

一份（不太）简短的 力学 笔记

或物理笔记之一

作者：wanzhao

日期：2026 年 2 月 14 日

课程：力学（H）

目录

目录	i	1.1.1 时间的计量	43
第零章 数学基础	1	1.1.2 空间的度量	44
0.1 一元函数微积分	1	1.2 物质世界的层次与数量级	44
0.1.1 微分	1	1.3 质点与运动	45
0.1.2 导数	2	1.3.1 质点与参考系	45
0.1.3 积分	6	第二章 狭义相对论	47
0.1.4 常微分方程	9	2.1 狭义相对论的基本原理	47
0.2 多元函数微积分	24	2.1.1 经典理论的危机	47
0.2.1 偏微商（偏导数）	24	2.1.2 狭义相对论基本原理	48
0.2.2 重积分	25	2.2 狭义相对论时空度量相对性	48
0.2.3 线积分、面积分与体积分	26	2.2.1 空间与时间的测量	48
0.3 行列式	29	2.2.2 惯性系间时空测量的相对性	49
0.3.1 行列式	29	2.3 狭义相对论时空变换及其推论	50
0.3.2 行列式的应用	30	2.3.1 狭义相对论时空变换	50
0.4 矢量的代数运算	31	2.3.2 狭义相对论时空变换的推论	52
0.4.1 矢量的叠加与分解	31	2.4 狭义相对论动力学	55
0.4.2 矢量的点积	33	2.4.1 引入	55
0.4.3 矢量的叉积	34	2.4.2 从力学的角度推导 ^[2]	56
0.4.4 矢量的多重积	35	2.4.3 从电动力学的角度推导 ^[3]	58
0.5 其他可能的数学工具	37	2.5 四维矢量与不变量	63
0.5.1 数列（序列）	37	2.5.1 时空间隔的不变性	63
0.5.2 不等式	37	2.5.2 闵可夫斯基空间与四维矢量	63
0.5.3 复数	38	2.5.3 四维速度	65
0.5.4 三角函数	39	2.5.4 四维动量	66
0.5.5 求和公式	39	2.5.5 四维力	67
0.5.6 双曲函数	40	2.5.6 不变量的应用	68
第一章 质点运动学	43	2.6 相对论的几何表示——「时空图」	70
1.1 时间与空间	43	参考文献	71

第零章 数学基础

物理的学习离不开数学，但学习时间的有限使得很多同学纠结，不知道数学应该学多深。在普通物理的范畴内，作者在这里提出一些小小的建议。

数学知识分两类。

第一类是需要完全理解和掌握的，比如数列、导数和积分、微分方程……它们在题目中出现的概率很高，当我们遇到时要能从容应对。

第二类是有助于物理解，但不要求完全掌握的，比如矢量分析与场论、张量代数……它们有助于加深你对物理情景的理解，但在普通物理的题目中出现的概率较低，这种知识需要学习，但学习的深度要视自身情况而定。

这一章虽然叫做数学基础，篇幅不长，但涵盖的内容并不少，因此并不需要通篇阅读。但是当你在物理学习的过程中遇到了数学上的困难时，可以回头查阅有关内容。倘若你能从我的注记中获得一些启发，那我的目的便达到了。

0.1 一元函数微积分

0.1.1 微分

一元函数可记为

$$y = y(x) \text{ 或 } y = f(x),$$

在它的连续区间内，如图 0.1

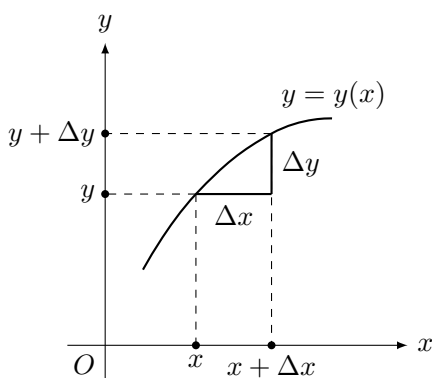


图 0.1

所示，自变量由 x 变到 $x + \Delta x$ ，相应的， y 由 $y(x)$ 变到 $y(x + \Delta x)$ ，则函数的增量定义为

$$\Delta y = y(x + \Delta x) - y(x).$$

自变量 $\Delta x \rightarrow 0$ 时, 称为自变量微分, 记为 dx , dx 是无穷小量, 不是零。但因为它是无穷小量, 它在与有限量的运算中, 在一些情况下可以视为 0 (但不是在所有情况下都可以视为 0)。在连续区间内, 自变量增量 $\Delta x \rightarrow dx$, 函数增量 $\Delta y \rightarrow dy$, 称为函数微分, 记为 dy , 它也是无穷小量。 dy 与 dx 之间的关系为

$$dy = y(x + dx) - y(x).$$

注释 0.1.1: 符号说明

显然, 因为 $y(x)$ 不一定是单调函数, 所以 Δy 和 dy 可以是正的、负的或零。但是只要你严格遵守定义与符号计算的基本规则, 就不会出错。这句话在一些微元分析的场景下尤其适用。

例 0.1.1: 重要近似

$$\sin x \sim \tan x \sim x, (x \rightarrow 0). \quad (0.1)$$

证明:

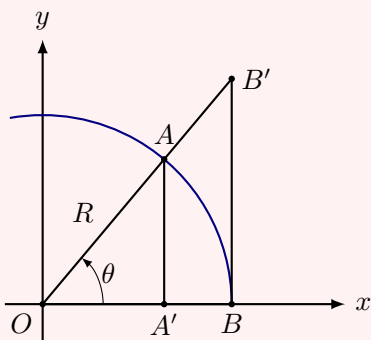


图 0.2: 重要近似的几何示意图

以 O 为原点建立直角坐标系, 绘出以 R 为半径的圆弧如图 0.2 所示, 其中圆心角 θ 对应的直线段 AA' , BB' , 圆弧 \widehat{AB} 的长度分别为

$$\overline{AA'} = R \sin \theta, \quad \overline{BB'} = R \tan \theta, \quad \widehat{AB} = R\theta.$$

$\theta \rightarrow 0$ 时, 有

$$\overline{AA'} \sim \overline{BB'} \sim \widehat{AB}.$$

化简得

$$\sin \theta \sim \tan \theta \sim \theta, (\theta \rightarrow 0).$$

□

0.1.2 导数

定义 0.1.1: 导数

函数 $y = f(x)$ 在点 x 处的导数定义为

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}, \quad (0.2)$$

如果极限存在, 则称函数在点 x 处可导。

显然, 导数的几何意义是函数图像在点 $(x, f(x))$ 处的切线的斜率。

在数学上可以证明导数与微分有如下关系:

定义 0.1.2: 微分与导数的关系

函数 $y = f(x)$ 在点 x 处可导, 则函数在该点的微分与导数的关系为

$$dy = f'(x) dx. \quad (0.3)$$

因此, 导数也称为函数的微商, 而求导和求微分实际上是等价的。

注释 0.1.2: 导数与微分的区别

导数是一个极限值, 是一个确定的数值, 而微分是函数增量的线性主部, 不是一个确定的数值。导数表示函数在某点处的变化率, 而微分表示函数在该点处变化量的线性近似。

并且导数不能视为 dy 与 dx 的比值, 这只是 Leibniz 记号下的一个形式上的表示方法。当然, 在非严格的物理推导中, 我们经常会把导数视为 dy 与 dx 的比值来进行计算。并且可以用类似的方法处理常微分中的微元计算, 这种做法 (形式计算) 在物理学中是被广泛接受的。

当然, 如果从更高级的数学角度来看, 我们的操作是有严格依据的, 这也保证了我们不会出错, 这里便不再深入了。

导数有一些重要性质, 举例如下:

示例 0.1.1: 导数的性质

$$(1) \quad (A_1 y_1 + A_2 y_2)' = A_1 y_1' + A_2 y_2'; \quad (0.4)$$

$$(2) \quad (y_1 y_2)' = y_1' y_2 + y_1 y_2'; \quad (0.5)$$

$$(3) \quad \left(\frac{y_1}{y_2} \right)' = \frac{y_1' y_2 - y_1 y_2'}{y_2^2}; \quad (0.6)$$

$$(4) \quad (y_1 \circ y_2)' = (y_1' \circ y_2) \cdot y_2'. \quad (0.7)$$

证明略去

注释 0.1.3: 复合函数记号的说明

公式 (0.7) 中的 $y_1 \circ y_2$ 表示复合函数, 即 $y_1(y_2(x))$ 。

其中, 式 公式 (0.7) 被称为复合函数的求导法则。在微分的形式下, 式 公式 (0.7) 可写为

$$\frac{dy_1}{dx} = \frac{dy_1}{dy_2} \cdot \frac{dy_2}{dx}. \quad (0.8)$$

这也在一定程度上为我们对导数执行的「消去」操作的正确性提供了依据。

常用的导数公式列举如下:

示例 0.1.2: 常用导数公式

$$(1) \quad (x^n)' = nx^{n-1}; \quad (0.9)$$

$$(2) \quad (a^x)' = a^x \ln a; \quad (0.10)$$

$$(3) \quad (\log_a x)' = \frac{1}{x \ln a}; \quad (0.11)$$

$$(4) \quad (\sin x)' = \cos x; \quad (0.12)$$

$$(5) \quad (\cos x)' = -\sin x; \quad (0.13)$$

此外, 还有几个常用的 n 阶导数公式:

示例 0.1.3: 常用 n 阶导数公式

$$(1) \quad (e^{ax})^{(n)} = a^n e^{ax}; \quad (0.14)$$

$$(2) \quad (\sin ax)^{(n)} = a^n \sin\left(ax + \frac{n\pi}{2}\right); \quad (0.15)$$

$$(3) \quad (\cos ax)^{(n)} = a^n \cos\left(ax + \frac{n\pi}{2}\right). \quad (0.16)$$

在数学上, 导数还可以用来讨论函数曲线的极值位置与凹凸性, 这在物理中经常用到, 而且有如下结论:

定义 0.1.3: 极值与凹凸性的判定

设函数 $y = f(x)$ 在点 x_0 处可导, 且 $f'(x_0) = 0$.

- 若 $f''(x_0) > 0$, 则 $f(x)$ 在点 x_0 处取得极小值, 且曲线在该点处向下凸, 称为「凸」;
- 若 $f''(x_0) < 0$, 则 $f(x)$ 在点 x_0 处取得极大值, 且曲线在该点处向上凸, 称为「凹」;
- 若 $f''(x_0) = 0$, 则不能确定 $f(x)$ 在点 x_0 处是否取得极值, 需考察更高阶导数。

注意到, 函数的凹凸与「凹」「凸」这两个字的字型是相反的, 这是因为函数的凸性与凸集有关, 而凸集的定义是明确的, 因此相反。

接下来, 我们介绍与导数有关的另一个重要工具——泰勒公式。

定理 0.1.1: 泰勒公式

设函数 $y = f(x)$ 在点 x_0 处具有 $n+1$ 阶导数, 则在点 x_0 附近, 函数 $f(x)$ 可展开为

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_n, \quad (0.17)$$

其中, 余项 R_n 有如下两种形式:

- 拉格朗日 (Lagrange) 余项形式:

$$R_n = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}, \quad \xi \text{ 在 } x \text{ 与 } x_0 \text{ 之间}; \quad (0.18)$$

- 皮亚诺 (Peano) 余项形式:

$$R_n = o((x-x_0)^n), \quad x \rightarrow x_0. \quad (0.19)$$

泰勒公式表明, 函数在某点处的值可以用该点处的各阶导数来近似表示. 上述定义中的余项 R_n 描述了近似的误差大小. 对于物理问题, 我们通常只需要前几项的近似, 因此余项的具体形式并不重要, 我们也不必过多纠结.

显然, 函数如果想展成泰勒级数, 至少要求通项趋于零, 即

$$\lim_{n \rightarrow \infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n = 0.$$

但事实上, 函数能否展成泰勒级数, 还需要满足更强的条件, 这里不再赘述.

函数 $f(x)$ 若能在点 $x = x_0$ 两侧某范围内展开成泰勒级数, 且级数在该范围内收敛于函数值, 便称这一范围为 $f(x)$ 的收敛区间. 例如, 在数学上可以证明:

示例 0.1.4: 常见函数的泰勒展开

$$(1) \quad e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in (-\infty, +\infty); \quad (0.20)$$

$$(2) \quad \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad x \in (-\infty, +\infty); \quad (0.21)$$

$$(3) \quad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \quad x \in (-\infty, +\infty); \quad (0.22)$$

$$(4) \quad \ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}, \quad x \in (-1, 1]; \quad (0.23)$$

$$(5) \quad (1+x)^m = \sum_{n=0}^{\infty} \binom{m}{n} x^n, \quad x \in (-1, 1). \quad (0.24)$$

注释 0.1.4: 二项式系数的说明

公式 (0.24) 中的二项式系数定义为

$$\binom{m}{n} = \frac{m(m-1)(m-2) \cdots [m-(n-1)]}{n!}.$$

特别地, 当 m 为非负整数时, 二项式系数可简化为

$$\binom{m}{n} = \begin{cases} \frac{m!}{n!(m-n)!}, & n \leq m; \\ 0, & n > m. \end{cases}$$

这与我们在组合数学中学过的排列数和组合数有如下关系:

$$A_m^n = n! \binom{m}{n}, \quad C_m^n = \binom{m}{n}.$$

其中 A_m^n 表示从 m 个不同元素中取出 n 个元素进行排列的总数, C_m^n 表示从 m 个不同元素中取出 n 个元素进行组合的总数. 需要满足 $m \geq n$.

$x_0 = 0$ 时, 称为麦克劳林 (Maclaurin) 级数, 即

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + R_n.$$

其中余项 R_n 同上.

这里还要引入另一个与导数有关的重要概念——曲率与曲率半径.

定义 0.1.4: 曲率与曲率半径

设函数 $y = f(x)$ 在点 x 处具有二阶导数, 则该点处的曲率定义为

$$\kappa = \frac{|f''(x)|}{[1 + (f'(x))^2]^{3/2}}. \quad (0.25)$$

曲率的倒数称为曲率半径, 记为 ρ , 即

$$\rho = \frac{1}{\kappa} = \frac{[1 + (f'(x))^2]^{3/2}}{|f''(x)|}. \quad (0.26)$$

对于物理问题, 曲率半径有时用来描述物体运动轨迹的弯曲程度。

在参数方程的形式下, 设曲线由参数 t 给出, 即

$$\begin{cases} x = \varphi(t); \\ y = \psi(t). \end{cases}$$

则曲率与曲率半径可表示为

$$\kappa = \frac{|\dot{x}\ddot{y} - \dot{y}\ddot{x}|}{(\dot{x}^2 + \dot{y}^2)^{3/2}}, \quad \rho = \frac{(\dot{x}^2 + \dot{y}^2)^{3/2}}{|\dot{x}\ddot{y} - \dot{y}\ddot{x}|}, \quad (0.27)$$

同时因极其常用, 给出椭圆的曲率半径公式, 设椭圆方程为

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

则椭圆上点 (x, y) 处的曲率半径为

$$\rho = \frac{(a^2b^2)^{3/2}}{(a^4y^2 + b^4x^2)^{3/2}}. \quad (0.28)$$

特别地, 椭圆的长短半轴处的曲率半径分别为

$$\rho_a = \frac{b^2}{a}, \quad \rho_b = \frac{a^2}{b}.$$

0.1.3 积分

有许多数学书上把积分分为了不定积分与定积分两类, 还有「积分是导数的逆运算」这样的说法。

但实际上, 积分的本质是求和, 而我更愿意把求不定积分看作是求原函数的问题, 而把定积分视为真正意义上的「积分」, 即求和。

定义 0.1.5: 不定积分

设函数 $F(x)$ 的导数为 $f(x)$, 即 $F'(x) = f(x)$, 则称函数 $F(x)$ 为函数 $f(x)$ 的一个原函数。函数 $f(x)$ 的所有原函数的集合称为函数 $f(x)$ 的不定积分, 记为

$$\int f(x) dx = F(x) + C,$$

其中, C 为任意常数。

根据 示例 0.1.1 中导数的性质, 可以得到不定积分的一些性质, 举例如下:

示例 0.1.5: 不定积分的性质

$$(1) \quad \int [A_1 f_1(x) + A_2 f_2(x)] dx = A_1 \int f_1(x) dx + A_2 \int f_2(x) dx; \quad (0.29)$$

$$(2) \quad \int f'(x)g(x) dx = f(x)g(x) - \int f(x)g'(x) dx; \quad (0.30)$$

$$(3) \quad \int f(g(x))g'(x) dx = \int f(u) du, \quad u = g(x). \quad (0.31)$$

同时, 根据 示例 0.1.2 中常用的导数公式, 可以得到一些常见函数的不定积分, 举例如下:

示例 0.1.6: 常见函数的不定积分

$$(1) \quad \int x^n dx = \frac{x^{n+1}}{n+1} + C, \quad (n \neq -1); \quad (0.32)$$

$$(2) \quad \int \frac{1}{x} dx = \ln |x| + C; \quad (0.33)$$

$$(3) \quad \int a^x dx = \frac{a^x}{\ln a} + C; \quad (0.34)$$

$$(4) \quad \int e^x dx = e^x + C; \quad (0.35)$$

$$(5) \quad \int \log_a x dx = x(\log_a x - \log_e a) + C; \quad (0.36)$$

$$(6) \quad \int \ln x dx = x(\ln x - 1) + C; \quad (0.37)$$

$$(7) \quad \int \sin x dx = -\cos x + C; \quad (0.38)$$

$$(8) \quad \int \cos x dx = \sin x + C. \quad (0.39)$$

有了这些, 我们就可以计算几乎所有可解 (可用初等函数表示) 的不定积分了。

定义 0.1.6: 定积分

设函数 $f(x)$ 在区间 $[a, b]$ 上有定义, 且在该区间上可积, 则称

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i) \Delta x_i, \quad (0.40)$$

其中, 区间 $[a, b]$ 被分成 n 个子区间 $[x_{i-1}, x_i]$, $\Delta x_i = x_i - x_{i-1}$, $\xi_i \in [x_{i-1}, x_i]$ 。若极限存在, 则称该极限为函数 $f(x)$ 在区间 $[a, b]$ 上的定积分。

这是定积分的黎曼 (Riemann) 定义, 它的几何意义是曲线 $y = f(x)$ 与 x 轴及直线 $x = a$ 和 $x = b$ 所围成的面积 (注意符号)。但是这个定义在实际计算中并不实用, 我们需要借助不定积分来计算定积分 (这也正是定积分这个概念被提出的原因)。

这就需要用到著名的微积分基本定理。

定理 0.1.2: 微积分基本定理

设函数 $f(x)$ 在区间 $[a, b]$ 上连续, 且 $F(x)$ 是 $f(x)$ 的一个原函数, 则

$$\int_a^b f(x) dx = F(b) - F(a). \quad (0.41)$$

证明:

设区间 $[a, b]$ 被分成 n 个子区间 $[x_{i-1}, x_i]$, $\Delta x_i = x_i - x_{i-1}$, $\xi_i \in [x_{i-1}, x_i]$ 。则有

$$\sum_{i=1}^n f(\xi_i) \Delta x_i = \sum_{i=1}^n [F(x_i) - F(x_{i-1})] = F(b) - F(a).$$

当 $n \rightarrow \infty$ 时, 右端不变, 因此

$$\int_a^b f(x) dx = F(b) - F(a).$$

□

上述证明中运用了微分中值定理, 亦即拉格朗日 (Lagrange) 中值定理, 可以自行查阅相关资料, 此处不再赘述。

由此, 我们可以通过求不定积分来计算定积分了。

类似于不定积分, 定积分也有一些性质, 举例如下:

示例 0.1.7: 定积分的性质

$$(1) \int_a^b [A_1 f_1(x) + A_2 f_2(x)] dx = A_1 \int_a^b f_1(x) dx + A_2 \int_a^b f_2(x) dx; \quad (0.42)$$

$$(2) \int_a^b f(x) dx = - \int_b^a f(x) dx; \quad (0.43)$$

$$(3) \int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (0.44)$$

注释 0.1.5: 积分变量的选择

在积分运算中, 积分变量是一个「虚拟变量」, 它可以是任何符号, 只要在积分式中前后一致即可。例如, 下面两个积分式是等价的:

$$\int_a^b f(x) dx = \int_a^b f(t) dt.$$

这是因为积分变量只是一个占位符号, 并不影响积分的结果。

下面给出一个定积分计算的例子, 即曲线段长度的计算。

例 0.1.2: 曲线段长度

设函数 $y = f(x)$ 在区间 $[a, b]$ 上连续且可导, 求曲线 $y = f(x)$ 在区间 $[a, b]$ 上的弧长。

证明:

设区间 $[a, b]$ 被分成 n 个子区间 $[x_{i-1}, x_i]$, $\Delta x_i = x_i - x_{i-1}$, $\xi_i \in [x_{i-1}, x_i]$ 。则曲线段在子区间 $[x_{i-1}, x_i]$ 上的弧长近似为

$$\Delta s_i = \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2} = \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i}\right)^2} \Delta x_i.$$

当子区间足够小时, 有

$$\frac{\Delta y_i}{\Delta x_i} \approx f'(\xi_i).$$

因此, 曲线段在区间 $[a, b]$ 上的弧长近似为

$$S \approx \sum_{i=1}^n \sqrt{1 + [f'(\xi_i)]^2} \Delta x_i.$$

当子区间无限细分时, 近似变为等于, 即

$$S = \int_a^b \sqrt{1 + [f'(x)]^2} dx.$$

□

注释 0.1.6: 微分形式下的曲线段长度

在一些书上也有类似于下面的表达式:

$$ds = \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$$

这种表达式其实是对微分形式的曲线段长度公式的非严格表达, 但也是形式计算奏效的一个例子。

注释 0.1.7: 有关于积分技巧

关于积分的计算, 笔者在写到这里时, 一直在思考是否需要介绍一些积分技巧, 比如双元法, 但最后还是选择放弃。一方面, 物理中的积分的计算往往并不复杂, 更多的是经验的积累, 而不是技巧的运用; 另一方面, 像双元法这样的技巧, 不论是深度还是广度, 都远远超出了普通物理的范畴, 这也违背了这份笔记的初衷。

因此笔者的建议是, 数值积分依靠计算器或计算机软件即可, 而符号积分则更多依赖于查表与经验积累。

另外还有一条经验, 就是面对一个困难的积分, 不妨把最复杂或者说「难看」的部分进行换元, 往往会有意想不到的效果。

0.1.4 常微分方程

定义 0.1.7: 常微分方程

设未知函数 $y = y(x)$ 及其各阶导数 $y', y'', \dots, y^{(n)}$ 之间存在某种关系, 可表示为

$$F(x, y, y', y'', \dots, y^{(n)}) = 0, \quad (0.45)$$

则称该方程为 n 阶常微分方程。方程中所含未知函数微商的最高阶数 n 称为微分方程的阶数。若方程关于 $y, y', y'', \dots, y^{(n)}$ 均是一次的, 则称该微分方程为 (n 阶) 线性微分方程。

一个函数 $y = y(x)$ 若在某区间内具有 n 阶导数, 且把该函数及其各阶导数代入微分方程后能使方程成立, 则称该函数为该微分方程在该区间内的一个解。因此当给定方程后, 最基本的事情当然是求出方程的解, 即求未知函数 $y = y(x)$ 。

从数学上看, 微分方程解的个数一般不唯一。例如, 最简单的一阶微分方程

$$\frac{dy}{dx} = f(x) \quad (0.46)$$

的通解为

$$y = \int f(x) dx = F(x) + C, \quad (0.47)$$

其中, $F(x)$ 是 $f(x)$ 的一个原函数, C 为任意常数。即对任意的常数 C , 函数 $y = F(x) + C$ 都是该微分方程的一个解。故称这样形式的解为该微分方程的「通解」。

