我们正在对所有多边形的内角 θ_j 求和，因此，特别是，我们正在对围绕该顶点的所有角度求和，从而得到 2π。 由于每个顶点贡献 2π，内角的总贡献为 $2\pi V$。 接下来，每条边都有一个 $*$（即 $-\pi$）写在它的两侧，每条边产生-2π。 因此 $-\pi$ 的总贡献是 $-2\pi E$。 最后，每个面带有一个 $\odot$（即 2π），总共产生 $2\pi F$。

因此，

$$\boxed{\sum_j \mathcal{E}(P_j) = 2\pi(V - E + F) = 2\pi\,\mathcal{X}(\mathcal{P})} \tag{18.6}$$

由于多面体网完全覆盖了球体的曲面，因此对（18.5）的左侧求和得到

$$\frac{1}{R^2}\sum_j \mathcal{A}(P_j) = \frac{1}{R^2}\left[\text{球面的面积}\right] = 4\pi$$

最后，(18.5) 由于这两个量相等：

$$2\pi\,\mathcal{X}(\mathcal{P}) = 4\pi \quad \Longrightarrow \quad \mathcal{X}(\mathcal{P}) = 2$$

从而完成欧拉多面体公式的勒让德证明。

尽管它显而易见的美丽，勒让德的证明在道德上感觉是错误的，因为它通过不断变化的几何角度来取得成功（反常地），而这些几何角度在拓扑中毫无意义。 但从微分几何的角度来看，这或许是一种正确的证明，因为它似乎以一种令人惊讶的方式将几何和拓扑联系起来，可能有助于解释 GGB。 我们很快就会看到，这种乐观是有道理的。

18.5 向曲面添加环柄增加其亏格

为了将这些最新的想法与 GGB 联系起来，我们的下一步是证明我们对欧拉特性的新定义（18.1）确实暗示（16.2）作为定理：

$$\boxed{\mathcal{X}(\mathcal{S}_g) = 2 - 2g} \tag{18.7}$$

我们相信这个结果还没有一个公认的名称，于是我们可以将其称为**欧拉–惠里尔公式**（Euler-L'Huilier formula），以纪念西蒙•安托万•让•惠里尔（Simon Antoine Jean L'Huilier），他在 1813 年第一个发表欧拉结果的这种概括; 详情参见 Richeson (2008, Ch. 15)。

我们刚刚在拓扑球面（$g = 0$）情况下给出了这个定理的两个证明，所以现在需要了解在曲面上打孔的效果。显然，应该从理解最简单的情况开始，即有一个孔的环面/甜甜圈（$g = 1$）。

[18.8] 一个咖啡杯在拓扑上等同于一个甜甜圈。 拓扑笑话，作者基南科瑞恩（Keenan Crane）和亨利塞格曼（Henry Segerman）； 参见塞格曼（2016 年第 101 页）。 照片由塞格曼（Segerman）教授提供并经其许可使用。

我们将通过一个笑话来解决这个问题，这是一个糟糕的笑话，“可能是第一次将这样的[东西] 用于建设性目的。”[7] 这个笑话是这样的：拓扑学家是一个无法区分咖啡杯和甜甜圈的人。

这个笑话背后的真相在图 [18.8] 中有说明。 在杯子变成甜甜圈的过程中，杯子的主体已经合并成一个团块，上面还附有一个把手。如果我们现在想象这个团块成长为一个球体，我们会得到菲利克斯・克莱因（Felix Klein） 在 1882 年首次提出的重要观察结果：[8]

> 环面在拓扑上等价于附有一个手柄的球面。 更一般地说，$\mathcal{S}_g$ 在拓扑上等价于一个有 g 个手柄的球面。 (18.8)

（有关曲面拓扑分类的所有看似合理但未经证实的主张的严格陈述和证明，请参见延伸阅读（本书末尾）。）

为了证明（18.7），现在足以证明

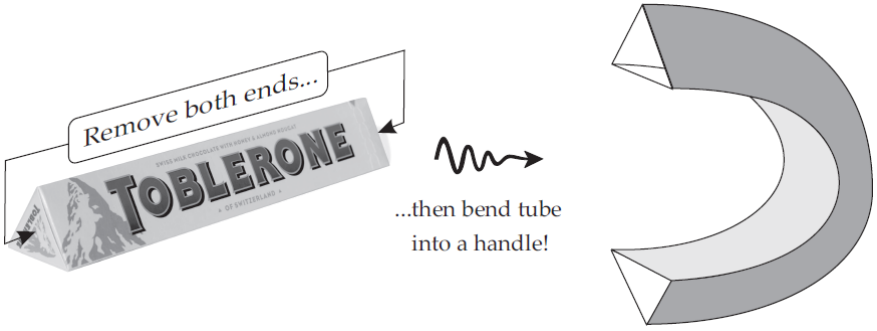
> 在闭合曲面 $\mathcal{S}$ 上添加手柄会使 $\mathcal{X}(\mathcal{S})$ 减少 2。 (18.9)

于是，添加 g 个手柄显然使 $\mathcal{X}$ 减少了 2g。 但是我们已经确定拓扑球面有 $\mathcal{X} = 2$，因此添加 g 个手柄将其简化为 $\mathcal{X}(\mathcal{S}_{\mathrm{g}}) = 2 - 2g$，从而证明（18.7）。

---

[7] 柯克船长在《末日机器》的结尾对斯波克先生讲话。 原话是“这样的武器”。

[8] 参见 Stillwell (1995, p. 60)。



[18.9] 要创建一个简单的手柄，请从三棱柱（这里用瑞士三角牌巧克力盒子表示）移除末端，然后弯曲管子以创建手柄（具有消失的欧拉示性数）。

为了证明（18.9），我们将首先构造一个把手，然后将它粘在给定的曲面 $\mathcal{S}$ 上。图 [18.9] 显示了用瑞士三角牌（Toblerone）巧克力盒构造一个特别简单的把手，即三角形直角棱镜。我们首先移除两个三角形端面（所以 $\mathrm{F} \rightsquigarrow \mathrm{F}-2$），从而创建一个空心管。 由欧拉多面体公式，或通过简单的计数[练习]，

$$\mathcal{X}(\text{三角牌巧克力盒}) = 2 \quad \Longrightarrow \quad \mathcal{X}(\text{空心管}) = 0$$

最后，如图所示，我们想象用橡胶制成的管子，然后将其弯曲成空心手柄的形式。 这不会改变它的欧拉示性数，所以

$$\boxed{\mathcal{X}(\text{环柄}) = 0}$$

接下来，如图 [18.10] 中所示，我们在曲面 $\mathcal{S}$ 上创建两个孔，该手柄的末端将粘在上面。 为此，假设我们已经对 $\mathcal{S}$ 进行了三角剖分[9]。接下来，移除这些三角形面中的两个，因此 $\mathrm{F} \rightsquigarrow \mathrm{F}-2$ 和 $\mathcal{X}(\mathcal{S}) \rightsquigarrow \mathcal{X}(\mathcal{S})-2$。

图 [18.10] 描绘了手柄向现在有两个孔的 $\mathcal{S}$ 移动，在它粘到 $\mathcal{S}$ 中的孔之前的片刻。手柄的末端有六个顶点和六个边，孔也是如此。但是在它们粘合在一起之后，只剩下六个顶点和六个边：净减少六个顶点和六个边。但这意味着总的欧拉示性数不会因粘上手柄的动作而改变。 （注意：显然，如果手柄的末端是 n 边形（$\mathrm{n} > 3$），粘在匹配的 n 边形孔上，这仍然是正确的。）

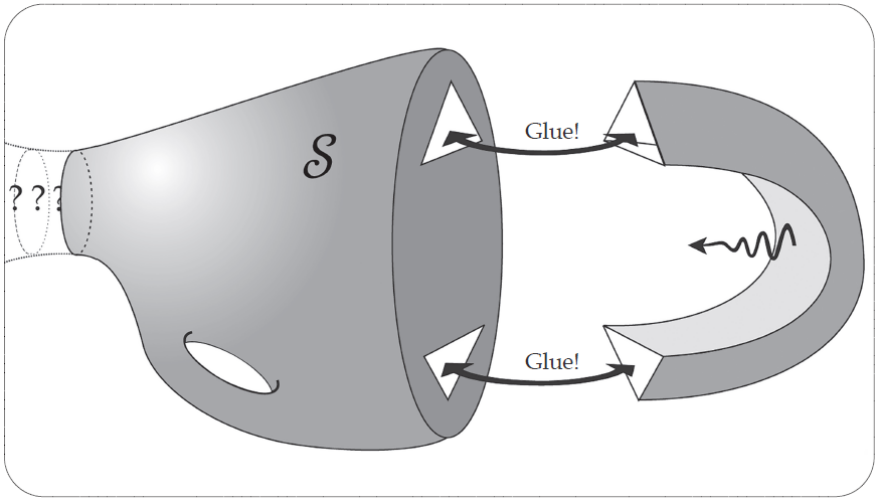
因此，如果 $\tilde{\mathcal{S}}$ 是通过粘合在手柄上获得的新曲面，

$$\begin{aligned}\mathcal{X}(\tilde{\mathcal{S}}) &= \mathcal{X}(\text{两孔曲面}\mathcal{S} + \text{粘上后的环柄}) \\ &= \mathcal{X}(\text{两孔曲面}\mathcal{S}) + \mathcal{X}(\text{粘合前的环柄}) \\ &= \mathcal{X}(\mathcal{S}) - 2 + 0\end{aligned}$$

从而证明（18.9），因此也证明（18.7）。

[9] 未显示三角剖分，但请注意，该论证并不要求我们从测地线三角形构建三角剖分。

有关霍普夫优雅的替代方法，请参见练习 27。



[18.10] 一般闭合曲面 $\mathcal{S}$ 以未知方式向左延伸，用???表示。 我们从 $\mathcal{S}$ 的三角剖分（未显示）中切出两个三角形，将 $\mathcal{X}(\mathcal{S})$ 减少 2，然后粘在手柄上，这对 $\mathcal{X}(\mathcal{S})$ 没有影响。 因此，粘在手柄上会使 $\mathcal{X}(\mathcal{S})$ 减少 2。

18.6 全局高斯-博内定理的角盈证明

让我们回到勒让德对欧拉多面体公式的证明，让我们从它所依赖的结果后退一步：哈里奥特结果的推广，（18.5）。 现在只关注右手边的“角盈”，由（18.4）给出：

$$\mathcal{E}(\mathrm{n}\ \text{边形}) = [\text{内角和}] - \mathrm{n}\pi + 2\pi$$

当然，这个表达式的全部意义在于测量测地 n 边形的内角超出欧氏几何的预测。 如果我们放弃 n 边形的边是测地线的要求，那么这种解释就不再适用。 此外，该表达式总是没有拓扑意义，因为它涉及角度。 尽管如此，即使 n 边形的边不是测地线，表达式本身仍然有意义。

尽管该表达式具有非拓扑性质，但我们在对 GGB 进行新解释的道路上的第一个关键步骤是认识到整个多边形网上该表达式的总和确实具有拓扑意义。 事实上，即使边不是测地线，我们看到 [18.7] 及其结论 (18.6) 仍然有效：

$$\boxed{\sum_i \mathcal{E}(\mathrm{P_j}) = 2\pi(\mathrm{V} - \mathrm{E} + \mathrm{F}) = 2\pi\,\mathcal{X}(\mathcal{P})}$$

证明这个结果所需要的只是一个事实，即图 [18.7] 中围绕顶点的角度和为 2π。

出于同样的原因，勒让德证明的这一部分不依赖于拓扑球形多面体膨胀成一个完美的圆形几何球面。 事实上，这是下一个关键点，它甚至不依赖于拓扑球面：（18.6）在任意亏格 g 的曲面 $\mathcal{S}_\mathrm{g}$ 上仍然有效。

就算我们一直（并将继续如此！）在拓扑上的再漫不经心，如果我们没有指出不能在 $\mathcal{S}_g$ 上投射一个完全任意的多边形网，仍是失职。相反，我们必然谨慎行事：我们不能允许我们的多边形连接到自己身上。 例如，在环面上，我们当然可以想象（但必然避免）一个多边形，它穿过孔洞并返回咬自己的尾巴！

随着我们的拓扑意识现在清晰（呃），我们回到证明 GGB 的努力，以及勒让德证明中的关键方程，即网络中任意测地多边形 P_j 的广义哈里奥特结果（18.5）：

$$\frac{1}{R^2}\mathcal{A}(P_j) = \mathcal{E}(P_j)$$

虽然刚刚看到右侧的总和在本质上是拓扑的——并且总是等于 $2\pi\,\mathcal{X}(\mathcal{S}_g)$——但与左侧的相等仅在我们 (i) 将多面体膨胀成半径为 R 的完美球体和 (ii) 确保多边形网的所有边都是该球面上大圆的测地弧时成立。

虽然哈里奥特在 1603 年无法认出它，但两个世纪后高斯告诉我们球面具有恒定曲率 $\mathcal{K} = (1 / R^2)$，并且上述方程的左侧实际上应该被视为 P_j 内的总曲率 $\mathcal{K}\,(P_j)$：

$$\mathcal{K}\,(P_j) = \mathcal{K}\,\mathcal{A}(P_j) = \frac{1}{R^2}\mathcal{A}(P_j)$$

高斯 1827 年局部高斯-博内定理的形式（2.6）是对这个结果的极大推广：一般曲面上测地三角形 Δ 的角盈同样由总曲率给出. 正如哈里奥特的原始结果推广到测地多边形一样，高斯的结果也是如此（以完全相同的方式）：

$$\mathcal{E}(\Delta) = \iint_{\Delta} \mathcal{K}\,\mathrm{d}\mathcal{A} \quad \Longrightarrow \quad \mathcal{E}(P_j) = \iint_{P_j} \mathcal{K}\,\mathrm{d}\mathcal{A} = \mathcal{K}(P_j)$$

最后，对所有测地多边形求和，并使用（18.6）和（18.7），我们得到了 GGB 的第二个证明：

$$\boxed{\mathcal{K}(\mathcal{S}_g) = \sum_j \mathcal{K}(P_j) = \sum_j \mathcal{E}(P_j) = 2\pi\mathcal{X}(\mathcal{S}_g) = 2\pi(2-2g)}$$

第十九章 全局高斯-博内定理的第三个（向量场）证明

19.1 引言

我们以几何、拓扑和向量场之间的美丽联系结束第三幕（我们戏剧的传统“高潮”）——高潮中的高潮！ 在这里，我们将仅勾勒出实现对 GGB 的新理解所需的那些想法； 有关与复分析和物理相关的向量场的更完整说明，请参阅 VCA 的第 10、11 和 12 章。 另见本书末尾的延伸阅读。

19.2 平面向量场

想象一层薄薄的流体在水平面上流动。 将这个平面视为复平面 $\mathbb{C}$ 对我们有利。因此，在 $\mathbb{C}$ 的每个点 z 处，我们都有一个速度向量，或者更确切地说，一个复数 $V(z)$，我们画成从 z 出发。 这个流 $V(z)$ 被称为 $\mathbb{C}$ 上的向量场。

我们将假设我们的向量场 $V(z)$ 表现得非常好，除了有限数量的孤立点外，它是连续的和可微的[143]。 因此，在普通或正则（regular）点，任何方向的微小移动都会导致 V 的方向和长度相应地发生微小的、最终成比例的变化。

与此相反，**奇点** s 是向量场遭受不连续性的特殊位置：在不同方向上远离 s 的无穷小运动导致 V 指向完全不同的方向或具有完全不同的长度。

图 [19.1] 展示了一个典型的向量场，其中奇点 a、b 和 c 可以立即被未经训练的眼睛识别出来，为了清楚起见，用黑点标记。

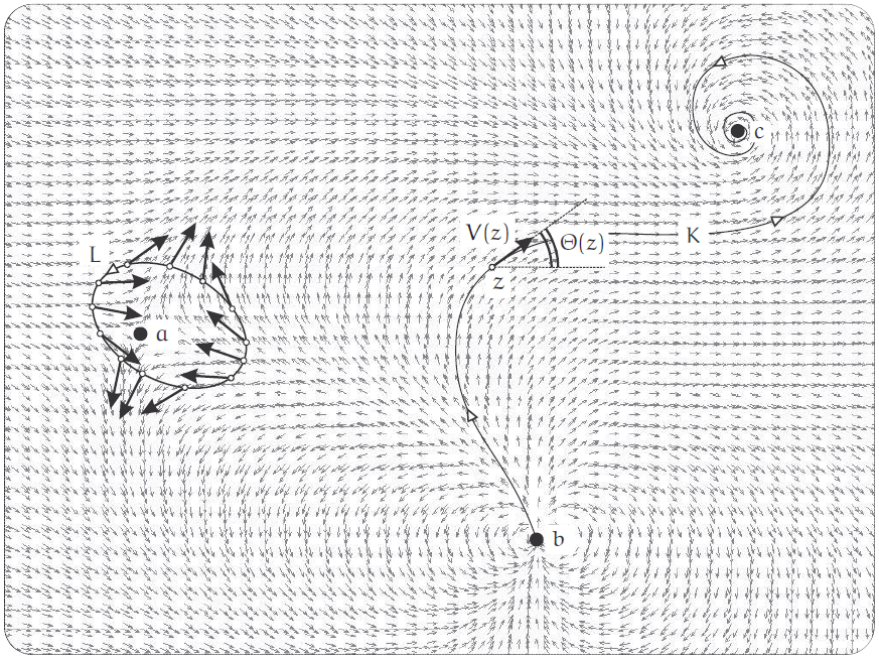
如果我们沿着流体的单个粒子的路径，我们得到流动的**流线** K（或**积分曲线**），这样 V 总是与 K 相切。如果我们跟踪所有这些流线，我们将在下面的例子中做，然后我们获得了向量场整体的生动描绘，称为**相图**。 注意，相图完全封装了 $V(z)$ 在每个点的方向 $\Theta(z) = \arg[V(z)]$ 的信息。

在 [19.1] 中，向量都具有相等的长度，但一般来说这肯定不是真的，我们将在下面的例子中说明。 这是相图的一个严重缺陷，因为它无法[144]表明大小 $|V(z)|$，而这个大小代表了重要的信息：例如流体流动的速度或磁场的强度。 然而，从拓扑的角度来看，相图是唯一重要的事情。

---

[143] 在较弱的真实意义上，而不是作为复解析意义上的伸扭。

[144] 在物理上最重要的矢量场的情况下，实际上有一种特殊的方式来绘制流线，以便流的强度变得可见，表示为流线的拥挤程度。 参见 VCA，第 494-502 页。



[19.1] 一个典型的向量场 V (z)。 清晰可见的是三个奇点：a 称为**鞍点**（或**交叉点**）； b 称为**偶极子**； c 称为**涡旋**（或**焦点**）。 还画出了一条**流线** K。**指数** $\mathcal{I}_V(a)$ 是当 z 沿 L 这样的环路绕 a 逆时针转一圈时 V (z) 旋转的净圈数。由于当我们围绕 a 转一圈时 V 顺时针（即负值）旋转一圈，我们看到 $\mathcal{I}_V(a) = -1$。

19.3 奇点的指数

正如奇点的位置是显而易见的，它们截然不同的特征也是如此。 如果我们想象一下在橡胶布上画出的流动模式[19.1]，然后以这种方式拉伸它，我们会直觉 a 的特征与 b 或 c 的特征完全不同，并且这种区别在拉伸下是不变的。换句话说，很明显，奇点 s 的特征与其说是几何特征，不如说是流的拓扑特征。

事实上，这些不同的特征导致了不同类型的奇点的不同名称。 例如，在[19.1]中，a 被称为**鞍点**（或**交叉点**），因为这是水倒在马鞍上时的流动方式； b 称为**偶极子**，因为这是磁场线在短条形磁铁的两个极之间流动的方式； c 被称为**涡旋**（或**焦点**），因为这是水在厨房水槽的排水管周围和下水道旋转的方式。

我们现在解释我们如何将奇点 s 的“特征”的这个模糊概念具体化为一个单一的、拓扑不变的整数 $\mathcal{I}_V(s)$，称为 s 的**指数**。（注意：若没有关于向量场 V 的任何歧义，我们可以去掉它并将符号缩写为 $\mathcal{I}(s)$。）

让我们立即陈述定义，尽管它甚至可能没有明确定义：

> 如果 s 是向量场 V (z) 的奇点，并且 L 是包含 s（并且没有其他奇点）的任何简单环路，那么 s 的指数 $\mathcal{I}_V(s)$ 是当 z 逆时针绕 L 转一圈时V (z) 的旋转的净圈数。 (19.1)

让我们立即用 [19.1] 中的奇点来说明这一点。 由于指数只关心 L 上向量的方向，为了清楚起见，我们冒昧地放大了这些特定的向量。 当我们遍历 L 时，很明显 V (z) 顺时针（即负）旋转一圈，所以

$$\mathcal{I}(\text{鞍点}) = -1$$

请自行检查 b 和 c 处的奇点，并验证

$$\mathcal{I}(\text{偶极子}) = +2, \qquad \mathcal{I}(\text{涡旋}) = +1$$

接下来，让我们更精确地定义（19.1）。令 $\Theta(z)$ 是 $V(z)$ 与水平面所成的角度。当然，这个角度有无限多个选择，彼此相差 2π 的倍数，但假设我们在 z 点选择一个。 假设我们要求 $\Theta(z)$ 连续变化，当 z 沿着有向曲线 J 移动时，V (z) 在 J 上连续变化也就意味着 J 不通过任何奇点。那么这时 $\Theta(z)$ 是唯一确定的。

我们现在可以将 $\delta_J\Theta$ 定义为角度 $\Theta(z)$ 在 z 沿 J 从开始到结束移动时的净变化：

$$\boxed{\delta_J\Theta \equiv \Theta(\text{结束}) - \Theta(\text{开始})}$$

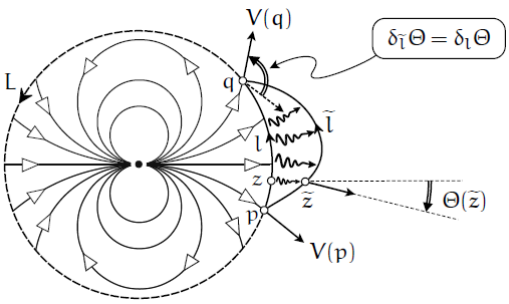
注意，这是角度的有符号变化，因此如果我们反转 J 的方向（表示为 $-J$），从而交换开始和结束，则 $\delta_J\Theta$ 的符号也会反转：

$$\boxed{\delta_{-J}\Theta = -\delta_J\Theta} \tag{19.2}$$

有了这样的理解和符号，我们可以如下重申指数的定义（19.1）：

$$\boxed{\mathcal{I}_V(s) = \frac{1}{2\pi}\mathcal{I}_L(s)} \tag{19.3}$$

到现在为止，你可能已经直觉地认为这个定义已经够明确了，但要证明它确实如此，请考虑 [19.2]，它用（最初）环绕偶极子的圆形环路 L 说明了一般论证。



[19.2] 圆形环路 L 环绕偶极子。 如果 L 的小段 l 变形为 $\tilde{l}$，把 z 带到 $\tilde{z}$，则角度 $\Theta(z)$ 连续变化，因此 $\Theta(\tilde{z})$ 将接近其原始值 $\Theta(z)$。 因此，沿该线段的 Θ 变化在变形下是不变的：$\delta_{\tilde{l}}\Theta = \delta_l\Theta$。

将注意力集中在 L 的图示小段 l，从 $p =$ 开始到 $q =$ 结束。 如图所示，

$$\delta_l\Theta = \Theta(q) - \Theta(p)$$

假设我们现在连续地将 l 变形到相邻的 $\tilde{l}$ 中，而 l 在此过程中没有穿过任何奇点。

随着 l 上的点 z 演化为 $\tilde{l}$ 上新的点 $\tilde{z}$，角度 $\Theta(z)$ 也会连续变化，因此 $\Theta(z)$ 将接近其原始值 $\Theta(z)$。 因此，当 $\tilde{z}$ 沿 $\tilde{l}$ 行进时，变化 $\delta_{\tilde{l}}\Theta$ 也必然接近 z 沿 l 行进时的变化 $\delta_l\Theta$。

现在是原型拓扑论证。 如果 $\delta_{\tilde{l}}\Theta$ 不完全等于 $\delta_l\Theta$，则必然与其相差 2π 的倍数，但这是不可能的，因为 $\delta_{\tilde{l}}\Theta$ 接近于 $\delta_l\Theta$。

因此，

$$\delta_{\tilde{l}}\Theta = \delta_l\Theta$$

我们可以类似地使 L 的任何其他部分变形，或一次变形所有部分，因此我们立即推断出指数确实与 L 的大小和形状无关：

> 如果 s 是向量场 V(z) 的一个奇点，并且 L 是任何包含 s（并且没有其他奇点）的简单环路，那么当 L 连续变形为任何其他简单环路时，只要它在过程中不跨越任何奇点，**指数** $\mathcal{I}_V(s)$ 就不会改变。 简而言之，$\mathcal{I}_V(s)$ 独立于L，是 V 的一个属性，可以附加到 s 本身。 (19.4)

如果我们将向量场视为一个复映射 $z \mapsto V(z)$ 从一个复平面到另一个复平面，取第一个平面中的点 z 到第二个平面中的点 V(z)，然后 $\mathcal{I}_V(s)$ 可以看作是像 V(L) 绕 0 的圈数[练习]。这称为 V(L) 的环绕数，它在复分析中起着基础性的作用； 参见 VCA，第 7 章。

19.4 原型奇点：复数的幂

为什么我们要把我们的平面变成复平面？ 有几个原因，但这里有一个重要的原因：原型奇点自然产生于复变量 z 的幂：

$$P_m(z) \equiv z^m = \left[r\, e^{i\theta}\right]^m = r^m e^{im\theta}$$

其中 m 是整数。

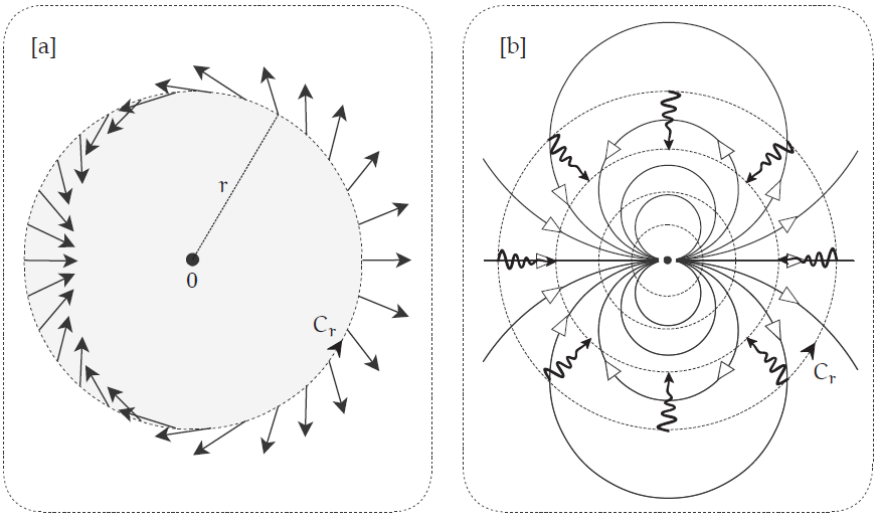
$P_m(z)$ 唯一的（有限）[145]奇点在原点，很容易确定它的指数。 沿着 θ 方向的射线，我们看到 P_m 指向 $m\theta$ 方向； 只有长度 $|P_m(z)| = r^m$ 随着我们沿着射线移动而变化。 因此，如果我们逆时针环绕任何包围原点的环路，使得 θ 从 0 到 2π，那么 P_m 的角度从 0 到 $2\pi m$。 换句话说，P_m 旋转 m（正或负）圈，所以 $\mathcal{I}(0) = m$。

例如，重新考虑我们在 [19.1] 中在 b 处遇到的“偶极子”。 图 [19.3] 说明这个偶极场就是 $P_2(z) = z^2$：[19.3a] 画出了半径为 r 的圆形环路 Cr 上的向量场本身，[19.3b] 画出了流线的相图。 确保你可以看到这两个图表至少定性是正确的（为了加分，证明[19.3b]中的流线确实是完美的圆！）

尽管指数的定义 (19.3) 本质上使用了包围奇点的环路 L，但我们已经在（19.4）中看到 L 的形状和大小实际上是混淆视线（原文红鲱鱼指掩人耳目的事物）：指数实际上是在奇点处的向量场自身的一个性质。

为了使这个想法更加生动，[19.3b] 说明了环路 C_r 向奇点收缩。 显然，向量场在奇点的无穷小邻域中的行为决定了该点的指数：

$$\mathcal{I}_V(s) = \frac{1}{2\pi}\lim_{r\to 0}\delta_{C_r}\Theta$$



[19.3] 当我们围绕 0 处的奇点（正向）环绕 [a] 中的圆一次时，$P_2(z) = z^2$ 正向旋转两次，因此 $\mathcal{I}(0) = +2$。 通过检查 [b] 中流的相图（即流线）可以获得相同的结果。 当我们将 Cr 缩小到奇点时，指数不会改变：指数表征奇点的无穷小邻域中的向量场。

综上所述，如果 m 是任意整数，

$$P_m(z) = z^m \quad \Longrightarrow \quad \mathcal{I}(0) = m$$

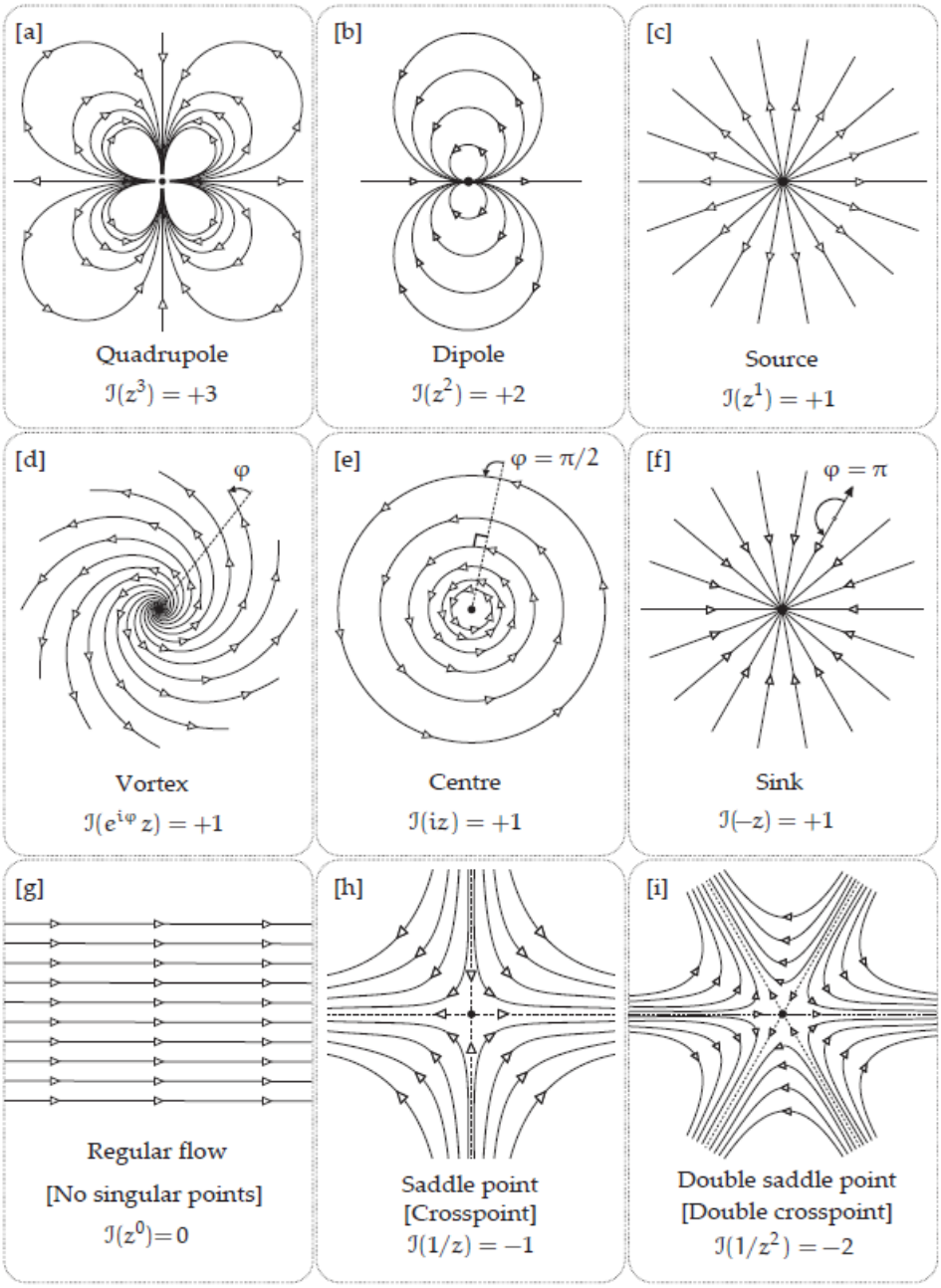
[145] 实际上在无穷远处还有第二个奇点。 参见 VCA，第 492 页。

图 [19.4] 画出了当 $m = +3, +2, +1, 0, -1, -2$ 时这些场的样子。 请花点时间对每个场进行视觉验证，(i) 它确实是它声称的样子，并且 (ii) 指数确实等于幂指数 m。

当我们以奇点“性格”的直观想法开始我们的讨论时，更精确的指数概念将我们推向了一个新的方向。 例如，一个源、一个涡旋和一个汇，从物理的角度来看，在性质上都大不相同，但从拓扑的角度来看，它们是*无法区分*的，因为这三者的指数都是+1。

这是由以下事实解释的：将复函数 $f(z)$ 乘以复常量 $k = Re^{i\varphi}$ 对其指数没有影响。对于 $f(z) \rightsquigarrow Re^{i\varphi}f(z)$ 将向量拉伸 R 并*旋转固定角度* φ。 因此，当 z 环绕任何环路时，$kf(z)$ 旋转的圈数与 $f(z)$ 相同，因此指数保持不变。 特别地，如果我们让 $f(z) = z$ 并且让 φ 从 0 逐渐增加到 $(\pi/2)$，然后从 $(\pi/2)$ 到 π，那么 [19.4c] 中源的场是这样演化的: [c] ⇝ [d] ⇝ [e] ⇝ [f]。

我们以一些一般性观察结束本节。 首先，请注意*反转任何流的方向对其奇点的位置或指数没有影响*。 当我们围绕一个奇点环绕一个环路时，V 和 −V 经历了旋转完全相等的圈数，就像罗盘针的两端一样，因此两个向量场的共同奇点对两者具有相同的指数。



[19.4] $\mathbb{C}$ 中复变量 z 的整数次幂的向量场，$P_m(z) = z^m$，$m = +3, +2, +1, 0, -1, -2$。 原点奇点的指数 $\mathcal{I}$ 等于 m。中间一行画出 $e^{i\varphi}z$ 的向量场，都有 $\mathcal{I} = +1$，对于[d] φ 为一般值的涡旋； [e]**圆心**，$\varphi = \pi/2$； [f] **汇**，$\varphi = \pi$。

替代名称：汇 = **稳定节点**； 源 = **不稳定节点**； 涡旋 =（**稳定或不稳定**）**焦点**— [d] **是不稳定的**； 鞍点=**交叉点**。

接下来，请注意，如果我们定义复共轭向量场为

$$\overline{P}_n(z) = \overline{z^n} = \bar{z}^n$$

那么 $P_m(z)$ 的流线与 $\overline{P}_{-m}(z)$ 的流线相同，因此它们在原点的公共奇点对于两个场都是零，并且在两个场中具有相同的指数。

更一般地，对于任何复函数 $f(z)$，向量场 $\overline{f(z)}$ 已被称为 $f(z)$ 的**波利亚向量场**，以纪念乔治·波利亚（George Pólya (1887 – 1985)）。 显然，波利亚向量场具有与 $f(z)$ 相同的奇点，但由于 $\overline{f(z)}$ 具有相反的角度，这些共同的奇点具有相反的指数。 波利亚向量场的主要优点是它允许对 $f(z)$ 的复轮廓积分进行简单、直观、视觉和物理解释。 有关详细信息，请参阅 VCA 第 11 章。

最后，观察如果我们取 $f(z)$ 的倒数，那么我们会得到一个新的向量场 $1/f(z)$，它在一个奇点处趋于无穷大（黎曼球的北极），其中 $f(z)=0$，并且 $1/f(z)$ 在这样一个奇点的指数是 $f(z)$ 的指数的负数[练习]：这一点在 $P_m(z)$ 的情况下尤其明显，对于那么$1/P_m(z)$ 指向与 $P_{-m}(z)$ 相同的方向。 如果我们现在考虑这个倒数的波利亚向量场，即 $1/\overline{f(z)}$，则指数第二次反转，因此与 $f(z)$ 的指数相同。

物理术语说明。 我们借用了物理学中的术语**源**、**涡旋**、**汇**、**偶极子**和**四极子**来描述 z、z^2 和 z^3 的流线。 虽然这在拓扑上是正确的，但在物理上并不正确：与这些对应的物理场实际上是波利亚向量场 $\overline{(1/z)}$、$\overline{(1/z^2)}$ 和 $\overline{(1/z^3)}$。

正如上一段所解释的，这些场中的每一个都与相应的正幂指向相同的方向，因此它们具有相同的奇点和指数，从而证明我们在当前拓扑设置中使用这些术语是合理的。 然而，物理场的大小是我们赋予相同名称的正幂大小的倒数。

为了问心无愧，让我们以最简单的源的例子来解释这一点。 假设水以恒定速率 s 通过一个非常窄的管子供应到水平面（视为复平面 $\mathbb{C}$）上的一个点。水将从这个源（现在用于物理意义上）径向向外、对称地流出，我们假设它位于原点。

虽然 [19.4c] 中描述的径向速度场 $P_1(z)=z=re^{i\theta}$ 的大小为 r，随着我们远离原点而加速，但直观上很明显，来自物理源的径向流的速度 v 必然水流得越远，速度越慢。 更准确地说，流出半径为 r 的圆（每单位时间）的总水量必然等于原点供应的水量，即 s。因此，

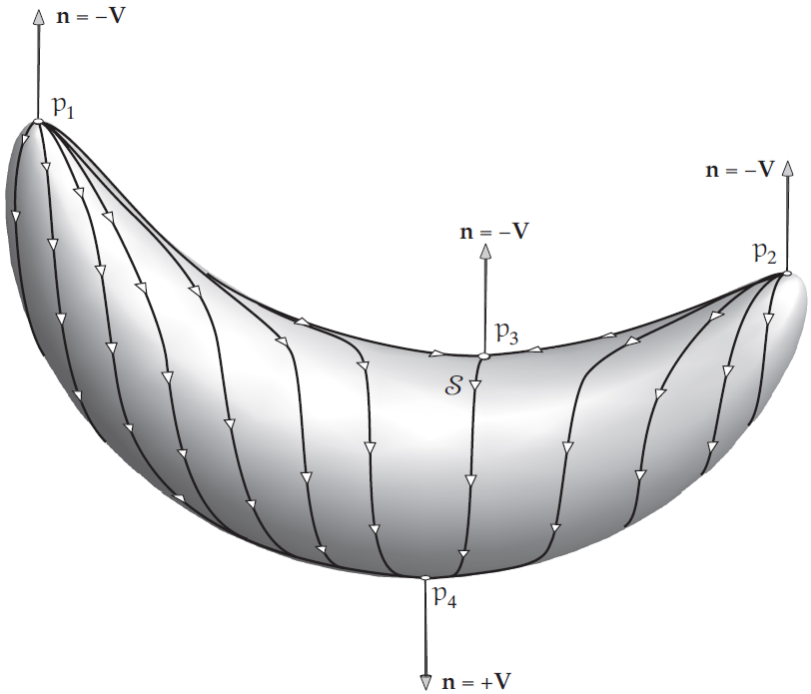
$$2\pi r\, v = s \Longrightarrow v = \frac{s}{2\pi r} \Longrightarrow \text{物理源} = v\, e^{i\theta} = \frac{s}{2\pi}\left[\frac{e^{i\theta}}{r}\right] = \frac{s}{2\pi}\overline{\left[\frac{1}{z}\right]}$$

有关此类场的物理解释以及它们与复分析的共生关系的更多信息，请参见 VCA，第 10、11 和 12 章。

19.5 曲面上的向量场

19.5.1蜂蜜流向量场

很容易想象在曲面上的向量场，而不是在平面上。 想象一下在暴雨中留在外边的任何光滑物体：雨水流过物体的表面 $\mathcal{S}$，向下流向地面，它在每个点 $p \in \mathcal{S}$ 处的速度 $v(p)$ 是**曲面上向量场**的一个例子，向量本身处处与曲面相切。



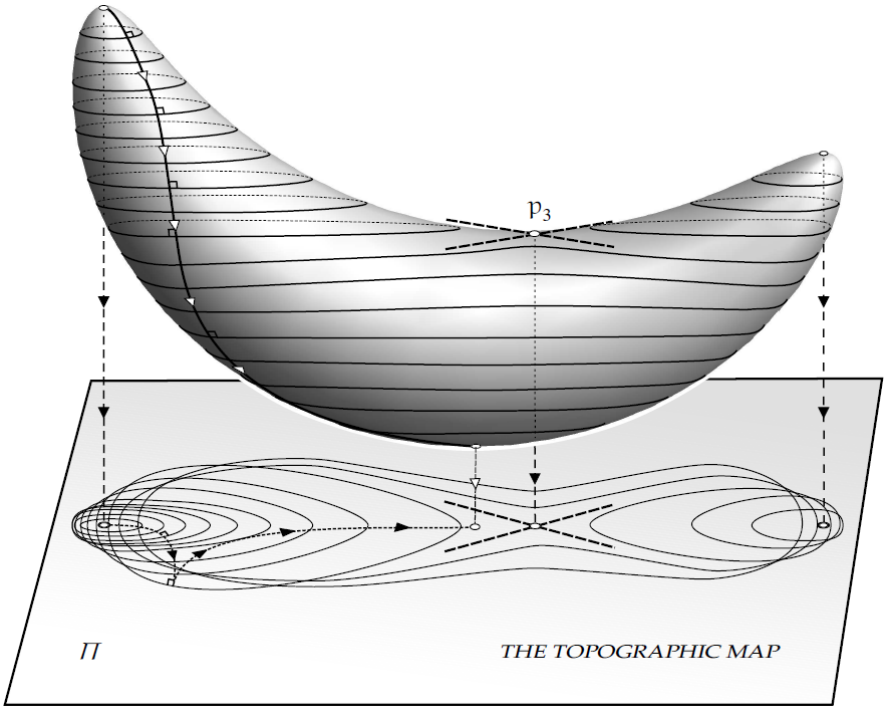
[19.5] **蜂蜜流**。 蜂蜜沿重力方向 $\mathbf{V}$ 下落到炸香蕉的曲面 $\mathcal{S}$ 上。 然后蜂蜜沿曲面往下流，在 $\mathcal{S}$ 上产生“蜂蜜流”速度向量场 $\mathbf{v}$。当 $\mathcal{S}$ 的外法线 $\mathbf{n}$ 为 $\mathbf{n} = -\mathbf{V}$（在 $\mathrm{p}_{1,2,3}$）时，就会出现 $\mathbf{v}$ 的奇点，或 $\mathbf{n} = +\mathbf{V}$（在 p_4）。 视觉直觉表明 p_1 和 p_2 是源，$\mathcal{I}(\mathrm{p}_1) = \mathcal{I}(\mathrm{p}_2) = +1$，$\mathrm{p}_3$ 是鞍点，$\mathcal{I}(\mathrm{p}_3) = -1$，$\mathrm{p}_4$ 是汇，$\mathcal{I}(\mathrm{p}4) = +1$。 这些直觉被精确的定义（19.5）所证实。

虽然这个雨流 $\mathbf{v}(\mathrm{p})$ 只是 $\mathcal{S}$ 上向量场的一个例子，但我们很快就会看到这个特定的向量场是一个强大的理论工具，也是 GGB 新证明的关键。

图 [19.5] 说明了在特定曲面上的这种流，引力的方向是 $\mathbf{V}$。实际上，流过这个曲面的雨水很可能会被拉离曲面，直接掉落到地上，即，一旦切平面变为垂直，包含 $\mathbf{V}$。

为了让物理直觉继续为数学直觉提供信息，我们必然确保流体附着在 $\mathcal{S}$ 上并始终流过它。 因此，我们创造了一个圣经奇迹：我们将水变成蜂蜜！ （通过附带的奇迹，让我们也将 [19.5] 变成炸香蕉！）

于是，流 $\mathbf{v}$（由空间中的力 $\mathbf{V}$ 将这种蜂蜜拖过曲面而产生）此后将被称为与 $\mathbf{V}$ 相关的**蜂蜜流向量场**。（警告：你可能会震惊地了解到“蜂蜜流”（还）不是标准的数学术语！此外，我们只会向物理学表示诗意的敬意，忠实地跟踪蜂蜜流的方向，而不是它的速度。）



[19.6] 曲面的**地形图**是通过将其与等距的水平面相交，然后将所得的**水平曲线**垂直向下投影到地图平面 Π 上而获得的。 （我们没有绘制曲面的下半部分，以避免混乱。）通过鞍点 p_3 的水平曲线是一个数字 8，其中只画出了它在 p_3 处的切线（虚线）。 曲面上蜂蜜流的流线与水平曲线正交，它们在投影到 Π 后保持正交（尽管地图不是共形的）。

19.5.2蜂蜜流与地形图的关系

我们将在第 19.7 节中给出蜂蜜流 **v** 的简明、纯数学定义，但现在，让我们继续物理思考，以获得新的见解。 为此，请回想一下，我们可以使用地理学中使用的平面地形图在 [19.5] 中表示炸香蕉的形状，或任何其它曲面。

为了构建这张地图，我们取了许多等距的水平面，并观察这些平面与其相交的曲面上的曲线，即所谓的水平曲线。 最后，我们将这些水平曲线垂直向下投影到下方的水平地图平面 Π 上，从而创建地形图。

这个过程在图 [19.6]中有说明，但只针对炸香蕉的上部；下半部分的地图会使图混乱，因此最好在单独的图中画出（未显示）。 如果假设水平曲线存在于透明平面上，那么地形图就是你从高处直接向下看时所看到的。 在图 [19.6]中的香蕉的情况下，你可以通过拿一把长刀将香蕉切成薄片，然后检查每片的形状来实验性地验证地图。

注意，当相交平面通过鞍点 p_3 时，水平曲线为 8 字形（未画出）； 该图确实画出了 8 字形经过 p_3 时的切线（虚线）。 回想一下，这些是曲面的渐近方向，$\kappa = 0$，并且鞍点附近的双曲线模式是那里的迪潘标形。 请参见第 162 页的第 15.8 节。

另请注意，水平曲线最靠近的地方，香蕉的曲面最陡峭，而它们相距最远的地方，曲面最浅。 沿水平曲线移动对应于在曲面上以恒定高度侧向移动。 因此，与水平曲线成直角移动对应于沿曲面上最陡下降（或上升）的方向旅行。 地图中（无限小分离的）相邻曲线相交的点对应于单个点上方的不同高度，即垂直切平面。

由于将蜂蜜拉下曲面的重力没有水平分量（沿着水平曲线），我们立即推断曲面上的蜂蜜流沿着水平曲线的正交轨迹。

这对地形图有影响。 由于 $\mathbf{v}$ 与水平曲线的切线 $\mathcal{T}$ 正交，因此它必然位于与 $\mathcal{T}$ 正交的平面内。但由于 $\mathcal{T}$ 是水平的（按构造），因此该正交平面是垂直的。 由于通过 $\mathcal{T}$ 和 $\mathbf{v}$ 的垂直平面是正交的，因此它们在正交方向上与水平地图平面 Π 相交。 所以，如 [19.6] 所示，

在曲面和地形图中，蜂蜜流的流线是水平曲线和地形图的正交轨迹。

请注意，从曲面向下到地图的垂直投影不是共形的：通常，地形图中没有忠实地表示曲面上的角度。 然而，尽管存在这种不共形性，我们刚刚证明了水平曲线和蜂蜜流的流线之间的直角是保持不变的。

如果两个平面向量场是正交的，那么当我们穿过一条曲线时，它们必然旋转相同的量。因此，两个场的公共奇点必然具有相同的指数。 例如，在地形图中，炸香蕉任一峰附近的水平曲线看起来像圆心向量场的流线，其中 $\mathcal{I} = +1$。 由于蜂蜜流的流线与水平曲线正交，因此（投影的）蜂蜜流也必然具有 $\mathcal{I} = +1$，就像它一样。

同样，在炸香蕉的鞍点附近，水平曲线看起来像鞍点的流线，$\mathcal{I} = -1$。 因此，蜂蜜流也有一个鞍点，$\mathcal{I} = -1$； 回想一下，这就是它被称为鞍点的原因！

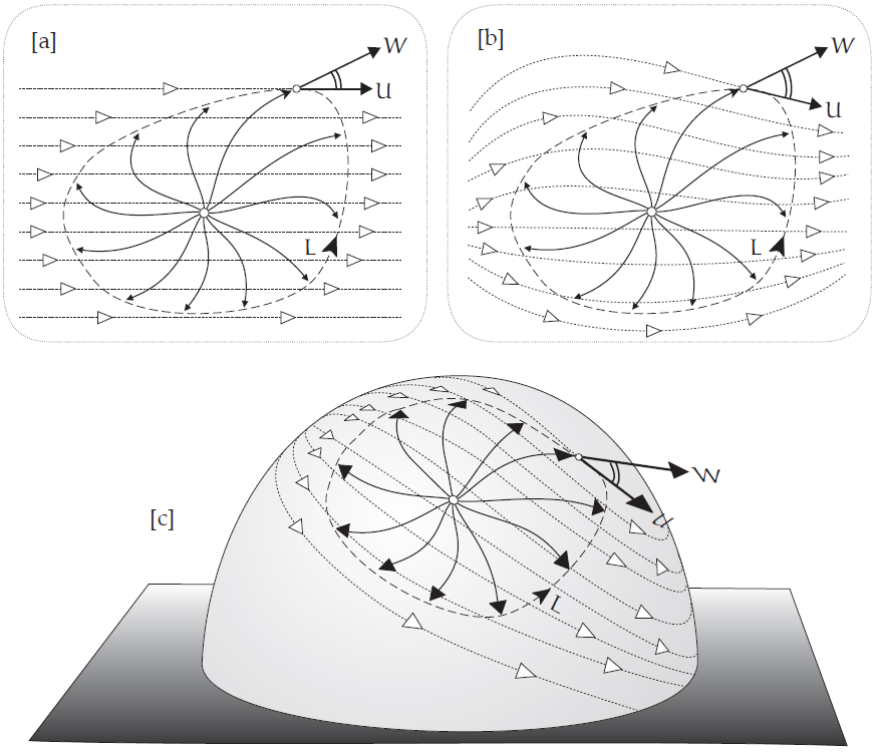
在蜂蜜流的情况下，我们因此有一种简单的方法可以通过曲面的地形图为其每个奇点分配一个“指数”。 我们现在直接在曲面上对指数的概念进行更一般的分析。

19.5.3定义曲面上的指数

给定曲面上的任何向量场，不一定是蜂蜜流，从直觉上看，显然必然可以将我们对指数的定义扩展到其奇点。

例如，在 [19.5] 中，我们立即确定了蜂蜜流的四个奇点：p_1、p_2、p_3 和 p_4。 这些点是外曲面法线的方向与 $-\mathbf{V}$ 重合，如 $\mathrm{p}_{1,2,3}$，或者与 $+\mathbf{V}$ 重合，如 p_4。 显然，$\mathcal{I}$ 的“正确”概括应该是 p_1 和 p_2 是源，$\mathcal{I}(\mathrm{p1}) = \mathcal{I}(\mathrm{p2}) = +1$，$\mathrm{p}_3$ 是鞍点，$\mathcal{I}(\mathrm{p}_3) = -1$，$\mathrm{p}_4$ 是汇，其中 $\mathcal{I}(\mathrm{p}_4) = +1$。

为了给这个新的、广义的“指数”概念赋予精确的含义，大概我们应该围绕曲面上的奇点绘制一个环路，然后在遍历环路时找到向量场的净旋转。 但是等等，相对于什么旋转？



[19.7] **在曲面上定义指数**。 我们通常测量平面向量场 W 相对于水平基准（或参考）场 U 的旋转，如 [a] 中所示。 但是我们可以改为测量相对于任何其他没有奇点的场的旋转，例如 [b] 中的场。 最后，在 [c] 中，我们推广到曲面：在包含奇点的区域上画出正则流 U； 那么指数是 W 相对于 U 的净旋转。

为了回答这个问题，我们首先重新审视熟悉的平面旋转概念。 图 [19.7a] 说明平面向量场 W (z) 沿环路 L 的旋转可以被认为是相对于具有水平流线的基准（或参考）向量场 U 发生的，比如 U (z) = 1 . 如果我们定义 ∠UW 为从 U 到 W 的角度，并令 $\delta_L(\angle UW)$ 为该角度沿 L 的净变化，那么我们旧的指数定义（19.3）可以写为

$$\boxed{\mathcal{I}_W(s) = \frac{1}{2\pi}\delta_L(\angle UW)} \tag{19.5}$$

如果我们使[19.7a]中U的直的水平流线连续地变形以产生[19.7b]中的弯曲流线，那么，根据通常的推理，（19.5）的右侧不会改变。 因此，我们得出结论，如果我们用在 L 上和内部非奇异的任何向量场替换 U，则该公式会产生正确的指数值。

接下来，假设 [19.7b] 绘制在橡胶布上。 如果我们继续将其拉伸成[19.7c]中的曲面形式，那么不仅（19.5）的右侧保持良好定义，而且其值不会改变。

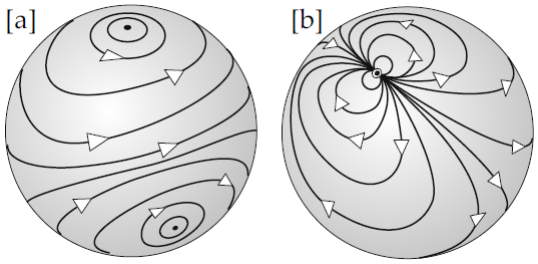
总而言之，如果 s 是曲面 $\mathcal{S}$ 上向量场 W 的奇点，我们定义其指数如下。 在覆盖 s

但没有其他奇点的 $\mathcal{S}$ 的曲面片上绘制任何非奇异向量场 U； 在这个曲面片上，绕 s 画一个简单环路 L； 最后，应用（19.5）——也就是说，当我们沿 L 环绕时，计算 W 相对于 U 的净转数。

19.6 庞加莱-霍普夫定理

19.6.1示例：拓扑球面

图 [19.8] 画出了球面上两种可能流的流线。 请注意，两者都有奇点：[19.8a] 有两个圆心，而 [19.8b] 有一个偶极子。 事实上，球面上没有全无奇点的向量场。



[19.8] [a] 指数和为 2 的两个圆心。 [b] 指数为 2 的偶极子。

这只是一个极其美丽而重要的结果的一个推论，称为**庞加莱–霍普夫定理**，我们将在稍后说明。

现在，为了初步了解结果，请注意，如果我们对 [19.8a] 中的所有指数求和，我们会得到

$$\mathcal{I}(\text{圆心}) + \mathcal{I}(\text{圆心}) = 1 + 1 = 2$$

而如果我们对 [19.8b] 做同样的事情，我们会得到

$$\mathcal{I}(\text{偶极子}) = 2$$

接下来，重新考虑 [19.5] 中描述的蜂蜜流的四个奇点。 这里指数的总和是 $(1) + (1) + (-1) + (1) = 2$，再一次！

尝试在桔子上绘制自己的流线。 例如，考虑从北极沿子午线到南极的流。 对两极处奇点的指数求和，

$$\mathcal{I}(\text{源}) + \mathcal{I}(\text{汇}) = 1 + 1 = 2$$

又一次！ 或许这一切都是些奇怪的巧合？

数学没有巧合！ 在球面的情况下，庞加莱-霍普夫定理指出，如果我们将其表面上任何向量场的指数相加，我们总是会得到 2 的答案。 事实上，它说我们将在拓扑上为球面的任

何曲面上得到这个答案，例如 [19.5] 中的炸香蕉。

以下是辉煌的普遍结果：

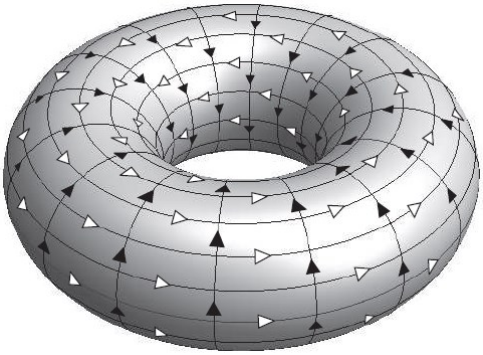
庞加莱–霍普夫定理：如果在亏格为 g 的光滑曲面 $\mathcal{S}_{\mathrm{g}}$ 上的向量场 $\mathbf{v}$ 只有有限数量的奇点 $\mathrm{p_i}$，那么它们的指数之和等于曲面的欧拉示性数：

(19.6)

$$\boxed{\sum_i \mathcal{I}_{\mathbf{v}}(\mathrm{p_i}) = \mathcal{X}(\mathcal{S}_{\mathrm{g}}) = 2 - 2g}$$

（19.6）的直接结果是没有任何奇点的向量场只能存在于欧拉示性数消失的曲面上，即拓扑甜甜圈。

即使这样，该定理实际上也不能保证这样的向量场一定存在，它只是要求拓扑甜甜圈上任何奇点的指数和为零。 然而，我们可以很容易地看到，在甜甜圈上确实存在没有任何奇点的向量场：[19.9] 说明了两个这样的场，一个用白色箭头表示，另一个用黑色箭头表示。



[19.9] 没有奇点的流只能存在于拓扑甜甜圈上。 这是圆环上两个这样的正则（拓扑不同）流：白色和黑色。

白色流中的所有流线在拓扑上都是等价的，因为任何一条都可以连续变形为另一条。黑流也是如此。 然而，白流和黑流在拓扑上是不同的：没有白色流线可以变形为黑色流线。

如果我们要将这两个流“相加”，我们可以想象第三个拓扑不同的正则流，每个流线绕甜甜圈一周，在再次连接之前穿过洞一次。

更一般地说，我们可以构建无限多的、拓扑不同的正则流，这样每个闭合流线围绕对称轴绕 m（白色）圈，并穿过孔绕 n（黑色）圈，然后自行闭合。通过将一个流线的末端连接到另一个流线的开始，可以将这样一个 (m,n)-流和一个 $(\mathrm{m}',\mathrm{n}')$-流“相加”，产生 $(\mathrm{m}+\mathrm{m}',\mathrm{n}+\mathrm{n}')$-流。这个想法很自然地导致了拓扑学家所说的曲面基本群，它是庞加莱在 1895 年发现的。但是我们离题了！

19.6.2庞加莱-霍普夫定理的证明

由于海因茨·霍普夫（Heinz Hopf）本人（见 Hopf 1956，第 13 页），我们现在可以给出定理 (19.6) 的一个非常优雅的推导。 论证分两步进行。 首先，我们证明在给定亏格的曲面上，所有向量场对其指数的总和产生相同的值。 其次，我们生成一个向量场的具体例子[146]，其总和等于欧拉示性数。 这就证明了结果。

考虑[19.10]中的曲面 $\mathcal{S}$，假设 $\mathbf{X}$ 和 $\mathbf{Y}$ 是 $\mathcal{S}$ 上的两个不同的向量场；为了避免混乱，我们只在一个点上画出了这些场。 如果 $\bullet$ 是 $\mathbf{X}$ 的奇点，$\odot$ 是 $\mathbf{Y}$ 的奇点，我们必然证明

$$\sum_{\bullet} \mathcal{I}_{\mathbf{X}}[\bullet] = \sum_{\odot} \mathcal{I}_{\mathbf{Y}}[\odot]$$

我们首先将 $\mathcal{S}$ 划分为曲线多边形(虚线)，使得每个多边形最多包含一个 $\bullet$ 和一个 $\odot$。

现在只关注这些多边形中的一个（深色阴影）及其边界 $\mathrm{K_j}$（从 $\mathcal{S}$ 外部逆时针看）。要找到 $\mathrm{K_j}$ 内 $\mathbf{X}$ 和 $\mathbf{Y}$ 的奇点的指数，画出任何非奇异基准向量场 U（再次仅画出在多边形上的单个点），然后使用（19.5）。 这些指数之间的差异是

$$\begin{aligned}\mathcal{I}_{\mathbf{Y}}[\mathrm{K_j}] - \mathcal{I}_{\mathbf{X}}[\mathrm{K_j}] &= \frac{1}{2\pi}\left[\delta_{\mathrm{K_j}}(\angle \mathrm{UY}) - \delta_{\mathrm{K_j}}(\angle \mathrm{UX})\right] \\ &= \frac{1}{2\pi}\delta_{\mathrm{K_j}}(\angle \mathrm{XY})\end{aligned}$$

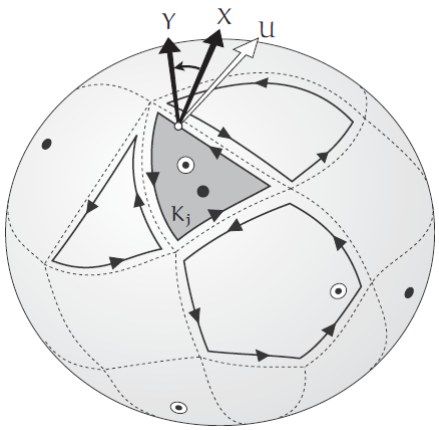
它明确独立于局部基准向量场 U； 它仅取决于从 $\mathbf{X}$ 到 $\mathbf{Y}$ 的图示（带符号）角度。

最后，由此我们推断出

$$\begin{aligned}&\sum_{\odot} \mathcal{I}_{\mathbf{Y}}[\odot] - \sum_{\bullet} \mathcal{I}_{\mathbf{X}}[\bullet] \\ &= \sum_{\mathrm{j}} \left(\mathcal{I}_{\mathbf{Y}}[\mathrm{K_j}] - \mathcal{I}_{\mathbf{X}}[\mathrm{K_j}]\right) \\ &= \frac{1}{2\pi} \sum_{\text{所有多边形}} \delta_{\mathrm{K_j}}(\angle \mathrm{XY}) \\ &= 0\end{aligned}$$

因为（见（19.2））每个多边形的每条边在每个方向上都经过一次，在（带符号的）角 $\angle$XY上产生相等和相反的变化。我们就这样完成了第一步：指数的总和独立于向量场。

[146] 实际上，我们将给出一个与上述参考文献（Hopf 1956）中 Hopf 给出的不同的例子。



[19.10] K_j 内的 **X** 的奇点（标 •）和 **Y** 的奇点（标⊙）的指数差仅取决于 **Y** 相对于 **X** 的旋转。但 K_j 的每条边都邻接相邻多边形的相反方向的边，因此所有多边形上的这些旋转的总和必然消失。

由于 [19.8] 中示例的指数和为 2，我们现在知道这是拓扑球面上每个向量场的值。一般论证的第二步同样是在任意亏格为 g 的曲面 $\mathcal{S}_g$ 上产生一个具体的例子，使得总和为 $\mathcal{X}(\mathcal{S}_g) = (2-2g)$。图 [19.11] 就是这样一个例子（这里 $g=3$）——即我们的老朋友蜂蜜流向量场。如图所示，顶部的源和底部的汇各有 $\mathcal{I}=+1$，g 个孔每个有两个鞍点，每个鞍点 $\mathcal{I}=-1$。因此，该特定流的指数总和确实是 $2-2g = \mathcal{X}(\mathcal{S}_g)$，因此这必然是 $\mathcal{S}_g$ 上每个流的指数总和。证毕。

19.6.3应用：欧拉-惠里尔公式的证明

庞加莱-霍普夫定理为欧拉-惠里尔公式(18.7)提供了惊人的直接证明。回想后者指出，如果我们将亏格为 g 的曲面 $\mathcal{S}_g$ 划分为多边形，具有 V 个顶点、E 个边和 F 个面，那么

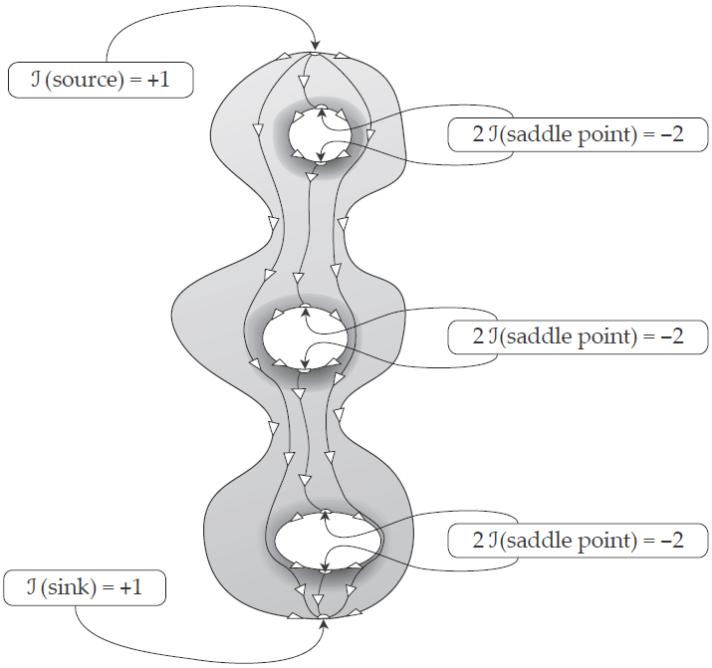
$$\boxed{V - E + F = 2 - 2g}$$

正如 [19.12] 所示，我们可以在 $\mathcal{S}_g$ 上构建一个一致的 **Stiefel 向量场**[147]，如下所示：在每个顶点（V 个）放置一个源（标记 ），在每条边（E 个）放置一个鞍点（标记⊙），以及在每个面（F 个）内放置一个汇（标记 •）。将庞加莱-霍普夫定理应用到我们构建的向量场中，我们立即推导出

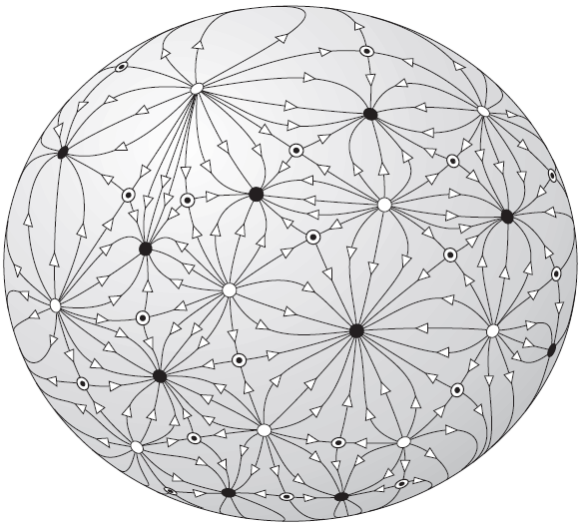
$$2-2g = \sum \mathcal{I} = V\,\mathcal{I}(\ ) + E\,\mathcal{I}(\odot) + F\,\mathcal{I}(\bullet)$$

完毕！

[147] 根据 Frankel (2012, §16.2b)，这是由于 Hopf 的学生 Eduard Stiefel (1909－1978)。



[19.11] 亏格为 g 曲面上的蜂蜜流向量场。 源（顶部）和汇（底部）每个都对指数总和贡献 $+1$，g 个孔每个都贡献 -2。 因此指数和为 $2-2g = \mathcal{X}$。



[19.12] 给定亏格为 g 的曲面 $\mathcal{S}_g$ 的任何多边形划分，我们可以在 $\mathcal{S}_g$ 上每个顶点（V 个）放置一个源（ ）——想想白洞！每条边（E 个）放置一个鞍点（⊙）；每个面（F 个）内放置一个汇（•）。 因此 $\sum \mathcal{I} = \mathrm{V} - \mathrm{E} + \mathrm{F}$。

19.6.4庞加莱微分方程与霍普夫线场

如上所述，二维庞加莱-霍普夫定理实际上是由庞加莱发现的。 然而，附加霍普夫的名字是正确的，因为他在两个不同的方向上极大地扩展了结果。 第一个方向是将结果推广到 n 维封闭流形上的向量场； 可悲的是，解释这一点会让我们走得太远，但请参阅延伸阅读。

然而，霍普夫扩展结果的第二个方向是相当初级的，我们现在将解释它。

庞加莱在微分方程定性理论方面的开创性工作导致他研究了向量场的拓扑行为。 这样的方程在 1687 年成为物理学的核心，当时牛顿宣布了他的第二运动定律：如果一个位置向量为 $\mathbf{X}$、质量为 m 的点粒子受到一个向量力 $\mathbf{F}$ 的作用，那么它通过在力的方向上加速来做出反应，根据该定律：

$$\ddot{\mathbf{X}} = \text{加速} = \frac{\mathbf{1}}{\mathrm{m}}\mathbf{F}$$

牛顿制定这条定律的动机是确定行星的轨道——现在称为天体力学的一部分。 正如我们在序言中所讨论的那样，牛顿在晚年避开了他自己年轻时在符号微积分方面的发现，而在《原理》中，鉴于太阳引力的已知（平方反比）定律，他转而采用优雅的几何推理来确定行星的轨道。

然而，从更传统的、现代的观点来看，（19.7）只是一个二阶微分方程，可以通过积分两次来求解（原则上）：一次是求速度 $\dot{\mathbf{X}}$，第二次是求轨道 $\mathbf{X}$ 本身。 在只有两个引力体（如地球和太阳）的情况下，牛顿找到了精确解。 通过引入月球作为第三个天体，牛顿能够解释地球上的潮汐，通过近似，他能够预测潮汐的变化。 但是牛顿和他的继任者无法找到这个三体问题的确切解。

最终，200 多年后，庞加莱——如图[19.13]——通过证明三体问题是不可解的[148]，取得了决定性的进展。 当然，如果我们试图分析整个太阳系，情况只会变得更糟，其中每颗行星不仅被太阳保持在其轨道上，而且还吸引着其他所有行星，并且它们相互决定了它们的集体轨道。 这就是所谓的 n 体问题。

庞加莱回归几何方法是自牛顿以来天体力学取得最大进步这种说法是恰当的（也许有些讽刺）。 在三卷（1892 年、1893 年和 1899 年出版）中，庞加莱发表了他的《天体力学新方法》[149]，其中他不再寻求明确求解微分方程（他已经证明这是不可能的），而是确定这些方程解的定性行为。

[148] 要了解我们所说的“不可解”的含义，以及对庞加莱发现之旅的生动描述，请参见 Diacu 和 Holmes (1996)。

[149] Poincaré (1899).

[19.13] 亨利 • 庞加莱 （1854 – 1912）

为了了解这与之前发生的事情有何联系，让我们像庞家莱所做的那样考虑 (x,y) 平面中的一阶微分方程，形式为

$$\frac{dy}{dx} = -\frac{\mathrm{P}(\mathrm{x},\mathrm{y})}{\mathrm{Q}(\mathrm{x},\mathrm{y})} \tag{19.8}$$

庞加莱实际上将这些写成

$$\mathrm{P}(\mathrm{x},\mathrm{y})\, dx + \mathrm{Q}(\mathrm{x},\mathrm{y})\, dy = 0$$

要研究积分曲线（“解），请考虑沿积分曲线的无穷小向量 $\begin{bmatrix} dx \\ dy \end{bmatrix}$，并注意前面的方程意味着

$$\begin{bmatrix} \mathrm{P} \\ \mathrm{Q} \end{bmatrix} \cdot \begin{bmatrix} dx \\ dy \end{bmatrix} = 0$$

换句话说，

（19. 8）的积分曲线处处与向量场 $\begin{bmatrix} \mathrm{P} \\ \mathrm{Q} \end{bmatrix}$ 正交。

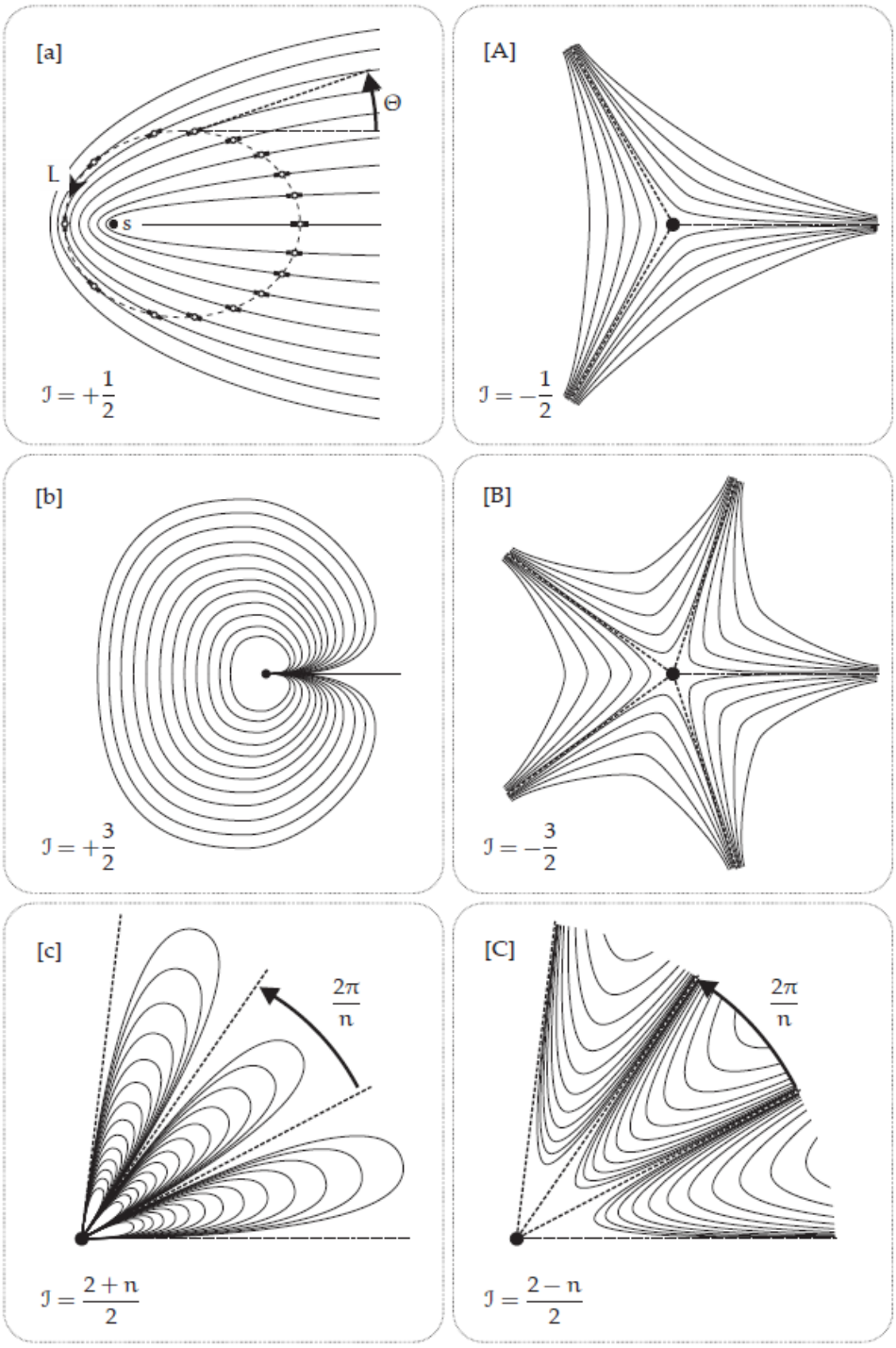
例如，考虑径向向量场 $\begin{bmatrix} \mathrm{P} \\ \mathrm{Q} \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$； 这是 [19.4c] 中的源。 显然，正交积分曲线是以原点为中心的圆，所以我们得到了 [19.4e] 中所示的圆心。 同样见 [练习]，$\begin{bmatrix} y \\ x \end{bmatrix}$ 在 [19.4h] 中

产生鞍点，$\begin{bmatrix} y \\ -x \end{bmatrix}$ 并产生源或汇（[19.4c, f]）。

现在让我们转向霍普夫对庞加莱发现的扩展。 除了推广到需要全新概念的 n 维向量场之外，霍普夫还意识到即使在二维中也存在我们可能称之为“无方向流”的东西，它不是由庞加莱考虑的那种微分方程产生的。

对于我们的第一个例子，考虑 [19.14a]。 我们将这种模式称为“无方向流”，但霍普夫实际上将这个概念命名为线元素场（field of line elements）。 然而，我们将采用的现代术语是线场（line field），类似于向量场。

如你所见，这些“流线”没有附加箭头，但我们仍然可以为切线指定一个连续变化的角度 Θ。 如果我们绕奇点 s 沿着一个简单环路 L 行进，则图示的线元素必然返回到它的原始位置。 以前，对于向量场，向量必然返回原点，指向与离开时相同的方向，因此它必然经历完整的旋转次数：旋转必然是 2π 的倍数，该倍数就是指数。

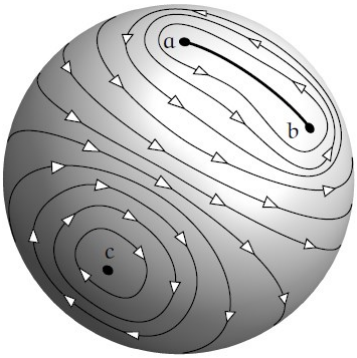


[19.14] **线场**没有与其“流线”相关联的方向。 因此，它们的奇点可以具有向量场不能具有的分数指数。 通过重复 [c] 和 [C] 中的模式，我们可以分别构造任意正或负指数的奇点。 如果正整数 n 是奇数，则指数是一个分数，但如果 n 是偶数，那么我们恢复 [19.4] 中所示的向量场的流，其中指数是整数。

但是，如果线元素旋转 π 的任意倍数，则它会返回其原始方向。 事实上，在 [19.14a] 中我们看到净旋转是π。 如果我们保留指数的原始定义（19.3），即旋转除以 2π，那么我们会看到这个奇点有一个分数指数：$\mathcal{I} = +\frac{1}{2}$。 一些作者将这种分数指数称为霍普夫指数。

请查看 [19.14] 中所有剩余的数字并直观地确认它们的指数。 特别是，[19.14c] 和 [19.14C] 分别展示了如何构建任意正或负指数的奇点，只需重复所示的模式，直到它们完全包围奇点。 如果 n 是奇数，则指数是一个分数，但如果 n 是偶数，那么我们恢复 [19.4] 中所示的向量场的流，其中指数是整数。

虽然我们不能将具有分数指数的线场解释为畅通无阻的流线，但我们可以将其解释为*存在一维障碍的流*。 例如，在 [19.14a] 中，如果我们将 s 右侧的水平射线视为这样的障碍，那么我们可以将箭头附加到曲线上，这种方式与围绕该障碍的流动一致，流向比如说，流向障碍物上方的左侧，然后流向障碍物下方的右侧。



[19.15] 通过安装连接 a 和 b 的障碍物，我们可以创建一个环绕这两个点的流，并且在圆心 c 处有第三个奇点。 尽管在 a 和 b 处有分数指数，但指数和仍然等于欧拉示性数。

从这个角度来看，障碍物的每个点都是向量场的一个奇点，因为当我们越过障碍时，向量的方向会不连续地翻转。 然而，对于图示的无方向线场，每个这样的点 p（除了 s）都是一个*正则*（regular）点。 因为如果 K 是围绕 p 的一个小环路，那么你可以看到 Θ 在 K 上不断变化，因此 $\mathcal{I}(\mathrm{p}) = \frac{1}{2\pi}\delta_{\mathrm{K}}\Theta = 0$。

现在回想一下以庞加莱-霍普夫定理告终的推理，关于向量场的奇点。 你会发现它同样*适用于这些更一般的线场*。 简而言之，

庞加莱–霍普夫定理（19.6）不仅适用于向量场，也适用于带分数指数的线场。

(19.9)

图 [19.15] 是球面上这个结果的一个例子。 在这里，我们看到一个包含三个奇点 a、b 和 c 的线场。 我们已经想象连接 a 和 b 的曲线是一个障碍物，我们在曲线上*添加了箭头*以创建围绕这个障碍物环绕的一致流。 注意，a 和 b 附近的流是 [19.14a] 中所示的流，其中 $\mathcal{I} = +(1/2)$。

庞家莱-霍普夫定理 (19.6) 预测球体上的指数之和应该是 $\mathcal{X} = 2$，实际上它是：

$$\mathcal{I}(\mathrm{a})+\mathcal{I}(\mathrm{b})+\mathcal{I}(\mathrm{c})=\frac{1}{2}+\frac{1}{2}+1=2$$

最后，注意，如果我们缩小障碍的长度并允许两个 $\mathcal{I}=(1/2)$ 奇点 a 和 b 彼此接近，那么它们最终将合并为一个 $\mathcal{I}=1$ 的中心，从而创建图 [19.8a] 所示的流。

线场的历史和未来：尽管霍普夫对他的想法做了非常清晰的描述（霍普夫 1956），但我不知道有任何关于拓扑、微分拓扑或微分几何的现代介绍性课本甚至提到了这一引人入胜的概念。 事实上，我们发现的唯一例子现在已有 50 多年的历史：斯托克 (Stoker) 的经典课本 (1969, p. 244)，他是霍普夫的博士生。 尽管事实上这样的线场在数学中自然出现。 例如，围绕曲面上的脐点（其中 $\kappa_1=\kappa_2$）的曲率线通常采用 [19.14 a 或 A] 的形式。 有关案例 [a] 的可爱说明，请参阅 Hilbert (1952, p. 189)。

虽然数学家似乎很少关注庞加莱-霍普夫定理在分数指数线场方面的适用性，但物理学家并非如此，因为大自然自己已经在物理学的多个领域将这种线场强加于他们，特别是在光学方面。 本书末尾的延伸阅读部分旨在引导你了解霍普夫思想的这些许多引人入胜的新物理应用。

19.7 全局高斯-博内定理的向量场证明

设 V 是 $\mathbb{S}^2$ 上的任意点，单位位置向量为 $\mathbf{V}$。现在想象一个具有法线 $\mathbf{V}$ 的平面，最初位于远离光滑闭合表面 $\mathcal{S}$ 的位置，并且 $\mathbf{V}$ 指向远离 $\mathcal{S}$。接下来，想象平面向 $\mathcal{S}$ 移动。最终平面必然与 $\mathcal{S}$ 接触，并且在那个时刻它将成为 $\mathcal{S}$ 在接触点 p 处的切平面，因此 $\mathbf{V}$ 必然与那里的曲面法线重合：$\mathbf{V}=\mathrm{n}(\mathrm{p})$。 换句话说，V 是 p 的球面像：$\mathrm{V}=\mathrm{n}(\mathrm{p})$。 由于这对于任意 $\mathbf{V}$ 是正确的，因此对于相反的向量 $-\mathbf{V}$ 也必然是正确的； 平面只是从 $\mathcal{S}$ 的另一边接近。总而言之，

> 如果 $\mathcal{S}$ 是一个任意的、封闭的、光滑的曲面，那么它的球面像至少覆盖 $\mathbb{S}^2$ 的每对对径点 $\pm\mathrm{V}$ 一次。 (19.10)

当然，V 和$-$V 可以各自被覆盖不止一次。 例如，在 [19.5] 中，我们看到 V 被覆盖了一次，而 $-$V 被覆盖了三次。

现在让我们回到 GGB，让我们简要回顾一下第 17 章中的第一个“启发式”证明。从总曲率 $\mathcal{K}(\mathcal{S})$ 的角度来看 V 的这些覆盖。 如果在 p 处 $\mathcal{K}(\mathrm{p})>0$，则 n 保留 p 周围小块面积 $\delta\mathcal{A}$ 的方向，所以它在 $\mathbb{S}^2$ 上围绕 V 的面积 $\delta\tilde{\mathcal{A}}\asymp\mathcal{K}(\mathrm{p})\,\delta\mathcal{A}$ 的球面像有相同的方向。 正如我们在 17.4 节中讨论的那样，我们将其视为 V 的正覆盖。同样，如果 $\mathcal{K}(\mathrm{p})<0$，则球面映射反转方向，并且覆盖计为负。 总曲率积分会自动以代数方式计算这些覆盖，同时考虑方向。 如前所述，令 $\mathcal{P}(\mathrm{V})$ 和 $\mathcal{N}(\mathrm{V})$ 分别表示 V 的正覆盖和负覆盖数，因此 V 的净覆盖数为 $[\mathcal{P}(\mathrm{V})-\mathcal{N}(\mathrm{V})]$。 因此，

令 V 为 $\mathbb{S}^2$ 上的一个点，并令其原像（在球面映射 n 下）为 $\mathcal{S}$ 的点 $\mathrm{p_i}$。考虑 V 周围的一个小（最终消失）块，并让 $\mathrm{p_i}$ 周围的小原像块的面积为 $\delta\mathcal{A}_\mathrm{i}$。 然后，

$$\boxed{\delta\mathcal{A}_{\mathrm{i}}\text{内的总曲率} \asymp \sum_i \mathcal{K}(\mathrm{p_i})\,\delta\mathcal{A}_{\mathrm{i}} \asymp [\mathcal{P}(\mathrm{V}) - \mathcal{N}(\mathrm{V})]\delta\tilde{\mathcal{A}}} \tag{19.11}$$

例如，在 [19.5] 中，我们看到 $\mathrm{p_1}$ 和 $\mathrm{p_2}$ 处的曲率是正的，并且源自这些点的 $-\mathrm{V}$ 的覆盖是正的。 另一方面，包含 $\mathrm{p_3}$ 的负弯曲贴片被映射到 $-\mathrm{V}$ 的方向反转的负覆盖。 因此，$-\mathrm{V}$ 的覆盖物的净代数数为 1。同样，V 的覆盖物的净数为 1。

回想一下，$\mathbb{S}^2$ 的这个净覆盖数是球面映射的拓扑度 $\deg(\mathrm{n})$，基本事实是 $\mathbb{S}^2$ 的所有[150]点的度都相同。 这是我们在第 17 章中对 GGB 的启发式证明的本质。该证明的最后一步是认识到度数仅取决于 $\mathcal{S}_{\mathrm{g}}$ 的亏格 g。这是 （16.7），在第 174 页，我们在此重申：

$$\boxed{\deg\left[\mathrm{n}\left(\mathcal{S}_{\mathrm{g}}\right)\right] = \mathcal{P} - \mathcal{N} = (1-\mathrm{g}) = \frac{1}{2}\mathcal{X}(\mathcal{S}_{\mathrm{g}})} \tag{19.12}$$

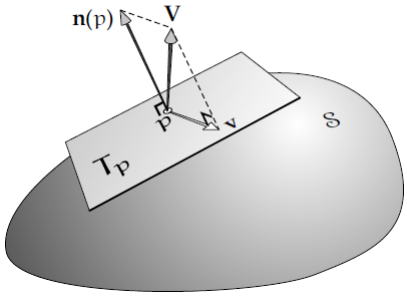
正如我们在 16.6 节中提到的，这是由沃尔特 • 戴克（Walther Dyck）在 1888 年首次证明的。

如果我们假定拓扑度存在，因此 $\mathbb{S}^2$ 的每个点都被覆盖相同的（代数）次数，那么庞家莱-霍普夫定理可用于给出戴克结果的优雅证明。 事实上，这就是吉列明（Guillemin）和波拉克（Pollack）(1974, p. 198) 等著名著作中证明 GGB 的方法。

我们也将很快重现该论点。 然而，拓扑度的存在绝不是显而易见的，所以我们现在提出一个不依赖于它的略有不同的论证。 该论证仍然依赖于庞加莱-霍普夫定理，但并不假设拓扑度的存在。 也就是说，如果 V 和 W 是 $\mathbb{S}^2$ 的两个点，我们不会假设它们必须被覆盖相同的次数（代数）； 相反，我们至少应考虑 $\deg[\mathrm{n}(\mathcal{S}),\mathrm{V}] \neq \deg[\mathrm{n}(\mathcal{S}),\mathrm{W}]$ 的可能性。从今以后，我们将理解为我们正在处理 $\mathcal{S}$ 的球面映射，因此这种（假设的和实际上不可能的）不等式可以缩写为 $\deg[\mathrm{V}] \neq \deg[\mathrm{W}]$。

为了在不假设拓扑度存在的情况下证明 GGB，重新考虑 [19.5] 中的蜂蜜流。 第一个重要的观察结果是，这种流的奇点恰好出现在垂直重力 $\mathbf{V}$ 在曲面内具有零分量的那些点上——也就是说，在垂直力向量直接指向（或指向）曲面的点上——也就是说，外表面法线 $\mathbf{n} = \pm\mathbf{V}$。 当然，从数学上讲，这个引力方向 $\mathbf{V}$ 并没有什么特别之处——我们可以想象引力向任意方向拉动； 或者，如果我们不想篡改地球的引力场，我们可以简单地旋转曲面！

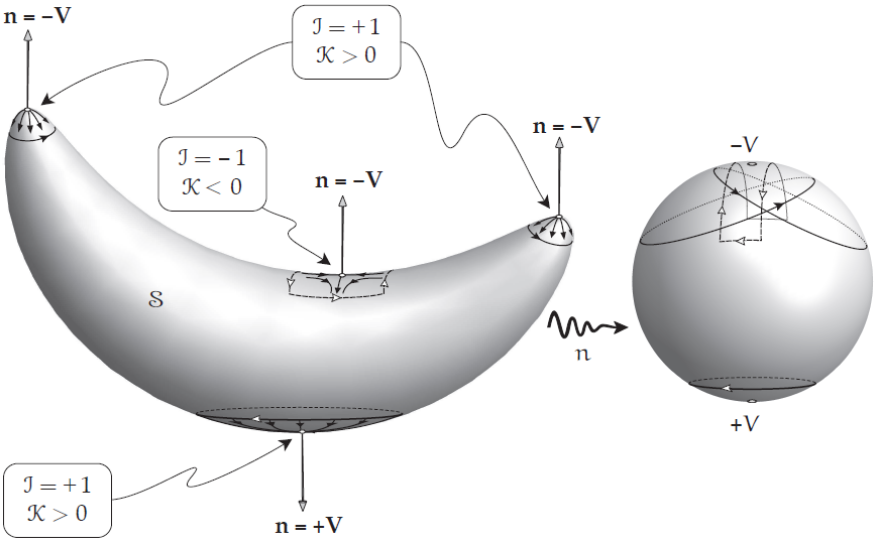
[150] 正如我们所讨论的，例外是曲率为零的点，它们对总曲率没有任何贡献。



[19.16] 曲面 $\mathcal{S}$ 上沿 **V** 方向的**蜂蜜流** **v** 是 **V** 到 p 处切平面 T_p 的正交投影。

更数学地说，我们的“蜂蜜流”向量场是通过将 **V** 正交投影到曲面上，或者更准确地说，到曲面的切平面 T_p 上来获得的。 见[19.16]。 让我们正式定义这个定义：

$$\boxed{\text{方向 V 上的蜂蜜流向量场为} \quad \boxed{v(p) \equiv \mathrm{proj}_{T_p}\mathbf{V}}} \tag{19.13}$$



[19.17] 方向 **V** 上 $\mathcal{S}$ 上蜂蜜流 **v** 的奇点由球面映射 n 映射到 $\mathbb{S}^2$ 上的 +V 或 −V，并且 $\mathbb{S}^2$ 上覆盖的方向由 $\mathcal{S}$ 上 **v** 的指数。这里 +V 被香蕉底部的水槽图像正面覆盖了一次 - 其中 $\mathcal{I} = +1$ 且 $\mathcal{K} > 0$。 另一方面，−V 被覆盖了三次：两次被两个源的图像正向覆盖——其中 $\mathcal{I} = +1$ 和 $\mathcal{K} > 0$——一次被鞍点的图像负向——其中 $\mathcal{I} = -1$ 和 $\mathcal{K} < 0$。

当 $n(p_i) = V$ 时出现奇点 $v(p_i) = 0$。 换句话说，

> 如果蜂蜜流过 $\mathcal{S}$，沿空间方向 **V** 拉动，蜂蜜流 **v** 的奇点由球面映射发送到 V 或 $\mathbb{S}^2$ 上的对径点 −V 的完整点集组成。

在曲面上所有可能的向量场中，为什么我们对蜂蜜流如此关注？ 答案是在曲面的几何

形状和蜂蜜流的拓扑结构之间存在着关键的联系。 这反过来将产生我们对 GGB 的第三种解释，作为庞加莱-霍普夫定理的结果。

重新考虑 [19.5] 中说明的蜂蜜流向量场 $\mathbf{v}$ 的奇点 $\mathrm{p_i}$。 这些在 [19.17] 中再次显示，但现在将注意力集中在球面映射 n 如何将它们发送到 $\mathbb{S}^2$。

决定性的观察是 v 具有正曲率的奇点和具有负曲率的奇点之间的区别。

回想 (12.9) 和 [12.7]，在第 136 页。如果 $\mathrm{K}(\mathrm{p_i}) > 0$，则曲面是局部朝上的圆顶（在这种情况下，它是具有 $\mathcal{I}(\mathrm{p_i}) = +1$ 的源）或直线向下（在这种情况下，它是一个汇，同样是 $\mathcal{I}(\mathrm{p_i}) = +1$）。 但是如果 $\mathrm{K}(\mathrm{p_i}) < 0$，则曲面局部是鞍点，在流动中产生相应的鞍点，在这种情况下 $\mathcal{I}(\mathrm{p_i}) = -1$。

这立即意味着如果 $\mathcal{I}(\mathrm{p_i}) = +1$，球形映射会保持方向，如果 $\mathcal{I}(\mathrm{p_i}) = -1$，则将其反转：

> 如果 p 是蜂蜜流 v 的奇点，那么如果 $\mathcal{I}(\mathrm{p}) = +1$，则其在球面映射 n 下的图像 (V) 为正覆盖，如果 $\mathcal{I}(\mathrm{p}) = -1$，则为负覆盖。

现在让我们将这个结果与庞家莱-霍普夫定理结合起来。 如果 $\mathrm{p_i}$ 是 $\mathbf{v}$ 在 $\mathcal{S}$ 上的奇点的集合，即那些球面像为 $+\mathrm{V}$ 或 $-\mathrm{V}$ 的点，则

$$\mathcal{X}(\mathcal{S}) = \sum_i \mathcal{I}_v(\mathrm{p_i}) = \{\mathcal{P}(+\mathrm{V}) - \mathcal{N}(+\mathrm{V})\} + \{\mathcal{P}(-\mathrm{V}) - \mathcal{N}(-\mathrm{V})\} \tag{19.14}$$

实际上，右边两个括号中的项彼此相等，是球面映射的拓扑度。 见（19.12）。 例如，在[19.17]中我们看到

$$\mathcal{P}(+\mathrm{V}) - \mathcal{N}(+\mathrm{V}) = 1 - 0 = 1$$
$$\text{且}$$
$$\mathcal{P}(-\mathrm{V}) - \mathcal{N}(-\mathrm{V}) = 2 - 1 = 1$$

然而，（19.14）可以在不假设这个事实的情况下证明 GGB。 为此，请注意单流 v 的奇点，对应于沿方向 $\mathbf{V}$ 流动的蜂蜜，实际上包括 $\mathbf{V}$ 和 $-\mathbf{V}$ 的奇点。 反转蜂蜜流的方向不会改变奇点的位置，也不会改变它们的指数。

现在，如果 V 仅在 $\mathbb{S}^2$ 的北半球漫游（其面积为 2π），则 $-\mathrm{V}$ 在南半球漫游，因此 V 覆盖整个 $\mathbb{S}^2$。 因此，由于（19.10），当 $\pm\mathrm{V}$ 仅在北半球漫游时，$\mathbf{v}$ 的奇点覆盖整个曲面 $\mathcal{S}$。 而且由于我们刚刚看到奇点的指数正确地计算了它们在下面的图像的覆盖球面图 (19.11) 和 (19.14) 产生了 GGB 的第三个证明，形式为

$$\boxed{\mathcal{K}(\mathcal{S}) = 2\pi\mathcal{X}(\mathcal{S})}$$

让我们回到拓扑度的话题，从而得到一个奖励结果。 如果我们假设拓扑度存在，那么 $\deg(\mathrm{V}) = \deg(\mathrm{W})$，对于 $\mathbb{S}^2$ 上的所有 V 和 W。特别地，它遵循 $\deg(+\mathrm{V}) = \deg(-\mathrm{V})$。在这种情况下，（19.14）变成，

$$\begin{aligned}\mathcal{X}(\mathcal{S}) &= \{\mathcal{P}(+\mathrm{V}) - \mathcal{N}(+\mathrm{V})\} + \{\mathcal{P}(-\mathrm{V}) - \mathcal{N}(-\mathrm{V})\} \\ &= \deg(+\mathrm{V}) + \deg(-\mathrm{V}) \\ &= 2\deg(\mathrm{V}) \\ &= 2\deg\end{aligned}$$

其中最终的等式反映了这样一个事实，即度数与 V 的选择无关。 因此，如果我们假设拓扑度数存在，我们将恢复戴克的结果，（19.12）：

$$\boxed{\deg = \frac{1}{2}\,\mathcal{X}(\mathcal{S})}$$

19.8 前路

迄今为止，GGB 的所有三个证明都依赖于将 $\mathcal{K}$ 解释为球面映射的局部区域扩展因子。这是一个外在的曲率概念，它依赖于法线 $\mathbf{n}$，它对 $\mathcal{S}$ 的居民来说是不可见和不可知的。虽然这三个证明在许多不同的方向上为我们提供了美妙的新见解，但还是缺乏一些东西，因为我们知道事实上，$\mathcal{K}$ 也是曲面的内蕴属性，它的居民可以知道和测量。 原则上，这些居民可以测量整个曲面的 $\mathcal{K}$，并且通过这些纯粹的局部几何测量，他们可以确定他们世界的拓扑结构！

我们即将开始第四幕，它引入了一种全新的、极其强大的、内蕴的理解和测量曲率的方法。 这将使我们能够最终解释迄今为止我们被迫假设的一些基本结果，例如局部高斯-博内定理 - 以及随之而来的绝妙定理 - 以及非凡的（“星际迷航移相器”）公式 （4.10），第 38 页，用于以度量表示的曲率。 此外，它将使我们能够弥补我们现有的三个 GGB 证明中提到的缺陷。 事实上，使用海因茨 •霍普夫的想法，我们最终将能够提供一个完全内蕴的 GGB 证明。

第二十章 第三幕练习

平面曲线的曲率

1. **曲率公式的计算证明**。图 [8.7] 提供了（8.7）的几何证明。用计算来证明这一点。（提示：如果曲线 $[x(t), y(t)]$ 在单位速度下被追踪，则 $\dot{x} = \cos\varphi$，并且 $\dot{y} = \sin\varphi$。）

3 维空间中的曲线

2. **副法线不能向运动方向倾斜**。计算证明当 Frenet 标架 $(\mathbf{T}, \mathbf{N}, \mathbf{B})$ 沿其曲线移动时，副法线 $\mathbf{B}$ 仅绕 $\mathbf{T}$ 旋转；它不能向 $\mathbf{T}$ 的方向倾斜。（提示：象征性地，必须证明 $\mathbf{B}'$ 在 $\mathbf{T}$ 方向上没有分量：$\mathbf{B}' \cdot \mathbf{T} = 0$。）

3. **达布向量**（Darboux Vector）。当 Frenet 标架沿其曲线移动时，它会旋转。计算验证其瞬时角速度向量为

$$\mathbf{A} \equiv \tau\mathbf{T} + \kappa\mathbf{B}$$

这个向量 $\mathbf{A}$ 有时被称为达布向量（参见 Stoker 1969，第 62 页），以纪念它的发现者让·加斯东·达布（Jean-Gaston Darboux (1842-1917)），微分几何的先驱和嘉当的老师。提示：用（9.3）证明

$$\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix}' = \mathbf{A} \times \begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix}$$

4. **变速 Frenet-Serret 方程**。Frenet-Serret 方程 (9.3) 假设导数是关于弧长的，如果以单位速度跟踪曲线，则该导数仅与时间导数相同。相反，假设粒子具有可变速度 $v = \dot{s}$。

(i) 证明 $[\Omega] \mapsto v[\Omega]$:

$$\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix}^{\cdot} = v\begin{bmatrix}0 & \kappa & 0\\ -\kappa & 0 & \tau\\ 0 & -\tau & 0\end{bmatrix}\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix}$$

(ii) 证明加速度为 $\dot{\mathbf{v}} = [v\mathbf{T}]^{\cdot} = \dot{v}\,\mathbf{T} + \kappa v^2\mathbf{N}$，并几何解释这两个项。

(iii) 证明 $\mathbf{B} = \frac{\mathbf{v}\times\dot{\mathbf{v}}}{|\mathbf{v}\times\dot{\mathbf{v}}|}$

(iv) 证明 $\kappa = \left|\frac{\mathbf{v}\times\dot{\mathbf{v}}}{v^3}\right|$

(v) 证明 $\tau = \frac{(\mathbf{v}\times\dot{\mathbf{v}})\cdot\ddot{\mathbf{v}}}{|\mathbf{v}\times\dot{\mathbf{v}}|^2}$

5. **螺旋**。 考虑在时间 t 的位置为 $(R\cos\omega t, R\sin\omega t, qt)$ 的粒子的路径。
   (i) 解释为什么这是一个螺旋，并说明 R、ω 和 q 的几何/物理解释。
   (ii) 证明 $v = \sqrt{(R\omega)^2 + q^2}$，并在几何学/物理上解释这一点。
   (iii) 使用上一个问题来计算 κ 和 τ。
   (iv) 从几何上解释为什么 $\lim_{\omega\to\infty}\kappa = (1/R)$ 和 $\lim_{q\to\infty}\kappa = 0$。
   (v) 使用（iii）来确认（iv）的预测。

6. **曲线的 Frenet-Serret 近似**。 设 $\mathbf{x}(t)$ 是一个粒子在时间 t 处的位置，该粒子以单位速度在空间中追踪曲线 C。 设$(\mathbf{T}_0, \mathbf{N}_0, \mathbf{B}_0)$为 $t = 0$ 时刻的 Frenet-Serret 坐标系，设 κ_0 和 τ_0 为此时的曲率和挠率。
   (i) 使用泰勒定理和 Frenet-Serret 方程 (9.3)，证明沿 C 的运动最初由 **Frenet 近似**给出：

$$\mathbf{x}(t) \asymp \mathbf{x}(0) + t\mathbf{T}_0 + \kappa_0\frac{t^2}{2}\mathbf{N}_0 + \kappa_0\tau_0\frac{t^3}{6}\mathbf{B}_0$$

   (ii) 几何解释前三个项。
   (iii) 第四项在几何上描述了什么？
   (iv) 根据你对 (iii) 的回答，为什么当 $\tau_0 = 0$ 时这一项必然消失？
   (v) 如果 $\kappa_0 = 0$，为什么第四项也应该消失是有意义的？

曲面的主曲率

7. **曲面的级数展开**。 对于以下每个方程，请使用计算机绘制曲面。 使用级数展开计算原点附近的二阶近似，然后使用（10.6），推导出主曲率和 $\mathcal{K}$。 目视确认你的计算至少在定性上是正确的。
   (i) $z = \exp(x^2 + 4y^2) - 1$
   (ii) $z = \ln\cos y - \ln\cos 2x$

8. **旋转曲面曲率的变速公式**。 在课本中，我们描绘了一个粒子以恒定单位速度沿着旋转曲面的母线运动，在这种情况下，主曲率由（10.11）和（10.12）给出。 如果速度 $v(t)$ 不保持恒定，请回想 (10.11) 变为 (8.6)：

$$\kappa_1 = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{v^3}$$

   (i) 通过修改第 114 页的 [10.6]，推导出第二个主曲率由下式给出

$$\kappa_1 = -\frac{\dot{x}}{yv}$$

   (ii) 推断出

$$\boxed{\mathcal{K} = -\frac{\dot{x}\,[\dot{x}\ddot{y} - \dot{y}\ddot{x}]}{y\,[\dot{x}^2 + \dot{y}^2]^2}}$$

(iii) 如果粒子只向右移动，从不回溯任何 x 值，那么我们可以沿着曲线自由调整速度，使 x 代表时间：$x(t) = t$，在这种情况下 $y(t) = y(x)$，时间导数变成 x 导数：$\dot{y} = y'$。 证明 (ii) 中的曲率公式然后简化为

$$\boxed{\mathcal{K} = -\frac{y''}{y\,[1 + (y')^2]^2}} \tag{20.1}$$

9. **旋转曲面曲率的极坐标公式**。 在前面的练习中，旋转曲面是使用生成曲线与旋转轴的距离定义的，旋转轴被视为 x 轴。 现在，让我们改为通过在垂直于其对称轴的平面上方给出其高度来描述这样的表面，我们现在将其视为 z 轴。 如果 r 是 (x,y) 平面中的常用极坐标，请考虑旋转表面 $z = f(r)$。

(i) 通过简单地适当改变名称，并使用链式法则，证明 (20.1) 变成

$$\boxed{\mathcal{K}(r) = \frac{f'(r)\,f''(r)}{r\,\{1 + [f'(r)]^2\}^2}}$$

(ii) 为以原点为中心的球体或半径 R 求 $f(r)$。 检查 (i) 中的公式是否给出了正确的 $\mathcal{K}$ 值。

(iii) 手动绘制曲面 $z = \exp[-r^2/2]$，计算 $\mathcal{K}$，并找到正、负和零曲率的区域。（计算机生成的表面图应该可以直观地确认你的答案。）

10. 在文本中，我们注意到表面 $z = r^4$ 在原点处有一个平面点（$\mathcal{K} = 0$），并且表面在其他地方具有正曲率，这在视觉上是显而易见的。 通过使用上一题中的公式计算 $K(r)$ 来证明这是正确的。

高斯绝妙定理

11. **为什么纸折叠成一条直线**。 我们理所当然地认为，当我们折叠一张纸时，折痕会自动形成一条直线，但为什么会发生这种情况？ 证明（无需任何计算）这是绝妙定理的直接结果！ （我要感谢我的同事罗伯特·沃尔夫博士，他曾是陈省身的学生。）

形状算子

12. **可视化线性代数**。 通过对线性变换本身进行直接和几何推理，而不是通过将“魔鬼机器”（参见序言）应用于表示它们的矩阵来完成以下所有操作！ 我们希望这些示例（连同课文中的示例）能够激励你在下次遇到线性代数时运用这种几何视角。
    (i) 验证 SVD 的几何解释 [15.4] 可以推广到 $\mathbb{R}^3$，因此 M^{T} 的解释 [15.5] 也可以。
    注意：虽然可视化变得更难，但在 $\mathbb{R}^n$ 中也是如此。
    (ii) 回想一下，正交线性变换是一种保持长度，因此保持角度（一般）和正交性（特别是）的变换——例子包括旋转和反射。 在 $\mathbb{R}^3$ 中，用 (i) 解释为什么任何正交变换 R 都具有 $\mathrm{R}^{-1} = \mathrm{R}^{\mathrm{T}}$ 的性质。
    (iii) 回想一下，如果 $[\mathrm{M}]^{\mathrm{T}} = -[\mathrm{M}]$，矩阵 $[\mathrm{M}]$ 被称为反对称矩阵。 在 $\mathbb{R}^2$ 中，使用 (15.12) 给出底层偏对称线性变换 M 的几何特征（根据扭转 τ）。
    (iv) 如果 $\mathrm{x} \cdot (\mathrm{Px}) > 0$，对于所有 x，对称线性变换 P 称为正定变换。用（15.13）证明当（且仅当）它自身的所有值都为正时，P 是正定的。
    (v) 直接论证，或者以 (iv) 为基础，表明对称的正定线性变换必须具有正行列式（因此是可逆的）。 通过在 $\mathbb{R}^3$ 中找到一个简单的反例，证明逆命题是错误的。 （提示：在几何上，行列式是线性变换的（有符号）体积膨胀因子，根据方向是保留还是反转，符号分别为正 (+) 或负 (-)。）
    (vi) 给定上面的 (iv)，从 (15.15) 得出 $\mathrm{M}^{\mathrm{T}}\mathrm{M}$ 和 $\mathrm{M}\mathrm{M}^{\mathrm{T}}$ 总是对称和正定的。相反，证明任何对称的正定线性变换 P 都可以以无限多种方式分解为 $\mathrm{P} = \mathrm{M}^{\mathrm{T}}\mathrm{M}$。
    (vii) 以 (vi) 为基础，证明正如任何正数都有实数平方根一样，任何正定线性变换 P 也有平方根 Q，使得 $\mathrm{P} = \mathrm{Q}^2$。 更详细地，证明 $\mathrm{Q} = \mathrm{R}^{-1}\mathrm{DR}$，其中 R 是正交的，$[\mathrm{D}]$ 是对角矩阵，其条目是 P 的特征值的平方根。
    (viii) 如果 $\mathrm{x} \cdot (\mathrm{Sx}) > 0$，对于所有 x，对称线性变换 S 称为半正定变换。 如果是 $\mathbb{R}^3$ 的任何线性变换，则证明 $\mathrm{M} = \mathrm{RS}$，其中 R 是正交的，S 是对称的和半正定的。 这称为 M 的极分解。

13. **消失形状运算符 $\Leftrightarrow$ 平坦**。 通过几何和计算表明，如果形状算子完全消失，则曲面是平面的一部分。

14. **形状算子中的高斯曲率**。 如果 $\mathbf{u}$ 和 $\mathbf{v}$ 是具有形状算子 S 的曲面上点 p 处的短切向量，请在几何上证明 $\mathrm{S}(\mathbf{u}) \times \mathrm{S}(\mathbf{v}) = \mathcal{K}(\mathrm{p})\,[\mathbf{u} \times \mathbf{v}]$。 （注意：即使切向量不短，方程本身也是有效的，但几何解释依赖于最终等式的思考。）

15. **形状算子、曲率和平均曲率的笛卡尔公式**。 令 $\mathrm{T_p}$ 为一般曲面 $\mathcal{S}$ 在点 p 处的切平面。 将笛卡尔坐标 (x, y) 引入 $\mathrm{T_p}$，取 $\mathrm{p} = (0, 0)$，使那里的法线 $\mathbf{n}$ 沿 z 轴指向。 然后，在局部，$\mathcal{S}$ 由下式给出

$$\mathrm{z} = \mathrm{f}(\mathrm{x}, \mathrm{y}) \quad \text{其中} \quad \mathrm{f}(0,0) = 0, \qquad \text{在}(0,0)\text{处}\ \partial_x \mathrm{f} = 0 = \partial_y \mathrm{f}$$

    (i) 请注意，如果我们定义 $\mathrm{F}(\mathrm{x}, \mathrm{y}, \mathrm{z}) \equiv \mathrm{f}(\mathrm{x}, \mathrm{y}) - \mathrm{z}$，则 F 在 $\mathcal{S}$ 上是常数。推导出

∇F 是 $\mathcal{S}$ 的法线，因此，

$$n = \frac{1}{1 + (\partial_x f)^2 + \left(\partial_y f\right)^2} \begin{pmatrix} \partial_x f \\ \partial_y f \\ -1 \end{pmatrix}$$

(ii) 推导出形状算子在 $(0,0)$ 处的矩阵为

$$\boxed{[S] = \begin{bmatrix} \partial_x^2 f & \partial_y \partial_x f \\ \partial_x \partial_y f & \partial_y^2 f \end{bmatrix}} \tag{20.2}$$

从中可以立即得出

$$\boxed{\mathcal{K} = \kappa_1 \kappa_2 = \det[S] = (\partial_x^2 f)\left(\partial_y^2 f\right) - \left(\partial_x \partial_y f\right)^2} \tag{20.3}$$

同样，$\bar{\kappa} \equiv$平均曲率$= \left[\frac{\kappa_1 + \kappa_2}{2}\right] = \frac{1}{2}\mathrm{Tr}[S] = \frac{1}{2}[\partial_x^2 f + \partial_y^2 f]$，可以写成

$$\boxed{\bar{\kappa} = \frac{1}{2}\nabla^2 f}$$

其中 $\nabla^2 = \partial_x^2 + \partial_y^2$ 是拉普拉斯算子，在第 41 页上首次遇到。

16. **非主坐标**。 让我们重新分析鞍座曲面，这次不要像我们在文本中所做的那样首先将坐标轴与主方向对齐。

(i) 在原点附近绘制表面 $z = xy$。（提示：回忆 $\sin 2\theta = 2\sin\theta\cos\theta$，求以原点为圆心半径为 r 的圆的点上方和下方的表面高度，其中 $x = r\cos\theta$ 和 $y = r\sin\theta$。）

(ii) 使用（10.4）识别主方向。

(iii) 用（20.2）推导出形状算子有矩阵

$$[S] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

并推导出 $\mathcal{K} = -1$ 和 $\kappa = 0$。

(iv) $[S]$ 表示的几何变换 S 是什么？

(v) 使用 (iv) 以几何方式确定 S 的特征向量和特征值（不是通过计算）。（注意：特征向量应指向您在（ii）中找到的相同方向。）

(vi) 在 (ii) 或 (v) 的指导下，旋转 (x,y) - 轴以获得与主方向对齐的新 (X,Y) − 轴，并显示原始方程现在变为 $z = \frac{1}{2}(X^2 - Y^2)$。

(vii) 通过将 (v) 中的新方程与第 111 页的 (10.6) 进行比较，确认您在 (iii) 中找到的 $\mathcal{K}$ 和 κ 的值。

(viii) 用（20.2）记下（vi）新的（X，Y）坐标系中的[S]，确认 $\mathcal{K}$ 和 κ 没有变化。

(ix) 验证 (iii) 和 (viii) 中的两个不同矩阵 [S]（代表相同的线性变换 S）符合通用矩阵公式 （15.20），第 161 页。

17. 使用（20.2）和（20.3）验证你在练习 7 中的结论。

18. **极坐标中的曲率公式**。 如果 $z = f(r,\theta)$，证明曲率是由这个令人失望的复杂公式给出的：

$$\boxed{\mathcal{K} = \frac{r^2\partial_r^2 f\left(\partial_\theta^2 f + r\partial_r f\right) - [\partial_\theta f - r\partial_r\partial_\theta f]^2}{\{r^2[1+(\partial_r f)^2] + (\partial_\theta f)^2 g\}^2}} \qquad (20.4)$$

19. **具有相等曲率的非等距曲面**。 正如我们在第一幕中讨论的那样，明丁证明了如果两个曲面具有相同的常数 $\mathcal{K}$，那么它们彼此局部等距。 因此，根据这个恒定曲率 $\mathcal{K} > 0$、$\mathcal{K} = 0$、$\mathcal{K} < 0$，表面分别与球面、平面或伪球面等距。 我们现在提供一个例子来证明反过来是错误的。

(i) 使用与上一个练习相同的符号，找出这两个曲面的形状：$\mathcal{S}_1$，$f_1(r,\ \theta) = \ln r$，$\mathcal{S}_2$，$f_2(r,\ \theta) = \theta$。（提示：$\mathcal{S}_2$ 被称为螺旋面。）通过使用计算机画出它们来检查您的答案。

(ii) 假设 $\mathcal{S}_1$ 中的一个点和 $\mathcal{S}_2$ 中的一个点“对应”，如果它们具有相同的 (r,θ) 坐标。 通过找到两个表面的度量公式，表明这种对应关系不是等距关系。

(iii) 然而，使用（20.4）推导出两个曲面的对应点具有相同的曲率！

$$\mathcal{K}_1(r,\theta) = \frac{-1}{(r^2+1)^2} = \mathcal{K}_2(r,\theta)$$

20. **n-鞍座的曲率**。 假设 n-鞍座的高度为 $z = f(r,\theta) = r^n\cos n\theta$，用 (20.4) 证明它的（负）曲率在每个圆上是常数 $r = \text{const.}$，并且由下式给出以下公式。

(i) 常见的 2-鞍座（围绕一个通用双曲点）有

$$\mathcal{K} = -\frac{4}{(1+4r^2)^2}$$

(ii) 猴子 3-鞍座有

$$\mathcal{K} = -\frac{36r^2}{(1+9r^4)^2}$$

(iii) n-鞍座有

$$\mathcal{K} = -\frac{(n-1)^2 n^2 r^{2n}}{(r^2+n^2r^{2n})^2}$$

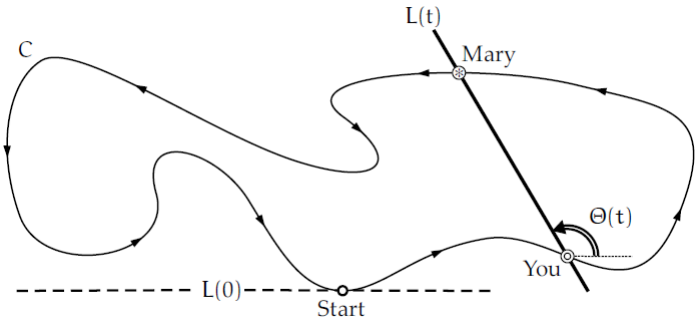
全局高斯-博内定理简介

21. GGB 的亏格相关预测。
   (i) 使用 GGB 推导出任何拓扑球体都必然具有正曲率的区域。
   (ii) 使用 GGB 推导出任何拓扑甜甜圈都必然具有正曲率的区域。
   (iii) 事实上，*每一个* $g \geq 2$ 的光滑闭合曲面 $\mathcal{S}_g$ 也必然有一个正曲率的点，尽管 GGB 不再保证这一点。设 $S(r)$ 是一个半径为 r 的球体，以 $\mathcal{S}_g$ 内的任意点为中心。想象一下，我们从一个足够大的 r 值开始，使 $S(r)$ 包含 Sg。现在缩小球体，直到它第一次与 $\mathcal{S}_g$ 接触，当 $r = R$，比如说。 论证在与 $S(R)$ 的接触点，$\mathcal{K} > (1/R^2) > 0$。 提示：考虑接触点的主方向。

22. **神圣破碎的高斯-博内，蝙蝠侠！** 回想我们在 [16.7] 中对虫洞的定义，并且还记得它具有总曲率，$\mathcal{K} = -4\pi$。
   (i) 如果将 n 个小百吉饼放在一个大的规则 n 边形的顶点，并沿 n 边形的边缘用虫洞连接器连接在一起形成表面 $\mathcal{S}$，解释为什么 $\mathcal{K}(\mathcal{S}) = -4\pi n$。 为什么这似乎并不违反 GGB？
   (ii) 假设 $n = 4$，因此百吉饼位于正方形的角上，沿四个边连接。 现在沿其中一条对角线添加*第五*个虫洞连接器，连接两个先前断开连接的百吉饼。 现在的总曲率是多少，为什么这看起来不违反 GGB？

GGB 的第一个（启发式）证明

23. **霍普夫环量定理**。 让我们尝试将 Hopf (1935) 巧妙的论证归结为它的本质。 想象一下，定理的（平滑且“简单”的）闭合曲线 C 是一条蜿蜒穿过沙丘的错综复杂的远足/跑步路径。 请参阅下面的地图。 入口位于最南端（页面下方）； C 的其余部分位于北部（向上）。如图所示，设 $\Theta(t)$ 是连接你（⊙）和你的朋友玛丽（⊛）的线 $L(t)$ 在时间 t 的角度。 最初，玛丽站在你旁边，所以 $L(0)$（虚线）是水平的，$\Theta(0) = 0$。



   (i) 玛丽决定做一个热身圈，但你留在起点，只看她跑。 当她逆时针跑完 C 的长度时，你转过头跟上。 解释为什么在 Mary 回到你身边时你的视线 $L(t)$ 必须执行 π 的净旋转：$\delta\Theta = \pi$。接下来，当你跑一圈时，玛丽保持不动并休息，

所以 L 旋转另一个 π。因此，在你和玛丽各自跑完一圈后，L 完成了一整圈：$\delta\Theta = 2\pi$。

(ii) 接下来，玛丽又跑了一圈，但这次你不要等到她跑完才开始移动。 你开始沿着她身后的轨道非常缓慢地走。 在玛丽回到起跑线之前，您只跑了一小段距离。 一旦她到了那里，你就沿着剩下的小路慢跑回到她身边。 解释为什么 L 的净旋转仍然必须是 $\delta\Theta = 2\pi$。

(iii) 接下来，您决定与玛丽赛跑，尽管她总是获胜。 果然，你永远不会靠近。 解释为什么在你和玛丽各自完成你的一圈后 L 的净旋转仍然是 $\delta\Theta = 2\pi$。（提示：想象在 (ii) 中逐渐增加你的速度，并想象 $\Theta(t)$ 图形的连续演化：(i) ⇝ (ii) ⇝ (iii)。）

(iv) 您决定进行最后一场比赛，这次您全力以赴，并且成功地始终紧跟玛丽。 当你盯着她的背影，眼睛沿着路径向前看时，你的视线 L 现在是 C 的切线。而且由于 L 的净旋转仍然必须是 $\delta\Theta = 2\pi$，霍普夫的证明是完整的！

GGB 的第二个（角盈）证明

24. **只有五个正多面体**。 如果正多面体的每个面都是正 n 边形，并且 m 个这样的面在每个顶点相遇，则使用以下步骤来证明唯一的可能性是 [18.2] 中所示的五个柏拉图多面体，即 $(n, m) = (3,3)$ 或 $(4,3)$ 或$(3,4)$ 或 $(5,3)$ 或 $(3,5)$。

(i) 解释为什么 $m > 3$。

(ii) 如果 θ_n 是规则 n 边形面的内角，解释为什么 $m\theta_n < 2\pi$。（提示：想象切出在顶点 v 处相交的 m 个面，现在沿着包含 v 的边之一切割，最后将表面展平到平面上。）

(iii) 写下 $n = 3$、4、5、6 的 θ_n 并穷尽所有的可能性。

25. **只有五个拓扑正多面体**。 让我们使用与前一个练习相同的 (n, m) 符号，只是现在我们的拓扑多面体的面可以是弯曲的、不规则的和/或带有波浪边的。像往常一样，让 V、E 和 F 表示顶点、边和面的数量。

(i) 解释为什么 $V = F_n / 2$。

(ii) 解释为什么 $E = F_n / m$。

(iii) 将前两个方程代入欧拉多面体公式，推导出

$$F = \frac{4m}{2(m+n) - mn}$$

(iv) 推导出 $\Omega(n, m) \equiv 2\,(m + n) - mn > 0$。

(v) 将 (n, m) 视为 $\mathbb{R}^2$ 中的网格点。 在相关区域（$n > 3$ 和 $m > 3$）中，用值 $\Omega(m, n)$ 标记每个网格点，注意并使用关于线 $n = m$ 的对称性。

(vi) 推断出唯一的解 (n, m) 与我们在之前的练习中发现的相同：五个柏拉图多面体，但现在拓扑变形了。

这个结果归功于 Simon Antoine Jean L’Huilier (1811)。 参见 James 1999 中的 Robin Wilson，第 516 页。

26. **凸多面体的勒让德投影**。 如文中所述，勒让德对欧拉多面体公式的证明假设多面体是凸的。 我们提出的证明不需要这个假设，但勒让德的凸性假设给了他一个优雅的优势：它允许他立即从原始多面体 $\mathcal{P}$ 跳到球面的拓扑等效测地多边形分区（见 [18.7]）。 想象一下，凸多面体 $\mathcal{P}$ 是一个线框，一个灯泡 B 放在它里面的某个地方。 如果 $\mathcal{S}$ 是一个以 B 为中心并包围 $\mathcal{P}$ 的球体，证明球体上框架的阴影是所需的测地线分区。

27. **霍普夫证明** $\mathcal{X}(\mathcal{S}_{\mathrm{g}}) = 2 - 2\mathrm{g}$。 在文中，我们首先证明 $\mathcal{X}(\mathcal{S}_{\mathrm{g}}) = 2$，然后证明粘贴在“Tobleroner handle”上（见 [18.9]）会使 $\mathcal{X}$ 减少 2（见 [18.10]）。 这是霍普夫（1956, pp. 8-10）提出的一个优雅的替代论点。

    (i) 将两个 Tobleroner 手柄粘合在一起以创建拓扑环面 $\mathcal{S}_1$，并推导出 $\mathcal{X}(\mathcal{S}_1) = 0$。
    (ii) 从这个圆环上去掉一个三角形，并从 g 属的一般曲面 $\mathcal{S}_{\mathrm{g}}$ 上去掉一个三角形。现在将两个三角形孔的边缘相互粘合，从而创建一个更高属的新封闭曲面 $\mathcal{S}_{\mathrm{g}+1}$。 推导出 $\mathcal{X}(\mathcal{S}_{\mathrm{g}+1}) = \mathcal{X}(\mathcal{S}_{\mathrm{g}}) - 2$。
    (iii) 使用 (i) 和 (ii) 推导出 $\mathcal{X}(\mathbb{S}^2) = 2$，且 $\mathcal{X}(\mathcal{S}_{\mathrm{g}}) = 2 - 2\mathrm{g}$。

28. **对偶多面体**。 在立方体每个面的中心放置一个顶点，并将它们连接在一起形成一个八面体； 这是立方体的对偶，每个面都变成了一个顶点，每个顶点都变成了一个面。 如果我们继续前进，取八面体的对偶，即立方体的对偶的对偶呢？ 在[18.2]的帮助下，对剩余的三个柏拉图固体重复这个练习。

GGB 的第三个（向量场）证明

29. **偶极流线是圆形的**。 证明 [19.3] 中的偶极流线是圆，
    (i) 通过计算
    (ii) 几何上。

30. $\mathcal{S}_{\mathrm{g}}$**上的蜂蜜流**. 取 [19.11] 中的 $\mathcal{S}_{\mathrm{g}}$ 并围绕垂直于页面的轴将其旋转直角，以便孔水平对齐，而不是垂直对齐。 画出新的蜂蜜流，推导出总共有 $6\mathrm{g} - 2$ 个奇异点。根据庞加莱-霍普夫定理，确认它们的指数之和确实是 $\mathcal{X} = 2 - 2\mathrm{g}$。

31. **Stiefel 向量场的存在**。尽可能明确地解释为什么 [19.12] 中的构造可以保证在 $\mathcal{S}_{\mathrm{g}}$ 上创建一致的向量场，而不管多边形的具体划分如何。

32. $\mathcal{S}_{\mathrm{g}}$ **上有一个奇异点的向量场**？ 根据庞加莱-霍普夫定理，球体[19.8b]上的偶极场只有一个奇异点，其指数为 2（$= \mathcal{X}$）。 在环面 ($\mathrm{g} = 1$) 上，我们已经看到流中不需要任何奇异点。 但是如果 $\mathrm{g} \geq 2$，则必定存在奇异点，因为它们的指数之和为 $\mathcal{X} \neq 0$。 在这种一般情况下，是否总是可以推广球体上的偶极场，以找到只有一个指数 $\mathcal{X}$ 的奇异点的流动？ （提示：可以通过将两个或多个奇异点合并为一个来创建新场。例如，我们可以想象通过合并 [19.8a] 中的两个漩涡来创建 [19.8b] 中的偶极子场。）

33. $\mathcal{S}_{\mathrm{g}}$ **上有** $-\mathcal{X}$ **个鞍点的向量场**。 证明（通过绘制！）可以在 $\mathcal{S}_{\mathrm{g}}$（$\mathrm{g} \geq 2$）上构造一个向量场，使得恰好有 $(2\mathrm{g}-2) = -\mathcal{X}$ 个奇异点，每个奇异点 $\mathcal{I} = -1$。（提示：（A）首先解决 $\mathrm{g} = 2$；结果证明对更高属的推广是显而易见的。（B）图片 $\mathcal{S}_2$ 两个百吉饼仍然连接在一起。接下来，想象这个表面在很大的重量下几乎被压扁，这样它类似于两个连接的环。最后，确认（通过绘制！）可以在“顶”表面放置一个鞍点，在“底”放置一个鞍点，并使顶场和底场在它们的位置一致在边缘相遇。）

第二十一章 历史谜题

平行移动[151]（Parallel transport）现在被认为是微分几何中最基本和最强大的概念之一，但它出现的时间太晚了。

到 1915 年，n 维空间中现代微分几何的大部分基本概念和计算工具已经到位，而且很快。

在过去的十年里，爱因斯坦一直在努力协调他 1905 年的狭义相对论与引力。 他从 1905 年就知道没有任何物理效应可以比光速传播得更快，但牛顿的平方反比定律——自 1687 年以来就无人质疑——坚持认为，如果一个巨大的太阳耀斑从太阳表面爆发，它对地球的微小引力牵引会立即被感觉到，尽管耀斑发出的光需要八分钟才能到达地球——这是一个赤裸裸的矛盾！

爱因斯坦的物理直觉只是逐渐地将他推向微分几何，在那里他奇迹般地发现，1901 年格雷戈里奥·里奇（Gregorio Ricci，1853-1925）和图利奥·列维-奇维塔（Tullio Levi-Civita，1873-1941）共同发表了“为揭示他的理论而精心准备且几乎是命中注定的工具”[152]——**张量分析**[153]（Tensor Calculus）。

1915 年 11 月 25 日，爱因斯坦终于结束了他长达十年的可怕斗争，利用张量微积分写下了他著名的**场方程**（Field Equation），该方程调和了引力与狭义相对论； 他将这种结合命名为**广义相对论**。 爱因斯坦成功地理解了引力的真正本质：它是物质和能量在 4 维时空上施加的**黎曼曲率**[154]（Riemannian curvature）； 自由落体的粒子然后通过沿着测地线[155]穿过弯曲的时空来响应引力场。

截至本文撰写之时，即 2020 年，该理论的每一个可检验预测都已得到证实，包括令人惊讶的诺贝尔奖获奖[156]确认书——2015 年 9 月 14 日——引力波的存在，爱因斯坦几乎在一个世纪前，即 1916 年就预言了这一点！

很少有人意识到，在一些情况下，这些实验确认已经达到了惊人的准确度，与以前的黄金标准：量子电动力学相匹敌甚至超越。 事实上，如果不考虑广义相对论预测的精确的重力时间扭曲效应，我们日常使用全球定位系统 (GPS) 技术[157]是不可能的！ 因此，爱因斯坦

---

[151] 有时也称为“平行位移”（parallel displacement）或“平行传播”（parallel propagation）。

[152] 列维-奇维塔 (1931)。 有关 1915 年之前张量分析被缓慢接受的历史，以及 1915 年之后的巨大变化，请参见 Bottazzini (1999)。

[153] 这个纯粹的数学发现最初被称为绝对微积分（Absolute Differential Calculus），但后来被称为里奇微积分（Ricci Calculus），最后被称为张量分析（Tensor Calculus）。埃利·嘉当（Élie Cartan）更强大、更优雅的张量分析版本构成了我们戏剧（第五幕）的结局，被称为外微分的微分形式（Exterior Calculus of Differential Forms）。

[154] 这种曲率的内蕴度量将在第四幕的结尾进行描述：它将高斯曲率推广到 n 维空间。

[155] 尽管测地线现在最大化使用弯曲时空度量测量的“距离”，它推广了闵可夫斯基时空间隔 λ； 见（6.15），第 75 页。

[156] 2017 年诺贝尔物理学奖授予了 RainerWeiss、Barry C. Barish 和 Kip S. Thorne，因为他们联合建造了激光干涉引力波天文台 (LIGO) 探测器，使这一发现成为可能。

[157] 参见 Taylor 和 Wheeler (2000, A-1)。

1915 年的发现不仅是人类智力的一项极其美丽的科学胜利，而且是我们拥有的经过最佳检验的物理理论之一。

[21.1] 图利奥・列维-奇维塔 (1873 – 1941)

但是这里有一个严重的谜题[158]，一个尚未广为人知或认可的谜题。 爱因斯坦的成功更加引人注目，而且仍然更加令人费解，因为他在列维-奇维塔（如图 [21.1] 中所示）发现[159] **平行移动**的概念之前实现了这一目标，而平行移动的概念直到 1917 年才出现！ 如果没有这个概念，我个人不知道有什么方法可以完全理解爱因斯坦 1915 年的发现。

那么什么是平行移动？ 就像测地线的概念一样，它是一种奇怪的两栖动物，在外在几何和内蕴几何世界间来去自如，我们现在将描述它在每个世界中的样子，并试图了解它如何能够跨越两者。

[158] 对于熟悉广义相对论的读者来说，实际上还有第二个更壮观的谜题：1915 年 11 月，爱因斯坦不知道微分比安基恒等式——后来被称为 (29.17)——因此他没有意识到他的定律可以自动得到能量-动量守恒！ 参见 Pais (1982, p. 256)。

[159] Levi-Civita (1917).

