# 译 者 的 话

对于**“马克思所遗留下来的极其重要的数学手稿”**(《反杜林论》)，恩格斯极为重视。但在马克思和恩格斯生前，都未能发表。

在马克思和恩格斯的书信中，也包含与数学手稿有关的内容。1881年8月18日恩格斯在给马克思的信中说：**“研究了你的数学手稿”**，**“我向你祝贺”**。1882年11月21日和22日恩格斯和马克思的相互通信也都对微分演算作了精辟的论述。上述书信均见《马克思恩格斯全集》第35卷。

1933年，苏联《在马克思主义旗帜下》杂志第一次刊登数学手稿部分内容的俄译文。

1971年，我们曾按这部分内容的日文本译成中文，印了试译本。1973年，我们根据德文原文对手稿的整个论文部分重新进行翻译，印了第二个试译本。接着，在上海《自然辩证法》杂志1974年第2、3两期上刊登了“导函数的概念”，“论微分”和“微分演算的历史发展过程”等主要论文。

这次刊出的译本，又对前两次的译稿进行了全面的修订和补充，目的是进一步征求对于译文的意见。这个译本收集了我们所见到的比较完整的手稿内容，整理成六个部分。前面五个部分都是马克思在八十年代(1880—1883)的手稿，最后第六部分标明是“七十年代的几份手稿”。这样，可以更好地领会马克思关于微分演算本质的思想发展和理解手稿的精神实质。

关于译文，有几点说明：

1. 在马克思的手稿中，变数 $x$, $y$ 变化以后写为 $x'$, $y'$，译文按现在通常的写法改为 $x_1$, $y_1$。

2. 凡有方括号“[ ]”的标题是译者所加的，译文中方括号内的文字也是译者所加的。

3. 译文下边的注都是译者加的。

参加翻译工作的有(按姓氏笔划为序)：复旦大学王福山、谷超豪、苏步青、陈少新、陈国亮、张开明、李立康、金若水、秦曾复、舒五昌，上海师范大学吕乃刚、茆诗松、周宝熙、程其襄等同志。

复旦大学理科资料组

一九七五年二月

# [导函数的概念。论文及草稿]

## [导函数的概念]①

### I

如果自变量 $x$ 增长到 $x_1$,那末因变量 $y$ 就增长到 $y_1$。

在 I 这里,我们将研究 $x$ 只以一次幂出现的这种最简单的情况。

1) $y = ax$; 如果 $x$ 增长到 $x_1$, 那末

$$y_1 = ax_1 \text{ 以及 } y_1 - y = a(x_1 - x)。$$

如果现在进行**微分运算**,也就是说,让 $x_1$ 减少到 $x$, 那末

$$x_1 = x; \quad x_1 - x = 0,$$

因而

$$a(x_1 - x) = a \cdot 0 = 0。$$

再者,只是由于 $x$ 变为 $x_1$,$y$ 才变为 $y_1$ 的,所以现在同样有

$$y_1 = y; \quad y_1 - y = 0。$$

因而

$$y_1 - y = a(x_1 - x)$$

变为 $0 = 0$。

先设差值,而后又把它扬弃,这种做法从字面上看来将导致**虚无**。在理解微分运算时所遇到的全部困难(就象一般理解**否定的否定**时一样),正在于要弄清楚它是**怎样**区别于这种简单的运算过程,以及**怎样**由此导出实际结果的。

如果我们用因子 $x_1 - x$ 去除 $a(x_1 - x)$, 并相应地去除等式的左边, 那末就得到

$$\frac{y_1 - y}{x_1 - x} = a。$$

由于 $y$ 是**因变量**, 它根本不能进行任何独立运动, 所以, 在 $x_1$ 没有变为 $= x$ 之前, $y_1$ 就不能变为 $= y$,因而 $y_1 - y$ 也就不能 $= 0$。

---

① 1881 年,马克思把这篇论文寄给了恩格斯。

另一方面我们已经看到，如果不使函数 $a(x_1-x)$ 变为 0，那末在函数中，$x_1$ 就不能变为 $=x$。因此当等式的两边用因子 $x_1-x$ 去除的时候，这因子**必然**是一个**有限差值**。所以在建立比值

$$\frac{y_1-y}{x_1-x}$$

的时候，$x_1-x$ 始终是一个有限差值。由此可知，

$$\frac{y_1-y}{x_1-x}$$

是一个**有限差值之比**；据此

$$\frac{y_1-y}{x_1-x}=\frac{\Delta y}{\Delta x}。$$

所以

$$\frac{y_1-y}{x_1-x}\text{或}\frac{\Delta y}{\Delta x}=a,$$

其中常量 $a$ 起着两个变量的有限差值之比的**极限值**的作用。

由于 $a$ 是常量，它不能有任何变化，所以化为这个常量的等式**右边**也就不能有任何变化。在这种情况下，**微分过程**在左边

$$\frac{y_1-y}{x_1-x}\text{或}\frac{\Delta y}{\Delta x}$$

进行，这是 $ax$ 这类简单函数的特点。

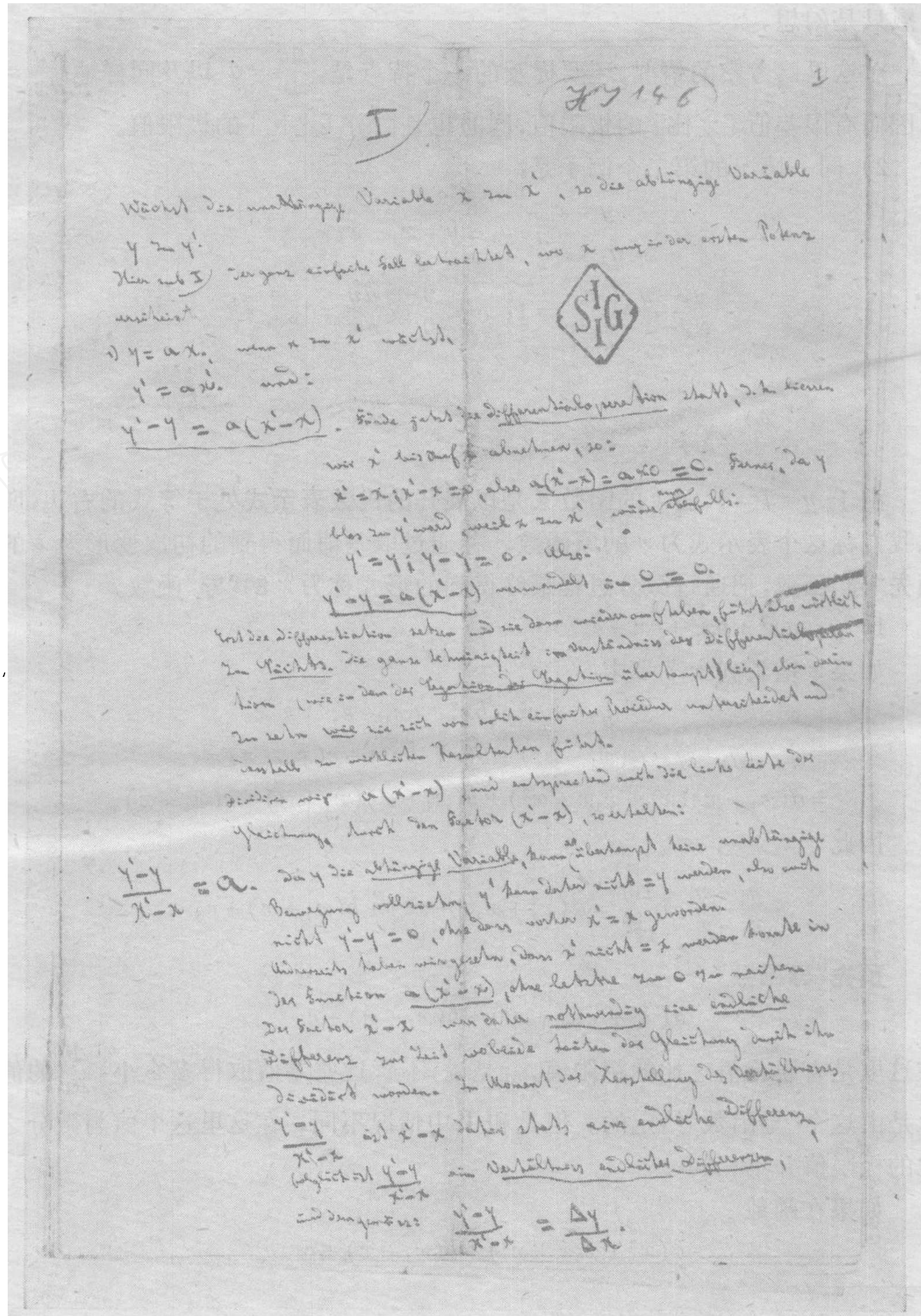
如果在这个比值的分母中 $x_1$ 减少，那末它将趋近于 $x$；一旦 $x_1$ 变为 $x$，它就达到了减少的极限。这样一来，就使差值 $x_1-x=x-x=0$，从而 $y_1-y$ 也就 $=y-y=0$。这样我们就得到

$$\frac{0}{0}=a。$$

由于在 $\frac{0}{0}$ 这个表示式中，$\frac{0}{0}$ 的来源和意义的任何痕迹都已消失，所以我们代之以 $\frac{dy}{dx}$，其中有限差值 $x_1-x$ 或 $\Delta x$ 以及 $y_1-y$ 或 $\Delta y$，都作为**扬弃了的或消失了的**差值而以符号化的形式出现，或者说 $\frac{\Delta y}{\Delta x}$ 变成了 $\frac{dy}{dx}$。因而

$$\frac{dy}{dx}=a。$$

一些唯理的数学家们，固执地认为 $dy$ 和 $dx$ 在量上实际只是无限小，$\left[\frac{dy}{dx}\right]$ 只是接近于 $\frac{0}{0}$。正如在 II 中将要更加清楚地指出的那样，他们借此聊以



马克思《导函数的概念》手稿的第一页

（戳记系原稿保存单位加的）

自慰只是幻想。

对这里已考察的情况，还要提及的一个特点是：$\frac{\Delta y}{\Delta x}=a$ 以及同样地 $\frac{dy}{dx}=a$，因而有限差值［之比］的极限值，同时也是微分［之比］的极限值。

2）同一情况的第二个例子是：

$$\begin{matrix} y=x \\ y_1=x_1 \end{matrix};\quad y_1-y=x_1-x;$$

$$\frac{y_1-y}{x_1-x}\text{或}\frac{\Delta y}{\Delta x}=1;\qquad \frac{0}{0}\text{或}\frac{dy}{dx}=1。$$

## II

由于 $y=f(x)$，而 $x$ 的函数又是以**展开的代数表示式**处于等式的右边，所以我们称这个表示式为 $x$ **的原函数**，称通过取差值而得到的初次变形为 $x$ **的预先"导"函数**，把通过**微分过程**最终得到的形式称为 $x$ **的"导"函数**。

1）$y=ax^3+bx^2+cx-e$。

如果 $x$ 增长到 $x_1$，那末

$$y_1=ax_1^3+bx_1^2+cx_1-e,$$

$$y_1-y=a(x_1^3-x^3)+b(x_1^2-x^2)+c(x_1-x)=$$

$$=a(x_1-x)(x_1^2+x_1x+x^2)+b(x_1-x)(x_1+x)+c(x_1-x)。$$

因此

$$\frac{y_1-y}{x_1-x}\text{或}\frac{\Delta y}{\Delta x}=a(x_1^2+x_1x+x^2)+b(x_1+x)+c。$$

**预先"导函数"**

$$a(x_1^2+x_1x+x^2)+b(x_1+x)+c$$

在这里是有限差值之比的**极限值**，就是说，不管这些差值取得多么小，$\frac{\Delta y}{\Delta x}$ 的值总是由这个"导函数"给定的。但是和 I 中情况不同，在这里这个值与微分之比的极限值并不相同①。

如果在函数

$$a(x_1^2+x_1x+x^2)+b(x_1+x)+c$$

---

① 在这篇论文的草稿中接着写道：

"另一方面，现在微分过程在 $x$ 的预先'导'函数中(右边)进行，而同一个过程必然在左边伴随着这个运动"。

中，变量 $x_1$ 减少，直到其减少的**极限**，也就是变为**等于** $x$，那末 $x_1^2$ 变为$x^2$，$x_1x$ 变为 $x^2$，以及 $x_1+x$ 变为 $2x$，从而我们得到$x$ 的"**导**"**函数**

$$3ax^2+2bx+c。$$

这里令人信服地表明：

**第一**，为了得到"导函数"，就必须令 $x_1=x$，所以是**严格数学意义上的** $x_1-x=0$，而无需任何只是无限接近之类的遁辞。

**第二**，由于令 $x_1=x$，于是 $x_1-x=0$。这样一来，就根本没有什么符号化的东西进入"导函数"①。原先通过$x$ 的变化而引入的那个量 $x_1$ 并没有消失，它只是**减少**到了它的最小极限值 $=x$，并且始终是$x$ 的原函数中新引进的元素，它通过部分地和自身相结合，部分地和原函数中的$x$ 相结合，给出了最终"导函数"，也就是给出了**减少到它的最小量的预先"导函数"**。

在最初的(预先)"导"函数中，把 $x_1$ 减少到 $x$，使左边的$\frac{\Delta y}{\Delta x}$变为$\frac{0}{0}$或$\frac{dy}{dx}$，因而

$$\frac{0}{0}\text{或}\frac{dy}{dx}=3ax^2+2bx+c,$$

所以，**导函数**显现为微分之比的**极限值**。

先验的②或符号化的不幸只发生在左边，但由于它现在只是作为一种过程的表示式，它的实际内容早已在等式的右边得到见证，所以已经失掉了它的吓人姿态。

在"导函数"

$$3ax^2+2bx+c$$

中，变量$x$ 处于与原函数(即与 $ax^3+bx^2+cx-e$) 中的$x$ 完全不同的条件之下。所以"导函数"本身又可以作为一个原函数出现，并且通过再一次的微分过程变成另一个"导函数"的母体。只要变量$x$ 并没有从某一个"导函数"中被干脆除掉，那末这种做法就可以一再重复，所以对那些只能用无穷级数表示的$x$ 的函数来说，这种做法可以无限地继续下去。大多数情况都是如此。

符号$\frac{d^2y}{dx^2}$，$\frac{d^3y}{dx^3}$等等只表明最初给定的$x$ 的原函数的"导函数"系谱。只

---

① 在草稿中写着下列句子：

"b)从$x$ 的原函数找出'导函数'的过程是这样进行的：我们先着手建立一个**有限差值**；这给我们提供了一个预先'导函数'，它是$\frac{\Delta y}{\Delta x}$的**极限值**。我们随后进行的微分过程，就是把这个**极限值减少到它的最小量**。在最初的差值中所引入的量$x_1$并没有消失……"。

② "先验的"一词，原文为 transzendentale。

有当人们把这些符号当作运动的**出发点**，而不把它们当作单纯的 $x$ **的逐阶导函数的表示式**看待时，它们才变得很神秘。这时确实显得很奇怪，怎么消失了的量的比值还得重新经历再次的消失呢？而当看到譬如 $3x^2$ 能够同它的母体 $x^3$ 一样经历一次微分过程，那末这里面就没有什么奇怪了。人们本来也可以把 $3x^2$ 当作 $x$ 的原函数而从它出发。

但是**必须注意**，事实上只有象在 I 中 $x$ 仅以一次幂出现的那些等式里，$\frac{\Delta y}{\Delta x}$才是**微分过程**的结局。可是在这里，正如 I 中所已指出的那样，结果是

$$\frac{\Delta y}{\Delta x} = a = \frac{dy}{dx}。$$

所以，在这里通过$\frac{\Delta y}{\Delta x}$所经历的微分过程，实际上并**没有找到新的极限值**；只有当预先"导函数"包含有变量 $x$ 时，因而只有当$\frac{dy}{dx}$保持为某个实在过程的符号时，才有可能找到新的极限值①。

当然，在微分演算中，这决不妨碍符号$\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$等等及其组合也能构成等式的右边。但这时人们也知道，这种纯粹符号的等式，仅仅表明以后对变量的实际函数应进行的那些**运算**。

2）$y = ax^m$。

如果 $x$ 变到 $x_1$，那末 $y_1 = ax_1^m$，以及

$$y_1 - y = a(x_1^m - x^m) = a(x_1 - x)(x_1^{m-1} + x_1^{m-2}x + x_1^{m-3}x^2 + \text{etc.②}$$

直到 $x_1^{m-m}x^{m-1}$ 这一项）。

因而

$$\frac{y_1 - y}{x_1 - x} \text{或} \frac{\Delta y}{\Delta x} = a(x_1^{m-1} + x_1^{m-2}x + x_1^{m-3}x^2 + \cdots + x_1^{m-m}x^{m-1})。$$

如果我们现在把微分过程应用到这个"**预先导函数**"上，以致

$$x_1 = x \text{ 或 } x_1 - x = 0,$$

那末

$x_1^{m-1}$ 变为 $x^{m-1}$；

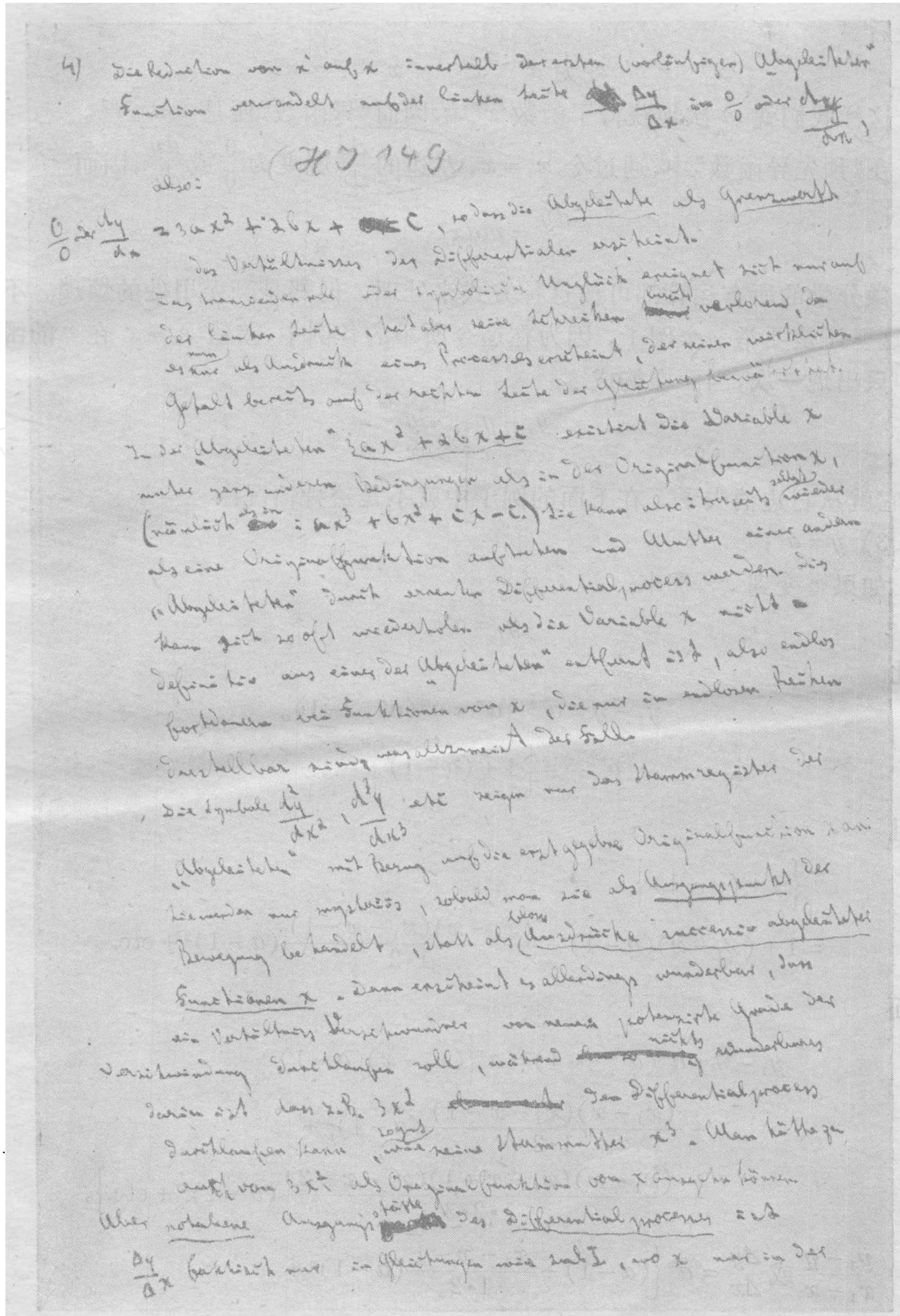
$x_1^{m-2}x$ 变为 $x^{m-2}x = x^{m-2+1} = x^{m-1}$；

---

① 在草稿中上面这句话是这样写的：

"这只能发生在这种地方，那里预先'导'函数含有变量 $x$，因而也能通过它的运动构成一个真正的新值，因而$\frac{dy}{dx}$是某个实在过程的符号。"

② "etc."表示"等等"。



**马克思《导函数的概念》手稿的第四页**

$$x_1^{m-3}x^2 \text{ 变为 } x^{m-3}x^2 = x^{m-3+2} = x^{m-1};$$

而最后

$$x_1^{m-m}x^{m-1} \text{ 变为 } x^{m-m}x^{m-1} = x^{0+m-1} = x^{m-1}。$$

这样我们就 $m$ 次地获得了函数 $x^{m-1}$，因而“导函数”便是 $max^{m-1}$。

在“预先导函数”中，通过令 $x_1 = x$，左边的 $\frac{\Delta y}{\Delta x}$ 就变为 $\frac{0}{0}$ 或 $\frac{dy}{dx}$，因而

$$\frac{dy}{dx} = max^{m-1}。$$

微分学的所有运算都可按这种方式来处理，但那是毫无用处的烦琐。不过，这里还是要举一个例子，因为在迄今所举的各例中，差值 $x_1 - x$ 在 $x$ 的函数中**只**出现**一次**，因而在构成

$$\frac{y_1 - y}{x_1 - x} \text{或} \frac{\Delta y}{\Delta x}$$

时，它就从右边消失了。在下面的例子中就不是这种情况：

3） $y = a^x$；

如果 $x$ 变到 $x_1$，那末

$$y_1 = a^{x_1}。$$

因此

$$y_1 - y = a^{x_1} - a^x = a^x(a^{x_1-x} - 1)。$$

$$a^{x_1-x} = \left\{1 + (a-1)\right\}^{x_1-x},$$

并且

$$\left\{1 + (a-1)\right\}^{x_1-x} =$$
$$= 1 + (x_1 - x)(a-1) + \frac{(x_1-x)(x_1-x-1)}{1\cdot 2}(a-1)^2 + \text{etc.}$$

