Chapter 19 PARTIAL DERVATIVES

19.1

19 PARTIAL DERIVATIVES

19.1 FUNCTIONS OF SEVERAL VARIABLES

Many of the functions that arise in mathematics and its applications involve two or more independent variables. We have already met functions of this kind in our study of solid analytic geometry. Thus, the equation $z = x^2 - y^2$ is the equation of a certain saddle surface, but it also defines z as a function of the two variables x and y, and the surface can be thought of as the graph of this function.

We usually denote an arbitrary function of the two variables x and y by writing $z = f(x, y)$, and we can visualize such a function by sketching—or imagining—its graph in xyz-space, as suggested in Fig. 19.1. In this figure, $P = (x, y)$ is a "suitable" point in the xy-plane—that is, a point in the domain D of the function—and z is the directed distance up or down to the corresponding point on the surface. This surface is thought of as lying "over" the domain D, even though part of it may actually be below the xy-plane.

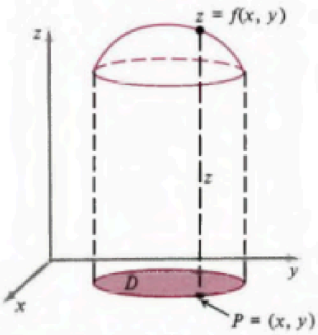


Figure 19.1 A surface in space.

By an obvious extension of the notation used here, $w = f(x, y, z, t, u, v)$ is a function of the six variables displayed in parentheses. For example, if the temperature T at a point P inside a solid iron sphere depends on the three rectangular coordinates x, y, and z of P, then we write $T = f(x, y, z)$; and if we also allow for the possibility that the temperature at a given point may vary with the time t, then T is a function of all four variables, $T = f(x, y, z, t)$.

In this chapter we shall see that the main themes of single-variable differential calculus—derivatives, rates of change, chain rule computations, maximum-minimum problems, and differential equations—can all be extended to functions of several variables. However, students should be prepared for the fact that there are striking differences between single-variable calculus and multivariable calculus. Since most of these differences already show up in functions of only two independent variables, we usually emphasize this case, and refer more briefly to functions of three or more variables. In the next chapter we turn to the integral calculus of functions of several variables.

DOMAIN

Just as in our previous work, the *domain* (or *domain of definition*) of a function $z = f(x, y)$ is the set of all points $P = (x, y)$ in the xy-plane for which there exists a corresponding z, and similarly for functions defined in xyz-space, $xyzt$-space, etc. Most of the functions we deal with are defined by formulas, and in these cases the domain is understood to be the largest set of points for which the formula makes sense. For example, the domain of

$$z = f(x, y) = \frac{1}{x - y}$$

is understood to be the set of all points (x, y) with $x \neq y$, that is, all points in the xy-plane that do not lie on the line $y = x$. The domain of

$$z = g(x, y) = \sqrt{9 - x^2 - y^2}$$

is the set of all points (x, y) for which $9 - x^2 - y^2 \geq 0$, that is, the circular disk $x^2 + y^2 \leq 9$ of radius 3 with center at the origin. And the domain of

$$w = h(x, y, z) = \frac{2x + 3y + 4z}{x^2 + y^2 + z^2}$$

is the set of all points (x, y, z) for which $x^2 + y^2 + z^2 \neq 0$, that is, all points of xyz-space except the origin.

In discussing a general function $z = f(x, y)$, we shall often require that this function be defined at a certain point P_0 and throughout some *neighborhood* of this point. This means that the domain of $f(x, y)$ must include not only P_0 itself, but also every point "sufficiently close" to P_0, that is, every point in some small circular disk centered on P_0. Similar remarks apply to functions defined in xyz-space, etc.

CONTINUITY

There are several places in this chapter where it will be necessary to mention continuity in order to state things correctly. This concept extends in a natural way from the one-variable case to functions $f(x, y)$, as follows.

A function $f(x, y)$ is said to be *continuous* at a point (x_0, y_0) in its domain if its value $f(x, y)$ can be made as close as we please to $f(x_0, y_0)$ by taking the point (x, y) close enough to (x_0, y_0), that is, if $|f(x, y) - f(x_0, y_0)|$ can be made as small as we please by making both $|x - x_0|$ and $|y - y_0|$ small enough. For example, $f(x, y) = xy$ is continuous at any point (x_0, y_0), because

$$\begin{aligned} |xy - x_0y_0| &= |xy - xy_0 + xy_0 - x_0y_0| \\ &= |x(y - y_0) + y_0(x - x_0)| \\ &\leq |x||y - y_0| + |y_0||x - x_0|, \end{aligned}$$

and it is easy to see that the quantity last written can be made as small as we please by making both $|x - x_0|$ and $|y - y_0|$ small enough.

On the other hand, the function defined by

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0), \end{cases} \tag{1}$$

is not continuous at the origin $(0, 0)$. For, if we let (x, y) approach $(0, 0)$ along a line $y = mx$ with $m \neq 0$, then

$$f(x, y) = \frac{mx^2}{x^2 + m^2x^2} = \frac{m}{1 + m^2}, \tag{2}$$

which is a nonzero constant, and these values cannot be made as close as we please to $f(0, 0) = 0$ by making (x, y) close enough to $(0, 0)$. To express this in another way, (2) shows that the values of the function approach different limiting values as the point (x, y) approaches the origin from different directions, and this is impossible if the function is continuous at the origin.

We shall not pursue the details of this topic any further, beyond making the rather loose statement that any finite combination of elementary functions is continuous at each point of its domain. Also, continuity is defined in essentially the same way for functions of three or more variables.

1. 标题与内容概述

《多变量函数与连续性的初步概念——从定义域到连续性》

本部分内容主要介绍了多变量函数（例如 $z = f(x, y)$ ）的基本概念，包括函数的图像、定义域与多变量函数在给定点及其邻域内如何被定义。同时还详细讨论了多变量函数的连续性，给出了一些典型例子来说明判定多变量函数连续与否的方法和思路。

---

2. 详细内容解析

下面我们按课件页面顺序逐点展开分析，帮助你从零开始理解多变量函数（主要是二维变量情况）如何被定义、它们的图像与定义域，以及多变量连续性的概念与例子。

2.1 多变量函数及其图像

1. **函数形式与变量**：
   - 当一个函数依赖两个或更多独立变量时，我们称之为多变量函数。例如，若函数 z 同时依赖于 x 和 y ，我们通常写作

$$z = f(x, y).$$

   - 从几何上看，若把 x, y 看作平面上的坐标，那么 z 则对应空间第三维的“高度”（可以是正也可以是负），这样就能在三维空间中画出这个函数的“图面（surface）”，如图 19.1 所示。
2. **示例：马鞍面**
   - 书中给了一个典型例子： $z = x^2 - y^2$ 。这对应一个马鞍面 (saddle surface)，可以视为依赖 x, y 的函数图像。
3. **通用的多变量函数**：
   - 作者也提到，我们可以将函数扩大到更多变量，例如：

$$w = f(x, y, z, t, u, v),$$

   用于表示更加广泛的物理量（如热力学、流体力学等），这些物理量可能依赖于多重空间坐标以及时间等其他参数。
4. **与单变量微积分的联系**：
   - 本章将要介绍的主要是把单变量微积分中的概念（例如导数、求导法则、极值问题、微分方程等等）推广到多变量的情形。但需要注意，多变量与单变量的微积分之间也存在一些重要差异。

2.2 定义域（Domain）

1. **定义域的含义**：
   - 对于一个多变量函数 $z = f(x, y)$ ，它的定义域是指所有使得该公式有意义或能正常计算的 (x, y) 的集合。
   - 有时候题目会直接给出函数表达式，但没有显式规定定义域，那么我们默认它的定义域就是使表达式“有意义”的最大集合。例如分母不为零、被开方的部分不为负等。
2. **具体示例**：

- **示例 1：**

$$z = f(x, y) = \frac{1}{x - y}.$$

该函数的定义域就是所有满足 $x \neq y$ 的点 (x, y)。也就是说，在 $x = y$ 这一条线上（$y = x$）函数没有定义。

- **示例 2：**

$$z = g(x, y) = \sqrt{9 - x^2 - y^2}.$$

该函数的定义域是满足

$$9 - x^2 - y^2 \geq 0$$

的所有 (x, y)。几何上，这就是以原点为圆心、半径为 3 的圆盘（包括圆周）。

- **示例 3：**

$$w = h(x, y, z) = \frac{2x + 3y + 4z}{x^2 + y^2 + z^2}.$$

定义域为所有满足 $x^2 + y^2 + z^2 \neq 0$ 的 (x, y, z)，也就是三维空间里除去原点的所有点。

3. **函数在某点及其邻域的定义：**
   - 要讨论函数在某点 P_0 是否定义良好，往往还要求它在这个点的"周围"也要有定义，也就是需要在某个小邻域内都可以计算。
   - 换言之，如果一个函数仅在离 P_0 非常远的地方定义，而在 P_0 的附近没有定义，那么我们就不能在 P_0 谈论它的相关性质（例如连续性或可微性）。

2.3 多变量函数的连续性

1. **连续性的直觉：**
   - 在单变量微积分中，我们说 $f(x)$ 在 $x = x_0$ 处连续，是指当 x 足够接近 x_0 时，$f(x)$ 足够接近 $f(x_0)$。
   - 对于两变量甚至多变量的函数，思路类似：如果 (x, y) 足够接近 (x_0, y_0)，那么 $f(x, y)$ 要足够接近 $f(x_0, y_0)$。
2. **形式化定义：**
   - 对函数 $f(x, y)$ 在其定义域的某点 (x_0, y_0) 连续，意味着：

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0).$$

   - 等价说法是：对任何给定的"误差" ε，都能在 (x_0, y_0) 附近找到一个"邻域" δ 使得只要 $\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$，就有 $|f(x, y) - f(x_0, y_0)| < \varepsilon$。
   - 这种定义可以自然地推广到更多变量的函数。
3. **例子：$f(x, y) = xy$ 的连续性**
   - 书中举例说明：$f(x, y) = xy$ 在它的定义域（即整个平面 $\mathbb{R}^2$）内处处连续。
   - 他们用不等式

$$|xy - x_0y_0| = |xy - x_0y + x_0y - x_0y_0| \leq |x(y - y_0)| + |y_0(x - x_0)|$$

展示了当 (x, y) 逼近 (x_0, y_0) 时，上式可以变得非常小，因此 xy 是连续的。

4. **不连续的例子：**
   - 书中给了一个两段式函数：

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{当 } (x, y) \neq (0, 0); \\ 0, & \text{当 } (x, y) = (0, 0). \end{cases}$$

   - 若要考察 $(0, 0)$ 处的连续性，就看极限 $\lim_{(x,y)\to(0,0)} f(x, y)$ 是否等于 $f(0, 0)$ 的值，即 0。
   - 书中令 (x, y) 沿不同直线靠近 $(0, 0)$，比如令 $y = mx$，可得到：

$$f(x, mx) = \frac{x(mx)}{x^2 + (mx)^2} = \frac{mx^2}{x^2(1 + m^2)} = \frac{m}{1 + m^2},$$

只要 $m \neq 0$，这是一个非零常数，与 m 有关，而不是随 (x, y) 靠近原点而趋于 0。

   - 这说明从不同方向逼近原点会有不同的极限值，因此总体不存在唯一的极限值。故该函数在 $(0, 0)$ 处不连续。

5. **结论与推广：**
   - 任何有限次加减乘除以及复合（只要不涉及到分母为零等非法运算）的"初等函数"在其定义域内都是连续的。
   - 高维函数在本质上和二维一样，也可以用相同的极限概念来定义连续性。

---

3. 本次内容小结

- **多变量函数的图像**
  在三维空间（或更高维空间）中，可以把 (x, y) 视为平面（或者更一般的域）中的点，把 $z = f(x, y)$ 视为该点在第三维度上的值，形成一个空间曲面（或超曲面）。
- **定义域**
  多变量函数的定义域是使表达式合法、有意义的所有点的集合。有些函数会因为分母为零、开平方数小于 0 等原因导致在部分区域不可定义。
- **连续性**
  多变量函数在某一点的连续性与单变量函数的概念相似：如果在那一点附近任意方向逼近，该函数值都一致地趋于函数在该点的值，则函数在该点连续；若不同方向逼近产生不同的极限值，则不连续。

---

4. 报错或补充说明

- 从截图中能看到的内容已全部覆盖并解析完毕；若有任何截图内容缺失或无法辨认，请及时告知。若无上述问题，请忽略此条。

以上就是本次关于"多变量函数定义、图像、定义域与连续性"的完整笔记。阅读完这些内容后，即使没有其他课本辅助，也能理解本章开头对多变量函数基本概念和连续性的引入与示例。

19.1.2

LEVEL CURVES

Many simple functions $z = f(x, y)$ have graphs that are much too difficult to sketch. Fortunately there is another way to understand and express the geometric nature of such a function.

The basic idea comes from the art of the mapmaker. In mapping terrain with valleys, hills, and mountains, it is common practice to draw curves joining points of constant elevation. When these curves are included on a map and properly labeled, the resulting topographical map enables an experienced user to obtain a clear mental picture of the contours of the land in three-dimensional space from this two-dimensional representation.

We can do the same thing to portray a function $z = f(x, y)$ of two variables. For any value c that $f(x, y)$ assumes, we can sketch the curve

$$f(x, y) = c$$

in the xy-plane, as shown in Fig. 19.2. Such a curve is called a *level curve*; it lies in the domain of the function, and on it $z = f(x, y)$ has the constant value c.

A collection of level curves is called a *contour map*; it can give a good idea of the shape of the graph, and is the next best thing to a three-dimensional sketch. For instance, the graph of $z = xy$ is difficult—though not impossible—to draw. However, a reasonably clear idea of the shape of this graph is given by the contour map shown in Fig. 19.3, which is easy to draw. Each level curve $xy = c$ is a hyperbola in the first and third quadrants if $c > 0$, a hyperbola in the second and fourth quadrants if $c < 0$, and the two axes taken together if $c = 0$. We ascend the surface as we leave the origin going into the first and third quadrants,

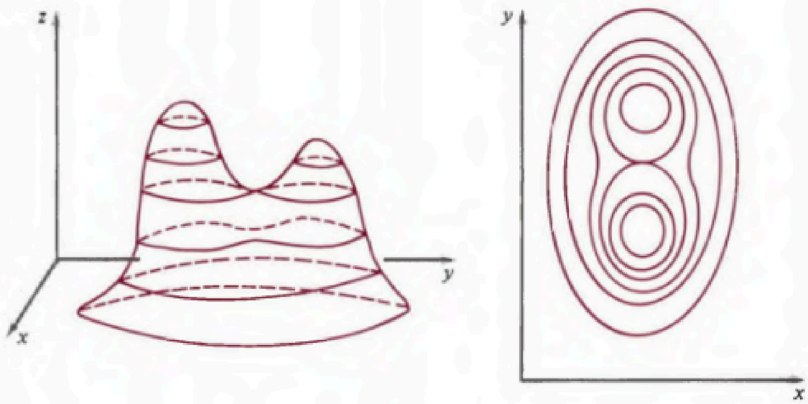


Figure 19.2 Level curves.

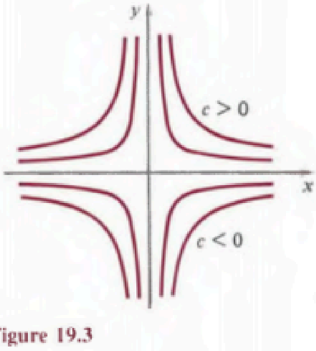


Figure 19.3

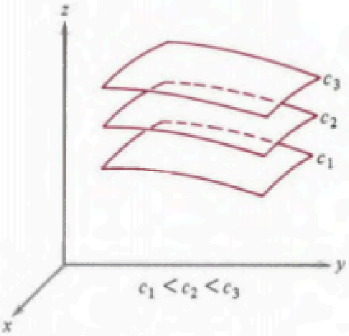


Figure 19.4 Level surfaces.

and descend it as we leave going into the second and fourth quadrants, and in this way we see that the origin is the saddle point of a saddle surface. Students should try to use this figure to visualize the shape of the surface as it appears in three-dimensional space, looking down on it from above.

LEVEL SURFACES

Drawing graphs for functions of two variables is often difficult, but drawing graphs for functions of three variables is always impossible. We would need a visible space of four dimensions to contain such a graph, and no such space is available.

However, the concept of level curves suggests a way to visualize the behavior of a function $w = f(x, y, z)$ of three variables: examine its *level surfaces.* These are the surfaces

$$f(x, y, z) = c \tag{3}$$

for various values of the constant c. Of course, level surfaces can be hard to draw, but a knowledge of what they are can help us form a useful intuitive idea of the nature of the function. In Fig. 19.4 we present a schematic view of three adjacent level surfaces of the form (3) for three values of the constant c, where $c_1 < c_2 < c_3$. As a point $P = (x, y, z)$ moves along the lowest surface, the value of $w = f(x, y, z)$ is constantly equal to c_1; but as this point hops to the next surface above it, the value of the function increases to c_2; and so on.

We consider two simple examples. In the case of the function $w = x + 2y + 3z$, the level surfaces are easily seen to be the planes

$$x + 2y + 3z = c$$

with normal vector $\mathbf{N} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$; and for $w = \sqrt{x^2 + y^2 + z^2}$, the level surfaces are the concentric spheres

$$x^2 + y^2 + z^2 = c^2.$$

In applications, if the function $w = f(x, y, z)$ represents the temperature at the point $P = (x, y, z)$, then the level surfaces are called *isothermal surfaces*; if it represents potential, they are called *equipotential surfaces.*

HIGHER DIMENSIONS

Level surfaces have a certain limited value, but in a sense they avoid the real question: How do we go about trying to obtain an intuitive understanding of the behavior of functions of three or more variables?

Briefly, what we do is work by analogy with the one- and two-variable cases. For example, there is nothing to prevent us from considering the set of all quadruples of numbers such as $(2, -3, 1, 4)$ as forming a perfectly legitimate four-dimensional space, with an origin $(0, 0, 0, 0)$, four coordinate axes, and a satisfactory concept of the distance from an arbitrary point (x, y, z, w) to the origin,

$$d = \sqrt{x^2 + y^2 + z^2 + w^2}.$$

We can now consider the graph of a function

$$w = f(x, y, z)$$

as forming a three-dimensional "surface" in this four-dimensional space, with the domain D of the function lying in the three-dimensional "coordinate plane" consisting of all points of the form $(x, y, z, 0)$.

In a similar way, if n is any positive integer we can think of the graph of a function of n variables as forming an n-dimensional "surface" in $(n + 1)$-dimensional space. It is true that for $n \geq 3$ we can no longer draw pictures, but we can still bolster our intuition by using geometric language, and we can still think geometrically, but in a looser way. However, as we move further away from the kind of mathematics that we can study and understand by drawing pictures, it is necessary to give more attention to the algebraic and analytic aspects of what we are doing, in order to avoid being misled by words and analogies. Nevertheless, the words, analogies, and geometric intuition remain indispensable, for they suggest worthwhile things to think about and prevent us from feeling totally lost among abstractions.

1. 标题与内容概述

《等高线与等值面：多变量函数的几何表示》

本次课件主要介绍了**等高线（level curves）**与**等值面（level surfaces）**这两个核心概念，分别用于二维和三维（或更高维）多变量函数的几何可视化。我们将从定义、典型示例以及在更高维情况的推广来逐步深入理解。

---

2. 详细内容解析

以下按照课件的逻辑顺序，对主要知识点作系统讲解。

2.1 等高线（Level Curves）

1. **引入背景：**
   当我们有函数 $z = f(x, y)$ 时，直接绘制它在三维空间里的图像往往比较困难（尤其是函数形状非常复杂的时候）。在制图学和地形图绘制中，人们常利用"等高线"来表示固定海拔（或高度）的一条曲线。类似地，我们在平面上画出满足

   $$f(x, y) = c$$

   的点所构成的曲线，就称之为**等高线（level curve）**。
2. **几何含义：**
   - 等高线位于函数定义域所在的 (x, y) 平面中（也可以是其子区域），而在这条曲线上， z 值恒等于某个常数 c 。
   - 若将不同的常数 c 取值所对应的等高线一起画在同一张图上，得到的便是**等值线图（contour map）**。它从二维图（ (x, y) 平面）帮助我们推测三维曲面的起伏和形状。
3. **典型例子：**
   - **例子 1：** $z = xy$ 。
     - 每条等高线即满足 $xy = c$ 。
       - 当 $c > 0$ 时，曲线是第一、三象限的双曲线。
       - 当 $c < 0$ 时，曲线是第二、四象限的双曲线。
       - 当 $c = 0$ 时，对应 $xy = 0$ ，也就是 $x = 0$ 或 $y = 0$ ，即坐标轴。
     - 由此，我们可以看出这个曲面在原点处表现为一个马鞍点（saddle point）；经过原点，函数符号随象限变化而改变，这也跟马鞍面的几何特征相一致。
   - **例子 2：** 书中右下方示意图（Fig. 19.2 和 Fig. 19.3）展示出不同函数可能生成的一组同心椭圆或双曲线、闭曲线等，可以帮助我们从上往下俯视去理解曲面。

2.2 等值面（Level Surfaces）

1. **概念与动机：**
   - 在三变量函数 $w = f(x, y, z)$ 的场景下，我们若想将四维情况（因为 (x, y, z) 再加一个输出量 w ）可视化，是非常困难的。
   - 但我们可以通过**等值面（level surfaces）**来帮助理解：令

     $$f(x, y, z) = c$$

     这些满足输出量 $w = c$ 的点集形成一个三维空间中的"面"（曲面）——这是等值面。
2. **可视化示意：**
   - 书中（Fig. 19.4）给出示意图：当我们选取不同常数 $c_1 < c_2 < c_3$ 时，对应三张水平（或倾斜）分布的"等值面"。
   - 随着 (x, y, z) 的移动，若你在最低等值面上， w 的值就是 c_1 ；一旦"跳"到更高的等值面上， w 就变为 c_2 ；再往上则变为 c_3 ；如此层层叠加。
3. **简单的例子：**
   - **例子 1：** $w = x + 2y + 3z$ 。
     - 等值面： $x + 2y + 3z = c.$
       这就是一族平行平面，法向量为 $\mathbf{N} = \langle 1,\, 2,\, 3\rangle$ 。不同的 c 值对应沿法向量方向平移得到不同的平面。
   - **例子 2：** $w = \sqrt{x^2 + y^2 + z^2}$ 。
     - 等值面： $\sqrt{x^2 + y^2 + z^2} = c \quad \Longrightarrow \quad x^2 + y^2 + z^2 = c^2.$
       这是一族以原点为中心、半径不等的同心球面，当 c 改变时球面半径相应变化。
4. **在实际应用中的命名：**
   - 如果 $f(x, y, z)$ 表示温度，则其等值面称为**等温面（isothermal surfaces）**。
   - 如果 $f(x, y, z)$ 表示势能，则其等值面称为**等势面（equipotential surfaces）**。

2.3 更高维的情况

1. **挑战：**
   - 当函数维度高于 3 时，我们很难再使用直观的图像去画出其"图面"（因为那至少要在 4 维或更高维空间中）。
   - 然而，**等值思想**仍然可以推广： $f(x_1, x_2, \ldots, x_n) = c$ 会是一个 $(n-1)$ 维的"超曲面"，嵌在 $\mathbb{R}^n$ 中。
2. **几何直觉与抽象思维：**
   - 当 $n \geq 3$ 时，我们只能依靠类比、代数和解析的方法来理解，而不再能简单地"画图"可视化。
   - 尽管如此，使用几何术语和类比仍然是重要的**直觉工具**，可以帮我们避免在抽象中迷失。但同时必须保持警惕，以免被过于"形象化"的描述所误导。
3. **总结观点：**
   - 等值曲线、等值面以及更高维的等值超曲面，都提供了一个在视图空间不足以绘制完整图形时，仍能分析多变量函数行为的方式。
   - 在进一步的学习中，我们会结合偏导数、梯度等概念来更深入地研究这些超曲面的局部性质。

---

3. 本次内容小结

- **等高线（level curves）：**
  在 (x, y) 平面上画出满足 $f(x, y) = c$ 的点所成的曲线，用于二维函数 $z = f(x, y)$ 的可视化；成组的等高线图（contour map）能在二维图中展示三维曲面的重要特征。
- **等值面（level surfaces）：**
  对三变量函数 $w = f(x, y, z)$ 而言，在三维空间里画出满足 $f(x, y, z) = c$ 的表面。通过不同的 c 值，能形成一系列相互并列或相互嵌套的曲面，从而帮助我们理解四维函数的分布。
- **推广到更高维：**
  依旧可以使用类似的"等值超曲面"概念，只是在维数较高时，无法直接通过"画图"去可视化，需要更抽象的代数和几何思维。

---

4. 报错或补充说明

从目前截图看，所有内容已得到完整讲解，无明显缺失或难以辨认之处。若有其他不清晰之处，请告知；否则本次笔记即为完整。

19.2

19.2 PARTIAL DERIVATIVES

Suppose that $y = f(x)$ is a function of only one variable. We know that its derivative, defined by

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x},$$

can be interpreted as the rate of change of y with respect to x. In the case of a function $z = f(x, y)$ of two variables, we shall need similar mathematical machinery for working with the rate at which z changes as both x and y vary. The key idea is to allow only one variable to change at a time, while holding the other fixed. For functions of more than two variables, we vary one of them while holding *all* the others fixed. Specifically, we differentiate with respect to only one variable at a time, regarding all the others as constants, and this gives us one derivative corresponding to each of the independent variables. These individual derivatives are the constituents from which we build the more complicated machinery that will be needed later.

To return to our function $z = f(x, y)$ of two variables, we first hold y fixed and let x vary. The rate of change of z with respect to x is denoted by $\partial z/\partial x$ and defined by

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}.$$

This limit (if it exists) is called the *partial derivative of z with respect to x*, and is read "partial z, partial x." The most commonly used notations for this partial derivative are

$$\frac{\partial z}{\partial x}, \qquad z_x, \qquad \frac{\partial f}{\partial x}, \qquad f_x, \qquad f_x(x, y),$$

and we shall use all of these from time to time in order to help students become accustomed to them. The symbol ∂ in the notation $\partial z/\partial x$ is called the "round-back d" or "curly d"; it is used to emphasize that there are other independent variables present during the process of differentiating with respect to x.

Similarly, if x is held fixed and y is allowed to vary, then the *partial derivative of z with respect to y* is defined by

$$\frac{\partial z}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y},$$

and the standard notations in this case are

$$\frac{\partial z}{\partial y}, \qquad z_y, \qquad \frac{\partial f}{\partial y}, \qquad f_y, \qquad f_y(x, y).$$

The actual calculation of partial derivatives for most functions is very easy: Treat every independent variable except the one we are interested in as if it were a constant, and apply the familiar rules.

Example 1 Calculate the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ of the function $f(x, y) = x^3 - 3x^2y^3 + y^2$.

Solution To find the partial of f with respect to x, we think of y as a constant and differentiate in the usual way,

$$\frac{\partial f}{\partial x} = 3x^2 - 6xy^3.$$

When we regard x as a constant and differentiate with respect to y, we obtain

$$\frac{\partial f}{\partial y} = -9x^2y^2 + 2y.$$

The notations $f_x(x, y)$ and $f_y(x, y)$ are useful for indicating the values of partial derivatives at specific points.

Example 2 (a) If $f(x, y) = xy^2 + x^3$, then

$$f_x(x, y) = y^2 + 3x^2, \qquad f_y(x, y) = 2xy,$$
$$f_x(2, 1) = 13, \qquad f_y(2, 1) = 4.$$

In the other notation, the numerical values given here by the simple and convenient symbols $f_x(2, 1)$ and $f_y(2, 1)$ would have to be written more clumsily as

$$\left(\frac{\partial f}{\partial x}\right)_{(2,1)} \quad \text{and} \quad \left(\frac{\partial f}{\partial y}\right)_{(2,1)}.$$

(b) If $g(x, y) = xe^{xy^2}$, then

$$g_x(x, y) = xy^2e^{xy^2} + e^{xy^2}, \qquad g_y(x, y) = 2x^2ye^{xy^2}.$$

(c) If $h(x, y) = \sin x^2 \cos 3y$, then

$$h_x(x, y) = 2x \cos x^2 \cos 3y, \qquad h_y(x, y) = -3 \sin x^2 \sin 3y.$$

These examples illustrate the fact that the partial derivatives of a function of x and y are themselves functions of x and y.

These ideas and notations apply just as easily to functions of any number of variables.

Example 3 If $w = f(x, y, z, u, v) = xy^2 + 2x^3 + xyz + zu + \tan uv$, then

$$\frac{\partial w}{\partial x} = y^2 + 6x^2 + yz, \qquad \frac{\partial w}{\partial y} = 2xy + xz, \qquad \frac{\partial w}{\partial z} = xy + u,$$
$$\frac{\partial w}{\partial u} = z + v \sec^2 uv, \qquad \frac{\partial w}{\partial v} = u \sec^2 uv.$$

In the one-variable case, we know that the derivative dy/dx can legitimately be thought of as a fraction, the quotient of the differentials dy and dx. The notation $\partial z/\partial x$ for the partial derivative $f_x(x, y)$ suggests that something similar might be done with ∂z and ∂x. However, it is not possible to treat partial derivatives as fractions. We give an example to emphasize this point.

Example 4 The *ideal gas law* states that for a given quantity of gas, the pressure p, volume V, and absolute temperature T are connected by the equation $pV = nRT$, where n is the number of moles of gas in the sample and R is a constant. Show that

$$\frac{\partial p}{\partial V}\frac{\partial V}{\partial T}\frac{\partial T}{\partial p} = -1.$$

Solution Since

$$p = \frac{nRT}{V}, \qquad V = \frac{nRT}{p}, \qquad T = \frac{pV}{nR},$$

we have

$$\frac{\partial p}{\partial V} = -\frac{nRT}{V^2}, \qquad \frac{\partial V}{\partial T} = \frac{nR}{p}, \qquad \frac{\partial T}{\partial p} = \frac{V}{nR}.$$

It follows that

$$\frac{\partial p}{\partial V}\frac{\partial V}{\partial T}\frac{\partial T}{\partial p}=\left(-\frac{nRT}{V^2}\right)\frac{nR}{p}\frac{V}{nR}=-\frac{nRT}{pV}=-1.$$

The fact that this result is -1 instead of $+1$ shows that we cannot treat the partial derivatives on the left as fractions.