如果事先要求解必须满足某些条件(初始条件或边界条件), 比如未知函数在一个特定的 x_0 处的值 $y(x_0) = y_0$, 则符合要求的解

$$y = \int_{x_0}^x f(x) dx + y_0 \quad (0.48)$$

是唯一的, 称为该微分方程的「特解」。

对于一般形式的 n 阶微分方程

$$F(x, y, y', y'', \dots, y^{(n)}) = 0, \quad (0.49)$$

它的通解通常包含 n 个独立的任意常数 C_1, C_2, \dots, C_n 。它的解无论是隐表示 $\Phi(x, y, C_1, C_2, \dots, C_n) = 0$ 还是显表示 $y = y(x, C_1, C_2, \dots, C_n)$, 都称为该微分方程的通解或通积分。常数 C_1, C_2, \dots, C_n 称为积分常数。而下列初值条件

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1} \quad (0.50)$$

一般来说决定了方程的一个特解。

从上面的讨论可以看出, 求解微分方程的过程实际上就是一个积分的过程。所以若微分方程的通解能用初等函数及初等函数的不定积分表示, 则称方程为可积微分方程, 而导出这种解的方法称为初等积分法。

然而, 能用初等积分法求解的微分方程只是微分方程中的一小部分, 大部分方程的求解都比较复杂, 甚至没有解析解。尽管如此这一基本方法在微分方程乃至物理学中仍然是非常重要的。我们将会简要介绍一些用初等积分法求解微分方程的解法。而对于(常)微分方程的一般理论, 这里便不再深入了。

首先, 我们介绍可分离变量的微分方程。

定义 0.1.8: 可分离变量的微分方程

设微分方程可化为如下形式:

$$\frac{dy}{dx} = g(x)h(y), \quad (0.51)$$

则称该微分方程为可分离变量的微分方程。

求解该类微分方程的方法如下:

1. 若 $h(y) = 0$ 有解 $y = y_0$, 则 $y = y_0$ 是该微分方程的一个解;
2. 当 $h(y) \neq 0$ 时, 可将方程变形为

$$\frac{1}{h(y)} dy = g(x) dx, \quad (0.52)$$

然后对两边分别积分, 得

$$\int \frac{1}{h(y)} dy = \int g(x) dx + C, \quad (0.53)$$

其中, C 为任意常数。由此可得该微分方程的通解。

其次, 我们介绍齐次微分方程。首先, 我们给出齐次函数的定义。

定义 0.1.9: 齐次函数

设函数 $f(x, y)$ 满足如下条件:

$$f(tx, ty) = t^n f(x, y), \quad (0.54)$$

其中, t 为任意常数, n 为非负整数, 则称函数 $f(x, y)$ 为 n 阶齐次函数。

基于齐次函数的定义, 我们可以给出齐次微分方程的定义。

定义 0.1.10: 齐次微分方程

对于一阶微分方程

$$P(x, y) dx + Q(x, y) dy = 0, \quad (0.55)$$

称为齐次的, 当且仅当 $P(x, y)$ 与 $Q(x, y)$ 均为同阶齐次函数。现在 $\frac{P(x, y)}{Q(x, y)}$ 为零阶齐次函数, 因此可化为 $\varphi\left(\frac{y}{x}\right)$ 的形式。于是该微分方程可化为如下形式:

$$\frac{dy}{dx} = \varphi\left(\frac{y}{x}\right), \quad (0.56)$$

而称该微分方程为齐次微分方程。

求解该类微分方程的方法如下:

1. 作变量代换 $u = \frac{y}{x}$, 则 $y = ux$, 从而有

$$\frac{dy}{dx} = u + x \frac{du}{dx}.$$

2. 将上式代入微分方程, 得

$$u + x \frac{du}{dx} = \varphi(u),$$

即

$$x \frac{du}{dx} = \varphi(u) - u.$$

3. 分离变量, 得

$$\frac{1}{\varphi(u) - u} du = \frac{1}{x} dx.$$

然后对两边分别积分, 得

$$\int \frac{1}{\varphi(u) - u} du = \int \frac{1}{x} dx + C,$$

其中, C 为任意常数。由此可得该微分方程的通解。

最后, 我们介绍一阶线性微分方程

定义 0.1.11: 一阶线性微分方程

对于一阶微分方程, 若未知函数 y 及其导数 $\frac{dy}{dx}$ 都是一次的, 则称该微分方程为一阶线性微分方程。它可化为标准形式:

$$\frac{dy}{dx} + P(x)y = Q(x), \quad (0.57)$$

其中, $P(x)$ 与 $Q(x)$ 为已知函数。若右端的 $Q(x) = 0$, 则称该微分方程为 (一阶) 线性齐次微分方程, 否则称为 (一阶) 线性非齐次微分方程。注意到, 这里的「齐次」与前面介绍的齐次微分方程的「齐次」是不同的概念。

求解该类微分方程的方法如下:

1. 求解对应的齐次方程

$$\frac{dy}{dx} + P(x)y = 0,$$

它的通解为

$$y_h = Ce^{-\int P(x)dx}, \quad (0.58)$$

其中, C 为任意常数。

2. 求该非齐次方程的一个特解 y_p , 可用如下方法求解, 即设

$$y_p = u(x)e^{-\int P(x)dx}, \quad (0.59)$$

其中, $u(x)$ 为待定函数。将 y_p 代入非齐次方程, 得

$$\frac{du}{dx} = Q(x)e^{\int P(x)dx}.$$

分离变量并积分, 得

$$u(x) = \int Q(x)e^{\int P(x)dx} dx.$$

因此, 非齐次方程的一个特解为

$$y_p = e^{-\int P(x)dx} \int Q(x)e^{\int P(x)dx} dx. \quad (0.60)$$

3. 由此可得该非齐次微分方程的通解为

$$y = y_h + y_p = Ce^{-\int P(x)dx} + e^{-\int P(x)dx} \int Q(x)e^{\int P(x)dx} dx. \quad (0.61)$$

比较, 公式 (0.58) 与 公式 (0.59) 可以发现, 解非齐次方程 公式 (0.57) 所做的代换, 可视为对齐次方程通解中的积分常数 C 进行函数化处理, 即将 C 替换为待定函数 $u(x)$, 因此, 这种方法称为「常数变易法」。

此外, 从通解 公式 (0.61) 可以看出: 非齐次微分方程的通解等于对应齐次微分方程的通解 $Ce^{-\int P(x)dx}$ 与非齐次方程的一个特解 $e^{-\int P(x)dx} \int Q(x)e^{\int P(x)dx} dx$ 之和。这件事情在更高阶的线性微分方程中也成立, 这一性质称为「叠加原理」, 我们之后会再提到它。

再此基础上, 我们介绍二阶微分方程中的特例——可降阶微分方程。

一般的二阶微分方程可表示为

$$F(x, y, y', y'') = 0. \quad (0.62)$$

这里介绍两种特殊类型的二阶微分方程, 它们都可以通过变量代换将方程降阶, 化为一阶微分方程来求解。

1. 若方程中不显含未知函数 y , 即方程可表示为

$$F(x, y', y'') = 0, \quad (0.63)$$

则可作变量代换 $p = y'$, 则 $y'' = \frac{dp}{dx}$ 。于是方程变为

$$F(x, p, \frac{dp}{dx}) = 0, \quad (0.64)$$

它是关于 p 的一阶微分方程, 求解出 $p = p(x)$ 后, 再积分即可得到 $y = \int p(x) dx + C$ 。

2. 若方程中不显含自变量 x , 即方程可表示为

$$F(y, y', y'') = 0, \quad (0.65)$$

则可作变量代换 $p = y'$, 则 $y'' = \frac{dp}{dx} = \frac{dp}{dy} \cdot \frac{dy}{dx} = p \frac{dp}{dy}$ 。于是方程变为

$$F(y, p, p \frac{dp}{dy}) = 0, \quad (0.66)$$

它是关于 p 的一阶微分方程, 求解出 $p = p(y)$ 后, 再积分即可得到 $x = \int \frac{1}{p(y)} dy + C$ 。

对于一般的二阶微分方程, 我们无法, 也没有必要做过于深入的讨论。但是我们可以再此处简单介绍一下二阶线性微分方程解的结构, 并给出它的解法。

定义 0.1.12: 二阶线性微分方程

设二阶微分方程可化为如下形式:

$$\frac{d^2 y}{dx^2} + P(x) \frac{dy}{dx} + Q(x)y = f(x), \quad (0.67)$$

其中, $P(x), Q(x), f(x)$ 为已知函数, 则称该微分方程为二阶线性微分方程。若右端的 $f(x) = 0$, 相应的齐次方程为

$$\frac{d^2 y}{dx^2} + P(x) \frac{dy}{dx} + Q(x)y = 0, \quad (0.68)$$

该微分方程被称为二阶线性齐次微分方程, 否则被称为二阶线性非齐次微分方程。

关于二阶线性微分方程初值问题的解的存在性与唯一性, 有如下定理:

定理 0.1.3: 二阶线性微分方程初值问题的解的存在性与唯一性

设函数 $P(x), Q(x), f(x)$ 在区间 I 上连续, 则对于任意 $x_0 \in I$ 及任意常数 α, β , 二阶线性微分方程初值问题

$$\begin{cases} \frac{d^2 y}{dx^2} + P(x) \frac{dy}{dx} + Q(x)y = f(x), \\ y(x_0) = \alpha, y'(x_0) = \beta, \end{cases}$$

在 x_0 附近存在唯一解。特别地, 对齐次方程 (即 $f(x) = 0$) 的初值问题, 满足初值 $y(x_0) = 0, y'(x_0) = 0$ 的唯一解为零解 $y = 0$ 。

这里略去定理的证明, 它超出了普通物理学, 乃至大多数高校物理专业数学分析课程的范围。

注释 0.1.8: 初值问题与边值问题的区别

看到这里, 读者可能会觉得奇怪, 为什么 $y(x)$ 与 $y'(x)$ 都要在相同的点 x_0 处给出初值, 而不能够分别在不同的点 x_0, x_1 处给出初值 $y(x_0) = \alpha, y'(x_1) = \beta$ 。

这就涉及到定解问题的类型划分了, 简单来说, 如果这 n 个条件给的点 (x 的值) 不完全相同, 这就不再是一个初值问题了, 而是一个边值问题。

关于边值问题, 有简单的结论如下:

1. 存在唯一性不再像初值问题那样有普适保证对于初值问题 (所有条件都在同一点 x_0 处给出), 只要系数函数 $P(x), Q(x), f(x)$ 在区间 I 上连续, 则在 x_0 附近总是存在唯一解。但对于边值问题 (条件分布在 x_0, x_1, \dots 处), 情况就复杂得多了:

- 可能没有解: 例如, 方程 $y'' + y = 0$ 在区间 $[0, \pi]$ 上满足边界条件 $y(0) = 0, y(\pi) = 1$ 的解不存在;

- 可能有唯一解。
- 可能有无穷多个解：例如，方程 $y'' + y = 0$ 在区间 $[0, \pi]$ 上满足边界条件 $y(0) = 0, y(\pi) = 0$ 的解有无穷多个。

2. 物理意义不同

- 初值问题通常用于描述系统的时间演化过程，例如经典力学中的运动方程，给定初始位置和速度，可以确定物体的运动轨迹。
- 边值问题通常用于描述空间分布问题，例如热传导方程和波动方程，给定边界条件，可以确定系统在空间中的状态分布。

3. 数学处理方法不同

- 初值问题通常使用常规的微分方程求解方法，如变量分离法、积分因子法等。
- 边值问题通常需要使用特殊的方法，如特征值问题、傅里叶级数展开等。
- 边值问题往往涉及到线性代数和函数空间的概念，例如希尔伯特空间和巴拿赫空间。

所以，结论是： n 个条件依然可以用来确定 n 个常数，但前提是这组条件必须是相容的。并不是随便在纸上画 n 个点就能找到一个解同时经过它们，这取决于微分方程本身的性质（特征值等）。

对于齐次方程 公式 (0.68) 很容易验证下面的结论：

定理 0.1.4: 齐次二阶线性微分方程解的叠加原理

设 $y_1(x)$ 与 $y_2(x)$ 是齐次二阶线性微分方程的两个解，则对于任意常数 C_1, C_2 ，函数

$$y(x) = C_1 y_1(x) + C_2 y_2(x)$$

也是该微分方程的解。

因此，齐次线性方程的解集具有线性结构。对于非齐次方程 公式 (0.67)，容易验证如下定理：

定理 0.1.5: 非齐次二阶线性微分方程解的结构

设 $y_p(x)$ 是非齐次二阶线性微分方程的一个特解， $y_h(x)$ 是对应齐次方程的通解，则非齐次方程的通解为

$$y(x) = y_h(x) + y_p(x).$$

除此之外，我么还需要给出如下两个定义：

定义 0.1.13: 线性无关

设函数 $y_1(x), y_2(x)$ 在区间 I 上有定义，若不存在常数 C_1, C_2 ，使得

$$C_1 y_1(x) + C_2 y_2(x) = 0$$

对区间 I 上的任意 x 都成立，则称函数 $y_1(x), y_2(x)$ 在区间 I 上线性无关，否则称为线性相关。

定义 0.1.14: 基本解组

若函数 $y_1(x)$ 与 $y_2(x)$ 是齐次方程 公式 (0.68) 的一对线性无关的解, 则该方程的任何解均可表示为

$$y(x) = C_1 y_1(x) + C_2 y_2(x)$$

的形式, 其中, C_1, C_2 为任意常数。

而且, 这样的一对线性无关的解 $y_1(x), y_2(x)$ 称为该齐次方程的一个基本解组。

但大多数时候, 两组解是否线性无关并不容易判断。这需要引入 Wronskian 行列式的概念, 它定义如下:

定义 0.1.15: Wronskian 行列式

设函数 $y_1(x), y_2(x)$ 在区间 I 上具有二阶导数, 则称如下行列式为函数 $y_1(x), y_2(x)$ 的 Wronskian 行列式:

$$W(y_1, y_2) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} = y_1(x)y_2'(x) - y_2(x)y_1'(x).$$

基于 Wronskian 行列式, 我们可以给出如下判别线性相关的定理:

定理 0.1.6: 判别线性无关的定理

设函数 $y_1(x), y_2(x)$ 是齐次方程 公式 (0.68) 的两个解, 则它们在区间 I 上线性相关的充分必要条件是它们的 Wronskian 行列式在区间 I 上恒为零。

注释 0.1.9: 关于二阶线性微分方程解的结构说明

关于以上给出的关于二阶线性微分方程解的结构定义与定理, 读者大可不必过于纠结这些定义与定理的细节, 毕竟它们超出了大多数物理学课程的范围。

况且在笔者学习普通物理学的实践中, 这些内容也极少被用到。这里只是为了让读者对二阶线性微分方程的解的结构有一个大致的了解, 以便在后续的物理学学习中遇到相关内容时不会感到陌生。

另外, 对于形如 公式 (0.68) 的齐次二阶线性微分方程, 实际上, 我们只要知道了它的一个解, 就可以通过 Liouville 公式求出另一个线性无关的解, 从而得到该方程的基本解组, 进而得到该方程的通解。具体的操作方法可参考相关数学分析教材, 此处不再赘述。

在上述基础上, 我们针对二阶常系数线性微分方程, 讨论如何求基本解组的问题。

定义 0.1.16: 二阶常系数线性微分方程

设二阶线性微分方程可化为如下形式:

$$\frac{d^2 y}{dx^2} + a \frac{dy}{dx} + by = f(x), \quad (0.69)$$

其中, a, b 为常数, $f(x)$ 为已知函数, 则称该微分方程为二阶常系数线性微分方程。若右端的 $f(x) = 0$, 相应的齐次方程为

$$\frac{d^2 y}{dx^2} + a \frac{dy}{dx} + by = 0, \quad (0.70)$$

该微分方程被称为二阶常系数线性齐次微分方程, 否则被称为二阶常系数线性非齐次微分方程。

求解该类微分方程的方法如下:

1. 求解对应的齐次方程 公式 (0.70)。设齐次方程的解为 $y = e^{\lambda x}$ 的形式, 将其代入齐次方程, 得特征方程

$$\lambda^2 + a\lambda + b = 0.$$

解该特征方程, 得两个根 λ_1, λ_2 。根据根的不同情况, 可分三种情形讨论:

(a) 若 $\lambda_1 \neq \lambda_2$, 则齐次方程的基本解组为

$$y_1 = e^{\lambda_1 x}, \quad y_2 = e^{\lambda_2 x}.$$

(b) 若 $\lambda_1 = \lambda_2 = \lambda$, 则齐次方程的基本解组为

$$y_1 = e^{\lambda x}, \quad y_2 = xe^{\lambda x}. \text{ (这是容易验证的)}$$

(c) 若 $\lambda_{1,2} = \alpha \pm \beta i$ (其中, $i = \sqrt{-1}$), 则齐次方程的基本解组为

$$y_1 = e^{\alpha x} \cos(\beta x), \quad y_2 = e^{\alpha x} \sin(\beta x).$$

由此可得齐次方程的通解为

$$y_h = C_1 y_1 + C_2 y_2,$$

其中, C_1, C_2 为任意常数。

2. 求该非齐次方程的一个特解 y_p , 可用如下方法求解:

(a) 若 $f(x)$ 为指数函数、多项式函数、三角函数或它们的有限和与积的形式, 则可用「待定系数法」求解, 即设

$$y_p = y_p^{(1)} + y_p^{(2)} + \dots + y_p^{(n)},$$

其中, 每一项 $y_p^{(i)}$ 对应 $f(x)$ 中的一项, 且 $y_p^{(i)}$ 的形式与该项相同, 但含有待定系数。将 y_p 代入非齐次方程, 解出待定系数, 即可得到该非齐次方程的一个特解。需要注意的是, 若 $y_p^{(i)}$ 中的某一项与齐次方程的某一解形式相同, 则需将该项乘以 x 的适当幂次, 以确保 $y_p^{(i)}$ 与齐次方程的解线性无关。

(b) 否则, 可用「常数变易法」求解, 即设

$$y_p = u_1(x)y_1(x) + u_2(x)y_2(x),$$

其中, $u_1(x), u_2(x)$ 为待定函数, $y_1(x), y_2(x)$ 为齐次方程的基本解组。将 y_p 代入非齐次方程, 得

$$\begin{cases} u_1'(x)y_1(x) + u_2'(x)y_2(x) = 0, \\ u_1'(x)y_1'(x) + u_2'(x)y_2'(x) = f(x). \end{cases}$$

解出 $u_1'(x), u_2'(x)$, 即

$$u_1'(x) = -\frac{y_2(x)f(x)}{W(y_1, y_2)}, \quad u_2'(x) = \frac{y_1(x)f(x)}{W(y_1, y_2)},$$

其中, $W(y_1, y_2)$ 为 $y_1(x), y_2(x)$ 的 Wronskian 行列式。然后对 $u_1'(x), u_2'(x)$ 分别积分, 得

$$u_1(x) = -\int \frac{y_2(x)f(x)}{W(y_1, y_2)} dx, \quad u_2(x) = \int \frac{y_1(x)f(x)}{W(y_1, y_2)} dx.$$

因此, 非齐次方程的一个特解为

$$y_p = -y_1(x) \int \frac{y_2(x)f(x)}{W(y_1, y_2)} dx + y_2(x) \int \frac{y_1(x)f(x)}{W(y_1, y_2)} dx.$$

3. 由此可得该非齐次微分方程的通解为

$$y = y_h + y_p = C_1 y_1 + C_2 y_2 + y_p.$$

类似的, 关于二阶常系数齐次线性微分方程的求解方法不难推广到 n 阶常系数齐次线性微分方程

$$\frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_1 \frac{dy}{dx} + a_0 y = 0, \quad (0.71)$$

按实根、重根与复根的不同情况, 可得如下结论:

定理 0.1.7: n 阶常系数齐次线性微分方程的基本解组

设 n 阶常系数齐次线性微分方程的特征方程为

$$\lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0 = 0,$$

它有 k 个互不相同的实根 $\lambda_1, \lambda_2, \dots, \lambda_k$, 其中, 第 i 个实根 λ_i 的重数为 m_i ($i = 1, 2, \dots, k$), 则该微分方程的基本解组为

$$\{e^{\lambda_1 x}, x e^{\lambda_1 x}, \dots, x^{m_1-1} e^{\lambda_1 x}, e^{\lambda_2 x}, x e^{\lambda_2 x}, \dots, x^{m_2-1} e^{\lambda_2 x}, \dots, e^{\lambda_k x}, x e^{\lambda_k x}, \dots, x^{m_k-1} e^{\lambda_k x}\}.$$

若 $\lambda_{k+1} = \alpha + \beta i$ 与 $\lambda_{k+2} = \alpha - \beta i$ ($\beta > 0$) 为一对共轭复根, 且它们的重数均为 m , 则该微分方程的基本解组还应包括

$$\{e^{\alpha x} \cos(\beta x), x e^{\alpha x} \cos(\beta x), \dots, x^{m-1} e^{\alpha x} \cos(\beta x), e^{\alpha x} \sin(\beta x), x e^{\alpha x} \sin(\beta x), \dots, x^{m-1} e^{\alpha x} \sin(\beta x)\}.$$

则方程 公式 (0.71) 的通解为

$$y = C_1 e^{\lambda_1 x} + C_2 x e^{\lambda_1 x} + \dots + C_{m_1} x^{m_1-1} e^{\lambda_1 x} + C_{m_1+1} e^{\lambda_2 x} + C_{m_1+2} x e^{\lambda_2 x} + \dots + C_{m_1+m_2} x^{m_2-1} e^{\lambda_2 x} + \dots$$

通解的系数 C_1, C_2, \dots, C_n 可由初值条件确定。

齐次方程通解的求解尚且如此, 对于非齐次方程, 也就是通解的计算, 类比二阶情形, 我们可以想象其复杂度会大幅提升。

为此, 我们引入一种计算方法——算子法。

定义 0.1.17: 微分算子

设 $D = \frac{d}{dx}$, 则称 D 为微分算子。由此可定义 $D^n = \frac{d^n}{dx^n}$ 。

注释 0.1.10: 关于微分算子的说明

在如下的推导中, 有诸多地方并不严格, 读者大可不必过于纠结这些细节。一来算子理论是有严格数学基础的, 亦即若尔当 (Jordan) 标准形等线性代数的内容; 二来在我学习普通物理学的实践中, 你可能会用到算子法的情形并不会触及这些所谓严格的边界。

于是, n 阶常系数线性微分方程可表示为

$$(D^n + a_{n-1} D^{n-1} + \dots + a_1 D + a_0) y = f(x). \quad (0.72)$$

设特征多项式 $P(\lambda) = \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0$, 则可将上式写为

$$P(D) y = f(x). \quad (0.73)$$

首先, 我们研究 $f(x) = e^{ax}$ 的情形, 其中, a 为常数。

若 $P(a) \neq 0$, 则有如下定理:

定理 0.1.8: 指数代换法则

设 $P(D)$ 为 n 阶常系数线性微分算子, a 为常数, 且 $P(a) \neq 0$, 则

$$P(D)e^{ax} = P(a)e^{ax}. \quad (0.74)$$

证明:

注意到 $P(D)$ 是由 D 的多项式构成的, 因此只需证明

$$D^n e^{ax} = a^n e^{ax}.$$

这是显然的。 □

由此可得如下结论:

定理 0.1.9: 指数输入定理

设 $P(D)$ 为 n 阶常系数线性微分算子, a 为常数, 且 $P(a) \neq 0$, 则方程

$$P(D)y = e^{ax}$$

的一个特解为

$$y_p = \frac{1}{P(a)} e^{ax}. \quad (0.75)$$

证明:

由指数代换法则, 有

$$P(D)y_p = P(D)\left(\frac{1}{P(a)}e^{ax}\right) = \frac{1}{P(a)}P(D)e^{ax} = \frac{1}{P(a)}P(a)e^{ax} = e^{ax}.$$

□

注意到三角函数是可以用指数函数表示的, 即欧拉 (Euler) 公式

$$e^{ix} = \cos x + i \sin x, \quad (0.76)$$

而这有什么用呢? 由于复变函数的求导可以简单理解为分别对实部与虚部求导, 于是就容易证明对于一个在复数域上的微分方程 (但 $P(x) \in \mathbb{R}[x]$, 就是 $P(x)$ 为实系数多项式)

$$P(D)y = f(x) \quad (0.77)$$

的特解的实部恰好是

$$P(D)y = \Re(f(x)) \quad (0.78)$$

的特解, 同样虚部恰好是

$$P(D)y = \Im(f(x)) \quad (0.79)$$

的特解。所以对于 $f(x)$ 为三角函数的情形完全可以转化为指数函数的情形来处理, 最后根据要求保留实部或虚部即可。所以我们讨论的 e^{ax} 中的 a 可以是任意复数, 即 $a \in \mathbb{C}$ 。

到现在其实已经得到了一个挺不错的结果了, 但这其实还不够。注意到指数输入 定理 0.1.9 中有一个限制条件 $P(a) \neq 0$, 因为分母显然不能为零。但是在实际中, $P(a) = 0$ 的情况比比皆是。所以有如下定理:

定理 0.1.10: 指数位移法则

设 $P(D)$ 为 n 阶常系数线性微分算子, a 为常数, $u(x)$ 为可微函数, 则

$$P(D)(e^{ax}u(x)) = e^{ax}P(D+a)u(x). \quad (0.80)$$

证明:

注意到 $P(D)$ 是由 D 的多项式构成的, 因此只需证明

$$D^n(e^{ax}u(x)) = e^{ax}(D+a)^n u(x).$$

那么很明显可以运用数学归纳法

1. 当 $n=1$ 时, 显然成立:

$$D(e^{ax}u(x)) = e^{ax}(D+a)u(x).$$

2. 假设当 $n=k$ 时成立, 即

$$D^k(e^{ax}u(x)) = e^{ax}(D+a)^k u(x).$$

则当 $n=k+1$ 时, 有

$$\begin{aligned} D^{k+1}(e^{ax}u(x)) &= D(D^k(e^{ax}u(x))) \\ &= D(e^{ax}(D+a)^k u(x)) \\ &= e^{ax}(D+a)(D+a)^k u(x) \\ &= e^{ax}(D+a)^{k+1} u(x). \end{aligned}$$

由数学归纳法原理, 定理得证。 □

这个命题的作用将会在后面揭晓。

一般我们只会讨论二阶常系数微分方程, 因为正常人是不会解三次及以上 (甚至五次及以上是没有根式解) 的方程的, 所以我们先由二阶微分方程入手。

显然如果 $P(a) = 0$, 则 a 必为 $P(x) = 0$ 的一个根。而学过待定系数法后也知道应该分两种情况讨论, 即 a 是单根与重根的情形。

定理 0.1.11: 单根指数输入定理

若 a 为 $P(x) = 0$ 的单根, 则有:

$$y_p = \frac{xe^{ax}}{P'(a)}. \quad (0.81)$$

证明:

不妨设 a, b ($a \neq b$) 为 $P(x) = 0$ 的两个根, 为方便起见, 我们假设 $P(x)$ 的最高次项系数为 1。那么有:

$$P(D) = (D-a)(D-b)$$

所以

$$P'(D) = 2D - (a+b)$$

故有 $P'(a) = a - b$, 于是将 y_p 代入方程可得:

$$\begin{aligned} P(D)y_p &= \frac{e^{ax}P(D+a)x}{P'(a)} \quad (\text{由指数位移法则}) \\ &= \frac{e^{ax}(D+a-b)Dx}{a-b} \\ &= \frac{e^{ax}(a-b)}{a-b} \\ &= e^{ax}. \end{aligned}$$

□

类似地, 有如下定理:

定理 0.1.12: 二重根指数输入定理

若 a 为 $P(x) = 0$ 的二重根, 则有:

$$y_p = \frac{x^2 e^{ax}}{P''(a)}. \quad (0.82)$$

证明类似单根情形, 此处略去。

由此, 我们可以推出如下结论:

定理 0.1.13: n 重根指数输入定理

若 $a \in \mathbb{C}$ 为 $P(x) = 0$ 的 n 重根, 则有:

$$y_p = \frac{x^n e^{ax}}{P^{(n)}(a)}. \quad (0.83)$$

证明:

由于 a 为 $P(D) = 0$ 的 n 重根, 因此 $P(D)$ 可表示为

$$P(D) = (D - a)^n \tilde{P}(D)$$

则:

$$P^{(n)}(D) = n! \tilde{P}(D) + (D - a)Q(D)$$

其中, $Q(D)$ 为某个微分算子。因此, $P^{(n)}(a) = n! \tilde{P}(a)$ 。于是, 将 y_p 代入方程可得:

$$\begin{aligned} P(D)y_p &= \frac{e^{ax}P(D+a)x^n}{P^{(n)}(a)} \quad (\text{由指数位移法则}) \\ &= \frac{e^{ax}\tilde{P}(D+a)(D)^n x^n}{P^{(n)}(a)} \\ &= \frac{e^{ax}\tilde{P}(a)n!}{n!\tilde{P}(a)} \\ &= e^{ax}. \end{aligned}$$

□

而当 $f(x)$ 为多项式函数时, 我们先考虑一个例子:

例 0.1.3: 算子法求特解——多项式函数情形

求微分方程

$$-y' + y = x$$

的一个特解。

解:

显然我们可以轻易的构造出一个特解 $y_p = x + 1$ 。这个方程用算子法可以写作:

$$(-D + 1)y = x.$$

如果大胆地把算子除到右边, 会得到

$$y = \frac{x}{-D + 1}.$$

如果联想到 $\frac{1}{1-D}$ 的泰勒级数:

$$\frac{1}{1-D} = 1 + D + D^2 + D^3 + \dots,$$

并更大胆地将其代入:

$$y = (1 + D + D^2 + D^3 + \dots)x = x + 1.$$

这正是我们之前构造的特解。 ■

但是根据我们在前面讨论的内容, 我们会知道

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad (0.84)$$

是有条件的, 即当且仅当 $|x| < 1$ 时才成立。但是 D 是一个算子, 根本没有什么度量概念! 所以其实我们是在考虑 $\frac{1}{P(D)}$ 的形式幂级数, 也就是不考虑其收敛性, 只关注它的形式。

关于这么做是否正确, 我可以得出一个简单的说明 (并非严格证明):

注释 0.1.11: 关于算子法中形式幂级数的说明

如果考虑 $1 + D + D^2 + \dots + D^{n-1}$ 这 n 项, 那么就有:

$$(1 - D)(1 + D + D^2 + \dots + D^{n-1}) = (1 - D)\frac{1 - D^n}{1 - D} = 1 - D^n.$$

如果记 $Q_n(D) = 1 + D + D^2 + \dots + D^{n-1}$, 那么我们期望的答案是

$$Q(D)f(x) = \lim_{n \rightarrow \infty} Q_n(D)f(x).$$

那么将其带回原方程, 并且注意到将算子 D 作为多项式不定元的多项式是满足结合律的:

$$\begin{aligned} P(D)(Q(D)f(x)) &= P(D)\left(\lim_{n \rightarrow \infty} Q_n(D)f(x)\right) \\ &= \lim_{n \rightarrow \infty} P(D)(Q_n(D)f(x)) \\ &= \lim_{n \rightarrow \infty} (1 - D^n)f(x) \\ &= f(x). \end{aligned}$$

所以现在问题转化成了求出 $\frac{1}{P(D)}$ 的形式幂级数展开式, 这可以利用泰勒展开或者利用多项式除法。个人觉得多项式除法更直观一些, 多项式的除法一般是长除法, 并且原则是从高位向低位以此消去高次项, 而在此处则反过来, 从低位向高位消。

$$\begin{array}{r}
x^3 + x^2 + 3x - 2 \\
x+1 \overline{) x^4 + 2x^3 + 4x^2 + x + 1} \\
\underline{-x^4 - x^3} \\
x^3 + 4x^2 \\
\underline{-x^3 - x^2} \\
3x^2 + x \\
\underline{-3x^2 - 3x} \\
-2x + 1 \\
\underline{2x + 2} \\
3
\end{array}$$

这是正常情况下的长除法（注意与部分数论教材稍有不同），长除法可以对给定的被除式 $f(x)$ ，除式 $g(x)$ ，求出商式 $h(x)$ 与余式 $r(x)$ ，使得

$$f(x) = g(x)h(x) + r(x),$$

其中 $r(x)$ 的次数小于 $g(x)$ 的次数。而对于一般的 $P(D)$ ，我们利用相反的方法，找到一个 $H(D)$ ，使得

$$1 = P(D)H(D) + R(D),$$

而 $R(D)$ 的最低次数是可以无限增大的，于是我们记长除法（反过来的）第 k 次的余式为 $R_k(D)$ ，商式为 $H_k(D)$ ，于是有

$$1 = P(D)H_k(D) + R_k(D).$$

在极限存在的情况下，记

$$\lim_{k \rightarrow \infty} H_k(D) = H(D), \lim_{k \rightarrow \infty} R_k(D) = R(D).$$

那么将 $H(D)f(x)$ 带回原方程：

$$\begin{aligned}
P(D)(H(D)f(x)) &= P(D) \left(\lim_{k \rightarrow \infty} H_k(D)f(x) \right) \\
&= \lim_{k \rightarrow \infty} P(D)(H_k(D)f(x)) \\
&= \lim_{k \rightarrow \infty} (1 - R_k(D))f(x) \\
&= f(x).
\end{aligned}$$

这就说明了算子法的正确性。

所以很自然的，我们有如下定义：

定义 0.1.18: 微分算子多项式的逆

若 $P(D)$ 中的常数项不为零，我们定义：

$$\frac{1}{P(D)} := H(D) \quad (0.85)$$

那么由上文的讨论，当 $f(x)$ 为多项式函数时，其解特解形式已经得到了：

定理 0.1.14: 多项式输入定理

若 $f(x)$ 为有限次多项式函数, 且 $P(D)$ 中的常数项不为零, 则方程 $P(D)y = f(x)$ 的一个特解为:

$$y_p = \frac{1}{P(D)} f(x). \quad (0.86)$$

类似地, 我们可以推广到 $f(x)$ 为指数函数与多项式函数乘积的情形:

定理 0.1.15: 指数-多项式输入定理

设 $a \in \mathbb{C}$, $f(x)$ 为有限次多项式函数, 则方程

$$P(D)y = e^{ax} f(x)$$

的一个特解为:

$$y_p = e^{ax} \frac{1}{P(D+a)} f(x). \quad (0.87)$$

当然到了这一步, 这个公式已经趋近于完美了, 只有一个小问题, 就是 $\frac{1}{P(D+a)}$ 的常数项可能为零。但这不足为虑, 因为位移法则其实在表明这样一个事实: 只要把 $P(D+a)y = f(x)$ 的特解乘上 e^{ax} 就是原方程的特解。所以就回到了一开始讨论的问题, 可以先把 D^k 提取出来, 最后在一步步积分即可。

为了让结论「更正确」, 我们对 $\frac{1}{P(D)}$ 做如下补充定义:

定义 0.1.19: 微分算子多项式的广义逆

设 $P(D) = D^k \tilde{P}(D)$, 其中, $\tilde{P}(D)$ 中的常数项不为零, 则定义

$$\frac{1}{P(D)} := D^{-k} \frac{1}{\tilde{P}(D)}, \quad (0.88)$$

其中, D^{-k} 表示对函数进行 k 次不定积分, 即

$$\frac{1}{D^k} f(x) = \underbrace{\int \int \dots \int}_{k \text{ 次}} f(x) dx. \quad (0.89)$$

而对于任意定义在 $\mathbb{R}[D]$ 上的多项式 $P(D) = D^k \tilde{P}(D)$, 其中 $D \nmid \tilde{P}(D)$, 我们定义:

$$\frac{1}{P(D)} := \frac{1}{D^k} \frac{1}{\tilde{P}(D)}. \quad (0.90)$$

下面给出一个最终的例子, 并附带完整过程

例 0.1.4: 算子法求特解——综合例子

求微分方程

$$y'' + 4y = x \cos x$$

的一个特解。

解: 复化:

$$(D^2 + 4)\tilde{y} = xe^{ix}.$$

故：

$$\begin{aligned}
 \tilde{y} &= \frac{1}{D^2 + 4} x e^{ix} \\
 &= e^{ix} \frac{1}{(D + i)^2 + 4} x \quad (\text{由指数-多项式输入定理}) \\
 &= e^{ix} \frac{1}{D^2 + 2iD + 3} x \\
 &= e^{ix} \left(\frac{1}{3} - \frac{2i}{9} D + \dots \right) x \quad (\text{由多项式输入定理}) \\
 &= (\cos x + i \sin x) \left(\frac{x}{3} - \frac{2i}{9} \right).
 \end{aligned}$$

故：

$$y_p = \Re(\tilde{y}) = \frac{x}{3} \cos x + \frac{2}{9} \sin x.$$

■

如下为升幂长除法的计算示例：

$$\begin{array}{r}
 \frac{1}{3} \quad -\frac{2}{9}iD \\
 \hline
 3 + 2iD + D^2 \quad 1 \\
 \quad -1 - \frac{2}{3}iD - \frac{1}{3}D^2 \\
 \quad \hline
 \quad \quad -\frac{2}{3}iD - \frac{1}{3}D^2 \\
 \quad \quad \quad \frac{2}{3}iD - \frac{4}{9}D^2 + \frac{2}{9}iD^3 \\
 \quad \quad \quad \hline
 \quad \quad \quad \quad -\frac{7}{9}D^2 + \frac{2}{9}iD^3
 \end{array}$$

注释 0.1.12: 关于微分算子法的致谢

上述关于微分算子法的叙述几乎全部来自 Hue-But 在知乎上发表的《常微分方程的微分算子法》^[1]一文，在此表示感谢。

0.2 多元函数微积分

多元函数微积分是对单变量微积分的推广，研究多个自变量的函数的微分与积分性质。但此处只能做简单介绍，一来篇幅有限，二来在普物理学中，多元函数微积分的应用并不多见，更多的是在电动力学与量子力学中才会频繁使用。更详细的介绍可以参考有关数学教材或者笔者后续的电磁学笔记。

0.2.1 偏微商（偏导数）

多元函数是由多个独立自变量构成的函数。将 k 个自变量记为 x_i ($i = 1, 2, \dots, k$)，则多元函数可表示为：

$$y = y(x_1, x_2, \dots, x_k). \quad (0.91)$$

仅由自变量 x_1 的无穷小变化引起的函数增量

$$y(x_1 + \Delta x_1, x_2, \dots, x_k) - y(x_1, x_2, \dots, x_k)$$

称为函数对 x_1 的偏微分, 将

$$y'_{x_1} = \lim_{\Delta x_1 \rightarrow 0} \frac{y(x_1 + \Delta x_1, x_2, \dots, x_k) - y(x_1, x_2, \dots, x_k)}{\Delta x_1} \quad (0.92)$$

称为函数对 x_1 的偏微商或偏导数。为了在形式上与一元函数导数 $\frac{dy}{dx}$ 区分开来, 偏导数通常记为

$$y'_{x_1} = \frac{\partial y}{\partial x_1}. \quad (0.93)$$

类似地, 可定义函数对其他自变量的偏导数:

定义 0.2.1: 偏导数的定义

设 $y = y(x_1, x_2, \dots, x_k)$, 则函数对自变量 x_i 的偏导数定义为

$$\frac{\partial y}{\partial x_i} = \lim_{\Delta x_i \rightarrow 0} \frac{y(x_1, \dots, x_i + \Delta x_i, \dots, x_k) - y(x_1, x_2, \dots, x_k)}{\Delta x_i} \quad (i = 1, 2, \dots, k). \quad (0.94)$$

类似的, 我们可以定义高阶偏导数:

定义 0.2.2: 高阶偏导数的定义

设 $y = y(x_1, x_2, \dots, x_k)$, 则函数对自变量 x_i 的 n 阶偏导数定义为

$$\frac{\partial^n y}{\partial x_i^n} = \frac{\partial}{\partial x_i} \left(\frac{\partial^{n-1} y}{\partial x_i^{n-1}} \right) \quad (i = 1, 2, \dots, k; n = 2, 3, \dots). \quad (0.95)$$

设 $i, j \in \{1, 2, \dots, k\}$ 且 $i \neq j$, 则函数对自变量 x_i 与 x_j 的混合偏导数定义为

$$\frac{\partial^2 y}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial y}{\partial x_j} \right). \quad (0.96)$$

注释 0.2.1: 关于偏导数的说明

偏导数的计算与一元函数的导数计算类似, 只是在计算时将其他自变量视为常数。

而对于混合偏导数, 我们有一个常用的性质: 若 $y = y(x_1, x_2, \dots, x_k)$ 在点 $(x_1^0, x_2^0, \dots, x_k^0)$ 的某邻域内具有连续的二阶偏导数, 则在该点处有

$$\frac{\partial^2 y}{\partial x_i \partial x_j} = \frac{\partial^2 y}{\partial x_j \partial x_i} \quad (i, j = 1, 2, \dots, k; i \neq j).$$

即交换偏导顺序, 结果不变, 该性质称为混合偏导数的对称性。

在此基础上, 我们可以定义多元函数的全微分:

定义 0.2.3: 全微分的定义

设 $y = y(x_1, x_2, \dots, x_k)$, 则函数 y 的全微分定义为

$$dy = \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 + \dots + \frac{\partial y}{\partial x_k} dx_k. \quad (0.97)$$

全微分表示函数在各个自变量发生微小变化时的总变化量。

0.2.2 重积分

重积分是对多元函数在某个区域上的积分运算, 是一元积分的推广。设 $f(x_1, x_2, \dots, x_k)$ 为定义在区域 $D \subseteq \mathbb{R}^k$ 上的多元函数, 则其在区域 D 上的重积分定义为

定义 0.2.4: 重积分的定义

$$\int \dots \int_D f(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k = \lim_{\max \Delta V_i \rightarrow 0} \sum_i f(\xi_i) \Delta V_i, \quad (0.98)$$

其中, 区域 D 被划分为若干小体积 ΔV_i , ξ_i 为 ΔV_i 内的某点, $\max \Delta V_i$ 表示所有小体积的最大值。

注释 0.2.2: 关于重积分的说明

这里的「体积」并非我们通常所说的三维空间中的体积, 而是 k 维空间中的「超体积」。同时上述定义实际上非常模糊且并不严谨, 关于 ΔV_i 的划分方式, 极限的存在性等问题都需要更为严谨的数学语言来描述, 限于篇幅, 此处便不做展开了。

重积分的计算通常可以通过迭代积分来实现, 即将多重积分转化为一系列一元积分。例如, 对于二重积分, 有:

$$\iint_D f(x, y) dx dy = \int_{y_1}^{y_2} \left(\int_{x_1(y)}^{x_2(y)} f(x, y) dx \right) dy,$$

其中, y_1 与 y_2 为区域 D 在 y 方向的边界, $x_1(y)$ 与 $x_2(y)$ 为区域 D 在 x 方向的边界。具体的操作及一些技巧可以参考相关数学教材。

0.2.3 线积分、面积分与体积分

物理学中有些标量是空间位置 \mathbf{r} 的函数, 例如非均匀物质场中的密度分布 $\rho(\mathbf{r})$, 电势分布 $V(\mathbf{r})$ 等。这类标量与位置之间的函数可一般地记作

$$\phi(\mathbf{r}) \text{ 或 } \phi(x, y, z). \quad (0.99)$$

它是一个由空间中的三个独立坐标 x, y, z 决定的三元函数。

类似地, 物理学中有些矢量是空间位置 \mathbf{r} 的函数, 例如速度场 $\mathbf{v}(\mathbf{r})$, 电场强度分布 $\mathbf{E}(\mathbf{r})$ 等。这类矢量与位置之间的函数可一般地记作

$$\mathbf{A}(\mathbf{r}) \text{ 或 } \mathbf{A}(x, y, z). \quad (0.100)$$

他可分解成

$$\mathbf{A}(\mathbf{r}) = A_x(\mathbf{r})\mathbf{e}_x + A_y(\mathbf{r})\mathbf{e}_y + A_z(\mathbf{r})\mathbf{e}_z, \quad (0.101)$$

其中, $A_x(\mathbf{r})$, $A_y(\mathbf{r})$, $A_z(\mathbf{r})$ 均为标量性三元函数。

注释 0.2.3: 关于单位矢量表示法的说明

关于 $Oxyz$ 中单位矢量的表达, 有一部分教材使用的是 $\mathbf{i}, \mathbf{j}, \mathbf{k}$, 但是笔者偏爱使用 $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$, 特在此说明。

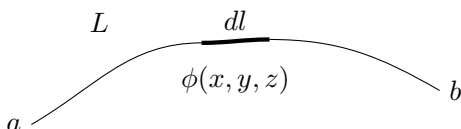


图 0.3

将图 0.3 中 a 到 b 的曲线段 L 上的任意一段曲线元记为 dl , 并忽略曲线元上各点的位置变化, 则曲线元上各点的函数值可视为常数, 记为 $\phi(\mathbf{r})$ 或 $\phi(x, y, z)$ 。将该位置处的函数值与曲线元长度的乘积 $\phi(x, y, z) dl$, 从曲线段 L 的起点 a 积分到终点 b , 即得到沿曲线段 L 的线积分:

定义 0.2.5: 线积分的定义

$$\int_L \phi(x, y, z) dl = \lim_{\max \Delta l_i \rightarrow 0} \sum_i \phi(\xi_i) \Delta l_i, \quad (0.102)$$

其中, 曲线段 L 被划分为若干小曲线元 Δl_i , ξ_i 为 Δl_i 上的某点, $\max \Delta l_i$ 表示所有小曲线元的最大值。

若取 $\phi(x, y, z) = 1$, 则线积分即为曲线段 L 的长度。如果 L 为闭合曲线, 则称该线积分为闭合线积分, 记为

$$\oint_L \phi(x, y, z) dl. \quad (0.103)$$

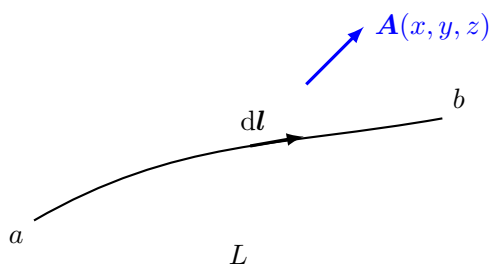


图 0.4

图 0.4 中从线元始端到终端的矢量 dl 称为线元矢量。各 dl 所在位置 (x, y, z) 均有函数值 $\mathbf{A}(x, y, z)$, 将该位置处的矢量函数值与线元矢量的标量积 $\mathbf{A}(x, y, z) \cdot d\mathbf{l}$, 从曲线段 L 的起点 a 积分到终点 b , 即得到沿曲线段 L 的矢量线积分:

定义 0.2.6: 矢量线积分的定义

$$\int_L \mathbf{A}(x, y, z) \cdot d\mathbf{l} = \lim_{\max \Delta l_i \rightarrow 0} \sum_i \mathbf{A}(\xi_i) \cdot \Delta \mathbf{l}_i, \quad (0.104)$$