因而

$$y_1 - y = a^x(a^{x_1-x} - 1) = a^x\left\{(x_1-x)(a-1) + \right.$$
$$+ \frac{(x_1-x)(x_1-x-1)}{1\cdot 2}(a-1)^2 +$$
$$\left. + \frac{(x_1-x)(x_1-x-1)(x_1-x-2)}{1\cdot 2\cdot 3}(a-1)^3 + \text{etc.}\right\}。$$

$$\therefore \quad \frac{y_1 - y}{x_1 - x} \text{或} \frac{\Delta y}{\Delta x} = a^x\left\{(a-1) + \frac{x_1-x-1}{1\cdot 2}(a-1)^2 + \right.$$
$$\left. + \frac{(x_1-x-1)(x_1-x-2)}{1\cdot 2\cdot 3}(a-1)^3 + \text{etc.}\right\}。$$

如果现在 $x_1=x$,因而 $x_1-x=0$,那末我们便得到"**导函数**"

$$a^x\left\{(a-1)-\frac{1}{2}(a-1)^2+\frac{1}{3}(a-1)^3-\text{etc.}\right\}。$$

因而

$$\frac{dy}{dx}=a^x\left\{(a-1)-\frac{1}{2}(a-1)^2+\frac{1}{3}(a-1)^3-\text{etc.}\right\}。$$

如果我们称花括号中的常量之和为 $A$,那末

$$\frac{dy}{dx}=Aa^x;$$

但是这个 $A$ = 底数 $a$①的耐普尔对数,所以

$$\frac{dy}{dx},\text{ 或当用 } y \text{ 的值代入时} \frac{da^x}{dx}=\log a\cdot a^x,$$

从而

$$da^x=\log a\cdot a^x dx。$$

**补充**

1）以前曾经研究过这样的情况：因子$(x_1-x)$在"**预先导函数**"中，即在有限差值等式中只出现一次。所以用 $x_1-x$ 除两边就构成

$$\frac{y_1-y}{x_1-x}\text{或}\frac{\Delta y}{\Delta x},$$

这个因子便从 $x$ 的函数中消去了。

2）(在例子 $d(a^x)$中)研究过这样的情况：在构成$\frac{\Delta y}{\Delta x}$之后，因子 $(x_1-x)$ 还保留在 $x$ 的函数中。

3）还要研究一下这种情况：因子 $x_1-x$ **不是直接**从（"预先导函数"的）最初差值等式中演化出来的。

$$y=\sqrt{a^2+x^2},$$

$$y_1=\sqrt{a^2+x_1^2},$$

$$y_1-y=\sqrt{a^2+x_1^2}-\sqrt{a^2+x^2};$$

我们用 $x_1-x$ 去除 $x$ 的函数,因而也用它去除左边。于是

$$\frac{y_1-y}{x_1-x}\left(\text{或}\frac{\Delta y}{\Delta x}\right)=\frac{\sqrt{a^2+x_1^2}-\sqrt{a^2+x^2}}{x_1-x}。$$

为了把根式从分子中消掉，分子和分母都用$\sqrt{a^2+x_1^2}+\sqrt{a^2+x^2}$去乘，得：

---

① 这里的"底数 $a$"表示 $a$ 是指数函数 $y=a^x$ 的底数。

$$\frac{\Delta y}{\Delta x}=\frac{a^2+x_1^2-(a^2+x^2)}{(x_1-x)(\sqrt{a^2+x_1^2}+\sqrt{a^2+x^2})}=$$
$$=\frac{x_1^2-x^2}{(x_1-x)(\sqrt{a^2+x_1^2}+\sqrt{a^2+x^2})}。$$

但是

$$\frac{x_1^2-x^2}{(x_1-x)(\sqrt{a^2+x_1^2}+\sqrt{a^2+x^2})}=$$
$$=\frac{(x_1-x)(x_1+x)}{(x_1-x)(\sqrt{a^2+x_1^2}+\sqrt{a^2+x^2})}。$$

因而

$$\frac{\Delta y}{\Delta x}=\frac{x_1+x}{\sqrt{a^2+x_1^2}+\sqrt{a^2+x^2}}。$$

如果 $x_1$ 变为 $=x$,或者 $x_1-x=0$,那末

$$\frac{dy}{dx}=\frac{2x}{2\sqrt{a^2+x^2}}=\frac{x}{\sqrt{a^2+x^2}}。$$

因此

$$dy\text{ 或 }d\sqrt{a^2+x^2}=\frac{xdx}{\sqrt{a^2+x^2}}。$$

# [关 于 符 号 $\frac{dy}{dx}$]①

## I

已经指出过，例如

1）当

$$y = x^m = f(x), \quad y_1 = x_1^m,$$

我们就得到

$$\frac{dy}{dx} \text{或} \frac{0}{0} = mx^{m-1}。$$

还曾指出过，导函数 $f'(x)$ 或 $mx^{m-1}$ 是由原先的

$$f(x) = x^m$$

通过令 $x_1 = x$ 即 $x_1 - x = 0$ 而得到的。

但是这同一个令 $x_1 - x = 0$ 或者令 $x_1 = x$，就把$\frac{y_1 - y}{x_1 - x}$变成了$\frac{0}{0}$，而我们用$\frac{dy}{dx}$来代替它，为的是要指明什么是这个$\frac{0}{0}$的来源，也就是说什么样的实际差值之比——在上述情况下就是$\frac{y_1 - y}{x_1 - x}$最终变成了$\frac{0}{0}$。

当我们得到结果

$$\frac{0}{0} = mx^{m-1} = f'(x),$$

并且等式左边的这个结果$\frac{0}{0}$是由于从右边的变量 $x$ 出发的运动所产生的时候，这样的代替就显得更加合理。

$\frac{0}{0}$可以 = 任何一个量 $X$，因为

$$0 = X \cdot 0 = 0。$$

① 这是《导函数的概念》这篇论文草稿中的最后一段。

由于这里的$\frac{0}{0}$不等于任意的$X$，而$=mx^{m-1}$，所以用$\frac{dy}{dx}$或$\frac{df(x)}{dx}$来表明，符号$\frac{0}{0}$是由确定的$f(x)$中自变量$x$的什么样运动得出的。

2）但是在这样一下子就全都把$\frac{dy}{dx}$的意义固定下来之后，它的特殊值自然会随着$f(x)$本身的确定形式而变化，从而一旦我们站在微分演算的地盘上，课题就反过来表现为，对于$\frac{dy}{dx}$我们应当通过微分去找它的特殊值，象上面的$=mx^{m-1}$，或者与它相应的导函数。

# [论微分。论文及底稿]

## [论　微　分]①

### I

1）设要加以微分的是 $f(x)$ 或 $y=uz$。$u$ 和 $z$ 是自变量 $x$ 的两个函数；而相对于依赖它们的函数 $y$ 来说，它们又是自变量。因此 $y$ 也依赖于 $x$。

$$y_1=u_1z_1,$$

$$y_1-y=u_1z_1-uz=z_1(u_1-u)+u(z_1-z),$$

$$\frac{y_1-y}{x_1-x}\text{或}\frac{\Delta y}{\Delta x}=z_1\frac{u_1-u}{x_1-x}+u\frac{z_1-z}{x_1-x}=\frac{z_1\Delta u}{\Delta x}+\frac{u\Delta z}{\Delta x}。$$

如果在右边 $x_1$ 变为 $=x$，因而 $x_1-x=0$，那末 $u_1-u=0$，$z_1-z=0$，所以 $z_1\frac{u_1-u}{x_1-x}$中的因子 $z_1$ 也就变为 $z$，最后在左边 $y_1-y=0$。因此：

A）$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$。

用各项的公分母 $dx$ 乘这个等式，它就变为

B）$dy$ 或 $d(uz)=zdu+udz$。

2）先来研究等式 A)：

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}。$$

在只有一个依赖于 $x$ 的因变量的那些等式中，最后结果总是

$$\frac{dy}{dx}=f'(x),$$

而在 $f'(x)$这个 $f(x)$的一阶导函数中，总是不出现任何符号表示式。例如，当 $x^m$ 为自变量 $x$ 的原函数时，$mx^{m-1}$就是这样。为了把 $f(x)$变为 $f'(x)$，它必须经历微分过程；正是由于这种微分过程，使得 $f'(x)$的化身$\frac{0}{0}$或$\frac{dy}{dx}$作为符

① 1881年，马克思把这篇论文寄给了恩格斯。

号等价物在左边，即在实在微系数 $f'(x)$ 的对面跳了出来。另一方面，$\frac{0}{0}$ 或 $\frac{dy}{dx}$ 也就在 $f'(x)$ 那里找到了自己的实在等价物。

在等式 A)中则相反，$uz$ 的一阶导函数 $f'(x)$ 本身就包含有符号微系数，因此两边都出现符号微系数而任何一边都不是实在值。但是，由于处理 $uz$ 的方法和以前处理只有一个自变量 $x$ 的函数的方法相同，因此结果上的这种明显差异，显然来源于起始函数即 $uz$ 本身的特殊性质。关于这方面详见 3)。

但是，还要先研究一下在等式 A)的**推导**中有没有问题。

在等式的右边，由于 $x_1 = x$，以及 $x_1 - x = 0$，所以

$$\frac{u_1 - u}{x_1 - x} \text{或} \frac{\Delta u}{\Delta x} \text{和} \frac{z_1 - z}{x_1 - x} \text{或} \frac{\Delta z}{\Delta x}$$

就变成了 $\frac{0}{0}$，$\frac{0}{0}$。但是，我们不假思索地用 $\frac{du}{dx}$，$\frac{dz}{dx}$ 代替了 $\frac{0}{0}$，$\frac{0}{0}$。是否允许这样做呢？问题在于两个 $\frac{0}{0}$ 在这里分别作为**变量 $u$ 和 $z$ 的乘数**；而在只有一个因变量的情况下，所得出的唯一的符号微系数——$\frac{0}{0}$ 或 $\frac{du}{dx}$——除了常量 1 以外是没有任何其他乘数的。

如果我们把 $\frac{du}{dx}$，$\frac{dz}{dx}$ 的原来的关键性形式代入右边，那末右边便变为 $z\frac{0}{0} + u\frac{0}{0}$。倘若我们分别用伴随着 $z$ 和 $u$ 的 $\frac{0}{0}$ 的分子去乘 $z$ 和 $u$，那末就会得到 $\frac{0}{0} + \frac{0}{0}$；而由于变量 $z$ 和 $u$ 本身已变成了 $=0$，所以它们的导函数也是这样，因而最后：

$$\frac{0}{0} = 0 \text{ 而不是 } z\frac{du}{dx} + u\frac{dz}{dx}。$$

然而这个演算过程在数学上是错误的。

例如，我们取

$$\frac{u_1 - u}{x_1 - x} = \frac{\Delta u}{\Delta x},$$

人们不是首先得到分子 $=0$，因为这样就要从设 $u_1 - u = 0$ 出发；而只是因为分母，即自变量 $x$ 的差值亦即 $x_1 - x$ 变为 $=0$，分子或 $u_1 - u$ 才变为 $=0$ 的。

所以面对着变量 $u$ 和 $z$ 所出现的不是 0，而是 $\left(\frac{0}{0}\right)$，**在这个形式中它的分子始终不能和它的分母分开**。因此作为乘数的 $\frac{0}{0}$，只有当

$$\frac{0}{0} = 0$$

时，才能使它的系数变成 0。

即使在普通代数中，如果一个乘积 $P\cdot\frac{m}{n}$ 采取了 $P\cdot\frac{0}{0}$ 的形式，就不假思索地作出结论说它**必须是** $=0$，这也是错误的，虽然这里它总**可以被**令 $=0$，因为我们可以任意地从分子或分母出发来使它变成零。

例如 $P\cdot\frac{x^2-a^2}{x-a}$。如果 $x^2=a^2$，因而 $x^2-a^2=0$，那末便得到 $P\cdot\frac{0}{0}=\frac{0}{0}$，这个 $\frac{0}{0}$ 可以令 $=0$。因为 $\frac{0}{0}$ 可以象它等于任何其他数一样，也可以等于 0。

反之，如果把 $x^2-a^2$ 分解因子，那末我们得到：

$$P\cdot\frac{x-a}{x-a}\cdot(x+a)=P(x+a)\text{，并由于}x^2=a^2\text{，[它又]}=2Pa\text{。}$$

逐次微分表明，只有在完全确定的条件下，$\frac{0}{0}$ 才会变为 $=0$。例如 $x^3$ 的逐次微分，在第三次求导中变量 $x$ 完全消失而代之以一个常量之后，第四次求导才使 $\frac{0}{0}$ 变成 $=0$。

但是在我们的情况中，已知这些 $\frac{0}{0}$，$\frac{0}{0}$ 的来源分别是差值表示式 $\frac{\Delta z}{\Delta x}$，$\frac{\Delta u}{\Delta x}$，所以一开始也理应给它们穿上制服 $\frac{dz}{dx}$，$\frac{du}{dx}$。

3）在以前讨论过的等式如 $y=x^m$，$y=a^x$ 等等中，处在 $x$ **的原函数**对面的是依赖于 $x$ 的“因变量” $y$。

$y=uz$ 的两边都为“因变量”所占据。如果说 $y$ 直接依赖于 $u$ 和 $z$，那末 $u$ 和 $z$ 则又依赖于 $x$。原函数 $uz$ 的这种特殊性质，必然也要给它的“导函数”打上自己的烙印。

$u$ 是 $x$ 的一个函数，而 $z$ 是 $x$ 的另一个函数，这可表示为：

$$\begin{matrix} u=f(x), \\ z=\varphi(x); \end{matrix}\quad\text{因此}\quad\begin{matrix} u_1-u=f(x_1)-f(x), \\ z_1-z=\varphi(x_1)-\varphi(x)\text{。} \end{matrix}$$

但是起始等式 [$y=uz$] 并没有为 $f(x)$ 和 $\varphi(x)$ 提供 $x$ 的原函数，也就是没有给出用 $x$ 表达的确定值。因此，$u$ 和 $z$ 只不过是依赖于 $x$ 的函数的名称和符号；所以通过对 $uz$ 的求导过程首先所能得出的，也不过是**这个依赖关系的一般形式**：

$$\frac{u_1-u}{x_1-x}=\frac{f(x_1)-f(x)}{x_1-x},\qquad\frac{z_1-z}{x_1-x}=\frac{\varphi(x_1)-\varphi(x)}{x_1-x}\text{。}$$

如果这个过程达到了令 $x_1=x$ 因而 $x_1-x=0$ 这样一点，那末这些一般形式就变为：

$$\frac{du}{dx}=\frac{df(x)}{dx}, \qquad \frac{dz}{dx}=\frac{d\varphi(x)}{dx},$$

于是，$\frac{du}{dx}$，$\frac{dz}{dx}$就作为符号微系数在“导函数”中出现了。

然而在只有一个因变量的那些等式中的$\frac{dy}{dx}$,同这里的$\frac{du}{dx}$，$\frac{dz}{dx}$相比，丝毫没有什么不同的内容。它也不过是

$$\frac{y_1-y}{x_1-x}=\frac{f(x_1)-f(x)}{x_1-x}$$

的符号微分表示式。

虽然当符号微系数$\frac{du}{dx}$，$\frac{dz}{dx}$出现**在导函数本身之内**，也就是出现在微分等式的右边时它们的性质决不会改变，但是它们的作用以及等式的特性却因此而发生了变化。

如果我们一般地用 $f(x)$ 表示原函数 $uz$，$f'(x)$ 就表示它的一阶“导函数”,那末

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

就显现为

$$\frac{dy}{dx}=f'(x)。$$

对于只有一个因变量的等式，我们得到与此相同的一般形式。在这两种情况下，$\frac{dy}{dx}$这个[作为微分演算的]起始形式都是在把 $f(x)$变为 $f'(x)$的求导过程中产生出来的。所以，一旦 $f(x)$ 变为 $f'(x)$，那末$\frac{dy}{dx}$就作为 $f'(x)$ 特有的符号表示式,作为它的化身或符号等价物出现在它的对面。

因此在这两种情况下，$\frac{dy}{dx}$都扮演着**同一个角色**。

$\frac{du}{dx}$，$\frac{dz}{dx}$则不一样。它们包含在$f'(x)$中,并和其他元素一道，找到了$\frac{dy}{dx}$作为它们的符号表示式或者符号等价物。但它们本身没有与之相对的$f'(x)$，$\varphi'(x)$，而实际上却是 $f'(x)$，$\varphi'(x)$ 的符号化身。它们生来是片面的,是没有实体的阴影,是没有实在微系数的符号微系数,也就是没有相应等价“导函数”的符号微系数。这样，符号微系数便成了**独立出发点**，而它的实在等价物却还有待于去寻找。所以主动性就从右边代数的一端移到了左边符号的一端。然而这样一来，微分演算也就表现为一种特殊的计算方法，这种计算方法早已

独立地在它固有的地盘上运用了。因为这种演算的出发点$\frac{du}{dx}$和$\frac{dz}{dx}$是仅属微分演算的、并表示其特性的数学量。这种方法上的转变，在这里是作为 $uz$ 的代数微分的结果而得出的。代数方法就自动地转变为与它对立的微分方法①。

那末与符号微系数$\frac{du}{dx}$，$\frac{dz}{dx}$相应的“导函数”是什么呢？起始等式 $y=uz$ 没有为解决这个问题提供任何资料。可是，如果我们用任意的 $x$ 的原函数来代替 $u$ 和 $z$，例如：

$$u=x^4,\quad z=x^3+ax^2,$$

那末这个问题总是可以回答的。

但这样一来，符号微系数$\frac{du}{dx}$，$\frac{dz}{dx}$也就立刻变为**运算符号**，即变为对 $x^4$ 和 $x^3+ax^2$ 求它们的“导函数”可以实行的过程的符号。符号微系数原先是作为“导函数”的符号表示式，因而是已经实行了微分运算的符号表示式而产生的，现在却起着还有待于去实行的微分运算的符号的作用。

同时，等式

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

变成了一般的符号运算等式。因为两边都有符号，所以它一开始就纯粹是符号化的。

我还要提到一点，从十八世纪初直到今天，微分演算的一般课题通常是这样表述的：要给符号微系数找出实在等价物。

4）

A）$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$。

---

① 在这篇论文的草稿中，整个一段是这样写的：

“$\frac{du}{dx}$，$\frac{dz}{dx}$则相反。它们产生于导函数之内，并和导函数中的其余元素一道，找到了$\frac{dy}{dx}$作为它们固有的符号表示式，从而找到了它们的符号等价物。它们本身存在着但没有等价的、实际的微系数，也就是说，没有导函数 $f'(x)$，$\varphi'(x)$，而实际上它们是 $f'(x)$，$\varphi'(x)$ 的符号表示式。它们都是一些现成的微分符号，其实在值表现为还有待去找的实体的阴影。所以问题就被悄悄地颠倒了过来。符号微系数变成了独立的**出发点**，而它的等价物，实际的微系数或相应的导函数倒是还有待去寻找的。这样一来，主动性便从右端移到了左端。由于这个方法上的转变是从函数 $uz$ 的代数运动中发生的，所以它本身是用代数方法论证的”。

这显然不是等式 A)的最简单表示式，因为它的各项都有共同的分母 $dx$。去掉这个分母后就得到：

B) $d(uz)$ 或 $dy = zdu + udz$。

在 B)中，它来源于 A) 的任何痕迹都已消失。所以 B) 既适用于 $u$ 和 $z$ 依赖于 $x$ 的情况，同样也完全适用于 $u$ 和 $z$ 与 $x$ 没有任何关系而只是相互依赖的情况。它一开始就是一个符号等式，并且一开始就能作为一个符号运算等式来使用。在后一种情况下它表明，如果

$$y = zu \text{ etc.},$$

即等于任意多个变量的乘积，那末 $dy$ = 一些乘积之和，在这些乘积中我们依次把其中的一个因子作为变量，而把其余因子作为常量来处理。

就我们进一步研究 $y$ 的微分的目的而言，形式 B)却并不适合。因此，如果令

$$u = x^4, \quad z = x^3 + ax^2,$$

那末正象以前在只有一个因变量的等式中所已证实了的那样，

$$du = 4x^3dx, \quad dz = (3x^2 + 2ax)dx。$$

把 $du, dz$ 的这些值代入等式 A)中，则

A) $\dfrac{dy}{dx} = (x^3 + ax^2)\dfrac{4x^3dx}{dx} + x^4\dfrac{(3x^2 + 2ax)dx}{dx}$；所以

$$\frac{dy}{dx} = (x^3 + ax^2)4x^3 + x^4(3x^2 + 2ax);$$

因而

$$dy = \left\{(x^3 + ax^2)4x^3 + x^4(3x^2 + 2ax)\right\}dx。$$

花括号中的表示式是 $uz$ 的一阶导函数；但是由于 $uz = f(x)$，它的导函数就 $= f'(x)$；如果我们把 $f'(x)$ 放在代数函数的位置上，那末就有

$$dy = f'(x)dx。$$

我们已经从只有一个因变量的任意等式中得到了同样的结果，例如：

$$y = x^m,$$

$$\frac{dy}{dx} = mx^{m-1} = f'(x),$$

$$dy = f'(x)dx。$$

一般地我们就有：如果 $y = f(x)$，那末无论这个 $x$ 的函数是一个用 $x$ 表达的原函数，或者它还含有因变量，总是 $dy = df(x)$ 以及 $df(x) = f'(x)dx$。因而：