1. 标题与内容概述

《多变量函数的偏导数：概念、记法与实例》

在单变量微积分中，我们用导数 $\frac{dy}{dx}$ 来表示函数 $y=f(x)$ 随着自变量 x 的变化率。而当函数依赖多个自变量时（例如 $z=f(x,y)$ ），我们希望依旧能刻画每个自变量对函数值的影响，这就引出了**偏导数（partial derivatives）**的概念。本次内容从偏导数的极限定义及符号记法出发，结合多个示例来阐释如何计算和理解偏导数，并通过一个“不能将偏导数简单视为分数”的实例（理想气体定律）来说明其中的细微差别。

---

2. 详细内容解析

下面按照课件顺序，逐点分析偏导数的定义、记法及计算方法，并结合示例做具体说明。

2.1 偏导数的定义与记法

1. **从单变量到多变量的延伸**
   - 在单变量情形下，导数的定义是
   $$\frac{dy}{dx} = \lim_{\Delta x\to 0}\frac{f(x+\Delta x) - f(x)}{\Delta x}.$$
   它表示函数 $y=f(x)$ 对于 x 的变化率。
   - 当函数依赖两个变量 x,y 时，为了探究 $z=f(x,y)$ 对“某个”变量的变化率，我们**在计算对 x 的偏导时保持 y 不变**，对 y 的偏导时保持 x 不变。
2. **定义：**
   - **偏导数** $\frac{\partial z}{\partial x}$ ：固定 y ，只让 x 变化，令
   $$\frac{\partial z}{\partial x} = \lim_{\Delta x\to 0}\frac{f(x+\Delta x,\ y) - f(x,\ y)}{\Delta x}.$$
   只要这个极限存在，即称之为“ z 对 x 的偏导数”。
   - **偏导数** $\frac{\partial z}{\partial y}$ ：固定 x ，只让 y 变化，定义同理。
   - 常见的记法有：
   $$\begin{aligned}&\frac{\partial z}{\partial x},\quad z_x,\quad \frac{\partial f}{\partial x},\quad f_x,\quad f_x(x,y),\\ &\frac{\partial z}{\partial y},\quad z_y,\quad \frac{\partial f}{\partial y},\quad f_y,\quad f_y(x,y).\end{aligned}$$
   其中“ ∂ ”通常读作“偏”或“曲线 d”（curly d），用来提醒我们：此时还有其他自变量存在。
3. **计算规则**
   - 计算偏导数时，**对正在求导的那个变量视为“自变量”**，而将**其他变量都当成常数**对待。
   - 这样就可以使用和单变量微积分中相同的微分规则（如幂函数求导、指数函数求导、三角函数求导等）。

2.2 计算示例

示例 1

- **题目：** 设

  $$f(x,y) = x^3 \;-\; 3x^2y^3 \;+\; y^2.$$

  计算 $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$。

- **解：**
  1. ** $\frac{\partial f}{\partial x}$ **：把 y 视为常数，

     $$\frac{\partial f}{\partial x} \;=\; 3x^2 \;-\; 6xy^3.$$

  2. ** $\frac{\partial f}{\partial y}$ **：把 x 视为常数，

     $$\frac{\partial f}{\partial y} \;=\; -9x^2y^2 \;+\; 2y.$$

示例 2

1. **(a)** 若

   $$f(x,y) \;=\; xy^2 + x^3,$$

   则

   $$f_x(x,y) = \frac{\partial}{\partial x}\bigl(xy^2 + x^3\bigr) = y^2 + 3x^2, \quad f_y(x,y) = \frac{\partial}{\partial y}\bigl(xy^2 + x^3\bigr) = 2xy.$$

   - 并且 $f_x(2,1) = 13$，$f_y(2,1) = 4$（把 $(x,y) = (2,1)$ 代入即可）。
2. **(b)** 若

   $$g(x,y) = x\,e^{xy^2},$$

   - 则

     $$g_x(x,y) = y^2\,e^{xy^2} + e^{xy^2} \cdot x \cdot 0 \;\;(\text{需用乘法和链式法则}),$$

     实际更精确写法：

     $$g_x(x,y) = \frac{\partial}{\partial x}\bigl(xe^{xy^2}\bigr) = e^{xy^2} + x \cdot (e^{xy^2} \cdot y^2) = e^{xy^2} + xy^2e^{xy^2} = e^{xy^2}(1 + xy^2).$$

     $$g_y(x,y) = \frac{\partial}{\partial y}\bigl(xe^{xy^2}\bigr) = x \cdot \frac{\partial}{\partial y}\bigl(e^{xy^2}\bigr) = x \cdot e^{xy^2} \cdot (2xy) = 2x^2y\,e^{xy^2}.$$

3. **(c)** 若

   $$h(x,y) = \sin\bigl(x^2\cos(3y)\bigr),$$

   # 则应用复合函数求导（链式法则）：

   $$h_x(x,y) = \cos\bigl(x^2\cos(3y)\bigr) \;\times\; \frac{\partial}{\partial x}\bigl(x^2\cos(3y)\bigr) = \cos\bigl(x^2\cos(3y)\bigr) \;\times\; 2x\cos(3y).$$

   - $h_y(x,y) = \cos\bigl(x^2\cos(3y)\bigr) \;\times\; \frac{\partial}{\partial y}\bigl(x^2\cos(3y)\bigr) = \cos\bigl(x^2\cos(3y)\bigr) \;\times\; \bigl(x^2 \cdot (-\sin(3y) \cdot 3)\bigr) = -3x^2\sin(3y)\;\cos\bigl(x^2\cos(3y)\bigr)$

这些例子说明：**偏导数本身还是函数**，它们依旧依赖 (x,y) 的数值。

2.3 多变量函数中的偏导数（更高维）

1. **示例 3：多变量** $w = f(x,y,z,u,v)$
   - 书中给出的例子：

$$w \;=\; f(x,y,z,u,v) \;=\; xy^2 + 2x^3 + xyz + zu + \tan(uv).$$

     - 分别对 x,y,z,u,v 求偏导数：

$$\frac{\partial w}{\partial x} = y^2 + 6x^2 + yz, \quad \frac{\partial w}{\partial y} = 2xy + xz, \quad \frac{\partial w}{\partial z} = xy + u, \quad \frac{\partial w}{\partial u} = z + v\,\sec^2(uv), \quad \frac{\partial w}{\partial v} = u\,\sec^2(uv).$$

     - 计算方法与二维情况相同：除要对的那个变量外，其他全部当作常数，然后使用熟悉的求导规则。
2. **偏导数≠可直接当作分数处理**
   - 在单变量时， $\frac{dy}{dx}$ 既可以视为极限定义，也常被“当作分数”在某些场合作形式操作。
   - 但在多变量里， $\frac{\partial z}{\partial x}$ 仅是一个符号，不能随意把 ∂z 和 ∂x 当可约分的“对象”处理。
   - 这在例 4（理想气体定律）里有典型的演示。

2.4 理想气体定律示例：不能简单把偏导数当分数

1. **理想气体定律：**

$$pV = nRT,$$

   其中 p 为压强， V 为体积， T 为绝对温度， n 为气体的摩尔数， R 为常数。
   - 从该方程可推出：

$$p = \frac{nRT}{V}, \quad V = \frac{nRT}{p}, \quad T = \frac{pV}{nR}.$$

2. **问题：**
   - 书中让我们验证

$$\frac{\partial p}{\partial V}\,\frac{\partial V}{\partial T}\,\frac{\partial T}{\partial p} \;=\; -1.$$

   - 直觉地，如果我们把这三项“看成分数”并做形式上的“ $\frac{\partial p}{\partial V} \times \frac{\partial V}{\partial T} \times \frac{\partial T}{\partial p} \equiv 1$ ”那就会以为结果是“ $+1$ ”。但真正算出来却是 -1 。这就表明了**偏导数不能像普通分数那样在多变量情况下直接相互抵消。**
3. **具体计算：**
   - 先求每个偏导：

$$p = \frac{nRT}{V} \implies \frac{\partial p}{\partial V} = -\frac{nRT}{V^2},$$
$$V = \frac{nRT}{p} \implies \frac{\partial V}{\partial T} = \frac{nR}{p},$$
$$T = \frac{pV}{nR} \implies \frac{\partial T}{\partial p} = \frac{V}{nR}.$$

   - 相乘：

$$\left(-\frac{nRT}{V^2}\right)\left(\frac{nR}{p}\right)\left(\frac{V}{nR}\right) = -\frac{nRT}{V^2}\cdot\frac{nR}{p}\cdot\frac{V}{nR} = -\frac{nRT}{V^2}\cdot\frac{V}{p}\cdot\frac{nR}{nR}.$$

     注意到 $\frac{nR}{nR} = 1$ ，还可用 $pV = nRT$ 替换，最终可得结果为 -1 。
   - 正是这个 -1 （而非 $+1$ ）证明了偏导数乘起来并不满足简单的“分数约分”原则。

---

3. 本次内容小结

- **偏导数的核心**：在多变量函数中，如果我们想知道函数对某个变量的变化率，就把其他变量都看作常数，对该变量进行单变量式求导；这就是偏导数。
- **记法与运算**：常见记法有 $\frac{\partial f}{\partial x}, f_x, z_x$ 等；运算时遵循单变量求导规则，但一定要明确哪些量是"常数"。
- **偏导数函数性**：偏导数本身仍然是函数，依赖于所在点的坐标。
- **高维推广**：任意个独立变量时，都可以逐个变量地固定其余，使用同样方法进行偏导计算。
- **偏导数并非普通分数**：在多变量场景下，$\frac{\partial z}{\partial x}$ 不能被简单视为 $\frac{\mathrm{d}z}{\mathrm{d}x}$ 去做分数的直观操作，理想气体定律例子很好地说明了这一点。

---

4. 报错或补充说明

- 已根据所给截图内容对偏导数的基本概念、计算示例以及理想气体定律中的特殊性质进行了完整讲解。若有任何截图遗失或无法辨认的部分，请告知；如无，则本次笔记即完整。

19.2.2

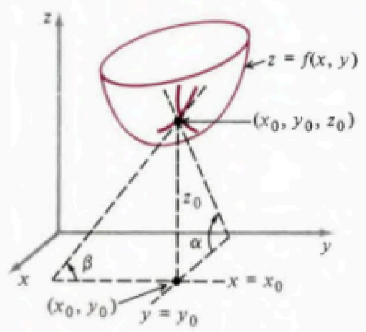


Figure 19.5

When we are working with a function $z = f(x, y)$ of only two variables, the partial derivatives have the following simple geometric interpretation. The graph of this function is a surface, as shown in Fig. 19.5. Let (x_0, y_0) be a given point in the xy-plane, with (x_0, y_0, z_0) the corresponding point on the surface. To hold y fixed at the value y_0 means to intersect the surface with the plane $y = y_0$, and the intersection is the curve

$$z = f(x, y_0)$$

in that plane. The number

$$\left(\frac{\partial z}{\partial x}\right)_{(x_0,y_0)} = f_x(x_0, y_0)$$

is the slope of the tangent line to this curve at $x = x_0$. Thus, in the figure we have

$$\tan \alpha = \left(\frac{\partial z}{\partial x}\right)_{(x_0,y_0)} = f_x(x_0, y_0).$$

Similarly, the intersection of the surface with the plane $x = x_0$ is the curve

$$z = f(x_0, y),$$

and the other partial derivative is the slope of the tangent to this curve at $y = y_0$,

$$\tan \beta = \left(\frac{\partial z}{\partial y}\right)_{(x_0,y_0)} = f_y(x_0, y_0).$$

No such interpretation is available when there are more than two independent variables.

We remarked that for a function $z = f(x, y)$ of two variables, the partial derivatives f_x and f_y are also functions of two variables, and may themselves have partial derivatives. As we might expect, these *second-order partial derivatives* are denoted by several symbols. If we start with the first derivatives

$$\frac{\partial f}{\partial x} = f_x \qquad \text{and} \qquad \frac{\partial f}{\partial y} = f_y,$$

then the derivatives with respect to x are

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} f_x = f_{xx}$$

and

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial}{\partial x} f_y = f_{yx};$$

and the derivatives with respect to y are

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\, \partial x} = \frac{\partial}{\partial y} f_x = f_{xy}$$

and

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y} f_y = f_{yy}.$$

This notation may seem a bit confusing at first, but it is actually quite reasonable. Observe that in f_{yx} we differentiate first with respect to the "inside" variable y, then with respect to the "outside" variable x. This is the natural order, since f_{yx} ought to mean $(f_y)_x$. Thus, in the symbols f_{yx} and f_{xy}, the subscript letters accumulate from left to right, because this is the order in which the differentiations are performed. For the same reason, in the symbols

$$\frac{\partial^2 f}{\partial x\,\partial y} \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y\,\partial x},$$

it is natural for the letters indicating the variable of differentiation to accumulate from right to left: first y, then x in the first of these; and first x, then y in the second.

The *pure* second partial derivatives,

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} \qquad \text{and} \qquad f_{yy} = \frac{\partial^2 f}{\partial y^2},$$

don't represent anything really new. Each is found by holding one variable constant and differentiating twice with respect to the other variable, and each gives the rate of change of the rate of change of f in the direction of one of the axes.

Example 5 If $f(x, y) = x^3 e^{5y} + y \sin 2x$, then

$$f_x = 3x^2 e^{5y} + 2y\cos 2x, \qquad f_y = 5x^3 e^{5y} + \sin 2x,$$

$$f_{xx} = 6xe^{5y} - 4y \sin 2x, \qquad f_{yy} = 25x^3 e^{5y}.$$

On the other hand, the *mixed* second partial derivatives,

$$f_{xy} = \frac{\partial^2 f}{\partial y\,\partial x} \qquad \text{and} \qquad f_{yx} = \frac{\partial^2 f}{\partial x\,\partial y},$$

represent new ideas. The mixed partial derivative f_{xy} gives the rate of change in the y-direction of the rate of change of f in the x-direction, and f_{yx} gives the rate of change in the x-direction of the rate of change of f in the y-direction. It is not at all clear how these two mixed partials are related to each other, if indeed they are related at all.

Example 5 (continued) For the function being considered, $f(x, y) = x^3 e^{5y} + y \sin 2x$, we easily see that

$$f_x = 3x^2 e^{5y} + 2y\cos 2x, \qquad f_y = 5x^3 e^{5y} + \sin 2x,$$

$$f_{xy} = 15x^2 e^{5y} + 2\cos 2x, \qquad f_{yx} = 15x^2 e^{5y} + 2\cos 2x.$$

For the particular function considered in this example, we obviously have

$$f_{xy} = f_{yx}, \tag{1}$$

or equivalently,

$$\frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial^2 f}{\partial x\,\partial y},$$

so the order of differentiation seems to be unimportant—at least in this case. But this is not an accident, and (1) is true for almost all functions that normally arise in applications. More precisely, *if both f_{xy} and f_{yx} exist for all points near* (x_0, y_0) *and are continuous at* (x_0, y_0), *then*

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$

A proof of this statement is given in Appendix A.17.

Partial derivatives of order greater than two, as well as higher-order derivatives of functions of more than two variables, are defined in the obvious way. For example, if $w = f(x, y, z)$, then

$$\frac{\partial^3 f}{\partial x\,\partial y\,\partial z} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y\,\partial z}\right) = (f_{zy})_x = f_{zyx},$$

$$\frac{\partial^4 f}{\partial z\,\partial y\,\partial x^2} = \frac{\partial}{\partial z}\left(\frac{\partial^3 f}{\partial y\,\partial x^2}\right) = (f_{xxy})_z = f_{xxyz},$$

etc. In general, with suitable continuity, it is immaterial in what order a sequence of partial differentiations is carried out, for by (1) we can reverse the order of any two successive differentiations. For example, $f_{xxyz} = f_{xyxz} = f_{xyzx} = f_{yxzx} = f_{yzxx}$.

1. 标题与内容概述

《多变量函数的几何解释与高阶偏导数——从斜率到混合偏导及其相等定理》

本次内容围绕函数 $z = f(x, y)$ 的几何解释和**二阶及更高阶偏导数**展开。首先，我们会看到，偏导数在几何上可以视为沿固定平面截出的曲线的斜率。随后介绍二阶偏导数（包括同类偏导 f_{xx}, f_{yy} 以及混合偏导 f_{xy}, f_{yx} ）的定义和记法，并通过示例说明在常见条件下 $f_{xy} = f_{yx}$ （即**混合偏导互换定理**）。最后也提到更高阶和更多变量的偏导数在定义和运算上与二元函数类似。

---

2. 详细内容解析

下面将按课件的顺序逐点展开，包括偏导数的几何理解、二阶偏导数的记法和计算示例等。

2.1 一阶偏导数的几何解释

1. **图像与平面截取：**
   - 对于一个二元函数 $z = f(x, y)$ ，我们可以在三维空间中画出它的曲面；其中 (x_0, y_0, z_0) 是曲面上的某个点，且 $z_0 = f(x_0, y_0)$ 。
   - 若我们固定 $y = y_0$ 不变，相当于用一个垂直于 y-轴（平行于 x - z 平面）的平面来**截**这条三维曲面，这个交线在三维图形上就是一条曲线：

     $$z = f(x,\, y_0).$$

     在该曲线中， $\frac{\partial z}{\partial x}\Big|_{(x_0,y_0)}$ 就是这条曲线在 $x = x_0$ 处的**切线斜率。**
   - 同样，若我们固定 $x = x_0$ ，得到曲线 $z = f(x_0, y)$ 。此时 $\frac{\partial z}{\partial y}\Big|_{(x_0,y_0)}$ 是这条曲线在 $y = y_0$ 处的切线斜率。
2. **斜率与切线：**
   书中提到，如果把这两个偏导数用几何的方式去看：
   - $\tan\alpha = \frac{\partial z}{\partial x}(x_0, y_0)$ 表示沿 x 方向截曲线处的切线斜率；
   - $\tan\beta = \frac{\partial z}{\partial y}(x_0, y_0)$ 表示沿 y 方向截曲线处的切线斜率。

3. **当超过两个自变量时（如三变量 $w=f(x,y,z)$ ），就难以有同样直观的“平面截曲面再看斜率”**，因为需要的维度更多。不过在数学原理上，这种“保持其余变量不变，考察对某一变量的变化率”的思路依然适用。

2.2 二阶偏导数

1. **定义和记法：**
   对 $f(x,y)$ 的一阶偏导 f_x, f_y ，还可以再分别对 x 或 y 做偏导，得到二阶偏导：
   - **纯二阶偏导数：**
   $$f_{xx}=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)=\frac{\partial^2 f}{\partial x^2},\quad f_{yy}=\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)=\frac{\partial^2 f}{\partial y^2}.$$
   这相当于：先把 f 对 x 求偏导一次，再对结果再求一次偏导（依旧对 x ）。 f_{yy} 类似。
   - **混合偏导数：**
   $$f_{xy}=\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)=\frac{\partial^2 f}{\partial y\,\partial x},$$
   $$f_{yx}=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)=\frac{\partial^2 f}{\partial x\,\partial y}.$$
   先对 x 求偏导，再对 y 求偏导；或先对 y ，再对 x 。
2. **符号说明：**
   - 例如， f_{xy} 表示“对 x 先导，再对 y ”，也可写作 $\left(f_x\right)_y$ 。
   - f_{yx} 则是“对 y 先导，再对 x ”。
   - 符号累加顺序： f_{xy} 左边的那个下标字母表示最外层偏导；右边的那个字母表示最内层偏导。
3. **物理或几何解释：**
   - f_{xx} 和 f_{yy} 称为“**纯二阶偏导**”，可以理解为：
   先固定 y 看 x 方向的变化率，再对得到的那个变化率本身再随 x 的变化率；或反之类似地固定 x 。
   - f_{xy} 或 f_{yx} 称为“**混合二阶偏导**”，代表先在一个方向上看变化率，再去看另一个方向对该变化率的影响。

2.3 示例讲解

示例 5
书中给出函数
$$f(x,y)=x^3e^{5y}\ +\ y\sin(2x).$$

1. # 先算一阶偏导
   - $$f_x=\frac{\partial}{\partial x}\left(x^3e^{5y}+y\sin(2x)\right)=3x^2e^{5y}+y\cdot(2\cos(2x))=3x^2e^{5y}+2y\cos(2x),$$
   $$f_y=\frac{\partial}{\partial y}\left(x^3e^{5y}+y\sin(2x)\right)=x^3\cdot(5e^{5y})+\sin(2x)=5x^3e^{5y}+\sin(2x).$$
2. **再算二阶偏导**
   - 纯二阶：
   $$f_{xx}=\frac{\partial}{\partial x}\left(f_x\right)=\frac{\partial}{\partial x}\left(3x^2e^{5y}+2y\cos(2x)\right)=6xe^{5y}-4y\sin(2x),$$
   $$f_{yy}=\frac{\partial}{\partial y}\left(f_y\right)=\frac{\partial}{\partial y}\left(5x^3e^{5y}+\sin(2x)\right)=25x^3e^{5y}.$$
   - 混合偏导：
   $$f_{xy}=\frac{\partial}{\partial y}\left(f_x\right)=\frac{\partial}{\partial y}\left(3x^2e^{5y}+2y\cos(2x)\right)=15x^2e^{5y}+2\cos(2x),$$

$$f_{yx} = \frac{\partial}{\partial x}\Big(f_y\Big) = \frac{\partial}{\partial x}\big(5x^3e^{5y} + \sin(2x)\big) = 15x^2e^{5y} + 2\cos(2x).$$

- 结果可见： $f_{xy} = f_{yx} = 15x^2e^{5y} + 2\cos(2x)$ 。

3. **混合偏导相等现象：**
   - 本例中 $\frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial^2 f}{\partial x\,\partial y}$ ，书中指出这并非偶然。对“足够光滑”的函数而言（即某些连续性条件满足）， **Clairaut 定理**（又称 Schwarz 定理）告诉我们

$$f_{xy}(x, y) = f_{yx}(x, y)$$

   在函数可微且混合偏导连续的区域内必然成立。

2.4 混合偏导互换定理的推广

1. **一般情形：**
   - 如果函数 f 具有三变量、四变量乃至更多个变量，那么同样可以定义各种高阶偏导，例如 f_{xyz} 、 f_{xxy} 等等。
   - 如果函数及其各阶偏导在相关区域内保持连续性，那么我们可以在**任何顺序**下对自变量进行偏导，最终结果相同。例如：

$$\frac{\partial^3 f}{\partial x\,\partial y\,\partial z} = \frac{\partial^3 f}{\partial y\,\partial x\,\partial z} = \dots$$

2. **适用条件：**
   - 只要符合足够的可微与连续性条件（“适合的连续性”，通常指某些光滑函数等级的要求），则混合偏导的顺序可以互换。

---

3. 本次内容小结

- **偏导数的几何解释：**
  对于 $z = f(x, y)$ ，固定一个变量后得到的截曲线，其切线斜率便是相应的偏导数值。
- **二阶偏导：**
  - **纯二阶偏导** f_{xx}, f_{yy} 表示同一个方向做两次偏导；
  - **混合偏导** f_{xy}, f_{yx} 表示先对一个方向，再对另一个方向。
- **Clairaut（Schwarz）定理：**
  在“正常”条件（函数及二阶偏导足够光滑）下，混合偏导顺序可互换，即 $f_{xy} = f_{yx}$ 。对于更高阶、更高维度，也可以类似地说明只要满足相应光滑性条件，多重偏导的顺序可以任意互换。
- **实用意义：**
  - 一方面，二阶偏导往往与曲面的弯曲性质、极值问题等密切相关；
  - 另一方面，知道“混合偏导可以互换”在实际运算中会让我们更灵活方便。

---

4. 报错或补充说明

- 从截图呈现的信息来看，已对几何解释、二阶偏导记法、混合偏导及其相等性做了全面的笔记与解析。如有任何遗漏或无法辨认之处请告知；若无，则以上内容即为本部分的完整讲解。

19.3

19.3 THE TANGENT PLANE TO A SURFACE

The concept of a tangent plane to a surface corresponds to the concept of a tangent line to a curve. Geometrically, the tangent plane to a surface at a point is the plane that "best approximates" the surface near the point. It will be necessary for us to think rather carefully about what this means, because—as we shall see in Sections 19.5 and 19.6—weighty practical consequences depend on it.

Consider a surface $z = f(x, y)$, as shown in Fig. 19.6. As we pointed out in Section 19.2, the plane $y = y_0$ intersects this surface in a curve C_1 whose equation is

$$z = f(x, y_0),$$

and the plane $x = x_0$ intersects it in a curve C_2 whose equation is

$$z = f(x_0, y);$$

and the slopes of the tangent lines to these curves at the point $P_0 = (x_0, y_0, z_0)$ are the partial derivatives

$$f_x(x_0, y_0) \qquad \text{and} \qquad f_y(x_0, y_0). \tag{1}$$

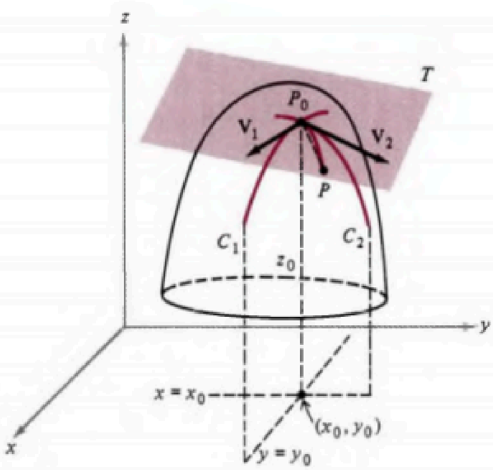


Figure 19.6 The tangent plane.

These two tangent lines determine a plane, and, as Fig. 19.6 suggests, if the surface is sufficiently smooth near P_0, then this plane will be tangent to the surface at P_0.

It is important to be quite clear about what we mean by a tangent plane, so we give a definition. In this context, where P_0 is a point on a surface $z = f(x, y)$, let T be a plane through P_0 and let P be any other point on the surface. If, as P approaches P_0 along the surface, the angle between the segment P_0P and the plane T approaches zero, then T is called the *tangent plane* to the surface at P_0.

It is easy to see that a surface need not have a tangent plane at a point P_0. A very simple example is provided by the half-cone $z = \sqrt{x^2 + y^2}$ shown in Fig. 19.7. It is clear that no plane is tangent to this surface at the origin. In this case the curves C_1 and C_2 have no tangent lines at the origin, and the partial derivatives (1) do not exist there. However, even when the curves C_1 and C_2 are smooth enough to have tangent lines at P_0, the surface may still not have a tangent plane at P_0, because of nonsmooth behavior near P_0 in the regions between C_1 and C_2. In Section 19.4 we discuss a vital lemma to the effect that this cannot happen if the partial derivatives $f_x(x, y)$ and $f_y(x, y)$ exist at all points in some neighborhood of (x_0, y_0) and are continuous at (x_0, y_0) itself.

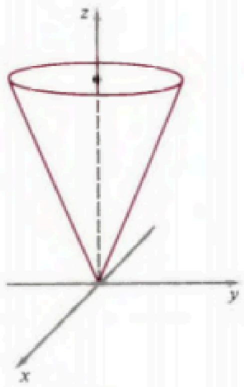


Figure 19.7

Meanwhile, we assume that the tangent plane exists at P_0, and we develop a method of finding its equation. Since the point $P_0 = (x_0, y_0, z_0)$ lies on this tangent plane, we know that the equation has the form

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0, \tag{2}$$

where $N = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ is any normal vector. It remains to find N, and to do this we use the cross product of two vectors $\mathbf{V}_1$ and $\mathbf{V}_2$ that are tangent to the curves C_1 and C_2 at P_0 (see Fig. 19.6). To find $\mathbf{V}_1$, we use the fact that along the tangent line to C_1, an increase of 1 unit in x produces a change $f_x(x_0, y_0)$ in z, while y does not change at all. Thus, the vector

$$\mathbf{V}_1 = \mathbf{i} + 0 \cdot \mathbf{j} + f_x(x_0, y_0)\mathbf{k}$$

is tangent to C_1 at P_0. Similarly, the vector

$$\mathbf{V}_2 = 0 \cdot \mathbf{i} + \mathbf{j} + f_y(x_0, y_0)\mathbf{k}$$

is tangent to C_2 at P_0. Since $\mathbf{V}_1$ and $\mathbf{V}_2$ lie in the tangent plane, we are now able to obtain our normal vector $\mathbf{N}$ by calculating

$$\mathbf{N} = \mathbf{V}_2 \times \mathbf{V}_1 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 1 & f_y(x_0, y_0) \\ 1 & 0 & f_x(x_0, y_0) \end{vmatrix} = f_x(x_0, y_0)\mathbf{i} + f_y(x_0, y_0)\mathbf{j} - \mathbf{k}. \tag{3}$$

(The order of factors in this cross product is chosen only for convenience, to produce one minus sign in the result instead of two.) When the components of (3) are inserted in (2), we see that the desired equation is

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0,$$

or equivalently,

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \tag{4}$$

1. 标题与内容概述

《曲面的切平面：几何直观与方程推导》

本次内容主要讨论**切平面（tangent plane）**在多变量函数 $z = f(x, y)$ 中的几何含义与计算方法。我们会看到，若曲面在某点够“平滑”，则可在该点拟合一个“最佳逼近”的平面，即所谓切平面。具体地，通过截曲线的切线斜率（偏导数）以及向量叉乘等方法，就能推导出切平面的方程。本节还探讨了可能出现“无切平面”的情况。

---

2. 详细内容解析

2.1 从截曲线切线斜率到切平面

1. **截曲线与偏导数：**
   - 当我们固定 $y = y_0$ 时，曲面 $z = f(x, y)$ 和平面 $y = y_0$ 的交线为

     $$C_1: \quad z = f(x, y_0),$$

     这是在三维空间里的一条曲线。它在 (x_0, y_0, z_0) 处（其中 $z_0 = f(x_0, y_0)$）的切线斜率就是 $\frac{\partial f}{\partial x}(x_0, y_0)$。
   - 当我们固定 $x = x_0$ 时，得到另一条曲线

     $$C_2: \quad z = f(x_0, y),$$

     在同一点 (x_0, y_0, z_0) 的切线斜率就是 $\frac{\partial f}{\partial y}(x_0, y_0)$。
   - 这两条截曲线的切线就各自延伸出两个方向向量，若曲面足够光滑，这两条切线就能共同确定一个平面；我们称之为曲面在该点的**切平面。**
2. **切平面的几何意义：**
   - 切平面是经过曲面上一点 (x_0, y_0, z_0) 的平面，且它在该点**最好地逼近**曲面的局部形状，正如在单变量函数中切线是曲线的“最佳线性逼近”一样。
   - 如果沿着曲面的邻近区域中的任何方向移动，曲面的高度变化跟切平面“非常贴近”（在函数可微条件下可以做严格定义）。
3. **曲面可能无切平面：**
   - 书中图 19.7 给了一个例子 $z = \sqrt{x^2 + y^2}$ 在原点的情形：由于尖点或不光滑导致无明确切平面。
   - 即使在原点处有两个截曲线，但如果它们本身没有良好的偏导数（或曲面在该点不具备可微性），就可能不存在真正意义上的切平面。