其中, 曲线段 L 被划分为若干小曲线元矢量 $\Delta \mathbf{l}_i$, ξ_i 为 $\Delta \mathbf{l}_i$ 上的某点, $\max \Delta l_i$ 表示所有小曲线元长度的最大值。

如果 L 为闭合曲线, 则称该矢量线积分为闭合矢量线积分, 记为

$$\oint_L \mathbf{A}(x, y, z) \cdot d\mathbf{l}. \quad (0.105)$$

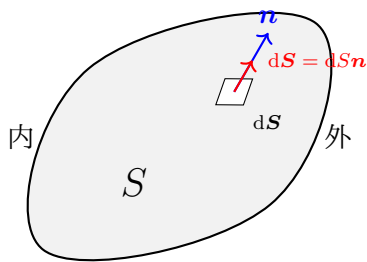


图 0.5

讲一个曲面 (包括平面) S , 如图 0.5 所示, 分割成一系列无穷小的小片, 称为面元, 面积一般地记为 dS 。在面元处引一个与面元法线方向相同的单位矢量 \mathbf{n} , 并规定曲面的一侧为内侧, 另一侧为外侧, \mathbf{n} 指向

外侧, 称

$$d\mathbf{S} = dS\mathbf{n}$$

为面元矢量。若面元处有函数值 $\phi(x, y, z)$, 则将该位置处的函数值与面元面积的乘积 $\phi(x, y, z) dS$, 从曲面 S 的各个面元上积分, 即得到曲面 S 上的面积分:

定义 0.2.7: 面积分的定义

$$\iint_S \phi(x, y, z) dS = \lim_{\max \Delta S_i \rightarrow 0} \sum_i \phi(\xi_i) \Delta S_i, \quad (0.106)$$

其中, 曲面 S 被划分为若干小面元 ΔS_i , ξ_i 为 ΔS_i 上的某点, $\max \Delta S_i$ 表示所有小面元的最大面积。

若取 $\phi(x, y, z) = 1$, 则面积分即为曲面 S 的面积。如果 S 为闭合曲面, 则称该面积分为闭合面积分, 记为

$$\oiint_S \phi(x, y, z) dS. \quad (0.107)$$

类似地, 若面元处有矢量函数值 $\mathbf{A}(x, y, z)$, 则将该位置处的矢量函数值与面元矢量的标量积 $\mathbf{A}(x, y, z) \cdot d\mathbf{S}$, 从曲面 S 的各个面元上积分, 即得到曲面 S 上的矢量面积分:

定义 0.2.8: 矢量面积分的定义

$$\iint_S \mathbf{A}(x, y, z) \cdot d\mathbf{S} = \lim_{\max \Delta S_i \rightarrow 0} \sum_i \mathbf{A}(\xi_i) \cdot \Delta \mathbf{S}_i, \quad (0.108)$$

其中, 曲面 S 被划分为若干小面元矢量 $\Delta \mathbf{S}_i$, ξ_i 为 $\Delta \mathbf{S}_i$ 上的某点, $\max \Delta S_i$ 表示所有小面元的最大面积。

如果 S 为闭合曲面, 则称该矢量面积分为闭合矢量面积分, 记为

$$\oiint_S \mathbf{A}(x, y, z) \cdot d\mathbf{S}. \quad (0.109)$$

将某空间区域 V 分割成一系列无穷小的小体积, 称为体元, 体积一般地记为 dV 。在体元处有函数值 $\phi(x, y, z)$, 则将该位置处的函数值与体元体积的乘积 $\phi(x, y, z) dV$, 从空间区域 V 的各个体元上积分, 即得到空间区域 V 上的体积分:

定义 0.2.9: 体积分的定义

$$\iiint_V \phi(x, y, z) dV = \lim_{\max \Delta V_i \rightarrow 0} \sum_i \phi(\xi_i) \Delta V_i, \quad (0.110)$$

其中, 空间区域 V 被划分为若干小体元 ΔV_i , ξ_i 为 ΔV_i 上的某点, $\max \Delta V_i$ 表示所有小体元的最大体积。

若取 $\phi(x, y, z) = 1$, 则体积分即为空间区域 V 的体积。

关于线积分、面积分与体积分的完整讨论与各种算例, 可以参考相关数学教材或者笔者后续的电磁学笔记。在普物理学的范围内, 就只介绍到这里了。

0.3 行列式

0.3.1 行列式

行列式用被 $||$ 符号包围的一个方阵表示，其中放置的数量或变量等称为行列式的元素，记为 a_{ij} ，其中 i 表示行号， j 表示列号。行列式的行数与列数相等，称为行列式的阶数， n 阶行列式表示为

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \quad (0.111)$$

限于篇幅，此处无法给出行列式的定义，但我们要知道的是，行列式是一个数值，可以通过行列式的元素计算出来。

在普物理学的范围内，最常用的是 2 阶与 3 阶行列式。2 阶行列式的计算较为简单，设

$$D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad (0.112)$$

则

$$D = a_{11}a_{22} - a_{12}a_{21}. \quad (0.113)$$

3 阶行列式的计算可以使用「对角线法则」。设

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad (0.114)$$

则

$$D = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \quad (0.115)$$

用对角线的方法，可以更直观地记忆 3 阶行列式的计算方法。将行列式的前两列复制粘贴到行列式的右侧，构成一个 3×5 的矩阵（在此处就是一个数阵），如下所示：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{bmatrix}.$$

然后从左上角到右下角画三条对角线，分别计算每条对角线上的元素乘积之和，记为 S_1 ；再从右上角到左下角画三条对角线，分别计算每条对角线上的元素乘积之和，记为 S_2 ；则行列式的值为 $D = S_1 - S_2$ 。

对于一般的 n 阶行列式，有如下的归递性质：

定理 0.3.1: 行列式的归递性质

设 D_n 为 n 阶行列式，则

$$D_n = \sum_{j=1}^n a_{ij}A_{ij} = \sum_{i=1}^n a_{ij}A_{ij}, \quad (0.116)$$

其中， $A_{ij} = (-1)^{i+j}M_{ij}$ ， M_{ij} 为由去掉行列式 D_n 的第 i 行与第 j 列所构成的 $(n-1)$ 阶行列式，称为 a_{ij} 的余子式，余子式就是去掉某行某列后剩下的行列式。

此外，笔者认为行列式的如下 4 个性质在普物理学中也较为常用，特此列出：

定理 0.3.2: 行列式的常用性质

设 D 为行列式, 则有如下性质:

1. 若行列式的某两行 (或两列) 互换位置, 则行列式的值变号;
2. 若行列式的某一行 (或某一列) 的各元素同时乘以一个常数 k , 则行列式的值也乘以 k ;
3. 若行列式的某一行 (或某一列) 的各元素同时加上另一行 (或另一列) 的各元素乘以一个常数 k , 则行列式的值不变;
4. 行列式经过转置, 也就是将行与列互换位置后, 行列式的值不变。
5. 如果行列式的某一行 (或某一列) 的各元素均为零, 则该行列式的值为零。
6. 如果行列式的两行 (或两列) 成比例关系, 则该行列式的值为零。

上述内容的证明此处从略。

0.3.2 行列式的应用

行列式在普物力学中的应用主要体现在解线性方程组与计算矢量的混合积等方面, 有关矢量的内容见后续 0.4 节。

线性方程组的一般形式为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases} \quad (0.117)$$

其中, a_{ij} 与 b_i 均为已知常数, x_j 为未知数。当 $m = n$ 时, 称该方程组为 n 元线性方程组。设 n 元线性方程组的系数行列式为

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \quad (0.118)$$

则有如下定理:

定理 0.3.3: 克拉默法则

设 $D \neq 0$, 则 n 元线性方程组有唯一解, 且该解可表示为

$$x_j = \frac{D_j}{D} \quad (j = 1, 2, \dots, n), \quad (0.119)$$

其中, D_j 为将系数行列式 D 的第 j 列替换为常数列 $(b_1, b_2, \dots, b_n)^T$ 后所得到的行列式, 即

$$D_j = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}. \quad (0.120)$$

例如对于三元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3, \end{cases}$$

的解及其展开式为

$$\begin{aligned} x_1 &= \frac{1}{D} \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix} = \frac{b_1(a_{22}a_{33} - a_{23}a_{32}) - b_2(a_{12}a_{33} - a_{13}a_{32}) + b_3(a_{12}a_{23} - a_{13}a_{22})}{a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22})}, \\ x_2 &= \frac{1}{D} \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix} = \frac{-b_1(a_{21}a_{33} - a_{23}a_{31}) + b_2(a_{11}a_{33} - a_{13}a_{31}) - b_3(a_{11}a_{23} - a_{13}a_{21})}{a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22})}, \\ x_3 &= \frac{1}{D} \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix} = \frac{b_1(a_{21}a_{32} - a_{22}a_{31}) - b_2(a_{11}a_{32} - a_{12}a_{31}) + b_3(a_{11}a_{22} - a_{12}a_{21})}{a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22})}. \end{aligned}$$

显然，克拉默法则可以极大的简化线性方程组的求解过程，尤其适用于 2 元与 3 元线性方程组的求解。

行列式在数学物理中的其他应用，后面陆续给出。

0.4 矢量的代数运算

0.4.1 矢量的叠加与分解

直观地说，没有方向的量称为标量，标量带有正负号，有大小，但没有方向，例如温度、质量、时间等。而有方向的量称为矢量，矢量既有大小，又有方向，例如位移、速度、力等，此处我们将其记为粗体字母形式，如 \mathbf{A} , \mathbf{B} 等。

注释 0.4.1: 关于矢量表示法的说明

关于矢量的表示法，印刷时常用粗体字母形式表示矢量，而手写时常用带箭头的字母形式表示矢量，例如 \vec{A} , \vec{B} 等。

矢量 \mathbf{A} 的大小称为 \mathbf{A} 的模，记为 $|\mathbf{A}|$ 或 A ，是一个非负的标量。标量 α 与矢量 \mathbf{A} 的乘积仍是一个矢量，记为

$$\alpha \mathbf{A} = \mathbf{B}. \quad (0.121)$$

\mathbf{B} 的模为

$$|\mathbf{B}| = |\alpha| |\mathbf{A}| = |\alpha| A. \quad (0.122)$$

此处 $|\alpha|$ 表示标量 α 的绝对值。若 $\alpha > 0$ ，则矢量 \mathbf{B} 与矢量 \mathbf{A} 方向相同；若 $\alpha < 0$ ，则矢量 \mathbf{B} 与矢量 \mathbf{A} 方向相反；若 $\alpha = 0$ ，则矢量 \mathbf{B} 为零矢量，记为 $\mathbf{0}$ ，其模为零，且没有确定的方向。

沿着 \mathbf{A} 的方向取一单位矢量，记为 \mathbf{a} 或 \mathbf{e}_A ，则有

$$\mathbf{A} = A \mathbf{e}_A. \quad (0.123)$$

矢量 \mathbf{A} 的单位矢量 \mathbf{e}_A 的模为 1。

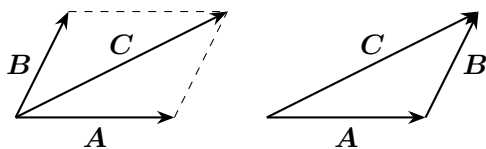


图 0.6

如图 0.6 所示, 设有两个矢量 A 与 B , 它们的和称为矢量 A 与矢量 B 的叠加, 记为

$$C = A + B. \quad (0.124)$$

矢量叠加的几何意义可通过「平行四边形法则」来理解, 即以矢量 A 与矢量 B 为邻边作平行四边形, 则从平行四边形的一个顶点到对角顶点的矢量即为矢量 C ; 也可通过「首尾相接法则」来理解, 即将矢量 B 的起点与矢量 A 的终点相接, 则从矢量 A 的起点到矢量 B 的终点的矢量即为矢量 C 。矢量叠加满足交换律与结合律, 即

$$A + B = B + A, \quad (0.125)$$

$$(A + B) + C = A + (B + C). \quad (0.126)$$

矢量的减法可视为加上一个方向相反的矢量, 即

$$A - B = A + (-B). \quad (0.127)$$

矢量的分解是将一个矢量表示为几个互相垂直的分量矢量之和。

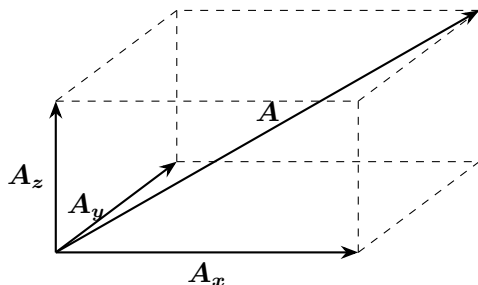


图 0.7

如图 0.7 所示, 对于三维空间中正交的 $Oxyz$ 坐标系, 任何矢量 A 都可以分解为沿 x 轴、 y 轴与 z 轴方向的三个分量矢量之和, 即

$$A = A_x + A_y + A_z, \quad (0.128)$$

其中,

$$A_x = A_x e_x, \quad A_y = A_y e_y, \quad A_z = A_z e_z. \quad (0.129)$$

这里, A_x , A_y , A_z 分别为矢量 A 在 x 轴、 y 轴与 z 轴方向的投影「长度」(这里的「长度是有向长度」), 称为矢量 A 在各坐标轴方向的分量。 e_x , e_y , e_z 分别为 x 轴、 y 轴与 z 轴的单位矢量。则矢量 A 可表示为

$$(A_x, A_y, A_z) = A_x e_x + A_y e_y + A_z e_z. \quad (0.130)$$

矢量 A 的模与各分量之间的关系为

$$A = \sqrt{A_x^2 + A_y^2 + A_z^2}. \quad (0.131)$$

两个三维矢量相加时, 可将它们分别在各坐标轴方向的分量相加, 得到新矢量在各坐标轴方向的分量。例如, 设有两个三维矢量 $\mathbf{A} = (A_x, A_y, A_z)$ 与 $\mathbf{B} = (B_x, B_y, B_z)$, 则它们的和为

$$\mathbf{C} = \mathbf{A} + \mathbf{B} = (A_x + B_x, A_y + B_y, A_z + B_z). \quad (0.132)$$

类似地, 对于 k 维空间中的正交坐标系, 任何矢量 \mathbf{A} 都可以分解为沿各坐标轴方向的 k 个分量矢量之和, 即

$$\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 + \dots + \mathbf{A}_k = A_1 \mathbf{e}_1 + A_2 \mathbf{e}_2 + \dots + A_k \mathbf{e}_k = \sum_{i=1}^k A_i \mathbf{e}_i, \quad (0.133)$$

且矢量 \mathbf{A} 的模与各分量之间的关系为

$$A = \sqrt{A_1^2 + A_2^2 + \dots + A_k^2} = \sqrt{\sum_{i=1}^k A_i^2}. \quad (0.134)$$

两个维数相同的矢量相加时, 可将它们分别在各坐标轴方向的分量相加, 得到新矢量在各坐标轴方向的分量。例如, 设有两个 k 维矢量 $\mathbf{A} = (A_1, A_2, \dots, A_k)$ 与 $\mathbf{B} = (B_1, B_2, \dots, B_k)$, 则它们的和为

$$\mathbf{C} = \mathbf{A} + \mathbf{B} = (A_1 + B_1, A_2 + B_2, \dots, A_k + B_k) = \sum_{i=1}^k (A_i + B_i) \mathbf{e}_i. \quad (0.135)$$

0.4.2 矢量的点积

注释 0.4.2: 关于矢量点积的名称说明

矢量的点积在物理学中也称为矢量的标积或数量积。所谓标积、标量积、点积、数量积、内积等, 实际上都是同一个概念, 只是名称不同而已。

矢量的点积定义如下:

定义 0.4.1: 矢量的点积

设有两个矢量 \mathbf{A} 与 \mathbf{B} , 它们的点积记为 $\mathbf{A} \cdot \mathbf{B}$, 定义为

$$\mathbf{A} \cdot \mathbf{B} = AB \cos \theta, \quad (0.136)$$

其中, A 与 B 分别为矢量 \mathbf{A} 与矢量 \mathbf{B} 的模, θ 为矢量 \mathbf{A} 与矢量 \mathbf{B} 之间的夹角。

矢量的点积是一个标量。

通过点积, 可将 \mathbf{A} 的模和分量各自表示成

$$A = \sqrt{\mathbf{A} \cdot \mathbf{A}}, \quad (0.137)$$

$$A_x = \mathbf{A} \cdot \mathbf{e}_x, \quad A_y = \mathbf{A} \cdot \mathbf{e}_y, \quad A_z = \mathbf{A} \cdot \mathbf{e}_z. \quad (0.138)$$

由点积的定义, 不难得到如下性质:

定理 0.4.1: 矢量点积的性质

设有三个矢量 \mathbf{A} , \mathbf{B} 与 \mathbf{C} , 以及一个标量 k , 则有如下性质:

1. 交换律: $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$;
2. 分配律: $\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$;

3. 结合律: $(k\mathbf{A}) \cdot \mathbf{B} = k(\mathbf{A} \cdot \mathbf{B})$ 。

4. 正定性: $\mathbf{A} \cdot \mathbf{A} \geq 0$, 且当且仅当 $\mathbf{A} = \mathbf{0}$ 时取等号。

在三维空间中有

$$\mathbf{e}_x \cdot \mathbf{e}_y = \mathbf{e}_y \cdot \mathbf{e}_z = \mathbf{e}_z \cdot \mathbf{e}_x = 0, \quad (0.139)$$

$$\mathbf{e}_x \cdot \mathbf{e}_x = \mathbf{e}_y \cdot \mathbf{e}_y = \mathbf{e}_z \cdot \mathbf{e}_z = 1. \quad (0.140)$$

借此可以将两个三维矢量 $\mathbf{A} = (A_x, A_y, A_z)$ 与 $\mathbf{B} = (B_x, B_y, B_z)$ 的点积表示为

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z. \quad (0.141)$$

类似地, 在 k 维空间中有

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad (0.142)$$

其中, δ_{ij} 称为克罗内克 (Kronecker) 符号。借此可以将两个 k 维矢量 $\mathbf{A} = (A_1, A_2, \dots, A_k)$ 与 $\mathbf{B} = (B_1, B_2, \dots, B_k)$ 的点积表示为

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i=1}^k A_i B_i. \quad (0.143)$$

0.4.3 矢量的叉积

注释 0.4.3: 关于矢量叉积的名称说明

矢量的叉积在物理学中也称为矢量的外积或向量积。所谓叉积、外积、向量积、矢积、矢量积等, 实际上都是同一个概念, 只是名称不同而已。

矢量的叉积定义如下:

定义 0.4.2: 矢量的叉积

设有两个三维矢量 \mathbf{A} 与 \mathbf{B} , 它们的叉积记为 $\mathbf{A} \times \mathbf{B}$, 定义为一个新矢量, 其模为

$$|\mathbf{A} \times \mathbf{B}| = AB \sin \theta, \quad (0.144)$$

其中, A 与 B 分别为矢量 \mathbf{A} 与矢量 \mathbf{B} 的模, θ 为矢量 \mathbf{A} 与矢量 \mathbf{B} 之间的夹角; 该新矢量的方向由右手定则确定, 即将右手的四指从矢量 \mathbf{A} 的方向弯向矢量 \mathbf{B} 的方向, 则大拇指所指的方向即为新矢量的方向。

矢量的叉积是一个矢量。

注释 0.4.4: 关于右手定则的说明

关于为什么选择右手, 而不是左手来确定矢量叉积的方向。这实际上是没有问题的, 因为这两个体系的结构是相同的, 但是如今大多数教材与文献都采用右手定则来确定矢量叉积的方向, 所以我们也沿用这一传统。

叉积也有一些性质:

定理 0.4.2: 矢量叉积的性质

设有三个矢量 \mathbf{A} , \mathbf{B} 与 \mathbf{C} , 以及一个标量 k , 则有如下性质:

1. 反交换律: $\mathbf{A} \times \mathbf{B} = -(\mathbf{B} \times \mathbf{A})$;
2. 分配律: $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$;
3. 结合律: $(k\mathbf{A}) \times \mathbf{B} = k(\mathbf{A} \times \mathbf{B})$;
4. 向量共线性: 若矢量 \mathbf{A} 与矢量 \mathbf{B} 共线, 则 $\mathbf{A} \times \mathbf{B} = \mathbf{0}$.