B) $dy = f'(x)dx$ 是 $y$ 的微分的普遍形式。即使所给的函数是 $f(x,z)$，也

就是**两个互不依赖的变量**的一个函数，这也是可以立刻予以证明的。但是对于我们的目的来说，这是多余的。

## II

1）微分

$$dy=f'(x)dx$$

比导出它的微系数

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

一开始就显得更为可疑。

在$\frac{dy}{dx}=\frac{0}{0}$中，分母和分子不可分地联系在一起，而在 $dy=f'(x)dx$ 中它们显然是分开的，以致只能得出这样一个结论：它无非是给“毫无办法”①的

$$0=f'(x)\cdot 0 \text{ 或 } 0=0$$

化了装的一个表示式。

十九世纪头三分之一年代里的一位法国数学家布夏拉，和那位熟知的“文雅的”法国人完全不同，他清楚地把微分方法和拉格朗日的代数方法联系起来。他说：

如果以$\frac{dy}{dx}=3x^2$为例，那末“$\frac{dy}{dx}$又称$\frac{0}{0}$，或者不如说它的值 $3x^2$ 是函数 $y$ 的微系数。由于$\frac{dy}{dx}$是代表极限 $3x^2$ 的符号，所以 $dx$ **必须总是处在** $dy$ **的下面**，但是**为了便于进行代数运算**，我们把$\frac{dy}{dx}$当做普通的分式，并把$\frac{dy}{dx}=3x^2$当做一个普通的等式来处理，于是从等式中消除分母后我们得到下面的结果：

$$dy=3x^2dx,$$

这个表示式叫做 $y$ 的微分”。

这样，为了“便于进行代数运算”，人们引进了一个分明是错误的公式，并称之为“微分”。

实际情况并不这样坏。

在$\frac{0}{0}$中分子不能和分母分开，这是为什么呢？因为只有两者不分开才表示一个比值，在这个情况下才表示减少到其绝对最小值的比值：

---

① “毫无办法”一词，原文为“nix ze wolle”。

$$\frac{y_1-y}{x_1-x}=\frac{f(x_1)-f(x)}{x_1-x},$$

其中分子因分母变为 0 而变为 0。如果把两个 0 分开，那就失去了它们作为符号的意义，失去了含义。

但是，一旦 $x_1-x=0$ 获得了以 $dx$ 表示的那种形式，而且始终把它看作自变量 $x$ 的消失了的差值，因而也把 $dy$ 看作 $x$ 的函数，即因变量 $y$ 的消失了的差值，那末分母与分子分开就完全是可以允许的运算了。现在不管 $dx$ 放在那里，这样一种位置的改变不会影响 $dy$ 跟它的关系。所以，$dy=f'(x)dx$ 是作为

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

的另一种形式出现在我们面前，并且总是可以变为后者。

2）微分 $dy=f'(x)dx$ 是从 A)通过直接的代数推导得来的(见 I，4)。等式 A)的代数推导早已表明，微分符号，在这里就是符号微系数，原来只是作为用代数方法完成了的微分过程的符号表示式而得出的，现在必须重新转变为独立的出发点，转变为只是尚待实行的运算的符号，或者说转变为运算符号。因此用代数方法得来的符号等式，也就转变为符号运算等式。

因此，我们有双重理由把 $dy=f'(x)dx$ 作为符号运算等式看待。对此，现在我们事前就知道，如果

$$y=f(x),\quad dy=df(x),$$

并对 $f(x)$完成了由 $df(x)$所指出的微分运算，那末其结果就是 $dy=f'(x)dx$。最后由此得出

$$\frac{dy}{dx}=f'(x)。$$

但也只有从微分充当演算出发点的那一时刻起，代数的微分方法的转变才告完成，因此微分演算本身就显现为一种对变量的完全独特的、专门的计算方法。

为了把这一点说清楚，我把我所应用的代数方法作一个一般性的总结。这里以 $f(x)$代替用 $x$ 表达的确定的代数表示式，并用 $f^1(x)$标记“预先导函数”(见第一篇稿子①)，以便同最终“导函数” $f'(x)$相区别。这样，如果

$$f(x)=y,\quad f(x_1)=y_1,$$

[那末]

① 即《导函数的概念》这篇论文。

$$f(x_1)-f(x)=y_1-y \text{ 或 } \Delta y,$$
$$f^1(x)(x_1-x)=y_1-y \text{ 或 } \Delta y。$$

预先导函数 $f^1(x)$完全象它的因子 $x_1-x$ 那样，**必须**包含用 $x_1$ 和 $x$ 表达的表示式

$$f^1(x)=\frac{y_1-y}{x_1-x} \text{或} \frac{\Delta y}{\Delta x}。$$

**唯一的例外是**：$f(x)$是**一次**的原函数。

现在如果在 $f^1(x)$ 中令

$$x_1=x, \text{ 因而 } x_1-x=0,$$

那末得到：

$$f'(x)=\frac{0}{0} \text{或} \frac{dy}{dx},$$

并且最后

$$f'(x)dx=dy \text{ 或 } dy=f'(x)dx。$$

所以 $y$ 的微分是代数演化的终点；它又将是在自己的地盘上活动的微分演算的出发点。孤立地考察，也就是说不和它的等价物联系起来，$dy$ 这个 $y$ 的微分元在这里立刻起着同 $\Delta y$ 在代数方法中相同的作用；$dx$ 这个 $x$ 的微分元则起着象那里的 $\Delta x$ 一样的作用。

如果我们从

$$\frac{\Delta y}{\Delta x}=f^1(x)$$

中解除分母，那末：

I) $\Delta y=f^1(x)\Delta x$

与此相反，如果我们把用代数推导得来的微分演算，作为已完成的、独特的计算方法并由此出发，那末我们就直接从 I)的微分表示式，即从：

II) $dy=f'(x)dx$ 开始。

3）由于在只有一个因变量的最初等函数的代数处理中，就已经出现了微分的符号等式，所以方法上的转换比在例子

$$y=uz$$

中所发生的似乎可以大为简化。

最初等的函数是一次函数。它们是：

a) $y=x$；它给出微系数$\frac{dy}{dx}=1$，因而微分 $dy=dx$。

b) $y=x\pm ab$；它给出微系数$\frac{dy}{dx}=1$，因而又是微分 $dy=dx$。

c) $y=ax$；它给出微系数 $\frac{dy}{dx}=a$，因而微分 $dy=adx$。

如果我们取最简单的情况 a)，那末：

$$y=x,$$

$$y_1=x_1;$$

$$y_1-y \text{ 或 } \Delta y=x_1-x \text{ 或 } \Delta x。$$

I) $\frac{y_1-y}{x_1-x}$ 或 $\frac{\Delta y}{\Delta x}=1$；因而也是 $\Delta y=\Delta x$。如果在 $\frac{\Delta y}{\Delta x}$ 中令 $x_1=x$ 或 $x_1-x=0$，那末：

II) $\frac{0}{0}$或$\frac{dy}{dx}=1$；因此 $dy=dx$。

一旦我们得到了 I)，即得到了$\frac{\Delta y}{\Delta x}=1$，那末我们一开始就只得在左边继续运算下去，因为右边是常量 1。这样一来，将主动性从右边移向左边的这种**方法上的转换**，似乎从一开始起就一下子全都证明了。事实上这是代数方法本身的头一个结论。

让我们仔细地研究一下这个问题。

实际结果曾经是：

I) $\frac{\Delta y}{\Delta x}=1$。

II) $\frac{0}{0}$或$\frac{dy}{dx}=1$。

由于 I) 和 II) 两者都导致同样的结果，所以我们可以任择其一。无论如何，看来令 $x_1-x=0$ 是多余的因而是任意的运算。再者，如果我们对 II) 从左边出发继续进行运算，那末由于右边已经“毫无办法”，所以就有

$$\frac{0}{0}\text{或}\frac{d^2y}{dx^2}=0。$$

最后的结论势必为$\frac{0}{0}=0$，所以这样获得$\frac{0}{0}$的方法是错误的。因为第一次政变①，它并没有得出什么新东西，而第二次则导致了虚无。

最后，我们从代数学知道，如果两个等式的右边相等，那末其左边也必须相等。由此得出

$$\frac{dy}{dx}=\frac{\Delta y}{\Delta x}。$$

但由于 $x$ 以及依赖于它的 $y$ 两者都是变量，所以 $\Delta x$ 虽然是一个有限差值，却可以无限地缩小，换句话说，要它怎样小就可以怎样小地**接近于** 0，也就是说

① “政变”原文为法文“coup”。

变为**无限小**；因而依赖于它的 $\Delta y$ 也是如此。此外，由于$\frac{dy}{dx}=\frac{\Delta y}{\Delta x}$，所以$\frac{dy}{dx}$实际上并不是毫无约束的$\frac{0}{0}$；相反，只要$\frac{\Delta y}{\Delta x}$起着与普通差值演算不同的无限小差值之比的作用，$\frac{dy}{dx}$就是$\frac{\Delta y}{\Delta x}$的节日盛装。

但微分 $dy=dx$ 本身并无任何意义，或者更确切些说，它的意义至多只有象我们在分析 $\frac{dy}{dx}$ 时对这两个微分元所发现的那么一些。如果我们在刚才所赋予的那个解释下采用它，那末我们就已经能够用微分进行许多巧妙的运算。正如 $adx$ 在决定抛物线的次切线①时的作用所表明的那样。为此不需要真正地了解 $dx$，$dy$ 的本性。

4）在我转到第 III 部分，即对微分演算的历史发展过程极其扼要地勾划一个轮廓以前，我还要为迄今所应用的代数方法再举一个例子。为了确切地说明这个方法，我把确定的函数放在左边，这一边总是主动的一边，因为我们是从左写到右的，因而一般的等式也就是：

$$x^m+Px^{m-1}+\text{etc.}+Tx+U=0,$$

而不是

$$0=x^m+Px^{m-1}+\text{etc.}+Tx+U。$$

如果函数 $y$ 和自变量 $x$ 分开在两个等式中，在第一个等式中 $y$ 表示为变量 $u$ 的函数，在第二个等式中 $u$ 表示为 $x$ 的函数，现在**要求出两者共同的符号微系数**。假定：

1）$3u^2=y$，　　于是　　$3u_1^2=y_1$，

2）$x^3+ax^2=u$；　　　　$x_1^3+ax_1^2=u_1$。

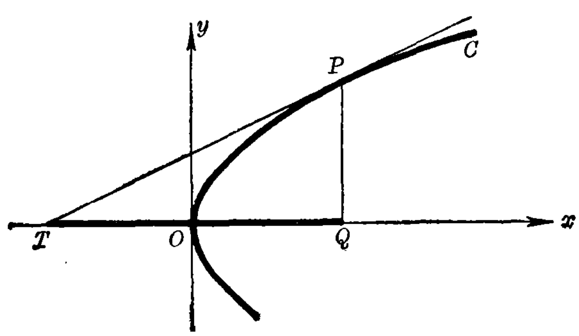
我们先处理等式 1)：

$$3u_1^2-3u^2=y_1-y,$$

$$3(u_1^2-u^2)=y_1-y,$$

---

① 亦称切线影。如下图所示，抛物线$C$在点$P$的次切线是指有向线段 $TQ$。



$$3(u_1-u)\ (u_1+u)=y_1-y,$$

$$3(u_1+u)=\frac{y_1-y}{u_1-u}\text{或}\frac{\Delta y}{\Delta u}。$$

如果在左边令 $u_1=u$，因而 $u_1-u=0$，那末

$$3\ (u+u)\ =\frac{dy}{du},$$

$$3(2u)=\frac{dy}{du},$$

$$6u=\frac{dy}{du}。$$

如果我们现在将 $u$ 的值 $x^3+ax^2$ 代入，那末：

3) $6\ (x^3+ax^2)\ =\frac{dy}{du}$。

现在我们转到等式 2)，于是：

$$x_1^3+ax_1^2-x^3-ax^2=u_1-u,$$

$$(x_1^3-x^3)+a(x_1^2-x^2)=u_1-u,$$

$$(x_1-x)\ (x_1^2+x_1x+x^2)+a\ (x_1-x)\ (x_1+x)=u_1-u,$$

$$(x_1^2+x_1x+x^2)+a(x_1+x)=\frac{u_1-u}{x_1-x}\text{或}\frac{\Delta u}{\Delta x}。$$

如果我们在左边令 $x_1=x$，于是 $x_1-x=0$；因而

$$(x^2+xx+x^2)+a(x+x)=\frac{du}{dx}。$$

4) $3x^2+2ax\ =\ \frac{du}{dx}$。

如果现在将 3)和 4)两等式相乘，则：

5) $6\ (x^3+ax^2)\ (3x^2+2ax)=\frac{dy}{du}\cdot\frac{du}{dx}=\frac{dy}{dx}$。

这样，就用代数方法找到了运算公式

$$\frac{dy}{dx}=\frac{dy}{du}\cdot\frac{du}{dx},$$

有时它也适用于带有两个自变量的等式。

---

一个在确定的函数中得到证实了的推导，可以转化为极其一般的形式，这并不是什么不可思议的事。上面的例子就说明了这一点。假定：

$$\begin{array}{lll}1)\ y=f(u), & \multirow{2}{*}{那末}\ y_1=f(u_1), & \multirow{2}{*}{因而}\ y_1-y=f(u_1)-f(u),\\ 2)\ u=\varphi(x), & u_1=\varphi(x_1), & u_1-u=\varphi(x_1)-\varphi(x)。\end{array}$$

从 1)中的差值得到

$$\frac{y_1-y}{u_1-u}=\frac{f(u_1)-f(u)}{u_1-u};\quad \frac{dy}{du}=\frac{df(u)}{du},$$

但由于 $df(u)=f'(u)du$，所以

$$\frac{dy}{du}=\frac{f'(u)du}{du};$$

由此得出：

3) $\frac{dy}{du}=f'(u)$。

从 2)中的差值得到

$$\frac{u_1-u}{x_1-x}=\frac{\varphi(x_1)-\varphi(x)}{x_1-x},\quad \frac{du}{dx}=\frac{d\varphi(x)}{dx},$$

但由于 $d\varphi(x)=\varphi'(x)dx$，所以

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

因此：

4) $\frac{du}{dx}=\varphi'(x)$。

如果我们用等式 4)乘 3)，则：

5) $\frac{dy}{du}\cdot\frac{du}{dx}$或$\frac{dy}{dx}=f'(u)\cdot\varphi'(x)$ 证毕。

---

III 同这第二部分的结尾，要在博物馆里查阅了约翰·兰登[的著作]之后才能续完。

---

## [一　稿]

当我们转到对 $f(u,z)$ [$=uz$] 进行微分时，其中变量 $u$ 和 $z$ 都是 $x$ 的函数，与以前只有一个**因变量**即 $y$ 的情况不同，我们在两边都将得到微分表示式，即：

**第一步**得到：

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx};$$

**第二步化为：**

$$dy=zdu+udz。$$

这后一个式子具有与一个因变量例如 $dy=max^{m-1}dx$ 的情况不一样的形式，因为那里$\frac{dy}{dx}$立刻给我们一个摆脱了微分符号的 $f'(x)=max^{m-1}$，而在 $dy=zdu+udz$ 中就绝不是这样。我们从只有一个因变量的等式中已经一下子就全都看到，如何通过实际的设置差值和随后的把它扬弃获得 $x$ 的导函数，在上述情况中就是 $x^m$ 的导函数，以及如何同时给这导函数得出符号等价物$\frac{0}{0}=\frac{dy}{dx}$。至于令$\frac{0}{0}=\frac{dy}{dx}$，这里非但是容许的，而且也是必要的。因为$\frac{0}{0}$按其固有的原始形式＝任何一个量，这由于$\frac{0}{0}=X$总是非得出 $0=0$ 不可。但在这里，$\frac{0}{0}$等于一个完全明确的特定值，$=mx^{m-1}$，并且本身就是从 $x^m$ 导出这个值的运算的符号结果；作为这样的结果，它用$\frac{dy}{dx}$来表示。所以在这里$\frac{dy}{dx}\left(=\frac{0}{0}\right)$就其来源而言表明它是已经导出的 $f'(x)$的符号值或微分表示式，而不是反过来利用符号$\frac{dy}{dx}$找到 $f'(x)$。

但是当我们一旦得到了这个结果，因而我们已经在微分演算的地盘上活动的时候，我们同时能够反过来做，例如要对

$$x^m=f(x)=y$$

进行微分的话，我们一开始就知道

$$dy=mx^{m-1}dx$$

或者

$$\frac{dy}{dx}=mx^{m-1}。$$

因此，在这里我们从符号出发；它不再作为对 $x$ 的函数求导的结果，而已作为**符号表示式**出现，这表示式指明，为要得到$\frac{dy}{dx}$的实在值即 $f'(x)$，应该对 $f(x)$ 进行哪些运算。在第一种情况中，我们得到$\frac{0}{0}$或$\frac{dy}{dx}$作为 $f'(x)$的符号等价物，而且为了揭示$\frac{dy}{dx}$的来源，首先必须这样做；在第二种情况中，我们得到 $f'(x)$ 作为符号$\frac{dy}{dx}$的实在值。但是当符号$\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$等等变为微分学的运算公式时，它们作为这样一些公式也可以出现在**等式的右边**，正象在最简单的情况 $dy=f'(x)dx$ 中早已如此。如果这种等式在它的最后形式中不象在这情况中立刻给我们$\frac{dy}{dx}=f'(x)$等等，那末正好表明它是这样的一个等式，只是以符号的形式表示当它应用到**确定的函数**上时要进行哪些运算。

$d(uz)$就是这种情况，并且是最简单的情况，其中 $u$ 和 $z$ 两者都是变量，而且两者都又是同一个第三变量例如 $x$ 的函数。

---

假定要对 $f(x)$ 或 $y=uz$ 进行微分，其中 $u$ 和 $z$ 是**两个都依赖于 $x$ 的变量**。那末

$$y_1=u_1z_1,$$

并且

$$y_1-y=u_1z_1-uz。$$

因此：

$$\frac{y_1-y}{x_1-x}=\frac{u_1z_1}{x_1-x}-\frac{uz}{x_1-x},$$

或

$$\frac{\Delta y}{\Delta x}=\frac{u_1z_1-uz}{x_1-x}。$$

然而

$$u_1z_1-uz=z_1(u_1-u)+u(z_1-z),$$

因为它等于

$$z_1u_1-z_1u+uz_1-uz=z_1u_1-uz。$$

所以：

$$\frac{u_1z_1-uz}{x_1-x}=z_1\frac{u_1-u}{x_1-x}+u\frac{z_1-z}{x_1-x}。$$

如果右边变为 $x_1-x=0$，或 $x_1=x$，那末 $u_1-u=0$，因而 $u_1=u$，以及

$z_1 - z = 0$,因而 $z_1 = z$; 所以我们得到

$$\frac{dy}{dx} = z\frac{du}{dx} + u\frac{dz}{dx},$$

从而得到

$$d(uz) \text{ 或 } dy = zdu + udz。$$

关于 $uz$ 的这种微分,现在必须指出——与我们以前**只有一个因变量**的情况不同——这里我们立刻发现在等式的两边都有微分符号,即:

**在第一步中**

$$\frac{dy}{dx} = z\frac{du}{dx} + u\frac{dz}{dx};$$

**在第二步中**

$$d(uz) \text{ 或 } dy = zdu + udz,$$

这个式子具有与一个自变量例如 $dy = f'(x)dx$ 的情况不一样的形式,因为那里用 $dx$ 去除,就会立刻给我们 $\frac{dy}{dx} = f'(x)$这个由 $x$ 的函数导出的 $f'(x)$的、摆脱了符号微系数的特定值,而在 $dy = zdu + udz$ 这里就绝不是这样。

在**只有一个因变量**的函数情况下已经指出,如何**利用实际的设置差值和随后的把它扬弃**,从 $x$ 的一个函数例如 $f(x) = x^m$ 导出 $x$ 的第二个函数 $f'(x)$,或者在当前情况下导出 $mx^{m-1}$,以及如何从这个过程中会同时对此导函数在等式左边得出符号等价物$\frac{0}{0} = \frac{dy}{dx}$。

再者,令$\frac{0}{0} = \frac{dy}{dx}$的做法,在这里非但是容许的,而且在数学上也是必要的,因为$\frac{0}{0}$按其固有的原始形式来说可以取任何一个值,这由于$\frac{0}{0} = X$总是非得出 $0 = 0$ 不可。但在这里,$\frac{0}{0}$显现为一个完全确定的实在值(例如上面的 $mx^{m-1}$)的符号等价物,而且它本身只是从 $x^m$ 导出这个值的那些运算的结果;作为这样的结果,它以$\frac{dy}{dx}$的形式固定了下来。

所以在这指明$\frac{dy}{dx}\left(=\frac{0}{0}\right)$的来源的地方,决不是借助于符号$\frac{dy}{dx}$找到$f'(x)$,恰恰相反,微分表示式$\frac{dy}{dx}$倒是已经导出的 $x$ 的函数的符号等价物。

但是一旦获得了这个结果,我们就可以反过来进行。如果要微分某个 $f(x)$,例如$x^m$,那末我们先找$dy$的值,并找到$dy = mx^{m-1}dx$,因而$\frac{dy}{dx} = mx^{m-1}$。符号表示式在这里就当作出发点了。这样,我们就已经在微分演算的地盘上

活动，也就是说，$\frac{dy}{dx}$等等已经当作这样的**公式**，指明要对 $x$ 的函数进行哪些已知的微分运算。在第一种情况中$\frac{dy}{dx}\left(=\frac{0}{0}\right)$是作为 $f'(x)$ 的符号等价物得到的，在第二种情况中 $f'(x)$倒是要找的，而且是作为符号$\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$等等的实在值得到的。

但是如果这些符号已经用作微分学的运算公式，那末作为这种公式，它们也就可以出现在等式的右边，犹如最简单的情况 $dy=f'(x)dx$ 就已是这样。如果这种等式在它的最后形式中不能象上述情况那样立刻可化为$\frac{dy}{dx}=f'(x)$等等，也就是说可化为一个实在值，那末正好表明是这样的一个等式，它只是用符号表示当**确定的函数**代替它们不确定的记号时应对之进行哪些运算。

出现这种现象的最简单情况是 $d(uz)$，其中 $u$ 和 $z$ 两者都是变量，但它们都又是同一个第三变量例如 $x$ 的函数。

如果这里我们在微分过程中立刻得到（参见笔记本 I 这方面的开始部分，它在这笔记本的第 10 页上又重复了一遍①）

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx},$$

那末不应忘记，这里 $u$ 和 $z$ 两者都是**依赖于 $x$ 的变量**，而由于 $y$ 依赖于 $z$ 和 $u$，所以 $y$ 就只依赖于 $x$。在**一个**因变量的情况下，这因变量是出现在符号化一边的，现在我们在右边有两个变量 $u$ 和 $z$，它们相对于 $y$ 是独立的，但两者都**依赖于** $x$，而作为依赖于 $x$ 的变量，它们的这一特性就体现在相应的符号微系数$\frac{du}{dx}$和$\frac{dz}{dx}$之中。如果因变量出现在右边，那末符号微系数也就必然出现在右边。