2.2 切平面的方程推导

1. **一般形式：**
   在三维空间中，一个平面可以写成

   $$a(x - x_0) \ + \ b(y - y_0) \ + \ c(z - z_0) \ = \ 0,$$

   其中向量 $\mathbf{N} = \langle a, \, b, \, c \rangle$ 是该平面的法向量（normal vector），(x_0, y_0, z_0) 是平面上的一点。
2. **如何求法向量 $\mathbf{N}$：**
   - 方法思路：我们先找到**位于切平面中的两个方向向量**，再用它们的叉乘（cross product）得到平面的法向量。
   - 在图 19.6，书中把与截曲线 C_1 相切的向量记作 $\mathbf{V}_1$，与另一条截曲线 C_2 相切的向量记作 $\mathbf{V}_2$。它们都在同一点 (x_0, y_0, z_0) 处。
3. **构造截曲线的切向量 $\mathbf{V}_1, \mathbf{V}_2$**
   - 沿着 C_1：当 x 变化 1 个单位时，z 随之变化 $\frac{\partial f}{\partial x}(x_0, y_0)$，而 y 不变。所以

     $$\mathbf{V}_1 = \langle 1, \ 0, \ f_x(x_0, y_0) \rangle.$$

   - 沿着 C_2：当 y 变化 1 个单位时，z 随之变化 $\frac{\partial f}{\partial y}(x_0, y_0)$，而 x 不变。所以

     $$\mathbf{V}_2 = \langle 0, \ 1, \ f_y(x_0, y_0) \rangle.$$

4. **求法向量 $\mathbf{N}$ 的叉乘**

   $$\mathbf{N} \ = \ \mathbf{V}_2 \ \times \ \mathbf{V}_1.$$

   - 书中在 (3) 式示范了具体计算顺序：

     $$\mathbf{V}_2 = \langle 0, \ 1, \ f_y(x_0, y_0) \rangle, \quad \mathbf{V}_1 = \langle 1, \ 0, \ f_x(x_0, y_0) \rangle.$$

     用行列式形式

$$\mathbf{N}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 0 & 1 & f_y(x_0,y_0)\\ 1 & 0 & f_x(x_0,y_0)\end{vmatrix}=(1)(f_x(x_0,y_0))\,\mathbf{i}+(f_y(x_0,y_0))(1)\,\mathbf{j}-(\dots)\mathbf{k}.$$

详细展开后

$$\mathbf{N} \;=\; f_x(x_0,y_0)\,\mathbf{i} \;+\; f_y(x_0,y_0)\,\mathbf{j} \;-\; \mathbf{k}.$$

- 当然，有时也会出现符号次序上的正负号区别，但最终我们关心的是这个向量的方向即可（正负倍数也代表同一个法向量方向）。

5. **得到切平面方程：**
   - 若 $\mathbf{N}=\langle a,b,c\rangle$ 是法向量，则平面方程可写为

$$a(x-x_0) \;+\; b(y-y_0) \;+\; c(z-z_0)=0.$$

   - 在这里，$a=f_x(x_0,y_0),\, b=f_y(x_0,y_0),\, c=-1$ 。所以整合得：

$$f_x(x_0,y_0)\,(x-x_0)+f_y(x_0,y_0)\,(y-y_0)-(z-z_0)=0,$$

   等价地写成：

$$z-z_0=f_x(x_0,y_0)\,(x-x_0)+f_y(x_0,y_0)\,(y-y_0).$$

   - 这正是我们常见的**切平面线性逼近**公式，和在单变量微积分中"切线"那句 $y-y_0=f'(x_0)(x-x_0)$ 十分相似。

2.3 方程形式总结与可微性的联系

1. **线性逼近形式：**
   - 切平面的方程

$$z-z_0 \;=\; f_x(x_0,y_0)\,(x-x_0) \;+\; f_y(x_0,y_0)\,(y-y_0)$$

   在多变量微分中（后续章节），也可写为

$$L(x-x_0,\; y-y_0) \;=\; \frac{\partial f}{\partial x}(x_0,y_0)\,(x-x_0) \;+\; \frac{\partial f}{\partial y}(x_0,y_0)\,(y-y_0),$$

   表示在 (x_0,y_0) 附近对 f 的一阶线性逼近。

2. **无切平面情形：**
   - 如果函数在某点的偏导数不存在，或者偏导数不连续导致无可微性，就可能像文中图 19.7 那样，在该点没有可定义的"单个平面"做线性逼近。
   - 在正常条件（ f 连续可微）下，切平面必然存在且其方程如上。

3. **物理/几何意义**
   - 切平面在该点附近能很好地"近似"曲面，实际应用中常用来研究局部性质，如在多变量极值判定、线性化近似等方面都极具重要性。

---

3. 本次内容小结

- **几何图像**：对于二元函数 $z=f(x,y)$ ，在点 (x_0,y_0,z_0) 处的截曲线 C_1 与 C_2 各有一条切线，其斜率分别是偏导数 f_x 和 f_y ；这两条切线通常能确定一个平面——即切平面。
- **切平面方程推导**：
  1. 找到平面内的两个方向向量（分别为 $\langle 1,0,f_x\rangle$ 和 $\langle 0,1,f_y\rangle$ ）。
  2. 用叉乘求法向量 $\mathbf{N}$ 。
  3. 代入平面一般方程 $a(x-x_0)+b(y-y_0)+c(z-z_0)=0$ 。
  4. 最终得出：

$$z-z_0=f_x(x_0,y_0)(x-x_0)+f_y(x_0,y_0)(y-y_0).$$

- **注意可微性与"无切平面"现象**：若函数在该点不光滑，可能并不存在惟一切平面。
- **与线性化的关系**：切平面方程在多变量微分理论里正是函数的线性化（近似）表达。

---

4. 报错或补充说明

- 截至目前，根据所给截图内容，已完整解析"切平面"的几何意义、方程推导及注意事项。若有遗漏或无法辨认之处，请告知；如无，则本次笔记已涵盖所有核心要点。

19.3.2

Example 1 Find the tangent plane to the surface

$$z = f(x, y) = 2xy^3 - 5x^2$$

at the point (3, 2, 3).

Solution The first step should be to check that this point actually lies on the given surface, and we assume that this has been done. Here we have $f_x = 2y^3 - 10x$ and $f_y = 6xy^2$, so $f_x(3, 2) = -14$ and $f_y(3, 2) = 72$. The equation of the tangent plane is therefore

$$z - 3 = -14(x - 3) + 72(y - 2).$$

Tangent planes to surfaces where z is not explicitly given as a function of x and y will be discussed in Section 19.5. However, we can get a preliminary idea of what to expect by applying our present method to simple cases.

Example 2 Find the tangent plane to the sphere

$$x^2 + y^2 + z^2 = 14 \tag{5}$$

at the point (1, 2, 3).

Solution Even though this sphere is not a surface of the form $z = f(x, y)$, it can be thought of as a combination of two such surfaces, the upper and lower hemispheres. By solving (5) for z, we see that the upper hemisphere is given by

$$z = f(x, y) = \sqrt{14 - x^2 - y^2},$$

so

$$f_x = \frac{-x}{\sqrt{14 - x^2 - y^2}} \qquad \text{and} \qquad f_y = \frac{-y}{\sqrt{14 - x^2 - y^2}}.$$

These formulas give

$$f_x(1, 2) = -\tfrac{1}{3} \qquad \text{and} \qquad f_y(1, 2) = -\tfrac{2}{3},$$

so the equation of the tangent plane is

$$z - 3 = -\tfrac{1}{3}(x - 1) - \tfrac{2}{3}(y - 2),$$

or

$$x + 2y + 3z = 14.$$

In this example we solved equation (5) explicitly for z, then proceeded as before. An alternative method that is often easier is to assume that the given equation defines z implicitly as a function of x and y, and to find the partial derivatives by implicit differentiation. With this method we use equation (4) in the slightly different form

$$z - z_0 = \left(\frac{\partial z}{\partial x}\right)_{P_0}(x - x_0) + \left(\frac{\partial z}{\partial y}\right)_{P_0}(y - y_0), \tag{6}$$

where the coefficients are written this way because $\partial z/\partial x$ and $\partial z/\partial y$ need not depend only on x and y.

Example 3 To find the tangent plane of Example 2 by the method just suggested, we first hold y fixed and differentiate (5) implicitly with respect to x, which gives

$$2x + 2z\frac{\partial z}{\partial x} = 0,$$

so $\partial z/\partial x = -x/z$. Similarly, $\partial z/\partial y = -y/z$. At the point $P_0 = (1, 2, 3)$, these partial derivatives have the numerical values

$$\left(\frac{\partial z}{\partial x}\right)_{P_0} = -\frac{1}{3} \quad \text{and} \quad \left(\frac{\partial z}{\partial y}\right)_{P_0} = -\frac{2}{3},$$

so by (6) the tangent plane is

$$z - 3 = -\tfrac{1}{3}(x - 1) - \tfrac{2}{3}(y - 2),$$

just as before. Of course, this method is of particular value when the equation of the surface is difficult or impossible to solve for z.

1. 标题与内容概述

《切平面的求法：显式与隐式两种情形的示例》

在前面小节中，我们已经了解了如何为一个“显式”给出的曲面 $z = f(x, y)$ 找到该曲面在某点处的切平面。本次课件的示例进一步展示了两种主要场景：

1. **显式函数** $z = f(x, y)$ 的切平面：直接对 f 求偏导数，然后套用线性化方程。
2. **隐式方程** $F(x, y, z) = 0$ 的切平面：利用**隐函数微分**求出 $\frac{\partial z}{\partial x}$ 和 $\frac{\partial z}{\partial y}$，同样可写出切平面方程。

通过对比，可以看到这两种方法在原理上是一致的，只是在处理形式上稍有不同；而当“显式解”不易或无法得到时，隐式微分法尤为重要。

---

2. 详细内容解析

下面分别对课件中两个示例及其解法思路做深入讲解。

2.1 示例 1：显式函数的切平面

> **题目：**
> 求函数
>
> $$z = f(x, y) = 2xy^3 - 5x^2$$
>
> 在点 $(3, 2, 3)$ 处的切平面。

1. **检验点是否在曲面上**
   - 一般先将 $(x, y) = (3, 2)$ 代入 f 中，查看 z 是否等于 3。课件中已假设检查无误。
2. ## 求偏导数
   - $$f_x(x, y) = \frac{\partial}{\partial x}(2xy^3 - 5x^2) = 2y^3 - 10x.$$
   - $$f_y(x, y) = \frac{\partial}{\partial y}(2xy^3 - 5x^2) = 6xy^2.$$
   - 在 $(x, y) = (3, 2)$ 处：

   $$f_x(3, 2) = 2 \cdot 2^3 - 10 \cdot 3 = 16 - 30 = -14,$$

   $$f_y(3, 2) = 6 \cdot 3 \cdot 2^2 = 6 \cdot 3 \cdot 4 = 72.$$

3. **写出切平面方程**
   - 通用形式（显式函数）：

$$z - z_0 = f_x(x_0, y_0)\,(x - x_0) \ + \ f_y(x_0, y_0)\,(y - y_0).$$

   - 这里 $(x_0, y_0, z_0) = (3, 2, 3)$ 。代入数值:

$$z - 3 = -14(x - 3) + 72(y - 2).$$

   - 这就是切平面的方程，也可根据需要化简或 rearrange。

2.2 示例 2：球面（隐式方程）的切平面

> **题目：**
> 已知球面方程
>
> $$x^2 + y^2 + z^2 = 14,$$
>
> 在点 $(1, 2, 3)$ 处求切平面。

本例展示了“**先把球面方程改写为显式**” 和 “**直接用隐式微分**” 两种方法的比较。

2.2.1 方法 A：先显式求 z 再求偏导

1. **将球面上半部分写为**

$$z = \sqrt{14 - x^2 - y^2}.$$

   （下半球则为 $-\sqrt{14 - x^2 - y^2}$ ，本题给出的点 $(1, 2, 3)$ 在上半球上，因此用正根号。）

2. # 求偏导

$$f(x, y) = \sqrt{14 - x^2 - y^2},$$

$$f_x(x, y) = \frac{\partial}{\partial x}\left(\sqrt{14 - x^2 - y^2}\right) = \frac{-x}{\sqrt{14 - x^2 - y^2}},$$

$$f_y(x, y) = \frac{-y}{\sqrt{14 - x^2 - y^2}}.$$

   - 代入 $(x, y) = (1, 2)$ ，并且 $z = 3$ 由题目给出。

$$f_x(1, 2) = \frac{-1}{\sqrt{14 - 1 - 4}} = \frac{-1}{\sqrt{9}} = -\frac{1}{3},$$

$$f_y(1, 2) = \frac{-2}{\sqrt{14 - 1 - 4}} = \frac{-2}{3} = -\tfrac{2}{3}.$$

3. **切平面方程**

$$z - z_0 = f_x(1, 2)(x - 1) + f_y(1, 2)(y - 2),$$

   其中 $z_0 = 3$ 。

$$z - 3 = -\tfrac{1}{3}(x - 1) \ - \ \tfrac{2}{3}(y - 2).$$

   若再整理，可以化成

$$x + 2y + 3z \ = \ 14.$$

2.2.2 方法 B：直接使用隐式微分

1. **球面隐式方程：**

$$F(x,y,z)=x^2+y^2+z^2-14=0.$$

2. **对 $F=0$ 作隐式微分**
   - 固定 y，对 x 做偏导：

   $$\frac{\partial}{\partial x}\left(x^2+y^2+z^2-14\right)=0 \quad\Longrightarrow\quad 2x+2z\frac{\partial z}{\partial x}=0.$$

   因而

   $$\frac{\partial z}{\partial x}=-\frac{x}{z}.$$

   - 固定 x，对 y 做偏导：

   $$2y+2z\frac{\partial z}{\partial y}=0 \quad\Longrightarrow\quad \frac{\partial z}{\partial y}=-\frac{y}{z}.$$

3. **将点 $(1,2,3)$ 代入**

$$\left(\frac{\partial z}{\partial x}\right)_{(1,2,3)}=-\frac{1}{3}, \qquad \left(\frac{\partial z}{\partial y}\right)_{(1,2,3)}=-\frac{2}{3}.$$

   这与方法 A 的结果一致。

4. **切平面方程**
   - 根据

   $$z-z_0=\left(\frac{\partial z}{\partial x}\right)_{P_0}(x-x_0)+\left(\frac{\partial z}{\partial y}\right)_{P_0}(y-y_0),$$

   代入数值：

   $$z-3=-\tfrac{1}{3}(x-1)\;-\;\tfrac{2}{3}(y-2).$$

   整理后同样得到

   $$x+2y+3z=14.$$

5. **小结：**
   - 当方程可以**轻松地显式化**（如把 z 表示出来）时，方法 A 直接把它当 $z=f(x,y)$ 来处理非常直观。
   - 当显式化困难或不可能时（比如很多复杂曲面），方法 B 的**隐式微分**更实用，只要能对 $F(x,y,z)=0$ 做偏导就行。

2.3 结论与使用场景

- **显式曲面 $z=f(x,y)$：**
  直接用

$$z-z_0=f_x(x_0,y_0)\,(x-x_0)+f_y(x_0,y_0)\,(y-y_0).$$

- **隐式曲面 $F(x,y,z)=0$：**
  用隐式微分求
  $\frac{\partial z}{\partial x}$ 和 $\frac{\partial z}{\partial y}$，
  再写出

$$z-z_0=\left(\tfrac{\partial z}{\partial x}\right)_{P_0}(x-x_0)+\left(\tfrac{\partial z}{\partial y}\right)_{P_0}(y-y_0).$$

- 本例（球面）两种方法都不算太困难，因此可以对比显示二者“殊途同归”。但在实际应用中，若方程 **“显式求解 $z=\ldots$”** 很麻烦，就往往选隐式法。

---

3. 本次内容小结

1. **切平面在显式情形**（ $z=f(x,y)$ ）的计算：先算偏导 f_x, f_y ，后套用线性近似。
2. **切平面在隐式情形**（ $F(x,y,z)=0$ ）的计算：利用隐式微分得 $\frac{\partial z}{\partial x}$ 和 $\frac{\partial z}{\partial y}$ 。
3. **两种思路本质相同**：都在讨论曲面在局部的线性逼近，只是写法有所不同。
4. **示例结果**：
   - 对 $z=2xy^3-5x^2$ 在 $(3,2,3)$ 的切平面：

$$z-3=-14(x-3)+72(y-2).$$

   - 对球面 $x^2+y^2+z^2=14$ 在 $(1,2,3)$ 的切平面：

$$x+2y+3z=14.$$

---

4. 报错或补充说明

- 本次示例的全部步骤和结论已经详尽说明；若有任何截图信息遗漏，请提示。无则表示笔记完整。

19.4

19.4 INCREMENTS AND DIFFERENTIALS. THE FUNDAMENTAL LEMMA

Most of calculus can be understood by using geometric intuition mixed with a little common sense, without getting bogged down in the underlying theory of the subject. In a few places, however, this theory is inescapable, because without it there is no way to grasp what is going on in the main developments of the subject itself. This is true for infinite series and the theory of convergence. It is also true for the topics of the next two sections—directional derivatives and the chain rule—which cannot be understood without a certain degree of attention to the theoretical issues that we now briefly discuss.

In order to see what these issues are, we begin by considering a function $y = f(x)$ of one variable that has a derivative at a point x_0. If Δx is an increment that carries x_0 to a nearby point $x_0 + \Delta x$ (see Fig. 19.8), we are interested in the corresponding increment in y,

$$\Delta y = f(x_0 + \Delta x) - f(x_0).$$

The definition of the derivative $f'(x_0)$ is

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}, \tag{1}$$

and this can be written in the equivalent form

$$\frac{\Delta y}{\Delta x} = f'(x_0) + \epsilon,$$

where $\epsilon \to 0$ as $\Delta x \to 0$. Accordingly, with no further hypotheses than the assumed existence of the derivative (1), we can write the increment Δy in the form

$$\Delta y = f'(x_0)\,\Delta x + \epsilon\,\Delta x, \qquad \text{where } \epsilon \to 0 \text{ as } \Delta x \to 0. \tag{2}$$

The situation is entirely different for a function of two (or more) variables, as we now explain.

Consider a function $z = f(x, y)$ and let (x_0, y_0) be a point at which the partial derivatives $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ both exist. The increment in z produced by moving from (x_0, y_0) to a nearby point $(x_0 + \Delta x, y_0 + \Delta y)$ is

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0),$$

as shown in Fig. 19.9. In order to develop the tools we shall need in Sections 19.5 and 19.6, it will be necessary to express Δz in a form analogous to (2),

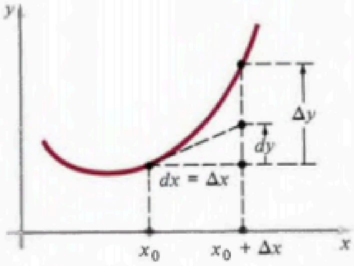


Figure 19.8 Differentials dx and dy.

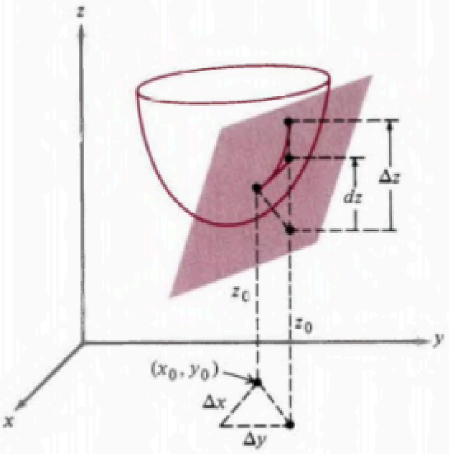


Figure 19.9 The differential dz.

$$\Delta z = f_x(x_0, y_0)\,\Delta x + f_y(x_0, y_0)\,\Delta y + \epsilon_1\,\Delta x + \epsilon_2\,\Delta y, \tag{3}$$

where ϵ_1 and $\epsilon_2 \to 0$ as Δx and $\Delta y \to 0$. Unfortunately, in sharp contrast to the one-variable case, the mere existence of the partial derivatives f_x and f_y at (x_0, y_0) is not enough to guarantee the validity of (3). Sufficient conditions for this conclusion are given in the

Fundamental Lemma *Suppose that a function $z = f(x, y)$ and its partial derivatives f_x and f_y are defined at a point (x_0, y_0), and also throughout some neighborhood of this point. Suppose further that f_x and f_y are continuous at (x_0, y_0). Then the increment Δz can be expressed in the form* (3), *where ϵ_1 and $\epsilon_2 \to 0$ as Δx and $\Delta y \to 0$.*

This statement is called a *lemma* for the usual reason: its significance lies not in itself, but rather in the use that can be made of it elsewhere. A proof is given in Appendix A.19.

We do not wish to linger on these matters, but nevertheless a few brief remarks are in order.

Remark 1 In the case of a function of one variable, (1) and (2) are equivalent, and if either condition holds it is customary to denote Δx by dx and to write $dy = f'(x_0)\,dx$, so that dy is the change in y along the tangent line. A function $z = f(x, y)$ for which $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ both exist is said to be *differentiable* at (x_0, y_0) if the conclusion of the lemma is valid—so that more is required than merely the existence of the partial derivatives. In this case—and *only* in this case!—we denote Δx and Δy by dx and dy, and we define the *differential dz* by*

$$dz = f_x(x_0, y_0)\,dx + f_y(x_0, y_0)\,dy.$$

Under these circumstances it can be proved that the surface $z = f(x, y)$ has a tangent plane at (x_0, y_0, z_0) and that *dz is the change in z* along this plane, as suggested in Fig. 19.9. The differential dz is usually written in the equivalent forms

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy \qquad \text{or} \qquad df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy.$$

Remark 2 A function $z = f(x, y)$ which is differentiable at a point is automatically continuous there. This follows at once from (3), which shows that $\Delta z \to 0$ if Δx and $\Delta y \to 0$. In the single-variable case, we know that if a function has a derivative at a point, then it is necessarily continuous there. However, this is not true for functions of more than one variable: the mere existence of the partial derivatives f_x and f_y at a point does not imply the continuity of $f(x, y)$ at that point. This is shown by the example of the bizarre function discussed in Section 19.1, for which $f_x(0, 0) = f_y(0, 0) = 0$ and yet the function is discontinuous at $(0, 0)$.

The concepts of a differentiable function and its differential, and also the Fundamental Lemma, can be extended in an obvious way to functions of any finite number of variables. This would involve much additional writing but no new ideas, and we shall not burden the reader with the details.

*Sometimes dz is called the *total differential*.

1. 标题与内容概述

《增量与微分：从单变量到多变量的可微性——基本引理的要点》

本次课件内容聚焦于“增量（increments）”和“微分（differentials）”的概念，以及它们与多变量函数可微性之间的关系。首先回顾了单变量情形中增量与导数的对应关系，然后在多变量情形下探讨为何光有偏导数并不一定保证函数可微，进而引入“基本引理（Fundamental Lemma）”的条件。最后给出当函数确实可微时，如何写出“全微分（total differential）”的形式。

---

2. 详细内容解析

以下按照课件顺序，对主要知识点和示例进行系统讲解和延伸。

2.1 单变量情形回顾：增量与导数

1. **增量 Δy 与导数 $\frac{dy}{dx}$**：
   - 对于单变量函数 $y = f(x)$，当自变量从 x_0 增加 Δx 到 $x_0 + \Delta x$ 时，函数值的增量为

     $$\Delta y = f(x_0 + \Delta x) - f(x_0).$$

   - 若在 x_0 处可导，则

     $$\left.\frac{dy}{dx}\right|_{x_0} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

   - 这等价于说：

     $$\Delta y = f'(x_0)\,\Delta x \;+\; \epsilon\,\Delta x, \quad \text{其中 } \epsilon \to 0 \text{当} \Delta x \to 0.$$

     或更简洁地写成

     $$\Delta y = f'(x_0)\,\Delta x, \quad \text{当}\Delta x\text{很小、忽略高阶小量时。}$$

2. **“微分” dy 的定义**：
   - 在单变量里，常把 Δx 记作 dx，并写

     $$dy = f'(x_0)\,dx.$$

   - 这样就把微分 dy 理解为“沿切线方向”函数值的改变量。

2.2 多变量情形：增量与偏导数

1. **增量 Δz**：
   - 对函数 $z = f(x, y)$，当从点 (x_0, y_0) 移动到 $(x_0 + \Delta x,\, y_0 + \Delta y)$ 时，函数值的增量

     $$\Delta z = f(x_0 + \Delta x,\, y_0 + \Delta y) \;-\; f(x_0, y_0).$$

2. **希望写成类比单变量的形式**：
   - 在单变量中，$\Delta y \approx f'(x_0)\,\Delta x$。
   - 对多变量，希望有

     $$\Delta z \approx f_x(x_0, y_0)\,\Delta x \;+\; f_y(x_0, y_0)\,\Delta y.$$

     其中 f_x, f_y 分别是对 x, y 的偏导数。

3. **问题：光有偏导不够！**
   - 书中指出：仅仅在 (x_0, y_0) 处存在偏导 f_x, f_y 并不能保证上式能被写成

     $$\Delta z = f_x(x_0, y_0)\,\Delta x + f_y(x_0, y_0)\,\Delta y \;+\; (\text{小量项}),$$

     其“小量项”要能控制到 $\epsilon_1 \Delta x + \epsilon_2 \Delta y$ 级别，并且 $\epsilon_1, \epsilon_2 \to 0$ 随 $\Delta x, \Delta y \to 0$。

- 在多变量情况下，存在偏导但函数不连续、或不能线性逼近的“怪函数”是可能的（课件曾提及某个在原点不连续的反例）。

2.3 基本引理（Fundamental Lemma）与可微性

1. **引理内容：**
   - 若 $z = f(x,y)$ 在 (x_0,y_0) 及其某邻域中定义，且 ** f_x, f_y 存在并且在 (x_0,y_0) 连续**，则增量 Δz 能写成
   $$\Delta z = f_x(x_0,y_0)\,\Delta x + f_y(x_0,y_0)\,\Delta y + \epsilon_1\,\Delta x + \epsilon_2\,\Delta y,$$
   其中 $\epsilon_1, \epsilon_2 \to 0$ 当 $\Delta x, \Delta y \to 0$。这实际上就是函数在该点“可微”的数学定义。
   - 换言之，**偏导数存在并连续**即可推出该函数在该点可微。
2. **可微性与切平面：**
   - 一旦满足可微性（即满足引理），便可证明曲面 $z = f(x,y)$ 在该点有惟一确定的切平面，其方程与之前我们学到的
   $$z - z_0 = f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$$
   相吻合。
   - 此时，“增量 Δz” 即可用**全微分（total differential）**近似。
3. **微分 dz 的写法：**
   - 若函数可微，则定义
   $$dz = f_x(x_0,y_0)\,dx + f_y(x_0,y_0)\,dy,$$
   并称之为该函数在 (x_0,y_0) 处的**微分**。
   - 也常写作
   $$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy, \quad \text{或} \quad df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy.$$

2.4 重要备注

1. **单变量情况 vs. 多变量情况**
   - 在单变量里，只要存在导数，就自动意味着函数在该点连续、可微，这些性质等价。
   - 在多变量里，仅有偏导数存在并不一定足以保证可微性或函数连续；需要补充“偏导数在邻域内存在且连续”这样的条件。
2. **可微性 $\implies$ 连续性**
   - 如果函数 $z = f(x,y)$ 在某点可微，那么它必然在该点连续。反之，若只知道偏导数存在，却无法断言函数连续或可微（可能是一个病态例子）。
3. **推广到更高维：**
   - 当函数有更多自变量时，情况类似。我们定义偏导数并要求其连续，以此保证可微。写出的全微分就会变成：
   $$dw = \frac{\partial w}{\partial x_1}dx_1 + \cdots + \frac{\partial w}{\partial x_n}dx_n.$$
   - 书中指出，虽然维度增加，但实质思想并无太大改变。

---

3. 本次内容小结

- **单变量回顾：** $\Delta y \approx f'(x_0)\Delta x$，且存在导数即意味着可微、连续等一系列特性。
- **多变量增量：** $\Delta z = f_x(x_0,y_0)\,\Delta x + f_y(x_0,y_0)\,\Delta y + \cdots$。要确保这能严格成立为线性逼近（附带小量）需要更强条件。
- **基本引理（可微性判定）**：若在某点及邻域内偏导数存在且连续，则该函数在该点可微。可微 $\implies$ 存在切平面 $\implies$ 函数在该点连续，并可定义全微分。
- **微分的写法**：
$$dz = f_x\,dx + f_y\,dy,$$

表示函数值沿着切平面的线性变化，是多变量微积分中非常重要的概念。

4. 报错或补充说明

- 根据截图信息，已对本节内容（增量、可微性、微分及“基本引理”）作了系统分析。若尚有截图遗漏或存在无法辨认之处，请告知；如无，则以上即为完整的学习笔记。

19.5

19.5 DIRECTIONAL DERIVATIVES AND THE GRADIENT

Let $f(x, y, z)$ be a function (of *three* variables!) defined throughout some region of three-dimensional space, and let P be a point in this region. At what rate does f change as we move away from P in a specified direction? In the directions of the positive x-, y-, and z-axes, we know that the rates of change of f are given by the partial derivatives $\partial f/\partial x$, $\partial f/\partial y$, and $\partial f/\partial z$. But how do we calculate the rate of change of f if we move away from P in a direction that is not a coordinate direction? In analyzing this problem, we will encounter the very important concept of the gradient of a function.

Suppose that the point P under consideration has coordinates x, y, and z, so that $P = (x, y, z)$; let $\mathbf{R} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ be the position vector of P, and let the specified direction be given by a unit vector $\mathbf{u}$, as shown in Fig. 19.10. If we move away from P in this direction to a nearby point $Q = (x + \Delta x, y + \Delta y, z + \Delta z)$, then the function f will change by an amount Δf. If we now divide this change Δf by the distance $\Delta s = |\Delta \mathbf{R}|$ between P and Q, then the quotient $\Delta f/\Delta s$ is the average rate of change of f (with respect to distance) as we move from P to Q. For instance, if the value of f at P is the temperature at this point, then $\Delta f/\Delta s$ is the average rate of change of temperature along the segment PQ. The limiting value of $\Delta f/\Delta s$ as Q approaches P, namely,

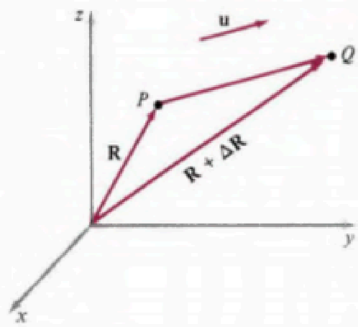


Figure 19.10

$$\frac{df}{ds} = \lim_{\Delta s \to 0} \frac{\Delta f}{\Delta s},$$

is called the *derivative of f at the point P in the direction* **u**, or simply the *directional derivative* of f. In the case of the temperature function, df/ds represents the instantaneous rate of change of temperature with respect to distance—roughly speaking, how fast it is getting hotter—at the point P as we move away from P in the direction specified by $\mathbf{u}$.