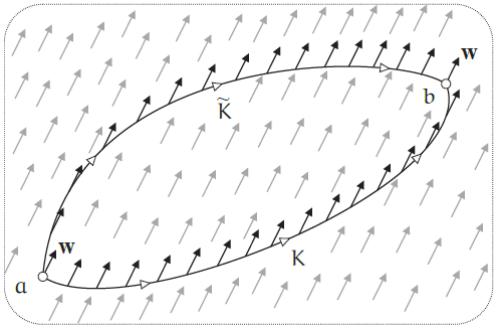
第二十二章 外在构造

22.1 边走边投影到曲面！

这是列维-奇维塔（Levi-Civita）在 1917 年的基本论文中寻求实现的启发式思想。要沿着曲面 $\mathcal{S}$ 中连接的两点 a 到 b 的曲线 K 将切向量 **w** ***平行移动***（英文 parallel transport，或译平行输运），我们希望始终保持 **w** 指向相同的方向（并具有相同的长度），同时在 **w** 沿 K 移动时保持与 $\mathcal{S}$ 相切，**w** 在每个时刻的方向始终与其前一刻的方向平行。

但这似乎是不可能的！ 如果我们沿着 K 从 p 到 q 移动很小的距离 ϵ，那么切平面 T_p 中的向量 **w**，当严格平行于自身（在 $\mathbb{R}^3$ 中）移动到 q 时，通常不会位于 Tq 中，而是会伸出 $\mathcal{S}$ 的。

在欧氏平面中不会出现这样的问题，并且有一个简单的全局平行感，这是由于通过任何不在直线 L 上（在 **w** 方向上）的点，存在一条与 L 平行的唯一直线，在相同的方向 **w**. 这种全局平行性可以被描述为穿过平面的均匀流动，速度为 **w**； 见[22.1]。 沿 K 的平行移动向量只是该恒速场对 K 的限制。 因此，很明显无论我们是沿 K还是沿其他路线 $\widetilde{K}$ 到达b 处，我们总是在那里得到相同的平行移动向量 **w**。



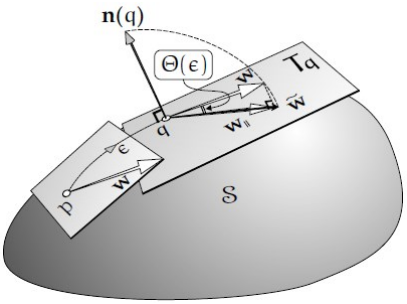
[22.1] 在欧氏平面上有一个全局的平行性概念，表现为等速向量 **w** 的流动。要沿 K 平行移动 **w**，我们只需将该向量场限制为 K。因此我们总是得到相同的向量 w 在 b 处，我们是否沿着 K 或 $\widetilde{K}$ 旅行。 这种路径无关性仅在曲率消失的情况下成立。

列维-奇维塔的关键观察是，这种平行移动向量的路径无关性仅在空间平坦时才会发生；如果空间是弯曲的，那么沿着两条不同路径的平行移动（如它将立刻被定义的样子）将产生不同的最终向量。 这个重要的现象被称为和乐（holonomy，又译完整），且是该名命名的即将到来章节的主题。 正如我们将看到的，这种和乐可用于测量曲率，不仅可以测量曲面的高斯曲率，还可以测量更高维弯曲空间的黎曼曲率，例如爱因斯坦的弯曲 4 维时空。

如果我们尝试在双曲平面 $\mathbb{H}^2$ 中重复欧氏构造，那么我们会立即遇到麻烦，因为如果给我们一个点和一条直线 L，我们知道现在有无限多条“平行”线穿过相邻点； 我们应该选择哪一个？！

稍后我们将回到如何在 $\mathbb{H}^2$ 中或等价地在伪球面上进行平行移动这个话题，但现在让

我们立即切入纵深，看看如何在一般曲面上进行平行移动。



[22.2] 为了将切向量 w 从 p **平行移动**到距离 ϵ 的相邻点 q，我们在 $\mathbb{R}^3$ 中将其平行于自身移动，然后 (i) 将其向下投影到 T_q 上以获得 $\mathbf{w}_\parallel$ 或者 (ii) 向下旋转得到 $\widetilde{\mathbf{w}}$。这两种结构是等价的，因为我们最终取了 ϵ 消失的极限。

图[22.2] 中使用了上面开头段落中引入的符号，说明了我们的第一个平行移动的外在方法。 如果我们将 p 处的切向量 $\mathbf{w}$ 平行于自身（作为 $\mathbb{R}^3$ 中的向量）沿曲线 K 移动一小段（最终消失）距离 ϵ，则 $\mathbf{w}$ 不再与 q 处的 $\mathcal{S}$ 相切：它仅以轻微的角度 $\Theta(\epsilon)$ 伸出切平面 T_q。 在 q 处与 $\mathcal{S}$ 相切的 $\mathbf{w}$ 的最佳近似是 $\mathbf{w}_\parallel$，通过将 w 正交向下投影到 T_q 上获得。 如果 $\mathcal{P}$ 表示该投影，并且 $\mathbf{n}(q)$ 是 q 处的单位法线，则如 [22.2] 所述，

$$\mathbf{w}_\parallel = \mathcal{P}[\mathbf{w}] = \mathbf{w} - (\mathbf{w}\cdot\mathbf{n})\mathbf{n} \tag{22.1}$$

为了沿着 K 平行移动 $\mathbf{w}$，我们必须想象将 K 分成大量长度为 ϵ 的微小步长，重复这个在 $\mathbb{R}^3$ 中平移并投影回曲面的过程，一遍又一遍，一次 ϵ 步长，直到我们最终到达 K 的末尾。即便如此，这也只是平行移动的一个近似——只有当我们取 ϵ 消失的极限时才能实现完美的平行移动，在这种情况下，随着我们的前进，连续不断地投影到曲面。

注意 [22.2] 中投影 $\mathbf{w}_\parallel$ 比原始向量 $\mathbf{w}$ 略短。如果我们只是简单地将 w 直接向下旋转到 T_q 上，那么我们得到 $\widetilde{\mathbf{w}}$，指向与 $\mathbf{w}_\parallel$ 相同的方向但 $|\widetilde{\mathbf{w}}| = |\mathbf{w}|$（长度 $|\mathbf{w}|$ 与构造无关，但这里我们选择它具有与 $\mathbf{n}$ 相同的单位长度，以便于可视化。）但是，由于 $\lim_{\epsilon\to 0}\Theta(\epsilon) = 0$，我们看到

$$\boxed{\mathbf{w}_\parallel \asymp \widetilde{\mathbf{w}}}$$

最后，因为完美的平行移动需要这个 ϵ 消失的极限，我们推断平行移动保持长度。

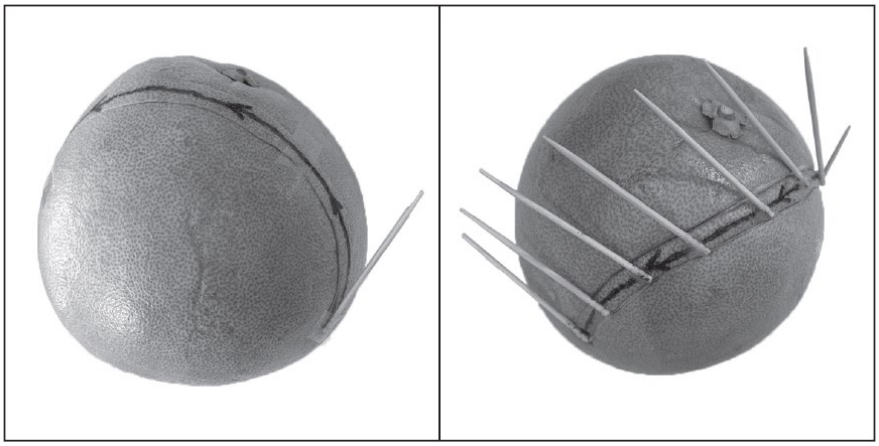
因此，我们得出了两种等价的平行移动外在方法：

> 要沿曲线 K 将切向量 $\mathbf{w}$ 在 $\mathcal{S}$ 上平行移动，请在 $\mathbb{R}^3$ 中平行于自身移动 $\mathbf{w}$，同时不断地（i）边走边投影到 $\mathcal{S}$ 或（ii）边走边向下旋转到 $\mathcal{S}$。 向量的长度在平行移动时保持不变。 (22.2)

方法 (ii) 在实践中要容易得多，我们强烈建议你在手头的任何弯曲物体上自己尝试一下——水果和蔬菜很方便，而且它们有一个很快就会显现出来的关键优势：它们可以去皮！

在你的曲面上画出你喜欢的任何曲线 K，将点 a 连接到点 b，将牙签的一端按你喜欢的任何切线方向压在曲面 a 处。现在将其平行于自身（在空间中）沿 K 移动一小段距离，然后将同一端直接向下压在曲面上（使其再次相切），. . . 重复，. . .重复，. . .重复，. . . 直到你到达 b!

图 [22.3] 在柚子上说明了这一点。 尽管我们可以选择与 $\mathcal{S}$ 相切的 a 处的任何方向，但这里我们选择 $\mathbf{w}$ 作为沿 K 的初始速度，以便说明这一点：通常，$\mathbf{w}_{\|}$ 不保持与 K 相切，尽管通过构造它确实保持与曲面相切。



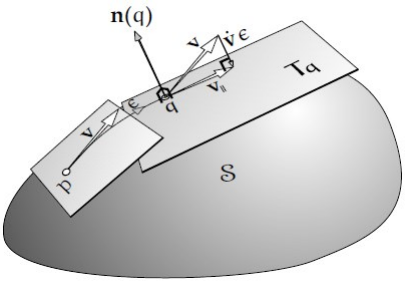
[22.3] 柚子上的平行移动：连续将牙签笔直向下压在曲面上，同时沿曲线平行移动牙签。

在这里，我们选择了路径 K 以接近柚子曲面的垂直截面，在右边的照片中，我们特意从几乎正上方拍摄的照片，使其看起来非常笔直。 因此，牙签从最初与 K 相切到最后几乎垂直于 K 的旋转如此之快可能非常令人惊讶和困惑。 这是因为我们的照片并没有揭示 K 相对于内蕴几何的弯曲程度。 这个特定路径 K 的内蕴测地曲率将在第 22.3 节中揭示，同时将揭示一种不同的平行移动方法。

22.2 测地线与平行移动

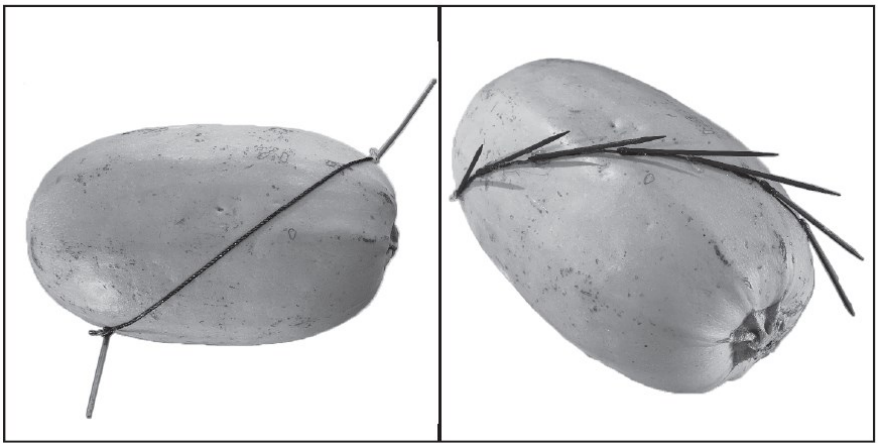
在什么情况下，初始切向量沿曲线 G 的平行移动会与 G 保持相切？ 答案是这恰好发生在 G 是测地线时！

要了解原因，回想一下测地线 G 的局部定义是一条单位速度曲线，对于该曲线，加速度始终沿着曲面法线 $\mathbf{n}$。这意味着如果我们查看相邻点 p 和 q 处的速度，距离/时间 ϵ 相隔，那么两个速度的差最终为 $\dot{\mathbf{v}}\epsilon$ 并最终沿 $\mathbf{n}$ 定向。图 [22.4] 描述了 [22.2] 的这个新的特殊情况，现在 $\mathbf{v}$ 作为速度，指向沿着线段 ϵ。 我们看到这意味着 q 处的新速度最终等于 p 处的原始速度 $\mathbf{v}$ 沿其自身平行移动的结果。



[22.4] 通过平行移动构建测地线。 从 p 以速度 $\mathbf{v}$ 发射一个粒子，并将 G 定义为沿自身平行移动 $\mathbf{v}$ 所获得的曲线，因此在相邻点 q（距离/时间 ϵ 远离）处的新速度为 $\mathbf{v}_{\parallel}$。 那么加速度 $\dot{\mathbf{v}}$ 由 $\dot{\mathbf{v}}\epsilon \asymp \mathbf{v}_{\parallel}$ 给出 -v。 但是 $\dot{\mathbf{v}} \propto \mathbf{n}$，因此 G 是测地线。

相反，假设我们给定一个点 p，我们从那里发射一个粒子，以 $\mathcal{S}$ 内的任意方向 $\mathbf{v}$ 穿过曲面。与之前相同的几何形状告诉我们，我们可以通过执行平行移动来构造测地线运动：在 $\mathcal{S}$ 内沿其自身携带 $\mathbf{v}$ 运动微小距离 ϵ，然后将其压到曲面上以获得那里的新速度. . . 然后重复！



[22.5] [a] 首先通过在黄色南瓜上拉紧一根绳子来构建其曲面上的测地线段 G。 [b] 移除绳子，在初始点留下一根牙签，与 G 相切。 正如预测，沿着自身平行移动初始牙签会产生处处与同一测地曲线 G 相切的牙签。

图 [22.5] 说明了这种构建测地线的新方法。 为了检查我们最终解决方案的正确性，我们首先使用我们原来的技巧在我们的曲面（一个南瓜）上构建一个测地线段 G，将一根绳子拉过它，这里在末端绑着两个（正常）牙签。 如果我们移除绳子，但留下一根与绳子起点相切的初始牙签，然后我们就可以拿起这个初始速度的牙签并沿着它自身平行移动。 如你所见，并且我们敦促你亲自尝试，此构造产生与细绳相同的测地线段 G。

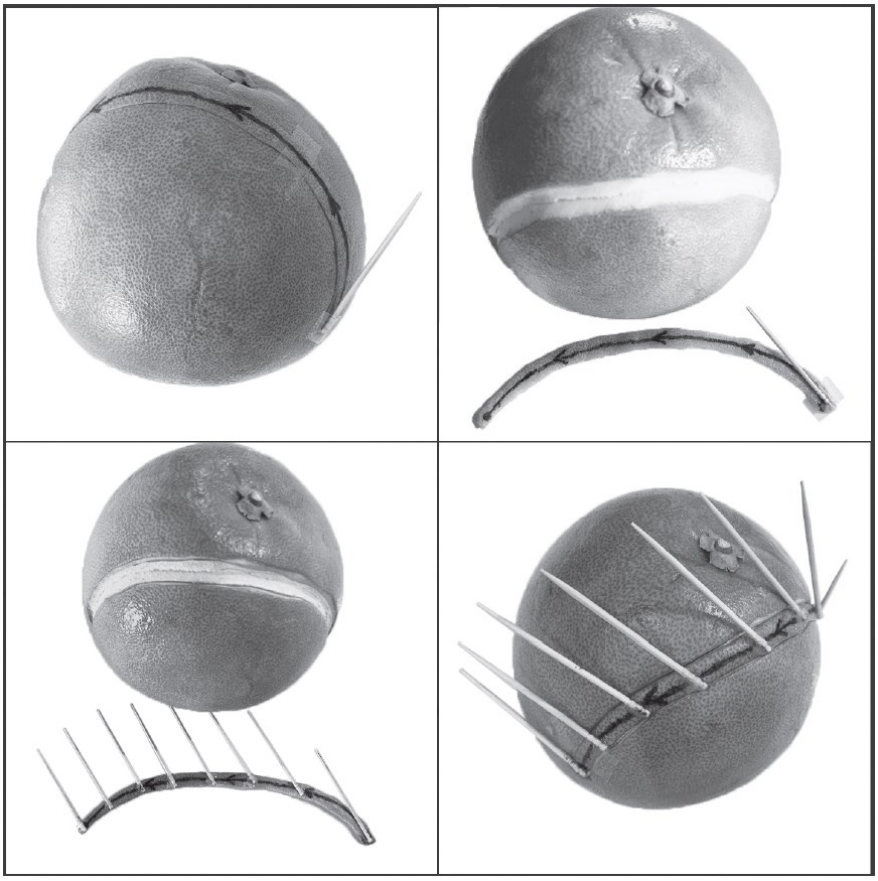
22.3 土豆削皮器移动

我们现在将描述第三种平行移动的外在方法，它将迅速证明其实力[160]，产生平行移动的

[160] 这种新构造在现代教科书中并不容易找到，我还记得 30 多年前第一次发现它时自己兴高采烈的样子。其他作者同样相信他们发现了一些新的东西：例如，Koenderink (1990)、Casey (1996) 和 Henderson (1998)。 但是到了写这本书的时候，我求助于原始资料，在那里我发现我们重新发现的东西是一个多世纪前由 Levi-Civita 本人首次发现的！ 要阅读 Levi-Civita 自己对他的想法的描述，翻译成英文，请参阅

重要特性，这些特性从我们在（22.2）中看到的前两个构造中并没有立即显现出来。

在 [22.2] 中，想象我们使用土豆削皮器剥下 $\mathcal{S}$ 上沿 K 运动 ϵ 周围的窄条。 现在，不要将 $\mathbf{v}$ 向下旋转到曲面，想象我们取下一小片果皮并将其弯曲，将其平压在 T_p 上，使 $\mathbf{v}_{\parallel}$ 到 $\mathbf{v}$，而不是相反。 如果我们沿着 K 一直剥离窄果皮，就会得到我们的新构造：



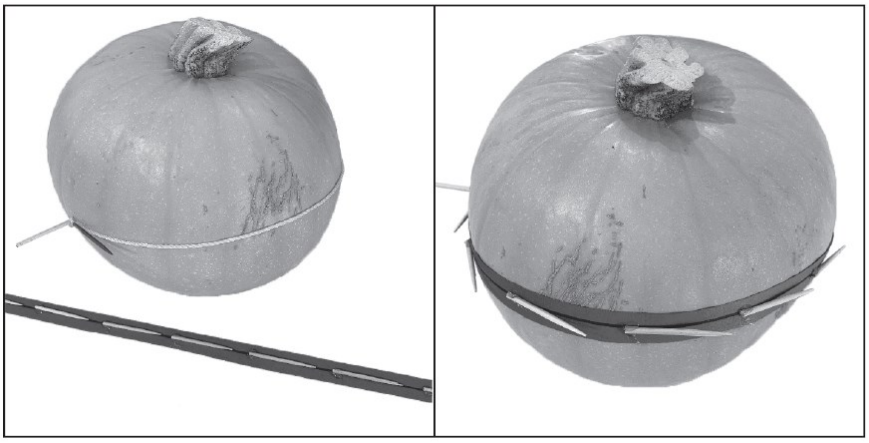
[22.6] 柚子上的**“土豆削皮器”平行移动**：沿曲线取下一小条果皮，在桌面上展平，在平坦的桌面内进行欧氏平行输送，最后将果皮和向量重新贴到曲面.

土豆削皮器平行移动。 要沿曲线 K 将切向量 $\mathbf{v}$ **平行移动**到 $\mathcal{S}$，请剥下一条狭窄的（最终非常薄的）$\mathcal{S}$ 条，其中包含 K 作为其中心线。 将这条带平放在平面上，然后沿着摊平的 K 对 $\mathbf{v}$ 做普通欧氏平行移动。最后，将 K（和构造的向量）重新粘到 $\mathcal{S}$ 的原始位置。 (22.3)

图 [22.6] 揭示了我们如何作弊并实际使用这种方法（而不是（22.2）中的方法）来执行先前在 [22.3] 中显示的平行移动。

Levi-Civita (1926, p. 102)。 有关这一发现历史的更多信息，请参阅 Goodstein (2018, Ch. 12)。

虽然土豆削皮器让人联想到正确的心理形象，但它实际上并不是完成这项工作的最佳工具。 要获得整齐、窄的果皮条，请拿一把锋利的小刀，沿着曲线在 K 的一侧做一个非常浅的切口，在 K 下方切割； 现在从另一侧做同样的事情，从而切出所需的窄条果皮，具有浅 V 形横截面。



[22.7] [a] 南瓜上的测地线段 G 首先通过在其曲面上拉紧一根绳子来构造。 与此同时，在桌面上，指向沿着直条带的牙签，被沿该条带平行移动，它自动保持与中心线相切。 [b] 取下绳子，将胶带（带有牙签）展开到曲面上，从同一点开始并朝同一方向展开。 正如预测的那样，胶带会自动沿着相同的测地曲线 G 展到曲面上，而牙签则沿着它平行输送。

平行移动的两个重要特性从新构造中立即变得清晰起来。

首先，如果两个向量沿欧氏平面中的曲线平行移动，那么显然它们之间的角度保持不变。新的构造立即暗示，因此在曲面上也必然如此：

> 如果两个向量沿曲面上的一条曲线平行移动，则它们之间的夹角保持不变。 (22.4)

相反，如果已知只有一个向量是平行移动的，并且与第二个向量的角度保持固定，那么第二个向量也必然是平行移动的。

其次，回忆第 13 页的（1.6）和第 14 页的（1.7）中的内容，如果将测地线从曲面剥离并平放在桌子上，它将变成一条直线； 相反，如果将一条又直又窄的胶带逐渐卷到曲面上，它会自动生成测地线。 在欧氏平面中，线的方向向量随着它沿线平行移动，继续沿线指向。 因此，我们有一个更直观、更直接、更可视的证据证明先前确定的事实，

> 如果有物体以初始速度向量 $\mathbf{v}$ 发射，沿测地线 G 以单位速率运动，那么任何未来时间的速度向量都可以通过沿 G 平行移动初始速度向量来得到。 相反，给定任意初始点和与 S 相切的任意初始速度向量 $\mathbf{v}$ ，$\mathbf{v}$ 沿 $\mathbf{v}$ 的平行移动产生由这些初始条件获得的唯一测地线。 (22.5)

这在 [22.7] 中有说明。 再次，我们敦促你亲自尝试。[161]

图 [22.8] 再次简单地说明了相同的想法。 条形 [22.7] 已从南瓜上取下，并沿与之前相同的测地线展开到 [22.5] 中所示的黄色南瓜的曲面上。 但是，在该图中，我们被迫手动执行平行移动和所示测地线的生成，而这次这两件事都是自动发生的。

[22.8] [22.7] 中所示的相同条带在这里展开到 [22.5] 中所示的曲面上。 与 [22.5] 中的手动构建不同，这里的测地线的生成和其速度向量沿其平行移动都是自动的。

[161]我们建议使用遮蔽胶带（又名画家胶带），因为它具有鲜艳的颜色，并且一旦创建了胶带，就可以轻松地反复拆除和重新粘贴。 制作窄条（从通常很宽的一卷胶带）的一种简单方法是将一段胶带粘在厨房砧板上，然后用锋利的刀沿胶带切开，想要多窄都行。 至于牙签，我们用热胶枪把它们粘起来，但只是在它们的底部，这样一旦贴上条带，它们就可以自如地与曲面相切。

第二十三章 内蕴构造

23.1 通过测地线的平行移动

在欧氏平面中，有一种简单、显而易见的方法可以沿直线 L（方向为 $\mathbf{v}$）平行移动向量 $\mathbf{w}$：保持 $\mathbf{w}_{\parallel}$ 和 $\mathbf{v}$ 之间的夹角在沿 L 移动时恒定不变。参见 [23.1a]。

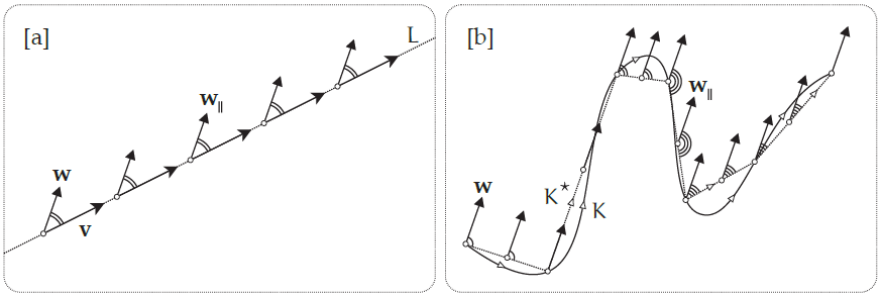
好吧，在曲面上，直线 L 的类似物是测地线 G，因此很自然地猜测欧氏构造推广如下：

> 要沿曲面 S 上测地线 G 以速度 $\mathbf{v}$ 将切向量 $\mathbf{w}$ 平行移动，只需保持 $\mathbf{w}_{\parallel}$ 和 $\mathbf{v}$ 之间的角度在沿 G 移动时恒定不变即可。 (23.1)

这种内蕴定义的方法确实产生与上述三种外在方法相同的平行移动向量，通过组合（22.4）和（22.5）立即可得[练习]。 这在 [23.2] 中有说明。

让我们重复我们的口头禅：请亲自尝试这种构造！ 在这里，我们使用了一根拉长的绳子来强调这种构造的内蕴本质，但实际上，用窄胶带结构来构造测地线通常更容易，因为即使在向你弯曲的曲面部分，胶带也会继续滚动成为测地线。 向你弯曲的曲面，不能在其上拉紧绳子，至少不能从曲面的外侧拉紧。

为了（内蕴地）沿着欧氏平面中的任意曲线 K 平行移动 $\mathbf{w}$，第一步是通过小直线段 $\{L_i\}$ 的序列 $K^\star$ 来近似 K，所有这些小段的长度都小于，比方说，ϵ（见[23.1b]）。 逐次沿每个 L_i 移动并保持角度不变，我们就有了沿 $K^\star$ 平行移动 $\mathbf{w}$ 的方法。 最后，取 ϵ 趋于零的极限，$K^\star$ 变为 K，我们就实现了沿 K 的内蕴的平行移动。



[23.1] [a] 在欧氏平面上，要沿直线 L 在方向 $\mathbf{v}$ 上平行输送 $\mathbf{w}$，保持 $\mathbf{w}_{\parallel}$ 和 $\mathbf{v}$ 之间的夹角不变。[b] 为了沿着一般曲线 K 平行移动，用一系列线段 $K^\star$ 来近似它，并保持 $\mathbf{w}_{\parallel}$ 与每个线段之间的夹角不变。 最后，让每个小段的长度趋于零，使 $K^\star$ 变为 K。

当然，在欧氏平面上这么做有点夸张，在那里我们已经有了全局平行性概念，但关键是现在很清楚我们在曲面上该怎么做：

为了沿着一般曲线 K 将切向量 $\mathbf{w}$ 在曲面 $\mathcal{S}$ 上平行移动，通过测地线段 $\{G_i\}$ 的序列 $K^\star$ 近似 K，每小段长度小于 ϵ，然后沿每个 Gi 移动 $\mathbf{w}$，保持 G_i 的方向和 $\mathbf{w}_\parallel$ 之间的角度不变。 最后，取 ϵ 趋于零的极限，使 $K^\star$ 变为 K。

23.2 内蕴（又名“协变”）导数

当我们沿着曲面 $\mathcal{S}$ 中的曲线 K 平行移动一个向量时，对于 $\mathcal{S}$ 上沿着 K 走路旁观的居民来说，它似乎是不变的。 这种不变性的思想为向量沿曲线移动时变化速度的内蕴测量打开了大门。

假设一个粒子在曲面 $\mathcal{S}$ 上沿曲线 K 以单位速度移动，其在时间 t 的位置为 $p(t)$，单位速度向量为 $\mathbf{v}(t)$。 此外，假设有一个与 $\mathcal{S}$ 相切的，从 $p(t)$ 发出的向量 $\mathbf{w}(t)$，因此 $\mathbf{w}$ 沿 K 处处有定义（尽管在其他地方不一定）。 $\mathbf{w}(t)$ 的内蕴变化率是多少？

好吧，如果 p 在短时间 ϵ 内移动到 $q = p(t+\epsilon)$，那么外在定义的变化率 $\nabla_{\mathbf{v}}\mathbf{w}$ 由下式给出

$$\epsilon\nabla_{\mathbf{v}}\mathbf{w} = \epsilon\mathbf{w}'(t) \asymp \mathbf{w}(q) - \mathbf{w}(p)$$

其中右侧两个向量的差在 $\mathbb{R}^3$ 中非常有意义。但是 $\mathbf{w}(q)$ 存在于 T_q 中，而 $\mathbf{w}(p)$ 存在于 T_p 中，所以这个差不存在于两者任何一个中，因此肯定不是 $\mathcal{S}$ 内蕴的。

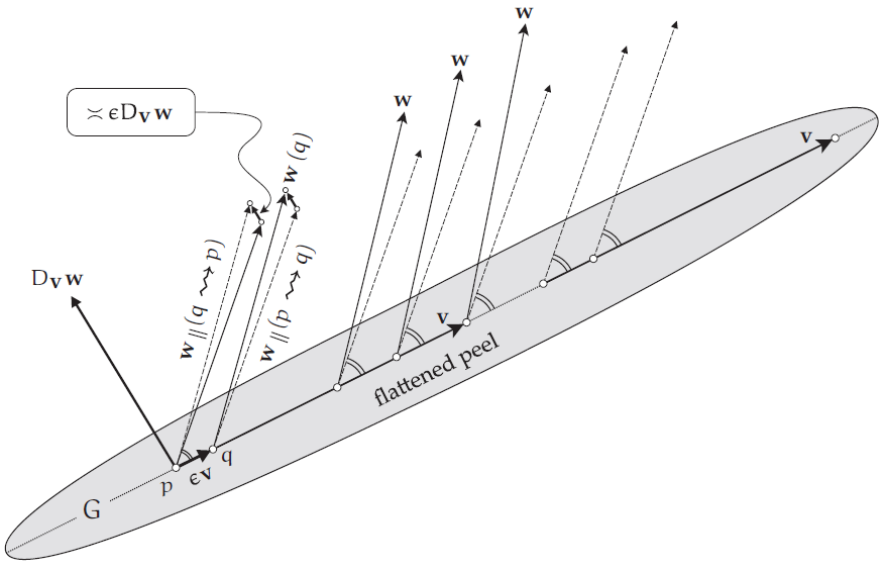
正是 [22.2] 中描述的平行移动允许我们将 $\mathbf{w}(p)$ 带到 q 成为 $\mathbf{w}_\parallel(p \rightsquigarrow q)$，保持它“不变”。 我们现在可以将这个旧值 $\mathbf{w}_\parallel(p \rightsquigarrow q)$ 与新值 $w(q)$ 进行比较，看看它改变了多少。令 $D_{\mathbf{v}}$ 表示沿 K 的内蕴变化率（算子），我们将其称为**内蕴导数**，于是

$$\epsilon D_{\mathbf{v}}\mathbf{w} \asymp \mathbf{w}(q) - \mathbf{w}_\parallel(p \rightsquigarrow q)$$

[23.2] 南瓜上的测地线段 G 首先通过在其表面上拉紧一根绳子来构造。 曲面的任何初始切向量都可以沿着 G 平行移动，只需保持它和 G 之间的角度不变。

由于这两个向量都在 T_q 中，因此它们的差也在。 由于 $D_{\mathbf{v}}\mathbf{w}$ 与 $\mathcal{S}$ 相切，因此可以将其视为 $\mathcal{S}$ 的内蕴特征。

警告：$D_{\mathbf{v}}$ 的标准名称是协变导数。 从历史上看，“协变”一词与 $D_{\mathbf{v}}$ 的坐标表达式在坐标变化下如何变换有关，但由于“协变”对现代学生来说意义不大或根本没有意义，我们会争辩（通过举例！） 是时候使用一个新的、不言自明的名称了。 也就是说，如果我们没有警告读者（截至 2019 年撰写本文时）基本上所有其他微分几何（和物理学）书籍都将 $D_{\mathbf{v}}$ 称为**协变导数**，将是我们的失职。 最后，我们注意到这也称为**列维-奇维塔**（Levi-Civita）**联络**。



[23.3] 内蕴（又名“协变”）导数 $D_{\mathbf{v}}\mathbf{w}$ 测量 $\mathbf{w}$ 沿 $\mathbf{v}$ 的内蕴变化率，如果 $\mathbf{v}$ 是测地线 G 的速度，则特别容易可视化，因为那样 G 周围的剥离细条带铺平成一条直线，而曲面内 $\mathbf{w}_{\parallel}$ 的平行移动变成平面内普通的欧氏平行移动。

为了生动地描述这个内蕴导数，为了简单起见，假设曲线是测地线 G。如果我们剥开 G 周围的条带并将其压在桌子上，它变成了 [23.3] 中所示的直线，G 的切向量 $\mathbf{v}(t)$ 和 $\mathcal{S}$ 的切向量场 $\mathbf{w}(t)$（沿 G 定义）都被压到这个桌面 $\mathbb{R}^2$ 上（就像在 [22.6] 和 [22.7] 中一样）。$\mathbf{w}$ 沿 G 的平行移动（如 [23.2] 中）现在相当于在 $\mathbb{R}^2$ 中沿这条直线的普通欧氏平行移动，保持 $\mathbf{w}_{\parallel}$ 和 $\mathbf{v}$ 之间的角度为常数（如[23.1a]所示）。

从图中可以明显看出，这里沿 G 移动时 $\mathbf{w}$ 的长度增加并在逆时针旋转。$D_{\mathbf{v}}\mathbf{w}$ 的工作是量化发生这种情况的速率，即当我们开始沿着 G 远离 p 时，$\mathbf{w}(t)$ 偏离其初始值（由 $\mathbf{w}_{\parallel}$ 表示）的速率。

或者，这实际上是更标准的定义，我们可以将 $\mathbf{w}(q)$ 并行传输回 p 以得到 $\mathbf{w}_{\parallel}(q \rightsquigarrow p)$，然后将其与原始向量 $\mathbf{w}(p)$ 进行比较（如 [23.3] 中所示）。因此，在 p 处的内蕴导数 $D_{\mathbf{v}}\mathbf{w}$ 从 p 发出，在 T_p 中，由下式给出

$$\boxed{D_{\mathbf{v}}\mathbf{w} \asymp \frac{\mathbf{w}_{\parallel}(q \rightsquigarrow p) - \mathbf{w}(p)}{\epsilon}}$$

(23.2)

图 [23.3] 说明了 $\epsilon = 0.1$ 的几何构造：我们画出了移动 $\mathbf{v}$ 的十分之一所导致的 $\mathbf{w}$ 的变化，然后将该变化向量拉伸十倍。 当然，$\mathrm{D}_{\mathbf{v}}\mathbf{w}$ 的准确值只有在我们将这个数字的极限取为 ϵ 趋于零时才能获得。

在一般情况下，如果 K 不是测地线，我们刚才想象的一切都保持不变，只要我们只剥掉 K 上连接 p 和 q 的非常短的一段周围的曲面。

注意 $\mathbf{w}_{\|}$ 沿 K 的内蕴不变性现在可以通过说它的内蕴导数趋于零来巧妙地表达：

$$\boxed{\mathrm{D}_{\mathbf{v}}\mathbf{w}_{\|} = 0 \quad \Longleftrightarrow \quad \mathbf{w}_{\|} \text{ 沿 } \mathbf{v} \text{ 平行移动}}$$

这是查看内蕴导数的外在方式。 与其将 $\mathrm{D}_{\mathbf{v}}$ 视为测量投影到切平面上的变化率，我们还可以取变化率 $\nabla_{\mathbf{v}}$ 本身的投影。 再次让 $\mathcal{P}$ 表示正交投影到切平面上，（22.1）的这种推广采用形式

$$\mathrm{D}_{\mathbf{v}}\mathbf{w} = \mathcal{P}[\nabla_{\mathbf{v}}\mathbf{w}] = \nabla_{\mathbf{v}}\mathbf{w} - (\mathbf{n}\cdot\nabla_{\mathbf{v}}\mathbf{w})\,\mathbf{n} \tag{23.3}$$

请注意，外在项也可以用形状运算符表示：

$$\boxed{\mathrm{D}_{\mathbf{v}}\mathbf{w} = \nabla_{\mathbf{v}}\mathbf{w} + [\mathbf{w}\cdot\mathrm{S}(\mathbf{v})]\,\mathbf{n}}$$

换句话说，为了获得 $\mathrm{D}_{\mathbf{v}}\mathbf{w}$，我们取 $\mathbb{R}^3$ 中的全变化率 $\nabla_{\mathbf{v}}\mathbf{w}$，然后减去与曲面不相切的部分，从而留下曲面内蕴的部分。

公式（23.3）可用于证明本征导数 $\mathrm{D}_{\mathbf{v}}$ 与全向量导数 $\nabla_{\mathbf{v}}$ 遵循所有相同的线性和莱布尼茨（乘积）规则[练习]，因此如果 a 和 b 是常数，而 f 和 g 是沿 K 定义的标量函数，并且 $\mathbf{x}$、$\mathbf{y}$、$\mathbf{z}$ 沿 K 与 $\mathcal{S}$ 相切，则

$$\mathrm{D}_{\mathbf{v}}[\mathrm{a}\mathbf{x} + \mathrm{b}\mathbf{y}] = \mathrm{a}\mathrm{D}_{\mathbf{v}}\mathbf{x} + \mathrm{b}\mathrm{D}_{\mathbf{v}}\mathbf{y}$$

$$\mathrm{D}_{[\mathrm{fx}+\mathrm{gy}]}\mathbf{z} = \mathrm{f}\mathrm{D}_{\mathbf{x}}\mathbf{z} + \mathrm{g}\mathrm{D}_{\mathbf{y}}\mathbf{z}$$

$$\mathrm{D}_{\mathbf{v}}[\mathrm{f}\mathbf{x}] = \mathrm{f}\mathrm{D}_{\mathbf{v}}\mathbf{x} + [\mathrm{D}_{\mathbf{v}}\mathrm{f}]\mathbf{x} = \mathrm{f}\mathrm{D}_{\mathbf{v}}\mathbf{x} + \mathrm{f}'\mathbf{x}$$

$$\mathrm{D}_{\mathbf{v}}[\mathbf{x}\cdot\mathbf{y}] = [\mathrm{D}_{\mathbf{v}}\mathbf{x}]\cdot\,\mathbf{y} + \,\mathbf{x}\cdot[\mathrm{D}_{\mathbf{v}}\mathbf{y}]$$

（这里的单引号表示沿 K 的时间或距离的微分。）然而，与其通过计算确定这些事实，考虑将 K 周围的条带与向量场 $\mathbf{x}, \mathbf{y}, \mathbf{z}$ 一起展平到桌面上要简单得多，因为那样 $\mathrm{D}_{\mathbf{v}}$ 就是 $\nabla_{\mathbf{v}}$。

内蕴导数为我们之前对测地曲率的讨论提供了新的思路。回忆一下 [11.2]，第 116 页，在曲面上以单位速度行进的粒子的完整加速度 $\kappa \equiv \nabla_{\mathbf{v}}\mathbf{v}$ 可以分解为两个分量：

$$\nabla_{\mathbf{v}}\mathbf{v} = \kappa = \kappa_g + \kappa_n$$

第一个分量 κ_g 是**测地曲率向量**：它是与 $\mathcal{S}$ 相切的加速度分量。它自动垂直于轨迹，并指向 $\mathcal{S}$ 居民感知的曲率中心，大小 $|\kappa_g|$ 是这个圆的曲率； 换句话说，它测量了 $\mathcal{S}$轨迹曲率的内蕴部分。与此相反，**法向曲率向量** κ_n 沿 $\mathbf{n}$ 指向并且对这些居民不可见。

直观地，κ_g 应该是 $\mathbf{v}$ 在曲面内的内蕴旋转速率：

$$\boxed{\kappa_g = D_{\mathbf{v}}\mathbf{v}}$$

从（23.3）可以看出这确实是真的，因为 $\kappa_n = (\mathbf{n}\cdot\nabla_{\mathbf{v}}\mathbf{v})\,\mathbf{n}$。

如果 $\theta_\parallel$ 表示 $\mathbf{v}$ 和沿轨迹平行移动的任意向量 $\mathbf{w}_\parallel$ 之间的角度，那么（见例 4）测地曲率只是 $\mathcal{S}$ 内速度向量相对于“恒定”向量 $\mathbf{w}_\parallel$ 的转动率：

$$\boxed{|\kappa_g| = |D_{\mathbf{v}}\theta_\parallel| = |\theta'_\parallel|} \tag{23.4}$$

测地线对于 $\mathcal{S}$ 的居民来说是一条直线：换句话说，$\kappa_g = 0$。 因此测地线方程采用以下形式，

$$\boxed{\kappa_g = 0 \quad \Leftrightarrow \quad D_{\mathbf{v}}\mathbf{v} = 0} \tag{23.5}$$

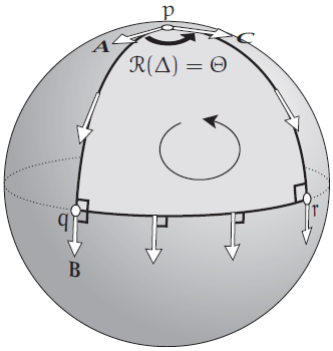
这只是看待 $\mathbf{v}$ 沿其自身平行移动事实的另一种方式，并且对于 $\mathcal{S}$ 的居民来说似乎是恒定的。

另一方面，如果 $\mathcal{S}$ 内的曲线不是测地线，那么（见例 4）当它周围的条带被剥离并压平到平面上时，它会显得弯曲，并且它在平面内的曲率就是 κ_g！

第二十四章 和乐

24.1 示例：球面

考虑 [24.1]，它在半径为 R 的球面上画出了一个测地三角形 Δ，两条子午线之间的角度为 Θ，第三条边是赤道的一部分。



[24.1] 向量 **A** 被围绕测地三角形 Δ 平行移动，它返回到 p 时的净旋转称为 Δ 的**和乐**，写作 $\mathcal{R}(\Delta) = \Theta$。

假设我们从赤道上 q 处的指向南的向量 **B** 开始，然后使用（23.1）将它沿两条不同的测地线平行移动到 p。 如果我们沿子午线测地线 qp 朝正北移动它，我们得到 **A**。而如果我们改为沿测地线赤道段 qr 向东平行移动它（与测地线保持直角），然后沿子午线正北移动 测地线 rp，我们得到了完全不同的向量 **C**。 沿着不同路线的平行移动结果之间的这种差异是我们之前提到的**和乐**（holonomy），它是列维-奇维塔（Levi-Civita）在 1917 年发现的。

以稍微不同的方式看待这个和乐是富有成效的。 假设我们不是沿着两条不同的路线将 **B** 运送到 p，而是假设我们从 p 处的 **A** 开始，然后绕着闭环 p ⇝ q ⇝ r ⇝ p 逆时针平行移动它。 该图表明它返回到 p 时经历了逆时针旋转 $\mathcal{R}(\Delta) = \Theta$。 这就是 Δ 的和乐，现在我们可以引入一般定义：

> 曲面 $\mathcal{S}$ 上的简单闭环 L 的和乐 $\mathcal{R}(\mathrm{L})$ 是 $\mathcal{S}$ 的切向量围绕 L 平行移动一周产生的净旋转。 (24.1)

注意，此定义未指定要平行移动哪个切向量。 原因是（22.4）意味着所有切向量必然一起刚性旋转相同的角度 $\mathcal{R}(\mathrm{L})$。 因此，

> 我们可以将和乐视为当围绕环路平行移动时整个切平面的旋转。

和乐的定义也没有指定我们应该从 L 上的哪里开始。 要了解为什么这也无关紧要，假设不是从 p 处的 **A** 开始，而是从 q 处的 **B** 开始，然后将它平行移动 q ⇝ r ⇝ p ⇝ q。

使用（23.1），你可以直观地证明，在返回 q 时，向量经历了与之前相同的旋转 $\mathcal{R}(\Delta) = \Theta$。更一般地，说服自己和乐与环路的起点（以及初始切向量）无关 [练习]。

接下来，注意在 [24.1] 中，球面上和乐的逆时针方向与我们绕 Δ 移动的方向相匹配。这将证明是因为球面具有正曲率。

如果我们在负曲率的曲面上（平行）移动向量，那么旋转方向将与移动方向相反。 在这一点上，我们强烈建议你凭经验验证这一点，使用测地线的粘条构造在合适的水果或蔬菜的负弯曲小片上创建测地三角形。 然后，你可以通过逐次与每条边保持恒定角度来轻松地将牙签绕三角形平行移动。

不仅 $\mathcal{R}(\mathrm{L})$ 的符号由 Δ 内的曲率决定，它的大小也是！ 球面的恒定曲率是 $\mathcal{K} = (1/\mathrm{R}^2)$，所以 Δ 内的总曲率是

$$\mathcal{K}(\Delta) = \iint_\Delta \mathcal{K}\,\mathrm{d}\mathcal{A} = \frac{1}{\mathrm{R}^2}\iint_\Delta \mathrm{d}\mathcal{A} = \frac{1}{\mathrm{R}^2}[\mathrm{R}^2\Theta] = \Theta$$

因此，对于 $\mathrm{L} = \Delta$，

$$\boxed{\mathcal{R}(\mathrm{L}) = \mathcal{K}(\mathrm{L})} \tag{24.2}$$

正如我们将要看到的，这并非偶然——对于任何曲面 $\mathcal{S}$ 上的任何简单环路 L 都是如此！ 证明这个结果（在下一章中）将为我们提供一个看似万能的钥匙，能够解开我们遇到的一些最深奥的谜团。 它将解锁绝妙定理。 它将解锁全局高斯-博内定理的内蕴本质。 它将解锁度量曲率公式（4.10），即我们未来提供的“星际迷航相位器”。 它对更高维度的推广将解锁位于爱因斯坦弯曲时空引力理论核心的黎曼曲率。

事实上，这个列表还在继续，远远超出了本书的范围。 它包括迈克尔·贝瑞（Michael Berry）爵士在 1983 年对量子力学中现在称为贝瑞相（以及物理学中的其他“几何相”）的非凡发现（参见 Shapere 和 Wilczek (1989) 以及 Berry (1990)）。 有关和乐在物理学中的应用的一些可爱甄选，请参阅 Berry (1991)（但请注意，我们称其为和乐（holonomy），物理学家通常称其为异和乐（anholonomy））。

24.2 一般测地三角形的和乐

图[24.2]画出了一般曲面上的一般测地三角形 Δ，内角为 θ_i，外角为 φ_i，所以

$$\theta_i + \varphi_i = \pi \tag{24.3}$$

我们知道和乐 $\mathcal{R}(\Delta)$ 与围绕它平行移动的向量无关，现在我们借用这种自由度来做出一个选择，这将使答案生动而清晰：我们选择 Δ 第一条边的切向量 $\mathbf{v}$。

平行移动保持 $\mathbf{v}$ 与第一条边相切，所以当它到达终点时，它与第二条边成 φ_2 角。 由于第二条边也是测地线，因此当 $\mathbf{v}$ 沿它平行移动时，角度 φ_2 保持不变。 因此，当它到达第二条边的末端时，它与第三条也是最后一条边成 $(\varphi_2+\varphi_3)$ 角，并且这个角度在它沿该边移动时保持不变，最终返回其起点时成为 $\mathbf{v}_{\parallel}$ ，与初始边成角度 $(\varphi_1+\varphi_2+\varphi_3)$。 因此我们可以看到和乐有

$$\boxed{\mathcal{R}(\Delta)=2\pi-(\varphi_1+\varphi_2+\varphi_3)} \tag{24.4}$$

在我们对霍普夫的环流定理的讨论中，我们首先注意到（参见 [17.1]，第 176 页）如果一个粒子绕欧氏三角形 Δ 运动，那么速度矢量的旋转是 $(\varphi_1+\varphi_2+\varphi_3)=2\pi$。 因此，和乐公式 (24.4) 测量速度的总旋转 $(\varphi_1+\varphi_2+\varphi_3)$ 与欧氏预测的 2π 相差多少。

在整本书中，我们使用了一种不同的方法来度量曲面上的测地三角形与欧氏三角形的偏离程度，即它的角盈 $\mathcal{E}(\Delta)$。 但实际上这两个概念上不同的 Δ 内曲率度量是相等的！ 为了看到这一点，我们将（24.3）与（24.4）结合起来：

$$\mathcal{R}(\Delta)=2\pi-[(\pi-\theta_1)+(\pi-\theta_2)+(\pi-\theta_3)]=\theta_1+\theta_2+\theta_3-\pi$$

所以，

$$\boxed{\mathcal{R}(\Delta)=\mathcal{E}(\Delta)} \tag{24.5}$$

以上所有内容都可以轻松地从测地线三角形推广到测地线 m 边形 $\mathrm{P_m}$。 首先，[24.2] 显然可以推广得到

$$\boxed{\mathcal{R}(\mathrm{P_m})=2\pi-\sum_{i=1}^{m}\varphi_{\mathrm{i}}} \tag{24.6}$$

另一方面，$\mathrm{P_m}$ 的角盈由（18.4）给出：

$$\mathcal{E}(\mathrm{P_m})=\sum_{i=1}^{m}\theta_{\mathrm{i}}-(m-2)\pi$$

再次使用（24.3），我们发现两个看似不同的 $\mathrm{P_m}$ 内曲率度量实际上是相等的：

$$\boxed{\mathcal{R}(\mathrm{P_m})=\mathcal{E}(\mathrm{P_m})} \tag{24.7}$$

24.3 和乐是可加的

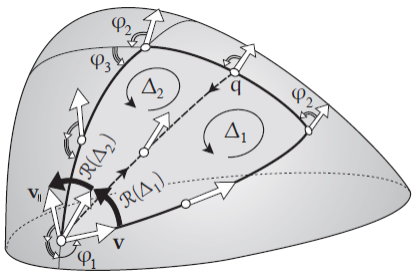
回忆一下 [2.8]，如果我们将 Δ 分成两个测地三角形 Δ_1 和 Δ_2，那么角盈 $\mathcal{E}$ 是可加的：

$$\mathcal{E}(\Delta) = \mathcal{E}(\Delta_1) + \mathcal{E}(\Delta_2)$$

从 (24.5) 可知，和乐 $\mathcal{R}$ 也是可加的。 然而，我们认为 $\mathcal{R}$ 在这两个概念中更基本，因此与其将其可加性视为继承自 $\mathcal{E}$，我们应该尝试直接理解这一点。

图 [24.3] 是一个直接的几何证明，证明 $\mathcal{R}$ 是可加的：

$$\boxed{\mathcal{R}(\Delta) = \mathcal{R}(\Delta_1) + \mathcal{R}(\Delta_2)} \tag{24.8}$$



[24.3] **和乐是可加的**。 通过插入虚线测地线，测地线三角形 Δ 被分成 Δ_1 和 Δ_2。 Δ_1 的第一条边的切向量 $\mathbf{v}$ 围绕 Δ_1 平行移动，然后围绕 Δ_2 平行移动。 我们看到沿着虚线测地线来回的平行移动“彼此抵消”，因此 $\mathcal{R}(\Delta) = \mathcal{R}(\Delta_1) + \mathcal{R}(\Delta_2)$。

在这里，我们将虚线测地线插入到原始测地线三角形 Δ 中，从而将其分成两个测地三角形 Δ_1 和 Δ_2。Δ_1 的第一条边的切向量 $\mathbf{v}$ 绕 Δ_1 平行移动，返回原点时旋转了 $\mathcal{R}(\Delta_1)$。然后它绕 Δ_2 平行移动，返回原点时旋转了 $\mathcal{R}(\Delta_2)$。 因此，围绕两者平行移动后的总旋转为 $[\mathcal{R}(\Delta_1) + \mathcal{R}(\Delta_2)]$，如图所示。

但是因为 Δ_1 的最后一条边（从 q 开始）也是 Δ_2 的第一条边（以 q 结束），我们沿着虚线测地线在相反的方向平行移动两次，所以向量返回 q 时不变。因此，围绕 Δ_1 和 Δ_2 依次平行移动向量等价于围绕 Δ 平行移动向量，如将要显示的。

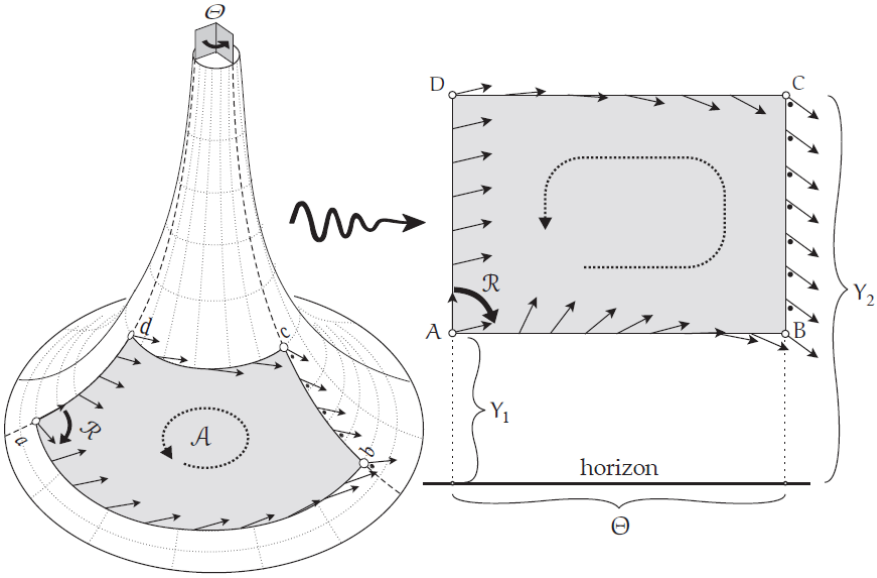
从这个意义上说，沿着同一条曲线来回的平行移动“彼此抵消”，即使该曲线不是测地线。

24.4 示例：双曲平面

我们通过将平行移动的概念应用于伪球（通过贝尔特拉米-庞加莱半平面模型）来结束

本章，以获得双曲平面恒定负曲率的新的、简单的内蕴几何演示[162]。 为了做到这一点，我们现在假设基本事实（24.2）成立，即环路的和乐度量它内部的总曲率（将在下一章证明）。

在我们开始演示之前，我们先做一个重要的观察：正如我们最初使用角盈给出一个点的曲率 $\mathcal{K}(\mathrm{p})$ 的内蕴定义一样（见（2.1）），所以现在可以以相同的方式使用（24.2）来找到点 p 处的曲率。



[24.4] 伪球面（左）上的“矩形”abcd（面积 $\mathcal{A}$）共形映射到贝尔特拉米-庞加莱上半平面（右）中的 ABCD。 a 处的图示向量绕 abcd 逆时针平行移动，它返回时顺时针旋转了 $\mathcal{R}$。地图的共形性保证了地图中的平行移动向量经历相同的旋转 $\mathcal{R}$。

如果 $\mathrm{L_p}$ 是围绕 p 的一个小环路，那么我们可以将（24.2）应用于 $\mathrm{L_p}$，因为它缩小到 p 以找到 p 处的曲率：

$$\boxed{\mathcal{K}(\mathrm{p}) = \lim_{\mathrm{L_p}\to\mathrm{p}} \frac{\mathcal{R}(\mathrm{L_p})}{\mathcal{A}(\mathrm{L_p})} = \mathrm{p}\text{ 处每单位面积的和乐}} \tag{24.9}$$

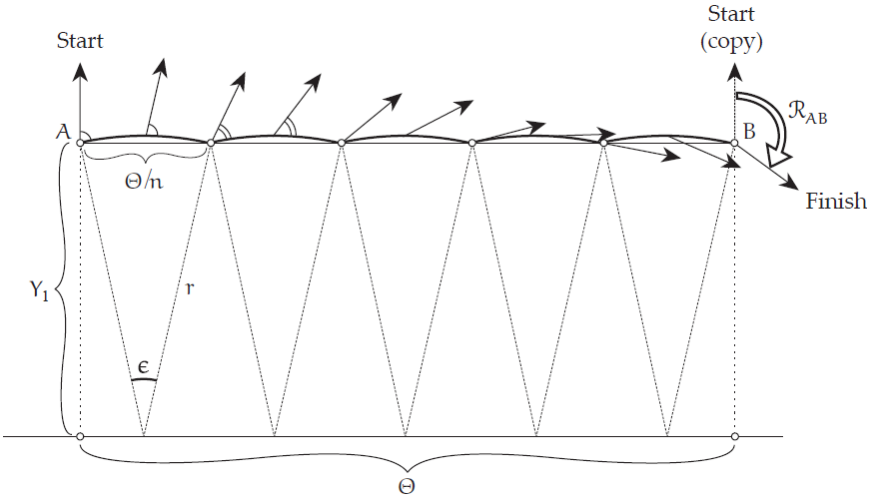
我们现在可以回到手头的问题。在半径为 R 的伪球面上，考虑由测地线（曳物线母线）的垂直线段 ad 和 bc（Θ 是从第一到第二个测地线的角度）以及非测地线水平圆弧 ab 和 cd 所围成的“矩形”abcd（逆时针方向）（参见 [24.4] 的左侧）。 如图所示，让我们绕 abcd 平行移动一个向量以查明其中的总曲率。

[24.4]的右侧描绘了贝尔特拉米-庞加莱模型中的共形像：abcd 被映射到顶点为 $\mathrm{A} = (\mathrm{x}, \mathrm{Y_1}), \mathrm{B} = (\mathrm{x}+\Theta, \mathrm{Y_1}), \mathrm{C} = (\mathrm{x}+\Theta, \mathrm{Y_2}), \mathrm{D} = (\mathrm{x}, \mathrm{Y_2})$。 因此，使用（5.6），伪球面上矩形 abcd 的面积 $\mathcal{A}$ 是

$$\mathcal{A} = \int_{x=0}^{x=\Theta} \int_{\mathrm{Y_1}}^{\mathrm{Y_2}} \frac{\mathrm{R}^2\,\mathrm{d}x\,\mathrm{d}y}{y^2} = \mathrm{R}^2\Theta\left[\frac{1}{\mathrm{Y_1}} - \frac{1}{\mathrm{Y_2}}\right] \tag{24.10}$$

[162]以下论证以前出现在 Needham (2014) 中。

在 a 处，我们选择了一个沿 ad 向上指向伪球面的初始向量。 当我们沿 ab 平行移动它时，它相对于运动方向顺时针旋转； 沿 bc 保持恒定的图示角度 • 与运动方向（因为它是测地线）；沿着 cd 它相对于运动方向逆时针旋转，但不像在 ab 上转得那样多； 最后，它与测地线 da 保持恒定的角度，经过 $\mathcal{R}$ 的负净旋转后返回到的 a。



[24.5] A 处的初始垂直向量沿贝尔特拉米-庞加莱半平面中的水平欧氏线段 AB 平行移动。 为此，我们用 n 个测地线段（以地平线 $y = 0$ 为中心的圆弧）来近似 AB，然后连续与每个测地线段保持恒定的角度。 最后，我们让 n 趋于无穷大。

由于贝尔特拉米-庞加莱映射是共形的，当向量围绕 ABCD 移动时，它会经历相同的净旋转 $\mathcal{R}$。但是，正如我们现在解释的，该映射的关键优势是它使我们能够看见[163]这种旋转实际上是什么。

将欧氏长度 Θ 的非测地水平线段 AB 分成 n 个长度 (Θ/n) 的小段。 接下来，如 [24.5] 中所示，用测地线段近似这些小段：回想一下，这些是以地平线为中心的圆弧。 令 ϵ 是每个这样的弧在地平线上对着的角度（如图所示）。

由于初始垂直的起始向量沿着第一个测地线段平行移动，它与该段的角度保持不变，因此它旋转了角度 −ϵ。 同样，对于每个连续的段，在遍历完所有 n 个段之后，从开始到结束的总旋转为 −nϵ。 但是由于

$$r\epsilon \asymp \frac{\Theta}{n} \quad 和 \quad r \asymp Y_1$$

我们推导出向量在地图中旋转的总角度是

$$\mathcal{R}_{AB} \asymp -n\epsilon \asymp -\frac{\Theta}{r} \asymp -\frac{\Theta}{Y_1}$$

[163] 需要进行少量计算，但由于我们能够像处理普通等式（涉及 =）一样操作“终极等式”（涉及 ≍），这大大简化了计算。

同样可以导出 $\mathcal{R}_{\mathrm{CD}} = (\Theta/\mathrm{Y}_2)$。由于向量不沿测地线 BC 或 DA 旋转，我们基于（24.10）推导出返回 A 时的净旋转为

$$\mathcal{R} = \mathcal{R}_{\mathrm{AB}} + \mathcal{R}_{\mathrm{CD}} = -\frac{\Theta}{\mathrm{Y}_1} + \frac{\Theta}{\mathrm{Y}_2} = \left[-\frac{1}{\mathrm{R}^2}\right]\mathcal{A}$$

因此，

$$\text{单位面积的旋转} = -\frac{1}{\mathrm{R}^2}$$

这个答案与矩形的大小、形状和位置无关的事实证明，通过（24.9），双曲平面确实具有恒定负的内蕴曲率 $-1/\mathrm{R}^2$，证明完毕。

如果你想亲自尝试在特定曲面（圆锥和球面）上执行平行移动，请尝试练习 5。

第二十五章 绝妙定理直观的几何证明

25.1 引言

自从高斯于 1827 年第一次用*绝妙定理*（Theorema Egregium）震惊世界以来，已经发现了几个它的证明。到目前为止，基本上所有教科书中都可以找到最常见的证明，是高斯原始证明的变体：执行冗长的计算，一路上昏昏欲睡，然后醒来发现自己正盯着外在曲率 $\mathcal{K}_{\text{ext}}$ 的公式，结果证明该公式*仅取决于度量*，从而确定 $\mathcal{K}_{\text{ext}} = \kappa_1\kappa_2$ 实际上是曲面内蕴的。

毫无疑问，这是一个滴水不漏的证明，就像高斯在 1827 年第一次写下它时一样，但我们对为什么结果为真的一无所知！

马塞尔·伯格（Marcel Berger）教授 (1927–2016) 是 20 世纪后期最重要的几何权威之一，在其权威（且出色）的著作《*黎曼几何的全景图*》（A Panoramic View of Riemannian Geometry）（Berger 2003，第 106 页）中，给出各种已知证明的摘要，并得出结论："*据我们所知，绝妙定理至今没有简单几何证明。*"[164] 本章的目的是通过提供这样一个简单的几何证明来纠正这种情况。[165]

与其试图直接理解绝妙定理，不如试图理解高斯的最初发现，即他所谓的"美丽定理"（13.1）。 为方便起见，我们在此重申结果，再次引用高斯自己的话，正如他在 1816 年的私人笔记本中记录的发现一样：

> "**美丽定理**。*如果曲面上的固定图形在空间中呈现不同的形状，则图形的球面像的面积总是相同的。*"

在第 13.3 节中，我们解释了局部*绝妙定理*可以轻松直观地从美丽定理中得出，只需让图形缩小到一个点即可。 所以，理解美丽定理就是理解*绝妙定理*。

但是，正如我们在同一部分中所指出的，东布罗夫斯基（Dombrowski (1979)）发现即使在他的私人笔记本中，高斯也既没有留下任何关于他如何发现此结果的痕迹，也没有给出任何证明的提示（假设他有证明） .

从目前的情况应该清楚，*我们将使用平行移动来证明美丽定理*，这个概念在 1816 年一

---

[164] 一旦我发现了我将要提出的简单几何证明，我就把它寄给了班乔夫、伯杰和彭罗斯，他们之前都没有见过它。 尤其是，母语是法语的伯杰回了一封电子邮件，"我很高兴能够祝贺你这个惊人的证明，它被错过了一个多世纪，这个结果不仅是几何中最美丽的结果之一，而且也开启了陈省身等人的整个事业。"

[165] 当这本书接近完成时，我了解到同样的见解之前由康奈尔大学的 DavidW. Henderson 教授（1998 年，问题 6.3）发表过。我正要写信给他时，才得知他被车撞死了，就在几天前，也就是 2018 年 12 月 20 日。所以，为了保持记录的正确性，这一发现应该归于已故的亨德森教授（除非有我不知道的先驱）。不管它的出处如何，我希望我在这里的可视化（亨德森没有提供）将有助于使这个简单直观的证明更广为人知。 （而且，在一个非常个人的观点上，优先权并不是一切：在被原始白雪包围的塞拉山脉中，令人吃惊的、出乎意料的清晰闪光是我一生中最快乐的时刻之一。）

个多世纪之后被列维-奇维塔（Levi-Civita）发现。因此，我们希望你会发现这样直观和简单的平行移动证明，不可能是高斯的原始方法。

我们强调，高斯在 1827 年开发用于分析一般曲面的计算装置之前整整 11 年就发现了美丽定理。 因此，这个谜仍然存在，随之而来的还有一种诱人的可能性，即可能存在一种完全不同的、甚至更简单的方法——但这可能必须等高斯本人重生才能揭晓了。

25.2 定义的一些符号和提示

设 $\mathcal{S}$ 为具有单位法向量场 $\mathbf{n}$ 的曲面，$\mathbb{S}^2$ 为单位球面，$\mathrm{n}:\mathcal{S}\mapsto\mathbb{S}^2$ 是球面映射。

令 $\mathcal{K}_{\mathrm{ext}}$ 为外在曲率，定义为球面图的局部（有符号）区域扩展因子，如（12.3）所述。

让波浪号（~）表示在 $\mathbb{S}^2$ 上定义的量。 因此，例如，$\mathcal{A}$ 表示 $\mathcal{S}$ 上的面积，而 $\tilde{\mathcal{A}}$ 表示 $\mathbb{S}^2$ 上的面积。

令 $\mathcal{E}(\Delta)$ 是 $\mathcal{S}$ 上测地三角形 Δ 的角盈，或者更一般地令 $\mathcal{E}(\mathrm{P_m})$ 表示测地 m 边形 $\mathrm{P_m}$ 的角盈，由（18.4）给出。

设 L 是 $\mathcal{S}$ 的一个区域 Ω 的简单环路，令 $\mathcal{R}(\mathrm{L})$ 表示 $\mathcal{S}$ 的切向量 $\mathbf{w}$ 绕 L 逆时针平行移动的净旋转（和乐角），旋转方向由 $\mathbf{n}$ 指着我们的眼睛确定。 为了便于可视化，让我们假设 $\mathcal{K}_{\mathrm{ext}}$ 在整个 Ω 中具有单一符号（总是正的，或者总是负的），因此 $\mathrm{n}(\Omega)$ 不包含任何折叠（请参阅第 17.5 节中的讨论）。

同样，让 $\tilde{\mathcal{R}}(\tilde{\mathrm{L}})$ 表示 $\mathbb{S}^2$ 的切向量绕 L 在 $\mathbb{S}^2$ 上的像 $\tilde{\mathrm{L}}\equiv\mathrm{n}(\mathrm{L})$ 平行移动的和乐。

最后，令 $\mathcal{K}_{\mathrm{ext}}(\Omega)$ 表示 Ω 内外在曲率总量，

$$\mathcal{K}_{\mathrm{ext}}(\Omega)=\iint_{\Omega}\mathcal{K}_{\mathrm{ext}}\,\mathrm{d}\mathcal{A}$$

并令 $\mathcal{K}(\Omega)$ 表示 Ω 内内蕴曲率的总量，

$$\mathcal{K}(\Omega)=\iint_{\Omega}\mathcal{K}_{\mathrm{ext}}\,\mathrm{d}\mathcal{A}=\mathcal{R}(\mathrm{L})$$

25.3 迄今为止的故事

让我们简要回顾一下我们迄今实际知道的内容。 这一点很重要，因为在这本书中，我们早在我们能够证明或解释未来的结果之前，就已经自由地引用（和使用）它们。 尽管如此，我们一直都小心翼翼地避免一点循环推理的味道。

然而，很有可能在我们的时间旅行阐述过程中，已经建立的和尚未证明的之间的界限可能变得模糊。 那么，这里是我们在证明美丽定理时需要的少数几个事实，所有这些事实都已经被正确地建立了：

- 由于 $\mathcal{K}_{\text{ext}}$ 被定义为球面映射的局部（有符号）面积扩展因子，因此 $\mathcal{S}$ 上 Ω 内的外在曲率总量是其像在 $\mathbb{S}^2$ 上的面积 $\widetilde{\Omega}$：

- 正如我们在（12.8）中证明的那样，外在曲率可以表示为主曲率的乘积：

$$\mathcal{K}_{\text{ext}} = \kappa_1\kappa_2$$

回想一下，我们最初通过检查 n 对与主方向对齐的小矩形的影响来证明这一点。 然而，后来（15.8）我们确定所有收缩形状最终都会经历相同的膨胀：$\kappa_1\kappa_2$。

因此像 $\widetilde{\Omega} = \mathrm{n}(\Omega)$ 在 $\mathbb{S}^2$ 上的面积也可以表示为，

$$\tilde{\mathcal{A}}(\widetilde{\Omega}) = \mathcal{K}_{\text{ext}}(\Omega) = \iint_{\Omega} \kappa_1\kappa_2 \, \mathrm{d}\mathcal{A}$$

- 哈里奥特（Harriot）1603 年关于在半径为 R 的球面上的结果（1.3）暗示（令$\mathrm{R} = 1$）在 $\mathbb{S}^2$ 上我们有 $\tilde{\mathcal{E}}(\widetilde{\Delta}) = \tilde{\mathcal{A}}(\widetilde{\Delta})$，这很容易推广到测地线 m 边形 $\tilde{\mathrm{P}}_{\mathrm{m}}$:

$$\tilde{\mathcal{E}}(\tilde{\mathrm{P}}_{\mathrm{m}}) = \tilde{\mathcal{A}}(\tilde{\mathrm{P}}_{\mathrm{m}}\text{的内部})$$

- 但是在 $\mathbb{S}^2$ 上，我们在（24.5）中也证明了 $\tilde{\mathcal{R}}(\widetilde{\Delta}) = \tilde{\mathcal{E}}(\widetilde{\Delta})$，并且在（24.7）中我们将其推广到测地线 m 边形。 因此，如果 $\tilde{\mathrm{P}}_{\mathrm{m}}$ 是 $\mathbb{S}^2$ 上的测地线 m 边形，那么围绕 $\tilde{\mathrm{P}}_{\mathrm{m}}$ 平行移动的向量的净旋转［和乐］与其包围的面积相同：

$$\tilde{\mathcal{R}}(\tilde{\mathrm{P}}_{\mathrm{m}}) = \tilde{\mathcal{A}}(\tilde{\mathrm{P}}_{\mathrm{m}}\text{的内部})$$

- 如果 $\tilde{\mathrm{L}}$ （通常是非测地线）是 $\mathcal{S}$ 上简单环路 L 在 $\mathbb{S}^2$ 上的像，我们可以用测地线 m 边形近似 $\tilde{\mathrm{L}}$，然后让 $\mathrm{m} \to \infty$。 因此，综合上述结果，我们可以将相关的已知事实总结如下：

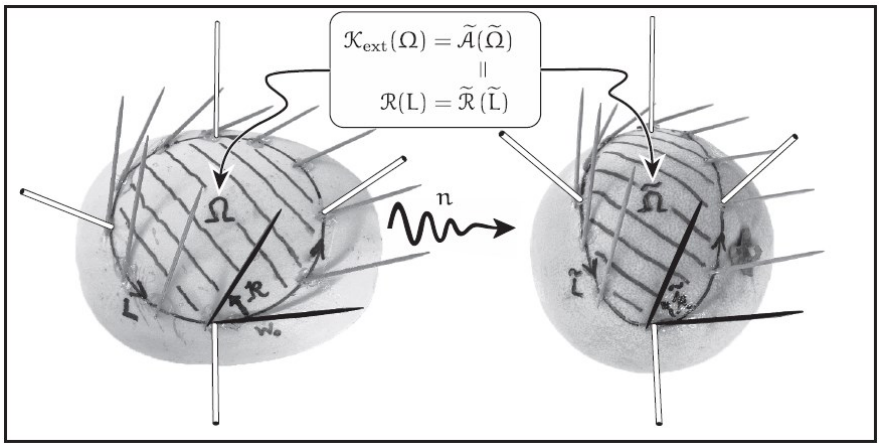
迄今为止的故事。 如果 $\mathcal{S}$ 上的简单环路 L（包围单符号外曲率区域 Ω）由 n 映射到 $\tilde{\mathrm{L}}$，则 $\mathcal{S}$ 上 Ω 内的外在曲率总量是 $\mathbb{S}^2$上 $\widetilde{\Omega}$ 的（有符号）面积，以 $\tilde{\mathrm{L}}$ 为界，而这反过来等于 $\mathbb{S}^2$ 的切向量围绕 $\tilde{\mathrm{L}}$ 平行移动的和乐：

$$\iint_{\Omega} \kappa_1\kappa_2 \, \mathrm{d}\mathcal{A} = \mathcal{K}_{\text{ext}}(\Omega) = \tilde{\mathcal{A}}(\widetilde{\Omega}) = \tilde{\mathcal{R}}(\tilde{\mathrm{L}}) \qquad (25.1)$$

25.4 球面投影保持平行移动

考虑 [25.1]，它画出了 $\mathcal{S}$ 上的简单环路 L（左侧的黄色南瓜）及其在 $\mathbb{S}^2$（右侧的柚子）上的像 $\tilde{\mathrm{L}} = \mathrm{n}(\mathrm{L})$。 p 和 n(p) 处的切平面是平行的，因此任何 $\mathcal{S}$ 上的沿 L 定义的切向量场 $\mathbf{w}$ 在从 p 移植到 n(p) 之后，自动也是沿 $\tilde{\mathrm{L}}$ 的 $\mathbb{S}^2$ 的切向量场。

虽然 L 的形状对下面的论证无关紧要，但在这里我们选择构建一个内蕴的圆，使用一根拉长的绳子（未显示）来画出与固定点相距恒定距离的点的轨迹。 为了帮助可视化球面映射对 L 的影响，我们在沿 L 的四个等距点建立了法线； 它们在 $\mathbb{S}^2$ 上的像点可以通过它们具有相同的法线来识别，如图所示。 由于主曲率的差异，请注意该圆的球面像在 $\mathbb{S}^2$ 上被相当显著地拉伸成椭圆形。



[25.1] **美丽定理的几何证明**。 $\mathcal{S}$（左）上 Ω 内的总外曲率 $\mathcal{K}_{\mathrm{ext}}(\Omega)$ 等于其在 $\mathbb{S}^2$（右）上的球面像 $\widetilde{\Omega}$ 的面积 $\widetilde{\mathcal{A}}(\widetilde{\Omega})$。 并且，$\widetilde{\mathcal{A}}(\widetilde{\Omega}) = \widetilde{\mathcal{R}}(\tilde{\mathrm{L}})$。 但是**球面映射保持平行移动**，所以 $\widetilde{\mathcal{R}}(\tilde{\mathrm{L}}) = \mathcal{R}(\mathrm{L})$。 因此，总**外在**曲率 $\mathcal{K}_{\mathrm{ext}}(\Omega)$ 等于 $\mathcal{S}$ 上**内蕴**定义的和乐 $\mathcal{R}(\mathrm{L})$。

现在假设 $\mathbf{w}$ 不是沿 L 的 $\mathcal{S}$ 的任意切向量场，而是如图所示，而是通过沿 L 平行移动初始切向量 $\mathbf{w}_0$ 以生成 $\mathbf{w}_\parallel$ 来获得的。 $\mathbf{w}_0$ 的选择对论证无关紧要，但在 [25.1] 中我们选择它与 L 相切。

现在，我们来到了这件事出人意料的简单而优雅的关键所在。 根据平行移动的定义，当我们沿 L 行进时，$\mathbf{w}_\parallel$ 的变化率总是垂直于 $\mathcal{S}$，即变化率沿 $\mathbf{n}$ 方向。 而这意味着 $\mathbf{w}_\parallel$ 的变化率在 n(p) 处也垂直于 $\mathbb{S}^2$，因为 p 处 $\mathcal{S}$ 的法线 $\mathbf{n}$ 与 n(p) 处 $\mathbb{S}^2$ 的法线相同。 换句话说，相同的向量 $\mathbf{w}_\parallel$——从 p 移植到 n(p)——在球面上自动沿着 $\tilde{\mathrm{L}}$ 平行移动！

让我们以不同的方式提出相同的观点，希望（想方设法）这个至关重要的结果将在此过程中变得非常清晰。 我们的平行移动方法 [22.2] 要求我们沿着 L 平行于 $\mathbb{R}^3$ 中其自身移动 $\mathbf{w}_\parallel$，同时随着我们的前进连续地投影回曲面（或者，技术上，投影到切平面 $\mathrm{T_p}$），从而获得 $\mathbf{w}_\parallel$。 现在从 $\mathcal{S}$ 上的 p 处迁移（$\mathbb{R}^3$ 中的简单平移）$\mathbf{w}_\parallel$ 和 $\mathrm{T_p}$ 到 $\mathbb{S}^2$ 上的 n(p) 处。 迁移后的切平面只是在 n(p) 处 $\mathbb{S}^2$ 的切平面，因此投影到其中也会产生 $\mathbb{S}^2$ 上的平行移动。

总之，如图所示，

球面映射保持平行传输。当 $\mathbf{w}_{\parallel}$ 在 p 处沿 L 相对于曲面 $\mathcal{S}$ 平行移动，完全相同的向量 $\mathbf{w}_{\parallel}$ 在 n(p) 处也自动相对于 $\mathbb{S}^2$ 沿 $\tilde{\mathrm{L}} = \mathrm{n}(\mathrm{L})$ 平行移动。 (25.2)

25.5 美丽定理与绝妙定理的证明

在围绕 L 转了一整圈后，我们现在可以从字面上看到曲面上的净旋转（和乐）等于球面上的净旋转：

$$\boxed{\mathcal{R}(\mathrm{L}) = \tilde{\mathcal{R}}(\tilde{\mathrm{L}})}$$

但是，根据“迄今为止的故事”（25.1），球面上的净旋转只是球面上围住的面积，而后者又是由 L 围住的曲面 $\mathcal{S}$ 上外在曲率的总量：

$$\boxed{\mathcal{R}(\mathrm{L}) = \tilde{\mathcal{R}}(\tilde{\mathrm{L}}) = \tilde{\mathcal{A}}(\tilde{\Omega}) = \mathcal{K}_{\mathrm{ext}}(\Omega) = \iint_{\Omega} \kappa_1\kappa_2 \, \mathrm{d}\mathcal{A}}$$

由于 $\mathcal{R}(\mathrm{L})$ 是在 $\mathcal{S}$ 上内蕴定义的，它在 $\mathcal{S}$ 的等距下是不变的，因此我们证明了高斯 1816 年的原始美丽定理。

如果我们把 $\mathrm{L_p}$ 看作是 $\mathcal{S}$ 上围绕 p 的一个小环路，并设 Ω_{p} 是它所围住的区域，然后将 $\mathrm{L_p}$ 缩小到 p，我们就可以得到更标准的绝妙定理的局部陈述：

$$\boxed{\kappa_1\kappa_2 = \lim_{\mathrm{L_p}\to\mathrm{p}} \frac{\mathcal{R}(\mathrm{L_p})}{\mathcal{A}(\Omega_{\mathrm{p}})}}$$

或者，假设我们取 $\mathrm{L_p}$ 为一个包含 p 的小测地三角形 Δ_{p}，让 $\mathcal{A}(\Delta_{\mathrm{p}})$ 表示 Δ_{p} 内部的面积。 然后，凭借（24.5），我们还恢复了绝妙定理的原始形式，根据曲率的原始内蕴定义（2.1）：

$$\boxed{\kappa_1\kappa_2 = \lim_{\Delta_{\mathrm{p}}\to\mathrm{p}} \frac{\mathcal{E}(\Delta_{\mathrm{p}})}{\mathcal{A}(\Delta_{\mathrm{p}})} = \mathcal{K}(\mathrm{p})}$$

由于前两个方程中任一右侧的量是 $\mathcal{S}$ 的内蕴量，因此左侧的外在定义曲率 $\kappa_1\kappa_2$ 也必

然——出乎意料地——在等距下保持不变。

因此，我们（终于！）对我们在 [13.1] 和 [13.2] 中首次观察到的经验现象得出了令人满意的几何证明。

第二十六章 全局高斯-博内定理的第四个（和乐）证明

26.1 引言

回想一下，随着第三幕的落幕，已经出现了不少于三种对 GGB 的不同证明，但所有这些证明都被迫依赖于外在几何。

直到现在，有了平行移动，我们终于能够实现内蕴证明 GGB 的承诺。 接下来的优雅论证完全归功于霍普夫（1956, pp. 112 – 113），但我们将解释一些被霍普夫的原始演示视为理所当然的重要细节，更重要的是，我们将以一种霍普夫的阐述没有的方式明确地形象化证明。

霍普夫对 GGB 内蕴证明的一个重要好处是，它同时为我们提供了庞家莱-霍普夫定理（19.6）的全新证明。

26.2 沿开曲线的和乐？

和乐（Holonomy），正如我们在（24.1）中定义的那样，是一个仅对闭环 L 有意义的概念。我们围绕 L 平行移动初始向量 $\mathbf{w}_0$，一路走来生成 $\mathbf{w}_\parallel$，当 $\mathbf{w}_\parallel$ 回到起点，我们可以将它与初始向量 $\mathbf{w}_0$ 进行比较，以查看它旋转的角度 $\mathcal{R}(\mathrm{L})$。

霍普夫论证的第一步是将和乐概念推广到开放曲线 K。显然我们应该将其定义为曲面内沿 K 平行移动时 $\mathbf{w}_\parallel$ 的旋转。但是相对于什么的旋转？

回想一下，我们之前在尝试定义曲面 $\mathcal{S}$ 上向量场 $\mathbf{F}$ 的奇点 s 的指数 $\mathcal{I}_\mathrm{F}(\mathrm{s})$ 时遇到了类似的问题。 正如我们在 [19.7] 中说明的那样，我们的答案是引入一个基准向量场 $\mathbf{U}$，唯一的要求是 $\mathbf{U}$ 在围绕 s 的环路 L 上或内没有任何奇点。 这允许我们将指数定义为当我们沿着环路 L 绕 s 转时 $\mathbf{F}$ 相对于 $\mathbf{U}$ 的转数，即（19.5）：

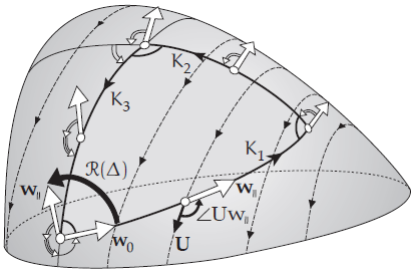
$$2\pi\mathcal{I}_\mathrm{F}(\mathrm{s}) = \delta_\mathrm{L}(\angle\mathrm{UF})$$

正如我们所证明的，这个指数的定义确实是明确定义的，也就是说向量场 $\mathbf{U}$ 的选择无关紧要，尽管沿 L 的角度 $\angle\mathrm{UF}$ 的精确变化确实取决于 $\mathbf{U}$ 场的具体选择。

现在让我们将同样的想法应用到和乐的概念上。 重新考虑 [24.2]，但现在考虑由其依次遍历的（测地）边 $\mathrm{K_j}$ 组成的测地三角形 Δ：

$$\Delta = \mathrm{K}_1 + \mathrm{K}_2 + \mathrm{K}_3$$

如[26.1]所示。（注意：我们重用此示例是因为具有测地边使平行移动易于可视化，但以下推理同样适用于任何（非测地）曲线。）



[26.1] 可以用非奇异基准向量场 $\mathbf{U}$ 覆盖 Δ，然后在边 $\mathrm{K_j}$ 上求和图中所示角度 $\angle \mathrm{Uw_{\parallel}}$ 中的变化 $\mathcal{R}_{\mathrm{U}}(\mathrm{K_j})$ 来找到和乐 $\mathcal{R}(\Delta)$。

图 [26.1] 复制了 [24.2]，但它引入了一个基准向量场 $\mathbf{U}$，使我们能够推广和乐的概念。我们将开放曲线 K 的 $\mathcal{R}(\mathrm{K})$ 定义当我们沿着 K 行进时角度 $\angle \mathrm{Uw_{\parallel}}$（从 $\mathbf{U}$ 到平行移动向量 $\mathbf{w}_{\parallel}$）的净变化：

$$\mathcal{R}(\mathrm{K}) \equiv \delta_{\mathrm{K}}(\angle \mathrm{Uw_{\parallel}}) \tag{26.1}$$

注意：这里下标 U 是必不可少的，因为 $\mathcal{R}_{\mathrm{U}}(\mathrm{K})$ 确实取决于我们任意选择的 $\mathbf{U}$；因此 $\mathcal{R}_{\mathrm{U}}(\mathrm{K})$ 没有真正的数学意义；它只是接下来讨论的垫脚石。也就是说，如果两个不同的向量沿 K 平行移动，它们之间的角度保持固定（由于（22.4）），因此它们都相对于 $\mathbf{U}$ 旋转相同的量。因此，如符号所示，$\mathcal{R}_{\mathrm{U}}(\mathrm{K})$ 与 $\mathbf{w}_{\parallel}$ 的选择无关。

闭三角形环路 Δ 的和乐 $\mathcal{R}(\Delta)$ 可以表示为其边的“和乐”之和：

$$\mathcal{R}(\Delta) = \mathcal{R}_{\mathrm{U}}(\mathrm{K_1}) + \mathcal{R}_{\mathrm{U}}(\mathrm{K_2}) + \mathcal{R}_{\mathrm{U}}(\mathrm{K_3})$$

虽然右侧的三个独立项中的每一项都取决于任意选择的 $\mathbf{U}$，但它们的总和 $\mathcal{R}(\Delta) = \mathcal{K}(\Delta)$ 与 $\mathbf{U}$ 无关。

我们注意到这种方法提供了比我们以前更好的和乐定义，解决了我们在最初讨论中忽略的潜在问题：如果 $\mathcal{R}(\Delta) > 2\pi$ 会怎样？那么如果我们天真地比较 $\mathbf{w}_{\parallel}$ 和 $\mathbf{w}_0$，我们只会（错误地）将旋转 $\mathcal{R}(\Delta)$ 视为超过 2π 的部分。但是使用基准向量场 $\mathbf{U}$ 允许我们连续跟踪 $\mathbf{w}_{\parallel}$ 的旋转，并且上面的公式将得到 $\mathcal{R}(\Delta)$ 的真实值。

26.3 全局高斯-博内定理的霍普夫内蕴证明

现在假设 $\mathcal{S}_{\mathrm{g}}$ 是亏格为 g 的闭合曲面，并假设 $\mathbf{F}$ 是 $\mathcal{S}$ 上的向量场，具有有限数量的奇点 $\mathrm{s_i}$。我们的最终目标是证明

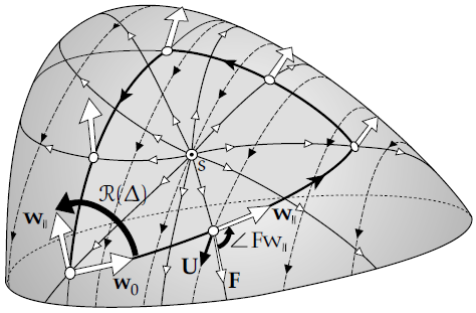
$$\boxed{\mathcal{K}(\mathcal{S}_{\mathrm{g}}) \equiv \iint_{\mathcal{S}_{\mathrm{g}}} \mathcal{K}\mathrm{d}\mathcal{A} = 2\pi \sum_i \mathcal{I}_{\mathrm{F}}(\mathrm{s_i})}$$

(26.2)

但是，在我们这样做之前，让我们解释一下庞家莱-霍普夫定理和 GGB 是如何从这里立即推出的。

首先，由于（26.2）的左侧与 **F** 无关，因此右侧的指数之和对于所有向量场必须具有相同的值。 但是，通过检查我们之前在 $\mathcal{S}_g$ 上的任何向量场示例，例如 [19.11]中画出的蜂蜜流，该普遍的指数和必然等于 $\mathcal{X}(\mathcal{S}_g) = 2 - 2g$，从而证明庞家莱-霍普夫定理。 因此，如果我们能证明（26.2），我们也将得到 GGB 的证明：

$$\mathcal{K}(\mathcal{S}_g) = 2\pi\mathcal{X}(\mathcal{S}_g)$$



[26.2] 差值 $\mathcal{R}(\Delta) - 2\pi\mathcal{I}_F(s)$ 可由对边 K_j 上所示角度 $\angle Fw_\parallel$ （即 $w_\parallel$ 相对于 F 的旋转）的变化 $\Phi(K_j)$ 求和得到。

要理解（26.2），请考虑 [26.2]，它画出了与之前相同的测地三角形 Δ，以及相同的向量 $w_\parallel$ 绕其边界平行移动。 但现在该图还想象了曲面上的向量场 **F**，在 Δ 内有一个奇点 s。（这里我们专门画出了一个源，但该论证适用于任何奇点。）

从这张图我们可以推导出，

$$\begin{aligned}
\mathcal{K}(\Delta) - 2\pi\mathcal{I}_F(s) &= \mathcal{R}(\Delta) - \delta_\Delta(\angle UF) \\
&= \sum_j \left[\mathcal{R}_U(K_j) - \delta_{K_j}(\angle UF)\right] \\
&= \sum_j \left[\delta_{K_j}(\angle Uw_\parallel) - \delta_{K_j}(\angle UF)\right] \\
&= \sum_j \delta_{K_j}(\angle Fw_\parallel)
\end{aligned}$$

最终表达式测量 $\mathbf{w}_\parallel$ 相对于 **F** 的旋转；它显然独立于任意基准向量场 **U**。

为简化起见，让我们采用霍普夫表示法，并将 $\Phi(\mathrm{K_j})$ 定义为 $\mathbf{w}_\parallel$ 沿 $\mathrm{K_j}$ 相对于 $\mathbf{F}$ 的净旋转，即当我们遍历 $\mathrm{K_j}$ 时向量场 $\mathbf{F}$ 和平行移动向量 $\mathbf{w}_\parallel$ 之间夹角的净变化：

$$\boxed{\Phi(\mathrm{K_j}) \equiv \delta_{\mathrm{K_j}}(\angle \mathrm{F w_\parallel})}$$

注意：$\Phi(\mathrm{K_j})$ 与 $\mathbf{w}_\parallel$ 的选择无关，原因与 $\mathcal{R}_\mathrm{U}(\mathrm{K_j})$ 相同。

那么之前的结果可以写成

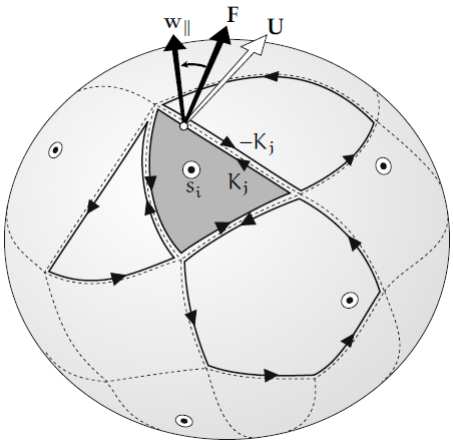
$$\boxed{\mathcal{K}(\Delta) - 2\pi \mathcal{I}_\mathrm{F}(s) = \sum_j \Phi(\mathrm{K_j})} \tag{26.3}$$

显然，如果 Δ 被多边形代替，这个结果同样适用，我们还提醒读者，这个结论并不要求边是测地线。

像往常一样，让 $(-\mathrm{K_j})$ 表示以 $\mathrm{K_j}$相反的方向遍历。 由于平行移动不依赖于曲线被遍历的方向，我们推导出

$$\boxed{\Phi(-\mathrm{K_j}) = -\Phi(\mathrm{K_j})}$$

(26.2) 证明的其余部分现在遵循与我们最初的庞家莱-霍普夫定理证明完全相同的路径。考虑 [26.3]，它只是 [19.10] 的一个重新标记的副本。我们将 $\mathcal{S}_\mathrm{g}$ 划分为具有边 $\mathrm{K_j}$ 的多边形 P_l（同样，不必是测地线），使得每个多边形最多包含 $\mathbf{F}$ 的一个奇点 $\mathrm{s_i}$。



[26.3] 曲面 $\mathcal{S}_\mathrm{g}$ 被划分成多边形 P_l，使得每个多边形中至多有 $\mathbf{F}$ 的一个奇点 $\mathrm{s_i}$。 差 $\mathcal{K}(\mathcal{S}_\mathrm{g}) - 2\pi \sum_\mathrm{i} \mathcal{I}_\mathrm{F}(\mathrm{s_i})$ 可以通过对（所有多边形 P_l 的）所有边 $\mathrm{K_j}$ 的所示角度 $\angle \mathrm{F w_\parallel}$ 的变化 $\Phi(\mathrm{K_j})$ 求和得到。但是 $\mathrm{K_j}$ 的每条边都与相邻多边形的一条相反方向的边相邻，因此所有多边形上的这些旋转的总和为零。因此，$\mathcal{K}(\mathcal{S}_\mathrm{g}) = 2\pi \sum_\mathrm{i} \mathcal{I}_\mathrm{F}(\mathrm{s_i})$。

如果我们现在对所有多边形求和（26.3），我们会发现

$$
\begin{aligned}
&\mathcal{K}(\mathcal{S}_g) - 2\pi \sum_i \mathcal{J}_F(s_i) \\
&= \sum_l \mathcal{K}(P_l) - 2\pi \sum_i \mathcal{J}_F(s_i) \\
&= \sum_{\text{所有多边形}} \Phi(K_j) \\
&= 0
\end{aligned}
$$

因为每条边 K_j 属于两个相邻的多边形，因此在相反的方向上遍历了两次，贡献 $\Phi(K_j) + \Phi(-K_j) = 0$。

这就完成了霍普夫对（26.2），以及随之而来的庞家莱-霍普夫定理和全局高斯-博内定理的内蕴证明。

第二十七章 度量曲率公式的几何证明

27.1 引言

本章的唯一目的是使用平行移动来给出（终于！）“星际迷航移相器”公式（4.10）的几何证明，对于度量的曲率，我们在这里重复：

$$\boxed{\mathcal{K} = -\frac{1}{\mathrm{AB}}\left(\partial_v \frac{\partial_v \mathrm{A}}{\mathrm{B}} + \partial_u \frac{\partial_u \mathrm{B}}{\mathrm{A}}\right)} \tag{27.1}$$

我们现在这个公式首次公布点的 20 多个章节之后，所以我们首先提醒读者 (u,v) 坐标以及度量分量 A 和 B 的含义。

曲面 $\mathcal{S}$ 上一般坐标系的构建（如 [4.3] 所示）总是可以特殊化为正交坐标系 (u,v)（如 [27.1] 所示）。我们首先绘制一族不相交的曲线以覆盖我们正在分析的曲面片，使得面片上的每个点 $\hat{\mathrm{p}}$ 有一条（且只有一条）曲线经过。 我们现在对这些曲线进行任意（但可微分）编号以创建 u-坐标，并将这些曲线及其分配的 u-值称为 u-曲线。

要完成正交坐标系的构建，我们现在画出 u-曲线的**正交轨迹**，并用 v-坐标对它们进行可区分的标记，并称这些为 v-曲线。

因此，如图所示，$\mathcal{S}$ 上的点 $\hat{\mathrm{p}}$ 可以由唯一的 u-曲线（比如 $\mathrm{u} = \mathrm{U}$）和 v-曲线（比如 $\mathrm{v} = \mathrm{V}$）在那里的交点标记。 故在图中 $\hat{\mathrm{p}}$ 现在可以表示为 $\mathrm{p} = (\mathrm{U},\mathrm{V})$。 因此，如图所示，u-曲线在地图中由垂直线表示，而 v-曲线由水平线表示。

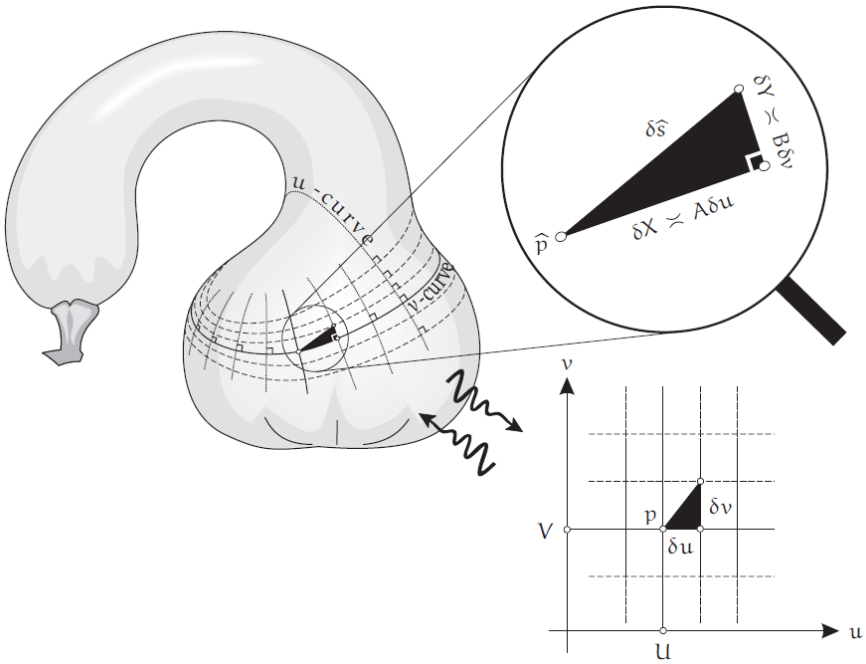
注意：这里我们恢复到我们之前的符号，通过将“帽子”（$\widehat{\ }$）附加到后者来区分地图中的对象与曲面中的相应对象。 因此，例如，地图中面积元素将表示为 $\delta\mathcal{A}$，而曲面上的相应区域将表示为 $\delta\hat{\mathcal{A}}$。

让 X 沿 $\mathcal{S}$ 上的 v-曲线测量距离，该曲线对应于地图中的水平线。 同样，让 Y 沿 u-曲线测量距离，该曲线对应于地图中的垂直线。 我们现在可以使用 X 和 Y 来解释 A 和 B 的含义。 如果我们沿地图中的水平线移动一个小的（最终趋于零）距离 $\delta\mathrm{u}$，那么曲面上的对应点沿对应的v-曲线移动一个最终成比例的距离 $\delta\mathrm{X}$：

$$\delta\mathrm{X} \asymp \mathrm{A}\delta\mathrm{u} \quad \Leftrightarrow \quad \mathrm{A}\text{是局部水平扩展因子}$$

同样，如果我们沿地图中的垂直线移动距离 $\delta\mathrm{v}$，则曲面上的相应点沿相应的 u-曲线移动距离 $\delta\mathrm{Y}$，并且

$$\delta\mathrm{Y} \asymp \mathrm{B}\,\delta\mathrm{v} \quad \Leftrightarrow \quad \mathrm{B}\text{是局部垂直扩展因子}$$



[27.1] **正交坐标中的度量**。 画出了一系列（不相交的）“u-曲线”（u=const.）后，我们将 v-曲线构建为它们的正交轨迹。 如果 $\hat{p}$ 是曲线 $u = U$ 和 $v = V$ 的交点，则它由图中的点 $p = (U, V)$ 表示。 地图中的一个小的水平运动 δu 会在曲面上产生最终成比例的运动 $\delta X \asymp A\delta u$（沿 v-曲线），而地图中的垂直运动 δv 在曲面上产生最终成比例的正交运动 $\delta Y \asymp B\delta v$（沿 u-曲线）。

然后，度量根据地图内的表观距离告诉我们曲面内的真实距离 $d\hat{s}$，而毕达哥拉斯定理意味着在我们的正交 (u, v) 坐标系中，它采用（4.9）的形式：

$$\boxed{d\hat{s}^2 = A^2 du^2 + B^2 dv^2}$$

27.2 围绕环路之向量场的环量

为了避免扰乱接下来的证明流程，我们首先解决一个引理，即如何计算向量场 $\mathbf{V}$ 围绕一个小（最终趋于零）环路的环量（circulation）。

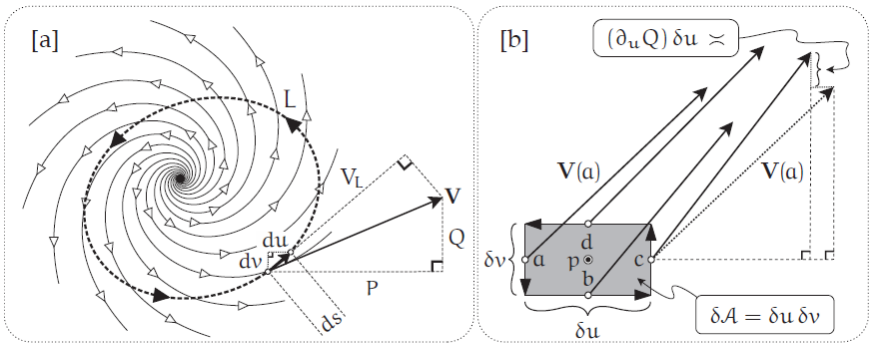
所要求的概念——称为 $\mathbf{V}$ 的旋度（curl）——对于那些学习过本科向量分析的人来说会很熟悉，但提醒一下也无妨； 此外，我们的说明将比通常的更具有几何意义。

令 $\mathbf{V} = \begin{bmatrix} P(u,v) \\ Q(u,v) \end{bmatrix}$ 为 (u, v) 平面中的向量场，令 $r = \begin{bmatrix} u \\ v \end{bmatrix}$ 为沿（逆时针）简单闭合环路 L 运动的点的位置向量。然后，如在 [27.2a] 中所阐明的，我们将 $\mathbf{V}$ 围绕 L 的环量 $\mathcal{C}_L(\mathbf{V})$ 定义为 $\mathbf{V}$ 在 L 方向上的分量的积分：

$$\boxed{\mathcal{C}_L(\mathbf{V}) \equiv \oint_L \mathbf{V} \cdot d\mathbf{r} = \oint_L V_L \, ds = \oint_L [P du + Q dv]}$$

(27.2)

其中 V_L 是 $\mathbf{V}$ 在 L 方向 $\mathbf{dr}$ 上的（有符号）投影，其中 $ds = |\mathbf{dr}|$。对于图示的涡流，很明显 V_L 在 L 上始终为正值，因此环量 $\mathcal{C}_L(\mathbf{V})$ 为正值。



[27.2] [a] **环量**$\mathcal{C}_L(\mathbf{V}) = \oint_L V_L\,ds = \oint_L [P du + Q dv]$。 [b] 随着阴影矩形的缩小，围绕其边界的环量最终等于 $\{\partial_u Q - \partial_v P\}\,\delta\mathcal{A}$。

现在考虑 [27.2b]，它说明了一个小的（最终趋于零）坐标矩形 R，以 p 为中心，边为 δu 和 δv，因此面积为 $\delta\mathcal{A} = \delta u\,\delta v$。 为了计算 $\mathcal{C}_R(\mathbf{V})$，我们求黎曼和，如图所示，我们选择使用边的中点。当然，在 R 趋于零的极限中，这个特定的选择无关紧要，但我们注意到，这种中点的选择使近似值比随机选择准确得多，每边的误差随着边的立方而消失。

中点的选择也使得环量可视化和计算变得特别容易。向上的通过 c 的垂直边对环量的贡献最终等于 $Q(c)\,\delta v$。 并且向下通过 a 的边贡献同样是 $[-Q(a)]\,\delta v$。 因此，来自两个垂直边的净贡献由 $\mathbf{V}$ 的垂直分量的差（图示）控制，即 $\{Q(c) - Q(a)\}\,\delta v$。 但是，如图所示，Q 的这种变化反过来又受它的偏导数支配，我们可以将其计算在 R 的中心 p 处，因为（取极限时） R 朝它收缩。 因此，两条垂直边对环量的净贡献最终等于 $(\partial_u Q\delta u)\,\delta v$。

完全相同的推理适用于两条水平边，因此围绕 R 的总环量由下式给出

$$
\begin{aligned}
\mathcal{C}_R(\mathbf{V}) &= \oint_R [P du + Q dv] \\
&\asymp Q(a)(-\delta v) + P(b)(\delta u) + Q(c)(\delta v) + P(d)(-\delta u) \\
&= \{Q(c) - Q(a)\}\delta v - \{P(d) - P(b)\}(\delta u) \\
&\asymp (\partial_u Q\delta u)\delta v - (\partial_v P\delta v)\delta u \\
&= \{\partial_u Q - \partial_v P\}\,\delta\mathcal{A}
\end{aligned}
$$

正如这个公式所示，我们现在可以将 $\mathbf{V}$ 的**旋度**定义为单位面积的局部环量：

$$
\boxed{\operatorname{curl}\begin{bmatrix} P \\ Q \end{bmatrix} \asymp \frac{\mathcal{C}_R(\mathbf{V})}{\delta\mathcal{A}} \asymp \partial_u Q - \partial_v P}
$$

27.3 演练：平面中的和乐

为了熟悉情况，让我们看看如何确定平面上一个小环的单位面积的和乐 ($\asymp \mathcal{K}$)。当然，答案最好是 $\mathcal{K} = 0$！

在前一章中，我们将**和乐**推广到开放曲线 $\hat{\mathrm{K}}$，在（26.1）中将其定义为任何平行移动向量 $\mathbf{w}_{\parallel}$ 沿 $\hat{\mathrm{K}}$ 相对于任意（但无奇点）基准向量场 $\mathbf{U}$ 的旋转：

$$\mathcal{R}_{\mathbf{U}}(\hat{\mathrm{K}}) \equiv \delta_{\hat{\mathrm{K}}}(\angle \mathrm{U}\mathrm{w}_{\parallel})$$

回想一下，虽然这与 $\mathbf{w}_{\parallel}$ 无关，但它并没有直接告诉我们关于曲面几何形状的任何信息——它只是测量 $\mathbf{w}_{\parallel}$ 相对于任意选择 $\mathbf{U}$ 场的旋转。 但是，如果 $\hat{\mathrm{K}}$ 变成闭环 $\hat{\mathrm{L}}$，则 $\mathcal{R}_{\mathbf{U}}(\hat{\mathrm{L}}) = \mathcal{R}(\hat{\mathrm{L}})$ 与 $\mathbf{U}$ 无关，并且等于 $\hat{\mathrm{L}}$ 内的总曲率。

让我们把我们的曲面 $\mathcal{S}$ 当作平面，让我们应用上面的想法，使用普通的极坐标：$\mathrm{u} = \mathrm{r}$，$\mathrm{v} = \theta$，$\mathrm{A} = 1$，和 $\mathrm{B} = \mathrm{r}$：

$$\mathrm{d}\hat{\mathrm{s}}^2 = \mathrm{A}^2\mathrm{du}^2 + \mathrm{B}^2\mathrm{dv}^2 = \mathrm{dr}^2 + \mathrm{r}^2\mathrm{d}\theta^2$$

在 [27.3] 的左侧，我们在极坐标地图平面中看到一个小的（最终趋于零）矩形环路 $\mathrm{L} = \mathrm{efghe}$，边为 $\delta\mathrm{r}$ 和 $\delta\theta$。 右边是它在（平坦！）曲面 $\mathcal{S}$ 上的像 $\hat{\mathrm{L}} = \hat{\mathrm{ef}}\hat{\mathrm{gh}}\hat{\mathrm{e}}$。让我们选择基准向量场 $\mathbf{U}$ 为径向，沿着射线 $\theta = \mathrm{const.}$ 向外指向，让我们沿着闭环 $\hat{\mathrm{L}}$ 的四个腿中的每一个找到和乐。 从 $\hat{\mathrm{e}}$ 开始，让我们选择初始向量 $\mathbf{w}$ 为 $\hat{\mathrm{e}}$ 处的 $\mathbf{U}$，现在让我们将它围绕环路平行移动来得到 $\mathbf{w}_{\parallel}$。

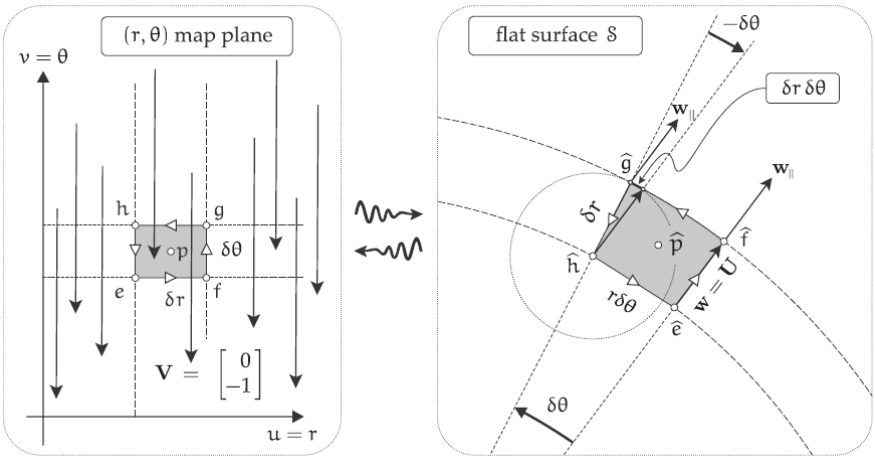
在我们旅程的第一站，$\mathbf{w}_{\parallel}$ 不相对于 $\mathbf{U}$ 旋转，所以 $\delta\mathcal{R}_{\mathrm{U}}(\hat{\mathrm{ef}}) = 0$。 在第二条腿上，正如我们在 [27.3] 中清楚地看到的，$\mathbf{w}_{\parallel}$ 的旋转是 $\delta\mathcal{R}_{\mathrm{U}}(\hat{\mathrm{fg}}) = -\delta\theta$。 在第三个径向腿 $\hat{\mathrm{gh}}$ 上没有旋转。 最后，当我们沿着 $\hat{\mathrm{he}}$ 回家时，$\mathbf{w}_{\parallel}$ 相对于 $\mathbf{U}$ 的旋转是 $\delta\theta$。

因此，$\mathbf{w}_{\parallel}$ 相对于 $\mathbf{U}$ 的旋转相互抵消，因此和乐（独立于 $\mathbf{U}$）——以及曲率——确实消失了，因为它应该：

$$\begin{aligned}\mathcal{R}(\hat{\mathrm{L}}) &= \delta\mathcal{R}_{\mathrm{U}}(\hat{\mathrm{ef}}) + \delta\mathcal{R}_{\mathrm{U}}(\hat{\mathrm{fg}}) + \delta\mathcal{R}_{\mathrm{U}}(\hat{\mathrm{gh}}) + \delta\mathcal{R}_{\mathrm{U}}(\hat{\mathrm{he}}) \\ &= 0 + (-\delta\theta) + 0 + \delta\theta \\ &= 0\end{aligned}$$

但是还有另一种方法可以确定 $\mathbf{w}_{\parallel}$ 相对于 $\mathbf{U}$ 的旋转，正是这种方法，我们将推广以推导出 $\mathcal{K}$ 的通用公式：我们将使用坐标网格内距离的局部测量值 - 换句话说，度量。

由于四边形半径为 $\mathrm{r} + \delta\mathrm{r}$ 的远边比半径为 r 的近边长，所以会出现沿 $\hat{\mathrm{ef}}$ 的明显旋转。长多少？嗯，远侧的长度为 $(\mathrm{r} + \delta\mathrm{r})\delta\theta$，而近侧的长度为 $\mathrm{r}\,\delta\theta$，因此，如图所示，长度的增加是 $\delta\mathrm{r}\delta\theta$。



[27.3] $\delta\mathcal{R}_U(\hat{h}\hat{e}) = \delta\theta$ 这一事实可以从这一边比对边短 $\delta r\delta\theta$ 的事实推导出来。 $\mathbf{V}$ 在地图平面中围绕环的环量在右侧的（平坦）曲面上产生（趋于零的）和乐。

如果我们认为这最终等于以 $\hat{h}$ 为中心通过 $\hat{g}$ 的图示圆（半径为 δr）的小弧线，那么这种长度的增加会在 $\hat{h}$ 处对着一个角度，这就是我们寻求的旋转：

$$-\delta\mathcal{R}_U(\hat{f}\hat{g}) \asymp \frac{\text{弧长}}{\text{半径}} \asymp \frac{\text{边长增加}}{\text{正交边}} = \frac{\delta r\delta\theta}{\delta r} = \delta\theta$$

更一般地，如果我们称边长为 $\delta Y \asymp B\,\delta v$，则 u 增加 δu 所导致的增量 $\delta^2 Y$ 将是

$$\delta^2 Y \asymp [\partial_u B\,\delta u]\,\delta v$$

在本例中，$B = r$，$u = r$，并且 $v = \theta$，因此

$$\delta^2 Y \asymp [\partial_u B\,\delta u]\,\delta v = [\partial_r r\,\delta r]\,\delta\theta = \delta r\,\delta\theta$$

本该如此。

在我们转向曲面的一般情况之前，让我们重新解释我们所做的，但现在在 [27.3] 左侧的 (r, θ) 地图平面中。 在环路 L 的每条腿上，$\mathcal{S}$ 上的旋转由下式给出

$$\delta\mathcal{R}_U(\text{边}) = 0\delta r + (-1)\delta\theta = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} \delta r \\ \delta\theta \end{bmatrix}$$

向量场 $\mathbf{V} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ 在 [27.3] 的地图平面中说明。 我们现在认识到 $\mathcal{S}$ 上环路的和乐等于 $\mathbf{V}$ 在地图平面中围绕环路的环量。 $\mathbf{V}$ 的流垂直于顶部和底部的边，因此它们对环量没有任何贡献。 但是 $\mathbf{V}$ 确实直接沿着边 he 流动，它直接逆着边 fg 流动，这两个贡献完全相等且彼此相反。

总之，

$$\boxed{\mathcal{R}(\hat{\mathrm{L}}) = \mathcal{C}_{\mathrm{L}}(\mathbf{V})}$$

正如我们现在将看到的，在一般情况下，也可以将曲面 $\mathcal{S}$ 上的环的和乐可视化为向量场 $\mathbf{V}$ 围绕 (u,v) 地图平面中的相应环的环量， 但在这种情况下，环量通常不会消失，对应于 $\mathcal{S}$ 上的环内曲率的存在

27.4 和乐作为地图中度量诱导向量场的环量

在本节中，我们将推广上述论证，并根据特殊向量场 $\mathbf{V}$（由度量确定）的围绕 (u,v) 地图上简单环路 L 的环量 $\mathcal{C}_{\mathrm{L}}(\mathbf{V})$ 推导出 $\mathcal{S}$ 上对应环路 $\hat{\mathrm{L}}$ 的和乐公式 $\mathcal{R}(\hat{\mathrm{L}})$。

令 R 表示 [27.3] 的地图平面（中心为 p）中显示的小矩形（边界 $\mathrm{L} = \mathrm{efgh}$），现在仅具有一般的 (u,v) 坐标，因此具有边 $\delta\mathrm{u}$ 和 $\delta\mathrm{v}$。 这映射到曲面 $\mathcal{S}$ 上的曲线四边形 $\hat{\mathrm{R}}$，具有相应的“中心”$\hat{\mathrm{p}}$。

由于 $\hat{\mathrm{R}}$ 的边最终是 $\delta\mathrm{X} \asymp \mathrm{A}\delta\mathrm{u}$ 和 $\delta\mathrm{Y} \asymp \mathrm{B}\,\delta\mathrm{v}$，所以 $\hat{\mathrm{R}}$ 的面积 $\delta\hat{\mathcal{A}}$ 最终与 R 的面积 $\delta\mathcal{A} = \delta\mathrm{u}\,\delta\mathrm{v}$ 相关联，

$$\boxed{\delta\hat{\mathcal{A}} \asymp \delta\mathrm{X}\delta\mathrm{Y} \asymp (\mathrm{A}\delta\mathrm{u})(\mathrm{B}\,\delta\mathrm{v}) = (\mathrm{AB})\delta\mathcal{A}} \tag{27.4}$$

换句话说，当我们从地图传递到曲面时，*局部面积扩展因子是* AB。

当 R 缩小到中心点 p 时，$\hat{\mathrm{R}}$ 缩小到点 $\hat{\mathrm{p}}$，因此肉眼看来它是一个真实的平面矩形。然而，为了研究曲率，我们必须将 $\hat{\mathrm{R}}$ 置于放大倍数为 $(1/\delta\mathrm{u})$ 或 $(1/\delta\mathrm{v})$ 的强大显微镜下，在这种情况下，很明显这个“矩形”的相对边的长度彼此不同，尽管只是非常轻微。

在这一点上，为了更生动地理解，我们强烈建议你在橙子、苹果或葡萄柚上画出一个小“矩形”$\hat{\mathrm{R}}$，也许可以使用普通的球面极坐标，通过绘制经度和纬度的圆圈来创建网格。 然后，你可以沿着 $\hat{\mathrm{R}}$ 的一条边平行移动一根牙签，并观察它如何相对于你选择在水果曲面画出的坐标网格旋转。

图 [27.4] 画出了$\hat{\mathrm{R}}$ 的顶点正交投影到 $\mathcal{S}$ 的密切平面上 $\hat{\mathrm{R}}$ 的任意内点（比如 $\hat{\mathrm{p}}$）处。我们用直线将它们连接起来形成一个四边形。 请注意，虽然曲面上的 (u,v) 坐标曲线始终是正交的，但该投影四边形的边并不正交。 但随着 R 缩小，$\hat{\mathrm{R}}$ 也缩小，它的形式最终是一个矩形。 在这里，我们特意从这个极限情况得出了一个相当极端的偏差，以便使几何推理更容易理解。