从等式

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

得出：

$$d(uz)\text{ 或 }dy=zdu+udz。$$

这等式只是指明，当 $u$ 和 $z$ 作为 $x$ 的确定函数而给定的时候所应进行的运算。

最简单的情况也许是例如

$$u=ax,\ z=bx。$$

---

① 见本文第 27 页。

这时

$$d(uz) \text{ 或 } dy = bx \cdot adx + ax \cdot bdx。$$

如果两边都用 $dx$ 去除，那末：

$$\frac{dy}{dx} = abx + bax = 2abx$$

以及

$$\frac{d^2y}{dx^2} = ab + ba = 2ab。$$

但如果我们一开始就取乘积

$$y \text{ 或 } uz = ax \cdot bx = abx^2,$$

那末

$$uz \text{ 或 } y = abx^2, \quad \frac{dy}{dx} = 2abx, \quad \frac{d^2y}{dx^2} = 2ab。$$

一旦我们得到例如 $z\frac{du}{dx}$这样的公式，那末很清楚，我们可称之为一般运算等式的那个等式，就是有待实行的微分运算的符号表示式。例如我们取表示式 $y\frac{dx}{dy}$，其中 $y$ 为纵坐标，$x$ 为横坐标，那末它就是任意一条曲线的次切线的一般符号表示式（完全象 $d(uz) = zdu + udz$ 是对两个变量的任意一个乘积的微分这样一种表示式，其中两个变量都依赖于同一个第三因变量①）。尽管我们对于 $dx$ 抱有它是横坐标的微分，对于 $dy$ 抱有它是纵坐标的微分这样感性的想象，然而只要我们放着这个表示式原封不动，那末它是不会再得出什么东西来的。

为了获得任何一个积极的结果，我们必须先取一条确定曲线的方程，它给我们 $y$ 用 $x$ 表达的确定值，因而也给我们 $dx$ 的确定值，例如取 $y^2 = ax$ 这样一个普通抛物线方程，而后通过微分得到 $2ydy = adx$；因此 $dx = \frac{2ydy}{a}$。如果把 $dx$ 的这个确定值代进次切线的一般公式 $y\frac{dx}{dy}$中去，那末我们得到

$$\frac{y\frac{2ydy}{a}}{dy} = \frac{y \cdot 2ydy}{ady} = \frac{2y^2}{a},$$

并且由于 $y^2 = ax$ 而

$$= \frac{2ax}{a} = 2x,$$

---

① 这是笔误。应是“变量”而不是因变量。

这就是普通抛物线的次切线的值；也就是说，它是 $=2\times$ 横坐标。但是如果我们把次切线叫做 $\tau$，那末一般的等式 $y\dfrac{dx}{dy}=\tau$ 只提供 $ydx=\tau dy$。所以从微分演算的观点来看，问题大多是这样提的(拉格朗日除外)：找 $\dfrac{dy}{dx}$ 的实在值。

困难似乎发生在我们对 $\dfrac{dy}{dx}$ 等等用它们的原始形式 $\dfrac{0}{0}$ 来替代的时候，这时

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

就显现为

$$\frac{0}{0}=z\cdot\frac{0}{0}+u\cdot\frac{0}{0},$$

这是一个正确的、然而得不出什么东西来的等式，尤其是因为这三个 $\dfrac{0}{0}$ 出自不同的微系数，而它们各别的推导已不再能看出；但要考虑到：

1）即使在一个自变量的最初表述中，我们先得到

$$\frac{0}{0}\text{或}\frac{dy}{dx}=f'(x)；\ \text{因而}\ dy=f'(x)dx。$$

但由于

$$\frac{dy}{dx}=\frac{0}{0},\ dy=0\ \text{以及}\ dx=0,\text{所以}\ 0=0。$$

当我们对 $\dfrac{dy}{dx}$ 再用它的不定表示式 $\dfrac{0}{0}$ 来代替的时候，我们在这里就犯了一个真正的错误，因为 $\dfrac{0}{0}$ 在这里只是作为实在值 $f'(x)$ 的符号等价物得出的，而作为这样的符号等价物它用表示式 $\dfrac{dy}{dx}$ 固定了下来，因而也用 $dy=f'(x)dx$ 固定了下来。

2）因为变量 $x_1$ 变为 $=x$，或 $x_1-x=0$，所以 $\dfrac{u_1-u}{x_1-x}$ 变为 $\dfrac{du}{dx}$ 或 $\dfrac{0}{0}$；因此我们对于 $\dfrac{u_1-u}{x_1-x}$ 立刻得到的不是 0 而是 $\dfrac{0}{0}$；然而在一般情况下我们知道，$\dfrac{0}{0}$ 可以取任意一个值，而在确定的情况下，即当 $u$ 为 $x$ 的一个确定的函数代替时，它就取一个特定的值；所以我们不但有理由把 $\dfrac{0}{0}$ 记为 $\dfrac{du}{dx}$，而且也必须这样做，因为 $\dfrac{du}{dx}$ 以及 $\dfrac{dz}{dx}$ 在这里都只起着有待实行的微分运算的符号的作用。只要我们停留在

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx},$$

也就是停留在

$$dy=zdu+udz$$

这个结果上，那末$\frac{du}{dx}$，$\frac{dz}{dx}$，$du$，$dz$ 也依然是一些不确定的值，正象能取任意一个值的$\frac{0}{0}$那样。

3）在普通代数学中，正因为$\frac{0}{0}$可以是任意一个量的符号，$\frac{0}{0}$也可以作为具有实在值的表示式的一种形式出现。例如给定$\frac{x^2-a^2}{x-a}$，若我们令 $x=a$，则 $x-a=0$，且 $x^2=a^2$，从而 $x^2-a^2=0$。于是我们得到

$$\frac{x^2-a^2}{x-a}=\frac{0}{0};$$

这结果到目前为止是正确的；但是它决不证明，由于$\frac{0}{0}$可以取任意一个值，以致$\frac{x^2-a^2}{x-a}$就没有一个实在值。

如果我们把 $x^2-a^2$ 分解成它的因子，那末它就 $=(x+a)(x-a)$；所以

$$\frac{x^2-a^2}{x-a}=(x+a)\frac{x-a}{x-a}=x+a;$$

于是当 $x-a=0$ 时，就得 $x=a$，因而 $x+a=a+a=2a$。

如果在一个普通的代数等式中我们有 $P(x-a)$这样一项，那末当 $x=a$，因而 $x-a=0$ 时，必然有 $P(x-a)=P\cdot 0=0$；在同样假定下，也就有 $P(x^2-a^2)=0$。纵使把 $x^2-a^2$ 分解成它的因子 $(x+a)(x-a)$，于此也改变不了什么，因为

$$P(x+a)(x-a)=P(x+a)\cdot 0=0。$$

但是决不能因此得出结论说，当令 $x=a$ 而演化出 $P\cdot\left(\frac{0}{0}\right)$这样一项时，它的值就必然$=0$。

由于从$\frac{0}{0}=X$ 总是得出 $0=X\cdot 0=0$，所以$\frac{0}{0}$可以取任意一个值；但是正因为$\frac{0}{0}$可以取任意一个值，它就不一定取 0 这个值，而如果我们知道它的来历，那末只要它后面隐藏着一个实在值，这个实在值也是可以找出来的。

所以例如 $P\cdot\frac{x^2-a^2}{x-a}$，当 $x=a$，$x-a=0$ 时，从而也就 $x^2=a^2$，$x^2-a^2=0$；

因此

$$P\cdot\frac{x^2-a^2}{x-a}=P\cdot\frac{0}{0}。$$

虽然我们数学上完全正确地得到了这个结果，但是如果不作进一步说明而假定 $P\cdot\frac{0}{0}=0$，那末数学上就会是全然错误的；因为这一假定包含着$\frac{0}{0}$一定不能取 0 以外的任何其他值，因而

$$P\cdot\frac{0}{0}=P\cdot 0。$$

倒不如来探讨一下，通过 $x^2-a^2$ 分解成它的因子$(x+a)(x-a)$，是否得不出别的结果；这样分解事实上把表示式变为

$$P\cdot(x+a)\frac{x-a}{x-a}=P\cdot(x+a)\cdot 1,$$

且当 $x=a$ 时，变为 $P\cdot 2a$ 或 $2Pa$。所以，在我们原先已证明了微分符号$\frac{dy}{dx}$，$\frac{dz}{dx}$等等是作为必须经过一定微分过程而得来的变量的导函数的符号等价物之后，一旦我们用变量进行计算，那末用微分符号$\frac{dy}{dx}$，$\frac{dz}{dx}$等等把$\frac{0}{0}$的来历固定下来，就更显得不仅是合理的，而且也是必须这样做的。如果它们原先是这种微分过程的结果，那末正因为如此，它们就能够**反过来**成为对变量有待实行的过程的符号，也就是成为**运算符号**，这些运算符号不再表示结果而表示出发点，这就是它们在微分演算中的主要作用。作为这种运算符号，它们本身能够变成不同变量之间等式的内容（在隐函数情况下，等式的右边一开始就是 0，而因变量和自变量连同它们的系数出现在左边）。

在我们获得的等式

$$\frac{d(uz)}{dx}\text{或}\frac{dy}{dx}=\frac{zdu}{dx}+\frac{udz}{dx}$$

中，情况就是这样。除了前面说过的以外，依赖于 $x$ 的函数 $z$ 和 $u$ 本身没有变，仍作为 $z$ 和 $u$ 出现在这里；不过它们中的每一个都配备着另一个的符号微系数作为因子。

这等式因此只具有一般等式的价值，它通过符号表明，一旦 $u$ 和 $z$ 各别地作为因变量而用两个 $x$ 的确定函数给定的时候，应该实行哪些运算。

只有当 $u$ 和 $z$ 是 $x$ 的确定函数时，$\frac{du}{dx}\left(=\frac{0}{0}\right)$和$\frac{dz}{dx}\left(=\frac{0}{0}\right)$从而$\frac{dy}{dx}\left(=\frac{0}{0}\right)$才能变为 0，因此$\frac{0}{0}=0$这个值不能预先假定，而必须从确定的函数等式本身

得出来。

例如设 $u=x^3+ax^2$,那末

$$\left(\frac{0}{0}\right)=\frac{du}{dx}=3x^2+2ax, \qquad \left(\frac{0}{0}\right)_1=\frac{d^2u}{dx^2}=6x+2a,$$

$$\left(\frac{0}{0}\right)_2=\frac{d^3u}{dx^3}=6, \qquad \left(\frac{0}{0}\right)_3=\frac{d^4u}{dx^4}=0,$$

因此$\frac{0}{0}$在这情况下$=0$。

总之,重要的在于,我们在这里利用微分本身获得了**以符号形式表示的微系数作为结果**,作为微分等式即等式

$$\frac{d(uz)}{dx}\text{或}\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

中$\left[\frac{dy}{dx}\right]$的值。但现在我们知道,$u=x$ 的确定函数,譬如 $f(x)$。因此$\frac{u_1-u}{x_1-x}$在其微分符号$\frac{du}{dx}$的情况下等于 $f'(x)$,即 $f(x)$ 的一阶导函数。与此相似,例如 $z=\varphi(x)$,因而同样得 $\frac{dz}{dx}=\varphi'(x)$,象上面一样它是 $\varphi(x)$ 的一阶导函数。但是原始等式本身现在既没有对 $u$ 也没有对 $z$ 为我们提供任何 $x$ 的确定函数,比如说

$$u=x^m,\ z=\sqrt{x}。$$

它只是为我们提供了 $u$ 和 $z$ 作为 2 个任意的 $x$ 的函数的一般表示式,而这两个函数的乘积是要加以微分的。

这等式表明,当对以 $uz$ 表示的任何两个 $x$ 的函数的乘积进行微分时,我们先给符号微系数$\frac{du}{dx}$找出相应的实在值,也就是譬如 $f(x)$的一阶导函数,并将此值乘以 $\varphi(x)=z$;然后同样找出$\frac{dz}{dx}$的实在值,并乘以 $f(x)=u$;最后把这两个如此得到的乘积相加。微分学的这些运算,这里认为是已经知道的。

因此,这等式只是所要实行的运算的一个符号指示,符号微系数$\frac{du}{dx}$,$\frac{dz}{dx}$在这里同时就变为在每一具体情况下所要实行的微分运算的符号,虽然它们原先是作为已经实行了的微分运算的符号公式而推导出来的。

一旦它们担任了这个角色,它们本身就能变成微分等式的内容,例如在**泰勒定理**

$$y_1=y+\frac{dy}{dx}h+\text{etc.}$$

中那样。但此时这些等式也不过是一般的符号运算等式而已。所以对 $uz$ 微分这种情况有趣的是，它不同于在自变量 $x$ 只有一个因变量 $y$ 的情况下演化的一种最简单情况，在这里通过应用原来的方法，微分符号也出现在等式的右边(它的展开表示式一边)，因此这些微分符号同时作为运算符号出现，并且它们自己就变成等式本身的内容。

它们指明所要实行的运算，因而充当出发点的这种作用，是它们已经在自己的地盘上活动的微分演算中特有的作用。但是可以肯定，没有一个数学家曾经注意到这种突变，作用的这种转换，更不用说觉得有必要用一个完全初等的微分等式去证明它了。作为一个事实，只是提一下：当微分演算的发明者及其大部分继承者把微分符号当作演算的出发点时，拉格朗日却反过来把对自变量的实际函数的代数推导作为出发点，而把微分符号作为已经导得的函数的纯粹符号表示式。

如果我们再一次回到 $d(uz)$，那末作为令 $x_1-x=0$ 的结果，也就是作为微分运算本身的结果，我们首先得到：

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}。$$

由于在这里分母相同，所以作为化简了的表示式我们得到

$$dy=zdu+udz。$$

这相当于在只有一个因变量的情况下，作为 $x$ 的导函数即 $f'(x)$（例如当 $ax^m=f(x)$ 时，$max^{m-1}$ 就是 $f'(x)$）的符号表示式，作为其符号表示式我们在左边得到了 $\frac{dy}{dx}$：

$$\frac{dy}{dx}=f'(x),$$

并且仅仅作为结果才由此得到

$$dy=f'(x)dx$$

（例如 $\frac{dy}{dx}=max^{m-1}$；$dy=max^{m-1}dx$，这是函数 $y$ 的微分）（我们立刻可以把这后一个式子再变成 $\frac{dy}{dx}=max^{m-1}$）。但是

$$dy=zdu+udz$$

的情况不同之处还在于，$du$，$dz$ 这些微分在这里作为运算符号出现在等式的右边，而且只有在完成了它们所指定的运算之后 $dy$ 才确定下来。如果

$$u=f(x)\ 和\ z=\varphi(x),$$

那末我们知道，对于 $du$ 我们得到

$$du=f'(x)dx,$$

而对于 $dz$

$$dz=\varphi'(x)dx。$$

因此：

$$dy=\varphi(x)f'(x)dx+f(x)\varphi'(x)dx,$$

而

$$\frac{dy}{dx}=\varphi(x)f'(x)+f(x)\varphi'(x)。$$

所以在第一种情况下，首先找到微系数

$$\frac{dy}{dx}=f'(x),$$

而后得到微分

$$dy=f'(x)dx。$$

在第二种情况下，首先找到微分 $dy$，而后得到微系数$\frac{dy}{dx}$。在第一种情况下，即在微分符号本身只是从对 $f(x)$ 实行的运算中得出的那种情况下，必须首先找出导函数这个实际的微系数，以便$\frac{dy}{dx}$作为其符号表示式出现在它的对面。只有在找到了实际的微系数之后，才能由此导出微分 $dy=f'(x)dx$。

在 $dy=zdu+udz$ 的情况下则相反。

由于 $du$，$dz$ 在这里起着运算符号的作用，而且它们都表示我们由微分演算已会实施的那些运算，所以为了找出$\frac{dy}{dx}$的实在值，我们就必须在每一具体情况下先对 $u$ 和 $z$ 各代之以它们用 $x$ 表达的值，以便找出

$$dy=\varphi(x)f'(x)dx+f(x)\varphi'(x)dx;$$

然后用 $dx$ 去除才得到

$$\frac{dy}{dx}=\varphi(x)f'(x)+f(x)\varphi'(x)$$

的实在值。

凡是对于$\frac{du}{dx}$，$\frac{dz}{dx}$，$\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$等等成立的，则对于**微分符号**本身作为一般符号运算等式的内容而出现的所有更复杂的公式也全都成立。

---

## [二　稿]

我们曾经从 $f'(x)$ 的代数推导出发，为的是用以同时指明它的符号微分表示式$\frac{0}{0}$或$\frac{dy}{dx}$的来源，从而也揭示它的意义。现在我们必须反过来，把符号微系数$\frac{du}{dx}$，$\frac{dz}{dx}$当作已经给定的公式，由此出发以找出分别与它们相应的实在等价物 $f'(x)$，$\varphi'(x)$。从相反两极出发的——并表征历史上两个不同学派的——微分演算的这些不同处理方式，其实在这里并不是由于我们主观方法的改变，却是由于所要处理的函数 $uz$ 的本性而产生的。我们以前处理它时，就象处理只有一个因变量的 $x$ 的函数那样，从右边的一端出发，并对这一边用代数方法进行了运算。我不相信有任何一个数学家，无论就 $uz$ 这样一个初等的函数，或者就任何一个别的函数，曾经证实过甚或注意过来自第一个代数推导方法(历史上是第二个)的这个必然的转换。他们过份地钻到微分演算的材料堆里去了。

事实上我们看到，在等式

$$\frac{0}{0}\text{或}\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

中，$\frac{dy}{dx}$完全象以前在一个因变量的 $x$ 的函数那里一样，也是从右边对 $uz$ 进行的推导得来的；但另一方面，微分符号$\frac{du}{dx}$，$\frac{dz}{dx}$又包含在 $f'(x)$或者 $uz$ 的一阶导函数中，因而构成了$\frac{dy}{dx}$的等价物的元素。

这样，符号微系数本身就已经变成微分运算的**对象**或**内容**，而不象以前那样只是呈现为微分运算的符号结果。

随着这两点，**第一**：符号微系数同变量一样，自身又变成推导的有内容的元素，成为微分运算的**对象**，**第二**：问题的提法颠倒了过来，这时不是为实在微系数 $(f'(x))$ 找符号表示式，而是为符号表示式求它的实在微系数，——随着这两点就得到了第三点，即符号微分表示式不是显现为对 $x$ 的实际函数进行的微分运算的符号结果，而现在反过来起着符号的作用，它指明对 $x$ 的实际函数应当完成的微分运算；也就是说它们变成了**运算符号**。

在我们的情况

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

下，只有在我们不但知道 $z$ 和 $u$ 是 $x$ 的两个函数，而且要象

$$y=x^m$$

那样，对于 $u$ 和 $z$ 给出了用 $x$ 表达的实际值例如

$$u=\sqrt{x}\text{，}z=x^3+2ax^2$$

的时候，我们才能继续运算下去。这样，$\frac{du}{dx}$，$\frac{dz}{dx}$ 事实上就成为运算的标志，而对于代入 $u$ 和 $z$ 的 $x$ 的每一个任意函数，运算的实施方法假定已经知道。

c）所找到的等式不但是符号运算等式，而且是纯粹预备性的符号运算等式。

由于在

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

的两边所有各项中都出现分母 $dx$，所以它的化简了的表示式为：

II) $$dy\text{ 或 }d(uz)=zdu+udz\text{。}$$

这等式直接说明，如果要对两个任意变量的乘积（在以后的应用中，这可推广到对任意多个变量的乘积）进行微分，那末两个因子的每一个都要与另一个因子的微分相乘，并把所得到的两个乘积相加。

因此，如果对两个任意变量的乘积进行微分，那末这第一个运算等式

$$\frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

在完成了它的使命，即提供了一个能直接达到目的的一般符号运算公式之后，它作为预备等式就成为多余的了。

但这里必须指出，原先代数的推导方法又转换成自己的对立面。在那里我们首先得到

$$\Delta y=y_1-y$$

作为 $f(x_1)-f(x)$ 的相应符号，$f(x_1)$ 和 $f(x)$ 两者都是普通的代数表示式（因为 $f(x)$ 和 $f(x_1)$ 都已给定为 $x$ 的确定的代数函数）。然后把 $\frac{f(x_1)-f(x)}{x_1-x}$ 表示成 $\frac{\Delta y}{\Delta x}$，接着把 $f'(x)$（$f(x)$ 的一阶导函数）表示成 $\frac{dy}{dx}$，并且只是从微系数的最后等式

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

我们才得到微分

$$dy=f'(x)dx。$$

与此相反，上面的等式为我们提供了微分 $dy$，$du$，$dz$ 作为出发点。所以如果 $u$ 和 $z$ 代之以 $x$ 的任意的、确定的代数函数，而这些函数我们只用

$$u=f(x) \text{ 和 } z=\varphi(x)$$

来表示，那末

$$dy=\varphi(x)df(x)+f(x)d\varphi(x),$$

而这些 $d$ 符号只是表明所要实行的微分运算。

这个微分的结果具有一般形式：

$$df(x)=f'(x)dx$$

和

$$d\varphi(x)=\varphi'(x)dx。$$

因此

$$dy=\varphi(x)f'(x)dx+f(x)\varphi'(x)dx。$$

最后

$$\frac{dy}{dx}=\varphi(x)f'(x)+f(x)\varphi'(x)。$$

这里，在微分已经起着现成的运算符号作用的地方，我们从微分导出微系数，而在原先的代数演化中则相反，微分是从微系数等式中推导出来的。

如果我们看**微分**本身，例如我们在它最简单的形式下，也就是从一次函数：

$$y=ax,\ \frac{dy}{dx}=a$$

演化它那样；那末这个微分

$$dy=adx。$$

这些微分的等式，比起它由之导出的微系数等式

$$\frac{0}{0}\text{或}\frac{dy}{dx}=a$$

看来可疑得多。

由于 $dy=0$ 和 $dx=0$，所以 $dy=adx$ 等同于 $0=0$。虽然如此，我们还是完全有理由使用 $dy$ 和 $dx$ 来代替消失了的、但在消失中用这些符号固定下来的差值 $y_1-y$ 和 $x_1-x$。