在三维空间中有

$$\mathbf{e}_x \times \mathbf{e}_y = \mathbf{e}_z, \quad \mathbf{e}_y \times \mathbf{e}_z = \mathbf{e}_x, \quad \mathbf{e}_z \times \mathbf{e}_x = \mathbf{e}_y, \quad (0.145)$$

$$\mathbf{e}_y \times \mathbf{e}_x = -\mathbf{e}_z, \quad \mathbf{e}_z \times \mathbf{e}_y = -\mathbf{e}_x, \quad \mathbf{e}_x \times \mathbf{e}_z = -\mathbf{e}_y, \quad (0.146)$$

$$\mathbf{e}_x \times \mathbf{e}_x = \mathbf{e}_y \times \mathbf{e}_y = \mathbf{e}_z \times \mathbf{e}_z = \mathbf{0}. \quad (0.147)$$

借此可以将两个三维矢量 $\mathbf{A} = (A_x, A_y, A_z)$ 与 $\mathbf{B} = (B_x, B_y, B_z)$ 的叉积表示为

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} = (A_y B_z - A_z B_y)\mathbf{e}_x + (A_z B_x - A_x B_z)\mathbf{e}_y + (A_x B_y - A_y B_x)\mathbf{e}_z. \quad (0.148)$$

这里的行列式见前面 0.3 节。

注释 0.4.5: 关于叉积定义维数的说明

关于叉积的定义, 肯定会有读者提出疑问: 为什么叉积只能定义在三维空间中, 而不能定义在二维空间或更高维空间中呢? 实际上, 叉积还可以定义在七维空间中, 至于为什么是三维与七维空间, 为什么要这样定义, 这里便不做展开了。但是我可以说的是这与四元数及李群和李代数有关, 感兴趣的读者可以自行查阅相关资料。

注释 0.4.6: 关于极矢量和轴矢量的说明

矢量根据其在空间中的变换性质, 可分为极矢量与轴矢量两种。极矢量 (polar vector) 也称为真矢量, 是我们通常所说的矢量, 例如位移、速度、力等, 极矢量在空间反射变换下, 与镜面垂直的分量会反向, 而与镜面平行的分量不变; 轴矢量 (axial vector) 也称为伪矢量, 是由两个极矢量通过叉积运算得到的矢量, 例如角动量、磁场强度等, 轴矢量在空间反射变换下, 与镜面垂直的分量不变, 而与镜面平行的分量会反向。

矢量的加法与点积运算不改变矢量的性质, 而矢量的叉积运算会改变矢量的性质, 即两个极矢量的叉积是一个轴矢量, 两个轴矢量的叉积是一个极矢量, 一个极矢量与一个轴矢量的叉积是一个极矢量。

0.4.4 矢量的多重积

矢量的多重积包括混合积与三重叉积等。

混合积定义如下:

定义 0.4.3: 矢量的混合积

设有三个三维矢量 \mathbf{A} , \mathbf{B} 与 \mathbf{C} , 它们的混合积记为 $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$, 定义为

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = ABC \sin \theta \cos \phi, \quad (0.149)$$

其中, A , B 与 C 分别为矢量 \mathbf{A} , \mathbf{B} 与 \mathbf{C} 的模, θ 为矢量 \mathbf{B} 与矢量 \mathbf{C} 之间的夹角, ϕ 为矢量 \mathbf{A} 与矢量 $\mathbf{B} \times \mathbf{C}$ 之间的夹角。

混合积是一个标量, 并且满足循环对称性与反对称性, 即

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}), \quad (0.150)$$

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = -\mathbf{A} \cdot (\mathbf{C} \times \mathbf{B}). \quad (0.151)$$

运用行列式可以将三个三维矢量 $\mathbf{A} = (A_x, A_y, A_z)$, $\mathbf{B} = (B_x, B_y, B_z)$ 与 $\mathbf{C} = (C_x, C_y, C_z)$ 的混合积表示为

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}. \quad (0.152)$$

三重叉积定义如下:

定义 0.4.4: 矢量的三重叉积

设有三个三维矢量 \mathbf{A} , \mathbf{B} 与 \mathbf{C} , 它们的三重叉积记为 $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$, 定义为一个新矢量, 其方向与矢量 \mathbf{B} 与矢量 \mathbf{C} 所在平面内的某个方向相同, 其模为

$$|\mathbf{A} \times (\mathbf{B} \times \mathbf{C})| = ABC \sin \theta \sin \phi, \quad (0.153)$$

其中, A , B 与 C 分别为矢量 \mathbf{A} , \mathbf{B} 与 \mathbf{C} 的模, θ 为矢量 \mathbf{B} 与矢量 \mathbf{C} 之间的夹角, ϕ 为矢量 \mathbf{A} 与矢量 $\mathbf{B} \times \mathbf{C}$ 之间的夹角。

三重叉积是一个矢量, 并且满足如下恒等式:

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}). \quad (0.154)$$

在 $Oxyz$ 坐标系中可以轻易的证明这个恒等式, 即

$$\begin{aligned} \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= (A_y B_z C_x - A_z B_y C_x) \mathbf{e}_x + (A_z B_x C_y - A_x B_z C_y) \mathbf{e}_y + (A_x B_y C_z - A_y B_x C_z) \mathbf{e}_z \\ &= [B_x (A_x C_x + A_y C_y + A_z C_z) - C_x (A_x B_x + A_y B_y + A_z B_z)] \mathbf{e}_x \\ &\quad + [B_y (A_x C_x + A_y C_y + A_z C_z) - C_y (A_x B_x + A_y B_y + A_z B_z)] \mathbf{e}_y \\ &\quad + [B_z (A_x C_x + A_y C_y + A_z C_z) - C_z (A_x B_x + A_y B_y + A_z B_z)] \mathbf{e}_z \\ &= \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}). \end{aligned} \quad (0.155)$$

下面给出几个例子, 帮助理解矢量的代数运算。

例 0.4.1: 点积的逆运算

设有两个矢量 \mathbf{A} 与 \mathbf{B} , 已知 \mathbf{B} , 以及它们的点积 $\alpha = \mathbf{A} \cdot \mathbf{B}$, 求矢量 \mathbf{A} 。解:

将 \mathbf{A} 分解为沿 \mathbf{B} 方向的分量 \mathbf{A}_{\parallel} 与垂直于 \mathbf{B} 方向的分量 \mathbf{A}_{\perp} , 即

$$\mathbf{A} = \mathbf{A}_{\parallel} + \mathbf{A}_{\perp}. \quad (0.156)$$

则有

$$\alpha = \mathbf{A} \cdot \mathbf{B} = \mathbf{A}_{\parallel} \cdot \mathbf{B} + \mathbf{A}_{\perp} \cdot \mathbf{B} = \mathbf{A}_{\parallel} \cdot \mathbf{B}, \quad (0.157)$$

因为 \mathbf{A}_\perp 与 \mathbf{B} 垂直, 所以 $\mathbf{A}_\perp \cdot \mathbf{B} = 0$ 。设 $|\mathbf{B}| = B$, 则有

$$\mathbf{A}_\parallel = \frac{\alpha}{B^2} \mathbf{B}. \quad (0.158)$$

因为 \mathbf{A}_\perp 与 \mathbf{B} 垂直, 但具有不确定性, 可表达成

$$\mathbf{A}_\perp = \mathbf{C} \times \mathbf{B}, \quad (0.159)$$

其中, \mathbf{C} 为任意矢量。因此, 矢量 \mathbf{A} 可表示为

$$\mathbf{A} = \frac{\alpha}{B^2} \mathbf{B} + \mathbf{C} \times \mathbf{B}. \quad (0.160)$$

例 0.4.2: 叉积的逆运算

设有两个矢量 \mathbf{A} 与 \mathbf{B} , 已知 \mathbf{B} , 以及它们的叉积 $\mathbf{D} = \mathbf{A} \times \mathbf{B}$, 求矢量 \mathbf{A} 。解:

将 \mathbf{A} 分解为沿 \mathbf{B} 方向的分量 \mathbf{A}_\parallel 与垂直于 \mathbf{B} 方向的分量 \mathbf{A}_\perp , 即

$$\mathbf{A} = \mathbf{A}_\parallel + \mathbf{A}_\perp. \quad (0.161)$$

则有

$$\mathbf{D} = \mathbf{A} \times \mathbf{B} = \mathbf{A}_\parallel \times \mathbf{B} + \mathbf{A}_\perp \times \mathbf{B} = \mathbf{A}_\perp \times \mathbf{B}, \quad (0.162)$$

因为 \mathbf{A}_\parallel 与 \mathbf{B} 共线, 所以 $\mathbf{A}_\parallel \times \mathbf{B} = \mathbf{0}$ 。设 $|\mathbf{B}| = B$, 则有

$$\mathbf{A}_\perp = \frac{1}{B^2} (\mathbf{B} \times \mathbf{D}). \quad (0.163)$$

因为 \mathbf{A}_\parallel 与 \mathbf{B} 共线, 但具有不确定性, 可表达成

$$\mathbf{A}_\parallel = \lambda \mathbf{B}, \quad (0.164)$$

其中, λ 为任意标量。因此, 矢量 \mathbf{A} 可表示为

$$\mathbf{A} = \lambda \mathbf{B} + \frac{1}{B^2} (\mathbf{B} \times \mathbf{D}). \quad (0.165)$$

这说明了在一定程度上, 点积与叉积是不可逆的运算。

0.5 其他可能的数学工具

0.5.1 数列 (序列)

有关于数列 (序列) 的技巧, 主要还是依靠高中的数学知识, 比如凑等比、等差, 裂项, 线性递推, 「不动点」等。此处便不详述了。

0.5.2 不等式

在物理中, 常用的不等式主要有基本不等式等。

定理 0.5.1: 基本不等式链

设有任意实数 a 与 b , 则有如下基本不等式链:

$$\frac{2}{\frac{1}{a} + \frac{1}{b}} \leq \sqrt{ab} \leq \frac{a+b}{2} \leq \sqrt{\frac{a^2+b^2}{2}}, \quad (0.166)$$

其中, 当且仅当 $a = b$ 时, 各等号成立。

类似地, 对于 n 个正实数 a_1, a_2, \dots, a_n , 则有

$$\frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_n}} \leq \sqrt[n]{a_1 a_2 \dots a_n} \leq \frac{a_1 + a_2 + \dots + a_n}{n} \leq \sqrt{\frac{a_1^2 + a_2^2 + \dots + a_n^2}{n}}, \quad (0.167)$$

其中, 当且仅当 $a_1 = a_2 = \dots = a_n$ 时, 各等号成立。

0.5.3 复数

复数在物理中有着广泛的应用, 尤其是在处理振动与波动现象时。设有一个复数 $z = x + iy$, 其中, x 与 y 分别为复数 z 的实部与虚部, 它们均为实数, i 为虚数单位, 满足 $i^2 = -1$ 。取实部与虚部分别作为横轴与纵轴, 则复数 z 可表示为平面上的一个点, 此平面称为复平面。复数 z 的共轭复数记为 $z^* = x - iy$ 。且有

$$zz^* = (x + iy)(x - iy) = x^2 + y^2. \quad (0.168)$$

取实部和取虚部的运算可表示为

$$\Re(z) = \frac{z + z^*}{2} = x, \quad (0.169)$$

$$\Im(z) = \frac{z - z^*}{2i} = y. \quad (0.170)$$

复数 z 也可以表示成极坐标形式, 即

$$z = r(\cos \theta + i \sin \theta), \quad (0.171)$$

其中, $r = \sqrt{x^2 + y^2}$ 为复数 z 的模, $\theta = \arctan \frac{y}{x}$ 为复数 z 的辐角。复数的乘法与除法在极坐标形式下有着简单的表示:

$$z_1 z_2 = r_1 r_2 [\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)], \quad (0.172)$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} [\cos(\theta_1 - \theta_2) + i \sin(\theta_1 - \theta_2)], \quad (0.173)$$

其中, 复数 $z_1 = r_1(\cos \theta_1 + i \sin \theta_1)$ 与复数 $z_2 = r_2(\cos \theta_2 + i \sin \theta_2)$ 。

欧拉公式是复数的一个重要公式, 内容为

$$e^{i\theta} = \cos \theta + i \sin \theta. \quad (0.174)$$

利用欧拉公式, 可以将复数 $z = r(\cos \theta + i \sin \theta)$ 表示成指数形式, 即

$$z = r e^{i\theta}. \quad (0.175)$$

复数的乘法与除法在指数形式下也有着简单的表示:

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}, \quad (0.176)$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)}, \quad (0.177)$$

其中, 复数 $z_1 = r_1 e^{i\theta_1}$ 与复数 $z_2 = r_2 e^{i\theta_2}$ 。

0.5.4 三角函数

三角函数在物理中有着广泛的应用，尤其是在处理周期性现象时。首先给出三倍角公式：

$$\sin 3\theta = 3 \sin \theta - 4 \sin^3 \theta, \quad (0.178)$$

$$\cos 3\theta = 4 \cos^3 \theta - 3 \cos \theta. \quad (0.179)$$

它们可以用复数法轻易的证明

$$\begin{aligned} \sin 3\theta + i \cos 3\theta &= (\cos \theta + i \sin \theta)^3 \\ &= \cos^3 \theta + 3i \cos^2 \theta \sin \theta - 3 \cos \theta \sin^2 \theta - i \sin^3 \theta \\ &= (4 \cos^3 \theta - 3 \cos \theta) + i(3 \sin \theta - 4 \sin^3 \theta). \end{aligned} \quad (0.180)$$

另外还有和差化积与积化和差公式：

$$\begin{aligned} \sin \alpha + \sin \beta &= 2 \sin \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}, \\ \sin \alpha - \sin \beta &= 2 \cos \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}, \\ \cos \alpha + \cos \beta &= 2 \cos \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}, \\ \cos \alpha - \cos \beta &= -2 \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}, \end{aligned} \quad (0.181)$$

$$\begin{aligned} \sin \alpha \cos \beta &= \frac{1}{2} [\sin(\alpha + \beta) + \sin(\alpha - \beta)], \\ \cos \alpha \sin \beta &= \frac{1}{2} [\sin(\alpha + \beta) - \sin(\alpha - \beta)], \\ \cos \alpha \cos \beta &= \frac{1}{2} [\cos(\alpha + \beta) + \cos(\alpha - \beta)], \\ \sin \alpha \sin \beta &= -\frac{1}{2} [\cos(\alpha + \beta) - \cos(\alpha - \beta)]. \end{aligned} \quad (0.182)$$

0.5.5 求和公式

在物理中，常用的一些求和公式如下：

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}, \quad (0.183)$$

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}, \quad (0.184)$$

$$\sum_{k=1}^n k^3 = \left[\frac{n(n+1)}{2} \right]^2. \quad (0.185)$$

这三个公式可以用数学归纳法证明。

另外，还有一些三角函数的求和公式：

$$\sum_{k=0}^n \sin(k\theta) = \frac{\sin \frac{(n+1)\theta}{2} \sin \frac{n\theta}{2}}{\sin \frac{\theta}{2}}, \quad (0.186)$$

$$\sum_{k=0}^n \cos(k\theta) = \frac{\sin \frac{(n+1)\theta}{2} \cos \frac{n\theta}{2}}{\sin \frac{\theta}{2}}, \quad (0.187)$$

三角函数的求和公式可以用复数法证明：

$$\sum_{k=0}^n \sin(k\theta) = \Im \left(\sum_{k=0}^n e^{ik\theta} \right) = \Im \left(\frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}} \right) = \Im \left(\frac{e^{i\frac{(n+1)\theta}{2}} (e^{-i\frac{(n+1)\theta}{2}} - e^{i\frac{(n+1)\theta}{2}})}{e^{i\frac{\theta}{2}} (e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}})} \right) = \frac{\sin \frac{(n+1)\theta}{2} \sin \frac{n\theta}{2}}{\sin \frac{\theta}{2}}, \quad (0.188)$$

$$\sum_{k=0}^n \cos(k\theta) = \Re \left(\sum_{k=0}^n e^{ik\theta} \right) = \Re \left(\frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}} \right) = \Re \left(\frac{e^{i\frac{(n+1)\theta}{2}} (e^{-i\frac{(n+1)\theta}{2}} - e^{i\frac{(n+1)\theta}{2}})}{e^{i\frac{\theta}{2}} (e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}})} \right) = \frac{\sin \frac{(n+1)\theta}{2} \cos \frac{n\theta}{2}}{\sin \frac{\theta}{2}}. \quad (0.189)$$

0.5.6 双曲函数

双曲函数在物理中也有着一定的应用，比如狭义相对论中的洛伦兹变换就可以用双曲函数来表示。双曲函数的定义如下：

定义 0.5.1: 双曲函数

双曲正弦函数与双曲余弦函数分别定义为

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad (0.190)$$

$$\cosh x = \frac{e^x + e^{-x}}{2}. \quad (0.191)$$

此外，还有双曲正切函数，定义为

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (0.192)$$

双曲函数满足如下恒等式：

$$\cosh^2 x - \sinh^2 x = 1, \quad (0.193)$$

它的导数公式如下：

示例 0.5.1: 双曲函数的导数

设 x 为自变量，则有

$$\frac{d}{dx} \sinh x = \cosh x, \quad (0.194)$$

$$\frac{d}{dx} \cosh x = \sinh x, \quad (0.195)$$

$$\frac{d}{dx} \tanh x = \frac{1}{\cosh^2 x}. \quad (0.196)$$

积分公式如下：

示例 0.5.2: 双曲函数的积分

设 x 为自变量，则有

$$\int \sinh x \, dx = \cosh x + C, \quad (0.197)$$

$$\int \cosh x \, dx = \sinh x + C, \quad (0.198)$$

$$\int \tanh x \, dx = \ln |\cosh x| + C. \quad (0.199)$$

其性质与三角函数类似，比如和差公式、倍角公式等，可以自行尝试推导。

注释 0.5.1: 关于双曲函数名称的说明

双曲函数的名称来源于它们与双曲线的关系。设有双曲线 $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, 则可以将其参数化为

$$x = a \cosh t, \quad y = b \sinh t, \quad (0.200)$$

其中, t 为参数。由此可见, 双曲函数与双曲线有着密切的关系。

再考虑由欧拉公式得到三角函数的表达式

$$\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}, \quad \cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}. \quad (0.201)$$

将其中的虚数单位 i 替换为实数单位 1, 即可得到双曲函数的表达式。

双曲三角函数的名称也是由此而来。

第一章 质点运动学

质点动力学是大多数普通物理学教材的第一章内容。同时这一章也是大多数人（包括笔者）将诸如微积分和矢量代数等较为困难的数学工具应用于物理学这种二级学科的第一次尝试，在学习这一章的过程中，你可能会感觉到一种「阵痛」。但请你相信，在翻过这座山峰之后，你会发现前方的道路变得平坦而宽广，你会更加自信地面对后续的物理学内容。

1.1 时间与空间

时间与空间是最基本的物理概念，没有严格的理论可以定义它们。在经典力学中，我们通常假设时间和空间是绝对的、均匀的，并且相互独立的。也就是说，时间的流逝不受空间位置的影响，空间的性质不受时间的影响。

注释 1.1.1: 时间与空间的绝对性

需要注意的是，时间与空间的绝对性假设在相对论中被推翻。在相对论框架下，时间和空间是相互关联的，并且它们的性质会受到观察者运动状态的影响。这些内容将会在最后的相对论章节，即第二章中进行介绍。

1.1.1 时间的计量

时间表征物质运动变化的顺序和持续时间。时间的计量主要是计数的过程，凡是已知其运动规律的物理过程都可以用来计量时间，例如：地球自转、公转，摆钟的摆动，石英晶体的振动，铯原子的跃迁等。国际单位制中时间的基本单位是秒（second，符号为 s），其最新定义是在 1967 年由第 13 届国际计量大会通过的，即：

定义 1.1.1: 秒的定义

秒是铯-133 原子基态的两个超精细能级之间跃迁所对应的辐射的 9,192,631,770 个周期的持续时间。

这一新的定义极大地提高了时间计量的精度，使得现代物理实验和技术应用能够达到前所未有的精确度。

注释 1.1.2: 一些非 SI 时间单位

下面列出了一些常用的时间单位（非 SI）及其换算关系：

- 分钟（minute，符号为 min）， $1 \text{ min} = 60 \text{ s}$ ；
- 小时（hour，符号为 h）， $1 \text{ h} = 60 \text{ min} = 3600 \text{ s}$ ；
- 天（day，符号为 d）， $1 \text{ d} = 24 \text{ h} = 86400 \text{ s}$ ；
- 年（year，符号为 a，也有使用 yr 或 y 的）， $1 \text{ a} = 365.25 \text{ d} = 31557600 \text{ s}$ 。

1.1.2 空间的度量

空间反应物质运动的广延性。在三维空间中的位置通常用三个相互独立的坐标来表示，例如笛卡尔坐标系中的 (x, y, z) 。空间中两点之间的距离为长度。国际单位制中长度的基本单位是米（meter，符号为 m），其最新定义是在 1983 年由第 17 届国际计量大会通过的，即：

定义 1.1.2: 米的定义

米是光在真空中在 $1/299,792,458$ 秒内传播的距离。

这一新的定义将长度单位与时间单位联系起来，利用了光速在真空中的恒定性，从而提高了长度测量的精度。

注释 1.1.3: 一些非 SI 长度单位

下面列出了一些常用的长度单位（非 SI）及其换算关系：

- 埃（angstrom，符号为 Å）， $1 \text{ Å} = 1 \times 10^{-10} \text{ m}$ ；
- 光年（light year，符号为 ly）， $1 \text{ ly} = 9.4607 \times 10^{15} \text{ m}$ ；
- 天文单位（astronomical unit，符号为 au）， $1 \text{ au} = 1.495\,978\,707 \times 10^{11} \text{ m}$ ；
- 秒差距（parsec，符号为 pc，被定义为周年视差等于 1 角秒时的距离）， $1 \text{ pc} = 3.085\,677\,581 \times 10^{16} \text{ m}$ 。

1.2 物质世界的层次与数量级

物质世界的层次结构极其丰富，从宇宙的宏观尺度到微观的基本粒子尺度，跨度达数十个数量级。这里主要是为了介绍一下国际单位制所用的词头，以便更好地理解物理量的数量级。表 1.1 列出了国际单位制中常用的词头及其对应的因数。

表 1.1: 国际单位制词头表

因数	词头名称		符号	因数	词头名称		符号
	英文	中文			英文	中文	
10^{-1}	deci	分	d	10^1	deca	十	da
10^{-2}	centi	厘	c	10^2	hecto	百	h
10^{-3}	milli	毫	m	10^3	kilo	千	k
10^{-6}	micro	微	μ	10^6	mega	兆	M
10^{-9}	nano	纳	n	10^9	giga	吉	G
10^{-12}	pico	皮	p	10^{12}	tera	太	T
10^{-15}	femto	飞	f	10^{15}	peta	拍	P
10^{-18}	atto	阿	a	10^{18}	exa	艾	E
10^{-21}	zepto	仄	z	10^{21}	zetta	泽	Z
10^{-24}	yocto	尤	y	10^{24}	yotta	尧	Y

它们与基本单位结合后，可以表示非常大或非常小的物理量。例如， $1 \text{ km} = 1 \times 10^3 \text{ m}$ ， $1 \text{ ms} = 1 \times 10^{-3} \text{ s}$ 。通过使用这些词头，我们可以更方便地表达和理解物理量的数量级，从而更好地描述物质世界的层次结构。

下面这一段摘自赵凯华老师的《新概念物理教程-力学》^[2]：

引用 1.2.1

物理学是一门定量程度很高的学科，它推理性强、逻辑严密，实验测量和理论计算都达到了很高的精度，如时间的计量就有 12 至 13 位有效数字。然而，理论物理学家在进行详细计算之前，为了选择和建立恰当的物理模型和数学模型，需要首先粗略的估计各参量的大小和各种可能效应的相对重要性，以判断什么是决定现象的主要机制；同样，实验物理学家在着手准备精密的测量之前，为了选择合适的仪器和测量方法，也需要对各个有关物理量的数量级先做一番估计。总之，掌握特征量的数量级，往往是研究一个物理问题时登堂入室的关键。学习物理学，就需要经常训练对各种事物作粗略的数量级估计，留心查看尺度大小的变化所产生的物理效应。

1.3 质点与运动

1.3.1 质点与参考系

在物理学中，为了突出研究对象的某些主要性质，暂不考虑一些次要因素，经常引入一些理想化的模型。质点模型就是这样一种理想化的模型，其定义如下：

定义 1.3.1: 质点

质点是指在研究物体的运动时，将物体视为一个没有大小和形状的点质量。质点突出了「物体具有质量」和「物体占有位置」这两个基本属性，而忽略了物体的空间扩展性和内部结构。

质点的位置及其运动与否，只有相对于实现选定的视为不动的物体（或彼此不做相对运动的物体群）才有明确的意义。这种视为不动的物体（或物体群） K 称为参考物，与参考物固连的空间称为参考空间。为了描述运动，还必须有计时装置——钟。参考空间和与之固连的钟的组合称为参考系。习惯上，我们把参考系简记为 K ，即把参考物简称为参考系。

在参考系选定之后，为了定量地描述质点相对于参考系的位置，还必须在参考系上建立适当的坐标系。坐标系的选取是完全任意的，可以根据具体问题的需要来选择。

第二章 狭义相对论

严格意义上，相对论并不属于普通物理学力学（Mechanics）的范畴，但多数高校的物理学力学课程都会包含狭义相对论的内容，并定性地介绍一下广义相对论，这份笔记自然也不例外。

狭义相对论这一章可以说是大多数人（包括笔者）重塑物理直觉的开始，也是真正接触现代物理学的开始。但不用担心，在严格的逻辑推导和时空图的指引下，你一定能理解狭义相对论的核心思想，并且能够熟练地运用它来分析物理问题。

2.1 狭义相对论的基本原理

2.1.1 经典理论的危机

19 世纪末，物理大厦上空有两朵乌云：

- 黑体辐射的「紫外灾难」
- 迈克耳孙-莫雷实验的「以太风」结果

这两朵乌云预示着经典物理学的危机，最终导致了量子力学和相对论的诞生。接下来我们将讨论第二朵乌云，也就是迈克耳孙-莫雷实验的结果。

库仑（Coulomb）在 1785 年提出了电荷之间的相互作用力遵循平方反比定律，半个世纪后，麦克斯韦（Maxwell）在 1865 年成功地将电学和磁学统一成了电磁学，建立了麦克斯韦方程组，预言了电磁波的存在。他发现空间电磁场可以以波的形式传播，当时假想的传播介质被称为「以太」，麦克斯韦方程组预言了电磁波在真空中的传播速度是一个常数，约为 $3 \times 10^8 \text{ m} \cdot \text{s}^{-1}$ ，这个速度被认为是以太中的光速。

$$\left(\nabla^2 - \mu\epsilon \frac{\partial^2}{\partial t^2} \right) \mathbf{E} = 0 \quad (2.1)$$

$$\left(\nabla^2 - \mu\epsilon \frac{\partial^2}{\partial t^2} \right) \mathbf{B} = 0 \quad (2.2)$$