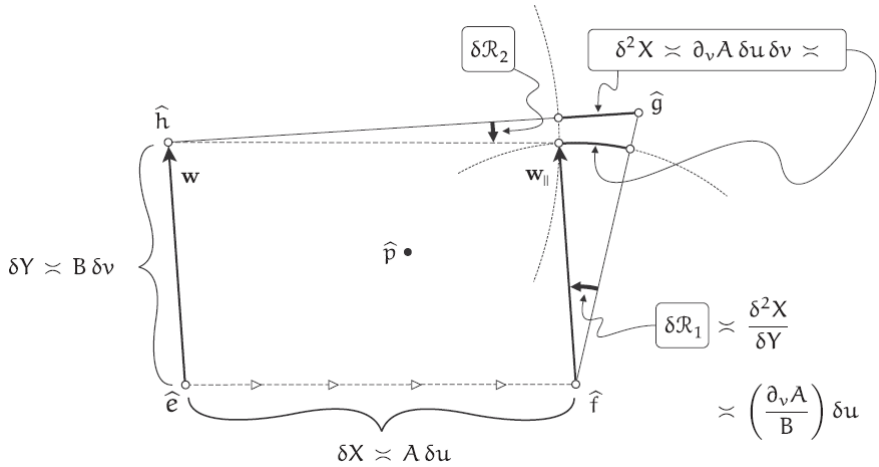
使用与上一小节相同的符号，让 $\delta^2\mathrm{X} = \delta[\delta\mathrm{X}]$ 表示 $\delta\mathrm{X}$ 的增加，如图所示。 这个 δ^2 符号提醒我们，我们现在正在处理一个二阶无穷小，如下所示：

$$\delta^2 X \asymp [\partial_v A \delta v]\, \delta u$$

如 [27.4] 所示，$\delta^2 X$ 最终等于半径为 $\delta Y \asymp B\,\delta v$ 的圆弧，在 $\hat{e}$ 对角 $\delta\mathcal{R}_1$。 因此，由地图中的变化 δu 产生的小和乐最终由下式给出

$$\delta\mathcal{R}_1 \asymp \frac{\text{弧长}}{\text{半径}} \asymp \frac{\delta^2 X}{\delta Y} \asymp \frac{[\partial_v A \delta v]\, \delta u}{B\,\delta v} = \left(\frac{\partial_v A}{B}\right)\delta u$$

其中 $\partial_v A$ 和 B 都被理解为在 R 的中心 p 处计算，因为 R （最终）缩小到 p。



[27.4] **度量曲率公式的几何证明**。 沿 $\hat{e}\hat{f}$ 的平行移动引起相对于 (u,v) 曲线的旋转 $\delta\mathcal{R}_1$——由 $\delta\mathcal{R}_1 \asymp \left(\frac{\partial_v A}{B}\right)\delta u$ 给出。

完全相同的推理适用于由地图中的变化 δv 产生的和乐，只需执行同时交换 $u \leftrightarrow v$ 和 $A \leftrightarrow B$。

然而，如果 $\partial_u B$ 是正的——正如我们在 [27.4] 中所画的那样——那么平行移动向量相对于坐标网格的旋转 $\delta\mathcal{R}_2$ 现在是顺时针旋转，如图所示，因此我们必须在公式中引入一个减号：

$$\delta\mathcal{R}_1 \asymp -\left(\frac{\partial_u B}{A}\right)\delta v$$

内蕴导数的线性意味着由变化 δu 和 δv 产生的净旋转是单独旋转的总和，因此由下式给出

$$\delta\mathcal{R} \asymp \left[\frac{\partial_v A}{B}\right]\delta u - \left[\frac{\partial_u B}{A}\right]\delta v = \begin{bmatrix} (\partial_v A)/B \\ -(\partial_u B)/A \end{bmatrix} \cdot \begin{bmatrix} \delta u \\ \delta v \end{bmatrix} \tag{27.5}$$

为了获得在 $\mathcal{S}$ 上一个简单闭环 $\hat{L}$ 的和乐，我们必须在 (u,v) 地图中沿着它的像 L 积分这个公式。 因此我们得到了这个重要的结果：

令 $\mathcal{S}$ 是一个具有度量 $\mathrm{d}\hat{s}^2 = \mathrm{A}^2\,\mathrm{du}^2 + \mathrm{B}^2\,\mathrm{dv}^2$ 的曲面，令 $\hat{\mathrm{L}}$ 是 $\mathcal{S}$ 上的一个简单环路，由 (u,v) 地图平面中的简单环路 L 表示。 如果在地图平面中我们定义了向量场

$$\mathbf{V} \equiv \begin{bmatrix} (\partial_{\mathrm{v}}\mathrm{A})/\mathrm{B} \\ -(\partial_{\mathrm{u}}\mathrm{B})/\mathrm{A} \end{bmatrix} \tag{27.6}$$

那么 $\hat{\mathrm{L}}$ 在 $\mathcal{S}$ 上的和乐等于 $\mathbf{V}$ 在地图中围绕 L 的环量：

$$\mathcal{R}(\hat{\mathrm{L}}) = \mathcal{C}_{\mathrm{L}}(\mathbf{V})$$

你可以很容易地验证这个一般结果与[27.3]中说明的特殊情况一致[练习]。

27.5 度量曲率公式的几何证明

现在让我们考虑环路 $\hat{\mathrm{L}}$（面积为 $\delta\hat{\mathcal{A}}$）收缩到 $\hat{\mathrm{p}}$ 点的极限。 曲率 $\mathcal{K}(\hat{\mathrm{p}})$ 最终等于每单位面积的和乐，因此通过结合（27.4）和（27.6）我们得到，

$$\mathcal{K}(\hat{\mathrm{p}}) = \lim_{\hat{\mathrm{L}}\to\hat{\mathrm{p}}} \frac{\mathcal{R}(\hat{\mathrm{L}})}{\delta\hat{\mathcal{A}}} \asymp \frac{1}{\mathrm{AB}}\left[\frac{\mathcal{C}_{\mathrm{L}}(\mathbf{V})}{\delta\mathcal{A}}\right]$$

但是（27.3）现在告诉我们，右边括号中的项——即单位面积的局部环量——正是 $\mathbf{V}$ 的旋度。

这使我们能够完成度量曲率公式（27.1）的几何证明，这是 1827 年高斯绝妙定理的最优雅和最明确的表现形式之一：

$$\begin{aligned}
\mathcal{K}(\hat{\mathrm{p}}) &= \frac{1}{\mathrm{AB}}\left\{\begin{bmatrix} \dfrac{(\partial_{\mathrm{v}}\mathrm{A})}{\mathrm{B}} \\ -\dfrac{(\partial_{\mathrm{u}}\mathrm{B})}{\mathrm{A}} \end{bmatrix} \text{的旋度}\right\} \\
&= \frac{1}{\mathrm{AB}}\left\{\partial_{\mathrm{u}}\left[-\frac{\partial_{\mathrm{u}}\mathrm{B}}{\mathrm{A}}\right] - \partial_{\mathrm{v}}\left[\frac{\partial_{\mathrm{v}}\mathrm{A}}{\mathrm{B}}\right]\right\} \\
&= -\frac{1}{\mathrm{AB}}\left\{\partial_{\mathrm{v}}\left[\frac{\partial_{\mathrm{v}}\mathrm{A}}{\mathrm{B}}\right] + \partial_{\mathrm{u}}\left[\frac{\partial_{\mathrm{u}}\mathrm{B}}{\mathrm{A}}\right]\right\}
\end{aligned}$$

当然在证明这个公式的过程中，我们也证明了一个重要的特例，即地图是共形的将使得 $\mathrm{A} = \mathrm{B} = \Lambda$。 在这种情况下，公式简化为更优雅的形式，（4.16）：

$$\boxed{\mathcal{K} = -\frac{\nabla^2 \ln \Lambda}{\Lambda^2}}$$

第二十八章 曲率作为相邻测地线之间的力

28.1 雅可比方程简介

本章将介绍对曲率的全新解释，这在爱因斯坦的弯曲时空引力理论中至关重要。 这个想法是查看相邻单位速度测地线的分离，并检查它们的相对加速度，这可能是相互靠近（吸引力），也可能是相互远离（排斥力）。

我们首先在常曲率的三个主要情况下研究这种现象，然后我们继续分析变曲率的一般曲面。在这种一般情况下，我们将给出控制该相对加速度的基本方程的两种不同的几何证明，称为**测地线偏差方程**或**雅可比方程**，以卡尔·古斯塔夫·雅各比·雅可比（如 [28.1] 中的照片）命名，他在 1837 年发现了它。

[28.1] 卡尔·古斯塔夫·雅各布·雅各比 (1804 – 1851)。

正如我们将看到的，基本的见解是，如果两个相邻的测地线穿过一个正曲率的区域，它们就会相互吸引。 如果这两个测地线从一个共同的原点 o 发射，在稍微不同的方向上，正曲率产生的这种吸引力可能会使它们聚焦回到第二个交点，称为 o 的**共轭点**。

另一方面，穿过负曲率区域的相邻测地线相互排斥，加速分离。

在这两种情况下，吸引力或排斥力与测地线的间距成正比，并且（局部）比例“常数”等于粒子所在位置的曲面曲率！

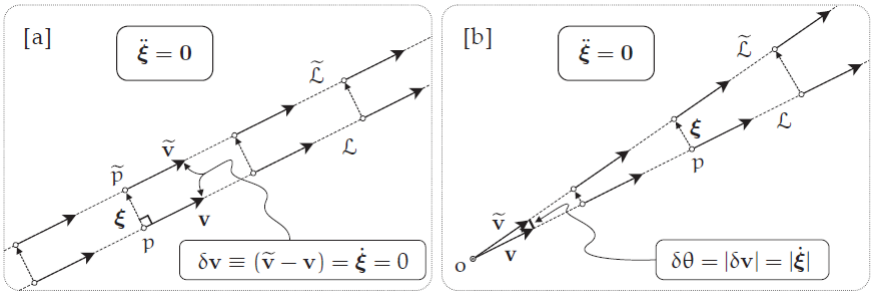
这是雅可比发现的本质。

28.1.1零曲率：平面

考虑 [28.2a]，它画出了位置为 $p(t)$ 的粒子沿平面中的测地直线 L 以单位速度行进，单位速度向量为 $\mathbf{v} = \dot{\mathbf{p}}$。 它还画出了一个相邻的平行测地线 $\tilde{L}$，由 $\tilde{p}$ 追踪，速度为 $\tilde{\mathbf{v}}$。

如图所示，

$$\boxed{\text{垂直连接向量}\ \xi \equiv \overrightarrow{p\tilde{p}} \quad \Longrightarrow \quad \delta v \equiv \tilde{v} - v = \dot{\xi}}$$



[28.2] [a] 在平面中，相邻平行线的间距 $\boldsymbol{\xi}$ 是恒定的，所以 $\dot{\boldsymbol{\xi}} = 0$，所以（平凡地）$\ddot{\boldsymbol{\xi}} = 0$。 [b] 如果直线改为发散，以角度 $\delta\theta$ 分隔，则 $|\dot{\boldsymbol{\xi}}| = \delta\theta$，因此再次 $\ddot{\boldsymbol{\xi}} = 0$。

显然在这种情况下 $\tilde{\mathbf{v}} = \mathbf{v}$，所以 $\delta\mathbf{v} = \dot{\boldsymbol{\xi}} = 0$，但在一般情况下，方程只不过说两个速度的差就是 $\tilde{L}$ 离开 L 的速度。

本章我们关注的重点不是相对速度，而是相对加速度：

$$\boxed{\ddot{\boldsymbol{\xi}} = \frac{d(\delta\mathbf{v})}{dt} = \text{相对加速度}} \tag{28.1}$$

在[28.2a]中，我们发现（微不足道）$\ddot{\boldsymbol{\xi}} = 0$。

接下来，考虑 [28.2b]，其中 L 和 $\tilde{L}$ 现在是从它们共同的原点 o 发出的射线，它们被一个小角度 $\delta\theta$ 分开。这些测地线正在散开，彼此远离，但（远离）以稳定的速率 $\delta\theta$ 进行。该图通过几何构造从 $\mathbf{v}$ 的尖端到 $\tilde{\mathbf{v}}$ 的尖端的连接向量 $\delta\mathbf{v}$ 来解释这一点，它最终等于半径 $|v| = 1$ 的圆的弧，对 o 的角为 $\delta\theta$。

我们也可以象征性地推断出这一点。 如果$r = op$，则p沿L的单位速度可以写成$|v| = .r = 1$。 由于$|\dot{\boldsymbol{\xi}}| \asymp r\,\delta\theta$，分离速度$|\dot{\boldsymbol{\xi}}| \asymp |(r\,\dot{\delta}\theta)| = \delta\theta$。

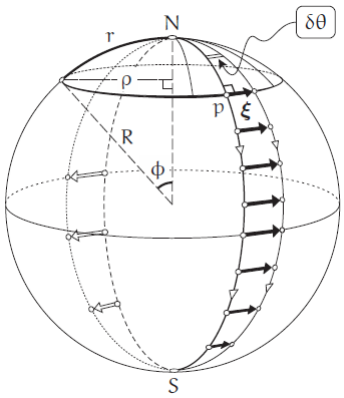
因此，即使这些测地线彼此发散，也没有任何力量推动它们分开——它们的相对加速度为零：

$$\boxed{\ddot{\boldsymbol{\xi}} = 0}$$

正如我们将要看到的，相邻测地线的[illegible]加速度为零是平面曲率为零的新表现。

28.1.2正曲率：球面

考虑[28.3]。我们从半径为 R 的球面的北极 N 发射两个粒子，它们之间的角度为 $\delta\theta$。回想一下，如果粒子粘在曲面上而在曲面内没有受到任何侧向力，它们将自动跟随曲面的测地线运动，在这种情况下，（曲面的测地线）即为大圆。



[28.3] **雅可比方程**。 两个粒子以单位速度从半径为 R 球面北极 N 发射，相隔一个小角度 $\delta\theta$。 在时间 t 之后，它们的间隔为 $\xi = R\delta\theta\sin(t/R)$。 因此，它们的相对加速度由雅可比方程给出：$\ddot{\boldsymbol{\xi}} = -\mathcal{K}\boldsymbol{\xi}$。

在时间 t 之后，位于这些测地线之一上的位置为 p(t) 的粒子将在曲面上移动距离 $r = t$，对球心的角度为 ϕ。 因此 p(t) 位于所示的内蕴半径 $r = t = R\phi$ 的纬度圆上，但具有所示的外在半径，

$$\rho = R\sin\phi = R\sin\left(\frac{t}{R}\right)$$

因此，图示的连接向量 $\boldsymbol{\xi}$ 的长度 ξ 为

$$|\boldsymbol{\xi}| = \xi = \rho\,\delta\theta = R\,\delta\theta\,\sin\left(\frac{t}{R}\right)$$

两次微分产生（第一个）相对速度和（第二个）相对加速度，如下所示：

$$\frac{d\xi}{dt} = \left[\frac{1}{R}\right] R\,\delta\theta\cos\left(\frac{t}{R}\right) \quad \Rightarrow \quad \frac{d^2\xi}{dt^2} = -\frac{1}{R^2} R\,\delta\theta\sin\left(\frac{t}{R}\right)$$

但是由于球体的曲率是 $\mathcal{K} = +(1/R^2)$，这个结果可以写成以下非常紧凑和优雅的形式，称为***测地线偏差方程***，或

$$\boxed{\ddot{\boldsymbol{\xi}} = -\mathcal{K}\boldsymbol{\xi}}$$

雅可比方程： (28.2)

我们仅仅证明了这个方程在 $\mathcal{K}$ 为常数（并且为正）的情况下的特例，但很快我们将能够证明它在变曲率的一般曲面上是正确的。 在一般情况下，我们必须将方程中的 $\mathcal{K}$ 取为 $\mathcal{K}(\mathrm{p})$，即粒子沿测地线穿过曲面时所在位置的曲率。

在我们继续之前，关于这个例子还有很多要说的。 首先，对于一些读者来说，（28.2）可能已经敲响了一个响亮、和谐的钟声！ 因为这个测地线偏差方程也有一个完全不同的名字——**谐振子**方程，它在在经典和量子物理学领域中无处不在。

在家里创建自己的个人谐波振荡器很容易。 拿一根橡皮筋（或弹簧），将一端系在桌子的下面。 现在将一个小物体（我们将其视为具有单位质量）连接到另一端，然后轻轻地放低直到它处于在平衡状态，重力的向下拉力被拉伸的橡皮筋或弹簧的向上拉力平衡。

现在从这个平衡位置直接向下拉动重物，从而拉伸橡皮筋或弹簧。 1676 年[166]，罗伯特•胡克（牛顿同时代的竞争对手）发现，将重物拉回原位的力与距离平衡位置的位移 ξ 成正比——这称为**胡克定律**，比例常数 k 称为橡皮筋或弹簧的**刚度**（stiffness）。

牛顿第二运动定律现在产生与雅可比方程 (28.2) 完全相同的运动方程，但曲面的曲率被弹簧的刚度取代：

$$\boxed{\ddot{\boldsymbol{\xi}} = -\mathrm{k}\boldsymbol{\xi} \quad \text{有解} \quad \boldsymbol{\xi} = \boldsymbol{\xi}_0 \sin\left(\sqrt{\mathrm{k}}\mathrm{t}\right)}$$

其中 $\boldsymbol{\xi}_0$ 表示最大伸展，发生在弹簧停止向下运动并开始将重物拉回的那一刻。

从这个实验和上面的数学解中可以清楚地看出谐振子这个名称的恰当性：重物以正弦曲线上下振荡，每次通过原始平衡位置 $\boldsymbol{\xi} = 0$ 时，都会达到其最大速度（等于其发射速度）。

现在，考虑到这个数学上完美的物理类比，让我们回到 [28.3]。 首先，请注意，在我们发射这两个粒子后，它们立即以与在平面中相同的（无加速度）速率分开，如 [28.2b] 所示。这在几何上是清楚的，但也可以通过注意到 $\sin\epsilon \asymp \epsilon$ 来证实，因此，当 r 趋于零时，

$$\xi = \mathrm{R}\delta\theta$$

尽管粒子在曲面上自由运动，但随着它们间隔的增加，曲率有效地施加了一种恰似弹簧一样的吸引力，而这种神秘的曲率“力”开始减缓它们的分离。

随着间隔的增加，吸引力也成正比增加，因此分离速度降低得更快，直到最终相对速度下降到零，就像粒子在赤道上达到最大间隔一样。 （这对应于重物处于最低点，弹簧处于

[166]事实上，胡克只是宣称他在 1676 年的发现了，并没有透露他发现了什么，而是将其作为一个难以理解的拉丁字谜：“ceiiinossttuv”发表！ 直到 1678 年，他才揭示了他的字谜的答案：ut tensio, sic vis（“作为延伸，所以力量”）。

最大伸展位置。）吸引力现在处于最大强度，它开始再次将测地线拉回到一起，直到它们最终会聚到第二个交点南极 S 处——这被称为 N 的共轭点。

当粒子在 S（其中 $\boldsymbol{\xi} = 0$）处会聚并相交时，它们会达到最大相对速度，就像我们的摆锤通过平衡点 $\boldsymbol{\xi} = 0$ 时达到最大速度一样。

从 N 到 S 的旅程只是完整 $\boldsymbol{\xi}$ 振荡的一半，其余一半在球面的背面完成，现在 $\boldsymbol{\xi}$ 的方向反转，如图所示，从指向 p 轨迹左侧切换为指向右侧。

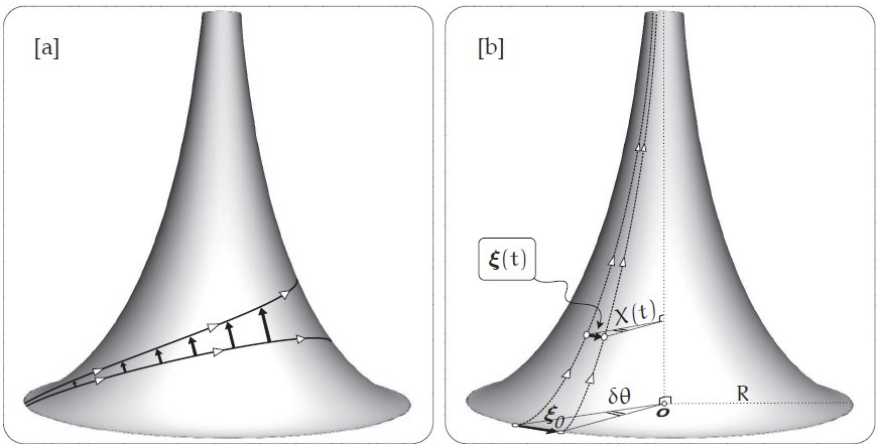
一旦粒子回到 N，ξ 振荡重新开始，并永远重复。

28.1.3负曲率：伪球面

正如测地线在正弯曲的球面上相互吸引，它们在负弯曲的伪球面上相互排斥。 为了让这尤其明显，[28.4a] 显示了从边缘沿“相同”方向（在 $\dot{\boldsymbol{\xi}}(0) = 0$ 的意义上）发射的两个相邻测地线。 然而，很快，相斥的负曲率就会显现出来，我们看到测地线飞散。

但是为了轻松计算相对加速度，我们将改为考虑 [28.4b]，它显示边缘上的两个相邻粒子以单位速度直接向上发射到伪球面，因此它们在沿着各自的曳物线母线行进时彼此接近，弧长等于时间：$\sigma = t$。

粒子的初始速度是水平的，都指向距离边缘 R 的中心 o。 由于它们以单位速度行进，因此在没有力的情况下，它们应该在时间 $t = R$ 相撞。 但显然它们不会相撞，. . . 永远不会！ 是的，两个曳物线母线彼此接近，但是由于负曲率产生的排斥力，它们的接近速度会随着它们沿伪球面向上运动而降低。



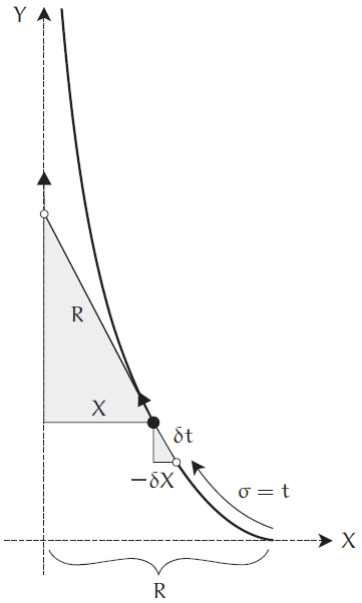
[28.4] [a] 两个相邻的测地线以相同的方向出发，但负曲率迫使它们分开。 [b] 相邻的曳物线母线之间的连接向量 $\boldsymbol{\xi}$ 服从 $\dot{\boldsymbol{\xi}} = -\left(\frac{1}{R}\right)\boldsymbol{\xi}$，因此 $\ddot{\boldsymbol{\xi}} = +\left(\frac{1}{R^2}\right)\boldsymbol{\xi} = -\mathcal{K}\boldsymbol{\xi}$。

如果 $\delta\theta$ 是这些母线之一必须绕对称轴旋转以达到另一母线的图示角度，那么它们的初始小间距将为 $|\boldsymbol{\xi}_0| \asymp R\,\delta\theta$。 如果 $X(t)$ 表示粒子在时间 t 到对称轴的距离，则

$$|\xi(t)| \asymp X(t)\,\delta\theta \asymp \frac{|\xi_0|}{R}X(t) \tag{28.3}$$

接下来，我们提醒读者曳物线的定义几何特性，如 [28.5] 所示（仅复制了我们原先的图 [5.2]）。 这个图意味着

$$\frac{-dX}{dt} = \frac{X}{R}$$



[28.5] **曳物线**的定义是沿其切线到固定线的距离具有恒定长度 R。

这个方程直接导致雅可比方程，甚至不需要先求解！

为了了解如何推导，我们根据粒子的相对速度重新表述方程： (28.3) 暗示

$$\dot{\xi} = -\left(\frac{1}{R}\right)\xi$$

因此初始接近速度是 $|\xi_0|/R$，因此（正如我们之前观察到的）在没有力的情况下初始间隙 $|\xi_0|$ 将在时间 $t = R$ 收缩为零。

粒子实际上没有碰撞的事实是由于负曲率产生的排斥力。这通过第二次微分得到证实，再次得到雅可比测地线偏差方程：

$$\boxed{\ddot{\xi} = -\left(\frac{1}{R}\right)\dot{\xi} = +\left(\frac{1}{R^2}\right)\xi = -\mathcal{K}\xi}$$

最后一个等式源于伪球具有恒定的负曲率，$\mathcal{K} = -(1/R^2)$。

上述论证可以很容易地推广到证明雅可比方程适用于所有旋转曲面，但我们将这一点留给您在练习 8 中证明。相反，我们现在继续建立一般情况下的结果，对于没有任何对称性的曲面。

28.2 雅克比方程的两个证明

28.2.1测地极坐标

图 [28.3] 可以被视为提供了对 $\mathbb{S}^2$ 上普通球面极坐标的新解释。

我们同时从北极向所有可能的方向发射粒子，因此 $\mathbb{S}^2$ 上的每一点（除了极点）都由哪个粒子撞击它，以及撞击时间唯一确定。 更详细地说，哪个粒子击中该点可以通过发射方向与某个特定[167]（但任意）方向（定义为 $\theta = 0$）的经度角 θ 来指定。粒子撞击该点的时间，由纬度角 $\phi = \sigma = t$ 指定（现在假设 $R = 1$）。

测地线极坐标是这个想法对任意曲面 $\mathcal{S}$ 的自然推广。参见 [28.6]，它说明了该想法在 $\mathcal{S}$ 是环面的情况。 从 $\mathcal{S}$ 上的任意点[168] o 向各个方向以单位速度发射粒子，使它们沿测地线向外运动，并选择其中一个为 $\theta = 0$。如果 o 周围的区域没有大到包含任何共轭点，那么其中的每个点都会在唯一的时间被一个唯一的粒子击中。 如果该粒子沿 θ 方向发射并在时间 t 撞击该点，那么我们可以为该点分配唯一的**测地极坐标** (t, θ)。

无论在平面上还是在球面上，我们都知道内蕴半径为 $\sigma = t = \text{const}$ 的内蕴圆周 $K(\sigma)$ 与其“半径”（即从其中心 o 发出的长度为 σ 的测地线段）正交。 事实证明，这在一般曲面上仍然成立，如 [28.6] 中所示。 高斯（1827, §15）在证明其他结果的过程中证明了这一点，因此它被称为

> **高斯引理。** 如果粒子从一般曲面上的一个点向各个方向发射，沿着测地线传播距离 σ，它们将形成（根据定义）具有内蕴半径 σ 的测地线圆周 $K(\sigma)$。 该测地线圆 $K(\sigma)$ 以直角切割其测地线半径。 (28.4)

向量分析（Vector Calculus）直到 1880 年代才由吉布斯（Gibbs）和海维赛德[169]（Heaviside）发展出来，因此在 1827 年高斯被迫通过整页计算（比现行标准计算更短、更透明的版本——基于测地线的加速度垂直于表面的外在事实——可以在例 9 中找到。）来证明这一点。但

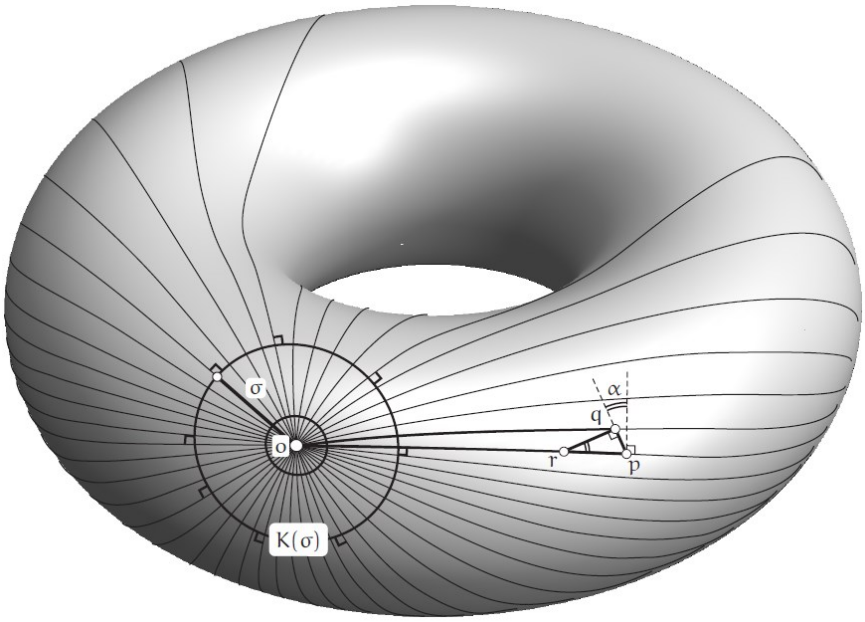
---

[167]在地球上，我们将 $\theta = 0$ 称为本初子午线，并将其定义为通过英国伦敦格林威治皇家天文台的那条子午线。 自 1999 年以来，这个方向就被向北发射划过伦敦夜空的光粒子（来自强大的绿色激光器）所标记！

[168] 现在我们不再考虑 $\mathcal{S}$ 上的点，因为我们不再需要将它们与地图中的点区分开来。

[169] 参见 Crowe 的年表 (1985, pp. 256 – 259)

是在完成他的计算证明后，高斯立即做了一些完全不同寻常的事情：他给出了第二个证明，反而是直观的和几何的。 此外，这第二个证明是内蕴的。



[28.6] **高斯引理**：内蕴半径为 σ 的测地圆 $K(\sigma)$ 以直角切割其测地半径。 高斯对此结果的第二个几何证明假设两个相邻点 p 和 q 位于 $K(R)$ 上，但 $\alpha \neq 0$。 正如文中所解释的，这会导致矛盾。

我们只能推测高斯这样做是因为他认为第二个证明特别重要，或者因为他对自己的几何独创性特别自豪。 无论如何，这就是证据，用高斯自己的话来说："我们认为从最短线的基本性质推导出这个定理是值得的； 只有通过以下推理，无需任何计算就可以证明定理的真实性。"

我们现在将展示这个推理，它似乎已经完全消失在时间的迷雾中——事实上，我们还没有在任何现代课本中发现它的转世实例。 这可能是因为高斯用无穷小来表达他的推理，从而误导了现代数学家。 取代高斯的无穷小，我们将转而采用牛顿最终趋于零的小量之间的终极等式。 此外，与不请自来的高斯（1827 年，第 15 节）相比，我们给出了一张图来看见高斯的真意！

假设我们从 o 发射两个相邻的测地线，它们之间的小（最终趋于零）角度是 $\delta\theta$。 这些是 [28.6] 中显示的两个较暗的测地线段，终止于 p 和 q。 假设 p 和 q 都位于 $K(R)$ 上——即测地线段 op 和 oq 具有相同的长度 $\sigma = R$。 但是现在想象（如图所示）pq 最终并不与 op 和 oq 正交（当 $\delta\theta \to 0$），这与高斯引理相反。 跟随高斯，我们现在证明这是不可能的，从而证明（28.4）。

假设 $\angle opq \asymp (\pi/2) - \alpha$，如图所示。 现在经过 q 正交于 pq 画测地线，设其与 op 的交点为 r，如图所示。 则 $\angle prq \asymp \alpha$，因此 $rq \asymp rp \cos\alpha$。 （在这里和下文，"rq"有双重含义，既代表线段，又代表线段的长度，上下文清楚地说明了这是什么意思。）

所以，

$$
\begin{aligned}
\mathrm{o} \rightsquigarrow \mathrm{r} \rightsquigarrow \mathrm{q} \text{ 的长度} &= \mathrm{or} + \mathrm{rq} \\
&\asymp (\mathrm{R} - \mathrm{rp}) + \mathrm{rp}\cos\alpha \\
&= \mathrm{R} - \mathrm{rp}(1 - \cos\alpha) \\
&< \mathrm{R}
\end{aligned}
$$

因此，这条从 o 到 q（通过 r）的间接路线比最短路线、直接测地线路线 oq 短。 高斯导出了矛盾！

确定了 (t, θ) 坐标的正交性后，度量的形式为

$$
d\hat{s}^2 = dt^2 + \rho^2(t, \theta)\, d\theta^2
$$

因此，如果 $\delta\theta$ 是从 o 发射的两个相邻测地线的角距，则

$$
|\boldsymbol{\xi}| \asymp \rho\, \delta\theta \quad \Longrightarrow \quad |\ddot{\boldsymbol{\xi}}| \asymp \ddot{\rho}\, \delta\theta \tag{28.5}
$$

但是如果我们将 $u = t,\ v = \theta,\ A = 1, \text{and } B = \rho$ 代入（27.1），我们得到

$$
\begin{aligned}
\mathcal{K} &= -\frac{1}{AB}\left(\partial_v \frac{\partial_v A}{B} + \partial_u \frac{\partial_u B}{A}\right) \\
&= -\frac{1}{\rho}\left(\partial_\theta \frac{\partial_\theta 1}{\rho} + \partial_t \frac{\partial_t \rho}{1}\right) \\
&= -\frac{\ddot{\rho}}{\rho}
\end{aligned} \tag{28.6}
$$

将此结果与（28.5）相结合，我们证明了雅可比方程（28.2）：

$$
\ddot{\boldsymbol{\xi}} = -\mathcal{K}\boldsymbol{\xi}
$$

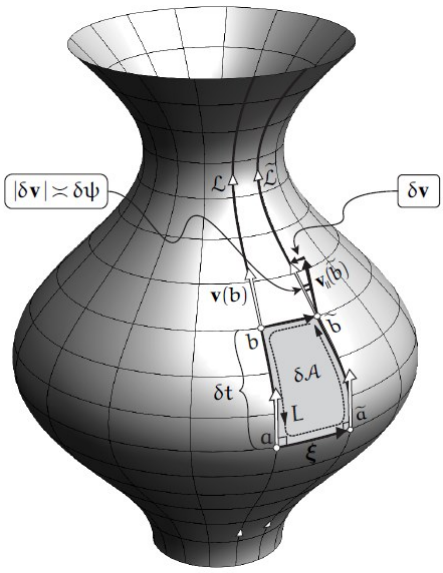
最后，我们注意到（28.6）还允许我们最终证明**明丁定理**，它指出如果两个曲面具有相同的恒定曲率，那么它们是局部等距的。 （详见例 7。）

28.2.2相对加速度=速度的和乐

虽然我们对雅可比方程的第一个证明非常简短（一旦高斯引理成立），它确实需要（“星际迷航移相器”）公式（27.1）的全部力量。 正如我们现在将看到的，实际上没有必要“设置相位器杀戮”！ 取而代之的是，我们现在提出第二个直接、几何和直观的证明。

事实上，本小节的标题就是那个证明，尽管是格言形式。 我们现在使用 [28.7] 详细说明细节，它使用与 [28.2] 中相同的符号。 该图重复使用了我们对克莱罗（Clairaut）定理

[11.7] 证明中的曲面，但现在以新的眼光看待。 具体来说，我们看到子午线测地线母线 $\mathcal{L}$ 和 $\tilde{\mathcal{L}}$ （经过 a 和相邻点 ã）最初在正曲率区域相互吸引，但当它们进入花瓶颈部的负曲率区域时相互排斥（注意：以下论证是完全通用的，不依赖或利用此特定曲面的对称性。）



[28.7] **雅可比方程的几何证明**。 围绕 L 的平行移动 $\mathbf{v}_{\parallel}(\tilde{b})$ 返回到 $\tilde{b}$ 旋转了和乐 $\delta\psi = \mathcal{R}(L) \asymp \mathcal{K}\delta\mathcal{A} \asymp \mathcal{K}|\boldsymbol{\xi}|\,\delta t$。 而 $\delta\psi \asymp |\delta\mathbf{v}| \asymp |\ddot{\boldsymbol{\xi}}|\,\delta t$，得到雅可比方程。

两个相邻的测地线 $\mathcal{L}$ 和 $\tilde{\mathcal{L}}$，以单位速度发射，垂直于短测地线段 aã，令 $\boldsymbol{\xi}$ 连接它们，如图所示。 这确保了 $\mathcal{L}$ 和 $\tilde{\mathcal{L}}$ 开始时是平行的，从某种意义上说，它们最初保持相同的距离，分离速率为零：$\dot{\boldsymbol{\xi}} = 0$。 （注意：为了让后面的几何推理更容易理解，我们已经把间隔画得相当大，但最终我们会让间隙 aã 变为零，所以 $\boldsymbol{\xi}$ 最终将是一个真正的切向量，可以被画在曲面内。）

在时间 δt 之后，两个粒子将在曲面上移动距离 δt，分别到达 b 和 $\tilde{b}$。 到那时，根据定义（28.1），$\mathcal{L}$ 和 $\tilde{\mathcal{L}}$ 之间的相对速度 $\dot{\boldsymbol{\xi}}$ 从零增加到 $\delta\mathbf{v} \asymp \ddot{\boldsymbol{\xi}}\,\delta t$。

$\delta\mathbf{v}$ 在这里想必应该被解释为新速度向量之差 $[\tilde{\mathbf{v}}(\tilde{b}) - \mathbf{v}(b)]$。 然而，这没有内蕴意义，从不同点发出并位于不同切平面的 $\mathcal{S}$ 的切向量不能相减。 为了比较它们，让我们用短测地线[170]段 $b\tilde{b}$ 将 b 连接到 $\tilde{b}$，然后沿 $b\tilde{b}$ 将 $\mathbf{v}(b)$ 平行移动 $\tilde{b}$ 以得到 $\mathbf{v}_{\parallel}(\tilde{b})$，如图所示。 由此，我们获得了相对加速度的内蕴测量：

$$\boxed{\ddot{\boldsymbol{\xi}}\,\delta t \asymp \delta\mathbf{v} = \tilde{\mathbf{v}}(\tilde{b}) - \mathbf{v}_{\parallel}(\tilde{b})}$$

如图所示，令 $\delta\psi$ 为这两个向量在 $\tilde{b}$ 处的夹角。 由于以 $\tilde{b}$ 为中心通过这些向量的尖端的单位圆的弧（未显示）最终等于连接这些尖端的弦，我们看到

[170] 这种选择使论证更容易理解，但我们最终将采用分段收缩为零的极限，因此分段的精确性质无关紧要。

$$|\delta\mathbf{v}| \asymp \delta\psi$$

现在我们来到了问题的关键。 如图所示，让我们将 L 定义为围绕阴影四边形边界的逆时针环路：$L \equiv \tilde{b}ba\tilde{a}\tilde{b}$。 从 $\tilde{b}$ 开始，让我们绕 L 平行移动 $\mathbf{v}_{\parallel}(\tilde{b})$。

首先，由于 $\mathbf{v}_{\parallel}(\tilde{b})$ 最初是通过沿 $b\tilde{b}$ 平行移动 $\mathbf{v}(b)$ 获得的，因此沿 $\tilde{b}b$ 再次将其平行移动回来将使其恢复到原始状态 $\mathbf{v}(b)$。 接下来，由于测地线的速度向量是沿其自身平行移动的，沿 ba 的平行移动在 a 处得到 $\mathbf{v}(a)$。 接下来，沿测地线 $a\tilde{a}$ 的平行移动与其保持正交，因此它到达 $\tilde{a}$ 时成为 $\tilde{\mathbf{v}}(\tilde{a})$。 最后，沿测地线 $\tilde{a}\tilde{b}$ 的平行移动将其返回到起点成为 $\tilde{\mathbf{v}}(\tilde{b})$，因此围绕 L 的平行移动已将 $\mathbf{v}$ 旋转了 $\delta\psi$。

换句话说，

> 测地线的相对加速度是通过速度向量的旋转 $\delta\psi$ 来测量的，但这正是速度向量绕 L 平行移动时的和乐。简而言之，$\delta\psi = \mathcal{R}(L)$！

由于 L 内的面积 $\delta\mathcal{A}$ 由 $\delta\mathcal{A} \asymp |\boldsymbol{\xi}|\delta t$ 给出，雅可比方程的证明归结为以下的最终等式序列，每个等式现在应该在视觉上很明显：

$$|\ddot{\boldsymbol{\xi}}|\,\delta t \asymp |\delta\mathbf{v}| \asymp \delta\psi = \mathcal{R}(L) \asymp \mathcal{K}\delta\mathcal{A} \asymp \mathcal{K}|\boldsymbol{\xi}|\delta t$$

消去 δt，并考虑到对于正 $\mathcal{K}$ 的图示情况下 $\delta\mathbf{v}$ 指向与 $\boldsymbol{\xi}$ 的方向相反这一事实，我们再次得出雅可比方程：

$$\ddot{\boldsymbol{\xi}} = -\mathcal{K}\,\boldsymbol{\xi}$$

28.3 小测地圆的周长和面积

在 2.2 节中，我们表明，在球面上，曲面上的居民可以通过仔细检查内蕴半径 r 的小圆 K(r) 的周长 C(r) 或面积 $\mathcal{A}(r)$ 来确定他们的世界的曲率。。 在这两种情况下，找到曲率的关键是检测这些量如何偏离欧氏值 $C(r) = 2\pi r$ 和 $\mathcal{A}(r) = \pi r^2$ 的。

具体来说，我们证明了在 r 趋于零的极限下，曲率由（2.4）和（2.5）给出，为了方便读者，我们在此重复：

$$\boxed{\mathcal{K} \asymp \frac{3}{\pi}\left[\frac{2\pi r - C(r)}{r^3}\right]} \tag{28.7}$$

和

$$\boxed{\mathcal{K} \asymp \frac{12}{\pi}\left[\frac{\pi r^2 - \mathcal{A}(r)}{r^4}\right]}$$

(28.8)

我们当时声称这些公式适用于所有曲面，而不仅仅是球面，现在我们终于可以证明这一点。

使用与之前相同的测地线极坐标，令 $\xi(r)$ 表示以角度 θ_0 和 $\theta_0+\delta\theta$ 发射的测地线的间隔，因此如果我们定义

$$g(r) \equiv \rho(\theta_0, r) \quad \text{那么} \quad \xi(r) \asymp g(r)\delta\theta$$

当 $r \to 0$ 时，我们知道这可以简化为欧氏公式，$\xi(r) \asymp r\,\delta\theta$，而这正是我们要测量非欧性（即曲率）必须检测到的与欧氏结果的微小偏离。

为此，我们将 $g(r)$ 展开为麦克劳林级数：

$$g(r) = r + \frac{1}{2}g''(0)r^2 + \frac{1}{6}g'''(0)r^3 + \cdots$$

但是（28.6）告诉我们 $g''(r) = -\mathcal{K}g(r)$，所以 $g''(0) = 0$，并且

$$g'''(0) = [-\mathcal{K}g]'(0) = -\mathcal{K}'(0)g(0) - \mathcal{K}(0)g'^{(0)} = -\mathcal{K}(0)$$

因此，

$$\boxed{C(r) = 2\pi g(r) = 2\pi r - \frac{\pi}{3}\mathcal{K}(0)r^3 + \cdots}$$

立即可得 （28.7）。

最后，宽度为 δr 的测地线环的面积 $\delta\mathcal{A}$ 由下式给出

$$\delta\mathcal{A} \asymp C(r)\delta r$$

因此，对 $C(r)$ 的公式进行积分很容易得出面积公式（28.8）[练习]。

第二十九章 黎曼曲率

29.1 导言和总结

在本章中，我们将看到前面关于二维曲面的见解如何自然地扩展到 n 维空间，称为**流形**（manifolds）。

正如我们的二维曲面是由它们的切平面局部描述的，每个切平面都有 $\mathbb{R}^2$ 的结构，所以 n 维流形中一个点的紧邻局部类似于 $\mathbb{R}^n$，但是相邻点之间的距离是用不是欧几里得的度量[171]来测量的。

- **首先**，我们将扩展到 **n-流形**（n 维流形的简写），这是我们对 2-曲面内蕴曲率 $\mathcal{K}(\mathrm{p})$ 的第一个度量，即当小测地三角形 Δ 收缩到点 p 时**每单位面积的局部角盈**：

$$\mathcal{K}(\mathrm{p}) = \lim_{\Delta\to\mathrm{p}} \frac{\mathcal{E}(\Delta)}{\mathcal{A}(\Delta)}$$

- **其次**，事实证明，角盈太钝了，无法直接洞察 n 流形的更微妙的曲率！ 然而，围绕一个收缩小环路的平行移动向量所产生的和乐能够以非常直接的方式充分揭示这种更复杂的曲率结构。 显然，定义 n-流形曲率的下一步必须是理解这样一个空间中的**平行移动**。

将平行移动从 2-曲面推广到 n-流形是非常重要的，但在我们研究过的几乎所有标准课本中，它都很少受到关注。 为了纠正这种不平衡，我们将给出三种不同的几何结构（都导致相同的结果），可用于将列维-奇维塔（Levi-Civita）的平行移动推广到这种 n-流形。

- **第三**，我们使用平行移动来定义 n-流形内的**内蕴**（又名“**协变**”）**导数**。 好消息是——当用内蕴术语表达时——从 2-曲面到 n-流形的过渡完全不需要改变我们原来的定义（23.2）； 只是符号发生了变化，用 $\nabla_{\mathbf{v}}$ 代替了 $\mathrm{D}_{\mathbf{v}}$。

- **第四**——这是本章的核心——我们使用平行移动将和乐从二维曲面推广到 n-流形。

黎曼，如图 [29.1]，发现代替高斯的单一数 $\mathcal{K}$ 来表征 2-曲面，他对 n-流形的广义内蕴曲率由一组[172]

---

[171] 流形的完全一般概念实际上并不要求在其中定义距离。 然而，就我们的目的而言，度量是要研究的中心结构，因此我们必须坚持我们的流形具有度量！ 具有正定度量的流形的技术名称是黎曼流形，而具有可以产生负（“距离”）^{2} 度量的流形，如时空的情况，称为伪黎曼流形。 与其不断明确地表明我们的流形要么是黎曼的要么是伪黎曼的，我们将仅仅说成我们的流形具有度量：它们总是要么是黎曼的要么是伪黎曼的。

[172] 这在练习 11 中得到证明。

$$\boxed{\frac{1}{12}n^2(n^2-1)} \tag{29.1}$$

不同的曲率分量指定：这些构成了一个称为**黎曼张量**（Riemann tensor）的几何对象的数值描述。

[29.1] 伯恩哈德・黎曼 (1826-1866)。

因此，在二维曲面中，黎曼张量减少（$n=2$）为单个分量，这就是 $\mathcal{K}$。你现在坐的房间看起来是 3 维的，因此你周围的空间（$n=3$) 由六个曲率分量描述。 然而，当你一动不动地坐在椅子上时，你实际上是在飞向未来，沿着第四维度，时间！ 在爱因斯坦的 4 维弯曲时空中，有 20 个曲率分量，我们将在下一章中看到，它们描述了引力场！

- **第五**，我们将看到如何将**雅可比方程**推广到 n-流形。

- **第六**，我们将描述一个特别重要的、具有几何意义的黎曼曲率平均值，称为里奇（Ricci）曲率张量（Ricci curvature tensor）。 在爱因斯坦弯曲的 4 维时空中，这恰好捕获了完整曲率信息的一半（总共 20 个分量中的 10 个）。

29.2 n-流形中的角盈

我们对 2-曲面的研究为理解 n-流形提供了必不可少的跳板。 但是 2-曲面实在是太特殊了，无法表现出可以而且确实发生在 n-流形中的各种现象。

然而，令人高兴的是，事实证明，如果我们仅仅上升一个维度，那么 n-流形的所有[173]

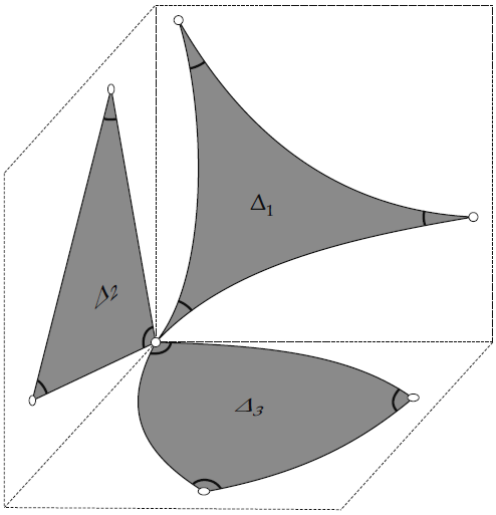
[173] 有一些现象只有在我们达到维度 n=4 时才会出现。 例如，练习 15 中描述的极其重要的 Weyl 曲率张量仅在 n>4 时存在。

基本新特征和概念都会被揭示出来，变成一个易于可视化的 3-流形，比如你现在坐的空间。因此，我们将首先将注意力集中在这个非常具体的案例上，但在下一章中，我们将再次将维度从三增加到四，以了解爱因斯坦的弯曲时空。

通过查看嵌入在 $\mathbb{R}^3$ 中的 2-曲面，我们获得了许多关于 2-曲面的重要见解。 虽然当然可以将我们的 3-流形设想为同样嵌入到更高维空间中，但这样做不再有任何视觉优势。 此外，由于我们着眼于理解爱因斯坦弯曲时空的奖赏，我们将只关注 n-流形的内蕴属性，由其内蕴度量决定。 因此，此后，*我们将始终假设我们是生活在 n−流形中的生物，只有在其中的内蕴测量才有意义。*

我们对内蕴曲率的第一个测量方法是通过测地三角形 Δ 的角盈 $\mathcal{E}(\Delta)$。 但是要构建测地三角形 Δ，我们首先需要的是测地线本身！ 在 2-曲面上，我们通过在曲面上[174]拉紧细绳来构建测地线。 在 3-流形中，我们只需在两点之间拉紧细绳，从而在它们之间获得一条看似直线。 或者，我们可以从一个点向另一个点发射激光束。 但我们不再假设我们的 3-流形是欧氏 $\mathbb{R}^3$； 相反，*它配备了一个通用的非欧度量*，因此不能再依赖拉紧的细绳和激光束来达到预期的效果。 事实上，正如我们将在下一章讨论的那样，“直的”光线的弯曲是爱因斯坦关于物理空间的几何形状是弯曲的预测的第一个实验证实（1919 年）。

在 2-曲面中，我们的 Δ 平面只有一种可能的选择。 但是在我们的 3-流形中，显然有无数个平面可供选择，每个平面对应于法向量的选择。 这种空间切片内的曲率称为**截面曲率**，是第 29.5.8 节的主题。 图 [29.2] 说明了在 3-流形中的三个正交坐标平面中的三个这样的截面曲率：它们在大小和符号方面都是独立的。 在这里，我们展示了三个测地三角形 Δ_i，它们具有一个公共顶点，但每个都位于不同的坐标平面中。 如图所示，这里 $\mathcal{K}(\Delta_1) < 0$，$\mathcal{K}(\Delta_2) = 0$ 和 $\mathcal{K}(\Delta_3) > 0$。



[29.2] 在 3-流形的三个垂直坐标平面中构建的测地三角形仅揭示了空间曲率结构的一部分。 这里，$\mathcal{E}(\Delta_1) < 0$，$\mathcal{E}(\Delta_2) = 0$ 和 $\mathcal{E}(\Delta_3) > 0$。

但是我们已经说过，一个 3-流形的黎曼张量有六个曲率分量，所以这三个坐标平面上的角盈只能检测到一半的曲率结构。 随着维度的增加，情况只会变得更糟：在 4-流形中，六

[174] 为了应对负曲率，回想一下我们想象的绳子被困在两个平行的曲面层之间。

个不同坐标平面中的角盈仅产生六个曲率，而黎曼张量有 20 个曲率分量。

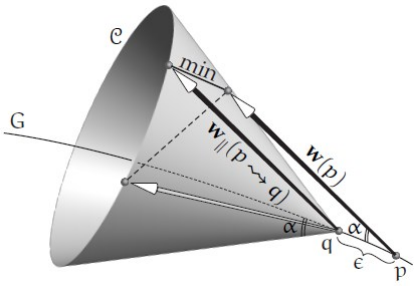
为了探索曲率，我们必须转向和乐。 为此，我们现在必须解决把平行移动从 2-曲面推广到 n-流形这一重要任务。 由于其根本重要性，我们现在给出不只一种，而是三种不同的几何构造！

29.3 平行移动：三种构造

29.3.1常角度圆锥上的最近向量

回想一下，为了在 2-曲面中沿曲线平行移动向量 $\mathbf{w}(p)$，基本的内蕴构造将此曲线视为测地线 G。（为了沿一般曲线平行移动，我们将其分解为测地线段）为了实现沿 G 的平行移动，只需要保持 $\mathbf{w}_{\parallel}$ 和 G 之间的角度 α 不变（并保持 $\mathbf{w}_{\parallel}$ 的长度不变）。

在 3-流形中，这种构造立即遇到麻烦。 见[29.3]。 假设我们试图沿着 G 从 p 到 q 携带 $\mathbf{w}(p)$ 一小段距离 ϵ 来创建平行向量 $\mathbf{w}_{\parallel}(p \rightsquigarrow q)$，我们所知道的是 α 保持不变，长度 $\mathbf{w}_{\parallel}$ 保持不变。 那么 $\mathbf{w}_{\parallel}(p \rightsquigarrow q)$ 可以位于图示的方向圆锥 $\mathcal{C}$ 上的任何位置； 我们应该选择 $\mathcal{C}$ 的哪个母线？



[29.3] 假设 $\mathbf{w}(p)$ 与测地线 G 成 α 角。为了将 $\mathbf{w}(p)$ 距离 ϵ 沿测地线 G 平行移动到 q，在 q 处构建与 G 成 α 角的向量圆锥 $\mathcal{C}$，然后选择尖端最接近 $\mathbf{w}(p)$ 尖端的那个。

对于这个构造，以及随后的每个其他构造，问自己如何在 $\mathbb{R}^3$ 中进行类似的结构就足够了，因为当 $\epsilon \to 0$ 时，两个构造必须一致。 好吧，在 $\mathbb{R}^3$ 中，与 $\mathbf{w}_o$ 平行的 $\mathcal{C}$ 的母线是最接近的。 也就是说，如图所示，

$\mathbf{w}_{\parallel}(p \rightsquigarrow q)$ 是 $\mathcal{C}$ 的尖端最接近 $\mathbf{w}(p)$ 的尖端的母线。

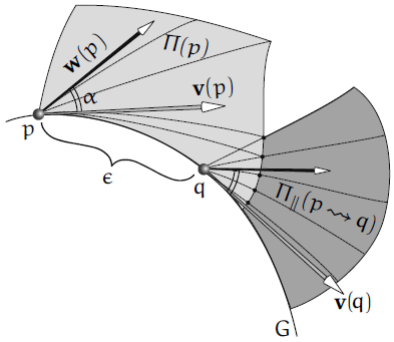
通过一遍又一遍地重复这个过程，我们可以沿着 G 将 $\mathbf{w}(p)$ 平行移动到任意距离。

29.3.2平行移动平面内的常角度

下一个构造，据我们所知是阿诺德（Arnol’d (1989, pp. 305-306)）描述的，再次从 $\mathbb{R}^3$ 中汲取灵感。

为了在 $\mathbb{R}^3$ 中沿欧氏直线 G （切线为 $\mathbf{v}$）平行移动 $\mathbf{w}(p)$，令 $\Pi(p)$ 为 $\mathbf{w}(p)$ 和 $\mathbf{v}(p)$ 所张的平面，两个向量之间的夹角为 α（和以前一样）。 现在沿 G 平行移动 $\Pi(p)$ 以获得 $\Pi_{\|}$。 最后，沿 G 平行移动 $\mathbf{w}(p)$ 成为 $\Pi_{\|}$ 中与 $\mathbf{v}$ 成 α 角的唯一向量。

当然，在 $\mathbb{R}^3$ 中这是浪费时间，因为我们有绝对的平行性，允许我们在没有 $\Pi_{\|}$ 的帮助下沿 G 移动 $\mathbf{w}(p)$。然而，现在这提出在弯曲的 3-流形（或 n-流形）中的新构造。见[29.4]。



[29.4] 我们通过从 p 向 $\mathbf{w}(p)$ 和 $\mathbf{v}(p)$ 所张的所有方向发射测地线来构造“平面”$\Pi(p)$。 然后，我们（大约）将 $\Pi(p)$ 平行移动到 q 以形成 $\Pi_{\|}(p \rightsquigarrow q)$，由从 q 发出并通过 $\Pi(p)$ 的相邻点的测地线生成。

在 $\mathbb{R}^3$ 中，$\Pi(p)$ 由从 p 发出的所有直线组成，其方向是 $\mathbf{v}(p)$ 和 $\mathbf{w}(p)$ 的线性组合。以此类推，在我们的弯曲 3-流形中，令 $\Pi(p)$ 现在表示由从 p 发出的所有测地线组成的“平面”，其方向是 $\mathbf{v}(p)$ 和 $\mathbf{w}(p)$ 的线性组合； 当然，这个“平面”实际上是一个弯曲的 2-曲面。 注意，这个曲面 $\Pi(p)$ 必然包含 G，即方向 $\mathbf{v}(p)$ 上的测地线。

和前面的构造一样，让 q 位于沿 G 距离 ϵ 处。通过构造，$\Pi(p)$ 必然包含 q 和 G 在 q 处的新方向，即 $\mathbf{v}(q)$。 如果我们放大 $\Pi(p)$ 围绕 q 的小区域，它看起来像一个欧氏平面。我们现在可以通过启动构建一个新的曲面 $\Pi_{\|}(p \rightsquigarrow q)$——$\Pi(p)$ 到 q 的平行平移 在围绕 q 的 $\Pi(p)$ 的这个小的平面区域内，从 q 在所有方向上的测地线。 图 [29.4] 说明了其中一些测地线（虚线），从 q 发射并通过 $\Pi(p)$ 的附近（黑点）点。 更准确地说，$\Pi_{\|}(p \rightsquigarrow q)$ 的测地线生成器是通过将黑点拉入 q 的极限来获得的。

一遍又一遍地重复这个构造，我们可以近似沿着 G 平行移动 $\Pi(p)$。最后，让 $\epsilon \to 0$，我们得到一个连续变化的、沿 G 平行移动的 2-曲面 $\Pi_{\|}$，于是，

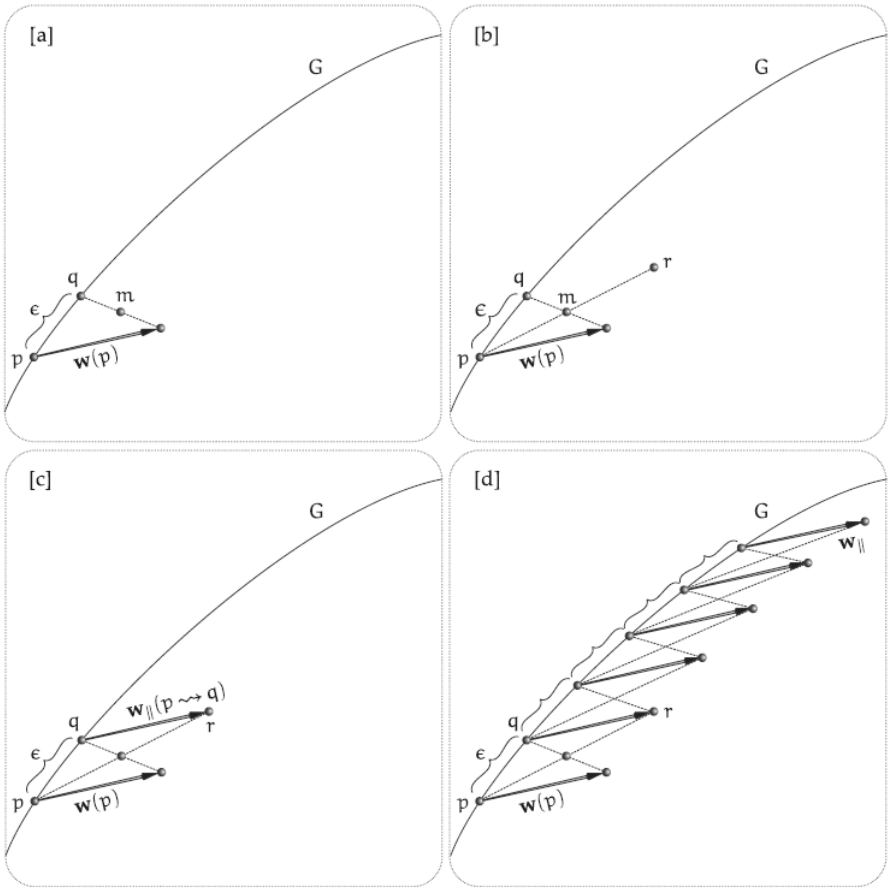
为了沿测地线 G 平行移动 $\mathbf{w}(p)$，我们保持它的长度不变，并保持它与 G 的角度 α，同时始终保持它与平行移动的 2-曲面 $\Pi_{\|}$ 相切。

29.3.3希尔德梯子

我们最终的构造也是最简单的。它首先由阿尔弗雷德·希尔德（Alfred Schild 1921-1977）在 1970 年（未发表的）普林斯顿讲座中提出，后来在米斯纳、索恩和惠勒（Misner、Thorne 和 Wheeler，1973 年）《引力论》中被命名为**希尔德的梯子**。

在 2-曲面上，想象位于曲面中的一个非常短的切向量（实际上位于曲面的 $\mathbb{R}^2$ 切平面中），通常对我们很有帮助。 同样，在 n-流形中，可以将 $\mathbb{R}^n$ 切空间中的一个非常短的向量想象成位于流形中。 因此，由于平行移动（根据定义）保持长度，如果我们希望在 p 处平行移动一个长切向量，我们可以首先在 p 处的 $\mathbb{R}^n$ 切空间中将其缩小某个大因子 N，然后将其想象成位于流形中，然后将其平行移动，最后在目的地的 $\mathbb{R}^n$ 切空间中将其扩展相同的大因子 N，从而恢复其原始长度。

希尔德的梯子构造利用这种自由来想象要沿测地线 G 平行移动的向量 $\mathbf{w}(p)$ 非常短，因此可以将其描绘为流形中非常短的测地线段。有了这种理解， [29.5] 的标题中详细说明了希尔德的梯子构造。



[29.5] **希尔德的梯子**。 [a] 沿测地线 G 行进距离 ϵ 以构造点 q。 用测地线段将 $\mathbf{w}(\mathrm{p})$ 的尖端连接到 q，并标记其中点 m。 [b] 用另一个测地线段将 p 连接到 m，然后将其延伸到 m 之外相等的距离，以构建以 r 为终点的线段。 [c] 将 q 连接到 r 以创建希尔德梯子的第一个梯级 $\asymp \mathbf{w}_{\parallel}(\mathrm{p} \rightsquigarrow \mathrm{q})$。 [d] 重复构造以向希尔德的梯子添加更多梯级。 最后，令 $\epsilon \to 0$，从而沿着 G 平行移动 $\mathbf{w}_{\parallel}$。

29.4 内蕴（又名“协变”）导数 $\boldsymbol{\nabla}_{\mathbf{v}}$

以前，我们用 $\nabla_{\mathbf{v}}$ 表示 $\mathbb{R}^3$ 中的方向导数，我们用 $\mathrm{D}_{\mathbf{v}}$ 表示 2-曲面内的内蕴导数。 在外在术语中，我们在（23.3）中看到 $\mathrm{D}_{\mathbf{v}}$ 可以看作是 $\nabla_{\mathbf{v}}$ 沿其法线 $\mathbf{n}$ 到曲面的投影 $\mathcal{P}$：

$$\mathrm{D}_{\mathbf{v}}\mathbf{w} = \mathcal{P}[\nabla_{\mathbf{v}}\mathbf{w}] = \nabla_{\mathbf{v}}\mathbf{w} - (\mathbf{n}\cdot\nabla_{\mathbf{v}}\mathbf{w})\mathbf{n}$$

然而，正如我们已经指出的，我们不会考虑将我们的 n-流形嵌入到更高维空间中，因此不会寻找或找到上述公式的类似物。 相反，我们将完全内蕴地思考。

幸运的是，我们已经在（23.2）中给出了这个内蕴导数的内蕴定义，我们甚至在 [23.3] 中画了它的图。 让我们简要回顾一下这种构造。

我们的任务是找到向量场 $\mathbf{w}$ 在单位向量 $\mathbf{v}$ 的方向上离开点 p 时的变化率。我们沿着 $\epsilon\mathbf{v}$ 离开 p 一个小的（最终趋于零）距离 ϵ，到达 q。为了查看新向量 $\mathbf{w}(\mathrm{q})$ 从其原始值 $\mathbf{w}(\mathrm{p})$ 变化了多少，我们将它从 q 平行移动回 p 以成为 $\mathbf{w}_{\parallel}(\mathrm{q} \rightsquigarrow \mathrm{p})$。然后我们找到变化 $[\mathbf{w}_{\parallel}(\mathrm{q} \rightsquigarrow \mathrm{p}) - \mathbf{w}(\mathrm{p})]$，最后我们除以 ϵ 来求变化率。

既然我们已经知道如何在 n-流形中进行平行移动 [三种方式！]，我们可以完全按照刚才描述的那样定义 n-流形中的内蕴（也称为“协变”）导数。 只是符号发生了变化：此后，我们采用标准的**粗体** Nabla 符号——$\boldsymbol{\nabla}$——来表示**内蕴导数**，

$$\boldsymbol{\nabla}_{\mathbf{v}}\mathbf{w} \asymp \frac{\mathbf{w}_{\parallel}(\mathrm{q} \rightsquigarrow \mathrm{p}) - \mathbf{w}(\mathrm{p})}{\epsilon} \tag{29.2}$$

对于我们即将到来的用途，证明内蕴导数不是根据变化率而是更直接地根据实际变化本身来考虑会更有用：

$$\mathbf{w}_{\parallel}(\mathrm{q} \rightsquigarrow \mathrm{p}) - \mathbf{w}(\mathrm{p}) \asymp \epsilon\boldsymbol{\nabla}_{\mathbf{v}}\mathbf{w} = \boldsymbol{\nabla}_{\epsilon\mathbf{v}}\mathbf{w}$$

令 $\delta_{\mathrm{pq}}\mathbf{w}$ 表示 $\mathbf{w}$ 沿向量 $\boldsymbol{\epsilon} \equiv \epsilon\mathbf{v}$ 从 p 到 q 的微小内蕴变化，则

$$\delta_{\mathrm{pq}}\mathbf{w} = \mathbf{w}\text{ 从尾部到尖端的内蕴变化}\boldsymbol{\epsilon} \asymp \boldsymbol{\nabla}_{\boldsymbol{\epsilon}}\mathbf{w}\text{。} \tag{29.3}$$

从这个定义可以直接得出，

$$\boxed{\nabla_{\mathbf{v}}\mathbf{w}_{\parallel} = 0 \quad \Leftrightarrow \quad \mathbf{w}_{\parallel} \text{ 沿 } \mathbf{v} \text{ 平行移动}}$$

如果 $\mathbf{v}$ 是测地线的速度，则它沿自身平行移动，因此**测地线方程** （23.5）现在采用以下形式，

$$\boxed{\nabla_{\mathbf{v}}\mathbf{v} = 0} \tag{29.4}$$

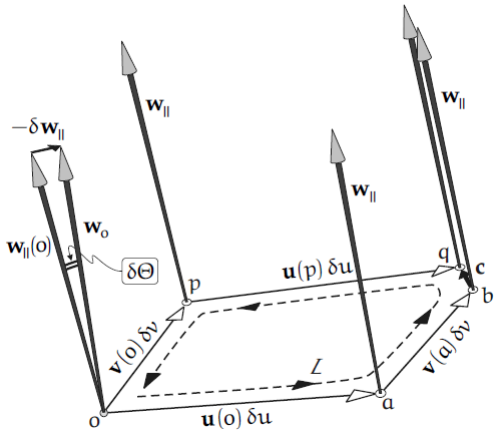
我们应该注意到，如果我们允许粒子在沿测地线行进时加速或减速，那么我们会得到测地线方程的更一般形式，即 $\nabla_{\mathbf{v}}\mathbf{v} \propto \mathbf{v}$，它表示 $\mathbf{v}$ 的方向内蕴地保持不变，但大小可以变化。

29.5 黎曼曲率张量

29.5.1围绕小“平行四边形”的平行移动

就像我们在 2-流形（曲面）的情况下所做的那样，我们现在可以通过围绕一个小环路 L 平行移动单位向量 $\mathbf{w}_{\parallel}$ 来研究 n-流形中的曲率。更具体地说，我们将尝试将 L 构造为一个平行四边形，其边由两个（单位）向量场 $\mathbf{u}$ 和 $\mathbf{v}$ 构成。

因此，如 [29.6] 中所示，从 o 开始，我们在这两个方向上布置短向量，$\mathbf{u}(\mathrm{o})\,\delta\mathrm{u}$（将 o 连接到 a）和 $\mathbf{v}(\mathrm{o})\,\delta\mathrm{v}$（将 o 连接到 p）。为了尝试创建平行四边形，我们现在布置 $\mathbf{u}(\mathrm{p})\,\delta\mathrm{u}$（将 p 连接到 q）和 $\mathbf{v}(\mathrm{a})\,\delta\mathrm{v}$（将 a 连接到 b）。 但问题是，一般来说，$\mathrm{q} \neq \mathrm{b}$：“平行四边形”无法闭合！



[29.6] **黎曼曲率的几何意义**。 从 o 开始，我们从两个向量场 $\mathbf{u}$ 和 $\mathbf{v}$ 中创建一个小的平行四边形，然后将一个初始向量 $\mathbf{w}_{\mathrm{o}}$ 绕环路 L 平行移动得到 $\mathbf{w}_{\parallel}$。 当它作为 $\mathbf{w}_{\parallel}(\mathrm{o})$ 返回到 o 时，它已经改变了所示的**向量和乐** $\delta\mathbf{w}_{\parallel} = -\mathcal{R}(\mathbf{u}\delta\mathrm{u}, \mathbf{v}\,\delta\mathrm{v})\mathbf{w}$，其中 $\mathcal{R}$ 是**黎曼曲率算子**。

在下一小节中，我们将找到连接 b 到 q 的极小“间隙闭合”向量 $\mathbf{c}$ 的公式。 不过，现在，只是假设我们知道如何缩小这个差距，因此能够沿闭环 $\mathrm{L} = \mathrm{oabqpo}$ 平行移动 $\mathbf{w}$。

在 2-曲面内，平行移动的初始向量 $\mathbf{w}_o$ 只能与 $\mathbf{u}$ 和 $\mathbf{v}$ 位于同一平面内。3-流形（或更一般地 n-流形）中的根本新特征是 $\mathbf{w}$ 现在可以伸出它平行移动的小环所在平面。

设置好场景后，让我们绕 L平行移动 $\mathbf{w}_o$ 以得到 $\mathbf{w}_\parallel$，将其返回到 o 成为 $\mathbf{w}_\parallel(o)$。然后我们可以将向量和乐[175]定义为由曲率引起的 $\mathbf{w}_\parallel$ 的净变化：

$$\boxed{\delta\mathbf{w}_\parallel \equiv \mathbf{w}_\parallel(\text{返回 o 时}) - \mathbf{w}(\text{离开 o 时})}$$

在 2-曲面内，我们引入了和乐算子 $\mathcal{R}(\mathrm{L})$，它（当应用于 $\mathbf{w}$ 时）给出了 $\mathbf{w}$ 在绕 L 平行移动之后的净旋转（和乐）。现在，在一个 3-流形中，或实际上是一个 n-流形，L 可能位于无数个不同的平面中。 此外，在 2-曲面中，整个切平面刚性旋转，因为其中的向量是平行移动的，因此我们不需要注意传输的是哪个向量 $\mathbf{w}$：它们都旋转了相同的量 $\mathcal{R}(\mathrm{L})$。 但是在 3-流形中，$\mathbf{w}$ 可以伸出 L 的平面，而在 n-流形中，它可以有许多独立的指向方式。至关重要的是，向量和乐现在确实取决于哪个向量在围绕环路平行移动。

出于这两个原因，我们必须改进和推广我们之前的符号，并引入与平行四边形 L 的边缘相关的黎曼曲率算子 R，然后作用于平行移动的向量，产生向量和乐：

$$\boxed{-\delta\mathbf{w}_\parallel \equiv \mathcal{R}(\mathbf{u}\,\delta\mathrm{u}, \mathbf{v}\,\delta\mathrm{v})\mathbf{w}.}$$

众所周知，2 –曲面的曲率 $\mathcal{K}$ 完全被单位面积的和乐捕获。 在 n-流形中，我们同样可以查看初始向量 $\mathbf{w}_o$ 和平行移动向量 $\mathbf{w}_\parallel(o)$ 返回到 o 时的角度 $\delta\Theta$，如 [29.6] 所示。 由于 $\mathbf{w}$ 是单位向量，因此距离 $|\delta\mathbf{w}|$ 它的尖端旋转最终等于旋转角度 $\delta\Theta$。

为简单起见，假设 $\mathbf{u}$ 和 $\mathbf{v}$ 是正交的，因此我们的平行四边形是面积为 $\delta\mathcal{A} = \delta\mathrm{u}\,\delta\mathrm{v}$ 的矩形。然后我们可以推广 $\mathcal{K}$ 并提取标量曲率 $\mathcal{K}(\mathbf{u},\mathbf{v};\mathbf{w})$，再次定义为每单位面积的旋转：

$$\mathcal{K}(\mathbf{u},\mathbf{v};\mathbf{w}) \asymp \frac{\delta\Theta}{\delta\mathcal{A}} \asymp \frac{|\delta\mathbf{w}_\parallel|}{\delta\mathcal{A}} \asymp |\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{w}|$$

然而，很明显这不再是一个令人满意的曲率测量，因为我们已经完全失去了关于向量和乐方向 $\delta\mathbf{w}_\parallel$ 的关键信息。 正如我们将看到的，黎曼张量是编码完整曲率信息的几何对象：$\mathbf{w}_\parallel$ 尖端远离 $\mathbf{w}_o$ 的角度 $\delta\Theta$，以及它的方向。

29.5.2用向量交换子闭合“平行四边形”

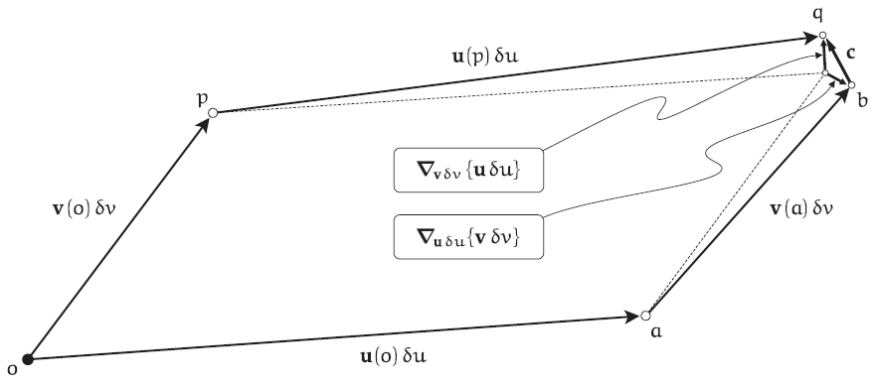
闭合我们的接近但未命中平行四边形的向量 $\mathbf{c}$ 非常小，（最少[176]）是 $\delta\mathrm{u}\,\delta\mathrm{v}$ 阶，所以

[175] 我们发明了这个术语，因为我们不知道文献中这个概念的标准名称。
[176] 对于绑定到坐标网格的矢量场，间隙完全消失。

忽略它仍然会产生一个很好的曲率近似值。 但是为了获得黎曼张量的数学上完美的描述，我们必须围绕一个闭环平行移动我们的向量，因此我们必须用 $\mathbf{c}$ 闭合“平行四边形”。

正如我们在几何上直接从[29.7]中看到的，这个闭合平行四边形间隙的非常短的向量可以表示为其边的**交换子**[177]（commutator）：

$$\boxed{\mathbf{c} \asymp [\mathbf{v}\delta v, \mathbf{u}\delta v] = [\mathbf{v}, \mathbf{u}]\delta u \delta v}$$



[29.7] 几何证明闭合“平行四边形”间隙的向量 $\mathbf{c}$ 是其边的交换子：$\mathbf{c} \asymp \nabla_{\mathbf{v}\delta v}\{\mathbf{u}\delta u\} - \nabla_{\mathbf{u}\delta u}\{\mathbf{v}\,\delta v\} = [\mathbf{v}, \mathbf{u}]\delta u \delta v$。

其中

$$\boxed{[\mathbf{v}, \mathbf{u}] \equiv \nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v}}$$

我们注意到一个我们很快就会需要的简单事实，即交换子是反对称的：

$$\boxed{[\mathbf{v}, \mathbf{u}] = -[\mathbf{u}, \mathbf{v}]}$$

29.5.3一般黎曼曲率公式

我们现在可以回到当 $\mathbf{w}_{\parallel}$ 绕闭环 $L = oabqpo$ 平行移动时，寻找变化 $\delta\mathbf{w}_{\parallel}$ 的问题。。

为了鼓励以下论点，让我们简要回顾一下霍普夫对 GGB 的内蕴证明的第一个关键步骤——引入开放曲线的“和乐”。 如 [26.1] 所示，我们引入了一个基准向量场 $\mathbf{U}$，使我们能够将开放曲线 K 的 $\mathcal{R}_U(K)$ 定义为当我们沿 K 行进时，从 $\mathbf{U}$ 到平行移动向量 $\mathbf{w}_{\parallel}$ 的角度 $\angle U\mathbf{w}_{\parallel}$ 的净改变：

$$\mathcal{R}_U(K) \equiv \delta_K(\angle U\mathbf{w}_{\parallel})$$

[177] 也称为李括号，以伟大的挪威数学家 Sophus Lie（发音为“lee”）（1842-1899）命名。

正如我们当时所指出的，下标 $\mathbf{U}$ 是必不可少的，因为 $\mathcal{R}_{\mathrm{U}}(\mathrm{K})$ 确实取决于我们对 $\mathbf{U}$ 的任意选择。因此 $\mathcal{R}_{\mathrm{U}}(\mathrm{K})$ 没有真正的数学意义； 它只是寻找具有几何意义的闭环和乐的垫脚石。

实际上，闭合多边形环路 L 的和乐 $\mathcal{R}(\mathrm{L})$ 可以表示为其边的“和乐”之和，虽然每个单独的和乐确实取决于 $\mathbf{U}$ 的任意选择，但它们的总和 $\mathcal{R}(\mathrm{L}) = \mathcal{K}(\mathrm{L})$ 测量内部的曲率，并且与 $\mathbf{U}$ 无关。

在 2-曲面中，我们不需要指定平行移动的向量 $\mathbf{w}$，因为当我们在曲面上移动时，整个切平面会刚性旋转。（这在 3-流形中不再正确：$\mathbf{w}$ 方向的初始选择确实会影响其围绕闭环的向量和乐。）

还是在一个 2-曲面中，目前我们可以自由选择 $\mathbf{w}$，所以让我们选择 $\mathbf{w} = \mathbf{U}$。 现在考虑这个和乐的负数——$\mathbf{U}$ 相对于平行移动向量 $\mathbf{U}_{\parallel}$ 的变化——沿着一个非常短的向量 ϵ。那么和乐就变成了 $\mathbf{U}$ 的内蕴导数：

$$\begin{aligned}-\mathcal{R}_{\mathrm{U}}(\epsilon) &= \delta_{\epsilon}(\angle \mathbf{U}_{\parallel}\mathbf{U}) \\ &\asymp \boldsymbol{\nabla}_{\epsilon}(\angle \mathbf{U}_{\parallel}\mathbf{U}) \asymp |\boldsymbol{\nabla}_{\epsilon}\mathbf{U}|\end{aligned}$$

我们现在可以回到手头的问题，并将这种见解应用到在 3-流形 [29.6] 内 $-\delta\mathbf{w}_{\parallel}$ 的评估中。如 [29.8] 所示，我们现在在包含平行四边形的整个区域中引入一个基准向量场 $\mathbf{w}$——这些是黑白向量，我们仅沿 oa 画出，以避免混乱。 这个向量场可以是完全任意的，除了现在哪个向量 $\mathbf{w}_{\mathrm{o}}$ 被平行移动确实很重要，所以我们坚持 $\mathbf{w}(\mathrm{o}) = \mathbf{w}_{\mathrm{o}}$。然后，使用（29.3），沿平行四边形第一条边的向量和乐的负数为[178]

$$-\delta_{\mathrm{oa}}\mathbf{w}_{\parallel} = \mathbf{w}(\mathrm{a}) - \mathbf{w}_{\parallel}(\mathrm{a}) \asymp \boldsymbol{\nabla}_{\mathbf{u}\delta\mathrm{u}}\mathbf{w}(\mathrm{o}) = \delta\mathrm{u}\boldsymbol{\nabla}_{\mathbf{u}}\mathbf{w}(\mathrm{o})$$

现在让我们重复应用（29.3）来求和任意向量场 $\mathbf{w}$ 中的这些变化，相对于平行移动的向量 $\mathbf{w}_{\parallel}$，沿着封闭的平行四边形的五个边。 当我们返回 o 时，当然 $\mathbf{w}(\mathrm{o})$ 与我们离开 o 时的值相同，所以 $\mathbf{w}$ 相对于 $\mathbf{w}_{\parallel}$ 的净变化完全是由于 $\mathbf{w}_{\parallel}$ 的绝对变化，并且与向量场 $\mathbf{w}$ 的任意选择无关——这个净变化就是我们寻求的向量和乐。

我们详细说明了这个求和的每一步，并鼓励你参考 [29.8] 来理解每条线的几何意义。吸收了这个关键论点的每一步之后，值得你花时间退后一步，把它作为一个单一的、完整的想法来理解：

$$\begin{aligned}-\delta\mathbf{w}_{\parallel} &= -\left[\delta_{\mathrm{oa}}\mathbf{w}_{\parallel} + \delta_{\mathrm{ab}}\mathbf{w}_{\parallel} + \delta_{\mathrm{bq}}\mathbf{w}_{\parallel} + \delta_{\mathrm{qp}}\mathbf{w}_{\parallel} + \delta_{\mathrm{po}}\mathbf{w}_{\parallel}\right] \\ &\asymp \delta\mathrm{u}\boldsymbol{\nabla}_{\mathbf{u}}\mathbf{w}(\mathrm{o}) + \delta\mathrm{v}\boldsymbol{\nabla}_{\mathbf{v}}\mathbf{w}(\mathrm{a}) + \boldsymbol{\nabla}_{\mathbf{c}}\mathbf{w}(\mathrm{b}) - \delta\mathrm{u}\boldsymbol{\nabla}_{\mathbf{u}}\mathbf{w}(\mathrm{p}) - \delta\mathrm{v}\boldsymbol{\nabla}_{\mathbf{v}}\mathbf{w}(\mathrm{o}) \\ &= \delta\mathrm{v}\{\boldsymbol{\nabla}_{\mathbf{v}}\mathbf{w}(\mathrm{a}) - \boldsymbol{\nabla}_{\mathbf{v}}\mathbf{w}(\mathrm{o})\} - \delta\mathrm{u}\{\boldsymbol{\nabla}_{\mathbf{u}}\mathbf{w}(\mathrm{p}) - \boldsymbol{\nabla}_{\mathbf{u}}\mathbf{w}(\mathrm{o})\} + \delta\mathrm{u}\delta\mathrm{v}\boldsymbol{\nabla}_{[\mathbf{v},\mathbf{u}]}\mathbf{w}(\mathrm{o}) \\ &\asymp \delta\mathrm{v}\{\delta\mathrm{u}\boldsymbol{\nabla}_{\mathbf{u}}\boldsymbol{\nabla}_{\mathbf{v}}\,\mathbf{w}\} - \delta\mathrm{u}\{\delta\mathrm{v}\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{u}}\mathbf{w}\} - \delta\mathrm{u}\delta\mathrm{v}\boldsymbol{\nabla}_{[\mathbf{u},\mathbf{v}]}\mathbf{w}\end{aligned}$$

---

[178] 在此以及在下文中，将通过评估段中点处的导数来获得更准确的近似值。 然而，由于我们将把整个图形缩小到点 o，并在这个极限内处理最终平等，我们认为如果不采取这个额外的步骤，这个论证可能更容易理解。

$$= \delta u \delta v \{\nabla_{\mathbf{u}}\nabla_{\mathbf{v}} - \nabla_{\mathbf{v}}\nabla_{\mathbf{u}} - \nabla_{[\mathbf{u},\mathbf{v}]}\}\mathbf{w}(o)$$

$$= \delta u \delta v\, \mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w}(o)$$

我们已经到达了我们戏剧中的一个重大事件——**用黎曼曲率表示的向量和乐：**

$$\boxed{-\frac{\delta \mathbf{w}_{\parallel}}{\delta u \delta v} \asymp \mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w} = \{[\nabla_{\mathbf{u}}, \nabla_{\mathbf{v}}] - \nabla_{[\mathbf{u},\mathbf{v}]}\}\mathbf{w}} \tag{29.5}$$

我们现在看到 [29.6] 中引入的**黎曼曲率算子**实际上是

$$\boxed{\mathcal{R}(\mathbf{u},\mathbf{v}) = [\nabla_{\mathbf{u}}, \nabla_{\mathbf{v}}] - \nabla_{[\mathbf{u},\mathbf{v}]}} \tag{29.6}$$

（29.5）左边的减号很重要，所以让我们提醒读者注意它的起源和几何意义。 右侧的运算符告诉我们在遍历循环时（任意）基准向量场 $\mathbf{w}$（相对于 $\mathbf{w}_{\parallel}$）的净变化。 但是我们寻求的几何意义的量是向量和乐 $\delta\mathbf{w}_{\parallel}$，它正好相反：$\mathbf{w}_{\parallel}$ 相对于 $\mathbf{w}$ 的变化。

请注意，$\mathcal{R}$ 是反对称的：

$$\boxed{\mathcal{R}(\mathbf{u},\mathbf{v}) = -\mathcal{R}(\mathbf{v},\mathbf{u}) \quad \Longrightarrow \quad \mathcal{R}(\mathbf{u},\mathbf{u}) = 0} \tag{29.7}$$

第一个方程的真实性在几何上是清楚的，因为它简单地断言，如果我们反转遍历平行四边形的方向，则向量和乐也会反转。 同样，第二个等式是说，如果任何向量沿任意向量 $\mathbf{u}$ 来回平行移动一小段距离，则它保持不变。

现在只需要稍微改变符号即可最终得出著名的**黎曼曲率张量 $\mathbf{R}$** 的标准定义——它是三个输入向量的向量值函数 $\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w})$：[179]

$$\boxed{\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w}) \equiv \mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w} = \{[\nabla_{\mathbf{u}}, \nabla_{\mathbf{v}}] - \nabla_{[\mathbf{u},\mathbf{v}]}\}\mathbf{w}} \tag{29.8}$$

约定警告：$\mathbf{R}$ 的槽中向量的排列因作者而异。 例如，Misner、Thorne 和 Wheeler (1973) 同意我们上面对 $\mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w}$ 的定义，但是在左侧，他们的 $\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w})$ 重新排列了这样他们的 $\mathbf{R}$ 就变成了我们的负数。

[179] 使用分号将平行四边形向量与围绕它并行传输的向量分开并不完全是标准，但我们认为它很有帮助（彭罗斯也如此认为！）。

29.5.4黎曼曲率是张量

正如我们将在第五幕中更详细地讨论的那样，**张量**（根据定义）是其多个向量[180]输入的线性函数。 正如我们现在解释的，**R** 确实是多线性的：

> 当其他两个保持固定时，$\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w})$ 是其三个输入中的每一个的线性函数。

此外，尽管公式（29.8）是由导数构成的，但黎曼曲率仅取决于计算 **R** 点处的三个向量的值——它与向量在该点附近如何变化无关。 在计算层面上，这看起来完全是自相矛盾的，但它确实具有几何意义，因为 **R** 告诉我们空间本身的曲率，而向量仅用于挑选出该曲率的特定部分（或**分量**）。

总结一下，这里有

> **张量的两个定义属性：**
>
> - 输出线性依赖于每个输入向量。
> - 输出仅取决于评估点的输入向量。

首先，让我们在 $\mathbf{w}$-槽 建立线性。 假设 $\mathbf{w}=\mathrm{k}_1\mathbf{w}_1+\mathrm{k}_2\mathbf{w}_2$，其中 k_1 和 k_2 是常数。 如果 $\mathbf{w}$、$\mathbf{w}_1$ 和 $\mathbf{w}_2$ 分别绕着一个小的（最终趋于零的）边沿 $\mathbf{u}$ 和 $\mathbf{v}$ 的平行四边形 L 平行移动（从 o 开始），则

$$\delta\mathbf{w}_{\parallel}=\mathrm{k}_1\delta[\mathbf{w}_1]_{\parallel}+\mathrm{k}_2\delta[\mathbf{w}_2]_{\parallel}$$

所以几何解释（29.5）意味着

$$\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathrm{k}_1\mathbf{w}_1+\mathrm{k}_2\mathbf{w}_2)=\mathrm{k}_1\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w}_1)+\mathrm{k}_2\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w}_2)$$

当然，这也可以通过计算立即确认：只需使用公式（29.8）和内蕴导数线性作用的简单事实，例如，

$$\boldsymbol{\nabla}_{\mathbf{v}}[\mathrm{k}_1\mathbf{w}_1+\mathrm{k}_2\mathbf{w}_2]=\mathrm{k}_1\boldsymbol{\nabla}_{\mathbf{v}}\mathbf{w}_1+\mathrm{k}_2\boldsymbol{\nabla}_{\mathbf{v}}\mathbf{w}_2$$

为了开始测试张量的第二个要求，假设我们将 o 附近的 $\mathbf{w}^{\mathrm{OLD}}$ 更改为 $\mathbf{w}^{\mathrm{NEW}}=$

---

[180] 在第五幕中，我们将把它推广到标准定义，即我们允许输入向量和 1-形式（在第五幕开始时介绍）。

$f\mathbf{w}^{OLD}$，其中 f 是位置的任意（可微分）函数，除了 $f(o) = 1$，因此 $\mathbf{w}^{NEW}(o) = \mathbf{w}^{OLD}(o)$。

在这一点上，我们强烈建议你暂停一下，以便实际执行有些冗长（但简单）的计算，证明当在 o 处评估时，

$$\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w}^{NEW}) = \mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w}^{OLD})$$

一旦你做这个计算，你会看到出现了四个涉及 f 的导数的项，但是，“奇迹般地”，它们都相互抵消了！ （注意：这不是值得庆祝的——请参阅序言中“虚假奇迹”的定义！）

幸运的是，这可以直接从几何角度来理解，因为以任何方式沿 L 改变基准向量场 $\mathbf{w}$（不仅仅是按 f 缩放）对初始向量围绕 L 的几何意义的平行移动绝对没有影响， 因此对向量和乐绝对没有影响！ 因此 $\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w})$ 确实只取决于 w 在 o 处的值。

现在让我们把注意力转向前两个槽，让我们从张量的第二个定义属性开始。 在环路收缩到 o 的极限下，只需要 $\mathbf{u}(o)$ 和 $\mathbf{v}(o)$ 的值来构建 $\mathbf{w}$ 绕其平行移动的环路，因此向量和乐以及 $\mathbf{R}$ 仅取决于在 o 处这些输入向量的值。

为了见证几何视角的决定性优势，同时作为公式（29.8）的计算使用中的字符构建练习，让我们用 $\mathbf{u}$ 玩我们之前用 $\mathbf{w}$ 做的同样的游戏。 也就是说，让我们将 o 附近的 $\mathbf{u}^{OLD}$ 变为 $\mathbf{u}^{NEW} = f\mathbf{u}^{OLD}$，其中 f 是任意函数。 现在尝试证明（直接从公式中）

$$\mathbf{R}(\mathbf{u}^{NEW},\mathbf{v};\ \mathbf{w}) = \mathbf{R}(f\mathbf{u}^{OLD},\mathbf{v};\ \mathbf{w}) = f\,\mathbf{R}(\mathbf{u}^{OLD},\mathbf{v};\ \mathbf{w})$$

因此，如果我们再次坚持 $f(o) = 1$，以确保 $\mathbf{u}^{NEW}(o) = \mathbf{u}^{OLD}(o)$，那么

$$\mathbf{R}(\mathbf{u}^{NEW},\mathbf{v};\ \mathbf{w})(o) = \mathbf{R}(\mathbf{u}^{OLD},\mathbf{v};\ \mathbf{w})(o)$$

正如张量所要求的。

最后，让我们转向前两个槽的线性。 由于反对称（29.7），证明任一槽的线性就足够了。在这里，几何和计算视角之间基本上没有区别，我们将非常简短的计算留给你，它确定，

$$\mathbf{R}(k_1\mathbf{u}_1 + k_2\mathbf{u}_2,\mathbf{v};\ \mathbf{w}) = k_1\mathbf{R}(\mathbf{u}_1,\mathbf{v};\ \mathbf{w}) + k_2\mathbf{R}(\mathbf{u}_2,\mathbf{v};\ \mathbf{w})$$

隆重结论：黎曼曲率确实是张量！

29.5.5黎曼张量的分量

为了给出几何对象的数值描述，我们引入了一组正交基向量 $\{\mathbf{e}_i\}$，用于 n-流形每个点的 $\mathbb{R}^n$ 切空间。 然后，例如，几何向量 $\mathbf{u}$ 由其数值分量 $\{u^i\}$ 表示，其中 $\mathbf{u} = \sum_i u^i\mathbf{e}_i$ 。警告：处理分量时，我们必须记住上标是标签，而不是幂！

更复杂的几何对象，例如黎曼张量，需要多个不同的索引来描述它们的分量。 这导致大量的求和。 但事实证明，我们总是可以安排被求和的索引一次作为上标出现，一次作为下标出现。

因此，爱因斯坦引入了一个简单的、化杂乱为有序的约定，称为**爱因斯坦求和约定**，由此可以轻松地理解这样一对索引的总和，而省略求和符号。 例如，

$$\boxed{\mathbf{u} = \sum_i u^i \mathbf{e}_i \quad \Leftrightarrow \quad \textbf{爱因斯坦求和约定：}\ \mathbf{u} = u^i \mathbf{e}_i}$$

为了找到黎曼张量的分量，我们将其三个输入向量分解为分量：

$$\mathbf{u} = u^i \mathbf{e}_i, \qquad \mathbf{v} = v^j \mathbf{e}_j, \qquad \text{和} \quad \mathbf{w} = w^k \mathbf{e}_k$$

于是

$$\begin{aligned} \mathbf{R}(\mathbf{u}, \mathbf{v};\ \mathbf{w}) &= \mathbf{R}\left(u^i \mathbf{e}_i, v^j \mathbf{e}_j;\ w^k \mathbf{e}_k\right) \\ &= \mathbf{R}\left(\mathbf{e}_i, \mathbf{e}_j;\ \mathbf{e}_k\right) u^i v^j w^k \end{aligned}$$

我们现在可以将**黎曼张量的分量** $R_{ijk}{}^l$ 定义为当黎曼张量作用于三个基向量时产生的向量的分量：

$$\boxed{\mathbf{R}\left(\mathbf{e}_i, \mathbf{e}_j;\ \mathbf{e}_k\right) \equiv R_{ijk}{}^l \mathbf{e}_l}$$

因此，**R** 对三个通用向量的影响可以很容易地用这些分量表示：

$$\mathbf{R}(\mathbf{u}, \mathbf{v};\ \mathbf{w}) = \left[R_{ijk}{}^l u^i v^j w^k\right] \mathbf{e}_l$$

为了以后使用，我们还定义，

$$\boxed{R_{ijkm} \equiv \mathbf{R}\left(\mathbf{e}_i, \mathbf{e}_j;\ \mathbf{e}_k\right) \cdot \mathbf{e}_m} \tag{29.9}$$

在我们选择的正交基中，如果 $l \neq m$，$\mathbf{e}_l \cdot \mathbf{e}_m = 0$，当 $l = m$ 时它等于 1，所以

$$\boxed{R_{ijkm} = R_{ijk}{}^m} \tag{29.10}$$

29.5.6对于给定的 $\mathbf{w}_o$，向量和乐仅取决于环路的平面及其面积

在 2-曲面内，小环 L 的和乐 $\mathcal{R}(\mathrm{L}) \asymp \mathcal{K}\delta\mathcal{A}$ 仅取决于其面积 $\delta\mathcal{A}$，而与其形状无关。在 n-流形中，现在 L 的平面有许多独立的选择。因此，让我们定义，

$$\Pi(\mathbf{u},\mathbf{v}) \equiv \text{由 } \mathbf{u} \text{ 和 } \mathbf{v} \text{ 张成的平面。}$$

对 $\Pi(\mathbf{u},\mathbf{v})$ 的选择肯定决定了向量全集。 然而，对于给定的 $\mathbf{w}_o$，向量和乐仅取决于 Π 本身，以及环路的面积：

> 如果我们在平面 Π 中围绕一个小的（最终趋于零的）平行四边形平行移动 $\mathbf{w}_o$，那么向量和乐与平行四边形的面积 $\delta\mathcal{A}$ 成正比，并且与它的形状无关。 (29.11)

我们将通过计算来证明这一点，以说明我们全新的张量分量工具的强大功能。让 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 在 L 的平面内选择。然后，使用爱因斯坦求和约定，

$$\mathbf{u}\,\delta u = \delta u^1\mathbf{e}_1 + \delta u^2\mathbf{e}_2 = \delta u^i\mathbf{e}_i \quad \text{和} \quad \mathbf{v}\,\delta v = \delta v^1\mathbf{e}_1 + \delta v^2\mathbf{e}_2 = \delta v^j\mathbf{e}_j$$

因此环的面积是

$$\delta\mathcal{A} = \begin{vmatrix} \delta u^1 & \delta v^1 \\ \delta u^2 & \delta v^2 \end{vmatrix} = \delta u^1\delta v^2 - \delta u^2\delta v^1$$

令 $\mathbf{w}_o = w_o^k\,\mathbf{e}_k$，且 $\delta\mathbf{w}_\parallel = \delta w_\parallel^l\mathbf{e}_l$。 所以

$$\begin{aligned} -\delta\mathbf{w}_\parallel &= -\delta w_\parallel^l\mathbf{e}_l \\ &= \mathbf{R}(\mathbf{u}\delta u, \mathbf{v}\delta v;\ \mathbf{w}_o) \\ &= \mathbf{R}\big(\delta u^i\mathbf{e}_i, \delta v^j\mathbf{e}_j;\ w_o^k\mathbf{e}_k\big) \\ &= \delta u^i\,\delta v^j\,w_o^k\,R_{ijk}{}^l\mathbf{e}_l \end{aligned}$$

但是，凭借（29.7），

$$R_{11k}{}^l = \mathbf{0} = R_{22k}{}^l \quad \text{和} \quad R_{21k}{}^l = -R_{12k}{}^l$$

因此，向量和乐的分量由下式给出

$$\boxed{-\delta\mathbf{w}_{\parallel} = -\delta w_{\parallel}^{l}\mathbf{e}_{l} \asymp \delta\mathcal{A}[R_{12k}{}^{l}w_{o}^{k}]\,\mathbf{e}_{l}} \tag{29.12}$$

完成（29.11）的证明。

29.5.7黎曼张量的对称性

天真地，人们会期望在（29.9）中定义的 3-流形中黎曼张量 $\mathbf{R}_{ijkm}$ 的独立分量的数量为 $3^4 = 81$。 然而我们已经说过，真正的数字只有六！ 这种显着减少是由于黎曼张量具有四个显著的代数对称性。

我们在 (29.7) 中已经遇到的第一个：$\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w})$ 在它的前两个槽中是反对称的，这意味着

$$\boxed{\mathbf{R}_{jikm} = -\mathbf{R}_{ijkm}}$$

这立即将独立分量的数量从 81 个减少到 54 个[练习]。

我们现在证明第二个不太明显的对称性，即 $\mathbf{R}_{ijmk}$ 在其后两个槽也是反对称的：

$$\boxed{\mathbf{R}_{ijmk} = -\mathbf{R}_{ijkm}} \tag{29.13}$$

这进一步将独立分量的数量从 54 个减少到 36 个[练习]。

由于平行移动保持长度，因此向量 $\mathbf{w}_o$ 作为稍微旋转的向量 $\mathbf{w}_{\parallel}(o)$ 返回到 o，因此连接向量 $\delta\mathbf{w}_{\parallel}$ 这些向量的尖端之间最终与它们都正交：

$$\delta\mathbf{w}_{\parallel}\cdot\mathbf{w}_o = 0 \quad\Longrightarrow\quad [\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\mathbf{w}_o]\cdot\mathbf{w}_o = 0$$

现在让 $\mathbf{x}$ 和 $\mathbf{y}$ 为任意向量，并设置 $\mathbf{w}_o = \mathbf{x}+\mathbf{y}$。所以，

$$\begin{aligned}
0 &= [\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})(\mathbf{x}+\mathbf{y})]\cdot(\mathbf{x}+\mathbf{y}) \\
&= [\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{x}]\cdot\mathbf{x} + [\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{x}]\cdot\mathbf{y} + [\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{y}]\cdot\mathbf{x} + [\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{y}]\cdot\mathbf{y} \\
&= 0 + [\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{x}]\cdot\mathbf{y} + [\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{y}]\cdot\mathbf{x} + 0
\end{aligned}$$

所以，

$$\boxed{[\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{x}]\cdot\mathbf{y} = -[\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{y}]\cdot\mathbf{x}} \tag{29.14}$$

所以由定义（29.9）立即可得（29.13）。

为了完整起见，我们现在说明黎曼张量的其余对称性。

首先，**代数比安基恒等式**[181]（Algebraic Bianchi Identity）指出，如果前三个向量循环置换，它们的和为零：

$$\boxed{\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{w} + \boldsymbol{\mathcal{R}}(\mathbf{v},\mathbf{w})\,\mathbf{u} + \boldsymbol{\mathcal{R}}(\mathbf{w},\mathbf{u})\,\mathbf{v} = \mathbf{0} \quad \Leftrightarrow \quad \mathrm{R}_{\mathrm{ijkm}} + \mathrm{R}_{\mathrm{jkim}} + \mathrm{R}_{\mathrm{kijm}} = 0} \tag{29.15}$$

练习 10 提供了这个结果的计算证明和几何证明。

接下来，黎曼张量在第一对和第二对向量的交换下是对称的：

$$\boxed{[\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{x}]\cdot\mathbf{y} = [\boldsymbol{\mathcal{R}}(\mathbf{x},\mathbf{y})\,\mathbf{u}]\cdot\mathbf{v} \quad \Leftrightarrow \quad \mathrm{R}_{\mathrm{ijkm}} = \mathrm{R}_{\mathrm{kmij}}} \tag{29.16}$$

尽管这非常有用，但它并不是真正的新对称，而是先前对称的结果。 有关计算证明，请参见练习 10。

使用这些对称性，你可以证明在 3-流形中，黎曼张量只有六个独立分量[练习][182]。 更一般地，练习 11 表明，在 n-流形中，分量的数量由 $\frac{1}{12}\mathrm{n}^2(\mathrm{n}^2-1)$ 给出，如先前在（29.1）中所述。

最后，除了上面的四个代数对称之外，还有第五个，微分对称，对爱因斯坦的引力理论至关重要——它被称为**微分比安基恒等式**[183]（Differential Bianchi Identity）：

$$\boxed{\boldsymbol{\nabla}_{\mathbf{x}}\,\boldsymbol{\mathcal{R}}(\mathbf{u},\mathbf{v})\,\mathbf{w} + \boldsymbol{\nabla}_{\mathbf{u}}\,\boldsymbol{\mathcal{R}}(\mathbf{v},\mathbf{x})\,\mathbf{w} + \boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\mathcal{R}}(\mathbf{x},\mathbf{u})\,\mathbf{w} = \mathbf{0}} \tag{29.17}$$

---

[181] 也称为第一比安基恒等式或比安基对称，但实际上是由里奇（Ricci）发现的！

[182] 尝试简单地列出六个独立的分量，然后说服自己可以使用对称性找到所有其他分量。

[183] 也称为第二比安基恒等式。 根据 Pais (1982, pp. 275 – 276)，这是首先由 Aurel Voss (1880)、Ricci (1889) 和 Bianchi (1902) 发现的。 然而，即使是 Voss 也不是第一个——见 [29.10]！

其中前三个向量循环置换，就像它们在代数比安基恒等式中一样。 这个结果的证明将被推迟，直到我们可以在第 38.12.4 节中使用曲率 2−形式优雅地推导出它。

29.5.8截面曲率

在 2-曲面中，高斯曲率 $\mathcal{K}$ 表现为向量的单位面积的和乐，该向量必须位于绕其平行移动的环路平面内。 但是在我们的 3-流形（或 n-流形）中，$\mathbf{w}_o$ 通常会伸出这个平面。 尽管如此，我们当然可以自由选择它位于 $\Pi(u,v)$ 内，以期恢复类似于二维情况的曲率概念。

然而，起初我们似乎遇到了障碍：

即使我们坚持初始向量 $\mathbf{w}_o$ 位于 $\Pi(u,v)$ 内，然后绕 Π 内一个环路平行移动返回 o 后它并不会仅在 Π 内旋转——返回的向量 $\mathbf{w}_\parallel(o)$ 通常会伸出 Π。

为了克服这个问题，让我们将注意力集中在 $\mathbf{w}_\parallel$ 向 Π 的正交投影 $\mathcal{P}[\mathbf{w}_\parallel]$ 上。更正式地说，如果 $\Pi = \Pi(\mathbf{e}_1, \mathbf{e}_2)$，那么这个正交投影算子 $\mathcal{P}$ 由下式给出，

$$\boxed{\mathcal{P}[a^1\mathbf{e}_1 + a^2\mathbf{e}_2 + a^3\mathbf{e}_3] = a^1\mathbf{e}_1 + a^2\mathbf{e}_2}$$

我们将证明以下内容：

> 令 $\mathbf{w}_o$ 为 Π 中的任意向量，令 $\mathbf{w}_\parallel$ 在 Π 中绕一个环路平行移动。将 $\mathcal{K}(u,v)$ 定义为 $\mathbf{w}_\parallel$ 到 Π 投影 $\mathcal{P}[\mathbf{w}_\parallel]$ 的单位面积旋转。 于是 $\mathcal{K}(u,v)$ 与 $\mathbf{w}_o$ 的具体选择无关。

因此曲率 $\mathcal{K}(\mathbf{u},\mathbf{v}) = \mathcal{K}(\Pi)$ 仅取决于平面 Π，称为 $\Pi(\mathbf{u},\mathbf{v})$ 的截面曲率（sectional curvature）。

为了验证这一点，我们简单地采用一般向量和乐公式（29.12）并坚持 $\mathbf{w}_o$ 位于 $\Pi(\mathbf{u},\mathbf{v}) = \Pi(\mathbf{e}_1, \mathbf{e}_2)$ 中：

$$\mathbf{w}_o = w_o^1\mathbf{e}_1 + w_o^2\mathbf{e}_2$$

为简单起见，让我们再次假设它是单位向量。将（29.12）写成列向量形式，并使用（29.10），我们看到 $\mathcal{P}[\delta\mathbf{w}_\parallel]$ 是

$$-\begin{bmatrix}\delta w_\parallel^1\\ \delta w_\parallel^2\end{bmatrix} = \begin{bmatrix}R_{12k}{}^1 w_o^k\\ R_{12k}{}^2 w_o^k\end{bmatrix}\delta\mathcal{A} = \begin{bmatrix}R_{121}{}^1 w_o^1 + R_{122}{}^1 w_o^2\\ R_{121}{}^2 w_o^1 + R_{122}{}^2 w_o^2\end{bmatrix}\delta\mathcal{A} = \begin{bmatrix}R_{1211} w_o^1 + R_{1221} w_o^2\\ R_{1212} w_o^1 + R_{1222} w_o^2\end{bmatrix}\delta\mathcal{A}$$

而由反对称（29.13）得到

$$\mathrm{R}_{1211} = 0 = \mathrm{R}_{1222} \quad 和 \quad \mathrm{R}_{1221} = 0 = -\mathrm{R}_{1212}$$

因此，我们现在将截面曲率正式定义为

$$\boxed{\mathcal{K}(\Pi) \equiv \mathcal{K}(\mathbf{e}_1, \mathbf{e}_2) \equiv [\mathcal{R}(\mathbf{e}_1, \mathbf{e}_2)\mathbf{e}_2] \cdot \mathbf{e}_1 = \mathrm{R}_{1221}} \tag{29.18}$$

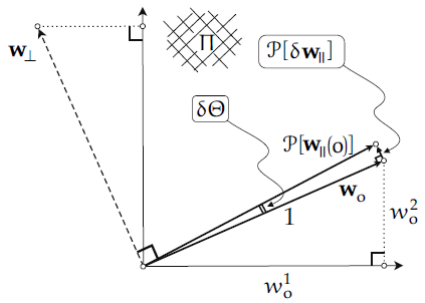
也就是说，$\mathbf{e}_2$ 绕 $\{\mathbf{e}_1, \mathbf{e}_2\}$-环路平行移动之后的和乐，投影到 $\mathbf{e}_1$ 上。于是

$$\boxed{\mathcal{P}[\delta\mathbf{w}_\parallel] = \begin{bmatrix} \delta \mathrm{w}_\parallel^1 \\ \delta \mathrm{w}_\parallel^2 \end{bmatrix} \asymp \begin{bmatrix} -\mathrm{w}_\mathrm{o}^2 \\ \mathrm{w}_\mathrm{o}^1 \end{bmatrix} \mathcal{K}(\Pi)\delta\mathcal{A} = \mathbf{w}_\perp \mathcal{K}(\Pi)\delta\mathcal{A}} \tag{29.19}$$

其中 $\mathbf{w}_\perp$ 只是 $\mathbf{w}_\mathrm{o}$ 在 Π 内旋转一个直角得到的，如 [29.9] 所示。

因此，如图所示，投影 $\mathcal{P}[\mathbf{w}_\parallel]$ 通过绕面积 $\delta\mathcal{A}$ 的环路平行移动所经过的旋转角度 $\delta\Theta$ 有 $\delta\Theta \asymp \mathcal{K}(\Pi)\,\delta\mathcal{A}$，与 $\mathbf{w}_\mathrm{o}$ 无关。 我们已经回到原来的曲率观点，即 2-曲面内每单位面积的和乐：

$$\boxed{\frac{\delta\Theta}{\delta\mathcal{A}} \asymp \mathcal{K}(\Pi) = \mathrm{R}_{1221}} \tag{29.20}$$



[29.9] 如果 $\mathbf{w}_\mathrm{o}$ 位于环路的平面 Π 内，则平行移动向量返回时到 Π 的投影旋转了 $\delta\Theta \asymp |\delta\mathcal{P}[\mathbf{w}_\parallel]| \asymp \mathcal{K}(\Pi)\,\delta\mathcal{A}$，其中 $\mathcal{K}(\Pi)$ 是 Π 的截面曲率。

只是现在有许多这样的截面曲率，这取决于我们在哪个平面 Π 内平行移动。

正如我们前面提到的，例如，仅知道正交坐标平面中的截面曲率只能给出空间黎曼曲率的非常不完整的[184]图像。但是，如果我们可以知道所有可能平面 Π 的 $\mathcal{K}(\Pi)$ 会怎么样？

事实证明，这足以重构完整的黎曼张量！ 很难在标准课本中找到这个讨论，但库内尔（Kühnel，2015, p. 247）给出了 $[\boldsymbol{\mathcal{R}}(\mathbf{u}, \mathbf{v})\,\mathbf{x}] \cdot \mathbf{y}$ 的明确公式，表示为由黎曼曲率的四个输入向量的线性组合所构造的 18 个平面的 18 个截面曲率之和！

[184] 在本章的最后一节中，我们将看到（令人惊讶！）这些截面曲率的总和仍然包含关键的几何信息！

29.5.9黎曼张量起源的历史笔记

黎曼于 1854 年 6 月 10 日在哥廷根大学的一次演讲中宣布了他的张量发现，尽管是间接的——这是他申请成为一名无薪教员的强制性部分！——主要面向非数学的听众。

讲座做出神谕（没有提供一丝证据），它只包含一个公式：恒定曲率 n-流形的度量。 但听众中肯定有一位数学家领悟了他的意思，首先迫使黎曼进行这项研究的那位：卡尔·弗里德里希·高斯！

按照规则的要求，黎曼为必修课提交了三个可能的主题，其中第一个是他最近完成的关于现在称为**黎曼曲面**的开创性博士工作。 惯例要求高斯选择这个，黎曼已经准备好解决这个问题。 高斯没有选！ 相反，他选择了黎曼的第三个备选主题："几何学基础"（译者注：《论作为几何学基础的假设》）。

因此，黎曼在他的博士杰作之后，被迫进行了一项全新的重大研究。 这项工作的压力，也许因生活贫困而加剧，导致黎曼再次遭受他反复出现的精神崩溃。 但他康复了，经过七个星期的努力，他准备好在 6 月 10 日发表演讲。

高斯是否知道或关心黎曼在演讲前的困境则不为人知； 众所周知的是，高斯对黎曼的演讲的反应。 戴德金（Dedekind）后来回忆说，高斯列席了讲座"黎曼提出想法的深度，超出了他所有的期望，非常惊讶，在从教员会议回来的一路上，他带着最大的赞赏和对他来说罕见的兴奋与威廉·韦伯（Wilhelm Weber）讨论着（这个想法）。"

1861 年，黎曼向法国科学院提交了一篇评奖论文，其中包括对黎曼张量的更明确的数学描述。 但是这篇论文——没有获奖！——主要是关于热传导的，关于它对几何学的意义黎曼只给出了最简短的暗示。

直到 1866 年黎曼去世（享年 39 岁，死于肺结核），这篇文章和 1854 年的原始演讲都没有发表。菲利克斯·克莱因（Felix Klein (1928)）总结了对 1868 年他们最终发表的反应："黎曼的思路给年轻数学家留下的非凡印象至今仍历历在目。 其中大部分似乎晦涩难懂，难以理解，且深不可测。"

现在我们面临着第四幕开始时同样的谜团，那里是关于爱因斯坦，但现在是关于黎曼：列维-奇维塔在黎曼演讲整整 63 年后发现了平行移动，所以我们提供的黎曼张量的现代几何解释对黎曼本人来说是遥不可及的。 那么黎曼是如何发现他的张量的呢？

答案是没有人知道！斯皮瓦克（Spivak (1999) ）提供了 1854 年演讲的英文翻译，以及斯皮瓦克推测黎曼可能是如何做到这一点的解释性说明。 斯皮瓦克还包括对黎曼 1861 年评奖论文的分析。

从黎曼自己的话可以清楚地看出，他对张量的最初解释是，它测量了 n−流形度量与欧氏度量的偏差。 更准确地说，黎曼在他的张量和度量之间发现了以下非常直接的联系。 如果 ds 是笛卡尔坐标为 x_i 的点 p 与坐标为 $(x_i + dx_i)$ 的相邻点之间的距离，则实际度量与欧氏度量之间的差异由以下非凡的公式给出：

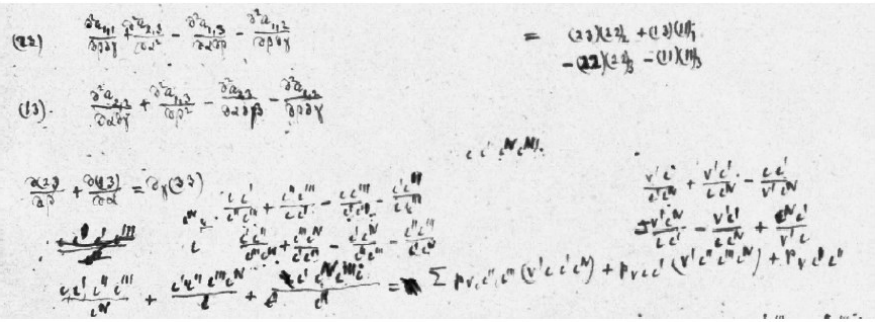
$$ds^2 - [dx_1^2 + \cdots + dx_n^2] \asymp \frac{1}{12} \sum_{i,j,k,m} R_{ijkm}(p)(x_i\, dx_j - x_j\, dx_i)(x_k\, dx_m - x_m\, dx_k)$$

有关这方面的更多信息，请参阅 Spivak (1999) 和 Berger (2003, §4.4)。

《黎曼曲率之谜》（The Mystery of Riemann's Curvature，Darrigol 2015）对这些问题做了引人入胜且发人深省的研究。达里戈尔（Darrigol）的分析部分基于黎曼之前未发表的私人笔记的照片，这些照片已在哥廷根大学的档案中保存了 150 多年。

[29.10] 中显示了一张这样的纸片。达里戈尔对黎曼符号的艰苦解码使他能够在数学史的编年史中进行一次引人入胜的更正。 对于那长长的、难以理解的最后一行，结果证明是微分比安基恒等式 (29.17)，几十年后它会被重新发现——首先[185]是由奥雷尔·沃斯 (Aurel Voss)（1880 年），然后是里奇（1889 年），然后是比安基（1902 年） ——成为爱因斯坦广义相对论的基石[186]之一！

有关黎曼的生平和科学的更多信息，我们推荐 Monastyrskiĭ (1999) 和 Laugwitz (1999)。



[29.10] 黎曼的草稿纸，由哥廷根大学档案馆提供。 最后一行是微分比安基恒等式（29.17），比其他人早了几十年！

29.6 n-流形中的雅克比方程

29.6.1截面雅克比方程的几何证明

在 2-曲面内，雅克比方程（测地线偏差方程）（28.2）分别描述了通过正（或负）曲率区域的相邻测地线的吸引力（或排斥力）。

在 2-曲面中，这种吸引力或排斥仅取决于测地线通过点处的曲面曲率：它不取决于测地线通过该点的方向。

---

[185] 参见 Pais (1982, pp. 275–276)。

[186] 见练习 14。

在 3-流形内，吸引力或排斥力再次仅取决于相邻测地线在其中行进的（瞬时）平面，而不取决于该平面内测地线的方向。该平面 $\Pi = \Pi(\mathbf{v}, \boldsymbol{\xi})$ 由速度向量 $\mathbf{v}$ 和连接向量 $\boldsymbol{\xi}$ 张成。但是，虽然对于 2-曲面的平面只有一种可能的“选择”，但现在在我们的 3-流形（或 n-流形）内有无限多个这样的平面，并且如 [29.2] 中已经说明的那样，这些平面中的截面曲率在一个平面上可以是正的，在另一个平面上可以是负的，在一个平面上引起吸引，在另一个平面上引起排斥。

因此很自然地猜测雅可比方程 (28.2) 到 n-流形的推广只是简单地将高斯曲率 $\mathcal{K}$ 替换为两个相邻测地线在其中行进的瞬时平面的*截面曲率* $\mathcal{K}(\Pi)$，并且我们将给出两个证明，但有一个转折：*我们必须关注位于 Π 内的相对加速度的分量*。

在我们使用牛顿上标点来表示沿测地线的导数之前，我们现在使用更标准的 $\boldsymbol{\nabla}_{\mathbf{v}}$。因此，相对速度现在写成 $\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}$，相对加速度写成 $\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}$。

重新考虑我们在 [28.7] 中对雅克比方程的原始证明，但现在假设我们平行移动速度所围绕的矩形位于 3-流形（或 n-流形）中的任意平面 Π 内。和以前一样，平行移动 $\mathbf{v}_{\parallel}$ 围绕具有正交边 $\boldsymbol{\xi}$ 和 $\mathbf{v}\,\delta t$ 的环路，产生

$$\delta\mathbf{v}_{\parallel} \asymp \delta t\,\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}$$

由于粒子的速度是恒定的，$\delta\mathbf{v}_{\parallel}$ 与 $\mathbf{v}$ 正交。在 2-曲面 [28.7] 中，这意味着它指向 $\pm\boldsymbol{\xi}$，但在 3-流形中*这不再正确*：它也可以具有与 Π 正交的分量[187]。总之，相对加速度 $\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}$ 在 Π 内有一个分量，负责将测地线拉在一起（或将它们推开），以及一个与 Π 正交的分量，它不影响测地线的间隔，但会导致它们围绕彼此旋转。

我们现在将注意力集中在 Π 内的吸引/排斥分量上，即 $\mathcal{P}[\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}]$。正如我们在（29.20）中看到的，如果我们将平行移动的向量投影到 Π 中，那么和乐再次由曲率乘以矩形面积给出，只是现在它是截面曲率 $\mathcal{K}(\Pi)$。

因此，我们在 [28.7] 中的原始论证基本上没有改变，除了现在（29.19）仅适用于 $\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}$ 到 Π 的投影，产生，

$$\delta t\,\mathcal{P}[\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}] \asymp \mathcal{P}[\delta\mathbf{v}_{\parallel}] \asymp \mathbf{v}_{\perp}\mathcal{K}(\Pi)\delta\mathcal{A} = \left[-\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\right]\mathcal{K}(\Pi)\,|\boldsymbol{\xi}|\,\delta t$$

所以我们得到了我们将命名[188]为

截面雅可比方程：

$$\boxed{\mathcal{P}[\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}] = -\mathcal{K}(\Pi)\,\boldsymbol{\xi}} \tag{29.21}$$

[187] 在 n-流形中，将有 (n-2) 个与 Π 正交的分量。

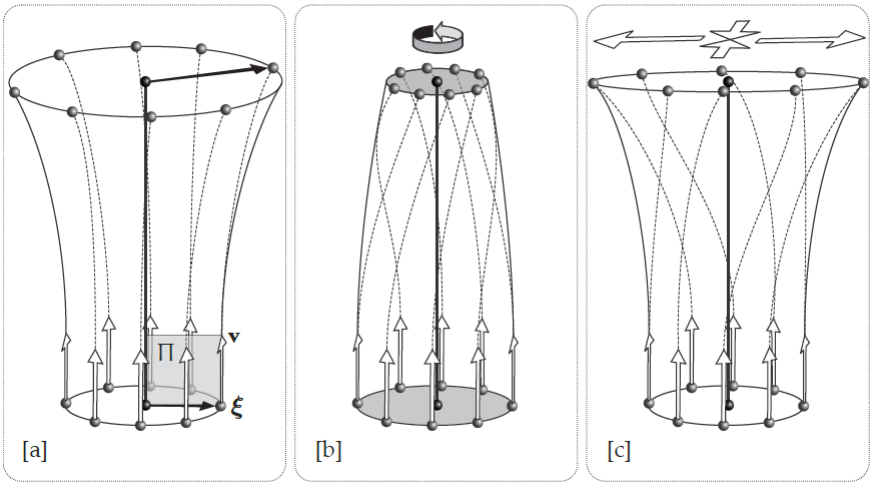
[188] 尽管它有两个优点——(1) 提供了对截面曲率的非常直接的解释，(2) 在形式上与 2-曲面中的完整雅可比方程 (28.2) 相同——我们还没有在任何标准课本中发现这个结果。因此，我们觉得有必要为这个公式发明一个合乎逻辑的名称。

注意：这不是标准的“雅可比方程”，它很快就会得到（见（29.24））。

29.6.2截面雅克比方程的几何含义

为了直观地掌握这个方程的重要性，请考虑[29.11]。 一组粒子围绕平面中的一个小圆圈的周边排列。 在圆的中心添加一个基准粒子，向量 $\boldsymbol{\xi}$ 将这个中心粒子连接到圆上的典型粒子。

现在同时以速度 $\mathbf{v}$ 垂直于平面发射这些粒子。最初，粒子将以刚性一致的方式移动，保持相同大小的圆形图案。 但随后空间的曲率开始显现，通过截面雅可比方程 (29.21)。 圆上每个单个粒子相对于中心粒子的运动现在由该特定粒子 $\boldsymbol{\xi}$ 的特定平面 $\Pi(\mathbf{v},\boldsymbol{\xi})$ 的特定 $\mathcal{K}(\Pi)$ 控制。



[29.11] 一小圈粒子垂直于该圆圈发射。 一些平面 $\Pi(\mathbf{v},\boldsymbol{\xi})$ 可能具有 $\mathcal{K}(\Pi)>0$，导致朝向中心粒子的吸引力，而其他平面可能具有 $\mathcal{K}(\Pi)<0$，导致与中心粒子的排斥。 [a] 所有平面中的 $\mathcal{K}(\Pi)<0$。 [b] $\mathcal{K}(\Pi)>0$ 在所有平面上。 [c] 某些平面的引力挤压圆，而另一些平面的斥力将圆拉长，从而使圆变形为椭圆。

因此，当我们从 2-曲面移动到 3-流形（或 n-流形）时，出现了一个全新的特征：当平面 $\Pi(\mathbf{v},\boldsymbol{\xi})$ 围绕中心轴（由中心粒子的轨迹组成）旋转时， 截面曲率 $\mathcal{K}(\Pi)$ 在大小和（可能）符号上都会发生变化。

图 [29.11a] 说明了所有截面曲率 $\mathcal{K}(\mathbf{v},\boldsymbol{\xi})$ 为负的情况，无论 $\boldsymbol{\xi}$ 的方向如何，导致粒子被排斥，在一个扩大的圆上加速分开。[189] 使用光学类比，负曲率就像一个发散透镜。

接下来，[29.11b] 说明了所有曲率都为正的情况。 除了由（29.21）控制的聚焦吸引力

[189]如果负曲率的大小发生变化，那么圆会随着它的增长而变成椭圆形。

之外，这里我们还说明了一种情况，其中相对加速度具有与 Π 正交的可感知的分量，导致测地线束围绕中心线旋转。 在我们的截面雅可比方程（29.21）中有意忽略了这种影响，但（即将推出的）完整雅可比方程（29.24）将其考虑在内。 这类似于通过放大镜对光进行正聚焦。

最后，[29.11c] 描述了最有趣的情况，正如我们将在下一章中看到的，这是对引力物理学最重要的情况。 在这里，截面曲率在某些平面上是正的（将粒子圈挤压在一起），而在其他平面上是负的（将圆拉开）。因此，最初的圆形粒子阵列变形为椭圆，其短轴小于原始圆，长轴大于它。 这类似于散光镜片的效果。