只要我们停留在表示式

$$dy=adx$$

或者一般地

$$dy=f'(x)dx,$$

那末它完全不外是

$$\frac{dy}{dx}=f'(x),$$

在上述情况下就是 $=a$ 这一事实的另一种表示，因此我们总是又能把它变为后者。但是这种可转变性已经使它成为运算符号。我们立刻看到，如果我们作为微分过程的结果已经得到 $dy=f'(x)dx$，那末为了找 $\frac{dy}{dx}=f'(x)$，也就是为了找微系数，我们只要用 $dx$ 去除两边就行。

这样，例如在 $y^2=ax$ 中，

$$d(y^2)=d(ax),\ 2ydy=adx。$$

后一微分等式给我们提供了两个微系数等式，即：

$$\frac{dy}{dx}=\frac{a}{2y}和\frac{dx}{dy}=\frac{2y}{a}。$$

但是 $2ydy=adx$ 也直接给我们提供了 $dx$ 的值 $\frac{2ydy}{a}$，例如当把这个值代入次切线的一般公式 $y\frac{dx}{dy}$ 中时，最后就会使我们得到 $2x$，即横坐标的两倍作为普通抛物线的次切线的值。

## II

我们现在要举一个例子。在这例子中符号表示式先是用作演算的现成运算公式，因而也为符号微系数求出它的实在值，而后再给出相反的初等的代数表述。

1）假定因变函数 $y$ 和自变量 $x$ 不是在单独一个等式中联系起来的，而是这样：$y$ 在第一个等式中直接表示为变量 $u$ 的函数，而 $u$ 在第二个等式中直接表示为变量 $x$ 的函数。课题：**求符号微系数 $\frac{dy}{dx}$ 的实在值。**

设

a）$y=f(u)$，b）$u=\varphi(x)$。

首先 1）$y=f(u)$ 给出：

$$\frac{dy}{du}=\frac{df(u)}{du}=\frac{f'(u)du}{du}=f'(u)。$$

2) $\frac{du}{dx}=\frac{d\varphi(x)}{dx}=\frac{\varphi'(x)dx}{dx}=\varphi'(x)$。

因此

$$\frac{dy}{du}\cdot\frac{du}{dx}=f'(u)\cdot\varphi'(x)。$$

然而

$$\frac{dy}{du}\cdot\frac{du}{dx}=\frac{dy}{dx}。$$

所以

$$\frac{dy}{dx}=f'(u)\cdot\varphi'(x)。$$

**例**。如果 a) $y=3u^2$, b) $u=x^3+ax^2$,那末根据公式

$$\frac{dy}{du}=\frac{d(3u^2)}{du}=6u(=f'(u));$$

而等式 b) 给出 $u=x^3+ax^2$。我们若把 $u$ 的这个值代入 $6u$,则

$$\frac{dy}{du}=6(x^3+ax^2)(=f'(u))。$$

再者:

$$\frac{du}{dx}=3x^2+2ax(=\varphi'(x))。$$

因此

$$\frac{dy}{du}\cdot\frac{du}{dx}\text{或}\frac{dy}{dx}=6(x^3+ax^2)(3x^2+2ax)(=f'(u)\cdot\varphi'(x))。$$

2) 我们现在把上面例子中的等式作为起始等式,以便现在用原先的代数方法对它们进行演化。

a) $y=3u^2$, b) $u=x^3+ax^2$。

由于 $y=3u^2$, $y_1=3u_1^2$, 从而

$$y_1-y=3(u_1^2-u^2)=3(u_1-u)(u_1+u)。$$

因此

$$\frac{y_1-y}{u_1-u}=3(u_1+u)。$$

如果现在 $u_1 - u$ 变为 $=0$，因而 $u_1 = u$，那末 $3(u_1+u)$ 就变为 $3(u+u)=6u$。

我们若用等式 b)给出的值代入 $u$,则

$$\frac{dy}{du}=6(x^3+ax^2)。$$

再者：由于

$$u=x^3+ax^2,\ u_1=x_1^3+ax_1^2;$$

所以

$$u_1-u=(x_1^3+ax_1^2)-(x^3+ax^2)=(x_1^3-x^3)+a(x_1^2-x^2),$$
$$u_1-u=(x_1-x)(x_1^2+x_1x+x^2)+a(x_1-x)(x_1+x);$$

因此

$$\frac{u_1-u}{x_1-x}=(x_1^2+x_1x+x^2)+a(x_1+x)。$$

如果现在 $x_1-x$ 变为 $=0$,因而 $x_1=x$,那末

$$x_1^2+x_1x+x^2=3x^2$$

以及

$$a(x_1+x)=2ax。$$

因此：

$$\frac{du}{dx}=3x^2+2ax。$$

现在我们若把右边的两个函数相乘,则得

$$6(x^3+ax^2)(3x^2+2ax),$$

并与此相应,左边得

$$\frac{dy}{du}\cdot\frac{du}{dx}=\frac{dy}{dx},$$

这样就同前面一样。

为了让演化的差别更加清楚地显示出来，我们把变量的确定函数放在左边，而把依赖于变量的函数放在右边，因为从一般等式右边只出现 0 这一点，人们已经习惯于将主动性设想在左边。因此：

a) $3u^2=y$; b) $x^3+ax^2=u$。

由于

$$3u^2=y,\quad 3u_1^2=y_1,$$

所以

$$3(u_1^2-u^2)=y_1-y$$

或者

$$3(u_1-u)(u_1+u)=y_1-y,$$

因而

$$3(u_1+u)=\frac{y_1-y}{u_1-u}。$$

如果现在 $u_1$ 变为 $=u$，因而 $u_1-u=0$，那末得到

$$3(u+u) \text{ 或 } 6u=\frac{dy}{du}。$$

我们若把由等式 b)来的值代入 $6u$,则

$$6(x^3+ax^2)=\frac{dy}{du}。$$

再者：如果

$$x^3+ax^2=u,$$

那末

$$x_1^3+ax_1^2=u_1$$

从而

$$x_1^3+ax_1^2-x^3-ax^2=u_1-u;$$

因此

$$(x_1^3-x^3)+a(x_1^2-x^2)=u_1-u。$$

我们把它进而分解成因子：

$$(x_1-x)(x_1^2+x_1x+x^2)+a(x_1-x)(x_1+x)=u_1-u。$$

于是

$$(x_1^2+x_1x+x^2)+a(x_1+x)=\frac{u_1-u}{x_1-x};$$

现在如果 $x_1$ 变为 $=x$，因而 $x_1-x=0$，那末

$$3x^2+2ax=\frac{du}{dx}。$$

若把这 2 个导函数相乘,则

$$6(x^3+ax^2)(3x^2+2ax)=\frac{dy}{dx},$$

而当我们把它换成惯常的次序时：

$$\frac{dy}{du}\cdot\frac{du}{dx}=\frac{dy}{dx}=6(x^3+ax^2)(3x^2+2ax)。$$

不言而喻,由于这个方法很烦琐,而且往往难于把最初的差值 $f(x_1)-f(x)$

分解成都含因子 $(x_1-x)$ 的各项，所以作为计算工具，它不能与历史上最早传下来的方法相比。

但另一方面，在后一方法中人们从 $dy$，$dx$，$\frac{dy}{dx}$ 作为已给定的运算公式出发，而在前一方法中则看到它们是用纯粹代数的方法产生出来的。此外我就不再说什么了。然而在第一个方法那里，微分符号作为运算公式的这个出发点是怎样获得的呢？是通过隐蔽的或者显露的形而上学的假设，而这些假设本身又导致形而上学的非数学的后果，那就是用暴力镇压掉某些对于推导碍手碍脚的，但却由推导本身产生出来的量。

现在为了举一个从相反两极出发的历史上的例子，我就上面演化过的情况 $d(uz)$ 的解法，以牛顿和莱布尼茨为一方、以拉格朗日为另一方作一比较。

1）**牛顿**。

首先告诉我们，当变量增长时，$\dot{x}$，$\dot{y}$ 等等表示它们流动的速度，或者说 $x$，$y$ 等等相应增长的速度。此外，由于一切可能的量在数值上的大小都可用直线表示，因而所产生的**瞬**或**无限小量**等于速度 $\dot{x}$，$\dot{y}$ 等等和它们所经历的无限小时间间隔 $\tau$ 的**乘积**，也就是 $=\dot{u}\tau$，$\dot{x}\tau$，$\dot{y}\tau$。

## [三　　稿]

如果我们现在考察一般形式下的 $y$ 的微分 $dy=f'(x)\,dx$，那末这里就有一个纯粹的符号运算等式出现在我们面前，即使在 $f'(x)$ 一开始就等于一个常量的情况下，如在 $dy=d(ax)=adx$ 中那样，也是如此。由$\frac{0}{0}$或$\frac{dy}{dx}=f'(x)$ 生出来的这个孩子，看上去比他的母亲更加可疑。因为在$\frac{dy}{dx}=\frac{0}{0}$中分母和分子**是不可分开地联系着的**；而在 $dy=f'(x)dx$ 中它们显然是分开的，所以势必得出结论：$dy=f'(x)dx$ 只是一个把 $0=f'(x)\cdot 0$，也就是把对之"毫无办法"的 $0=0$ 化了装的表示式。我们这世纪比较细心的分析学家们，如法国人布夏拉，也已经感觉到这里有些不妥。他说：

在"譬如$\frac{dy}{dx}=3x^2$ 中，$\frac{0}{0}$或者说$\frac{dy}{dx}$，或者不如说它的值 $3x^2$，是函数 $y$ 的微系数。由于$\frac{dy}{dx}$是代表极限 $3x^2$ 的符号，所以 $dx$ 必须总是在 $dy$ 的下面。但是为了**便于代数运算**，我们把$\frac{dy}{dx}$当作普通分式而把$\frac{dy}{dx}=3x^2$ 当作普通等式来处理，通过摆脱这等式的分母 $dx$ 就得到结果 $dy=3x^2dx$。这个表示式称为 **$y$ 的微分**"。

为了"便于代数运算"，我们就这样引进了一个错误的公式。

事实上情况并非如此。在$\frac{0}{0}$$\left(\text{本来应写成}\left(\frac{0}{0}\right)\right)$中，$y_1-y$ 或 $f(x_1)-f(x)$ 的最小表示式或者 $f(x)$ 的增量，与 $x_1-x$ 的最小表示式或与自变量 $x$ 的增量之比，具有分子与分母不可分开的形式。但这是为什么呢？是为了保持$\frac{0}{0}$作为消失了的差值之比。但是一旦 $x_1-x=0$ 借助于 $dx$ 获得了表明它是 $x$ 的消失了的差值这个形式，因而 $y_1-y=0$ 也就显现为 $dy$ 之后，把分子和分母分开将是一种完全许可的运算。不管现在 $dx$ 在什么地方，它和 $dy$ 的关系不会因这种位置的调换而受到影响。$dy=df(x)$，因而 $=f'(x)dx$，只是$\frac{dy}{dx}$的另一种表示式，而为了得到独立的 $f'(x)$，最后必然要出现这个$\frac{dy}{dx}$。但是这个公式 $dy=df(x)$ 立即作为**运算公式**将多么有用，下面的例子就可说明：

$$y^2=ax,$$

$$d(y^2)=d(ax),\ 2ydy=adx;$$

因此

$$dx=\frac{2ydy}{a}。$$

把 $dx$ 的这个值代入次切线的一般公式 $y\frac{dx}{dy}$,就得到

$$\frac{y\frac{2ydy}{a}}{dy}=\frac{2y^2dy}{ady}=\frac{2y^2}{a},$$

并且由于 $y^2=ax$ 而

$$=\frac{2ax}{a}=2x;$$

所以 $2x$,即普通抛物线的横坐标的两倍,是它的次切线的值。

但是如果把 $dy=df(x)$ 看做最初的出发点,甚至 $\frac{dy}{dx}$ 也是尔后才由此导出的,那末为了使 $y$ 的这个微分具有任何意义,微分元 $dy,dx$ 就必须**假定**为具有确定意义的符号。这种假定如果不是来自数学的形而上学,而是也许直接从一次函数如 $y=ax$ 中导出的话,那末正象我们前面已经看到的,它就给出 $\frac{y_1-y}{x_1-x}=a$,而这变为 $\frac{dy}{dx}=a$。然而从这里先验地也抓不到什么确定的东西。这是因为:由于 $\frac{\Delta y}{\Delta x}$ 和 $\frac{dy}{dx}=a$ 一样,也是 $=a$,并且由于 $\Delta x$,$\Delta y$ 虽是有限差值或增量,但却是能无限减小的有限差值或增量,所以可把 $dx,dy$ 同样地想象为无限小的,能任意接近 0 的量,正如它们从实际上令 $x_1-x=0$,因而 $y_1-y=0$ 中得出的那样。右边的结果两次都是同一的,在这一边根本无需令 $x_1=x$,因而也无所谓 $x_1-x=0$。所以在另一边令它 $=0$,就象假定 $dx,dy$ 是无限小量一样,似乎是一个任意的假说。我将在 IV)中以 $d(uz)$ 为例简短地说明这个历史过程,但在此以前还要在 III)中举一个例子,对它第一步在符号演算的地盘上用一个现成的运算公式进行处理,第二步用代数方法来叙述。象 II)所已经说明了的那样,这后一方法本身通过把它应用到象两个变量的乘积这样初等的函数,借助于它自己的结果,必然会给从相反一极出发进行运算的方法提供出发点。

对 IV 的补充。

最后(**根据拉格朗日**)还必须指出,在牛顿那里有时已经相当于微系数见到的,并且他还是从纯粹几何想象中推导出来的这个**极限**或**极限值**,即使在今天仍然起着特别重要的作用,不管符号表示式作为 $f'(x)$ 的极限,或者反过来

$f'(x)$作为符号的极限出现，或者两者都作为极限出现。这个尤其为拉克鲁瓦在分析上长篇论述的范畴，作为“最小表示式”这范畴的替代物，无论作为与“预先导函数”相对立的导函数的替代物，或者作为比值$\frac{y_1-y}{x_1-x}$的替代物，只有当它涉及到把演算应用于曲线的时候，才是重要的。它在几何上**比较容易想象**，所以它**在古老的几何学家那里**已经可以看到。但是在有些近代几何学家那里，它还一直隐藏在微分元和微系数仅仅表示近似值这种说法的后面。

# [补　　遗]

**A）关于 $uz$ 微分的补遗。**

1）在上面手稿中演化 $d(uz)$ 的时候，关于等式

$$\text{A）}\quad \frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}$$

对我来说重要的是，证明了这里所用的代数方法自行转变成了微分方法。因为它在导函数内部，也就是在其右边演化出了没有实在系数这个相应等价物的**符号微系数**，而这样一来，这些符号本身就变成了**独立的出发点**和现成提供的**运算公式**。

等式 A）这个形式对此目的而言显得更为适当，因为它使在导函数 $f'(x)$ 内部产生的$\frac{du}{dx}$，$\frac{dz}{dx}$和左边与之对立的$\frac{dy}{dx}$之间能作一比较，而这$\frac{dy}{dx}$是 $f'(x)$ 的符号微系数，所以构成了它的符号等价物。

关于作为运算公式的$\frac{du}{dx}$，$\frac{dz}{dx}$的性质，我只限于指出，当我们给 $u$ 代入任何一个 $f(x)$，例如 $3x^2$，给 $z$ 代入任何一个 $\varphi(x)$，例如 $x^3+ax^2$时，对于这些符号微系数就可以找到任意的“导函数”作为它们的实在值。

我当然也可以指出这些运算公式在几何上的可应用性，因为例如**曲线的次切线的一般公式** $=y\frac{dx}{dy}$，这公式和 $z\frac{du}{dx}$，$u\frac{dz}{dx}$具有完全相同的形式，它们都是一个变量和一个符号微系数的乘积。

最后当然还可以指出，$y=uz$ 是我们的论题可在其上发展的**最简单的初等函数**（$y$ 在这里 $=y^1$，而 $uz$ 是二次幂的最简单形式）。

**A）$\frac{u}{z}$的微分。**

1）假定在 $\frac{u}{z}$中 $z$ 是自变量，$u$是因变量。

为多样化起见，我们这一次来处理——象经常会碰到的那样——用代数表示式（其中包括用正弦、余弦等等，用对数，用象 $a^x$ 那样的指数表示式所给出的这种表示式）给出的、不具有任何等式形式的函数①。

① 在另一稿中写着：“为多样化起见，我们这一次把$\frac{u}{z}$作为 $u$ 和 $z$ 的函数来处理，而不把它与第三个依赖于$\frac{u}{z}$的变量置于等式形式之中。”

a) $\frac{u}{z}$；若 $z$ 增长到 $z_1$，则①：

b) $\frac{u_1}{z_1}$；从 b)中减去 a)：

c) $\frac{u_1}{z_1}-\frac{u}{z}$；通分后：$\frac{zu_1-uz_1}{z_1z}$；将分子分解：

d) $\frac{z(u_1-u)-u(z_1-z)}{z_1z}$。如果 $z_1$ 变为 $=z$，因而 $z_1-z=0$，那末②：

e) $\frac{zdu-udz}{z^2}$③。

这个表示式显得有点奇怪，事实上它是借助于**方法的完全改变**而得来的。因为——见 d)——$z_1-z$ 不是出现在分母中而是出现在分子中，并且 d)之变为它的微分表示式 e)，只是由于我们把处于**分子**中的 $z_1-z$ 减少到了 $z_1-z=0$ 才得以实现的④。

此外，虽然假定了$\frac{u}{z}$中的 $u$ 是因变量，$z$ 是自变量，但是如果我们在d)的分子 $z(u_1-u)-u(z_1-z)$中（在这里 $z_1-z$ 绝非处于与 $u_1-u$ 不同的地位）令 $u_1=u$，$u_1-u=0$，并假定 $z$ 依赖于 $u$，那末我们还会达到同样的结果。

实际上这个过程我们已经能够这样来解释。在 d)中分母 $z_1z$ 变为 $zz$，即令 $z_1=z$，也就是 $z_1-z=0$，因此一方面在分子中 $z_1-z$ 变为 $dz$，另一方面 $u_1-u$ 变为 $du$（这是因为 $u$ 依赖于 $z$，所以一旦分母中的 $z_1$ 变为 $=z$，那末 $u_1=u$，因而 $u_1-u=u-u=du$）。

就分子来说这个方法也许就已得救，但只是为了将它完全牺牲掉。因为它的一般结果曾经是，一个变量对另一个变量的依赖关系必须表示为$\frac{y_1-y}{x_1-x}$，如果假定 $y$ 是因变量，$x$ 是自变量的话⑤。

---

① 在另一稿中："若 $z$ 增长到 $z_1$，则 $u$ 增长到 $u_1$，因此："。

② 在另一稿中为："如果 $z_1$ 变为 $=z$，那末 $z_1z$ 变为 $zz$ 或 $z^2$，$z_1-z=dz$，$u_1-u=du$，因此："。

③ 在另一稿中，这后面还加上："因此 $d\frac{u}{z}=\frac{zdu-udz}{z^2}$"。

④ 在另一稿中这一段写成："使你惊奇的是这个结果的出现。我假定这一点，因为否则你就不会想到，$\frac{u}{z}$的微分提出了一种特殊情况，在这特殊情况的演化中方法经历了某些变更"。

⑤ 在另一稿中上面两段是这样写的："事实上，$dz$(e)（它的有限形式为 $z_1-z$(d)）这个自变量 $z$ 的微分元作为 $u$ 的乘数处于分子中，而 $z$ 本身则以肯定的形式 $z^2$（其有限形式为 $z_1z$(d)）处于分母中。所以这似乎是，由于我们在 d)中的分子 $z(u_1-u)-u(z_1-z)$ 那里令 $z_1=z$ 即令$z_1-z=0$，我们就从 d)的有限比值到达了它的微分表示式 e)。这好象是方法的完全改变，尤其因为这样一来**分母**不是变为扬弃了的差值 $z_1-z=dz$，而是从 $z_1z$ 变为 $z^2$"。

2）我们再来应用等式形式①。

a）$y=\frac{u}{z}$；

b）$y_1=\frac{u_1}{z_1}$；$y_1-y=\frac{u_1}{z_1}-\frac{u}{z}$；

c）$$\frac{y_1-y}{z_1-z}=\frac{\frac{u_1}{z_1}-\frac{u}{z}}{z_1-z}=\frac{\frac{zu_1-uz_1}{z_1z}}{z_1-z}=\frac{(zu_1-uz_1)\cdot\frac{1}{z_1z}}{z_1-z};$$

d）$$\frac{y_1-y}{z_1-z}=\frac{(z(u_1-u)-u(z_1-z))\cdot\frac{1}{z_1z}}{z_1-z}。$$

现在如果在右边令 $z_1=z$，因此 $z_1-z=0$ 等等，等等，那末：

e）$$\frac{dy}{dz}=\frac{(zdu-udz)\cdot\frac{1}{z^2}}{dz}$$

从而

f）$dy$ 或 $d\frac{u}{z}=(zdu-udz)\cdot\frac{1}{z^2}$；

由此：

$$dy=\frac{zdu-udz}{z^2}。$$

所以困难只是由于**微分**被误解为**微系数**而发生的②。

如果我们做一个比较（见前稿），那末在对 $uz$ 微分的情况下找到：

$$\text{A)}\ \frac{dy}{dx}=z\frac{du}{dx}+u\frac{dz}{dx}\ \text{和 B)}\ dy=zdu+udz。$$

存在于 a）$d(uz)=zdu+udz$ 与 b）$d\frac{u}{z}=(zdu-udz)\cdot\frac{1}{z^2}$之间的差别仅仅来源于被微分的函数的差别本身。