This is all very well, but how do we actually calculate df/ds in a specific case? To discover how to do this, we assume that $f(x, y, z)$ has continuous partial derivatives with respect to x, y, and z. Indeed, to avoid the tedious repetition of hypotheses, we make this a blanket assumption for every function we discuss, unless we explicitly state otherwise. With this, the Fundamental Lemma enables us to write Δf in the form

$$\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z + \epsilon_1 \Delta x + \epsilon_2 \Delta y + \epsilon_3 \Delta z, \tag{1}$$

where $\epsilon_1, \epsilon_2, \epsilon_3 \to 0$ as Δx, Δy, and $\Delta z \to 0$, that is, as $\Delta s \to 0$. Dividing (1) by Δs now gives

$$\frac{\Delta f}{\Delta s} = \frac{\partial f}{\partial x}\frac{\Delta x}{\Delta s} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta s} + \frac{\partial f}{\partial z}\frac{\Delta z}{\Delta s} + \epsilon_1 \frac{\Delta x}{\Delta s} + \epsilon_2 \frac{\Delta y}{\Delta s} + \epsilon_3 \frac{\Delta z}{\Delta s}, \tag{2}$$

and by taking the limit as $\Delta s \to 0$, we see that the last three terms in (2) approach zero and we obtain the formula

$$\frac{df}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} + \frac{\partial f}{\partial z}\frac{dz}{ds}. \tag{3}$$

This formula should be recognized as a special kind of chain rule, in the sense that as we move along the line through P and parallel to $\mathbf{u}$, f is a function of x, y, and z, where x, y, and z are in turn functions of the distance s, and (3) shows how to differentiate f with respect to s.

We observe that the first factor in each product on the right of (3) depends only on the function f and the coordinates of the point P at which the partial derivatives of f are evaluated, while the second factor in each product is independent of f and depends only on the direction in which df/ds is being calculated. These facts suggest that the right side of (3) ought to be thought of—and written—as the dot product of two vectors, as follows:

$$\begin{aligned}\frac{df}{ds} &= \left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}\right)\cdot\left(\frac{dx}{ds}\mathbf{i} + \frac{dy}{ds}\mathbf{j} + \frac{dz}{ds}\mathbf{k}\right) \\ &= \left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}\right)\cdot\frac{d\mathbf{R}}{ds}. \end{aligned} \tag{4}$$

The first factor here is a vector called the *gradient* of f. It is denoted by the symbol grad f, so that by definition

$$\operatorname{grad} f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}. \tag{5}$$

With this notation, (4) can be written as

$$\frac{df}{ds} = (\operatorname{grad} f)\cdot\frac{d\mathbf{R}}{ds}. \tag{6}$$

But we know that $d\mathbf{R}/ds$ is a unit vector, and since it has the same direction as $\mathbf{u}$, it equals $\mathbf{u}$. Formula (6) is therefore equivalent to

$$\frac{df}{ds} = (\operatorname{grad} f)\cdot\mathbf{u}. \tag{7}$$

This tells us how to calculate df/ds, because (5) is presumably simple to compute from the given function f, and then to evaluate at the given point P, and the dot product (7) of two known vectors is easy to find.

For a given function f and a given point P, grad f is a fixed vector which can be placed so that its tail lies at P. We also place the tail of $\mathbf{u}$ at P, as shown in Fig. 19.11. To understand the significance of grad f, we use the definition of the dot product and the fact that $\mathbf{u}$ is a unit vector to write (7) in the form

$$\frac{df}{ds} = |\operatorname{grad} f|\cos\theta, \tag{8}$$

where θ is the angle between grad f and $\mathbf{u}$. Since the direction of $\mathbf{u}$ can be chosen to suit our convenience, (8) immediately yields the first fundamental property of the gradient:

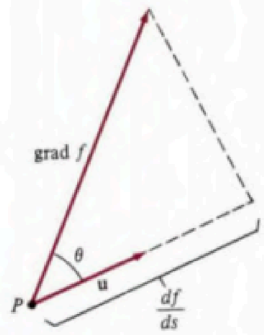


Figure 19.11 Directional derivative.

Property 1 The directional derivative df/ds in any given direction is the scalar projection of grad f in that direction (see Fig. 19.11).

In this sense, the single vector grad f contains within itself the directional derivatives of f at P in all possible directions.

Next, if $\mathbf{u}$ is chosen to point in the same direction as grad f, so that $\theta = 0$ and $\cos 0 = 1$, then (8) shows that df/ds has its maximum value—that is, f increases most rapidly—in this direction. Also, this maximum value equals $|\operatorname{grad} f|$. These remarks give the next two fundamental properties of the gradient:

Property 2 The vector grad f points in the direction in which f increases most rapidly.

Property 3 The length of the vector grad f is the maximum rate of increase of f.

1. 标题与内容概述

《方向导数与梯度：多变量函数的变化率与几何意义》

在三维空间里，如果有一个函数 $f(x, y, z)$（例如温度场、位势场等），我们不仅可以讨论它沿着坐标轴方向的变化率（即偏导数），还常需要考察**在任意给定方向**上函数的变化快慢程度。**方向导数（directional derivative）**正是用来描述函数在某方向上的瞬时变化率，而**梯度（gradient）**向量则能在几何上统一体现各方向上的变化行为。本节的主要结论包括：

1. **方向导数公式** $\frac{df}{ds} = f_x \frac{dx}{ds} + f_y \frac{dy}{ds} + f_z \frac{dz}{ds}$。
2. **方向导数可写成点积** $(\nabla f) \cdot \mathbf{u}$，其中 $\mathbf{u}$ 是单位方向向量。
3. **梯度** ∇f 是一个矢量，指向函数增长最快的方向，其大小 $\|\nabla f\|$ 则是函数在该点的最大增加率。

---

2. 详细内容解析

下面分步骤讲解方向导数的定义、推导，以及梯度向量的引入与重要几何性质。

2.1 方向导数的引入

1. **问题背景：**
   - 在二维或三维场景中，仅知道 $\frac{\partial f}{\partial x}$、$\frac{\partial f}{\partial y}$、$\frac{\partial f}{\partial z}$ 只能告诉我们函数在与坐标轴平行方向上的变化率。
   - 然而，更一般地，我们想知道："若从一点 P 出发，沿任意指定方向前进，函数值 f 的变化率是多少？" 这就是**方向导数**要解决的问题。
2. **形式化定义：**
   - 设 $f = f(x, y, z)$，在点 $P = (x_0, y_0, z_0)$ 给定一个方向向量 $\mathbf{u}$，且要求 $\mathbf{u}$ 是单位向量（长度为 1）。
   - 如果我们从 P 沿着 $\mathbf{u}$ 前进一个非常小的距离 Δs，到达点 $Q \approx P + \Delta s\,\mathbf{u}$，则 f 的增量为

   $$\Delta f = f(Q) - f(P).$$

   考虑平均变化率 $\Delta f / \Delta s$，再取极限：

   $$\left.\frac{df}{ds}\right|_{P,\mathbf{u}} = \lim_{\Delta s \to 0} \frac{\Delta f}{\Delta s}.$$

   这称为**沿方向 $\mathbf{u}$ 的方向导数**。

2.2 方向导数公式推导

1. **坐标变化与增量表达式：**
   - 在前文的**基本引理**（多变量可微性）中指出：如果 f_x, f_y, f_z 存在并连续，则小的增量可以写为

   $$\Delta f = f_x(x_0, y_0, z_0)\,\Delta x + f_y(x_0, y_0, z_0)\,\Delta y + f_z(x_0, y_0, z_0)\,\Delta z + (\text{小量}),$$

   并且当 $\Delta x, \Delta y, \Delta z \to 0$，这就像一个线性近似。
2. **与"沿着 $\mathbf{u}$ 方向前进"相对应：**
   - 若把 $\Delta x = \frac{dx}{ds}\Delta s,\ \Delta y = \frac{dy}{ds}\Delta s,\ \Delta z = \frac{dz}{ds}\Delta s$，那么

   $$\frac{\Delta f}{\Delta s} = f_x \frac{dx}{ds} + f_y \frac{dy}{ds} + f_z \frac{dz}{ds} + \dots$$

   在极限 $\Delta s \to 0$ 下，高阶小量归于 0，得到

   $$\frac{df}{ds} = f_x \frac{dx}{ds} + f_y \frac{dy}{ds} + f_z \frac{dz}{ds}.$$

   - 因为我们沿 $\mathbf{u}$ 前进，$\mathbf{u} = \langle \frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds} \rangle$ 是一个单位向量，所以这三个分量恰好是 $\mathbf{u}$ 的坐标。

2.3 梯度向量（Gradient）与方向导数

1. **点积形式：**
   书中指出，可以把

   $$\frac{df}{ds} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \cdot \left(\frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds}\right).$$

   将其视为两个向量的**点积**：

   $$\frac{df}{ds} = (\nabla f) \cdot \mathbf{u},$$

   其中

   $$\nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle$$

   称为**梯度向量**（gradient）。

2. **几何含义**
   - **梯度向量** ∇f 在点 P 是一个三维向量，方向与函数最强上升方向一致，长度则是“最大上升速率”。
   - 若沿 $\mathbf{u}$ 方向前进，则方向导数即为 $(\nabla f) \cdot \mathbf{u} = \|\nabla f\| \, \|\mathbf{u}\| \cos\theta$，其中 θ 是 ∇f 与 $\mathbf{u}$ 的夹角。由于 $\mathbf{u}$ 是单位向量，$\|\mathbf{u}\| = 1$。
   - 由此可见，$\theta = 0$（即 $\mathbf{u}$ 沿着梯度方向）时，$\cos\theta = 1$，得到最大方向导数值 $\|\nabla f\|$；$\theta = 180^\circ$ 时（逆梯度方向）最小方向导数值 $-\|\nabla f\|$；$\theta = 90^\circ$ 时（正交于梯度）方向导数为 0（函数不升不降）。

2.4 梯度的三大性质

在本段最后，课件列出了梯度的 3 个关键几何/物理意义：

1. **Property 1:**
   方向导数 $\frac{df}{ds}$ 在任何方向 $\mathbf{u}$ 上都是 ∇f 在该方向的投影，即 $\frac{df}{ds} = (\nabla f) \cdot \mathbf{u}$。
2. **Property 2:**
   梯度向量 ∇f 的方向是**函数在该点增长最快的方向。**
3. **Property 3:**
   梯度向量 ∇f 的长度 $\|\nabla f\|$ 是**函数在该点最大增长率。**

这些性质在解析几何和物理中应用广泛。例如：

- 在温度场中，∇T 指向温度上升最迅速的方向。
- 在等势场中，$\nabla \phi$ 指向力（或流）的方向与强度。

---

3. 本次内容小结

- **方向导数**：描述在给定方向上的变化率，可以用偏导数线性组合来计算，也可以在矢量形式上用点积写作 $(\nabla f) \cdot \mathbf{u}$。
- **梯度（gradient）**：将偏导数组合成的向量，其方向是“最陡峭上升”方向，长度是“最大上升率”。
- **方向导数与梯度的关系**：

  $$\frac{df}{ds} = \nabla f \cdot \mathbf{u}.$$

  这个公式在解析几何、物理场理论、优化等领域都十分重要。

---

4. 报错或补充说明

- 截图范围内的全部核心概念和公式已涵盖，如有任何缺失或需要额外详细之处，请告知；否则以上笔记即为完整。

19.5.2

Example 1 If $f(x, y, z) = x^2 - y + z^2$, find the directional derivative df/ds at the point $(1, 2, 1)$ in the direction of the vector $4\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$.

Solution At the point $(1, 2, 1)$, we have $\text{grad}\, f = 2x\mathbf{i} - \mathbf{j} + 2z\mathbf{k} = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$. We obtain a unit vector $\mathbf{u}$ in the desired direction by dividing the given vector by its own length,

$$\mathbf{u} = \frac{4\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}}{\sqrt{16 + 4 + 16}} = \frac{2}{3}\mathbf{i} - \frac{1}{3}\mathbf{j} + \frac{2}{3}\mathbf{k}.$$

Formula (7) now gives

$$\begin{aligned}\frac{df}{ds} &= (\text{grad}\, f) \cdot \mathbf{u} \\ &= (2\mathbf{i} - \mathbf{j} + 2\mathbf{k}) \cdot (\tfrac{2}{3}\mathbf{i} - \tfrac{1}{3}\mathbf{j} + \tfrac{2}{3}\mathbf{k}) = 3.\end{aligned}$$

Thus, the function f is increasing at the rate of 3 units per unit distance as we leave $(1, 2, 1)$ in the given direction.

Example 2 Let the temperature of the air at points in space be given by the function $f(x, y, z) = x^2 - y + z^2$. A mosquito located at $(1, 2, 1)$ wishes to get cool as soon as possible. In what direction should it fly?

Solution We saw in Example 1 that $\text{grad}\, f = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$ at the point $(1, 2, 1)$. Since the direction of $\text{grad}\, f$ is that in which the temperature increases most rapidly, the mosquito should fly in the opposite direction, that of $-\text{grad}\, f = -2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.

The fourth fundamental property of the gradient is useful in geometry. In order to explain what it is, we denote the point under consideration by $P_0 = (x_0, y_0, z_0)$ to emphasize that it is fixed in this discussion, and we let c_0 be the value of our function f at the point P_0. Then the set of all points in space at which $f(x, y, z)$ has the same value c_0 constitutes, in general, a level surface through P_0 whose equation is $f(x, y, z) = c_0$. We wish to show that the vector $\text{grad}\, f$ is normal (perpendicular) to this level surface at the point P_0, as suggested on the left in Fig. 19.12. To this end, we consider a curve that lies on the surface and passes through P_0. If we move to a nearby point Q on this curve and measure s along the curve, then $\Delta f = 0$ because f has the same value at all points on the surface, and therefore $df/ds = 0$ at P_0 in the direction of the tangent to the curve. Formula (6) remains valid and implies that

$$(\text{grad}\, f) \cdot \frac{d\mathbf{R}}{ds} = 0, \tag{9}$$

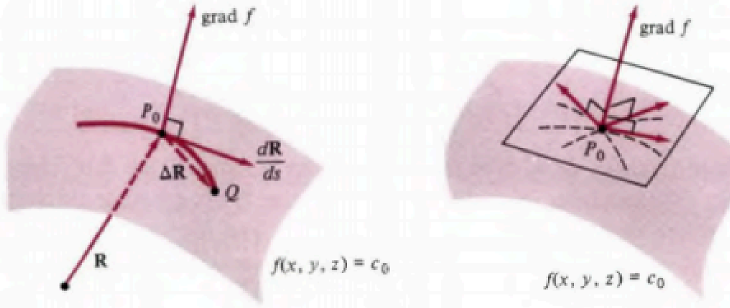


Figure 19.12 Gradient is normal to level surface.

where $d\mathbf{R}/ds$ is the unit tangent vector to the curve at P_0. The vanishing of the dot product in (9) tells us that grad f is perpendicular to this tangent vector. But the same reasoning applies to every curve on the surface that passes through P_0, so grad f is perpendicular to the tangent vectors to all these curves (Fig. 19.12, right). Since these tangent vectors determine the tangent plane at P_0, and being normal to the surface means being normal to this tangent plane, we have:

Property 4 The gradient of a function $f(x, y, z)$ at a point P_0 is normal to the level surface of f that passes through P_0.

In the context of this discussion, we point out that the equation of any surface can be written in the form $f(x, y, z) = c_0$, and can therefore be regarded as a level surface of the function $f(x, y, z)$. If $P_0 = (x_0, y_0, z_0)$ is a point on this surface, then Property 4 tells us that the vector

$$\mathbf{N} = \text{grad } f = \left(\frac{\partial f}{\partial x}\right)_{P_0}\mathbf{i} + \left(\frac{\partial f}{\partial y}\right)_{P_0}\mathbf{j} + \left(\frac{\partial f}{\partial z}\right)_{P_0}\mathbf{k}$$

is normal to the tangent plane at P_0, so if $\mathbf{N} \neq \mathbf{0}$, the equation of this tangent plane is

$$\left(\frac{\partial f}{\partial x}\right)_{P_0}(x - x_0) + \left(\frac{\partial f}{\partial y}\right)_{P_0}(y - y_0) + \left(\frac{\partial f}{\partial z}\right)_{P_0}(z - z_0) = 0. \qquad (10)$$

We observe that this equation includes equation (4) in Section 19.3 as a special case; for if the surface is given in the form $z = g(x, y)$, then this can be written as $g(x, y) - z = 0$, so the surface is a level surface of the function $f(x, y, z) = g(x, y) - z$, and this makes the coefficients in (10) equal to $g_x(x_0, y_0)$, $g_y(x_0, y_0)$, -1.

Example 3 Find the equation of the tangent plane to the surface $xy^2z^3 = 12$ at the point $(3, -2, 1)$.

Solution This surface is a level surface of the function $f(x, y, z) = xy^2z^3$. The vector grad f at the point $(3, -2, 1)$ is normal to the surface at this point. This vector is

$$\begin{aligned} \text{grad } f &= y^2z^3\mathbf{i} + 2xyz^3\mathbf{j} + 3xy^2z^2\mathbf{k} \\ &= 4\mathbf{i} - 12\mathbf{j} + 36\mathbf{k} = 4(\mathbf{i} - 3\mathbf{j} + 9\mathbf{k}). \end{aligned}$$

Therefore the equation of the tangent plane is

$$(x - 3) - 3(y + 2) + 9(z - 1) = 0$$

or

$$x - 3y + 9z = 18.$$

Remark 1 The main uses of directional derivatives and gradients are found in the geometry and physics of three-dimensional space. However, these concepts can also be defined in two dimensions, and they have similar (but thinner) properties. Thus, a curve $f(x, y) = c_0$ can be thought of as a level curve of the function $z = f(x, y)$; and if the gradient of this function is defined by

$$\text{grad } f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j},$$

then the value of this gradient at a point $P_0 = (x_0, y_0)$ on the curve is a vector that is normal to the curve.

Remark 2 The gradient of a function $f(x, y, z)$ can be written in "operational form" as

$$\text{grad } f = \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) f.$$

The *del operator* preceding the function f is usually denoted by the symbol ∇ (an inverted delta, read "del") so that

$$\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}.$$

This del operator is similar to, but more complicated than, the familiar differentiation operator d/dx. When del is applied to a function f, it produces a vector, namely, the vector grad f. In this notation, formulas (5), (6), and (7) become

$$\text{grad } f = \nabla f, \qquad \frac{df}{ds} = \nabla f \cdot \frac{d\mathbf{R}}{ds}, \qquad \text{and} \qquad \frac{df}{ds} = \nabla f \cdot \mathbf{u}.$$

We will make very substantial use of the operator ∇ in Chapter 21.

1. 标题与内容概述

《方向导数与梯度的应用：示例与梯度法向量之几何意义》

在前一部分里我们已经了解了如何求**方向导数**以及为什么梯度 ∇f 可以帮助我们用点积的形式快速得到方向导数。本次内容通过若干示例进一步加深对方向导数和梯度的应用理解，并引入梯度与**水平面（或水平曲面/等值面）**之间的“正交（垂直）”关系，从而得到曲面的切平面方程。

---

2. 详细内容解析

2.1 方向导数的实际计算——示例

示例 1

> **题意：** 对函数
>
> $$f(x,y,z) = x^2 \; - \; y \; + \; z^2,$$
>
> 在点 $(1,2,1)$ 沿向量 $\langle 4,\; -2,\; 4\rangle$ 的方向求方向导数 $\frac{df}{ds}$ 。

1. # 计算梯度 ∇f

$$\nabla f = \left(\frac{\partial f}{\partial x},\; \frac{\partial f}{\partial y},\; \frac{\partial f}{\partial z}\right) = \left(2x,\; -1,\; 2z\right).$$

   - 在给定点 $(x,y,z) = (1,2,1)$:

$$\nabla f(1,2,1) = (2\cdot 1,\; -1,\; 2\cdot 1) = (2,\; -1,\; 2).$$

2. **标准化方向向量**
   - 方向向量 $\mathbf{v} = \langle 4,\; -2,\; 4\rangle$ ，其长度 $\|\mathbf{v}\| = \sqrt{4^2 + (-2)^2 + 4^2} = \sqrt{16+4+16} = \sqrt{36} = 6$ 。
   - **单位向量** $\mathbf{u}$ 即为

$$\mathbf{u} = \frac{1}{6}\,\langle 4,\; -2,\; 4\rangle = \left(\tfrac{2}{3},\; -\tfrac{1}{3},\; \tfrac{2}{3}\right).$$

3. **计算方向导数**
   - 方向导数公式：

$$\frac{df}{ds} = \nabla f \cdot \mathbf{u}.$$

   - 在该点：

$$(2,\; -1,\; 2)\; \cdot\; \left(\tfrac{2}{3},\; -\tfrac{1}{3},\; \tfrac{2}{3}\right) = 2\cdot\tfrac{2}{3} \;+\; (-1)\cdot\left(-\tfrac{1}{3}\right) \;+\; 2\cdot\tfrac{2}{3} = \tfrac{4}{3} + \tfrac{1}{3} + \tfrac{4}{3} = 3.$$

   - **结论：** 沿该方向，每移动一单位距离， f 增加 3 个单位值。

示例 2

> **题意：** 若同一函数
>
> $$f(x,y,z) = x^2 \; - \; y \; + \; z^2$$
>
> 表示空间中某点的温度（假设在点 $(1,2,1)$ 处），且蚊子想尽快变得凉快，那么它应向哪个方向飞？

- 上例中在点 $(1,2,1)$ 的梯度 $\nabla f = (2,\; -1,\; 2)$ 。
- 梯度方向是函数**增长**最快的方向。若这是温度场，则顺着 ∇f 是“温度上升最快”的方向。
- 蚊子想要“温度下降最快”，就应朝着与梯度相反的方向 $-\nabla f$ 飞，**即** $\langle -2,\; 1,\; -2\rangle$ 的方向。

2.2 梯度正交于等值面（或水平面）

1. **梯度与等值面**
   - 回想若 $f(x,y,z) = c$ 表示某个**水平面（或等值面、level surface）**，那么在这个曲面上 f 的值处处相同。
   - 若我们在曲面上沿任意切线（曲面上的方向）移动， $\Delta f = 0$ 。因此沿曲面的切线方向的方向导数也为 0。
   - 由方向导数公式 $\nabla f \cdot \mathbf{t} = 0$ 可知： ∇f 必**垂直**（正交）于每一条切向线 $\mathbf{t}$ ；因此它就是**曲面的法向量**。
2. **引申为“切平面”方程**
   - 若点 (x_0, y_0, z_0) 落在曲面 $f(x,y,z) = c$ 上，则切平面的法向量就是

$$\mathbf{N} = \nabla f(x_0, y_0, z_0).$$

- 切平面方程：

$$\nabla f \cdot \langle x - x_0,\ y - y_0,\ z - z_0 \rangle = 0,$$

即

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0)\,(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0, z_0)\,(y - y_0) + \frac{\partial f}{\partial z}(x_0, y_0, z_0)\,(z - z_0) = 0.$$

性质总结（Property 4）：

> 对于任意三元函数 $f(x, y, z)$，它在点 P_0 所生成的等值面 $f(x, y, z) = c_0$ 的**法向量**就是该点处的梯度 $\nabla f(P_0)$。

2.3 示例：由梯度求切平面方程

示例 3

> **题意**：在曲面
>
> $$xy^2z^3 = 12$$
>
> 上的点 $(3,\ -2,\ 1)$ 处，求该曲面的切平面方程。

1. **视作"隐式"水平面**
   - 将曲面方程写成

   $$f(x, y, z) = xy^2z^3 - 12 = 0,$$

   这里 $f \equiv 0$ 即一条等值面（值是 0）。
2. ## 求梯度 ∇f

   $$f_x = \frac{\partial}{\partial x}\big(xy^2z^3 - 12\big) = y^2z^3,$$

   $$f_y = \frac{\partial}{\partial y}\big(xy^2z^3 - 12\big) = 2xy\,z^3,$$

   $$f_z = \frac{\partial}{\partial z}\big(xy^2z^3 - 12\big) = 3xy^2\,z^2.$$

   - 在 $(x, y, z) = (3,\ -2,\ 1)$：

   $$f_x(3, -2, 1) = (-2)^2 \cdot 1^3 = 4, \quad f_y(3, -2, 1) = 2 \cdot 3 \cdot (-2) \cdot 1^3 = -12, \quad f_z(3, -2, 1) = 3 \cdot 3 \cdot (-2)^2 \cdot 1^2 = 3 \cdot 3 \cdot 4 = 36.$$

   - 故 $\nabla f\big(3, -2, 1\big) = (4,\ -12,\ 36)$。
3. **切平面方程**
   - 用梯度做法向量：

   $$4(x - 3)\ +\ (-12)(y - (-2))\ +\ 36(z - 1) = 0.$$

   - 也可写作

   $$4(x - 3) - 12(y + 2) + 36(z - 1) = 0,$$

   或简化成：

   $$x - 3y + 9z = 18.$$

   - 这就是该曲面在 $(3, -2, 1)$ 处的切平面方程。

2.4 补充：运算符 ∇ （del 运算符）

- 在课件中，梯度常写作

$$\nabla f = \frac{\partial f}{\partial x}\,\mathbf{i} \;+\; \frac{\partial f}{\partial y}\,\mathbf{j} \;+\; \frac{\partial f}{\partial z}\,\mathbf{k}.$$

- 这也可视为一个“操作符”

$$\nabla = \frac{\partial}{\partial x}\,\mathbf{i} + \frac{\partial}{\partial y}\,\mathbf{j} + \frac{\partial}{\partial z}\,\mathbf{k},$$

  作用在函数 f 上即产生梯度向量 ∇f 。

- 用这种**算子**写法方便记忆和运算，比如：

$$\frac{df}{ds} = \nabla f \;\cdot\; \frac{d\mathbf{R}}{ds},$$

  其中 $\mathbf{R}(s)$ 是移动路径的参数表示。

---

3. 本次内容小结

1. **方向导数计算**：将所给方向归一化后，与 ∇f 取点积即可；若问题仅问“最大/最小增长方向”，则直接用梯度或其相反方向。
2. **梯度与等值面**： ∇f 在空间中正交于该点所在的“等值面”或“水平面”（ $f(x,y,z)=c$ ）。
3. **切平面方程**：若曲面由 $f(x,y,z)=c$ 给出，则在点 (x_0,y_0,z_0) 的切平面法向量是 $\nabla f(x_0,y_0,z_0)$ 。可写出：

$$\frac{\partial f}{\partial x}\Big|_0 (x - x_0) + \frac{\partial f}{\partial y}\Big|_0 (y - y_0) + \frac{\partial f}{\partial z}\Big|_0 (z - z_0) = 0.$$

4. **算子** ∇ ：在高等微积分中又称“del”运算符，其作用在标量函数上产生梯度，作用在向量函数上可以生成散度、旋度等更深层次概念（后续会见到）。

---

4. 报错或补充说明

- 本次截图的全部内容（包括示例以及梯度与等值面的垂直性）已完整讲解。若有任何无法辨认的部分，请告知；如无，则以上为本节笔记的总结与补充。

19.6

19.6 THE CHAIN RULE FOR PARTIAL DERIVATIVES

The single-variable chain rule for ordinary derivatives tells us how to differentiate composite functions. It says that if w is a function of x where x is in turn a function of a third variable t, say $w = f(x)$ where $x = g(t)$, then

$$\frac{dw}{dt} = \frac{dw}{dx}\frac{dx}{dt}. \tag{1}$$

We know from ample experience that this is an indispensable tool of calculus; it is used more frequently than any other differentiation rule.

The simplest multivariable chain rule involves a function $w = f(x, y)$ of two variables x and y, where x and y are each functions of another variable t, $x = g(t)$ and $y = h(t)$. Then w is a function of t,

$$w = f[g(t), h(t)] = F(t),$$

and we shall prove that the derivative of this composite function is given by the formula

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}. \tag{2}$$

This is the *chain rule* for this situation.

The proof of (2) is easy. We begin by changing t to $t + \Delta t$, where $\Delta t \neq 0$. This increment in t produces increments Δx and Δy in x and y, which in turn produce an increment Δw in w. Since all the functions we discuss are assumed to have continuous partial derivatives, the Fundamental Lemma enables us to write Δw in the form

$$\Delta w = \frac{\partial w}{\partial x}\Delta x + \frac{\partial w}{\partial y}\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y, \tag{3}$$

where ϵ_1 and $\epsilon_2 \to 0$ as Δx and $\Delta y \to 0$. On dividing (3) by Δt, we obtain

$$\frac{\Delta w}{\Delta t} = \frac{\partial w}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial w}{\partial y}\frac{\Delta y}{\Delta t} + \epsilon_1\frac{\Delta x}{\Delta t} + \epsilon_2\frac{\Delta y}{\Delta t}. \tag{4}$$

If we now form the limit as $\Delta t \to 0$, then Δx and Δy also $\to 0$, so ϵ_1 and $\epsilon_2 \to 0$, and (4) immediately yields (2).

Example 1 If $w = 3x^2 + 2xy - y^2$ where $x = \cos t$ and $y = \sin t$, find dw/dt.

Solution Formula (2) tells us that

$$\frac{dw}{dt} = (6x + 2y)(-\sin t) + (2x - 2y)\cos t.$$

By substituting $x = \cos t$ and $y = \sin t$, we can express this in terms of t alone,

$$\begin{aligned}\frac{dw}{dt} &= (6\cos t + 2\sin t)(-\sin t) + (2\cos t - 2\sin t)(\cos t)\\ &= -6\sin t\cos t - 2\sin^2 t + 2\cos^2 t - 2\sin t\cos t\\ &= 2(\cos^2 t - \sin^2 t) - 8\sin t\cos t = 2\cos 2t - 4\sin 2t.\end{aligned}$$

We can check this result by first substituting and then differentiating, which gives

$$w = 3\cos^2 t + 2\sin t\cos t - \sin^2 t$$

and

$$\begin{aligned}\frac{dw}{dt} &= 6\cos t(-\sin t) + 2\sin t(-\sin t) + 2\cos^2 t - 2\sin t\cos t\\ &= 2(\cos^2 t - \sin^2 t) - 8\sin t\cos t = 2\cos 2t - 4\sin 2t,\end{aligned}$$

as before.

In the situation of formula (2), it is convenient to call w the *dependent variable*, x and y the *intermediate variables,* and t the *independent variable.* We notice that the right side of (2) has two terms, one for each intermediate variable, and that each of these terms resembles the right side of the single-variable chain rule (1).

Formula (2) extends in an obvious way to any number of intermediate variables. For instance, if $w = f(x, y, z)$ where x, y, and z are each functions of t, then

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}. \tag{5}$$

The proof of this is essentially the same as the proof of (2), except that it uses the Fundamental Lemma for three variables instead of two.

Further, x, y, and z here need not be functions of only one independent variable, but can be functions of two or more variables. Thus, if x, y, and z are each functions of the variables t and u, then w is also a function of t and u, and its partial derivatives are given by

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t} \tag{6}$$

and

$$\frac{\partial w}{\partial u} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial u}. \tag{7}$$

We use roundback d's everywhere here because every function depends on more than one variable. It is necessary to be very clear about the meanings of the letters in formulas like these. For example, on the left side of (6), w is considered a function of t and u, while on the right side it is considered a function of x, y, and z. The proofs are the same as before, and all of these formulas—(2), (5), (6), (7), and their extensions to any number of intermediate and independent variables—are collectively called the *chain rule.*

In Section 19.4 we defined the differential dw of a function $w = f(x, y, z)$ by the formula

$$dw = \frac{\partial w}{\partial x}\,dx + \frac{\partial w}{\partial y}\,dy + \frac{\partial w}{\partial z}\,dz. \tag{8}$$

The chain rule (5) tells us that if x, y, z are themselves functions of a single independent variable t, then it is permissible to calculate dw/dt by formally dividing (8) by dt. Similarly, if x, y, z are functions of the independent variables t, u and we want to calculate $\partial w/\partial t$, then the chain rule (6) tells us that we can find $\partial w/\partial t$ by dividing (8) by dt and writing roundback ∂'s in place of ordinary d's to show that there is another independent variable present which is being held fixed.