在下一章中，我们将看到光学类比根本不是类比——它是现实！ 也就是说，当光穿过它时，时空的引力曲率确实像一个透镜！

29.6.3雅克比方程和截面雅克比方程的计算证明

对于 Π 的正交基，我们取 $\mathbf{e}_1 = \hat{\boldsymbol{\xi}} = \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}$ 和 $\mathbf{e}_2 = \mathbf{v}$。 那么 $\mathbf{e}_3 \equiv \mathbf{e}_1 \times \mathbf{e}_2$ 将是 Π 的单位法线。 现在可以写出截面曲率的定义（29.18），

$$\mathcal{K}(\Pi) = \mathrm{R}_{1221} = [\mathcal{R}(\mathbf{e}_1, \mathbf{e}_2)\, \mathbf{e}_2] \cdot \mathbf{e}_1 = \left[\mathcal{R}\left(\hat{\boldsymbol{\xi}}, \mathbf{v}\right) \mathbf{v}\right] \cdot \hat{\boldsymbol{\xi}} \tag{29.22}$$

根据定义，$\boldsymbol{\xi}$ 总是连接两个相邻的测地线，因此具有边 $\mathbf{v}$ 和 $\boldsymbol{\xi}$ 的平行四边形闭合：

$$[\mathbf{v}, \boldsymbol{\xi}] = 0 \quad \Leftrightarrow \quad \boldsymbol{\nabla}_{\mathbf{v}} \boldsymbol{\xi} = \boldsymbol{\nabla}_{\boldsymbol{\xi}} \mathbf{v} \tag{29.23}$$

注意（稍后使用）这意味着黎曼曲率算子（29.6）简化为

$$\mathcal{R}(\mathbf{v}, \boldsymbol{\xi}) = [\boldsymbol{\nabla}_{\mathbf{v}}, \boldsymbol{\nabla}_{\boldsymbol{\xi}}]$$

沿测地线取（29.23）的内蕴导数 $(\boldsymbol{\nabla}_{\mathbf{v}})$，回想 $\mathbf{v}$ 满足测地线方程，$\boldsymbol{\nabla}_{\mathbf{v}} \mathbf{v} = 0$，我们发现

$$\begin{aligned} \boldsymbol{\nabla}_{\mathbf{v}}\, \boldsymbol{\nabla}_{\mathbf{v}}\, \boldsymbol{\xi} &= \boldsymbol{\nabla}_{\mathbf{v}}\, \boldsymbol{\nabla}_{\boldsymbol{\xi}} \mathbf{v} \\ &= \left[\boldsymbol{\nabla}_{\mathbf{v}}, \boldsymbol{\nabla}_{\boldsymbol{\xi}}\right] \mathbf{v} + \boldsymbol{\nabla}_{\boldsymbol{\xi}} (\boldsymbol{\nabla}_{\mathbf{v}} \mathbf{v}) \\ &= \mathcal{R}(\mathbf{v}, \boldsymbol{\xi}) \mathbf{v} + \boldsymbol{\nabla}_{\boldsymbol{\xi}}(0) \end{aligned}$$

因此，我们得到了***测地线偏差方程***，或

$$\text{雅克比方程：} \qquad \boxed{\boldsymbol{\nabla}_{\mathbf{v}}\, \boldsymbol{\nabla}_{\mathbf{v}}\, \boldsymbol{\xi} = -\mathcal{R}(\boldsymbol{\xi}, \mathbf{v}) \mathbf{v}} \tag{29.24}$$

这是雅可比方程的标准形式，可在所有课本中找到。

为了恢复我们更直观的“截面”形式，我们将等式右侧的向量分解为其分量，并发现，

$$\begin{aligned}\mathcal{R}(\boldsymbol{\xi},\mathbf{v})\mathbf{v} &= \left\{[\mathcal{R}(\boldsymbol{\xi},\mathbf{v})\mathbf{v}]\cdot\hat{\boldsymbol{\xi}}\right\}\hat{\boldsymbol{\xi}} + \{[\mathcal{R}(\boldsymbol{\xi},\mathbf{v})\mathbf{v}]\cdot\mathbf{v}\}\,\mathbf{v} + \left\{\left[\mathcal{R}\left(|\boldsymbol{\xi}|\hat{\boldsymbol{\xi}},\mathbf{v}\right)\mathbf{v}\right]\cdot\mathbf{e}_3\right\}\mathbf{e}_3 \\ &= \left\{\left[\mathcal{R}\left(\hat{\boldsymbol{\xi}},\mathbf{v}\right)\mathbf{v}\right]\cdot\hat{\boldsymbol{\xi}}\right\}\boldsymbol{\xi} + 0 + \{[\mathcal{R}(\mathbf{e}_1,\mathbf{e}_2)\mathbf{e}_2]\cdot\mathbf{e}_3\}\,|\boldsymbol{\xi}|\mathbf{e}_3 \\ &= \mathcal{K}(\Pi)\,\boldsymbol{\xi} + \mathrm{R}_{1223}|\boldsymbol{\xi}|\mathbf{e}_3\end{aligned}$$

根据 （29.22）。

第二项 $\mathrm{R}_{1223}|\boldsymbol{\xi}|\mathbf{e}_3$ 是与 Π 正交的相对加速度的分量，它负责测地线束的旋转，如 [29.11b] 所示。

第一项表示 Π 内的吸引或排斥。实际上，将投影算子 $\mathcal{P}$ 应用于雅克比方程（29.24）的两侧，我们恢复了截面雅克比方程（29.21）：

$$\boxed{\mathcal{P}[\boldsymbol{\nabla}_{\mathbf{v}}\boldsymbol{\nabla}_{\mathbf{v}}\,\boldsymbol{\xi}] = -\mathcal{P}[\mathcal{R}(\boldsymbol{\xi},\mathbf{v})\mathbf{v}] = -\mathcal{K}(\Pi)\,\boldsymbol{\xi}}$$

29.7 里奇张量

29.7.1被测地线束包围面积的加速度

重新考虑[29.11b]。 当粒子被正截面曲率拉在一起时，圆盘的阴影区域 $\delta\mathcal{A}$（在粒子圆圈内）显然会收缩。 但根据什么规律？

为简单起见，我们先假设当 Π 围绕中心粒子的轨迹旋转时，所有截面曲率具有相同的正值 $\mathcal{K}$。现在令 $\mathrm{r(t)} = |\boldsymbol{\xi}(\mathrm{t})|$，使得 $\delta\mathcal{A}(\mathrm{t}) = \pi[\mathrm{r(t)}]^2$ 是发射后 t 时刻的面积。 回到牛顿（上标）点的时间导数，很明显，面积的初始变化率趋于零，因为粒子以刚性一致的方式前进，所以 $\dot{\mathrm{r}}(0) = 0$，因此

$$(\dot{\delta\mathcal{A}})(0) = 2\pi\mathrm{r}(0)\dot{\mathrm{r}}(0) = 0$$

截面雅可比方程（29.21）告诉我们粒子向中心的加速度是 $\ddot{\mathrm{r}}\,(0) = -\mathcal{K}\mathrm{r}(0)$。 因此曲率表现在该面积的初始加速度中：

$$(\ddot{\delta\mathcal{A}})(0) = 2\pi[\dot{\mathrm{r}}^2(0) + \mathrm{r}(0)\ddot{\mathrm{r}}(0)] = -2\pi\mathcal{K}[\mathrm{r}(0)]^2 = -2\mathcal{K}\delta\mathcal{A}(0)$$

因此，对于较小的 t，麦克劳林展开表明面积与曲率成正比并随着时间的平方而减小：

$$\text{时间 t 后面积的变化} = \delta\mathcal{A}(\mathrm{t}) - \delta\mathcal{A}(0) \asymp -\mathcal{K}\delta\mathcal{A}(0)\mathrm{t}^2$$

当然，如果 $\mathcal{K}$ 为负值，如 [29.11a] 中所示，那么右侧将为正值，相应地加速了该面积

的增长。

既然我们对正在发生的事情有了一个大致的了解，让我们继续讨论一般情况，其中 $\mathcal{K}(\Pi)$ 随着 Π 围绕中心轴旋转而在大小上（可能还有符号）发生变化。

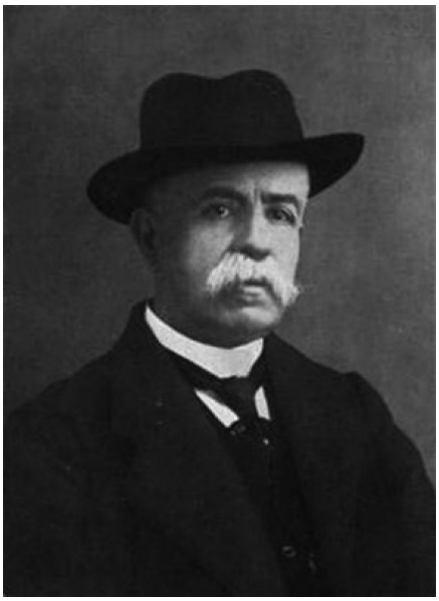
让 θ 表示围绕该轴的角度，从任意初始 $\boldsymbol{\xi}$ 方向开始。 考虑位于角度 θ 和角宽度 $\delta\theta$ 的圆盘的狭窄（最终趋于零）扇形区域。 如果我们用在该方向上的特定平面 $\Pi(\theta)$ 的截面曲率 $\mathcal{K}(\theta)$ 代替 $\mathcal{K}$，上述分析仍然适用于该扇区。

因此，在一般情况下，前面的方程变为

$$\delta\mathcal{A}(\mathrm{t}) - \delta\mathcal{A}(0) \simeq -\mathcal{K}_{\mathrm{mean}}\mathcal{A}(0)\,\mathrm{t}^2$$

其中

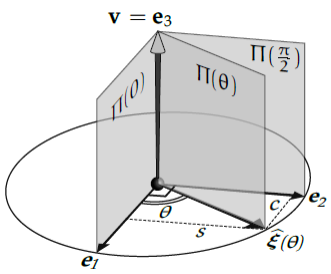
$$\mathcal{K}_{\mathrm{mean}} \equiv \frac{1}{2\pi}\int_0^{2\pi}\mathcal{K}(\theta)\mathrm{d}\theta \tag{29.25}$$



[29.12] 格雷戈里奥・里奇-库巴斯特罗 (1853 – 1925),
后来被简单地称为格雷戈里奥・里奇（发音为“reechee”）。

这里，$\mathcal{K}_{\mathrm{mean}}$ 是包含中心粒子轨迹的所有平面中在方向 $\mathbf{v}$ 上的截面曲率的平均值。

接下来，我们将根据黎曼张量找到 $\mathcal{K}(\theta)$ 的公式。 为此，让我们选择 $\mathbf{e}_3 = \mathbf{v}$ 作为中心轴的方向，即粒子发射的方向。 见[29.13]。 那么 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 是粒子圆平面上的正交单位向量。 任意选择 $\mathbf{e}_1$ 的方向，从这个方向测量 θ，如图所示。



[29.13]

平面 $\Pi(\theta)$ 由 $\mathbf{v}$ 和单位向量 $\hat{\boldsymbol{\xi}}(\theta)$ （在角度 θ 的粒子方向）张成：

$$\hat{\boldsymbol{\xi}}(\theta) = \cos\theta\,\mathbf{e}_1 + \sin\theta\,\mathbf{e}_2 \equiv c\,\mathbf{e}_1 + s\,\mathbf{e}_2$$

简而言之。$\Pi(\theta)$ 的截面曲率由（29.22）给出：

$$\begin{aligned}
\mathcal{K}(\theta) &= \left[\mathbf{R}\left(\hat{\boldsymbol{\xi}}, \mathbf{v}; \mathbf{v}\right)\right] \cdot \hat{\boldsymbol{\xi}} \\
&= \left[\mathbf{R}\left((c\,\mathbf{e}_1 + s\,\mathbf{e}_2), \mathbf{v}; \mathbf{v}\right)\right] \cdot (c\,\mathbf{e}_1 + s\,\mathbf{e}_2) \\
&= [c\,\mathbf{R}(\mathbf{e}_1, \mathbf{v}; \mathbf{v}) + s\,\mathbf{R}(\mathbf{e}_2, \mathbf{v}; \mathbf{v})] \cdot (c\,\mathbf{e}_1 + s\,\mathbf{e}_2) \\
&= c^2\,\mathbf{R}(\mathbf{e}_1, \mathbf{v}; \mathbf{v}) \cdot \mathbf{e}_1 + s^2\,\mathbf{R}(\mathbf{e}_2, \mathbf{v}; \mathbf{v}) \cdot \mathbf{e}_2 + sc[\mathbf{R}(\mathbf{e}_1, \mathbf{v}; \mathbf{v}) \cdot \mathbf{e}_2 + \mathbf{R}(\mathbf{e}_2, \mathbf{v}; \mathbf{v}) \cdot \mathbf{e}_1]
\end{aligned}$$

其中最后一项来自对称性 $R_{2331} = R_{1332}$ [练习]。

要计算 $\mathcal{K}_{\text{mean}}$，首先验证

$$\int_0^{2\pi} \cos^2\theta\,\mathrm{d}\theta = \int_0^{2\pi} \sin^2\theta\,\mathrm{d}\theta = \pi, \qquad \int_0^{2\pi} \sin\theta\cos\theta\,\mathrm{d}\theta = 0$$

然后，

$$\boxed{\mathcal{K}_{\text{mean}} = \frac{\mathcal{K}(0) + \mathcal{K}(\frac{\pi}{2})}{2}} \tag{29.26}$$

回想一下，方向 $\theta = 0$ 是任意选择的，因此我们已经证明了一个显著的事实，即所有方向上的截面曲率的平均值是通过在任何两个正交方向上对其进行平均来获得的。

结合（29.25）和（29.26），我们可以将我们的发现总结如下：

> 如果一小圈粒子以垂直于圆的速度 $\mathbf{v}$ 发射，则圆的面积 $\delta\mathcal{A}$ 的加速度由包含 $\mathbf{v}$ 的所有平面的截面曲率的平均值控制，而这是由包含 $\mathbf{v}$ 的任意两个正交平面上的截面曲率之和给出的：
>
> $$(\delta\ddot{\mathcal{A}})(0) = -\left[\mathcal{K}(0) + \mathcal{K}\left(\frac{\pi}{2}\right)\right]\delta\mathcal{A}(0)$$
>
> 所以

(29.27)

29.7.2里奇张量的定义和几何意义

让我们将注意力集中在包含粒子速度 $\mathbf{v}=\mathbf{e}_3$ 的正交坐标平面中的截面曲率总和上。回想 $\mathrm{R}_{3333}=0$，根据反对称性，我们可以写成

$$\begin{aligned}\mathcal{K}(0)+\mathcal{K}\left(\frac{\pi}{2}\right) &= \mathbf{R}(\mathbf{e}_1,\mathbf{e}_3;\mathbf{e}_3)\cdot\mathbf{e}_1+\mathbf{R}(\mathbf{e}_2,\mathbf{e}_3;\mathbf{e}_3)\cdot\mathbf{e}_2\\ &= \mathrm{R}_{1331}+\mathrm{R}_{2332}+\mathrm{R}_{3333}\\ &= \mathrm{R}_{\mathrm{m}33}{}^{\mathrm{m}}\end{aligned}$$

稍后，这个表达式将引导我们引入一个全新的张量，里奇（Ricci）曲率张量。 黎曼张量将三个向量作为输入，并输出一个向量，而新的里奇张量将两个向量作为输入，并输出一个标量。

注意：黎曼（Riemann）和里奇（Ricci）的名字都以字母“R”开头，这是一个不幸的历史事故！ 这导致了普遍接受的符号，其中里奇张量的分量表示为 R_{jk}。我们不敢违背这一传统——甚至不希望！——但请高度警惕，具有四个索引的 R 是 黎曼张量的一个分量，而具有两个索引的 R 是里奇张量的一个分量。也就是说，用相同的符号 **R** 来表示这两个（非常不同的）几何张量实在是太令人困惑了！

因此，我们将采用（基本上）Misner、Thorne 和 Wheeler (1973) 使用的符号，将 **R** 保留为黎曼张量，并将几何里奇张量写为本名 Ricci。

里奇曲率张量的定义是这样的：

$$\boxed{\mathbf{R}\mathrm{icci}(\mathbf{v},\mathbf{w})\equiv\sum_{m=1}^{n}\mathbf{R}(\mathbf{e}_{\mathrm{m}},\mathbf{v};\mathbf{w})\cdot\mathbf{e}_{\mathrm{m}}\quad\Leftrightarrow\quad\mathrm{R}_{\mathrm{jk}}=\mathbf{R}\mathrm{icci}\left(\mathbf{e}_{\mathrm{j}},\mathbf{e}_{\mathrm{k}}\right)=\mathrm{R}_{\mathrm{mjk}}{}^{\mathrm{m}}}$$

约定警告：许多作者改为定义 $\mathrm{R}_{\mathrm{jk}}\equiv\mathrm{R}_{\mathrm{jmk}}{}^{\mathrm{m}}=-\mathrm{R}_{\mathrm{mjk}}{}^{\mathrm{m}}$，即我们的负数。 值得注意的是，Misner、Thorne 和 Wheeler (1973) 也是如此。 然而，正如我们之前提到的，他们对 **R**

的定义也与我们的相反，因此两个符号约定的分歧抵消：他们的 Ricci 的符号与我们的 Ricci 相同！正如我们稍后将看到的，正 Ricci 曲率对应于吸引力（对我们和 Misner、Thorne 和 Wheeler 都是如此）。另一方面，仅举一个重要的例子，Penrose (2005) 分享了我们对 **R** 的定义，但使用了相反的 Ricci 定义，因此对于 Penrose 来说，引起吸引力的是负 Ricci 曲率。

R 是一个张量这一事实立即可得 Ricci 本身确实是一个张量[练习]。从 **R** 的对称性也可以看出，Ricci 是对称的：

$$\boxed{\mathbf{R}\mathrm{icci}(\mathbf{w},\mathbf{v}) = \mathbf{R}\mathrm{icci}(\mathbf{v},\mathbf{w}) \quad \Leftrightarrow \quad \mathrm{R}_{\mathrm{kj}} = \mathrm{R}_{\mathrm{jk}}} \tag{29.28}$$

要看到这一点，请将其黎曼张量定义中的第一对向量与第二对交换，然后交换每对中的向量。

我们现在可以看到

$$\mathcal{K}(0) + \mathcal{K}\left(\frac{\pi}{2}\right) = \mathrm{R}_{\mathrm{m33}}{}^{\mathrm{m}} = \mathbf{R}\mathrm{icci}(\mathbf{v},\mathbf{v}) = \mathrm{R}_{\mathrm{jk}}\mathrm{v}^{\mathrm{j}}\mathrm{v}^{\mathrm{k}}$$

因此，我们得到了支配以速度 **v** 发射的测地线束面积加速度的非常简单的雅可比型方程：

$$\boxed{\delta\ddot{\mathcal{A}} = -\mathbf{R}\mathrm{icci}(\mathbf{v},\mathbf{v})\,\delta\mathcal{A}}$$

2-曲面中高斯曲率在雅克比方程中的作用（或 n-流形中的截面曲率）在这里由 Ricci 曲率取代。我们强调正 Ricci 曲率会引起吸引力，导致测地线收敛时面积缩小。

现在想象将上述分析推广到 4-流形。再次令 **v** 表示我们将发射一组粒子的速度向量。该组再次被选择为与中心粒子等距，并位于与 **v** 正交的空间中。但这种构造现在不是在平面上产生一个圆，而是产生一个粒子构成的小球面，包围体积 $\delta\mathcal{V}$，在 3 维空间垂直于 **v**!

虽然我们将不计算细节，但上述分析不变，具体而言，里奇曲率现在决定了以速度 **v** 同时发射的小粒子球面包围体积的加速度：

$$\boxed{\delta\ddot{\mathcal{V}} = -\mathbf{R}\mathrm{icci}(\mathbf{v},\mathbf{v})\,\delta\mathcal{V}} \tag{29.29}$$

为了将来使用我们强调，正 Ricci 曲率会引起吸引力，随着测地线的收敛导致体积缩小。

和以前一样，这意味着（最初）体积与 Ricci 曲率成比例变化，并且作为时间的平方：

$$\boxed{\delta\mathcal{V}(\mathrm{t}) - \delta\mathcal{V}(0) \asymp -\frac{1}{2}\,\mathbf{R}\mathrm{icci}(\mathbf{v},\mathbf{v})\,\delta\mathcal{V}(0)\,\mathrm{t}^2}$$

(29.30)

29.8 尾声

本章介绍了许多新的想法和结果，但有一个公式因它的缺失而显著：即用度量来表示 n-流形的黎曼曲率的（27.1）的类似物。 这种遗漏的主要原因是严格专注于给出理解爱因斯坦引力理论所需的最少概念和结果。 然而，第二个原因是，根据度量系数的 **R** 公式非常复杂，至少在以标准张量形式表示时是这样。 在第五幕中，我们将看到如何通过使用形式（Forms）在理论和实用计算层面上大大简化 **R** 的计算。

当你开始本章时，你可能已经被空间可以弯曲的想法先发制人地弄糊涂了——谁又能怪你呢！ 但是我们非常希望我们给出的多种几何解释（尤其是通过截面雅可比方程和里奇张量）不仅让你对 3-流形内的黎曼曲率感到轻松，而且让你对这个曲率的作用有了一个非常具体、切实的把握。

但 3-流形只是热身。 因为你正坐在一个三维房间里是一种心理错觉。 在实在中，你和你的房间正在爱因斯坦的 4 维时空中飞向未来。 我们现在转向那个流形（我们称之为“实在”，英文 reality）的曲率

好好研究公式（29.29），因为我们会看到它是宇宙的钥匙！

第三十章 爱因斯坦的弯曲时空

30.1 引言："我一生中最幸福的想法"

本章降下了第四幕的帷幕——这是我们对（可视化）微分几何的独立介绍的结束。 我们通过回到第四幕开始的主题来纪念这一里程碑：爱因斯坦极其美妙的广义相对论，它宣称引力是 4 维时空的曲率，它构成了我们所谓 "实在"（reality）的框架。

伟大的美国物理学家约翰·阿奇博尔德·惠勒（John Archibald Wheeler，1911-2008）将这一理论提炼成一句话："空间告诉物质如何运动，物质告诉空间如何弯曲。" 更详细地说，自由落体的物质沿着弯曲时空的测地线运动：这些仍然是最直的路径，但它们现在最大化（而不是最小化）由时空度量测量的"距离"。 惠勒格言的后半部分是第四幕的压轴戏：爱因斯坦在 1915 年发现的精确定律，它描述了物质和能量如何弯曲时空——广义相对论的引力场方程。

我们从历史的开端开始，从牛顿发现平方反比万有引力定律的（可能过于简单化[1]）的故事开始，我们可以想象他在 1666 年坐在他的花园里，看着一个苹果从树上掉下来。 突然之间，故事如此发展，牛顿意识到将苹果拉向地球的同样的力可能会伸向月球，将其拉向地球，使其保持在轨道上。

爱因斯坦的几何引力理论的起源故事非常相似，但爱因斯坦坠落的物体不是苹果，而是一个人！

1907 年，在他发现狭义相对论整整两年后，爱因斯坦仍然没有担任学术职务。 过去五年，他在瑞士伯尔尼联邦知识产权局担任专利文员（三等，晋升二等）。

正如我们在第四幕开始时所指出的，爱因斯坦知道，尽管牛顿的万有引力定律非常精确，但它的瞬时远距离作用从根本上与有限的光速不相容，因此与他在 1905 年发现的狭义相对论是不相容的。爱因斯坦为理解万有引力而进行的长期斗争就这样开始了，这一努力在几乎整整八年后结束，即 1915 年 11 月 25 日，他终于写下了引力场方程。

爱因斯坦在 1907 年 11 月迈出了漫长旅程的第一步，也是至关重要的一步，当时他有了后来所说的"我一生中最幸福的想法"[2]：

> 我坐在伯尔尼专利局的椅子上，突然想到：如果一个人自由落下，他就不会感觉到自己的重量... 我吓了一跳。 这个简单的想法给我留下了深刻的印象。 它促使我走向万有引力理论。

当我们在电视上观看宇航员在他们的轨道空间站内漂浮时，似乎重力已完全消除——这

---

[1] 然而，牛顿本人确实在四个不同的场合向四个不同的人讲述了这个故事。 参见 Westfall (1980, pp. 154 – 155)。

[2] 见（Pais 1982 §9）。

正是爱因斯坦洞察力的重要意义。 这种看似完全消除重力的原因是宇航员和他们的航天器在地球的引力场中一起自由下落——这并不是因为航天器离地球太远以至于它逃脱了重力！

是的，牛顿的平方反比定律告诉我们，那里的引力比这里的要弱，但快速计算表明，牛顿引力场对国际空间站（轨道高度为 250 英里）的拉力仅约比我们这里弱 12%。 不，重力的看似完全消除与这种适度的影响无关。

相反，重力的完全消除是显著经验事实的结果——伽利略在 1590 年左右首次发现，通过（据说）从比萨斜塔掉落物体——所有物体都受到重力的同等加速，无论它们的质量或组成。[3] 因此，宇航员和他们的航天器都沿着同一个自由落体的轨道一起运动。

这一经验事实同时被牛顿万有引力定律所铭记和解释。 如果一个质量为 m 的粒子与一个质量为 M 的粒子的距离为 r，那么 m 对 M 的引力作用可以忽略，我们可以认为 m 只是被后者的引力场拉向 M（力为 F）。 那么 m 对 M 的加速度 a 由牛顿第二运动定律给出：

$$F = ma$$

但是如果 r 是 m 到 M 的距离，那么**牛顿的平方反比万有引力定律**指出

$$\boxed{F = \frac{GmM}{r^2},}$$

其中 G 是**引力常数**。 结合这些事实，

$$m\ddot{r} = -\frac{GmM}{r^2} \quad \Longrightarrow \quad \boxed{\ddot{r} = -\frac{GM}{r^2},} \tag{30.1}$$

它与 m 无关。

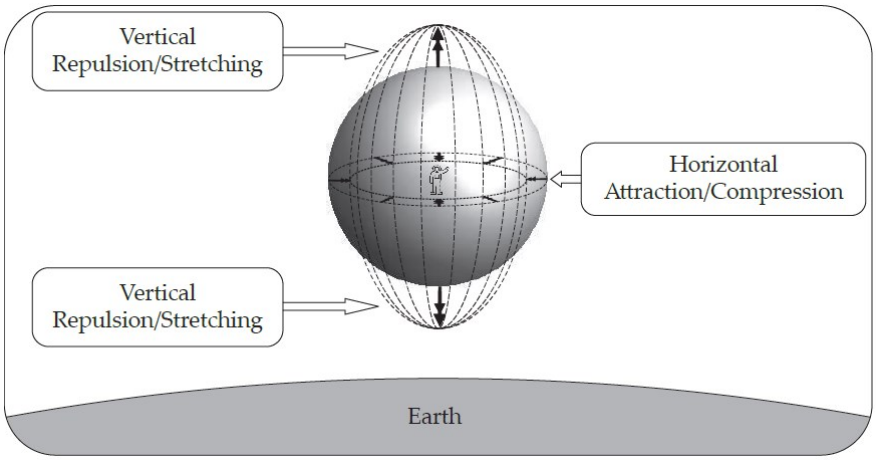
所以牛顿的万有引力定律解释了伽利略的另一个非常神秘的经验事实，即所有物体（以相同方式从同一个地方发射）都遵循相同的轨迹，而不管它们的成分如何。

但有可能更深入吗？ 我们能反过来解释牛顿万有引力定律的这个方面吗？ 爱因斯坦的几何理论正是这样做的！ 如果一个粒子注定要沿着几何确定的时空测地线行进——这与粒子没有任何关系！——这解释了伽利略的经验事实，同样给出了引力 F 必然与质量（m）成正比的原因！

[3] 只有在可以忽略空气阻力的情况下，这才是正确的。 有一部精彩的电影（来自 1971 年的阿波罗 15 号任务），宇航员大卫斯科特站在月球表面同时放下一根羽毛和一把锤子——他们同时击中了月球表面！

30.2 引力潮汐力

但是，如果对于航天器内一起自由落体的宇航员来说，重力似乎消失了，那么重力究竟还剩下什么？当然它必然留下一些它存在过的痕迹？！ 这是爱因斯坦惊人的洞察力引出的一个全新的问题。



[30.1] **地球的引力潮汐力作用在一个自由落体的粒子球体上**，导致在平行于地球表面的水平“赤道”平面压缩，并在引力场本身的垂直方向上拉伸。 随着球体下落，它因此开始变形为椭圆形蛋。 由于平方反比定律，垂直拉伸的强度正好是赤道压缩的两倍。

好吧，想象一下你自己就是那个宇航员，在地球上空，但你不是安全地在轨道上的空间站内，而是在外面，穿着宇航服，使用你的喷气背包在一个固定的位置盘旋。 进一步想象你的方向是直上直下的，你的头离地球最远，你的脚离地球最近。 围绕着你，均匀地分布在以你为中心的球体表面上，是数百个闪亮的金属珠子，像你一样暂时静止（见[30.1]）。

现在你关掉你的喷气背包。 你和你周围的珠子球面从静止状态开始一起自由落体，全部向下朝向地球加速。 当你坠落时，你会在粒子球面中观察到什么变化（如果有的话）？

如果引力场是完全均匀的，具有恒定的大小和固定的方向，你的答案将是，“我看不到任何变化！” 重力对你来说真的是不可见的。 而且由于地球的实际引力场在小球体上几乎是均匀的，所以你也不会看到任何变化… 至少一开始不会。

然而，慢慢地，你会开始注意到你头上和脚下的球体部分正在加速远离你，而“赤道”粒子正在加速接近你。 最终结果是球面开始变成鸡蛋！

要理解为什么会发生这种情况，请参阅 [30.1]，让我们立即澄清，该图画出了将驱使球体变形成为鸡蛋的力，而非鸡蛋本身，它是在向地球坠落一段时间后球面的新形状。

首先，要意识到赤道粒子，在与你相同的高度，并没有在与你完全相同的方向上加速。是的，你和粒子正在以相同的速率向下加速，但你们都在向地球中心加速[4]，因此你们的轨

[4]实际上，你正被拉向地球的每一个小部分，其作用力与该特定部分与你的距离的平方成反比。 但事

迹正在缓慢收敛。 由于你只看到相对加速度，赤道粒子似乎直接朝你加速，好像被你吸引了，如 [30.1] 所示。

现在看头顶上球面顶部的粒子，以及脚下球面底部的粒子。 与之前的情况不同，你和粒子现在都朝着同一个方向前进，笔直向下朝向地球中心。 然而，你脚下粒子的加速度大于你的加速度，因为它们比你更靠近吸引力中心。 因此，你会看到它们开始加速远离你，向下朝向地球运动。 同样，你头顶上的粒子的加速度小于你的加速度，因为它们离吸引力中心更远，所以你向下加速的速率比它们快，所以它们似乎也在远离你，向上加速运动。 因此，这些粒子似乎被你排斥，如 [30.1] 中所示。

这种在一个平面内压缩并在正交方向上拉伸，将球体变形为鸡蛋的现象称为重力**潮汐力**。 这是对我们受爱因斯坦启发的问题的答案，“自由落体中还剩下什么引力？”

让我们更直接地将这些观察结果与我们在地球轨道上的原始宇航员联系起来。假设你和你的粒子球都以任意但共享的速度一起爆炸，而不是从静止中坠落；鸡蛋诱导力将与之前一模一样！特别是，假设这个速度被选择为水平指向，平行于地球表面，并假设发射速度被选择为使你和你的球体进入一个围绕地球的圆形轨道。最后，想象一下你和你的球体在轨道空间站内，而金属珠子被替换为其他宇航员！如果你仔细观察，我们会发现重力并没有完全消除：垂直（即径向）分离的宇航员会逐渐加速远离你，而水平分离的宇航员会逐渐加速接近你。然而，正如我们将在下一节中看到的，这些潮汐力与粒子的间隔成正比，并且在空间站内的小距离内它们是无法检测到的。但在更大的范围内，这些潮汐力的影响变得显而易见，这将我们带到了下一个话题…

为什么这种引力畸变力场被称为“潮汐”？ 这种现象与我们海洋的潮汐之间有什么联系？

想象在[30.1]中地球是太阳，而我们的小球体是地球，或者更确切地说是地球的薄皮——它的海洋。 太阳对地球的潮汐力现在导致地球面向太阳的一侧的海洋向太阳隆起，而相反的一侧的海洋远离太阳隆起，表现为在这两个相反的方向海平面上升。 这是 [30.2] 四幅图中画出的固定的、虚线的椭圆形蛋。

随着地球每 24 小时自转一次，这些凸起的方向保持不变，因此每个沿海地区每 12 小时就会遇到一次这样的凸起——这就是涨潮！ 好吧，正如 [30.2] 中所说明的，它实际上比这要复杂得多，因为月球对我们的海洋施加完全相同的引力畸变，而且，尽管它的质量与太阳相比微不足道，但因其靠近地球故产生潮汐力大约是太阳的两倍[5]！ 实际的潮汐是太阳和月亮这两个潮汐力叠加的复杂[6]结果。 [30.2] 说明并解释了它们在一个月球轨道上的相互作用。 这是牛顿最先了解的，是《原理》的又一次胜利（牛顿，1687 年）。

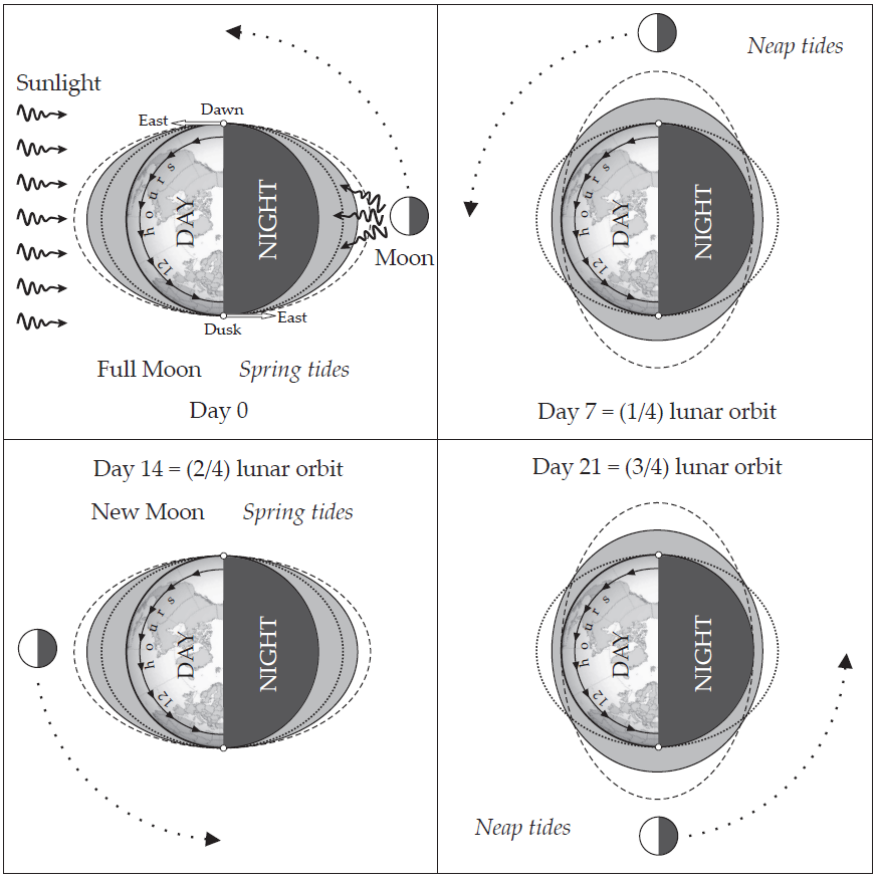
当我们看到满月（[30.2] 的左上角）或新月（[30.2] 的左下角）时，几何关系是最简单的，因为这意味着月球、地球和太阳排成一条直线，在这种情况下，月球和太阳在地球海洋

实证明，这种极其复杂的力总和的净效应恰恰就像地球的所有质量都集中在它的中心！ 牛顿本人惊讶地发现他的平方反比定律的这种真正神奇的性质，他在《原理》中提供了一个优雅的几何证明（牛顿 1687，定理 31）。 请注意，这是平方反比定律的另一个签名——其他类型力的定律没有这种性质。

[5] 见练习 13。

[6] 参见 Schutz (2003, §5)。

上引起的潮汐隆起也在一条直线上，相互加强，产生最大的高潮，也产生最低的低潮； 这些被称为**大潮**（Spring tides）。 （注意：这些全年都在运作，与我们称之为“春天”的季节无关。）



[30.2] **海洋潮汐的解释**。 月球（短线虚线）和太阳（点虚线椭圆）的潮汐力共同扭曲了海洋（阴影椭圆区域），从而产生了潮汐。 （注意：并非按比例缩放的！海洋和潮汐被大大夸大了，到月球的距离也被大大缩小了！）左上角，月球被照亮的一半面向地球，反射的阳光被我们视为满月。 由于太阳和月亮的潮汐力现在对齐，它们相互加强以产生最大的高潮和最低的低潮； 这些是春潮。 右上角，更强大的月球潮汐力被正交的太阳潮汐力部分抵消，导致高潮和低潮之间的差异最小； 这些是小潮。

当月球和太阳与地球成直角时（[30.2] 的右上角和右下角），它们会在正交方向上产生凸起，因此较小的太阳凸起部分抵消了较大的月球凸起，导致涨潮和退潮两者之间的差异最小； 这些被称为**小潮**（Neap tides）。

因为——根据定义！——月球每个月绕地球运行一圈，它每月与地球和太阳对齐两次：新月对应于月亮位于太阳和地球之间，而满月对应于地球位于太阳和月亮之间。 因此，大潮每两周发生一次，小潮也是如此，两者之间的交替是每周一次——月球绕地球运行四分之一轨道所需的时间，如 [30.2] 所示 .

要了解更多关于重力物理学（尤其是潮汐）的信息，我们强烈推荐 Gravity From The Ground Up (Schutz 2003)。

30.3 几何形式的牛顿万有引力定律

（注意：以下对爱因斯坦引力几何理论的阐述受到彭罗斯美丽的论文《宇宙的几何》（Penrose 1978）以及他杰出著作《通往实在之路》（Penrose 2005，§§17-19）的三个相关章节的启发。然而，我们将明确证明彭罗斯申明但作为练习留给读者的几个关键结果，我们还将包括一些彭罗斯未提及的额外几何观察。）

我们刚刚描述的潮汐力对于任何可想象的随距离减小的引力定律定性都一样。 但是，由牛顿发现并发表在他 1687 年的《原理》中的实际万有引力定律随着距离的平方而减小。正如我们现在将要展示的，这条定律，而且仅这条定律，会产生具有非常独特和非常漂亮几何特征的潮汐力。

设 r 是你到地球中心的距离，设 ξ 是环绕你的粒子球的水平“赤道”圆的半径，设 $\delta\varphi$ 是 ξ 在地球中心所对的小角度 . 那么 $\xi \asymp r\,\delta\varphi$。由于 $\delta\varphi$ 在你和赤道粒子径向朝地球中心下落时保持固定，所以这个圆朝你的向心加速度由下式给出

$$\ddot{\xi} \asymp \ddot{r}\,\delta\varphi = -\frac{GM}{r^2}\,\delta\varphi = -\frac{GM}{r^3}\,r\,\delta\varphi \asymp -\frac{GM}{r^3}\,\xi,$$

凭借（30.1）可得。

如果前两章已经完成了他们的工作，那么你对这个公式的直接的，巴甫洛夫式的反应[7]会是惊呼：“这是一个雅可比方程！” 我们很快就会明确这种联系，但是水平分割粒子的这种吸引力确实是由于正截面曲率 $\mathcal{K}_+$，其中

$$\boxed{\ddot{\xi} = -\mathcal{K}_+\xi \quad \Longrightarrow \quad \mathcal{K}_+ = +\frac{GM}{r^3}} \tag{30.2}$$

让 Ξ 表示从你到你头上和脚下的粒子的距离——我们选择了一个希腊字母来描述这一点！ 如 [30.1] 所示，并且如前所述，这些粒子远离你的相对加速度由在你的高度 r 和高度 $r+\delta r = r+\Xi$ 的加速度 $\ddot{r}$ （由（30.1）给出）之差 $\delta\ddot{r}$ 给出。 因此，

$$\ddot{\Xi} = \delta\ddot{r} \asymp [\partial_r \ddot{r}]\,\delta r = \partial_r\left[-\frac{GM}{r^2}\right]\Xi = \left[+\frac{2GM}{r^3}\right]\Xi$$

垂直分割的粒子的这种排斥是由于负截面曲率 $\mathcal{K}_-$，其中

$$\boxed{\ddot{\Xi} = -\mathcal{K}_-\,\Xi \quad \Longrightarrow \quad \mathcal{K}_- = -\frac{2GM}{r^3}} \tag{30.3}$$

[7] 见 (28.2) 和 (29.21)。

因此，垂直斥力恰好是水平引力的两倍，如 [30.1] 所示。我们现在准备承认这一事实的真正几何意义——牛顿平方反比潮汐力的几何特征。 关键是要跟踪球体落入地球时体积 $\delta\mathcal{V}$ 的演变，潮汐力将其变形为椭圆形蛋。

想象一下，最初的粒子球体被封闭在一个立方体中。 当球体下落并变形为鸡蛋时，封闭的立方体及其内容物会发生线性变换，将其变形为一个边长为 2ξ、2ξ 和 2Ξ 的盒子，因此体积为 $8\,\xi^2\,\Xi$。 鸡蛋所占的体积是这个盒子的一个固定比例，所以

$$\boxed{\delta\mathcal{V} = \frac{4\pi}{3}\xi^2\,\Xi}$$

由于你和粒子球从静止开始，显然这个体积的变化率最初必然为零：$\delta\dot{\mathcal{V}} = 0$。 这可以通过计算来证实：

$$\delta\dot{\mathcal{V}} = \frac{4\pi}{3}\left[2\xi\dot{\xi}\Xi + \xi^2\dot{\Xi}\right] = 0$$

因为 $\dot{\xi} = 0 = \dot{\Xi}$。

但是潮汐力确实会立即加速球体的粒子，所以现在让我们计算鸡蛋体积的加速度：

$$\delta\ddot{\mathcal{V}} = \frac{4\pi}{3}\left[2\left(\dot{\xi}\right)^2\Xi + 2\xi\ddot{\xi}\,\Xi + 2\xi\dot{\xi}\dot{\Xi} + 2\xi\dot{\xi}\dot{\Xi} + \xi^2\,\ddot{\Xi}\right] = \frac{4\pi}{3}\left[2\xi\ddot{\xi}\,\Xi + \xi^2\,\ddot{\Xi}\right]$$

但是，最初，$\xi = \delta r = \Xi$，我们知道水平引力和垂直斥力分别由（30.2）和（30.3）控制，所以，

$$\delta\ddot{\mathcal{V}} = \frac{4\pi}{3}\left[2\ddot{\xi} + \ddot{\Xi}\right](\delta r)^2 = -\frac{4\pi}{3}\left(2\mathcal{K}_+ + \mathcal{K}_-\right)(\delta r)^3 = 0\,!$$

我们到达了美丽的

平方反比潮汐力的几何特征：
由平方反比定律产生的潮汐力——并且只有平方反比定律——是**保持体积的**，在精确意义上，体积的加速度为零，因此体积 对 t^2 阶保持不变：

$$\delta\ddot{\mathcal{V}} = 0 \qquad (30.4)$$

我们还没有证明这个命题的“唯一”部分。 为此，假设 $\ddot{r} = f(r)$，其中 f 是未知函数。用它代替上述分析中的（30.1），我们发现 [练习]，

$$\delta\ddot{\mathcal{V}} = 0 \quad \Longrightarrow \quad \frac{df}{dr} + \frac{2f}{r} = 0 \quad \Longrightarrow \quad f(r) \propto \frac{1}{r^2}$$

正如将要证明的。

在前面的分析中，可以理解的是，我们忽略了你的微观引力场及其对你周围粒子的影响。但假设我们将你从球体内部移开，取而代之的是用密度为 ρ 的极其致密的物质填充它。 然后半径为 $\xi = \Xi$ 的粒子球体将向中心加速，就像在中心有一个点质量 $\rho\,\delta\mathcal{V}$ 一样，以加速度将它们拉入，

$$\ddot{\xi} = -\frac{G\,\rho\,\delta\mathcal{V}}{\xi^2}.$$

现在让我们计算在这种新情况下体积的加速度。 由于现在 $\delta\mathcal{V} = \frac{4\pi}{3}\xi^3$，

$$\delta\dot{\mathcal{V}} = 4\pi\xi^2\dot{\xi} \quad \Longrightarrow \quad \delta\ddot{\mathcal{V}} = 8\pi\xi(\dot{\xi})^2 + 4\pi\xi^2\,\ddot{\xi} = 8\pi\xi(0)^2 + 4\pi\xi^2\left(-\frac{G\rho\,\delta\mathcal{V}}{\xi^2}\right).$$

因此我们发现了这个

反平方吸引力的几何特征：
考虑一个体积为 $\delta\mathcal{V}$ 的球体，其中填充有密度为 ρ 的物质，并假设其表面正上方是微小的测试粒子，从静止状态释放出来。 它们立即开始向中心加速，平方反比定律导致它们所包围的体积内向爆炸，加速度受此几何定律支配：

$$\delta\ddot{\mathcal{V}} = -4\pi G\rho\,\delta\mathcal{V} \tag{30.5}$$

如果我们现在想象这个物质球体（以及它的测试粒子表皮）在地球引力场中以任意速度发射，那么我们得到这两种效应的叠加：地球的潮汐力开始将球体变形为等体积的蛋，但球内物质的吸引力也会发挥其缩小体积的作用。 最终结果是最初的球体演变成一个蛋，但它的体积与内部质量和时间平方成正比收缩。

30.4 时空度量

很快，我们将能够根据爱因斯坦的 4 维时空重新构建上述分析，然后我们将很自然地被引导到爱因斯坦的引力场方程。 但在此之前，我们必须先讨论时空的度量结构，而且我们还必须学习如何绘制时空图。

首先要理解的是，局部切线空间结构不是标准欧几里得度量的 $\mathbb{R}^4$； 取而代之的是闵可夫斯基时空，我们在（6.15）中介绍的时空间隔给出了**闵可夫斯基度量**。（我们提醒读者，我们选择的单位是光在一个单位时间内传播一个单位的距离，所以 它的速度是 $c = 1$。在

这个单位中，爱因斯坦著名的方程 $E = mc^2$ 简化为 $E = m$，即“能量就是质量”!）

因此，

$$ds^2 = dt^2 - (dx^2 + dy^2 + dz^2) = [dx^0]^2 - ([dx^1]^2 + [dx^2]^2 + [dx^3]^2) \qquad (30.6)$$

这类似于切平面内到曲面点的距离，就像切平面是平的一样，闵可夫斯基时空也是平的——它的黎曼张量同样为零。

曲面本身相邻点之间真实距离的类似物由弯曲时空的**度量张量**给出，通常表示为 $\mathbf{g}$——它以两个向量作为输入，并输出一个标量。 它推广了**点积**（或**标量积**），并且是对称的：

$$g(\mathbf{u},\mathbf{v}) \equiv \mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u} = g(\mathbf{v},\mathbf{u}).$$

就像在 2-曲面中一样，度量构成了关于时空的最基本信息：它定义了距离，一旦我们知道了距离，我们就知道了一切：测地线、平行移动和黎曼曲率。 如果 ε 是时空中两个相邻事件之间的一个小（最终趋于零）连接向量，那么该度量告诉我们它们之间的爱因斯坦距离：

$$ds^2 = g(\varepsilon,\varepsilon)$$

但回想一下，只有当 ε 连接具有非零质量的物质粒子的时空轨迹（也称为**世界线**）上的两个点时，$ds^2 > 0$，该物质粒子的行进速度被限制为小于光速； 在这种情况下，我们说间距是**类时间的**（timelike）。 正如我们在（6.16）中所述，在这种情况下，ds 只是测量环球线路的旅行者经过的手表时间。 如果 ε 是沿着光线（光子的世界线），则 $ds^2 = 0$； 在这种情况下，我们说间距是空的（null）。 最后，如果 $ds^2 < 0$，我们说间距是**类空间的**（spacelike）。

如果 $\{\mathbf{e}_i\}$ 是一组四个基向量（称为四分体），它不需要是正交的，那么我们可以用与获得里奇（Ricci）张量分量完全相同的方式获得度量张量的分量——只需将张量应用于基向量对：

$$g_{ij} \equiv \mathbf{g}(\mathbf{e}_i,\mathbf{e}_j) = \mathbf{g}(\mathbf{e}_j,\mathbf{e}_i) = g_{ji}$$

因此，

$$\boxed{ds^2 = \mathbf{g}(\varepsilon,\varepsilon) = \mathbf{g}(dx^i\,\mathbf{e}_i, dx^j\,\mathbf{e}_j) = \mathbf{g}(\mathbf{e}_i,\mathbf{e}_j)\,dx^i\,dx^j = g_{ij}\,dx^i\,dx^j} \qquad (30.7)$$

例如，在闵可夫斯基时空中，如（30.6）所示，$g_{00} = +1$，而 $g_{11} = g_{22} = g_{33} = -1$，如果 $i \neq j$，则 $g_{ij} = 0$。 然而，一般来说，这些 g_{ij} 是时空中的函数，如果 $i \neq j$，则 $g_{ij} \neq 0$。 重要的是要认识到，即使在平坦的闵可夫斯基时空中，这些分量看起来也可能截然不同，这取决于坐标的选择。 例如，如果我们使用标准球极坐标作为空间坐标，那么

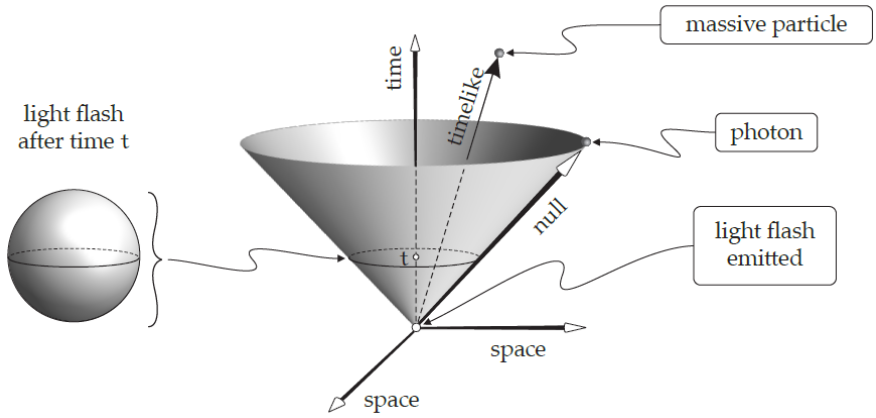
$$ds^2 = dt^2 - dr^2 - r^2(d\varphi^2 + \sin^2\varphi\, d\theta^2), \qquad (30.8)$$

所以，$g_{tt} = 1$，$g_{rr} = -1$，$g_{\varphi\varphi} = -r2$，和 $g_{\theta\theta} = -r^2 \sin^2\varphi$，但这些函数描述与之前完全相同的平面几何。

注解：这个 g_{ij}-符号已被普遍接受：它可以在所有现代课本中找到，无论是数学还是物理。 所以我们也将使用这个符号，它是高斯原始对 2-曲面的 (E, F, G) 符号的 n 维推广；请参见（4.8）。另一方面，至少在二维情况下，我们希望我们已经多次说服你——我们的替代 (A^2, B^2) 表示法的概念和计算优势。 尽管如此，我们的符号不敢涉足 n 维...

30.5 时空图

绝大多数人类（包括作者）无法直接可视化 4-维时空。 那么，我们如何画图，使我们能够对时空进行几何推理？！ 这个难题最常见和最有用的答案是简单地隐藏三个空间方向之一。 更具体地说，我们将向上画出时间，并将（三个中的）两个空间方向表示为形成一个水平面，垂直于竖直的时间方向。



[30.3] **时空中的零锥（光锥）**。 时间由竖直方向表示，（三个中的）两个空间方向由正交的水平方向表示。 在时空中的一个事件中发出一道闪光，在每个时间 t 处，膨胀的光球由其圆形横截面表示。 因此，闪光的整个未来由一个圆锥表示，光子的世界线是圆锥的零母线。 物质粒子被限制在锥体内以低于光速的速度沿类时的世界线行进。

尽管这种信息丢失看起来很严重，但对称性通常会帮助我们：如果两个空间方向在物理上是等价的，那么丢弃其中一个方向不会对我们造成任何损失。 例如，在 [30.1] 中，所有水平方向（在与地球表面平行的平面中）都无法区分，因此如果我们想象这个平面由两个正交向量张成，那么我们可以安全地隐藏其中一个， 不丢失信息。 与此形成鲜明对比的是，竖直方向是引力场的方向，它与水平方向在物理上是截然不同的，所以在时空图中，我们不能把那个方向扔掉！

但让我们从一个更加对称的情况开始：空的、平坦的闵可夫斯基时空。 在这里，所有空间方向在物理上都无法区分。 在时空图中画出的最有用的对象之一是在特定时间从特定

点发出的闪光——一个事件。 这导致光球每单位时间扩展一个单位距离。 由于爱因斯坦告诉我们物质粒子的传播速度不能超过光速，任何从与闪光相同的事件中喷出的大质量粒子都必须留在这个膨胀的球体内。 所有这些在时空图中看起来如何？

隐藏一个空间维度，膨胀的光球现在由 [30.3] 中所示的膨胀圆表示，它产生一个圆锥。因此，闪光的时空描述被称为**光锥**或**零锥**。 如图所示，物质粒子的世界线停留在该圆锥内。零锥的根本重要性在于，它的内部因此代表了可以受原始事件影响的事件集合——它告诉我们所谓的时空的**因果结构**，即这个事件，此时此地，是否可以导致彼时彼地的那个事件发生？

粒子世界线的切向量称为粒子的 **4–速度**。 静止的粒子具有非零的 4-速度：它直线向上指向时间轴！ 请注意，对于大质量粒子，可以按照我们通常坚持我们的粒子以单位速度在 2-曲面上行进的相同方式归一化 4-速度，但这对于光子的 4-速度是不可能的 ，因为光子 4-速度的“长度”始终为零。

30.6 几何形式的爱因斯坦真空场方程

为了“推导出”[8]爱因斯坦的真空场方程，让我们回到地球表面上方的空间真空，以及在那里运行的牛顿平方反比潮汐力，如 [30.1] 中所述。 但现在让我们从时空的角度来分析一下。

我们在[30.1]中已经注意到所有水平方向都是不可区分的，所以如果我们画出一个只保留一个这样的方向的时空图，就不会丢失信息。 因此，考虑由一个这样的水平方向和地球引力的竖直方向张成的球体的任何竖直大圆。 随着粒子球体的下落，潮汐力将球体扭曲成一个椭圆形的蛋，但仅通过观察原始球体的这个单一的竖直大圆横截面的演变，就可以完全忠实地捕捉到这种变化。 如果我们只在这些粒子下落时跟随它们，潮汐力会将它们的圆扭曲成一个椭圆，但随后可以简单地通过围绕其竖直对称轴旋转椭圆来恢复完整的物理椭球体。

再一次想象一下，你是这个球体中心的宇航员。 让我们从你的角度画一个时空图。 关闭喷气背包后，你将处于自由落体状态，周围的粒子也是如此。 所以，只要你不低头看地球，你就觉得你只是在漂浮，一动不动，而以你为中心的粒子球也几乎是一动不动的… 除了你逐渐见证了球体的潮汐演变成为一个蛋。 如果你向下看地球，你会看到它向你移动，但你仍然有权认为自己一动不动——当然，是地球向你移动！

正是从这个角度，我们画出了[30.4]，你是中心的黑色世界线。 它画出了球体（由垂直圆圈表示）在落向地球时变成椭圆形蛋的时空演变。 在这里，地球由标记为 MASSIVE 的对象示意性地表示。

这张图应该看起来眼熟！ 它看起来几乎与 [29.11c] 完全一样，描绘了在一个方向为正

[8] 在我们目前的理解状态下，物理定律是不可简化的：它们不能从任何更原始的事物逻辑推导出来。然而，从牛顿开始，与其他物理定律的一致性，再加上数学之美，已经成为非常有效的指导原则，使我们能够正确猜测自然界中一些最深奥的秘密。

（引起吸引力）而在垂直方向为负（引起排斥）的截面曲率存在的情况下圆形测地线束的演变。 如果我们考虑时空中的截面曲率，并将（30.2）和 （30.3）与（现在是 4 维的）截面雅可比方程（29.21）进行比较，这个类比就变得精确了。

粒子圈（从静止开始）的 4-速度 $\mathbf{v}$ 是垂直的，沿时间轴。 当包含这个 4-速度的垂直时空平面围绕它旋转时，它通过与地球表面平行的方向，具有正的、吸引的截面曲率（30.2）：

$$\mathcal{K}_{+} = +(\mathrm{GM}/\mathrm{r}^3)$$

旋转另一个直角，时空平面包含地球引力场的方向，产生负的、排斥的截面曲率 (30.3)：

$$\mathcal{K}_{-} = -(2\mathrm{GM}/\mathrm{r}^3)$$

但是，当然，还有与这两个方向正交的第三个空间方向，但在我们的时空图中被隐藏了——它是与时间轴正交的第三个方向，位于与地球表面平行的平面上。

让我们用适应这种情况的正交四分体来说明这一点。 让 $\mathbf{e}_0 = \mathbf{v}$ 沿（垂直）时间轴指向，让 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 张成平行于地球表面的平面，让 $\mathbf{e}_3$ 在重力场方向上径向远离地球。

（29.26）从 3-流形到 4-流形的推广是，在包含 $\mathbf{e}_0 = \mathbf{v}$ 的所有平面上，时空中截面曲率的平均值 $\mathcal{K}_{\mathrm{mean}}$ 可以通过仅对包含 $\mathbf{e}_0 = \mathbf{v}$ 的这三个正交平面求平均来获得：

$$\Pi_1 \equiv \Pi(\mathbf{e}_0, \mathbf{e}_1), \Pi_2 \equiv \Pi(\mathbf{e}_0, \mathbf{e}_2), \text{and } \Pi_3 \equiv \Pi(\mathbf{e}_0, \mathbf{e}_3)$$

所以，

$$\mathcal{K}_{\mathrm{mean}} = \frac{\mathcal{K}(\Pi_1) + \mathcal{K}(\Pi_2) + \mathcal{K}(\Pi_3)}{3} = \frac{1}{3}\mathbf{R}\mathrm{icci}(\mathbf{e}_0, \mathbf{e}_0) = \frac{1}{3}\mathrm{R}_{00}$$

非常值得注意的是，以这种几何方式来看，从牛顿的平方反比潮汐力到爱因斯坦的弯曲时空的通路是完全无缝衔接的！

我们在 （30.4）中看到牛顿平方反比潮汐力的特征是保持体积，我们在（29.29）中看到体积的加速度受里奇曲率控制。 结合这两个结果，我们发现

$$\boxed{\mathbf{R}\mathrm{icci}(\mathbf{v}, \mathbf{v})\, \delta\mathcal{V} = -\delta\ddot{\mathcal{V}} = 0.} \tag{30.9}$$

更明确地说，从（30.2）和（30.3），

$$\mathrm{R}_{00} = \mathcal{K}(\Pi_1) + \mathcal{K}(\Pi_1) + \mathcal{K}(\Pi_1) = \mathcal{K}_{+} + \mathcal{K}_{+} + \mathcal{K}_{-} = \frac{\mathrm{GM}}{\mathrm{r}^3} + \frac{\mathrm{GM}}{\mathrm{r}^3} - \frac{2\mathrm{GM}}{\mathrm{r}^3}$$

所以

$$\boxed{\mathrm{R}_{00} = 0.} \tag{30.10}$$

这里我们假设粒子球面从静止开始，所以 $\mathbf{v} = \mathbf{e}_0$ 是一个纯垂直的 4-速度，沿时间轴。但是我们知道，如果我们以任何速度发射球体，牛顿潮汐力的保持体积特征将继续起作用，因此 (30.9) 对于任何类时 4-速度 $\mathbf{v}$ 都成立。

推导爱因斯坦真空场方程的最后一步取决于 **R**icci 的对称性 (29.28)：

$$\mathbf{R}\mathrm{icci}(\mathbf{x}, \mathbf{y}) = \mathbf{R}\mathrm{icci}(\mathbf{y}, \mathbf{x})$$

现在让 $\mathbf{v} = \mathbf{x} + \mathbf{y}$，对于任意（类时的）$\mathbf{x}$ 和 $\mathbf{y}$。 于是，

$$\begin{aligned}
0 &= \mathbf{R}\mathrm{icci}(\mathbf{v}, \mathbf{v}) \\
&= \mathbf{R}\mathrm{icci}([\mathrm{x} + \mathrm{y}], [\mathrm{x} + \mathrm{y}]) \\
&= \mathbf{R}\mathrm{icci}(\mathrm{x}, \mathrm{x}) + \mathbf{R}\mathrm{icci}(\mathrm{x}, \mathrm{y}) + \mathbf{R}\mathrm{icci}(\mathrm{y}, \mathrm{x}) + \mathbf{R}\mathrm{icci}(\mathrm{y}, \mathrm{y}) \\
&= 0 + \mathbf{R}\mathrm{icci}(\mathrm{x}, \mathrm{y}) + \mathbf{R}\mathrm{icci}(\mathrm{y}, \mathrm{x}) + 0 \\
&= 2\mathbf{R}\mathrm{icci}(\mathrm{x}, \mathrm{y})
\end{aligned}$$

我们就这样到达了

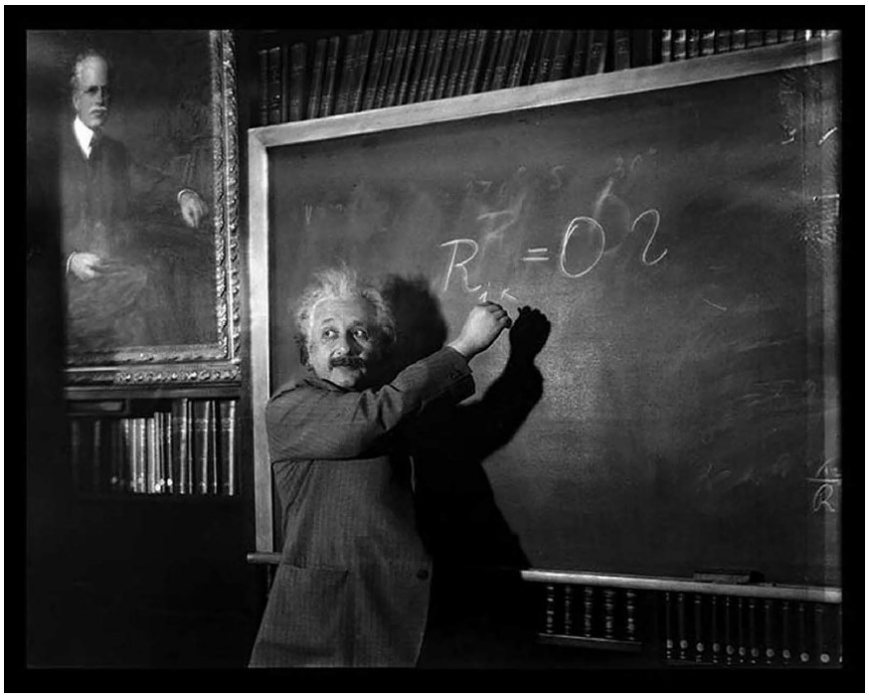
$$\boxed{\textbf{爱因斯坦真空场方程：}\mathbf{R}\mathrm{icci} = 0 \quad \Leftrightarrow \quad \mathrm{R}_{\mathrm{ik}} = 0。} \tag{30.11}$$

让我们不要忘记它的来源和含义：

> 在真空中，为了让潮汐力保持体积，正截面曲率必须与负截面曲率完美平衡，使得以平均值中完全抵消。

虽然里奇（Ricci） 曲率在真空中同样为零，但这只是正截面曲率和负截面曲率的体积减小平均值。 黎曼张量本身当然一般不会为零。 在（30.2）和（30.3）中见证正向和负向潮汐曲率。

一般来说，可以将黎曼曲率拆分为体积减小的里奇部分，加上一个纯潮汐的体积保持部分，称为**外尔曲率**（Weyl curvature）； 见习题 15。这种拆分通过彭罗斯的 2-旋量形式最优雅自然地完成的； 参见 Penrose 和 Rindler (1984, §4.6) 或 Wald (1984, §13.2)。



[30.5] 阿尔伯特·爱因斯坦（1879-1955）大约在 1931 年，在加利福尼亚州帕萨迪纳的威尔逊山天文台的黑尔图书馆发表演讲。 黑板上，爱因斯坦刚刚写下了他的真空场方程：$R_{ik} = 0$（加上一个问号）。 图片由位于加利福尼亚州圣马力诺市亨廷顿图书馆的卡内基科学研究所天文台提供。

免得你怀疑（30.11）实际上是爱因斯坦的方程，[30.5] 显示爱因斯坦自己自豪地写下了它，虽然他谦虚地添加了一个问号，但他的方程现在包含了一系列惊人的、可检验的预测，所有这些已经通过实验得到证实，而且通常非常精确。

我们将概述其中的几个实验成果，但它们的数学推导必须留给我们将提供的技术参考资料，以及本书末尾延伸阅读的广义相对论部分中的那些参考资料。

30.7 史瓦西解和理论的第一次检验

当我们谈到爱因斯坦方程的解时，我们指的是满足该方程的时空几何（由其度量定义）。

唯一最重要的此类解（至少在历史上）描述了球对称（非自旋）质量 M 之外的真空区域中的时空。该解是有史以来发现的第一个精确解，由卡尔· 史瓦西[9]（Karl Schwarzschild，1873-1916，如图 [30.6] 所示）几乎是在爱因斯坦宣布他的理论之后立即发现。史瓦西在第一次世界大战期间作为一名 42 岁的炮兵左撇子在俄罗斯前线服役时发现了它。他在 1915 年 12 月 22 日的一封信中将解发送给爱因斯坦：

> 如你所见，尽管枪声猛烈，战争对我还是很友善的，使我能够摆脱这一切，并在你的想法的土地上漫步。

[9]史瓦西是一个神童，16 岁时就发表了有关天体力学的论文。 后来，从 1901 年到 1909 年，他在哥廷根担任教授，他的同事包括希尔伯特、克莱因和闵可夫斯基。

[30.6] 卡尔・史瓦西 (1873 – 1916)。

几十年后，物理学家会慢慢意识到，史瓦西的解不仅描述了诸如太阳或地球等球形质量周围的几何形状，还描述了黑洞的纯真空引力场！

史瓦西解建立在闵可夫斯基（Minkowski）时空的球极度量公式 (30.8) 之上：

<u>史瓦西解：</u>

$$ds^2 = (1-\frac{2GM}{r})dt^2 - \frac{dr^2}{\left(1-\frac{2GM}{r}\right)} - r^2(d\varphi^2 + \sin^2\varphi\, d\theta^2) \tag{30.12}$$

注意：

- 这里，径向坐标 r 被定义为保持其与半径为 r 的球体的面积 $\mathcal{A}(r)$ 的欧氏关系，即 $r=\sqrt{\mathcal{A}(r)/4\pi}$。 因此 dr 不度量径向距离； 相反，径向距离等于 $dr/\sqrt{(1-\frac{2GM}{r})}$。 只有当我们从大质量物体（或黑洞）后退时，径向坐标才会渐近地重新获得其欧氏解释为径向距离。

- 如果我们不选择光速 $c=1$ 的单位，那么 $(2GM/r)$ 必须替换为 $(2GM/c^2r)$。

- 在我们的戏剧结束时，我们将使用**曲率 2–形式**来证实该几何确实满足爱因斯坦真空方程（30.11）。

- 当 g_{rr} 度量分量爆破时

$$r = r_s \equiv \textbf{史瓦西半径} \equiv \frac{2GM}{c^2} \tag{30.13}$$

爱因斯坦和其他人最初对此感到困惑，并认为这意味着当 $r = r_s$ 时时空本身存在某种奇异之处。 然而，这最终被证明是特定坐标选择的非物理人工产物。 如果一个长度为 l 的物体位于史瓦西半径，它会经历普通的、非奇异的有序 $(GM/r_s^3)\,l$ 潮汐力，换句话说，有序 $(c^6 l/G^2)\,(1/M^2)$。

- 如果质量 M 极大，则史瓦西半径处的潮汐力极小。

- 如果球形恒星或行星的半径为 R，则该解仅适用于其外部的真空区域，即 $r > R$。在内部，$r < R$，应用一个完全不同的度量，它满足完整的爱因斯坦场方程（有物质），我们很快就会推导出来。在均匀密度球体的情况下，同一个卡尔 •史瓦西（Karl Schwarzschild）几乎在找到真空解之后立即找到了这个非真空方程的精确解； 从逻辑上讲，它被称为**内部史瓦西解**。

- 对于半径为 430,000 英里的太阳来说，史瓦西半径只有 2 英里！ 地球的史瓦西半径只有大约三分之一英寸！ 因此，我们必须在接近史瓦西半径之前很久就从真空史瓦西解切换到不同的内部解。

尽管地球和太阳确实自转，但史瓦西解仍然是两者时空几何的极好近似[10]。 事实上，这个解足以对我们的太阳系进行完整的分析，并对太阳系中偏离牛顿理论的现象做出三个重要的预测。

为了历史的准确性，我们必须指出，所有这三个最初的预测都是爱因斯坦在他知道史瓦西的精确解之前做出的。 爱因斯坦通过对他的方程的近似解来做到这一点，这个古怪而不优雅的方程是他使用直角坐标获得的！

以下是爱因斯坦对其理论的三个初步预测/测试：

1. 当光通过引力场向上传播时，它的频率应该降低——我们说它是**红移的**。 爱因斯坦在 1907 年做出了这个预测，早在发现场方程之前，它就不是对该方程的直接检验。然而，如果实验否定了这一预测，爱因斯坦理论的整个框架就会失效。 可悲的是，直到 1959 年，也就是爱因斯坦去世四年后，庞德（Pound）和雷布卡（Rebka）才对这一预测进行了第一次权威性的地面证实。

2. 水星椭圆轨道的长轴随着每个轨道旋转一个微小的量。牛顿理论预测，每世纪的自转应该是 532 角秒——小于 1 度。然而，于尔班 • 勒威耶（Urbain Le Verrier）和后来的西蒙 • 纽科姆（Simon Newcomb）对自 1697 年以来收集的观测数据进行了极其准确的分

[10]快速旋转的恒星和黑洞必须由新西兰数学家罗伊 • 克尔（Roy Kerr，生于 1934 年）于 1963 年发现的另一种至关重要的解来描述； 因此，它被称为克尔解。 有关 Kerr 对其非凡发现的个人描述，请参见 Kerr (2008) 和 Wiltshire 等人。 （2009 年）。

析，到 1882 年，他们确定轨道的实际自转为每世纪 575 角秒——与每世纪 43 角秒的差异：非常小（但也不可磨灭且神秘的）牛顿理论的缺陷。

当爱因斯坦根据自己的万有引力定律计算修正时，它做出了一个明确的、毫不含糊的预测——无论判决如何，都不能上诉。

*爱因斯坦的公式得出每个世纪 43 角秒！*在那一刻，爱因斯坦意识到大自然已经对他说话了——他的法则被建立在世界的永恒结构中。水星的轨道在人类存在之前 40 亿年就以这种速度自转，并且在未来数十亿年还会继续如此。

爱因斯坦告诉一位朋友，在那一刻，他经历了身体上的心悸；他告诉另一位朋友，他觉得有什么东西真的在他体内“啪”了一声。参见 Pais (1982, p. 253)。

3. 爱因斯坦还计算出靠近太阳边缘的光线应该弯曲 1.75 角秒。 但是只有在日全食期间才能看到掠过太阳的星光，因此只有在 1919 年 5 月 29 日发生日全食之时，才能对预测进行检验。亚瑟·爱丁顿爵士组织了前往巴西和西非的观测队，只为在短暂的日食期间拍摄关键照片，并进行偏转测量，其实，看看是否有任何偏转。

   爱因斯坦对光弯曲的预测的戏剧性证实，以及它弯曲的精确程度，引起了国际轰动，在世界各地报纸的头版上都引起了轰动。 一夜之间，这位名不见经传的德国科学家家喻户晓，“爱因斯坦”成为天才的代名词。

在这些早期的胜利之后，接下来是几十年的休耕期，物理学家将注意力转移到量子力学。然而，从 1960 年代开始，广义相对论经历了某种复兴，并一直持续到今天，以许多新的理论预测和许多新的实验证实吸引了最优秀的理论家和实验家。

其中一些较新的测试仍然依赖于 1915 年的简单史瓦西解。我们对 GPS 导航的日常成功使用——参见 Taylor 和 Wheeler (2000, §A-1)——是对爱因斯坦理论和史瓦西解的新证实之一！

可悲的是，史瓦西本人并没有看到这些胜利中的任何一个：就在他发现了爱因斯坦方程的两个非凡解几个月后，他于 1916 年 5 月 11 日死于一种罕见的自身免疫性疾病。 为了纪念他，亚瑟·爱丁顿爵士写道：“… 他的乐趣是在知识的牧场上不受限制地游荡，而且，就像游击队的领袖一样，他的攻击落在了他们最意想不到的地方。”

30.8 引力波

爱因斯坦方程的其他证实来自真空场方程（30.11）的完全不同类型的解。 正如我们在第 231 页首次概述的那样，引力波的存在是由爱因斯坦在 1916 年预测的，而引力波的第一次实验检测发生在几乎整整一个世纪后的 2015 年 9 月 14 日。

我们习惯于这样的观念，即电磁波——比如我们用来观察世界的光——是穿越空间的光子。引力波是惊人的不同：它们是时空结构本身的曲率波纹！但不是任何旧的涟漪都可能：

就像史瓦西解一样，它们必须满足爱因斯坦真空方程。 它们是纯粹潮汐力的振荡波，已经摆脱了产生它们的剧烈事件，以光速穿越空间（作为空间！）。

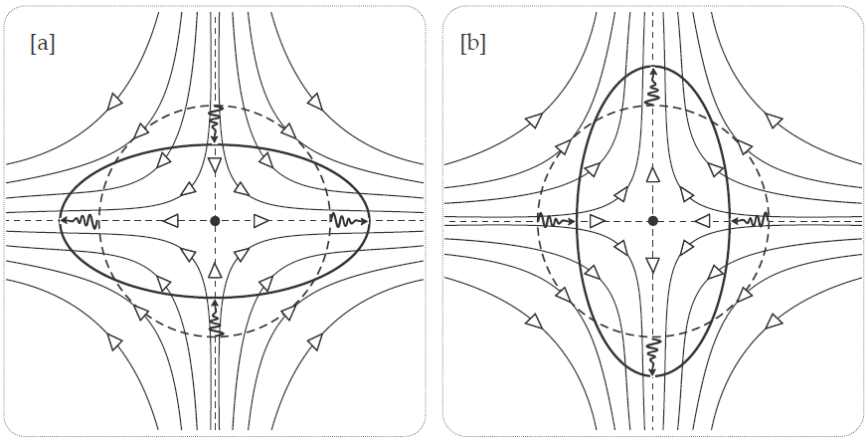
当引力波在一个固定位置穿过一个粒子球体时，它对在波的传播方向上分离的粒子没有影响。 但在垂直于波浪方向的平面上，它会引起振荡、潮汐、蛋形变形，使球体在一个方向上拉伸，并在正交方向上压缩它。 参见[30.7a]，它显示了某一特定时刻的**场线**（即潮汐力场的流线）。

引力波的潮汐力场与地球上空真空中的潮汐力场有两个重要方面的不同：

- 首先，引力波只在与拉伸方向正交的一个方向上压缩，而地球的潮汐力场在两个正交方向上压缩。

- 其次，爱因斯坦真空方程告诉我们，位于与地球表面平行的平面内的任意两个正交方向上的正曲率必须完全抵消地球引力场径向方向上的负曲率。 因此，这两个正曲率都必须是径向负曲率的一半，如 [30.1] 所示。 但在引力波的情况下，与行进方向相关的曲率消失了。 因此，爱因斯坦方程告诉我们，剩下的两个曲率必须抵消，因此必须具有相等的强度。 换句话说，在[30.7a]中，水平方向的拉伸力必须等于垂直方向的压缩力，如图所示。

回到[30.7a]，粒子的短线虚线圆最初变形为实线椭圆。 但是，半个波长后，潮汐力完全翻转，如 [30.7b] 所示：现在在波浪先前引起拉伸的方向上存在压缩，反之亦然。 当波通过时，力场在这两种相反的力模式之间来回振荡。



[30.7] **引力波的振荡潮汐力场**。 当引力波穿过一个粒子球体时，它对那些在波的传播方向上分离的粒子没有影响，这里被认为是垂直于纸面的。 在与该传播方向正交的平面中，波会在两个垂直方向（此处显示为水平方向和垂直方向）上引起振荡膨胀和收缩。 图 [a] 显示了测试粒子的（短线虚线）圆变形为（实线）椭圆。 图 [b] 显示了半个波长后波的潮汐力场的反转，现在拉伸与原始拉伸正交。 当波通过时，力场在这两个相反的模式之间来回振荡。

正如我们在 19.2 节中讨论的那样，一般来说，相图不能表示底层向量场的大小。 在 [30.7a] 中，我们仅仅画出了几条随机的场线，并且无法通过观察来判断潮汐力的强度。 然

而，假设向量场是无散度的，就像真空中的电场一样。 如果我们以这样一种方式画出场线，即它们的密度/拥挤程度与一个地方的电场大小成正比，那么，当我们沿着远离那个地方的场线移动时，新地方的场线密度将自动继续[11]以忠实地表示新位置的电场大小。 因此，虽然我们在 [30.7a] 中没有这样做，但引力波潮汐力场的流线也可以用这种方式画出，因为（正如我们现在解释的）它也是无散度的。

如果我们将平面 [30.7] 作为复平面，那么 [30.7a] 中的潮汐力场看起来就像在 [19.4h] (p. 200) 中首次遇到的鞍点（又名交叉点），代表复函数 (1/z)。 然而，雅可比方程告诉我们，引力波的潮汐力实际上随着距原点的距离而线性增加，因此用 $\bar{z} = x - iy$ 表示，即用“z 的波利亚向量场”表示； 见第 201 页。因此，

$$\nabla \cdot \begin{bmatrix} x \\ -y \end{bmatrix} = 0$$

注意：人们也可以从更高级的角度来理解这一点：所有复分析函数的波利亚向量场都是自动无散度（和无旋度）的； 参见 VCA，第 483 页。

但是，波与测试粒子球体之间的上述相互作用微弱得令人难以置信。 爱因斯坦本人怀疑我们是否能够探测到他的波。 尽管爱因斯坦确实低估了我们未来的技术实力，但更重要的是，他无法想象能够产生引力波的宇宙事件的非凡暴力。我们现在将描述一个这样的事件，它转换了惊人数量的能量（在几分之一秒内！）转化为强大的引力波——时空中名副其实的“潮汐波”[原文如此]，在扩散和传播超过 10 亿年之后，它仍能触发地球上的探测器！

2015 年对第一个引力波的探测（在两个独立的探测器上）本身就是划时代的，但科学家们能够走得更远：他们使用爱因斯坦方程从探测到的信号的细节逆推来非常详细地确定造成引力波的灾难性事件。

我们将让其中一位诺贝尔奖获得者基普·索恩 (Kip Thorne) 现身说法。（请注意，$M_\odot$ 表示我们太阳的质量。）

> 2015 年 9 月 14 日，先进的 LIGO 引力波探测器首次探测到：名为 GW150914 的波阵，振幅为 1.0×10^{-21}，持续时间 ~ 150ms，频率嘁嘁喳喳从~ 50Hz（进入 LIGO 波段时）到 240Hz．通过将观测到的波形与数值相对论模拟的波形进行比较，LIGO-VIRGO 科学家推断出这些波来自距离地球 12 亿光年的一个 $29M_\odot$ 黑洞与一个 $36M_\odot$ 黑洞合并形成一个 $62M_\odot$ 黑洞，在引力波中释放出 $3M_\odot c^2$ 的能量。—（Thorne 和 Blandford 2017，第 1346 页）

我们不能让能量释放的巨大程度不加评论地过去。 爱因斯坦的 $E = mc^2$ 将这一说法翻译为说，在黑洞碰撞的最后几分之一秒内，相当于我们太阳质量三倍的能量被转化为引力波能量。 这意味着什么？ 让我们从一个冷酷但熟悉的 $E = mc^2$ 的具现化开始：1945 年摧毁广岛的原子弹，造成 75,000 人死亡。 在那次原子弹爆炸中，转化为能量的物质总量还不到一个葡萄干。 牢记这点。

---

[11]参见 (VCA, pp. 494 – 502) 和 (Thorne and Blandford, 2017, §27.3.2)

那么，我们怎么能想象太阳的全部质量都被转化为引力波能量呢？！ Thorne 和 Blandford (2017, §27.5.5) 提供了这样的解释：在观测到的黑洞碰撞的最后十分之一秒内，产生 GW150914 引力波的功率输出是可观测宇宙中所有恒星光度之和的 100 倍！

关于引力波的产生、传播和探测的最全面、最新的技术处理方法，请参见 Thorne and Blandford (2017, §27) 和 Schutz (2021)。

自 GW150914.12[12] 以来，又发现了几个引力波。 2017 年，同样的科学家再创第一：他们发现了一个爆发，GW170817，爱因斯坦的方程允许他们将其解码为由两颗中子星合并产生的！ 利用 Schutz（1986） 早在 30 年前的有先见之明的理论工作[13]，这些引力波科学家随后能够使用这些数据（与光学数据相结合）以一种全新的方式计算哈勃常数，为宇宙膨胀研究带来新的曙光！ 显然，我们已经进入了**引力波天文学**的新时代。

同样，亚瑟•爱丁顿爵士在 1919 年首次见证了光的引力弯曲和聚焦，现在已经导致了**引力透镜**：一种利用光（和无线电波）的引力聚焦作为研究非常微弱和遥远物体的工具。

因此，事后看来，我们对爱因斯坦理论的测试类似于俯视望远镜的错误一端，检查它的镜头是否配置正确。 科学家们现在正从正确的角度审视这一理论，凝视时空，并使用爱因斯坦方程来见证和破译原本无法看到和无法解释的现象。

到目前为止，我们所描述的一切都只取决于爱因斯坦的真空场方程 (30.11)。 甚至产生 GW150914 的黑洞也可以用真空方程的解来描述，引力波本身也是如此。 但是如果没有亿万年前坍缩的恒星内部的物质，就没有黑洞，没有黑洞碰撞，没有引力波！

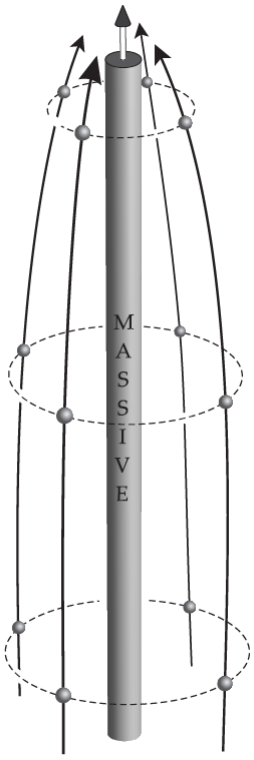
因此，我们现在转向关于物质的完整的爱因斯坦场方程。

30.9 几何形式的爱因斯坦场方程（含物质）

图 [30.4] 说明了当我们的粒子球仅围住真空时潮汐力的体积保持效应。 图 [30.8] 是当物质存在于球体中时体积减小效应的时空描述。 再一次，这张照片应该看起来眼熟！ 它几乎与 [29.11b] 完全相同！ 而且，这又不是偶然的：[30.8] 通过截面雅可比方程 (29.21) 显示了物质存在时纯正截面曲率的影响。

[12]这些事件的命名很简单：GW 代表“引力波”； 15 表示 2015 年； 09 表示月份，九月； 14 表示月份中的日期。

[13]对于这项工作，皇家天文学会授予伯纳德舒茨 2019 年爱丁顿奖章。



[30.8] **包含物质的自由落体球体体积减小的时空描述**。 像往常一样，时间是垂直运行的。

如果体积 $\delta\mathcal{V}$ 充满密度为 ρ 的物质，则牛顿的体积减小加速度由（30.5）给出，但在时空方面，这再次由截面曲率的平均值来描述，即由里奇曲率公式，（29.29）。 因此，结合这些结果，我们发现

$$\mathbf{R}\mathrm{icci}(\mathbf{v},\mathbf{v})\,\delta\mathcal{V} = -\delta\ddot{\mathcal{V}} = 4\pi G\rho\delta\mathcal{V}$$

因此

$$\mathbf{R}\mathrm{icci}(\mathbf{v},\mathbf{v}) = 4\pi G\rho \tag{30.14}$$

为了取得进一步的进展，我们必须用在狭义和广义相对论中描述物质和能量的张量来表达这个方程的右手边。 与里奇（Ricci）张量一样，这个新的张量用 **T** 表示，将两个向量作为输入，并输出一个标量。 它被称为**能量-动量张量**，或者同样常见的，***应力-能量张量***。 像里奇（Ricci）张量一样，它是对称的：

$$\boxed{\mathbf{T}(\mathbf{w},\mathbf{v}) = \mathbf{T}(\mathbf{v},\mathbf{w}) \quad \Leftrightarrow \quad T_{ki} = T_{ik}} \tag{30.15}$$

尽管看起来相反，但这不是（至少主要是！）一本物理书，所以对于 **T** 的完整描述，我们必须让读者参考 Misner、Thorne 和 Wheeler（1973 年）或 Thorne 和 Blandford（2017 年）。 对于我们的目的，**T** 的关键特征是对于具有（归一化）4-速度 **v** 的观察者，物质和能量的总密度由下式给出

$$\mathbf{T}(\mathbf{v},\mathbf{v}) = \rho_{\text{matter}} + \rho_{\text{energy}} \equiv \rho_{\text{total}}$$

回想一下，爱因斯坦发现物质和能量是等价的，所以它们都会弯曲时空。 例如，麦克斯韦发现电磁场的能量密度与 $(|\mathbf{E}|^2 + |\mathbf{B}|^2)$ 成正比，其中 $\mathbf{E}$ 是电场，$\mathbf{B}$ 是磁场。 这种电磁能量就像行星一样可靠地弯曲时空。为了获得完整的能量-动量张量，也必须加入其他场。

因此，考虑到这些相对论效应，（30.14）变为

$$\mathbf{R}\text{icci}(\mathrm{v},\mathrm{v}) = 4\pi G\rho_{\text{total}} = 4\pi\mathbf{T}(\mathbf{v},\mathbf{v}) \tag{30.16}$$

请注意，为了将来的比较，如果我们使用与以前相同的四分体，且 $\mathbf{v} = \mathbf{e}_0$，则

$$\boxed{R_{00} = 4\pi G\rho_{\text{total}}} \tag{30.17}$$

但是，如果 $\mathbf{v}$ 是一个任意的类时向量，则 (30.16) 为真。 因此，使用与之前相同的技巧，写成 $\mathbf{v} = \mathbf{x} + \mathbf{y}$ 并诉诸 $\mathbf{T}$ 的对称性，我们推导出 [练习]

$$\boxed{\mathbf{R}\text{icci} = 4\pi G\,\mathbf{T} \quad \Leftrightarrow \quad R_{ik} = 4\pi G\,T_{ik}.} \tag{30.18}$$

这实际上是爱因斯坦从 1913 年开始对他的场方程提出的原始建议之一。（当然，如果 $\mathbf{T} = 0$，那么它会简化为（物理上正确的）真空方程，（30.11）。）确实，爱因斯坦仍然相信（30.18）直到 1915 年 11 月 11 日，也就是他在 11 月 25 日发现正确的广义相对论最终场方程之前的两周。有关爱因斯坦十年来为正确方程而奋斗的更多信息，请参阅 Misner、Thorne 和 Wheeler (1973, §17.7)，或者更好的是 Pais (1982, §14)，它记录了（基本上每天）爱因斯坦在那个决定性的十一月里颠簸不断的心理活动。

识别（30.18）中固有的致命缺陷的现代方法是观察如果这个方程是真的，微分比安奇恒等式（29.17）将意味着能量不守恒！ 要完全解释这一点会让我们走得太远，但基本的数学事实在练习 14 中导出。为了正确理解物理学，我们建议你参考 Misner、Thorne 和 Wheeler (1973, §17) 或 Penrose (2005, §19.6)，或 Thorne 和 Blandford (2017, §25.8)，或 Wald (1984, §4.3)。 正如我们前面提到的，我们还应该指出，爱因斯坦最终不是这样得出真相的，因为在 1915 年他还不知道微分比安奇恒等式！ 参见 Pais (1982, p. 256)。

结果是微分比安奇恒等式几乎唯一地确定了必须对（30.18）进行的更正，以便恢复能量守恒。

为了进行这种校正，我们引入了能量-动量张量的**迹**：

$$\mathbf{T} \equiv \mathbf{T}^{\mathrm{m}}{}_{\mathrm{m}} = \rho_{\text{total}} - (P_1 + P_2 + P_3)\,,$$

其中 P_1、P_2、P_3 是我们三个正交空间方向 $\mathbf{e}_1$、$\mathbf{e}_2$、$\mathbf{e}_3$ 上物质内部的压力。 请注意，虽然压力和密度可能看起来非常不同的概念，但可以看到如下联系：

$$\text{pressure} = \frac{\text{force}}{\text{area}} = \frac{\text{force} \times \text{distance}}{\text{volume}} = \frac{\text{energy}}{\text{volume}} = \frac{\text{mass}}{\text{volume}} = \text{density}.$$

我们现在可以陈述**爱因斯坦场方程**的正确的最终形式：

$$\boxed{\mathbf{R}\text{icci} = 8\pi G\left[\mathbf{T} - \frac{1}{2}\,T_g\right] \quad \Leftrightarrow \quad R_{ik} = 8\pi G[T_{ik} - \frac{1}{2}Tg_{ik}].} \tag{30.19}$$

我们注意到，这不是你会发现几乎所有广义相对论课本中写的方程的形式。 然而，这就是爱因斯坦本人在 1915 年 11 月 25 日最初写出方程的形式！ 参见 Pais (1982, p. 256)。

让我们申明更标准的、数学上等价的形式，尽管我们将简短的证明留给上面的技术参考：

$$\boxed{\mathbf{R}\text{icci} - \frac{1}{2}Rg = 8\pi G\,\mathbf{T} \quad \Leftrightarrow \quad R_{ik} - \frac{1}{2}Rg_{ik} = 8\pi G T_{ik}} \tag{30.20}$$

其中 $R \equiv R^m{}_m$ 是 $\mathbf{R}$icci 的迹，称为**曲率标量**。 左边的张量，

$$\boxed{\mathbf{G} \equiv \mathbf{R}\text{icci} - \frac{1}{2}R\,g,} \tag{30.21}$$

称为**爱因斯坦张量**[14]。这个符号导致（到目前为止）最常见的爱因斯坦方程[15] [16]的写法：

$$\boxed{\mathbf{G} = 8\pi G\,\mathbf{T}} \tag{30.22}$$

最终的爱因斯坦场方程（30.19）看起来与原始方程（30.18）非常不同，我们自然而然地被牛顿平方反比定律的几何推理所引导。 实际上，它们差别很小。 事实上，我们现在证明，在物理上，爱因斯坦方程只对原始牛顿方程增加了一个修正，在正常情况下这个修正是非常小的。

---

[14] 有关爱因斯坦张量几何的更多信息，请参阅 Frankel (2011, §4)。

[15] 我们提醒读者，Ricci 曲率的定义各不相同，影响其符号； 因此，例如，Penrose (2005, §19.6) 改为 $\mathbf{G} = -8\pi G\mathbf{T}$。

[16] 使用几何单位也很常见，其中 c=1 和 G=1，在这种情况下，爱因斯坦方程简单地变为 $\mathbf{G} = 8\pi\mathbf{T}$。

为了看到这一点，再次选择沿着观察者的 4-速度的时间轴：$\mathbf{e}_0 = \mathbf{v}$。 然后，由于 $g_{00} = 1$，我们发现

$$R_{00} = 8\pi G[T_{00} - \frac{1}{2}Tg_{00}] = 8\pi G\left[\rho_{total} - \frac{1}{2}\{\rho_{total} - (P_1 + P_2 + P_3)\}\right]$$

所以

$$\boxed{R_{00} = 4\pi G(\rho_{total} + P_1 + P_2 + P_3).} \tag{30.23}$$

将其与我们最初的基于牛顿的方程（30.17）进行比较，我们看到爱因斯坦方程的不同之处仅在于添加了三个压力项，对于正常情况下的物质，这些项与质量/能量项 ρ_{total} 相比是很小的。

重要的是要记住里奇（Ricci） 张量的重要性，以及（30.23）的重要性——它通过 (29.29) 告诉我们时空曲率的体积压缩效应。 我们终于得到了**几何形式的爱因斯坦场方程：**

$$\boxed{\delta\ddot{\mathcal{V}} = -R_{00}\,\delta\mathcal{V} = -4\pi G\,(\rho_{total} + P_1 + P_2 + P_3)\,\delta V.} \tag{30.24}$$

30.10 引力坍塌到黑洞

在恒星的正常生命周期中，正是其内部持续不断的氢弹般的爆炸使恒星发光，并防止恒星的物质在自身重量下向内坍塌。 但恒星最终耗尽了其主要核燃料氢的供应，随着核火熄灭，恒星开始失去抵抗引力的能力。让我们简要地考虑一下普通恒星在终局开始时的命运。

在引力坍缩过程中，温度和压力会上升到新的核反应开始的程度，通过燃烧氦和更重的元素来重新对抗引力。 此外，引力坍缩的极端情况甚至会导致原子本身被压碎，此时，凭借泡利不相容原理，新的量子力学力会抵抗坍缩。

确切地说，哪些核反应和哪些量子力学力占主导地位，以及以什么顺序，是支配恒星垂死挣扎的复杂舞蹈[17]。但在很大程度上决定这条最终路径的关键信息是恒星的初始质量 M。

$M < 8\,M_\odot$ 的小恒星（如我们的太阳）最终可能成为稳定的白矮星，其中量子力学电子简并压力防止进一步坍缩。 但是更大的恒星，$M > 10M_\odot$，最终可能会变成超新星，恒星的大部分物质会爆炸到太空中，只留下一个核心，然后会发生引力坍缩。如果原始恒星的 $M <$

[17]有关这种舞蹈的迷人细节，请参阅 Schutz (2003, §12)。

$30M_{\odot}$，那么后超新星核心的坍缩可能会被中子简并压力（和其他力）阻止，从而形成稳定、快速旋转的中子星。

但是爱因斯坦的方程预测，如果坍缩的核心足够大，就会发生一些违反直觉且几乎自相矛盾的事情，从而导致黑洞的产生。

在这种引力坍缩的极端情况下，爱因斯坦方程（30.24）中的压力项会变得非常重要，因为物质被压缩并且原子的速度接近光速。 普通的物理直觉告诉我们，这些内部压力将对抗坍塌，甚至可能阻止坍塌。 但是，相反，爱因斯坦方程告诉我们，增长的压力只会增加重力的体积破碎力！ 这颗恒星越是努力抵抗它的坍塌，引力就越会加强它的束缚。

如果最终发生引力坍缩的后超新星核心的质量大于约[18] $2\ M_{\odot}$（托尔曼-奥本海默-沃尔科夫极限，Tolman–Oppenheimer–Volkoff limit，简称 TOV），这种纯粹的爱因斯坦效应会导致一个不归路：现在在宇宙中没有力可以阻止坍塌。 见[30.9]。 重力会无情地将整个核心压碎到 $r = 0$ 处的无限密度和无限潮汐力的点，称为时空奇点。 坍缩结束后，剩下的是一个纯粹的真空引力场。[19]

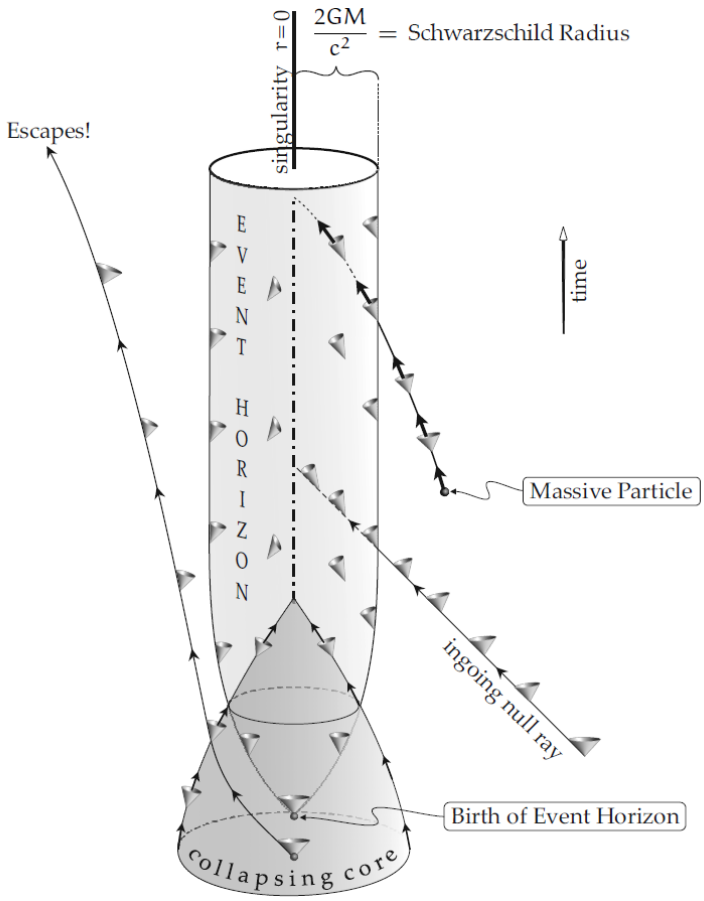
想象一下，在坍缩过程中从核心中心发出闪光，并进一步想象这些闪光可以穿过核心的物质，就好像它们是中微子一样。 单个闪光的命运取决于它何时发出。 如果闪光发出得足够早，那么，如图所示，它能够逃离引力场。

但是有一个临界时刻，来自中心的闪光首先扩散，然后减慢并最终在史瓦西半径处盘旋，$r_s = (\frac{2GM}{c^2})$。 这个悬浮的光球就是**事件视界**，它的内部是一个**黑洞**：一旦形成，任何物质或信息都无法逃离这个区域。 零锥与视界相切，因此它们允许物质和光穿入，但从不允许向外穿出，因为物质必须始终在零锥内行进。

---

[18] TOV 限值的精确值正在进行中； $2.2M_{\odot}$ 是 2019 年撰写本文时的最佳估计。

[19] Birkhoff 定理告诉我们，当球形质量 M 缩小并保持球形时，其外部真空的几何形状必然是史瓦西解，并且度量中的 M 值不会随着物体半径 R 的缩小而改变； 唯一的变化是解适用的真空区域 r>R 的大小。



[30.9] **黑洞的诞生：后超新星核心的引力坍缩**。爱因斯坦的场方程告诉我们，一个足够大质量的核心坍缩将达到一个不归路：重力将无情地将整个核心压碎到一个点，在 $r = 0$ 时具有无限的密度和无限的潮汐力，称为时空奇点；剩下的是一个纯粹的真空引力场。如果足够早地从坍缩核心的中心发出闪光，它就可以逃离引力场。但是有一个临界时刻，来自中心的闪光首先扩大，然后减慢并最终在史瓦西半径 $rs = (2GM / c^2)$ 处盘旋。这个悬浮的光球就是**事件视界**，它的内部是一个**黑洞**：一旦形成，任何物质或信息都无法逃离这个区域。零锥与视界相切，因此它们允许物质和光向内穿入，但从不向外穿出，因为物质必须始终在零锥内行进。

因此，一旦物体落入黑洞，它就会被无情地拖向中心奇点，施加在其上的潮汐力与 $(1/r^3)$ 成正比增长。 任何固态物最终都必然屈服于这些快速增长的潮汐力，在径向方向上被拉伸，在垂直方向上被压缩。 由于潮汐力趋于无穷大，这个过程会将物体拉伸/压缩成细长的形状，就像一块意大利面条。 出于这个原因，物理学家实际上将这个过程称为**意大利面条化**！

黑洞也可以通过其他方式形成。 我们已经提到，2017 年探测到的引力波（GW170817）起源于两颗中子星盘旋并彼此接近，最终碰撞并融合。 合并后，最终物体的质量超过了 TOV 极限。 值得注意的是，引力波科学家能够在空间上将源定位到南部天空的一小块区域，传统的 X 射线天文学家随后能够在那里观察，成功地识别和研究了中子星碰撞的后果。 这个 X 射线数据现在似乎证实了合并的中子星立即完全坍缩形成了一个黑洞！

我们现在知道，宇宙中到处都是这些类型的黑洞，每个都有几个太阳质量。 然而，此

外，天文学家现在已经证实，几乎每个星系（包括我们自己的银河系）的中心都有一个完全不同的黑洞——一个比我们的太阳质量大数百万甚至数兆亿倍的**超大质量黑洞！** 尽管天体物理学家假设了它们可能形成的不同机制，但超大质量黑洞的真正起源仍然是个谜。

如前所述，如果一个线性尺寸为 l 的天体穿过质量为 M 的黑洞的事件视界，则施加在其上的潮汐力为 $(c^6 l/G^2)\,(1/M^2)$，随着 M 变大而变小。 对于一个 $M = 10^8\,M_\odot$ 的超大质量黑洞来说，这个力远小于我们在地球表面所经历的潮汐力！ 因此，如果你让你的宇宙飞船飞入这样的黑洞，你甚至不会意识到你已经这样做了——当你穿过事件视界时，你什么都感觉不到。

然而，不管你的宇宙飞船引擎的功率如何（以及你如何尝试部署它们），在你穿越视界之后，你到达奇点的最长时间与黑洞的质量成正比，并且由下式给出[20]

$$t_{max} = \frac{M}{M_\odot} \times 15.5 \times 10^{-6}\ \text{秒}$$

如果你“足够幸运”将你的宇宙飞船飞入一个质量足够大的黑洞——几个星系的质量就够了！——那么你甚至可以在黑洞中生活多年； 请记住，你在洞外的朋友仍然可以将护理包送入洞中。

然而，对于更真实的超大质量黑洞，前景更加严峻。 例如，我们银河系中心的黑洞（称为人马座 A*）的估计质量[21] $M = 3.6 \times 10^6\,M_\odot$，所以如果你要越过它的视界，你的最长生存时间将不到一分钟！

不管 t_{max} 如何，一旦你进入黑洞，就无法逃脱——可悲的是，你最终的、完全不可避免的宿命是意大利面化！

30.11 宇宙常数：“我一生中最大的错误。”

我们以爱因斯坦最快乐的想法开始这一章，我们将以他最不快乐的想法结束这一章。

但是这部看似悲伤的电影在最后却突然出现了一个奇怪的、超级快乐的情节转折！ 的确，当大幕落下，你终于走出第四幕的出口时，你可能会自言自语，“这在现实生活中是不可能发生的！”

1916 年，普遍的信念（更多地基于哲学而非科学）是宇宙是静止的、不变的。 但是爱因斯坦的场方程很快被发现与这个想法相冲突——它自动导致了一个正在膨胀或收缩的宇宙。

---

[20] 参见 Taylor 和 Wheeler (2000, §3–21)。

[21] 2018 年的测量值在早先估计的 $M = 2.6 \times 10^6\,M_\odot$ 基础上向上做了修正。

请注意，此时，“宇宙”被认为仅由我们的银河系组成——埃德温·哈勃于 1924 年首次发现了我们自身之外的星系。

为了挽救一个静止的宇宙[22]，爱因斯坦采用了一种大胆而绝望的策略——尽管这是十年辛勤劳动的成果，他改变了他的场方程，（30.22）!

我们已经注意到，微分比安奇恒等式（29.17）几乎唯一地确定了必须对爱因斯坦的原始方程（30.18）进行的更改才能恢复能量守恒。 然而，爱因斯坦意识到他的方程中还有一个，而且只有一个额外的自由度：他可以在左侧添加度量张量 g 的恒定倍数 Λ。

爱因斯坦在 1917 年对他的方程进行了这种修改，为了将这个新方程与 1915 年的原始方程区分开来，我们将其称为

宇宙爱因斯坦场方程：
$$\boxed{\mathbf{G} + \Lambda g = 8\pi G\,\mathbf{T}} \tag{30.25}$$

如果为真，这个常数 Λ（称为宇宙常数）将需要非常小，以避免与观测证据发生冲突。在太阳系甚至银河系的尺度上，新的宇宙场方程将做出与原始场方程相同的预测。 只有在宇宙学尺度上，它的预测才会截然不同。

1929 年，也就是爱因斯坦对他的方程做出这一改变的十二年后，埃德温·哈勃发现了第二个非凡的发现。 在他早先发现我们之外的星系的基础上，他试图测量它们与我们的距离，然后将这些距离与星系的径向速度（使用它们的红移测量）相关联。 他发现所有星系的速度都指向远离我们，而且这些星系的速度与它们与我们的距离成正比。 这现在被称为**哈勃定律**。

因此，宇宙毕竟不是静止的——它在膨胀！ 爱因斯坦意识到，如果他只忠实于 1915 年的原始方程式，他本可以做出人类历史上最非凡的科学预测！ 他告诉乔治·伽莫夫，他引入宇宙常数是“我一生中最大的错误”。 此后，爱因斯坦撤回了宇宙场方程以支持原来的。

在随后的几十年中，许多（大多数？）专家倾向于支持爱因斯坦的撤回，认为 Λ 完全为零，并且更优雅的原始爱因斯坦方程是正确的。 然而，一旦 Λ 精灵从瓶子里拿出来，就不可能再放回去了。因此，研究论文总是倾向于对冲他们的赌注，至少研究如果 $\Lambda \neq 0$，他们的结论可能会受到怎样的影响。

我们现在跨越时间到 1998 年，与 70 年前哈勃发现一样伟大的宇宙学发现[23]。 尽管哈勃发现宇宙正在膨胀，但人们一直期望内部物质的正向引力会随着时间的推移减缓宇宙的膨胀。 然而，通过观察 **Ia 型超新星**（type Ia supernovas）——它们正在爆炸的白矮星已经超过了它们的稳定性极限，因此它们具有相似的质量和标准化的内在光度——科学家们惊人地

[22]爱因斯坦动机的完整故事要复杂和微妙得多。 参见 Pais (1982, §15c)。

[23]这一发现由两个独立的研究团队共同完成，2011 年诺贝尔物理学奖被瓜分，一半授予 Saul Perlmutter，另一半共同授予 Brian P. Schmidt 和 Adam G. Riess。

发现宇宙的膨胀正在加速！

但是我们如何解释这个令人震惊的发现呢？！ 自 1917 年以来，爱因斯坦的答案就一直待在那里，耐心等待！

如果 $\Lambda > 0$，那么（将 Λg 移到等式的右侧）其效果在数学上等效于虚构的[24]负能量密度 $\rho = -(\Lambda/8\pi G)$，从而导致引力排斥。 虽然 [30.8] 说明了普通物质的吸引性质和正截面曲率，但在存在负能量和负截面曲率时的相反行为在 [29.11a]，第 301 页中说明。当被视为膨胀的时空图片时 在宇宙中，这显示了由正值 Λ 引起的负能量的排斥效应，显然导致宇宙加速膨胀。

简而言之，观测到的宇宙加速膨胀与宇宙学爱因斯坦场方程（详细）一致，$\Lambda > 0$。 那么，这就是令人愉快的、难以置信的最终情节转折：爱因斯坦的“最大错误”结果却是他最伟大的天才之笔之一！

30.12 结束

我们希望第四幕的最后两章能激发你更深入地了解黎曼曲率和爱因斯坦弯曲时空的相关奇迹。 然而，我们必须停在这里——完整的解释需要第 2 卷和第 3 卷！

如果我们已经完成了我们的工作，你就迫不及待地想了解更多——让本书末尾的延伸阅读部分成为你的指南。 在那里，我们推荐了许多优秀的作品，每件作品都有自己独特的特色。

话虽如此，让我们不耐烦地断言，我们想不出比彭罗斯的《通往实在之路》（2005）更顺畅或更快的启蒙之路，因为它包含了他对数学和物理学的无数原创性和启示性的见解，都以惊人的清晰度表达 通过他异常美丽的手绘。

对于更详细的技术细节，我们也强烈推荐——参见 Misner、Thorne 和 Wheeler 的 Prologue——Gravitation (1973)。对于更多最新的发展，特别是在引力波和现代宇宙学方面，我们强烈推荐 Thorne 和 Blandford（2017 年）的大规模权威著作《现代经典物理学》。

最后，在阅读第五幕之后，我们推荐 Dray (2015)，它使用形式（Forms）来阐明广义相对论的几何。 事实上，我们正在转向形式，因为它们是我们最后一幕的主题。

但是，至于我们对纯可视化微分几何的处理，终于结束了。

[24]虽然我个人满足于接受 $\Lambda > 0$ 的宇宙场方程，但许多著名科学家反而认为存在一种真实的、非虚构的东西，渗透到整个宇宙中，导致了这种排斥——他们称之为暗能量。

第三十一章 第四幕练习

外在构造

1. **球体上的测地线曲率**。在半径为 R 的球面上，假设一个粒子以单位速度沿纬度 ϕ 的圆运动。
   (i) 粒子加速度的大小和方向是什么？
   (ii) 画出该加速度在球体切面上的投影，并推断测地曲率是 $\kappa_g = \frac{\cot\phi}{R}$。
   (iii) 验证 κ_g 的这个公式产生的几何正确答案为 $\phi \to 0$ 和 $\phi \to (\pi/2)$。
   (iv) 证明指向球心的加速度分量的大小为 $(1/R)$，与 ϕ 无关。几何地解释这个。

2. **圆锥上的测地线曲率**。重新考虑在第 145 页 [14.3] 中所示的半垂直角 α 的圆锥。假设一个粒子以单位速度沿该圆锥上半径为 r 的水平圆移动。
   (i) 粒子加速度的大小和方向是什么？
   (ii) 画出该加速度在圆锥切面上的投影，并推导出测地曲率是 $\kappa_g = \frac{\sin\alpha}{r}$ 。
   (iii) 令 s 为圆锥内（沿测地线发生器）从顶点到圆的距离。沿发电机切割圆锥并将其平压在平面上，如 [14.3]，第 145 页，这样圆在半径为 s 的平面内变成圆弧，因此在半径内的曲率 $(1/s)$飞机。验证该曲率是否与 κ_g 相同。
   （注意：从内在的角度来看，这个圆的半径是多少？好吧，如果圆的相邻点的两只蚂蚁开始沿着半径向内行走——沿着与圆正交的测地线发生器——它们将在顶点相遇，所以固有半径就是顶点到圆的距离 s。因此，圆曲率的自然固有定义再次是 $(1/s)$。）
   (iv) 设 ρ 为沿法线从粒子到对称轴的距离。证明加速度沿法线的分量是$(1/\rho)$，并用几何学解释。

3. **沿接触曲面的测地线曲率**。 假设两个曲面沿公共曲线 $\mathcal{C}$ 相接触。
   (i) 如果 $\mathcal{C}$ 是一个曲面的测地线，从几何上解释为什么它也必须是另一个曲面的测地线。
   (ii) 如果 $\mathcal{C}$ 的一个曲面具有测地曲率 κ_g，请从几何上解释为什么它也必须具有另一个曲面的测地曲率 κ_g。
   (iii) 将 (i) 推导出为 (ii) 的特例。
   (iv) 使用前两次练习中 κ_g 的公式来确认 (ii) 在一个曲面是球体，另一个曲面是圆锥体的情况下，$\mathcal{C}$ 是球体上的一个纬度圆。

内蕴构造

4. **通过内蕴导数的测地线曲率**。 设 $\mathbf{v}$ 是一个粒子以单位速率沿着轨迹 K 在一些可剥皮水果的表面 $\mathcal{S}$ 上移动的速度向量。 如果 o 是发射点，令 $(\mathbf{e}_1, \mathbf{e}_2)$ 为 T_o 的正交基。 现在沿 K 平行移动 $\mathbf{e}_1$，并让 $\theta_\parallel$ 表示这个平行移动的 $\mathbf{e}_1$ 和 $\mathbf{v}$ 之间的角度。
   (i) 解释为什么 $\mathbf{e}_2$ 也自动沿 K 平行移动：

$$D_{\mathbf{v}}\mathbf{e}_1 = 0 \quad \Longrightarrow \quad D_{\mathbf{v}}\mathbf{e}_2 = 0$$

   (ii) 令 $\mathbf{v}_\perp$ 为在 T_p 内将 $\mathbf{v}$ 旋转 $(\pi/2)$ 得到的单位向量。按照 $(\mathbf{e}_1, \mathbf{e}_2)$ 写出 $\mathbf{v}$ 和 $\mathbf{v}_\perp$。 现在通过计算证明

$$\kappa_g = D_{\mathbf{v}}\mathbf{v} = [D_{\mathbf{v}}\theta_\parallel]\mathbf{v}_\perp$$

   从而证明 (23.4)：$|\kappa_g| = |D_{\mathbf{v}}\theta_\parallel|$。
   (iii) 在从 $\mathcal{S}$ 中剥离围绕 K 的窄条并平压到平面上之后，在沿 K 的几个点处画草图说明 $\mathbf{v}$、$(\mathbf{e}_1, \mathbf{e}_2)$ 和 $\theta_\parallel$ 。
   (iv) 推论 K 的测地曲率确实是将 K 周围的窄带压平得到平面曲线的普通曲率。

和乐

5. **圆锥和球体的和乐**。
   (i) 重新考虑在 [14.3] 中所示的半垂直角 α 的圆锥。在那里我们表明，如果圆锥的尖端是钝的，它的球面像允许我们将确定的曲率 K 分配给原本固有的扁锥，由 (14.2) 给出：

$$\mathcal{K}(\text{尖刺}) = \beta = \text{尖刺铺平后的分割角} = 2\pi(1 - \sin\alpha)$$

   通过在 [14.3] 中的扁平圆锥内进行平行移动，表明和乐将相同的总曲率分配给尖刺。

   (ii) 在 $\mathbb{S}^2$ 上，让我们使用外在的马铃薯剥皮机平行移动来找到一个固定角度 ϕ 的纬度圆的和乐。想象一个半垂直角 α 的圆锥体放在球体上，使其沿着这个圆接触球体。因此，沿该圆的球体剥离带（最终）与接触锥体的带相同。证明$\alpha = \frac{\pi}{2} - \phi$，并使用 (i) 来确认极冠内的总曲率确实是围绕它的纬度圈的和乐。

   (iii) 在 $\mathbb{S}^2$ 的赤道上，想象一个指向正东的向量。现在将其沿测地赤道平行向东传输，使其返回家中看似不变，即消失的全息。但是这个环包围了球体

的一半，总曲率为 2π！使用 (ii) 通过将 ϕ 从 0 逐渐增加到 $(\pi/2)$ 来调和这些事实。

6. **一般局部高斯-博内定理**。

(i) 通过用测地线 m 边形逼近一个光滑的闭环 L，然后让 $m \to 1$，推导出和乐公式（24.6）变为，

$$\mathcal{R}(L) = 2\pi - \oint_L \kappa_g ds$$

其中 κ_g 是沿 L 的测地曲率，s 是沿 L 的距离。

(ii) 如果 P 是一个外角为 φ_i 的封闭“多边形”，但边不是测地线（即 $\kappa_g \neq 0$），则证明 (24.6) 的推广是，

$$\boxed{\mathcal{R}(P) = 2\pi - \left[\oint_L \kappa_g ds + \sum_i \varphi_i\right]}$$

(iii) 如果 R 表示 P 的内部，则推导出**一般局部高斯-博内定理**：

$$\boxed{\iint_R \mathcal{K} d\mathcal{A} = 2\pi - \left[\oint_L \kappa_g ds + \sum_i \varphi_i\right]}$$

曲率作为相邻测地线之间的力

7. **明丁定理**。 使用与第 28.3 节中相同的测地极坐标，度量采用以下形式

$$ds^2 = dr^2 + g^2(r)\, d\theta^2$$

以便

$$g'' = -\mathcal{K} g$$

利用 $g(r) \simeq r$（当 r 趋于零）这一事实，在常数 $\mathcal{K}$ 的三种情况求解此微分方程：

(i) 如果 $\mathcal{K} = 0$ 处处，则推论曲面与欧几里得平面局部等距。

(ii) 如果 $\mathcal{K} = (1/R^2)$ 在整个曲面上是常数，则推论该空间与半径为 R 的球体局部等距。

(iii) 如果 $\mathcal{K} = -(1/R^2)$ 在整个曲面上是恒定的，则推断该空间与半径为 R 的伪球面局部等距。

8. **一般旋转曲面上的雅可比方程**。 如第 89 页 练习 22，想象一个粒子以单位速率沿 (x,y) 平面中的曲线行进，并设其在时间 t 的位置为 $[x(t),y(t)]$。 现在想象一下这个平面绕 x 轴旋转 θ 角。 当 θ 从 0 变化到 2π 时，曲线扫过一个旋转曲面。

(i) 解释为什么$\dot{x}^2 + \dot{y}^2 = 1$，其中的点代表时间导数。
(ii) 从几何上证明曲面的度量是 $ds^2 = dt^2 + y^2\, d\theta^2$。
(iii) 通过考虑两个相邻子午线测地线的相对加速度，从雅可比方程 (28.2) 推导出 $\mathcal{K} = -\ddot{y}/y$。（回想一下，这是我们之前在证明旋转曲面的*绝妙定理*过程中获得的相同的公式 (13.4)。）

9. **通过计算导出高斯引理**。如[28.6]，考虑从 o 发射的两个相邻的单位速率测地线，角度间隔为 δθ，并让 $\mathbf{v}$ 表示它们的单位速度向量。 令 $\boldsymbol{\xi}$ 为测地线之间的连接向量，连接距 o 相同距离 $\sigma = t$ 的两个点，因此两者都位于测地线圆 $K(\sigma)$ 上。 然后，当 $\delta\theta \to 0$ 时，$\boldsymbol{\xi}$ 与 $K(\sigma)$ 相切。 为了建立高斯引理 (28.4)，我们必须证明 $\mathbf{v}\cdot\boldsymbol{\xi} = 0$。 （注：在下文中，∇v 是普通的 $\mathbb{R}^3$ 导数，不是本征曲面导数，$D_{\mathbf{v}} = \boldsymbol{\nabla}_{\mathbf{v}}$。）

(i) 解释为什么 $\lim_{\sigma\to 0} \mathbf{v}\cdot\boldsymbol{\xi} = 0$。
(ii) 推导出证明 $\mathbf{v}\cdot\boldsymbol{\xi} = 0$，只需证明$\boldsymbol{\nabla}_{\mathbf{v}}\,[\mathbf{v}\cdot\boldsymbol{\xi}] = 0$。
(iii) 解释为什么 $\mathbf{v}\cdot\nabla_{\boldsymbol{\xi}}\mathbf{v} = 0$。
(iv) 解释为什么 $[\mathbf{v},\boldsymbol{\xi}] = 0$。
(v) 利用 $\mathbf{v}$ 是测地线的速度这一事实，推导出 $\boldsymbol{\xi}\cdot\boldsymbol{\nabla}_{\mathbf{v}}\mathbf{v} = 0$。
(vi) 结合前面的三个结果证明

$$\boldsymbol{\nabla}_{\mathbf{v}}\,[\mathbf{v}\cdot\boldsymbol{\xi}] = 0$$

从而完成了高斯引理的计算证明。

黎曼曲率

10. **黎曼张量的两个对称性**。

(i) 让我们确认第一（代数）比安基恒等式 (29.15)，也称为**比安基对称**：

$$\mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w} + \mathcal{R}(\mathbf{v},\mathbf{w})\mathbf{u} + \mathcal{R}(\mathbf{w},\mathbf{u})\mathbf{v} = 0 \quad \Leftrightarrow \quad R_{ijkm} + R_{jkim} + R_{kijm} = 0$$

为了证明一般结果，在所有三个矢量场都是坐标矢量场的情况下（通过线性）

证明它实际上就足够了，在这种情况下它们的交换子都消失了。 证明在这种情况下，

$$\mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w}+\mathcal{R}(\mathbf{v},\mathbf{w})\mathbf{u}+\mathcal{R}(\mathbf{w},\mathbf{u})\mathbf{v}=\boldsymbol{\nabla}_{\mathbf{u}}[\mathbf{v},\mathbf{w}]+\boldsymbol{\nabla}_{\mathbf{v}}[\mathbf{w},\mathbf{u}]+\boldsymbol{\nabla}_{\mathbf{w}}[\mathbf{u},\mathbf{v}]=0$$

使用曲率 2-形式的更优雅的证明可以在第 38.12.4 节中找到。

(ii) 使用与 (i) 中相同的交换向量场，我们可以对比安基对称性进行几何解释，如下所示。 （注意：我当然不相信我是第一个发现这个证明的人，但我无法在印刷品中找到它，因此不知道应该归功于谁。）从某个点开始，画 $\epsilon\mathbf{u}$， $\epsilon\mathbf{v}$, $\epsilon\mathbf{w}$，其中 ϵ 很小并且最终消失。通过完成每对这些边开始构建一个多面体“盒子”，创建三个平行四边形面，这些面由于它们的消失换向器而闭合。 现在构造（并画出）这个盒子的两个新的向量边 $\mathbf{A}$ 和 $\mathbf{B}$，通过首先沿着 $\epsilon\mathbf{v}$ 然后沿着 $\epsilon\mathbf{u}$ 平行移动 $\epsilon\mathbf{w}$（产生 $\mathbf{A}$），并且以相反的顺序，首先沿着 $\epsilon\mathbf{u}$ 然后沿着 $\epsilon\mathbf{v}$（产生 乙）。 接下来将 $\mathbf{B}$ 的末端连接到 $\mathbf{A}$ 的末端，从而创建盒子新三角形面的新矢量边。 推断这个新边由下式给出，

$$\mathbf{A}-\mathbf{B}\ \asymp\ \epsilon^3\,\mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w}$$

使用其他两个面重复此构造，从而构造盒子的另外两个三角形面。 你的图现在应该显示 $\epsilon^3\,\mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w}$, $\epsilon^3\,\mathcal{R}(\mathbf{v},\mathbf{w})\mathbf{u}$ 和 $\epsilon^3\,\mathcal{R}(\mathbf{w},\mathbf{u})\mathbf{v}$ 形成了闭合盒子的第四个三角形面的向量边，从而证明比安奇对称。

(iii) 让我们确认黎曼张量在第一对和第二对向量的交换下是对称的，(29.16)：

$$[\mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{x}]\cdot\mathbf{y}=[\mathcal{R}(\mathbf{x},\mathbf{y})\mathbf{u}]\cdot\mathbf{v}\quad\Leftrightarrow\quad \mathrm{R}_{\mathrm{ijkl}}=\mathrm{R}_{\mathrm{klij}}$$

在撰写本文时，我还没有对这个结果做出几何意义，所以我们现在必须求助于通过撒旦代数（见序言）来证明它。 如果我们定义

$$\mathbf{B}(\mathbf{u},\mathbf{v},\mathbf{x},\mathbf{y})\equiv[\mathcal{R}(\mathbf{u},\mathbf{v})\,\mathbf{x}]\cdot\mathbf{y}+[\mathcal{R}(\mathbf{v},\mathbf{x})\,\mathbf{u}]\cdot\mathbf{y}+[\mathcal{R}(\mathbf{x},\mathbf{u})\,\mathbf{v}]\cdot\mathbf{y}$$

所以

$$\mathbf{B}(\mathbf{u},\mathbf{v},\mathbf{x},\mathbf{y})=0$$

凭借 (i) 中的比安基对称性。 回顾黎曼张量在前两个时隙和后两个时隙中都是反对称的，通过从以下（显然是微不足道的!）恒等式中取消相似项来证明结果：

$$\mathbf{B}(\mathbf{u},\mathbf{v},\mathbf{x},\mathbf{y})+\mathbf{B}(\mathbf{v},\mathbf{x},\mathbf{y},\mathbf{u})=0=\mathbf{B}(\mathbf{x},\mathbf{y},\mathbf{u},\mathbf{v})+\mathbf{B}(\mathbf{y},\mathbf{u},\mathbf{v},\mathbf{x})$$

（你能闻到硫磺的味道吗？）

11. **计算黎曼张量的分量**。（我们在 Lightman 等人（1975）中发现的以下证明比我们见过的其他标准证明更短更简单。）

(i) 鉴于在 n-流形中黎曼张量的分量 R_{ijkl} 在 ij 和 kl 中是反对称的，推断有 $P = \frac{1}{2}n(n-1)$ 种非平凡的选择对 ij 的方法，同样有 P 种选择对 kl 的方法。

(ii) 给定第一对和第二对索引交换下的对称性 (29.16)，$R_{ijkl} = R_{klij}$（参见前面的练习），推断如果我们只考虑这些对对称性，则有 $\frac{1}{2}P(P+1)$选择ijkl的独立方式。

(iii) 定义 $B_{ijkl} \equiv R_{ijkl} + R_{iklj} + R_{iljk}$，如在前面的练习中，比安基对称性（29.15）声明 $B_{ijkl} = 0$（参见前面的练习）。验证对对称性现在确保 B_{ijkl} 在所有四个索引上都是完全反对称的，因此约束 $B_{ijkl} = 0$ 很容易满足，除非所有四个索引都是不同的。

(iv) 如果 $n < 4$，比安基对称因此不会施加任何新的约束。使用 (ii) 推导出如果 $n = 2$ 黎曼张量有一个分量（高斯曲率），如果 $n = 3$ 它有六个分量。