---

3）由于 $d\frac{u}{z}$是 $d(uz)$的相逆情况，后者是乘，前者是除，所以很容易想到

---

① 在另一稿中："一旦我们再来应用原先的等式形式，神秘就此解开"。

② 在另一稿中这一段写为："整个的谜已经解开。**微系数**由 e）来表示，而在 1）e）以及这里的 f）下所得到的结果是**微分**"。

直接利用由**代数方法**求得的运算公式

$$d(uz)=zdu+udz$$

去找 $d\frac{u}{z}$。我现在就来做这件事，以此来清楚地突出推导方法和以前已经求得而现在却用作运算公式的微分结果的单纯应用之间的区别。

a) $y=\frac{u}{z}$；

b) $u=yz$。

由于 $y=\frac{u}{z}$，所以

$$yz=\frac{u}{z}\cdot z=u。$$

这样，我们只在形式上把 $u$ 化装成 2 个因子的乘积。虽然如此，但事实上课题就此已经解决，因为问题已从分式的微分变成了乘积的微分，而对此我们口袋里有着一个魔术公式。根据这公式：

c) $du=zdy+ydz$。

对于右边的第一项即 $zdy$，我们立刻看到，它必须**静止不动地**在其位置上留到最后一刹那，因为课题恰恰在于要找 $y\left(=\frac{u}{z}\right)$ 的微分，也就是要找它的用 $u$ 和 $z$ 的微分元表达的表示式。另一方面，由于这个原因，就要把 $ydz$ **移到**左边。因此：

d) $du-ydz=zdy$。

如果我们现在把 $y$ 的值即 $\frac{u}{z}$ 代入 $ydz$，那末

$$du-\frac{u}{z}dz=zdy;$$

因此

$$\frac{zdu-udz}{z}=zdy。$$

现在已到了把 $dy$ 从它的不起作用的伙伴 $z$ 那里解放出来的时刻，于是我们得到

$$\frac{zdu-udz}{z^2}=dy=d\frac{u}{z}。$$

# [微分演算的历史发展过程]

## 历史发展过程

1）**神秘的微分演算**。$x_1=x+\Delta x$，一开始就变为 $x_1=x+dx$ 或 $x+\dot{x}$，其中 $dx$ 是通过形而上学的**解释**来假定的。先是它存在，然后对它进行解释。

于是也就有 $y_1=y+dy$ 或 $y_1=y+\dot{y}$。从这个任意的假定，就得出这样的结论：为了得到正确的结果，我们在二项式 $x+\Delta x$ 或 $x+\dot{x}$ 的展开中，必须把例如与一阶导函数一起获得的、用 $x$ 和 $\Delta x$ 表达的那些项**魔术般地丢掉**，等等，等等。由于在实际建立微分演算的时候，是从上述结果出发的，也就是从那些预料到的、不是推导得来而是用解释来**假定的微分元**出发的，所以符号微系数 $\frac{dy}{dx}$ 或 $\frac{\dot{y}}{\dot{x}}$ 也是为这种解释所**预料到**的。

如果 $x$ 的增量 $=\Delta x$，而依赖于它的变量的增量 $=\Delta y$，那末不言而喻，$\frac{\Delta y}{\Delta x}$ 表示 $x$ 和 $y$ 的增量之比。至于 $\Delta x$ 出现在分母中，即自变量的增量出现在分母中而不是反过来出现在分子中，这是因为微分形式演化的最后结果本身，即**微分**，也是一开始就由那些假定的微分元所给定的缘故。

如果我取因变量 $y$ 和自变量 $x$ 的最简单的关系，即 $y=x$，那末我知道 $dy=dx$ 或 $\dot{y}=\dot{x}$。但是由于我要找自变量 $x$ 的导函数，在这里它 $=\dot{x}$，所以我必须用 $\dot{x}$ 或 $dx$ 去除两边，因而

$$\frac{dy}{dx} \text{ 或 } \frac{\dot{y}}{\dot{x}}=1。$$

所以我一下子就全都知道，在符号微系数中自变量的增量必须出现在分母中而不是出现在分子中。

但从 $x$ 的二次幂函数开始，用二项式定理就立刻可以找到**导函数**，它完全现成地以第二项出现在其中，伴随着 $dx$ 或 $\dot{x}$，即一次幂的增量，再加上要魔术般地丢掉的各项。这种**魔术般地丢掉**不知不觉地在数学上倒是正确的，因为它只是丢掉了由最初的魔术一开始就产生的那个计算误差。

把 $x_1=x+\Delta x$ 变为

$$x_1=x+dx \text{ 或 } x+\dot{x},$$

于是就可对这微分二项式象对普通的二项式一样进行处理，从技术观点来看，这是很有成效的。

唯一还可能提出的问题是：为什么要把那些碍手碍脚的项用暴力镇压掉？这就假定了大家都已知道它们是碍手碍脚的，并且实际上是不属于导函数的。

**回答**很简单：这纯粹是从经验得来的。不但对许多更加发展了的 $x$ 的函数，以及作为曲线方程的它们的解析形式等等，人们早已知道了实际的导函数，而且就在最先可能的决定性试验中，即在对最简单的二次代数函数的处理中发现了这一点，例如：

$$y = x^2,$$

$$y + dy = (x + dx)^2 = x^2 + 2xdx + dx^2,$$

$$y + \dot{y} = (x + \dot{x})^2 = x^2 + 2x\dot{x} + \dot{x}^2。$$

如果两边都减去原来的函数 $x^2(y = x^2)$，那末

$$dy = 2xdx + dx^2,$$

$$\dot{y} = 2x\dot{x} + \dot{x}\dot{x};$$

如果我从两个式子的右边镇压掉最后一项，那末

$$dy = 2xdx, \qquad \dot{y} = 2x\dot{x},$$

进而得到

$$\frac{dy}{dx} = 2x,$$

或

$$\frac{\dot{y}}{\dot{x}} = 2x。$$

但是从 $(x + a)^2$ 知道，$x^2$ 是第一项；第二项是 $2xa$；如果我用 $a$ 去除这表示式，犹如用 $dx$ 去除上面的 $2xdx$，或者用 $\dot{x}$ 去除 $2x\dot{x}$，那末得到 $2x$ 作为 $x^2$ 的一阶导函数，作为二项式给 $x^2$ 添加的用 $x$ 表达的增长。因此，为了找出导函数，必须把 $dx^2$ 或 $\dot{x}\dot{x}$镇压掉；而根本不管对 $dx^2$ 或 $\dot{x}\dot{x}$ 原来是无可奈何的。

所以，人们通过试验的方法——就在第二步中——必然会认识到：不但为了得到一个正确的结果，甚至为了得到任何一个结果，都必须把 $dx^2$ 或 $\dot{x}\dot{x}$ 魔术般地丢掉。

其次，人们已在

$$2xdx + dx^2 \text{ 或 } 2x\dot{x} + \dot{x}\dot{x}$$

中看到，它们是二项式 $(x + dx)^2$ 或 $(x + \dot{x})^2$ 的正确的数学表示式（第二和第

三项)。至于**这个数学上正确的结果建立在数学上根本错误的假定之上**,即一开始就把 $x_1-x=\Delta x$ 当作 $x_1-x=dx$ 或 $\dot{x}$,这是人们所不知道的。不然的话,人们不用魔术般地丢掉而用最简单格式的代数运算也会获得同样的结果,并把它提供给数学界。

所以,人们自己就相信了这种新发现的算法的神秘性质。这种算法用数学上肯定是错误的方法得出了正确的(尤其在几何应用中惊人的)结果。这样,人们就把自己神秘化了,越加高估这个新发现,也就越加引起了一群旧式正统派数学家的恼怒,并激起了敌对的叫嚣,这种叫嚣甚至在数学界以外得到了共鸣,而这也是为新事物开拓道路所必需的。

2)**理性的微分演算**。达兰贝尔直接从**牛顿和莱布尼茨的出发点**:$x_1=x+dx$ 开始。但是他立刻做了一个根本的修正:$x_1=x+\Delta x$,也就是给 $x$ 加上一个**不确定的**,但初看起来是**有限的增量**,他把这个增量叫做 $h$。而这个 $h$ 或 $\Delta x$ 变为 $dx$(他和所有法国人一样,都采用莱布尼茨的写法),只是作为演化的最后结果,或者至少发生在最后一刹那之前,但神秘主义者和这种演算的创始者,却把它作为出发点(达兰贝尔本人从符号的一边出发,然而是在这一边变为符号之前出发的)。这样就立刻得到两种结果。

a)构成差值之比

$$\frac{f(x+h)-f(x)}{h}=\frac{f(x+h)-f(x)}{x_1-x}$$

的出发点是

1)$f(x+h)-f(x)$,它相应于一个用 $x$ 给定的代数函数,这个代数函数是在用 $x$ 表达的原函数例如 $x^3$ 中,以 $x$ 和它的增量即 $x+h$ 代替 $x$ 而得出的。这种形式(如果 $y=f(x)$,它就 $=y_1-y$)是**函数差值**的形式,为了变成函数增量对自变量增量的比值,这种形式还需要进行演化,因而它起着实在的、而不象在神秘主义者们那里仅仅是有名无实的作用;因为,如果我象这些人那样有

$$f(x)=x^3,$$
$$f(x+h)=(x+h)^3=x^3+3x^2h+3xh^2+h^3,$$

那末我一开始就知道,在

$$f(x+h)-f(x)=x^3+3x^2h+3xh^2+h^3-x^3$$

中对立着的两边,都已归结为增量。这甚至不必写出来,因为我在右边已看到 $x^3$ 的增量 = 其后面的三项,同样在 $f(x+h)-f(x)$ 中剩下的只是 $f(x)$ 的增量或 $dy$。因此,这个最初的差值等式只起着一开始就重新消失的作用。这些增

量一开始就处在对立的两边，如果我有了它们，那末从 $dx$，$dy$ 的定义我就得出$\frac{dy}{dx}$或$\frac{\dot{y}}{\dot{x}}$是比值等等；所以为了形成$\frac{dy}{dx}$或$\frac{\dot{y}}{\dot{x}}$，我用不到从（以 $x+h$ 代替 $x$）变化了的函数（增长了的函数）中减掉用 $x$ 表达的原函数而形成的那个最初的差值。

在达兰贝尔那里必须抓住这个差值，因为演化运动要在它身上进行。所以，在左边处于突出地位的就不是差值的肯定表示式，即不是增量，而是增量的否定表示式，即差值，也就是 $f(x+h)-f(x)$。这种强调差值而不强调增量（牛顿的流数）的做法，至少在莱布尼茨所用的、与牛顿的 $\dot{y}$ 相对立的写法 $dy$ 中，已经预感到了。

2) $f(x+h)-f(x)=3x^2h+3xh^2+h^3$。

用 $h$ 去除两边，便得

$$\frac{f(x+h)-f(x)}{h}=3x^2+3xh+h^2。$$

由此，在左边形成了

$$\frac{f(x+h)-f(x)}{h}=\frac{f(x+h)-f(x)}{x_1-x},$$

它（本身）就**显现为有限差值的导出的比值**，而在神秘主义者们那里，它是由 $dx$ 或 $\dot{x}$ 和 $dy$ 或 $\dot{y}$ 的定义所给出的增量的现成的比值。

3) 现在，如果在

$$\frac{f(x+h)-f(x)}{h}=\frac{f(x+h)-f(x)}{x_1-x}$$

中令 $h=0$，或 $x_1=x$，即 $x_1-x=0$，那末这表示式就变为$\frac{dy}{dx}$；而由于令 $h=0$，$3xh+h^2$ 这两项也就同时变为 0，并且确是通过正确的数学运算得到的。所以现在不用魔术就把它们消除了。我们得到：

4) $\frac{0}{0}$或$\frac{dy}{dx}=3x^2=f'(x)$。

一旦 $x$ 变为 $x+h$，这个式子就同在神秘主义者们那里一样，已作为给定的东西存在了，因为 $(x+h)^3$ 代替 $x^3$ 而给出了 $x^3+3x^2h+\text{etc.}$，其中 $3x^2$ 已经在这个展开级数的第二项中作为 $h$ 一次幂的**系数**出现。因此，这种推导本质上与莱布尼茨和牛顿相同，但是这个完全现成的导函数 $3x^2$ 是用严格的代数方法从与它相联系的其余各项那里**解脱出来**的。这不是**演化**，而是把 $f'(x)$，在这里就是 $3x^2$，从它的因子 $h$ 以及与它排在一起的其余各项那里**解脱出来**。而实际上演化的，倒是在左面符号一边，就是 $dx$，$dy$ 和它们的比值，即符号微

系数$\frac{dy}{dx}=\frac{0}{0}$(反过来,$\frac{0}{0}=\frac{dy}{dx}$更为恰当)。这个符号微系数本身还是引起了某些形而上学的恐惧,尽管这个符号是用数学方法导出的。

达兰贝尔给微分演算撕下了神秘的外衣,从而向前迈进了一大步。虽然他的《**流体论**》已于1744年出版(参见15页①),但莱布尼茨的方法很多年来仍然在法国占着优势。至于牛顿的方法在英国一直统治到十九世纪的头几十年,这几乎是没有必要提及的。但是在这里,正象以前在法国一样,达兰贝尔奠定的基础,经过一些修改,一直到现在还占着统治地位。

3) **纯粹代数的微分演算**。拉格朗日,《**解析函数论**》(1797年和1813年)。因此在1)中和在2)中一样,最初的出发点都是增长了的$x$;如果

$$y \text{ 或 } f(x)=\text{etc.},$$

那末在神秘的方法中是$y_1$或$f(x+dx)$,在理性的方法中是$y_1$或$f(x+h)$($=f(x+\Delta x)$)。这个二项式出发点立刻在等式的另一边提供了一个二项式展开,例如:

$$x^m+mx^{m-1}h+\text{etc.},$$

这里的第二项$mx^{m-1}h$已完全现成地提供了所要找的实在微系数$mx^{m-1}$。

a) 在给定的$x$的原函数中,当用$x+h$代替$x$时,左边的$f(x+h)$和它对面的展开级数的关系,完全象代数中**未展开的一般表示式**,尤其是二项式,和与它对应的**展开级数**的关系一样,例如,象在

$$(x+h)^3=x^3+3x^2h+\text{etc.}$$

中,$(x+h)^3$与其等价的展开级数$x^3+3x^2h+\text{etc.}$的关系一样。因此,$f(x+h)$就处于同整个代数学中一般表示式与其展开式的关系一样的代数关系(只应用于变量)之中,例如,同

$$\frac{a}{a-x}=1+\frac{x}{a}+\frac{x^2}{a^2}+\frac{x^3}{a^3}+\text{etc.}$$

中$\frac{a}{a-x}$与展开级数$1+\text{etc.}$一样的代数关系之中;或同

$$\sin(x+h)=\sin x\cos h+\cos x\sin h$$

中$\sin(x+h)$与其对面的展开一样的代数关系之中。

达兰贝尔仅仅把$(x+dx)$或$(x+\dot{x})$代数化了,使之变为$(x+h)$,因而$f(x+h)$也从$y+dy$,$y+\dot{y}$变成了$f(x+h)$。但是,当拉格朗日把整个表示式作为**一般的未展开的表示式**放在从它导出的展开级数对面时,他就给了这整

---

① 见本文第59页。

个表示式一个纯粹代数的性质。

b) 在第一个方法1)中和在理性的方法2)中一样，所要找的实在系数都是完全现成地由二项式定理产生出来的，而且已经出现在展开式的第二项，也就是必然附有 $h^1$ 的那一项中。因此，整个往后的微分过程，无论在1)或者在2)中，都是奢侈品。所以我们把这无用的压舱物抛在一边。从二项式的展开中我们一下子就全都知道，一阶的实在微系数是 $h$ 的因子，二阶的是 $h^2$ 的因子等等。这些实在微系数无非就是用 $x$ 表达的**原函数**依次按二项式展开的**导函数**(而引进**导函数**这一范畴是最重要的引进之一)。就各别的微分形式而论，我们知道 $\Delta x$ 变为 $dx$，$\Delta y$ 变为 $dy$，知道一阶导函数得到符号形象$\frac{dy}{dx}$，二阶导函数，即 $\frac{1}{2}h^2$ 的系数得到符号形象$\frac{d^2y}{dx^2}$等等。因此，为对称起见，我们可以把这些纯粹用代数方法得到的结果同时也用它们的符号微分等价物来表示——这是微分演算本身唯一遗留下来的名称问题。整个实际的课题于是就化为寻找(代数)方法，“把 $x+h$ 的各种函数按 $h$ 的整数升幂加以展开，而在许多情况下，没有非常冗长的运算这是不可能完成的”。

到此为止，在拉格朗日那里没有什么直接从达兰贝尔方法出发不能得到的东西(因为这方法也包括了神秘主义者们的整个演化，不过是以修正的形式罢了)。

c) 因此，当 $y_1$ 的展开或 $f(x+h)=$ etc. 代替以前的微分演算的时候〔这样一来，事实上清楚地暴露了从

$$y+dy \text{ 或 } y+\dot{y}, x+dx \text{ 或 } x+\dot{x}$$

出发的那些方法的秘密，即它们的实际演化是以二项式定理的应用为基础的，因为它们一开始就对增长了的 $x$ 用 $x+dx$，增长了的 $y$ 用 $y+dy$ 来表示，从而把单项式变为二项式〕①，就会有这样一个课题：由于 $f(x+h)$在这里是没有幂次的 $x$ 的函数，只是它的**一般的未展开的表示式**，所以要从这个未展开的表示式自己，用代数方法把一般的，因而也就是对任何幂次的 $x$ 的函数都适用的**展开级数**推导出来。

为了把微分演算代数化，拉格朗日用了**牛顿学派以及牛顿生前的泰勒**的定理作为他的直接出发点。这个定理实际上是最一般最概括的定理，同时也是微分学的运算公式，即用符号微系数表示的 $y_1$ 或 $f(x+h)$ 的展开级数，亦即：

---

① 这对括号是在马克思的手稿里原来就有的。

$$y_1 \text{ 或 } f(x+h)=y(\text{或 } f(x))+$$

$$+\frac{dy}{dx}h+\frac{d^2y}{dx^2}\frac{h^2}{[2]}+\frac{d^3y}{dx^3}\frac{h^3}{[2\cdot3]}+\frac{d^4y}{dx^4}\frac{h^4}{[2\cdot3\cdot4]}+\text{etc.}。$$

d）这里要插进对于麦克劳林和泰勒定理的研究。

e）拉格朗日用代数方法把 $f(x+h)$ 展开成等价级数，代替泰勒的 $\frac{dy}{dx}$ 等等，并且只让它们作为用代数方法导来的 $x$ 的函数的符号微分表示式而存在。（这一点以后再详细讨论。）

---

## ［初　　稿］

**牛顿**：1642年生，1727年死（85岁）。《**自然哲学的数学原理**》（1687年初版，参见**引理Ⅰ**和**引理XI，注释**）。

而后特别是《**使用级数，流数的分析**…》，1711年第一次出版，但在1665年已经完成，而莱布尼茨于1676年才获得同样的发现。

**莱布尼茨**：1646年生，1716年死（70岁）。

**拉格朗日**：1736年生，帝制时期（拿破仑一世）才死，**变分法**的发明者。《**解析函数论**》（1797年和1813年）。

**达兰贝尔**：1717年生，1783年死（66岁）。《**流体论**》，1744年。

1）**牛顿**。速度或流数，例如 $x$，$y$ 等**变量**的速度或流数，**用** $\dot{x}$，$\dot{y}$ 等**表示**。例如，若 $u$ 和 $x$ 是**由连续运动所产生的相关联的量**（**流量**），则 $\dot{u}$ 和 $\dot{x}$ 表示它们的增长率，因此 $\frac{\dot{u}}{\dot{x}}$ 就是它们的增量据以生成的增长率之间的**比值**。

由于各种各样的量的数值都可用直线表示，所以**瞬**或者生成量的无限小部分＝它们的速度和这些速度所经历的无限小时间间隔的乘积，因此如用 $\tau$ 表示这无限小的时间间隔，那末 $x$ 和 $y$ 的瞬就分别用 $\tau\dot{x}$ 和 $\tau\dot{y}$ 来表示。

例如 $y=uz$；若 $\dot{y},\dot{z},\dot{u}$ 分别表示 $y,z,u$ 增长的速度，则 $\dot{y},\dot{z},\dot{u}$ 的**瞬**便是 $\tau\dot{y},\tau\dot{z},\tau\dot{u}$，于是我们得到：

$$y=uz,\ y+\tau\dot{y}=(u+\tau\dot{u})(z+\tau\dot{z})=uz+u\tau\dot{z}+z\tau\dot{u}+\tau^2\dot{u}\dot{z};$$

因此

$$\tau\dot{y}=u\tau\dot{z}+z\tau\dot{u}+\tau^2\dot{u}\dot{z}。$$

由于 $\tau$ 无限小，所以它会自行消失，就更不用说作为乘积的 $\tau^2\dot{u}\dot{z}$ 了，这不是出于无限小的时间间隔 $\tau$，而是它的2次方。$\left(\text{假如 }\tau=\frac{1}{\text{百万}}\text{，那末 }\tau^2=\frac{1}{\text{百万}\times\text{百万}}\text{。}\right)$

因此我们得到

$$\dot{y}=\dot{u}z+\dot{z}u,$$

或者说 $y=uz$ 的流数是 $\dot{u}z+\dot{z}u$。

2）**莱布尼茨**。假定要找的是 $uz$ 的微分。

$u$ 变为 $u+du$，$z$ 变为 $z+dz$；因此

$$uz+d(uz)=(u+du)(z+dz)=uz+udz+zdu+dudz。$$

如果我们从这式子中减去给定的量 $uz$，那末剩下来的是作为增量的 $udz+zdu+dudz$；由无限小的 $du$ 乘上另一个无限小的 $dz$ 组成的乘积 $dudz$，是一个二阶无限小，它的消失将先于一阶无限小 $udz$ 和 $zdu$，所以

$$d(uz)=udz+zdu。$$

3）**达兰贝尔**。一般这样来提问题。设

$$y=f(x),$$

$$y_1=f(x+h);$$

要确定：当 $h$ 这个量消失时，$\frac{y_1-y}{h}$ 的值将变为什么，也就是说，$\frac{0}{0}$ 的值将变为什么。

---

牛顿和莱布尼茨，以及他们的大多数继承者，一开始就在微分演算的地盘上活动，因此微分表示式一开始就当作尔后去找实在等价物的运算公式。全部奥妙就在于此。如果自变量 $x$ 变为 $x_1$，那末因变量就变为 $y_1$。但是 $x_1-x$ 必然等于某个差值，例如 $=h$。这包含在变量概念本身之中。然而决不能由此得出这样的结论：这个差值，$=dx$，是消失着的，因而事实上 $=0$。它也可以表示一个有限差值。可是如果我们一开始就假定，当 $x$ 增长时它变为 $x+\dot{x}$（牛顿的 $\tau$ 在他对基本函数的分析中不起任何作用，所以可镇压掉），或者象莱布尼茨那样，变为 $x+dx$，那末微分表示式就立刻变为运算符号，而没有显示出它们的代数起源。