The individual terms on the right side of (8) are sometimes called the *partial differentials* of w with respect to x, y, z. From this point of view, the quantity dw defined by (8) deserves the name *total differential,* as we remarked in Section 19.4.

1. 标题与内容概述

《偏导数的链式法则：从单变量扩展到多变量的复合函数求导》

在单变量微积分中，若函数 $y = f(x)$ 而 x 又是 t 的函数，即 $x = g(t)$，则复合函数 $y = f(g(t))$ 的导数由**链式法则**给出

$$\frac{dy}{dt} = \frac{dy}{dx}\,\frac{dx}{dt}.$$

本节介绍了多变量场景下的相应推广：当 $w = f(x, y, z, \ldots)$，而 x, y, z 各自又依赖某些其他独立变量（如 t, u 等），则我们如何计算 $\frac{dw}{dt}$ 或 $\frac{\partial w}{\partial t}$ 等导数？这便是**多变量的链式法则**（the chain rule for partial derivatives）。

---

2. 详细内容解析

2.1 最简单的多变量链式法则

1. **情形：** $w = f(x, y)$，其中 $x = g(t)$，$y = h(t)$。
   于是 $w = f\big(g(t), h(t)\big)$ 成为单变量 t 的复合函数，想要求 $\frac{dw}{dt}$。
2. **公式：**

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\,\frac{dx}{dt} + \frac{\partial w}{\partial y}\,\frac{dy}{dt}.$$

   - 其中 $\frac{\partial w}{\partial x}$ 和 $\frac{\partial w}{\partial y}$ 都在相同的点处计算（即把 $x = g(t)$，$y = h(t)$ 代入）。
   - 它可以看作是单变量链式法则的自然扩展：每个中间变量都会贡献一项。
3. **证明思路：**
   - 类似在单变量的做法：把 Δt 引起的 Δx 与 Δy 带入增量公式

$$\Delta w = f_x\,\Delta x + f_y\,\Delta y + (\text{小量}).$$

   - 令 $\Delta x = \frac{dx}{dt}\Delta t$，$\Delta y = \frac{dy}{dt}\Delta t$，再除以 Δt、取极限即得结果。

2.2 三变量情形及更多一般化

1. **三变量情形：** $w = f(x, y, z)$，其中 x, y, z 各自为 t 的函数（例如 $x = g(t), y = h(t), z = k(t)$）。则

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}.$$

这就是课件中的公式 (5)。

2. **多独立变量的情形：**
   - 如果 x, y, z 还依赖多个独立变量（例如 t, u），那么在想计算 $\frac{\partial w}{\partial t}$ 时，就把 u 当作常数不变，只对 t 的变化进行追踪。这类似于多变量隐函数的偏导运算。
   - 例如对两独立变量 t, u 而言，

$$w = f(x, y, z) \text{ 且 } x = x(t, u),\ y = y(t, u),\ z = z(t, u),$$

   则

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t},$$
$$\frac{\partial w}{\partial u} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial u}.$$

   这就是课件中的公式 (6) 和 (7)。

3. **总微分的观点：**
   - 课件里提到的 (8) 式：

$$dw = \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy + \frac{\partial w}{\partial z}dz,$$

   若 x, y, z 都是某个 t 的函数，我们“形式上”可以对 (8) 除以 dt 来得到 $\frac{dw}{dt}$。
   - 这个做法在多变量微积分中非常常见，也就是把微分形式与偏导数相结合来理解复合函数的导数。

2.3 示例计算

Example 1（课件中的示范）

- **题意：** 令

$$w = 3x^2 + 2xy - y^2, \quad x = \cos t, \quad y = \sin t.$$

  求 $\frac{dw}{dt}$。

1. **应用链式法则**：

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}.$$

2. **先求偏导数**：

$$\frac{\partial w}{\partial x} = 6x + 2y, \quad \frac{\partial w}{\partial y} = 2x - 2y.$$

3. **求** $\frac{dx}{dt}, \frac{dy}{dt}$：
   - $x = \cos t \Rightarrow \frac{dx}{dt} = -\sin t, \quad y = \sin t \Rightarrow \frac{dy}{dt} = \cos t.$
4. **代入：**

$$\frac{dw}{dt} = (6\cos t + 2\sin t)(-\sin t) + (2\cos t - 2\sin t)\cos t.$$

课件给出的步骤将结果化简为可在纯三角函数形式中表达，并可通过“先把 $x = \cos t, y = \sin t$ 代入 w 再直接微分”来验证，两种结果相同。

2.4 关键要点和总结

1. **多变量链式法则精要**：要找 w 对某个独立变量的导数，就把所有中间变量的偏导数分别乘上各自对该独立变量的导数，再相加。
2. **总微分理解**：

$$dw = \frac{\partial w}{\partial x} dx + \frac{\partial w}{\partial y} dy + \cdots$$

如果每个 $x, y, \ldots$ 又都是函数 t 的函数，则

$$\frac{dw}{dt} = \frac{dw}{dx}\frac{dx}{dt} + \frac{dw}{dy}\frac{dy}{dt} + \cdots.$$

3. **层层嵌套**：若有更多层的复合（比如 $x = x(u, v), y = y(u, v)$ 等），可继续使用同理：对需要的偏导"把路径分开"，对所有中间依赖方向做加和。
4. **在应用时**：
    - 明确哪些是"**自变量**"（可以改变），哪些是"**中间变量**"（本身是函数），以及哪一个/哪几个是"**依赖变量**"（要对其进行求导）。
    - 小心记号：$\frac{\partial w}{\partial x}$ 是在固定其他自变量时对 x 的偏导，而 $\frac{dx}{dt}$ 是 x 作为 t 函数的普通导数。

---

3. 本次内容小结

- **单变量链式法则**：$\frac{dy}{dx}\frac{dx}{dt} = \frac{dy}{dt}$ 推广到多变量后，需要对每个中间变量的偏导与其对独立变量的导数相乘，并把所有结果相加。
- **多变量链式法则公式**
    1. $w = f(x, y)$，$x = x(t), y = y(t)$

    $$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}.$$

    2. $w = f(x, y, z)$，$x = x(t), y = y(t), z = z(t)$

    $$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}.$$

    3. 更多独立或中间变量时，依旧是"**每个路径都算贡献然后求和**"。
- **总微分与微分形式**：$dw = \dfrac{\partial w}{\partial x}\,dx + \dfrac{\partial w}{\partial y}\,dy + \cdots$ 可以在更抽象的层面统一理解链式法则。

---

4. 报错或补充说明

- 根据截图与课件内容，已完整介绍单变量与多变量的链式法则、公式证明思路及应用示例。若有任何不清晰之处请告知，否则本次笔记即为完整。

19.6.2

Example 2 A function of several variables is said to be *homogeneous of degree n* if multiplying each variable by t (where $t > 0$) has the same effect as multiplying the original function by t^n. Thus, $f(x, y)$ is homogeneous of degree n if

$$f(tx, ty) = t^n f(x, y). \tag{9}$$

For example, $f(x, y) = x^2 + 3xy$ is homogeneous of degree 2, because $f(tx, ty) = (tx)^2 + 3(tx)(ty) = t^2(x^2 + 3xy) = t^2 f(x, y)$. Similarly, $f(x, y) = (x + y)/(x - y)$ is homogeneous of degree 0, $f(x, y) = (xy - x^2 e^{x/y})/y$ is homogeneous of degree 1, and $f(x, y, z) = \sqrt{x^3 - 3xy^2 + 2z^3}$ is homogeneous of degree $\frac{3}{2}$. Most functions, for instance $f(x, y) = y^2 + x \sin y$, are not homogeneous at all.

There is a theorem of Euler about homogeneous functions that has several important applications: *If $f(x, y)$ is homogeneous of degree n, then*

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = nf(x, y). \tag{10}$$

To prove this, we hold x and y fixed and differentiate both sides of (9) with respect to t. We can clarify this process by writing $u = tx$ and $v = ty$, so that (9) becomes

$$f(u, v) = t^n f(x, y).$$

Then by using the chain rule to differentiate with respect to t, we obtain

$$\frac{\partial f}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial t} = nt^{n-1} f(x, y)$$

or

$$x \frac{\partial f}{\partial u} + y \frac{\partial f}{\partial v} = nt^{n-1} f(x, y),$$

and putting $t = 1$ yields (10). Similarly, if $f(x_1, x_2, \ldots, x_m)$ is homogeneous of degree n, then the same argument shows that

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \cdots + x_m \frac{\partial f}{\partial x_m} = nf(x_1, x_2, \ldots, x_m).$$

Euler's theorem has some interesting consequences for economics. As an example, suppose that $f(x, y)$ is the production (measured in dollars) of x units of capital and y units of labor. If the amounts of capital and labor are doubled, then it is reasonable to expect that the resulting production will also double, that is, that $f(2x, 2y) = 2f(x, y)$. More generally, we expect that

$$f(tx, ty) = tf(x, y),$$

so the production function is homogeneous of degree 1. [In economics, this property of $f(x, y)$ is called *constant returns to scale.*] Euler's theorem now says that

$$f(x, y) = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y}. \tag{11}$$

The partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ are called the *marginal product of capital* and the *marginal product of labor,* respectively. In this language, (11) is a theorem of quantitative economics whose verbal statement is, "The total value of production equals the cost of capital plus the cost of labor if each is paid for at the rate of its marginal product." Under these circumstances there are no surplus earnings, and in the real world this is a very bad thing.*

Example 3 Many applications of the chain rule involve calculating the effect on some equation or expression when new variables are introduced. As an illustration of a method that will be useful for solving the wave equation in Section 19.9, we now solve the partial differential equation

$$a\frac{\partial w}{\partial x} = \frac{\partial w}{\partial y}, \qquad a \neq 0. \tag{12}$$

That is, we find the most general function $w = f(x, y)$ that satisfies this equation. To do this, we introduce new independent variables u, v by writing

$$u = x + ay, \qquad v = x - ay. \tag{13}$$

We think of w as a function of u and v,

$$w = F(u, v),$$

and we find the uv-equation equivalent to (12) by using the chain rule to write

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial w}{\partial u} + \frac{\partial w}{\partial v},$$

$$\frac{\partial w}{\partial y} = \frac{\partial w}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial y} = a\frac{\partial w}{\partial u} - a\frac{\partial w}{\partial v}.$$

By substituting in these expressions, we see that (12) transforms into the partial differential equation

$$2a\frac{\partial w}{\partial v} = 0 \qquad \text{or} \qquad \frac{\partial w}{\partial v} = 0.$$

*For further information on these matters, see pp. 81–84 of J. M. Henderson and R. E Quandt, *Microeconomic Theory* (McGraw-Hill, 1971); or Chapter 12, "Homogeneous Functions and Euler's Theorem," in D. E. James and C. D. Throsby, *Quantitative Methods in Economics* (Wiley, 1973). For an application of Euler's theorem to advanced theoretical mechanics, see p. 531 of the present writer's text, *Differential Equations,* 2nd ed. (McGraw-Hill, 1991).

This equation is very easy to solve, because it says that the function $w = F(u, v)$ is constant when u is held fixed and v is allowed to vary, and therefore is a function of u alone. This means that our desired solution of (12) is

$$w = g(u) = g(x + ay),$$

where $g(u)$ is a *completely arbitrary* (continuously differentiable) function of u. We apologize to students for introducing "out of the blue" the apparently unmotivated transformation equations (13). However, some of the developments of Section 19.9 will make this procedure seem fairly natural.

Example 4 Partial derivatives are the main mathematical tools used in thermodynamics. It is the universal practice in this science to avoid confusion by using subscripts on partial derivatives to specify the variable (or variables) held fixed in the differentiation. Thus, if $w = F(x, y)$ then $\partial w/\partial x$ would be denoted by

$$\left(\frac{\partial w}{\partial x}\right)_y.$$

This notation tells us that w is being thought of as a function of x and y, and that y is held fixed and x is the variable of differentiation. This usage may seem superfluous, but the following situation—which is quite common in thermodynamics—shows that it is not.

If $w = f(x, y)$ where y is a function $g(x, t)$ of x and another variable t, so that w is a composite function of x and t, we find its partial derivative with respect to x.

This is a typical chain rule situation with x and y the intermediate variables and x and t the independent variables:

$$w = f(x, y) \qquad \text{where} \quad \begin{cases} x = x, \\ y = g(x, t). \end{cases}$$

The chain rule therefore gives

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial x}, \tag{14}$$

so

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial x}. \tag{15}$$

Unfortunately this equation contains two partial derivatives of w with respect to x. By an effort of thought one can keep in mind that $\partial w/\partial x$ on the left of (15) is the derivative of the composite function, while $\partial w/\partial x$ on the right is the derivative of $w = f(x, y)$. Nevertheless, this ambiguous notation invites confusion and is contrary to the overall spirit of mathematical symbols, which are intended to make it easy to be correct with a minimum of thought. However, if we use the subscript notation of thermodynamics, then (14) can be written as

$$\left(\frac{\partial w}{\partial x}\right)_t = \left(\frac{\partial w}{\partial x}\right)_y\left(\frac{\partial x}{\partial x}\right)_t + \left(\frac{\partial w}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_t.$$

Since $(\partial x/\partial x)_t = 1$, this becomes

$$\left(\frac{\partial w}{\partial x}\right)_t = \left(\frac{\partial w}{\partial x}\right)_y + \left(\frac{\partial w}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_t, \tag{16}$$

which is somewhat clumsy but much less vulnerable to misunderstanding than (15).*

*Students whose main interest is physics may wish to read discussions of these matters in some of the standard treatises. See, for example, p. 19 of Enrico Fermi, *Thermodynamics* (Dover, 1956); p. 28 of Philip M. Morse, *Thermal Physics* (W. A. Benjamin, 1969); or pp. 30–33, 52–55 of F. W. Sears, *Thermodynamics* (Addison-Wesley, 1953).

1. 标题与内容概述

《多变量链式法则的拓展与应用：Euler 同次定理、经济学解释及变量变换实例》

在之前的内容里我们学习了多变量链式法则的基本形式和应用方法。本次课件则在此基础上进一步探讨了以下几方面：

1. **同次齐次函数（Homogeneous Functions）与 Euler 定理**：如果某个多变量函数满足 $f(tx, ty, \ldots) = t^n f(x, y, \ldots)$，则称它是同次齐次函数；Euler 定理给出该函数与其偏导数的一条重要关系式，在经济学等领域有重大应用。
2. **经济学背景：** 把生产函数作为同次齐次函数时，可用 Euler 定理解析资本和劳动力的边际产量之和。
3. **引入新变量的链式法则应用**：例如，用新独立变量 u, v 替换原来 (x, y)，可将偏微分方程简化。
4. **热力学记号及部分微分符号的潜在歧义**：在物理或化学中常用下标表示“保持哪些变量不变”，从而更明确地强调当下的偏导数所对应的情形。

---

2. 详细内容解析

下面围绕课件主要例子和定理展开，从同次齐次函数与 Euler 定理，到更一般的变量变换及在物理、经济学上的解释。

2.1 同次齐次函数及 Euler 定理

1. **齐次函数的定义**
   - 若对某函数 $f(x, y, \ldots)$ 存在一个正实数 n，使得对任意 $t > 0$，
   $$f(tx,\ ty,\ \ldots) = t^n\, f(x, y, \ldots),$$
   则称 f 是**同次（齐次）函数（homogeneous function）**，且“阶”为 n。
   - 例如：
     - $f(x, y) = x^2 + 3xy$ 是二阶齐次函数：$f(tx, ty) = t^2(x^2 + 3xy)$。
     - $f(x, y) = (xy - x^2 e^{-y})/y$ 是一阶齐次。
     - 大部分函数并非齐次，例如 $y^2 + \sin(y)$ 并不满足此性质。
2. **Euler 定理（多元情形）**
   - 对二元函数：若 $f(x, y)$ 是同次 n 阶，则有
   $$x\,\frac{\partial f}{\partial x} \;+\; y\,\frac{\partial f}{\partial y} \;=\; n\, f(x, y). \tag{10}$$
   - **证明思路简要**：
     1. 写 $u = tx,\ v = ty$，那么 $f(u, v) = f(tx, ty) = t^n f(x, y)$。
     2. 对 t 求导（保持 x, y 不变）并在 $t = 1$ 时代入，可得到上述结果。
   - 更一般地，对 m 个变量的 n 阶齐次函数 $f(x_1, \ldots, x_m)$，

$$\sum_{i=1}^{m} x_i \frac{\partial f}{\partial x_i} = n f(x_1, \ldots, x_m).$$

3. **经济学解释：**
   - 若 $f(x,y)$ 表示生产函数（产出）随资本 x 与劳动力 y 的投入而变化，并且其满足**"规模报酬不变（constant returns to scale）**"的假设（即同次齐次度 $n=1$），则

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = f(x,y).$$

   - 其中 $\frac{\partial f}{\partial x}$ 称为**资本的边际产量**，$\frac{\partial f}{\partial y}$ 称为**劳动力的边际产量**。该结果在经济学中也有直观含义："产出总值 = 资本报酬 + 劳动力报酬"，每项各按其边际产量计价。

2.2 变量变换与链式法则——偏微分方程示例

1. **示例：波动方程简化**
   - 给定偏微分方程（PDE）：

$$a\frac{\partial w}{\partial x} = \frac{\partial w}{\partial y}, \quad a \neq 0. \tag{12}$$

   - 通过定义新变量

$$u = x + a\,y, \quad v = x - a\,y, \tag{13}$$

   可将 w 视为 $w = F(u,v)$。

   - 使用链式法则，对 $\frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}$ 进行转换：

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial x}, \quad \frac{\partial w}{\partial y} = \frac{\partial w}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial y}.$$

   代入后，(12) 归结为

$$2a\frac{\partial w}{\partial v} = 0,$$

   即 $\frac{\partial w}{\partial v} = 0$，说明 w 不依赖 v，故

$$w = g(u) = g(x + ay),$$

   其中 g 是某个任意可微函数。这种方法对后续更复杂的波动方程、超曲面简化很有用。

2.3 热力学记号与下标含义

1. **热力学与物理中的下标**
   - 在物理、化学、工程等领域，常见记号如 $\left(\frac{\partial w}{\partial x}\right)_y$ 表示在微分运算时"将 y 视为常量"，而 $\left(\frac{\partial w}{\partial x}\right)_t$ 则表示在微分时"将 t 视为常量"等。
   - 这与我们在数学中常用的 $\frac{\partial w}{\partial x}$ 略有区别，因为在多变量场景下，你必须明确"除了 x 之外，哪些变量被固定？"。
2. **课件中的示例：**
   - 若 $w = f(x,y)$ 但 y 又是另一个函数 $y = g(x,t)$，我们想要 $\frac{\partial w}{\partial x}$ 而把 t 视为不变，就必须在公式里分清哪部分是对 x 的"直接"偏导，哪部分是对 y 而后 y 又对 x 的影响。
   - 这种情况会产生类似

$$\frac{\partial w}{\partial x} = \left(\frac{\partial w}{\partial x}\right)_y + \left(\frac{\partial w}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_t,$$

   并且需要在记号中表明对谁保持不变，以免混淆。
3. **(15) 与 (16) 公式差异：**
   - 教材中提到，从数学角度用 $\frac{\partial w}{\partial x}$ 有时有歧义，因为不清楚"其余变量是否保持固定？"

- 在热力学中，会写得更加明确：

$$\left(\frac{\partial w}{\partial x}\right)_t = \left(\frac{\partial w}{\partial x}\right)_y + \left(\frac{\partial w}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_t.$$

- 虽然显得更繁琐，但避免了误解，是物理化学中普遍使用的“下标记法”。

---

3. 本次内容小结

1. **Euler 同次定理**：
   若 f 为 n 阶齐次函数，则

$$\sum_{i=1}^{m} x_i \frac{\partial f}{\partial x_i} = n\, f(x_1, \dots, x_m).$$

   在经济学的产出函数里，若 $n = 1$ 则表示“规模报酬不变”，有重要经济学含义。

2. **变量变换与链式法则**：
   - 在偏微分方程中，引入新变量（如 $u = x + ay, v = x - ay$）后，用链式法则可化简原方程。
   - 一旦化简，可得出相对容易求解的形式，比如 $\frac{\partial w}{\partial v} = 0$ 意味着 w 与 v 无关。

3. **热力学记号**：
   - 在应用学科中，为标明“谁被保持不变”，往往使用 $\left(\frac{\partial w}{\partial x}\right)_y$ 或 $\left(\frac{\partial w}{\partial x}\right)_t$ 等记号。
   - 这是因为在实际问题里，多数变量可能互相依赖，需要明确哪一个变量是“真”的独立量。

---

4. 报错或补充说明

- 当前截图的主要知识点（Euler 同次定理、经济学解释、偏微分方程中的变量替换、热力学记号）均已详细讲解。若有任何遗失或难辨之处请指出；否则以上就是本次内容的完整笔记。

19.7

19.7 MAXIMUM AND MINIMUM PROBLEMS

In the case of functions of a single variable, one of the main applications of derivatives is to the study of maxima and minima. In Chapter 4 we developed various tests involving first and second derivatives, and we used these tests for graphing functions and attacking a wide variety of geometric and physical problems. Maximum and minimum problems for functions of two or more variables can be much more complicated. We confine ourselves here to an introduction to such problems, including a two-variable version of the second derivative test (Remark 3, Section 4.2).

Suppose that a function $z = f(x, y)$ has a maximum value at a point $P_0 = (x_0, y_0)$ in the interior of its domain. This means that $f(x, y)$ is defined, and also $f(x, y) \le f(x_0, y_0)$, throughout some neighborhood of P_0, as shown on the left in Fig. 19.13.* If we hold y fixed at the value y_0, then $z = f(x, y_0)$ is a function of x alone, and since it has a maximum value at $x = x_0$, its derivative must be zero there, as in Chapter 4. That is, $\partial z/\partial x = 0$ at this point. In just the same way, $\partial z/\partial y = 0$ at this point. The equations

$$\frac{\partial z}{\partial x} = 0 \quad \text{and} \quad \frac{\partial z}{\partial y} = 0 \qquad (1)$$

*In this discussion we are considering only a so-called *relative* (or *local*) maximum, which takes into account only points near to P_0, but for simplicity we drop the adjective.

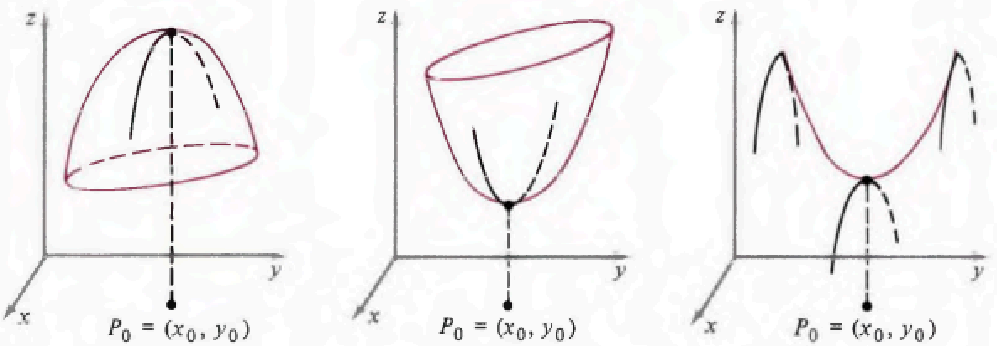


Figure 19.13

are therefore two equations in two unknowns that are satisfied at the maximum point (x_0, y_0). In many cases we can solve these equations simultaneously to find the point (x_0, y_0), and thus the actual maximum value of the function.

Exactly the same considerations apply to the minimum value shown in the center of the figure. However, when we try to locate maximum or minimum values of a function by solving equations (1), it is necessary to keep in mind that these equations are also satisfied at a saddle point like that shown on the right, where the function has a maximum in one direction and a minimum in some other direction. Equations (1) mean only that the tangent plane is horizontal, and it is then up to us to decide what significance this fact has.

By analogy with our earlier definition in Chapter 4 for functions of one variable, we call a point (x_0, y_0) where both partial derivatives are zero a *critical point* of $f(x, y)$.

Example 1 Find the dimensions of the rectangular box with open top and a fixed volume of 4 ft^3 which has the smallest possible surface area.

Solution If x and y are the edges of the base, and z is the height, then the area is

$$A = xy + 2xz + 2yz.$$

Since $xyz = 4$, we have $z = 4/xy$, and the area to be minimized can be expressed as a function of the two variables x and y,

$$A = xy + \frac{8}{y} + \frac{8}{x}. \tag{2}$$

We seek a critical point of this function, that is, a point where

$$\frac{\partial A}{\partial x} = y - \frac{8}{x^2} = 0, \qquad \frac{\partial A}{\partial y} = x - \frac{8}{y^2} = 0.$$

To solve these equations simultaneously, we first write them as

$$x^2y = 8, \qquad xy^2 = 8.$$

Dividing gives $x/y = 1$, so $y = x$ and either equation becomes $x^3 = 8$. Therefore $x = y = 2$, and it follows from this that $z = 1$, so the box with minimum surface area has a square base and a height one-half the edge of the base.

In this example it is geometrically clear that the critical point (2, 2) is actually a minimum point, and not a maximum or saddle point. However, in a more complicated situation we might find a critical point and yet be completely unable to state its nature, based on commonsense considerations alone. A useful tool for classifying critical points is provided by the *second derivative test*:

If $f(x, y)$ has continuous second partial derivatives in a neighborhood of a critical point (x_0, y_0), and if a number D (called the discriminant) is defined by

$$D = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - [f_{xy}(x_0, y_0)]^2, \tag{3}$$

then (x_0, y_0) is

(i) *a maximum point if $D > 0$ and $f_{xx}(x_0, y_0) < 0$;*
(ii) *a minimum point if $D > 0$ and $f_{xx}(x_0, y_0) > 0$;*
(iii) *a saddle point if $D < 0$.*

Further, if $D = 0$, then no conclusion can be drawn, and any of the behaviors described in (i) *to* (iii) *can occur.*

A complete proof of this theorem requires machinery that is not available to us. We refer interested students to more advanced books.*

Example 1 (continued) As an illustration of the use of the second derivative test, we apply it to verify that the critical point (2, 2) found in Example 1 is a minimum point of the function (2). Here we have

$$A_{xx} = \frac{16}{x^3}, \qquad A_{yy} = \frac{16}{y^3}, \qquad A_{xy} = 1,$$

so the discriminant (3) has the value $D = 2 \cdot 2 - 1^2 = 3 > 0$. Since A_{xx} is also positive at the point (2, 2), the test tells us that this critical point is indeed a minimum point, as claimed.

Example 2 Find the critical points of the function

$$z = 3x^2 + 2xy + y^2 + 10x + 2y + 1,$$

and use the second derivative test to classify them.

Solution Here we have

$$\frac{\partial z}{\partial x} = 6x + 2y + 10 = 0 \qquad \text{and} \qquad \frac{\partial z}{\partial y} = 2x + 2y + 2 = 0,$$

so the system of equations we must solve is

$$3x + y = -5,$$
$$x + y = -1.$$

By simple manipulations we easily see that $x = -2$ and $y = 1$, so there is a single critical point $(-2, 1)$. At this point we have

$$D = z_{xx}z_{yy} - {z_{xy}}^2 = 6 \cdot 2 - 2^2 = 8 > 0,$$

and since $z_{xx} = 6 > 0$, the critical point is a minimum point.

*See, for example, pp. 157–159 of R. C. Buck, *Advanced Calculus* (McGraw-Hill, 1978).

Success in finding maximum and minimum points for a function $z = f(x, y)$ clearly depends on our ability to solve the two simultaneous equations $f_x = 0$ and $f_y = 0$. In Examples 1 and 2 these equations were very easy to solve. However, as students can readily imagine, there are many complicated situations that arise in which routine methods of solving simultaneous equations are quite useless. The only general advice we can give is to try to solve one of the equations for one of the unknowns in terms of the other, substitute this in the second equation, and try to solve the result. Apart from this, make good guesses and be ingenious —advice that is easier to give than to follow!