(v) 如果$n > 4$，推断由比安基对称性产生的附加约束的数量等于从 n 个对象中选择 4 个对象的方式的数量。

(vi) 推导出黎曼张量的独立分量的个数为

$$\frac{1}{2}P(P+1) - \frac{n!}{(n-4)!\,4!}$$

(vii) 验证这确实产生公式 (29.1)： 黎曼张量的独立分量的数量是

$$\frac{1}{2}n^2(n^2-1)$$

12. **指数运算符和曲率**。 回忆一下指数级数：

$$\exp(x) = e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots$$

现在将其与一般函数 f(x) 的泰勒级数进行比较：

$$\begin{aligned} f(a+\delta x) &= f + \delta x\frac{df}{dx} + \frac{1}{2!}(\delta x)^2\frac{d^2f}{dx^2} + \frac{1}{3!}(\delta x)^3\frac{d^3f}{dx^3} + \cdots\Big|_a \\ &= \left[1 + \delta x\frac{d}{dx} + \frac{1}{2!}(\delta x)^2\left(\frac{d}{dx}\right)^2 + \frac{1}{3!}(\delta x)^3\left(\frac{d}{dx}\right)^3 + \cdots\right] f\Big|_a \\ &\equiv \exp\left[\delta x\frac{d}{dx}\right] f\Big|_a \end{aligned}$$

其中最后一行作为**指数算子**的定义。 向量场的扩展自然而然。 如果 **w** 是定义在

a 附近的向量场，那么当我们在单位向量 $\mathbf{u}$ 的方向上移动距离 δu 时，

$$\mathbf{w}(a + \mathbf{u}\,\delta u) = \exp[\delta u \boldsymbol{\nabla}\mathbf{u}]\mathbf{w}|_a$$

(i) 重新考虑 [29.8] 中黎曼张量的推导，但为了简单起见，假设 $[\mathbf{u}, \mathbf{v}] = 0$，使得平行四边形闭合，在这种情况下，曲率算子简化为

$$\mathcal{R}(\mathbf{u}, \mathbf{v}) = [\boldsymbol{\nabla}_{\mathbf{u}}, \boldsymbol{\nabla}_{\mathbf{v}}]$$

解释为什么平行四边形的向量完整由下式给出

$$-\delta\mathbf{w}_{\parallel} = [\exp(\delta u \boldsymbol{\nabla}_{\mathbf{u}}), \exp(\delta v \boldsymbol{\nabla}_{\mathbf{v}})]\mathbf{w}$$

(ii) 推断

$$-\delta\mathbf{w}_{\parallel} = (\delta u\,\delta v)\mathcal{R}(\mathbf{u}, \mathbf{v})\mathbf{w} + (\text{三阶误差})$$

其中“三阶误差”由涉及 $(\delta u)^p (\delta v)^q$ 的项组成，其中 $(p + q) > 3$。

爱因斯坦弯曲时空

13. **日食和潮汐**。 月球对海洋的潮汐影响是太阳的两倍多，尽管事实上太阳对地球的引力是月球的 200 倍。 让我们试着理解这个看似矛盾的逆转。

(i) 日全食的发生是一个了不起的事实。 这仅仅是因为经验上的巧合：月球和太阳看起来几乎与从地球上看到的（角度）大小完全相同。 如果月球和太阳的半径分别为 r_m 和 r_s，它们与地球的距离分别为 R_m 和 R_s，则推导出 $(R_m/R_s) \approx (r_m/r_s)$。

(ii) 我们已经看到，距离地球 R 的质量为 M 的物体对海洋施加的潮汐力与 $\frac{M}{R^3}$ 成正比。 通过诉诸 (i)，推导出月球与太阳潮汐力的比率是月球与太阳的密度之比。

(iii) (iii) 月球的平均密度约为每立方米 3300 公斤，而太阳的平均密度约为每立方米 1400 公斤。 使用 (ii) 来解释奇怪的开场事实！

14. **爱因斯坦张量守恒**。（注意：除非你已经熟悉张量收缩，以及提高/降低张量指数，否则最好推迟这个练习，直到你学习了第五幕的相关部分，即 33.7 和 33.8。）回忆 (30.21) 爱因斯坦张量 **G** 是

$$\mathbf{G} \equiv \mathbf{R}\text{icci} - \frac{1}{2}\mathrm{R}\,\mathrm{g}$$

这个练习证明了爱因斯坦张量是“守恒的”，就像能量-动量**守恒**一样：

$$\boxed{\boldsymbol{\nabla}^{a}\mathrm{G}_{ab} = 0}$$

爱因斯坦在 1915 年 11 月 25 日写下他的场方程（30.19）时，并不知道这个至关重要的纯数学事实（爱因斯坦的原始形式在数学上等同于现代形式，$\mathrm{G} = 8\pi\mathrm{T}$）。随后，人们认识到爱因斯坦发现的几何与物质之间的联系实际上意味着能量-动量必须是守恒的：$\boldsymbol{\nabla}^{a}\mathrm{T}_{ab} = 0$！

(i) 检查第二个（微分）比安基恒等式 (29.17) 是否可以写成

$$\boldsymbol{\nabla}_{a}\mathrm{R}_{bcd}{}^{e} + \boldsymbol{\nabla}_{b}\mathrm{R}_{cad}{}^{e} + \boldsymbol{\nabla}_{c}\mathrm{R}_{abd}{}^{e} = 0$$

(ii) 现在执行这个方程的两个收缩，通过 (1) 提高 a 并将 d 重命名为 a，以便将两个 a 相加； (2) 将 e 重命名为 c，使两个 c 相加。 确认这会产生

$$\boldsymbol{\nabla}^{a}\mathrm{R}_{bca}{}^{c} + \boldsymbol{\nabla}_{b}\mathrm{R}_{c}{}^{a}{}_{a}{}^{c} + \boldsymbol{\nabla}_{c}\mathrm{R}^{a}{}_{ba}{}^{c} = 0$$

(iii) 证明前面的方程可以改写为

$$-\boldsymbol{\nabla}^{a}\mathrm{R}_{ba} + \boldsymbol{\nabla}_{b}\mathrm{R} - \boldsymbol{\nabla}_{c}\mathrm{R}_{b}{}^{c} = 0$$

(iv) 证明这可以反过来重写为

$$\boldsymbol{\nabla}^{a}\mathrm{R}_{ba} - \frac{1}{2}\boldsymbol{\nabla}_{b}\mathrm{R} = 0$$

(v) 推导出 $\boldsymbol{\nabla}^{a}\mathrm{G}_{ab} = 0$，如将要显示的那样。

15. **外尔曲率**。 黎曼张量在时空中的 20 个分量可以分为里奇张量的 10 个分量（通过爱因斯坦场方程由物质和能量产生）和代表真空中纯引力自由度的 10 个分量。这些引力自由度完全由**外尔曲率张量**编码：

$$\boxed{\mathrm{C}_{ij}{}^{kl} \equiv \mathrm{R}_{ij}{}^{kl} - 2\mathrm{R}_{[i}{}^{[k}\,\mathrm{g}_{j]}{}^{l]} + \frac{1}{3}\,\mathrm{R}\,\mathrm{g}_{[i}{}^{k}\,\mathrm{g}_{j]}{}^{l}}$$

在这里，包含成对索引的方括号表示反对称操作，如稍后在第 369 页的 (33.9) 中定义的那样。

(i) 显式执行这些反对称，从而推导出以下笨拙（且难以记住）的公式：

$$C_{ijkl} = R_{ijkl} + \frac{1}{2}(R_{il}\, g_{jk} + R_{jk}\, g_{il} - R_{ik}\, g_{jl} - R_{jl}\, g_{ik}) + \frac{1}{6} R\,(g_{ik}\, g_{jl} - g_{jk}\, g_{il})$$

(ii) 从 (i) 推导出外尔张量与全黎曼张量具有相同的对称性。

(iii) 验证爱因斯坦场方程立即暗示，在真空中，外尔张量是黎曼张量。

(iv) 证明外尔张量的所有迹都为零：特别是，

$$\boxed{C_{ij}{}^{ki} = 0}$$

因此，里奇张量的等价物——黎曼张量的物质-能量部分——对于外尔张量为零。

(v) （注意：本练习的其余部分更高级，并且确实需要比我们迄今为止解释的更多。解决方案可以在本练习结束时引用的参考文献中找到。）假设我们对时空度量进行**共形变换** ：

$$g \longrightarrow \Omega^2\, g$$

其中 Ω 是一个在整个时空中随点变化的函数。 但是，在每一点，所有局部距离都被拉伸相同的因子 Ω，在该点进行评估。 这种变换保留了角度，也保留了小物体的形状。 黎曼张量在这种共形变换下以极其复杂的方式进行变换。然而，证明外尔张量——纯粹的引力/真空部分——具有它仅缩放的显著且至关重要的属性：

$$\boxed{C_{ijkl} \longrightarrow \Omega^2 C_{ijkl}.}$$

(vi) 如果我们提出一个索引，表明我们可以将其改写为 $C_{ijk}{}^{l}$ 是**共形不变量**：

$$\boxed{C_{ijk}{}^{l} \longrightarrow C_{ijk}{}^{l}}$$

注意：外尔张量最自然和优雅的表达是根据彭罗斯的 2-旋量形式，它采用外尔共形旋量 Ψ_{ABCD} 的形式，它是完全对称且共形不变的：

$$\boxed{\Psi_{ABCD} = \Psi_{(ABCD)} \rightarrow \Psi_{ABCD}}$$

Ψ_{ABCD} 也简称为**引力旋量**。 参见 Penrose 和 Rindler (1984, §4.6, §6.8) 或 Wald (1984, §13.2)。 有关外尔曲率的直观讨论，以及彭罗斯关于其在表征大爆炸异常特殊性质方面至关重要的猜想，请参见 Penrose (2005, §19.7, §28.8)。

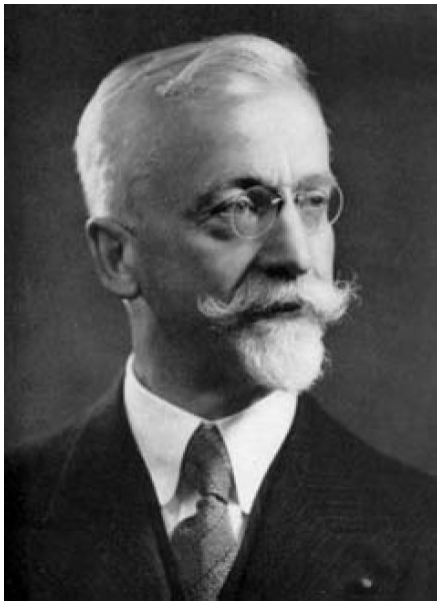
第三十二章 1-形式

32.1 引言

第五幕从严格几何规则的连续四幕中释放出来。

正如前言所预言的，我们现在的目标是建造“魔鬼机器”，并提供给本科生使用。我们说的是一种强大而优雅的计算方法，它的全名是**微分形式的外微分**（The Exterior Calculus of Differential Forms）， 我们在此将其简写为**形式**（Forms）。

大约在一个多世纪前的 1900 年左右，非常深刻、独创和广博的埃利·嘉当（如图 [32.1]）发现了形式，但即使嘉当本人也又花了 40 年的时间来驾役该发现的全部力量。



[32.1] 埃利·嘉当 (1869 – 1951)。

我们在第五幕中的目标是——简洁明了地直面盎格鲁撒克逊人——一个长达一个世纪的丑闻，即绝大多数本科生（数学和物理专业）将在没有见过嘉当之形式的情况下获得学位。

尽管我们的主要目的是开发一种新的计算方法，但我们也希望鱼和熊掌两者兼得。 也就是说，我们将努力为嘉当的形式提供尽可能多的几何之光——比标准的[1]处理方式要多得多，标准的处理经常进入一个顽固形式化的、极度抽象的平行宇宙，在其中形式的意义已经被完全剥夺 ！

---

1 Misner、Thorne 和 Wheeler (1973) 和 Schutz (1980) 是两个极好的例外，我们强烈建议将它们作为第五幕的配套（或后续）阅读。其他值得注意的努力可以在这本书的结尾的延伸阅读部分找到。

相反，我们的方法将是极其具体和生动的——有时是彻头彻尾骇人听闻的。 我们将把嘉当的（柏拉图式的！）形式直接捧在你的眼前，它们的侧面闪闪发光； 我们会把它们放在你的手中，让你感受它们错综复杂的形状和令人满意的重量； 最后，你将见证它们可怕的力量。

但请留意前言预示的警告：不要被这种力量蒙蔽了双眼！ 让这台恶魔机器为几何学服务，而不是取代它！

——课程到此结束。 阿门！——

虽然嘉当形式的应用很多，但我们的主要应用（在第五幕的最后一章）自然是微分几何，使我们能够用符号的手段来重新证明前四幕用几何证明的结果。

然而，首先，我们将充分发展嘉当的思想，提供完全独立于前四幕的形式（Forms）的自成一体的介绍。 我们这样做是因为形式（Forms）在数学、物理和其他学科的不同领域中找到了卓有成效的应用。 简而言之，我们的目标是让尽可能多的读者容易理解形式，即使他们的主要兴趣不是微分几何。

32.2 1-形式的定义

嘉当的恶魔般巧妙的机器的起点和基本建筑模块是 **1−形式**（1-form）的概念。 让我们立即说明它的定义：

1−形式是单个向量作为输入的线性实值函数。

注："1-"是指单向量输入； 稍后我们将遇到以两个向量作为输入的 **2−形式**（2-forms），以三个向量作为输入的 **3−形式**（3-forms），依此类推。 因此，1-形式是一种特别简单的张量。 较早的作品将此概念称为**协变向量**或**协向量**[2]。我们将用小写粗体希腊字母表示 1-形式，同时继续用小写粗体罗马字母表示向量。

更明确地说，如果 k_1 和 k_2 是任意常数，并且 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 是任意向量，那么

$$\omega \text{ 是个 } 1\text{−形式} \quad \Leftrightarrow \quad \omega(k_1\mathbf{v}_1 + k_2\mathbf{v}_2) = k_1\,\omega(\mathbf{v}_1) + k_2\,\omega(\mathbf{v}_2) \tag{32.1}$$

当检查一个特定的 ω 是一个 1-形式时，将这个单一的要求分解成两个更简单的要求在概念上是有帮助的：

2该术语在练习 3 中进行了解释。

$$\boxed{\omega(\mathbf{v}_1 + \mathbf{v}_2) = \omega(\mathbf{v}_1) + \omega(\mathbf{v}_2),} \tag{32.2}$$

和

$$\boxed{\omega(k\mathbf{v}) = k\omega(\mathbf{v}).} \tag{32.3}$$

自行检查（32.2）和（32.3）一起蕴涵（32.1），反之亦然：线性的两个定义是等价的。

1-形式由其对向量的作用定义：两个 1-形式相等当且仅当它们对所有向量具有相同的效果。 给定两个不同的 1-形式，ω 和 φ，因此有一种自然的方法来定义它们的和，$(\omega + \varphi)$，通过它对一般向量 $\mathbf{v}$ 的作用：

$$(\omega + \varphi))(\mathbf{v}) \equiv \omega(\mathbf{v}) + \varphi(\mathbf{v}).$$

并且使用（32.2）和（32.3）很容易检查这个和本身是 1-形式 [练习]。 同样，我们可以将 1-形式 ω 乘以常数 k 以获得新的 1-形式 $[k\omega]$ ，定义为

$$[k\omega](\mathbf{v}) \equiv k[\omega(\mathbf{v})].$$

因此，1-形式的集合在常数的加法和乘法下是封闭的，因此构成了所谓的**向量空间**。 这个 1-形式的向量空间被认为是它所作用的向量空间的**对偶**。 这个术语的原因是这两个空间之间存在对称关系：我们也可以将向量空间视为 1-形式空间的“对偶”。

为了看到这种对称性，让我们将向量 $\mathbf{v}$ 视为作用于 1−形式 ω 的函数，该作用由下式定义

$$\boxed{\mathbf{v}(\omega) \equiv \omega(\mathbf{v}).}$$

向量和 1-形式相互之间的这种对称作用通常称为一种与另一种的**缩并**（contraction），有时表示为 $\langle\omega, \mathbf{v}\rangle$，以强调两种对象的平等地位。

由此得出向量 $\mathbf{v}$ 是 1-形式的线性函数：

$$\begin{aligned}\mathbf{v}(\omega + \varphi) &= (\omega + \varphi)(\mathbf{v}) \\ &= \omega(\mathbf{v}) + \varphi(\mathbf{v}) \\ &= \mathbf{v}(\omega) + \mathbf{v}(\varphi)\end{aligned}$$

和，

$$\mathbf{v}(k\omega) = k\omega(\mathbf{v}) = k\,\mathbf{v}(\omega).$$

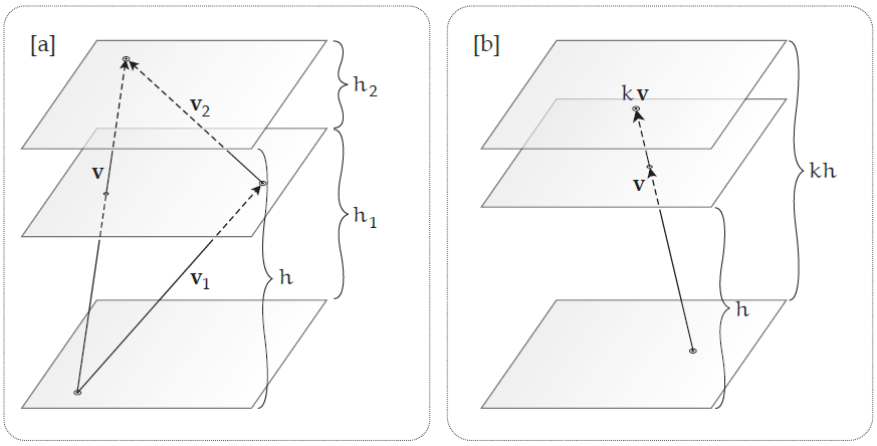
正如 T_p 表示点 p 处的向量空间一样，T_p^* 表示 p 处的 1-形式的对偶向量空间。 正如向量场为每个点 p 分配一个向量 $\mathbf{v}_p$，一个 1-形式的场为每个点 p 分配一个 1-形式 ω_p。

下一部分通过揭示一个秘密来证明 1-形式的引入（尽管是在事实之后！）有道理：你已经在 1-形式包围中度过了你的成年生活——你只是不知道而已。

32.3 1-形式的示例

32.3.1引力功

令 F 为施加在靠近地球表面的单位质量上的重力的大小。 如果我们只移动质量很短的距离，我们将近似认为 F 是常数。 如果我们将质量提升到垂直距离 h，那么所做的功（即我们必须消耗的能量）就是 $\omega = Fh$。



[32.2] [a] 当质量沿 $\mathbf{v} = (\mathbf{v}_1 + \mathbf{v}_2)$ 移动时，其高度的增加 h 只是沿 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 移动分别导致的单独高度增加的总和 $(h_1 + h_2)$。 [b] 将运动 $\mathbf{v}$ 拉伸 k 倍也会将垂直运动拉伸相同的因子。

现在假设我们沿着向量 $\mathbf{v}$ 移动质量，并定义

$$\boxed{\omega(\mathbf{v}) \equiv \text{沿 } \mathbf{v} \text{ 移动质量所做的功。}}$$

那么ω是一个 1-形式！

为了验证这一点，我们必须检查（32.2）和（32.3）。 如 [32.2a] 所示，当质量沿 $v = (\mathbf{v}_1 + \mathbf{v}_2)$ 移动时，其高度的增加 h 只是沿 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 移动分别导致的单独高度增加的总

和 $(h_1 + h_2)$。 同样，如 [32.2b] 所示，将运动 $\mathbf{v}$ 拉伸 k 倍也会将垂直运动（以及功）拉伸相同的因子。 证毕！

32.3.2可视化引力功 1-形式

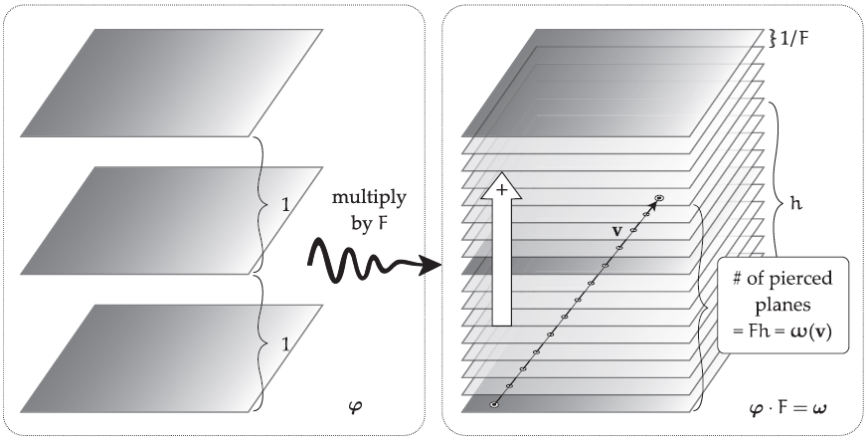
这个例子自然地暗示了一种可视化 1-形式的方法。 想象一族等距且高度不变的曲面。这些是以地球中心为中心的球面，不过，在重物的局部附近，这些球面看起来像一摞等距的水平面。 见[32.3]的右边。

当我们沿着 $\mathbf{v}$ 移动重物时，高度的变化于是做功将与 $\mathbf{v}$ 穿过的平面数量成正比。通过将曲面/平面的间距调整为 $(1/F)$，我们因此可以让

$$\boxed{\omega(\mathbf{v}) = \text{被 } \mathbf{v} \text{ 刺穿的 } \omega \text{ 曲面的数量。}}$$

当然，一般来说，$\mathbf{v}$ 的尖端会位于两个平面之间，但我们可以很容易地想象用插值填充两个平面之间的空间，例如，如果 $\mathbf{v}$ 的尖端位于平面 17 和平面 18 的中间，我们可以说它穿透了 17.5 张平面。

因此，1-形式的 ω 本身可以被可视化为具有恒定重力势能的等间距曲面的堆叠，并且由 ω 表示的力 F 越大，ω 的曲面越密集。



[32.3] **1-形式的可视化**。 右边是代表引力功 1-形式 ω 的栈（stack），其平面相距 $(1/F)$，方向向上。 因此 $\omega(\mathbf{v})$ 可以可视化为被 $\mathbf{v}$ 穿透的（有符号的）平面数。左边是单位间隔的 1-形式 φ。 当它乘以 F 时，其平面的密度增加了 F，将它们的间距缩小到 $(1/F)$，得出 $\varphi \cdot F = \omega$。

为了说明这一点，[32.3] 显示了一个 1-形式的 φ（在左侧），单位间距乘以 F 得到我们的做功 1-形式，$\varphi \cdot F = \omega$，在右侧。

请注意，向量与 1-形式缩并的这种可视化具有非常清楚地将两种对象置于同等地位的优点。 我们称这组等间距的平行平面为**栈**（stack）。 请注意，为了完成 [32.3] 中的解释，

我们必须将方向附加到栈，在我们的例子中是向上的，如图所示。 如果 $\mathbf{v}$ 的方向与 ω 的方向一致，则将穿孔计为正，而如果向量与堆栈的方向相反，则将穿孔计为负。

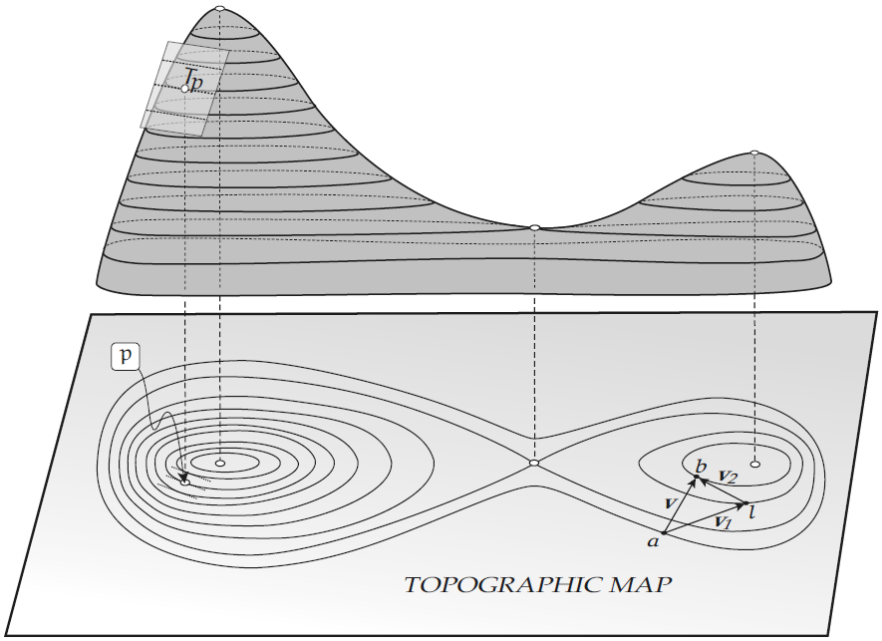
这种可视化可以更普遍地应用于任意 1-形式。 在点 p 处，ω 的代表曲面 $\mathcal{S}$ 由以下事实定义：从 p 出发与 $\mathcal{S}$ 相切的向量满足 $\omega(\mathbf{v}) = 0$，对应于零张曲面被刺穿。 将其放入线性代数的语言中，$\mathcal{S}$ 的切向量集满足 $\omega(\mathbf{v}) = 0$，因此它们构成了 ω 的**核**（kernel）。

通常，只能以这种方式将 1-形式场表示为一个点处的一摞平面——通常相邻点处的小块平面将无法啮合在一起以创建一个平滑的空间填充曲面族，称为**叶状结构**（foliation），尽管在我们的引力示例中它们确实做到了。 有关这一点的讨论，请参见 Penrose (2005, §12.3)、Bachman (2012, §5.7) 或 Dray (2015, §13.8)。

在 $\mathbb{R}^3$ 中，满足 $\omega(\mathbf{v}) = 0$ 的向量集张成二维平面，但在 n 维空间中，ω 将改为由 $(n-1)$ 维空间表示。

32.3.3地形图与梯度 1-形式

如果你计划在崎岖多山的地形上进行长途徒步旅行，建议你使用地形图规划路线，以避免攀登过于艰难，或爬下危险陡峭的斜坡。 回想一下第 203 页的示例 [19.6]，这样的地图显示了恒定高度 $h(x,y)$ 的等高线：当地图中的点 $p = (x,y)$ 沿着这样的等高线行进时，$h(p) = \text{constant}$.。



[32.4] **地形图和梯度 1-形式**。 当我们放大曲面地形图中点 p 周围的小区域时，轮廓看起来越来越直，间距也越来越均匀。 最终，它们成为在 p 正上方点处与曲面切平面 T_p 的地形图的代表，后者是对梯度 1-形式 ζ 的描述。

考虑 [32.4] 中描述的两座山丘和山口，以及下面的地形图。 如果你沿着这些等高线之

一徒步旅行，地面是平坦的：你既不攀爬也不下降。 另一方面，如果你沿着与等高线正交的方向出发，你上升（或下降）将达到最快：与等高线正交的方向是最陡峭的上升/下降方向。 等高线是在高度 h 上按相等增量 δh 画出的，所以如果你沿着与等高线正交的方向行走，地图中的等高线越挤在一起，地形就越陡峭。

更详细地说，假设我们画出了许多等高线，其中一个和下一个之间具有固定的、小的高度差 δh。 在点 p 附近，等高线的密度大致恒定； 让我们将地图中相邻等高线之间的大致恒定水平距离称为 δr。 那么那个小区域的地形坡度（slope）是

$$\text{slope} \asymp \frac{\delta h}{\delta r}$$

换句话说，地形的陡峭程度与地图中相邻等高线之间的距离成反比。

这很容易让人联想到我们对引力功 1-形式的可视化：引力越强，代表相应功 1-形式的曲面的堆积越密。考虑到这一点，让我们尝试明确识别与我们的地形图相关的陡度 1-形式。

在地图平面中，设 $v = \overrightarrow{ab}$ 是从 a 发出的向量，其尖端位于 b，如图所示。 假设我们尝试如下定义陡度 1-形式：

$$\begin{aligned}\eta(\mathbf{v}) &\equiv (\mathbf{v}\text{ 穿透的等高线数})\cdot(\delta h)\\ &= h(b)-h(a)\\ &= \mathbf{v}\text{ 从尾部到尖端的高度变化}\end{aligned}$$

这是否满足线性条件（32.2）和（32.3）？

令 $\mathbf{v}_1 = \overrightarrow{al}$，$\mathbf{v}_2 = \overrightarrow{lb}$，并且 $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$，如图所示； 那么，乍一看，似乎一切顺利：

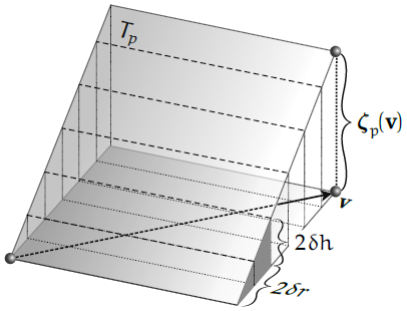
$$\eta(\mathbf{v}_1) + \eta(\mathbf{v}_2) = [h(l) - h(a)] + [h(b) - h(l)] = h(b) - h(a) = \eta(\mathbf{v}).$$

然而，与真正的 1-形式不同，$\eta(\mathbf{v})$ 的值取决于向量 $\mathbf{v}$ 的绘制位置：如果我们想象 $\mathbf{v}_2$ 在 a 而不是 l 处绘制，那么显然（32.2）不再满足！ 因此，我们在 p 处假定的 1-形式 η 实际上并没有通过第一次测试。 此外，即使我们限制为从 a 发出的向量，它也无法通过第二次测试，因为很明显，如果我们将所示向量 $\mathbf{v}$ 的长度加倍，那么我们最终会在山的另一边，并且高度实际上降低了，而不是像它应该的那样加倍。 因此 η 不满足（32.3）。

然而，这个讨论为真正的 1-形式指明了道路。 我们检查的点 p 周围的区域越小，等高线图案就越精确地一致：它们看起来像间隔均匀的平行线。 因此，如果我们只将 η 应用于从 p 发出的非常短（最终趋于零）的向量，那么 η 确实（最终）满足线性要求！

现在让我们尝试将 η 在 p 处的局部作用从 p 发出的微小向量扩展到大向量，绘制在地图平面的任何位置。 我们需要做的就是在 p 的紧邻区域（均与通过 p 的等高线的切线平行）采用这种局部平行、等距线图案，并将其延伸到整个平面。 如果我们仔细看[32.4]，

我们会发现我们刚刚构建的实际上是在 p 正上方点处曲面切平面 T_p 的地形图！



[32.5] 当我们沿 **v** 行进时，曲面高度的**梯度 1-形式** ζ_p 应用于平面中的向量 **v** 时，会产生 T_p 高度的变化，其中 T_p 是曲面点 p 正上方点处的切平面。

为了获得一个真正的 1-形式场 ζ，我们只需要将旧的 η 定义应用于 T_p 的地形图。见[32.5]。我们重申，这张地形图由等间距的平行线组成，所有平行线都与通过 p 的原始轮廓的方向相同，一条线与另一条线之间的间隙为 $\delta r \asymp (\delta h)/(T_p$ 的斜率$)$：T_p 的斜率越大，其地形图中的平行线越密集。 与这些线正交的方向是在点 p 处曲面最陡上升的方向。

p 处的 1-形式场 ζ 的值 ζ_p 是由位于 p 正上方的曲面点处的切平面 T_p 定义的 1-形式。 如[32.5]所示，

$$\begin{aligned}\zeta_p(\mathbf{v}) &\equiv \big(\mathbf{v}\text{ 穿过的线数}\big)\cdot(\delta h)\\ &= T_p\text{ 从 }\mathbf{v}\text{ 尾部到尖端的高度变化。}\end{aligned} \tag{32.4}$$

随着 p 的变化，围绕 p 的原始等高线的方向和间距都发生变化，反映了与曲面的切平面 T_p 变化的事实，因此我们在每个点都有不同的 T_p 地形图，因此不同 1-形式 域 ζ 的值 ζ_p。

如果我们将 ζ_p 应用于从 p 发出的非常短（最终消失）的向量，那么它的效果最终等于 η 的效果； 也就是说，$\zeta_p(\mathbf{v}) \asymp \eta_p(\mathbf{v})$。 但是随着向量变大，$\eta(\mathbf{v})$ 告诉我们实际地形/曲面高度的非线性变化，而 1-形式 $\zeta_p(v)$ 告诉我们曲面在那个点的切平面高度的完美线性变化。

这个 1-形式场 ζ 称为 h 的**梯度**。你可能会问自己，这是否与向量分析（Vector Calculus）中熟悉的同名概念有关。 它的确是！ 关联的解释是第 32.6 节的主题。 梯度 1-形式场 ζ 的这个想法非常重要，我们将看到它如何推广到依赖于两个以上变量的函数 h，因此不能可视化为曲面关于 (x,y) 平面的高度 $z = h(x,y)$。

32.3.4行向量

考虑一个特定的二维行向量，例如 $\omega = [-3, 2]$。我们可以通过左侧的标准矩阵乘法来定义 ω 对二维列向量 $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ 的作用：

$$\omega(\mathbf{v}) = [-3, 2]\begin{bmatrix} x \\ y \end{bmatrix} = -3x + 2y$$

请自行检查这确实满足（32.2）和（32.3）。 因此，这样看，行向量是 1-形式。

我们可以将这个例子与引力做功的例子联系起来。 为此，让我们选择笛卡尔坐标 (x, y, z)，z 轴垂直，在地球引力场的方向上。 然后我们可以将引力做功表示为 1-形式（即行向量），

$$\omega = [0, 0, F], \qquad \text{因为那时} \quad \omega(v) = [0, 0, F]\begin{bmatrix} x \\ y \\ h \end{bmatrix} = Fh = \text{功}$$

32.3.5狄拉克的左矢

注意：这个例子假设你已经熟悉量子力学； 如果你不是，请随意跳过它。

尽管我们已经将 1-形式定义为实值的，并且将继续这样做——除了一个例外——在本书的其余部分，扩大定义以包括复值 1-形式通常是有用的。 事实上，这种推广在量子力学中自然出现，且在这种情况下必不可少。

如果我们将**狄拉克右矢** $|v\rangle$（即量子态）视为“向量”，那么**狄拉克左矢** $\langle\omega|$ 是 1−形式，前提是我们将 1-形式和向量的收缩定义为标准（复）中括号（bra-ket）内积：

$$\omega(\mathbf{v}) \equiv \langle\omega|\mathbf{v}\rangle.$$

要确认左矢 $\langle\omega|$ 是一个 1-形式，我们只需检查定义，（32.1）：

$$\begin{aligned} \omega(k_1\mathbf{v}_1 + k_2\mathbf{v}_2) &= \langle\omega|k_1\mathbf{v}_1 + k_2\mathbf{v}_2\rangle \\ &= k_1\langle\omega|\mathbf{v}_1\rangle + k_2\langle\omega|\mathbf{v}_2\rangle \\ &= k_1\,\omega(\mathbf{v}_1) + k_2\,\omega(\mathbf{v}_2). \end{aligned}$$

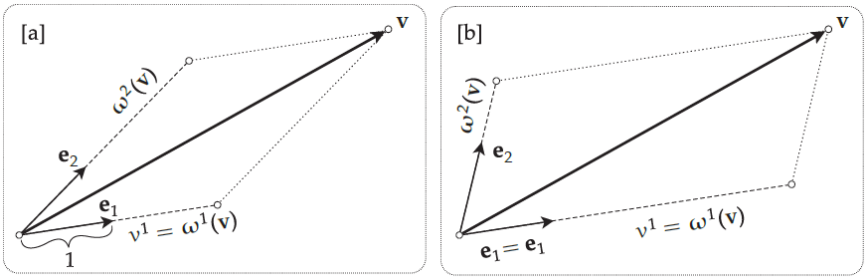
32.4 基 1-形式

在 n-流形中的点 p 处，假设我们为 T_p 选择了一个基 $\{\mathbf{e}_j\}$，因此可以使用爱因斯坦求和约定写出一个通用向量，如

$$\mathbf{v} = v^j\,\mathbf{e}_j$$

我们不应假设这个基是正交的。 给定这个基，有一种自然的方法可以将其与 1-形式在 p 处空间 T_p^* 的**对偶基** $\{\boldsymbol{\omega}^i\}$ 与相关联：

$$\boxed{\boldsymbol{\omega}^i \text{ 挑选出 } \mathbf{v} \text{ 的第 i 个分量} \quad \Leftrightarrow \quad \boldsymbol{\omega}^i(\mathbf{v}) = v^i\text{。}} \tag{32.5}$$



[32.6] [a] 第一个基 1-形式 $\boldsymbol{\omega}^1$ 挑出 $\mathbf{v}$ 的第一个分量 v^1； 同样，第二个基 1-形式 $\boldsymbol{\omega}^2$ 挑选出第二个分量。 [b] 使用与之前相同的 $\mathbf{e}_1$，仅更改 $\mathbf{e}_2$ 会导致两个基 1-形式都发生更改。

这些 $\{\boldsymbol{\omega}^i\}$ 确实是 1-形式，这可以通过检查（32.2）和（32.3）来确认：两个向量之和的第 i 个分量是它们第 i 个分量的和， $k\mathbf{v}$ 的第 i 个分量就是 kv^i。 完毕！

图 [32.6a] 在切平面 $T_p = \mathbb{R}^2$ 的 2-流形（即曲面）的情况下说明了这个定义。 在这里，如图所示，我们选择了非正交单位向量的基 $\{\mathbf{e}_1, \mathbf{e}_2\}$，并且

$$\mathbf{v} = v^1\,\mathbf{e}_1 + v^2\,\mathbf{e}_2 = \boldsymbol{\omega}^1(\mathbf{v})\,\mathbf{e}_1 + \boldsymbol{\omega}^2(\mathbf{v})\,\mathbf{e}_2$$

因此，在这种情况下，$\boldsymbol{\omega}^1(\mathbf{v})$ 只是 $\mathbf{v}$ 在 $\mathbf{e}_1$ 上的投影（平行于 $\mathbf{e}_2$）的长度，对应于 $\boldsymbol{\omega}^2(\mathbf{v})$ 相似，如图所示。

让我们将以下常见的误解扼杀在萌芽状态。 虽然基 1-形式的集合 $\{\boldsymbol{\omega}^1, \boldsymbol{\omega}^2\}$ 与基向量集合 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 是对偶的，但认为 $\boldsymbol{\omega}^1$ 与 $\mathbf{e}_1$ 对偶，而 $\boldsymbol{\omega}^2$ 与 $\mathbf{e}_2$ 对偶是完全错误的。

图[32.6b]揭示了这种想法的错误。 在这里，我们已经说明了相同的向量 $\mathbf{v}$ 分解成相对于新基 $\{\widetilde{\mathbf{e}_1}, \widetilde{\mathbf{e}_2}\}$ 的分量，其中 $\widetilde{\mathbf{e}_1} = \mathbf{e}_1$，但 $\widetilde{\mathbf{e}_2} \neq \mathbf{e}_2$。 如你所见，即使我们只更改了一个基向量，两个基 1-形式都已改变！ 同样，在一般的 n 维情况下，改变单个基向量可以改变

1-形式的整个对偶基。

我们应该注意，我们对对偶基的首选定义（32.6）不是传统的，尽管它是舒茨（Schutz，1980）使用的定义，并且正如我们现在解释的那样，它在数学上等同于标准定义。

首先回忆一下定义

$$\textbf{克罗内克 delta 函数:} \qquad \delta^i_j \equiv \left\{ \begin{array}{c} 1 \text{ 若 } i = j \\ 0 \text{ 若 } i \neq j \end{array} \right\}$$

很容易证明我们的定义（32.6）等价于以下标准定义：

$$\text{基}\{\boldsymbol{\omega}^i\}\text{的等价定义：} \qquad \boxed{\boldsymbol{\omega}^i(\mathbf{e}_j) = \delta^i_j \quad \Leftrightarrow \quad \boldsymbol{\omega}^i(\mathbf{v}) = v^i\text{。}}$$

从右到左的蕴涵是直接的，从左到右的蕴涵是这样的：

$$\boldsymbol{\omega}^i(\mathbf{v}) = \boldsymbol{\omega}^i(\mathbf{v})(v^j\,\mathbf{e}_j) = v^j\,\boldsymbol{\omega}^i(\mathbf{e}_j) = v^j\,\delta^i_j = v^i$$

32.5 1-形式的分量

一直保持警惕的读者会注意到，我们实际上并没有证明集合 $\{\boldsymbol{\omega}^i\}$ 是 1-形式的基。 与其直接证明它们是线性独立的，不如遵循舒茨（Schutz，1980， §2.20）采用的更优雅、更富有启发性的方法。

让一般 1-形式 φ 作用于一般向量 $\mathbf{v}$：

$$\begin{aligned} \varphi(\mathbf{v}) &= \varphi(v^j\,\mathbf{e}_j) \\ &= v^j\,\varphi(\mathbf{e}_j) \\ &= \boldsymbol{\omega}^j(\mathbf{v})\,\varphi(\mathbf{e}_j) \end{aligned}$$

我们现在定义

$$\boxed{\varphi\text{ 的分量 }\varphi_j\text{：}\ \varphi_j \equiv \varphi(\mathbf{e}_j)\text{。}}$$

因此，

$$\varphi(\mathbf{v}) = \varphi_j\,\boldsymbol{\omega}^j(\mathbf{v})$$

但是由于 1-形式是由它对一般向量的作用来定义的，我们可以“抽象掉”这个等式两边的向量 $\mathbf{v}$，从而使 1-形式本身相等，并在这样做时将任意 1-形式 φ 分解为与向量基 $\{\mathbf{e}_j\}$ 对偶的 1-形式 基 $\{\boldsymbol{\omega}^i\}$ 的唯一分量：

$$\boxed{\varphi = \varphi_j\, \boldsymbol{\omega}^j = \varphi(\mathbf{e}_j)\boldsymbol{\omega}^j} \tag{32.6}$$

32.6 作为 1-形式的梯度：df

32.6.1梯度作为向量的回顾：∇f

回想一下向量分析（Vector Calculus），在 $\mathbb{R}^2$ 中，函数 f 的**梯度**被定义为向量

$$\nabla f \equiv \begin{bmatrix} \partial_x f \\ \partial_y f \end{bmatrix}$$

这个向量的意义在于，

> ∇f 指向 f 增长最快的方向，其大小 $|\nabla f|$ 等于我们朝那个方向移动时 f 随距离增加的最大速率。 (32.7)

这种解释源于我们现在得出的一个更原始的事实。 令 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 为沿 $(x^1, x^2) = (x, y)$ 轴的标准正交基，令 δf 为沿短（最终趋于零）向量移动导致的 f 的微小变化

$$\mathbf{v} = \delta x^1\, \mathbf{e}_1 + \delta x^2\, \mathbf{e}_2 = \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$$

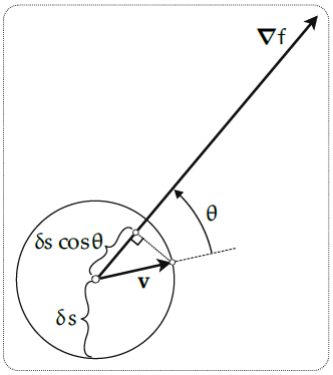
根据定义，$\partial_x f$ 是我们沿 x 方向移动时 f 的变化率，因此移动 δx 导致的 f 变化最终等于 $\partial_x f\, \delta x$；对于 y 同样。 所以，

$$\boxed{\delta f \asymp (\partial_x f)\, \delta x + (\partial_y f)\, \delta y = (\nabla f) \cdot \mathbf{v}} \tag{32.8}$$

在 18 世纪的古典符号中，这被表示为

$$df = (\partial_x f)\, dx + (\partial_y f)\, dy, \tag{32.9}$$

其中 df、dx 和 dy 被理解为无穷小——我们在本书中避开了这个概念。 然而，正如我们将很快看到的那样，当通过现代的 1-形式这一棱镜观察时，这个古老的公式有了新的生命，具有精确和严格的含义。



[32.7] **梯度向量** ∇f。 沿短向量 **v** 移动产生的微小变化 δf 最终等于 |∇f| (δs cosθ)。

为了导出（32.8），考虑 [32.7]，它以几何方式解释（32.9）。 如果我们保持长度 $|v| \equiv \delta s$ 固定，那么当 **v** 绕圆旋转时，

$$\delta f \asymp |\nabla f|\,(\delta s \cos\theta),$$

其中 θ 是 **v** 与梯度 ∇f 的夹角，因此 (δs cos θ) 只是 **v** 在 ∇f 方向上的投影，如图所示。

令 $\hat{\mathbf{v}} = \mathbf{v}/|\mathbf{v}|$ 表示 **v** 方向上的**单位向量**。则上述结果可以改为表示为

$$\boxed{|\nabla_{\hat{\mathbf{v}}} f| = \frac{df}{ds} = |\nabla f| \cos\theta.}$$

这立即证实了（32.8）中给出的 ∇f 的解释。 它还表明，如果我们沿与 ∇f 正交的方向移动，那么 f 的变化率就会消失。 这是由 f = constant 定义的曲线的切线方向。

在 $\mathbb{R}^3$ 中，分析与以前完全相同，但现在有一个与 ∇f 正交的方向的整个平面：这是曲面的切平面，f = constant。

32.6.2作为 1-形式的梯度：df

虽然我们习惯于将梯度视为一个向量，但实际上它更自然地是一个 1-形式。 让我们立即定义这个 1-形式，然后继续理解它。

函数 f 的梯度 1−形式 $\mathbf{d}f$ 由其对向量的作用定义：

$$\boxed{\mathbf{d}f(\mathbf{v}) \equiv \boldsymbol{\nabla}_{\mathbf{v}} f.} \tag{32.10}$$

加粗的 $\mathbf{d}$ 算子称为**外导数**（exterior derivative）；它将在接下来的工作中发挥核心作用。

让我们通过验证（32.2）和（32.3）来检查 $\mathbf{d}f$ 是否真的是 1-形式：

$$(\mathbf{d}f)\,(\mathbf{v}_1 + \mathbf{v}_2) = \boldsymbol{\nabla}_{\mathbf{v}_1 + \mathbf{v}_2} f = \boldsymbol{\nabla}_{\mathbf{v}_1} f + \boldsymbol{\nabla}_{\mathbf{v}_2} f = \mathbf{d}f(\mathbf{v}_1) + \mathbf{d}f(\mathbf{v}_2)$$

和，

$$(\mathbf{d}f)(k\mathbf{v}) = \boldsymbol{\nabla}_{k\mathbf{v}} f = k\,\boldsymbol{\nabla}_{\mathbf{v}} f = k\,\mathbf{d}f(\mathbf{v}).$$

请注意，由于 $\boldsymbol{\nabla}_{\mathbf{v}}$ 遵循莱布尼茨法则（又名乘积法则），因此外导数也遵循：

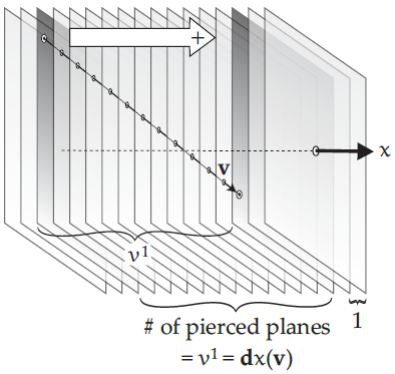
$$\boxed{\mathbf{d}(fg) = f\,\mathbf{d}g + g\,\mathbf{d}f.} \tag{32.11}$$

32.6.3笛卡尔 1-形式基：$\{\mathbf{d}x^j\}$

$\mathbb{R}^2$ 中 1-形式 $\mathbf{d}x = \mathbf{d}x^1$ 和 $\mathbf{d}y = \mathbf{d}x^2$ 是什么意思？ 要回答这个问题，我们必须确定它们对一般向量 $\mathbf{v} = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix}$ 的影响。定义（32.11）告诉我们

$$\begin{aligned}
(\mathbf{d}x)\begin{bmatrix} v^1 \\ v^2 \end{bmatrix} &= (\mathbf{d}x)\mathbf{v} \\
&= \boldsymbol{\nabla}_{\mathbf{v}}\, x \\
&= \boldsymbol{\nabla}_{v^1\mathbf{e}_1 + v^2\mathbf{e}_2}\, x \\
&= \left(v^1\boldsymbol{\nabla}_{\mathbf{e}_1} + v^2\boldsymbol{\nabla}_{\mathbf{e}_2}\right)x \\
&= \left(v^1\,\partial_x + v^2\,\partial_y\right)x \\
&= v^1
\end{aligned}$$

换句话说，$\mathbf{d}x$ 是挑选出向量的 x 分量的 1−形式。 同样，我们看到 $(\mathbf{d}y)\mathbf{v} = v^2$。 但是根据定义（32.6），这意味着$\{\mathbf{d}x, \mathbf{d}y\}$ 是 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 的对偶基。 等效地，$(\mathbf{d}x^i)\mathbf{e}_j = \delta^i_j$。



[32.8] **笛卡尔基 1-形式 $\mathbf{d}x$**。 当应用于向量 $\mathbf{v}$ 时，1-形式 $\mathbf{d}x$ 挑选出 x 分量：$\mathbf{d}x(\mathbf{v}) = v^1$。

在几何上，$\mathbf{d}x$ 可以被描绘为垂直于 x 轴的单位间距线族，并且 $(\mathbf{d}x)\mathbf{v}$ 是 $\mathbf{v}$ 穿过的这些线的数量。 上述推理同样适用于 $\mathbb{R}^3$，在这种情况下，$\mathbf{d}x$ 是垂直于 x 轴的单位间距平面族，如 [32.8] 所示。

显然，所有这些都可以立即推广到任意数量的维度。 令 $\{\mathbf{e}_i\}$ 为 $\mathbb{R}^n$ 的正交基，具有笛卡尔坐标 $\{x^j\}$，因此 $\mathbf{v} = v^i\,\mathbf{e}_i$ 是一般向量。 然后 $\mathbf{d}x^j$ 挑选出向量的第 j 个分量，即 $(\mathbf{d}x^j)\,\mathbf{v} = v^j$。 特别是，

$$\boxed{(\mathbf{d}x^i)\mathbf{e}_j = \delta^i_j.} \tag{32.12}$$

我们称之为**笛卡尔基**：

$$\boxed{\{\mathbf{d}x^j\} = \{\boldsymbol{\omega}^j\} \text{ 是 } \{\mathbf{e}_j\} \text{ 对偶的 1−形式的笛卡尔基。}}$$

从（32.7）可以看出，一般的 1-形式 φ 可以在这个对偶基中分解成分量：[3]

$$\boxed{\varphi = \varphi_j\,\boldsymbol{\omega}^j = \varphi(\mathbf{e}_j)\,\mathbf{d}x^j.} \tag{32.13}$$

32.6.4 $\mathbf{d}f = (\partial_x f)\,\mathbf{d}x + (\partial_y f)\,\mathbf{d}y$ 的 1-形式解释

取（32.14）中的 $\varphi = \mathbf{d}f$，我们绕了一圈，将一般函数 f 的梯度 1-形式 $\mathbf{d}f$ 分解为其笛卡尔基 1-形式分量，如下所示：

$$\boxed{\mathbf{d}f = \left[(\mathbf{d}f)\,\mathbf{e}_j\right]\mathbf{d}x^j = \left[\partial_{x^j} f\right]\mathbf{d}x^j} \tag{32.14}$$

在形式上，这与经典公式（32.10）相同，但现在它具有精确、严格的含义，不需要无穷小。然而，它与[32.7] 中说明的具有几何意义的终极等式（32.9）非常直接和直观地相关联。

看到这一点，再次将 $\mathbf{v}$ 视为一个短的，最终趋于零的向量，

$$\mathbf{v} = \delta x^1\, \mathbf{e}_1 + \delta x^2\, \mathbf{e}_2 = \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$$

于是，

$$\mathbf{d}f(\mathbf{v}) = \left\{[\partial_x f]\mathbf{d}x + [\partial_y f]\mathbf{d}y\right\}\begin{bmatrix} \delta x \\ \delta y \end{bmatrix} = (\partial_x f)\, \delta x + (\partial_y f)\, \delta y \asymp \delta f,$$

正好是（32.9）。 因此，1-形式 $\mathbf{d}f$ 为我们提供了两全其美的效果！

上述分析试图将新的 1-形式 $\mathbf{d}f$ 与你对向量梯度 ∇f 的先验知识联系起来。 但是，如果你回顾我们对地形图的讨论，你会发现我们已经仅通过几何推理就已经到达了这个新的 1-形式。

确实，如果你查看与地形图相关的 1-形式 ζ 的定义（32.5）（在 [32.5] 中说明），它描述了当我们沿任意方向移动时切平面的高度如何变化，你会看到确实是高度函数 h 的梯度：

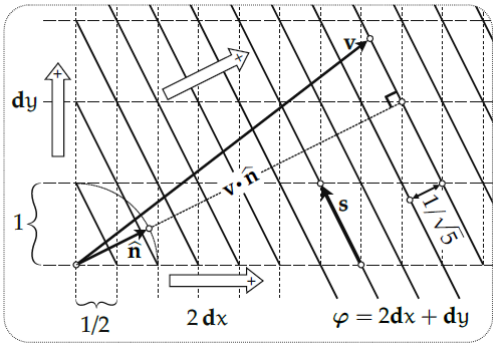
$$\boxed{\zeta = \mathbf{d}h.}$$

特别是，如果 $\mathbf{v}$ 是通过该点的恒定高度的等高线在 p 处的方向，则 $\zeta(\mathbf{v}) = \mathbf{d}h(\mathbf{v}) = 0$，这是理所当然的。

32.7 以几何方式加入 1-形式

我们已经知道将 1-形式乘以常数 k 的几何意义：它将栈缩并 k。 但是 1-形式的加法是什么意思呢？

图 [32.9] 将代表 1-形式 $2\mathbf{d}x$ 和 $\mathbf{d}y$ 的栈叠加，然后连接得到的交点以创建一个新栈。这是它们和的几何构造，$\varphi = 2\mathbf{d}x + \mathbf{d}y$！



[32.9] **1-形式的几何加法**。 要将 1-形式 $2\,\mathbf{d}x$ 和 $\mathbf{d}y$ 相加，叠加它们的栈，然后连接它们的交点以创建 $\varphi = 2\mathbf{d}x + \mathbf{d}y$ 的栈。

如果 $\tilde{\varphi}$ 表示对应于构造栈的 1-形式，那么我们必须证明 $\tilde{\varphi} = \varphi$。

我们将给出三个证明，并不是因为结果特别重要，而只是作为思考和使用 1-形式的宝贵实践。

首先注意，垂直于 $\tilde{\varphi}$ 的单位向量是[练习]

$$\hat{n} = \frac{1}{\sqrt{5}}\begin{bmatrix}2\\1\end{bmatrix}$$

另请注意，这用于指定栈的方向，必须沿要相加的两个栈具有正分量[练习]。

如果 $\mathbf{v} = \begin{bmatrix}a\\b\end{bmatrix}$ 表示一个通用向量，那么它在 $\hat{\mathbf{n}}$ 方向上的投影就是$\mathbf{v}\cdot\hat{\mathbf{n}}$，如图。但是 $e\varphi$ 的间距是$1/\sqrt{5}$ [练习]，所以被 $\mathbf{v}$ 穿过的栈的线数 $\tilde{\varphi}$ 是

$$\tilde{\varphi}(\mathbf{v}) = \text{穿过线数} = v\cdot bn\left(\frac{1}{\sqrt{5}}\right) = 2a + b = (2dx + dy)\begin{bmatrix}a\\b\end{bmatrix} = \varphi(\mathbf{v})$$

由于 $\mathbf{v}$ 是一个通用向量，我们可以将其抽象出来，从而证实我们的主张 $\tilde{\varphi} = \varphi$。

这是第二个证明，它更简单，但不太直接。我们知道，对于任何 1-形式，例如我们的 φ，位于栈中的方向 $\mathbf{s}$ 是 1-形式的核（kernel）：$\mathbf{s}$ 不穿过栈的任何线（或平面，在 $\mathbb{R}^3$ 中），所以

$$0 = \varphi(\mathbf{s}) = (2dx + dy)\begin{bmatrix}p\\q\end{bmatrix} = 2p + q \quad\Longrightarrow\quad \mathbf{s} \propto \begin{bmatrix}-1\\2\end{bmatrix}$$

但是，如图所示，这也是 $\tilde{\varphi}$ 栈的方向，连接网格点。 因此，这两个堆栈是平行的，并且只能在间距方面有所不同。 用代数来说，这两个 1-形式必须成比例：对于某个常数 K，$\tilde{\varphi} = K\varphi$。但是如果我们将这个方程应用于一个特定的向量，它的穿过线数是明显的——尝试 $\begin{bmatrix}0\\1\end{bmatrix}$ ——那么我们立即推断 $K = 1$. 完成！

第三个证明可能是最简单的。 在 [32.9] 中，将 $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ 视为一个位置向量，从一个固定的原点发出。在 φ 栈中一条线的方程是什么？嗯，当 $\mathbf{v}$ 的尖端沿着这条线移动时， $\mathbf{v}$ 穿透固定数量的堆栈行，所以

$$\text{constant} = \varphi(v) = (2dx + dy)\begin{bmatrix} x \\ y \end{bmatrix} = 2x + y \implies y = -2x + \text{constant}.$$

但是这些斜率为 -2 的线平行于 $\tilde{\varphi}$ 栈。 其余证明与之前相同。

很容易推广[练习]概括这些论点，以确认这种几何构造适用于 $p\,\mathbf{d}x + q\,\mathbf{d}y$，产生一个堆栈

$$\text{法线向量}\begin{bmatrix} p \\ q \end{bmatrix}\text{和间距}\frac{1}{\sqrt{p^2 + q^2}}$$

实际上，在 $\mathbb{R}^3$ 中，我们可以 [练习] 通过将三个正交平面栈相加来构造 $p\,\mathbf{d}x + q\,\mathbf{d}y + r\,\mathbf{d}z$，通过交叉点构造一个新的平面栈，

$$\text{法线向量}\begin{bmatrix} p \\ q \\ r \end{bmatrix}\text{和间距}\ \frac{1}{\sqrt{p^2 + q^2 + r^2}}$$

第三十三章 张量

33.1 张量的定义：价

在我们对黎曼张量的讨论中，我们暂时将张量定义为多个向量输入的多线性函数（multilinear function）。 这是我们当时能做的最好的事情，因为我们缺乏 1-形式的概念。但实际上，一个完全一般的张量是向量和 1-形式的多线性函数，而**价**（valence）告诉我们每个有多少：

> 在点 p 的 $\begin{Bmatrix} \mathfrak{f} \\ \mathfrak{v} \end{Bmatrix}$ **价张量** **H** 是 $\mathfrak{f}$ 个 1−形式和 $\mathfrak{v}$ 个向量的多线性实值函数，因此在 p 处 H 的值仅取决于 1−形式和向量在 p 的值。

因此 1-形式是 $\begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ 价的张量，因为它有一个“槽”，可以将单个向量输入其中，输出是该向量与 1-形式的缩并。 同样，向量是 $\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$ 价的张量，因为它有一个槽，可以将单个 1-形式输入其中，输出是该 1-形式与向量的缩并。

在一般情况下，我们可以将 **H** 的输入槽分为两组：$\mathfrak{f}$ 个槽，用于输入 $\mathfrak{f}$ 个 1-形式，$\varphi_1,\ldots,\varphi_{\mathfrak{f}}$ 和第二组 $\mathfrak{v}$ 个槽来输入 $\mathfrak{v}$ 个向量 $\mathbf{v}_1,\ldots,\mathbf{v}_{\mathfrak{v}}$。 用 ‖ 表示这两组之间的边界，我们有

$$\mathbf{H}(\varphi_1,\ldots,\varphi_{\mathfrak{f}} \parallel \mathbf{v}_1,\ldots,\mathbf{v}_{\mathfrak{v}}).$$

一般来说，我们将 1-形式和向量输入这些槽的顺序很重要：如果我们在这些槽中交换一对 1-形式或一对向量，则输出通常与原始值完全无关。

让我们从这个新的一般观点重新审视黎曼张量 **R**。 在（29.8）中，我们将其定义为三个向量输入的向量值多线性函数：**w** 是围绕平行四边形平行移动的向量，边由 **u** 和 **v** 定义，我们现在认为它们很短，输出是向量和乐 $\delta\mathbf{w}$：

$$\delta\mathbf{w} = \mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w}) = \{[\boldsymbol{\nabla}_{\mathbf{u}}, \boldsymbol{\nabla}_{\mathbf{v}}] - \boldsymbol{\nabla}_{[\mathbf{u},\mathbf{v}]}\}\mathbf{w}$$

为了使它成为新意义上的张量，它必须输出一个实数，而不是向量 $\delta\mathbf{w}$。 为了从 $\delta\mathbf{w}$ 中提取一个实数，我们必须用 1-形式缩并它。 出于这个原因，取代我们原来的定义，黎曼张量的标准定义是一个 $\begin{Bmatrix} 1 \\ 3 \end{Bmatrix}$ 价张量：

$$\mathbf{R}(\varphi \parallel \mathbf{u},\mathbf{v},\mathbf{w}) \equiv \langle\varphi,\delta\mathbf{w}\rangle = \varphi(\delta\mathbf{w}).$$

注：我们将为具有多个输入的张量保留大写粗体罗马字母，例如黎曼张量 $\mathbf{R}$、能量动量张量 $\mathbf{T}$ 和爱因斯坦张量 $\mathbf{G}$。另一方面，一些这样的高价张量传统上用小写粗体罗马字母表示，最著名的例子是度量张量 $\mathbf{g}$，别将其与向量混淆！

33.2 示例：线性代数

应该已经很清楚了，大量的数学和物理对象属于我们新的、广义的张量类别。 为了明白这一点，请考虑线性代数这一庞大且极其重要的主题。 从我们的新观点来看，这“仅仅是”对 $\begin{Bmatrix}1\\1\end{Bmatrix}$ 价的张量 $\mathbf{L}(\varphi \parallel \mathbf{v})$ 的研究！

要了解这是为什么，请考虑 $\mathbf{L}(\ldots \parallel \mathbf{v})$，其中我们将第一个槽留空。为了产生一个数字，我们必须重新插入 1-形式 φ。 因此，$\mathrm{L}(\mathbf{v}) \equiv \mathbf{L}(\ldots \parallel \mathbf{v})$ 必须是一个向量，准备好与 1-形式 φ 缩并以产生一个数。 这完全类似于我们最初对黎曼张量的定义：我们有一个由三个向量组成的向量值函数； 现在我们有一个单一向量的向量值函数 L，它只有在我们用 1-形式缩并它时才会产生一个数字。

但是，根据张量的定义，$\mathbf{L}(\varphi \parallel \mathrm{v})$ 是 φ 和 $\mathbf{v}$ 的线性函数。因此，

$$\begin{aligned}\mathrm{L}(\mathrm{k}_1\mathbf{v}_1 + \mathrm{k}_2\mathbf{v}_2) &= \mathbf{L}(\ldots \parallel \mathrm{k}_1\mathbf{v}_1 + \mathrm{k}_2\mathbf{v}_2)\\ &= \mathrm{k}_1\,\mathbf{L}(\ldots \parallel \mathbf{v}_1) + \mathrm{k}_2\,\mathbf{L}(\ldots \parallel \mathbf{v}_2)\\ &= \mathrm{k}_1\,\mathrm{L}(\mathbf{v}_1) + \mathrm{k}_2\,\mathrm{L}(\mathbf{v}_2).\end{aligned}$$

因此 L 是一个线性变换，将向量发送到向量——线性代数的基本研究对象！ 当然，出于同样的原因，如果我们让向量槽打开，我们可以将 $\mathbf{L}(\varphi \parallel \ldots)$ 解释为输入为 1-形式的，线性的，1-形式值的函数。

33.3 张量：从旧得新

33.3.1加法

显然，将具有不同价的张量（需要完全不同类型的输入）相加是没有意义的。 但是，如果两个张量具有相同的价，则加法的定义显而易见。 例如，如果 $\mathbf{H}$ 和 $\mathbf{J}$ 都是 $\begin{Bmatrix}2\\1\end{Bmatrix}$ 价，那么

$$(\mathbf{H} + \mathbf{J})(\varphi, \psi \parallel \mathbf{v}) \equiv \mathbf{H}(\varphi, \psi \parallel \mathbf{v}) + \mathbf{J}(\varphi, \psi \parallel \mathbf{v}).$$

很容易检查 [练习] $(\mathbf{H}+\mathbf{J})$ 确实是另一个 $\begin{Bmatrix}2\\1\end{Bmatrix}$ 价张量。

33.3.2乘法：张量积

给定两个 1-形式 φ 和 ψ，每个都作用于单个向量，很自然地定义它们的**张量积**（tensor product），作用于两个向量，因此为 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价，如下所示：

$$\boxed{(\varphi\otimes\psi)(\mathbf{v},\mathbf{w})\equiv\varphi(\mathbf{v})\psi(\mathbf{w}).} \tag{33.1}$$

请注意，顺序很重要：$\varphi\otimes\psi\neq\psi\otimes\varphi$。

术语：张量积通常也称为**直积**（direct product）或**外积**（outer product）。

更高价张量的张量积也是如此。 例如，我们可以将 $\begin{Bmatrix}2\\1\end{Bmatrix}$ 价的张量 $\mathbf{J}(\varphi,\psi\parallel\mathbf{u})$ 和 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价的张量 $\mathbf{T}(\mathbf{v},\mathbf{w})$ 相乘，得到 $\begin{Bmatrix}2\\3\end{Bmatrix}$ 价的张量 $\mathbf{J}\otimes\mathbf{T}$：

$$(\mathbf{J}\otimes\mathbf{T})(\varphi,\psi\parallel\mathbf{u},\mathbf{v},\mathbf{w})\equiv\mathbf{J}(\varphi,\psi\parallel\mathbf{u})\cdot\mathbf{T}(\mathbf{v},\mathbf{w}).$$

注意在张量乘法下，价的相加与向量类似：在我们的示例中，$\begin{Bmatrix}2\\1\end{Bmatrix}+\begin{Bmatrix}0\\2\end{Bmatrix}=\begin{Bmatrix}2\\3\end{Bmatrix}$。 由于乘以标量不会改变价，因此必须将标量的价视为 $\begin{Bmatrix}0\\0\end{Bmatrix}$。

33.4 分量

像往常一样，令 $\{\mathbf{e}_i\}$ 为正交向量基，令 $\{\mathbf{d}x^j\}$ 为 1-形式的对偶笛卡尔基。 就像我们获得向量和 1-形式的分量一样，我们通过用基 1-形式和向量填充它的槽来获得更一般的张量的**分量**。 例如，$\mathbf{T}(\mathbf{v},\mathbf{w})$ 的分量是

$$T_{ij}=\mathbf{T}(\mathbf{e}_i,\mathbf{e}_j)$$

我们可以将完整的张量 $\mathbf{T}$ 分解为张量分量，如下所示：

$$
\begin{aligned}
\mathbf{T}(\mathbf{v},\mathbf{w}) &= \mathbf{T}\left(v^i\mathbf{e}_i, w^j\mathbf{e}_j\right) \\
&= \mathbf{T}\left(\mathbf{e}_i,\ \mathbf{e}_j\right)v^i w^j \\
&= T_{ij}\, v^i w^j \\
&= T_{ij}\left[\mathbf{d}x^i(\mathbf{v})\right]\left[\mathbf{d}x^j(\mathbf{w})\right] \\
&= T_{ij}\,(\mathbf{d}x^i \otimes \mathbf{d}x^j)(\mathbf{v},\mathbf{w})
\end{aligned}
$$

但是由于 $\mathbf{v}$ 和 $\mathbf{w}$ 都是一般向量，我们可以将它们抽象出来，发现张量本身是

$$
\boxed{\mathbf{T} = T_{ij}\,(\mathbf{d}x^i \otimes \mathbf{d}x^j).} \tag{33.2}
$$

因此我们看到

张量集合 $\{\mathbf{d}x^i \otimes \mathbf{d}x^j\}$ 构成了 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价张量的一组基。

请注意，如果我们将这个想法应用于（33.1），那么我们会发现

$$
\varphi \otimes \psi = \varphi_i\psi_j\,(\mathbf{d}x^i \otimes \mathbf{d}x^j).
$$

然而，这样的张量是特殊的：一般的 $\mathbf{T}$ 不能以这种方式分解。

我们可以从以下事实中看出这一点：虽然 $\mathbf{T}$ 在 n 维中具有 n^2 个分量 T_{ij}，但 $(\varphi \otimes \psi)$ 仅由 2n 个数字唯一确定：φ 的 n 个分量 φ_i 和 ψ 的 n 个分量 ψ_i。

以完全相同的方式，我们可以取两个向量的张量积：

$$
(\mathbf{v} \otimes \mathbf{w})(\varphi,\psi) = \mathbf{v}(\varphi)\mathbf{w}(\psi) = v^i\, w^j(\mathbf{e}_i \otimes \mathbf{e}_j)(\varphi,\psi).
$$

同样，$\begin{Bmatrix}2\\0\end{Bmatrix}$ 价的一般张量 $K(\varphi,\psi)$ 可以分解为基张量：

$$
\boxed{\mathbf{K} = K^{ij}(\mathbf{e}_i \otimes \mathbf{e}_j).}
$$

显然，上述技术可以应用于任意价 $\begin{Bmatrix}f\\g\end{Bmatrix}$ 的张量，使用基张量将它们分解为分量，

$$(\mathbf{e}_{i1} \otimes \mathbf{e}_{i2} \otimes \cdots \otimes \mathbf{e}_{i\mathfrak{f}}) \otimes (\mathbf{d}x^{j1} \otimes \mathbf{d}x^{j2} \otimes \cdots \otimes \mathbf{d}x^{j\mathfrak{v}}).$$

请注意，$\left\{\begin{matrix}\mathfrak{f}\\ \mathfrak{v}\end{matrix}\right\}$ 价张量的分量有 $\mathfrak{f}$ 个上标和 $\mathfrak{v}$ 个下标[1]。

33.5 度量张量与经典线元的关系

前面我们讨论了现代梯度 1-形式 $\mathbf{d}f$ 与经典的、基于无穷小的公式的微分 df （33.9）密切相关的事实。 类似地，现在让我们考虑现代度量张量如何与高斯（1827）的经典的、基于无穷小的**线元**（line element） ds 相关联。

高斯考虑了曲面上一点坐标中的极小变化 du 和 dv，导致曲面内的极小线元 ds。 然后，在高斯的原始[2]符号中，

$$ds^2 = E\,du^2 + 2F\,du\,dv + Gdv^2.$$

代替无穷小，我们改为使用终极等式，将其表示为

$$\delta s^2 \asymp E\,\delta u^2 + 2F\,\delta u\,\delta v + G\delta v^2.$$

现在让

$$x^1 = u, x^2 = v \quad 和 \quad g_{11} = E, \qquad g_{12} = g_{21} = F, \qquad g_{22} = G.$$

然后，使用（33.2），度量张量可以表示为

$$\mathbf{g} = g_{ij}(\mathbf{d}x^i \otimes \mathbf{d}x^j) = E\,(\mathbf{d}u \otimes \mathbf{d}u) + F\,(\mathbf{d}u \otimes \mathbf{d}v) + F\,(\mathbf{d}v \otimes \mathbf{d}u) + G(\mathbf{d}v \otimes \mathbf{d}v),$$

这看起来与高斯的 ds^2 公式非常相似。

为了使连接更明确，令 $\delta r = \begin{bmatrix}\delta u\\ \delta v\end{bmatrix}$ 是坐标图平面中的一个短（最终消失）向量，从而导致表面内有一个小的移动 δs。然后，由于

$$\mathbf{d}u\begin{bmatrix}\delta u\\ \delta v\end{bmatrix} = \delta u, \qquad 和 \qquad \mathbf{d}v\begin{bmatrix}\delta u\\ \delta v\end{bmatrix} = \delta v,$$

我们确实从度量张量中恢复了线元公式：

---

[1]在较早的文献中，较高的指数称为逆变，较低的指数称为协变。

[2]实际上，正如我们之前提到的，Gauss (1827) 使用 p 和 q 而不是 u 和 v。

$$
\begin{aligned}
\delta s^2 &\asymp \mathbf{g}(\delta\mathbf{r}, \delta\mathbf{r}) \\
&= [E\,(\mathbf{d}u \otimes \mathbf{d}u) + F\,(\mathbf{d}u \otimes \mathbf{d}v) + F\,(\mathbf{d}v \otimes \mathbf{d}u) + G(\mathbf{d}v \otimes \mathbf{d}v)]\left(\begin{bmatrix}\delta u\\ \delta v\end{bmatrix}, \begin{bmatrix}\delta u\\ \delta v\end{bmatrix}\right) \\
&= E\,\delta u^2 + F\,\delta u\,\delta v + F\,\delta v\,\delta u + G\delta v^2.
\end{aligned}
$$

33.6 示例：线性代数（再次）

传统线性代数课程中熟悉的矩阵与作为研究 $\begin{Bmatrix}1\\1\end{Bmatrix}$ 价张量的主题看似抽象的张量视图之间有什么联系？ 本节使用分量和张量基的概念来解释连接。

首先，我们的张量 $\mathbf{L}(\varphi \parallel \mathbf{v})$ 有分量

$$L^i{}_j = \mathbf{L}(\mathbf{d}x^i \parallel \mathbf{e}_j).$$

为了使事情尽可能简单和具体，我们将自己限制在两个维度上。

$$\{\mathbf{e}_1, \mathbf{e}_2\} = \left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\},$$

然后，如第 32.3.4 节所述，对偶 1-形式基对应于行向量：

$$\{\mathbf{d}x^1, \mathbf{d}x^2\} = \{[1\ 0], [0\ 1]\}.$$

因此，

$$\mathbf{d}x^1(\mathbf{v}) = [1\ 0]\begin{bmatrix}v_1\\v_2\end{bmatrix} = v_1,$$

本该如此。 同样适用于 $\mathbf{d}x^2$。

因此，$\mathbf{L}$ 确实由一个熟悉的 2×2 矩阵表示，其条目是张量的分量：

$$
\begin{aligned}
\mathbf{L} &= L^i{}_j\,\mathbf{e}_i \otimes \mathbf{d}x^j \\
&= L^1{}_1\begin{bmatrix}1\\0\end{bmatrix}[1 \quad 0] + L^1{}_2\begin{bmatrix}1\\0\end{bmatrix}[0 \quad 1] + L^2{}_1\begin{bmatrix}0\\1\end{bmatrix}[1 \quad 0] + L^2{}_2\begin{bmatrix}0\\1\end{bmatrix}[0 \quad 1] \\
&= L^1{}_1\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + L^1{}_2\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + L^2{}_1\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} + L^2{}_2\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} \\
&= \begin{bmatrix}L^1{}_1 & L^1{}_2\\L^2{}_1 & L^2{}_2\end{bmatrix}.
\end{aligned}
$$

在线性代数中，这个矩阵通过标准矩阵乘法描述了一个线性变换：

$$\begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \to \begin{bmatrix} L^1{}_1 & L^1{}_2 \\ L^2{}_1 & L^2{}_2 \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} = \begin{bmatrix} L^1{}_1 v^1 + L^1{}_2 v^2 \\ L^2{}_1 v^1 + L^2{}_2 v^2 \end{bmatrix} = \begin{bmatrix} L^1{}_j v^j \\ L^2{}_j v^j \end{bmatrix}$$

换句话说，

$$v^i \longrightarrow L^i{}_j \; v^j \quad \Leftrightarrow \quad \mathbf{v} = v^i \mathbf{e}_i \longrightarrow (L^i{}_j \; v^j) \, \mathbf{e}_i.$$

以这种形式表示，我们现在可以看到这与我们在第 33.2 节中的讨论是等价的，其中我们使用张量来定义线性变换 $L(\mathbf{v})$：

$$\mathbf{v} \longrightarrow L(\mathbf{v}) = \mathbf{L}(\ldots \| \; \mathbf{v}) = L^i{}_j \, (\mathbf{e}_i \otimes \mathbf{d}x^j) \, (\ldots, \mathbf{v}) = L^i{}_j \, ei(\ldots) \, \mathbf{d}x^j(\mathbf{v}) = (L^i{}_j v^j) \, \mathbf{e}_i(\ldots).$$

33.7 缩并

从分量的角度考虑 1-形式和向量的缩并：

$$\varphi(\mathbf{v}) = (\varphi_i \, \mathbf{d}x^i)(v^j \, \mathbf{e}_j) = (\varphi_i \, v^j) \, \mathbf{d}x^i(\mathbf{e}_j) = (\varphi_i \, v^j) \, \delta^i_j = \varphi_j \, v^j.$$

正如我们所知，这种缩并是一种几何运算，答案与 φ 和 $\mathbf{v}$ 的特定分量无关，尽管这些分量确实取决于对偶基 $\{\mathbf{e}_j\}$ 和 $\{\mathbf{d}x^i\}$ 的选择。

接下来，考虑上一节中张量的分量 $L^i{}_j$。 类似地，我们现在定义这个张量的**缩并**为

$$L^i{}_j \; = \mathbf{L}(\mathbf{d}x^j \; \| \; \mathbf{e}_j) = L11 \; + L22 \; = \mathrm{Tr} \begin{bmatrix} L^1{}_1 & L^1{}_2 \\ L^2{}_1 & L^2{}_2 \end{bmatrix}.$$

正如我们从线性代数中知道的那样，这条迹也是与基无关的[3]。

我们将证明这个基独立结果的一个非常广义的版本——使用纯张量方法，并且不将自己限制在二维——但首先我们必须解释在一般情况下缩并意味着什么。

缩并的想法可以应用于任何张量，其输入包括至少一个 1-形式和至少一个向量：我们对一个上指标和一个下指标求和。我们已经遇到了这种缩并的一个非常重要和自然的例子，即黎曼张量的缩并产生里奇（Ricci）张量：

$$R_{mij}{}^m \; = R_{ij}.$$

[3] 这有一个几何原因，尽管在线性代数课本中很难找到。 参见 Arnol'd (1973, §16.3)。

请注意，此缩并过程将价从 $\begin{Bmatrix}1\\3\end{Bmatrix}$ 更改为 $\begin{Bmatrix}0\\2\end{Bmatrix}$。 一般来说，新的缩并张量的输入槽接受少一个 1-形式和少一个向量，从而消除一个上指标和一个下指标。

事实上，缩并在张量积的背景下具有更广泛的含义。 假设我们形成 $\mathbf{A}\otimes\mathbf{B}$，然后对 $\mathbf{A}$ 的上指标和 $\mathbf{B}$ 的下指标进行求和。让我们证明这个缩并的结果是另一个张量，独立于特定的 $\{\mathbf{e}_j\}$ 和 $\{\mathbf{d}x^i\}$ 导致分量被求和。

按照 Schutz (1980, §2.25) 的例子，我们将通过一个具体的例子来传达证明的思想。令 $\mathbf{A}$ 为 $\begin{Bmatrix}2\\0\end{Bmatrix}$ 价，$\mathbf{B}$ 为 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价。 让我们证明缩并

$$A^{ij}B_{jk} \equiv C^i{}_k$$

产生 $\begin{Bmatrix}1\\1\end{Bmatrix}$ 价的新张量 $\mathbf{C}$ 的分量，使得

$$\mathbf{C}(\varphi \parallel \mathbf{v}) = C^i{}_k\, \varphi_i\, v^k.$$

为了证明这一点，首先观察

$$\begin{aligned} C^i{}_k\, \varphi_i\, v^k &= A^{ij}\, B^{jk}\, \varphi_i\, v^k \\ &= [\varphi_i\, \mathbf{A}(\mathbf{d}x^i, \mathbf{d}x^j)]\, [\mathbf{B}(\mathbf{e}_j, \mathbf{e}_k)\, v^k] \\ &= \mathbf{A}(\varphi_i\, \mathbf{d}x^i, \mathbf{d}x^j)\, \mathbf{B}(\mathbf{e}_j, v^k\, \mathbf{e}_k) \\ &= \mathbf{A}(\varphi, \mathbf{d}x^j)\, \mathbf{B}(\mathbf{e}_j, \mathbf{v}). \end{aligned}$$

但是 $\mathbf{B}(\mathbf{e}_j, \mathbf{v})$ 是数字，因此 $\mathbf{A}$ 在第二个槽中的线性意味着

$$C^i{}_k\, \varphi_i\, v^k = \mathbf{A}(\varphi, \mathbf{B}(\mathbf{e}_j, \mathbf{v})\, \mathbf{d}x^j).$$

但是，对于固定 $\mathbf{v}$，

$$\mathbf{B}(\mathbf{e}_j, \mathbf{v})\, \mathbf{d}x^j = \mathbf{B}(\ldots, \mathbf{v})$$

是一个 1-形式，需要插入一个向量来产生一个数字。 所以，

$$\mathbf{C}(\varphi \parallel \mathbf{v}) = \mathbf{A}(\varphi, \mathbf{B}(\ldots, \mathbf{v}))$$

确实是张量。

33.8 使用度量张量更改价

我们已经知道 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价的度量张量 $\mathbf{g}$ 是流形的基本结构：它为我们提供了测地线、平行移动和曲率。 但是 $\mathbf{g}$ 还起着另一个至关重要的作用：它允许我们改变张量的价。 第一步是查看度量如何允许我们将特定的 1-形式与特定的向量相关联，反之亦然。

如果我们让 $\mathbf{g}$ 的一个槽打开，并将向量 $\mathbf{n}$ 插入到另一个槽中，我们会得到一个对应于 $\mathbf{n}$ 的唯一 1-形式 ν：

$$\boxed{\text{向量 } \mathbf{n} \longrightarrow \text{1-形式 } \boldsymbol{\nu}\text{，其中 } \nu(\mathbf{w}) \equiv \mathbf{g}(\mathbf{w},\mathbf{n})\text{。}} \tag{33.3}$$

1-形式 $\boldsymbol{\nu}$ 的分量与原始向量 $\mathbf{n}$ 的分量有何关系？ 我们只需要将 ν 应用于基向量。但首先…

符号警告：通用约定是对应于向量 $\mathbf{n}$ 的 1-形式的分量表示为 n_i，这违反了我们在 1-形式和向量之间的希腊/罗马二分法。从此以后，读者必须比以前更加警惕：如果指标在上，则它属于向量，如果指标在下，则它属于 1-形式。

有了这个约定，

$$\nu = n_i\, \mathbf{d}x^i.$$

因此，

$$n_i = \boldsymbol{\nu}(\mathbf{e}_i) = \mathbf{g}(\mathbf{e}_i,\mathbf{n}) = \mathbf{g}(\mathbf{e}_i, n^j\, \mathbf{e}_j) = \mathbf{g}(\mathbf{e}_i,\mathbf{e}_j) n^j,$$

所以，

$$\boxed{n_i = g_{ij}\, n^j.} \tag{33.4}$$

例如，如果 $\mathbf{n} = n^j\, \mathbf{e}_j$ 是闵可夫斯基时空中的一个向量，度量为（30.6），那么 [练习] 对应的 1-形式是

$$\boldsymbol{\nu} = n_i\, \mathbf{d}x^i = n^0\, \mathbf{d}t - n^1\, \mathbf{d}x - n^2\, \mathbf{d}y - n^3\, \mathbf{d}z.$$

该度量将向量转换为 1-形式的能力使我们能够改变任何张量的价。 例如，考虑 $\begin{Bmatrix}1\\3\end{Bmatrix}$ 价的黎曼张量 $\mathbf{R}(\psi \parallel \mathbf{u},\mathbf{v},\mathbf{w})$，其分量为 $R_{ijk}{}^m$。 让我们演示如何通过将单个 1-形式输入更改为向量输入，将其更改为 $\begin{Bmatrix}0\\4\end{Bmatrix}$ 价张量。 按照惯例，这个新张量仍用 $\mathbf{R}$ 表示，它的分量

写成 R_{ijkl}。

为了评估具有四个向量输入的新张量，我们只需将额外的向量输入 $\mathbf{n}$ 替换为其在黎曼张量的原始定义中对应的 1-形式 ν：

$$\mathbf{R}(\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{n}) \equiv \mathbf{R}(\nu \parallel \mathbf{u},\mathbf{v},\mathbf{w}).$$

在分量形式中，这个方程变为

$$R_{ijkl}u^i v^j w^k n^l = R_{ijk}{}^m u^i v^j w^k n_m = R_{ijk}{}^m u^i v^j w^k (g_{ml}\, n^l).$$

所以，

$$\boxed{R_{ijkl} = R_{ijk}{}^m\, g_{ml}.}$$

这个过程被称为（从逻辑上讲）指标下降（index lowering）。

也可以朝相反的方向前进，将向量输入变为 1-形式输入； 这导致指标上升（index raising）。 为了遵循与之前相同的路径 a，我们需要从 1-形式 ν 到向量 $\mathbf{n}$ 的映射。 所需要的是例如度量张量的模拟，现在是 $\left\{\begin{matrix}2\\0\end{matrix}\right\}$ 价，采用两个 1-形式作为输入。

然后像以前一样获得映射，将 ν 插入一个槽，使另一个槽打开，从而得到一个向量：

$$\boxed{1\text{−形式 } \nu \rightarrow \text{向量 } \mathbf{n}\text{，其中 } \mathbf{n}(\varphi) \equiv \tilde{\mathbf{g}}(\varphi,\nu)\text{。}} \tag{33.5}$$

在分量方面，我们有，

$$\boxed{n^i = \tilde{g}^{ik}\, n_k.} \tag{33.6}$$

而且，出于同样的原因，我们现在可以提升任何我们希望的张量指标，例如，

$$\boxed{R_{ijkl}\, \tilde{g}^{km} = R_{ij}{}^m{}_l}$$

最后一步是认识到 $\tilde{g}$ 是由度量 g 唯一确定的，如下所示。 降低和提升指标必须是逆

操作：如果我们降低一个指标然后再次提高它，我们应该获得与开始时相同的张量。 换句话说，使用 (33.3) 将向量更改为 1-形式，然后使用 (33.5) 再次将其更改回来，应该会产生恒等式。 换句话说，$\tilde{g}$ 是 g 的倒数：

$$\mathbf{n} \overset{g}{\to} \boldsymbol{\nu} \overset{\tilde{g}}{\to} \mathbf{n}.$$

将（33.4）代入（33.6），这意味着

$$n^i = \tilde{g}^{ik} n_k = \tilde{g}^{ik} g_{kj} n^j.$$

因此，例如的分量必须通过以下方式与度量的分量相关联

$$\boxed{\tilde{g}^{ik} g_{kj} = \delta^i_j.} \tag{33.7}$$

正如我们总是使用相同的符号 $\mathbf{R}$ 来表示黎曼张量，不管它接受作为输入的 1-形式和向量的数量，所以传统的（如果令人困惑）简单地写 $\mathbf{g}$ 而不是 $\tilde{\mathbf{g}}$，并相应地写其分量为 g^{ik}。因此，(33.7) 通常写成 $g^{ik} g_{kj} = \delta^i_j$。

33.9 对称与反对称

回想偶函数 $f^+(x)$ 和奇函数 $f^-(x)$ 的定义属性：

$$f^+(-x) = +f^+(x) \quad 和 \quad f^-(-x) = -f^-(x).$$

我们可以说 $f^+(x)$ 是对称的，而 $f^-(x)$ 是反对称的来描述这一点。 在几何上，图 $y = f^+(x)$ 在 y 轴上具有镜像对称性，例如 $y = x^2$ 或 $y = \cos x$，而 $y = f^-(x)$ 在 x 及其镜像 $-x$ 处相反， 例如，$y = x^3$ 或 $y = \sin x$。

假设我们尝试将任意函数 $f(x)$（缺乏任何特殊对称性）拆分为对称函数和反对称函数之和：

$$f(x) = f^+(x) + f^-(x).$$

如果这是可能的，那么，根据定义，

$$f(-x) = f^+(x) - f^-(x).$$

通过将前两个方程相加然后相减，我们推断拆分总是可能的，并且它明确地由下式给出

$$f^+(x) = \frac{f(x) + f(-x)}{2} \quad 和 \quad f^-(x) = \frac{f(x) - f(-x)}{2}.$$

例如，如果我们取 $f(x) = e^x$，那么这种拆分自然会导致我们得到双曲函数：e^x 的对称部分是 $f^+(x) = \cosh x$，反对称部分是 $f^-(x) = \sinh x$.

现在让我们尝试用 $\begin{Bmatrix} 0 \\ 2 \end{Bmatrix}$ 价的一般张量 $\mathbf{E}(\mathbf{v}, \mathbf{w})$ 做一些类似的事情。 以此类推，让**对称张量** $\mathbf{E}^+$ 和**反对称**[4]**张量** $\mathbf{E}^-$ 的定义性质为

$$\mathbf{E}^+(\mathbf{w}, \mathbf{v}) = +\mathbf{E}^+(\mathbf{v}, \mathbf{w}) \quad 和 \quad \mathbf{E}^-(\mathbf{w}, \mathbf{v}) = -\mathbf{E}^-(\mathbf{v}, \mathbf{w}).$$

通过与之前相同的推理，我们发现总是可以将 $\begin{Bmatrix} 0 \\ 2 \end{Bmatrix}$ 价的一般张量拆分为对称部分和反对称部分：

$$\mathbf{E}(\mathbf{v}, \mathbf{w}) = \mathbf{E}^+(\mathbf{v}, \mathbf{w}) + \mathbf{E}^-(\mathbf{v}, \mathbf{w}), \tag{33.8}$$

其中

$$\mathbf{E}^+(\mathbf{v}, \mathbf{w}) = \left[\frac{\mathbf{E}(\mathbf{v}, \mathbf{w}) + \mathbf{E}(\mathbf{w}, \mathbf{v})}{2}\right] \quad 和 \quad \mathbf{E}^-(\mathbf{v}, \mathbf{w}) = \left[\frac{\mathbf{E}(\mathbf{v}, \mathbf{w}) - \mathbf{E}(\mathbf{w}, \mathbf{v})}{2}\right].$$

如果我们将基向量插入槽中，我们将获得这些方程的分量形式。 在我们讨论的同时，让我们还介绍以下标准符号：**圆括号表示对称，方括号表示反对称：**

$$\boxed{E_{(ij)} \equiv E^+_{ij} = \frac{1}{2}\left[E_{ij} + E_{ji}\right] \quad 和 \quad E_{[ij]} \equiv E^-_{ij} = \frac{1}{2}\left[E_{ij} - E_{ji}\right]} \tag{33.9}$$

所以，

$$\boxed{E_{ij} = E_{(ij)} + E_{[ij]}.}$$

[4] 也通常称为斜对称（skew symmetric）。

第三十四章 2-形式

34.1 2-形式和 p 形式的定义

我们已经遇到了几个至关重要的 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价对称张量：度量张量、里奇（Ricci）张量、能量-动量张量和爱因斯坦张量。

因此，得知将推动我们的戏剧走向结局的嘉当形式的*秘密动力来源是反对称*，这可能会让人感到惊讶！

让我们立即向你介绍第二个嘉当形式，出现在 1-形式之后第一个：

$$\boxed{\begin{array}{c}2\text{−形式 }\Psi\text{ 是 }\begin{Bmatrix}0\\2\end{Bmatrix}\text{ 价的反对称张量。 即，}\\ \Psi(\mathbf{v},\mathbf{u})=-\Psi(\mathbf{u},\mathbf{v}).\end{array}} \tag{34.1}$$

作为简单的家务，让我们立即陈述 p-形式的一般定义，充分了解如果没有具体示例，这对你来说基本上没有意义：

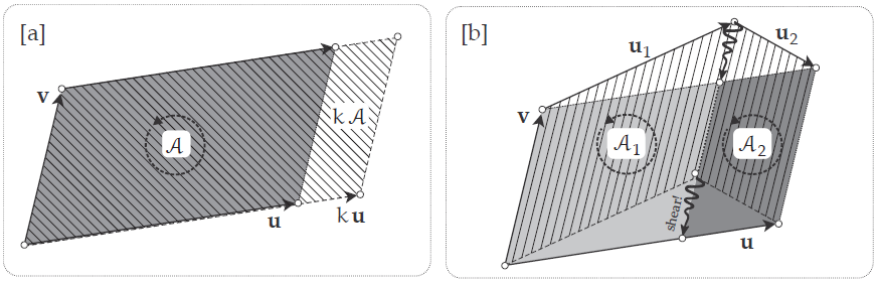
$$\boxed{\begin{array}{l}\text{p−形式（也称为 p 阶的微分形式）是 }\begin{Bmatrix}0\\p\end{Bmatrix}\text{ 价的完全反对称}\\ \text{张量，这意味着交换任意两个向量输入会反转其符号。}\end{array}}$$

自然地，如果我们取 $p=2$，那么我们就恢复了 2-形式的定义（34.1）。

我们现在的计划是通过逐步增加 p 来建立对 p-形式的直观的几何理解，一次一步。第五幕以一整章专门讨论 1-形式开始，这一章同样完全致力于理解 2-形式。 之后，我们将转向 3-形式，…

这听起来像是（写出）一本无限长的书的秘诀！ 幸运的是，当我们到达 3-形式时，理解一般 p-形式所需的所有基本思想都将发挥作用。 此外，正如我们将在下一章中解释的那样，可以存在于 n-流形中的最高 p-形式是 n-形式，并且由于时空只有四个维度，因此对于本书而言，4-形式就足够了。 （注意：然而，在形式的一些重要应用中需要更高的 p 值，例如在哈密顿力学中自然出现的辛流形；参见阿诺德（Arnol'd，1989，Ch. 8）。

但是，为什么首先要在这种反对称张量上投入宝贵的时间和精力呢？ 毕竟，除了黎曼张量之外，到目前为止我们遇到的每个重要张量都是对称的！



[34.1] **定向面积是 2-形式的几何证明**。 [a] 将一条边扩大 k 倍，面积扩大 k 倍：$\mathcal{A}(\mathrm{k}\mathbf{u},\mathbf{v}) = \mathrm{k}\mathcal{A}(\mathbf{u},\mathbf{v})$。 [b] 沿 $\mathbf{v}$ 方向剪切 $\mathcal{A}_1 = \mathcal{A}(\mathbf{u}_1,\mathbf{v})$ 和 $\mathcal{A}_2 = \mathcal{A}(\mathbf{u}_2,\mathbf{v})$ 保持它们的面积，所以 $\mathcal{A} = \mathcal{A}(\mathbf{u},\mathbf{v}) = \mathcal{A}_1 + \mathcal{A}_2$。

只有逐渐地，埃利·嘉当像狐狸一样疯狂的程度才会变得明显。 就像 1-形式的情况一样，2-形式也是如此：嘉当认出他们一直躲在视线之外！

34.2 示例：面积 2-形式

在 $\mathbb{R}^2$ 中，让我们定义

$$\boxed{\mathcal{A}(\mathbf{u},\mathbf{v}) = \text{具有边 } \mathbf{u} \text{ 和 } \mathbf{v} \text{ 的平行四边形的定向面积。}}$$

那么 $\mathcal{A}$ 是一个 2-形式！

首先，很明显 $\mathcal{A}(\mathbf{u},\mathbf{v})$ 是反对称的：如果我们交换 $\mathbf{u}$ 和 $\mathbf{v}$，那么面积的大小不会改变，但平行四边形的方向是相反的。 仍有待验证 $\mathcal{A}$ 是张量，即它在每个槽中是线性的。 正如我们对 1-形式所做的那样，我们可以将线性要求分解为两部分：（32.2）和 （32.3），应用于每个插槽。

（32.3）的真值由[34.1a] 解释，它说明如果我们将平行四边形的任一边扩大 k，那么它的面积也扩大 k：

$$\mathcal{A}(\mathrm{k}\mathbf{u},\mathbf{v}) = \mathcal{A}(\mathbf{u},\mathrm{k}\mathbf{v}) = \mathrm{k}\mathcal{A}(\mathbf{u},\mathbf{v}).$$

请注意，如果 Ψ 是任意 2-形式，那么如果（32.3）对一个插槽为真，那么它对另一个插槽也必须为真：

$$\Psi(\mathrm{k}\mathbf{u},\mathbf{v}) = \mathrm{k}\Psi(\mathbf{u},\mathbf{v}) \implies \Psi(\mathbf{u},\mathrm{k}\mathbf{v}) = -\Psi(\mathrm{k}\mathbf{v},\mathbf{u}) = -\mathrm{k}\Psi(\mathbf{v},\mathbf{u}) = \mathrm{k}\Psi(\mathbf{u},\mathbf{v}).$$

（32.2）的真值不太明显，但刚刚给出的论证意味着我们只需要证明第一个插槽的真值——那么第二个插槽的真值就来自反对称。

令 $\mathbf{u}=\mathbf{u}_1+\mathbf{u}_2$，并定义

$$\mathcal{A}=\mathcal{A}(\mathbf{u},\mathbf{v}),\mathcal{A}_1=\mathcal{A}(\mathbf{u}_1,\mathbf{v}),\mathcal{A}_2=\mathcal{A}(\mathbf{u}_2,\mathbf{v}).$$

那么（32.2）要求

$$\boxed{\mathcal{A}=\mathcal{A}_1+\mathcal{A}_2.}$$

这确实是真实的，在[34.1b]中得到了几何证明。 因此 $\mathcal{A}$ 确实是一个 2-形式。

请注意，如果 Ψ 是一般的 2-形式，则两个向量输入交换下的反对称意味着 $\Psi(\mathbf{u},\mathbf{u})=-\Psi(\mathbf{u},\mathbf{u})$，因此对于任意 $\mathbf{u}$，

$$\boxed{\Psi(\mathbf{u},\mathbf{u})=0.} \tag{34.2}$$

在面积 2-形式的情况下，这具有直接的几何意义：如果在 $\mathcal{A}(\mathbf{u},\mathbf{v})$ 中我们让 $\mathbf{v}$ 变为 $\mathbf{u}$，则平行四边形坍缩为面积为零的线段，因此 $\mathcal{A}(\mathbf{u},\mathbf{v})=0$。

事实上，（34.2）等价于我们最初基于反对称的 2-形式定义（34.1）。 为了看到这一点，我们回收了我们之前在黎曼张量的上下文中使用的参数。 令 $\mathbf{x}$ 和 $\mathbf{y}$ 为任意向量，令 $\mathbf{u}=\mathbf{x}+\mathbf{y}$。 然后，

$$\begin{aligned}0 &= \Psi(\mathbf{u},\mathbf{u})\\ &= \Psi([\mathbf{x}+\mathbf{y}],[\mathbf{x}+\mathbf{y}])\\ &= \Psi(\mathbf{x},\mathbf{x})+\Psi(\mathbf{x},\mathbf{y})+\Psi(\mathbf{y},\mathbf{x})+\Psi(\mathbf{y},\mathbf{y})\\ &= 0+\Psi(\mathbf{x},\mathbf{y})+\Psi(\mathbf{y},\mathbf{x})+0.\end{aligned}$$

因此，$\Psi(\mathbf{y},\mathbf{x})=-\Psi(\mathbf{x},\mathbf{y})$。 证毕。

34.3 两个 1-形式的楔积

回忆（33.8）中的任何 $\left\{\begin{matrix}0\\2\end{matrix}\right\}$ 价张量都可以拆分为对称张量和反对称张量之和。 如果我们将此过程应用于两个任意 1-形式的张量积，那么通过交换张量积中的 1-形式（而不是交换向量），我们发现 [练习]

$$\varphi \otimes \psi = \frac{1}{2}[\varphi \otimes \psi + \psi \otimes \varphi] + \frac{1}{2}[\varphi \otimes \psi - \psi \otimes \varphi].$$

根据定义，它的反对称部分是通过将两个 1-形式相乘而创建的 2-形式。这是一种新的乘法，称为楔积[1]（wedge product），表示为 ∧：

$$\boxed{\varphi \wedge \psi \equiv \varphi \otimes \psi - \psi \otimes \varphi.} \tag{34.3}$$

正如张量积允许我们系统地从低价张量构建高价张量一样，楔积也允许我们从低价张量构建高价形式。

让我们强调一下，使 $(\varphi \wedge \psi)$ 成为 2-形式的原因是，当它作用于一对向量时，它在两个插槽中都是线性的，并且交换两个向量输入只会反转输出的符号，而保持其大小固定不变：

$$(\varphi \wedge \psi)(\mathbf{v}_1, \mathbf{v}_2) = \varphi(\mathbf{v}_1)\psi(\mathbf{v}_2) - \psi(\mathbf{v}_1)\varphi(\mathbf{v}_2) = -(\varphi \wedge \psi)(\mathbf{v}_2, \mathbf{v}_1).$$

但请注意存在第二个反对称，即楔积本身的反对称。保持向量输入的顺序相同，但交换楔积的 1-形式，我们看到

$$(\varphi \wedge \psi)(\mathbf{v}_1, \mathbf{v}_2) = -(\psi \wedge \varphi)(\mathbf{v}_1, \mathbf{v}_2).$$

抽象出向量，我们得到了楔积本身的反对称，这在（34.3）的定义中很明显：

$$\boxed{(\varphi \wedge \psi) = -(\psi \wedge \varphi).}$$

请注意，这意味着，对于任意 ψ，

$$\boxed{\psi \wedge \psi = 0.}$$

另请注意，楔积和加法服从分配率：

$$\boxed{\varphi \wedge (\psi + \sigma) = \varphi \wedge \psi + \varphi \wedge \sigma.}$$

在 $\mathbb{R}^2$ 中，考虑笛卡尔基 1-形式的所有可能的楔积。由于 $\mathbf{d}\mathrm{x} \wedge \mathbf{d}\mathrm{x} = 0$ 和 $\mathbf{d}\mathrm{y} \wedge \mathbf{d}\mathrm{y} =$

[1]也称为外积（exterior product）。

0，唯一的非零楔积是 $\mathbf{d}x \wedge \mathbf{d}y = -\mathbf{d}y \wedge \mathbf{d}x$。 而且，事实上，这正是面积 2-形式！

$$\boxed{\mathcal{A} = \mathbf{d}x \wedge \mathbf{d}y.} \tag{34.4}$$

请注意，这再次与面积的经典表达式有惊人的相似之处：当我们进行二重积分时，我们将面积的元素写为 $dx\,dy$。 显式连接遵循与我们之前的微分 df 和线元 ds 示例相同的模式。 如果我们把 $\mathbf{u} = \delta x\,\mathbf{e}_1$ 和 $\mathbf{v} = \delta y\,\mathbf{e}_2$ 作为一个小矩形的边，那么

$$(\mathbf{d}x \wedge \mathbf{d}y)(\mathbf{u}, \mathbf{v}) = (\mathbf{d}x \wedge \mathbf{d}y)(\delta x\,\mathbf{e}_1, \delta y\,\mathbf{e}_2) = \delta x\,\delta y.$$

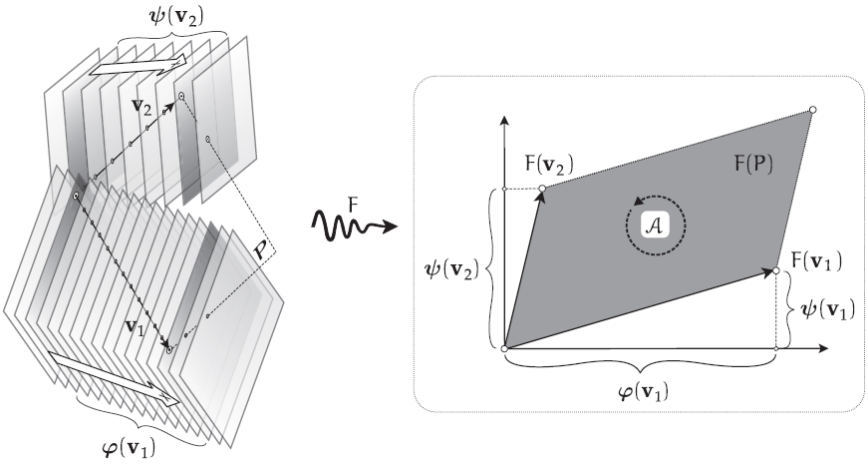
使用分量很容易实现（34.4）的一般证明：

$$\begin{aligned}(\mathbf{d}x \wedge \mathbf{d}y)(\mathbf{u}, \mathbf{v}) &= (\mathbf{d}x \otimes \mathbf{d}y - \mathbf{d}y \otimes \mathbf{d}x)\left(\begin{bmatrix} u^1 \\ u^2 \end{bmatrix}, \begin{bmatrix} v^1 \\ v^2 \end{bmatrix}\right) \\ &= u^1 v^2 - u^2 v^1 \\ &= \det\begin{bmatrix} u^1 & v^1 \\ u^2 & v^2 \end{bmatrix} \\ &= \mathcal{A}(\mathbf{u}, \mathbf{v}).\end{aligned}$$

现在让我们描述作用于 $\mathbb{R}^n$ 中向量的两个一般 1-形式的楔积 $(\varphi \wedge \psi)$ 的几何意义。该论证将适用于任何 n，但为了使这个想法生动，让我们在 $\mathbb{R}^3$ 的情况下对其进行说明。在这种情况下，φ 和 ψ 都由平面堆栈表示，它们对 $\mathbb{R}^3$ 中向量 $\mathbf{v}$ 的作用是计算每个堆栈的多少平面被穿透，如 [34.2] 所示。

现在关键的想法来了：让我们使用这两个数字 $\varphi(\mathbf{v})$ 和 $\psi(\mathbf{v})$ 作为 $\mathbb{R}^2$ 中一个点的 x 和 y 坐标。 因此，这两个 1-形式定义了 $\mathbb{R}^3$（或任何 $\mathbb{R}^n$）中的向量到 $\mathbb{R}^2$ 中的向量的映射 F：

$$\boxed{v \longrightarrow F(\mathbf{v}) \equiv \begin{bmatrix} \varphi(\mathbf{v}) \\ \psi(\mathbf{v}) \end{bmatrix}} \tag{34.5}$$



[34.2] **楔积的几何意义**。 $\mathbb{R}^3$ 中的平行四边形 P 有边 $\mathbf{v}_1$ 和 $\mathbf{v}_2$，$\varphi(\mathbf{v}_k)$ 和 $\psi(\mathbf{v}_k)$ 计算每个堆栈的多少平面被边 $\mathbf{v}_k$ 穿透。 映射 F 将这些穿孔数对 $(\varphi(\mathbf{v}_k), \psi(\mathbf{v}_k))$ 融合到 $\mathbb{R}^2$ 中的笛卡尔坐标中，将 P 映射到平行四边形 F(P)。 则 $(\varphi \wedge \psi)(\mathbf{v}_1, \mathbf{v}_2)$ 为 F(P) 的定向面积 $\mathcal{A}$。

在这个映射 F 下，$\mathbb{R}^n$ 中具有边 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 的平行四边形映射到 $\mathbb{R}^2$ 中具有边 $\mathrm{F}(\mathbf{v}_1)$ 和 $\mathrm{F}(\mathbf{v}_2)$ 的平行四边形，如 [34.2] 所示。 我们现在可以陈述楔形积的精美简单的含义[2]：

> 当楔积 $(\varphi \wedge \psi)$ 应用于 $\mathbb{R}^n$ 中的任意平行四边形时，它在映射 F 下输出 $\mathbb{R}^2$ 中图像平行四边形的定向面积：
>
> $$(\varphi \wedge \psi)(\mathbf{v}_1, \mathbf{v}_2) = \mathrm{A}[\mathrm{F}(\mathbf{v}_1), \mathrm{F}(\mathbf{v}_2)].$$

(34.6)

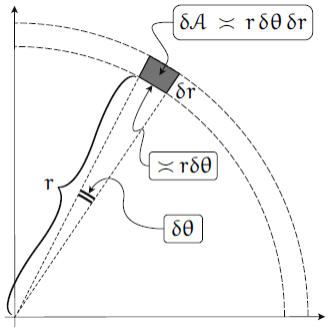
证明很简短：

$$\begin{aligned}
(\varphi \wedge \psi)(\mathbf{v}_1, \mathbf{v}_2) &= \varphi(\mathbf{v}_1)\psi(\mathbf{v}_2) - \psi(\mathbf{v}_1)\varphi(\mathbf{v}_2) \\
&= \det \begin{bmatrix} \varphi(\mathbf{v}_1) & \varphi(\mathbf{v}_2) \\ \psi(\mathbf{v}_1) & \psi(\mathbf{v}_2) \end{bmatrix} \\
&= \mathcal{A}[\mathrm{F}(\mathbf{v}_1), \mathrm{F}(\mathbf{v}_2)].
\end{aligned}$$

34.4 极坐标中的面积 2-形式

当我们在 $\mathbb{R}^2$ 中进行二重积分时，我们将面积元素写为 $\mathrm{d}\mathcal{A} = \mathrm{dx\,dy}$，我们之前已经解释了它如何与面积 2-形式 (34.4) 连接：$\mathcal{A} = \mathrm{dx} \wedge \mathrm{dy}$。 但是当我们切换到极坐标时，我们将面积元素写为 $\mathrm{d}\mathcal{A} = \mathrm{r\,dr\,d\theta}$。 在这种情况下，这个公式有一个简单的几何推导，如[34.3]所示。 因此相应的 2-形式是 $\mathcal{A} = \mathrm{r\,dr} \wedge \mathrm{d\theta}$。

[2] Arnol'd (1989, §32) 更进一步，将 (34.6) 作为楔积的定义。



[34.3] $\delta\mathcal{A} \asymp r\,\delta\theta\,\delta r$ 的几何证明。

但是，当我们进行更复杂的坐标变化时，传统的方法（你无疑是在多元微积分中学过的）需要我们使用雅可比来找到新坐标系中面积元素的表达式。 正如我们现在演示的那样，以极坐标为例，使用面积 2-形式允许我们抛弃雅可比行列式，并将整个事情变成（无意识的）代数儿童游戏！

应用莱布尼茨规则（32.12），并记住 $dr \wedge dr = 0$，$d\theta \wedge d\theta = 0$，$d\theta \wedge dr = -dr \wedge d\theta$，我们发现

$$\begin{aligned}
\mathbf{d}x \wedge \mathbf{d}y &= \mathbf{d}(r\cos\theta) \wedge \mathbf{d}(r\sin\theta) \\
&= [(\mathbf{d}r)\cos\theta - r\sin\theta\,\mathbf{d}\theta] \wedge [(\mathbf{d}r)\sin\theta + r\cos\theta\,\mathbf{d}\theta] \\
&= \cos^2\theta\, r\,\mathbf{d}r \wedge \mathbf{d}\theta - \sin^2\theta\, r\,\mathbf{d}\theta \wedge \mathbf{d}r \\
&= r\,\mathbf{d}r \wedge \mathbf{d}\theta.
\end{aligned}$$

34.5 基 2-形式和投影

很明显，两个 2-形式的和是另一个 2-形式，同样乘以一个常数也会产生另一个 2-形式。因此，2-形式构成了一个向量空间，自然要为这个空间寻找一个基。

回忆（33.2）张量集合 $\{\mathbf{d}x^i \otimes \mathbf{d}x^j\}$ 形成 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价张量的基。 这适用于所有此类张量，包括 2-形式，但在 2-形式的情况下，我们更进一步：

$$\boxed{\text{2-形式的集合 } \{\mathbf{d}x^i \wedge \mathbf{d}x^j\}\text{，其中 } i < j\text{，是 2-形式的基。}} \tag{34.7}$$

条件 $i < j$ 只是为了避免列出重复的 2-形式：例如，$(\mathbf{d}x^3 \wedge \mathbf{d}x^2) = -(\mathbf{d}x^2 \wedge \mathbf{d}x^3)$。

由于 $\mathbf{d}x^i \wedge \mathbf{d}x^i = 0$，非零（非冗余）基 2-形式是通过选择一对无序的不同 $\mathbf{d}x^i$，（从

n 中）然后形成它们的楔积。 所以，如果我们假设（34.7）为真，那么，

$$\boxed{\text{在 } \mathbb{R}^n \text{ 中，2–形式的集合是维数为 } \frac{1}{2}n(n-1) \text{ 的向量空间。}} \tag{34.8}$$

为了确认（34.7），让我们从 $\mathbb{R}^2$ 开始。 如果 Ψ 只是 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价的一般张量，那么我们从（33.2）知道它可以展开为

$$\boldsymbol{\Psi} = \Psi_{11}(\mathbf{d}x \otimes \mathbf{d}x) + \Psi_{12}(\mathbf{d}x \otimes \mathbf{d}y) + \Psi_{21}(\mathbf{d}y \otimes \mathbf{d}x) + \Psi_{22}(\mathbf{d}y \otimes \mathbf{d}y).$$

但是，因为 $\boldsymbol{\Psi}$ 是 2-形式

$$\Psi_{11} = 0 = \Psi_{22} \quad \text{和} \quad \Psi_{21} = -\Psi_{12}.$$

所以，

$$\boxed{\boldsymbol{\Psi} = \Psi_{12}\,(\mathbf{d}x \wedge \mathbf{d}y) = \Psi_{12}\,\mathcal{A}\text{。}}$$

换句话说，$(\mathbf{d}x \wedge \mathbf{d}y)$ 是 $\mathbb{R}^2$ 中 2-形式的一个基，并且每个 2-形式都与面积 2-形式成正比。

如果我们现在上升一维到 $\mathbb{R}^3$，那么完全相同的推理得到[练习]

$$\boxed{\boldsymbol{\Psi} = \Psi_{23}\,(\mathbf{d}y \wedge \mathbf{d}z) + \Psi_{31}\,(\mathbf{d}z \wedge \mathbf{d}x) + \Psi_{12}\,(\mathbf{d}x \wedge \mathbf{d}y),} \tag{34.9}$$

确认（34.7）对于这种情况。（注意：以我们所拥有的方式以及我们所拥有的奇怪顺序编写这三个术语的原因将在下一节中解释。）

$\mathbb{R}^n$ 中的证明是这些示例的直接扩展。 展开式（33.2）中的非零项总是成对出现：

$$\Psi_{ij}(\mathbf{d}x^i \otimes \mathbf{d}x^j) + \Psi_{ji}(\mathbf{d}x^j \otimes \mathbf{d}x^i) = \Psi_{ij}(\mathbf{d}x^i \otimes \mathbf{d}x^j) - \Psi_{ij}(\mathbf{d}x^j \otimes \mathbf{d}x^i) = \Psi_{ij}(\mathbf{d}x^i \wedge \mathbf{d}x^j).$$

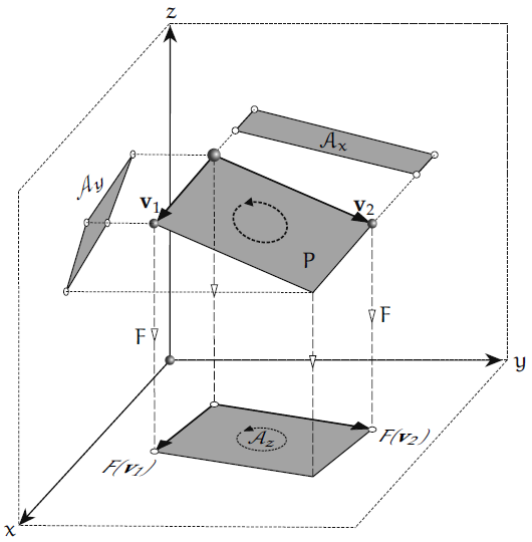
让我们回到 $\mathbb{R}^3$ 并找出（34.9）中出现的基 2-形式的含义，从 $(\mathbf{d}x \wedge \mathbf{d}y)$ 开始。 在 $\mathbb{R}^2$ 中，这只是 (x, y) 平面中的面积 2-形式，与 $\mathbb{R}^3$ 中的含义密切相关，如 [34.4] 所示：

> 如果将 $(\mathbf{d}x \wedge \mathbf{d}y)$ 应用于 $\mathbb{R}^3$ 中的平行四边形 P，则输出是 P 在 (x, y) 平面上的正交投影（在 z 方向上）的定向面积 $\mathcal{A}_z$。

这是两个任意 1-形式的楔积的几何解释（34.6）的直接结果。 关键是要认识到，在这种情况下，映射 F 只是简化为 z 方向上的正交投影到 (x, y) 平面上。 简单比较一下[34.2]中楔积的一般构造：这里 $\varphi = \mathbf{d}x$ 是一个垂直于 x 轴的单位空间栈，同样 $\psi = \mathbf{d}y$ 是一个垂直于 y 轴的单位空间栈。 （请花点时间确保你能真正看到这一切。）

如[34.4]所示，其他两种基 2-形式的含义完全类似：$(dy \wedge dz)$是在(y, z)-平面上的投影（在 x 方向上）的面积 Ax； $(dz \wedge dx)$ 是在 (z, x) 平面上的投影面积 Ay（在 y 方向上）。

在 $\mathbb{R}^4$ 中，(34.9) 的推广包含六个基 2-形式，$(dxi \wedge dxj)$，但每个形式与上面的含义相同：例如，$(dx1 \wedge dx3)$ 产生 $(x1, x3)$-沿 x2 和 x4 轴投影后的平面。



[34.4] **基 2-形式的几何意义**。 每个基 2-形式产生 P 在相关坐标平面上的投影面积。 例如，$(\mathbf{d}x \wedge \mathbf{d}y)(\mathbf{v}_1, \mathbf{v}_2) = \mathcal{A}_z$，P 在 z 方向上在$(x, y)$平面上的投影面积。

34.6 将 2-形式与$\mathbb{R}^3$中的向量关联：通量

正如我们将看到的，可以在任意维数中使用形式（Forms）进行微积分。 但是我们作为本科生在大学里学习的向量分析（Vector Calculus）只有在三个维度上才有可能，在这一节和下一节我们开始理解为什么。

在几何上，2-形式和向量之间肯定不会混淆。 甚至在代数上，区别也很明显：在 n 维中，向量有 n 个分量，而 2-形式有 $\frac{1}{2}n(n-1)$ 个分量，正如我们在（34.8）中看到的。

但这意味着当 $n = 3$ 时会发生一些值得注意的事情：*在三个维度中，并且仅在三个维度中，2−形式具有与向量相同数量的分量。*

向量分析在 1880 年代由物理学家成功开创，在 21 世纪仍被认为是现代科学不可或缺

的工具，现在可以看出它是建立在这个奇异的数值巧合之上的。 虽然向量分析的先驱们无法理解他们的数学引擎为什么会发出如此强大的声音，原因在于它实际上隐藏了 2-形式引擎盖下面，只是伪装成了向量！

为了开始深入了解，让我们用单个上升指标重写一般 2-形式 Ψ 的分量（34.9），并立即将它们与对应向量的分量标识，我们将用与 2-形式同样的符号（但下划线）表示：即 $\underline{\Psi}$，

$$\boxed{\mathbf{\Psi} = \Psi^1\,(\mathbf{d}x^2 \wedge \mathbf{d}x^3) + \Psi^2\,(\mathbf{d}x^3 \wedge \mathbf{d}x^1) + \Psi^3\,(\mathbf{d}x^1 \wedge \mathbf{d}x^3) \;\leftrightarrows\; \underline{\mathbf{\Psi}} = \begin{bmatrix}\Psi^1\\ \Psi^2\\ \Psi^3\end{bmatrix}.} \tag{34.10}$$

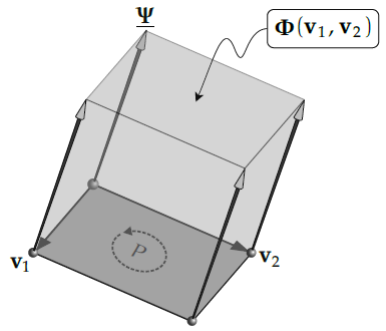
理解这种对应关系的最自然的数学方法是通过霍奇星对偶算子 $(\star)$，它是练习 15 的主题。它是一个纯数学运算，（在 n 维中）将 p-形式映射到 (n− p)-形式，称其为霍奇（Hodge）对偶（通常简写为对偶），以重要的英国数学家霍奇[3]（SirW.V.D. Hodge 1903-1975）命名。

然而，在这里，我们将寻求一个合乎逻辑的和令人信服的物理原因，以将 2-形式与向量相关联，为了理解它，我们现在必须引入物理学家强大的通量[4]（flux）概念。

重新考虑[34.4]，并想象均匀流动的流体以速度 $\underline{\mathbf{\Psi}}$ 通过空间。 现在定义，

$$\boxed{\Phi(\mathbf{v}_1, \mathbf{v}_2) \equiv \text{单位时间通过 P 的流体量} = \underline{\mathbf{\Psi}} \text{ 通过 P 的通量。}} \tag{34.11}$$

如果 P 的方向是逆时针围绕 $\underline{\mathbf{\Psi}}$，我们将通量计为正。 如果我们描绘在单位时间内通过 P 的所有流体，它将填充一个具有边 $\mathbf{v}_1$、$\mathbf{v}_2$、$\underline{\mathbf{\Psi}}$ 的平行六面体，并且通量将是它的体积，如 [34.5] 所示。 它遵循通量 $\mathbf{\Phi}$ 实际上是 2-形式[练习]。



[34.5] $\underline{\mathbf{\Psi}}$ 通过 P 的**通量** =体积 $\mathbf{\Phi}(\mathbf{v}_1, \mathbf{v}_2)$

[3]霍奇是我的数学祖师：我师从彭罗斯，彭罗斯师从霍奇。

[4]我们将使用流体流动来解释通量的概念，但它最初是由法拉第直观地引入来描述他关于电场和磁场穿过空间曲面流动的巧妙实验（始于 1820 年代）； 在随后的几十年里，麦克斯韦赋予了这个概念当前的数学形式。 参见福布斯和马洪 (2014, §10)。

现在很自然地要问以下问题：这种通量 2-形式 $\mathbf{\Phi}$ 与流速 $\underline{\mathbf{\Psi}}$ 相关的 2-形式 $\mathbf{\Psi}$ 有何关系？

我们从一个简单的例子开始。 假设在 [34.4] 中 $\underline{\Psi} = \Psi^3\,\mathbf{e}_3$ 是纯粹向上的，平行于 z 轴。 显然，通过 P 的通量与通过其投影 $\mathcal{A}_z = \mathcal{A}_3$ 的通量相同。 但是，如果我们想象流体以 Ψ^3 的速度向上流动通过 $\mathcal{A}_3$ 一个单位时间，那么流体将充满一个底面积为 $\mathcal{A}_3$、高度为 Ψ^3 的固体，因此通量为 $\Psi^3\mathcal{A}_3$。

在一般情况下，同样的推理适用于 $\underline{\Psi}$ 的每个分量，所以总通量就是简单的总和：

$$\begin{aligned}\mathbf{\Phi}(\mathbf{v}_1,\mathbf{v}_2) &= \Psi^1\mathcal{A}_1 + \Psi^2\mathcal{A}_2 + \Psi^3\mathcal{A}_3\\ &= \Psi^1\,(\mathbf{d}\mathrm{x}^2 \wedge \mathbf{d}\mathrm{x}^3)(\mathbf{v}_1,\mathbf{v}_2) + \Psi^2\,(\mathbf{d}\mathrm{x}^3 \wedge \mathbf{d}\mathrm{x}^1)(\mathbf{v}_1,\mathbf{v}_2) + \Psi^3\,(\mathbf{d}\mathrm{x}^1 \wedge \mathbf{d}\mathrm{x}^2)(\mathbf{v}_1,\mathbf{v}_2)\\ &= \mathbf{\Psi}(\mathbf{v}_1,\mathbf{v}_2).\end{aligned}$$

所以，

$$\mathbf{\Phi} = \mathbf{\Psi}!$$

因此，我们已经获得了通过（34.10）将 2-形式与向量相关联的令人满意的物理理由：

> 如果流体以速度 $\underline{\mathbf{\Psi}}$ 流经 3 维空间，则其通量 2−形式为 $\mathbf{\Psi}$，反之亦然。

(34.12)

除了使用分量将 2-形式 $\mathbf{\Psi}$ 连接到它的关联向量 $\underline{\mathbf{\Psi}}$ 之外，就像我们最初在（34.10）中所做的那样，我们还可以在几何上识别关联向量的方向。 当且仅当它包含方向 $\underline{\mathbf{\Psi}}$ 时，$\underline{\mathbf{\Psi}}$ 穿过平行四边形 P 的通量为零，因此我们可以通过通量为零唯一地表征流动方向：

$$\boxed{\mathbf{\Psi}(\underline{\mathbf{\Psi}},\ldots) = 0.} \tag{34.13}$$

34.7 $\mathbb{R}^3$中向量与楔积的关系

将向量分析（Vector Calculus）从标记为“形式”的大礼帽中提取出来的魔术技巧有第二个分量：我们还将 1-形式与向量相关联，就像我们之前在第 33.8 节中所做的那样。 给定一个 1-形式 φ，我们将再次用与 1-形式相同的符号表示相应的向量，但带下划线：

$$\boxed{\varphi = \varphi_1\,\mathbf{d}x^1 + \varphi_2\,\mathbf{d}x^2 + \varphi_3\,\mathbf{d}x^3 \quad \rightleftarrows \quad \underline{\varphi} = \begin{bmatrix}\varphi_1\\ \varphi_2\\ \varphi_3\end{bmatrix}.} \tag{34.14}$$

向量微积分中的两种基本乘法是标量积（又称点积）和向量积（又称叉积）。 前者很容易推广到任意维度，并且可以表示为

$$\boxed{\varphi(\mathbf{v}) = \underline{\varphi}\cdot\mathbf{v}.} \tag{34.15}$$