---

补充 15①（**牛顿**）。

对要进行微分的乘积 $uz$，如果我们取牛顿的起始等式，那末：

$$y=uz,$$

$$y+\tau\dot{y}=(u+\dot{u}\tau)(z+\dot{z}\tau)。$$

如果我们丢掉 $\tau$，就象牛顿在展开第一个微分等式之后自己所喜欢做的那样，便会得到：

$$y+\dot{y}=(u+\dot{u})(z+\dot{z}),$$

$$y+\dot{y}=uz+\dot{u}z+\dot{z}u+\dot{u}\dot{z},$$

$$y+\dot{y}-uz=\dot{u}z+\dot{z}u+\dot{u}\dot{z}。$$

① 指《论微分》底稿中的页码，见本文第 34 页。

因此，由于 $uz=y$，

$$\dot{y}=\dot{u}z+\dot{z}u+\dot{u}\dot{z}。$$

并且为了得到正确的结果，必须把 $\dot{u}\dot{z}$ 镇压掉。

那末这个要用暴力去镇压掉的项 $\dot{u}\dot{z}$ 是从哪里来的呢？

非常简单，是这样来的：$y$ 的微分 $\dot{y}$，$u$ 的微分 $\dot{u}$ 和 $z$ 的微分 $\dot{z}$，一开始就是通过定义，当作与产生它们的那些变量相分开的、独立的存在而引进来的，并不是用任何一种数学方法推导出来的。

我们一方面看到，这种预先假定 $dy$，$dx$ 或 $\dot{y}$，$\dot{x}$ 的存在具有什么样的好处：一旦变量增长，我只须一开始就把二项式 $y+\dot{y}$，$x+\dot{x}$ 等代入代数函数中去，就可以把它们作为普通代数量来处理。

例如当 $y=ax$ 时，我便得到

$$y+\dot{y}=ax+a\dot{x};$$

从而

$$y-ax+\dot{y}=a\dot{x};$$

因此

$$\dot{y}=a\dot{x}。$$

这样，我立刻就得到下面的结果：因变量的微分等于 $ax$ 的增长，即 $a\dot{x}$，也就是等于从 $ax$ **导出的实在值** $a$（至于它在这里是个常量，属于偶然，一点也不影响所得结果的普遍性，因为出现这种情况，只应归之于变量 $x$ 在这里是一次幂）乘上 $\dot{x}$。如果我把这结果加以普遍化，那末我知道 $y=f(x)$，因为它表明 $y$ 是一个依赖于 $x$ 的变量。如果我把从 $f(x)$ 导出的量，也就是增量的实在元素叫做 $f'(x)$，那末一般的结果便是

$$\dot{y}=f'(x)\dot{x}。$$

因此我一开始就知道，因变量 $y$ 的微分的等价物，等于自变量的一阶导函数乘上它的微分，即乘上 $dx$ 或 $\dot{x}$。

所以一般地说，如果

$$y=f(x),$$

那末

$$dy=f'(x)dx$$

或者 $\dot{y}=$ 用 $x$ 表达的实在系数（除了因 $x$ 处于一次幂而出现常量的情况外）乘上 $\dot{x}$。

但 $\dot{y}=a\dot{x}$ 立刻给了我 $\frac{\dot{y}}{\dot{x}}=a$，而在一般情况下：

$$\frac{\dot{y}}{\dot{x}}=f'(x)。$$

这样，我就给微分和微系数找到了两个进一步发展了的运算公式，它们构成了整个微分演算的基础。

此外，一般说来，由于我把先验地①假定的 $dx$，$dy$ 等等，或者 $\dot{x}$，$\dot{y}$ 等等，当作 $x$ 和 $y$ 的独立的、孤立的增量而得到了微分演算所特有的极大好处，这就是说，变量的一切函数从一开始就可以用微分形式来表示。

如果我用这种方法对变量的一些主要的函数，如 $ax$，$ax\pm b$，$xy$，$\frac{x}{y}$，$x^n$，$a^x$，$\log x$，以及初等的圆函数进行了推导，那末当要找 $dy$，$\frac{dy}{dx}$时，我就完全可以象算术中的乘法表那样来利用它们。

但是，如果我们现在看一下事情的反面，那末我们立刻发现，原先的全部运算在数学上都是错误的。

我们举一个非常简单的例子：$y=x^2$。如果 $x$ 增长，那末它就得到一个不确定的增量 $h$，因而依赖于它的因变量 $y$ 也得到了一个不确定的增量 $k$，于是我们有

$$y+k=(x+h)^2=x^2+2hx+h^2,$$

这是由二项式定理给我们的一个公式。因此

$$y+k-x^2 \text{ 或 } y+k-y=2hx+h^2;$$

所以

$$(y+k)-y \text{ 或 } k=2hx+h^2;$$

如果用 $h$ 除两边，那末

$$\frac{k}{h}=2x+h。$$

现在令 $h=0$，就得到

$$2x+h=2x+0=2x。$$

但另一方面，$\frac{k}{h}$却变为$\frac{k}{0}$；又由于仅当 $x$ 变为 $x+h$ 时 $y$ 才变为 $y+k$，所以当 $h$ 变为 0，因而 $x+h$ 又变为 $x+0$，即变为 $x$ 时，$y+k$ 重新变为 $y$。因此，$k$ 也将变为 0 而 $\frac{k}{0}=\frac{0}{0}$，这可用 $\frac{dy}{dx}$或$\frac{\dot{y}}{\dot{x}}$来表示。于是我们得到

$$\frac{0}{0} \text{ 或 } \frac{\dot{y}}{\dot{x}}=2x。$$

① “先验地”一词，原文为 a priori。

如果我们不这样做,而在

$$y+k-x^2=2hx+h^2 \text{ 或 } (y+k)-y=2xh-h^2$$

中令 $h=0$ ($h$ 只有事先在它的原来形式中被令为 0 之后才变为符号 $dx$),那末得到 $k=0+0=0$,而我们所获得的唯一结果,是看清楚了我们的假定:当 $x$ 变为 $x+h$ 时 $y$ 才变为 $y+k$,……所以当 $x+h=x+0=x$ 时,$y+k=y$,或者 $k=0$。

但是我们绝不会象牛顿所做的那样得到:

$$k=2xdx+dxdx,$$

或者按牛顿的写法得到:

$$\dot{y}=2x\dot{x}+\dot{x}\dot{x};$$

只有当 $h$ 经历了通过 0 的地狱之行以后,也就是说,在 $x_1-x$(或$(x+h)-x$)这个差值,因而 $y_1-y(=(y+k)-y)$这个差值也被减少到它们的绝对最小表示式 $x-x=0$ 和 $y-y=0$ 以后,$h$ 才能化为 $\dot{x}$,从而 $k$ 化为 $\dot{y}$。

但由于牛顿不是用数学推导来确定变量 $x,y$ 等等的增量,而是立刻给它们打上了微分 $\dot{x},\dot{y}$ 等等的标记,这些增量就不能 $=0$;否则结果势将为 0,因为用代数表达的话,一开始令这些增量 $=0$,就象上面在等式

$$(y+k)-y=2xh+h^2$$

中一样立刻令 $h$ 等于 0,所以 $k=0$,从而最终就会得出 $0=0$。在用除法使 $x$ 的一阶导函数,即这里的 $2x$,从因子 $h$ 那里解放出来之前,也就是在获得

$$\frac{y_1-y}{h}=2x+h$$

之前,是不允许将 $h$ 化为零的。只有在这以后,有限差值才能被扬弃。因此在我们能够获得微分

$$dy=2xdx$$

之前,原来就必须先导出微系数

$$\frac{dy}{dx}=2x。$$

所以没有其他办法,只有把变量的增量 $h$ 想象为无限小增量,并赋予它们以**独立的存在**,例如在 $\dot{x},\dot{y}$ 或 $dx,dy$ 等等符号中那样。然而,无限小量和无限大量一样,也是量(无限这个词实际上只意味着不确定地小),因此这些 $dy,dx$ 或 $\dot{y},\dot{x}$ 等等在计算中也起着和普通的代数量一样的作用,并且在上述等式

$$(y+k)-y \text{ 或 } k=2xdx+dxdx$$

中，$dxdx$ 具有和 $2xdx$ 一样的存在权利。但是最使人惊奇的是用暴力把它镇压掉的那个理由，恰恰在于利用了无限小概念的相对性。$dxdx$ 要镇压掉，因为它与 $dx$ 相比，从而也与 $2xdx$ 或与 $2x\dot{x}$ 相比是无限小……

或者，如果在

$$\dot{y}=\dot{u}z+\dot{z}u+\dot{u}\dot{z}$$

中，$\dot{u}\dot{z}$ 由于与 $\dot{u}z$ 和 $\dot{z}u$ 相比是无限小而被镇压掉，那末在数学上只能这样来补救，那就是认为 $\dot{u}z+\dot{z}u$ 是人们愿意怎样逼近就能怎样逼近的一个近似值。这样一种技巧在普通代数里也会遇到。但这时就出现了一个更大的奇迹，即通过这种方法得到的，绝不是用 $x$ 表达的导函数的近似值，而是它的严格的准确值（尽管象上面那样仅仅符号上是正确的），象在

$$\dot{y}=2x\dot{x}+\dot{x}\dot{x}$$

这个例子中那样。如果我们在这里镇压掉 $\dot{x}\dot{x}$，那末就得到

$$\dot{y}=2x\dot{x}$$

和

$$\frac{\dot{y}}{\dot{x}}=2x,$$

正如二项式定理已经表明的那样，这是由 $x^2$ 正确地导出的第一个函数。

但这奇迹并非什么奇迹。如果通过暴力镇压掉 $\dot{x}\dot{x}$ 而**得不出准确的结果**，那倒真是奇迹了。**因为镇压掉的只是一个计算误差**，而它却是这样一种方法不可避免的结果，这种方法把变量的不确定的增长，例如 $h$，立刻作为微分 $dx$ 或 $\dot{x}$，作为现成的运算符号引了进来，从而在微分演算中一开始也就获得了一个独特的、不同于普通代数的计算方法。

---

我们所应用的代数方法的过程，一般可以表述如下。

设给定 $f(x)$，那末首先演化出“预先导函数”，我们称之为 $f^1(x)$：

1) $$f^1(x)=\frac{\Delta y}{\Delta x}，\text{或}\ \frac{\Delta y}{\Delta x}=f^1(x)。$$

从这等式得出：

$$\Delta y=f^1(x)\Delta x。$$

因而也得出

$$\Delta f(x)=f^1(x)\Delta x$$

（因为 $y=f(x)$，$\Delta y=\Delta f(x)$）。

通过令 $x_1-x=0$，因而 $y_1-y=0$，我们得到

$$\frac{dy}{dx} = f'(x)。$$

于是

$$dy = f'(x)dx;$$

所以也得

$$df(x) = f'(x)dx$$

(因为 $y = f(x), dy = df(x)$)。

如果我们一旦演化出

1) $$\Delta f(x) = f^1(x)\Delta x,$$

那末

2) $$df(x) = f'(x)dx$$

不过是1)的微分表示式而已。

1) 如果 $x$ 变到 $x_1$,那末

A) $x_1 - x = \Delta x$;

由此得出下列结论:

Aa) $\Delta x = x_1 - x$;  a) $x_1 - \Delta x = x$;

所以,$\Delta x$ 这个 $x_1$ 和 $x$ 之间的**差值**被肯定地表示时,它是 $x$ 的**增量**;因为当重新从 $x_1$ 中减去它时,$x_1$ 就回到了其原来状态,即回到了 $x$。

因此,这个差值可用两种方式来表示:**直接作为**增长了的变量与其增长前的状态之间的**差值**,这是它的**否定表示式**;肯定地把它作为增量①,**作为结果**、**作为**对于尚未增长时的 $x$ 的**增量**,这是肯定表示式。

我们将会看到,这种双重解释在微分演算历史中起了什么样的作用。

b) $x_1 = x + \Delta x$。

$x_1$ 是增长了的 $x$ 本身,$x$ 的增长没有和它相分开;$x_1$ 是 $x$ 增长的完全不确定的形式;这形式把增长了的 $x$,即 $x_1$,同增长前的原来形式 $x$ 区别了开来,但是这形式并没有把 $x$ 同它的增量本身相区别。因此,$x_1$ 和 $x$ 之间的关系只能否定地作为**差值**,即作为 $x_1 - x$ 来表示。与此相反,在

$$x_1 = x + \Delta x$$

中:

1) 差值是**肯定地**作为 $x$ 的增量来表示的。

---

① 马克思在这里用铅笔加上了"或减量"。

2）因此 $x$ 的增长不是作为**差值**，而是作为处在原来状态中的 $x$ 本身 + 它的增量的**和式**来表示的。

3）从技术上来表述时，$x$ 将从单项式变为二项式，凡是原函数中有 $x$ 的任何幂次出现的地方，都要用 $x$ **自己和它的增量**组成的二项式来代替增长了的 $x$，一般地说，用二项式 $(x+h)^m$ 来代替 $x^m$。这样，$x$ 的增长的展开事实上就是**二项式定理**的简单应用。由于在这个二项式中 $x$ 是第一项，$\Delta x$ 是第二项——这是由它们的相互关系本身所给定了的，因为在增量 $\Delta x$ 产生之前 $x$ 就必须已经存在——所以事实上用二项式只是导出了 $x$ 的一些函数，而 $\Delta x$ 则以升幂方式出现在它们的旁边作为因子，并且 $\Delta x$ 的一次幂即 $\Delta x^1$，应该是展开级数第二项的因子，即由二项式定理导出的 $x$ 的一阶导函数的因子。当 $x$ 以二次幂给定时，就立刻可以看出这一点。$x^2$ 将变为 $(x+\Delta x)^2$，这无非是 $(x+\Delta x)$ 和它自己**相乘**而给出 $x^2+2x\Delta x+\Delta x^2$。这就是说，$x$ 的原函数应该是第一项；$x^2$ 的一阶导函数，即这里的 $2x$，和因子 $\Delta x^1$ 一起构成第二项，而在第一项中 $\Delta x$ 仅以 $\Delta x^0=1$ 的形式出现。所以导函数不是通过取差值而是通过应用二项式定理即通过**乘法**找到的，并且正是因为增长了的 $x_1$ 本身一开始就表现为二项式 $x+\Delta x$ 的缘故。

4）虽然在 $x+\Delta x$ 中，$\Delta x$ 就其大小而言是不确定的，象不确定的变量 $x$ 自己一样；但是 $\Delta x$ 作为与 $x$ 有区别的、与之相分离的量又是确定的，如同胎儿与怀孕前的母亲并列着一样。

$x+\Delta x$ 不只是不确定地表明 $x$ 作为变量有所增长，而且它表明 $x$ 增长了**多少**，即增长了 $\Delta x$。

5）$x$ 从不显现为 $x_1$；当应用二项式定理，也就是把 $x+\Delta x$ 代入 $x$ 的一定幂次中而找到了导函数的时候，整个演化就围绕着增量 $\Delta x$ 进行。只是在左边，当 $\frac{y_1-y}{\Delta x}$ 中的 $\Delta x$ 变为

$$\Delta x=0$$

时，它最终又以 $=x_1-x$ 的形式出现，以致

$$\frac{y_1-y}{\Delta x}=\frac{y_1-y}{x_1-x}\left(=\frac{\Delta y}{\Delta x}\right)。$$

所以寓于 $x_1-x=0$ 中肯定的一面，即 $x_1$ 之变为 $=x$，决不能在演化中出现，因为 $x_1$ 本身决不会出现在展开级数的一边；这样，微分演算的真正秘密也就决不会暴露出来。

6）如果 $y=f(x)$ 且 $y_1=f(x+\Delta x)$，那末我们可以说，在这个方法中 $y_1$ 的

**展开解决了找导函数的问题。**

c) $x+\Delta x=x_1$（因而 $y+\Delta y=y_1$）。$\Delta x$ 在这里只能以 $\Delta x=x_1-x$ 的形式出现，也就是以 $x_1$ 和 $x$ 的**差值**这种**否定**形式，而不是象在 $x_1=x+\Delta x$ 中那样作为 $x$ 的增量那种**肯定**形式出现。

1）在这里，增长了的 $x$，即 $x_1$，有别于增长前的**它本身**，即有别于 $x$，但 $x_1$ 并不是作为增长了 $\Delta x$ 的 $x$ 而出现的；因此 $x_1$ 实际上完全象 $x$ 一样是不确定的。

2）再者：正象 $x$ 出现在一个原函数中一样，作为增长了的 $x_1$ 也就出现在由于这个增长而变化了的原函数中。例如 $x$ 出现在函数 $x^3$ 中时，$x_1$就出现在函数 $x_1^3$ 中。以前，在原函数中凡有 $x$ 出现的地方，都用$(x+\Delta x)$来代替，从而由二项式定理完全现成地得出了导函数，尽管这个导函数附有因子 $\Delta x$，而且是作为附有 $\Delta x^2$ 等等因子并用 $x$ 表达的其他各项的先行者出现的；而现在，从 $x_1^3$ 也就是从增长了的 $x$ 的单项式这一直接形式中，就象从 $x^3$ 中一样，是直接推导不出什么的。但是由此给出的是**差值** $x_1^3-x^3$。我们从代数学知道，所有形如 $x^3-a^3$ 的差值都可用$(x-a)$除尽，在现在的情况下就是可用$(x_1-x)$除尽。所以用 $x_1-x$ 去除 $x_1^3-x^3$（而不是象以前那样让$(x+\Delta x)$按函数的给定幂次自乘若干次），我们就获得一个形如$(x_1-x)P$ 的表示式，不管 $x$ 的原函数是多项式（即含有 $x$ 的不同幂次），还是象在我们这例子中那样是单项式，在这里都不会有什么改变。通过除法，这个$(x_1-x)$变为左边 $y_1-y$ 的分母，因而产生了$\frac{y_1-y}{x_1-x}$这个用抽象的差值形式表示的函数的差值和自变量 $x$ 的差值之比。要把用 $x_1$ 表达的函数和用 $x$ 表达的函数之间的差值分解成都含有$(x_1-x)$作为因子的各项，就要根据 $x$ 的原函数的性质，或多或少用到一些代数技巧，不会总是象在 $x_1^3-x^3$ 中那样容易。但这在方法上并没有什么改变。

凡是原函数按其性质不能直接分解为$(x_1-x)P$ 的地方，如 $f(x)=uz$（两个依赖于 $x$ 的变量）就是这样，$(x_1-x)$ 便出现在因子$\frac{1}{x_1-x}$中。再者，用$(x_1-x)$去除两边而在左边把它消掉以后，如果在 $P$ 中还保留有 $x_1-x$ 的话（例如对 $y=a^x$ 求导时，我们求得

$$\frac{y_1-y}{x_1-x}=a^x\left\{(a-1)+\frac{(x_1-x)-1}{1\cdot 2}(a-1)^2+\text{etc.}\right\},$$

若在这里令 $x_1-x=0$，则得

$$=a^x\left\{(a-1)-\frac{1}{2}(a-1)^2+\frac{1}{3}(a-1)^3-\text{etc.}\right\}),$$

那末象在刚才引用的例子中那样，总是只能这样做：通过令 $x_1-x=0$ 来使它消失，从而在它的位置上总会留下一个肯定的结果。换句话说，在 $P$ 中还保留下来的 $x_1-x$ 是不可能作为因子（作为乘数）与 $P$ 的其他成份联结在一起的。不然的话，就可把 $P$ 分解成 $P=p(x_1-x)$，而由于已令 $x_1-x=0$，它就分解成 $p\cdot 0$；因而 $P=0$……

---

如果 $y=x^3$ 从而 $y_1=x_1^3$，那末一阶有限差值 $x_1^3-x^3$ 就演化成

$$y_1-y=(x_1-x)P,$$

于是

$$\frac{y_1-y}{x_1-x}=P。$$

$P$ 这个由 $x_1$ 和 $x$ 组成的表示式 $=f^1$，也就是一阶有限差值的导函数，其中 $x_1-x$ 象高次幂 $(x_1-x)^2$ 等等一样已被消去。因此 $x_1$ 和 $x$ 只能结合在象 $x_1+x$，$x_1x$，$\frac{x_1}{x}$，$\sqrt{x_1x}$ 等等这样一些肯定的表示式中。所以如果现在令 $x_1=x$，那末这些表示式就相应地变为 $2x$，$x^2$，$\frac{x}{x}$ 或 $1$，$\sqrt{xx}$ 或 $x$ 等等，而只是在 $x_1-x$ 作为分母的左边才出现 $0$，因而才出现符号微系数等等。

---

## [续　　稿]

c) 续 25 页①。

我们原先把 $x_1 - x = \Delta x$ 当作差值 $x_1 - x$ 的表示式；这个**差值在这里只是在它的形式上作为差值而存在**（正象当 $y$ 为 $x$ 的因变量时，通常都写成 $y_1 - y$)。通过令 $x_1 - x = \Delta x$，我们就给了这个差值一个与它本身不同的表示式。即使用的是不确定的形式，我们还是把**这个差值的数值**作为与这个差值本身不同的量来表示。例如 $4-2$ 是 4 和 2 的差值的纯粹表示式；但是 $4-2=2$ 是用(右边的) 2 表示的差值：a)因而在肯定的形式中，它已经不再作为差值；b)当减法已做完，差值已算出，$4-2=2$ 就给我提供了 $4=2+2$。第二个 2 在这里以**原来的** 2 的**增量的**肯定**形式**出现，因而直接以**一个与差值形式相对立的形式**出现。（正如 $a-b=c$，$a=b+c$，其中 $c$ 表现为 $b$ 的增量；又如 $x_1 - x = \Delta x$，$x_1 = x + \Delta x$，其中 $\Delta x$ 直接表现为 $x$ 的增量。）

原先仅仅设置 $x_1 - x = \Delta x =$ 某个东西的做法，也就设置了代替差值形式的另一个形式，即和的形式 $x_1 = x + \Delta x$，同时把只表示差值的 $x_1 - x$ 看作是这个差值的，即量 $\Delta x$ 的数值等价物。