其中， μ 和 ϵ 分别是真空的磁导率和电容率， \mathbf{E} 和 \mathbf{B} 分别是电场和磁场。根据麦克斯韦方程组，电磁波的传播速度 c 满足

$$c = \frac{1}{\sqrt{\mu\epsilon}} \approx 3 \times 10^8 \text{ m} \cdot \text{s}^{-1} \quad (2.3)$$

与参考系无关 (?) 这个结果与经典力学的预期相矛盾（在经典力学中，速度是相对的）。

于是，人们假想了一个「静止」的以太参考系（否则，地心说复活，太猎奇了...），麦克斯韦方程组在这个参考系中成立。而在其他参考系中，电磁波的传播速度应该是 $c \pm v$ ，其中 v 是参考系相对于以太的速度。

为了验证这个假说，迈克耳孙（Michelson）和莫雷（Morley）设计了一个精密的干涉仪，试图测量地球相对于以太的运动速度。在这里，我简单介绍一下迈克耳孙-莫雷干涉仪的工作原理。

如图 2.1 所示，干涉仪由一个分束器和两个垂直放置的反射镜组成。光源发出的光束经过分束器被分成两束，分别沿两个垂直方向传播，经过反射镜反射后重新汇合，形成干涉图样。

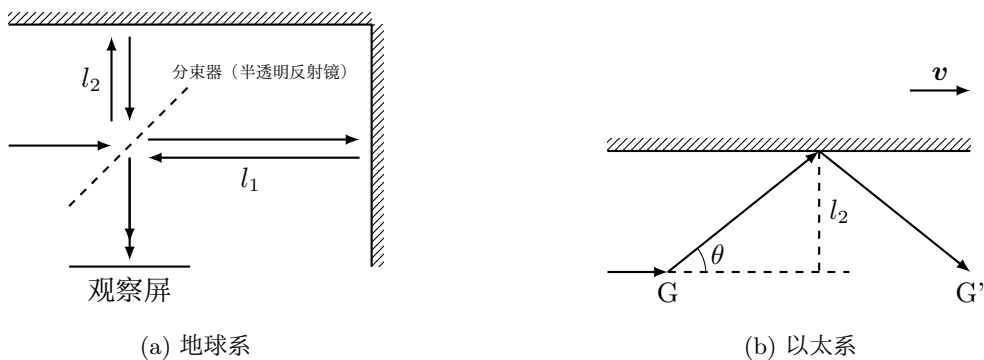


图 2.1: 迈克耳孙-莫雷干涉仪：在地球系与以太系中的光路示意图

设地球相对于以太的速度为 v ，记 $\beta = \frac{v}{c}$ 。在以太系中看，光束在 G 处分开，在 G' 处重新汇合。对于图中水平方向的光束，在地球系中计算其往返时间：

$$t_1 = \frac{l_1}{c-v} + \frac{l_1}{c+v} = \frac{2l_1/c}{1-\beta^2} \quad (2.4)$$

即光顺着地球运动的方向传播。

另一条垂直方向的光束，在以太系中计算其往返时间：

$$t_2 = 2 \frac{l_2}{\sin \theta} \frac{1}{c} = \frac{2l_2/c}{\sqrt{1-\beta^2}} \quad (2.5)$$

两束光到达 G' 点的时间差为

$$\Delta t = t_1 - t_2 = \frac{2l_1/c}{1-\beta^2} - \frac{2l_2/c}{\sqrt{1-\beta^2}} \quad (2.6)$$

相位差为

$$n = \frac{c\Delta t}{\lambda} = \frac{2l_1/\lambda}{1-\beta^2} - \frac{2l_2/\lambda}{\sqrt{1-\beta^2}} \quad (2.7)$$

而经过 1/4 年（约 91 天）后，地球运动方向改变，重新测量相位差，得到

$$n' = \frac{2l_2/\lambda}{1-\beta^2} - \frac{2l_1/\lambda}{\sqrt{1-\beta^2}} \quad (2.8)$$

即干涉条纹要移动！但是，迈克耳孙-莫雷实验的结果却是**没有**观测到任何干涉条纹的移动，这与以太假说的预期完全相反。

2.1.2 狭义相对论基本原理

为了解释迈克耳孙-莫雷实验的结果，爱因斯坦在 1905 年提出了狭义相对论，其基本原理包括两个假设：

定理 2.1.1: 狭义相对论基本原理

- 相对性原理：所有惯性参考系中的物理定律形式相同。
- 光速不变原理：在所有惯性参考系中，真空中的光速 c 是一个常数，与光源和观察者的运动状态无关。

2.2 狭义相对论时空度量相对性

2.2.1 空间与时间的测量

在介绍狭义相对论的主要内容之前，我们需要明确如何在不同参考系中测量空间和时间。

1. 空间测量：在空间中建立一个坐标系 $Oxyz$ ，制作一把标准尺（如米尺），定义单位长度。用标准尺测量物体的位置，记录其坐标值。

2. 时间测量：

- 考虑某个稳定的周期性现象（如摆钟、原子钟等），定义一个时间单位（如秒）。
- 零点校准：在 O 处设置一个时钟，定义 $t = 0$ 时刻，在此时刻发出一个光信号。光信号传播到空间中某个位置 P 处的时刻 t_P ，则定义该位置的时间为 $t_P = \frac{\overline{OP}}{c}$ 。这样就能将此参考系中任意的时钟与 O 处的时钟同步。
- 走时率校准：在 O 处的时钟与 P 处的时钟同步后。从 O 处在 t_0 时刻发出一束光信号，照射到 P 处的时钟，记录 P 处时钟显示的时间 t_P ，然后立即反射光信号回到 O 处，记录 O 处时钟显示的时间 t_O 。如果 P 处的时钟走时正确，则应满足

$$t_P - t_0 = \frac{1}{2}(t_O - t_0) \quad (2.9)$$

2.2.2 惯性系间时空测量的相对性

时间零点校准的差异（「同时」的相对性）

如图 2.2，考虑两个惯性参考系 S 和 S' ，其中 S' 以速度 v 相对于 S 运动。当 S 的坐标原点 O 和 S' 的坐标原点 O' 重合时，两参考系对表，即 $t = t' = 0$ 。

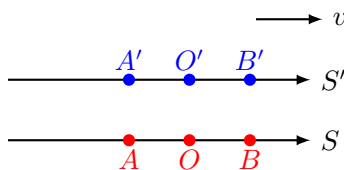


图 2.2: 「同时」的相对性

同时，位于 $O(O')$ 处的光源向左右两侧各发射一束光信号。

在 S 系中， $\overline{OA} = \overline{OB} = ct_0$ ， A 和 B 据此对表。

在 S' 系中， $\overline{O'A} = \overline{O'B} = ct_0$ ， A' 和 B' 据此对表。

然而在 S 系中的观者认为 S' 系中的对表有问题，因为 S' 系中的光源在向右运动，光信号到达 A' 的时间为

$$t = \frac{\overline{OA} - vt}{c} \quad (2.10)$$

类似地，光信号到达 B' 的时间为

$$t = \frac{\overline{OB} + vt}{c} \quad (2.11)$$

即光信号到达 A' 的时间早于到达 B' 的时间，同时是相对的。

时间间隔的差异

在不同参考系中，观察到的光的传播距离不同，而光速不变，因此观察到的时间间隔也不同。

空间测量的差异

考虑将一把标准尺放置在 S' 系中，沿 x' 轴方向静止。在 S' 系中测量该标准尺的长度为 L_0 。而在 S 系中测量该标准尺的长度时，需要在同一时刻测量标准尺的两个端点的位置。因为同时的相对性，空间测量也出现了差异。

2.3 狭义相对论时空变换及其推论

2.3.1 狭义相对论时空变换

考虑两个惯性参考系 S 和 S' ，其中 S' 以速度 v 相对于 S 运动。在不失一般性的情况下，不妨设 v 沿 x 轴正方向。

在开始推导之前，我们要定义一下什么是事件 (Event)。

定义 2.3.1: 事件

事件是指在时空中发生的一个特定的现象，可以用一组时空坐标 (t, x, y, z) 来描述。即在参考系 S 中，事件发生的时间为 t ，空间位置为 (x, y, z) 。

由于只考虑 x 轴方向的运动， y 和 z 坐标不变，可以认为 $y' = y$ ， $z' = z$ 。设事件在 S 系中的坐标为 (t, x) ，在 S' 系中的坐标为 (t', x') 。

首先， S 系中的匀速直线性运动在 S' 系中仍然是匀速直线性运动，即

$$a = 0 \iff a' = 0 \quad (2.12)$$

因此，时空变换应为线性变换（否则与实验相悖）：

$$\begin{cases} x' = a_{11}x + a_{12}t \\ t' = a_{21}x + a_{22}t \end{cases} \quad (2.13)$$

接下来，考虑 S 系与 S' 系地位互换的逆变换 $x' \rightarrow -x$ ， $x \rightarrow -x'$ ，则有

$$\begin{cases} -x = a_{11}(-x') + a_{12}t' \\ t = a_{21}(-x') + a_{22}t' \end{cases} \quad (2.14)$$

将公式 (2.14) 代入公式 (2.13)，得到

$$\begin{cases} x' = a_{11}^2 x' + a_{11}a_{12}t + a_{12}a_{21}x + a_{12}a_{22}t \\ t' = -a_{21}a_{11}x' - a_{21}a_{12}t + a_{22}a_{21}x + a_{22}^2 t \end{cases} \quad (2.15)$$

比较两式，得到

$$\begin{cases} a_{11}^2 - a_{12}a_{21} = 1 \\ a_{11} = a_{22} \end{cases} \quad (2.16)$$

再考虑 S' 系中的原点 O' 在 S 系中的运动轨迹，由于 O' 在 S' 系中静止，有

$$x' = 0 \implies x = -\frac{a_{12}}{a_{11}}t \quad (2.17)$$

因此， S' 系相对于 S 系的速度为

$$v = -\frac{a_{12}}{a_{11}} \quad (2.18)$$

即

$$a_{11}v + a_{12} = 0 \quad (2.19)$$

最后考虑光速不变原理，在 S 系中，光沿 x 轴正方向传播时，满足 $x = ct$ ，在 S' 系中应满足 $x' = ct'$ 。将 $x = ct$ 和 $x' = ct'$ 代入公式 (2.13)，得到

$$\begin{cases} ct' = a_{11}(ct) + a_{12}t \\ t' = a_{21}(ct) + a_{22}t \end{cases} \quad (2.20)$$

即

$$a_{12} = a_{21}c^2 \quad (2.21)$$

综上所述, 得到

$$\begin{cases} a_{11}^2 - a_{21}^2c^2 = 1 \\ a_{22} - a_{11} = 0 \\ a_{11}v + a_{21} = 0 \\ a_{12} - a_{21}c^2 = 0 \end{cases} \quad (2.22)$$

解得

$$\begin{cases} a_{11} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \\ a_{22} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \\ a_{21} = -\frac{v}{c^2} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \\ a_{12} = -v \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \end{cases} \quad (2.23)$$

因此有如下定理:

定理 2.3.1: 狭义相对论时空变换 (洛伦兹变换)

设两个惯性参考系 S 和 S' , 其中 S' 以速度 v 相对于 S 运动, 且 v 沿 x 轴正方向。事件在 S 系中的坐标为 (t, x, y, z) , 在 S' 系中的坐标为 (t', x', y', z') , 则有如下时空变换关系:

$$\begin{cases} x' = \frac{(x - vt)}{\sqrt{1 - \frac{v^2}{c^2}}} \\ t' = \frac{(t - \frac{vx}{c^2})}{\sqrt{1 - \frac{v^2}{c^2}}} \\ y' = y \\ z' = z \end{cases} \quad (2.24)$$

在这里, 我们顺便引入一个重要的概念, 即固有时 (proper time), 即

定义 2.3.2: 固有时

固有时是指在与事件发生的物体共动的参考系中测量到的时间, 常用符号 τ 表示。对应的时间间隔 $\Delta\tau$ 被称为固有时间间隔。

例如, 在 S' 系中测量到的时间间隔 $\Delta t'$ 就是事件发生的物体的固有时间间隔, 即 $\Delta\tau = \Delta t'$ 。

注释 2.3.1: 洛伦兹变换的历史背景

洛伦兹变换是洛伦兹 (Lorentz) 为了解决麦克斯韦方程组的协变性问题 (即电磁现象在不同惯性参考系中的形式不变性) 而提出的。后来由庞加莱 (Poincaré) 和爱因斯坦 (Einstein) 进一步发展, 成为狭义相对论的核心内容之一。

在实践中, 我们经常将上式写成如下形式:

$$\begin{cases} x' = \gamma(x - vt) \\ t' = \gamma\left(t - \frac{vx}{c^2}\right) \\ y' = y \\ z' = z \end{cases} \quad (2.25)$$

其中, $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ 称为洛伦兹因子 (Lorentz factor)。而 $\beta = \frac{v}{c}$ 则是一个无量纲的速度参数, 表示物体速度与光速之比。

此外, 洛伦兹变换的逆变换为

$$\begin{cases} x = \gamma(x' + vt') \\ t = \gamma\left(t' + \frac{vx'}{c^2}\right) \\ y = y' \\ z = z' \end{cases} \quad (2.26)$$

推导过程是简单的, 即将 v 替换为 $-v$ 即可。

2.3.2 狭义相对论时空变换的推论

时间膨胀

考虑 S' 系中的一个时钟静止在 $x' = 0$ 处, 测量其两个滴答之间的时间间隔为 $\Delta t' = t'_2 - t'_1$ 。在 S 系中测量该时钟的时间间隔为 $\Delta t = t_2 - t_1$ 。

根据洛伦兹变换, 有

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) \quad (2.27)$$

由于时钟在 S' 系中静止在 $x' = 0$ 处, 因此 $x = vt$ 。代入上式, 得到

$$t' = \gamma\left(t - \frac{v(vt)}{c^2}\right) = \gamma t \left(1 - \frac{v^2}{c^2}\right) = \frac{t}{\gamma} \quad (2.28)$$

因此, 有

$$\Delta t' = \frac{\Delta t}{\gamma} \implies \Delta t = \gamma \Delta t' \quad (2.29)$$

这表明, 在 S 系中测量到的时间间隔 Δt 大于在 S' 系中测量到的时间间隔 $\Delta t'$, 即时间膨胀效应。

长度收缩

考虑 S' 系中的一把标准尺静止在 x' 轴上, 测量其长度为 $l_0 = x'_2 - x'_1$ 。在 S 系中测量该标准尺的长度为 $l = x_2 - x_1$ 。根据洛伦兹变换, 有

$$x' = \gamma(x - vt) \quad (2.30)$$

为了在 S 系中测量标准尺的长度, 需要在同一时刻测量其两个端点的位置, 即 $t_1 = t_2$ 。因此, 有

$$x'_1 = \gamma(x_1 - vt_1) \quad (2.31)$$

$$x'_2 = \gamma(x_2 - vt_2) \quad (2.32)$$

两式相减, 得到

$$l_0 = x'_2 - x'_1 = \gamma(x_2 - x_1) = \gamma l \quad (2.33)$$

因此, 有

$$l = \frac{l_0}{\gamma} \quad (2.34)$$

这表明, 在 S 系中测量到的长度 l 小于在 S' 系中测量到的长度 l_0 , 即长度收缩效应。

因果律与时序关系的保持

考虑两个事件 A 和 B ，在 S 系中的时空坐标分别为 (t_A, x_A) 和 (t_B, x_B) 。定义事件 A 发生在事件 B 之前，即 $t_A < t_B$ 。根据洛伦兹变换，有

$$t' = \gamma \left(t - \frac{vx}{c^2} \right) \quad (2.35)$$

因此，有

$$t'_A = \gamma \left(t_A - \frac{vx_A}{c^2} \right) \quad (2.36)$$

$$t'_B = \gamma \left(t_B - \frac{vx_B}{c^2} \right) \quad (2.37)$$

两式相减，得到

$$t'_B - t'_A = \gamma \left((t_B - t_A) - \frac{v(x_B - x_A)}{c^2} \right) \quad (2.38)$$

为了保证因果律不被破坏，即事件 A 发生在事件 B 之前在所有惯性参考系中都成立，需要满足

$$(t_B - t_A) - \frac{v(x_B - x_A)}{c^2} > 0 \quad (2.39)$$

即

$$t_B - t_A > \frac{v(x_B - x_A)}{c^2} \quad (2.40)$$

这表明，只要事件 A 和事件 B 之间的时空间隔是类时的（即 $c^2(t_B - t_A)^2 > (x_B - x_A)^2$ ），则因果律在所有惯性参考系中都得到保持。只要观者所在的惯性参考系的速度 v 小于光速 c ，事件的时序关系不会被改变。

速度变换公式

考虑质点相对 S, S' 系的速度分别为

$$\mathbf{u} : u_x = \frac{dx}{dt}, \quad u_y = \frac{dy}{dt}, \quad u_z = \frac{dz}{dt} \quad (2.41)$$

$$\mathbf{u}' : u'_x = \frac{dx'}{dt'}, \quad u'_y = \frac{dy'}{dt'}, \quad u'_z = \frac{dz'}{dt'} \quad (2.42)$$

根据洛伦兹变换，有

$$x' = \gamma(x - vt) \quad (2.43)$$

$$t' = \gamma \left(t - \frac{vx}{c^2} \right) \quad (2.44)$$

对两式分别微分，得到

$$dx' = \gamma(dx - v dt) \quad (2.45)$$

$$dt' = \gamma \left(dt - \frac{v dx}{c^2} \right) \quad (2.46)$$

因此，有

$$u'_x = \frac{dx'}{dt'} = \frac{\gamma(dx - v dt)}{\gamma \left(dt - \frac{v dx}{c^2} \right)} = \frac{dx - v dt}{dt - \frac{v dx}{c^2}} \quad (2.47)$$

将 $dx = u_x dt$ 代入上式，得到

$$u'_x = \frac{u_x - v}{1 - \frac{vu_x}{c^2}} \quad (2.48)$$

类似地，有

$$u'_y = \frac{dy'}{dt'} = \frac{dy}{\gamma \left(dt - \frac{v dx}{c^2} \right)} = \frac{u_y dt}{\gamma \left(dt - \frac{vu_x dt}{c^2} \right)} = \frac{u_y}{\gamma \left(1 - \frac{vu_x}{c^2} \right)} \quad (2.49)$$

$$u'_z = \frac{dz'}{dt'} = \frac{dz}{\gamma \left(dt - \frac{v dx}{c^2} \right)} = \frac{u_z dt}{\gamma \left(dt - \frac{vu_x dt}{c^2} \right)} = \frac{u_z}{\gamma \left(1 - \frac{vu_x}{c^2} \right)} \quad (2.50)$$

综上所述，有如下定理：

定理 2.3.2: 狭义相对论速度变换公式

设质点相对 S, S' 系的速度分别为 \mathbf{u} 和 \mathbf{u}' , 则有如下速度变换关系:

$$\begin{cases} u'_x = \frac{u_x - v}{1 - \frac{vu_x}{c^2}} \\ u'_y = \frac{u_y}{\gamma \left(1 - \frac{vu_x}{c^2}\right)} \\ u'_z = \frac{u_z}{\gamma \left(1 - \frac{vu_x}{c^2}\right)} \end{cases} \quad (2.51)$$

其中, $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ 。

类似地, 速度变换的逆变换为

$$\begin{cases} u_x = \frac{u'_x + v}{1 + \frac{vu'_x}{c^2}} \\ u_y = \frac{u'_y}{\gamma \left(1 + \frac{vu'_x}{c^2}\right)} \\ u_z = \frac{u'_z}{\gamma \left(1 + \frac{vu'_x}{c^2}\right)} \end{cases} \quad (2.52)$$

注释 2.3.2: 速度变换公式的意义

速度变换公式表明, 在狭义相对论中, 速度的叠加不再是简单的代数和, 而是受到光速不变原理的限制。这确保了无论在哪个惯性参考系中, 物体的速度都不会超过光速 c 。

注释 2.3.3: 速度变换公式与双曲函数的关系

在这里, 你可能会发现这个速度变换公式与 $\tan x$ 的加法公式非常相似, 这并非巧合。实际上, 狭义相对论中的速度变换可以类比为双曲正切函数的加法公式, 这揭示了狭义相对论中速度空间的几何结构。具体来说, 设 $u = c \cdot \tanh \phi$ 和 $v = c \cdot \tanh \psi$, 则速度变换公式可以写成

$$\tanh(\phi + \psi) = \frac{\tanh \phi + \tanh \psi}{1 + \tanh \phi \tanh \psi} \quad (2.53)$$

这表明, 狭义相对论中的速度变换实际上对应于双曲角的加法, 这进一步强调了狭义相对论中时空的双曲几何性质。而此处的 ϕ 和 ψ 被称为快度角 (boost angle) 或快度参数 (rapidity)。

下面给出一个例子, 即加速度变换公式的推导。

例 2.3.1: 加速度变换公式的推导

设质点相对 S, S' 系的加速度分别为

$$\mathbf{a} : a_x = \frac{du_x}{dt}, \quad a_y = \frac{du_y}{dt}, \quad a_z = \frac{du_z}{dt} \quad (2.54)$$

$$\mathbf{a}' : a'_x = \frac{du'_x}{dt'}, \quad a'_y = \frac{du'_y}{dt'}, \quad a'_z = \frac{du'_z}{dt'} \quad (2.55)$$

解:

根据速度变换公式, 有

$$u'_x = \frac{u_x - v}{1 - \frac{vu_x}{c^2}} \quad (2.56)$$

对上式分别对 t 和 t' 微分, 得到

$$du'_x = \frac{(1 - \frac{v^2}{c^2}) du_x}{(1 - \frac{vu_x}{c^2})^2} \quad (2.57)$$

$$dt' = \frac{(1 - \frac{vu_x}{c^2}) dt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.58)$$

因此, 有

$$a'_x = \frac{du'_x}{dt'} = \frac{\frac{(1 - \frac{v^2}{c^2}) du_x}{(1 - \frac{vu_x}{c^2})^2}}{\frac{(1 - \frac{vu_x}{c^2}) dt}{\sqrt{1 - \frac{v^2}{c^2}}}} = \frac{(1 - \frac{v^2}{c^2})^{3/2} a_x}{(1 - \frac{vu_x}{c^2})^3} \quad (2.59)$$

类似地, 有

$$du'_y = \frac{\partial u'_y}{\partial u_y} du_y + \frac{\partial u'_y}{\partial u_x} du_x = \frac{du_y}{\gamma(1 - \frac{vu_x}{c^2})} + \frac{vu_y du_x}{c^2 \gamma (1 - \frac{vu_x}{c^2})^2} \quad (2.60)$$

因此, 有

$$a'_y = \frac{du'_y}{dt'} = \frac{\frac{du_y}{\gamma(1 - \frac{vu_x}{c^2})} + \frac{vu_y du_x}{c^2 \gamma (1 - \frac{vu_x}{c^2})^2}}{\frac{(1 - \frac{vu_x}{c^2}) dt}{\sqrt{1 - \frac{v^2}{c^2}}}} = \frac{a_y}{\gamma^2 (1 - \frac{vu_x}{c^2})^2} + \frac{vu_y a_x}{c^2 \gamma^2 (1 - \frac{vu_x}{c^2})^3} \quad (2.61)$$

同理, 有

$$a'_z = \frac{a_z}{\gamma^2 (1 - \frac{vu_x}{c^2})^2} + \frac{vu_z a_x}{c^2 \gamma^2 (1 - \frac{vu_x}{c^2})^3} \quad (2.62)$$