1. 标题与内容概述

《多变量极值问题：临界点与二阶导数判定》

在单变量微积分中，我们知道找极值（最大值或最小值）通常先求导数并令其为零，得到所谓的"临界点"，再结合二阶导数测试等方法进行判定。对于二元函数 $z = f(x, y)$，寻找局部极值也需要找出**偏导数同时为零**的点（临界点），然后再使用相应的**二阶偏导数判定法**（second derivative test）加以区分是极大值、极小值还是马鞍点。

---

2. 详细内容解析

以下内容围绕本节主要知识点，分步骤进行详尽说明，从如何求临界点，到如何用二阶导数判别极值，最后辅以例题。

2.1 临界点的定义与求法

1. **临界点（critical point）**
   - 对一个二元函数 $z = f(x, y)$，如果在某点 (x_0, y_0) 处偏导数同时满足
     $$\frac{\partial z}{\partial x}\Big|_{(x_0,y_0)} = 0, \quad \frac{\partial z}{\partial y}\Big|_{(x_0,y_0)} = 0,$$
     则称 (x_0, y_0) 为函数的临界点。
   - 物理意义：沿着坐标轴方向的"瞬时变化率"都为 0，暗示切平面水平，函数可能处于局部高峰、低谷或鞍形拐点等。
2. **极值点必为临界点（必要条件）**
   - 若在函数定义域内部有一点 (x_0, y_0) 是局部最大值或最小值，那么它必定是临界点。
   - 这类似单变量的情况：极值处导数为零，但需要注意这只是"必要不充分"的条件，因有些临界点也会对应马鞍点（既非极大也非极小）。
3. **求解临界点的基本步骤**
   - **Step 1**：求偏导数 f_x 和 f_y。
   - **Step 2**：解方程组
     $$f_x(x, y) = 0, \quad f_y(x, y) = 0.$$
   - 这有时容易，有时会很复杂，需要观察、代入或数值方法等。

2.2 二阶导数判定法（Second Derivative Test）

1. **二阶偏导数与判定判别式** D
   - 令
     $$f_{xx} = \frac{\partial^2 f}{\partial x^2}, \quad f_{yy} = \frac{\partial^2 f}{\partial y^2}, \quad f_{xy} = \frac{\partial^2 f}{\partial x\,\partial y}.$$
   - 定义

$$D(x_0,y_0) = f_{xx}(x_0,y_0)\, f_{yy}(x_0,y_0) \;-\; \big[f_{xy}(x_0,y_0)\big]^2. \tag{3}$$

这被称为判别式（discriminant）。

2. **判定规则**
假设 f_{xx}，f_{yy}，f_{xy} 在点 (x_0, y_0) 连续，并且 (x_0, y_0) 是临界点，则
   1. 若 $D > 0$ 且 $f_{xx}(x_0,y_0) < 0$，则 (x_0, y_0) 为**局部最大值**；
   2. 若 $D > 0$ 且 $f_{xx}(x_0,y_0) > 0$，则 (x_0, y_0) 为**局部最小值**；
   3. 若 $D < 0$，则 (x_0, y_0) 为**鞍点**（既非极大也非极小）；
   4. 若 $D = 0$，则该测试不确定，需要借助更高级方法或综合其他信息来判定。
3. **与单变量对比**
   - 单变量时，如果二阶导数 $f''(x_0) > 0$，则最小；若 $f''(x_0) < 0$，则最大；若 $f''(x_0) = 0$，不确定。
   - 二变量时也有类似的“正”“负”“零”的区分，但因为多了交叉项 f_{xy}，需要借助判别式 D。

2.3 示例与应用

示例 1：开口箱子最小表面积

> **题目：** 体积固定为 4 立方英尺的无盖长方体箱子，求使表面积最小的边长。

1. **设定变量**
   - 设底边长 x，y，高度 z。体积 $xyz = 4$。
   - 箱子表面积（无盖）$A = xy + 2xz + 2yz$。由 $z = \frac{4}{xy}$，可得

$$A(x,y) = xy \;+\; 2x\Big(\frac{4}{xy}\Big) \;+\; 2y\Big(\frac{4}{xy}\Big) = xy \;+\; \frac{8}{y} \;+\; \frac{8}{x}.$$

2. **找临界点**
   - $A_x = y - \frac{8}{x^2} = 0 \Rightarrow xy^2 = 8.$
   - $A_y = x - \frac{8}{y^2} = 0 \Rightarrow x^2y = 8.$
   - 解出 $x = y = 2$，则 $z = \frac{4}{2\cdot 2} = 1$。
   - 直觉 + 二阶导数计算（或现实判断）可知这是最小表面积解。最终结论：箱子尺寸 $2 \times 2 \times 1$。

示例 2：判定点类型

> **题目：** 给函数
>
> $$z = 3x^2 + 2xy + y^2 + 10x + 2y + 1,$$
>
> 求临界点并用二阶导数测定其性质。

1. **求偏导并设为零**

$$z_x = 6x + 2y + 10 = 0, \quad z_y = 2x + 2y + 2 = 0.$$

   - 方程组：$\{3x + y = -5,\ x + y = -1\}$。解得 $(x, y) = (-2, 1)$。
2. **计算二阶偏导及判别式**

$$z_{xx} = 6, \quad z_{yy} = 2, \quad z_{xy} = 2.$$

   - 判别式 $D = z_{xx}\, z_{yy} - (z_{xy})^2 = 6 \cdot 2 - 2^2 = 12 - 4 = 8 > 0$。且 $z_{xx} = 6 > 0$。
   - 按判定法则：$D > 0$ 且 $z_{xx} > 0$，故该临界点是**局部最小值**。

2.4 总结与常见问题

1. **找到临界点后**，我们必须再进行二阶判定或其他方式来区分“极大、极小、鞍点”。
2. **判别式为零或无穷大**时，需要补充分析，包括直接观察函数形式、利用更高级技巧（例如方向导数、三阶项等）。
3. **解方程组难度**：在实际问题里，常常方程组相当复杂，可能需要数值方法或其他特殊技巧。

---

3. 本次内容小结

- **临界点**： $f_x = 0,\ f_y = 0$ 必须同时成立。
- **二阶导数测试**：计算

$$D = f_{xx} f_{yy} - (f_{xy})^2,$$

  - 若 $D > 0$ 且 $f_{xx} < 0$，是局部最大；
  - 若 $D > 0$ 且 $f_{xx} > 0$，是局部最小；
  - 若 $D < 0$，是鞍点；
  - 若 $D = 0$，需进一步方法判定。
- **几何理解**：在点 (x_0, y_0)，“水平”切平面表明偏导为零，但表面曲率可能朝上（最小）、朝下（最大）、或在相反方向有不同曲率（鞍点）。

---

4. 报错或补充说明

- 从截图所示内容来看，已全面覆盖“临界点定义、二阶偏导数测试及示例”的要点。若有任何图示或文字缺失，请告知；否则本次笔记即为完整。

19.8

19.8 CONSTRAINED MAXIMA AND MINIMA. LAGRANGE MULTIPLIERS

In this section we explain the method of Lagrange multipliers by means of intuitive ideas that depend on the geometric meaning of gradients. This method is used for maximizing or minimizing functions of several variables subject to one or more constraints. It is an important tool in economics, differential geometry, and advanced theoretical mechanics.

We begin with the simplest case, that of two variables and one constraint.

In Section 19.7 we learned how to calculate maximum and minimum values of a function $z = f(x, y)$ of two independent variables x and y. However, in many problems x and y are not independent, but instead are connected by a *side condition* or *constraint* in the form of an equation

$$g(x, y) = 0. \tag{1}$$

In Chapter 4 we became thoroughly familiar with situations of this kind. The following is a simple illustration.

Example 1 Find the dimensions of the rectangle of maximum area that can be inscribed in a semicircle of radius a.

Solution It is clear from Fig. 19.15 that the problem is to maximize the function

$$A = 2xy \tag{2}$$

subject to the constraint

$$x^2 + y^2 = a^2. \tag{3}$$

In Example 3 of Section 4.3 we solved this problem by using the constraint (3) to express A as a function of only one variable,

$$y = \sqrt{a^2 - x^2}, \qquad \text{so that} \qquad A = 2x\sqrt{a^2 - x^2}.$$

We then calculated dA/dx, set it equal to zero, solved the resulting equation, and so on. This example will be continued after some remarks and explanations.

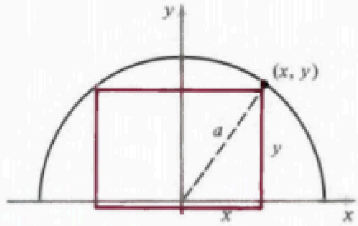


Figure 19.15

The procedure we have just described works well enough for this problem, but as a general method it has two defects. First, in this particular case equation (3) is easy to solve for y, but in another problem the constraint (1) might be so complicated that it would be difficult or impossible to solve. The other defect lies in the fact that even though the variables x and y play equal roles in the problem, they are handled differently in the solution: We singled out one variable, x, to be the independent variable, and the other, y, to be the dependent variable. It is often more convenient, and certainly more elegant, to treat such problems in a symmetric form, in which no preference is given to any one of the variables over the others.*

We now return to the general problem of maximizing a function $f(x, y)$ subject to the constraint $g(x, y) = 0$. To understand what is going on, we sketch the graph of $g(x, y) = 0$ (Fig. 19.16) together with several level curves $f(x, y) = c$ of the function $f(x, y)$, noting the direction in which c increases. In the figure, for instance, we suppose that $c_1 < c_2 < c_3 < c_4$. To find the maximum value of $f(x, y)$ along the curve $g(x, y) = 0$, we look for the largest c for which $f(x, y) = c$ intersects $g(x, y) = 0$. At such an intersection point (P_0 in the figure) the two curves have the same tangent line, so they also have the same normal line. But the vectors

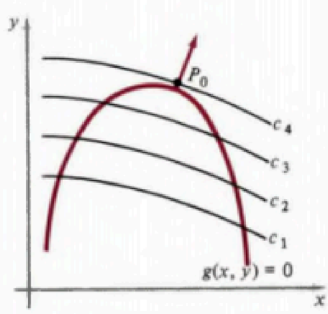


Figure 19.16

$$\text{grad}\, f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

and

$$\text{grad}\, g = \frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j}$$

are normal to these curves, and are therefore parallel to each other at the point P_0. Hence one vector is a scalar multiple of the other at P_0, that is,

$$\text{grad}\, f = \lambda\, \text{grad}\, g \tag{4}$$

for some number λ. (This argument assumes that grad $g \neq \mathbf{0}$ at P_0, so that the curve $g(x, y) = 0$ actually has a tangent at this point.)

*The great physicist Einstein once said—probably in a fit of impatience with mathematicians and their ways—that "Elegance is for tailors," but he was wrong. For mathematicians and theoretical physicists alike, the aesthetic factor in their thinking is as indispensable as the senses of taste and smell are for a master chef.

The vector equation (4), together with $g(x, y) = 0$, yields the three scalar equations

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \qquad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \qquad g(x, y) = 0. \tag{5}$$

Accordingly, we have three equations that we can try to solve simultaneously for the three unknowns x, y, and λ. The points (x, y) that we find are the only possible locations for the maximum (or minimum) values of $f(x, y)$ with the constraint $g(x, y) = 0$. The corresponding values of λ may emerge from the process of solving (5), but they are usually not of much interest to us. The final step is to calculate the value of $f(x, y)$ at each of the solution points (x, y) in order to distinguish maximum values from minimum values.

The *method of Lagrange multipliers* is simply the following handy device for obtaining equations (5): Define a function $L(x, y, \lambda)$ of the three variables x, y, and λ by

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y), \tag{6}$$

and observe that equations (5) are equivalent, in the same order, to the equations

$$\frac{\partial L}{\partial x} = 0, \qquad \frac{\partial L}{\partial y} = 0, \qquad \frac{\partial L}{\partial \lambda} = 0. \tag{7}$$

The variable λ is called the *Lagrange multiplier*. Thus, to find the *constrained* maximum or minimum values of $f(x, y)$ with the constraint $g(x, y) = 0$, we look for the *unconstrained* (or *free*) maximum or minimum values of the function L defined by (6). We emphasize that this method has two major features that can be of practical value, and are often important for theoretical work: It does not disturb the symmetry of the problem by making an arbitrary choice of the independent variable, and it removes the constraint at the small expense of introducing λ as another variable.

1. 标题与内容概述

《拉格朗日乘数法：在约束条件下求多变量函数的极值》

前面我们讨论了多变量函数 $f(x, y)$（或更多变量）在**无额外约束**下如何寻找极值：令偏导数同时为 0 即可找出临界点，然后结合二阶偏导数等方法判断极值性质。而在很多实际问题中，自变量还要满足某个（或多个）**约束条件**，例如 $g(x, y) = 0$。这时候，拉格朗日乘数法（Lagrange multipliers）为我们提供了一种**优雅而普适**的求解思路：

1. 将约束方程 $g(x, y) = 0$ 视为一条曲线（或更高维的超曲面）；
2. 在满足约束的同时，要使 $f(x, y)$ 达到最大或最小值。
3. 几何解释表明：极值点处，∇f 与 ∇g **平行**（相差一个常数因子），这就是拉格朗日乘数法的核心原理。

---

2. 详细内容解析

下面将按逻辑顺序介绍拉格朗日乘数法的几何直观、核心方程，以及应用示例，帮助你从零开始理解该方法的用途与实现步骤。

2.1 约束极值问题的提出

1. **无约束 vs. 有约束**
   - **无约束**：在平面或空间中自由地找 f 的极大、极小值，只需令各偏导数为 0 即可。

- **有约束**：如今假设 (x,y) 并非任意取，而是要满足某个条件 $g(x,y)=0$ 。这在几何上意味着 (x,y) 必须在曲线 $g(x,y)=0$ 上移动。

2. **示例：半圆中矩形最大面积**
   - 我们想要最大化矩形面积 $A=2xy$ ，前提是矩形顶点 (x,y) 落在半径为 a 的半圆上，即约束 $x^2+y^2=a^2$ （并且 $y\geq 0$ ）。
   - 过去的做法是用约束表达 y 为 $\sqrt{a^2-x^2}$ ，然后代入 A 求导。但若约束方程复杂，很难解出表达式时就不方便。

2.2 拉格朗日乘数法的几何基础

1. **等高线（水平曲线）与约束曲线**
   - 若我们画出 $f(x,y)=c$ 的一族等值曲线，以及 $g(x,y)=0$ 的曲线，在某点 P_0 两者“恰好相切”——即形成极值。
   - 从微分几何的角度，曲线“相切”意味着它们有**相同法线方向。**
   - 在函数语言里，“曲线法线方向”就是各自梯度： ∇f 与 ∇g 。若两条曲线在点 P_0 处相切，则 ∇f 与 ∇g **平行**于同一条法线。
2. **向量形式**
   - $\nabla f(x_0,y_0)$ 和 $\nabla g(x_0,y_0)$ 同方向，即

     $$\nabla f(x_0,y_0)=\lambda\,\nabla g(x_0,y_0)$$

     对某个常数 λ （这就是“拉格朗日乘数”）。
   - 同时还需满足约束： $g(x_0,y_0)=0$ 。
   - 这样，我们得到三元方程组：

     $$\begin{cases} f_x=\lambda\,g_x, \\ f_y=\lambda\,g_y, \\ g(x,y)=0. \end{cases}$$

2.3 拉格朗日乘数法的一般公式

1. **三元方程组**
   - 用记号写成：

     $$\frac{\partial f}{\partial x}=\lambda\frac{\partial g}{\partial x},\quad \frac{\partial f}{\partial y}=\lambda\frac{\partial g}{\partial y},\quad g(x,y)=0. \tag{5}$$

   - 这里 λ 是未知数之一，与 x,y 同时求解。
2. **拉格朗日函数** $L(x,y,\lambda)$
   - 有人喜欢将上面方程组写为：

     $$L(x,y,\lambda)=f(x,y)\;-\;\lambda\,g(x,y).$$

     然后令

     $$\frac{\partial L}{\partial x}=0,\quad \frac{\partial L}{\partial y}=0,\quad \frac{\partial L}{\partial \lambda}=0,$$

     它们与 (5) 等价。
   - 解得所有 (x,y,λ) 后，把满足约束的 (x,y) 代入 f 计算即可，判断最大值或最小值。
3. **适用情形扩展**
   - 更高维：若 $f(x_1,\ldots,x_n)$ 在约束 $g(x_1,\ldots,x_n)=0$ 下求极值，同理可写

     $$\nabla f=\lambda\nabla g,\quad g=0.$$

   - 多重约束：若有 $g_1=0,\,g_2=0,\ldots$ ，就引入多个拉格朗日乘数 $\lambda_1,\lambda_2,\ldots$ ，形成更大方程组。

2.4 示例：最大矩形在半圆中的面积

> **题目：** 寻找半圆 $x^2+y^2=a^2$ （$y\ge 0$）中所能内接的最大矩形（对称于 x -轴）的宽和高。

1. **方法一（传统）：**
   - 约束：$x^2+y^2=a^2 \implies y=\sqrt{a^2-x^2}$。
   - 面积函数：$A=2xy=2x\sqrt{a^2-x^2}$。
   - 用单变量方法解即可得到 $x=\frac{a}{\sqrt{2}}$，$y=\frac{a}{\sqrt{2}}$。
2. **方法二（拉格朗日乘数）**
   - 令 $f(x,y)=2xy,\ g(x,y)=x^2+y^2-a^2=0$。
   - 则方程

$$f_x=2y=\lambda\cdot 2x,\quad f_y=2x=\lambda\cdot 2y,\quad x^2+y^2=a^2.$$

   - 从 $2y=\lambda\cdot 2x$ 得 $\lambda=\frac{y}{x}$；从 $2x=\lambda\cdot 2y$ 得 $\lambda=\frac{x}{y}$。合并可知 $\frac{y}{x}=\frac{x}{y} \implies x^2=y^2 \implies x=\pm y$（考虑正半圆则 $x=y>0$）。
   - 再由约束 $x^2+x^2=a^2 \implies x=y=\frac{a}{\sqrt{2}}$。
   - 面积即 $A_{\max}=2x^2=2\cdot\frac{a^2}{2}=a^2\sqrt{2}$（或更精确地 $2\cdot\frac{a}{\sqrt{2}}\cdot\frac{a}{\sqrt{2}}=a^2$）。

此方法不必先解 y 为 $\sqrt{a^2-x^2}$，更可推广到复杂情形中。

2.5 方法优点与后续步骤

1. **方法优点**
   - 无需先行解出约束 $\Rightarrow$ 避免可能的代数难题。
   - 在多元或多个约束下仍能使用同一思路，十分通用。
   - 幾何解释清晰：极值处梯度向量彼此平行。
2. **找到候选点后**
   - 需判定是否为极大值或极小值：可以借助二阶偏导、或者直接在应用背景下判断（如几何直觉、测试函数值大小等）。

---

3. 本次内容小结

- **约束极值的必要条件**：在 (x_0,y_0) 处，∇f 与 ∇g 必须平行，且约束 $g(x_0,y_0)=0$ 要成立。
- **拉格朗日乘数** λ：由 $\nabla f=\lambda\nabla g$ 引入，是求解该方程组的辅助变量。
- **拉格朗日函数**：$L(x,y,\lambda)=f(x,y)-\lambda\, g(x,y)$；令 $\partial L/\partial x=0$、$\partial L/\partial y=0$、$\partial L/\partial\lambda=0$ 即可求所有解。
- **应用范围**：
  1. 优化问题（极大值、极小值）常见于物理、经济、工程；
  2. 多约束则多乘数；
  3. 同样原理可推广到更高维情形。

---

4. 报错或补充说明

- 截至目前，已对本节截图中的主要概念和公式（拉格朗日乘数法及其几何意义、方程组写法、示例等）做了完整讲解。如有任何图片或文字内容无法辨认，请告诉我；如无，则本次笔记便涵盖了本节的核心精华。

19.8.2

Example 1 (continued) To solve the inscribed rectangle problem by this new method, we first express the constraint (3) in the form $x^2 + y^2 - a^2 = 0$ and then write down the function

$$L = 2xy - \lambda(x^2 + y^2 - a^2).$$

The equations (7) are

$$\frac{\partial L}{\partial x} = 2y - 2\lambda x = 0, \tag{8}$$

$$\frac{\partial L}{\partial y} = 2x - 2\lambda y = 0, \tag{9}$$

$$\frac{\partial L}{\partial \lambda} = -(x^2 + y^2 - a^2) = 0. \tag{10}$$

Equations (8) and (9) yield $y = \lambda x$ and $x = \lambda y$, and substituting in (10) gives

$$\lambda^2(x^2 + y^2) = a^2.$$

But (10) tells us that $x^2 + y^2 = a^2$, so $\lambda^2 = 1$ and $\lambda = \pm 1$. The value $\lambda = -1$ would imply that $y = -x$, which is impossible because both x and y are positive numbers, so $\lambda = 1$ and $y = x$. This gives the shape of the largest inscribed rectangle, namely, twice as long as it is wide, because

$$\text{length} = 2x = 2y = 2(\text{width}).$$

If we want the actual dimensions of this largest rectangle, we substitute $y = x$ into $x^2 + y^2 = a^2$ to find that $x = y = \frac{1}{2}\sqrt{2}a$, so the length $= 2x = \sqrt{2}a$ and the width $= y = \frac{1}{2}\sqrt{2}a$.

One of the merits of the method of Lagrange multipliers is that it extends very easily to situations with more variables or more constraints. For instance, to maximize $f(x, y, z)$ subject to the constraint $g(x, y, z) = 0$, the gradient

$$\text{grad } f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}$$

must be normal to the surface $g(x, y, z) = 0$ (Fig. 19.17), so grad f must be parallel to grad g, and again we have

$$\text{grad } f = \lambda \text{ grad } g.$$

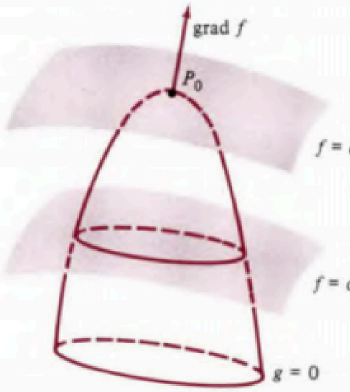


Figure 19.17

The four equations

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \qquad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \qquad \frac{\partial f}{\partial z} = \lambda \frac{\partial g}{\partial z}, \qquad g(x, y, z) = 0,$$

in the four unknowns x, y, z, λ are again equivalent to the simpler equations

$$\frac{\partial L}{\partial x} = 0, \qquad \frac{\partial L}{\partial y} = 0, \qquad \frac{\partial L}{\partial z} = 0, \qquad \frac{\partial L}{\partial \lambda} = 0,$$

where $L = f(x, y, z) - \lambda g(x, y, z)$.

Similarly, suppose we want to maximize or minimize $f(x, y, z)$ subject to *two* constraints $g(x, y, z) = 0$ and $h(x, y, z) = 0$. Each constraint defines a surface, and in general these two surfaces have a curve of intersection. As before, a point P_0 where $f(x, y, z)$ has a maximum or minimum value on this curve is a point where a level surface of f is tangent to the curve, that is, a point where grad f is normal to the curve (Fig. 19.18). But the vectors grad g and grad h determine the normal plane to the curve at P_0, and since grad f lies in this plane, there must be scalars λ and μ (*two* Lagrange multipliers this time) with the property that

$$\text{grad } f = \lambda \text{ grad } g + \mu \text{ grad } h.$$

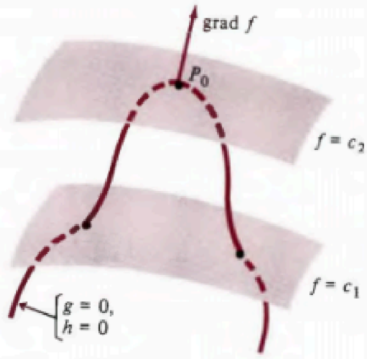


Figure 19.18

(This argument assumes that grad $g \neq \mathbf{0}$ and grad $h \neq \mathbf{0}$, and that these vectors are not parallel.) Just as before, this vector equation and the two constraint equations are easily seen to be equivalent to the following five equations in five unknowns:

$$\frac{\partial L}{\partial x} = 0, \qquad \frac{\partial L}{\partial y} = 0, \qquad \frac{\partial L}{\partial z} = 0, \qquad \frac{\partial L}{\partial \lambda} = 0, \qquad \frac{\partial L}{\partial \mu} = 0,$$

where $L = f(x, y, z) - \lambda g(x, y, z) - \mu h(x, y, z)$.

We illustrate these methods with two examples.

Example 2 Find the point on the plane $x + 2y + 3z = 6$ that is closest to the origin.

Solution We want to minimize the distance $\sqrt{x^2 + y^2 + z^2}$ subject to the constraint $x + 2y + 3z - 6 = 0$. If the distance is a minimum, its square is a minimum, so we simplify the calculations a bit by minimizing $x^2 + y^2 + z^2$ with the same constraint. Let

$$L = x^2 + y^2 + z^2 - \lambda(x + 2y + 3z - 6).$$

Then the equations we must solve are

$$\frac{\partial L}{\partial x} = 2x - \lambda = 0,$$

$$\frac{\partial L}{\partial y} = 2y - 2\lambda = 0,$$

$$\frac{\partial L}{\partial z} = 2z - 3\lambda = 0,$$

$$\frac{\partial L}{\partial \lambda} = -(x + 2y + 3z - 6) = 0.$$

Substituting the values for x, y, and z from the first three equations into the fourth gives

$$\tfrac{1}{2}\lambda + 2\lambda + \tfrac{9}{2}\lambda = 6 \qquad \text{or} \qquad \tfrac{14}{2}\lambda = 6 \qquad \text{or} \qquad \lambda = \tfrac{6}{7}.$$

It now follows that $x = \frac{3}{7}$, $y = \frac{6}{7}$, and $z = \frac{9}{7}$, so the desired point is $(\frac{3}{7}, \frac{6}{7}, \frac{9}{7})$.

Example 3 Find the point on the line of intersection of the planes $x + y + z = 1$ and $3x + 2y + z = 6$ that is closest to the origin.

This time we want to minimize $x^2 + y^2 + z^2$ subject to the two constraints $x + y + z - 1 = 0$ and $3x + 2y + z - 6 = 0$. If we write

$$L = x^2 + y^2 + z^2 - \lambda(x + y + z - 1) - \mu(3x + 2y + z - 6),$$

then our equations are

$$\frac{\partial L}{\partial x} = 2x - \lambda - 3\mu = 0,$$

$$\frac{\partial L}{\partial y} = 2y - \lambda - 2\mu = 0,$$

$$\frac{\partial L}{\partial z} = 2z - \lambda - \mu = 0,$$

$$\frac{\partial L}{\partial \lambda} = -(x + y + z - 1) = 0,$$

$$\frac{\partial L}{\partial \mu} = -(3x + 2y + z - 6) = 0.$$

The first three equations give

$$x = \tfrac{1}{2}(\lambda + 3\mu), \qquad y = \tfrac{1}{2}(\lambda + 2\mu), \qquad z = \tfrac{1}{2}(\lambda + \mu).$$

When these expressions are substituted in the fourth and fifth equations and the results are simplified, we get

$$3\lambda + 6\mu = 2,$$

$$3\lambda + 7\mu = 6,$$

so $\mu = 4$ and $\lambda = -\frac{22}{3}$. These values give $x = \frac{7}{3}$, $y = \frac{1}{3}$, $z = -\frac{5}{3}$, so the desired point is $(\frac{7}{3}, \frac{1}{3}, -\frac{5}{3})$.

Remark 1 In economics, Lagrange multipliers are used to analyze the problem of maximizing the total production of a manufacturing firm subject to the constraint of fixed available resources. For example, let

$$P = f(x, y) = Ax^{\alpha}y^{\beta}, \qquad \alpha + \beta = 1,$$

be the production (measured in dollars) resulting from x units of capital and y units of labor. Then this function—known to economists as the *Cobb-Douglas production function*—is homogeneous of degree 1 in the sense explained in Example 2 of Section 19.6. If the cost of each unit of capital is a dollars, and of each unit of labor is b dollars, and if a total of c dollars is available to cover the combined costs of capital and labor, then we want to maximize the production $P = f(x, y)$ subject to the constraint $ax + by = c$. In Problem 23 we ask students to show that production is maximized when $x = \alpha c/a$ and $y = \beta c/b$.*

Remark 2 At the beginning of this section we said that Lagrange multipliers also have applications to differential geometry and advanced theoretical mechanics. These applications are too complicated to describe here, but the details can be found on pp. 521–523 and 529 of the present writer's text, *Differential Equations*, 2nd ed. (McGraw-Hill, 1991).

*For further details, interested students can look up Cobb-Douglas production functions in the indexes of the books mentioned in the first footnote of Section 19.6.

1. 标题与内容概述

《拉格朗日乘数法的应用示例与多重约束：从二元到三元，再到经济学中的意义》

在上一节已经介绍了拉格朗日乘数法的基本原理和单个约束的方程组写法。本次内容则通过几个具体示例展示如何利用**拉格朗日乘数法**在更复杂的场景下求解带一个或多个约束的极值问题，包括三维空间下的最短距离点、两平面交线上的最短距离点，以及经济学中的生产函数最大化等。还补充说明了拉格朗日乘数法在高维、多重约束情形以及在经济学、微分几何、理论力学中的重要应用。

---

2. 详细内容解析

2.1 续示例：半圆内接矩形问题的解法

1. **简要回顾：**
   - 我们要在半圆 $x^2 + y^2 = a^2$ （且 $y \geq 0$）里，找能容纳的最大矩形。
   - 拉格朗日函数设为

     $$L(x, y, \lambda) = 2xy \; - \; \lambda\,(x^2 + y^2 - a^2).$$

   - 解方程

     $$\begin{cases} L_x = 2y - 2\lambda x = 0, \\ L_y = 2x - 2\lambda y = 0, \\ L_\lambda = -(x^2 + y^2 - a^2) = 0, \end{cases}$$

     即
     $2y = 2\lambda x,\ 2x = 2\lambda y,\ x^2 + y^2 = a^2.$
2. **求得结果：**

- 由前两式可知 $y=\lambda x$ 与 $x=\lambda y$，合并得 $x=y$ 或 $\lambda^2=1$。
- 约束给出 $x^2+x^2=a^2 \Rightarrow x=y=\frac{a}{\sqrt{2}}$。
- 因而矩形长宽比 $2x:2y=2:2=1:1$ （即边长相等），各边长 $\frac{a}{\sqrt{2}}$ 对应的半长是 $\sqrt{2}\,\frac{a}{2}$。故该矩形为“高=宽”，或者说长=2x, 宽=2y, 都为 $\sqrt{2}a$ 的半量度等。
- 细化：若题意要求“宽”为 x 方向尺寸，则长 = 2x = $\sqrt{2}a$，高 = 2y 同理。

（课本也提到另一种表达：$x=\frac{a}{\sqrt{2}}, y=\frac{a}{\sqrt{2}}$，长度 = $2x=\sqrt{2}a$，宽度 = $2y=\sqrt{2}a$。两边相等说明内接矩形其实是个正方形（但放置在半圆中）。）

2.2 三维情形示例

示例：最接近平面的点

1. **问题：** 寻找与平面 $x+2y+3z=6$ 距离最近的点（最近即最短的欧几里得距离）。
2. **函数与约束：**
   - 目标函数：最小化距离 $\sqrt{x^2+y^2+z^2}$。
   - 通常等效地最小化其平方 $x^2+y^2+z^2$ （更简便）；
   - 约束：$g(x,y,z)=x+2y+3z-6=0$。
3. **拉格朗日函数：**

$$L(x,y,z,\lambda)=x^2+y^2+z^2\ -\ \lambda\,(x+2y+3z-6).$$

   - 设各偏导=0：

$$\frac{\partial L}{\partial x}=2x-\lambda=0,\quad \frac{\partial L}{\partial y}=2y-2\lambda=0,\quad \frac{\partial L}{\partial z}=2z-3\lambda=0,\quad \frac{\partial L}{\partial \lambda}=-(x+2y+3z-6)=0.$$

4. **解方程组：**
   - 从前 3 个方程可得 $x=\frac{\lambda}{2},\ y=\lambda,\ z=\frac{3\lambda}{2}.$
   - 约束：$\frac{\lambda}{2}+2\lambda+\frac{9\lambda}{2}=6 \implies \frac{1}{2}\lambda+2\lambda+\frac{9}{2}\lambda=6 \implies \frac{14}{2}\lambda=6 \implies \lambda=\frac{6}{7}.$
   - 故 $x=\frac{3}{7},\ y=\frac{6}{7},\ z=\frac{9}{7}.$
   - 这点即与平面最近，距离则是 $\sqrt{\left(\frac{3}{7}\right)^2+\left(\frac{6}{7}\right)^2+\left(\frac{9}{7}\right)^2}$ （如题只问坐标就停止）。

2.3 多重约束场景

示例：与交线最近的点

1. **问题：** 若有两平面：$x+y+z=1,\ 3x+2y+z=6$。它们相交成一条直线。问哪点在此交线上离原点最近?
2. **要做的：** 在三维里，最小化 $x^2+y^2+z^2$，带约束

$$g(x,y,z)=x+y+z-1=0,\quad h(x,y,z)=3x+2y+z-6=0.$$

3. **拉格朗日函数**

$$L(x,y,z,\lambda,\mu)=x^2+y^2+z^2-\lambda(x+y+z-1)-\mu(3x+2y+z-6).$$

4. **方程组：**
   - $\frac{\partial L}{\partial x}=2x-\lambda-3\mu=0$,
   - $\frac{\partial L}{\partial y}=2y-\lambda-2\mu=0$,
   - $\frac{\partial L}{\partial z}=2z-\lambda-\mu=0$,
   - $\frac{\partial L}{\partial \lambda}=-(x+y+z-1)=0$,
   - $\frac{\partial L}{\partial \mu}=-(3x+2y+z-6)=0$.