但是[34.6]中说明的向量积是三维的。 给定两个向量 $\underline{\varphi}$ 和 $\underline{\sigma}$，从第一个到第二个的角度为 θ，回想一下，我们将向量积 $\underline{\varphi}\times\underline{\sigma}$ 定义为与两者正交的向量，其方向由右手定则给出，其长度作为它们张成的平行四边形的面积：

$$|\underline{\varphi}\times\underline{\sigma}| \equiv \mathcal{A}(\underline{\varphi},\underline{\sigma}) = |\underline{\varphi}|\,|\underline{\sigma}|\sin\theta.$$

我们应该立即产生怀疑：向量的长度等于面积？！ 现实情况是我们正在处理第二个花招：向量积并不是它看起来的那样——它实际上是对应的 1-形式的楔积！ 对此的完整解释在于前面提到的霍奇对偶。 见练习 15。

$$\boxed{\underline{\varphi}\times\underline{\sigma} \quad \rightleftarrows \quad \varphi\wedge\sigma.} \tag{34.16}$$

在证明这种对应之前，让我们停下来看看它有多么特别：在任意维数上，我们可以自由地将两个向量与两个 1-形式相关联，然后我们可以形成它们的楔积，但是得到的 2-形式将有 $\frac{1}{2}n(n-1)$ 个分量，因此通常不能将其重新解释为向量。 只有在三个维度上才能将一个 2-形式识别为矢量流 $\underline{\mathbf{\Psi}}$ 的通量 2-形式 $\mathbf{\Psi}$。

为了证明（34.16），让我们定义 2-形式 $\mathbf{\Psi}\equiv\varphi\wedge\sigma$。 然后我们必须证明 $\underline{\mathbf{\Psi}} = \underline{\varphi}\times\underline{\sigma}$。我们将分两个阶段证明这一点。首先，我们将证明向量 $\underline{\mathbf{\Psi}}$ 和 $\mathbf{n}\equiv\underline{\varphi}\times\underline{\sigma}$ 指向相同的方向，然后我们将证明它们具有相同的长度。

由于根据定义，$\mathbf{n}\equiv\underline{\varphi}\times\underline{\sigma}$ 与 $\underline{\varphi}$ 和 $\underline{\sigma}$ 正交，

$$\varphi(\mathbf{n}) = \underline{\varphi}\cdot\mathbf{n} = 0 \quad \text{和} \quad \sigma(\mathbf{n}) = \underline{\sigma}\cdot\mathbf{n} = 0.$$

这意味着对于任何包含 $\mathbf{n}$ 的平行四边形，通量都为零：

$$\mathbf{\Psi}(\mathbf{n},\ldots) = (\varphi \wedge \sigma)(\mathbf{n},\ldots) = \varphi(\mathbf{n})\sigma(\ldots) - \sigma(\mathbf{n})\varphi(\ldots) = 0.$$

因此，正如（34.13）中所讨论的，这意味着 $\mathbf{n}$ 与流对齐：$\underline{\mathbf{\Psi}} \propto \mathbf{n}$。

仍有待证明两个流的速率匹配：$|\underline{\mathbf{\Psi}}| = |\mathbf{n}| = \mathcal{A}(\underline{\varphi},\underline{\sigma}) = |\underline{\varphi}|\,|\underline{\sigma}|\sin\theta$。我们首先注意到，由于我们现在知道 $\underline{\mathbf{\Psi}}$ 与具有边 $\underline{\varphi}$ 和 $\underline{\sigma}$ 的平行四边形 P 正交，因此它在 P 上的通量由下式给出

$$\text{通量} = |\underline{\mathbf{\Psi}}|\,\mathcal{A}(\underline{\varphi},\underline{\sigma})$$

但，

$$\begin{aligned}
\text{通量} &= \Psi(\underline{\varphi},\underline{\sigma}) \\
&= (\varphi \wedge \sigma)(\underline{\varphi},\underline{\sigma}) \\
&= \varphi(\underline{\varphi})\sigma(\underline{\sigma}) - \sigma(\underline{\varphi})\varphi(\underline{\sigma}) \\
&= |\underline{\varphi}|^2\,|\underline{\sigma}|^2 - |\underline{\varphi}|^2\,|\underline{\sigma}|^2\cos^2\theta \\
&= [\mathcal{A}(\underline{\varphi},\underline{\sigma})]^2.
\end{aligned}$$

结合前两个结果，

$$|\underline{\mathbf{\Psi}}|\mathcal{A}(\underline{\varphi},\underline{\sigma}) = \text{通量} = [\mathcal{A}(\underline{\varphi},\underline{\sigma})]^2.$$

所以，

$$|\underline{\mathbf{\Psi}}| = \mathcal{A}(\underline{\varphi},\underline{\sigma}) = |\underline{\varphi} \times \underline{\sigma}| \implies \underline{\mathbf{\Psi}} = \underline{\varphi} \times \underline{\sigma},$$

证明（34.16）。

在几何上建立（34.16）后，我们现在可以使用它从楔积的基本公式推导出向量积的标准代数表达式：

$$\begin{aligned}
\mathbf{\Psi} &= \varphi \wedge \sigma \\
&= (\varphi_1\,\mathbf{d}x^1 + \varphi_2\,\mathbf{d}x^2 + \varphi_3\,\mathbf{d}x^3) \wedge (\sigma_1\,\mathbf{d}x^1 + \sigma_2\,\mathbf{d}x^2 + \sigma_3\,\mathbf{d}x^3)
\end{aligned}$$

$$= (\varphi_2\sigma_3 - \varphi_3\sigma_2)(\mathbf{d}x^2 \wedge \mathbf{d}x^3) + (\varphi_3\sigma_1 - \varphi_1\sigma_3)(\mathbf{d}x^3 \wedge \mathbf{d}x^1) + (\varphi_1\sigma_2 - \varphi_2\sigma_1)(\mathbf{d}x^1 \wedge \mathbf{d}x^2).$$

因此，相关向量 $\underline{\mathbf{\Psi}} = \underline{\varphi} \times \underline{\sigma}$ 由向量分析（Vector Calculus）中熟悉的公式给出：

$$\begin{bmatrix}\varphi_1\\ \varphi_2\\ \varphi_3\end{bmatrix} \times \begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\end{bmatrix} = \begin{bmatrix}\varphi_2\sigma_3 - \varphi_3\sigma_2\\ \varphi_3\sigma_1 - \varphi_1\sigma_3\\ \varphi_1\sigma_2 - \varphi_2\sigma_1\end{bmatrix}$$

最后，参考[34.5]和（34.12），我们还推导出边为$\mathbf{u}$、$\mathbf{v}$、$\underline{\mathbf{\Omega}}$ 的平行六面体体积的标准公式：

$$\begin{aligned} \text{体积} = \text{通量} = \mathbf{\Omega}(\mathbf{u},\mathbf{v}) &= \begin{bmatrix}\Omega^1\\ \Omega^2\\ \Omega^3\end{bmatrix} \cdot \begin{bmatrix}u^2v^3 - u^3v^2\\ u^3v^1 - u^1v^3\\ u^1v^2 - u^2v^1\end{bmatrix} \\ &= \det\begin{bmatrix}\Omega^1 & u^1 & v^1\\ \Omega^2 & u^2 & v^2\\ \Omega^3 & u^3 & v^3\end{bmatrix} \end{aligned} \tag{34.17}$$

34.8 法拉第和麦克斯韦电磁 2-形式

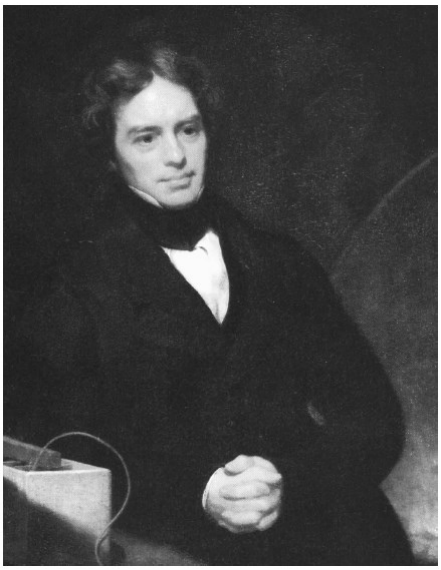
介绍性说明：本节完全是可选的； 不需要理解后面的数学（尽管这是第 36.6 节的先决条件）。 也就是说，这一部分非常重要，因为这是我们第一次证明形式不仅是柏拉图式的数学宇宙的居民——它们还伸出手，编织到物理宇宙法则的结构中！ 还有一件事：为了便于读者从我们的处理过渡到更高级的作品——例如 Misner、Thorne 和 Wheeler (1973)、Schutz (1980)、Baez 和 Muniain (1994)，以及 Frankel (2012）——我们（不情愿地）采用了他们的约定，其中时空的度量系数是我们迄今为止所采用的那些的负值。 所以，从今以后，闵可夫斯基度量是

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2.$$

在 4 维时空中有几种 2-形式在物理定律中发挥着基本作用。 在本节中，我们将描述其中的两个，它们共同引出对电磁定律的非常紧凑和优雅的描述（详见第 36.6 节）。 这些 2-形式以迈克尔 •法拉第（Michael Faraday，1791-1867）和詹姆斯 •克拉克 •麦克斯韦（James Clerk Maxwell，1831-1879）（分别在 [34.7] 和 [34.8] 中画出）命名。

自学成才的迈克尔 • 法拉第（Michael Faraday）在他位于伦敦皇家学院地下室的实验室[5]中进行的一系列绝顶巧妙的实验中，首先凭经验发现了电磁定律。 这些实验始于 1820 年代，并在 1830 年代达到高潮。参见福布斯和马洪（Forbes and Mahon，2014， § § 4 – 5）。

[5] 我们强烈建议你参观法拉第实验室的遗迹，在那里你可以跪下（恭敬地！）研究人类建造的第一台发电机——精心保存的，距离你只有几英寸的玻璃后面！



[34.7] 迈克尔·法拉第 (1791-1867)。

[34.8] 詹姆斯·克拉克·麦克斯韦 (1831-1879)。

麦克斯韦很快就认识到法拉第的力线和通量线思想的解释性价值，而且与法拉第不同，麦克斯韦拥有数学技能，可以给它们精确的形式。 事实上，自从牛顿以来，世界上还没有见过如此深刻的物理洞察力和原始数学力量的炼金术融合。

到 1873 年，这项工作完成了：麦克斯韦宣布了四个[6]方程，它们完全描述了电、磁以及两者之间的相互作用。 特别是，麦克斯韦用他的方程计算出了纯理论的电磁波穿过真空的速度。 他的数字答案与实验测量的光速相吻合——麦克斯韦是第一个了解光是什么的人——一种电磁波！

阅读这两位科学巨人之间的信件[7]令人动容。 麦克斯韦对法拉第物理洞察力的深刻性和

[6]实际上是 Oliver Heaviside 在 1885 年将麦克斯韦最初的 20 个方程组简化为我们现在所知的四个“麦克斯韦方程”！ 参见 Forbes and Mahon (2014, §16) 和 Mahon (2017)。

[7]参见 Forbes and Mahon (2014) 和 Jones (1870)。

他的实验独创性深表钦佩，同时法拉第对麦克斯韦在将物理直觉结晶为精确的数学定律方面的非凡能力感到敬畏。

爱因斯坦 1905 年对狭义相对论的发现与法拉第和麦克斯韦的这些电磁学发现直接相关。 事实上，爱因斯坦那篇划时代的论文的标题没有提及空间或时间，而是写着《论动体的电动力学》[8]。

爱因斯坦取得突破的动力是他意识到电场和磁场并不是独立存在的——它们只是单个电磁场的两个依赖于观察者的方面，并且这个场的波（光!）对于所有观察者以相同的速度传播。 这使他（通过闵可夫斯基）认识到空间和时间不能单独存在绝对存在——它们是时空几何的依赖观察者的方面，它确实具有绝对存在。

让我们简单地引用爱因斯坦 1905 年论文的开头一段：

众所周知，麦克斯韦的电动力学——正如目前通常理解的那样——当应用于运动物体时，会导致不对称现象，这似乎不是现象中固有的。以磁铁和导体的相互电动作用为例。此处可观察到的现象仅取决于导体和磁体的相对运动，而通常的观点在这些物体中的一个或另一个运动的两种情况之间做出了鲜明的区分。因为如果磁体在运动而导体处于静止状态，则在磁体附近会产生具有一定能量的电场，在导体部分所在的位置产生电流。但是，如果磁铁是静止的而导体在运动，则磁铁附近不会产生电场。然而，在导体中，我们发现了一种电动势，它本身没有相应的能量，但它会产生——假设在所讨论的两种情况下相对运动相等——产生与产生的电流具有相同路径和强度的电流由前一种情况下的电力。

这些场的经典描述是两个 3 维向量场，共有六个分量，每个分量都是空间和时间的函数：

$$\text{电场} = \underline{\mathbf{E}} = \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} \quad \text{和 磁场} = \underline{\mathbf{B}} = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix}。$$

这些中的每一个都可能与通量 2-形式（通过（34.10））和 1-形式通过（34.14）相关联。对于电场，我们将通量称为 2-形式 **E** 和 1-形式 ε：

$$\boxed{\mathbf{E} = E_x\,(\mathbf{d}y \wedge \mathbf{d}z) + E_y\,(\mathbf{d}z \wedge \mathbf{d}x) + E_z\,(\mathbf{d}x \wedge \mathbf{d}y),} \tag{34.18}$$

和

$$\boxed{\varepsilon = E_x\,\mathbf{d}x + E_y\,\mathbf{d}y + E_z\,\mathbf{d}z.} \tag{34.19}$$

[8]整篇论文（翻译成英文）可在网络上免费获取：只需搜索其标题即可。

我们对磁场做同样的事情，将通量 2-形式写为 **B**，将 1-形式写为 β：

$$\boxed{B = B_x\,(\mathbf{d}y \wedge \mathbf{d}z) + B_y\,(\mathbf{d}z \wedge \mathbf{d}x) + B_z\,(\mathbf{d}x \wedge \mathbf{d}y)} \tag{34.20}$$

和

$$\boxed{\beta\ = B_x\,\mathbf{d}x + B_y\,\mathbf{d}y + B_z\,\mathbf{d}z.} \tag{34.21}$$

正如爱因斯坦所解释的（见上文），这些 2-形式和 1-形式都没有绝对的、独立于观察者的存在。 也就是说，相对运动的观察者会不同意电场和磁场，但他们也会不同意空间和时间，以使物理现象独立于观察者的运动：只能有一个物理现实！

非常显著——非常漂亮！——大自然告诉我们，*电和磁 2–形式和 1–形式结合成一个电磁 2-形式，在时空中确实具有绝对意义！*

它被称为[9]*法拉第*（Faraday）2-*形式*，用 **F** 表示：

$$\boxed{\mathbf{F} = \text{Faraday} = \varepsilon \wedge \mathbf{d}t + \mathbf{B}.} \tag{34.22}$$

让我们立即描述 **F** 的物理（独立于观察者的）意义。

如果我们将向量 **u** 插入到 **F** 的第二个槽中，则我们得到 $\mathbf{F}(\ldots,\mathbf{u})$。 当向量插入第一个槽时，这将产生一个数字； 因此，就目前而言，它必须是 1-形式。

假设我们在 **F** 描述的电磁场中放置一个电荷 q。设粒子在时空中的 4-速度矢量为 **u**，让它的 4-动量用 1-形式 π 来描述，并让它的固有时间（又名“手表时间”）为 τ。 *那么* $q\,\mathbf{F}(\ldots,\mathbf{u})$ *描述了施加在粒子上的电磁力*

$$\boxed{\frac{d\pi}{d\tau} = q\,\mathbf{F}(\ldots,\mathbf{u}).} \tag{34.23}$$

这是一个 1-形式，等待插入向量，并且（物理上）得到的数字是作用在插入向量方向上的力的大小。

---

[9]这是 Misner、Thorne 和 Wheeler (1973) 使用的术语，但并不通用：例如，Penrose (2005) 将 F 称为“麦克斯韦 2-形式”。 正如我们稍后将解释的，我们将名称麦克斯韦 2-形式保留为不同的 2-形式。

将此与更复杂的公式（洛伦兹力定律）进行对比，后者在介绍性电动力学中学习了具有空间矢量速度 $\mathbf{v}$ 的粒子的空间矢量动量 $\mathbf{p}$ 的变化率：

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{dt}} = \mathrm{q}\,(\underline{\mathbf{E}} + \mathbf{v} \times \underline{\mathbf{B}}).$$

在这个经典公式中，相对运动的不同观察者都会对 $\mathbf{p}$、$\underline{\mathbf{E}}$、$\mathbf{v}$ 和 $\underline{\mathbf{B}}$ 的值有不同的看法！（有关这一点的更全面的讨论，请参见 Misner、Thorne 和 Wheeler (1973, §3.3)。）将此与几何的、独立于观察者的时空定律（34.23）的优雅进行对比！

在我们介绍麦克斯韦 2-形式之前，让我们先陈述一下麦克斯韦电磁方程，我们将把它分成两对方程。

第一对描述了无源时的电磁场（意味着（标量）电荷密度 ρ 和（矢量）电流密度 j）：

$$\boxed{\text{麦克斯韦无源方程：}\qquad \begin{aligned} \nabla \cdot \underline{\mathbf{B}} &= 0, \\ \nabla \times \underline{\mathbf{E}} + \partial_{\mathrm{t}} \underline{\mathbf{B}} &= 0. \end{aligned}} \tag{34.24}$$

正如我们将在第 36.6 节中看到的，当以法拉第 2 形式表示时，这两个方程简化为一个单一的、优雅的方程。

第二对麦克斯韦方程描述有源产生的电磁场：

$$\boxed{\text{麦克斯韦源方程：}\qquad \begin{aligned} \nabla \cdot \underline{\mathbf{E}} &= 4\pi\rho, \\ \nabla \times \underline{\mathbf{B}} - \partial_{\mathrm{t}} \underline{\mathbf{E}} &= 4\pi\, \underline{\mathbf{j}}. \end{aligned}} \tag{34.25}$$

这第二对也简化为一个单一的、优雅的方程——将在第 36.6 节中推导——但这次是与 $\mathbf{F}$ 密切相关的不同的 2-形式，并表示为 $\star\mathbf{F}$。 这是麦克斯韦（Maxwell）2-形式，它几乎（但不完全）是法拉第 2-形式中电场和磁场角色互换的结果：

$$\boxed{\star\mathbf{F} = \mathrm{Maxwell} = \beta \wedge \mathbf{d}\mathrm{t} - \mathbf{E}.} \tag{34.26}$$

星算子（$\star$）是前面提到的霍奇对偶算子； 见习题 15。这是一个纯粹的数学运算，（在 n 维中）将 p-形式映射到 (n-p)-形式。 由于我们在 4 维中，$\star$ 将 2-形式 $\mathbf{F}$ 映射到另一个 2-形式 $\star\mathbf{F}$，即 $\mathbf{F}$ 的（霍奇）对偶。如果我们再次应用星算子，换句话说，采用麦克斯韦

（Maxwell）的对偶，则我们恢复法拉第（只差一个减号）：

$$\star\star\,\mathbf{F} = -\mathbf{F}. \tag{34.27}$$

因此，我们可以说麦克斯韦和法拉第是彼此的对偶——我们猜测两位科学家都会喜欢这种提法（要是他们能活着看到它就好了）。

第三十五章 3-形式

35.1 3-形式需要 3 个维度

让我们首先认识到 3-形式（3-forms）至少需要三个维度才能存在。

在 $\mathbb{R}^2$ 中，很容易创建 $\begin{Bmatrix}0\\3\end{Bmatrix}$ 价的张量 $\mathbf{H}$。 例如，取任何 1-形式的 φ，并构造，

$$\mathbf{H} = \varphi \otimes \varphi \otimes \varphi \implies \mathbf{H}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \varphi(\mathbf{v}_1)\varphi(\mathbf{v}_2)\varphi(\mathbf{v}_3).$$

正如我们所看到的，这个张量是完全对称的，这意味着交换任何一对输入向量都会返回相同的输出。

但是要创建一个 3-形式 Ξ，我们要求（根据定义）我们的张量是完全反对称的，这是不可能的！ 考虑 Ξ 的分量：

$$\Xi_{\mathrm{ijk}} = \Xi(\mathbf{e}_{\mathrm{i}}, \mathbf{e}_{\mathrm{j}}, \mathbf{e}_{\mathrm{k}}).$$

在 $\mathbb{R}^2$ 中只有两个基向量 $\{\mathbf{e}_1, \mathbf{e}_2\}$，这意味着 Ξ 的两个槽包含相同的向量，但反对称要求输出消失。

为了使 Ξ_{ijk} 不会完全消失，因此所有三个索引都必须是不同的：换句话说，我们必须至少在三个维度中。显然，完全相同的推理非常普遍地表明，p-形式的存在至少需要 p 维。

我们也可以反过来观察，可以存在于 n 维中的最高阶形式是 n−形式。

35.2 2-形式和 1-形式的楔积

两个 1-形式的楔积产生 2-形式，我们现在寻求扩展楔积的定义，以确保 2-形式 Ψ 和 1-形式 σ 的乘积自动产生 3-形式，$\Psi \wedge \sigma$。

让我们从最简单的（不幸的）猜测开始：怎么样

$$(\Psi \wedge \sigma)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)??? = \Psi(\mathbf{v}_1, \mathbf{v}_2)\, \sigma(\mathbf{v}_3)\,?$$

嗯，这在 $\mathbf{v}_1 \leftrightarrow \mathbf{v}_2$ 下显然是反对称的，但在 $\mathbf{v}_2 \leftrightarrow \mathbf{v}_3$ 下不是，所以让我们使用我们在 (33.8) 中使用的相同技巧来施加反对称。 也就是说，让我们用 $\mathbf{v}_2 \leftrightarrow \mathbf{v}_3$ 减去相同的表达式：

$$(\Psi \wedge \sigma)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)??? = \Psi(\mathbf{v}_1, \mathbf{v}_2)\sigma(\mathbf{v}_3) - \Psi(\mathbf{v}_1, \mathbf{v}_3)\sigma(\mathbf{v}_2)\,?$$

但是现在第二项破坏了 $\mathbf{v}_1 \leftrightarrow \mathbf{v}_2$ 下的反对称，所以重复这个技巧并减去相同的项，但使用 $\mathbf{v}_1 \leftrightarrow \mathbf{v}_2$：

$$(\Psi \wedge \sigma)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)??? = \Psi(\mathbf{v}_1, \mathbf{v}_2)\sigma(\mathbf{v}_3) - [\Psi(\mathbf{v}_1, \mathbf{v}_3)\sigma(\mathbf{v}_2) - \Psi(\mathbf{v}_2, \mathbf{v}_3)\sigma(\mathbf{v}_1)]?$$

成功！

$$\boxed{(\Psi \wedge \sigma)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \Psi(\mathbf{v}_1, \mathbf{v}_2)\sigma(\mathbf{v}_3) + \Psi(\mathbf{v}_3, \mathbf{v}_1)\sigma(\mathbf{v}_2) + \Psi(\mathbf{v}_2, \mathbf{v}_3)\sigma(\mathbf{v}_1).} \tag{35.1}$$

注意向量的循环排列，这使得公式很容易记住。

我们可以玩同样的游戏来定义一个看似不同的 3-形式 $(\sigma \wedge \Psi)$，但结果却是相同的 3-形式 $(\Psi \wedge \sigma)$！ 要了解原因，只需将每个 σ (vi) 移动到 (35.1) 右侧的每个项的前面，并观察这保留了循环置换：

$$(\sigma \wedge \Psi)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \sigma(\mathbf{v}_1)\Psi(\mathbf{v}_2, \mathbf{v}_3) + \sigma(\mathbf{v}_2)\Psi(\mathbf{v}_3, \mathrm{v}1) + \sigma(\mathbf{v}_3)\Psi(\mathbf{v}_1, \mathbf{v}_2). \tag{35.2}$$

因此，虽然楔积在结合两个 1-形式时是反对称的，即

$$(\varphi \wedge \psi) = -(\psi \wedge \varphi),$$

当它结合一个 2-形式 Ψ 和一个 1-形式 σ 时，它是对称的：

$$\boxed{\Psi \wedge \sigma = \sigma \wedge \Psi.}$$

这是一般规则，我们将很快证明这一点：

$$\boxed{\text{如果 } \boldsymbol{\Psi} \text{ 是 p-形式且 } \boldsymbol{\Omega} \text{ 是 q-形式，则 } \boldsymbol{\Psi} \wedge \boldsymbol{\Omega} = (-1)^{\mathrm{pq}}\boldsymbol{\Omega} \wedge \boldsymbol{\Psi}\text{。}} \tag{35.3}$$

您可以轻松检查前两个方程是否符合此规则。

假设 (35.3) 为真，我们还推导出如果 $\boldsymbol{\Omega}$ 的度数是奇数，那么 [练习]，

$$\boldsymbol{\Omega} \wedge \boldsymbol{\Omega} = 0. \tag{35.4}$$

当然，1-形式与自身的楔积是我们这样观察的第一个例子：$\psi \wedge \psi = 0$。

35.3 体积 3-形式

现在让我们取 $\mathbf{\Psi} = \mathbf{d}x^1 \wedge \mathbf{d}x^2$ 和 $\sigma = \mathbf{d}x^3$，并将体积 3-形式定义为

$$\boxed{\mathcal{V} \equiv (\mathbf{d}x^1 \wedge \mathbf{d}x^2) \wedge \mathbf{d}x^3.} \tag{35.5}$$

我们现在证明 $\mathcal{V}$ 的命名是合理的。

再次查看具有边 $\mathbf{u}$、$\mathbf{v}$ 和 $\underline{\mathbf{\Omega}}$ 的平行六面体体积的标准公式（34.17）。 如果我们将 3-形式 $\mathcal{V}$ 应用于这三个向量，则 (35.1) 产生

$$\begin{aligned}
\mathcal{V}(\mathbf{u},\mathbf{v},\underline{\mathbf{\Omega}}) &= [(\mathbf{d}x^1 \wedge \mathbf{d}x^2) \wedge \mathbf{d}x^3](\mathbf{u},\mathbf{v},\underline{\mathbf{\Omega}}) \\
&= (\mathbf{d}x^1 \wedge \mathbf{d}x^2)(\mathbf{u},\mathbf{v})\, \mathbf{d}x^3(\underline{\mathbf{\Omega}}) + (\mathbf{d}x^1 \wedge \mathbf{d}x^2)(\mathbf{v},\underline{\mathbf{\Omega}})\, \mathbf{d}x^3(\mathbf{u}) + (\mathbf{d}x^1 \\
&\qquad \wedge \mathbf{d}x^2)(\underline{\mathbf{\Omega}},\mathbf{u})\, \mathbf{d}x^3(\mathbf{v}) \\
&= (u^1 v^2 - u^2 v^1)\Omega^3 + (v^1 \Omega^2 - v^2 \Omega^1)\, u^3 + (\Omega^1 u^2 - \Omega^2 u^1)\, v^3 \\
&= \Omega^1(u^2 v^3 - u^3 v^2) + \Omega^2(u^3 v^1 - u^1 v^3) + \Omega^3(u^1 v^2 - u^2 v^1) \\
&= \mathbf{\Omega}(\mathbf{u},\mathbf{v}) \\
&= \text{具有边 u、v 和 } \Omega \text{ 的平行六面体的体积.}
\end{aligned}$$

我们将看到这三个 1-形式的楔积满足结合律，因此体积 3-形式可以表示为

$$\boxed{(\mathbf{d}x^1 \wedge \mathbf{d}x^2) \wedge \mathbf{d}x^3 = \mathbf{d}x^1 \wedge \mathbf{d}x^2 \wedge \mathbf{d}x^3 = \mathbf{d}x^1 \wedge (\mathbf{d}x^2 \wedge \mathbf{d}x^3).}$$

35.4 球极坐标中的体积 3-形式

（35.5）再次提醒我们一个经典表达式：在空间中进行三重积分时，我们将体积元写为 $d\mathcal{V} = dx\, dy\, dz$。 这种联系很容易明确，就像以前一样：将 $\mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}z$ 应用于边为 δx、δy、δz 的小盒子的向量边，我们得到它的体积，$\delta\mathcal{V} = \delta x\, \delta y\, \delta z$。

如果我们切换到球极坐标，(r, ϕ, ϑ)，那么经典的体积元变为，

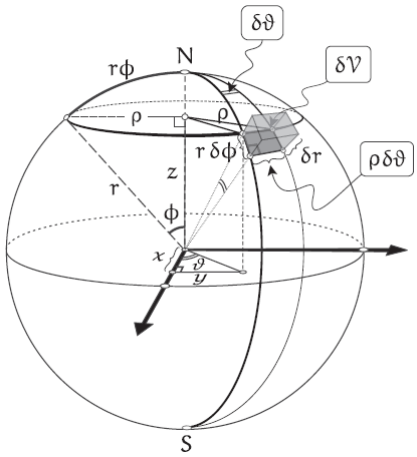
$$d\mathcal{V} = r^2 \sin\phi\, dr\, d\phi\, d\vartheta.$$

图[35.1]包含一个直接的几何证明，使我们能够理解这个结果。 注意，该图还提供了坐标变化的公式：由于球面上的一点到 NS，z 轴的距离为 $\rho = r\sin\phi$，

$$x = \rho\cos\vartheta = (r\sin\phi)\cos\vartheta,$$

$$y = \rho\sin\vartheta = (r\sin\phi)\sin\vartheta,$$

$$z = r\cos\phi.$$



[35.1] $\delta\mathcal{V} \asymp \delta r\,(r\delta\phi)(\rho\delta\vartheta) = r^2\sin\phi\,\delta r\,\delta\phi\,\delta\vartheta$的几何证明，以及 x、y 和 z 公式的推导。

最后，请注意已选择坐标[1]的顺序 (r, φ, ϑ)，以便坐标的小幅增加导致的运动是*右手的*。

在更复杂的坐标变化的情况下，我们传统上会被迫计算雅可比行列式。 好吧，不再！

让我们使用球极坐标的例子来演示（在更困难的情况下）体积 3-形式如何节省我们的时间、记忆，甚至是思考的需要！

好吧，稍微思考一下肯定会有所帮助：当乘以下面的楔积时，我们只是在寻找获得 $\mathbf{d}r \wedge \mathbf{d}\phi \wedge \mathbf{d}\vartheta$ 的方法，因为所有其他项都消失了。 在计算接近尾声时，我们写成 $c_\vartheta \equiv \cos\vartheta$ 等等，以节省空间，但我们也建议在私人的纸笔计算中将其作为澄清符号。

$$\mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}z = \mathbf{d}(r\sin\phi\cos\vartheta) \wedge \mathbf{d}(r\sin\phi\sin\vartheta) \wedge \mathbf{d}(r\cos\phi)$$

$$= [(\mathbf{d}r)\sin\phi\cos\vartheta + r\cos\phi\mathbf{d}\phi\cos\vartheta - r\sin\phi\sin\vartheta\,\mathbf{d}\vartheta]$$

$$\wedge\ [(\mathbf{d}r)\sin\phi\sin\vartheta + r\cos\phi\mathbf{d}\phi\sin\vartheta + r\sin\phi\cos\vartheta\,\mathbf{d}\vartheta]$$

[1] 我们正在使用美国数学家的（符合逻辑！）惯例，其中 ϑ 表示 (x, y) 平面内的角度，就像它在 $\mathbb{R}^2$ 中所做的那样。 然而，英国的数学家，基本上所有的科学家，都互换了 ϑ 和 ϕ 的含义，因此它们的坐标顺序和它们的体积 3-形式似乎与我们的不同。 采用物理学家的选择实际上有令人信服的历史原因——球谐函数立即浮现在脑海——只要我们都同意将 $\mathbb{R}^2$ 中的角度更改为 ϕ，冲突就可以解决！ Dray (2015) 是这种解决方案的罕见而勇敢的采用者。

$$\wedge\ [(\mathbf{d}r)\cos\phi - r\sin\phi\mathbf{d}\phi]$$

$$= r^2\sin\phi[s_\phi^2\, c_\vartheta^2 + s_\phi^2\, s_\vartheta^2 + c_\phi^2\, s_\vartheta^2 + c_\phi^2\, c_\vartheta^2]\mathbf{d}r\wedge\mathbf{d}\phi\wedge\mathbf{d}\vartheta$$

$$= r^2\sin\phi\mathbf{d}r\wedge\mathbf{d}\phi\wedge\mathbf{d}\vartheta.$$

几何思维让我们理解了这个结果，而“魔鬼机器”只是给了我们答案，不需要任何思考。注意打开前言的警告：“…我们灵魂的危险就在那里，因为当你进入代数计算时，你基本上停止了思考：你停止了几何思考，你停止了思考意义。”

这些只是“魔鬼机器”力量的第一眼。 因此，你必须保持坚定，时刻保持警惕，以免你屈服于它的诱惑，从而失去你的灵魂！

35.5 三个 1-形式和 p 个 1-形式的楔积

为了证实我们刚刚声称的结合性，让我们概括讨论，并定义三个任意 1-形式的楔积。我们希望系统地标记这三个 1-形式，但我们也希望避免与分量发生任何可能的混淆，所以我们称它们为 $\overset{\sigma}{1}, \overset{\sigma}{2}, \overset{\sigma}{3}$。

$\overset{\sigma}{1}\wedge\overset{\sigma}{2}\wedge\overset{\sigma}{3}$ 的公式可以通过在 (35.1) 中取 $\Psi = \overset{\sigma}{1}\wedge\overset{\sigma}{2}$ 来找到，但是更简单地推广几何映射（34.5）和 [34.2] 中给出的结果几何解释。

也就是说，让我们使用三个 1-形式来定义 $\mathbb{R}^n$ 中的向量 $\mathbf{v}$ 到 $\mathbb{R}^3$ 中的向量 $F(\mathbf{v})$ 的映射 F：

$$\boxed{v \longrightarrow F(v) \equiv \begin{bmatrix} \overset{\sigma}{1}(\mathbf{v}) \\ \overset{\sigma}{2}(\mathbf{v}) \\ \overset{\sigma}{3}(\mathbf{v}) \end{bmatrix}} \tag{35.6}$$

因此，$\mathbb{R}^n$ 中具有边 $\mathbf{v}_1$、$\mathbf{v}_2$ 和 $\mathbf{v}_3$ 的平行六面体被 F 映射到 $\mathbb{R}^3$ 中具有边 $F(\mathbf{v}_1)$、$F(\mathbf{v}_2)$ 和 $F(\mathbf{v}_3)$ 的平行六面体，并且与 [34.2] 类似， 我们可以将 $\overset{\sigma}{1}\wedge\overset{\sigma}{2}\wedge\overset{\sigma}{3}$ 定义为产生这个平行六面体像体积的算子：

$$\boxed{\left[\overset{\sigma}{1}\wedge\overset{\sigma}{2}\wedge\overset{\sigma}{3}\right](\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3) = \det\begin{bmatrix} | & | & | \\ F(\mathbf{v}_1) & F(\mathbf{v}_2) & F(\mathbf{v}_3) \\ | & | & | \end{bmatrix}} \tag{35.7}$$

显然，这可以立即推广到定义 p 个 1-形式的楔积：

$$\left[\overset{\sigma}{1}\wedge\cdots\wedge\overset{\sigma}{\mathrm{p}}\right](\mathbf{v}_1,\cdots,\mathbf{v}_\mathrm{p})=\det\begin{bmatrix}| & \cdots & | \\ \mathrm{F}(\mathbf{v}_1) & \cdots & \mathrm{F}(\mathbf{v}_\mathrm{p}) \\ | & \cdots & |\end{bmatrix}.$$

等式（35.7）描述了 3-形式在三个向量上的作用，但是我们可以抽象出向量来揭示 3-形式本身吗？ 是的，但要这样做，我们需要直接处理张量积，就像我们在两个 1-形式的楔积的原始定义（34.3）中所做的那样。

为了实现我们的目标，我们必须尝试写出在任意一对 1-形式交换下反对称的表达式。可以通过从（35.7）向后工作来做到这一点 [练习]，但实际上正面解决问题更简单。 在研究答案之前自己尝试一下，这里是：

$$\begin{aligned}&\overset{\sigma}{1}\wedge\overset{\sigma}{2}\wedge\overset{\sigma}{3}=\overset{\sigma}{1}\otimes\overset{\sigma}{2}\otimes\overset{\sigma}{3}+\overset{\sigma}{2}\otimes\overset{\sigma}{3}\otimes\overset{\sigma}{1}+\overset{\sigma}{3}\otimes\overset{\sigma}{1}\otimes\overset{\sigma}{2}-\\&\overset{\sigma}{2}\otimes\overset{\sigma}{1}\otimes\overset{\sigma}{3}-\overset{\sigma}{3}\otimes\overset{\sigma}{2}\otimes\overset{\sigma}{1}-\overset{\sigma}{1}\otimes\overset{\sigma}{3}\otimes\overset{\sigma}{2}\end{aligned}\tag{35.8}$$

请注意，第一行的正项是 123 的循环排列，而第二行的负项是非循环排列。 第二行中的每一项都是通过在其正上方的循环项中交换前两个 1-形式而获得的。

我们让您检查（35.8）是否确实等价于（35.7）。

35.6 基 3-形式

正如 $\{\mathbf{d}\mathrm{x}^\mathrm{i}\wedge\mathbf{d}\mathrm{x}^\mathrm{j}\}$ 形成了 2-形式的基，所以 $\{\mathbf{d}\mathrm{x}^\mathrm{i}\wedge\mathbf{d}\mathrm{x}^\mathrm{j}\wedge\mathbf{d}\mathrm{x}^\mathrm{k}\}$ 形成了 3-形式的基。在 $\mathbb{R}^3$ 中只有一个这样的基 3-形式，即体积 3-形式，$\mathcal{V}=\mathbf{d}\mathrm{x}^1\wedge\mathbf{d}\mathrm{x}^2\wedge\mathbf{d}\mathrm{x}^3$； $\mathbb{R}^3$ 中的所有 3-形式都是 $\mathcal{V}$ 的倍数。

然而，在 $\mathbb{R}^4$ 中，有四种不同的基 3-形式，一般的 3-形式 $\mathbf{\Omega}$ 可以分解为如下分量：

$$\begin{aligned}\mathbf{\Omega}=\Omega_{123}\,(\mathbf{d}\mathrm{x}^1\wedge\mathbf{d}\mathrm{x}^2\wedge\mathbf{d}\mathrm{x}^3)+\Omega_{234}\,(\mathbf{d}\mathrm{x}^2\wedge\mathbf{d}\mathrm{x}^3\wedge\mathbf{d}\mathrm{x}4)+\Omega_{341}\,(\mathbf{d}\mathrm{x}^3\wedge\mathbf{d}\mathrm{x}^4\wedge\mathbf{d}\mathrm{x}^1)\\+\Omega_{412}\,(\mathbf{d}\mathrm{x}^4\wedge\mathbf{d}\mathrm{x}^1\wedge\mathbf{d}\mathrm{x}^2).\end{aligned}$$

现在让我们使用这些分量来解释规则（35.3）。 为了使论证生动，让我们取 $\mathrm{p}=2$ 和 $\mathrm{q}=3$，因此 $\mathbf{\Psi}$ 是 2-形式，$\mathbf{\Omega}$ 是 3-形式。 当我们形成 $\mathbf{\Psi}\wedge\mathbf{\Omega}$ 时，$\mathbf{\Psi}$ 的典型分量 $(\mathbf{d}\mathrm{x}^{\mathrm{j}_1}\wedge$

$\mathbf{d}x^{j_2}$) 乘以 $\mathbf{\Omega}$ 的典型分量 $(\mathbf{d}x^{k_1} \wedge \mathbf{d}x^{k_2} \wedge \mathbf{d}x^{k_3})$。（我们假设我们至少在五个维度中，否则这个积同样为零！）

现在我们希望将其转换为逆积 $\mathbf{\Omega} \wedge \mathbf{\Psi}$ 的项。为此，我们必须首先将 $\mathbf{d}x^{j_2}$ 通过三个 1-形式向右移动，每次连续交换都会产生符号变化：

$$
\begin{aligned}
(\mathbf{d}x^{j_1} \wedge \mathbf{d}x^{j_2}) \wedge (\mathbf{d}x^{k_1} \wedge \mathbf{d}x^{k_2} \wedge \mathbf{d}x^{k_3}) &= -\mathbf{d}x^{j_1} \wedge \mathbf{d}x^{k_1} \wedge \mathbf{d}x^{j_2} \wedge \mathbf{d}x^{k_2} \wedge \mathbf{d}x^{k_3} \\
&= +\mathbf{d}x^{j_1} \wedge \mathbf{d}x^{k_1} \wedge \mathbf{d}x^{k_2} \wedge \mathbf{d}x^{j_2} \wedge \mathbf{d}x^{k_3} \\
&= -\mathbf{d}x^{j_1} \wedge (\mathbf{d}x^{k_1} \wedge \mathbf{d}x^{k_2} \wedge \mathbf{d}x^{k_3}) \wedge \mathbf{d}x^{j_2}.
\end{aligned}
$$

在 $q = 3$ 的情况下，这导致了 $(-1)^3 = -1$ 的因子，并且在一般情况下，因子将是 $(-1)^q$。为了完成 $\mathbf{\Psi} \wedge \mathbf{\Omega}$ 的反转以创建 $\mathbf{\Omega} \wedge \mathbf{\Psi}$，我们现在必须通过 $\mathbf{\Omega}$ 的 1-形式移动 $\mathbf{d}x^{j_1}$ 以重新形成右边的 $(\mathbf{d}x^{j_1} \wedge \mathbf{d}x^{j_2})$。这样做，我们选择了 $(-1)^q$ 的第二个因子。在一般情况下，我们必须对 $\mathbf{\Psi}$ 的 p 个 1-形式中的每一个重复此操作，选取 $(-1)^q$ 的 p 个因子。因此，净因子为 $(-1)^{pq}$，完成了（35.3）的证明。

35.7 $\mathbf{\Psi} \wedge \mathbf{\Psi} \neq \mathbf{0}$ 可能么？

我们现在已经证明了规则（35.3）的一个结果是（35.4）：如果 $\mathbf{\Psi}$ 是奇数次，则 $\mathbf{\Psi} \wedge \mathbf{\Psi} = 0$。

因此，如果 $\mathbf{\Psi} \wedge \mathbf{\Psi} \neq 0$ 是可能的，那么 $\mathbf{\Psi}$ 必须是偶数。因此，最简单的可能性是 Ψ 是 2-形式，在这种情况下 $\mathbf{\Psi} \wedge \mathbf{\Psi}$ 将是 4-形式。4-形式可以存在的最小维数是 4，因此让我们将 $\mathbf{\Psi}$ 视为时空的 2-形式，坐标为 (t, x, y, z)。

考虑，

$$\mathbf{\Psi} = \mathbf{d}t \wedge \mathbf{d}x + \mathbf{d}y \wedge \mathbf{d}z.$$

如果 $\mathcal{V}$ 表示时空的体积 4-形式，那么 [练习]，

$$\mathbf{\Psi} \wedge \mathbf{\Psi} = 2\, \mathbf{d}t \wedge \mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}z = 2\mathcal{V}.$$

第三十六章 微分

36.1 1-形式的外导数

在形式（Forms）的上下文中，函数 f 被视为 0-形式。 当外导数 d 应用于 f 时，它会产生梯度 1-形式 $\mathbf{d}\mathrm{f}$，它编码 f 在所有可能方向上的变化速度。

如果我们想知道 f 在特定方向 $\mathbf{u}$ 上的变化率，我们必须将它作为一个问题提交给 神谕 1-形式 $\mathbf{d}\mathrm{f}$，然后它将答案声明为 （32.11）：

$$\mathbf{d}\mathrm{f}(\mathbf{u}) = \boldsymbol{\nabla}_{\mathbf{u}}\mathrm{f}.$$

总而言之，$\mathbf{d}$ 测量所有可能方向的变化率，它将形式的度数增加一，以允许输入一个额外的向量，沿着该向量来测量形式的变化率。

以这种方式表达，我们的任务很明确：我们必须扩展 d 的定义，以便当它应用于 p-形式的 Ψ 时，它会产生一个 (p+1)-形式的 $\mathrm{d}\Psi$，它在所有可能的方向测量 Ψ 的变化率。

按照我们的习惯，我们将从最简单的情况开始，然后逐步向上。 因此，让我们尝试微分一个 1-形式的 φ 以获得一个 2-形式的 $\mathrm{d}\varphi$，作用于向量对 $\mathbf{u}$ 和 $\mathbf{v}$。

为了让变化率变得有意义，我们必须想象 1-形式和它所作用的向量至少定义在我们寻求变化率的点的附近。 因此我们取 φ 为 1-形式场（field），同样取 $\mathbf{u}$ 和 $\mathbf{v}$ 为向量场（fields）。

让我们首先假设 $\mathbf{u}$ 是我们想知道 φ 变化率的方向，它在 $\mathbf{d}\mathrm{f}(\mathbf{u}) = \boldsymbol{\nabla}_{\mathbf{u}}\mathrm{f}$ 中所起的作用相同。 由于 $\mathbf{v}$ 是唯一起作用的其他向量，因此将寻求其变化率的函数只能是 $\varphi(\mathbf{v})$。 换句话说，我们基本上被迫考虑 $\varphi(\mathbf{v})$ 沿 $\mathbf{u}$ 的变化率：

$$\boldsymbol{\nabla}_{\mathbf{u}}\varphi(\mathbf{v}).$$

我们有了一个好的开始，但我们要求 $\mathbf{d}\varphi$ 是 2-形式，因此它对向量对的作用 $\mathbf{d}\varphi(\mathbf{u},\mathbf{v})$ 在这些向量中必须是反对称的，而这个表达式不是。 因此，我们使用了一个现在应该非常熟悉的技巧：我们通过减去相同的表达式，但交换向量 $\mathbf{u} \leftrightarrow \mathbf{v}$ 来强加反对称性：

$$\boldsymbol{\nabla}_{\mathbf{u}}\varphi(\mathbf{v}) - \boldsymbol{\nabla}_{\mathbf{v}}\varphi(\mathbf{u}). \tag{36.1}$$

这个表达式现在在两个向量输入中是反对称的，所以我们肯定越来越近了，但这仍然不是正确的答案。

问题是 $\mathbf{d}\varphi$ 应该测量 1-形式 φ 的变化率（在所有方向上）——向量场 $\mathbf{u}$ 和 $\mathbf{v}$ 应该与此无关。 然而，即使在我们详细研究这个问题之前，很明显 $\varphi(\mathbf{v})$ 的变化，例如，将取决

于 φ 的变化（这是我们关心的事情）和 $\mathbf{v}$ 的变化（我们不关心）。

因此，为了定义外导数 $\mathbf{d}\varphi$，我们必须发现 $\mathbf{u}$ 和 $\mathbf{v}$ 的变化对（36.1）的贡献，然后我们必须通过外科手术切除它们不需要的贡献！

为了实现这一点，让我们用对应的向量 $\underline{\varphi}$ 来表达 φ 对向量的作用，就像我们在 (34.15) 中所做的那样：

$$\varphi(\mathbf{u}) = \underline{\varphi} \cdot \mathbf{u}.$$

然后，

$$\begin{aligned}
\boldsymbol{\nabla}_{\mathbf{u}}\varphi(\mathbf{v}) - \boldsymbol{\nabla}_{\mathbf{v}}\varphi(\mathbf{u}) &= \boldsymbol{\nabla}_{\mathbf{u}}(\underline{\varphi} \cdot \mathbf{v}) - \boldsymbol{\nabla}_{\mathbf{v}}(\underline{\varphi} \cdot \mathbf{u}) \\
&= \{\mathbf{v} \cdot \boldsymbol{\nabla}_{\mathbf{u}}\underline{\varphi} - \mathbf{u} \cdot \boldsymbol{\nabla}_{\mathbf{v}}\underline{\varphi}\} + \{\underline{\varphi} \cdot \boldsymbol{\nabla}_{\mathbf{u}}\mathbf{v} - \underline{\varphi} \cdot \boldsymbol{\nabla}_{\mathbf{v}}\mathbf{u}\} \\
&= \{\mathbf{v} \cdot \boldsymbol{\nabla}_{\mathbf{u}}\underline{\varphi} - \mathbf{u} \cdot \boldsymbol{\nabla}_{\mathbf{v}}\underline{\varphi}\} + \{\underline{\varphi} \cdot [\mathbf{u}, \mathbf{v}]\} \\
&= \{\mathbf{v} \cdot \boldsymbol{\nabla}_{\mathbf{u}}\underline{\varphi} - \mathbf{u} \cdot \boldsymbol{\nabla}_{\mathbf{v}}\underline{\varphi}\} + \varphi([\mathbf{u}, \mathbf{v}]),
\end{aligned}$$

其中 $[\mathbf{u}, \mathbf{v}]$ 是交换子（commutator），如图 [29.7]，第 288 页。

最后一项 $\varphi([\mathbf{u}, \mathbf{v}])$ 包含向量场的变化——这是必须通过手术切除的，只留下第一个括号内的项作为我们对 φ 的外导数 $\mathbf{d}\,\varphi$ 的定义。

使用向量对应作为垫脚石后，我们现在可以省去它，并纯粹用 1-形式 φ 表示最终结果：

$$\boxed{\mathbf{d}\varphi(\mathbf{u}, \mathbf{v}) = \boldsymbol{\nabla}_{\mathbf{u}}\varphi(\mathbf{v}) - \boldsymbol{\nabla}_{\mathbf{v}}\varphi(\mathbf{u}) - \varphi([\mathbf{u}, \mathbf{v}]).} \tag{36.2}$$

这个公式的一个优点是它是几何定义的，独立于基和坐标的任何选择。 然而，我们也发现这个公式有两个问题，一个是美学问题，一个是哲学问题：（i）它复杂得令人失望，（ii）它的含义非常不清楚。 我们将立即解决第一个问题，本章稍后将讨论第二个问题，更深层次的问题将在下一章中讨论。

简化（36.2）的关键是将向量完全抽象出来，留下 2-形式 $\mathbf{d}\varphi$ 本身。 既然我们费了很大的力气让（36.2）独立于向量场的变化，让我们简单地选择向量场是常数，在这种情况下它们的交换子就消失了。 接下来，假设已经选择了一个基，使我们能够用笛卡尔基 1-形式 $\{\mathbf{d}x^k\}$ 来表达公式。

在这种情况下，写作 $\varphi = \varphi_i\,\mathbf{d}x^i$，并使用缩写 $\partial_k \equiv \partial_{x^k}$，我们发现（36.2）简化为，

$$\mathbf{d}\varphi(\mathbf{u}, \mathbf{v}) = \boldsymbol{\nabla}_{\mathbf{u}}\varphi(\mathbf{v}) - \boldsymbol{\nabla}_{\mathbf{v}}\varphi(\mathbf{u})$$

$$
\begin{aligned}
&= u^i\,\partial_i(v^j\,\varphi_j) - v^i\,\partial_i(u^j\,\varphi_j)\\
&= u^i v^j\ \partial_i\varphi_j\ - u^j v^i\ \partial_i\varphi_j\\
&= \partial_i\varphi_j\{\mathbf{d}x^i(\mathbf{u})\mathbf{d}x^j(\mathbf{v}) - \mathbf{d}x^j(\mathbf{u})\mathbf{d}x^i(\mathbf{v})\}\\
&= \partial_i\varphi_j\,(\mathbf{d}x^i\ \wedge \mathbf{d}x^j)(\mathbf{u},\mathbf{v}).
\end{aligned}
$$

因此，抽象出任意向量场，

$$
\boxed{\mathbf{d}\varphi = \partial_i\varphi_j\,(\mathbf{d}x^i\ \wedge \mathbf{d}x^j).} \tag{36.3}
$$

好多了！

但我们还可以做得更好！ 由于 $\mathbf{d}f = \partial_i f\,\mathbf{d}x^i$，我们现在为 $\mathbf{d}$ 的作用推导出以下优雅、自然、实用性强的表达式：

$$
\boxed{\mathbf{d}\varphi = d(\varphi_j\,\mathbf{d}x^j) = \mathbf{d}\varphi_j\ \wedge \mathbf{d}x^j.} \tag{36.4}
$$

36.2 2-形式和 p-形式的外导数

显而易见的下一步是尝试探索 $\mathbf{d}$ 在 2-形式 Ψ 上的作用。 这应该是一个 3-形式，将三个向量场作为输入：$\mathbf{v}_1$、$\mathbf{v}_2$、$\mathbf{v}_3$。 显然，$\mathbf{d}\Psi(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)$ 应该由诸如 $\boldsymbol{\nabla}_{\mathbf{v}_1}\Psi(\mathbf{v}_2,\mathbf{v}_3)$ 之类的元素构建，但是如何构建呢？

由于 $\mathbf{d}$ 将形式的度数增加 1，因此它类似于 1-形式。 因此，$\mathbf{d}\Psi(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)$ 的构造的答案可以在我们早期关于构造 1-形式和 2-形式的楔积的工作中找到，如（35.2）中给出的。从一开始就假设向量场恒定，因此我们发现

$$
\boxed{\mathbf{d}\Psi(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3) = \boldsymbol{\nabla}_{\mathbf{v}_1}\Psi(\mathbf{v}_2,\mathbf{v}_3) + \boldsymbol{\nabla}_{\mathbf{v}_2}\Psi(\mathbf{v}_3,\mathbf{v}_1) + \boldsymbol{\nabla}_{\mathbf{v}_3}\Psi(\mathbf{v}_1,\mathbf{v}_2).}
$$

同样，我们可以通过将向量抽象出来以显示 3-形式本身来以令人满意的方式简化这个公式。 我们发现（36.4）以一种明显是更大模式的一部分的方式进行推广[练习]：

$$
\boxed{\mathbf{d}\Psi = d(\Psi_{ij}\,\mathbf{d}x^i\ \wedge \mathbf{d}x^j) = d\Psi_{ij}\ \wedge \mathbf{d}x^i\ \wedge \mathbf{d}x^j.} \tag{36.5}
$$

现在很清楚 $\mathbf{d}$ 如何作用于 p-形式 $\boldsymbol{\Phi}$ 以创建 (p+1)-形式 $\mathbf{d}\boldsymbol{\Phi}$：

$$\boxed{\mathbf{d\Phi} = \mathbf{d}(\Phi_{i_1\ldots i_p}\,\mathbf{d}x^{i_1} \wedge\cdots\wedge \mathbf{d}x^{i_p}) = \mathbf{d}\Phi_{i_1\ldots i_p} \wedge \mathbf{d}x^{i_1} \wedge\cdots\wedge \mathbf{d}x^{i_p}.} \tag{36.6}$$

36.3 形式的莱布尼茨规则

验证 [练习] 很简单，如果 f 是一个函数（0-形式），那么

$$\boxed{\mathbf{d}(f\mathbf{\Psi}) = (\mathbf{d}f) \wedge \mathbf{\Psi} + f\,\mathbf{d\Psi}.} \tag{36.7}$$

但是，形式的一般莱布尼茨规则（又名乘积规则）不太明显：

$$\boxed{\mathbf{d}(\mathbf{\Phi} \wedge \mathbf{\Psi}) = (\mathbf{d\Phi}) \wedge \mathbf{\Psi} + (-1)^{\deg\Phi}\mathbf{\Phi} \wedge (\mathbf{d\Psi}),} \tag{36.8}$$

其中 $\deg\mathbf{\Phi}$ 表示 $\mathbf{\Phi}$ 的度数。

例如，如果 $\deg\mathbf{\Phi} = 1$，则 $\mathbf{\Phi} = \phi$ 是 1-形式，则

$$\boxed{\mathbf{d}(\phi \wedge \Psi) = (\mathbf{d}\phi) \wedge \Psi - \phi \wedge (\mathbf{d}\Psi).} \tag{36.9}$$

我们首先证明这个简单的例子，因为它的证明将立即推广以提供对完整规则（36.8）的解释。

具体来说，假设 Ψ 是 2-形式。 然后，使用函数的莱布尼茨规则（32.12），我们发现，

$$\begin{aligned}
\mathbf{d}(\phi \wedge \Psi) &= \mathbf{d}\left(\left[\phi_i\,\mathbf{d}x^i\right] \wedge \left[\Psi_{jk}\,\mathbf{d}x^j \wedge \mathbf{d}x^k\right]\right) \\
&= \mathbf{d}(\phi_i\,\Psi jk) \wedge \mathbf{d}x^i \wedge \mathbf{d}x^j \wedge \mathbf{d}x^k \\
&= (\Psi_{jk}\,\mathbf{d}\phi_i + \phi_i\,\mathbf{d}\Psi_{jk}) \wedge \mathbf{d}x^i \wedge \mathbf{d}x^j \wedge \mathbf{d}x^k \\
&= \left(\mathbf{d}\phi_i \wedge \mathbf{d}x^i\right) \wedge [\Psi_{jk}\,\mathbf{d}x^j \wedge \mathbf{d}x^k] + (\phi_i\,\mathbf{d}\Psi_{jk}) \wedge \mathbf{d}x^i \wedge \mathbf{d}x^j \wedge \mathbf{d}x^k \\
&= (\mathbf{d}\phi) \wedge \Psi - \left[\phi_i \wedge \mathbf{d}x^i\right] \wedge [\mathbf{d}\Psi_{jk} \wedge \mathbf{d}x^j \wedge \mathbf{d}x^k] \\
&= (\mathbf{d}\phi) \wedge \Psi - \phi \wedge (\mathbf{d}\Psi).
\end{aligned}$$

理解通式（36.8）的关键是从第四行到第五行的段落，其中我们将 $\mathbf{d}\Psi_{jk}$ 推过 $\mathbf{d}x^i$（属于 ϕ）以便将其返回到它应有的位置，靠近它自己的 $\mathbf{d}x^j \wedge \mathbf{d}x^k$，获得

$$\mathbf{d}x^i \wedge \mathbf{d}\Psi_{jk} = -\mathbf{d}\Psi_{jk} \wedge \mathbf{d}x^i.$$

在一般情况下，我们必须将 $\mathbf{d}\Psi_{jk}$ 推过所有属于 Φ 的 $\mathbf{d}x^i$——编号 $\deg\Phi$——导致 $\deg\Phi$ 符号的变化。 换句话说，$\Phi \wedge (d\Psi)$ 的符号是$(-1)^{\deg\Phi}$，这就完成了证明。

记住一般公式的一种有用方法是将 $\mathbf{d}$ 视为 1-形式。 然后，在 $\mathbf{d}$ 到达 Ψ 之前，在 $d(\Phi \wedge \Psi)$ 的右边，它必须推过 Φ 的所有 1-形式，导致 $\deg\Phi$ 符号的变化。 正如我们刚才看到的，这个助记词其实非常接近事情的真相！

36.4 闭的和恰当的形式

36.4.1基本结果：$\mathbf{d}^2 = 0$

重新考虑公式（36.3）：

$$\mathbf{d}\varphi = \partial_i\varphi_j\,(\mathbf{d}x^i \wedge \mathbf{d}x^j).$$

现在让我们使用这个 2-形式（36.5）并再次应用 $\mathbf{d}$：

$$\begin{aligned}\mathbf{d}^2\varphi &= \mathbf{d}[\partial_i\varphi_j\,(\mathbf{d}x^i \wedge \mathbf{d}x^j)]\\ &= \mathbf{d}\{\partial_i\varphi_j\} \wedge \mathbf{d}x^i \wedge \mathbf{d}x^j\\ &= \{[\partial_k\partial_i\varphi_j]\mathbf{d}x^k \wedge \mathbf{d}x^i\} \wedge \mathbf{d}x^j\\ &= 0,\end{aligned}$$

因为 $\partial_k\partial_i\varphi_j = \partial_i\partial_k\varphi_j$，但 $(\mathbf{d}x^k \wedge \mathbf{d}x^i) = -(\mathbf{d}x^i \wedge \mathbf{d}x^k)$。

凭借（36.6），很明显上述论证可以推广到任意 p-形式：$\mathbf{d}^2\boldsymbol{\Phi} = 0$。因此我们可以抽象出 p-形式，获得这个绝对基本的结果：

$$\text{每种形式都被 } \mathbf{d} \text{ 的两次应用所消灭：} \quad \boxed{\mathbf{d}^2 = 0.} \tag{36.10}$$

计算证明当然简短明了，但结果本身究竟意味着什么？ 这个谜团的解决必须等待我们对积分的讨论。 （如果你坚持破坏悬念，请参阅第 37.4 节。）

36.4.2闭的和恰当的形式

如果一个形式的外导数为零，则称该形式为闭的（closed）：

$$\boxed{\Upsilon \text{ 是闭的} \iff \mathbf{d}\Upsilon = 0.}$$

我们最终将看到，闭形式是不可压缩流体流动的通量 2-形式的高维模拟。

如果一个 p-形式 是 (p−1)-形式的外导数，则它被称为恰当的（exact）：

$$\boxed{\Upsilon \text{ 是恰当的} \iff \Upsilon = \mathbf{d}\Psi(\text{对于某些 } \Psi).}$$

在物理学的语言中，我们称 Ψ 为 Υ 的一个势（potential）。 如果存在这样的势，那么它远非唯一，因为如果 Θ 是任意的 (p-2)-形式，并且我们改变

$$\Psi \rightsquigarrow \tilde{\Psi} = \Psi + \mathrm{d}\Theta,$$

那么 [练习] $\tilde{\Psi}$ 也是一个势：$\Upsilon = \mathbf{d}\tilde{\Psi}$。 这种选择势的自由度称为规范自由度（gauge freedom），变换 $\Psi \rightsquigarrow \tilde{\Psi}$ 称为规范变换（gauge transformation）。

从（36.10）可以立即得出

$$\boxed{\text{每个恰当的形式都是闭的：} \Upsilon = \mathbf{d}\Psi \implies \mathbf{d}\Upsilon = 0}$$

有鉴于此，很自然地要问一个可能的逆向：闭形式总是恰当的吗？ 事实证明，这是一个非常有趣、深刻的问题。

简短的回答是，它取决于定义形式的区域的拓扑。

$$\boxed{\textbf{庞加莱引理。}\text{ 如果在整个简单连通区域中 } \mathrm{d}\Upsilon = 0\text{，则对于某些 } \Psi\text{，} \Upsilon = \mathrm{d}\Psi\text{。}} \tag{36.11}$$

因此，势（potential）在局部总是存在的。 但是，如果该区域不是简单连通的，则会出现全局问题，在这种情况下，可能不存在这样的 Ψ。 于是对不恰当的闭形式的研究编码有关空间拓扑的详细信息，称为德拉姆上同调（de Rham cohomology）。 这是第 37.9 节的主

题。

让我们通过对这些概念的一些实践来结束本小节。 我们要求你自行证明以下每个有趣的事实：

$$\boxed{\text{如果 } \Upsilon \text{ 和 } \Phi \text{ 是闭的，那么 } \Upsilon \wedge \Phi \text{ 也是闭的。}}$$

$$\boxed{\text{如果 } \Upsilon \text{ 是闭合的，则 } \Upsilon \wedge d\Phi \text{ 对于所有 } \Phi \text{ 都是闭合的。}}$$

$$\boxed{\text{如果 } \deg \Phi \text{ 是偶数，则 } \Phi \wedge d\Phi \text{ 是闭的。}}$$

提示：（35.4）。

36.4.3复分析：柯西-黎曼方程

在 VCA（第 216-217 页）中，我们从几何上证明了复函数 $f(z) = u + iv$ 局部是伸扭（amplitwist）——如 4.6 节所定义的——当且仅当

$$\boxed{i\ \partial_x f = \partial_y f.} \tag{36.12}$$

将实部和虚部相等，我们恢复了著名的柯西-黎曼方程（Cauchy-Riemann equations）的更常见形式，表征了共形（又名复解析）映射：

$$\boxed{\partial_x u = \partial_y v \quad \text{和} \quad \partial_x v = -\partial_y u.}$$

我们现在表明，这些方程可以用形式（Forms）语言优雅地（并且随之而来）重新表述。

考虑复数值 1-形式，

$$f\,\mathbf{d}z = (u + iv)(\mathbf{d}x + i\mathbf{d}y) = [u\,\mathbf{d}x - v\,\mathbf{d}y] + i[v\,\mathbf{d}x + u\,\mathbf{d}y].$$

然后，

$$
\begin{aligned}
\mathbf{d}(\mathrm{f}\,\mathbf{d}\mathrm{z}) &= \mathbf{d}\mathrm{f} \wedge \mathbf{d}\mathrm{z} \\
&= (\partial_{\mathrm{x}}\mathrm{f}\,\mathbf{d}\mathrm{x} + \partial_{\mathrm{y}}\mathrm{f}\,\mathbf{d}\mathrm{y}) \wedge (\mathbf{d}\mathrm{x} + \mathrm{i}\,\mathbf{d}\mathrm{y}) \\
&= (\mathrm{i}\ \partial_{\mathrm{x}}\mathrm{f} - \partial_{\mathrm{y}}\mathrm{f})\mathcal{A},
\end{aligned}
$$

其中，像往常一样，$\mathcal{A} = \mathbf{d}\mathrm{x} \wedge \mathbf{d}\mathrm{y}$ 是面积 2-形式。

因此，凭借柯西-黎曼（Cauchy-Riemann）方程 (36.12) 的紧致形式，我们得出结论：

$$
\boxed{\text{复函数 } \mathrm{f} \text{ 是局部伸扭当且仅当 1−形式 } \mathrm{f}\,\mathbf{d}\mathrm{z} \text{ 是闭的：} \mathbf{d}(\mathrm{f}\,\mathbf{d}\mathrm{z}) = 0\text{。}} \tag{36.13}
$$

关于形式积分的中心结果被称为（在本书中[1]）外微分基本定理； 这是下一章的主要成果。 有了这个定理，我们将看到（在第 37.7 节中），（36.13）立即暗示了复分析的关键结果——柯西定理！

即使我们对不是局部伸扭（即非解析函数）的复函数感兴趣，上述公式也很有价值。 特别是，我们注意到将来使用实轴上的反射是反共形的，因此映射 $\mathrm{f}(\mathrm{z}) = \mathrm{z} = \mathrm{x} - \mathrm{i}\mathrm{y}$ 不受 $\mathbf{d}(\mathrm{f}\,\mathbf{d}\mathrm{z}) = 0$ 的制约。 事实上，我们发现 [练习]

$$
\boxed{\mathbf{d}(\bar{\mathrm{z}}\,\mathbf{d}\mathrm{z}) = 2\mathrm{i}\mathcal{A}.} \tag{36.14}
$$

36.5 通过形式的向量分析

在本节中，我们简单地将外导数应用于 $\mathbb{R}^3$ 中的形式场（form fields），并发现向量分析的基本运算和恒等式自动出现，就像魔术一样——如魔法般自然发生！

正如我们在第 34.6 节和第 34.7 节中解释的那样，外微分在所有维度和所有 p-形式上都可以完美运行，但是只有在 $\mathbb{R}^3$ 中，2-形式才能伪装成向量，并且只有在 $\mathbb{R}^3$ 中，两个 1-形式的楔积才能伪装成两个向量的向量积——这些是使向量分析存在的双重咒语！

首先，我们提醒读者注意符号：$\underline{\varphi}$ 是对应于（⇄） 1-形式 φ 的向量，$\underline{\Psi}$ 是对应于通量 2-形式 Ψ 的向量，$\mathcal{V} = \mathbf{d}\mathrm{x}^1 \wedge \mathbf{d}\mathrm{x}^2 \wedge \mathbf{d}\mathrm{x}^3$ 是体积 3-形式。

我们已经看到，向量分析概念梯度向量 $\nabla\mathrm{f}$ 对应于 0-形式 f 的梯度 1-形式 $\mathbf{d}\mathrm{f}$。 现在让我们上一层，取 1-形式的外导数，明确写出 $\mathbb{R}^3$ 中出现的（36.3）的三个结果分量； 最

[1]相反，更常见的名称是广义斯托克斯定理，其扭曲的起源在第 37.3.2 节中描述。

后，让我们将答案解释为空间中流的通量 2-形式：

$$
\begin{aligned}
\mathbf{d}\varphi &= \partial_i\varphi_j\,(\mathbf{d}x^i \wedge \mathbf{d}x^j) \\
&= (\partial_2\varphi_3 - \partial_3\varphi_2)(\mathbf{d}x^2 \wedge \mathbf{d}x^3) + (\partial_3\varphi_1 - \partial_1\varphi_3)(\mathbf{d}x^3 \wedge \mathbf{d}x^1) + (\partial_1\varphi_2 - \partial_2\varphi_1)(\mathbf{d}x^1 \wedge \mathbf{d}x^2) \\
&\rightleftarrows \begin{bmatrix} \partial_2\varphi_3 - \partial_3\varphi_2 \\ \partial_3\varphi_1 - \partial_1\varphi_3 \\ \partial_1\varphi_2 - \partial_2\varphi_1 \end{bmatrix}
\end{aligned}
$$

凭借（34.10）。

瞧瞧，对应于 1-形式的向量场的旋度（curl）出现了！ 总结一下，

$$
\boxed{\mathbf{d}\varphi \quad \rightleftarrows \quad \nabla \times \underline{\varphi} = \operatorname{curl} \underline{\varphi} = \begin{bmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{bmatrix} \times \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix}.} \tag{36.15}
$$

因此，如果 φ 是闭的，即 $\mathbf{d}\varphi = 0$，则 $\underline{\varphi}$ 旋度为零：如果设想 $\underline{\varphi}$ 为空间中流体流动的速度，那么如果我们将一个小球（而不是点粒子）插入流，它将以速度 $\underline{\varphi}$ 被带走，但它不会自旋。 在这种情况下，流称为无旋（irrotational）的。 如果 $\mathbf{d}\varphi \neq 0$，那么球会自旋，$\operatorname{curl}\underline{\varphi}$ 指向自旋轴，它的大小给出了双倍[2]的自旋速率。 在这种情况下，$\operatorname{curl}\underline{\varphi}$ 称为涡量向量（vorticity vector）。

或者，我们可以将闭形式设想为对应于保守力场，随之而来的优势是 1-形式 φ 本身可以立即被赋予意义，做功，就像在第 32.3.1 节的引力示例中一样。 在那里，由于它是常数这个相当无趣的原因，φ 是闭的。 在那种引力情况下，通过将 φ 应用于沿 $\overrightarrow{pq}$ 的直接路径，可以简单地获得沿路径从 p 到 q 传输粒子所做的与路径无关的功。

根据庞加莱引理，在这个引力例子中，封闭的 1-形式 φ 是恰当的：如果 h 再次表示高度，则 gh 是单位质量的势能，则 $\varphi = \mathbf{d}(\mathrm{gh}) = \mathrm{g}\,\mathbf{d}\mathrm{h}$。 这自然地与经典表达式 $\mathrm{g}\,\mathrm{dh}$ 联系起来，因为如果粒子通过沿短向量 $\mathbf{v}$ 移动而提升 $\delta\mathrm{h}$，那么所做的功是

$$
\mathrm{g}\,\mathbf{d}\mathrm{h}(\mathbf{v}) = \mathrm{g}\,\delta\mathrm{h}.
$$

正如我们将在下一章中讨论的那样，对于非常量 φ，必须通过将路径分解为许多小向量，对每个小向量应用 φ 并求和来求值沿路径所做的功。 正如我们将要讨论的，如果 φ 是闭的，那么围绕闭合环路移动粒子所做的功就为零。

接下来，让我们取 2-形式 Ψ 的外导数，但用相应的向量 $\underline{\Psi}$ 标记它的分量，就像我们

[2]见费曼等人。 （1963 年，第 1 卷，第 40 节）。

在（34.10）中所做的那样：

$$\mathbf{\Psi} = \Psi^1\,(\mathbf{d}x^2 \wedge \mathbf{d}x^3) + \Psi^2\,(\mathbf{d}x^3 \wedge \mathbf{d}x^1) + \Psi^3\,(\mathbf{d}x^1 \wedge \mathbf{d}x^2) \rightleftarrows \underline{\Psi} = \Psi^j\,\mathbf{e}_j = \begin{bmatrix}\Psi^1\\ \Psi^2\\ \Psi^3\end{bmatrix}$$

于是，

$$\begin{aligned}\mathbf{d\Psi} &= \mathbf{d}\Psi^1 \wedge \mathbf{d}x^2 \wedge \mathbf{d}x^3 + \mathbf{d}\Psi^2 \wedge \mathbf{d}x^3 \wedge \mathbf{d}x^1 + \mathbf{d}\Psi^3 \wedge \mathbf{d}x^1 \wedge \mathbf{d}x^2\\ &= \partial_1\Psi^1\,\mathbf{d}x^1 \wedge \mathbf{d}x^2 \wedge \mathbf{d}x^3 + \partial_2\Psi^2\mathbf{d}x^2 \wedge \mathbf{d}x^3 \wedge \mathbf{d}x^1 + \partial_3\Psi^3\,\mathbf{d}x^3 \wedge \mathbf{d}x^1 \wedge \mathbf{d}x^2\\ &= (\partial_1\Psi^1 + \partial_2\Psi^2 + \partial_3\Psi^3)\mathbf{d}x^1 \wedge \mathbf{d}x^2 \wedge \mathbf{d}x^3\\ &= \left\{\begin{bmatrix}\partial_1\\ \partial_2\\ \partial_3\end{bmatrix} \cdot \begin{bmatrix}\Psi^1\\ \Psi^2\\ \Psi^3\end{bmatrix}\right\}\mathcal{V}.\end{aligned}$$

瞧瞧，对应于通量 2-形式的向量场流的散度（divergence）出现了！ 总结一下，

$$\boxed{\mathbf{d}\Psi = (\mathrm{div}\underline{\Psi})\mathcal{V} = (\nabla \cdot \underline{\Psi})\mathcal{V}.} \tag{36.16}$$

因此，如果 Ψ 是闭的，即 $\mathbf{d}\Psi = 0$，则 $\underline{\Psi}$ 的散度为零，对应于不可压缩流体在没有流体被泵入或吸出的区域中的流动，因此恰好有同样多的流体流入和流出闭合曲面。 通常，将 $\mathbf{d}\Psi$ 应用于非常小的平行六面体最终会产生流体的净通量。 这将在下一章中解释。

外微分的基本恒等式是（36.10）：

$$\mathbf{d}^2 = 0,$$

这对 $\mathbb{R}^3$ 中的向量场具有重要意义。 如果在（36.15）中我们取 $\varphi = \mathbf{d}f$，那么我们立即得到向量分析的经典结果：

$$\boxed{\mathbf{d}^2\varphi = 0 \quad \Leftrightarrow \quad \mathrm{curl}(\mathrm{grad}\ f) = \nabla \times \nabla f = 0.}$$

如果在 (36.16) 中我们取 $\Psi = \mathbf{d}\varphi$，那么我们立即得到向量分析的第二个经典结果：

$$\boxed{\mathbf{d}^2\,\Psi = 0 \quad \Leftrightarrow \quad \mathrm{div}(\mathrm{curl}\underline{\varphi}) = \nabla \cdot \nabla \times \underline{\varphi} = 0.}$$

使用形式（Forms）无需记住这些恒等式，以及工作科学家在使用向量分析时经常使用的许多更复杂的身份。此外，如果我们确实希望使用这样的恒等式，它可以使用形式（Forms）导出，而且通常比直接向量方法更快、更优雅。我们用两个微分恒等式的例子来说明这一点，

其他的留给练习 14。

然而，首先，我们将申明形式世界和（$\mathbb{R}^3$）向量世界之间一个非常有用的代数联系，并将简单的证明留给你：

$$\boxed{\varphi \wedge \Psi = (\underline{\varphi} \cdot \underline{\Psi})\mathcal{V}} \tag{36.17}$$

作为我们的第一个示例，假设我们希望找到 $\boldsymbol{\nabla} \cdot \left[\mathrm{f}\,\underline{\Psi}\right]$ 的恒等式。 根据（36.16），因此我们应该使用（36.7）计算 $\mathbf{d}\,[\mathrm{f}\,\Psi]$ ，如下所示：

$$\left(\boldsymbol{\nabla} \cdot \left[\mathrm{f}\,\underline{\Psi}\right]\right)\mathcal{V} = \mathbf{d}\,[\mathrm{f}\,\Psi] = (\mathbf{d}\mathrm{f}) \wedge \Psi + \mathrm{f}\,\mathbf{d}\Psi = \left[(\nabla \mathrm{f}) \cdot \underline{\Psi} + \mathrm{f}\,\nabla \cdot \underline{\Psi}\right]\mathcal{V},$$

得到熟悉的恒等式，

$$\boxed{\nabla \cdot \left[\mathrm{f}\,\underline{\Psi}\right] = (\nabla \mathrm{f}) \cdot \underline{\Psi} + \mathrm{f}\,\nabla \cdot \underline{\Psi}.}$$

在这种情况下，直接推导出恒等式几乎同样容易[练习]。因此，让我们继续看一个例子，在这个例子中，形式（Forms）的优势开始明显显现。

假设我们希望找到 $\nabla \times \left[\mathrm{f}\,\underline{\varphi}\right]$ 的恒等式。 根据（36.15），这对应于通量 2-形式

$$\mathbf{d}[\mathrm{f}\,\varphi] = \mathbf{d}\mathrm{f} \wedge \varphi + \mathrm{f}\,\mathbf{d}\varphi,$$

所以我们毫不费力地推导出向量分析的另一个熟悉的恒等式：

$$\boxed{\nabla \times \left[\mathrm{f}\,\underline{\varphi}\right] = \nabla \mathrm{f} \times \underline{\varphi} + \mathrm{f}\,\nabla \times \underline{\varphi}.}$$

随着向量恒等式的复杂性增加，形式（Forms）的简化能力也在增加； 见练习 14。

36.6 麦克斯韦方程

我们以对上述思想的优美而深刻的应用结束本章：将麦克斯韦的完整电磁定律封装在两个极其优美的方程中。

我们从无源方程（34.24）开始，为方便起见，我们在此重申：

$$\nabla \cdot \underline{\mathbf{B}} = 0,$$

$$\nabla \times \underline{\mathbf{E}} + \partial_t \underline{\mathbf{B}} = 0.$$

仿照贝兹和穆尼亚因（Baez and Muniain，1994，§5）的例子，让我们用 $\mathbf{d}_S$ 来表示时空外导数的空间部分，因此

$$\mathbf{d}f = \mathbf{d}_S f + \partial_t f\, \mathbf{d}t.$$

现在考虑法拉第 2-形式（34.22）的外导数：

$$\mathbf{dF} = \mathbf{d}(\varepsilon \wedge \mathbf{d}t) + \mathbf{dB}.$$

让我们对这两项分别求值。

第一，

$$\begin{aligned} \mathbf{d}(\varepsilon \wedge \mathbf{d}t) &= \mathbf{d}(\varepsilon) \wedge dt \\ &= \mathbf{d}_S(\varepsilon) \wedge \mathbf{d}t \\ &= [\nabla \times \underline{\mathbf{E}} \text{ 的通量 2 形式}] \wedge \mathbf{d}t, \end{aligned}$$

凭借 (36.15)。

第二，

$$\begin{aligned} \mathbf{dB} &= \mathbf{d}_S\mathbf{B} + [\partial_t B_x\, \mathbf{d}t \wedge (\mathbf{d}y \wedge \mathbf{d}z) + \cdots] \\ &= (\nabla \cdot \mathbf{B})\mathcal{V} + \mathbf{d}t \wedge [\partial_t B_x\, (\mathbf{d}y \wedge \mathbf{d}z) + \cdots] \\ &= (\nabla \cdot \mathbf{B})\mathcal{V} + \mathbf{d}t \wedge \partial_t \mathbf{B} \end{aligned}$$

凭借（36.16）。

但是 $\mathbf{d}t \wedge (\partial_t \mathbf{B}) = (\partial_t \mathbf{B}) \wedge \mathbf{d}t$，由于（35.3），因为 $\partial_t \mathbf{B}$ 是 2-形式。 因此，由于

$$\partial_t \mathbf{B} = \text{通量 2 形式 } \partial_t \underline{\mathbf{B}},$$

结合前两个结果得到

$$\mathbf{dF} = (\nabla \cdot \underline{\mathbf{B}})\mathcal{V} + \{\text{通量 2 形式的}[\nabla \times \underline{\mathbf{E}} + \partial_t \underline{\mathbf{B}}]\} \wedge \mathbf{d}t。$$

因此，向上伸手，我们现在可以触摸上帝的脸：

麦克斯韦无源方程表明法拉第 2−形式是闭的：

$$\mathbf{dF} = 0$$

正如我们已经说过的，并将在下一章进一步讨论，在局部，每个闭的形式都是恰当的。换句话说，存在一个 1-形式[3]势 $\mathbf{A}$，使得

$$\boxed{\mathbf{F} = \mathbf{d}\mathbf{A}。}$$

注：关于电动力学的经典文献改为采用相应的向量场 $\underline{\mathbf{A}}$，称为向量势。

接下来，我们将注意力转向第二对麦克斯韦方程（34.25），它描述了有源场——电荷密度 ρ 和电流密度 $\underline{\mathbf{j}}$：

$$\nabla \cdot \underline{\mathbf{E}} = 4\pi\rho,$$
$$\nabla \times \underline{\mathbf{B}} - \partial_t \underline{\mathbf{E}} = 4\pi\, \underline{\mathbf{j}}.$$

为了了解这些方程真正试图向我们展示什么，我们必须首先引入源的时空版本，组合成一个单一的 1-形式：

$$\mathbf{J} = -\rho\, \mathbf{d}t + \mathbf{j},$$

其中（空间）1-形式[4] $\mathbf{j}$ 对应于（空间）向量电流密度 $\underline{\mathbf{j}}$：

$$\underline{\mathbf{j}} = \begin{bmatrix} j^1 \\ j^2 \\ j^3 \end{bmatrix} \quad \rightleftarrows \quad \mathbf{j} = j^1\, \mathbf{d}x^1 + j^2\, \mathbf{d}x^2 + j^3\, \mathbf{d}x^3.$$

因此，当 $\mathbf{J}$ 作用于纯空间 4-向量时，它产生在该空间方向上流动的电流量，但当它作用于静止观察者的时间轴 4-速度时，它描述了随观察者流动的电荷 时间，即观察者所在位置的电荷密度 ρ 。更一般地说，如果 u 是观察者的 4 速度，那么她测量的电荷密度为 J(u)。

在练习 15 中，你将看到在 4 维时空中，霍奇星算子 (⋆) 将 p 型映射到对偶 (4−p) 型。特别是，源密度 1-形式 J 的对偶是 3-形式，

$$\star\, \mathbf{J} = -\rho\mathcal{V} + [\,\underline{\mathbf{j}} \text{ 的通量 2 形式}] \wedge \mathbf{d}t, \tag{36.18}$$

[3]这种表示法是通用的，因此在这种情况下，我们必须放弃使用小写希腊字母表示 1-形式的惯例。
[4]这种表示法是通用的，因此，我们必须再次放弃使用小写希腊字母表示 1-形式的惯例。

法拉第 2-形式的对偶是麦克斯韦 2-形式，（34.26）：

$$\star\,\mathbf{F} = \text{Maxwell} = \beta\ \wedge \mathbf{d}\mathrm{t} - \mathbf{E}.$$

由于麦克斯韦可以通过交换电场和磁场从法拉第获得，并改变第一项的符号，所以上面的 $\mathbf{dF}$ 的整个计算可以被用来简单地写下 $\mathbf{d}\star\mathbf{F}$ 的答案：

$$\begin{aligned}\mathbf{d}\ \star\ \mathbf{F}\ &=\ -(\nabla\cdot\underline{\mathbf{E}})\mathcal{V} + \{\left[\nabla\times\underline{\mathbf{B}} - \partial_{\mathrm{t}}\underline{\mathbf{E}}\right]\text{的通量 2 形式}\}\wedge\mathbf{d}\mathrm{t}\\ &=\ 4\pi\left[-\rho\mathcal{V} + \left\{\underline{\mathbf{j}}\text{的通量 2 形式}\right\}\wedge\mathrm{dt}\right],\end{aligned}$$

凭借上面的两个麦克斯韦有源方程。

我们得出了一个非常了不起且非常漂亮的结论：

> 麦克斯韦 2−形式服从这个自然法则：
>
> $$\mathbf{d}\star\mathbf{F} = 4\pi\star\mathbf{J}.$$

有关形式语言中麦克斯韦方程组的进一步讨论，请参阅 Hubbard 和 Hubbard (2009, §6.11)、Baez 和 Muniain (1994, §§1,5) 以及 Misner、Thorne 和 Wheeler (1973, § 4)。

第三十七章 积分

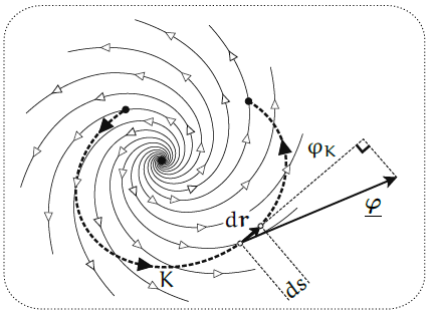
37.1 1-形式的线积分

37.1.1环量与功

以速度 $\underline{\varphi}$ 沿有向曲线 K 流动的流体的环量 $\mathcal{C}_{\mathrm{K}}(\underline{\varphi})$ 先前被定义为（27.2）为流体沿 K 流动分量的积分：

$$\mathcal{C}_{\mathrm{K}}(\underline{\varphi}) \equiv \int_{\mathrm{K}} \underline{\varphi} \cdot \mathrm{d}\mathrm{r},$$

如[37.1]所示。 或者，同样重要的是，我们可以将 $\underline{\varphi}$ 视为一个力场，在这种情况下，相同的积分现在被解释为力场沿 K 移动粒子时所做的功。



[37.1] 粒子沿路径 K 移动距离 ds 所做的功为 $\varphi_{\mathrm{K}}\,\mathrm{ds}$，其中 φ_{K} 是力场 $\underline{\varphi}$ 在 K 方向上的分量。

为了简单起见，我们选择在 $\mathbb{R}^2$ 中说明这个想法，但我们应该强调，这个概念本质上不局限在二维：想象流体是海洋，或力场是重力。 事实上，我们将要讨论的一切都适用于（不变）作用于 $\mathbb{R}^{\mathrm{n}}$ 中向量的 1-形式。

将其翻译成形式（Forms）的语言，我们可以这样写：

$$\varphi(\mathrm{d}\mathbf{r}) = \underline{\varphi} \cdot \mathrm{d}\mathbf{r} = \varphi_{\mathrm{K}}\,\mathrm{ds}.$$

这里，如[37.1]所示，$\varphi_{\mathrm{K}}\,\mathrm{ds}$ 是 $\underline{\varphi}$ 在 K 方向上的投影，$\mathrm{ds} = |\mathrm{d}\mathbf{r}|$。 换言之，1-形式 φ 作用于有向曲线的一小段[1] $\mathrm{d}\mathbf{r}$，产生沿运动方向的力的分量，乘以该小运动的长度，即在执行该运动时所做的功。 然后可以将完整积分视为完成的总功，通过对功的增量求和获得，

[1]在这里我们很懒惰：我们应该写 δr，并使用终极等式。

因为 1-形式的 φ 沿着有向曲线 K 前进，一次只咬 $\mathbf{dr}$ 大的一小口。

我们强调*只有在运动方向上的力的分量才做功*。因此，例如，如果 K 是场线的正交轨迹，则不会做任何功，即积分为零。

以这种方式思考，我们得到了 1-形式 φ 沿 K 的积分的定义，写成

$$\boxed{\int_{\mathrm{K}} \varphi \equiv \mathcal{C}_{\mathrm{K}}(\underline{\varphi}) \equiv \int_{\mathrm{K}} \varphi_{\mathrm{K}}\, \mathrm{ds}.} \tag{37.1}$$

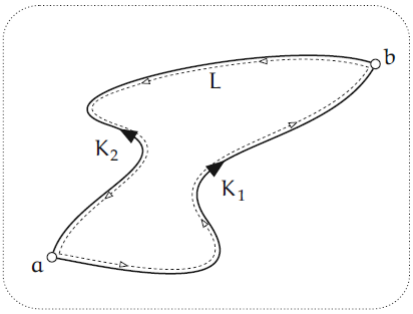
在这个定义中，积分符号下 d（*某物*）的缺失（在左侧）是不熟悉的，起初可能令人不安：请记住，这不是必需的，因为可以*理解*为 φ 沿着 K 一路吃下去，一次一小口 $\mathbf{dr}$。

在图示的情况下，很明显 K 总是在与流场或力场大致相同的方向上移动，因此积分显然是正的。但是假设我们穿过相同的曲线，但方向相反，表示为 $-\mathrm{K}$。现在我们逆流而上，或者，使用另一种解释，我们正在推动粒子抵抗力场的阻力。因此积分现在是负的。更准确地说，由于 $\varphi(-\mathbf{dr}) = -\varphi(\mathbf{dr})$，因此得出

$$\boxed{\int_{-\mathrm{K}} \varphi = -\int_{\mathrm{K}} \varphi.} \tag{37.2}$$

37.1.2路径无关⇔环路积分为零

如果 K 连接两个固定点 a 和 b，那么一般来说，$\mathcal{C}_{\mathrm{K}}(\underline{\varphi})$ 的值将取决于所选择的从 a 到 b 的特定路径 K。然而，对于许多重要的物理流体流和力场，我们有*路径无关性*（path independence），顾名思义：从 a 到 b 的所有路径 K 产生相同的积分值。



[37.2] 如果积分与 a 和 b 之间的路径 K 无关，则沿闭环 L 的积分为零。相反，如果所有环路积分都为零，则积分是路径无关的。

这个概念可以用一种重要且有用的方式重新表述。 如 [37.2] 中所示，考虑连接 a 到 b 的任意两条路径 K_1 和 K_2。 然后我们可以创建一个闭环 L，首先沿着 K_1 从 a 到 b，然后从 b 沿着 $-K_2$ 回到 a，因此 $L = K_1 - K_2$。

而路径无关意味着环路积分为零：

$$\int_L \varphi = \int_{K_1 - K_2} \varphi = \int_{K_1} \varphi - \int_{K_2} \varphi = 0.$$

相反，假设我们知道所有的环路积分都为零。 由此可见，积分是路径无关的。 因为如果 K_1 和 K_2 是连接 a 到 b 的任意两条路径，那么我们可以形成闭环 $L = K_1 - K_2$，并且 $\mathcal{C}_L(\underline{\varphi})$ 的事实意味着

$$\int_{K_1} \varphi = \int_{K_2} \varphi.$$

简单来说，

$$\boxed{\text{路径无关} \quad \Leftrightarrow \quad \text{环路积分为零.}} \tag{37.3}$$

再看[37.1]，想象我们扩展 K，使它闭合成为一个环绕涡旋的椭圆环。 显然，围绕这个环路的积分并不是零，实际上是个正数。 因此，这里我们有一个例子，其中 φ 的积分确实取决于连接两点的路径。

37.1.3恰当形式的积分：$\boldsymbol{\varphi = \mathbf{d}f}$

如果 φ 是恰当的，那么根据定义，对于某个函数 f，$\varphi = \mathbf{d}f$。 如果 $\delta\mathbf{r}$ 是沿积分路径 K 的一个短暂的、最终趋于零的运动，那么它对积分的贡献为

$$\varphi(\delta\mathbf{r}) = \mathbf{d}f(\delta\mathbf{r}) \asymp \delta f \equiv f \text{ 从 } \delta r \text{ 的尾部到尖端的变化。}$$

因此，完整积分意味着 f 从 K 的起点到 K 的终点的净变化。总而言之，

$$\boxed{\begin{array}{l}\text{恰当的 1-形式 } \varphi = \mathbf{d}f \text{ 的积分是路径无关的，由 f 沿 K 的变化给出：} \\ \qquad\qquad \displaystyle\int_K \mathbf{d}f = f(b) - f(a).\end{array}} \tag{37.4}$$

用物理学的语言来说，$\varphi = \mathbf{d}\mathrm{f}$ 对应一个保守力场，$\underline{\varphi} = \nabla\mathrm{f}$，f 是势能（potential energy）。那么势能的变化 $[\mathrm{f(b)} - \mathrm{f(a)}]$ 是力场将粒子从 a 带到 b 过程中所做的与路径无关的功，因此（37.4）只是下列熟悉事实的另一种书写方式

$$\int_{\mathrm{K}} (\nabla\mathrm{f}) \cdot \mathrm{d}\mathbf{r} = \mathrm{f(b)} - \mathrm{f(a)}. \tag{37.5}$$

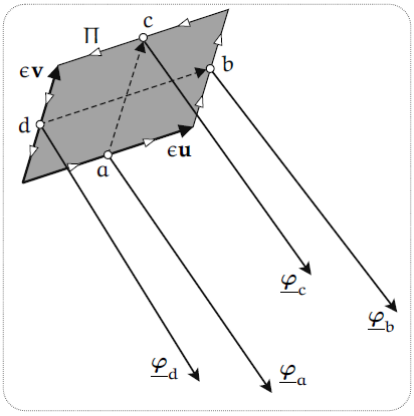
正如费曼（Feynman）等人所讨论的那样（1963 年，第 1 卷，第 14-4 节），在最基本的层面上，根据能量守恒的基本原理，自然界中的所有力都是保守的。

37.2 作为积分的外导数

37.2.1d(1-形式)

从我们第一次接触微积分开始，我们就被教导微分和积分互为逆运算——此即微积分基本定理。 事实上，我们刚刚在（37.4）中看到（或者看起来）这个想法的一个新实例。

因此，当得知到了形式这里，外微分是（围绕一个小闭环或一个小封闭曲面的）积分，而不是它的逆运算，这可能会让人感到迷惑不解！



[37.3] 围绕小平行四边形 Π 的 1-形式 φ 的积分最终等于将其外导数 $\mathbf{d}\varphi$ 应用于 Π 的边界。

令 $\Pi(\epsilon\mathbf{u}, \epsilon\mathbf{v})$ 是以 $\epsilon\mathbf{u}$ 为第一条边和 $\epsilon\mathbf{v}$ 为第二条边的小平行四边形的有向边界，这个有向边界随着 ϵ 趋于零，而最终趋于零，如 [37.3] 所示。 让我们将 $\Omega(\epsilon\mathbf{u}, \epsilon\mathbf{v})$ 定义为 1-形式 φ 围绕 Π 的积分：

$$\boxed{\Omega(\epsilon\mathbf{u}, \epsilon\mathbf{v}) \equiv \oint_{\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v})} \varphi}$$

我们现在将证明这个积分是一个 2-形式，而又不是普通 2-形式——它是 φ 的外导数！就是说，

$$\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v}) \asymp \mathbf{d}\varphi(\epsilon\mathbf{u},\epsilon\mathbf{v}).$$

抽象出向量，

$$\boxed{\Omega \asymp \mathbf{d}\varphi.} \tag{37.6}$$

Ω 确实具有 2-形式所需的反对称性，这直接源于 $\Pi(\epsilon\mathbf{v},\epsilon\mathbf{u}) = -\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v})$，因为这意味着

$$\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v}) = \oint_{\Pi(\epsilon\mathbf{v},\epsilon\mathbf{u})} \varphi = \oint_{-\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v})} \varphi = -\oint_{\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v})} \varphi = -\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v}).$$

也可以直接根据定义从几何上确认满足线性要求，但我们将直接切入正题，证明完整的结果，（37.6），从而在这过程证实线性，并（重新）证实反对称性。 为此，我们只需要推广我们之前在 [27.2] 中给出的论证。

然而，首先让我们试着想象一面旗帜在我们的最终目的地上空飘扬，这样我们就知道应该朝哪个方向前进。 为了使 $\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v})$ 成为一个封闭的平行四边形环，正如在[29.7] 和即将到来的 [37.4] 中解释的那样，它的边的交换子必然为零：

$$[\epsilon\mathbf{u}\,,\epsilon\mathbf{v}] = 0.$$

这很容易通过选择恒定的向量场来实现。 在这种情况下，我们对外导数的第一个表达式（36.2）简化为

$$\mathbf{d}\varphi(\epsilon\mathbf{u},\epsilon\mathbf{v}) = \nabla_{\epsilon\mathbf{u}}\,\varphi(\epsilon\mathbf{v}) - \nabla_{\epsilon\mathbf{v}}\,\varphi(\epsilon\mathbf{u}), \tag{37.7}$$

这是我们的旗帜。

目标明确后，让我们回到积分 $\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v})$ 的求值。 由于我们最终将通过让 ϵ 变为零来将平行四边形缩小到零，因此我们可以使用黎曼和，平行四边形的每边只取一个项。 通常使用黎曼和来求值沿某个固定区间的积分，将其分解为越来越多的越来越小的子区间，因此我们在每个子区间上用于求值 φ 的精确点位置在此极限下变得无关紧要。

然而，这不是我们在这里所做的：我们寻求当 ϵ 趋于零时，积分随之趋于零的极限形

式（双关语）。 如果我们只是在每一边选一个随机点，那么误差会随着 ϵ^2 而趋于零，这还不够好。 因为很明显，Π 对边对 Ω 的贡献将抵消到 ϵ 阶，因此 Ω 本身将是 ϵ^2 阶——与随机黎曼和产生的误差阶数相同。

答案是选择边的中点，因为这样误差趋于零快得多——实际上是 ϵ^3 阶。（这在 VCA, p. 382 中有几何证明。）因此，通过使用中点，我们立即达到了我们需要[2]的精度，以找到 Ω 的净结果，即 ϵ^2 阶。

因此，如 [37.3] 中所示，让我们使用这些中点 a、b、c、d 来近似 $\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v})$，并让 a 处的对应向量为 $\underline{\varphi}_{\mathrm{a}}$，依此类推，如图所示。 然后，在平行四边形的中心取所有要计算的导数，

$$\begin{aligned}\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v}) &\asymp \varphi_{\mathrm{a}}(\epsilon\mathbf{u})+\varphi_{\mathrm{b}}(\epsilon\mathbf{v})+\varphi_{\mathrm{c}}(-\epsilon\mathbf{u})+\varphi_{\mathrm{d}}(-\epsilon\mathbf{v})\\ &= [\varphi_{\mathrm{b}}(\epsilon\mathbf{v})-\varphi_{\mathrm{d}}(\epsilon\mathbf{v})]-[\varphi_{\mathrm{c}}(\epsilon\mathbf{u})-\varphi_{\mathrm{a}}(\epsilon\mathbf{u})]\\ &\asymp \nabla_{\epsilon\mathbf{u}}\,\varphi(\epsilon\mathbf{v})-\nabla_{\epsilon\mathbf{v}}\,\varphi(\epsilon\mathbf{u})\\ &= \mathbf{d}\varphi(\epsilon\mathbf{u},\epsilon\mathbf{v}),\end{aligned}$$

凭借（37.7）。 抽象掉向量，这就完成了（37.6）的证明。

我们强调前面的论点本质上没有限制在二维：将 [37.3] 中的 $\underline{\varphi}$ 向量想象为伸出书页，然后拿起并倾斜书本，使 Π 占据空间中的任何位置！

如果我们现在将向量场 $\underline{\varphi}$ 视为空间中的流体流动，那么结果（37.6）为外导数附加了一个非常具体而生动的含义：

> 如果流体以速度 $\underline{\varphi}$ 流动，那么它围绕一个小的（最终趋于零的）平行四边形的环量 Ω 最终等于将通量 2-形式 $\mathbf{d}\varphi$ 应用于其边的结果。

(37.8)

为了将来的使用，让我们用经典向量分析的语言将这个关键结果表达为一个明确的公式。令 $\hat{\mathbf{n}}$ 为平行四边形的法线单位，其方向由右手定则确定，将手指从 $\mathbf{u}$ 卷曲到 $\mathbf{v}$。另外，设 $\delta\mathcal{A}$ 为平行四边形的面积：

$$\delta\mathcal{A} = |\epsilon\,\mathbf{u}\times\epsilon\,\mathbf{v}|,$$

[2] 我们首先在 VCA（第 412-413 页）中使用这种推理来提供柯西定理的直接几何证明。

以便

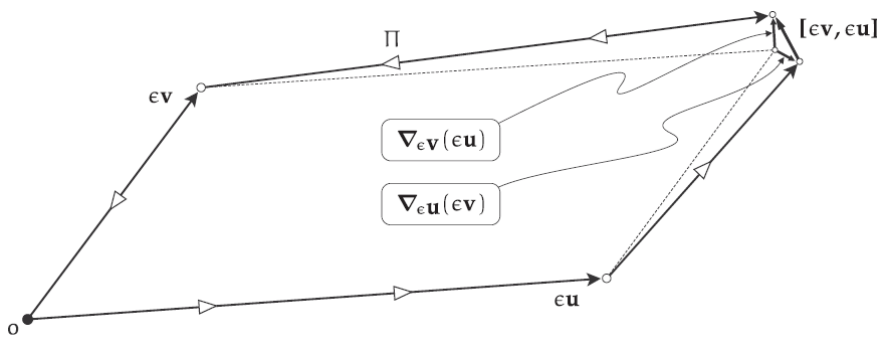
$$\epsilon\,\mathbf{u}\times\epsilon\,\mathbf{v} = \hat{\mathbf{n}}\,\delta\mathcal{A}。$$

$\nabla\times\underline{\varphi}$ 通过平行四边形的通量是它垂直于平行四边形的分量（即它在 $\hat{\mathbf{n}}$ 方向上的分量）乘以平行四边形的面积：

$$\boxed{\text{通量} \asymp (\nabla\times\varphi)\cdot\hat{\mathbf{n}}\,\delta\mathcal{A}.} \tag{37.9}$$

然后，结合 （36.15）、（37.8）和（37.9），

$$\boxed{\oint_{\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v})} \varphi = \Omega(\epsilon\mathbf{u},\epsilon\mathbf{v}) \asymp \mathbf{d}\varphi(\epsilon\mathbf{u},\epsilon\mathbf{v}) = (\nabla\times\underline{\varphi})\cdot\hat{\mathbf{n}}\,\delta\mathcal{A}.} \tag{37.10}$$



[37.4] 一般向量场不会创建闭合平行四边形。 然而，我们可以通过用交换子桥接间隙来创建一个闭环 Π：$[\epsilon\mathbf{v},\epsilon\mathbf{u}] = \nabla_{\epsilon\mathbf{v}}\,\epsilon\mathbf{u} - \nabla_{\epsilon\mathbf{u}}\epsilon\mathbf{v}$。

我们以另一个奖励结束：外导数（36.2）的原始完整公式的几何解释，包括它以前神秘的交换子项。 回想一下，我们首先通过纯粹的形式计算（尽管是一种开明的计算）得出了以下 dϕ 公式：

$$\mathbf{d}\varphi(\mathbf{u},\mathbf{v}) = \nabla_{\mathbf{u}}\varphi(\mathbf{v}) - \nabla_{\mathbf{v}}\varphi(\mathbf{u}) - \varphi([\mathbf{u},\mathbf{v}]).$$

我们选择通过坚持恒定向量场来简化前面的阐述，但是如果 $\mathbf{u}$ 和 $\mathbf{v}$ 不是恒定的呢？然后，如 [37.4] 中所示，向量场通常不会创建闭环。 然而，我们可以通过用交换子 $[\epsilon\mathbf{v},\epsilon\mathbf{u}]$ 桥接间隙来创建一个闭环 Π。 上述分析的大部分内容没有改变，但现在积分 $\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v})$ 从间隙闭合交换子边缘获得了一个额外的贡献：

$$\varphi([\epsilon\mathbf{v},\epsilon\mathbf{u}]) = -\varphi([\epsilon\mathbf{u},\epsilon\mathbf{v}]),$$

从而证明 $\mathbf{d}\varphi$ 的公式！

37.2.2d(2-形式)

让我们重复上面的分析，在以 $\epsilon\mathbf{u}$ 为第一条边，$\epsilon\mathbf{v}$ 为第二条边，$\epsilon\mathbf{w}$ 为第三条边的右手平行六面体的有向二维边界 $\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w})$ 上对 2-形式 Ψ 的积分，随着 ϵ 趋于零，而最终趋于零。

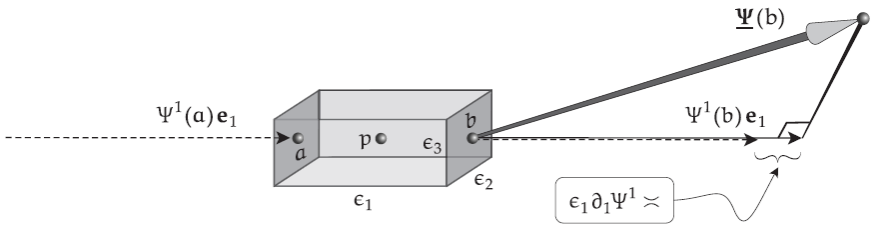
让我们将 $\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w})$ 定义为 Ψ 在 Π 上的积分：

$$\boxed{\Omega(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w}) \equiv \iint\limits_{\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w})} \Psi.}$$

在经典向量分析的语言中，这是 $\underline{\Psi}$ 通过 Π 的向外的净通量：

$$\boxed{\Omega(\Pi) = \iint\limits_{\Pi} \underline{\Psi}\cdot\hat{\mathbf{n}}\,\delta\mathcal{A},}$$

其中 $\hat{\mathbf{n}}$ 是平行六面体指向外的单位法线。



[37.5] 从右侧面流出的通量$\asymp\ \epsilon_2\epsilon_3\,\Psi^1(\mathrm{b})$。 从左面流出的通量 $\asymp\ -\epsilon_2\epsilon_3\,\Psi^1(\mathrm{a})$。 因此，通过这对面的净流出量为 $\asymp\ [\partial_1\Psi^1\epsilon_1]\,\epsilon_2\epsilon_3\ =\ [\partial_1\Psi^1]\mathcal{V}$。

根据 1-形式的结果（37.6），我们预计这个积分是 3-形式，而又不仅仅是任何 3-形式——它应该是 Ψ 的外导数：

$$\iint\limits_{\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w})} \Psi = \Omega(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w}) \ \asymp\ \mathbf{d}\Psi(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w}) = (\nabla\cdot\underline{\Psi})\mathcal{V}(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w}),$$

凭借（36.16）。 因此，将向量抽象出来，我们被引导相信

$$\boxed{\Omega \asymp \mathbf{d}\Psi = (\nabla \cdot \underline{\Psi})\mathcal{V}.} \tag{37.11}$$

这确实是正确的，并且在平行六面体的边缘（我们认为它以 p 为中心）是正交的情况下，证明是最简单和最透明的，沿着正交基向量指向：$\{\epsilon_1\mathbf{e}_1, \epsilon_2\mathbf{e}_2, \epsilon_3\mathbf{e}_3\}$，所以它的体积是 $\mathcal{V} = \epsilon_1\epsilon_2\epsilon_3$。

如 [37.5] 所示，考虑 Ψ 通过与 $\mathbf{e}_1$ 正交的两个面的通量，这两个面相隔距离 ϵ_1，其中心为 a 和 b。只有分量 $\Psi^1\mathbf{e}_1$ 携带流体穿过这些面；其他两个分量平行于面流动。（为明确起见，在此图中，我们假设 $\Psi^1 > 0$。）

为了计算右侧面的通量，我们将其面积 $\epsilon_2\epsilon_3$ 乘以面中心速度的正交分量 $\Psi^1(\mathrm{b})$，因此

$$\text{从右侧面流出的通量} \asymp \epsilon_2\epsilon_3\,\Psi^1(\mathrm{b})。$$

对于左侧面，向外的法线 $\hat{\mathbf{n}} = -\mathbf{e}_1$ 与右侧面的向外法线相反，所以

$$\text{左侧面流体向外的速度} = \underline{\Psi} \cdot \hat{\mathbf{n}} = -\Psi^1(\mathrm{a})。$$

所以，

$$\text{流出左侧面的通量} = -(\text{流入左侧面的通量}) \asymp -\epsilon_2\epsilon_3\,\Psi^1(\mathrm{a})。$$

请注意，在 [37.5] 中，我们绘制了 $\Psi^1(\mathrm{a})\,\mathbf{e}_1$（虚线），其箭头位于 a，既是为了整理图片，又是为了清楚地表明这表示流入盒子。 请注意，我们还在 b 处绘制了 $\Psi^1(\mathrm{a})\,\mathbf{e}_1$ 的副本，以便于即将与 $\Psi^1(\mathrm{b})\,\mathbf{e}_1$ 进行比较

加上这些流出通量，

$$\begin{aligned}\text{通过这对面的净流出量} &\asymp [\,\Psi^1(\mathrm{b}) - \Psi^1(\mathrm{a})\,]\epsilon_2\epsilon_3 \\ &\asymp [\partial_1\Psi^1\epsilon_1]\,\epsilon_2\epsilon_3 \\ &= [\partial_1\Psi^1]\mathcal{V},\end{aligned}$$

其中导数在盒子的中心 p 进行求值。

当然，完全相同的推理也适用于与 $\mathbf{e}_2$ 正交的一对面，通过该对产生 $\left[\partial_2\Psi^2\right]\mathcal{V}$ 的净流出量。 最后一对同样贡献 $\left[\partial_3\Psi^3\right]\mathcal{V}$。 因此，

$$\Omega(\Pi) = \text{流出 } \Pi \text{ 的总通量} \asymp [\partial_1\Psi^1 + \partial_2\Psi^2 + \partial_3\Psi^3]\mathrm{V} = (\nabla \cdot \underline{\Psi})\mathcal{V},$$

从而证明（37.11）。

结果（37.6）和（37.11）是继续进入更高维度的模式的一部分。 尽管可视化让我们失望，但将 $\mathbf{d}$(3-形式) 应用于 4 维空间的一个小的、紧致的、最终趋于零的区域确实会产生 3-形式在其 3 维边界上的积分，依此类推。有关这些高维结果的更多信息，请参阅本书末尾的延伸阅读部分。

37.3 外微分的基本定理（广义斯托克斯定理）

37.3.1外微分的基本定理

你在向量分析中学到的所有积分定理——（37.5）、格林定理、斯托克斯定理和高斯定理——只是关于形式的一个优雅定理的特例。

外微分基本定理：

$$\boxed{\int_{\mathrm{R}} \mathbf{d}\Phi = \int_{\partial \mathrm{R}} \Phi} \tag{37.12}$$

在左侧，(p+1)-形式 $\mathbf{d}\Phi$ 在紧致（定向）（p+1）维区域 R 上积分，在右侧，p-形式 Φ 在（定向）R 的 p 维边界 ∂R 上积分。请注意，我们必须放弃多重积分符号，而在等式两侧使用单重积分符号，因为度数 p 是一般且未知的。

警告：我们所称的外微分基本定理（Fundamental Theorem of Exterior Calculus，以下简称 FTEC）实际上在所有其他书籍中被称为广义斯托克斯定理（Generalized Stokes's Theorem），通常缩写为 GST。

37.3.2历史旁白

毫无疑问，斯托克斯（Stokes）对科学做出了巨大贡献，但这个定理以他且仅以他的名字命名，是非常不合适的。 Vladimir Arnol’d (1989, p. 192) 将（37.12）称为“牛顿-莱布尼茨-高斯-格林-奥斯特罗格拉茨基-斯托克斯-庞家莱公式”（Newton-Leibniz-Gauss-Green-Ostrogradskii-Stokes-Poincaré formula）！ 然而，在这个列表中，斯托克斯实际上与这个定理的关系最小。

1854 年，斯托克斯设计了剑桥大学的史密斯奖考试。 第 8 题要求学生证明原始的、预先推广的结果——我们现在称之为斯托克斯定理； 这是该结果在印刷品中的首次亮相。

参加那次考试的候选人之一正是詹姆斯·克拉克·麦克斯韦（James Clerk Maxwell）——顺便说一句，他并列第一——此定理最初以斯托克斯命名可能正是源于这一事件！

但斯托克斯只知道结果，因为威廉·汤姆森（开尔文勋爵）在 1850 年 7 月 2 日的一封信中把它寄给了他！ 但开尔文似乎是通过研究格林得出了这个结果…！你开始明白我们现在的处境是多么的麻烦； 如需额外的麻烦，请参阅卡茨（Katz，1979）和克劳（Crowe，1985）。

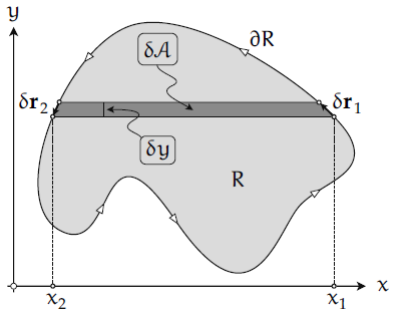
伍德豪斯（N.M.J. Woodhouse）向罗杰·彭罗斯（Roger Penrose）提出了一个亚历山大砍断戈耳狄俄斯之结式的方案来解决这个命名难题：抛弃所有这些名称，取而代之以外微分基本定理。 这是彭罗斯和林德勒（Penrose and Rindler, 1984）以及随后在 Penrose (2005) 中采用的解决方案，我们在此遵循这些先例。

至于一般结果的历史，（37.12），我们只说：第一个完整的陈述发生在嘉当于 1936-1937 年在巴黎的一次讲座中，他在其中明确指出向量分析的所有结果只是特例； 这些讲座首次出现在嘉当（Cartan , 1945）中。

37.3.3示例：面积

回到结果本身，让我们从这个定理的一个初等几何示例开始。

如 [37.6] 中所示，在 $\mathbb{R}^2$ 中，考虑 1-形式 $\mathrm{x}\,\mathbf{d}\mathrm{y}$ 逆时针围绕（面积为 $\mathcal{A}(\mathrm{R})$ 的）区域 R 的边界 $\partial\mathrm{R}$ 的积分。



[37.6] $\delta\mathbf{r}_1$ 和 $\delta\mathbf{r}_2$ 对 $\oint_{\partial\mathrm{R}} \mathrm{x}\,\mathbf{d}\mathrm{y}$ 的贡献总和为 $\delta\mathcal{A}$。

让我们计算夹在高度 y 和 $\mathrm{y}+\delta\mathrm{y}$ 的水平线之间的图示匹配运动对 $\delta\mathbf{r}_1$ 和 $\delta\mathbf{r}_2$ 对积分的贡献：

$$(\mathrm{x}\,\mathbf{d}\mathrm{y})(\delta\mathbf{r}_1) + (\mathrm{x}\,\mathbf{d}\mathrm{y})(\delta\mathbf{r}_2) = (\mathrm{x}_1 - \mathrm{x}_2)\,\delta\mathrm{y} \asymp \delta\mathcal{A}.$$

因此，

$$\oint_{\partial \mathrm{R}} \mathrm{x}\,\mathbf{d}\mathrm{y} = \mathcal{A}(\mathrm{R}).$$

（检查我们向下滑动阴影水平条纹直到它分成两条时，这个论证是否继续正常工作。）

这确实符合外微分基本定理（FTEC）（37.12），因为它预测

$$\oint_{\partial \mathrm{R}} \mathrm{x}\,\mathbf{d}\mathrm{y} = \iint_{\mathrm{R}} \mathbf{d}(\mathrm{x}\,\mathbf{d}\mathrm{y}) = \iint_{\mathrm{R}} \mathbf{d}\mathrm{x} \wedge \mathbf{d}\mathrm{y} = \iint_{\mathrm{R}} \mathcal{A} = \mathcal{A}(\mathrm{R}).$$

这也具有物理意义。根据对应关系（34.10），$\mathrm{dx} \wedge \mathrm{dy}$ 对应于 z 方向上的流动，与 (x,y) 平面正交，具有单位速度。 因此，它通过面积为 $\mathcal{A}$ 的区域的通量等于 $\mathcal{A}$。

37.4 边界的边界是零！

在我们证明外微分基本定理或 FTEC 之前，让我们看看它是如何解决我们长期以来一直存在的一个谜团：为什么 $\mathbf{d}^2 = 0$？ 这个基本结果的计算证明很简单（见第 36.4.1 节），但结果本身究竟意味着什么？！

如果我们应用 FTEC 两次，我们发现

$$0 = \int_{\mathrm{R}} \mathbf{d}^2\Phi = \int_{\partial \mathrm{R}} \mathbf{d}\Phi = \int_{\partial(\partial \mathrm{R})} \Phi.$$

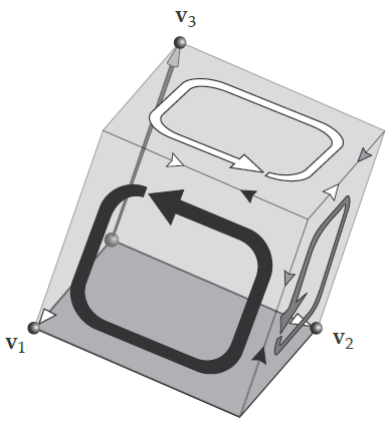
但由于 Φ 和 R 是任意的，这只能在下式成立时为真

$$\boxed{\partial^2 = 0 \iff \text{边界的边界为零！}}$$

反过来，如果我们能理解这个几何命题，我们就已经证明了 $\mathbf{d}^2 = 0$ 的几何意义：

$$\boxed{\mathbf{d}^2 = 0 \iff \partial^2 = 0\text{。}}$$

让我们从 3-2-1 维度来解释这一点。 考虑 [37.7]，它显示了一个右手的 3 维平行六面体，边为 $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$。 它的边界由六个二维面组成，具有指定的方向。 最后，边界的边界由 12 条边组成，每条边在一个方向上遍历一次，在相反方向上遍历一次。



[37.7] 边界的边界为零：$\partial^2 = 0$。

因此，如图所示，我们看到 3 维实体的 2 维边界的 1 维边界确实为零！

有关 4-3-2 维度及其他维度的证明，请参阅 Misner、Thorne 和 Wheeler (1973, §15)。正如他们所解释的，但我们不会，这也是理解微分比安基恒等式（29.17）的关键。

37.5 向量分析的经典积分定理

当然，上述 $\mathbf{d}^2 = 0$ 的几何解释取决于 FTEC，我们还没有解释为什么这是真的！ 但在我们提供 FTEC 的证明之前，我们将进一步增加悬念，以使你相信外微分基本定理确实是“基本的”！ 为此，我们现在验证我们之前的主张，即这个单一定理包含向量分析的所有经典定理作为其特例。

37.5.1 Φ=0-形式

取 $\Phi = \mathrm{f}$ 为 FTEC 中的函数（0-形式），$\mathbf{d}\Phi = \mathbf{d}\mathrm{f}$ 为 1-形式，所以必须取 R 为一维曲线 K，取其边界 $\partial\mathrm{K}$ 为其端点，比如说，a 和 b。 由于 K 指向 b 并远离 a，因此我们将到达点计为 $+\mathrm{b}$，将出发点计为 $-\mathrm{a}$。 然后 FTEC 得到先前建立的（37.4）及其经典向量等价物（37.5）：

$$\int_{\mathrm{K}} \mathbf{d}\mathrm{f} = \mathrm{f}(\mathrm{b}) - \mathrm{f}(\mathrm{a}) \iff \int_{\mathrm{K}} (\nabla \mathrm{f}) \cdot \mathbf{d}\mathrm{r} = \mathrm{f}(\mathrm{b}) - \mathrm{f}(\mathrm{a}).$$

37.5.2 Φ=1-形式

取 $\Phi = \varphi$ 为 FTEC 中的 1-形式，$\mathbf{d}\Phi = \mathbf{d}\varphi$ 为向量场 $\nabla \times \underline{\varphi}$ 的通量 2-形式，所以 R

必须取为一个有向的二维曲面 $\mathcal{S}$，其边界为 $\partial\mathcal{S}$ 必须取为其一维边界曲线。

让我们从最简单的情况开始，也是历史上的第一种情况，其中曲面 $\mathcal{S}$ 只是 (x,y) 平面的平面区域 R，正如我们上面的区域示例 [37.6]。 这里的向量场也是平面的，所以 1-形式由下式给出

$$\varphi = \varphi_{\mathrm{x}}\,\mathbf{d}\mathrm{x} + \varphi_{\mathrm{y}}\,\mathbf{d}\mathrm{y}.$$

那么对应的通量 2-形式是

$$\begin{aligned}\mathbf{d}\varphi &= \mathbf{d}\varphi_{\mathrm{x}} \wedge \mathbf{d}\mathrm{x} + \mathbf{d}\varphi_{\mathrm{y}} \wedge \mathbf{d}\mathrm{y}\\ &= \partial_{\mathrm{y}}\varphi_{\mathrm{x}}\,\mathbf{d}\mathrm{y} \wedge \mathbf{d}\mathrm{x} + \partial_{\mathrm{x}}\varphi_{\mathrm{y}}\,\mathbf{d}\mathrm{x} \wedge \mathbf{d}\mathrm{y}\\ &= (\partial_{\mathrm{x}}\varphi_{\mathrm{y}} - \partial_{\mathrm{y}}\varphi_{\mathrm{x}})\mathcal{A}。\end{aligned}$$

根据（34.10），这个 $\mathbf{d}\varphi$ 对应于 z 方向上，与 (x,y) 平面正交，速度为 $(\partial_{\mathrm{x}}\varphi_{\mathrm{y}} - \partial_{\mathrm{y}}\varphi_{\mathrm{x}})$ 的流。 根据 FTEC，它穿过 R 的通量等于 $\underline{\varphi}$ 围绕 $\partial\mathrm{R}$ 的环量。 换句话说，如果我们以经典符号开始和结束，我们会看到 FTEC 蕴含格林定理[3]（Green's Theorem）：

$$\begin{aligned}\oint_{\partial\mathrm{R}} \varphi_{\mathrm{x}}\,\mathrm{d}\mathrm{x} + \varphi_{\mathrm{y}}\,\mathrm{d}\mathrm{y} &= \oint_{\partial\mathrm{R}} \varphi\\ &= \iint_{\mathrm{R}} \mathrm{d}\varphi\\ &= \iint_{\mathrm{R}} (\partial_{\mathrm{x}}\varphi_{\mathrm{y}} - \partial_{\mathrm{y}}\varphi_{\mathrm{x}})\mathcal{A}\\ &= \iint_{\mathrm{R}} \left(\partial_{\mathrm{x}}\varphi_{\mathrm{y}} - \partial_{\mathrm{y}}\varphi_{\mathrm{x}}\right)\mathrm{d}\mathrm{x}\,\mathrm{d}\mathrm{y}.\end{aligned}$$

这是[37.6]中面积示例的通用版本。

现在转向更一般的空间中曲面 $\mathcal{S}$ 的情况，设 $\hat{\mathbf{n}}$ 为曲面的单位法线，方向的选择通过右手定则由曲面的方向确定。 然后，凭借（37.9），FTEC 得出斯托克斯定理：

$$\oint_{\partial\mathcal{S}} \underline{\varphi} \cdot \mathrm{d}\mathbf{r} = \oint_{\partial\mathcal{S}} \varphi$$

[3] 格林从未真正写下现在以他的名字命名的 2 维公式，但这确实是他发现的 3 维公式的逻辑结果。格林定理的第一个已知出版物（但没有证明）是柯西于 1846 年出版的。参见 Katz (1979)。

$$= \iint_{\mathcal{S}} \mathbf{d}\varphi$$

$$= \iint_{\mathcal{S}} (\nabla \times \underline{\varphi}) \cdot \hat{\mathbf{n}} \, \mathrm{d}\mathcal{A}.$$

这个定理在法拉第在他的实验室中观察到的宏观电磁现象与麦克斯韦最终在数学上将其简化为的微观微分方程之间提供了关键的桥梁。

以法拉第在 1831 年发现的法拉第电磁感应定律为例：

> 如果曲面 $\mathcal{S}$ 跨越一圈导线 $\partial\mathcal{S}$，并且变化磁场的磁力线穿过它，则在环中感应出一个电动势，该电动势等于通过 $\mathcal{S}$ 的磁通量变化率的负数。

以数学形式表示，

$$\oint_{\partial\mathcal{S}} \underline{\mathbf{E}} \cdot \mathrm{d}\mathbf{r} = -\partial_{\mathrm{t}} \iint_{\mathcal{S}} \underline{\mathbf{B}} \cdot \hat{\mathbf{n}} \, \mathrm{d}\mathcal{A}.$$

但随后斯托克斯定理允许我们将其重新表达为

$$\iint_{\mathcal{S}} \left[\boldsymbol{\nabla} \times \underline{\mathbf{E}} + \partial_{\mathrm{t}} \underline{\mathbf{B}}\right] \cdot \hat{\mathbf{n}} \, \mathrm{d}\mathcal{A} = 0.$$

但是，如果这对所有曲面 $\mathcal{S}$ 都成立，那么被积函数本身必须为零，因此我们发现 1831 年法拉第定律等价于 1873 年麦克斯韦方程之一：

$$\boldsymbol{\nabla} \times \underline{\mathbf{E}} + \partial_{\mathrm{t}} \underline{\mathbf{B}} = 0. \tag{37.13}$$

37.5.3 Φ=2-形式

如果我们将 $\Phi = \Psi$ 视为 FTEC 中的通量 2-形式，则凭借 (37.11)，$\mathbf{d}\Psi = (\nabla \cdot \underline{\Psi})\mathcal{V}$ 是描述源密度的 3-形式。 现在必须将 R 视为一个定向体积 V，并且必须将其边界 $\mathcal{S} = \partial\mathcal{V}$ 视为其二维边界曲面。 因此，如果我们以经典符号开始和结束，我们会看到 FTEC 蕴含高斯定理（Gauss's Theorem），也称为散度定理（Divergence Theorem）：

$$\iiint_{\mathrm{V}} (\boldsymbol{\nabla}\cdot\underline{\boldsymbol{\Psi}})\mathrm{d}\mathcal{V} = \iiint_{\mathrm{V}} \mathbf{d}\boldsymbol{\Psi} = \oiint_{\partial\mathrm{V}} \boldsymbol{\Psi} = \oiint_{\mathcal{S}} \underline{\boldsymbol{\Psi}}\cdot\hat{\mathbf{n}}\,\mathrm{d}\mathcal{A}. \tag{37.14}$$