逆变换类似, 有

$$a_x = \frac{(1 - \frac{v^2}{c^2})^{3/2} a'_x}{(1 + \frac{vu'_x}{c^2})^3} \quad (2.63)$$

$$a_y = \frac{a'_y}{\gamma^2 (1 + \frac{vu'_x}{c^2})^2} - \frac{vu'_y a'_x}{c^2 \gamma^2 (1 + \frac{vu'_x}{c^2})^3} \quad (2.64)$$

$$a_z = \frac{a'_z}{\gamma^2 (1 + \frac{vu'_x}{c^2})^2} - \frac{vu'_z a'_x}{c^2 \gamma^2 (1 + \frac{vu'_x}{c^2})^3} \quad (2.65)$$

在相对论中, 加速度的变换关系比经典力学中的更为复杂, 反映了狭义相对论对时间和空间测量的影响。仅当 $\mathbf{a}' = 0$ 时, 才有 $\mathbf{a} = 0$, 即匀速直线运动在所有惯性参考系中保持不变。此外, 它还有三个特征:

- 一般情况下, $\mathbf{a} \neq \mathbf{a}'$;
- 加速度分量之间存在耦合;
- 质点加速度的变换与质点速度有关。

2.4 狭义相对论动力学

2.4.1 引入

牛顿力学中物体的质量 m 和物体所受力 \mathbf{F} 一起由关系式 $\mathbf{F} = m\mathbf{a}$ 定义, 大量实验对定义的认可使得关系式 $\mathbf{F} = m\mathbf{a}$ 成为定律。定律中的 m 不随物体的运动状态发生变化, 是一个参考系不变量。经典理论中 \mathbf{a} 是惯性系不变量, 于是牛顿定律的惯性系不变性要求 \mathbf{F} 也是惯性系不变量。经典力学涉及的诸多真实力如牛顿万有引力、重力、浮力、摩擦力、弹性力等在当时的实验精度范围内已被证实都是惯性系不变量。

爱因斯坦狭义相对论时空度量关系的建立，揭示出除非加速度为零，否则 a 不是惯性系不变量。面对这一结论，显然不能不加论证地判定在狭义相对论中 m 、 \mathbf{F} 仍然可以是惯性系不变量，而是需要一般地考察 m 、 \mathbf{F} 的惯性系变换关系。另一方面，在狭义相对论中也不能不加论证地认可，以 $\mathbf{F} = m\mathbf{a}$ 形式表述的牛顿第二定律仍然满足相对性原理的要求。总之，在狭义相对论的动力学内容中，一是要找出 m 、 \mathbf{F} 的惯性系变换式，二是要建立符合相对性原理要求的牛顿第二定律。

完成上述工作的途径不是唯一的。例如可在改造牛顿第二定律之前，先设定与外界无相互作用的物质系统其总能量（包括机械能、热能...）和动量分别守恒或者改为设定这样的系统其总质量和动量分别守恒。而后，通过两个全同质点的弹性碰撞或完全非弹性碰撞，简捷地得到质点质量 m 随其运动速度 v 的变化关系，实质上也就是得到了质量的惯性系变换式。牛顿第二定律表达式 $\mathbf{F} = m\mathbf{a}$ 中 m 的不变性既然已被否定，便改取表达式 $\mathbf{F} = \frac{d\mathbf{p}}{dt}$ 来考察它的惯性系不变性和 \mathbf{F} 的惯性系变换式。

狭义相对论的创建源于电作用理论。如前所述，爱因斯坦第一篇相对论论文的标题定为《论动体的电动力学》，全篇宗旨就是以新建的狭义相对论时空变换为基础，从理论上证明麦克斯韦场方程满足相对性原理要求即具有惯性系不变性，同时导得了电磁场量的惯性系变换式。

经过综合考量，笔者便在此处忠实的记录下这两种推导方法。

2.4.2 从力学的角度推导^[2]

动量、质量与速度的关系

首先，我们仍然定义一个质点的动量 \mathbf{p} 是一个与它的速度 \mathbf{v} 成正比的矢量，即

$$\mathbf{p} = m\mathbf{v} \quad (2.66)$$

其中，式中的比例系数仍定义为质点的质量 m 。

不过由于在数量上，质点的动量 \mathbf{p} 不再与速度 \mathbf{v} 成简单的正比关系，而是与速度 \mathbf{v} 的大小有关。因此，质点的质量 m 也不再是一个常量，而是与质点的速度 \mathbf{v} 的大小有关。又由于空间的各向同性，我们认为质点的质量 m 只能与速度 \mathbf{v} 的大小 v 有关，即 $m = m(v)$ 。接下来，我们要确定质点质量 m 随速度 v 的变化关系。

我们考虑一个例子——全同粒子的完全非弹性碰撞。考虑 A 、 B 两个全同粒子正碰之后成为一个复合粒子。我们从两个参考系 S 和 S' 来分析这个碰撞过程。在 S 系中，粒子 B 静止不动，粒子 A 以速度 v 向右运动。它们的质量分别为 $m_B = m_0$ 和 $m_A = m(v)$ 。碰撞之后，复合粒子以速度 u 向右运动。

在 S' 系中，粒子 A 静止不动，粒子 B 以速度 $-v'$ 向左运动。它们的质量分别为 $m_A = m_0$ 和 $m_B = m(v)$ 。碰撞之后，复合粒子以速度 u' 向右运动。

显然， S 系相对于 S' 系的速度为 v 。设碰撞之后复合粒子的质量为 M ，在 S 系中的速度为 u ，在 S' 系中的速度为 u' ，依对称性可以看出 $u = -u'$ 。

根据动量守恒定律，我们有

$$m(v)v = M(u)u \quad (2.67)$$

根据质量（实际上是能量）守恒定律，我们有

$$m(v) + m_0 = M(u) \quad (2.68)$$

由此得

$$\frac{M(u)}{m(v)} = \frac{m(v) + m_0}{m(v)} = \frac{v}{u} \quad (2.69)$$

另一方面，根据相对论速度变换公式，我们有

$$u' = -u = \frac{u - v}{1 - \frac{vu}{c^2}} \implies \left(\frac{v}{u}\right)^2 - 2 \cdot \frac{v}{u} + \frac{v^2}{c^2} = 0 \quad (2.70)$$

解得

$$\frac{v}{u} = 1 \pm \sqrt{1 - \frac{v^2}{c^2}} \quad (2.71)$$

因为 $\frac{v}{u} > 1$, 所以取正号, 有

$$\frac{M(u)}{m(v)} = \frac{m(v) + m_0}{m(v)} = 1 + \sqrt{1 - \frac{v^2}{c^2}} \quad (2.72)$$

解得

$$m(v) = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.73)$$

因此, 有如下定理:

定理 2.4.1: 狭义相对论质量公式

设质点的静止质量为 m_0 , 则当质点以速度 v 运动时, 其质量 m 为

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.74)$$

在此之后, 我们称式中的 m 为质点的相对论质量或动质量, 而 m_0 则称为质点的静止质量或固有质量。相应的, 质点的动量 \mathbf{p} 为

$$\mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma m_0 \mathbf{v} \quad (2.75)$$

力、功与动能

再考虑动能的表达式

$$\begin{aligned} E_k &= \int_0^l \mathbf{F} \cdot d\mathbf{s} = \int_0^v \mathbf{F} \cdot \frac{d\mathbf{s}}{dt} dt \\ &= \int_0^v \frac{d\mathbf{p}}{dt} \cdot \mathbf{v} dt = \int_0^v \mathbf{v} \cdot d\mathbf{p} \\ &= \int_0^v v d\left(\frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}}\right) \\ &= \frac{m_0 \mathbf{v} \cdot \mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \Big|_0^v - m_0 \int_0^v \frac{\mathbf{v} \cdot d\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \\ &= \frac{m_0 v^2}{\sqrt{1 - \frac{v^2}{c^2}}} + m_0 c^2 \left(\sqrt{1 - \frac{v^2}{c^2}} - 1 \right) \\ &= m_0 c^2 \left(\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right) \end{aligned} \quad (2.76)$$

因此, 有如下定理:

定理 2.4.2: 狭义相对论动能公式

设质点的静止质量为 m_0 , 当质点以速度 v 运动时, 其动能 E_k 为

$$E_k = (\gamma - 1)m_0 c^2 \quad (2.77)$$

其中, $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ 。

在 $v^2/c^2 \ll 1$ 的情况下, 将上式展开, 有

$$E_k = \frac{1}{2}m_0v^2 + \frac{3}{8}m_0\frac{v^4}{c^2} + \cdots \quad (2.78)$$

显然, 在低速极限下, 狭义相对论动能公式退化为经典力学动能公式 $E_k = \frac{1}{2}m_0v^2$ 。

质能关系

在此基础上, 我们可以定义质点的总能量 E 为

$$E = E_k + m_0c^2 = \gamma m_0c^2 \quad (2.79)$$

定理 2.4.3: 狭义相对论能量公式

设质点的静止质量为 m_0 , 当质点以速度 v 运动时, 其总能量 E 为

$$E = \gamma m_0c^2 \quad (2.80)$$

其中, $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ 。

注释 2.4.1: 质能关系

质点的总能量 E 和质量 m 之间存在如下关系:

$$E = mc^2 \quad (2.81)$$

这表明质量和能量是等价的, 可以相互转换。这一结论在核反应和粒子物理中得到了广泛的验证。

能量与动量的关系

最后我们介绍一下动质能关系

定理 2.4.4: 狭义相对论动质能关系

设质点的静止质量为 m_0 , 当质点以速度 v 运动时, 其总能量 E 和动量 p 满足如下关系:

$$E^2 = p^2c^2 + m_0^2c^4 \quad (2.82)$$

2.4.3 从电动力学的角度推导^[3]

力的定义与变换

麦克斯韦场方程是以 \mathbf{E} 和 \mathbf{B} 已有度量定义为前提的, 逻辑上可以将带电质点在电作用场中的受力公式

$$\mathbf{F} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \quad (2.83)$$

处理为对 \mathbf{E} 和 \mathbf{B} 的定义式, 式中 q 是带电质点的电量, \mathbf{u} 是带电质点的运动速度。公式 (2.83) 的含义是将电作用场分解成电场和磁场两部分。其中电场给电荷 q 的作用力 \mathbf{F}_e 与带电质点运动状态无关, 力的结构式 $\mathbf{F}_e = q\mathbf{E}$ 中的 \mathbf{E} 即为电场参与力构造的因素量。磁场仅对运动电荷施以附加的作用力 \mathbf{F}_m , 力的结构式 $\mathbf{F}_m = q\mathbf{u} \times \mathbf{B}$ 中的 \mathbf{B} 即为磁场参与力构造的因素量。经典理论中, \mathbf{F} 已在质点动力学中给出度量定义。

狭义相对论中, 在尚未完成对牛顿三定律的检查和修正工作时逻辑上也可以在电作用理论框架内, 将公式 (2.83) 处理成同时给出 \mathbf{F} 、 q 、 \mathbf{E} 、 \mathbf{B} 度量定义的基本公式。电作用理论中认定 q 是运动不变量这也是可

以理解的。宏观物体均由原子构成原子中电子的电量若因电子绕核高速运动而发生变化原子便不再是电中性的原子间会相互排斥，宏观物体很难稳定，这与事实不符。

公式 (2.83) 是惯性系不变式，符号右边的 q 是惯性系不变量，速度 \mathbf{u} 的惯性系变换式已经给出，那么在获得 \mathbf{E} 和 \mathbf{B} 的惯性系变换式后，便可导得力 \mathbf{F} 的变换式。

我们首先推导 \mathbf{E} 和 \mathbf{B} 的惯性系变换式。

先考虑麦克斯韦方程组中的

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.84)$$

即

$$\begin{cases} \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0 \\ \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -\frac{\partial B_x}{\partial t} \\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -\frac{\partial B_y}{\partial t} \\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -\frac{\partial B_z}{\partial t} \end{cases} \quad (2.85)$$

由公式 (2.24) 可得

$$\begin{cases} \frac{\partial}{\partial x} = \gamma \left(\frac{\partial}{\partial x'} - \frac{v}{c^2} \frac{\partial}{\partial t'} \right) \\ \frac{\partial}{\partial y} = \frac{\partial}{\partial y'} \\ \frac{\partial}{\partial z} = \frac{\partial}{\partial z'} \\ \frac{\partial}{\partial t} = \gamma \left(\frac{\partial}{\partial t'} - v \frac{\partial}{\partial x'} \right) \end{cases} \quad (2.86)$$

代入公式 (2.85)，可得

$$\begin{cases} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial B_x}{\partial x'} - \frac{v}{c^2 \sqrt{1-\frac{v^2}{c^2}}} \frac{\partial B_x}{\partial t'} + \frac{\partial B_y}{\partial y'} + \frac{\partial B_z}{\partial z'} = 0 \\ \frac{\partial E_z}{\partial y'} - \frac{\partial E_y}{\partial z'} = -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial B_x}{\partial t'} + \frac{v}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial B_x}{\partial x'} \\ \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial E_x}{\partial z'} - \frac{\partial E_z}{\partial x'} = -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial B_y}{\partial t'} + \frac{v}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial B_y}{\partial x'} \\ \frac{\partial E_y}{\partial x'} - \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial E_x}{\partial y'} + \frac{v}{c^2 \sqrt{1-\frac{v^2}{c^2}}} \frac{\partial E_x}{\partial t'} = -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial B_z}{\partial t'} + \frac{v}{\sqrt{1-\frac{v^2}{c^2}}} \frac{\partial B_z}{\partial x'} \end{cases} \quad (2.87)$$

从前两个公式中，先消去含 $\frac{\partial}{\partial t'}$ 的项，再消去含 $\frac{\partial}{\partial x'}$ 的项，可以得出两个等价的表达式。对后两个公式，合并同类项，也可以得出两个等价的表达式。由此得出

$$\begin{cases} \frac{\partial B_x}{\partial x'} + \frac{\partial}{\partial y'} \left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} (B_y + \frac{v}{c^2} E_z) \right] + \frac{\partial}{\partial z'} \left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} (B_z - \frac{v}{c^2} E_y) \right] = 0 \\ \frac{\partial}{\partial y'} \left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} (E_z + v B_y) \right] - \frac{\partial}{\partial z'} \left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} (E_y - v B_z) \right] = -\frac{\partial B_x}{\partial t'} \\ -\frac{\partial}{\partial x'} \left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} (E_z + v B_y) \right] + \frac{\partial E_x}{\partial z'} = -\frac{\partial}{\partial t'} \left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} (B_y + \frac{v}{c^2} E_z) \right] \\ \frac{\partial}{\partial x'} \left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} (E_y - v B_z) \right] - \frac{\partial E_x}{\partial y'} = -\frac{\partial}{\partial t'} \left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} (B_z - \frac{v}{c^2} E_y) \right] \end{cases} \quad (2.88)$$

比较上式与 公式 (2.85) 在 S' 系中的形式, 可得

$$\begin{cases} B'_x = B_x \\ B'_y = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} (B_y + \frac{v}{c^2} E_z) \\ B'_z = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} (B_z - \frac{v}{c^2} E_y) \\ E'_x = E_x \\ E'_y = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} (E_y - v B_z) \\ E'_z = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} (E_z + v B_y) \end{cases} \quad (2.89)$$

为了导出力 \mathbf{F} 的惯性系变换式, 先用 \mathbf{u} 点乘 公式 (2.83), 得

$$\mathbf{F} \cdot \mathbf{u} = q \mathbf{E} \cdot \mathbf{u} + q(\mathbf{u} \times \mathbf{B}) \cdot \mathbf{u} = q \mathbf{E} \cdot \mathbf{u} \quad (2.90)$$

同理, 在 S' 系中, 有

$$\mathbf{F}' \cdot \mathbf{u}' = q \mathbf{E}' \cdot \mathbf{u}' \quad (2.91)$$

考虑 x' 方向的力分量 F'_x :

$$F'_x = q (E'_x + u'_y B'_z - u'_z B'_y) \quad (2.92)$$

代入 \mathbf{E}' 、 \mathbf{B}' 和 \mathbf{u}' 的惯性系变换式, 得

$$\begin{aligned} F'_x &= q \left[E_x + \frac{u_y (B_z - \frac{v}{c^2} E_y)}{1 - \frac{v}{c^2} u_x} - \frac{u_z (B_y + \frac{v}{c^2} E_z)}{1 - \frac{v}{c^2} u_x} \right] \\ &= \frac{q}{1 - \frac{v}{c^2} u_x} \left(E_x + u_y B_z - u_z B_y - \frac{v}{c^2} (u_x E_x + u_y E_y + u_z E_z) \right) \\ &= \frac{1}{1 - \frac{v}{c^2} u_x} \left(F_x - \frac{v}{c^2} q \mathbf{u} \cdot \mathbf{E} \right) \end{aligned} \quad (2.93)$$

即

$$F'_x = \frac{F_x - \frac{v}{c^2} \mathbf{F} \cdot \mathbf{u}}{1 - \frac{v}{c^2} u_x} \quad (2.94)$$

类似地, 有

$$F'_y = \frac{\sqrt{1 - \frac{v^2}{c^2}} F_y}{1 - \frac{v}{c^2} u_x} \quad (2.95)$$

$$F'_z = \frac{\sqrt{1 - \frac{v^2}{c^2}} F_z}{1 - \frac{v}{c^2} u_x} \quad (2.96)$$

综上所述, 力 \mathbf{F} 的惯性系变换式为

$$\begin{cases} F'_x = \frac{F_x - \frac{v}{c^2} \mathbf{F} \cdot \mathbf{u}}{1 - \frac{v}{c^2} u_x} \\ F'_y = \frac{\sqrt{1 - \frac{v^2}{c^2}} F_y}{1 - \frac{v}{c^2} u_x} \\ F'_z = \frac{\sqrt{1 - \frac{v^2}{c^2}} F_z}{1 - \frac{v}{c^2} u_x} \end{cases} \quad (2.97)$$

逆变换为

$$\begin{cases} F_x = \frac{F'_x + \frac{v}{c^2} \mathbf{F}' \cdot \mathbf{u}'}{1 + \frac{v}{c^2} u'_x} \\ F_y = \frac{\sqrt{1 - \frac{v^2}{c^2}} F'_y}{1 + \frac{v}{c^2} u'_x} \\ F_z = \frac{\sqrt{1 - \frac{v^2}{c^2}} F'_z}{1 + \frac{v}{c^2} u'_x} \end{cases} \quad (2.98)$$

牛顿三定律的修正

- 牛顿第一定律 牛顿第一定律在狭义相对论中仍然成立，即在没有外力作用下，质点保持匀速直线运动状态。
- 牛顿第二定律 狭义相对论中，改取牛顿第二定律的原始形式，即

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} \quad (2.99)$$

其中， \mathbf{p} 是质点的动量。 dt' 和 \mathbf{F} 的惯性系变换式已经给出，对于给定的 S 系中的质点动量 \mathbf{P} ，如果能够找到对应的 S' 系中的质点动量 \mathbf{p}' ，使得

$$\mathbf{F}' = \frac{d\mathbf{p}'}{dt'} \quad (2.100)$$

成立，那么在 公式 (2.99) 表述形式上便是惯性系不变的。当然，这样修正的牛顿第二定律是否正确，还需经过实验验证。

- 牛顿第三定律 狭义相对论中，如果牛顿第三定律仍然成立，便要求在 S 系中的一对力若是 $\mathbf{F}_1 + \mathbf{F}_2 = \mathbf{0}$ ，那么变换到 S' 系也必有 $\mathbf{F}_1 + \mathbf{F}_2 = \mathbf{0}$ ，据力的变换式 公式 (2.97) 很易举出反例。

狭义相对论中牛顿第三定律虽然不能成立，但在狭义相对论之前已经建立的普遍的动量守恒定律（即任何一个与外界无相互作用的系统其动量守恒）仍然成立。

动量及能量变换

根据牛顿第二定律的修正形式 公式 (2.99)，可得质点动量定理

$$\mathbf{F} dt = d\mathbf{p} \quad (2.101)$$

同时补充定义，假设力 \mathbf{F} 对质点所做的功仍然等于质点能量的增量，即

$$\mathbf{F} \cdot d\mathbf{l} = dE \quad (2.102)$$

假设 公式 (2.101) 和 公式 (2.102) 都是惯性系不变的，则有

$$dp'_x = F'_x dt' = \frac{F_x - \frac{v}{c^2} \mathbf{F} \cdot \mathbf{u}}{1 - \frac{v}{c^2} u_x} \cdot \frac{dt - \frac{v}{c^2} dx}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.103)$$

将 $dx = u_x dt$ 和 $\mathbf{u} dt = d\mathbf{l}$ 代入上式，得

$$dp'_x = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left(dp_x - \frac{v}{c^2} dE \right) \quad (2.104)$$

积分后得

$$p'_x = \frac{p_x - \frac{v}{c^2} E}{\sqrt{1 - \frac{v^2}{c^2}}} + C_1 \quad (2.105)$$

其中 C_1 是积分常数, 可以令 $C_1 = 0$, 再在实验中验证这个选择的正确性。

同样地, 有

$$dy' = F'_y dt' = \frac{\sqrt{1 - \frac{v^2}{c^2}} F_y}{1 - \frac{v}{c^2} u_x} \cdot \frac{dt - \frac{v}{c^2} dx}{\sqrt{1 - \frac{v^2}{c^2}}} = F_y dt = dp_y \quad (2.106)$$

$$dz' = F'_z dt' = \frac{\sqrt{1 - \frac{v^2}{c^2}} F_z}{1 - \frac{v}{c^2} u_x} \cdot \frac{dt - \frac{v}{c^2} dx}{\sqrt{1 - \frac{v^2}{c^2}}} = F_z dt = dp_z \quad (2.107)$$

积分后, 令积分常数为零, 得

$$p'_y = p_y \quad (2.108)$$

$$p'_z = p_z \quad (2.109)$$

同样地, 对于能量的变换, 有

$$dE' = \mathbf{F}' \cdot d\mathbf{l}' = F'_x dx' + F'_y dy' + F'_z dz' \quad (2.110)$$

代入 \mathbf{F}' 的变换式和 $d\mathbf{l}'$ 的变换式, 得

$$\begin{aligned} dE' &= \frac{F_x - \frac{v}{c^2} \mathbf{F} \cdot \mathbf{u}}{1 - \frac{v}{c^2} u_x} \cdot \frac{dx - v dt}{\sqrt{1 - \frac{v^2}{c^2}}} + F_y dy + F_z dz \\ &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} (F_x u_x + F_y u_y + F_z u_z - v F_x) dt \end{aligned} \quad (2.111)$$

积分后, 得

$$E' = \frac{E - vp_x}{\sqrt{1 - \frac{v^2}{c^2}}} + C_2 \quad (2.112)$$

其中 C_2 是积分常数, 可以令 $C_2 = 0$, 再在实验中验证这个选择的正确性, 及其是否与 C_1 的选择相容。

综上所述, 质点的动量和能量的惯性系变换式为

$$\begin{cases} p'_x = \frac{p_x - \frac{v}{c^2} E}{\sqrt{1 - \frac{v^2}{c^2}}} \\ p'_y = p_y \\ p'_z = p_z \\ E' = \frac{E - vp_x}{\sqrt{1 - \frac{v^2}{c^2}}} \end{cases} \quad (2.113)$$

逆变换为

$$\begin{cases} p_x = \frac{p'_x + \frac{v}{c^2} E'}{\sqrt{1 - \frac{v^2}{c^2}}} \\ p_y = p'_y \\ p_z = p'_z \\ E = \frac{E' + vp'_x}{\sqrt{1 - \frac{v^2}{c^2}}} \end{cases} \quad (2.114)$$