5. **解出** (x,y,z,λ,μ)
    - 先利用前 3 个方程表达 x,y,z 的线性关系；再带入后 2 个约束；可解得最终坐标。
    - 书中结果是 $(x,y,z)=\left(\frac{7}{3},\,\frac{1}{3},\,-\frac{5}{3}\right)$。

2.4 经济学背景：Cobb–Douglas 生产函数

1. **Cobb–Douglas**：
    - 形式 $P(x,y)=A\,x^{\alpha}\,y^{\beta}$，其中 $\alpha+\beta=1$ 表示一阶同次（即规模报酬不变）。
    - 若资本 x 和劳动力 y 的成本分别为 a,b；给定总费用 c，则约束 $ax+by=c$。
    - 想令 P 极大，则拉格朗日函数：$L(x,y,\lambda)=A\,x^{\alpha}\,y^{\beta}-\lambda(ax+by-c)$。
    - 解出 $x=\frac{\alpha c}{a},\ y=\frac{\beta c}{b}$。这正是经济学中"要素最佳投入"之解。

2.5 总结与优势

- **拉格朗日乘数法**：只要我们能表示约束为 $g(x_1,\ldots,x_n)=0$，就令 $\nabla f=\lambda\nabla g$，加上 $g=0$，再解方程组，得到候选点；可扩展到多个约束时则增加相应拉格朗日乘数。
- **广泛用途**：在几何（找最短、最大等）、物理（稳定配置、极值原理）、经济学（最优生产、资源配置）等众多领域发挥关键作用。

---

3. 本次内容小结

- **单约束情形**

$$\nabla f=\lambda\nabla g,\quad g(x,y,\ldots)=0.$$

  同时解出 $(x,y,\ldots)$ 和 λ。
- **多约束情形**

$$\nabla f=\lambda\nabla g+\mu\nabla h+\ldots,\quad g=0,\ h=0,\ldots$$

- **示例**
    1. **矩形内接于半圆**：快速求最大矩形；
    2. **平面上最短距离**：找与给定平面最近点；
    3. **两平面交线最近点**：多约束情形；
    4. **经济学：Cobb–Douglas** 生产函数最大化等。
- 之后根据需要再判断是否是最大或最小，并结合背景知识或二阶条件来分类结果。

---

4. 报错或补充说明

- 以上笔记已涵盖截图中的主要内容：拉格朗日乘数法用于约束极值的一系列方程、例题的求解过程，以及在经济学中的简单示例。若有任何文字或方程难以辨认，请告知；无则本次讲解即完整。

19.9

19.9 (OPTIONAL) LAPLACE'S EQUATION, THE HEAT EQUATION, AND THE WAVE EQUATION. LAPLACE AND FOURIER

A very large part of mathematical physics is concerned with three classic partial differential equations: *Laplace's equation*,

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} = 0; \tag{1}$$

the *heat equation*,

$$a^2\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right) = \frac{\partial w}{\partial t}; \tag{2}$$

and the *wave equation*,

$$a^2\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right) = \frac{\partial^2 w}{\partial t^2}. \tag{3}$$

As the notation indicates, in (2) and (3) the variable w is understood to be a function of the time t and the three space coordinates x, y, z of a point P, and in (1) w depends only on x, y, z and is independent of t. The quantity a is a constant. Each of our three equations also has simpler two- and one-dimensional versions, depending on whether two space coordinates are present, or only one. Thus,

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = 0 \tag{4}$$

is the two-dimensional Laplace equation, and

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial w}{\partial t} \quad \text{and} \quad a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial^2 w}{\partial t^2}$$

are the one-dimensional heat equation and wave equation.

A full study of these equations can occupy years, because their physical meaning is extraordinarily rich, and also much concentrated thought is necessary to master the various branches of advanced mathematics that are needed to solve and interpret them. In this section we consider several aspects of these equations that do not require too much technical background.

LAPLACE'S EQUATION

If a number of particles of masses $m_1, m_2, \ldots, m_n$, attracting according to the inverse square law of gravitation, are placed at points $P_1, P_2, \ldots, P_n$, then the *potential* due to these particles at any point P (that is, the work done against their attractive forces in moving a unit mass from P to an infinite distance) is

$$w = \frac{Gm_1}{PP_1} + \frac{Gm_2}{PP_2} + \cdots + \frac{Gm_n}{PP_n}, \tag{5}$$

where G is the gravitational constant.* If the points $P, P_1, P_2, \ldots, P_n$ have rectangular coordinates $(x, y, z), (x_1, y_1, z_1), (x_2, y_2, z_2), \ldots, (x_n, y_n, z_n)$, so that

$$PP_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2},$$

with similar expressions for the other distances, then it is quite easy to verify that the potential w satisfies Laplace's equation (1). This equation does not involve either the particular masses or the coordinates of the points at which they are located, so it is satisfied by the potential produced in empty space by an arbitrary discrete or continuous distribution of mass.

The function w defined by (5) is called a *gravitational potential.* If we work instead with electrically charged particles of charges $q_1, q_2, \ldots, q_n$, then their *electrostatic potential* has the same form as (5) with the m's replaced by q's and G by Coulomb's constant, so it also satisfies Laplace's equation. In fact, this equation has such a wide variety of applications that its study is a branch of mathematics in its own right, known as *potential theory.*

1. 标题与内容概述

《拉普拉斯方程、热方程与波动方程——物理与数学中的经典偏微分方程》

在物理学、工程学与更广泛的自然科学里，三个非常重要的偏微分方程（PDE）被大量研究和应用，分别是：

1. **拉普拉斯方程（Laplace's equation）**
2. **热方程（heat equation）**
3. **波动方程（wave equation）**

本节简要介绍它们的形式、部分物理背景，以及一些低维简化形式。对它们的深入研究往往需要更多高等数学工具，但这里给出其核心思想与部分示例应用。

---

2. 详细内容解析

2.1 三个经典方程的基本形式

1. **拉普拉斯方程**

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} = 0. \tag{1}$$

   - 这里 $w = w(x, y, z)$ 只与空间坐标 (x, y, z) 有关，不依赖时间 t。
   - 若只在二维平面上，则对应简化为

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = 0, \tag{4(a)}$$

   称为**二维拉普拉斯方程**；如果再简化到一维，就变成 $\frac{d^2 w}{dx^2} = 0$。

2. **热方程**

$$a^2\Big(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\Big) = \frac{\partial w}{\partial t}. \tag{2}$$

   - 其中 $w = w(x, y, z, t)$ 表示温度（或类似扩散量）随时间、空间的变化；a^2 是扩散系数或导热系数的常量。
   - 若只在一维空间，则方程变为

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial w}{\partial t}. \tag{4(b)}$$

3. **波动方程**

$$a^2\Big(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\Big) = \frac{\partial^2 w}{\partial t^2}. \tag{3}$$

   - 这里 $w = w(x, y, z, t)$ 常表示波的位移（如弦振动、声波、波浪等）随时间、空间的演化；a 是波速。
   - 若一维场景，可写为

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial^2 w}{\partial t^2}. \tag{4(c)}$$

2.2 物理背景简介

1. **拉普拉斯方程：静态势场**
   - 在静态条件下的势函数（如引力势、静电势）往往满足 $\nabla^2 w = 0$（即拉普拉斯方程）。
   - 例如，引力势能或静电势能若只在空域分布并无时间演化，在没有分布质量或电荷的空域内，就满足 $\nabla^2 w = 0$。
   - 当质量/电荷分布存在时，则会得到更一般的泊松方程 $\nabla^2 w = \rho$，可视为拉普拉斯方程的推广。
2. **热方程：导热与扩散**
   - 描述热量在介质中的流动过程或浓度扩散过程：温度（或浓度）随时间和空间坐标的变化。
   - 对应物理图景中"温度会从热区向冷区逐渐传导"，使得整体趋于平衡。
3. **波动方程：传播与振动**
   - 若一个介质可以传播扰动（例如声波、弦振动、电磁波等），其位移/压力/场强随空间、时间变化满足类似的二阶时间导数方程。
   - 线性波动现象常满足 $\frac{\partial^2 w}{\partial t^2} = a^2 \nabla^2 w$，其中 a 是波在介质中的传播速度。

2.3 低维简化版

正如上式 (4) 所示，每个方程都可在二维或一维中再简化：

- **二维拉普拉斯**：$\partial^2 w/\partial x^2 + \partial^2 w/\partial y^2 = 0$；
- **一维热方程**：$a^2\,\partial^2 w/\partial x^2 = \partial w/\partial t$；
- **一维波动方程**：$a^2\,\partial^2 w/\partial x^2 = \partial^2 w/\partial t^2$ 。

在很多基础物理模型和入门教材中，这些低维形式也是人们初次接触 PDE 的重要例子。

2.4 拉普拉斯方程的典型示例：引力势或静电势

1. **引力势的形式**
   - 若空间内有若干点质量 $m_1, m_2, \ldots, m_n$，其引力势能函数（针对单位质量）可写为

     $$w(P) = \frac{Gm_1}{PP_1} + \frac{Gm_2}{PP_2} + \cdots + \frac{Gm_n}{PP_n},$$

     其中 PP_i 表示点 P 与质量 m_i 的距离；G 为引力常数。
   - 在空域中无质量分布处，这个势函数会满足 $\nabla^2 w = 0$（拉普拉斯方程）。
2. **电势的类似性**
   - 把质量 m_i 换成电荷 q_i，引力常数 G 换成库仑常数，就得到静电势。该势在空域无电荷处也同样满足拉普拉斯方程。
   - 这表明**拉普拉斯方程**在描述引力势场或静电势场等物理系统上有广泛应用，相关理论发展成“**势理论**(potential theory)**”。

2.5 学习和研究的延伸方向

1. **数值方法与边界条件**
   - 实际应用中，这些方程通常需要配合边界条件/初始条件，方能求得独特解。
   - 常见方法有分离变量法、傅里叶级数、有限差分、有限元等。
   - 每个方程都有丰富而成熟的数学理论，例如热方程的平滑性与向后不稳定性，波动方程的因果性与达朗贝尔解，拉普拉斯方程的调和函数性质等。
2. **高等数学分支**
   - 这些 PDE 研究衍生出庞大的分支：如调和分析、傅里叶分析、Sobolev 空间理论、椭圆型/抛物型/双曲型方程的一般解理论等。
3. **实际应用**
   - **拉普拉斯/泊松方程**：电磁场、流体静力学、引力势、图像处理(如图像平滑/修复)。
   - **热方程**：热传导、扩散过程、财务数学中的 Black–Scholes 方程也有相似结构。
   - **波动方程**：声学、光学、量子力学（薛定谔方程与之类似但复数形式）、地震波、弦振动等。

---

3. 本次内容小结

- **三大方程**：
  1. **拉普拉斯方程** $\nabla^2 w = 0$：静态势场/调和函数；
  2. **热方程** $a^2\nabla^2 w = \frac{\partial w}{\partial t}$：扩散/热传导；
  3. **波动方程** $a^2\nabla^2 w = \frac{\partial^2 w}{\partial t^2}$：振动/波动现象。
- **低维形式 (4)**：在一维、二维情形下，这些方程可以写成更加简单的 PDE。
- **Laplace 方程的例子**：点质量引力势、点电荷静电势在无源区满足拉普拉斯方程；在存在质量或电荷密度时，则满足泊松方程。
- **后续需要更深入的数学工具**来解这些方程，尤其是边界值或初值问题，这里仅作概述性质的介绍。

---

4. 报错或补充说明

- 已针对课件的最后一节内容，对“拉普拉斯方程、热方程、波动方程”之形式和例子做了完整介绍；若有额外图表或文字需进一步解读，请告知；如无，则以上为本次笔记的全部。

19.9.2

THE HEAT EQUATION

When we study the flow of heat in thermally conducting bodies, we encounter an entirely different type of problem leading to a partial differential equation.

---

*See Example 2 in Section 7.7. In this example we show that if two particles of masses M and m are separated by a distance a, then the work done in separating them to an infinite distance is GMm/a.

In the interior of a body where heat is flowing from one region to another, the temperature generally varies from point to point at any one time, and from time to time at any one point. Thus, the temperature w is a function of the space coordinates x, y, z and the time t, say $w = f(x, y, z, t)$. The precise form of this function naturally depends on the shape of the body, the thermal characteristics of its material, the initial distribution of temperature, and the conditions maintained on the surface of the body. The French physicist-mathematician Fourier studied this problem in his classic treatise of 1822, *Théorie Analytique de la Chaleur (Analytic Theory of Heat)*. He used physical principles to show that the temperature function w must satisfy the heat equation (2).* We shall retrace his reasoning in a simple one-dimensional situation, and thereby derive the one-dimensional heat equation.

The following are the physical principles that will be needed.

(a) Heat flows in the direction of decreasing temperature, that is, from hot regions to cold regions.

(b) The rate at which heat flows across an area is proportional to the area and to the rate of change of temperature with respect to distance in a direction perpendicular to the area. (This proportionality factor is denoted by k and called the *thermal conductivity* of the substance.)

(c) The quantity of heat gained or lost by a body when its temperature changes, that is, the change in its thermal energy, is proportional to the mass of the body and to the change of temperature. (This proportionality factor is denoted by c and called the *specific heat* of the substance.)

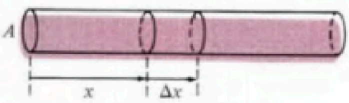


Figure 19.19

We now consider the flow of heat in a thin cylindrical rod of cross-sectional area A (Fig. 19.19) whose lateral surface is perfectly insulated so that no heat flows through it. This use of the word "thin" means that the temperature is assumed to be uniform on any cross section, and is therefore a function only of the time and the position of the cross section, say $w = f(x, t)$. We examine the rate of change of the heat contained in a thin slice of the rod between the positions x and $x + \Delta x$.

If ρ is the density of the rod, that is, its mass per unit volume, then the mass of the slice is

$$\Delta m = \rho A \, \Delta x.$$

Furthermore, if Δw is the temperature change at the point x in a small time interval Δt, then (c) tells us that the quantity of heat stored in the slice in this time interval is

$$\Delta H = c \, \Delta m \, \Delta w = c\rho A \, \Delta x \, \Delta w,$$

so the rate at which heat is being stored is approximately

$$\frac{\Delta H}{\Delta t} = c\rho A \, \Delta x \, \frac{\Delta w}{\Delta t}. \tag{6}$$

We assume that no heat is generated inside the slice—for instance, by chemical or electrical processes—so that the slice gains heat only by means of the flow

*The same partial differential equation also describes a more general class of diffusion processes, and is sometimes called the *diffusion equation*.

of heat through its faces. By (b) the rate at which heat flows into the slice through the left face is

$$-kA \left.\frac{\partial w}{\partial x}\right|_x.$$

The negative sign here is chosen in accordance with (a), so that this quantity will be positive if $\partial w/\partial x$ is negative. Similarly, the rate at which heat flows into the slice through the right face is

$$kA \left.\frac{\partial w}{\partial x}\right|_{x+\Delta x},$$

so the total rate at which heat flows into the slice is

$$kA \left.\frac{\partial w}{\partial x}\right|_{x+\Delta x} - kA \left.\frac{\partial w}{\partial x}\right|_x. \tag{7}$$

If we equate the expressions (6) and (7), the result is

$$kA \left.\frac{\partial w}{\partial x}\right|_{x+\Delta x} - kA \left.\frac{\partial w}{\partial x}\right|_x = c\rho A\, \Delta x \frac{\Delta w}{\Delta t},$$

or

$$\frac{k}{c\rho}\left[\frac{\partial w/\partial x|_{x+\Delta x} - \partial w/\partial x|_x}{\Delta x}\right] = \frac{\Delta w}{\Delta t}.$$

Finally, by letting Δx and $\Delta t \to 0$ we obtain the desired equation,

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial w}{\partial t},$$

where $a^2 = k/c\rho$. This is the physical reasoning that leads to the one-dimensional heat equation. The three-dimensional equation (2) can be derived in essentially the same way.

1. 标题与内容概述

《热方程的物理推导——从热流原理到一维偏微分方程》

本小节深入展示了**热方程（heat equation）**在一维情形下的推导思路。它建立在物理上的若干关键假设之上：

- 热量流动总是从温度较高的区域流向温度较低的区域；
- 单位时间内通过某截面的导热流速与截面积和温度梯度成正比（传导系数 k）；
- 温度变化所需的热量与物体质量及其比热容 c 成正比。

通过对细薄环形截面（或称“切片”）的能量平衡分析，可以导出一维的热方程：

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial w}{\partial t},$$

其中 $a^2 = \frac{k}{c\,\rho}$（若截面积 A 归入常数或将 ρA 视为“单位面积”密度，也可写作 $a^2 = \frac{k}{c\,\rho}$）。

---

2. 详细内容解析

以下按照课件顺序，结合物理假设依次说明一维导热方程的推导过程。

2.1 物理背景与基本假设

1. **温度函数与热流方向**
   - 在一个导热介质内部，不同位置、不同时间的温度可记为 $w(x,t)$ 。
   - 根据热传导的基本规律：热量从高温处向低温处流动，故若温度沿 x 轴递减，则热流也沿负 x 方向（或反之）。
2. **主要比例系数**
   - **热传导系数** k ：描述导热能力；热流量与温度梯度成正比。
   - **比热容** c ：描述物体在单位温度变化所需要吸收/放出的热量与质量的关系。
   - **密度** ρ ：物体单位体积的质量。
   - 实际计算中，这些参数均可看作常数（对均匀材料）。
3. **无热源假设**
   - 在推导中假设没有其他热源或外部功输入，即"切片"内的温度变化仅由热流的出入所致，不存在化学能、电能等内生热源。

2.2 对细薄截面的能量守恒分析

为方便，我们考虑如图 19.19 所示的**一维棒状**介质：

- 截面积 A ，在 x 和 $x+\Delta x$ 之间取一薄片（厚度 Δx ）。
- 温度在该薄片内可随时间、空间而变动，但在横截面上假设均匀（即横截面处温度近似相同）。

1. **热量存储变化率** $\Delta H/\Delta t$
   - 薄片质量： $\Delta m = \rho\, A\, \Delta x$ 。
   - 若在极短时间 Δt 内，温度改变 Δw ，则吸收（或放出）的热量

$$\Delta H = c\,\Delta m\,\Delta w = c\,\rho\,A\,\Delta x\,\Delta w.$$

   - 单位时间内的热量变化率：

$$\frac{\Delta H}{\Delta t} = c\,\rho\,A\,\Delta x\,\frac{\Delta w}{\Delta t}. \tag{6}$$

2. **热流量出入平衡**
   - 棒子的左右两端截面上热量流入和流出决定薄片本身温度升降。
   - 若把热流强度记为 Φ ，则由傅里叶导热定律：

$$\Phi = -\,k\,A\,\frac{\partial w}{\partial x}$$

（负号因为当 $\frac{\partial w}{\partial x} > 0$ 时，热流向负 x 方向。书中在 (7) 式用 " $-kA\,\frac{\partial w}{\partial x}\big|_x$ " 表示左端流入量，右端流入量则换符号。）
   - 左端截面（在 x 处）热流量： $-kA\frac{\partial w}{\partial x}\big|_x$ ；
   - 右端截面（在 $x+\Delta x$ 处）热流量： $+kA\frac{\partial w}{\partial x}\big|_{x+\Delta x}$ （符号相反）。
   - 因此净热流**进入**薄片的速率为：

$$kA\,\tfrac{\partial w}{\partial x}\big|_{x+\Delta x} \;-\; kA\,\tfrac{\partial w}{\partial x}\big|_x. \tag{7}$$

3. **平衡方程**
   - 令"热流入速率" = "热量变化率"，即：

$$kA\left[\tfrac{\partial w}{\partial x}\Big|_{x+\Delta x} - \tfrac{\partial w}{\partial x}\Big|_x\right] \;=\; c\,\rho\,A\,\Delta x\,\frac{\Delta w}{\Delta t}.$$

   - 整理得到

$$\frac{k}{c\,\rho}\,\frac{\frac{\partial w}{\partial x}\big|_{x+\Delta x} - \frac{\partial w}{\partial x}\big|_x}{\Delta x} \;=\; \frac{\Delta w}{\Delta t}.$$

- 当 $\Delta x \to 0$ 及 $\Delta t \to 0$，就变成偏微分形式：

2.3 推导得到的一维热方程

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial w}{\partial t}, \quad \text{其中 } a^2 = \frac{k}{c\,\rho}.$$

- 该方程说明：一维空间中温度随时间的变化速率等于温度在空间方向的二阶导数，乘以系数 a^2 。
- 当扩展到三维，形式即

$$a^2 \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} \right) = \frac{\partial w}{\partial t}.$$

---

3. 本次内容小结

- **热传导的物理原理**：热量在介质中流动与温度梯度成正比；系统内部温度变化需考虑本身质量、比热容；无外部热源时，热量改变源自热流的出入。
- **一维热方程推导**：借助对一小段棒子的热量平衡分析，利用傅里叶定律、比热容原理以及能量守恒，将热流入 - 热流出 = 储存热量变化联立得出

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial w}{\partial t}.$$

- **更高维度**：若考虑三维空间、或材料各向异性、或有内生热源等情况，就会得到相应更一般或更复杂的热方程形式。

---

4. 报错或补充说明

- 截止目前，已完整解析“热方程”在一维情形下的物理推导过程，如有其他截图内容不清晰或疑问，请告知；若无，则上述为本节笔记的全部。

19.9.3

THE WAVE EQUATION

All phenomena of wave propagation, for example, of light or sound or radio waves, are governed by the wave equation (3). We shall consider the simple case of a one-dimensional wave described by the one-dimensional wave equation

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial^2 w}{\partial t^2}. \tag{8}$$

Such a wave involves some property $w = f(x, t)$, such as the position of a particle, the intensity of an electric field, or the pressure in a column of air, that depends not only on the position x but also on the time t.

In order to understand the connection between waves and equation (8), we consider a function $w = F(x - at)$. At $t = 0$, it defines the curve $w = F(x)$, and at any later time $t = t_1$, it defines the curve $w = F(x - at_1)$. It is easy to see that these curves are identical except that the latter is translated to the right through a distance at_1, and therefore with velocity

$$v = \frac{at_1}{t_1} = a.$$

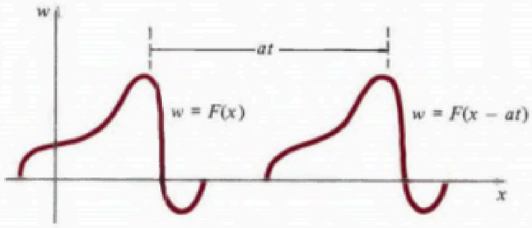


Figure 19.20 A traveling wave.

This shows that the function $w = F(x - at)$ represents a traveling wave that moves to the right with velocity a, as suggested in Fig. 19.20. If we assume that $w = F(u)$ has a second derivative, then by the chain rule applied to $w = F(u)$ where $u = x - at$, we have

$$\frac{\partial w}{\partial x} = F'(u), \qquad \frac{\partial w}{\partial t} = F'(u) \cdot (-a) = -aF'(u),$$

$$\frac{\partial^2 w}{\partial x^2} = F''(u), \qquad \frac{\partial^2 w}{\partial t^2} = -aF''(u) \cdot (-a) = a^2 F''(u).$$

It is clear from this that $w = F(x - at)$ satisfies the one-dimensional wave equation (8).

Similarly, the function $w = G(x + at)$ represents a traveling wave that moves to the left with velocity a, and it is equally easy to show that this function is a solution of (8). By the linearity of differentiation, it follows that the sum

$$w = F(x - at) + G(x + at) \tag{9}$$

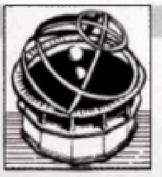
is also a solution. In fact, it can be shown (see Problem 8) that if F and G are *arbitrary* twice-differentiable functions, then (9) is the *general* solution of (8), in the sense that every solution of (8) has the form (9). It is fairly clear that the function (9) represents the most general one-dimensional wave, and this result confirms it.

NOTE ON LAPLACE

Pierre Simon de Laplace (1749–1827) was a French mathematician and theoretical astronomer who was so famous in his own time that he was known as the Newton of France. His main scientific interests throughout his life were celestial mechanics and the theory of probability.

At the age of 24 he was already deeply engaged in the detailed application of Newton's law of gravitation to the solar system as a whole, in which the planets and their satellites are not governed by the sun alone but interact with one another in a bewildering variety of ways. Even Newton had been of the opinion that divine intervention would occasionally be needed to prevent this complex mechanism from degenerating into chaos. Laplace decided to seek reassurance elsewhere, and succeeded in proving that the ideal solar system of mathematics is a stable dynamical system that will endure unchanged for all time. This achievement was only one of the long series of triumphs recorded in his monumental treatise *Mécanique Céleste* (published in five volumes from 1799 to 1825), which summed up the work on gravitation of several generations of illustrious mathematicians. Many anecdotes are associated with this work. One of the best known describes the occasion on which Napoleon tried to get a rise out of Laplace by protesting that he had written a huge book on the system of the world without once mentioning God as the author of the universe. Laplace is supposed to have replied, "Sire, I had no need of that hypothesis." The principal legacy of the *Mécanique Céleste* to later generations lay in Laplace's wholesale development of potential theory, with its far-reaching implications for a dozen different branches of physical science ranging from gravitation and fluid mechanics to electromagnetism and atomic physics. Even though the concept of the potential is due to Lagrange, Laplace exploited it so extensively that ever since his time the fundamental differential equation of potential theory has been known as Laplace's equation.

His other masterpiece was the treatise *Théorie Analytique des Probabilités* (1812), in which he incorporated his own discoveries in probability from the preceding 40 years. This book is generally agreed to be the greatest contribution to this part of mathematics ever made by one man. In the introduction he says, "At bottom, the theory of probability is only common sense reduced to calculation." This may be so, but the following 700 pages of intricate analysis—in which he freely used Laplace transforms, generating functions, and many other highly nontrivial tools—has been said by some to surpass in complexity even the *Mécanique Céleste.*



NOTE ON FOURIER

Jean Baptiste Joseph Fourier (1768–1830), an excellent mathematical physicist, was a friend of Napoleon (so far as such people have friends) and accompanied his master to Egypt in 1798. On his return he became prefect (governor) of the district of Isère in southeastern France, and in this capacity built the first real road from Grenoble to Turin. He also befriended the boy Champollion, who later deciphered the Rosetta Stone as the first long step toward understanding the hieroglyphic writing of the ancient Egyptians.

During these years he worked on the theory of the conduction of heat, and in 1822 published his famous *Théorie Analytique de la Chaleur,* in which he made extensive use of the series that now bear his name. These series were of profound significance in connection with the evolution of the concept of a function. The general attitude at that time was to call $f(x)$ a function if it could be represented by a single expression like a polynomial, a finite combination of elementary functions, a power series $\sum_{n=0}^{\infty} a_n x^n$, or a trigonometric series of the form

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx).$$

If the graph of $f(x)$ were "arbitrary"—for example, a polygonal line with a number of corners and even a few gaps—then $f(x)$ would not have been accepted as a genuine function. Fourier claimed that "arbitrary" graphs can be represented by trigonometric series and should therefore be treated as legitimate functions, and it came as a shock to many that he turned out to be right. It was a long time before these issues were completely clarified, and it was no accident that the definition of a function that is now almost universally used was first formulated by Dirichlet in 1837 in a research paper on the theory of Fourier series. Also, the classical definition of the definite integral due to Riemann was first given in his fundamental paper of 1854 on the subject of Fourier series. Indeed, many of the most important mathematical discoveries of the nineteenth century are directly linked to the theory of Fourier series, and the applications of this subject to mathematical physics have been scarcely less profound.

Fourier himself is one of the fortunate few: his name has become rooted in all civilized languages as an adjective that is well known to physical scientists and mathematicians in every part of the world.

1. 标题与内容概述

《一维波动方程与旅行波解：从“移动函数”到一般性解决方案》

本节介绍了**一维波动方程**（单变量空间 x 与时间 t ），

$$a^2 \frac{\partial^2 w}{\partial x^2} = \frac{\partial^2 w}{\partial t^2},$$

并说明其通用解的形式实际上是“向左与向右传播”的两种**旅行波（traveling waves）**的线性叠加。这一结果揭示了波动现象的本质：**任何形状都可以以速度 a 向左或向右传播**，或两者叠加形成更复杂的波形。

---

2. 详细内容解析

下面将分步骤说明为何波动方程具有“沿正负方向移动”的解，最后给出这一方程的一般解是两个独立旅行波之和。

2.1 一维波动方程与物理含义

1. **方程形式**

$$a^2\,\frac{\partial^2 w}{\partial x^2} \;=\; \frac{\partial^2 w}{\partial t^2}. \tag{8}$$

   - 这里 $w = f(x,t)$ 可以代表弦上质点的位移、空气中声压的偏离量、电磁波中的场强等。
   - 常数 a 通常解释为**波的传播速度**。

2. **物理解读**
   - 该方程说明**二阶时间变化**与**二阶空间变化**成正比，揭示了扰动会以速度 a 在介质中传播。
   - 同样在三维中也有波动方程，但一维是最基本、易理解的情况。

2.2 旅行波解：$w = F(x - at)$

1. **右移波函数**
   - 假设有个函数形状 $F(u)$，其中 $u = x - at$。当 $t = 0$ 时，波形为 $w = F(x)$。
   - 在任意时刻 $t > 0$，这同样的波形只是在 x -坐标轴上向右平移了 $a\,t$ 距离，因此可写成 $w = F(x - at)$。
   - 以图 19.20 为例，$w = F(x)$ 在 $t = 0$ 时可能是一道脉冲曲线，当 $t \neq 0$ 时变为 $w = F(x - at)$，表示该波形以速度 a 向正 x 方向移动。

2. **验证满足波动方程**
   - 令 $u = x - at$。利用链式法则：

$$\frac{\partial w}{\partial x} = F'(u), \quad \frac{\partial w}{\partial t} = F'(u)\cdot(-a),$$
$$\frac{\partial^2 w}{\partial x^2} = F''(u), \quad \frac{\partial^2 w}{\partial t^2} = F''(u)\cdot(-a)^2 = a^2\,F''(u).$$

   - 由此可见
     $a^2\,\frac{\partial^2 w}{\partial x^2} = a^2\,F''(u), \quad \frac{\partial^2 w}{\partial t^2} = a^2\,F''(u),$
     二者相等，故 $w = F(x - at)$ 满足方程 (8)。

3. **左移波函数**
   - 同理，若定义 $G(x + at)$，它则代表**以速度 a 向负 x 方向传播**的波形，亦满足 (8)。

2.3 通用解：两旅行波之和

1. **线性叠加**
   - 波动方程是线性的，因此若 $F(x - at)$ 和 $G(x + at)$ 都是解，则它们的任意线性组合

$$w(x,t) = F(x - at) + G(x + at)$$

     也是解。

2. **一般解的形态**
   - 事实上，可以证明（见教材配套习题或更高级理论）任意足够光滑的一维波方程解都可以写成两个独立旅行波的叠加。
   - 这意味着在一维中，初始波形在后续演化时，会“分裂”为向左、向右移动的波，各自独立传播，且速度均为 a。

2.4 小结及拓展

1. **在更高维度**
   - 在二维、三维中，解的结构更为复杂，但波动方程仍反映了扰动以速度 a 向外传播的特性。
2. **物理实例**
   - 弦的振动：初始形状可视作两个相向或反向行进的波；
   - 声波在气体中：扰动以声速 a 传播；
   - 电磁波在真空：a 就是光速 c。

---

3. 本次内容小结

- **一维波动方程** $a^2\,\partial^2 w/\partial x^2 = \partial^2 w/\partial t^2$ 的解可以用两"旅行波"之和来表达：

$$w(x,t) = F(x-at) \;+\; G(x+at).$$

- **旅行波解**：$F(x-at)$ 表示以速度 a 向右传播的波形；$G(x+at)$ 表示以速度 a 向左传播。
- **常数** a：波速，由介质或系统性质决定。
- **物理含义**：该结论在声波、光波、电磁波、振动弦等现象中普遍适用，为波动现象分析的基础。

---

4. 报错或补充说明

- 上述笔记完整讲解了课件针对"一维波方程"与"旅行波解"提供的思路和结论。如有截图或文字缺失之处，请指出；若无，则本次笔记已涵盖全部要点。

19.10

19.10 (OPTIONAL) IMPLICIT FUNCTIONS

In Section 3.5 we stated that when we are given an equation

$$F(x, y) = 0, \tag{1}$$

there usually exists at least one function

$$y = f(x) \tag{2}$$

that "solves" (1), in the sense that (2) reduces (1) to an identity in x. With the idea in mind that y in (1) stands for this function of x, we then differentiated the identity (1) with respect to x and went on to solve the resulting equation for dy/dx, calling the process "implicit differentiation." For instance, if we have the equation

$$x^2y^5 - 2xy + 1 = 0, \tag{3}$$

then by differentiating with respect to x we obtain

$$x^2 \cdot 5y^4 \frac{dy}{dx} + 2xy^5 - 2x \frac{dy}{dx} - 2y = 0, \tag{4}$$

so

$$\frac{dy}{dx} = \frac{2y - 2xy^5}{5x^2y^4 - 2x}. \tag{5}$$

Most students feel slightly uncomfortable about implicit differentiation, and with good reason. For one thing, in this particular case we have no idea whether (3) actually defines y as a function of x or not; and if it doesn't, then the subsequent calculation leading to (5) has no meaning at all. Also, the procedure itself is a bit clumsy, because it requires us to keep in mind the different roles played by the variables x and y. We are now in a position to clarify the meaning of this process, and also to give a precise statement of the conditions under which an equation of the form (1) defines a differentiable function (2).