37.6 外微分基本定理的证明

我们将证明 FTEC 的二维形式（斯托克斯定理），我们还将证明（基本不变）扩展到 3 维情况（高斯定理）。 但是我们不会发展以统一的方式将证明明确扩展到所有维度所需的概念、术语和符号。 若想了解这些，请参阅本书末尾的延伸阅读部分。

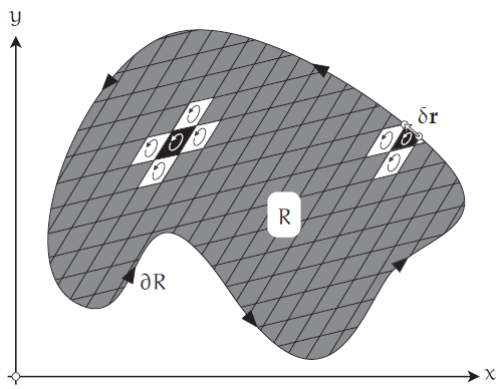
考虑 [37.8]，它显示了 (x,y) 平面的区域 R 被分成小的（最终趋于零的）平行四边形，每个平行四边形都有第一边 $\epsilon\mathbf{u}$ 和第二边 $\epsilon\mathbf{v}$。 与此排序相关的方向由每个单元格内的漩涡指示。 请注意，邻接边界 $\partial\mathrm{R}$ 的平行四边形被切割成不规则形状。

现在让我们对 R 被划分成的所有平行四边形和截断平行四边形的 1-形式 φ 的积分求和。 我们将把平行四边形和截断平行四边形统称为 R 的单元格（cells）。

考虑图示的黑色内部平行四边形，以及与之相邻的四个白色副本。 每个黑色边都与相邻单元格的相反方向的白色边匹配，因此沿这两个边的积分抵消。 唯一未被抵消的边是边界 $\partial\mathrm{R}$ 的小块，例如所示的 $\delta\mathbf{r}$。 所以，

$$\oint_{\partial\mathrm{R}} \varphi = \sum_{\text{cells in R}} \oint_{\partial(\text{cell})} \varphi \tag{37.15}$$

人们可能会天真地猜测每个单元周围的积分应该是 ϵ 阶，因为每条边都是那个阶。 但这不可能是正确的，从以下粗略的论证中可以看出。



[37.8] 围绕内部的，黑色平行四边形边的环量，都被相邻的白色平行四边形相反方向的边抵消。 对所有单元格求和，唯有边界边（例如所示的 $\delta\mathbf{r}$）的贡献未被抵消而幸存下来。

我们知道，对于一般的 1-形式，围绕 $\partial\mathrm{R}$ 的积分——因此是和（37.15）——将是非零

和有限的。 这使我们相信，每一项都必须随着对 ϵ 的相互依赖而趋于零，因为它控制着序列中项数的增长。但项的数量增长为（R 的固定面积，除以每个单元格的面积），即为 $(1/\epsilon^2)$。因此，每个项的大小预计会随着 ϵ^2 而趋于零。如果我们最初的猜测是正确的，那么（37.15）中的和的阶将是 $\epsilon(1/\epsilon^2)$，随着平行四边形的缩小，会产生无限的结果。 相反，任何对涉及 ϵ 大于 2 次幂的项的贡献都不会对最终结果产生任何影响。

事实上，没有必要诉诸这种粗略的推理，因为我们之前已经导出了关于围绕这种收缩平行四边形的 φ 积分的精确结果（37.8），这是理解和证明 FTEC 的关键。 我们看到，正如预期的那样，每个平行四边形的贡献确实是 ϵ^2 阶的，并且是通过将 2-形式 $\mathbf{d}\varphi$ 应用于平行四边形的边而给出的。 因此，（37.15）变为

$$\oint_{\partial \mathrm{R}} \varphi = \sum_{\text{cells in R}} \mathbf{d}\varphi(\epsilon\mathbf{u}, \epsilon\mathbf{v})$$

但是，在 ϵ 趋于零的极限下，这个和正是 $\mathbf{d}\varphi$ 在 R 上的积分的定义，从而完成了我们对外微分基本定理的证明：

$$\oint_{\partial \mathrm{R}} \varphi = \iint_{\mathrm{R}} \mathbf{d}\varphi$$

现在让我们推广这个论点来证明 FTEC 的斯托克斯定理形式。 为此，我们只需要假设画出平行四边形的区域 R 实际上是一个橡胶膜，跨越一根坚硬但可弯曲的线 $\partial\mathrm{R}$。 我们现在可以将这条线 $\partial\mathrm{R}$ 从 $\mathbb{R}^2$ 变形到空间中，这样 R 在 $\mathbb{R}^3$ 中呈现出一些曲面的形式。 然后，我们可以根据需要进一步变形曲面 R，同时保留边界线。

上述证明的两个关键要素是所有内部边的抵消，和（37.8）。 在这个新的 3 维上下文中，这两个都是正确的，后者采用（37.10）形式：

$$\oint_{\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v})} \varphi \asymp \mathbf{d}\varphi(\epsilon\mathbf{u}, \epsilon\mathbf{v}) = (\nabla \times \underline{\varphi}) \cdot \hat{\mathbf{n}}\, \delta\mathcal{A}.$$

至此，我们在这个例子中证明了外微分基本定理，从而在这个过程中也证明了斯托克斯定理：

$$\oint_{\partial\mathcal{S}} \underline{\varphi} \cdot \mathrm{d}\mathbf{r} = \oint_{\partial\mathcal{S}} \varphi = \iint_{\mathcal{S}} \mathbf{d}\varphi = (\nabla \times \underline{\varphi}) \cdot \hat{\mathbf{n}}\, \delta\mathcal{A}$$

最后，让我们提升一个维度，看看为什么上面的 FTEC 证明基本上没有变化，并且在这个过程中得出了对高斯定理的证明。 令 V 为 $\mathbb{R}^3$ 的紧致区域，边界曲面 $\mathcal{S} = \partial\mathrm{V}$。 现在将 V 填充为具有边 $\{\epsilon\mathbf{u}, \epsilon\mathbf{v}, \epsilon\mathbf{w}\}$ 和体积 $\mathcal{V}(\epsilon\mathbf{u}, \epsilon\mathbf{v}, \epsilon\mathbf{w})$ 的平行六面体。 曲面的那些被截断，

就像[37.8]的平行四边形发生的那样。 再次，让我们将内部平行六面体和在曲面处被截断平行六面体曲面统称为 V 的单元格。

现在考虑 2-形式 Ψ 流出内部单元格的通量。 它的每个面也是相邻单元格的面，方向相反：我们单元格的外向法线指向相邻单元格。 因此，如果我们将所有细胞的通量相加，则所有内部面的通量都会抵消：从物理上讲，从一个面流出的流体会流入它的相邻面。 因此，唯一未被抵消的面是那些边界曲面包含的部分面，因此，

$$\oiint_{\partial \mathrm{V}} \Psi = \sum_{\text{cells in V}} \oint_{\partial(\text{cell})} \Psi.$$

但是我们从（37.11）知道，从一个小的、最终趋于零的平行六面体中流出的通量最终是 ϵ^3 阶的，并且是通过将 3-形式 $\mathbf{d}\Psi$ 应用到它的边界来获得的：

$$\oiint_{\Pi(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w})} \Psi \asymp \mathbf{d}\Psi(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w}) = \left(\nabla\cdot\underline{\Psi}\right)\mathcal{V}\,(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w})$$

因此，

$$\oiint_{\partial \mathrm{V}} \Psi \asymp \sum_{\text{cells in V}} \mathbf{d}\Psi(\epsilon\mathbf{u},\epsilon\mathbf{v},\epsilon\mathbf{w})$$

最后，当 ϵ 趋于零时，右边成为 $\mathbf{d}\Psi$ 在 V 上的积分。因此，我们在这种情况下证明了外微分基本定理。

以经典向量分析的记号开始和结束，我们也在这个过程中证明了高斯定理：

$$\oiint_{\mathcal{S}} \underline{\Psi}\cdot\hat{\mathbf{n}}\,\delta\mathcal{A} = \oiint_{\partial \mathrm{V}} \Psi = \iiint_{\mathrm{V}} \mathbf{d}\Psi = \iiint_{\mathrm{V}} (\nabla\cdot\underline{\Psi})\,\mathrm{d}\mathcal{V}.$$

37.7 柯西定理

假设 [37.8] 中描述的区域 R 不在 $\mathbb{R}^2$ 中，而是在复平面中，并且复函数 f(z) 在整个区域中是解析的，因此其局部几何效果是个伸扭。 然后，正如我们在（36.13）中看到的那样，f(z) 是局部伸扭的事实等价于 1-形式 $\mathrm{f}\,\mathbf{d}\mathrm{z}$ 是闭的：$\mathbf{d}(\mathrm{f}\,\mathbf{d}\mathrm{z}) = 0$。

因此，FTEC 立即得出柯西定理（Cauchy’s Theorem）：

$$\boxed{\oint_{\partial \mathrm{R}} \mathrm{f}\,\mathbf{d}\mathrm{z} = \iint_{\mathrm{R}} \mathbf{d}(\mathrm{f}\,\mathbf{d}\mathrm{z}) = 0}$$

当 $f(z)$ 不是局部伸扭时，FTEC 也可用于获得有趣的结果。例如，在反共形映射 $f(z) = \bar{z}$ 的情况下，我们在（36.14）中看到 $\mathbf{d}(\bar{z}\,\mathbf{d}z) = 2i\mathcal{A}$，所以现在 FTEC 得到

$$\boxed{\oint_{\partial R} \bar{z}\,\mathbf{d}z = \iint_R \mathbf{d}(\bar{z}\,\mathbf{d}z) = 2i\iint_R \mathcal{A}}$$

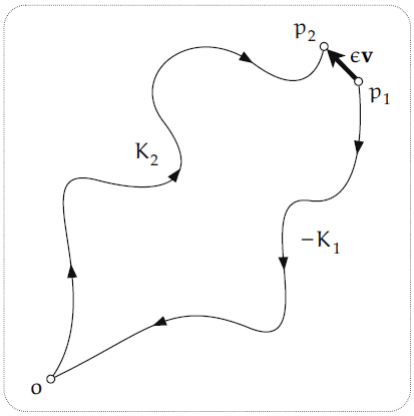
有关此结果的几何和物理证明，请参阅 VCA。

37.8 1-形式的庞加莱引理

假设 φ 在 $\mathbb{R}^n$ 的一个简单连通区域 R 中是闭的，即 $\mathbf{d}\varphi = 0$。然后，通过 FTEC，如果 L 是 R 中的一个闭环，并且我们用曲面 S 跨越它，使得 $L = \partial S$，那么

$$\oint_L \varphi = \int_S \mathbf{d}\varphi = 0$$

这意味着，凭借（37.3），$\int \varphi$ 是路径无关的。



[37.9] 如果 $\mathbf{d}\varphi = 0$，则 $f(p) \equiv \int_o^p \varphi$ 是路径无关的，并且 $\varphi = \mathbf{d}f$。

现在让我们选择一个任意固定点 o 并将其视为我们的原点。将 $\underline{\varphi}$ 视为力场，我们可以将势能函数 $f(p)$ 定义为将粒子从 o 带到 p 时所做的与路径无关的功：

$$\boxed{f(p) \equiv \int_o^p \varphi}$$

我们现在将证明 φ 是恰当的，并且由下式给出

$$\boxed{\varphi = \mathbf{d}f,} \tag{37.16}$$

从而在 1-形式的情况下证明了庞家莱引理。

令 p_1 和 p_2 是两个相邻点，如 [37.9] 所示，并让它们之间的短（最终趋于零）连接向量为 $\epsilon\mathbf{v} = \overrightarrow{p_1p_2}$。 则两点之间的能量差 δf 为

$$\delta f = f(p_2) - f(p_2) \asymp \mathbf{d}f(\epsilon\mathbf{v}).$$

但是，如 [37.9] 所示，如果 K_1 是从 o 到 p_1 的路径，而 K_2 是从 o 到 p_2 的路径，则 $-K_1 + K_2$ 是从 p_1 到 p_2 的路径。 并且由于 $\int\varphi$ 是路径无关的，我们可以用沿着 $\epsilon\mathbf{v}$ 的短的、直接的路径来代替这条长的、间接的路径。 所以，

$$\mathbf{d}f(\epsilon\mathbf{v}) \asymp \int_{-K_1+K_2} \varphi = \int_{\epsilon V} \varphi \asymp \varphi(\epsilon\mathbf{v}).$$

由于 $\mathbf{v}$ 是任意的，我们可以将它抽象出来，完成（37.16）的证明，并用它完成庞家莱引理（36.11）对 1-形式的证明。

37.9 德拉姆上同调入门

37.9.1引言

如果区域 R 不是简单连通的怎么办？ 那么事实证明，闭形式不一定是恰当的，对闭的但不恰当的形式的研究会得到有关 R 拓扑的详细信息。

这些信息被编码在所谓的德拉姆上同调群（de Rham cohomology groups） $H^k(R)$ 中，以乔治·德拉姆（Georges de Rham，1903-1990）的名字命名，他在 1931 年发现了它们。这里的 k 指的是在 R 上定义的 k-形式，它们是闭的，但不恰当。

我们将仅给出该主题的一小部分，从闭的但不恰当的 1-形式开始，得到第一个德拉姆上同调群 $H^1(R)$。

37.9.2一种特殊的二维涡旋向量场

在 $\mathbb{R}^2$ 中，使用极坐标 (r,θ)，考虑以原点为中心的特殊圆形逆时针涡旋[4]，如 [37.10] 所示，其中流动速度选择为 $|\underline{\varphi}| = (q/2\pi r)$ ，其中 q 是衡量涡旋强度的常数。 这一特殊选择的原因将在稍后揭晓。

当我们接近原点时，速度变为无穷大，所以原点是一个奇点，我们排除这一点：流定义在穿孔平面（punctured plane） $\mathbb{R}^2 - 原点$ 中。

由于沿半径为 r 的圆周 K_r 的速度为 $(q/2\pi r)$，且 K_r 的长度为 $2\pi r$，因此围绕它的环量（circulation）为

$$\boxed{\mathcal{C} = \oint_{K_r} \varphi = q,}$$

与半径 r 无关。 事实上，正是为了使环量与半径无关，我们需要选择与半径成反比的流速。

$\mathcal{C} \neq 0$这一事实意味着

> 涡旋的 1-形式 φ 不可能是恰当的：对于任何 f，都有 $\varphi \neq \mathbf{d}f$。

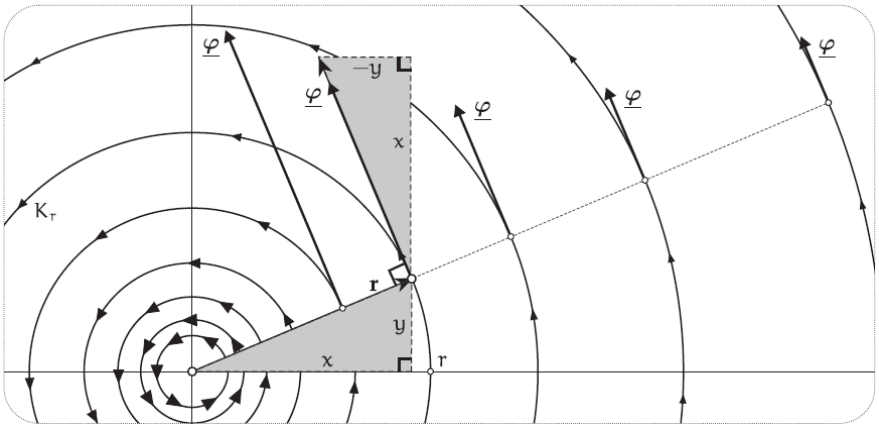
然而，稍后我们将证明

> 旋涡的 1−形式 φ 是闭的：$\mathbf{d}\varphi = 0$。

这种闭的但不恰当的 1-形式的存在性是因为平面的穿孔而成为可能的。 在原点环绕涡旋奇点的环路不能在不越过原点的情况下收缩到一个点，因此穿孔平面不是简单连通的，因此与庞加莱引理不冲突。

---

[4]在这里，我们更喜欢更令人回味的词涡旋（vortex），而不是技术上更精确的术语中心（centre）。 参见 [19.4]，第 200 页。



[37.10] 一个特殊的圆形涡旋向量场 $\underline{\varphi}$，速度为 $|\underline{\varphi}| = (q/2\pi r)$。 注意当我们远离原点时速度下降。 $\underline{\varphi}$ 的方向是通过将半径向量 r 旋转一个直角获得的，但由于其长度为 $r = |r|$，我们必须将其乘以 $(q/2\pi r^2)$。 因此涡旋的 1-形式是 $\varphi = (q/2\pi r^2)[-y\,\mathbf{d}x + x\,\mathbf{d}y]$。 原点是一个速度无限的奇点，必须排除。

37.9.3涡旋 1-形式是闭的

图 [37.10] 画出了一个典型的长度为 r 的半径向量 $\mathbf{r}$，然后将其旋转 $(\pi/2)$ 以获得涡流方向向量。 但是这个向量的长度是 r，而我们希望构建一个速度为 $(q/2\pi r)$ 的涡旋，所以我们必须将这个向量乘以 $(q/2\pi r^2)$。

表示为 1-形式，因此涡旋向量场为

$$\boxed{\varphi = \left(\frac{q}{2\pi r^2}\right)[-y\,\mathbf{d}x + x\,\mathbf{d}y],} \tag{37.17}$$

其中 $r^2 = x^2 + y^2$。

所以，

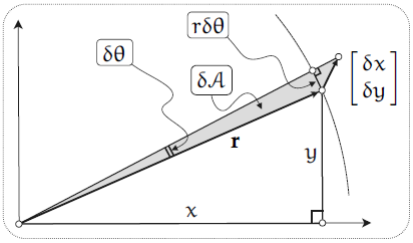
$$\begin{aligned}
\left[\frac{2\pi}{q}\right]\mathbf{d}\varphi &= -\frac{1}{r^4}\,(\mathbf{d}r^2)\wedge[-y\,\mathbf{d}x + x\,\mathbf{d}y] + 2\frac{1}{r^2}\,\mathbf{d}x\wedge\mathbf{d}y \\
&= -\frac{1}{r^4}\,2\,(x\,\mathbf{d}x + y\,\mathbf{d}y)\wedge[-y\,\mathbf{d}x + x\,\mathbf{d}y] + 2\frac{1}{r^2}\,\mathbf{d}x\wedge\mathbf{d}y \\
&= -\frac{1}{r^4}\,2\,(x^2 + y^2)\,\mathbf{d}x\wedge\mathbf{d}y + 2\frac{1}{r^2}\,\mathbf{d}x\wedge\mathbf{d}y \\
&= 0,
\end{aligned}$$

如声称的那样。

37.9.4涡旋 1-形式的几何意义

我们提出了上述计算，以便为外微分机制提供有用的实践。 然而，这个计算实际上是不必要的，因为可以给出一个简单的几何原因（为什么 φ 是闭的）。

考虑 [37.11]，它画出了一个短的、最终趋于零的向量 $\begin{bmatrix}\delta x\\ \delta y\end{bmatrix}$ 从 $\mathbf{r}=\begin{bmatrix}x\\ y\end{bmatrix}$ 的尖端发出，在原点处对向一个角度 $\delta\theta$。 设有这两条边的三角形的阴影面积为 $\delta\mathcal{A}$。



[37.11] 阴影三角形的面积 $\delta\mathcal{A}$ 可以通过其边的行列式进行求值，得出 $\delta A=\frac{1}{2}(-y\delta x+x\delta y)$，或者从图中几何得出 $\delta A \asymp \frac{1}{2}r^2\,\delta\theta$。 将这些表达式等同起来，我们发现涡旋 1-形式是 $\varphi=\frac{q}{2\pi}\,\mathbf{d}\theta$。

一方面，我们可以将 $\delta\mathcal{A}$ 计算为三角形边行列式的一半，另一方面，它最终等于底为 r，高为 $r\delta\theta$ 的三角形面积，所以，

$$\begin{aligned}\frac{1}{2}(-y\,\delta x+x\,\delta y)&=\frac{1}{2}\det\begin{bmatrix}x & \delta x\\ y & \delta y\end{bmatrix}\\ &=\delta A \asymp \frac{1}{2}r(r\,\delta\theta).\end{aligned}$$

因此，涡旋 1-形式（37.17）具有非常简单的几何含义：

$$\boxed{\varphi=\frac{q}{2\pi}\,\mathbf{d}\theta,}$$

清楚地说明为什么 φ 是闭的：$\mathbf{d}\varphi=\frac{q}{2\pi}\,\mathbf{d}^2\theta=0$。

但是等等——这个例子的重点不就是 φ 不恰当的吗？ 可爱的小牛顿，这里发生了什么？！

这个表面上的悖论的解决方法有些微妙。 尽管 $\mathbf{d}\theta$ 具有完美定义的几何意义，但它不能解释为（$\mathbf{d}$ 作用于 θ），因为 θ 甚至不是函数！ 也就是说，任何给定的点都必须分配无限多个角度。

如果我们将注意力限制在一个不包含原点处奇点的简单连通区域 S，例如 [37.12] 中所示的区域，那么我们可以（非唯一地）在 S 上定义一个真正的单值角函数 θ，$\theta_1 \leq \theta \leq \theta_2$，如图所示。 现在庞家莱引理适用，所以闭的 1-形式 φ 必须是恰当的，并且我们已经知道了明确的公式：$\varphi = \frac{q}{2\pi}\,\mathbf{d}\theta$。

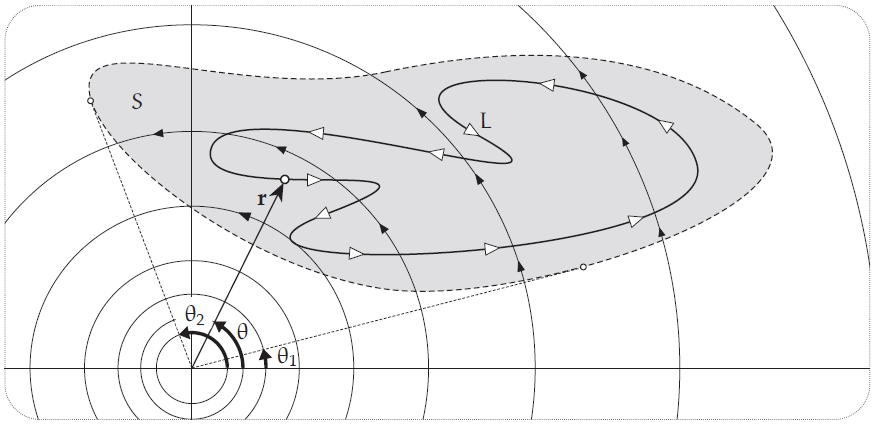
此外，S 中包含的围绕任何环路 L 的涡旋环量确实趋于零了。 因为，当图示的粒子穿过 L 时，位置向量 r 来回摆动，但当它返回到其起点时，其角度 θ 的净变化为零，因此 $\mathcal{C} = \oint_L \frac{q}{2\pi}\,\mathbf{d}\theta = 0$。

但是如果我们试图在整个穿孔平面上这样做，那么我们马上就会遇到麻烦。 是的，我们可以尝试通过坚持 $0 \leq \theta < 2\pi$ 来定义 θ，但是 θ 甚至不是连续的，更不用说可微了。因此，正是在其上定义 φ 的区域的全局拓扑阻止了它在整个区域中的恰当性。

37.9.5闭 1-形式环量的拓扑稳定性

我们之前看到，我们的特殊涡旋的环量 $\mathcal{C} = \oint_{K_r} \varphi$ 与圆周 K_r 的大小无关，而对 φ 的新几何解释清楚地说明了原因：当我们沿任何 K_r 遍历时，角度增加了 2π，因此 $\int_0^{2\pi} \frac{q}{2\pi} d\theta = q$。

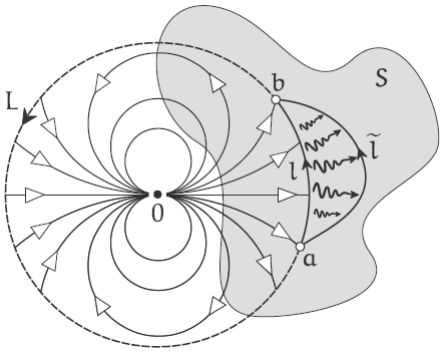
事实上，这种几何解释清楚地表明，$\mathcal{C}$ 的大小无关实际上是一个更普遍现象的非常特殊的情况：如果我们逐渐将这样的圆变形为任何形状，但要小心在它演化时不要穿过原点处的奇点，那么积分的值不会改变！



[37.12] 简单连通区域 S 不包含原点，因此我们可以在其上定义一个单值角函数θ，其中 $\theta_1 \le \theta \le \theta_2$。 那么 $\varphi = \frac{q}{2\pi}\mathbf{d}\theta$ 在 S 上是恰当的。当 $\mathbf{r}$ 沿环路 L 环绕运动时，它来回摆动，但是当它返回到其起点时，角度 θ 的净变化为零，因此 $\mathcal{C} = \oint_L \frac{q}{2\pi}\mathbf{d}\theta = 0$ 。

换句话说，

所有包围涡旋原点处奇点的简单环路都具有相同的环量，$\mathcal{C} = q$。另一方面，任何不包含奇点的环路环量为零：$\mathcal{C} = 0$。	(37.18)



[37.13] 简单连通区域 S 不包含闭的偶极子场 $\mathcal{X}$ 的奇点。 庞加莱引理告诉我们 $\oint_L \mathcal{X}$ 在 l 到 $\tilde{l}$ 的形变下是不变的。

请注意，这两种环路之间的区别纯粹是拓扑的。

注意：如果一个环路绕原点 m 次，则说它的环绕数（winding number）为 m，在这种情况下 $\mathcal{C} = mq$。 有关环绕数概念及其许多应用的更多信息，请参阅 VCA。

事实上，积分的拓扑稳定性，如（37.18）所示，适用于所有闭的 1-形式，而不仅仅是我们特定的涡旋。 为了理解这个更广泛的结果，考虑 [37.13] 中所示的偶极场 $\mathcal{X}$，假设是闭的：$\mathbf{d}\mathcal{X} = 0$。 这个特定的场没有特殊意义； 我们只是想明确一点，我们可以将以下推理应用于任何闭的 1-形式。

现在令 S 是不包含原点处奇点的任何简单连通区域，例如 [37.13] 中所示的阴影区域。由于 $\mathcal{X}$ 在 S 内是闭的，庞加莱引理暗示 $\mathcal{X} = \mathbf{d}\mathrm{f}$ 在 S 内是恰当的，因此它的积分在 S 内是路径无关的。

现在，如图所示，考虑一个以原点为中心的圆 L，它的一部分包含在 S 中。如果我们采用图示的弧 l（连接 a 到 b）并将其变形为替代路线 $\tilde{l}$，那么路径独立性告诉我们

$$\int_{\tilde{l}} \mathcal{X} = \int_{l} \mathcal{X} = \mathrm{f(b)} - \mathrm{f(a)}.$$

因此，如果 $\tilde{\mathrm{L}}$ 表示新的变形环，其中 l 已变形为 $\tilde{l}$，则

$$\boxed{\oint_{\tilde{\mathrm{L}}} \mathcal{X} = \oint_{\mathrm{L}} \mathcal{X}}$$

直观上很清楚（我们在 VCA 中证明了这一点），我们现在可以继续将 L 变形为围绕奇点的任何简单环路，并且由于环量在整个变形过程中不会改变，我们已经证明了（37.18）的广义版本：

闭的 1−形式的变形定理。 闭 1−形式沿环路的环量（称为（德拉姆）周期）在环路的连续形变下保持不变的。	(37.19)

鉴于环量与环绕奇点的环路的特定形状无关，我们有权将其视为奇点本身的属性。 这是复解析函数 f(z) 在其每个奇点处的复分析概念留数（residue）的推广。 事实上，在 VCA（第 11 章）中，我们明确地根据波利亚向量场 $\overline{\mathrm{f(z)}}$ 的环量和通量来描述留数。

37.9.6第一德拉姆上同调群

现在让我们回到第一德拉姆上同调群的定义，对于当前情况，$\mathrm{R} = (\mathbb{R}^2 - \text{原点})$，写作 $\mathrm{H}^1(\mathbb{R}^2 - \text{原点})$。 这个群的一个元素不是单个的 1-形式，而是 1-形式的等价类，其中

> 如果两个 1−形式围绕所有环路具有相同的环量，则它们根本上是相同的，即“等价的”。

（注意：技术术语（我们不会使用）是说两个 1-形式是上同调的（cohomologous）。）

例如，假设我们在涡旋 1-形式 φ 中加上一个常数 1-形式，对应于某个方向上的稳定

流，它本身显然没有围绕任何环路的环量。 那么新流的环量是旧流的环量和稳定流的环量（为零）的之和。 因此，围绕所有环路的环量都没有改变，因此这个新的 1-形式与 φ 属于相同的等价类。

更一般地说，我们可以加上任何围绕所有环路的环量都为零的流。 换句话说，我们可以加上任何恰当的 1-形式。总而言之，

> 令 $\tilde{\varphi} \sim \varphi$ 表示 1−形式 $\tilde{\varphi}$ 和 φ 的等价性。 那么，
> $\tilde{\varphi} \sim \varphi \iff$ 对于某些 f，$\tilde{\varphi} - \varphi = \mathbf{d}\mathrm{f}$。

每个等价类由其环量值 $\mathcal{C}$ 定义。（注意：在德拉姆上同调的标准课本中，环量 $\mathcal{C}$ 称为等价类的周期（period），而类本身称为上同调类（cohomology class）。）

术语警告：从历史上看，在这种情况下使用术语周期（period）源于椭圆积分（elliptic integrals）理论。 但是在物理学中，我们习惯于用周期来指代钟摆来回摆动所需的时间，或者行星绕太阳一周所需的时间。 因此，在查看 [37.10] 时，我们应该警告你，德拉姆周期与流体中一个粒子围绕奇点绕一个轨道运行一周所需的时间无关。 实际上，绕 $\mathrm{K_r}$ 一周所需的时间 $\mathrm{T_r}$ 取决于 r，由 $\mathrm{T_r} = 4\pi^2(\mathrm{r}^2/\mathrm{q})$ 给出，而固定的德拉姆周期为 $\mathcal{C} = \mathrm{q}$。

上述所有推理仅取决于 FTEC，而不取决于我们用来说明这一思想的特定形式的涡旋。因此，这一切都可推广到更一般的，闭的但不恰当的， 1-形式 φ。

如果我们将两个不同等价类中的两个这样的 1-形式 φ_1 和 φ_2 相加，那么我们得到有如下周期的类中的一个新的 1-形式

$$\mathcal{C}(\varphi_1 + \varphi_2) = \mathcal{C}(\varphi_1) + \mathcal{C}(\varphi_2).$$

因此，群 $\mathrm{H}^1(\mathbb{R}^2 - \text{原点})$ 在加法下同构于实的周期集合，我们可以简单地写成

$$\boxed{\mathrm{H}^1(\mathbb{R}^2 - \text{原点}) = \mathbb{R}.}$$

接下来，假设我们提升一个维度并考虑 $\mathbb{R}^3$ 中的 1-形式。 如果我们从 $\mathbb{R}^3$ 中删除原点，那么我们仍可将每个环路收缩到一个点，因为现在我们有足够的空间在原点周围滑动一个收缩的环路，而不必穿过它——所以现在空间再次简单地连通起来。 因此不存在不恰当的闭的 1-形式，我们将这一事实记录为

$$\boxed{\mathrm{H}^1(\mathbb{R}^3 - \text{原点}) = 0。} \tag{37.20}$$

另一方面，如果我们从 $\mathbb{R}^3$ 中移除（例如）整个 z 轴，那么我们就有了与穿孔的 $\mathbb{R}^2$ 类似的情况：它不是简单连通的。 事实上，为了加深类比，想象流 [37.10] 向上延伸出页面，产生一个围绕 z 轴旋转的 3 维涡流，流体在以奇异的 z 轴为中心的同心圆柱中旋转，其中 z 轴处速度达到无穷大。 确实，与穿孔平面的类比是完美的：

$$\boxed{H^1(\mathbb{R}^3 - \text{z 轴}) = \mathbb{R}。}$$

37.9.7 $\mathbb{R}^3$ 中的平方反比点源

作为我们讨论第二德拉姆上同调群 $H^2(R)$ 的前奏，我们现在定义一个在 $\mathbb{R}^3$ 的原点具有奇点的特殊向量场。 设 $\hat{\mathbf{r}}$ 为单位径向向量：

$$\hat{\mathbf{r}} = \frac{1}{r}\mathbf{r} = \frac{1}{r}\begin{bmatrix} x \\ y \\ z \end{bmatrix}, \qquad \text{其中} \quad r = \sqrt{x^2 + y^2 + z^2}$$

我们现在向你介绍可能是所有物理学中最重要的向量场，即*平方反比点源*（inverse-square point source）：

$$\boxed{\underline{\Psi} = \frac{q}{4\pi}\frac{\hat{\mathbf{r}}}{r^2},}$$

其中 q 是决定场强的常数。 由于该场在原点具有奇点，因此必须排除该点：该场在 $\mathbb{R}^3 - \text{原点}$ 上定义。

这个场有（至少！）三个非常重要的，物理上不同的解释：

- 如果流体以速率 q 被泵入 0，并径向对称地向外流动，则 $\underline{\Psi}$ 是流体流动的速度向量。
- 如果一个点电荷 q 位于0，库仑定律说 $\underline{\Psi}$ 是它的电场。
- 如果我们颠倒场的方向，那么牛顿万有引力定律说 $-\underline{\Psi}$ 是位于 0 处质点 q 的引力场。

正如牛顿在《原理》中所证明的，这个场也代表（非常显著！）任何球对称质量 q 的物体（例如地球或太阳）的外部引力场。

在 $\mathbb{R}^3$ 中，我们可以将任何向量场解释为 1-形式或 2-形式。 下一节我们的重点将放在 2-形式上，但首先让我们讨论对应于向量场 $\underline{\Psi}$ 的 1-形式 ψ：

$$\boxed{\psi = \frac{q}{4\pi r^3}[x\,\mathbf{d}x + y\,\mathbf{d}y + z\,\mathbf{d}z].}$$

让我们从第三种物理解释的角度来看待这一点：万有引力。 那么，与其将沿开放路径 J 的线积分视为代表环量，不如将其视为代表功：

$$W_J = \text{work} = \int_J \psi.$$

我们知道万有引力是保守的：在这个万有引力场内围绕一个环路携带一个质量运动净做功必然为零，否则我们可以构造一个永动机！ 因此，我们预计 ψ 是闭的。 我们把它作为一个练习，通过直接计算来证实 $\mathbf{d}\psi = 0$。

然而，我们可以通过证明 ψ 是恰当的来证明这一点：$\psi = \mathbf{d}f$。 更明确地说，让我们证明势能是

$$\boxed{f = \text{势能} = -\frac{q}{4\pi r}}$$

作为引理，首先观察到[练习]

$$\mathbf{d}r = \frac{1}{r}(x\,\mathbf{d}x + y\,\mathbf{d}y + z\,\mathbf{d}z).$$

所以，

$$\begin{aligned}\mathbf{d}f &= -\frac{q}{4\pi}\mathbf{d}\left[\frac{1}{r}\right]\\ &= \frac{q}{4\pi r^2}\,\mathbf{d}r\\ &= \frac{1}{4\pi r^3}(x\,\mathbf{d}x + y\,\mathbf{d}y + z\,\mathbf{d}z)\\ &= \psi.\end{aligned}$$

因此，如果 J 是连接 a 到 b 的任何路径，则

$$W_J = \int_J \psi = \int_J \mathbf{d}f = f(b) - f(a) = \frac{q}{4\pi}\left[\frac{1}{|a|} - \frac{1}{|b|}\right].$$

请注意，与我们之前的涡旋示例不同，这里的势 f 是一个在整个($\mathbb{R}^3$ – 原点)明确定义的函数。 因此，该结果与（37.20）宣称的一致，即第一德拉姆上同调群在这种情况下为零。

37.9.8第二德拉姆上同调群

接下来，让我们计算第二德拉姆上同调群 $H^2(R)$ 的具体示例，它衡量 R 上闭的但不恰当的 2-形式。让我们留在 $\mathbb{R}^3$ 中，并考虑从闭合曲面 $\mathcal{S}$ 穿出的闭的 2-形式 Ψ 的通量。

让我们继续考虑定义在 $R = (\mathbb{R}^3 - 原点)$ 上的平方反比点源，但现在让我们将其转换为流体流动，在这种情况下，通过（34.10）和（34.12）可将该场很自然地表示为通量 2-形式：

$$\boxed{\Psi = \frac{q}{4\pi}\frac{1}{r^3}(x\,\mathbf{d}y \wedge \mathbf{d}z + y\,\mathbf{d}z \wedge \mathbf{d}x + z\,\mathbf{d}x \wedge \mathbf{d}y).}$$

通过查看以原点为中心半径为 r 的球面 $\mathcal{S}_r$ 的通量，可以看出该场不是恰当的。 球面面积为 $4\pi r^2$，场与它正交，所以通量 Ω 只不过是面积和场强的乘积：

$$\boxed{\Omega = \oiint \Psi = q,}$$

与 r 无关。

尽管 Ψ 不是恰当的，我们现在将证明它是闭的：

$$\begin{aligned}
\left[\frac{4\pi}{q}\right]\mathbf{d}\Psi &= \frac{3}{r^3}\mathcal{V} - \frac{3}{r^4}\,\mathbf{d}r \wedge (x\,\mathbf{d}y \wedge \mathbf{d}z + y\,\mathbf{d}z \wedge \mathbf{d}x + z\,\mathbf{d}x \wedge \mathbf{d}y) \\
&= \frac{3}{r^3}\left[\mathcal{V} - \frac{1}{r^2}(x\,\mathbf{d}x + y\,\mathbf{d}y + z\,\mathbf{d}z) \wedge (x\,\mathbf{d}y \wedge \mathbf{d}z + y\,\mathbf{d}z \wedge \mathbf{d}x + z\,\mathbf{d}x \wedge \mathbf{d}y)\right] \\
&= \frac{3}{r^3}\left[\mathcal{V} - \frac{1}{r^2}\,(x^2 + y^2 + z^2)\mathcal{V}\right] \\
&= 0.
\end{aligned}$$

接下来，让我们证明（37.21）推广为：

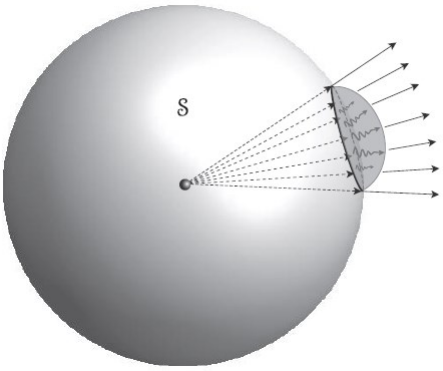
闭 2−形式的变形定理。 闭合 2−曲面的闭的 2−形式的通量（周期）在曲面的连续形变下是不变的。

(37.21)

图 [37.14] 显示流体从原点处的平方反比点源径向向外流过球面 $\mathcal{S}$。 $\mathcal{S}$ 的圆形曲面片向外形变（由波浪形箭头表示）以在 $\mathcal{S}$ 的表面上产生图示的凸起或泡。

所有流入水泡底部的流体也流出它的外表面，所以如果 $\tilde{\mathcal{S}}$ 表示新的变形曲面，那么，

$$\oiint_{\tilde{\mathcal{S}}} \Psi = \oiint_{\mathcal{S}} \Psi$$



[37.14] 因为 2-形式是闭的，流入泡的（流体）必然流出。 所以从 $\mathcal{S}$ 流出的通量在形变下保持不变。

这个论证很容易推广以适用于任何闭的 2-形式 Ψ。由于泡内的 $\mathbf{d}\Psi = 0$，FTEC 意味着来自泡的净外向通量为零。但这可以像我们刚才所做的那样重新表述：流入泡的（流体）必然流出。因此我们证明了（37.21）：$\mathcal{S}$ 的通量在 $\mathcal{S}$（不会迫使它通过 Ψ 之奇点）的形变下保持不变。

回到具体的平方反比点源，包围原点处奇点的封闭曲面无法在不与原点相交的情况下收缩到一点，因此所有包围奇点的曲面都具有相同的通量 q。这种拓扑稳定的通量衡量原点处源的强度，具有上面列出的三种物理上不同的解释。 （我们再次重申，关于德拉姆上同调理论的更高级的课本将这个源的强度称为周期（period）。）

2-形式的等价类的定义完全类似于 1-形式；也就是说，如果两个 2-形式 $\tilde{\Psi}$ 和 Ψ 从每个闭合曲面流出的通量相同，则它们是等价的。用更传统的术语表示，

$$\tilde{\Psi} \sim \Psi \iff \text{对于某些 1-形式 } \varphi \text{ 有：} \tilde{\Psi} - \Psi = \mathbf{d}\varphi。$$

如果我们将两个 2-形式相加，那么它们的通量也会增加。 所以，

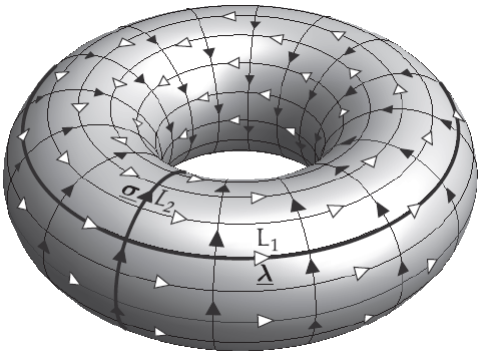
$$\mathrm{H}^2\left(\mathbb{R}^3 - \text{原点}\right) = \mathbb{R}.$$

37.9.9环面的第一德拉姆上同调群

如果在 $\mathbb{R}^n$ 以外的流形（例如亏格 g 的闭合 2-曲面）上定义非恰当且闭的形式会怎样？

考虑 [37.15] 中所示的环面。 在这里，我们看到曲面上有两种不同的可能的流：白色流 $\underline{\lambda}$ 围绕环面的对称轴，沿赤道方向，而黑色流 $\underline{\sigma}$ 穿过洞。 我们将假设两个流都是闭的：

$$\boxed{\mathbf{d}\lambda = 0 = \mathbf{d}\sigma。}$$



[37.15] 围绕环面赤道的环路 L_1 和通过环面洞的环路 L_2 在拓扑上是不同的：我们不能将一个变形为另一个。 因此，环面上闭的 1-形式流的等价类需要指定两个独立的周期。

现在令 L_1 为白色流 $\underline{\lambda}$ 的闭合环路流线，令 L_2 为黑色流 $\underline{\sigma}$ 的闭合环路流线，如图所示。 正如我们在 [19.9] 中所讨论的，L_1 和 L_2 在拓扑上是不同的：我们不能连续地将一个变形为另一个。 然而，任何围绕对称轴一圈（通过洞 0 次）的环路在拓扑上等价于 L_1，任何围绕洞一圈（绕对称轴 0 圈）的环路在拓扑上等价于 L_2 .

接下来，观察到这两个流都不是恰当的。 由于白色向量场 $\underline{\lambda_1}$ 直接沿 L_1 流动，它沿它有一个非零环量。 让我们将此称为周期 $\omega_1(\lambda)$，下标标识了正在考虑哪些拓扑不同的环路：

$$\omega_1(\lambda) = \oint_{L_1} \lambda \neq 0$$

然后，变形定理（37.19）告诉我们，这个周期对于我们可以将 L_1 变形到的任何环路都是相同的。 也就是说，周期 $\omega_1(\lambda)$ 是白色向量场的拓扑性质，与任何绕对称轴绕一圈（并且穿过洞 0 次）的环路相关联。

另一方面，$\underline{\lambda}$ 处处与 L_2 正交，因此沿 L_2 的环量为零：

$$\omega_2(\lambda) = \oint_{L_2} \lambda = 0$$

而且，根据变形定理，任何穿过洞一次（绕对称轴零次）的环路都是如此。

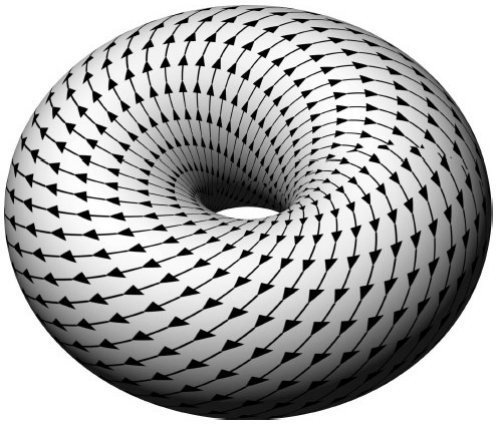
完全对称的分析适用于通过洞的黑色向量场 $\underline{\sigma}$：

$$\omega_1(\sigma) = \oint_{L_1} \sigma = 0, \quad 和 \quad \omega_2(\sigma) = \oint_{L_2} \sigma \neq 0,$$

同样，对于可能将 L_1 和 L_2 变形成为的任何环路，这些都是正确的。

现在考虑通过对 $\underline{\lambda}$ 和 $\underline{\sigma}$ 进行线性组合获得的，环面上更一般的向量场 $\underline{\phi}$：

$$\phi = a\lambda + b\sigma, \tag{37.22}$$



[37.16] 环面上闭的 1-形式 ϕ 的等价类由两个独立的拓扑周期决定：围绕对称轴的环量 $\omega_1(\phi)$ 和通过洞的环量 $\omega_2(\phi)$。 由 RokerHRO 提供，由 SA 3.0 提供。

其中 a 和 b 是常数。 如[37.16]所示，这个流既穿过洞又围绕对称轴。 但它仍然是闭的，因为

$$\mathbf{d}\phi = a\,\mathbf{d}\lambda + b\,\mathbf{d}\sigma = 0,$$

因此它的周期仍然是拓扑定义的。

但与构建它的任何一个场不同，$\underline{\phi}$ 有两个不为零的独立周期：

$$\omega_1(\phi) = \mathrm{a}\,\omega_1(\lambda) \neq 0,$$

$$\omega_2(\phi) = \mathrm{b}\,\omega_2(\sigma) \neq 0.$$

这种通过两个独立周期对 ϕ 的表征也适用于环面上完全一般的闭的 1-形式，即便 ϕ 碰巧没有特殊形式（37.22）。 总而言之，

> 环面上闭的 1−形式 ϕ 的等价类由两个拓扑定义的周期确定：$\omega_1(\phi)$ 衡量围绕对称轴的环量，而 $\omega_2(\phi)$ 衡量穿过洞的环量。

因此，我们可以用二维向量巧妙地标识 ϕ 的等价类，我们将其称为周期向量（period vector） $\omega(\phi)$，由下式给出

$$\omega(\phi) \equiv \begin{bmatrix} \omega_1(\phi) \\ \omega_2(\phi) \end{bmatrix}$$

最后，观察到如果我们将来自两个不同等价类的两个流相加，它们的周期向量相加，如下所示：

$$\omega(\tilde{\phi} + \phi) = \begin{bmatrix} \omega_1(\tilde{\phi} + \phi) \\ \omega_2(\tilde{\phi} + \phi) \end{bmatrix} = \begin{bmatrix} \omega_1(\tilde{\phi}) \\ \omega_2(\tilde{\phi}) \end{bmatrix} + \begin{bmatrix} \omega_1(\phi) \\ \omega_2(\phi) \end{bmatrix} = \omega(\tilde{\phi}) + \omega(\phi)$$

因此，我们找到了环面的第一德拉姆上同调群：

$$\boxed{\mathrm{H}^1(\text{环面}) = \mathbb{R}^2.}$$

整本书都致力于德拉姆的理论，但我们必须在这里画上句号，以免这本书成为其中之一！我们瞥见了一座巨大而美丽的冰山一角，我们必须不情愿地转身。 但是以延伸阅读部分为指导，你应该可以随意地深入探索冰面之下的美景。

第三十八章 途经形式的微分几何

38.1 引言：嘉当的活动标架法

终于，我们终于准备好让“魔鬼机器”[1]为几何学服务了！

在这最后一幕的最后一章中，我们将把嘉当形式的机制应用于微分几何，就像嘉当自己所做的那样，通过计算来验证许多在前四幕中被几何证明的基本结果。 此外，我们将看到形式自然地引导我们获得许多新的结果。 我们首先概述我们的作战计划及其历史渊源。

回顾一下 1847 年的弗勒内（Frenet）和 1851 年的（独立的）塞雷（Serret），他们都通过将正交标架$(\mathbf{T},\mathbf{N},\mathbf{B})$ 附加到曲线来分析 $\mathbb{R}^3$ 中的曲线，发现其沿曲线的变化率由弗勒内-塞雷公式（Frenet-Serret）方程 （9.3）给出，我们在此重复：

$$\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix}' = \begin{bmatrix}0 & \kappa & 0\\ -\kappa & 0 & \tau\\ 0 & -\tau & 0\end{bmatrix}\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix} = [\Omega]\begin{bmatrix}\mathbf{T}\\ \mathbf{N}\\ \mathbf{B}\end{bmatrix} \tag{38.1}$$

因此，当这个标架沿着曲线移动时，它的旋转速率被完全编码为矩阵 $[\Omega]$，它是斜对称的（即，$[\Omega]^{\mathrm{T}} = -[\Omega]$）。

该事业的成功取决于两个想法：

想法 1：以几何上有意义的方式使标架适应曲线，称为弗勒内（Frenet）标架：$\mathbf{T}$ 与曲线相切，$\mathbf{N}$ 是*主法线*，位于运动的瞬时平面内，并指向瞬时曲率中心。然后 $\mathbf{B}$ 由右手定则确定。

想法 2：用标架本身来表达标架的变化率。

大约在 1880 年，让-加斯顿 • 达布（Jean-Gaston Darboux）迈出了下一步，将这些想法扩展到曲面：

想法 3：通过选择一个向量作为曲面的*法线*，使标架适应*曲面*； 剩下的两个向量会自动成为曲面的切向量。

从 1900 年左右开始，埃利 • 嘉当开发了形式并将它们应用于微分几何，他的努力在嘉当（1928 年）已基本完成，尽管他的想法的最终完美可能是嘉当（1945 年）：

想法 4：完全释放标架，并允许它在整个$\mathbb{R}^3$ 中以*任何（可微的）方式变化*。 我们现在不再称其为弗勒内标架，而是将其称为*嘉当的活动标架场*（Cartan’s moving frame field），将

[1] 见前言。

这种方法作为一个整体称为*嘉当的活动标架方法*（Cartan's method of moving frames）。 正如我们将很快看到的，这再次产生了一个斜对称矩阵，但它现在是一个 1-形式的矩阵。

想法 5：将活动标架替换为它的 1-*形式的对偶标架*，观察它在空间中任意方向的变化率。推广想法 2，用对偶基*本身*来表达对偶基的这种变化率。

让我们停下来评论一下想法 4，这似乎是一个命运多舛的*倒退*。 弗勒内、塞雷和达布都以几何上有意义的方式调整了他们的标架以适应几何对象，因此他们对其变化率的检查将产生关于该对象的几何上有意义的信息是有道理的。

但是（恶魔般的）*嘉当*反而指示我们做以下事情。拿起你的右手，将你的拇指（想象一下被撒旦拉长！）、食指和中指摆成正交标架。现在在空间中挥动你的手，一边走一边转动你的手！[2]

即使转瞬即逝，你也已经至少在你的手穿过的空间区域的每个点上，创建了一个任意角度的活动框架场（moving frame field）。 现在想象整个空间区域（可微分地）*填充*有这样一个任意活动标架场。 看看你的三个手指的任意旋转，到底怎么能产生任何数学上有趣的东西？！

非常值得注意的是，嘉当发现这样一个任意的活动标架场受制于两个极其优雅的定律，称为*嘉当的第一和第二结构方程*[3]。这些是第 38.4 节的主题。

大多数作者都默默地忽略了这些定律的存在，仿佛在说："继续前进！这没东西看。"但是当我作为学生第一次看到这两个方程时，我感到*很震惊*——认为它们是神奇的，而且可能是*黑魔法*！即使是现在，我也无法为这些美丽法则的存在提供真正令人信服的解释，但考虑到这一点...

要指定 $\mathbb{R}^3$ 中的单位向量，或者等价地指定单位球面上的一个点，我们需要两个数字——例如，经度和纬度。因此，三个任意单位向量需要六个数字。但是现在考虑嘉当的*正交*活动标架场，假设我们只知道他的*一个*向量（使用两个数字）。然后我们知道标架的其余两个正交向量的尖端位于与给定向量正交的单位圆上的某个位置。 因此，剩余的两个向量可以用*单个*角度（任一向量的）的指定来固定。

换句话说，嘉当的活动标架场的*正交性*提供的额外结构使数据量减少了*一半*。反过来说，嘉当的*正交*活动标架场的结构是非正交场的*两倍*。

想法 6：在超越了战斗以获得任意活动标架场的完全一般的结构方程之后，嘉当然后下降到地球，并使它们适应特定的几何对象。

当嘉当使它们适应曲线时，他只是恢复了（38.1）。 但是当他将他的活动标架场调整到一个曲面时——达布的想法 3——嘉当的结构方程立即产生了一个曲面的所有基本方程，但其形式比高斯、达布、科达齐（Codazzi）和其他早期的微分几何大师——都归功于他们在嘉

[2] 我们谦恭地建议您不要在公共场合尝试此操作！
[3] 这些通常也称为嘉当结构方程。

当的形式语言中的表达。 这些形式方程是第 38.5 节的主题。

在本章中，我们将紧跟加州大学洛杉矶分校巴雷特·奥尼尔（Barrett O'Neill，1924-2011）教授的开创性著作《微分几何基础》（第二修订版）的脚步，第一版于 1966 年出版。

半个多世纪后的今天，在我们看来，奥尼尔（O'Neill）这本仍然是本科阶段可用的，对该主题最清晰、优雅和（讽刺地）现代的处理方式（当前戏剧除外！）的著作——事实上，在撰写本书时，它在将形式作为向本科生观众传达微分几何的主要手段的开创性拥护方面仍然是独树一帜[4]的。

我们将比奥尼尔的课本进行更深入的研究，使用曲率 2-形式计算 n 流形的黎曼张量，并提供许多他没有提供的几何证明。 然而，奥尼尔研究了我们不会涉及的许多非常有趣的话题。 因此，我们强烈推荐奥尼尔的开创性著作作为最后一章的配套或后续课本。

38.2 联络 1-形式

38.2.1符号约定和两个定义

在开始之前，让我们提醒你一些（正在进行的）符号约定：

- 向量用小写粗体罗马字母表示，例如 $\mathbf{v}$。
- 1-形式由小写粗体希腊字母表示，例如 $\boldsymbol{\theta}$。
- 2-形式用大写粗体希腊字母表示，例如 $\boldsymbol{\Psi}$。
- 矩阵用方括号表示，例如，$[A] = [a_{ij}]$，其中 a_{ij} 表示第 i 行和第 j 列中的矩阵元素。

为避免在你适应嘉当的想法时出现任何混淆的可能性，我们将背离通常的做法，放弃爱因斯坦求和约定，而是写出求和符号，甚至明确写出总和。这实际上是一个很小的牺牲，因为我们将看到，一旦我们专门研究 2-曲面，许多“和”就会减少到一个元素！

最后，这里有两个定义，一个旧的，一个新的：

- 像往常一样，令 $\{\mathbf{e}_j\}$ 为固定欧几里得标准正交基，具有对偶 1-形式 $\{\mathbf{d}x^i\}$。因此，正如我们在（32.12）中看到的，$\mathbf{d}x^i(\mathbf{e}_j) = \delta^i_j$。
- 令 $\{\mathbf{m}_j\}$ 为嘉当（任意但可微）的标准正交活动标架场，并令其对偶 1-形式基为 $\{\boldsymbol{\theta}^i\}$。那么，根据定义，$\boldsymbol{\theta}^i(\mathbf{m}_j) = \delta^i_j$。

请注意，虽然 [32.6] 表明一般的非正交基不允许我们将单个基向量与匹配的 1-形式配对，但我们的标准正交活动标架场不再适用：我们可以将 $\boldsymbol{\theta}^i$ 视为 $\mathbf{m}_i$ 的对偶。

[4] Darling (1994) 是我们所知道的最接近例外的东西，但它更像是研究生课本。

38.2.2联络 1-形式

让我们将想法 2 应用于嘉当的活动标架场。 也就是说，让我们用标架场本身来表示活动标架场的变化率。

在 Frenet-Serret 方程的情况下，弗勒内（Frenet）标架仅沿曲线定义，我们只需要查看其沿该曲线的变化率，但现在我们必须检查活动标架场当我们在空间中的某个一般方向 $\mathbf{v}$ 上离开时如何旋转：

$$\begin{aligned}\nabla_{\mathbf{v}}\mathbf{m}_1 &= c_{11}\mathbf{m}_1 + c_{12}\mathbf{m}_2 + c_{13}\mathbf{m}_3\\ \nabla_{\mathbf{v}}\mathbf{m}_2 &= c_{21}\mathbf{m}_1 + c_{22}\mathbf{m}_2 + c_{23}\mathbf{m}_3\\ \nabla_{\mathbf{v}}\mathbf{m}_3 &= c_{31}\mathbf{m}_1 + c_{32}\mathbf{m}_2 + c_{33}\mathbf{m}_3.\end{aligned}$$

如果我们将这些系数收集到矩阵 $[C] \equiv [c_{ij}]$ 中，并进行缩写，

$$[\mathbf{m}] \equiv \begin{bmatrix}\mathbf{m}_1\\ \mathbf{m}_2\\ \mathbf{m}_3\end{bmatrix},$$

那么我们可以将这些方程更简洁紧凑地改写为

$$\nabla_{\mathbf{v}}[\mathbf{m}] = [C][\mathbf{m}].$$

由于活动标架场（根据定义）是正交的，因此这些系数（构成矩阵 $[C]$ 的元素）由下式给出

$$c_{ij} = (\nabla_{\mathbf{v}}\mathbf{m}_i)\cdot\mathbf{m}_j.$$

但是这些系数取决于 $\mathbf{v}$ 的选择，所以我们改变我们的符号[5]并将它们写成 $\mathbf{v}$ 的函数：

$$\boxed{\omega_{ij}(\mathbf{v}) \equiv (\nabla_{\mathbf{v}}\mathbf{m}_i)\cdot\mathbf{m}_j \text{ 是当框架沿 } \mathbf{v} \text{ 移动时 } \mathbf{m}_i \text{ 向 } \mathbf{m}_j \text{ 倾斜的初始速率。}} \tag{38.2}$$

得知这些 ω_{ij} 是 1-形式，你可能不会感到惊讶！

$$\boxed{\omega_{ij} \text{ 称为联络 1-形式，且} \omega_{ji} = -\omega_{ij}\text{。}}$$

[5] 令人沮丧的是，索引的顺序因作者而异。 我们将使用 O'Neill (2006) 的符号，而 Dray (2015) 例如将 i 和 j 颠倒过来：他的 ω_{ji} 就是我们的 ω_{ij}。

警告：ω_{ij} 的两个下标立即（但错误地！）让人想到 $\begin{Bmatrix}0\\2\end{Bmatrix}$ 价张量的分量。 因此，让我们强调这些都是 1-形式，位于 1-形式的矩阵中，并且下标用于标识该矩阵中每个 1-形式元素的行和列。

我们首先验证它们确实是 1-形式：

$$\begin{aligned}
&\omega_{ij}(a\,\mathbf{v}+b\mathbf{w}) = (\nabla_{a\,\mathbf{v}+b\mathbf{w}}\mathbf{m}_i)\cdot\mathbf{m}_j\\
&= (a\nabla_{\mathbf{v}}\mathbf{m}_i + b\nabla_{\mathbf{w}}\mathbf{m}_i)\cdot\mathbf{m}_j\\
&= a(\nabla_{\mathbf{v}}\mathbf{m}_i)\cdot\mathbf{m}_j + b(\nabla_{\mathbf{w}}\mathbf{m}_i)\cdot\mathbf{m}_j\\
&= a\omega_{ij}(\mathbf{v}) + b\omega_{ij}(\mathbf{w}).
\end{aligned}$$

接下来，为了证明 $\omega_{ji} = -\omega_{ij}$，我们必须证明对于每个向量 $\mathbf{v}$，$\omega_{ji}(\mathbf{v}) = -\omega_{ij}(\mathbf{v})$：

$$0 = \nabla_{\mathbf{v}}\delta^i_j = \nabla_{\mathbf{v}}(\mathbf{m}_i\cdot\mathbf{m}_j) = \nabla_{\mathbf{v}}(\mathbf{m}_i)\cdot\mathbf{m}_j + (\nabla_{\mathbf{v}}\mathbf{m}_j)\cdot\mathbf{m}_i = \omega_{ij}(\mathbf{v}) + \omega_{ji}(\mathbf{v}).$$

这个结果的几何证明隐含在即将到来的图 [38.2] 中。

它遵循

$$\omega_{11} = \omega_{22} = \omega_{33} = 0.$$

因此，系数矩阵 [C] 是一个 1-形式的斜对称矩阵，只有三个独立的元素 ω_{12}、ω_{13}、ω_{23}：

$$[\omega] \equiv \begin{bmatrix}\omega_{11} & \omega_{12} & \omega_{13}\\ \omega_{21} & \omega_{22} & \omega_{23}\\ \omega_{31} & \omega_{32} & \omega_{33}\end{bmatrix} = \begin{bmatrix}0 & \omega_{12} & \omega_{13}\\ -\omega_{12} & 0 & \omega_{23}\\ -\omega_{13} & -\omega_{23} & 0\end{bmatrix}.$$

如果我们已理解 $[\omega](\mathbf{v}) = [\omega(\mathbf{v})]$ 意味着矩阵中的每个 1-形式元素都作用于 $\mathbf{v}$，那么活动标架场随着它在一般方向 $\mathbf{v}$ 上移开的旋转可以用所谓的联络方程（Connection Equations）优雅地描述：

$$\boxed{\nabla_{\mathbf{v}}[\mathbf{m}] = [\boldsymbol{\omega}(\mathbf{v})][\mathbf{m}].} \tag{38.3}$$

在分量形式中，这表明

$$\boxed{\nabla_{\mathbf{v}}\mathbf{m}_i = \sum_j \omega_{ij}(\mathbf{v})\mathbf{m}_j\,.} \tag{38.4}$$

由于担心优雅成为清晰的敌人，让我们朴素地完整说明这些方程式：：

$$\begin{aligned}
\nabla_{\mathbf{v}}\mathbf{m}_1 &= \qquad\qquad\quad \omega_{12}(\mathbf{v})\mathbf{m}_2 + \omega_{13}(\mathbf{v})\mathbf{m}_3 \\
\nabla_{\mathbf{v}}\mathbf{m}_2 &= -\omega_{12}(\mathbf{v})\mathbf{m}_1 \qquad\qquad + \omega_{23}(\mathbf{v})\mathbf{m}_3 \\
\nabla_{\mathbf{v}}\mathbf{m}_3 &= -\omega_{13}(\mathbf{v})\mathbf{m}_1 - \omega_{23}(\mathbf{v})\mathbf{m}_2.
\end{aligned}$$

我们可以使用想法 6 从这些一般方程中恢复 Frenet-Serret 方程，如下所示：我们简单地通过下述选择将活动标架适合曲线，

$$\mathbf{m}_1 = \mathbf{T}, \qquad \mathbf{m}_2 = \mathbf{N}, \qquad \mathbf{m}_3 = \mathbf{B}, \qquad \text{和} \quad \mathbf{v} = \mathbf{T}.$$

在将这些方程与（38.1）进行比较时，唯一的谜团是缺少对应于 $\omega_{13}(\mathbf{v})$ 的元素。

为了理解这一点，回想一下 $\mathbf{B}$ 是曲线瞬时运动平面（密切平面）的法线；正交对 $(\mathbf{T},\mathbf{N})$ 在该平面内一起围绕轴 $\mathbf{B}$ 旋转。这意味着 $\mathbf{T} = \mathbf{m}_1$ 不会在 $\mathbf{B} = \mathbf{m}_3$ 的方向上倾斜，因此（38.2）意味着 $\omega_{13}(\mathbf{v}) = 0$，从而解释 谜。

38.2.3警告：前方有符号欺凌仪式！！

首先是好消息：联络 1-形式（几乎）普遍写成 ω_{ij}，就像我们写它们一样，尽管为了能够使用爱因斯坦求和约定，通常会提升一个指标 $(\omega^i{}_j)$——我们也将在适当的时候采用这个符号。

我们使用 θ^i 来表示 1-形式的对偶基遵循奥尼尔（2006）。 令人高兴的是，这种表示法很普遍，但存在其他完全合理的替代方案，例如，德雷（Tevian Dray，2015）将 θ^i 写为 σ^i。

现在是坏消息：几位备受尊敬的数学家[6]（包括不止一位菲尔兹奖得主）已经选择使用一种在我们看来被明确和不正当设计，以使倒霉的学生最大程度地糊涂的符号——作为 θ^i 的替换，这些数学家写成 ω^i！是的，他们使用相同的希腊字母来表示联络 1-形式 $\omega^i{}_j$ 和基 1-形式 ω^i，只有指标的数量作为符号代表完全不同概念的线索！

虽然比较少见，但存在同样乖张的替代符号——陈省身等人（1999）给出了一个示例——其中对偶基写成 θ^i，就像我们写的一样，但使用相同的希腊字母表示联络 1-形式，写成 $\theta^i{}_j$！

很难不把这些非凡的符号愚蠢看作是一种欺凌仪式，旨在考验胆大妄为的新手们敢于掌

[6] 可悲的是，我心爱的米斯纳、索恩和惠勒 (1973) 也必须算在这些罪人之列！ 最可悲的是，这个符号的起源正是嘉当 (1927) 本人！ 嘉当的作品一直“享有”难以理解的名声； 我认为他选择的符号没有帮助！

握长辈和长辈的博学论文的决心！

我们不得不引用我们的一位科学英雄科内利乌斯•兰佐斯（Cornelius Lanczos，1893－1974）。在他（精彩的）力学变分原理的序言中，兰佐斯（1970）写道，

今天的许多科学论文都是用一种半神秘的语言来表述的，好像是为了给读者留下一种不舒服的感觉，那就是他永远在一个超人的面前。 本书以谦卑的精神构思，是为谦卑的人而写的。

好吧，蹑手蹑脚地追随兰佐斯的脚步，我们也将谦卑地与你交谈，尽可能简单明了，欢迎你进入嘉当发现的辉煌花园。 因此，我们将用不同的字母表示不同的概念。

38.3 姿态矩阵

38.3.1途径姿态矩阵的联络形式

联邦星舰企业号[7]（或者，不那么有趣的，任何刚体）在空间中的姿态意味着它在 $\mathbb{R}^3$ 内的方向。

为了描述嘉当活动标架场 $\{\mathbf{m}_j\}$ 的姿态，我们只需要指定使固定的欧几里得坐标系 $\{\mathbf{e}_j\}$ 与 $\{\mathbf{m}_j\}$ 重合所需旋转 A。

这种旋转由姿态矩阵 $[A] = [a_{ij}]$ 描述，其中矩阵元素是位置的函数，因为整个想法是 m-标架的姿态在空间中变化。 因此，如果我们写

$$[\mathbf{e}] \equiv \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix},$$

那么

$$\boxed{[\mathbf{m}] = [A]\,[\mathbf{e}].}$$

根据我们对矩阵转置的几何解释（在 Ex. 12, p. 221 中给出），旋转矩阵必须满足

$$\boxed{[A]^T = [A]^{-1}.} \tag{38.5}$$

[7] 为避免任何歧义，我们在此参考 NCC-1701。

回想一下线性代数，这样的矩阵称为正交矩阵。

直观上很明显，[A] 的变化率决定了 m-标架的变化率，因此必然决定了联络 1-形式。为了推导出一个显式公式，让我们首先同意矩阵的外导数是其元素的外导数的矩阵：

$$\mathbf{d}[\mathrm{A}] = \mathbf{d}[\mathrm{a}_{\mathrm{ij}}] = [\mathbf{d}\mathrm{a}_{\mathrm{ij}}].$$

我们现在将证明联络 1-形式的矩阵由下式给出

$$\boxed{[\boldsymbol{\omega}] = (\mathbf{d}[\mathrm{A}])[\mathrm{A}]^{\mathrm{T}}\,.} \tag{38.6}$$

由于$[\mathrm{a}_{\mathrm{ij}}]^{\mathrm{T}} = [\mathrm{a}_{\mathrm{ij}}]$，矩阵方程（38.6）等价于这个分量方程：

$$\boxed{\omega_{\mathrm{ij}} = \sum\nolimits_{\mathrm{k}} (\mathbf{d}\mathrm{a}_{\mathrm{jk}})\mathrm{a}_{\mathrm{jk}}.} \tag{38.7}$$

为了证明这一点，我们首先计算活动标架本身的旋转速率：

$$\nabla_{\mathbf{v}}\mathbf{m}_{\mathrm{i}} = \nabla_{\mathbf{v}}\sum_{\mathrm{k}} \mathrm{a}_{\mathrm{ik}}\mathbf{e}_{\mathrm{k}} = \sum_{\mathrm{k}} (\mathbf{d}\mathrm{a}_{\mathrm{ik}})\,(\mathbf{v})\,\mathbf{e}_{\mathrm{k}},$$

因为$\nabla_{\mathbf{v}}\mathbf{e}_{\mathrm{k}} = 0$。

根据联络 1-形式的定义（38.2），我们现在必须确定 $\mathbf{m}_{\mathrm{i}}$ 的倾斜有多少是在 $\mathbf{m}_{\mathrm{j}}$ 的方向上。 既然

$$\mathbf{m}_{\mathrm{j}} = \sum_{\mathrm{l}} \mathrm{a}_{\mathrm{jl}}\,\mathbf{e}_{\mathrm{l}},$$

我们发现

$$\mathbf{m}_{\mathrm{j}} \cdot \mathbf{e}_{\mathrm{k}} = \sum_{\mathrm{l}} \mathrm{a}_{\mathrm{jl}}\,\mathbf{e}_{\mathrm{l}} \cdot \mathbf{e}_{\mathrm{k}} = \mathrm{a}_{\mathrm{jk}},$$

因此，

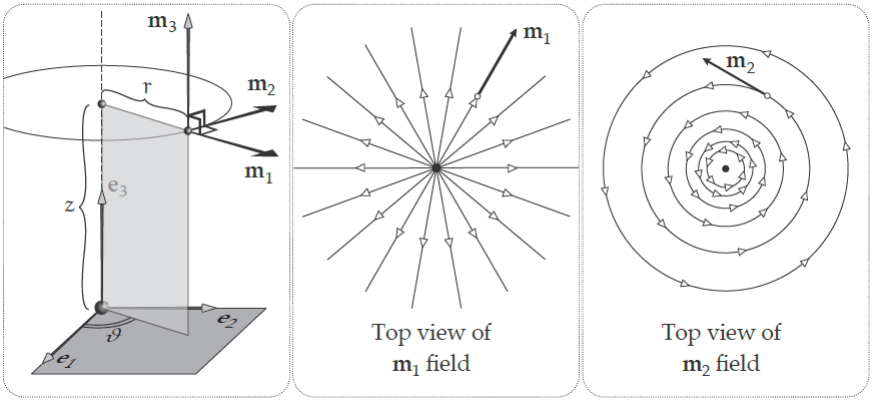
$$\omega_{\mathrm{ij}}(\mathbf{v}) \equiv (\nabla_{\mathbf{v}}\mathbf{m}_{\mathrm{i}}) \cdot \mathbf{m}_{\mathrm{j}} = \sum_{\mathrm{k}} (\mathbf{d}\mathrm{a}_{\mathrm{ik}})\,(\mathbf{v})\,\mathbf{e}_{\mathrm{k}} \cdot \mathbf{m}_{\mathrm{j}} = \sum_{\mathrm{k}} ((\mathbf{d}\mathrm{a}_{\mathrm{ik}})(\mathbf{v}))\,\mathrm{a}_{\mathrm{jk}}.$$

抽象出任意向量 $\mathbf{v}$，我们因此证明了（38.7），并由此证明了矩阵方程（38.6）。

38.3.2示例：圆柱形标架场

[38.1] 的左边画出了圆柱标架场（cylindrical frame field）的定义，该定义源自图示的圆柱极坐标系[8] (r, ϑ, z)。 从这个图中，人们很容易推导出 [练习]

$$\begin{aligned}\mathbf{m}_1 &= \cos\vartheta\,\mathbf{e}_1 + \sin\vartheta\,\mathbf{e}_2\\ \mathbf{m}_2 &= -\sin\vartheta\,\mathbf{e}_1 + \cos\vartheta\,\mathbf{e}_2\\ \mathbf{m}_3 &= \mathbf{e}_3.\end{aligned}$$



[38.1] 在左边，圆柱标架场的定义，基于图示的圆柱极坐标，(r, ϑ, z)。 从上方向下看 z 轴，中间画出 $\mathbf{m}_1$ 场，右侧画出 $\mathbf{m}_2$ 场。

需要在整个空间中感受圆柱标架场，而不是在单个图示点。 为此，[38.1] 的中间画出了整个 $\mathbf{m}_1$ 向量场的顶视图（向下看 z 轴），它向外（水平）辐射，远离 z 轴。 同样，右边画出了 $\mathbf{m}_2$ 向量场的顶视图，它围绕 z 轴旋转。我们没有绘制它，但 $\mathbf{m}_3$ 向量场只是一个均匀的向上流动，单位速度，平行于 z 轴。

从上面的公式中，我们很容易读出姿态矩阵：

$$[A] = \begin{bmatrix}\cos\vartheta & \sin\vartheta & 0\\ -\sin\vartheta & \cos\vartheta & 0\\ 0 & 0 & 1\end{bmatrix}.$$

因此，（38.6）得出

$$[\omega] = (\mathbf{d}[A])[A]^T = \begin{bmatrix}-\sin\vartheta\,\mathbf{d}\vartheta & \cos\vartheta\,\mathbf{d}\vartheta & 0\\ -\cos\vartheta\,\mathbf{d}\vartheta & -\sin\vartheta\,\mathbf{d}\vartheta & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\cos\vartheta & -\sin\vartheta & 0\\ \sin\vartheta & \cos\vartheta & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$= \begin{bmatrix}0 & \mathbf{d}\vartheta & 0\\ -\mathbf{d}\vartheta & 0 & 0\\ 0 & 0 & 0\end{bmatrix}.$$

[8] 在这里，我们从传统的 θ 切换到 ϑ，以避免与对偶基 1-形式 θ_j 产生任何可能的混淆。

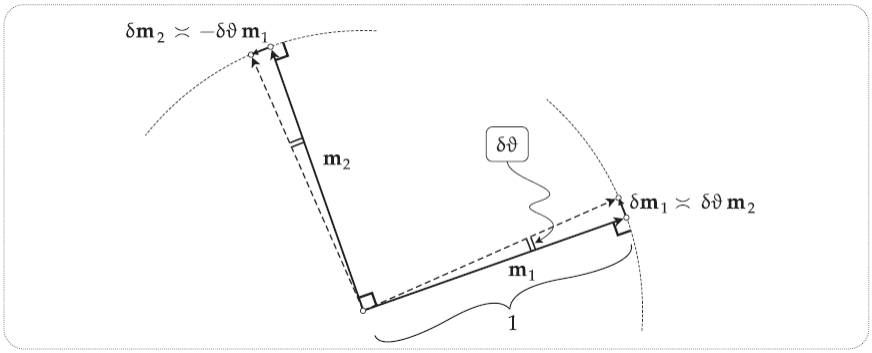
因此联络方程（38.3）变为

$$\nabla_{\mathbf{v}}\mathbf{m}_1 = \mathrm{d}\vartheta(\mathbf{v})\mathbf{m}_2 = (\nabla_{\mathbf{v}}\vartheta)\mathbf{m}_2$$

$$\nabla_{\mathbf{v}}\mathbf{m}_2 = -\mathrm{d}\vartheta(\mathbf{v})\mathbf{m}_1 = -(\nabla_{\mathbf{v}}\vartheta)\mathbf{m}_1$$

$$\nabla_{\mathbf{v}}\mathbf{m}_3 = 0.$$

我们将这个例子作为对姿态矩阵机制、联络 1-形式和联络方程的非常有用的实践。 然而，我们实际上可以直接和几何地获得这最后三个方程，而不需要任何计算。



[38.2] 在空间中沿一个短的（最终趋于零的）向量 $\mathbf{v}$ 移动会导致正交对 $(\mathbf{m}_1, \mathbf{m}_2)$ 的一个小的（最终趋于零的）旋转 $\delta\vartheta$。 $\mathbf{m}_1$ 的尖端沿着图示的单位圆弧移动距离 $\delta\vartheta$。 由于尖端最初在与其自身正交的方向上移动，$\delta\mathbf{m}_1 \asymp \delta\vartheta\mathbf{m}_2$。 同样，$\delta\mathbf{m}_2 \asymp -\delta\vartheta\mathbf{m}_1$。

最后一个方程很简单：$\mathbf{m}_3 = \mathbf{e}_3$ 是常数，所以它的导数当然为零。前两个方程更有趣，我们现在用几何推导它们。

首先，我们观察到如果 $\mathbf{v}$ 非常短，并且最终趋于零，那么

$$\mathrm{d}\vartheta(\mathbf{v}) \asymp \delta\vartheta \equiv \text{从 } \mathbf{v} \text{ 的尾部到尖端的 } \vartheta \text{ 变化。}$$

接下来，同理，

$$\begin{aligned}\nabla_{\mathbf{v}}\mathbf{m}_{\mathrm{j}} &\asymp \delta\mathbf{m}_{\mathrm{j}} \\ &\equiv \text{从 v 的尾部到尖端的 } \mathbf{m}_{\mathrm{j}} \text{ 变化} \\ &= \text{变化 } \delta\vartheta \text{ 带来的 } \mathbf{m}_{\mathrm{j}} \text{ 变化，由沿 } \mathbf{v} \text{ 的运动产生.}\end{aligned}$$

现在考虑 [38.2]，它显示正交对 $(\mathbf{m}_1, \mathbf{m}_2)$ 通过 $\delta\vartheta$ 一起刚性旋转。关注变化 $\delta\mathbf{m}_1$。$\mathbf{m}_1$ 的尖端沿着图示的单位圆弧移动距离 $\delta\vartheta$。最初，尖端在与其自身正交的方向上移动，换句话说，在 $\mathbf{m}_2$ 的方向上，如图所示。因此，我们证明了两个联络方程中的第一个，形

式为

$$\delta\mathbf{m}_1 \asymp \delta\vartheta\mathbf{m}_2$$

同一张图同时证明了第二个联络方程：

$$\delta\mathbf{m}_2 \asymp -\delta\vartheta\mathbf{m}_1$$

38.4 嘉当的两个结构方程

38.4.1按照 $\mathbf{e}_j$ 的对偶 $\mathbf{d}x^j$ 得 $\mathbf{m}^i$ 的对偶 $\boldsymbol{\theta}^i$

回想一下（嘉当的）想法 5 是用活动标架场 $\mathbf{m}^i$ 的对偶 1-形式基 $\boldsymbol{\theta}^i$ 取代它。

我们知道（根据定义）姿态矩阵 $[a_{ij}]$ 将固定欧几里得基 $\{\mathbf{e}_j\}$ 转换为嘉当的移动标架场，根据

$$\mathbf{m}_i = \sum_k a_{ik}\,\mathbf{e}_k.$$

我们现在证明相同的旋转矩阵将与 $\{\mathbf{e}_j\}$ 对偶的 1-形式基 $\{\mathbf{d}x^j\}$ 转换到与 $\{\mathbf{m}_i\}$ 对偶的 1-形式基 $\{\theta^i\}$：

$$\boxed{[\theta] = [A][\mathbf{d}x].} \tag{38.8}$$

用分量表示，

$$\boxed{\theta^i = \sum_k a_{ij}\,\mathbf{d}x^j.} \tag{38.9}$$

为了证明这一点，首先回顾（32.14），第 357 页，通过将 φ 应用于欧几里得基向量，可以将任何 1-形式 φ 分解为 $\{\mathbf{d}x_j\}$ 基中的分量。 因此，特别是，

$$\theta^i = \sum_j \theta^i(\mathbf{e}_j)\,\mathbf{d}x^j.$$

但是因为 θ^i 是 $\mathbf{m}_i$ 的对偶，

$$\theta^i(\mathbf{e}_j) = \mathbf{m}_i \cdot \mathbf{e}_j = [\sum_k a_{ik}\,\mathbf{e}_k] \cdot \mathbf{e}_j = a_{ij},$$

完成（38.9）的证明。

38.4.2嘉当第一结构方程

联络方程（38.3）告诉我们当我们在任意方向 $\mathbf{v}$ 上移动时，活动标架场 $\mathbf{m}_i$ 的旋转。因此很自然地要问对偶 1-形式 θ^i 如何在空间中变化。 由于外导数测量所有方向的变化率（同时！），我们寻求的是 $\mathbf{d}[\theta]$。

非常优雅的答案是*嘉当第一结构方程*（Cartan's First Structural Equation）：

$$\boxed{\mathbf{d}[\theta] = [\omega] \wedge [\theta],} \tag{38.10}$$

其中

$$[\theta] \equiv \begin{bmatrix} \theta^1 \\ \theta^2 \\ \theta^3 \end{bmatrix}$$

证明是一个简单的计算。写作

$$[\mathbf{d}x] \equiv \begin{bmatrix} \mathbf{d}x_1 \\ \mathbf{d}x_2 \\ \mathbf{d}x_3 \end{bmatrix}$$

（38.9）变为

$$[\theta] = [A][\mathbf{d}x].$$

微分，并首先使用（38.5），然后使用（38.6），

$$\mathbf{d}[\theta] = \mathbf{d}[A] \wedge [\mathbf{d}x] = \mathbf{d}[A][A]^T \wedge [A][\mathbf{d}x] = [\omega] \wedge [\theta].$$

完毕！

在进行具体计算时，有时根据其分量重新表述（38.10）是有用的：

$$\boxed{\mathbf{d}\theta^i = \sum_i \omega_{ij} \wedge \theta^j} \tag{38.11}$$

38.4.3嘉当第二结构方程

现在让我们问一下联络 1-形式如何在空间中变化。 同样，外导数同时编码所有可能方向的变化，因此我们寻求的是 $\mathbf{d}[\omega]$。

非常优雅的答案是*嘉当第二结构方程*（Cartan's Second Structural Equation）：

$$\boxed{\mathbf{d}[\omega] = [\omega] \wedge [\omega].} \tag{38.12}$$

正如我们将在第 38.12 节中看到的那样，这个方程的更深层次的几何意义是它的特征是没有曲率，即我们迄今为止一直在其中操作的欧几里得空间 $\mathbb{R}^3$ 的*平坦性*（flatness）。 当我们继续讨论弯曲的流形时，例如爱因斯坦的引力扭曲时空，这个方程的两侧不再相等，实际上会发现两侧的*差异*编码了流形的曲率。

为了证明这个等式，我们需要三个简单的引理。 首先，注意如果 f 和 g 是函数，那么

$$\mathbf{d}(\mathbf{d}\mathrm{f}\,\mathrm{g}) = \mathbf{d}(\mathrm{g}\,\mathbf{d}\mathrm{f}) = \mathbf{d}\mathrm{g} \wedge \mathbf{d}\mathrm{f} = -\mathbf{d}\mathrm{f} \wedge \mathbf{d}\mathrm{g}.$$

接下来，回想转置的几何解释（在 Ex. 12, p. 221 中给出）意味着 $(\mathrm{AB})^{\mathrm{T}} = \mathrm{B}^{\mathrm{T}}\,\mathrm{A}^{\mathrm{T}}$。 最后，回想一下 $[\omega]$ 是斜对称的：$[\omega]^{\mathrm{T}} = -[\omega]$。

因此，微分（38.6），

$$\begin{aligned}
\mathbf{d}[\omega] &= \mathbf{d}[(\mathbf{d}[\mathrm{A}])[\mathrm{A}]^{\mathrm{T}}] \\
&= -(\mathbf{d}[\mathrm{A}]) \wedge \mathbf{d}[\mathrm{A}]^{\mathrm{T}} \\
&= -(\mathbf{d}[\mathrm{A}])[\mathrm{A}]^{\mathrm{T}} \wedge [\mathrm{A}]\,\mathrm{d}[\mathrm{A}]^{\mathrm{T}} \\
&= -(\mathbf{d}[\mathrm{A}])[\mathrm{A}]^{\mathrm{T}} \wedge [(\mathrm{d}[\mathrm{A}])[\mathrm{A}]^{\mathrm{T}}]^{\mathrm{T}} \\
&= -[\omega] \wedge [\omega]^{\mathrm{T}} \\
&= [\omega] \wedge [\omega].
\end{aligned}$$

完毕！

在进行具体计算时，根据其组成部分重新表述（38.12）有时是有用的：

$$\boxed{\mathbf{d}\omega_{\mathrm{ij}} = \sum_{\mathrm{k}} \omega_{\mathrm{ik}} \wedge \omega_{\mathrm{kj}}} \tag{38.13}$$

最后，请注意，一次记住两个结构方程很容易，只需记住以下内容：

$$\mathbf{d}[?] = [\omega] \wedge [?].$$

38.4.4示例：球面标架场

在本节中，我们将找到与第 35.4 节的球极坐标[9] (r, φ, ϑ) 相关的对偶基和联络形式，我们将使用它们来测试/说明嘉当结构方程。

在下文中，我们将写成 $c_\vartheta \equiv \cos\vartheta$、$s_\phi \equiv \sin\phi$ 等，部分是为了节省篇幅，但也因为我们发现这是在私人的纸笔计算中的一种澄清符号。

从[38.3]，我们看到，

$$[x] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \rho c_\vartheta \\ \rho s_\vartheta \\ r c_\phi \end{bmatrix} = \begin{bmatrix} r s_\phi c_\vartheta \\ r s_\phi s_\vartheta \\ r c_\phi \end{bmatrix}$$

从同一张图中，我们还推导出 [练习] 所示的球面标架场（spherical frame field）由下式给出

$$[\mathbf{m}] = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix} = \begin{bmatrix} s_\phi[c_\vartheta \mathbf{e}_1 + s_\vartheta \mathbf{e}_2] + c_\phi \mathbf{e}_3 \\ c_\phi[c_\vartheta \mathbf{e}_1 + s_\vartheta \mathbf{e}_2] - s_\phi \mathbf{e}_3 \\ -s_\vartheta \mathbf{e}_1 + c_\vartheta \mathbf{e}_2 \end{bmatrix} = [A] \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix}$$

从中我们立即推导出姿态矩阵：

$$[A] = \begin{pmatrix} s_\phi c_\vartheta & s_\phi s_\vartheta & c_\phi \\ c_\phi c_\vartheta & c_\phi s_\vartheta & -s_\phi \\ -s_\vartheta & c_\vartheta & 0 \end{pmatrix}$$

此时我们可以使用（38.6）计算联络形式：

$$[\omega] = (\mathbf{d}[A])[A]^T.$$

显然，这将需要冗长而艰巨的计算，我们在此将其指定为（严酷的）[10]练习！ 但是，让我们至少提供一些关于如何进行此类计算的具体指导。

---

[9] O'Neill (2006, pp. 86, 97) 是从赤道测量 ϕ，因此我们的 ϕ 是通过从 $(\pi/2)$ 中减去他而获得的。这具有互换 $\sin\phi$ 和 $\cos\phi$ 的效果。 此外，为了使坐标系为右手坐标系，奥尼尔坐标的顺序为 (r, ϑ, φ)，因此他的球坐标系场记为 $(F1, F2, F3)$，因此与我们的坐标系相关为 $(\mathbf{m}_1 = \mathbf{F}_1, \mathbf{m}_2 = -\mathbf{F}_3, \mathbf{m}_3 = \mathbf{F}_2)$。

[10] 当被迫为我的旧金山大学的学生安排如此残酷的计算任务时，我同时安排他们观看塔伦蒂诺的《杀死比尔》，第 2 卷，第 8 章：《派美的残酷教导》，因为它证明了这种“残酷”最终如何将女主（Beatrix Kiddo）从坟墓中拯救出来！

首先，我们当然不需要将整个矩阵 $\mathbf{d}[A]$ 乘以整个矩阵 $[A]$。 例如，计算对角线元素是浪费时间：我们知道它们必然为零！ 实际上我们只需要计算产品右上角的三个元素：ω_{12}、ω_{13}、ω_{23}。

为了仅针对这三个元素，使用矩阵乘积的分量形式更简单，即（38.7）：

$$\omega_{ij} = \sum_k (\mathbf{d}a_{ik})\, a_{jk}$$

因此，我们永远不需要实际形成 $[A]$ 的转置。 相反，我们将这个公式解释为告诉我们

> 为了获得 ω_{ij}，我们取 $[A]$ 的第 i 行的 d，然后将每个元素乘以 $[A]$ 的第 j 行中的对应元素。

例如，要找到 ω_{12}，我们必须取第一行的 $\mathbf{d}$，并乘以第二行：

$$\omega_{12} = [\mathbf{d}(s_\phi\, c_\vartheta)]c_\phi\, c_\vartheta + [\mathbf{d}(s_\phi\, s_\vartheta)]c_\phi\, s_\vartheta - [\mathbf{d}(c_\phi)]s_\phi.$$

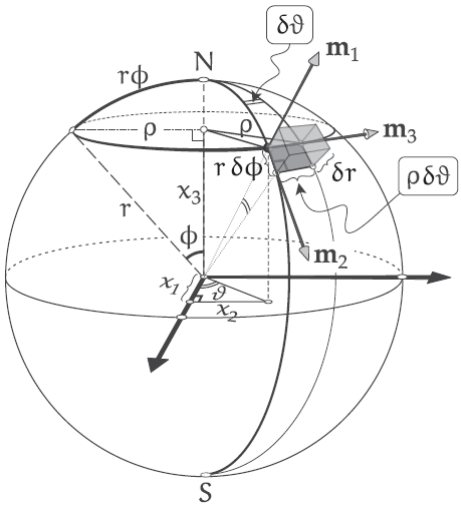
即使你不想计算所有三种联络形式，我们也敦促你至少做这一种。 求导后，你会遇到五个项，每个项都是四个三角函数的乘积。 但是，“奇迹般地”，一切都抵消或简化了，我们得到了一个非常简单的答案：$\omega_{12} = \mathbf{d}\phi$！

同样“神奇”的抵消和简化出现在其他两种联络形式的计算中，我们得到以下简单的公式：

$$\boxed{\begin{aligned} \omega_{12} &= \mathbf{d}\phi \\ \omega_{13} &= \sin\phi\, \mathbf{d}\vartheta \\ \omega_{23} &= \cos\phi\, \mathbf{d}\vartheta. \end{aligned}} \tag{38.14}$$

回顾序言中所说的，我们这里有三个我们称之为“虚假奇迹”的例子：如果所有这些项都被抵消，它们一开始就不应该存在——我们一定是以错误的方式看待数学 ！

在我们的戏剧中，到了这一点，知道看待这个问题的正确方法是用几何学，这肯定不足为奇。 从而避免了你刚刚执行的所有痛苦计算，并直接直观地获得结果！我们将在本节末尾提供这个牛顿几何解释，但现在我们继续努力测试嘉当的结构方程。



[38.3] $\theta_1 = \mathbf{d}r$, $\theta_2 = r\,\mathbf{d}\phi$, 和 $\theta_3 = r\sin\phi\mathbf{d}\vartheta$ 的几何证明。

为此，让我们找到对偶基 1-形式 θ^i。 同样，我们当然可以通过盲目计算获得这些：我们知道 [A] 并且我们知道 [x]，因此我们可以计算 $\mathbf{d}[x] = [\mathbf{d}x]$，因此我们可以使用（38.8）：

$$[\theta] = [A]\,[\mathbf{d}x].$$

但是，再一次，这显然需要一个冗长的计算，我们在此将其指定为另一个严酷的练习！（如果你真的进行这些计算，你会更加感激接下来会发生什么，这些计算充满了许多抵消和简化。成功地穿过这片密集的符号丛林后，你应该出现在一片空地上，在其中你将 找到这些简单的公式：（38.16）。）

在这里，牛顿的几何也允许我们在空中跳跃，飞越我们下方冗长而不透明的计算，直接降落在答案上！

考虑 [38.3] 中的阴影盒，它会缩小并最终趋于零。 它的边最终等于

$$\delta r\mathbf{m}_1, \qquad r\,\delta\phi\mathbf{m}_2, \qquad \rho\,\delta\vartheta\mathbf{m}_3 = r\,s_\phi\,\delta\vartheta\mathbf{m}_3,$$

如图所示。 因此，指向任意方向的盒子的对角线是短（最终消失）向量 $\mathbf{v}$，由下式给出

$$\boxed{\mathbf{v} \asymp \delta r\mathbf{m}_1 + r\,\delta\phi\mathbf{m}_2 + r\,s_\phi\,\delta\vartheta\mathbf{m}_3.} \tag{38.15}$$

根据对偶基的定义，$\theta^i(\mathbf{m}_j) = \delta^i_j$ ，所以，

$$\theta^1(\mathbf{v}) \asymp \delta r \asymp \mathbf{d}r(\mathbf{v}) \implies \theta^1 = \mathbf{d}r.$$

这是因为 $\mathbf{v}$ 是任意的，因此可以被抽象掉。

我们以完全相同的方式找到 θ^2 和 θ^3。 因此，[38.3] 的几何立即产生

$$\boxed{[\boldsymbol{\theta}] = \begin{bmatrix}\theta^1\\ \theta^2\\ \theta^3\end{bmatrix} = \begin{bmatrix}\mathbf{d}r\\ r\,\mathbf{d}\phi\\ r\,s_\phi\,\mathbf{d}\vartheta\end{bmatrix}.} \tag{38.16}$$

将此几何与你刚刚执行的计算进行对比！

现在我们知道了对偶基和联络形式，让我们检查嘉当第一结构方程（38.10），

$$\mathbf{d}[\boldsymbol{\theta}] = [\omega] \wedge [\boldsymbol{\theta}].$$

在左手边，我们发现

$$\mathbf{d}[\boldsymbol{\theta}] = \mathbf{d}\begin{bmatrix}\mathbf{d}r\\ r\,\mathbf{d}\phi\\ r\,s_\phi\,\mathbf{d}\vartheta\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{d}r \wedge \mathbf{d}\phi\\ s_\phi \mathbf{d}r \wedge \mathbf{d}\vartheta + r\,c_\phi \mathbf{d}\phi \wedge \mathbf{d}\vartheta\end{bmatrix}$$

在右手边，我们发现

$$\begin{aligned}
[\omega] \wedge [\boldsymbol{\theta}] &= \begin{bmatrix}0 & \omega_{12} & \omega_{13}\\ -\omega_{12} & 0 & \omega_{23}\\ -\omega_{13} & -\omega_{23} & 0\end{bmatrix} \wedge \begin{bmatrix}\theta^1\\ \theta^2\\ \theta^3\end{bmatrix}\\
&= \begin{bmatrix}0 & \mathbf{d}\phi & s_\phi\,\mathbf{d}\vartheta\\ -\mathbf{d}\phi & 0 & c_\phi\,\mathbf{d}\vartheta\\ -s_\phi\,\mathbf{d}\vartheta & -c_\phi\,\mathbf{d}\vartheta & 0\end{bmatrix} \wedge \begin{bmatrix}\mathbf{d}r\\ r\,\mathbf{d}\phi\\ r\,s_\phi\,\mathbf{d}\vartheta\end{bmatrix}\\
&= \begin{bmatrix}0\\ \mathbf{d}r \wedge \mathbf{d}\phi\\ s_\phi \mathbf{d}r \wedge \mathbf{d}\vartheta + r c_\phi \mathbf{d}\phi\end{bmatrix},
\end{aligned}$$

从而验证球标架场的嘉当第一结构方程。

现在让我们检查嘉当的第二个结构方程（38.12），

$$\mathrm{d}[\omega] = [\omega] \wedge [\omega].$$

在左手边，我们发现

$$\mathbf{d}\begin{bmatrix}0 & \mathbf{d}\phi & s_\phi\,\mathbf{d}\vartheta\\ -\mathbf{d}\phi & 0 & c_\phi\,\mathbf{d}\vartheta\\ -s_\phi\,\mathbf{d}\vartheta & -c_\phi\,\mathbf{d}\vartheta & 0\end{bmatrix} = \begin{bmatrix}0 & 0 & c_\phi \mathbf{d}\phi \wedge \mathbf{d}\vartheta\\ 0 & 0 & -s_\phi \mathbf{d}\phi \wedge \mathbf{d}\vartheta\\ -c_\phi \mathbf{d}\phi \wedge \mathbf{d}\vartheta & s_\phi \mathbf{d}\phi \wedge \mathbf{d}\vartheta & 0\end{bmatrix}$$

在右手边，我们发现

$$
[\omega]\wedge[\omega]=\begin{bmatrix}0 & \mathbf{d}\phi & s_\phi\,\mathbf{d}\vartheta\\ -\mathbf{d}\phi & 0 & c_\phi\,\mathbf{d}\vartheta\\ -s_\phi\,\mathbf{d}\vartheta & -c_\phi\,\mathbf{d}\vartheta & 0\end{bmatrix}\wedge\begin{bmatrix}0 & \mathbf{d}\phi & s_\phi\,\mathbf{d}\vartheta\\ -\mathbf{d}\phi & 0 & c_\phi\,\mathbf{d}\vartheta\\ -s_\phi\,\mathbf{d}\vartheta & -c_\phi\,\mathbf{d}\vartheta & 0\end{bmatrix}
$$

$$
=\begin{bmatrix}0 & 0 & c_\phi\mathbf{d}\phi\wedge\mathbf{d}\vartheta\\ 0 & 0 & -s_\phi\mathbf{d}\phi\wedge\mathbf{d}\vartheta\\ -c_\phi\mathbf{d}\phi\wedge\mathbf{d}\vartheta & s_\phi\mathbf{d}\phi\wedge\mathbf{d}\vartheta & 0\end{bmatrix}
$$

从而验证球框架场的嘉当第二结构方程。

正如所承诺的，我们通过返回联络形式结束本节，表明导致（38.14）的冗长计算可以用简单、清晰的牛顿几何代替。

为此，我们将牛顿终极等式直接应用于定义（38.2）：

$\omega_{ij}(\mathbf{v})\equiv(\nabla_{\mathbf{v}}\mathrm{mi})\cdot\mathrm{mj}$ 是当标架沿 $\mathbf{v}$ 移动时 $\mathbf{m}_i$ 向 $\mathbf{m}_j$ 倾斜的初始速率。

在这里，我们将 $\mathbf{v}$ 是由 (38.15) 给出的任意的、短的（最终趋于零的）向量：

$$
\mathbf{v}\asymp\delta r\,\mathbf{m}_1+r\delta\phi\,\mathbf{m}_2+r\,s_\phi\,\delta\vartheta\mathbf{m}_3\text{。}
$$

以下分析中的关键工具是，由于 $\mathbf{v}$ 很短，并且最终趋于零，

$\nabla_{\mathbf{v}}\mathbf{m}_i\asymp\mathbf{m}_i$ 从 $\mathbf{v}\equiv\delta\mathbf{m}_i$ 的尾部到尖端的变化。

请注意，为了方便地可视化变化 $\delta\mathbf{m}_i$，我们将把所有向量 $\mathbf{m}_i$ 绘制为从一个公共点发出； 那么 $\delta\mathbf{m}_i$ 只是 $\mathbf{m}_i$ 的尖端的运动。

我们从 $\omega_{12}(\mathbf{v})\equiv(\delta\mathbf{m}_1)\cdot\mathbf{m}_2$ 开始。 分别考虑沿 $\mathbf{v}$ 的三个分量中的每一个移动导致的 $\mathbf{m}_1$ 中的变化 $\delta\mathbf{m}_1$：

- 沿 $\delta r\,\mathbf{m}_1$ 径向向外移动使 $\mathbf{m}_1$ 保持不变：它不会向 $\mathbf{m}_2$（或其他任何东西！）倾斜。

- 沿 $r\,s_\phi\,\delta\vartheta\,\mathbf{m}_3$ （正东，沿纬度圆）移动导致 $\mathbf{m}_1$ 围绕[38.5] 的上半部分所示的圆锥体旋转，该圆锥体具有半垂直角 ϕ 和倾斜高度 1，因此底半径为 s_ϕ。 显然，$\delta\mathbf{m}_1$ 与 $\mathbf{m}_2$ 正交，因此这对 ω_{12} 的贡献也为零。

更详细地说，我们注意到该图表明，以备将来使用

$$\delta\mathbf{m}_1 = s_\phi\, \delta\vartheta\, \mathbf{m}_3. \tag{38.17}$$

- 顺便说一句，为了将来的使用，我们注意到这种旋转导致 $\mathbf{m}_2$ 围绕 [38.5] 下半部分所示的锥体旋转，半垂直角 $(\pi 2 - \phi)$ 和倾斜高度为 1，因此底半径为 c_ϕ。 因此，如图所示，

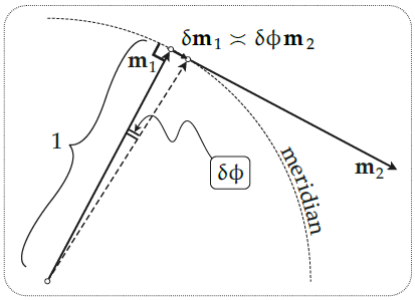
$$\delta\mathbf{m}_2 = c_\phi\, \delta\vartheta\, \mathbf{m}_3. \tag{38.18}$$

- 将我们的焦点返回到 $\mathbf{m}_1$，$\mathbf{v}$ 中唯一导致 $\mathbf{m}_1$ 向 $\mathbf{m}_2$ 倾斜的部分是沿子午线向赤道的南移 $r\,\delta\phi\mathbf{m}_2$。 图 [38.4] 绘制在该子午线的垂直平面上，包含 $\mathbf{m}_1$ 和 $\mathbf{m}_2$。 当 $\mathbf{m}_1$ 旋转角度 $\delta\phi$ 时，其尖端最终沿 $\delta\mathbf{m}_1 \asymp \delta\phi\mathbf{m}_2$ 移动，如图所示。 所以，

$$\omega_{12}(\mathbf{v}) \asymp \delta\mathbf{m}_1 \cdot \mathbf{m}_2 \asymp (\delta\phi\mathbf{m}_2) \cdot \mathbf{m}_2 = \delta\phi \asymp d\phi(\mathbf{v}) \implies \omega_{12} = d\phi.$$

接下来，考虑

$$\omega_{13}(\mathbf{v}) \equiv (\nabla_{\mathbf{v}}\mathbf{m}_1) \cdot \mathbf{m}_3.$$



[38.4] $\omega_{12} = \mathbf{d}\phi$ 的几何证明。

在这里，我们正在考虑与上述相同的变化 $\delta\mathbf{m}_1$ 在 $\mathbf{m}_1$ 中，但现在我们想知道有多少 $\delta\mathbf{m}_1$ 指向 $\mathbf{m}_3$（而不是 $\mathbf{m}_2$）的方向。

正如我们所见，$\mathbf{v}$ 中唯一沿 $\mathbf{m}_3$ 产生 $\delta\mathbf{m}_1$ 的部分是沿 $r\,s_\phi\,\delta\vartheta\,\mathbf{m}_3$ 的运动。 从（38.17）我们立即推导出

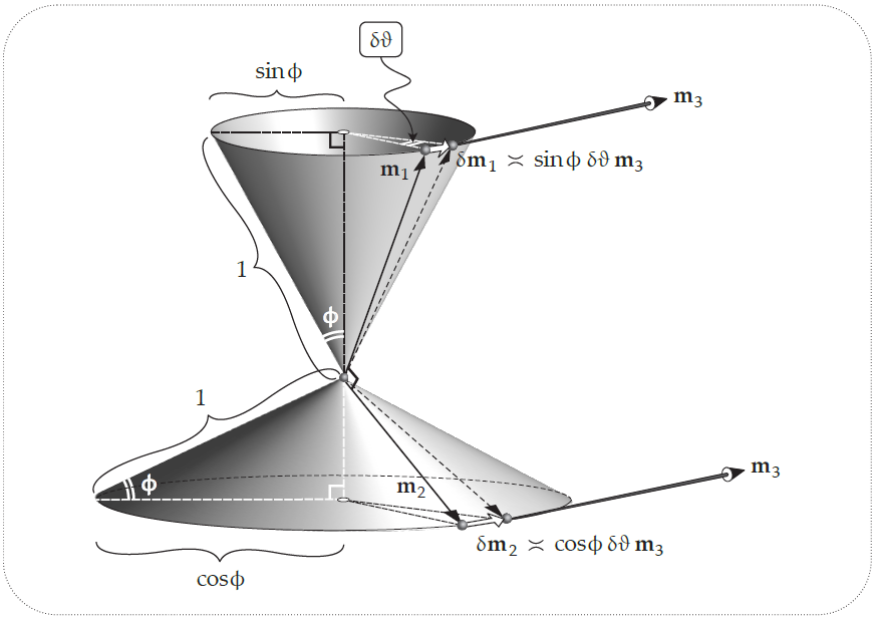
$$\omega_{13}(\mathbf{v}) \asymp \delta\mathbf{m}_1 \cdot \mathbf{m}_3 \asymp (s_\phi\,\delta\vartheta\mathbf{m}_3) \cdot \mathbf{m}_3 = s_\phi\,\delta\vartheta \asymp s_\phi\, d\vartheta(\mathbf{v}) \implies \omega_{13} = s_\phi\, d\vartheta.$$

最后，考虑 $\omega_{23}(\mathbf{v}) \equiv (\nabla_{\mathbf{v}}) \cdot \mathbf{m}_3$。 很容易看出 [练习]，导致 $\mathbf{m}_2$ 向 $\mathbf{m}_3$ 方向倾斜的唯一分量是沿 $r\,s_\phi\,\delta\vartheta\mathbf{m}_3$（正东，沿纬度圆）的运动。 因此，使用（38.18），如[38.5]的底部圆锥所示，我们推导出

$$\omega_{23}(\mathbf{v}) \asymp \delta\mathbf{m}_2 \cdot \mathbf{m}_3 \asymp (c_\phi\delta\vartheta\mathbf{m}_3) \cdot \mathbf{m}_3 = c_\phi\,\delta\vartheta \asymp c_\phi\, d\vartheta(\mathbf{v}) \implies \omega_{23} = c_\phi\, d\vartheta.$$

这完成了我们对（38.14）的几何证明。 再次，将牛顿几何推理的启发性简单性和直接

性与你最初为了得出这三个公式而被迫执行的三个冗长的、毫无启发性的计算进行对比！



[38.5] $\omega_{13} = \sin\phi d\vartheta$（顶锥）和 $\omega_{23} = \cos\phi d\vartheta$（底锥）的几何证明。

38.5 曲面的六个基本形式方程

38.5.1嘉当活动标架适用于曲面：形状算子与外在曲率

我们现在结合嘉当的想法 6 来实现达布（Darboux）的想法 3（使标架适应曲面）：使用形式，将嘉当的两个结构方程专门用于这个适配的标架场。

给定具有法线 $\mathbf{n}$ 的曲面 $\mathcal{S}$，我们选择 $\mathbf{m}_3 = \mathbf{n}$； 这称为适应标架场（adapted frame field）。 然后 $(\mathbf{m}_1, \mathbf{m}_2)$ 成为向量 $\mathbf{v}$ “在” $\mathcal{S}$ 内的内蕴（即切线）基。

此后，我们将只考虑与 $\mathcal{S}$ 相切的向量 $\mathbf{v}$。当应用于这种切向量时，联络方程（38.4）为形状算子 S（15.4）产生以下简单表达式联络形式：

$$\boxed{\mathrm{S}(\mathbf{v}) = \omega_{13}(\mathbf{v})\mathbf{m}_1 + \omega_{23}(\mathbf{v})\mathbf{m}_2.} \tag{38.19}$$

证明如下：

$$\mathrm{S}(\mathbf{v}) = -\nabla_{\mathbf{v}}\mathbf{n} = -\nabla_{\mathbf{v}}\mathbf{m}_3 = -[\nabla_{\mathbf{v}}\mathbf{m}_3 \cdot \mathbf{m}_1]\mathbf{m}_1 - [\nabla_{\mathbf{v}}\mathbf{m}_3 \cdot \mathbf{m}_2]\mathbf{m}_2.$$

最后一步是插入联络 1-形式（38.2）的定义，注意它们的斜对称性。

表示形状算子的矩阵 [S] 的第一列是第一个基向量 S 下的图像，即 $S(\mathbf{m}_1)$。 同样，第二列是第二个基向量的图像。 所以，

$$\boxed{[S] = \begin{bmatrix} \omega_{13}(\mathbf{m}_1) & \omega_{13}(\mathbf{m}_2) \\ \omega_{23}(\mathbf{m}_1) & \omega_{23}(\mathbf{m}_2) \end{bmatrix}.} \tag{38.20}$$

现在让我们用它来表达外在曲率 $\mathcal{K}_{\text{ext}} = \kappa_1\kappa_2$ 的联络形式。既然我们打算重新证明前四幕中的基本结果，我们必须避免基于这些先验知识做出假设。因此，我们必须重新区分曲率的外在测量和内蕴测量。只有这样我们才能证明（并欣赏！）高斯绝妙定理，它揭示了这两种曲率测量实际上是相同的！

另一方面，我们当然不打算使用形式从头开始重建整个微分几何大厦！只要我们明确指出调用了哪些先前的结果，并且严格避免循环推理，这种混合方法就不会造成任何损害。

因此，现在回想一下（15.8），$\mathcal{S}$ 的外在曲率 $\mathcal{K}_{\text{ext}}$ 测量法向量 $\mathbf{m}_3$ 在一小块曲面上的扩展（spreading）。更准确地说，它是 S 的局部面积扩展因子。换句话说，它是表示 S 的矩阵 [S] 的行列式：

$$\mathcal{K}_{\text{ext}} = \det[S] = \omega_{13}(\mathbf{m}_1)\omega_{23}(\mathbf{m}_2) - \omega_{13}(\mathbf{m}_2)\omega_{23}(\mathbf{m}_1).$$

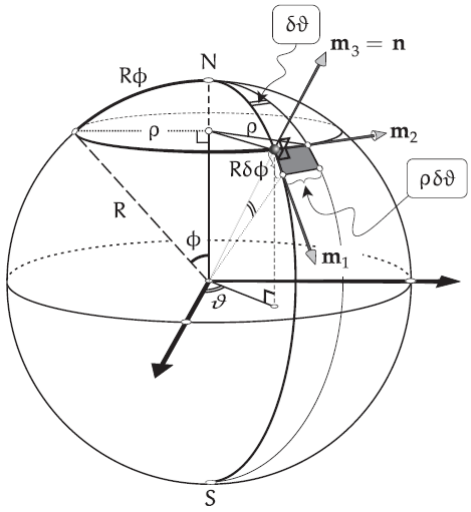
我们已经得出了一个外在曲率的表达式，它很快就会证明它的重要性：

$$\boxed{\mathcal{K}_{\text{ext}} = (\omega_{13} \wedge \omega_{23})(\mathbf{m}_1, \mathbf{m}_2).} \tag{38.21}$$

最后，考虑与适应标架场 $(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3)$ 对偶的 1-形式基 θ^i。 这些给出了任何切向量 $\mathbf{v}$ 相对于适应标架场的坐标 $\theta^i(\mathbf{v}) = \mathbf{v} \cdot \mathbf{m}_i$。 与联络形式一样，这些对偶基形式将仅应用于 $\mathcal{S}$ 的切向量，因此它们成为 $\mathcal{S}$ 上的形式。奥尼尔（O'Neill，2006）完美地说明了含义："这种限制对 θ^3，因为如果 $\mathbf{v}$ 与 $\mathcal{S}$ 相切，它与 $\mathbf{m}_3$ 正交，所以 $\theta^3(v) = v \cdot \mathbf{m}_3 = 0$。 因此 θ^3 在 $\mathcal{S}$ 上同样为零。"

38.5.2示例：球面

为了使球面标架场适应特定的球面，$r = R$，我们必须重新标记[38.3]中所示的向量。 径向向量（它是 $\mathbf{m}_1$）是球体的法线，因此（根据我们的约定）它必须重新标记为 $\mathbf{m}_3$。 与球体相切的原始框架域的其余两个成员同样必须重新标记，如 [38.6] 所示，遵守右手规则：$\mathbf{m}_1$ 指向南，$\mathbf{m}_2$ 指向东。



[38.6] 适应标架场（根据定义）具有 $\mathbf{m}_3 = \mathbf{n}$。 由此产生的名称变化意味着，现在，$\theta^1 = R\mathbf{d}\phi$，并且 $\theta^2 = \rho\mathbf{d}\vartheta = R\sin\phi\mathbf{d}\vartheta$。

因此，先前获得的 1-形式基 θ^i 和联络形式 ω_{ij} 现在由下式给出[11]，

$$\begin{aligned} \theta^1 &= R\,\mathbf{d}\phi & \omega_{12} &= c_\phi\,\mathbf{d}\vartheta \\ \theta^2 &= R\,s_\phi\,\mathbf{d}\vartheta & \omega_{13} &= -\mathbf{d}\phi \\ & & \omega_{23} &= -s_\phi\,\mathbf{d}\vartheta \end{aligned}$$

作为我们刚刚解释的几何方法的宝贵实践，尝试检查这五个结果中的每一个，只需检查 [38.6] - 不需要计算！

最后，注意阴影面元是 $\mathcal{A} = (R\,\delta\phi)(\rho\,\delta\vartheta) = R^2\sin\vartheta\,\delta\phi\,\delta\vartheta$。 因此，球体的面积 2-形式——或者，实际上，任何曲面的面积——是

$$\boxed{\mathcal{A} = \theta^1 \wedge \theta^2.} \tag{38.22}$$

38.5.3基分解的唯一性

让我们简要回顾一下 1-形式和 2-形式分解相对于基的唯一性； 短期内将需要这两个事实。

对于 $\mathcal{S}$ 中的 1-形式，我们的基是 (θ^1, θ^2)。 回想一下，两个 1-形式 φ 和 ψ 相等当

[11]同样，这些与你在 O'Neill (2006, p. 267) 中找到的结果不匹配，因为他定义 ϕ 的不同方法改变了一切。

且仅当

$$\varphi(\mathbf{m}_1) = \psi(\mathbf{m}_1) \quad 和 \quad \varphi(\mathbf{m}_2) = \psi(\mathbf{m}_2).$$

紧接着 [练习] 可以将 φ 唯一地分解为

$$\boxed{\varphi = \varphi(\mathbf{m}_1)\,\theta^1 + \varphi(\mathbf{m}_2)\,\theta^2.} \tag{38.23}$$

同样，两个 2-形式 Φ 和 Ψ 相等当且仅当

$$\Phi(\mathbf{m}_1, \mathbf{m}_2) = \Psi(\mathbf{m}_1, \mathbf{m}_2).$$

紧接着[练习] Φ可以唯一地分解为

$$\boxed{\Phi = \Phi(\mathbf{m}_1, \mathbf{m}_2)\mathcal{A} = \Phi(\mathbf{m}_1, \mathbf{m}_2)\,\theta^1 \wedge \theta^2.} \tag{38.24}$$

特别是，将此结果应用于（38.21），我们发现

$$\boxed{\omega_{13} \wedge \omega_{23} = (\omega_{13} \wedge \omega_{23})(\mathbf{m}_1, \mathbf{m}_2)\,\theta^1 \wedge \theta^2 = \mathcal{K}_{\text{ext}}\,\theta^1 \wedge \theta^2.} \tag{38.25}$$

38.5.4曲面的六个基本形式方程

让我们立即说明所有这六个方程，并一举说出它们的名称[12]。 然后我们将通过计算来证明它们，最后逐步提取它们真正的几何意义。

$$\left.\begin{aligned} \mathbf{d}\theta^1 &= \omega_{12} \wedge \theta^2 \\ \mathbf{d}\theta^2 &= \omega_{21} \wedge \theta^1 \end{aligned}\right\} \quad 第一结构方程$$

$$\omega_{31} \wedge \theta^1 + \omega_{32} \wedge \theta^2 = 0 \quad 对称方程$$

$$\mathbf{d}\omega_{12} = \omega_{13} \wedge \omega_{32} \quad 高斯方程$$

$$\left.\begin{aligned} \mathbf{d}\omega_{13} &= \omega_{12} \wedge \omega_{23} \\ \mathbf{d}\omega_{23} &= \omega_{21} \wedge \omega_{13} \end{aligned}\right\} \quad \text{Peterson–Mainardi–Codazzi 方程}$$

[12] 在这 … 些名称被广泛使用，但 … zzi 方程，但我们试图 … 管 Mainardi 在 … 1853 年，一位年 … 由明丁指导！）最终 … hillips (1979)。

(38.26)

在我们对曲面 $\mathcal{S}$ 的描述中，只有 $\mathbf{m}_3$ 在几何上由曲面本身唯一地确定为它的法线。同时，我们可以在曲面内绕轴 $\mathbf{m}_3$ 任意旋转 $(\mathbf{m}_1, \mathbf{m}_2)$。 所有这些任意不同的标架场产生了对曲面的不同描述——不同的对偶基 1-形式和不同的联络 1-形式。 然而，尽管有这些自由，上面的六个方程包含了关于曲面可以知道或说的一切。

因此，值得注意的是，从嘉当结构方程中得出所有六个方程是多么快速和容易——我们需要做的就是设置 $\theta^3 = 0$。

因为 $\omega_{11} = 0$，第一个结构方程（38.11）立即产生

$$\mathbf{d}\theta^1 = \sum_{\mathrm{j}} \omega_{1\mathrm{j}} \wedge \theta^{\mathrm{j}} = \omega_{12} \wedge \theta^2 + \omega_{13} \wedge \theta^3 = \omega_{12} \wedge \theta^2.$$

$\mathbf{d}\theta^2$ 的计算基本相同。

对称方程以相同的方式遵循：

$$0 = \mathbf{d}\theta^3 = \sum_{\mathrm{j}} \omega_{3\mathrm{j}} \wedge \theta^{\mathrm{j}} = \omega_{31} \wedge \theta^1 + \omega_{32} \wedge \theta^2.$$

接下来，由于 $\omega_{11} = 0 = \omega_{22}$，高斯方程紧随第二个结构方程 (38.13)：

$$\mathbf{d}\omega_{12} = \sum_{\mathrm{k}} \omega_{1\mathrm{k}} \wedge \omega_{\mathrm{k}2} = \omega_{13} \wedge \omega_{32}.$$

最后两个 Peterson-Mainardi-Codazzi 方程以完全相同的方式从第二个结构方程得出。

38.6 对称方程和 Peterson – Mainardi – Codazzi 方程的几何意义

但是这些基本形式方程实际上意味着什么？

让我们从对称方程开始。 如果我们将这个消失的 2 形式应用于基向量，我们得到

$$\begin{aligned} 0 &= [\omega_{31} \wedge \theta^1 + \omega_{32} \wedge \theta^2](\mathbf{m}_1, \mathbf{m}_2) \\ &= \omega_{31}(\mathbf{m}_1)\theta^1(\mathbf{m}_2) - \omega_{31}(\mathbf{m}_2)\theta^1(\mathbf{m}_1) + \omega_{32}(\mathbf{m}_1)\theta^2(\mathbf{m}_2) - \omega_{32}(\mathbf{m}_2)\theta^2(\mathbf{m}_1) \\ &= \omega_{32}(\mathbf{m}_1) - \omega_{31}(\mathbf{m}_2). \end{aligned}$$

因此，调用联络形式的斜对称，

$$\boxed{\omega_{23}(\mathbf{m}_1) = \omega_{13}(\mathbf{m}_2).}$$

因此，回顾 (38.20)，我们推导

$$\boxed{\text{对称方程等价于形状算子对称的事实：} S^T = S\text{。}}$$

这种对称性的几何解释先前在（15.13）中得出： S 的特征向量是正交的。

如果我们允许自己先前知道这些特征向量实际上是最大和最小法向曲率的主方向，这些曲率是相关的特征值 κ_1 和 κ_2，那么我们已经重新证明了欧拉发现的主要部分：这些最大和最小曲率发生在正交方向。

让我们开辟新天地，细化我们适应标架场，使 $(\mathbf{m}_1, \mathbf{m}_2)$ 在 $\mathcal{S}$ 的每个点与这些主方向对齐； 这称为主标架场（principal frame field）。 在这种情况下，$S(\mathbf{m}_1) = \kappa_1 \mathbf{m}_1$ 和 $S(\mathbf{m}_2) = \kappa_2 \mathbf{m}_2$，所以我们推导出（38.20）的形式为

$$[S] = \begin{bmatrix} \omega_{13}(\mathbf{m}_1) & \omega_{13}(\mathbf{m}_2) \\ \omega_{23}(\mathbf{m}_1) & \omega_{23}(\mathbf{m}_2) \end{bmatrix} = \begin{bmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{bmatrix}.$$

因此，凭借 (38.23)，

$$\boxed{\omega_{13} = \kappa_1\, \theta^1 \quad \text{和} \quad \omega_{23} = \kappa_2\, \theta^2.} \tag{38.27}$$

我们现在可以推导出 Peterson-Mainardi-Codazzi 方程的几何解释——这是我们在之前的幕中从未见过的真正的新数学。将方程(38.27)代入（38.26)的第一个 Peterson-Mainardi-Codazzi 方程，我们发现

$$\mathbf{d}(\kappa_1\,\theta^1) = \omega_{12}\ \wedge \kappa_2\,\theta^2.$$

因此

$$\mathbf{d}\kappa_1\ \wedge \theta^1\ + \kappa_1\,\mathrm{d}\theta^1\ = \kappa_2\,\omega_{12} \wedge \theta^2.$$

使用（38.26）的第一结构方程，即 $\mathrm{d}\theta^1 = \omega_{12} \wedge \theta^2$，我们推导出

$$\mathbf{d}\kappa_1\ \wedge \theta^1\ + \kappa_1\,\omega_{12}\ \wedge \theta^2\ = \kappa_2\,\omega_{12}\ \wedge \theta^2,$$

因此，

$$\mathbf{d}\kappa_1 \wedge \theta^1\ = (\kappa_2\ - \kappa_1)\,\omega_{12} \wedge \theta^2.$$

最后，将这些 2-形式（这些指等式的两侧）应用于基向量 $(\mathbf{m}_1, \mathbf{m}_2)$，我们得到

$$0 - \mathbf{d}\kappa_1(\mathbf{m}_2) = (\kappa_2 - \kappa_1)\omega_{12}(\mathbf{m}_1) - 0.$$

因此，Peterson-Mainardi-Codazzi 方程现在采用几何上有意义的形式，

$$\boxed{\nabla_{\mathbf{m}_2}\kappa 1\ = (\kappa_1\ - \kappa_2)\omega_{12}(\mathbf{m}_1).} \tag{38.28}$$

对另一个 Peterson-Mainardi-Codazzi 方程重复此计算，我们得到

$$\boxed{\nabla_{\mathbf{m}_1}\kappa_2\ = (\kappa_1\ - \kappa_2)\omega_{12}(\mathbf{m}_2).} \tag{38.29}$$

在这两个方程中，$\omega_{12}(\mathbf{v})$ 告诉我们，当我们沿着曲面在方向 $\mathbf{v}$ 上移动时，主方向在 $\mathcal{S}$ 内旋转的速度有多快。因此，将这两个方程变成文字，当我们与主方向成直角移开时，该主曲率的变化率与两个主曲率的差以及主方向旋转的速率成正比。

38.7 高斯方程的几何形式

在六个形式方程（38.26）中，高斯方程无疑是最重要的，因为它既掌握了内蕴曲率的关键，也掌握了它与外在曲率关系的关键。 与后者的联系直接来自（38.25），因为这允许我们

以一种新的、几何上有意义的形式重写高斯方程。

$$\text{高斯方程：}\quad \boxed{d\omega_{12} = -\mathcal{K}_{\text{ext}}\,\theta^1 \wedge \theta^2.} \tag{38.30}$$

例如，考虑在第 38.5.2 节中分析的半径为 R 的球面：

$$\mathbf{d}\omega_{12} = d(c_\phi\,\mathbf{d}\vartheta) = -s_\phi\,\mathbf{d}\phi \wedge \mathbf{d}\vartheta = -\frac{1}{R^2}\,[R\,\mathbf{d}\phi] \wedge [R\,s_\phi\,\mathbf{d}\vartheta] = -\frac{1}{R^2}\,\theta^1 \wedge \theta^2.$$

由此我们立即（正确地！）推导出 $\mathcal{K}_{\text{ext}} = +(1/R^2)$。

38.8 度量曲率公式和绝妙定理的证明

38.8.1引理：$\boldsymbol{\omega_{12}}$ 的唯一性

$$\boxed{\begin{array}{l}\text{联络形式}\omega_{12} = -\omega_{21}\text{是满足(38.26)嘉当第一结构方程的唯一 1-形式：}\\[2ex] \qquad \mathbf{d}\theta^1 = \omega_{12} \wedge \theta^2 \ \text{和}\ \mathbf{d}\theta^2 = -\omega_{12} \wedge \theta^1\text{。}\end{array}} \tag{38.31}$$