质量变换、质能关系与能量分解

狭义相对论中, 质点的动量 \mathbf{p} 、质量 m 和速度 \mathbf{u} 之间的关系与经典一致, 且应具有惯性系不变性, 即

$$\mathbf{p} = m\mathbf{u}, \quad \mathbf{p}' = m'\mathbf{u}' \quad (2.115)$$

考虑上式的 y 分量, 有

$$m' = \frac{p'_y}{u'_y} = \frac{p_y}{u_y} \frac{1 - \frac{v}{c^2} u_x}{\sqrt{1 - \frac{v^2}{c^2}}} = m \frac{1 - \frac{v}{c^2} u_x}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.116)$$

逆变换为

$$m = \frac{p_y}{u_y} = \frac{p'_y}{u'_y} \frac{1 + \frac{v}{c^2} u'_x}{\sqrt{1 - \frac{v^2}{c^2}}} = m' \frac{1 + \frac{v}{c^2} u'_x}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.117)$$

考虑任一瞬时, S 系中的质点速度 \mathbf{u} 沿 x 轴方向, 再取相对质点瞬时静止的 S' 系, 此时 $u'_x = 0$, 在 S' 系中记 $m' = m_0$, 上式变为

$$m = \frac{m'}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.118)$$

即

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.119)$$

其余内容与前一节的推导相同, 此处不再赘述。

2.5 四维矢量与不变量

2.5.1 时空间隔的不变性

在前面的推导中, 我们介绍了狭义相对论中的动量、能量、质量等物理量的惯性系变换式。这些物理量的变换式表明, 在不同的惯性系中, 这些物理量的数值可能不同, 但它们之间存在某些组合形式在所有惯性系中保持不变。例如, 动量和能量的变换式表明, $E^2 - p^2 c^2$ 在所有惯性系中都是相同的。这启发我们引入四维矢量的概念, 以便更自然地描述狭义相对论中的物理量及其变换性质。

首先, 我们从时空间隔引入四维矢量的概念。

我们考虑时空中的两个事件 A 和 B , 在某个惯性系 S 中, 它们的坐标分别为 (t_A, x_A, y_A, z_A) 和 (t_B, x_B, y_B, z_B) 。定义事件 A 和 B 之间的时空间隔 Δs^2 为

$$\Delta s^2 = c^2(t_B - t_A)^2 - (x_B - x_A)^2 - (y_B - y_A)^2 - (z_B - z_A)^2 \quad (2.120)$$

然后, 我们对 Δs^2 做洛伦兹变换, 有

$$\begin{aligned} \Delta s'^2 &= c^2(t'_B - t'_A)^2 - (x'_B - x'_A)^2 - (y'_B - y'_A)^2 - (z'_B - z'_A)^2 \\ &= c^2 \left(\gamma(t_B - t_A) - \frac{\gamma v}{c^2} (x_B - x_A) \right)^2 - (\gamma(x_B - x_A) - \gamma v(t_B - t_A))^2 - (y_B - y_A)^2 - (z_B - z_A)^2 \\ &= c^2(t_B - t_A)^2 - (x_B - x_A)^2 - (y_B - y_A)^2 - (z_B - z_A)^2 \\ &= \Delta s^2 \end{aligned} \quad (2.121)$$

我们惊喜地发现, 时空间隔 Δs^2 在洛伦兹变换下保持不变, 即 $\Delta s'^2 = \Delta s^2$ 。这表明 Δs^2 是一个不变量。

2.5.2 闵可夫斯基空间与四维矢量

这启发我们引入如下定义

定义 2.5.1: 四维矢量

相对论下, 时间和空间不再是相互无关的。用四维矢量表示时空, 形式上非常优雅、简洁。约定:

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) \quad (2.122)$$

其中, x^0 是时间分量, x^1 、 x^2 、 x^3 是空间分量。

注释 2.5.1: 一些相对超前的内容

时空间隔 ds^2 (这里采用微分形式) 可以写成

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad (2.123)$$

其中矩阵 $\eta_{\mu\nu}$ 是闵可夫斯基 (Minkowski) 度量张量, 定义为

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (2.124)$$

令固有时空间隔为 $c^2 d\tau^2 = ds^2$, 则: $d\tau = \sqrt{1 - \frac{v^2}{c^2}} dt$ 。一般的洛伦兹变换可以表示为

$$x'^\mu = \Lambda^\mu_\nu x^\nu + a^\mu \quad (2.125)$$

其中 Λ^μ_ν 是洛伦兹变换矩阵, 满足

$$\Lambda^\mu_\alpha \Lambda^\nu_\beta \eta_{\mu\nu} = \eta_{\alpha\beta} \quad (2.126)$$

如果规定 $t = t' = 0$ 时, 两个惯性系的空间坐标重合, 则 $a^\mu = 0$, 此时洛伦兹变换简化为

$$x'^\mu = \Lambda^\mu_\nu x^\nu \quad (2.127)$$

注释 2.5.2: 为何避免使用「ict」

在一些教材中, 可能有将时间写成虚变量 $\omega = ict$ 的习惯, 这样时空间隔 ds^2 就可以写成 $ds^2 = d\omega^2 + dx^2 + dy^2 + dz^2$ 的形式, 这种定义方式被称为“欧几里得度规” (Euclidean Metric)。这种定义方式的优点是可以将时空看作一个四维欧几里得空间, 数学上更为简洁。但这种定义方式的缺点是时间分量变成了虚数, 物理意义不太直观。而且我们应该避免在物理学中引入不必要的虚数, 以免造成混淆。

我们在这里引入了四维矢量模长的概念, 即

定义 2.5.2: 四维矢量模长的定义

先引入四维矢量模方的定义

$$A^2 = \eta_{\mu\nu} A^\mu A^\nu \quad (2.128)$$

也就是

$$A^2 = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 \quad (2.129)$$

则四维矢量模长的定义为

$$|A| = \sqrt{|A^2|} \quad (2.130)$$

而且我们可以证明, 四维矢量如果在洛伦兹变换下保持不变, 则其模长和与其他同样保持不变的四维矢量的内积也保持不变, 即

定理 2.5.1: 四维矢量模长与内积的不变性

如果 A^μ 和 B^μ 是两个在洛伦兹变换下保持不变的四维矢量, 则 A^2 、 B^2 和 $A \cdot B = \eta_{\mu\nu} A^\mu B^\nu$ 都是洛伦兹不变量。

如果 S 相对于 S' 沿 $x(x')$ 轴以速度 v 运动, 则洛伦兹变换可写成矩阵形式

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \quad (2.131)$$

若令 $\beta = \tanh \phi$, 则 $\gamma = \cosh \phi$, $\gamma\beta = \sinh \phi$, 上式可以写成

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh \phi & -\sinh \phi & 0 & 0 \\ -\sinh \phi & \cosh \phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \quad (2.132)$$

注释 2.5.3: 西海岸度规与东海岸度规

在上述内容中, 我们定义时空间隔 Δs^2 时, 采用了 $c^2t^2 - x^2 - y^2 - z^2$ 的形式, 这种定义方式被称为“西海岸度规”(West Coast Metric)。这样可以保证对于真实发生的事件, 时空间隔 Δs^2 是正数。常用于粒子物理和高能物理领域的文献中。还有一种定义方式是 $-c^2t^2 + x^2 + y^2 + z^2$, 这种定义方式被称为“东海岸度规”(East Coast Metric)。这种定义方式的优点是可以自然的退回到经典物理中的空间距离定义, 但对于真实发生的事件, 时空间隔 Δs^2 是负数。常用于广义相对论领域的文献中。两者的区别在于时间分量的符号不同, 但它们描述的物理内容是相同的。在本章中, 我们采用西海岸度规的定义方式。

2.5.3 四维速度

先前面定义的质点速度 \mathbf{u} 的惯性系变换式为

$$\begin{cases} u'_x = \frac{u_x - v}{1 - \frac{v}{c^2}u_x} \\ u'_y = \frac{\sqrt{1 - \frac{v^2}{c^2}}u_y}{1 - \frac{v}{c^2}u_x} \\ u'_z = \frac{\sqrt{1 - \frac{v^2}{c^2}}u_z}{1 - \frac{v}{c^2}u_x} \end{cases} \quad (2.133)$$

可以看出, 质点速度 \mathbf{u} 的惯性系变换式与坐标 x^μ 的惯性系变换式非常相似, 这启发我们引入四维速度的概念。

定义 2.5.3: 四维速度

质点的四维速度定义为

$$u^\mu = (u^0, u^1, u^2, u^3) = (\gamma c, \gamma u_x, \gamma u_y, \gamma u_z) \quad (2.134)$$

其中, u^0 是时间分量, u^1 、 u^2 、 u^3 是空间分量。注意此处的 γ 不是两个惯性系之间的相对速度的洛伦兹因子, 而是质点速度 \mathbf{u} 的洛伦兹因子, 即 $\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}$ 。

它实际上可以由坐标 x^μ 的四维导数定义, 即

$$u^\mu = \frac{dx^\mu}{d\tau} = \gamma(c, u_x, u_y, u_z) \quad (2.135)$$

其中 τ 是质点的固有时间。

定理 2.5.2: 四维速度的模长

质点的四维速度的模长恒为 c , 即

$$u^2 = (u^0)^2 - (u^1)^2 - (u^2)^2 - (u^3)^2 = c^2 \quad (2.136)$$

在惯性系 S 和 S' 中, 四维速度的服从洛伦兹变换, 即

$$\begin{cases} u'^0 = \gamma(u^0 - \beta u^1) \\ u'^1 = \gamma(u^1 - \beta u^0) \\ u'^2 = u^2 \\ u'^3 = u^3 \end{cases} \quad (2.137)$$

或

$$\begin{cases} \gamma' c = \gamma_v(\gamma c - \beta_v \gamma u_x) \\ \gamma' u'_x = \gamma_v(\gamma u_x - \beta_v \gamma c) \\ \gamma' u'_y = \gamma_v \gamma u_y \\ \gamma' u'_z = \gamma_v \gamma u_z \end{cases} \quad (2.138)$$

其中 v 是两个惯性系之间的相对速度, $\beta = \frac{v}{c}$, $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ 。

由公式 (2.138) 的第一个式子可以得出

$$\frac{\gamma_v \gamma}{\gamma'} = \frac{1}{1 - \frac{v}{c^2} u_x} \quad (2.139)$$

将上式代入公式 (2.138) 的后三个式子, 得

$$\begin{cases} u'_x = \frac{u_x - v}{1 - \frac{v}{c^2} u_x} \\ u'_y = \frac{\sqrt{1 - \frac{v^2}{c^2}} u_y}{1 - \frac{v}{c^2} u_x} \\ u'_z = \frac{\sqrt{1 - \frac{v^2}{c^2}} u_z}{1 - \frac{v}{c^2} u_x} \end{cases} \quad (2.140)$$

即四维速度的空间分量的惯性系变换式与质点速度 \mathbf{u} 的惯性系变换式相同。

2.5.4 四维动量

根据质点动量 \mathbf{p} 和四维速度 u^μ 的定义, 我们可以定义质点的四维动量为

定义 2.5.4: 四维动量

质点的四维动量定义为

$$p^\mu = m_0 u^\mu = \left(\frac{E}{c}, p_x, p_y, p_z \right) \quad (2.141)$$

其中, p^0 是能量分量, p^1 、 p^2 、 p^3 是动量分量。

定理 2.5.3: 四维动量的模长

质点的四维动量的模长恒为 $m_0 c$, 即

$$p^2 = (p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = m_0^2 c^2 \quad (2.142)$$

四维动量的惯性系变换式与四维速度的惯性系变换式相同，即

$$\begin{cases} p'^0 = \gamma(p^0 - \beta p^1) \\ p'^1 = \gamma(p^1 - \beta p^0) \\ p'^2 = p^2 \\ p'^3 = p^3 \end{cases} \quad (2.143)$$

或

$$\begin{cases} \frac{E'}{c} = \gamma_v \left(\frac{E}{c} - \beta_v p_x \right) \\ p'_x = \gamma_v \left(p_x - \beta_v \frac{E}{c} \right) \\ p'_y = p_y \\ p'_z = p_z \end{cases} \quad (2.144)$$

四维动量还可以用于推导光的多普勒效应的频率变换式。一般的机械波的多普勒公式既与波源的速度 v_S 有关，也与观察者的速度 v_D 有关。但对于光的多普勒效应，频率的变换式只与观察者相对于波源的相对速度 v 有关，而与观察者和波源各自的速度无关。

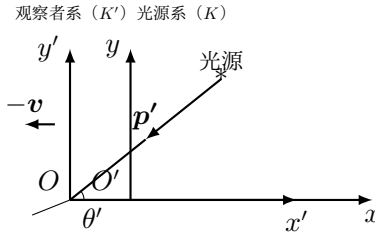


图 2.3: 光的多普勒效应

如图 2.3 所示，设观察者相对光源速度为 $-v$ ，沿 x 轴方向运动。以光源为 S 系，观察者为 S' 系，四维动量的洛伦兹逆变换给出

$$E = \gamma_v(E' - \beta_v p'_x c) \quad (2.145)$$

其中负号抵消。设 S' 系中光源到观察者连线与 x 轴的夹角为 θ' ，则 $p'_x = p' \cos \theta' = \frac{E'}{c} \cos \theta'$ ，此外 $E = h\nu$ ， $E' = h\nu'$ ，代入上式，得

$$\nu = \gamma_v(1 - \beta_v \cos \theta')\nu' = \frac{1 - \frac{v}{c} \cos \theta'}{\sqrt{1 - \frac{v^2}{c^2}}}\nu' \quad (2.146)$$

或

$$\frac{\nu'}{\nu} = \frac{\sqrt{1 - \frac{v^2}{c^2}}}{1 - \frac{v}{c} \cos \theta'} \quad (2.147)$$

$\theta' = 0$ （靠近）和 $\theta' = \pi$ （远离）时，得到光的纵向多普勒效应的频率变换式

$$\frac{\nu'}{\nu} = \sqrt{\frac{1 \mp \frac{v}{c}}{1 \pm \frac{v}{c}}} \approx 1 \pm \frac{v}{c} \quad \text{或} \quad \frac{\Delta\nu}{\nu} = \frac{\nu' - \nu}{\nu} \approx \pm \frac{v}{c} \quad (2.148)$$

$\theta' = \frac{\pi}{2}$ 时，得到光的横向多普勒效应的频率变换式

$$\frac{\nu'}{\nu} = \sqrt{1 - \frac{v^2}{c^2}} \approx 1 - \frac{v^2}{2c^2} \quad \text{或} \quad \frac{\Delta\nu}{\nu} = \frac{\nu' - \nu}{\nu} \approx -\frac{v^2}{2c^2} \quad (2.149)$$

从上式可以看出横向多普勒效应的频率变换式中没有一阶项，只有二阶项，经典情况下将其忽略了。

2.5.5 四维力

根据质点动量 \mathbf{p} 和四维动量 p^μ 的定义，我们可以定义质点的四维力为

定义 2.5.5: 四维力

质点的四维力定义为

$$f^\mu = \frac{dp^\mu}{d\tau} = \left(\frac{1}{c} \frac{dE}{dt}, \frac{dp_x}{dt}, \frac{dp_y}{dt}, \frac{dp_z}{dt} \right) \quad (2.150)$$

其中, f^0 是能量分量, f^1 、 f^2 、 f^3 是动量分量。

定理 2.5.4: 四维力的模长

质点的四维力的模长恒为零, 即

$$f^2 = (f^0)^2 - (f^1)^2 - (f^2)^2 - (f^3)^2 = 0 \quad (2.151)$$

四维力的惯性系变换式与四维动量的惯性系变换式相同, 即

$$\begin{cases} f'^0 = \gamma(f^0 - \beta f^1) \\ f'^1 = \gamma(f^1 - \beta f^0) \\ f'^2 = f^2 \\ f'^3 = f^3 \end{cases} \quad (2.152)$$

或

$$\begin{cases} \frac{1}{c} \frac{dE'}{dt} = \gamma_v \left(\frac{1}{c} \frac{dE}{dt} - \beta_v \frac{dp_x}{dt} \right) \\ \frac{dp'_x}{dt} = \gamma_v \left(\frac{dp_x}{dt} - \beta_v \frac{1}{c} \frac{dE}{dt} \right) \\ \frac{dp'_y}{dt} = \frac{dp_y}{dt} \\ \frac{dp'_z}{dt} = \frac{dp_z}{dt} \end{cases} \quad (2.153)$$

2.5.6 不变量的应用

在核物理和粒子物理中经常需要在质心系 (CM) 和实验室系 (L) 之间变来变去, 而核反应和粒子反应的最基本规律是动量和能量的守恒定律。所以把守恒定律写成与参考系无关的不变形式是很方便的。设反应式为

$$A_1 + A_2 + \dots \rightarrow A'_1 + A'_2 + \dots \quad (2.154)$$

其中 A_i 和 A'_i 分别是反应前后参与反应的粒子或核, 相应的动量、能量守恒定律为

$$\mathbf{p}_1 + \mathbf{p}_2 + \dots = \mathbf{p}'_1 + \mathbf{p}'_2 + \dots \quad (2.155)$$

$$E_1 + E_2 + \dots = E'_1 + E'_2 + \dots \quad (2.156)$$

将上面两式合成一个四维动量守恒定律, 即

$$p_1^\mu + p_2^\mu + \dots = p_1'^\mu + p_2'^\mu + \dots \quad (2.157)$$

且从它导出许多洛伦兹不变式来。例如取两端的模方, 即有

$$(p_1^\mu + p_2^\mu + \dots)^2 = (p_1'^\mu + p_2'^\mu + \dots)^2 \quad (2.158)$$

就是一个很常用的洛伦兹不变式, 在下面的讨论中还会用到一些其他的洛伦兹不变式。

两体反应的阈能

利用加速器使 A_1 、 A_2 两粒子碰撞，以产生某个或某些静质量为 M_0 的新粒子 A_3 ：

$$A_1 + A_2 \rightarrow A_1 + A_2 + A_3 \quad (2.159)$$

这是相当典型的一类粒子反应。在所有产物 (A_1 、 A_2 、 A_3) 都相对静止的情况下所需能量最少，这能量称为反应的阈能 (threshold)。此时与产物相对静止的参考系显然是质心系。

过去实现粒子反应的惯用的办法是，将一种粒子 (譬如 A_1) 加速到很高的能量去撞击静止的靶粒子 A 。为了分析阈能问题，首先根据动量、能量守恒定律写出此反应的洛伦兹不变式：

$$(p_1^\mu + p_2^\mu)^2 = (p_1'^\mu + p_2'^\mu + p_3'^\mu)^2 \quad (2.160)$$

如前所述，在刚达到阈能条件时，质心系里所有产物都是静止的。运用上式右端于质心系：

$$\begin{aligned} (p_1'^\mu + p_2'^\mu + p_3'^\mu)^2 &= (m_{10} + m_{20} + M_0)^2 c^2 \\ &= (m_{10}^2 + m_{20}^2 + M_0^2 + 2m_{10}m_{20} + 2(m_{10} + m_{20})M_0)c^2 \end{aligned} \quad (2.161)$$

式中 m_{10} 和 m_{20} 分别是 A_1 和 A_2 的静质量， M_0 是新粒子 A_3 的静质量。在实验室系中， A_2 静止，所以 $p_2^\mu = (m_{20}c, 0, 0, 0)$ ，而 $p_1^\mu = (\frac{E_1}{c}, p_{1x}, p_{1y}, p_{1z})$ 。因此，式中左端为

$$\begin{aligned} (p_1^\mu + p_2^\mu)^2 &= (p_1^\mu)^2 + (p_2^\mu)^2 + 2p_1^\mu p_{2\mu} \\ &= m_{10}^2 c^2 + m_{20}^2 c^2 + 2m_{20}E_1 \end{aligned} \quad (2.162)$$

这里的 E_1 就是在实验室系加速器所提供的阈能 $E_{\text{阈}}^L$ 。将左、右端的表达式代回原式，得

$$E_{\text{阈}}^L = \frac{(m_{10} + m_{20} + M_0)^2 - m_{10}^2 - m_{20}^2}{2m_{20}} c^2 \quad (2.163)$$

而现代大加速器多采用对撞机的形式。假定 A_1 、 A_2 是同种粒子， $m_{10} = m_{20} = m_0$ ， $\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{0}$ ， $E_1 + E_2 = 2E$ ，对撞机所提供的阈能 $E_{\text{阈}}^{CM} = 2E$ 。在质心系中有

$$\begin{aligned} (p_1^\mu + p_2^\mu)^2 &= (p_1^\mu)^2 + (p_2^\mu)^2 + 2p_1^\mu p_{2\mu} \\ &= m_0^2 c^2 + m_0^2 c^2 + 2(E^2/c^2 - \mathbf{p}_1 \cdot \mathbf{p}_2) \\ &= 4E^2/c^2 \\ &= (E_{\text{阈}}^{CM})^2/c^2 \end{aligned} \quad (2.164)$$

在 A_1 、 A_2 相同的情况下，原式右端简化为 $(2m_0 + M_0)^2 c^2$ ，于是有

$$E_{\text{阈}}^{CM} = (2m_0 + M_0)c^2 \quad (2.165)$$

在高能情况下可以产生静质量很大的新粒子，这时 $M_0 \gg m_0$ ，所以上述式子可以近似为

$$E_{\text{阈}}^L \approx \frac{M_0^2 c^2}{2m_0}, \quad E_{\text{阈}}^{CM} \approx M_0 c^2 \quad (2.166)$$

二者之比为

$$\frac{E_{\text{阈}}^L}{E_{\text{阈}}^{CM}} \approx \frac{M_0}{2m_0} \gg 1 \quad (2.167)$$

上述式子表明加速器的能量越高，如果不采用对撞机形式，资用能也就是有效能量的比例就越小。

粒子的衰变

一个粒子 A 衰变为两个粒子, 即

$$A \rightarrow A_1 + A_2 \quad (2.168)$$

最简单, 也是最常见的衰变。下面我们试用洛伦兹不变式来推导衰变物的能量。动量、能量守恒定律 $p^\mu = p_1^\mu + p_2^\mu$ 可改写成

$$p^\mu - p_1^\mu = p_2^\mu \quad (2.169)$$

两端取模方, 得

$$(p^\mu - p_1^\mu)^2 = (p_2^\mu)^2 \quad (2.170)$$

即

$$p^2 + p_1^2 - 2p^\mu p_{1\mu} = p_2^2 \quad (2.171)$$

在质心系中 $\mathbf{p} = \mathbf{0}$, $E = m_0 c^2$, $p^\mu p_{1\mu} = m_0 E_1$, 于是

$$E_1 = \frac{m_0^2 + m_{10}^2 - m_{20}^2}{2m_0} c^2 \quad (2.172)$$

同样的, E_2 的表达式为

$$E_2 = \frac{m_0^2 + m_{20}^2 - m_{10}^2}{2m_0} c^2 \quad (2.173)$$

其中 m_0 是衰变前粒子的静质量, m_{10} 和 m_{20} 分别是衰变后两个粒子的静质量。

2.6 相对论的几何表示——「时空图」

参考文献

- [1] Hue-But. 常微分方程的微分算子法[EB/OL]. 知乎. 2021 [2026-01-27]. <https://zhuanlan.zhihu.com/p/299733023> (引用页: 24).
- [2] 赵凯华, 罗蔚茵. 力学[M]. 2 版. 北京: 高等教育出版社, 2004 (引用页: 44, 56).
- [3] 舒幼生. 力学[M]. 2 版. 北京: 北京大学出版社, 2005 (引用页: 58).