We broaden the discussion slightly, and instead of (1) consider an equation of the form

$$F(x, y) = c, \tag{6}$$

whose graph is a level curve of the function $z = F(x, y)$. For example, the graph of

$$x^2 + y^2 = 1 \tag{7}$$

is a circle about the origin (Fig. 19.21, left), and this is a level curve of the function $F(x, y) = x^2 + y^2$. Generally, as in this case, the graph of (6) will be some sort of curve that is not the graph of a single function. However, even though the entire graph of (7) is not the graph of a single function, it is clear that every point (x_0, y_0) on this graph with $y_0 \neq 0$ lies on a *portion* of the graph that *is* the graph of a function—indeed, of a differentiable function. Specifically, if $y_0 > 0$ then (x_0, y_0) lies on the graph of the function

$$y = f_1(x) = \sqrt{1 - x^2}, \tag{8}$$

and if $y_0 < 0$ then (x_0, y_0) lies on the graph of the function

$$y = f_2(x) = -\sqrt{1 - x^2}. \tag{9}$$

Similarly, the graph of (6) might consist of the graphs of two or more differentiable functions $y = f(x)$, as suggested on the right in the figure.

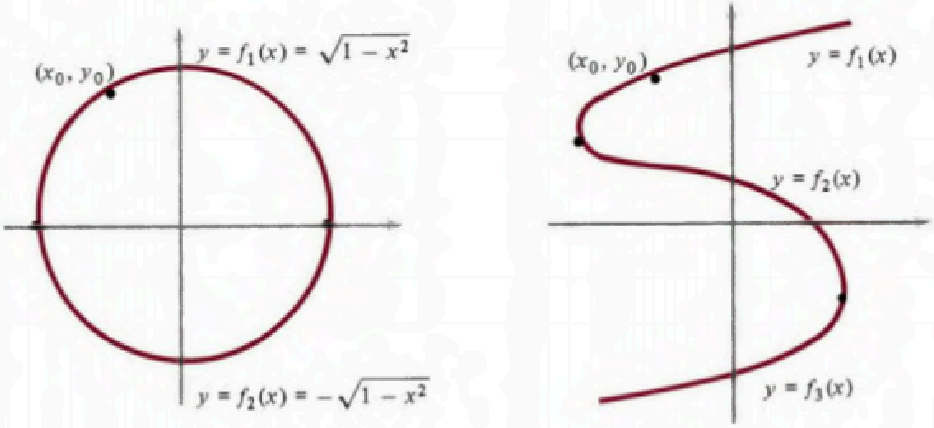


Figure 19.21

We next point out that the function $z = F(x, y)$ has the constant value c along the graph of any such function $y = f(x)$,

$$z = F[x, f(x)] = c.$$

As usual, we assume that $F(x, y)$ has continuous partial derivatives, so it is permissible to write

$$\frac{dz}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0,$$

or equivalently,

$$\frac{dz}{dx} = F_x(x, y) + F_y(x, y)\frac{dy}{dx} = 0. \tag{10}$$

The middle term here is just the chain rule evaluation of dz/dx when $z = F(x, y)$ and y is a function of x, and the result is zero because z is constant as a function of x. If $F_y(x, y) \neq 0$, equation (10) can be solved for dy/dx,

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)}. \tag{11}$$

In the language of Section 3.5, any differentiable function $y = f(x)$ with the property that

$$F[x, f(x)] = c$$

is an *implicit function* defined by (6), and (11) provides a general formula for the derivative of such a function.

If we apply formula (11) to equation (7), where $F(x, y) = x^2 + y^2$, we obtain

$$F_x = 2x \quad \text{and} \quad F_y = 2y, \quad \text{so} \quad \frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{x}{y}, \quad y \neq 0. \tag{12}$$

In this case we know from (8) and (9) that (7) actually determines two implicit functions $y = f(x)$, so the calculations (12) are legitimate and apply to either function as long as we avoid points where $y = 0$. However, suppose that instead of (7) we have one of the equations

$$x^2 + y^2 = -1 \quad \text{or} \quad x^2 + y^2 = 0. \tag{13}$$

By acting blindly without thinking, we can write down the calculations (12) just as before and "find" dy/dx. The difficulty with this is obvious: Since the graph of (13) is either empty or consists of a single point, no implicit function $y = f(x)$ exists, and these calculations would be nothing more than a kind of mathematical doubletalk, which seems to be saying something but really says nothing at all.

In order to avoid committing such nonsense, it is necessary to have definite knowledge that implicit functions exist. This is the purpose of the

Implicit Function Theorem *Let $F(x, y)$ have continuous partial derivatives throughout some neighborhood of a point (x_0, y_0), and assume that $F(x_0, y_0) = c$ and $F_y(x_0, y_0) \neq 0$. Then there is an interval I about x_0 with the property that there exists exactly one differentiable function $y = f(x)$ defined on I such that $y_0 = f(x_0)$ and*

$$F[x, f(x)] = c.$$

Further, the derivative of this function is given by the formula

$$\frac{dy}{dx} = -\frac{F_x}{F_y},$$

and is therefore continuous.

1. 标题与内容概述

《隐函数与隐函数定理：从“方程 $F(x, y) = 0$ ”到“函数 $y = f(x)$ ”的条件与推论》

在前面章节中，我们时常看到形如 $F(x, y) = 0$ 的方程，有时我们能把它“解”成显式函数 $y = f(x)$ 。然而，这种显式求解往往并不简单，甚至无法用初等函数表达。但微分学告诉我们，**只要满足一定条件**，我们依然可以把 $F(x, y) = 0$ **局部**看作 $y = f(x)$ 并进行求导。这种“隐式”地定义函数的技术，被称作**隐函数法**，而其在多变量情形下更具一般性，称为**隐函数定理（Implicit Function Theorem）**。

---

2. 详细内容解析

以下按课件中的逻辑，对隐式函数的产生背景、求导方式以及隐函数定理做系统讲解。

2.1 从“隐式方程”到“显式函数”的必要性

1. **隐式方程的范例**
   - 例如， $x^2y^5 - 2xy + 1 = 0$
     并未直接给出 y 等于什么，但在某些区域里，可能可以把这方程“翻译”为 $y = f(x)$ 。
   - 在之前（单变量微积分）的学习中，我们曾用隐式微分得到

$$\frac{dy}{dx} = \frac{2y - 2xy^5}{5x^2y^4 - 2x},$$

   但做这些运算前提是：该方程**真的**可以在邻域内视为函数 $y = f(x)$ 。
2. **何时能解成函数?**
   - 如果方程 $F(x, y) = 0$ 在附近只是某条“曲线”（而不是一条封闭曲线或自交图形），并且在考察点 (x_0, y_0) 处满足一个非退化条件（如 $\frac{\partial F}{\partial y} \neq 0$ ），就能局部地把它表示成 $y = f(x)$ 。
   - 反之，如果在某点 $\frac{\partial F}{\partial y} = 0$ ，可能导致无法以 y 作为“因变量”唯一地表达出来（或者需要把 x 视为 y 的函数，等等）。

2.2 例子： $x^2 + y^2 = 1$

1. **约束方程**
   - 圆周方程 $x^2 + y^2 = 1$ 。在直观几何上，这是一条闭合曲线，并非某个单值函数 $y = f(x)$ 。
   - 然而，在上半圆段 $(y > 0)$ 可以写为 $y = \sqrt{1 - x^2}$ ；在下半圆段 $(y < 0)$ 则是 $y = -\sqrt{1 - x^2}$ 。
   - 这说明：**在不同子区域**（如上半或下半），能把隐式方程拆分成两个不同的显式函数 $f_1(x)$ 和 $f_2(x)$ 。
2. **隐式微分**
   - 若我们在上半圆， $y > 0$ ，由公式 (11) 可得

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)}.$$

   以 $F(x, y) = x^2 + y^2$ 为例， $F_x = 2x,\ F_y = 2y$ ，所以

$$\frac{dy}{dx} = -\frac{2x}{2y} = -\frac{x}{y}.$$

   - 这与直接对 $y = \sqrt{1 - x^2}$ 求导所得结果一致（需要 $y \neq 0$ ）。

2.3 隐函数定理及其要点

1. **隐函数定理的陈述**
   - 课件（或教材）里给出的简化版是：

$$\text{若 } F(x,y) \text{ 在某邻域内有连续偏导数，并且 } F(x_0,y_0)=c,\ F_y(x_0,y_0)\neq 0,$$

   则存在一条区间 I （以 x_0 为中心），使得在这个区间上能找到一个**可微函数** $y=f(x)$ ，满足

$$F\big(x,f(x)\big)=c,\quad \text{且 } f(x_0)=y_0,$$

   并且其导数

$$\frac{dy}{dx}=-\frac{F_x(x,y)}{F_y(x,y)}.$$

   - 这告诉我们：只要 $\frac{\partial F}{\partial y}\neq 0$ （非退化条件），就能把该方程看作 y 关于 x 的函数。
2. **可微性与连续性**
   - 定理还保证 f 不仅存在，而且是可微的， $\frac{dy}{dx}$ 连续可得。
   - 如果要把 x 视为 $\phi(y)$ 的函数，则要求 $\frac{\partial F}{\partial x}\neq 0$ 类似地成立。
3. **为什么要这样严格?**
   - 若 $\frac{\partial F}{\partial y}=0$ 或者方程根本没在附近形成曲线（例如 $x^2+y^2=-1$ 不存在实解），就无法应用该定理；结果可能就是“空”或非函数曲线等情况。
   - 隐函数定理可推广到更高维度，处理 $F(x_1,\ldots,x_n)=c$ 等复杂场景。

2.4 避免盲目的“隐式微分”

1. **错误示例**
   - 若我们看方程 $x^2+y^2=-1$ ，在实数范围内根本无解，但如果我们强行用公式 $\frac{dy}{dx}=-\frac{2x}{2y}$ 就得到一个空洞说法，因为 (x,y) 不存在实解也就谈不上函数 $y=f(x)$ 。
   - 说明**必须先确认方程确在邻域里定义了一条真实可微曲线**，再谈隐式微分才有意义。
2. **小结**
   - **先判定**是否在点 (x_0,y_0) 附近有可微函数 $y=f(x)$；
   - **再**做隐式微分，即使用公式 (11)。
   - 这样避免“在不存在解的场合”做无意义运算。

---

3. 本次内容小结

- **隐函数：** 当二元方程 $F(x,y)=c$ 在某处满足 $\frac{\partial F}{\partial y}\neq 0$ ，就可以在邻域内把它视作可微函数 $y=f(x)$ 。
- **隐式微分：**

$$\frac{dy}{dx}\ =\ -\frac{F_x}{F_y}.$$

- **关键：** 不要盲目使用此公式，必须先确保“隐函数”确实存在（例如图形在那个点附近不出现竖直切线或自交）。
- **更一般情形：** 还可换位思考，用 $\frac{\partial F}{\partial x}\neq 0$ 则把 x 当作 $x=g(y)$；或推广到更多维度（隐函数定理在高维中也常用）。

---

4. 报错或补充说明

- 截止该截图内容，已对隐函数及其微分原理、隐函数定理的主要要点作了全面讲解。若有图片或文字难以辨认，请告知；否则本次笔记已涵盖相关关键知识。

19.10.2

It should be understood that this theorem is a purely theoretical statement to the effect that the specified implicit function $y = f(x)$ does in fact exist, and it has no bearing on the issue of whether a simple formula can be found for this function. A proof is given in Appendix A.20.

Example 1 We consider once more the equation mentioned earlier,

$$F(x, y) = x^2y^5 - 2xy + 1 = 0. \tag{3}$$

It is clear that the point (1, 1) lies on the graph, so the graph is not empty. Since $F_x = 2xy^5 - 2y$ and $F_y = 5x^2y^4 - 2x$, our theorem guarantees that equation (3) determines an implicit function $y = f(x)$ about any point of the graph where $F_y = 5x^2y^4 - 2x \neq 0$, for instance the point (1, 1). It is instructive to write down equation (10) for this case,

$$(2xy^5 - 2y) + (5x^2y^4 - 2x)\frac{dy}{dx} = 0, \tag{14}$$

and to compare the result with (4), where implicit differentiation is carried out by the old method. Equation (14) evidently yields

$$\frac{dy}{dx} = \frac{2y - 2xy^5}{5x^2y^4 - 2x},$$

just as before.

The simplicity of our present method is even more clearly visible when there are three variables in the given equation.

Thus, suppose an equation $F(x, y, z) = c$ defines a certain implicit function $z = f(x, y)$, and let us find $\partial z/\partial x$ in terms of the function $F(x, y, z)$. The equations

$$w = F(x, y, z),$$

$$x = x, \qquad y = y, \qquad z = f(x, y),$$

give w as a composite function of x and y. Also,

$$w = F[x, y, f(x, y)] = c,$$

so if we differentiate this with respect to x, the chain rule yields

$$\frac{\partial w}{\partial x} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0.$$

We therefore obtain

$$\frac{\partial z}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z}, \tag{15}$$

and this formula is valid wherever $\partial F/\partial z \neq 0$. In just the same way, we also have

$$\frac{\partial z}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial z}. \tag{16}$$

As students have surely guessed, there is also an Implicit Function Theorem that covers this situation. Briefly, it says that if $\partial F/\partial z \neq 0$ at a point (x_0, y_0, z_0) on a surface $F(x, y, z) = c$, then in a neighborhood of this point the surface defines a unique implicit function $z = f(x, y)$ such that $z_0 = f(x_0, y_0)$, and furthermore the partial derivatives of this function are given by (15) and (16).

Example 2 It is easy to verify that the point $(1, 2, -1)$ lies on the graph of the equation

$$x^2z + yz^5 + 2xy^3 = 13, \tag{17}$$

so this graph is not empty. If the equation defines an implicit function $z = f(x, y)$ in a neighborhood of this point, then we can calculate $\partial z/\partial x$ by implicit differentiation in the old way. This means we differentiate (17) implicitly with respect to x, thinking of y as a constant, which gives

$$x^2\frac{\partial z}{\partial x} + 2xz + y \cdot 5z^4\frac{\partial z}{\partial x} + 2y^3 = 0,$$

so

$$\frac{\partial z}{\partial x} = -\frac{2xz + 2y^3}{x^2 + 5yz^4}.$$

This procedure is unsatisfactory because we don't know in the beginning whether any such function $z = f(x, y)$ actually exists—after all, (17) is a fifth-degree equation in z—and also because in the implicit differentiation each of the three variables has to be treated in a different way, and it is quite easy to lose track of what is going on. Our present ideas provide a much better method. We have

$$F(x, y, z) = x^2z + yz^5 + 2xy^3,$$

so

$$\frac{\partial F}{\partial x} = 2xz + 2y^3 \qquad \text{and} \qquad \frac{\partial F}{\partial z} = x^2 + 5yz^4.$$

It is easy to see that $\partial F/\partial z = 11 \neq 0$ at $(1, 2, -1)$, so the Implicit Function Theorem guarantees that $z = f(x, y)$ exists. Also, by (15) we have

$$\frac{\partial z}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z} = -\frac{2xz + 2y^3}{x^2 + 5yz^4},$$

which avoids the messy implicit differentiation.

Remark The two-variable version of the Implicit Function Theorem enables us to complete a long-standing piece of unfinished business. In the earlier chapters of this book we gave quite a bit of attention to the important problem of finding the inverse function of a given function $g(y) = x$, in other words, the problem of solving the equation

$$F(x, y) = g(y) - x = 0 \tag{18}$$

for the variable y. Specifically, this is the way the familiar functions $y = \ln x$, $y = \sin^{-1} x$, and $y = \tan^{-1} x$ were defined. Each of these inverse functions was discussed earlier in an *ad hoc* but perfectly legitimate way. We are now in a position to draw the general inference that when $g(y)$ has a continuous derivative and $\partial F/\partial y = g'(y) \neq 0$, then (18) can indeed be solved for y, $y = f(x)$, and also that this function has a continuous derivative given by

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y} = -\frac{-1}{g'(y)} = \frac{1}{dx/dy}.$$

This completes the line of thought that was briefly described in Remark 2 of Section 9.5.

1. 标题与内容概述

《多变量隐函数定理：从二元到三元的推广与应用》

在前面，我们已看到当方程 $F(x,y)=c$ 在某一点满足 $\partial F/\partial y \neq 0$ 时，可以在这点邻域将方程视为一条可微曲线——即 $y=f(x)$ 隐式定义；并且可用

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

来计算其导数。本节进一步展示如何把这个思想推广到**更多变量和更多方程**，以及为什么隐式函数定理能让我们避免杂乱无章的"隐式微分"。同时，通过示例说明在三元方程 $F(x,y,z)=c$ 下，如何把其中一个变量（如 z）隐式地视为另一个两个变量（x,y）的可微函数。

---

2. 详细内容解析

2.1 多变量隐式函数： $F(x,y,z)=c$

1. **三元方程下的隐式函数**
   - 如果我们有一个方程 $F(x,y,z)=c$，其图形在三维空间中一般是一个曲面（或集合）。
   - 在合适条件（比如 $\partial F/\partial z \neq 0$）下，就能在 (x_0,y_0,z_0) 附近把这曲面视作" $z=f(x,y)$ "的一张图，从而在局部可用显式方式描述。
   - 这就是**三变量情形的隐函数定理。**
2. **导数公式**
   - 若 $\frac{\partial F}{\partial z}(x_0,y_0,z_0) \neq 0$，则在点 (x_0,y_0,z_0) 存在邻域，使我们能写 $z=f(x,y)$ 并满足 $F(x,y,f(x,y))=c$。
   - 通过类似的链式法则，如课件(15) 与 (16) 给出：

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}}, \quad \frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}.$$

3. **为什么更优雅?**
   - 过去如果试图直接对 $F(x,y,z)=c$ 做隐式微分，常要把 z 视作 $z=f(x,y)$，结果牵涉到众多变量之间的关系，容易让人迷糊。
   - 有了隐函数定理的保证：我们先确认" $\partial F/\partial z \neq 0$ "，然后才能放心地说" z 可以看作 $z=f(x,y)$ "，从而套用公式(15)(16) 安全求出 $\frac{\partial z}{\partial x}$ 与 $\frac{\partial z}{\partial y}$。

2.2 示例： $x^2z+y^2\,5^z+2xy^3=13$

1. **设定方程**

$$F(x,y,z) = x^2\,z + y^2\,5^z + 2xy^3 - 13 = 0.$$

   若要在点 $(1,2,-1)$ 尝试把它视为" $z=f(x,y)$ "，必须检查 $\partial F/\partial z$ 在该点是否非零。
2. **计算并检验**

- $\frac{\partial F}{\partial z} = 2x + (y^2 \cdot 5^z \ln 5)$；带入 $(1,2,-1)$ 得 $\partial F/\partial z\big|_{(1,2,-1)} = 2\cdot 1 + (2^2 \cdot 5^{-1} \cdot \ln 5) = 2 + \frac{4}{5}\ln 5$.
- 假设结果不为0（课件中算得 11 $\neq 0$ ），则可以说明该点处确实存在局部隐函数 $z = f(x,y)$ 。

3. **求** $\partial z/\partial x$
   - 由课件(15):

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z} = -\frac{\frac{\partial}{\partial x}[\,x^2 z + y^2\,5^z + 2xy^3\,]}{\frac{\partial}{\partial z}[\,x^2 z + y^2\,5^z + 2xy^3\,]}$$

在 $(1,2,-1)$ 带入后即可得到数值，不必再做繁琐的"把 y 看作常量"之类的操作。

2.3 隐函数定理的价值与推广

1. **一般陈述**
   - 如果有方程组 $F(x_1,\ldots,x_n) = \cdots = 0$，并且满足对某些变量的偏导子矩阵非奇异，就能在该点把若干变量隐式地当作另外的自变量的函数。
   - 这种思想广泛用在微分几何、优化、数值分析等领域。
2. **与反函数定理关系**
   - 隐函数定理可视作反函数定理的推广：反函数定理谈" $y = g^{-1}(x)$ "可以局部存在，隐函数定理则谈" $y = f(x)$ "可以局部存在。
3. **总结**
   - **先**检查：目标方程/方程组在所关心的点，是否能满足相应偏导（或雅可比矩阵）的非零/非奇异条件；
   - **再**从定理得知：可局部表示成一组可微函数；
   - **最后**用链式法则做微分，即可得到那些隐函数的偏导数公式。

2.4 一些小注解

1. **Remark**: 二元情况 (x,y) 常见；三元情况 (x,y,z) 亦类同。
2. **最终应用**:
   - 例如在经济学、工程学中，把约束方程 $\Phi(x,y,z) = 0$ 视为 $z = f(x,y)$ 后，就可用 $\partial z/\partial x, \partial z/\partial y$ 来分析系统行为。
   - 在更高级章节，还会出现多重隐式方程及其导数（可与拉格朗日乘数法对照）。

---

3. 本次内容小结

- **多变量隐函数定理**：若 $\frac{\partial F}{\partial z}(x_0,y_0,z_0) \neq 0$，则在该点附近可以把 z 看作可微函数 $z = f(x,y)$，并可用

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}, \quad \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}.$$

- **避免盲目隐式微分**：只有在确认"隐式函数存在"后，才有意义地使用这些公式。
- **应用范畴**：从二元到三元、乃至多元方程都可依赖同样思路；关键在于满足非退化（非奇异）条件。
- **示例收获**：通过实际例子（ $x^2 z + y^2\,5^z + 2xy^3 = 13$ 等）可以省去繁琐的"对 $\partial/\partial x$ "提取过程中保持 y 不变、 z 亦不变"的混乱，而以更直接的分块来计算。

---

4. 报错或补充说明

- 截至此，已结合截图完整讲解多元隐函数定理及其在三维下的具体运用。如尚有任何不清晰处，请告诉我；若无，则本节笔记即为完备。

CHAPTER 19 REVIEW

CHAPTER 19 REVIEW: DEFINITIONS, METHODS

Think through the following.

1 Domain, continuity, and level curves for $z = f(x, y)$.
2 Definition and geometric meaning of partial derivatives of $z = f(x, y)$.
3 Equality of mixed partial derivatives.
4 Equation of tangent plane to $z = f(x, y)$.
5 Directional derivative and gradient.
6 The del operator.
7 The chain rule.
8 Method of Lagrange multipliers for constrained maxima and minima.

---

1. 定义域、连续性与等高线（level curves）

1. **定义域**
   - 对函数 $z = f(x, y)$ 而言，它的定义域是满足函数表达式有意义的所有 (x, y) 。例如分母不可为 0，开平方数需非负等。
   - 在涉及高维的情形，也可类似地定义： $w = f(x, y, z, \ldots)$ 在满足表达式可计算的所有点处有定义域。
2. **连续性**
   - 若在点 (x_0, y_0) 附近，令 $(x, y) \to (x_0, y_0)$ 时 $f(x, y) \to f(x_0, y_0)$ ，则说 f 在 (x_0, y_0) 连续。
   - 直观解释： (x, y) 越靠近 (x_0, y_0) ，对应的函数值就越靠近 $f(x_0, y_0)$ 。
3. **等高线（Level Curves）**
   - 定义：在 (x, y) 平面上画出所有满足 $f(x, y) = c$ 的点所组成的曲线。
   - 实际应用：可以在二维图中观察函数在不同高度 (即不同 c 值) 所形成的轮廓线（contour map）。

---

2. 偏导数的定义及几何意义

1. **一阶偏导数**
   - 对 $z = f(x, y)$ ，固定一个变量，让另一个变量变化并求单变量导数。
   - 例如：

$$f_x(x, y) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}, \quad f_y(x, y) = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}.$$

2. **几何意义**
   - 将曲面 $z = f(x, y)$ 与平面 $y = y_0$ 截取，产生一条曲线，其在 (x_0, y_0) 处的切线斜率就是 $f_x(x_0, y_0)$ ；
   - 同理，与 $x = x_0$ 截面上的切线斜率是 $f_y(x_0, y_0)$ 。

---

3. 混合偏导数的相等性

1. **二阶偏导**
   - 包含纯二阶： f_{xx}, f_{yy} ，以及混合二阶： f_{xy}, f_{yx} 。
   - 若 f 及其二阶偏导数在邻域内都连续（满足一定光滑条件），则

$$f_{xy}(x, y) = f_{yx}(x, y).$$

   - 称此为**Clairaut（Schwarz）定理。**

---

4. 切平面的方程

1. **空间曲面 $z = f(x,y)$ 在 (x_0, y_0, z_0) 的切平面**
   - 若 $z_0 = f(x_0, y_0)$ ，则切平面方程可写作：

$$z - z_0 = f_x(x_0, y_0)\,(x - x_0) \;+\; f_y(x_0, y_0)\,(y - y_0).$$

2. **法向量**
   - 该平面的法向量可由 $\bigl(f_x(x_0,y_0),\, f_y(x_0,y_0),\, -1\bigr)$ 给出。

---

5. 方向导数与梯度（gradient）

1. **方向导数**
   - 若 $\mathbf{u}$ 是单位向量，方向导数 $\frac{df}{ds}\big|_{\mathbf{u}}$ 表示在点 (x_0, y_0) 沿 $\mathbf{u}$ 方向的瞬时变化率。
   - 公式：

$$D_{\mathbf{u}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \mathbf{u}$$

   - $\mathbf{u}$ 归一化是关键。
2. **梯度 ∇f**
   - 向量 $\bigl(f_x, f_y\bigr)$ （在三维时 $\bigl(f_x, f_y, f_z\bigr)$ ）。
   - 性质：指向函数增长最快方向，其大小 $\|\nabla f\|$ 是最大方向导数值。

---

6. del 运算符（ ∇ 算子）

1. **定义**
   - (\displaystyle \nabla = \frac{\partial}{\partial x}\mathbf{i}
     - \frac{\partial}{\partial y}\mathbf{j}
     - \frac{\partial}{\partial z}\mathbf{k})。
   - 对标量场 f 作用产生梯度 ∇f 。
   - 在向量场分析中还可产生散度div、旋度curl等概念。

---

7. 链式法则（Chain Rule）

1. **多变量版本**
   - 若 $w = f(x,y)$ 而 $x = g(t),\, y = h(t)$ ，则

$$\frac{dw}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt}.$$

   - 推广到三变量 $w = f(x,y,z)$ 依赖于单一 t ，则

$$\frac{dw}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt}.$$

   - 对多个自变量 t, u 等也有类似思路。

---

8. 拉格朗日乘数法（约束下的极值）

1. **问题类型**
   - 在求 $\max$ 或 $\min$ $f(x,y)$ 时，若受制于约束 $g(x,y) = 0$ ，可使用拉格朗日乘数 λ 并解：

$$\nabla f = \lambda \nabla g, \quad g(x,y) = 0.$$

    - 这在三元及更高元情况亦适用。
2. **方程组**
    - 二元例子：

$$f_x = \lambda g_x, \quad f_y = \lambda g_y, \quad g(x, y) = 0.$$

    - 解出 (x, y)（及 λ）；再带回 f 判断极值大小、类型。
3. **几何解释**
    - 在极值点处，f 的等值线 $f(x, y) = \text{const}$ 与 $g(x, y) = 0$ 的曲线相切，故 ∇f 与 ∇g 平行。

额外说明

- 上述 8 个专题几乎覆盖了多变量微积分中的核心基本操作：从函数定义域与连续性，到求偏导数、方向导数、应用梯度处理几何与优化问题，再到链式法则与拉格朗日乘数，都对后续内容（如多重积分、向量分析）打下坚实基础。
- 建议在复习时辅以一定量的例题演练，如：求切平面方程、方向导数、梯度最大增加率、链式法则计算，以及拉格朗日乘数法在几何或经济学中的典型运用。

总结

这 8 个要点在多变量微积分中相互关联：

1. 在掌握多变量函数定义与连续性后，才能理解如何取偏导数、何时可以说函数在一开口区域内可微；
2. 偏导数不仅有一阶、还可探讨二阶与混合偏导，进而得到切平面或更高维的线性近似；
3. 梯度向量与方向导数让我们深入理解“函数如何在各方向增长”；
4. 链式法则、del 算子等是将单变量技巧平滑迁移到多变量领域的关键工具；
5. 最后，拉格朗日乘数法处理有约束的极值问题，为更多实际场景（经济学、物理、工程）提供统一算法。

只要把这些要点融会贯通，就能自信地处理多变量微积分的各类问题。祝学习顺利！