为了证明这一点，我们将这些 2-形式应用于基向量对 $(\mathbf{m}_1, \mathbf{m}_2)$，并发现 [练习]

$$\omega_{12}(\mathbf{m}_1) = \mathbf{d}\theta^1(\mathbf{m}_1, \mathbf{m}_2) \text{ 和 } \omega_{12}(\mathbf{m}_2) = \mathbf{d}\theta^2(\mathbf{m}_1, \mathbf{m}_2).$$

从（38.23）可以看出，ω_{12} 由下式唯一且明确地给出

$$\boxed{\omega_{12} = [\mathbf{d}\theta^1(\mathbf{m}_1, \mathbf{m}_2)]\theta^1 + [\mathbf{d}\theta^2(\mathbf{m}_1, \mathbf{m}_2)]\theta^2} \tag{38.32}$$

从而完成 (38.31) 的证明。

38.8.2度量曲率公式的证明

从数学上讲，我们现在生活在 23 世纪——“星际迷航相位器”公式 (4.10) 被传递给我们的未来，很久以前，在第 38 页。

嘉当的形式现在使这个度量曲率公式的计算证明成为可能，它的简洁性令人吃惊，与我们在 [27.4] 中的几何演示相当。

从

$$ds^2 = (Adu)^2 + (B\,dv)^2,$$

我们立即推导出适用于曲面的 1-形式基是

$$\boxed{\theta^1 = A\mathbf{d}u \quad 和 \quad \theta^2 = B\,\mathbf{d}v,} \tag{38.33}$$

并且曲面的面积 2-形式是

$$\boxed{\mathcal{A} = \theta^1 \wedge \theta^2 = AB\,\mathbf{d}u \wedge \mathbf{d}v.}$$

我们现在计算对偶基 1-形式的外导数：

$$\mathbf{d}\theta^1 = \mathbf{d}A \wedge \mathbf{d}u = \partial_v A\,\mathbf{d}v \wedge \mathbf{d}u = -\frac{\partial_v A}{B}\mathbf{d}u \wedge \mathbf{d}\theta^2$$
$$\mathbf{d}\theta^2 = \mathbf{d}B \wedge \mathbf{d}v = \partial_u B\,\mathbf{d}u \wedge \mathbf{d}v = -\frac{\partial_u B}{A}\mathbf{d}v \wedge \mathbf{d}\theta^1$$

将这些方程与嘉当第一结构方程（38.31）进行比较，我们立即看到 ω_{12} 的一个可能解是

$$\boxed{\omega_{12} = -\frac{\partial_v A}{B}\mathbf{d}u + \frac{\partial_u B}{A}\mathbf{d}v.} \tag{38.34}$$

但是引理（38.31）告诉我们，如果这是一个解，那么它就是唯一的解——它就是解！

取外导数，

$$\begin{aligned}\mathbf{d}\omega_{12} &= -\mathbf{d}\left[\frac{\partial_v A}{B}\right] \wedge \mathbf{d}u + d\left[\frac{\partial_u B}{A}\right] \wedge \mathbf{d}v \\ &= -\partial_v[\frac{\partial_v A}{B}]\mathbf{d}v \wedge \mathbf{d}u + \partial_u[\frac{\partial_u B}{A}]\mathbf{d}u \wedge \mathbf{d}v \\ &= \frac{1}{AB}(\partial_v[\frac{\partial_v A}{B}] + Bu[\frac{\partial_u B}{A}])\theta^1 \wedge \theta^2.\end{aligned}$$

将其与高斯方程（38.30）进行比较，我们看到，仅用一页纸，我们的工作就完成了！我们已经（重新）证明了非常不同凡响的公式 (4.10)：

$$\boxed{\mathcal{K}_{\text{ext}} = -\frac{1}{AB}\left(\partial_v\left[\frac{\partial_v A}{B}\right] + \partial_u\left[\frac{\partial_u B}{A}\right]\right).} \tag{38.35}$$

该方程左侧的外在曲率 $\mathcal{K}_{\text{ext}} = \kappa_1\kappa_2$ 来自（38.21）——它测量了法线向量在空间曲面上的扩展。但是这个方程右边的表达式只取决于曲面的内蕴度量几何。因此，这个公式是高斯绝妙定理的一个非常明确和非常纯粹的表达（和证明！）。

这个计算证明的简单性与我们在 [27.4] 中的几何证明相媲美是有原因的——从根本上说，它是相同的证明！

为了开始理解这一点，首先观察内蕴联络形式 ω_{12} 告诉我们 $(\mathbf{m}_1, \mathbf{m}_2)$ 在曲面内的旋转，相对于平行移动的向量。由于这与平行移动向量相对于标架的旋转 $\delta\mathcal{R}$ 相反，这就解释了为什么（38.34）只是我们最初的几何推导公式（27.5）的负值。

几何证明的最后一步是找到一个围绕一个小闭环平行移动的向量的和乐。但是，正如我们在（37.8）中了解到的，ω_{12} 围绕一个小环的积分正是由 $\mathbf{d}\omega_{12}$ 测量的。

38.9 一种新的曲率公式

我们现在推导出一个新的曲率公式，它将使我们能够获得新的结果。随着绝妙定理（重新）建立，我们可以恢复将曲面的曲率表示为 $\mathcal{K}$，因为外在曲率和内蕴曲率实际上是一个，并且相同的。

我们现在将证明

$$\boxed{\mathcal{K} = \nabla_{\mathbf{m}_2}[\omega_{12}(\mathbf{m}_1)] - \nabla_{\mathbf{m}_1}[\omega_{12}(\mathbf{m}_2)] - [\omega_{12}(\mathbf{m}_1)]^2 - [\omega_{12}(\mathbf{m}_2)]^2.} \tag{38.36}$$

为避免混乱，让我们将 ω_{12} 的分量写为 f_1 和 f_2，这样（38.32）的形式为

$$\omega_{12} = f_1\,\theta^1 + f_2\,\theta^2,$$

其中

$$f_1 = \omega_{12}(\mathbf{m}_1) = \mathbf{d}\theta^1(\mathbf{m}_1, \mathbf{m}_2) \quad \text{和} \quad f_2 = \omega_{12}(\mathbf{m}_2) = \mathbf{d}\theta^2(\mathbf{m}_1, \mathbf{m}_2).$$

如果我们将 2-形式 $\mathbf{d}\omega_{12}$ 应用于基向量对 $(\mathbf{m}_1, \mathbf{m}_2)$，则高斯方程（38.30）得出，

$$
\begin{aligned}
\mathcal{K} &= -\mathbf{d}\omega_{12}(\mathbf{m}_1, \mathbf{m}_2) \\
&= -[\mathbf{d}f_1 \wedge \theta^1 + f_1\, d\theta^1 + \mathbf{d}f_2 \wedge \theta^2 + f_2\, \mathbf{d}\theta^2](\mathbf{m}_1, \mathbf{m}_2) \\
&= \mathbf{d}f_1(\mathbf{m}_2) - \mathbf{d}f_2(\mathbf{m}_1) - f_1^2 - f_2^2 \\
&= \nabla_{\mathbf{m}_2} f_1 - \nabla_{\mathbf{m}_1} f_2 - f_1^2 - f_2^2.
\end{aligned}
$$

这确实是所需的公式（38.36）。

38.10 希尔伯特引理

希尔伯特证明，如果曲面上的点 p 具有以下三个属性，则曲面在 p 处的曲率不能为正：

$$
\boxed{\begin{Bmatrix} \kappa_1 \text{ 在 p 处具有局部最大值} \\ \kappa_2 \text{ 在 p 处具有局部最小值} \\ \kappa_1(p) > \kappa_2(p) \end{Bmatrix} \quad \Longrightarrow \quad \mathcal{K}(p) \le 0.} \tag{38.37}
$$

例如，这些条件在环面内侧赤道上的每个点都实现，并且这些点的曲率确实是负的，符合希尔伯特引理。

因为 p 是 κ_1 和 κ_2 的临界点，所以在 p 处我们有

$$
\nabla_{\mathbf{m}_2}\kappa_1 = 0 = \nabla_{\mathbf{m}_1}\kappa_2. \tag{38.38}
$$

此外，由于 κ_1 在 p 处达到局部最大值，而 κ_2 在 p 处达到局部最小值，

$$
\nabla_{\mathbf{m}_2}\nabla_{\mathbf{m}_2}\kappa_1 \le 0 \quad \text{和} \quad \nabla_{\mathbf{m}_1}\nabla_{\mathbf{m}_1}\kappa_2 \ge 0. \tag{38.39}
$$

由于 $(\kappa_1 - \kappa_2) > 0$，将（38.38）代入 Peterson-Mainardi-Codazzi 方程（38.28）和（38.29），我们发现在 p 处，

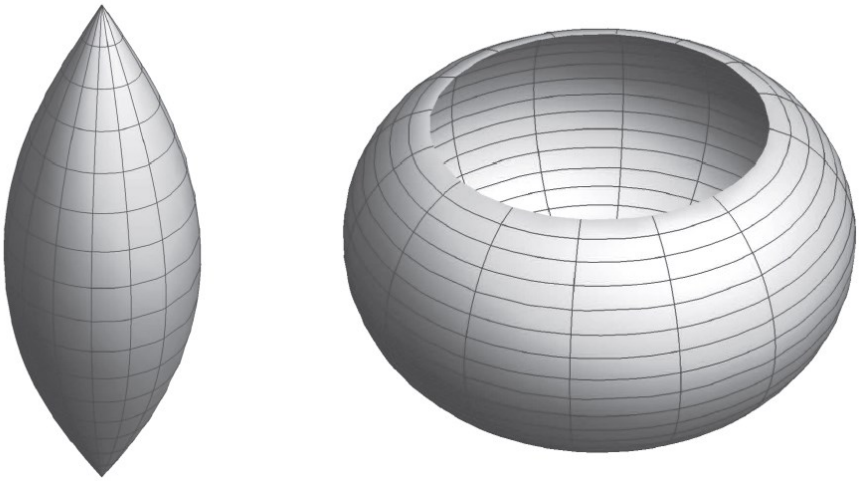
$$
\omega_{12}(\mathbf{m}_1) = 0 = \omega_{12}(\mathbf{m}_2). \tag{38.40}
$$

因此，曲率公式（38.36）简化为

$$
\mathcal{K} = \nabla_{\mathbf{m}_2}[\omega_{12}(\mathbf{m}_1)] - \nabla_{\mathbf{m}_1}[\omega_{12}(\mathbf{m}_2)]. \tag{38.41}
$$

微分（38.28）和插入（38.40），

$$\nabla_{\mathbf{m}_2}\nabla_{\mathbf{m}_2}\kappa 1 = [\nabla_{\mathbf{m}_2}(\kappa_1-\kappa_2)]\omega_{12}(\mathbf{m}_1)+(\kappa_1-\kappa_2)\nabla_{\mathbf{m}_2}[\omega_{12}(\mathbf{m}_1)]$$
$$= (\kappa_1-\kappa_2)\nabla_{\mathbf{m}_2}[\omega_{12}(\mathbf{m}_1)].$$



[38.7] 这两种曲面具有恒定的正曲率，以及与球面相同的内蕴几何，但它们显然不是球面！

最后，由于 $(\kappa_1-\kappa_2)>0$, （38.39）意味着

$$\nabla_{\mathbf{m}_2}[\omega_{12}(\mathbf{m}_1)]\le 0.$$

将相同的逻辑应用于另一个 Peterson-Mainardi-Codazzi 方程 (38.29)，我们发现

$$\nabla_{\mathbf{m}_1}[\omega_{12}(\mathbf{m}_2)]>0.$$

将前两个不等式代入曲率公式（38.41），证明引理。

38.11 利布曼刚性球定理

回想一下 1839 年的明丁（Minding）定理（在 Ex. 7, p. 336 中证明）意味着如果曲面 $\mathcal{S}$ 具有恒定的正曲率 $\mathcal{K}$，那么它的内蕴度量几何必然与半径为 $\frac{1}{\sqrt{\mathcal{K}}}$ 的球面相同，至少局部如此。

更进一步，推测 $\mathcal{S}$ 一定是球面的一部分是很有诱惑力的，体现了球体完全对称的外在几何，其中每个点都是一个脐点：$\kappa_1=\kappa_2$。事实上，从此以后，我们将假定这个条件确实表征了球面。

然而，这样的猜测是不正确的：[38.7] 描述了两种非常肯定不是球面的曲面，但它们都

具有恒定的正曲率，并且具有与球面相同的内蕴几何！ （有关这些曲面的方程式，请参见示例 7，第 336 页。）

我们也可以通过实验看到这一点。拿一个乒乓球，把它切成两半；现在轻轻地将赤道上的对径点挤压在一起。该曲面将在不拉伸的情况下弯曲，从而创建一个具有恒定正曲率的新非球面，与原始球面具有相同的内蕴几何。

但是如果我们尝试弯曲一个完整的乒乓球呢？看来我们做不到！ 明丁（Minding）本人在 1839 年首次推测了这个问题，但又过了 60 年，海因里希・利布曼（Heinrich Liebmann）终于能够证明（在 1899 年）嵌入在 $\mathbb{R}^3$ 中的恒定曲率的闭合（有界）曲面只能是球面。

因此，如果几何是内蕴均匀的（在 $\mathcal{K}$ 恒定的意义上），那么外在几何也只能是外在均匀的球面几何，因此球面不能变形为任何其他外在形状——一个完整的球面是刚性的。

不久之后，大卫・希尔伯特[13]（David Hilbert）发现了利布曼定理最著名的证明，他证明的关键成分是（38.37），因此得名——“希尔伯特引理”。

我们首先观察到，由于 $\mathcal{S}$ 是有界的，所以它的恒定曲率必须是正曲率。根据定义，有界意味着 $\mathcal{S}$ 可以包含在一个足够大的球面中。现在缩小球面，直到它第一次接触到 $\mathcal{S}$，例如在 c 处。然后围绕 c 的小块 $\mathcal{S}$ 都位于 $\mathcal{S}$ 和 c 处的球面的公共切平面的一侧。因此 c 必须是一个正曲率的椭圆点，并且由于曲率被假定为常数，所以它在 $\mathcal{S}$ 上的任何地方都是正的。

让我们假设在封闭（有界）曲面上，函数 κ_1 必然能取到最大值。（可以在 O'Neill (2006, pp. 185 – 186) 中找到证明。）由于假设 $\mathcal{K} = \kappa_1\kappa_2$ 是常数，因此 κ_2 必然在 κ_1 达到最大值的点处达到最小值。但这意味着 $\kappa_1 > \kappa_2$ 是不可能的。因为，如果是这样的话，希尔伯特引理会要求 $\mathcal{K} \leq 0$，这与曲面具有（恒定）正曲率的事实相矛盾。

因此 $\kappa_1(\mathrm{p}) = \kappa_2(\mathrm{p})$。由于 $\kappa_1(\mathrm{p})$ 是 κ_1 在 $\mathcal{S}$ 上的最大值，而 $\kappa_2(\mathrm{p})$ 是 κ_2 在 $\mathcal{S}$ 上的最小值，这意味着在 $\mathcal{S}$ 的每个点上 $\kappa_1 = \kappa_2$，因此 $\mathcal{S}$ 是一个球面。

38.12 n-流形的曲率 2-形式

38.12.1 导言和总结

回想一下我们戏剧第一幕拉开帷幕的场景。 欧几里得平面 $\mathbb{R}^2$ 的平坦性由在其中构建的任意三角形 Δ 的以下性质表征：

[13]利布曼在哥廷根完成了他的博士工作，但不是作为希尔伯特的学生。 然而，也许有人会感觉到希尔伯特有些嫉妒，因为他写道，利布曼“在我的敦促下”证明了这个定理（希尔伯特 1902，第 197 页）。 无论如何，是希尔伯特对利布曼定理的证明（在这里给出）活了下来！

$$\boxed{\Delta\text{的内角和} = \pi.} \tag{38.42}$$

如果不满足这个方程，则意味着 Δ 存在于弯曲的表面内。

这种情况类似于牛顿第一运动定律和第二运动定律的并用：如果我们观察到一个物体不在做匀速直线运动，则一定有一个力作用在它上面。 力的大小表现为加速度，即实际运动与无力、恒速、直线运动之间的差异。

同样，曲面内的曲率量由上述等式两侧之间的差异量化。 这是角盈，$\mathcal{E}$：

$$\boxed{\mathcal{E}(\Delta) \equiv (\Delta\text{的内角和}) - \pi.} \tag{38.43}$$

当应用于一个小的（最终趋于零的）三角形时，这允许我们通过这个基本公式来量化面积 $\mathcal{A}(\Delta)$ 的三角形 Δ 处的曲率 $\mathcal{K}$：

$$\boxed{\mathcal{E}(\Delta) \asymp \mathcal{K}\mathcal{A}(\Delta).} \tag{38.44}$$

现在让我们上升一个维度，到 3-维欧几里得空间 $\mathbb{R}^3$。正如我们将证明的那样，（38.42）的类比结果正是嘉当第二结构方程 （38.12）——这个方程表征了空间的平坦性：

$$\mathbf{d}[\omega] = [\omega] \wedge [\omega].$$

在弯曲的 n-流形中，这个方程不成立，方程两边的差再次量化了空间的曲率。（38.43）的类似物现在是 2-形式的矩阵 $[\Omega]$，称为曲率矩阵（curvature matrix）：

$$[\Omega] \equiv \mathbf{d}[\omega] - [\omega] \wedge [\omega]. \tag{38.45}$$

该矩阵中的各个元素称为[14]

$$\text{曲率 2-形式：}\Omega_{ij} = \mathbf{d}\omega_{ij} - \sum_k \omega_{ik} \wedge \omega_{kj}. \tag{38.46}$$

正是这个（完全一般的）方程（在更高级的作品中）通常被称为嘉当第二结构方程；在 $\mathbb{R}^3$ 中，它简化为我们的原始版本（38.13），因为在这种情况下 $\Omega_{ij} = 0$。

重要的符号说明：文献中存在两种不同的符号，用于联络 1-形式和曲率 2-形式。在我们很快将采用的更常见的表示法中，这两个对象都有一个上升的指标和一个下降的指标。正

[14] 曲率 2-形式存在多种符号。 我们的 Ω_{ij} 可能是最常见的，它最初是嘉当(1927) 自己使用的符号，但 Flanders (1989) 和 Chern 等人(1999)使用了Θ_{ij}，而 Misner、Thorne 和 Wheeler(1973) 使用 R_{ij}。

如我们将要解释的，符号之间的关系是，$\omega^i{}_j \equiv \omega_{ji}$ 和 $\Omega^i{}_j \equiv \Omega_{ji}$。指标顺序的这种颠倒具有在（38.45）和（38.46）中“-”变为“+”的效果。

我们将证明这些曲率 2-形式包含与黎曼张量完全相同的信息，但包装成更紧凑和优雅的形式 [原文如此]。

为方便起见，我们在这里重复黎曼张量 $\mathbf{R}$ 和黎曼曲率算子 $\mathcal{R}$ 的定义，这两者都在第（29.8）中说明：

$$\boxed{\mathbf{R}(\mathbf{u},\mathbf{v};\ \mathbf{w}) \equiv \mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w} = \left\{[\boldsymbol{\nabla}_{\mathbf{u}},\boldsymbol{\nabla}_{\mathbf{v}}] - \boldsymbol{\nabla}_{[\mathbf{u},\mathbf{v}]}\right\}\mathbf{w}} \tag{38.47}$$

回想一下，在第 29 章中，我们实际上从未计算出 n-流形的“星际迷航移相器”公式（27.1）的推广。 也就是说，我们从未找到直接用度量系数明确表示的黎曼张量公式。 有这样一个公式——它既丑陋又复杂，涉及所谓的克里斯托弗符号（Christoffel symbols），它可以在每一本关于黎曼几何的标准教科书中找到——但我们选择不推导出这个公式是有充分理由的：我们可以做得更好！

曲率 2-形式使我们能够比任何正面出击有望达到的更有效和更清晰地计算黎曼张量。公式（38.46）允许我们计算 Ω_{ij}，然后我们简单地读出黎曼张量的分量，因为我们将证明它们只不过是曲率 2-形式的分量！ 那是，

$$\boxed{\Omega_{ij} = R_{ijkl}\,\theta^k \wedge \theta^l.} \tag{38.48}$$

注意：在这里，以及在下文中，我们已经恢复到爱因斯坦求和约定：对匹配的 k 指标和 l 指标的求和已经懂了。 但是，我们假设对不同的基 2-形式求和，否则我们将需要附加 (1/2) 因子。

38.12.2 广义外导数

通过曲率 2-形式计算黎曼张量，如（38.46）中所定义，只需要联络 1-形式的常规外微分。 然而，为了达到公式（38.48），我们需要将 $\mathbf{d}$ 的作用推广到形式（Forms）之外。

不过，首先，让我们为联络 1-形式引入另一种表示法，其中一个指标升高，一个指标降低：$\omega^i{}_j$。 这种符号不仅可以让我们恢复使用爱因斯坦求和约定，更重要的是，它还可以让你轻松过渡到研究 Misner、Thorne 和 Wheeler (1973)、Frankel (2012)、Dray (2015) 和 Tu (2017) 等著作，他们都使用这种表示法。

定义这些 $\omega^i{}_j$ 使得（38.4）现在采用形式

$$\boxed{\nabla_{\mathbf{v}}\mathbf{m}_j = \sum_i \omega_{ji}(\mathbf{v})\mathbf{m}_i \equiv \sum_i \omega^i{}_j(\mathbf{v})\mathbf{m}_i = \omega^i{}_j(\mathbf{v})\mathbf{m}_i,} \tag{38.49}$$

其中最终的等式表示返回到爱因斯坦求和约定——对匹配对或 i 指标（一个上和一个下）的求和已完全理解。

注意指标顺序的反转[15]！

$$\boxed{\omega^i{}_j \equiv \omega_{ji}.} \tag{38.50}$$

现在让我们尝试推广外导数 $\mathbf{d}$，定义它对活动标架场向量 $\mathbf{m}_i$ 的作用。 以此类推

$$\mathbf{d}f(\mathbf{v}) = \nabla_{\mathbf{v}}f,$$

让我们定义，

$$\mathbf{d}\mathbf{m}_j(\mathbf{v}) \equiv \nabla_{\mathbf{v}}\mathbf{m}_j = \omega^i{}_j(\mathbf{v})\mathbf{m}_i.$$

抽象出任意向量 $\mathbf{v}$，

$$\boxed{\mathbf{d}\mathbf{m}_j = \omega^i{}_j\mathbf{m}_i.}$$

这现在允许我们将 $\mathbf{d}$ 应用于一般向量场 $\mathbf{w} = w^j\mathbf{m}_j$。 重命名指标，我们发现

$$\mathbf{d}\mathbf{w} = \mathbf{d}[w^j\mathbf{m}_j] = [\mathbf{d}w^j]\mathbf{m}_j + w^j\,\mathbf{d}\mathbf{m}_j = \mathbf{m}_i(\mathbf{d}w^i + \omega^i{}_j\,w^j).$$

现在让我们再次微分。

等等！——不是说 $\mathbf{d}^2 = 0$吗？！ 不，$\mathbf{d}^2$ 在应用于形式时会消失，但现在不会！ 重命名指标，我们发现，

$$\begin{aligned}\mathbf{d}^2\mathbf{w} &= \mathbf{d}\mathbf{m}_k \wedge (\mathbf{d}w^k + \omega^k{}_j\,w^j) + \mathbf{m}_i(\mathbf{d}^2w^i + w^j\,\mathbf{d}\omega^i{}_j - \omega^i{}_j \wedge \mathbf{d}w^j)\\ &= \omega^i{}_k\mathbf{m}_i \wedge (\mathbf{d}w^k + \omega^k{}_i\,w^i) + \mathbf{m}_i(w^j\,\mathbf{d}\omega^i{}_j - \omega^i{}_k \wedge \mathbf{d}w^k)\\ &= \mathbf{m}_i(\mathbf{d}\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_i)w^j.\end{aligned}$$

因此，

[15]我不明白这个选择的由来，也不欢迎，但它是标准符号，所以我们要遵守它！

$$\boxed{\mathbf{d}^2\mathbf{w} = \mathbf{m}_i \Omega_{ij}\, w^j,} \tag{38.51}$$

其中 Ω_{ij} 又是

曲率 2−形式：
$$\boxed{\Omega^i{}_j = \mathbf{d}\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_i.} \tag{38.52}$$

重要的注解：观察到这个 2-形式的公式与我们的原始定义（38.46）之间存在不幸的符号差异。 这种差异源于（38.50）：

$$\boxed{\Omega_{ij} = \mathbf{d}\omega_{ij} - \sum_k \omega_{ik} \wedge \omega_{kj} = \mathbf{d}\omega_{ji} - \sum_k \omega^k{}_i \wedge \omega^j{}_k = \mathbf{d}\omega^j{}_i + \sum_k \omega^j{}_k \wedge \omega^k{}_i = \Omega^j{}_i.}$$

同样，请注意，曲率矩阵的公式（38.45）中的符号也会发生变化：

$$\boxed{\left[\Omega^i{}_j\right] = \mathbf{d}[\omega] + [\omega] \wedge [\omega].} \tag{38.53}$$

同样，新符号翻转了嘉当第一结构方程中的符号，因此（38.11）现在变为

$$\boxed{\mathbf{d}\theta^i = -\omega^i{}_j \wedge \theta^j.} \tag{38.54}$$

此后，我们将只使用由（38.52）给出的 $\Omega^i{}_j$，因为这是文献中发现的曲率 2-形式最常见的表达式。

38.12.3 从曲率 2-形式中提取黎曼张量

回想一下我们的第一个公式，用于 1-形式的外导数（36.2）：

$$\mathbf{d}\varphi(\mathbf{u},\mathbf{v}) = \nabla_{\mathbf{u}}\varphi(\mathbf{v}) - \nabla_{\mathbf{v}}\varphi(\mathbf{u}) - \varphi([\mathbf{u},\mathbf{v}]).$$

我们现在通过用 $\mathbf{dw}$ 代替 φ 来推广这一点，回想一下

$$\mathbf{dw}(\mathbf{v}) = \nabla_{\mathbf{v}}\mathbf{w}.$$

因此，

$$
\begin{aligned}
\mathbf{d}^2\mathbf{w}(\mathbf{u},\mathbf{v}) &= \nabla_{\mathbf{u}}\mathbf{d}\mathbf{w}(\mathbf{v}) - \nabla_{\mathbf{v}}\mathbf{d}\mathbf{w}(\mathbf{u}) - \mathbf{d}\mathbf{w}([\mathbf{u},\mathbf{v}]) \\
&= \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\mathbf{w} - \nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\mathbf{w} - \nabla_{[\mathbf{u},\mathbf{v}]}\mathbf{w} \\
&= \mathcal{R}(\mathbf{u},\mathbf{v})\mathbf{w},
\end{aligned}
$$

凭借（38.47）。

将其与（38.52）相结合，我们得到了依据曲率 2-形式得到黎曼张量的非凡公式：

$$
\boxed{R^i{}_{jkl}\,\theta^k \wedge \theta^l = \Omega^i{}_j = \mathbf{d}\omega^i{}_j + \omega^i{}_m \wedge \omega^m{}_j.} \tag{38.55}
$$

我们通过广义外导数的梯子得到了这个公式，但是，已经到了应许之地，我们现在可以把梯子扔到一边了！ 对于这个最终公式（38.55）仅使用应用于联络 1-形式的普通外导数。

因此，我们现在拥有了一种使用曲率 2-形式计算黎曼张量的优雅且异常强大的方法。我们将通过将这种方法应用于具有巨大物理重要性的特定示例来结束我们的戏剧。

38.12.4 重温比安奇恒等式

使用我们闪亮的新曲率 2-形式，第一（代数）比安奇恒等式（29.15）和第二（微分）比安奇恒等式（29.17）都可以非常优雅地导出和紧凑地表示。

由于基 1-形式 θ^i 是普通的 1-形式，$\mathbf{d}^2$ 湮灭它们，产生等于零的 3-形式。 因此，对嘉当第一结构方程（38.54）进行微分，

$$
\begin{aligned}
0 &= -\mathbf{d}\mathbf{d}\theta^i \\
&= \mathbf{d}\left(\omega^i{}_j \wedge \theta^j\right) \\
&= \mathbf{d}\omega^i{}_j \wedge \theta^j - \omega^i{}_k \wedge \mathbf{d}\theta^k \\
&= \mathbf{d}\omega^i{}_j \wedge \theta^j + \omega^i{}_k \wedge \omega^k{}_j \wedge \theta^j \\
&= (\mathbf{d}\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_j) \wedge \theta^j \\
&= \Omega^i{}_j \wedge \theta^j.
\end{aligned}
$$

因此，采用更紧凑的矩阵表示法，我们得到了

$$\boxed{\text{第一(代数)比安奇恒等式：}[\Omega]\wedge[\theta]=0.} \tag{38.56}$$

在目前的形式中，这肯定无法识别为原始的第一（代数）比安奇恒等式（29.15）！ 为了看到这两个恒等式实际上是一回事，我们调用（38.55）：

$$0=\Omega^{i}{}_{j}\wedge\theta^{j}=R^{i}{}_{jkl}\,\theta^{j}\wedge\theta^{k}\wedge\theta^{l},$$

这确实意味着与（29.15）等价 [练习]：

$$R^{i}{}_{jkl}+R^{i}{}_{klj}+R^{i}{}_{ljk}=0.$$

第二（微分）比安奇恒等式（29.17）可以类似地通过将 $\mathbf{d}^2$ 应用于联络 1-形式来获得。因此，微分嘉当第二结构方程，（38.52），

$$\begin{aligned}
0 &= \mathbf{d}\mathbf{d}\omega^{i}{}_{j}\\
&= \mathbf{d}(\Omega^{i}{}_{j}-\omega^{i}{}_{k}\wedge\omega^{k}{}_{j})\\
&= \mathbf{d}\Omega^{i}{}_{j}-\mathbf{d}\omega^{i}{}_{k}\wedge\omega^{k}{}_{j}+\omega^{i}{}_{k}\wedge\mathbf{d}\omega^{k}{}_{j}\\
&= \mathbf{d}\Omega^{i}{}_{j}-(\Omega^{i}{}_{k}-\omega^{i}{}_{m}\wedge\omega^{m}{}_{k}k)\wedge\omega^{k}{}_{j}+\omega^{i}{}_{k}\wedge(\Omega^{k}{}_{j}-\omega^{k}{}_{m}\wedge\omega^{m}{}_{j})\\
&= \mathbf{d}\Omega^{i}{}_{j}-\Omega^{i}{}_{k}\wedge\omega^{k}{}_{j}+\omega^{i}{}_{k}\wedge\Omega^{k}{}_{j}.
\end{aligned}$$

因此，采用更紧凑的矩阵表示法，我们得到了

$$\boxed{\text{第二(微分)比安奇恒等式：}\mathbf{d}[\Omega]=[\Omega]\wedge[\omega]-[\omega]\wedge[\Omega].} \tag{38.57}$$

同样，（38.55）给出了与黎曼张量的显式链接，使人们能够看到这确实与我们最初版本的微分比安奇恒等式（29.17）等价[练习]。

38.13 史瓦西黑洞的曲率

本节拉开第五幕的帷幕，结束我们漫长的戏剧。 它通过将上述想法应用于具有宇宙重要性的数学和物理对象：黑洞的时空曲率！

具体来说，我们将努力验证由[16]（38.58）给出的史瓦西黑洞几何（我们在此重复），

[16]正如我们在麦克斯韦方程的上下文中解释的那样，我们现在已经颠倒了度量系数的符号，以便使我们的工作与我们在本书末尾的进一步阅读部分中推荐的大多数广义相对论课本保持一致 .

史瓦西黑洞

$$ds^2 = -(1-\frac{2GM}{r})dt^2 + \frac{dr^2}{(1-\frac{2GM}{r})} + r^2(d\phi^2 + \sin^2\phi d\vartheta^2), \tag{38.58}$$

确实是爱因斯坦真空场方程（30.11）的解（我们也在此重复）：

$$\text{爱因斯坦真空场方程：}\mathbf{R}\text{icci} = 0 \quad \Leftrightarrow \quad R_{ik} = 0\text{。} \tag{38.59}$$

让我们定义[17]

$$f(r) = 1 - \frac{2GM}{r},$$

这样（供以后使用），

$$f' = \frac{2GM}{r^2}, \qquad \text{和 } f'' = -\frac{4GM}{r^3}, \qquad \text{和 } 1 - f = \frac{2GM}{r}. \tag{38.60}$$

让我们用其相关的时空坐标标记每个基 1-形式。 那么史瓦西度量（38.58）可以写成：

$$\mathbf{g} = g_{ij}(\theta^i \otimes \theta^j) = -\theta^t \otimes \theta^t + \theta^r \otimes \theta^r + \theta^\phi \otimes \theta^\phi + \theta^\vartheta \otimes \theta^\vartheta,$$

其中

$$\theta^t = \sqrt{f}\,\mathbf{d}t, \qquad \theta^r = \frac{\mathbf{d}r}{\sqrt{f}}, \qquad \theta^\phi = r\,\mathbf{d}\phi, \qquad \theta^\vartheta = r\sin\phi\mathbf{d}\vartheta.$$

在每个方程的左边写下第一个结构方程——在 m 上用爱因斯坦求和——然后实际求值右边的每个导数，最后用基 1-形式重写每个结果，我们发现

$$\omega^t{}_m \wedge \theta^m = -\mathbf{d}\theta^t = -\frac{f'}{2\sqrt{f}}\,\mathbf{d}r \wedge \mathbf{d}t = \frac{f'}{2\sqrt{f}}\sqrt{f}\,\mathbf{d}r \wedge \frac{\mathbf{d}r}{\sqrt{f}} = \frac{f'}{2\sqrt{f}}\theta^t \wedge \theta^r,$$

[17]在下文中，我们主要遵循 Dray (2015, pp. 259–261) 的符号，尽管他将基 1-形式写为 σ^i，而我们将继续将它们写为 θ^i。 另请注意，我们的 ϕ 和 ϑ 与他的相反！ 请参见第 388 页的脚注。

$$\omega^{r}{}_{m}\wedge\theta^{m}=-\mathbf{d}\theta^{r}=-\mathbf{d}\left[\frac{1}{\sqrt{f}}\right]\wedge\mathbf{d}r=-\partial_{r}\left[\frac{1}{\sqrt{f}}\right]\mathbf{d}r\wedge\mathbf{d}r=0,$$

$$\omega^{\phi}{}_{m}\wedge\theta^{m}=-\mathbf{d}\theta^{\phi}=-\mathbf{d}r\wedge\mathbf{d}\phi=\frac{\sqrt{f}}{r}r\mathbf{d}\phi\wedge\frac{\mathbf{d}r}{\sqrt{f}}=\frac{\sqrt{f}}{r}\theta^{\phi}\wedge\theta^{r}$$

$$\omega^{\vartheta}{}_{m}\wedge\theta^{m}=-\mathbf{d}\theta^{\vartheta}=-s_{\phi}\mathbf{d}r\wedge\mathbf{d}\vartheta-rc_{\phi}\mathbf{d}\phi\wedge\mathbf{d}\vartheta=\frac{\sqrt{f}}{r}\theta^{\vartheta}\wedge\theta^{r}+\frac{\cot\phi}{r}\theta^{\vartheta}\wedge\theta^{\phi}$$

代替更著名的奥卡姆剃刀，让我们援引（并同时命名[18]）逻辑上站不住脚但非常有效的[19]

> 嘉当剃刀：如果一项没有被明确禁止不能为零，那么它就会为零。

检查上面的方程，只写下明确强加给我们的非零项，并假设所有其他项都为零（如嘉当剃刀所规定的那样！），我们发现，

$$\omega^{t}{}_{r}=\frac{f'}{2\sqrt{f}}\theta^{t},\qquad \omega^{\phi}{}_{r}=\frac{\sqrt{f}}{r}\theta^{\phi},\qquad \omega^{\vartheta}{}_{r}=\frac{\sqrt{f}}{r}\theta^{\vartheta},\qquad \omega^{\vartheta}{}_{\phi}=\frac{\cot\phi}{r}\theta^{\vartheta}.$$

唯一性引理（38.31）很容易推广，所以这些就是解！

接下来，我们必须对这些联络 1-形式求微分以获得曲率 2-形式，记住它们的反对称性和度量系数 $g_{tt}=-1$，一起意味着：

$$\omega^{r}{}_{t}=\omega^{t}{}_{r},\qquad \omega^{r}{}_{\phi}=-\omega^{\phi}{}_{r},\qquad \omega^{r}{}_{\vartheta}=-\omega^{\vartheta}{}_{r},\qquad \omega^{\phi}{}_{\vartheta}=-\omega^{\vartheta}{}_{\phi},$$

和

$$\omega^{t}{}_{t}=\omega^{r}{}_{r}=\omega^{\phi}{}_{\phi}=\omega^{\vartheta}{}_{\vartheta}=0.$$

有六个独立的曲率 2-形式，我们现在（详细）说明其中第一个的计算，剩下的五个由你来验证。

有六个独立的曲率 2-形式，我们现在（详细）说明其中第一个的计算，剩下的五个由您来验证。

计算 $\mathbf{d}\omega^{i}{}_{j}$ 的第一步是回到用 $\mathbf{d}x^{i}$ 来表示 $\omega^{i}{}_{j}$；这使得计算外导数变得容易。最终结果然后用 $\theta^{i}\wedge\theta^{j}$ 表示。从（38.52），

---

[18] 需要明确的是，嘉当本人从未说过任何这样的话（据我所知）——尽管如此，我还是选择以他的名字命名这个原则！

[19] 毫无疑问，存在反例，但我个人不记得曾经见过任何一个。

$$
\begin{aligned}
\Omega^t{}_r &= \mathbf{d}\omega^t{}_r + \omega^t{}_m \wedge \omega^m{}_r \\
&= \mathbf{d}[\frac{f'}{2}] \wedge \mathbf{d}t + \omega^t{}_t \wedge \omega^t{}_r + \omega^t{}_r \wedge \omega^r{}_r + \omega^t{}_\phi \wedge \omega^\phi{}_r + \omega^t{}_\vartheta \wedge \omega^\vartheta{}_r \\
&= \frac{f''}{2}\mathbf{d}r \wedge \mathbf{d}t + 0 + 0 + 0 + 0 \\
&= -\frac{f''}{2}\,\theta^t \wedge \theta^r.
\end{aligned}
$$

我们现在列出所有六种曲率 2-形式，从我们刚刚计算的那一种开始：

$$
\begin{aligned}
\Omega^t{}_r &= -\frac{f''}{2}\,\theta^t \wedge \theta^r, \\
\Omega^t{}_\phi &= -\frac{f'}{2r}\theta^t \wedge \theta^\phi, \\
\Omega^t{}_\vartheta &= -\frac{f'}{2r}\theta^t \wedge \theta^\vartheta, \\
\Omega^\phi{}_r &= -\frac{f'}{2r}\theta^\phi \wedge \theta^r, \\
\Omega^\vartheta{}_r &= -\frac{f'}{2r}\theta^\vartheta \wedge \theta^r, \\
\Omega^\vartheta{}_\phi &= [\frac{1-f}{r^2}]\theta^\vartheta \wedge \theta^\phi.
\end{aligned}
$$

同样，纯空间的曲率 2-形式是反对称的，但具有时间分量的曲率 2-形式是对称的，因为 $g_{tt} = -1$：

$$
\Omega^\vartheta{}_r = -\Omega^r{}_\vartheta, \quad \Omega^\phi{}_r = -\Omega^r{}_\phi, \quad \Omega^\vartheta{}_\phi = -\Omega^\phi{}_\vartheta,
$$

和

$$
\Omega^t{}_r = \Omega^r{}_t, \quad \Omega^\phi{}_t = \Omega^t{}_\phi, \quad \Omega^\vartheta{}_t = \Omega^t{}_\vartheta.
$$

将上面的 $\Omega^i{}_j$ 公式与（38.55）进行比较，即 $\Omega^i{}_j = R^i{}_{jkl}\,\theta^k \wedge \theta^l$，并诉诸（38.60），我们可以立即读出黎曼张量的分量：

$$
\boxed{
\begin{aligned}
&R^t{}_{rtr} = -\frac{f''}{2} = +\frac{2GM}{r^3} \\
&R^t{}_{\phi t\phi} = R^t{}_{\vartheta t\vartheta} = R^\phi{}_{r\phi r} = R^\vartheta{}_{r\vartheta r} = -\frac{f'}{2r} = -\frac{GM}{r^3} \\
&R^\vartheta{}_{\phi\vartheta\phi} = [\frac{1-f}{r^2}] = +\frac{2GM}{r^3}.
\end{aligned}
}
$$

虽然 r 坐标的含义不再像“径向距离”那么简单，但请注意，这些公式与我们从由牛顿平方反比万有引力定律引起的潮汐力推导出来的公式（30.2）和（30.3）完全一致！

我们已经到了戏剧性的时刻：这种假定的黑洞时空几何在物理上是否可能——它是否满足爱因斯坦真空场方程？ 让我们一探究竟！

将相关的黎曼张量分量相加以找到里奇（Ricci）张量，我们很容易看到里奇（Ricci）张量的非对角线分量为零，并且使用（38.60），我们现在计算其对角线分量，如下所示：

$$R_{tt} = R^{r}{}_{trt} + R^{\phi}{}_{t\phi t} + R^{\vartheta}{}_{t\vartheta t} = +\frac{f''}{2} + \frac{f'}{2r} + \frac{f'}{2r} = -\frac{2GM}{r^3} + \frac{GM}{r^3} + \frac{GM}{r^3} = 0,$$

$$R_{rr} = R^{t}{}_{rtr} + R^{\phi}{}_{r\phi t} + R^{\vartheta}{}_{r\vartheta r} = -\frac{f''}{2} - \frac{f'}{2r} - \frac{f'}{2r} = +\frac{2GM}{r^3} - \frac{GM}{r^3} - \frac{GM}{r^3} = 0,$$

$$R_{\phi\phi} = R^{t}{}_{\phi t\phi} + R^{r}{}_{\phi r\phi} + R^{\vartheta}{}_{\phi\vartheta\phi} = -\frac{f'}{2r} - \frac{f'}{2r} + \frac{1-f}{r^2} = -\frac{GM}{r^3} - \frac{GM}{r^3} + \frac{2GM}{r^3} = 0,$$

$$R_{\vartheta\vartheta} = R^{t}{}_{\vartheta t\vartheta} + R^{r}{}_{\vartheta r\vartheta} + R^{\phi}{}_{\vartheta\phi\vartheta} = -\frac{f'}{2r} - \frac{f'}{2r} + \frac{1-f}{r^2} = -\frac{GM}{r^3} - \frac{GM}{r^3} + \frac{2GM}{r^3} = 0.$$

因此，史瓦西中尉于 1915 年 12 月发现的黑洞几何，写在第一次世界大战他在俄罗斯前线的战壕中，确实是爱因斯坦一个月前发现的方程的精确解！

需要明确的是，施瓦西当时并不知道他的发现代表了一个黑洞。对他和爱因斯坦来说，它只代表了普通天体（如地球或太阳）周围的时空几何。又过了半个世纪，物理学家才最终领悟到，这个解在整体上也呈现了黑洞的纯真空引力场。

1965 年 1 月，爱因斯坦去世 10 年后，罗杰·彭罗斯第一次证明了爱因斯坦的理论在数学上暗示，足够大的物体的坍缩必然导致黑洞的形成[20]。由于这一发现，彭罗斯获得了 2020 年诺贝尔物理学奖的一半；另一半由莱因哈德·根泽尔（Reinhard Genzel）和安德里亚·盖兹（Andrea Ghez）平分，因为他们独立实验发现了我们银河系中心的超大质量黑洞（人马座 A*）。

在我们漫长的数学戏剧终于结束，大幕落下之前，我们还有最后一招。为什么我们坚持用 f、f′和 f″ 来表示曲率，尽管我们从一开始就知道它们的值，并在（38.60）中得到充分展示？

我们可以从爱因斯坦方程中推导出来史瓦西几何，而不是“仅仅”验证它的确是爱因斯

[20]事实上，彭罗斯证明的是，坍缩必然导致时空奇点的形成。 但是，原则上，这个奇点可能是比黑洞更可怕的东西——它可能是所谓的裸奇点，对外界可见！ 1969 年，彭罗斯推测这永远不会发生，并且“宇宙审查员”会坚持认为每个奇点都被体面地包裹在事件视界内，从而形成黑洞！ 许多理论论据证明了这一宇宙审查假说，但它在 2020 年仍未得到证实。

坦方程的解！

爱因斯坦的场方程意味着

$$\boxed{R_{\phi\phi} = R_{\vartheta\vartheta} = -\frac{f'}{r} + \frac{1-f}{r^2} = 0 \implies \frac{df}{dr} = \frac{1-f}{r} \implies f(r) = 1 - \frac{C}{r},}$$

其中 C 是一个常数。

但随后里奇张量的剩余对角线分量也自动为零，从而证实我们已经成功求解了爱因斯坦方程：

$$\boxed{R_{tt} = -R_{rr} = \frac{f''}{2} + \frac{f'}{r} = -\frac{C}{r^3} + \frac{C}{r^3} = 0.}$$

仍有待理解为什么常数 C 通过 $C = 2GM$ 与黑洞的质量[21]相关联。 但要解释这一点，我们需要比这部数学戏剧所允许的程度大大深入研究迷人的引力物理学世界。

[21] 这种联系可以通过各种论证来建立，我们在本书末尾的进一步阅读部分推荐的所有广义相对论文本中都有这种联系。 然而，也许可以在 Schutz (2003, §§18, 21) 中找到对该主题的最佳、最物理处理。

第三十九章 第五幕练习

1-形式

1. **作为 1-形式的狄拉克 Delta 函数**。（此示例改编自 Schutz 1980。）考虑定义在区间 $-1 \le x \le +1$ 上的无限可微实值函数 $f(x)$ 的向量空间 $C[-1,+1]$。 1-形式的对偶空间称为分布（distributions）。这种分布的一个例子是狄拉克三角函数 $\delta(x)$，它是 1-形式，它对“向量”$f(x)$ 的作用在 $x=0$ 处产生它的值：

$$\langle \delta(x), f(x) \rangle = f(0).$$

正如我们现在所解释的，这根本不是保罗·狄拉克（1930）自己最初将他的 delta 函数引入量子力学的方式。

(i) 对于 $C[-1,+1]$ 中的任何函数 $g(x)$，定义一个对应的 1-形式 $\tilde{g}(x)$，使其对向量 $f(x)$ 的作用由下式给出

$$\langle \tilde{g}(x), f(x) \rangle \equiv \int_{-1}^{+1} g(x)\, f(x)\, dx.$$

证明 $\tilde{g}$ 确实是一个 1-形式。

(ii) 因此，狄拉克实际上定义了他的 delta 函数，他坚持认为

$$\int_{-1}^{+1} \delta(x)\, f(x)\, dx = f(0). \tag{39.1}$$

没有普通函数 $\delta(x)$ 可以做到这一点，但是考虑一个以 0 为中心的平滑的钟形曲线，并想象它随着它变得越来越窄而越来越高，同时始终服从 (39.1)，其中 $f(x) = 1$:

$$\int_{-1}^{+1} \delta(x)\, 1\, dx = 1.$$

解释为什么在钟形曲线宽度趋于零的极限下，我们恢复（39.1）。

2. 在 $\mathbb{R}^3$ 中，给出向量 $\mathbf{v}$ 上 1-形式 $\omega = 2\,\mathbf{d}x + \mathbf{d}y + 2\,\mathbf{d}z$ 的作用 $\omega(\mathbf{v})$ 的显式几何解释。

3. **协变和逆变的解释**。 令 R_θ 表示 $\mathbb{R}^2$ 旋转 θ。 然后，如第 153 页 [15.3] 中所

述，其矩阵为

$$[R_\theta] = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}, \text{逆矩阵} \quad [R_\theta]^{-1} = [R_{-\theta}] = [R_\theta]^T = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

其中 $c \equiv \cos\theta$ 和 $s \equiv \sin\theta$。令 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 为标准笛卡尔基，令 $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2\}$ 为将旧基旋转 θ 得到的新基，因此 $\tilde{\mathbf{e}}_j = R_\theta \mathbf{e}_j$。

(i) 在同一张图中画两个基，也画一个通用向量 $\mathbf{v} = v^j \mathbf{e}_j$，在旋转基上画出它的分量 $\tilde{v}^j$：$\mathbf{v} = \tilde{v}^j \tilde{\mathbf{e}}_j$。使用你的图片来解释为什么通过将相反的旋转矩阵 $[R_\theta]^{-1}$ 应用于原始分量来获得新分量：

$$\begin{bmatrix} \tilde{v}^1 \\ \tilde{v}^2 \end{bmatrix} = [R_\theta] \begin{bmatrix} v^1 \\ v^2 \end{bmatrix}$$

因为分量以与基向量相反（“相反”）的方式变化，所以在较早的文献中，向量被称为逆变。

(ii) $\{\mathbf{e}_1, \mathbf{e}_2\}$ 的对偶基是 $\{dx1, dx2\}$。设 $\varphi = \varphi_j \, \mathbf{d}x^j$ 为一般 1-形式，设 $\{\tilde{\omega}^1, \tilde{\omega}^2\}$ 为 $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2\}$ 的对偶基，因此新分量由 $\varphi = \tilde{\varphi}_j \tilde{\omega}^j$ 给出。证明 1-形式变换的分量与基向量的方式相同（“co-”）：

$$\begin{bmatrix} \tilde{\varphi}_1 \\ \tilde{\varphi}_2 \end{bmatrix} = [R_\theta] \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix}$$

出于这个原因，1-形式曾经被称为协变向量或协向量。（注意：现代微分几何关注的是（坐标无关的）几何对象，这些对象不会被诸如此类的基变换所改变。因此，术语逆变和协变在现代文献中已基本消失。对于一个优秀的、完全通用的有关该主题的讨论，请参见 Schutz (1980, §2.26)。）

张量

4. **矩阵乘法**。来自舒茨（1980）。证明在矩阵代数中，一个矩阵对第二个矩阵的作用，以产生它们的矩阵乘积，是一个 $\begin{Bmatrix} 2 \\ 2 \end{Bmatrix}$ 价张量。

（提示：回想一下，每个矩阵本身就是一个 $\begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$ 价张量；参见第 33.2 节。）

5. **张量缩并**。（摘自 Schutz (1980)。）有多少个不同的 $\begin{Bmatrix} 2 \\ 1 \end{Bmatrix}$ 价张量可以通过对 $\begin{Bmatrix} 3 \\ 2 \end{Bmatrix}$ 价张量 $Q^{ijk}{}_{lm}$？第二次缩并可以产生多少个 $\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$ 价张量（即向量）？

6. **矩阵乘法作为缩并的张量视图**。回到第 33.2 节和第 33.6 节，令 A 为线性变换，表示为 $\begin{Bmatrix}1\\1\end{Bmatrix}$ 价的向量值张量：$\mathbf{v} \to A(\mathbf{v})$。令 B 为第二个线性变换，令 C 为它们的组合：$C(\mathbf{v}) \equiv B[A(\mathbf{v})]$。证明 $\begin{Bmatrix}1\\1\end{Bmatrix}$ 张量 C 的分量由缩并（也称为矩阵乘法）给出：$C^i{}_j = B^i{}_k A^k{}_j$。

2-形式

7. $\mathbb{R}^3$**中因式分解**。在 $\mathbb{R}^3$ 中，令 Ψ 是一般的 2-形式。证明 Ψ 总是可以分解为两个 1-形式的楔积：$\Psi = \alpha \wedge \beta$。

8. $\mathbb{R}^4$**中无需因式分解**。在 $\mathbb{R}^4$ 中，令Ψ是一般的 2-形式。
   (i) 证明 Ψ 不能总是被分解为两个 1-形式的楔积。提示：如果可以分解，考虑 $\Psi \wedge \Psi$。但现在回头看 § 35.7。
   (ii) 如果 $\Psi \wedge \Psi = 0$，证明 Ψ 可以分解为两个 1-形式的楔积。
   (iii) 证明 Ψ 总是可以表示为两个楔积之和：

$$\Psi = \alpha \wedge \beta + \gamma \wedge \delta.$$

9. **曲面 $f = \text{const}$ 的面积公式**。令 $\mathcal{S}$ 为方程 $f(x,y,z) = \text{const}$ 的曲面。

   (i) 证明 $\mathcal{S}$ 的单位法线 $\mathbf{n}$ 是

$$\mathbf{n} = \frac{1}{\sqrt{(\partial_x f)^2 + \left(\partial_y f\right)^2 + (\partial_z f)^2}} \begin{bmatrix}\partial_x f\\ \partial_y f\\ \partial_z f\end{bmatrix}.$$

   (ii) 通过将 $\mathbf{n}$ 视为以单位速度垂直流过 $\mathcal{S}$ 的流体的速度，推导出 $\mathcal{S}$ 上的面积 2-形式为

$$\mathcal{A} = \frac{\partial_x f\, \mathbf{d}y \wedge \mathbf{d}z + \partial_y f\, \mathbf{d}z \wedge \mathbf{d}x + \partial_z f\, \mathbf{d}x \wedge \mathbf{d}y}{\sqrt{(\partial_x f)^2 + \left(\partial_y f\right)^2 + (\partial_z f)^2}}.$$

3-形式

10. $\mathbb{R}^4$中因式分解。证明 $\mathbb{R}^4$ 中的任何 3-形式都可以表示为三个 1-形式的楔积。

微分

11. **闭的和恰当的形式**。 在 $\mathbb{R}^3$ 中，让

$$\varphi = 2x\,\mathbf{d}x + 2y\,\mathbf{d}y + 2z\,\mathbf{d}z,$$

$$\chi = xy\,\mathbf{d}z,$$

$$\Psi = x\,\mathbf{d}y \wedge \mathbf{d}z + y\,\mathbf{d}x \wedge \mathbf{d}z.$$

(i) 验证 φ 是闭的，并通过证明它是恰当的来证明这一点： $\varphi = \mathbf{d}r^2$，其中 r 是到原点的距离。

(ii) 验证 Ψ 是闭的，并通过证明它是恰当的来解释这一点： $\Psi = \mathbf{d}\chi$。

12. **闭的和恰当的形式**。 继续练习 11 的形式，直接计算以下每个外导数，(A) 直接地； (B) 使用莱布尼茨乘积法则，（36.9）。

(i) $\mathbf{d}(\chi \wedge \chi)$.

(ii) $\mathbf{d}(\varphi \wedge \chi)$.

(iii) $\mathbf{d}(\varphi \wedge \Psi)$.

(iv) $\mathbf{d}(\chi \wedge \Psi)$.

13. **闭的和恰当的形式**。 如果你之前没做过，请证明以下每个事实：

(i) 如果 Υ 和 Φ 是合的，那么 $\Upsilon \wedge \Phi$ 也是闭合的。

(ii) 如果 Υ 是闭的，那么 $\Upsilon \wedge \mathbf{d}\Phi$ 对于所有 Φ 都是闭的。

(iii) 如果 $\deg\Phi$ 是偶数，则 $\Phi \wedge \mathbf{d}\Phi$ 是闭的。 提示：见（35.4）。

14. **途径形式的向量分析恒等式**。 使用形式导出向量分析的以下恒等式：

(i) $\nabla \cdot \underline{\varphi} \times \nabla f = \nabla \times \underline{\varphi} \cdot \nabla f$.

(ii) $(\nabla \cdot \underline{\varphi} \times \underline{\psi} = \nabla \times \underline{\varphi} \cdot \underline{\psi} - \underline{\varphi} \cdot \nabla \times \underline{\psi}$.

（提示：（36.9）。）请注意，（i）现在可以导出为（ii）的特例。

15. **霍奇星对偶算子** $(\star)$（有关此概念的全面讨论，请参阅 Schutz (1980)、Baez 和 Muniain (1994) 或 Dray (2015)。特别是 Baez 和 Muniain (1994, §§1, 5) ）包含对霍奇对偶性与麦克斯韦方程的关系的精彩讨论。）

(i) 在 n 维中，证明 p-形式的空间与 (n-p)-形式的空间具有相同的维度。 因此，这两个空间可以进行一一对应，而霍奇对偶是实现这一点的一种特定方法。

(ii) 在 $\mathbb{R}^3$ 中，令 $\mathcal{V}_3 = \mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}z$ 为体积 3-形式。给定一个基 p-形式 σ，我们将 $\star$ 定义为填充 $\mathcal{V}_3$ 的“缺失部分”的线性算子：

$$\boxed{\sigma \wedge \star\sigma = \mathcal{V}_3.}$$

推导

$$\star\,\mathbf{d}\mathrm{x} \;=\; \mathbf{d}\mathrm{y}\wedge\mathbf{d}\mathrm{z},$$

$$\star\,\mathbf{d}\mathrm{y} \;=\; \mathbf{d}\mathrm{z}\wedge\mathbf{d}\mathrm{x},$$

$$\star\,\mathbf{d}\mathrm{z} \;=\; \mathbf{d}\mathrm{x}\wedge\mathbf{d}\mathrm{y},$$

并且对称地推导出

$$\begin{aligned}\star\,(\mathbf{d}\mathrm{y}\wedge\mathbf{d}\mathrm{z}) \;&=\; \mathbf{d}\mathrm{x},\\ \star\,(\mathbf{d}\mathrm{z}\wedge\mathbf{d}\mathrm{x}) \;&=\; \mathbf{d}\mathrm{y},\\ \star\,(\mathbf{d}\mathrm{x}\wedge\mathbf{d}\mathrm{y}) \;&=\; \mathbf{d}\mathrm{z}.\end{aligned}$$

因此，在这种情况下，霍奇星算子的两个应用产生恒等式：$\star\star = 1$。 最后，为了完整起见，自然定义

$$\star\,1 = \mathcal{V}_3 \quad 和 \quad \star\,\mathcal{V}_3 \;= 1.$$

(iii) 继续在 $\mathbb{R}^3$ 中，我们提醒你我们之前的符号：$\underline{\varphi}$ 是对应于 $(\rightleftarrows)$ 1-形式 φ 的向量。 证明霍奇对偶产生两种形式的向量（叉）积：

$$\underline{\alpha}\times\underline{\beta} \;=\underline{\gamma} \quad\Leftrightarrow\quad \alpha\wedge\beta \;=\star\,\gamma \quad\Leftrightarrow\quad \star\,(\alpha\wedge\beta) = \gamma.$$

(iv) 仍然在 $\mathbb{R}^3$ 中，回忆每个 2-形式的 Ψ 可以被认为是与速度流 $\underline{\Psi}$ 相关的通量，通过（34.10）：

$$\Psi = \Psi^1(\mathbf{d}\mathrm{x}^2\wedge\mathbf{d}\mathrm{x}^3) + \Psi^2(\mathbf{d}\mathrm{x}^3\wedge\mathbf{d}\mathrm{x}^1) + \Psi^3(\mathbf{d}\mathrm{x}^1\wedge\mathbf{d}\mathrm{x}^2) \quad\rightleftarrows\quad \underline{\Psi} = \begin{bmatrix}\Psi^1\\ \Psi^2\\ \Psi^3\end{bmatrix}.$$

如果 ψ 是对应于 Ψ^3 的 1-形式，则推论 Ψ 和 ψ 之间的对应是霍奇对偶：

$$\star\,\Psi = \psi \quad 和 \quad \Psi =\star\,\psi。$$

(v) 现在考虑闵可夫斯基时空，配备度量

$$\mathrm{ds}^2 = \mathrm{g}_{\mathrm{ij}}\,\mathrm{dx}^{\mathrm{i}}\,\mathrm{dx}^{\mathrm{j}} = -\mathrm{dt}^2 \;+\mathrm{dx}^2 \;+\mathrm{dy}^2 \;+\mathrm{dz}^2.$$

注意：这里和以后的度量系数 g_{ij} 是之前第二幕使用的那些的负数，因为这是我们将要参考的大多数更高级的物理课本所青睐的约定，继续前进。

令体积 4-形式为 $\mathcal{V}_4 = \mathbf{d}\mathrm{x}\wedge\mathbf{d}\mathrm{y}\wedge\mathbf{d}\mathrm{z}\wedge\mathbf{d}\mathrm{t}$，定义 $\star\,\mathbf{d}\mathrm{x}^{\mathrm{i}}$ 为

$$\boxed{\mathbf{d}x^i \wedge \star\, \mathbf{d}x^i = g_{ii}\, \mathcal{V}_4.}$$

推导

$$\star\, \mathbf{d}t = \mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}z = \mathcal{V}_3,$$

$$\star\, \mathbf{d}x = \mathbf{d}y \wedge \mathbf{d}z \wedge \mathbf{d}t,$$

$$\star\, \mathbf{d}y = \mathbf{d}z \wedge \mathbf{d}x \wedge \mathbf{d}t,$$

$$\star\, \mathbf{d}z = \mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}t.$$

(vi) 仍然在闵可夫斯基时空中，我们定义一个基 2-形式 Ψ 的对偶 $\star\Psi$

$$\boxed{\Psi \wedge \star\, \Psi = \pm\mathcal{V}_4,}$$

如果 Ψ 包含 $\mathbf{d}t$，我们选择减号 $(-)$，否则我们选择加号 $(+)$。（请注意，此规则包含上一部分中的规则作为特例。）推导出

$$\star\,(\mathbf{d}x \wedge \mathbf{d}t) = -\mathbf{d}y \wedge \mathbf{d}z, \qquad \star\,(\mathbf{d}y \wedge \mathbf{d}z) = \mathbf{d}x \wedge \mathbf{d}t,$$

$$\star\,(\mathbf{d}y \wedge \mathbf{d}t) = -\mathbf{d}z \wedge \mathbf{d}x, \qquad \star\,(\mathbf{d}z \wedge \mathbf{d}x) = \mathbf{d}y \wedge \mathbf{d}t,$$

$$\star\,(\mathbf{d}z \wedge \mathbf{d}t) = -\mathbf{d}x \wedge \mathbf{d}y, \qquad \star\,(\mathbf{d}x \wedge \mathbf{d}y) = \mathbf{d}z \wedge \mathbf{d}t.$$

(vii) 验证对于闵可夫斯基时空中的 2-形式，$\star\star = -1$。（事实上，所有 p-形式都是如此。）

(viii) 回忆法拉第 2-形式由（34.22）给出：

$$\mathbf{F} = \mathbf{F}\text{araday} = \varepsilon \wedge \mathbf{d}t + \mathbf{B}.$$

使用前面的部分推导出麦克斯韦 2-形式（34.26）确实是法拉第 2-形式的霍奇对偶：

$$\star\, \mathbf{F} = \mathbf{M}\text{axwell} = \beta \wedge \mathbf{d}t - \mathbf{E}.$$

(ix) 同样，验证时空电流 1-形式 $\mathbf{J} = -\rho\, \mathbf{d}t + \mathbf{j}$ 的对偶确实由电流密度 3-形式的公式（36.18）给出：

$$\star\, \mathbf{J} = -\rho\mathcal{V}_3 + \left[\underline{\mathbf{j}}\text{的通量 2 形式}\right] \wedge \mathbf{d}t$$

16. 电荷守恒

(i) 假设电荷是守恒的，即它既不能被创造也不能被消灭，推论如果 V 是封闭曲面 $\mathcal{S} = \partial \mathrm{V}$ 的内部，则

$$\frac{\mathrm{d}}{\mathrm{dt}} \iiint_{\mathrm{V}} \rho \, \mathrm{d}\mathcal{V} = -\oiint_{\mathcal{S}} \underline{\mathbf{j}} \cdot \mathbf{n} \, \mathrm{d}\mathcal{A}.$$

(ii) 用高斯定理推导出电荷守恒的局部表达式为

$$\frac{\mathrm{d}\rho}{\mathrm{dt}} + \nabla \cdot \underline{\mathbf{j}} = 0.$$

(iii) 通过对第二对麦克斯韦方程组的外导数，形式为

$$\mathbf{d} \star \mathbf{F} = 4\pi \star \mathbf{J},$$

推断电荷守恒是麦克斯韦方程组的一个逻辑推论，其形式为

$$\boxed{\mathbf{d} \star \mathbf{J} = 0.}$$

历史注释：法拉第 1831 年对电磁感应的实验发现最终被麦克斯韦结晶成数学形式：变化的磁场在电场中产生旋量——见 (37.13)。 三十年后，在没有任何实验证据的情况下，这项练习的纯理论考虑迫使麦克斯韦在 1861 年得出结论，对称地，变化的电场必然在磁场中产生旋量。没有这种对称性，电磁波就不可能存在—— 要有光！

17. 旋转光子和自对偶。 在闵可夫斯基时空中，一个复 2-形式 Ψ 称为自对偶，如果

$$\star \Psi = \mathrm{i}\Psi,$$

和反自对偶如果

$$\star \Psi = -\mathrm{i}\Psi,$$

其中 ⋆ 是练习 15 中介绍的霍奇对偶算子。

(i) 让我们验证法拉第 2-形式 **F**（或任何 2-形式，就此而言）总是可以拆分为自对偶部分 ${}^{+}\mathbf{F}$ 和反自对偶部分 ${}^{-}\mathbf{F}$，因此

$$\mathbf{F} = {}^{+}\mathbf{F} + {}^{-}\mathbf{F}.$$

从这个公式开始，并利用 ${}^{\pm}\mathbf{F} = \pm i\,{}^{\pm}\mathbf{F}$这一事实，推导出

$$ {}^{+}\mathbf{F} = \frac{1}{2}[\mathbf{F} + i \star \mathbf{F}] \quad 和 \quad {}^{-}\mathbf{F} = \frac{1}{2}[\mathbf{F} - i \star \mathbf{F}]. $$

(ii) 事实证明，在量子力学中，这些复共轭分量 ${}^{+}\mathbf{F}$ 和 ${}^{-}\mathbf{F}$ 分别描述了右旋和左旋光子（电磁场的量子）。验证所有四个麦克斯韦方程都可以封装成一个复方程：

$$ \boxed{\mathbf{d}\,{}^{+}\mathbf{F} = -2\pi i \star \mathbf{J}.} $$

积分

18. **闭的和恰当的 1-形式**。用物理学的语言来说，$\varphi = \mathbf{d}f$ 对应一个保守的力场，$\underline{\varphi} = \nabla f$，f 是势能。通过证明以下每一个 1-形式都是闭的，证明每一个都是保守的：$\mathbf{d}\varphi = 0$。然后通过显式找到匹配的势能函数 f 来证明每个都是恰当的。

    (i) $\varphi = x\,\mathbf{d}x + y^2\,\mathbf{d}y + z^3\,\mathbf{d}z$。

    (ii) $\varphi = 3x^2y^2z\,\mathbf{d}x + 2x^3yz\,\mathbf{d}y + x^3y^2\,\mathbf{d}z$。

    (iii) $\varphi = (2xy + z)\,\mathbf{d}x + (x^2 + 3y^2)\,\mathbf{d}y + x\,\mathbf{d}z$。

19. **闭的和恰当的 2-形式**。验证以下 2-形式中的每一种是闭的：$\mathbf{d}\Psi = 0$。然后通过明确地找到一个 1-形式势 φ 使得 $\Psi = \mathbf{d}\varphi$ 来证明每个都是恰当的。

    (i) $\Psi = 3x^2y^4\,\mathbf{d}x \wedge \mathbf{d}y$。

    (ii) $\Psi = 2xyz^3\,\mathbf{d}x \wedge \mathbf{d}y - 3x^2yz^2\,\mathbf{d}y \wedge \mathbf{d}z$。

20. **闭的和恰当的 3-形式**。在 $\mathbb{R}^3$ 中，每个 3-形式 Υ 都与体积 3-形式成正比，$\mathcal{V} = \mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}z$，并且由于没有 4-形式，所以每个 3-形式都是闭的：$\mathbf{d}\Upsilon = 0$。通过明确地找到一个 2-形式的势 Ψ 使得 $\Upsilon = \mathbf{d}\Psi$，证明每个 Υ 都是恰当的。

    (i) $\Upsilon = yz\,\mathcal{V}$。

    (ii) $\Upsilon = 2(x + y + z)\,\mathcal{V}$。

21. **齐次函数**。（来自 do Carmo (1994)。）如果 $g(tx, ty, tz) = t^k\,g(x, y, z)$，则函数 $g: \mathbb{R}^3 \to \mathbb{R}$ 称为齐次 k 次函数。

    (i) 证明这样的 g 满足欧拉关系：

$$x\,\partial_x g + y\,\partial_y g + z\,\partial_z g = kg.$$

提示：对定义的关系关于 t 求微分。

(ii) 如果 1-形式，

$$\psi = a\,\mathbf{d}x + b\,\mathbf{d}y + c\,\mathbf{d}z$$

是这样的，a，b，c 是 k 次齐次的，并且 $\mathbf{d}\psi = 0$，则证明 $\psi = \mathbf{d}f$，其中

$$f = \frac{xa + yb + zc}{k+1}.$$

提示：以分量形式写出 $\mathbf{d}\psi = 0$，然后应用欧拉关系。

(iii) 考虑对应于 ψ 的通量 2-形式：

$$\Psi = a\,\mathbf{d}y \wedge \mathbf{d}z + b\,\mathbf{d}z \wedge \mathbf{d}x + c\,\mathbf{d}x \wedge \mathbf{d}y.$$

证明如果 $\mathbf{d}\Psi = 0$，则 $\Psi = \mathbf{d}\gamma$，其中

$$\gamma = \frac{(zb - yc)\,\mathbf{d}x + (xc - za)\,\mathbf{d}y + (ya - xb)\,\mathbf{d}z}{k+2}.$$

(iv) 如果 V 是 $\mathbb{S}^2$ 的内部，因此 $\partial V = \mathbb{S}^2$，使用（37.14）形式的 FTEC 来证明

$$\iiint_V [\nabla^2 g]\,d\mathcal{V} = k \oiint_{\mathbb{S}^2} g\,d\mathcal{A}.$$

提示：如果 $\mathbf{n}$ 是 $\mathbb{S}^2$ 的单位法线，验证欧拉关系可以写成 $\mathbf{n}\cdot\nabla g = kg$。

途径形式的微分几何

22. **毛球定理**。（来自 do Carmo (1994)。）正如我们结合 [19.8] 所讨论的，庞家莱-霍普夫定理（19.6）意味着在 $\mathbb{S}^2$ 上不存在无奇点向量场。这通常被称为毛球定理：我们不能在椰子上梳理头发！
这是第二个证明，使用 FTEC。假设存在这样一个不灭场 $\mathbf{v}$，并用它构造正交基 $\{\mathbf{m}_1, \mathbf{m}_2\}$ 的第一个向量 $\mathbf{m}_1 = \mathbf{v}/|\mathbf{v}|$。然后高斯方程（38.30）产生

$$\mathbf{d}\omega_{12} = -\theta^1 \wedge \theta^2 = -\mathcal{A}.$$

现在用 FTEC 在 $\mathbb{S}^2$ 上积分这个方程，从而得到一个矛盾。

23. **共形曲率公式**。 如果曲面的共形度量为 $ds^2 = (du^2 + dv^2)/\Omega^2$，则证明

$$\boxed{\mathcal{K} = \Omega\nabla^2\Omega - (\partial_u\Omega)^2 + (\partial_v\Omega)^2.}$$

作为一个简单的检查，在双曲平面的贝尔特拉米-庞加莱半平面模型上尝试这个，其中 $\Omega = v$。

24. **曲率为零表征欧几里得几何**。 （来自 do Carmo (1994)。）本练习将确立

$$\boxed{\text{当且仅当 2-曲面局部是欧几里得的时，高斯曲率才为零。}}$$

“局部欧几里得”是指在每个点周围我们可以找到 (u, v) 坐标，使得度量为 $ds^2 = du^2 + dv^2$。

(i) 在曲面的切向量中选择一个正交标架 $\{\mathbf{m}_1, \mathbf{m}_2\}$。 由高斯方程 (38.30)，

$$\mathbf{d}\omega_{12} = -\mathcal{K}\,\theta^1 \wedge \theta^2 = 0。$$

推导出存在一个函数 ϕ——我们将其解释为一个角度——使得 $\omega_{12} = \mathbf{d}\phi$。

(ii) 考虑一个新的正交标架 $\{\tilde{\mathbf{m}}_1, \tilde{\mathbf{m}}_2\}$，通过将 $\{\mathbf{m}_1, \mathbf{m}_2\}$ 旋转角度 ϕ 获得。证明在新标架中，$\tilde{\omega}_{12} = 0$。

(iii) 推导

$$\mathbf{d}\tilde{\theta}^1 = 0 = \mathbf{d}\tilde{\theta}^2.$$

(iv) 推导出度量为欧几里得的坐标的存在性。

25. **贝尔特拉米-庞加莱半平面的联络形式**。 贝尔特拉米-庞加莱半平面的度量（5.8）立即产生基 1-形式：

$$ds^2 = \frac{dx^2 + dy^2}{y^2} \Longrightarrow \theta^1 = \frac{1}{y}\mathbf{d}x \quad 和 \quad \theta^2 = \frac{1}{y}\mathbf{d}y.$$

(i) 计算 $\mathbf{d}\theta^1$ 和 $\mathbf{d}\theta^2$。

(ii) 使用嘉当第一结构方程，

$$\mathbf{d}\theta^1 = -\omega^1{}_2 \wedge \theta^2, \qquad \mathbf{d}\theta^2 = -\omega^2{}_1 \wedge \theta^1 \ = \omega^1{}_2 \wedge \theta^1,$$

推断出（唯一的）解决方案是

$$\omega^1{}_2 = -\frac{1}{y}\mathbf{d}x.$$

(iii) 计算 $\mathbf{d}\omega^1{}_2$，并将结果与高斯方程进行比较，

$$\mathbf{d}\omega^1{}_2 \ = \mathcal{K}\theta^1 \wedge \theta^2,$$

推导出双曲平面的曲率确实是 $\mathcal{K} = -1$。

26. **2-曲面的曲率 2-形式**。 令 $\{\mathbf{m}_1, \mathbf{m}_2\}$ 为 2-曲面的切向量 $\mathbf{v}$ 的标准正交基场，因此

$$\begin{aligned} \nabla_{\mathbf{v}}\mathbf{m}_1 \ &= -\omega^1{}_2(\mathbf{v})\mathbf{m}_2 \\ \nabla_{\mathbf{v}}\mathbf{m}_2 \ &= \ \omega^1{}_2(\mathbf{v})\mathbf{m}_1. \end{aligned}$$

(i) 写下联络矩阵 $[\omega]$，从而证明

$$[\omega] \wedge [\omega] = 0.$$

(ii) 推导出曲率矩阵（38.53）为

$$[\Omega] = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{d}\omega^1{}_2.$$

因此，曲率矩阵完全由单曲率 2-形式 $\Omega^1{}_2$ 描述，由高斯方程控制：

$$\Omega^1{}_2 = \mathbf{d}\omega^1{}_2 = \mathcal{K}\mathcal{A}.$$

27. **双曲 3-空间 $\mathbb{H}^3$ 曲率**。 回想一下，$\mathbb{H}^3$ 的度量由（6.23）给出，在第 80 页的 [6.6] 中描述：

$$ds^2 \ = \frac{dx^2 \ + dy^2 \ + dz^2}{z^2} \implies [\theta] = \begin{bmatrix} \theta^1 \\ \theta^2 \\ \theta^3 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} \mathbf{d}x \\ \mathbf{d}y \\ \mathbf{d}z \end{bmatrix}$$

(i) 回想一下，从我们原来的 (O'Neill) 记号 ω_{ij} 到我们新的、更标准的记号 $\omega^i{}_j$ 的转换会导致嘉当第一结构方程中的符号发生变化，因此它现在采用以下形式（见 (38.54)）：

$$\mathbf{d}[\theta] = -[\omega] \wedge [\theta].$$

推导出联络矩阵是

$$[\omega] = \frac{1}{z}\begin{matrix} 0 & 0 & -\mathbf{d}x \\ 0 & 0 & -\mathbf{d}y \\ \mathbf{d}x & \mathbf{d}y & 0 \end{matrix}$$

(ii) 使用（38.53）推导出曲率矩阵为

$$[\Omega] = \frac{1}{z^2}\begin{bmatrix} 0 & -\mathbf{d}x \wedge \mathbf{d}y & -\mathbf{d}x \wedge \mathbf{d}z \\ \mathbf{d}x \wedge \mathbf{d}y & 0 & -\mathbf{d}y \wedge \mathbf{d}z \\ \mathbf{d}x \wedge \mathbf{d}z & \mathbf{d}y \wedge \mathbf{d}z & 0 \end{bmatrix}$$

(iii) 验证 $[\Omega]$ 是否可以重写为

$$[\Omega] = \begin{bmatrix} 0 & -\theta^1 \wedge \theta^2 & -\theta^1 \wedge \theta^3 \\ \theta^1 \wedge \theta^2 & 0 & -\theta^2 \wedge \theta^3 \\ \theta^1 \wedge \theta^3 & \theta^2 \wedge \theta^3 & 0 \end{bmatrix}$$

并用（38.55）推导出黎曼张量有分量

$$R^1{}_{212} = R^1{}_{313} = R^2{}_{323} = -1.$$

(iv) 将黎曼张量的这些分量解释为截面曲率。因此，截面雅可比方程（29.21）告诉我们关于在该空间内最初以相同方向发射的相邻测地线的行为？得出结论，说“$\mathbb{H}^3$ 具有恒定曲率 -1”是有意义的。

28. **宇宙曲率**。（这个练习的详细解可以在 Dray (2015, §A.9) 和 Sternberg (2012, §6.7) 中找到。）现在考虑爱因斯坦场方程的弗里德曼-勒梅特-罗伯逊-沃克（Friedmann-Lemaître-Robertson-Walker (FLRW)）解我们膨胀的宇宙的大尺度几何的标准模型。它们于 1922 年由亚历山大•弗里德曼首次发现，1927 年由耶稣会神父乔治•勒梅特重新发现，1935 年罗伯逊和沃克共同证明了它们作为宇宙可能的空间均匀和各向同性几何形状的独特性。

弗里德曼寻找并找到了原始爱因斯坦方程的解，而我们现在知道宇宙学爱因斯坦方程（30.25）实际上是正确的，宇宙学常数 $\Lambda > 0$。幸运的是，可以使弗里德曼的发现适应这一新现实。

FLRW 度量由下式给出

$$ds^2 = -dt^2 + R^2(t)\left[\frac{dr^2}{(1-Kr^2)} + r^2(d\phi^2 + \sin^2\phi d\vartheta^2)\right],$$

其中 $R(t)$ 测量大爆炸后 t 时刻膨胀宇宙的大小，K 是宇宙一直以来的恒定空间曲率。 参见 Penrose (2005, §§27.11, 27.12) 对 $K > 0$, $K = 0$, $K < 0$, 包括 $\Lambda > 0$ 的影响这三种情况的精彩讨论（和可爱的图画）。

跟随我们计算史瓦西黑洞曲率的脚步，写出基 1-形式 θ^i，取其外导数求出联络 1-形式 $\omega^i{}_j$，然后计算 $\mathbf{d}\omega^i{}_j$，最后推导出这样一个宇宙的曲率 2-形式 $\Omega^i{}_j$。

延伸阅读

下面，我仅仅列出标题和作者； 完整细节可以在参考文献中找到。某些作品被列进来仅仅因为他们与主题高度相关，而我相信他们值得引起你的注意（即使它们并不是我的菜）。然而大多数作品被列进来是因为我认为它们如珍宝一般，而向你强烈推荐。许多其他极好的书躺在我的书架上，我也经常寻求它们的建议，然而它们并没有被包括进来，仅仅因为我努力把这个（已经很长的）书单削减到可管理的长度—我向所有这些优秀作品的作者道歉。最后，我也向尚未发现的精彩作品的作者们表示歉意！

整体推荐

首先，让我用我认为非常有价值的六部作品来做铺垫，每部作品的内容都涵盖了本书的多个章节。

- 《通向实在之路》，罗杰·彭罗斯（Roger Penrose）著。

一部几乎涵盖所有物理学和大部分数学的非凡全景图，由这两个领域的大师创作。许多见解只有在这里才能找到，而且这些见解被彭罗斯出色的（和美丽的！）手绘赋予了生命。

- 《引力论》，米斯纳（Misner）、索恩（Thorne）和惠勒（Wheeler）著。

在 1973 年首次出版后的近 50 年里，这部经典之作仍然是对爱因斯坦的几何引力理论（广义相对论）和它所依据的微分几何的最好介绍之一。它也包含了对微分形式的最好的、最具几何性的介绍之一，包括允许人们有效计算黎曼张量的曲率 2-形式。普林斯顿大学出版社的新版（2017 年）做得很好，并包含了 Charles Misner 和 Kip Thorne 的新介绍，讨论了自该书出版以来该领域令人兴奋的发展。

- 《整体微分几何》，海因茨·霍普夫（Heinz Hopf）著。

霍普夫不仅是二十世纪数学界的杰出人物之一，也是一位论述的大师。在这里，微分几何学和拓扑学的思想（其中许多是由霍普夫本人提出的）以一种美妙的方式汇集在一起，并以非常清晰和简单的方式进行了解释。每当我再次阅读这本经典作品，我就觉得有位仁慈的魔术师在书中插入了更多美妙的想法，因为我发誓，在我上次看的时候，这个美丽的想法还没有出现在书页上！

- 《微分几何基础》（修订版第 2 版），巴雷特·奥尼尔（Barrett O'Neill）著。

这本开创性的书于 1966 年首次出版，开创了在本科阶段使用微分形式的先河。 半个多世纪后的今天，在我看来，奥尼尔的作品仍然是本科阶段对这一主题唯一最清晰、最优雅、最现代（讽刺吧？）的处理方式（除了当代的同类书籍！）！

- 《数学物理的几何方法》，伯纳德·舒茨（Bernard Schutz）著。

这部作品（现在已经有 40 年的历史了！）是一个永恒的宝库，涵盖了流形、张量、李导数和李群、形式、黎曼几何、规范理论以及其他大量的物理学应用。为了实现这一目标，舒茨教授成为了《星际迷航》中的斯波克先生。他的另一半瓦肯人格使他能够建立一个完美的

晶体逻辑结构，其中每一件事都被简明而严格地证明，而且－与我在本书中的（我希望是美味的）芝士汉堡方法不同－犹太教规被严格遵守：依赖于公制存在的概念被严格而明确地与那些不依赖于公制的概念分开。但是，与此同步，舒茨能够利用他的另一半人格来提供丰富的直觉，揭示出潜在的几何现实。

- 《数学及其历史》（第 3 版），约翰·斯蒂尔威尔（John Stillwell）著。

通过历史的视角，对所有的数学进行了一次出色的全景展示。但不要误解，这主要不是一本关于历史的书，相反，它从根本上是一本关于数学本身的相互联系和意义的作品，所有这些都以一种相当简洁的风格（相对于我来说！）来解释，具有深刻的洞察力和清晰的思路。

一般几何

以下作品涉及一般的几何学，尤其是双曲几何学（微分几何有自己的类别）。

- 《直观几何》，大卫·希尔伯特与康福森（S. Cohn-Vossen）著。

20 世纪最伟大的数学家之一，对几何学进行了宏伟而深刻的研究，侧重于直观的理解。这些图（由 K. H. Naumann 和 H. Bödeker 绘制）非常漂亮，让我羡慕不已！

- 《体验几何》，DavidW. Henderson and Daina Taimina 著。

一种非常不寻常的方法，在哲学上类似于我的（但使用摩尔方法），专注于对几何的直观、实验研究。 它包含对平行移动和和乐论的重要讨论。 序言中的这句话清楚地表明了他们的方法与我的方法的重叠：“这本书基于一种证明观点，即一种令人信服的交流，它回答了——为什么？” （他们的斜体。）

- 《几何导论》（第 2 版），H.S.M. Coxeter 著。

一位现代大师进行的非常清晰的概述。

- 《几何》，Brannan, Esplen, and Gray 著。

基于克莱因对变换群的看法，对几何进行了出色而现代的概述。

- 《欧几里德和非欧几里德几何：发展和历史》，Marvin J. Greenberg 著。

双曲几何发展的宝贵而详细的历史，包括对高斯、博利亚和许多其他人的批评性私人信件的冗长引用。

- 《庞加莱半平面》，索尔·斯塔尔（Saul Stahl）著。

标题说明了一切。

- 《几何揭示》，马塞尔·伯杰（Marcel Berger）著。

由 20 世纪最伟大的几何学家之一对几何进行更高级的概述，重点是概念证明和未解决的问题。

拓扑

- 《直观拓扑》，V. V. Prasolov 著。
  不到 100 页长，充满图画，这个超级友好的介绍不辜负它的标题！

- 《欧拉的宝石》，David S. Richeson 著。
  对欧拉的多面体公式、它的历史和相关的数学思想的精湛、数学准确但引人入胜的描述。

- 《曲面拓扑》，菲尔比和加德纳（P. A. Firby and C. F. Gardiner）著。
  对拓扑学基本几何概念的非常温和、插图精美的介绍。

- 《拓扑的第一概念》，钦和斯滕罗德（W. G. Chinn and N. E. Steenrod）著。
  另一个非常温和、插图精美的拓扑基本几何概念介绍。

- 《拓扑：一个非常简短的介绍》，理查德·厄尔（Richard Earl）著。
  这本非凡的小书名副其实，仅 140 页就涵盖了大量的基本思想，而且它以非常清晰、基本、非正式的风格进行。 这是我最喜欢的对该主题的介绍。

- 《空间的形状（第 3 版）》，杰弗里·威克斯（Jeffrey R.Weeks）著。
  对二维和三维空间的拓扑结构进行了非常清晰、引人入胜的基本处理。 本书四部分的最后一部分讨论了探测宇宙拓扑的可能性！ 附录包含 John Horton Conway 著名的曲面分类定理 ZIP 证明，由 George K. Francis 精美插图。

- 《三维几何和拓扑》，威廉·瑟斯顿（William P. Thurston）著。
  瑟斯顿因发现 3-流形从根本上是由双曲几何构造而成，因此获得了菲尔兹奖。 在这本书中，你将听到瑟斯顿以独特的声音讲述自己的发现，尽管拓扑学的难度迅速增加，但前 100 页给出了不容错过的对双曲几何的相对基本、高度原创的介绍。

霍普夫的线场和物理学中的庞家莱-霍普夫定理

在本书中，我试图引起人们对线场（line fields）和霍普夫将庞家莱-霍普夫定理应用于它们得到的美妙结果的关注（（19.9），第 213 页）。 这些想法几乎从现代数学教科书中消失了，我坚信现在是复兴的时候了。 话虽如此，物理学家从未忘记这些想法的价值，并且他们用奇妙的新发现来支撑着这些想法。

在这里，我被迫列出研究论文而不是说明性教科书。 我所知道的唯一的说明性数学处理是霍普夫(1956) 本人的讲座（见前文），以及他的学生 Stoker (1969) 的书。 尽管对光学的最新贡献明确引用了霍普夫的工作（甚至将分数指标称为霍普夫指标”）但这些想法是由 JF Nye 和后来的 Michael Berry 爵士率先提出的，他们似乎都不知道霍普夫（1956）。 （注意：在阅读这些论文时，重要的是要了解数学家称之为和乐（holonomy），物理学家有时称之为异和乐（anholonomy）！）

The Topology of Ridge Systems，罗杰·彭罗斯著。

使用指纹和掌纹作为示例，对这一思想圈进行完整的基本介绍（对于非数学家），因为这些确实是线场！

The Fingerprint of the Weyl Tensor, Spinors and Space-Time, Vol. 2, pp. 229–233，罗杰·彭罗斯著。

彭罗斯（Penrose） 对外尔（Weyl）曲率张量的“指纹”描述揭示了具有具有分数指标的奇点的线场。

I. Liquid Crystals. On the Theory of Liquid Crystals，F. C. Frank 著。

物理学中最早出现的线场之一。请注意，在这里它们被称为向列场，它们的奇点被称为拓扑缺陷。

The Poincaré–Hopf Theorem for Line Fields Revisited，Diarmuid Crowley 和 Mark Grant 著。

对定理的历史和应用的一个很好的回顾；它还包括所有维度的一般证明。

Generic Singularities of Line Fields on 2D Manifolds，Ugo Boscain, Ludovic Sacchelliy, and Mario Sigalotti 著。

对自然界中线场的许多外观及其与脐点相关的数学分类进行了很好的回顾。

Natural Focusing and Fine Structure of Light: Caustics and Wave Dislocations，J. F. Nye 著。

Nye 率先将线场（偏振场）引入光学：通用奇点是 2D 中的圆偏振点或 3D 中的线。分析了另外两种类型的奇点：强度（焦散）和相位（涡流），它们是由 Nye 和 Michael Berry 爵士开创的。

- The Quantum Phase, Five Years After，M. V. Berry, in Shapere and Wilczek (1989) 著。

  量子相发现者本人的评论，明确地将发现与微分几何联系起来，讨论围绕脐点的曲率线，并做出关键观察，即围绕这一点的电路导致 $\pm\pi$ 的旋转，对应于 $\pm\frac{1}{2}$ 的指标，如 [19.14 a & A] 所示。对于这种现象的一个很好的说明，请参见 Hilbert (1952, p. 189)。

- Geometry of Phase and Polarization Singularities, Illustrated by Edge Diffraction and the Tides，M. V. Berry 著。

  边缘衍射和潮汐也会产生带有分数指标的奇异点。

- Index Formulae for Singular Lines of Polarization，M. V. Berry 著。

  对于单色光场的偏振为纯圆形（C）或纯线性（L）的空间线的指标，获得了公式。指标（C 线为$\pm\frac{1}{2}$，L 线为±1）涉及线上的电场或磁场及其导数。

- Polarization Singularities in the Clear Sky，M. V. Berry, M. R. Dennis, and R. L. Lee 著。

  日光下偏振方向模式的说明。 奇点（两个靠近太阳，两个靠近反太阳）是天空中的点，其中极化线场图案的指数为 $+\frac{1}{2}$，极化强度为零。

- A Half-Century of Physical Asymptotics and Other Diversions: Selected Works by Michael Berry，M. V. Berry 著

Michael Berry 爵士对光学的开创性贡献，专注于揭示日常生活和前沿基础光子研究中各种光学奇点的普遍表现。

- Singularities and Poincaré Indices of Electromagnetic Multipoles，作者：陈伟金、陈云天和刘伟。

作者绘制了不同阶多极辐射的所有奇点，确定了它们的指标，并明确表明整个动量球上的指标总和始终为 2，这与庞家莱-霍普夫定理一致。

- Global Mie Scattering: Polarization Morphologies and the Underlying Topological Invariant，作者 Weijin Chen、Qingdong Yang、Yuntian Chen 和 Wei Liu。

庞家莱-霍普夫定理（带分数指标）用于表明，如果你将任何形式的相干偏振光照射到任何粒子或粒子簇上，则必须存在一个散射为零或圆偏振的方向。

- Line Singularities and Hopf Indices of Electromagnetic Multipoles，作者：陈伟进、陈云天、刘伟

本文研究了半整数霍普夫指标、电磁多极子和 Bloch 模式之间的联系。

微分几何

- 《微分几何：几何介绍》，DavidW. Henderson 著。

在我将列出的所有微分几何书籍中，这本在精神上与我的工作最接近，这就是我将它放在首位的原因。我在写 VDGF 的后期才发现它，并立即确信 Henderson 教授是志同道合的人，是数学上的兄弟。关于我试图与他联系的悲伤故事，请参见第 252 页的脚注。

- 《高斯“关于曲面的一般研究”的 150 年后》，Peter Dombrowski 著。

使用高斯给朋友的信函、私人笔记本和未发表的手稿草稿，这部具有独特价值和极富洞察力的作品追溯了高斯在微分几何方面的发现的演变，最终导致了他革命性的《*关于曲面的一般研究*》，向世界宣布了*绝妙定理*。作为额外的奖励，Dombrowski 完整地收录了高斯 1827 年的杰作，原版拉丁文，对页有英文翻译。

- 《微分几何综合介绍（第 2 卷）》，迈克尔·斯皮瓦克（Michael Spivak）著。

斯皮瓦克的巨著实际上包括五卷，但就我而言，第二卷是最相关的，因为它包括对高斯和黎曼开创性研究的详细而深刻的分析。这本书还包括嘉当的活动标架方法，以及他的两个结构方程。

- 《曲线和曲面的微分几何（第 2 版）》，作者：Thomas Banchoff 和 Stephen Lovett。

虽然这里使用的数学机制是标准的，具有克里斯托费尔符号等，但它包含一些你在其他地方找不到的高度原创的几何见解（由于 Banchoff）。此外，随附的网站包含一系列精彩的 Java 小程序，用于交互式探索这些概念。

- 《曲线和曲面的微分几何》，作者 Kristopher Tapp。

近年来最好的，最几何的介绍之一。虽然数学机制是标准的——用克里斯托弗符号填充的丑陋的方程！——但我的方法仍有许多概念上的联系点。它有许多做得很好的彩色图片，许多出色而有趣的应用程序，并且（就像这本书一样）它强调了平行移动和和乐的重要性。强烈推荐！ （注：它只处理 2-曲面，因此没有讨论黎曼张量或广义相对论。）

- 《弯曲空间》，P.M.H. Wilson 著。

本书的前三分之二是对欧几里得、球面和双曲几何的简洁、现代、优雅的处理，采用了默比乌斯变换的统一力量（就像我一样）。最后三分之一是对微分几何的介绍。

- 《微分几何第一门课程：欧几里得空间中的曲面》，L. M.Woodward 和 J. Bolton 著。

这本作品的最后一章是我所知道的（在本科阶段）对本书中我选择完全忽略的一个重要而美丽的主题（极小曲面）的最佳介绍。回想一下，这些是平均曲率为零的曲面，由肥皂膜实现：$H \equiv \frac{\kappa_1+\kappa_2}{2} = 0$。讨论包括 1982 年 Celso Costa 的突破性发现，即闭合的且没有自相交的新的（平面、链状曲面和螺旋面之外）极小曲面。此外，不同寻常的是，它还对推广到 $H = \text{const} \neq 0$ 的称为恒定平均曲率曲面（surfaces of constant mean curvature，CMC 曲面）的情形进行了彻底的讨论。

- 《初等微分几何（第 2 版）》，作者 Andrew Pressley。

一个优秀的、初等的介绍，和上一本书一样，它包含了对极小曲面的非常好的介绍。

- 《微分几何和拓扑课程》，作者 A. Mishchenko 和 A. Fomenko 。

一个优秀的、有点进阶的、内容广泛的介绍，包括黎曼几何和同调理论。它以大量明显福缅科式漫画风格的出色插图而著称。

- 《微分几何：连接、曲率和特征类》，作者杜武亮 (Loring W. Tu)。

这是我所知道的从 VDGF 过渡到处理陈类（陈省身的示性类）的更高级作品的最好的书，这些作品在纯数学和物理学中都变得很重要。这本书确实假设读者已经熟悉微分形式，而我的第五幕提供了足够多的准备。杜（武亮）还假设读者熟悉德拉姆（de Rham）上同调，在这里我在第五幕中的“入门”可能还不够，尽管我希望我的物理和几何处理（这在其他地方很难找到）将大大方便对该主题的进一步研究。

- 《微分几何讲义》，陈省身, 陈维桓和 K. S. Lam。

这本书有 356 页，包含不到 10 张图。我将这项进阶的作品纳入有且仅有一个理由：陈（他的讲座是本书的基础）是 20 世纪最伟大的几何学家之一，因此，他要说的任何话，我们都应该专心聆听。特别是，陈对芬斯勒几何进行了长篇论述，这是对黎曼几何的广泛推广，由陈本人开创。与本书的唯一联系是陈使用微分形式作为他的主要工具——这并不奇怪，因为陈实际上是在嘉当本人的手下学习的！

黎曼曲率

- 《黎曼几何全景图》，马塞尔·伯杰（Marcel Berger ）著。

来自 20 世纪最伟大的几何学家之一的无价全景视图。 它侧重于结果和概念，解释清楚直观，除了最重要的外，它（有意）省去了所有细节。

- 《黎曼几何：初学者指南（第 2 版）》，弗兰克·摩根（Frank Morgan）著。

这篇文笔优美、插图精致的介绍仅 129 页就包含了很多内容。 尽管有它的副标题，我猜想一个真正的初学者在阅读我的作品后会更容易应付这件作品！

- 《黎曼几何》，曼弗雷多·P·多·卡尔莫（Manfredo P. do Carmo）著。

对许多深层定理的出色(但高级)处理，其中一些清楚地讨论很难在其他任何地方找到。

- 《半黎曼几何：与相对论的应用》，巴雷特·奥尼尔（Barrett O'Neill）著。

半黎曼几何是时空几何，这是我所知道的关于这个主题的最好的数学作品。

爱因斯坦的弯曲时空（广义相对论）

- 《从头开始的重力》，伯纳德·舒茨（Bernard Schutz）著。

第二杰作！舒茨在他的第一部杰作中引导了斯波克先生（见上文），而在这里他引导了柯克船长！他采用真实的、费曼式的、物理的、直观的推理——仅使用高中数学——来处理重力的各个方面。同时，他成功地描绘了我们对自然基本力量不断发展的理解的历史和人文背景。

- 《重力和时空之旅》，约翰·阿奇博尔德·惠勒著。

20 世纪最伟大的物理学家之一，以及 1960 年代广义相对论重生的关键人物之一，对爱因斯坦的愿景进行了深刻的、主要是非数学的、概念性的解释：例如，惠勒创造了黑洞这个术语。我与这部作品有个人联系：阅读它后，我发现并寄给了惠勒一个牛顿的纯几何证明，证明了他至关重要的双曲率定律（第 142 页）。正如序言中所解释的，惠勒（1911-2008），作为彭罗斯的学生，我有幸多次见到他，他是我生命中的关键人物，所以当他回信说我的证明让他愉快时，我激动不已！

- 《广义相对论第一课（第 3 版）》，伯纳德·舒茨（Bernard Schutz）著。

在撰写本文时，第三版还不存在，但由于我与作者的通信，我知道它即将出版。第二版已经很好了，第三版只会更好！特别是，我预测新版本将大大扩展对引力波的处理，因为在第二版时甚至还没有发现，而 Schutz 在该领域已经有 30 多年的历史了。2019 年，舒茨因其具有先见之明的理论工作（Schutz 1986）获得了爱丁顿奖章，该工作现在提供了一种从引力波数据计算哈勃常数的全新方法。

- 《地心引力：爱因斯坦引力导论》，James B. Hartle 著。

爱因斯坦理论最好、最物理（较少数学）的介绍之一，由该理论的大师级实践者撰写。

- 《广义相对论》，Robert M.Wald 著。

我认为这部出色的作品是米斯纳、索恩和惠勒万有引力的重大更新继任者。它优雅地平衡了物理直觉与高等数学，包括第一门课程中很少涉及的主题，例如彭罗斯关于 2-旋量的开创性工作，以及彭罗斯和霍金的奇点定理。

- 《广义相对论：一种几何方法》，马尔科姆·路德维森（Malcolm Ludvigsen）著。

如果你急于掌握广义相对论，你几乎没有比这更好的了。这是对理论的非常干净、优雅、几何和简洁的介绍，但是（奇迹般地）实现了这一点而不会感到匆忙。

- 《简述爱因斯坦引力》，A. Zee 著。

广义相对论的精彩介绍，涵盖了许多不寻常的主题，以生动、非正式、引人入胜和高度固执己见的风格写成。 （我们有关系吗？）

形式（数学）

- 《微分形式的几何方法》，David Bachman 著。

少数几本旨在将微分形式带给本科生读者的勇敢书籍之一。虽然它确实包括了对微分几何的非常简短的应用，但它不包括任何关于黎曼张量或广义相对论的讨论。

- 《对流形上的微分形式和微积分的可视化介绍》，Jon Pierre Fortney 著。

尝试将微分形式带给本科生读者的最新（也是最长的）书籍。虽然它确实包括了对电磁学的讨论，但它不包括任何微分几何，因此不包括任何关于黎曼张量或广义相对论的讨论。

- 《微分形式：启发式介绍》，M. Schreiber 著。

将微分形式带给大众的最早尝试之一。它包含一些不错的见解，但几乎没有图片！它不包括任何关于黎曼张量或广义相对论的讨论。

- 《微分形式：向量分析的补充》，Steven H. Weintraub 著。

正如标题所示，这是一次将微分形式带入高年级本科生多元微积分课堂的尝试。虽然非常清晰，但它不是几何的，而且图片很少。

- 《向量分析、线性代数和微分形式：一种统一的方法（第 4 版）》，作者：John H. Hubbard 和 Barbara Burke Hubbard。

与上一本书一样，这是将微分形式带入高年级本科生多元微积分课堂的尝试，不过这里的方法更加几何化，与我的擦肩而过。但是，必须注意，微分形式仅在第 558 页首次出现。

- 《微分形式和应用》，Manfredo P. do Carmo 著。

这本只有 118 页的小册子（由陈省身的一位非常受人尊敬的学生所著）包含了丰富的想法和信息，以及一些可爱的练习，其中一些我已在此处挪用（已注明出处！）。

形式（物理学）

- 《经典力学的数学方法（第 2 版）》，V. I. Arnol'd 著。

二十世纪最杰出的数学家之一的杰作。第 7 章是对 微分形式的自成一体、优雅且富有

洞察力的介绍，然后将其应用于哈密顿力学。

- 《规范场、纽结与引力》，John Baez 和 Javier P. Muniain 著。

对电磁学、规范场和引力的广泛、高度原创、现代的介绍，其中大部分内容以形式语言表达。在它的许多优秀特征中，有一个对霍奇对偶在麦克斯韦方程中的作用的深刻讨论。不要忽视总结本书三个部分的注释：它们包含进一步研究的建议（非常像本节！），它们还包含引人入胜的历史小插曲和精辟的引文。作者以一种友好、非正式的方式直接与读者交谈，就像对坐在他们旁边的一位聪明的朋友讲话，而不是把枯燥的定理空洞地宣布，这非常令人耳目一新并且很有帮助。 （当然这也正是我在本书中试图做的！）

- 《微分形式和广义相对论的几何》，Tevian Dray 著。

对广义相对论的唯一介绍，专门使用 微分形式 作为其数学引擎。在推荐它的许多特性中，我会注意到它包含比我提供的更完整的霍奇对偶性讨论。

- 《物理学的几何学（第 3 版）》，西奥多·弗兰克尔（Theodore Frankel）著。

一本令人印象深刻的巨著，几乎涵盖了上述所有书籍所涵盖的所有主题结合起来还更多！微分形式是本书的主要语言。然而，特别值得注意的是它对爱因斯坦张量几何的独特处理。

- 《数学和物理曲率》，Shlomo Sternberg 著。

微分形式是本书的主要语言。正如标题所示，它包含许多有趣的数学和物理应用。特别是，它深入探讨了以下物理主题：霍奇对偶和电磁学、史瓦西解的几何和轨道的显式计算，以及极其重要的克尔解（代表一个旋转的黑洞）的几何，尽管他没有真正计算曲率 2 形式。但是这份清单几乎不能公正地涵盖所涵盖的大量材料。警告：作者将这本书描述为适合高年级本科生——事实并非如此！尽管如此，如果你已经掌握了我的第五幕，那么你就可以从这本书中学到很多东西。

参考文献

Aleksandrov, A. D.
1969. *Mathematics: Its Content, Methods, Meaning*, Vol. 2. Chapter 7: *Curves and Surfaces*, 57–117.A. D. Aleksandrov, A. N. Kolmogorov, M. A. Lavrent'ev, eds. Translated from the Russian by S. H. Gould. 2nd ed., paperback. Cambridge, Mass.: MIT.

Arnol'd, V. and V. Vasil'ev
1991. *Newton's Principia read 300 years later.* Current Science, 61(2).

Arnol'd, V. I.
1973. *Ordinary Differential Equations.* Cambridge, Mass.: MIT.

Arnol'd, V. I.
1989. *Mathematical Methods of Classical Mechanics*, Vol 60 of *Graduate Texts in Mathematics*, 2nd ed. New York: Springer.

Arnol'd, V. I.
1990. *Huygens and Barrow, Newton and Hooke.* Basel: Birkhäuser Verlag.
Pioneers in mathematical analysis and catastrophe theory from evolvents to quasicrystals, Translated from the Russian by Eric J. F. Primrose.

Bachman, D.
2012. *A Geometric Approach to Differential Forms*, 2nd ed. New York: Birkhäuser. Baez, J. C. and J. P. Muniain 1994. Gauge Fields, Knots, and Gravity, Vol. 4. Singapore:World Scientific.

Banchoff, T.
1970. *Critical points and curvature for embedded polyhedral surfaces.* American Mathematical Monthly, 77(5):475–485.

Banchoff, T. and S. Lovett
2016. *Differential Geometry of Curves and Surfaces*, 2nd ed. Natick, Mass.: AK Peters.

Berger, M.
2003. *A Panoramic View of Riemannian Geometry.* Berlin: Springer.

Berger, M.
2010. *Geometry Revealed: A Jacob's Ladder to Modern Higher Geometry*. Berlin: Springer.

Berry, M.
1991. Bristol anholonomy calendar. In Sir Charles Frank, OBE, FRS: *An eightieth birthday tribute*. Bristol: A. Hilger.

Berry, M.
2017. *A Half-Century of Physical Asymptotics and Other Diversions: Selected Works by Michael Berry*. Singapore:World Scientific.

Berry, M. V.
1990. *Anticipations of the geometric phase*. Physics Today, 43(12):34–40.

Berry, M. V.
2001. Geometry of phase and polarization singularities illustrated by edge diffraction and the tides. *In Second International Conference on Singular Optics (Optical Vortices):* Fundamentals and Applications, M. S. Soskin and M. V. Vasnetsov, eds. 4403:(1–12). Bellingham,Wash.: International Society for Optics and Photonics.

Berry, M. V.

2004. Index formulae for singular lines of polarization. *Journal of Optics A: Pure and Applied Optics*,6(7):675–678.

Berry, M. V., M. R. Dennis, and R. L. Lee

2004. *Polarization singularities in the clear sky.* New Journal of Physics, 6:162–162.

Blaschke,W.

1929. *Vorlesungen über Differentialgeometrie und geometrische Grundlagen von Einsteins Relativitätstheorie* [2 volumes in one]. Chelsea Publishing. Bloye, N. and S. Huggett 2011. Newton, the Geometer. Newsletter of the European Mathematical Society, (82):19–27.

Boscain, U., L. Sacchelli, and M. Sigalotti

2016. *Generic singularities of line fields on 2D manifolds.* arxiv.org: 1605.06295.

Bottazzini, U.

1999. *Ricci and Levi-Civita: From Differential Invariants to General Relativity.* In The Symbolic Universe: Geometry and Physics 1890-1930, J. Gray, ed. Oxford: Oxford University.

Boy,W.

1903. *Über die Curvatura integra und die Topologie geschlossener Flächen.* Mathematische Annalen, 57(2):151–184.

Brackenridge, J. B. and M. Nauenberg

2002. *Curvature in Newton's Dynamics.* In The Cambridge Companion to Newton, I. B. Cohen and G. E. Smith, eds., Cambridge Companions to Philosophy, Chapter 3, pp. 85–137. Cambridge: Cambridge University.

Brannan, D. A., M. F. Esplen, and J. Gray

2012. *Geometry*, 2nd ed. Cambridge: Cambridge University.

Carathéodory, C.

1937. *The most general transformations of plane regions which transform circles into circles*. Bulletin of the American Mathematical Society, 43(8):573–579.

Cartan, E.

1927. *Riemannian Geometry in an Orthogonal Frame: from lectures delivered by Élie Cartan at the Sorbonne in 1926-1927.* [Translated into Russian in 1960 by S. P. Finikov, and from Russian into English by V. V. Golberg in 2001.] Singapore:World Scientific.

Cartan, E.

1928. *Geometry of Riemannian Spaces*, Volume 13. Brookline, Mass.: Math Sci Press. Translation of 2nd ed., 1951, published 1983.

Cartan, E.

1945. *Les Systèmes Differentiels Extérieure et leur Applications Géométriques*. Paris: Hermann.

Casey, J.

1996. *Exploring Curvature.* Braunschweig: F. Vieweg und Sohn.

Chandrasekhar, S.

1995. *Newton's Principia for the Common Reader.* Oxford: The Clarendon Press.

Chen,W., Y. Chen, and W. Liu

2019. *Singularities and Poincaré indices of electromagnetic multipoles.* Phys. Rev. Lett., 122:153907.

Chen,W., Y. Chen, and W. Liu

2020a. *Line singularities and Hopf indices of electromagnetic multipoles.* Laser & Photonics Reviews, 2000049.

Chen,W., Q. Yang, Y. Chen, and W. Liu
2020b. *Global Mie scattering: Polarization morphologies and the underlying topological invariant.* ACS Omega, 5(23):14157–14163.
Chern, S.-S., W.-H. Chen, and K. S. Lam
1999. *Lectures on Differential Geometry*, Volume 1. Singapore:World Scientific.
Coxeter, H.S.M.
1967. *The Lorentz group and the group of homographies*. In Proc. Internat. Conf. Theory of Groups, L. G. Kovacs and B. H. Neumann, eds., New York: Gordon and Breach.
Coxeter, H.S.M.
1969. *Introduction to Geometry*, 2nd ed. New York: Wiley.
Crowe, M. J.
1985. *A History of Vector Analysis: the Evolution of the Idea of a Vectorial System.* New York: Dover.
Crowley, D. and M. Grant
2017. *The Poincaré–Hopf theorem for line fields revisited.* Journal of Geometry and Physics, 117:187–196.
Darling, R.W.R.
1994. *Differential Forms and Connections.* Cambridge: Cambridge University.
Darrigol, O.
2015. *The Mystery of Riemann's Curvature.* Historia Mathematica, 42(1):47–83.
de Gandt, F.
1995. *Force and Geometry in Newton's Principia.* Princeton, N.J.: Princeton University. [Translated from the French original and with an introduction by Curtis Wilson.]
Diacu, F. and P. Holmes
1996. *Celestial Encounters: the Origins of Chaos and Stability.* Princeton, N.J.: Princeton University.
Dirac, P.A.M.
1930. *The Principles of Quantum Mechanics* [The International Series Of Monographs On Physics]. Oxford: The Clarendon Press.
do Carmo, M. P.
1992. *Riemannian Geometry.* Basel: Birkhäuser.
do Carmo, M. P.
1994. *Differential Forms and Applications*, Universitext. Berlin: Springer.
Dombrowski, P.
1979. *150 Years after Gauss' "Disquisitiones generales circa superficies curvas,"* Vol. 62 of Astérisque. Paris: Société Mathématique de France. With the original text of Gauss.
Dray, T.
2015. *Differential Forms and the Geometry of General Relativity*. Boca Raton: CRC Press.
Earl, R.
2019. *Topology: A Very Short Introduction*. Oxford: Oxford University.
Feynman, R. P.
1985. *QED: the Strange Theory of Light and Matter.* Princeton, N.J.: Princeton University.
Feynman, R. P., R. B. Leighton, and M. L. Sands
1963. *The Feynman Lectures on Physics.* Reading, Mass.: Addison-Wesley.
Firby, P. and C. F. Gardiner

2001. *Surface Topology*, 3rd ed.: Cambridge:Woodhead.
Flanders, H.
1989. *Differential Forms with Applications to the Physical Sciences.* Mineola, N.Y.: Dover.
Forbes, N. and B. Mahon
2014. *Faraday, Maxwell, and the Electromagnetic Field: How Two Men Revolutionized Physics*. Amherst, N.Y.: Prometheus.
Fortney, J. P.
2018. *A Visual Introduction to Differential Forms and Calculus on Manifolds.* Basel: Birkhäuser.
Frank, F. C.
1958. *I. Liquid crystals: On the theory of liquid crystals.* Discuss. Faraday Soc., 25:19–28.
Frankel, T.
2011. *Gravitational Curvature: An Introduction to Einstein's Theory.* Mineola, N.Y.: Dover.
Frankel, T.
2012. *The Geometry of Physics: An Introduction*, 3rd ed. Cambridge: Cambridge University.
Fricke, R.
1926. *Lehrbuch der Algebra*, Vol. 2. Braunschweig: F. Vieweg und Sohn.
Gauss, C. F.
1827. *General Investigations of Curved Surfaces*. [Translated from the Latin and German by Adam Hiltebeitel and James Morehead, 1965 edition.] Hewlett, N.Y.: Raven Press.
Gauß, C. F.
1973. Werke. Band VIII. Hildesheim: Georg Olms Verlag. [Reprint of the 1900 original.]
Gindikin, S.
2007. *Tales of Mathematicians and Physicists*, 2nd English ed. Berlin: Springer.
Goodstein, D. L. and J. R. Goodstein
1996. *Feynman's Lost Lecture: The Motion of Planets Around the Sun*, 1st ed. New York: Norton.
Goodstein, J. R.
2018. *Einstein's Italian Mathematicians: Ricci, Levi-Civita, and the Birth of General Relativity*. Providence, R.I.: American Mathematical Society.
Gray, J.
1989. *Ideas of Space: Euclidean, Non-Euclidean, and Relativistic*, Oxford: Clarendon Press.
Greenberg, M. J.
2008. *Euclidean and Non-Euclidean Geometries*, 4th ed. New York:W.H. Freeman.
Guicciardini, N.
1999. *Reading the Principia.* Cambridge: Cambridge University.
The debate on Newton's mathematical methods for natural philosophy from 1687 to 1736.
Guicciardini, N.
2009. *Isaac Newton on Mathematical Certainty and Method* [Transformations: Studies in the History of Science and Technology]. Cambridge, Mass.: MIT.
Guillemin, V. and A. Pollack
1974. *Differential Topology.* Englewood Cliffs, N.J.: Prentice Hall.
Hankins, T. L.
1970. *Jean d'Alembert: Science and the Enlightenment*. Oxford: Clarendon Press.
Hayes, D. F., T. Shubin, G. L. Alexanderson, and P. Ross
2004. *Mathematical Adventures for Students and Amateurs.* Washington, D.C.: Mathematical

Association of America.
Henderson, D. W.
1998. *Differential Geometry: A Geometric Introduction.* Upper Saddle River, N.J.: Prentice Hall.
Henderson, D. W. and D. Taimina
2005. *Experiencing Geometry: Euclidean and non-Euclidean with History*, 3rd ed. Upper Saddle River, N.J.: Pearson Prentice Hall.
Hilbert, D.
1902. *The Foundations of Geometry.* [2nd English ed. of 10th German ed. Translated by Leo Unger. Revised and enlarged by Dr. Paul Bernays, published 1971.] Chicago: The Open Court.
Hilbert, D.
1952. *Geometry and the Imagination.* New York: Chelsea.
Hirsch, M. W.
1976. *Differential Topology*, Vol. 33. Berlin: Springer.
Hopf, H.
1935. *Über die drehung der tangenten und sehnen ebener kurven*. Compositio Mathematica, 2:50–62.
Hopf, H.
1946–1956. *Differential Geometry In The Large*: Seminar Lectures, New York University, 1946 And Stanford University, 1956, 2nd ed., 1983, Vol. 1000 of Lecture Notes In Mathematics. Berlin: Springer.
Hubbard, J. and B. Hubbard
2009. *Vector Calculus, Linear Algebra, and Differential Forms*, 4th ed. Ithaca, N.Y.: Matrix Editions.
James, I. M.
1999. *History of Topology*, 1st ed. Amsterdam: Elsevier.
Jones, B.
1870. *The Life and Letters of Faraday. London*: Longmans, Green and Co.
Katz, V.
1979. *The History of Stokes's Theorem*. Mathematics Magazine, 52(3):146–156.
Kerr, R. P.
2008. *Discovering the Kerr and Kerr-Schild metrics*. arXiv:0706.1109v2.
Klein, F.
1928. *Development of Mathematics in the 19th Century*
[English translation by M. Ackerman of the German original, with an extensive appendix by R. Hermann, published 1979]. Vol. 9 of Lie groups. Brookline, Mass.: Math Sci Press.
Knoebel, A.
2007. *Mathematical Masterpieces: Further Chronicles by the Explorers*, Undergraduate texts mathematics. Readings in mathematics. Berlin: Springer.
Koenderink, J. J.
1990. *Solid Shape*, Artificial intelligence. Cambridge, Mass.: MIT.
Kolmogorov, A. N. and A. P. Yushkevich
1996. *Mathematics of the 19th Century: Geometry, Analytic Function Theory*. Basel: Birkhäuser Verlag.
Kühnel, W.
2015. *Differential Geometry: Curves - Surfaces - Manifolds*, Vol. 77, 3rd ed. Providence, R.I.:

American Mathematical Society.
Lanczos, C.
1970. *The Variational Principles of Mechanics*, 4th ed. Mineola, N.Y.: Dover.
Laugwitz, D.
1999. *Bernhard Riemann*, 1826-1866: Turning Points In The Conception Of Mathematics. Boston: Birkhäuser.
Levi-Civita, T.
1917. Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana. In *Opere Matematiche: Memorie e Note, Pubblicate a cura dell'Accademia Nazionale dei Lincei*, Vol. IV (1917 to 1928), published 1960, pp. 1–39. Bologna: N. Zanichelli.
Levi-Civita, T.
1926. *The Absolute Differential Calculus*: (Calculus of Tensors). Mineola, N.Y.: Dover, 1977.
Levi-Civita, T.
1931. Calcolo differenzale assoluto. In *Opere Matematiche: Memorie e Note, Pubblicate a cura dell'Accademia Nazionale dei Lincei*, Vol. V (1929 to 1937), published 1970, pp. 219–229. Bologna: N. Zanichelli. Lightman, A. P.,W. H. Press, R. H. Price, and S. A.
Teukolsky
1975. *Problem Book in Relativity and Gravitation*. Princeton, N.J.: Princeton University.
Ludvigsen, M.
1999. *General Relativity: A Geometric Approach*. Cambridge: Cambridge University.
Mahon, B.
2017. *The Forgotten Genius of Oliver Heaviside: A Maverick of Electrical Science.* Buffalo, N.Y.: Prometheus.
Mahoney, M. S.
1994. *The Mathematical Career of Pierre de Fermat*, 1601–1665, 2nd ed. Princeton, N.J.: Princeton University.
Maxwell, J. C.
2002. *The Scientific Letters and Papers of James Clerk Maxwell*. Cambridge: Cambridge University.
Maxwell, J. C.
2003. *On the Transformation of Surfaces by Bending.* The Scientific Papers Of James Clerk Maxwell, Vol. 1, pp. 80–114. Mineola, N.Y.: Dover.
Mishchenko, A. S. and A. T. Fomenko
1988. *A Course of Differential Geometry and Topology*. Moscow: Mir Publishers.
Misner, C. W., K. S. Thorne, and J. A. Wheeler
1973. *Gravitation*. San Francisco: W. H. Freeman. A marvelous new hardback edition was published by Princeton University Press in 2017. It contains a new introduction by Charles Misner and Kip Thorne, discussing exciting developments in the field since the book's original publication.
Monastyrskiˇı, M. I.
1999. *Riemann, Topology, and Physics*, 2nd ed. Boston: Birkhäuser. Morgan, F. 1998. Riemannian Geometry: A Beginner's Guide, 2nd ed.Wellesley, Mass.: AK Peters.
Needham, T.
1993. *Newton and the transmutation of force.* Amer. Math. Monthly, 100(2):119–137.
Needham, T.

1997. *Visual Complex Analysis*. Oxford: The Clarendon Press.

Needham, T.

2014. *Visual Differential Geometry and Beltrami's Hyperbolic Plane.* In The Art of Science: From Perspective Drawing to Quantum Randomness, R. Lupacchini and A. Angelini, eds., pp. 71–99. Berlin: Springer.

Newton, I.

1687. *The Principia: Mathematical Principles of Natural Philosophy*, 1999 edition. Berkeley, Calif.: University of California. A new translation by I. Bernard Cohen and Anne Whitman, assisted by Julia Budenz, Preceded by "A guide to Newton's *Principia*" by Cohen.

Nye, J.

1999. *Natural Focusing and Fine Structure of Light: Caustics and Wave Dislocations*. Boca Raton: Taylor & Francis.

O'Neill, B.

2006. *Elementary Differential Geometry*, Revised 2nd ed. Amsterdam: Elsevier.

Pais, A.

1982. "*Subtle is the Lord. . .* ": The Science and the Life of Albert Einstein. Oxford: Oxford University.

Penrose, R.

1978. *The geometry of the universe*. In Mathematics Today: Twelve Informal Essays, L. A. Steen, ed., pp. 83–127. New York: Vintage Books.

Penrose, R.

1979. *The topology of ridge systems*. Annals of Human Genetics, 42(4):435–444. [Also reprinted in his Collected Works, Oxford University, 2010: Vol. 3, pp. 347–357.]

Penrose, R.

2005. *The Road to Reality*. New York: Alfred A. Knopf. A complete guide to the laws of the universe.

Penrose, R. and W. Rindler

1984. *Spinors and Space-Time*. Cambridge: Cambridge University.

Phillips, E. R.

1979. *Karl M. Peterson: The earliest derivation of the Mainardi–Codazzi equations and the fundamental theorem of surface theory*. Historia Mathematica, 6(2):137–163.

Poincaré, H.

1892, 1893, 1899. *New Methods of Celestial Mechanics*, Vol. 13. American Institute of Physics, published 1993.

Pólya, G.

1954. *Mathematics and Plausible Reasoning*. Princeton, N.J.: Princeton University.

Prasolov, V. V.

1995. *Intuitive Topology*, Vol. 4. Providence: American Mathematical Society.

Pressley, A.

2010. *Elementary Differential Geometry* [Springer Undergraduate Mathematics Series], 2nd ed. Berlin: Springer.

Richeson, D. S.

2008. *Euler's Gem: The Polyhedron Formula and the Birth of Topology.* Princeton, N.J.: Princeton University.

Rosenfeld, B. A.

1988. *A History of Non-Euclidean Geometry: Evolution of the Concept of a Geometric Space*, Vol 12. Berlin: Springer.

Schreiber, M.
1977. *Differential Forms: A Heuristic Introduction*, Universitext. Berlin: Springer.

Schutz, B. F.
1980. *Geometrical Methods of Mathematical Physics*. Cambridge: Cambridge University.

Schutz, B. F.
1986. *Determining the Hubble constant from gravitational wave observations*. Nature, 323(6086):310–311.

Schutz, B. F.
2003. *Gravity from the Ground Up*. Cambridge: Cambridge University.

Schutz, B. F.
2021. *A First Course in General Relativity,* 3rd ed. Cambridge: Cambridge University.

Segerman, H.
2016. *Visualizing Mathematics with 3D Printing*. Baltimore: Johns Hopkins.

Shapere, A. and F.Wilczek
1989. *Geometric Phases in Physics*, Vol. 5. Singapore: World Scientific.

Shenitzer, A. and J. Stillwell
2002. *Mathematical Evolutions*. Washington, D.C.: Mathematical Association of America.

Spivak, M.
1999. *A Comprehensive Introduction to Differential Geometry*, Vol. 2, 3rd ed. Houston: Publish or Perish.

Stahl, S.
1993. *The Poincaré Half-Plane: A Gateway to Modern Geometry*. Boston: Jones and Bartlett.

Sternberg, S.
2012. *Curvature in Mathematics and Physics*, Dover books on mathematics. Dover.

Stewart, G. W.
1993. *On the early history of the singular value decomposition*. SIAM Review, 35(4):551–566.

Stillwell, J.
1995. *Classical Topology and Combinatorial Group Theory*, Vol. 72, 2nd ed. Berlin: Springer.

Stillwell, J.
1996. *Sources of Hyperbolic Geometry*, Vol. 10 of History of Mathematics. Providence: American Mathematical Society.

Stillwell, J.
2005. *The Four Pillars of Geometry*. Berlin: Springer.

Stillwell, J.
2010. *Mathematics and its History,* 3d ed. Berlin: Springer.

Stoker, J. J.
1969. *Differential Geometry*, Vol. 20 of Pure and Applied Mathematics. New York:Wiley-Interscience.

Taylor, E. F. and J. A. Wheeler
1992. *Spacetime Physics: Introduction to Special Relativity*, 2nd ed. New York:W.H. Freeman.

Taylor, E. F. and J. A. Wheeler
2000. *Exploring Black Holes: Introduction to General Relativity*. San Francisco: Addison Wesley

Longman.
Thorne, K. S. and R. D. Blandford
2017. *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*. Princeton, N.J.: Princeton University.
Thurston, W. P.
1997. *Three-Dimensional Geometry and Topology,* Vol. 35. Princeton, N.J.: Princeton University.
Tu, L. W.
2017. *Differential Geometry: Connections, Curvature, and Characteristic Classes*. Berlin: Springer.
Wald, R. M.
1984. *General Relativity*. Chicago: University of Chicago.
Watson, G. N.
1917. *A problem of analysis situs*. Proceedings of the London Mathematical Society, 2(1): 227–242.
Weeks, J. R.
2020. *The Shape of Space*, ed. Boca Raton: CRC Press.
Weintraub, S. H.
1997. *Differential Forms: A Complement to Vector Calculus.* San Diego: Academic Press.
Westfall, R. S.
1980. *Never at Rest: A biography of Isaac Newton.* Cambridge: Cambridge University.
Wilson, P.M.H.
2008. *Curved Spaces: From Classical Geometries to Elementary Differential Geometry*. Cambridge: Cambridge University.
Wiltshire, D. L., M. Visser, and S. Scott
2009. *The Kerr Spacetime: Rotating Black Holes in General Relativity.* Cambridge: Cambridge University.
Woodward, L. M. and J. Bolton
2019. *A First Course in Differential Geometry: Surfaces in Euclidean Space.* Cambridge: Cambridge University.