同样，从 $x_1 - x = \Delta x$ 可以得到 $x_1 - \Delta x = x$。这里，我们在左边又有一个差值形式，但这是作为增长了的 $x_1$ 和它自己的、独立地出现在它旁边的增量之间的差值形式。$x_1$ 和 $x$ 的增量 $\Delta x$ 之间的差值是这样的一个差值，它虽然不确定，但现在已表示 $x$ 的一个确定值。

然而如果从神秘的微分演算出发，在那里 $x_1 - x$ 立刻以 $x_1 - x = dx$ 的形式出现，又如果一开始就把 $dx$ 修正为 $\Delta x$，那末这就是从 $x_1 - x = \Delta x$ 出发，因而也就是从 $x_1 = x + \Delta x$ 出发；但是这式子本身又可以转变为 $x + \Delta x = x_1$，以致 $x$ 的增长又获得不确定的形式 $x_1$，而且作为这样的增长直接出现在微分演算中。这就是我们所应用的代数方法的出发点。

d) 从这些简单的形式上的差别，立刻产生了对待微分演算的一个根本差别。这个根本差别，我们在分析达兰贝尔方法时已经详细指出过(见所附的几张活页)。这里只是一般地作些评述。

---

① 指微分演算的历史发展过程《初稿》中的页码，见本文第 64 页。

1）如果**差值** $x_1 - x$（因而 $y_1 - y$）立即作为它的对立物，**作为和式** $x_1 = x + \Delta x$ 出现，因而其数值大小立即以**增量** $\Delta x$ **的肯定形式**出现，那末当在用 $x$ 表达的**原函数**中处处以 $x + \Delta x$ 代替 $x$ 时，就要展开一定幂次的二项式，而 $x_1$ 的演化便变为**二项式定理的应用**。二项式定理无非是一次的二项式自**乘** $m$ 次后所给出的表示式的一般表述。因此，如果我们一开始就把**差值**表示为**它的对立物**，即表示为**和式**，那末**乘法**就成了 $x_1$（即 $x + \Delta x$）的演化方法。

2）由于在一般的表示式 $x_1 = x + \Delta x$ 中，处于肯定形式 $\Delta x$，即**增量**形式下的差值 $x_1 - x$ 是这表示式的**后一项或第二项**，所以，一旦用 $x$ 表达的原函数换成用 $x + \Delta x$ 表达的函数，$x$ 就是它的第一项而 $\Delta x$ 是第二项。但是我们从二项式定理知道，这第二项附在第一项的旁边只表现为升幂的因子，表现为乘数，因而用 $x$ 表达的（由二项式的幂次决定的）第一个表示式的因子是 $(\Delta x)^0 = 1$，第二项的乘数是 $(\Delta x)^1$，第三项的是 $(\Delta x)^2$ 等等。因此以增量的肯定形式表示的差值，只是作为乘数而出现，而且第一次实际上作为乘数而出现（由于 $(\Delta x)^0 = 1$）是在二项式 $(x + \Delta x)^m$ 展开的第二项中。

3）另一方面，如果我们现在来看用 $x$ 本身表达的函数的展开，那末二项式定理就为这里的第一项 $x$，给我们提供了一连串导函数。例如，如果我们有 $(x + h)^4$，在这代数二项式中，$h$ 是已知量，$x$ 是未知量，那末我们得到

$$x^4 + 4x^3h + \text{etc.},$$

因此，处于第二项中的、并以 $h$ 的一次幂为其因子的 $4x^3$，是 $x$ 的一阶导函数，或者用代数方法来表述：如果我们有**二项式** $(x + h)^4$ **的未展开的表示式**，那末它的展开级数给我们提供了 $4x^3$ 作为 $x^4$ 的第一个增长（作为它的增量），这个 $4x^3$ 是作为 $h$ 的系数出现的。但如果 $x$ 是个变量，且我们有 $f(x) = x^4$，那末这函数将由于它的增长本身而变为 $f(x + h)$，或者在第一种形式下变为：

$$f(x + \Delta x) = (x + \Delta x)^4 = x^4 + 4x^3\Delta x + \text{etc.}。$$

在普通的代数二项式 $(x + h)^4$ 中，$x^4$ 是作为这二项式的第一项给出的；而在变量 $x$ 的二项式表示中，亦即在 $(x + \Delta x)^4$ 中，这 $x^4$ 现在就显现为在 $x$ 增长并且变为 $x + \Delta x$ 之前用 $x$ 表达的原函数的再生产。由二项式定理的本性，我们一开始就很清楚，如果 $f(x) = x^4$ 变为 $f(x + h) = (x + h)^4$，那末 $(x + h)^4$ 的第一项等于 $x^4$，亦即必须 = 用 $x$ 表达的原函数；$(x + h)^4$ 必须包括两部分，即用 $x$ 表达的原函数（这里是 $x^4$）+ 因 $x^4$ 变为 $(x + h)^4$ 而得来的一切附加项，因此二项式 $(x + h)^4$ 的第一项……

4）再者，二项式展开的第二项 $4x^3h$ 立刻**完全现成地**给我们提供了 $x^4$ 的一阶导函数，即 $4x^3$。所以这种推导是通过

$$f(x+\Delta x)=(x+\Delta x)^4$$

的展开而获得的，是通过把差值 $x_1-x$ 一开始就表示为它的**对立物**，表示为**和式** $x+\Delta x$ 而获得的。

所以由于 $x$ 增长而从 $f(x)$ 获得的 $f(x+\Delta x)$ 或 $y_1$ 的二项式展开，给我们提供了一阶导函数，它是(二项式级数中)$h$ 的系数，而且就在二项式展开的开头，即在其第二项中。因此，这种推导决不是通过取差值得来的，而是直接通过把 $f(x+h)$ 或 $y_1$ 展开成一个确定的、由简单乘法产生的表示式而得来的。

因此这方法的关键，在于把不确定的表示式 $y_1$ 或 $f(x+h)$ 展开成确定的二项式形式，而绝不是把 $x_1-x$，从而也不是把 $y_1-y$ 或 $f(x+h)-f(x)$ 作为差值来加以展开。

5）由于我们立刻得到

$$f(x+\Delta x)=(x+\Delta x)^4=x^4+4x^3\Delta x+6x^2\Delta x^2+4x\Delta x^3+\Delta x^4,$$

所以这方法中出现的唯一的差值等式，如果我们把它写下的话，就是

$$x^4+4x^3\Delta x+6x^2\Delta x^2+4x\Delta x^3+\Delta x^4-x^4,$$

也就是把级数开端的原函数 $x^4$ 在后面重新减掉，于是在我们的面前就出现一个**增量**，它是用 $x$ 表达的原函数通过二项式展开而得到的。因此牛顿也是这样写的。所以我们得到增量

$$4x^3\Delta x+6x^2\Delta x^2+4x\Delta x^3+\Delta x^4,$$

原函数 $x^4$ 的增量。因而我们在对立的一边就不需要**任何一种差值**了。当

$$y \text{ 或 } f(x)=x^4$$

时，相应于 $x$ 的增量有 $y$ 的增量。因而牛顿也立刻写为：

$$dy\text{，在他那里是 } \dot{y}=4x^3\dot{x}+\text{etc.}。$$

6）此后的整个演化，就在于把完全现成的导函数 $4x^3$ 从其因子 $\Delta x$ 和各相邻项那里解放出来，从其环境中摆脱出来。所以这个方法不是演化法而是**解脱法**。

e）**(作为一般表示式的)$f(x)$的微分**。

首先我们指出，“导函数”这个概念作为符号微系数的逐次实在等价物，是微分演算原先发明人及其最初继承者们所完全不知道的，事实上是首先由拉格朗日引进的。在前者那里，只有因变量例如 $y$ 才显现为 **$x$ 的函数**，它完全相应于**函数**的原来代数意义，这种函数最初用于未知量多于方程个数的那些所谓不定方程，所以在那里随着 $x$ 取不同的值，例如 $y$ 也就取不同的值。但在拉格朗日这里，原函数是应予微分的、$x$ 的确定的代数表示式；因此，如果 $y$ 或

$f(x)=x^4$,那末 $x^4$ 是原函数，$4x^3$ 是一阶导函数等等。所以为避免混乱起见，应把因变量 $y$ 或者 $f(x)$称为 $x$ **的函数**,而相反地把拉格朗日意义下的原函数称为**用 $x$ 表达的原函数**，并相应地把那些“导”函数称为**用 $x$ 表达的**“导”函数。

在代数方法中，我们首先演化出预先导函数或有限差值之比 $f^1$,而后才由此导出最终导函数 $f'$。在这个方法中我们一开始就知道：$f(x)=y$，所以 a) $\Delta f(x)=\Delta y$,因而反过来 $\Delta y=\Delta f(x)$。于是首先要演化的，恰恰是 $\Delta f(x)$，即 $f(x)$ 的有限差值。

我们求得：

$$f^1(x)=\frac{\Delta y}{\Delta x}，\text{所以}\frac{\Delta y}{\Delta x}=f^1(x)。$$

因而也得：

$$\Delta y=f^1(x)\Delta x,$$

并且由于 $\Delta y=\Delta f(x)$,所以

$$\Delta f(x)=f^1(x)\Delta x。$$

这微分表示式的进一步演化，给我们最后提供了

$$df(x)=f'(x)dx,$$

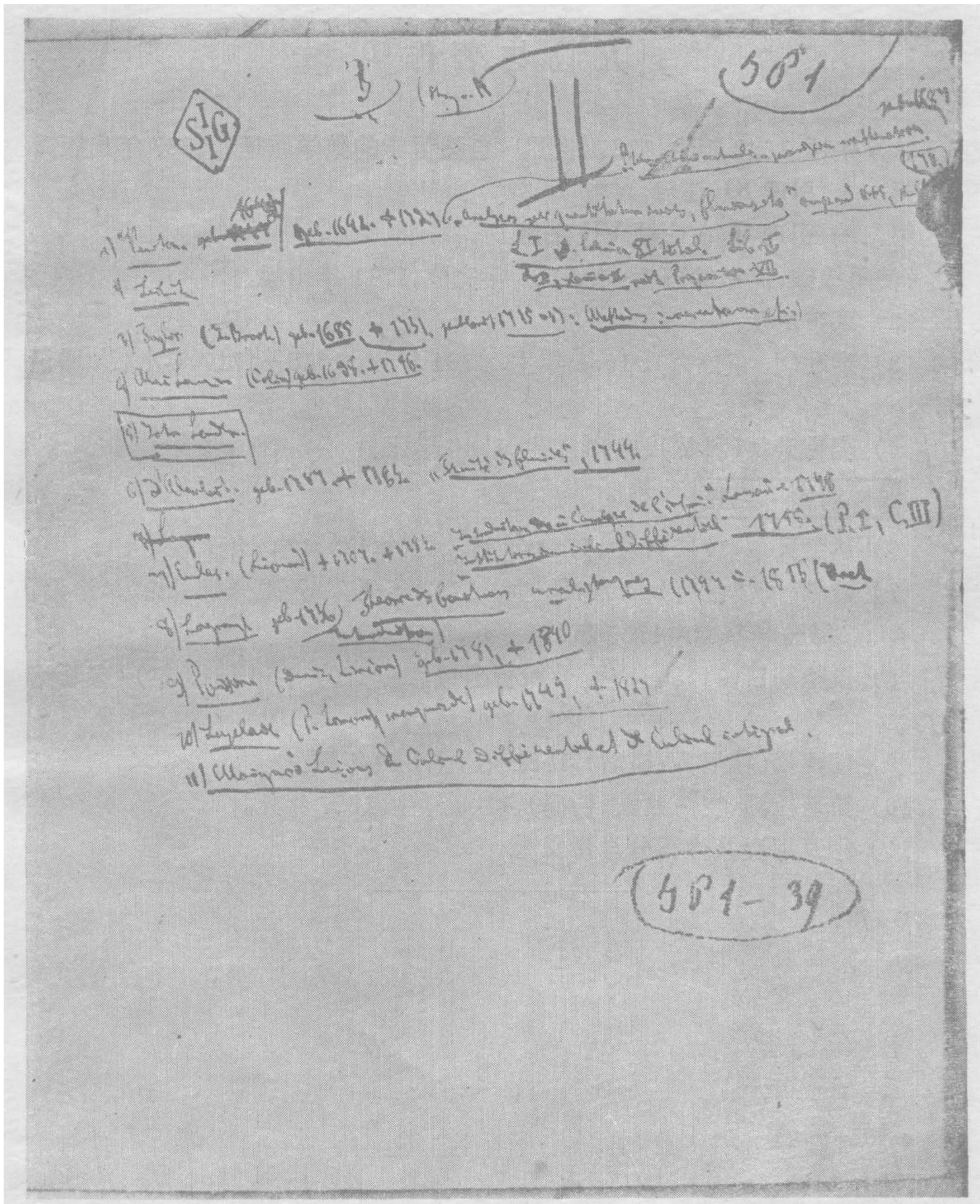
它不过是以前演化得到的有限差值的微分表示式而已。

在普通方法中，

$$dy\text{ 或 }df(x)=f'(x)dx$$

根本不是演化得来，而是如上所见，只是把由二项式$(x+\Delta x)$或$(x+dx)$完全现成地提供的 $f'(x)$从它的因子和相邻各项那里**解脱出来**。

---



马克思《人物与著作》手稿一页

# [人物与著作][①]

1）**牛顿**，1642年生，1727年死。《**自然哲学的数学原理**》，1687年出版。

第一编，**引理 XI，注释**。

第二编，**引理 II，在命题 VII 之后**。

《**使用级数、流数的分析…**》，1665年完成，1711年出版。

2）**莱布尼茨**。

3）**泰勒**（J.布鲁克），1685年生，1731年死，1715—1717年出版《**增量方法…**》。

4）**麦克劳林**（科林），1698年生，1746年死。

**5）约翰·兰登**

6）**达兰贝尔**，1717年生，1783年死。《**流体论**》，1744年。

7）**欧勒**（列奥纳特），1707年生，1783年死。《**无限分析引论**》，洛桑，1748年。《**微分学基础**》，1755年（第一编，第三章）。

8）**拉格朗日**，1736年生。《**解析函数论**》（1797年和1813年）（**参看绪论**）。

9）**普阿松**（德尼，西梅翁），1781年生，1840年死。

10）**拉普拉斯**（皮·西蒙，侯爵），1749年生，1827年死。

11）**穆瓦尼奥**，《**微积分学讲义**》。

---

① 这是附在笔记本《B（续A）II》中的一页。

# [达兰贝尔方法分析]

## [关 于 极 限]

I) $x^3$;

a) $(x+h)^3 = x^3 + 3hx^2 + 3h^2x + h^3$;

b) $(x+h)^3 - x^3 = 3x^2h + 3xh^2 + h^3$;

c) $\dfrac{(x+h)^3 - x^3}{h} = 3x^2 + 3xh + h^2$。

如果 $h$ 变为 $=0$,那末

$\dfrac{(x+0)^3 - x^3}{0}$ 或 $\dfrac{x^3 - x^3}{0} = \dfrac{0}{0}$ 或 $\dfrac{dy}{dx}$,而其右边 $=3x^2$,因此

$$\frac{dy}{dx} = 3x^2。$$

---

$$y = x^3; \qquad y_1 = x_1^3;$$

$$y_1 - y = x_1^3 - x^3 = (x_1 - x)(x_1^2 + x_1x + x^2);$$

$$\frac{y_1 - y}{x_1 - x} \text{ 或 } \frac{dy}{dx} = x^2 + xx + x^2;$$

$$\frac{dy}{dx} = 3x^2。$$

---

II) 如果我们令 $x_1 - x = h$,那末:

1) $(x_1 - x)(x_1^2 + x_1x + x^2) = h(x_1^2 + x_1x + x^2)$;

2) 因此

$$\frac{y_1 - y}{h} = x_1^2 + x_1x + x^2。$$

在 1)中,$h$ 的系数**不是**象上面 $f'$ 那样的**现成导函数**,而是 $f^1$;所以用 $h$ 去除两边,给出的也不是 $\dfrac{dy}{dx}$,而是

$$\frac{\Delta y}{h} \text{ 或 } \frac{\Delta y}{\Delta x} = x_1^2 + x_1x + x^2$$

等等，等等。

如果在Ic)的另一边，即在

$$\frac{f(x+h)-f(x)}{h} \text{ 或 } \frac{y_1-y}{h}=3x^2+3xh+h^2$$

中，我们从这样一种想象出发，以为右边 $h$ 的值越减少，$3xh+h^2$ 这两项的值就会越加减少，因而整个右边的值 $3x^2+3xh+h^2$ 也就越来越接近于值 $3x^2$，那末我们必须补充一句：无论什么时候都不会同它相吻合。

这样，$3x^2$ 就成为级数 $3x^2+3xh+h^2$ 能够不断接近，但永远不能达到，因而更不能超过的一个值。在这个意义下，$3x^2$ 就成为级数 $3x^2+3xh+h^2$ 的**极限值**。

另一方面，量 $\frac{y_1-y}{h}$ $\left(\text{或}\frac{y_1-y}{x_1-x}\right)$ 也将随着分母 $h$ 的减少而越来越减少。但因 $\frac{y_1-y}{h}$ 是 $3x^2+3xh+h^2$ 的等价物，所以这一级数的极限值就是它固有的**极限值**，正象在同样意义下这个极限值就是与它等价的级数的极限值一样。

但是一旦我们令 $h=0$，那末右边的各项就此消失，而使 $3x^2$ 成为右边的值的极限；$3x^2$ 现在是 $x^3$ 的一阶导函数，因而 $=f'(x)$。作为 $f'(x)$ 它表明：从它又可以导出 $f''(x)$（在当前情况下 $=6x$）等等；因而增量 $f'(x)$ 或 $3x^2$ 不等于从展开 $f(x)=x^3$ 而得到的那些增量之和。假如 $f(x)$ 本身是一个无穷级数，那末从它展开而得到的增量的级数自然也是无穷级数。而在这个意义下，一旦我把展开得来的增量的级数截断，就变为它的展开的**极限值**，从而在这里就变为一个通常代数或通常算术意义下的**极限值**，正如一个无限十进位小数的展开部分是它可能的展开的**极限**，是一个在实践或理论中够用了的极限。这和第一种意义下的极限值绝无共同之处。

在这里第二种意义下的极限值**可以**随意使之**增大**，而在那里则只能**减少**。再者，只要 $h$ 在减少，

$$\frac{y_1-y}{h}=\frac{y_1-y}{x_1-x}$$

只会接近于表示式 $\frac{0}{0}$；后者是它永远达不到而更不能超过的极限，在这情况下可以把 $\frac{0}{0}$ 看作是它的**极限值**。

但是一旦 $\frac{y_1-y}{h}$ 变为 $\frac{0}{0}=\frac{dy}{dx}$，后者就不再是 $\frac{y_1-y}{h}$ 的极限值，因为后者本身在它的极限中已经消失。关于它先前的形式 $\frac{y_1-y}{h}$ 或 $\frac{y_1-y}{x_1-x}$，我们只能

说，$\frac{0}{0}$是它的绝对最小表示式，孤立地看，这个绝对最小表示式不是一个数值表示式；但是现在$\frac{0}{0}\left(或\frac{dy}{dx}\right)$有 $3x^2$ 即 $f'(x)$ 在其对面作为它的实在等价物。

这样一来，在等式

$$\frac{0}{0}\left(或\frac{dy}{dx}\right)=f'(x)$$

的两边，没有一边是另一边的极限值。它们彼此并没有**极限关系**，而只有**等价关系**。

如果我有$\frac{6}{3}=2$，那末既非 2 是$\frac{6}{3}$的极限，亦非$\frac{6}{3}$是 2 的极限。一个量的值 = 它的值的极限这种说法，只能导致淡而无味的同义反复。

因此，极限值概念是可能被曲解的，而且经常被曲解。应用到微分等式时，它是为令 $x_1-x$ 或 $h=0$ 作准备的、并使之易于想象的手段，是来源于最初神秘的和故弄玄虚的演算方法的一种幼稚行为。

在微分等式对曲线等等的应用中，极限值概念确实有助于几何上的直观了解。

---

# [达兰贝尔方法和代数方法的比较]

我们把达兰贝尔方法同代数方法作一比较。

I) $f(x)$ 或 $y=x^3$;

a) $f(x+h)$ 或 $y_1=(x+h)^3=x^3+3x^2h+3xh^2+h^3$;

b) $f(x+h)-f(x)$ 或 $y_1-y=3x^2h+3xh^2+h^3$;

c) $\frac{f(x+h)-f(x)}{h}$ 或 $\frac{y_1-y}{h}=3x^2+3xh+h^2$;

当 $h=0$ 时,

d) $\frac{0}{0}$ 或 $\frac{dy}{dx}=3x^2=f'(x)$。

II) $f(x)$ 或 $y=x^3$;

a) $f(x_1)$ 或 $y_1=x_1^3$;

b) $f(x_1)-f(x)$ 或 $y_1-y=x_1^3-x^3=(x_1-x)(x_1^2+x_1x+x^2)$;

c) $\frac{f(x_1)-f(x)}{x_1-x}$ 或 $\frac{y_1-y}{x_1-x}=x_1^2+x_1x+x^2$。

如果 $x_1$ 变为 $=x$,那末 $x_1-x=0$,因此

d) $\frac{0}{0}$ 或 $\frac{dy}{dx}=(x^2+xx+x^2)=3x^2$。

在两种方法中,这样一点是相同的:如果自变量 $x$ 增长,那末因变量 $y$ 也增长。整个问题在于如何表示 $x$ 的增长。如果 $x$ 变为 $x_1$,那末 $x_1-x=\Delta x=h$(这是一个不确定的、可以无限缩小而总是有限的差值)。

$\Delta x$ 或 $h$ 是 $x$ 增长的增量,因为:

a) $x_1=x+\Delta x$,但也可以反过来 b) $x+\Delta x$ 或 $x+h=x_1$。

历史上微分演算是从 a)出发的,也就是说从下面一点出发的:差值 $\Delta x$ 或增量 $h$(两者都表示同一个东西,一个是否定地作为差值 $\Delta x$,另一个是肯定地作为增量 $h$)在量 $x$ 的旁边**独立地存在**着,它就是 $x$ 的增量,表示 $x$ 已**增长**,而且增长了 $h$。由此一开始就得到了好处,那就是一旦这个变量增长,那末相应于这个一般表示式的变量的原函数,就可用一定幂次的二项式来表示,因而二项式定理一开始就可应用于它。事实上在一般的一边即左边,我们已经有一个二项式,即 $x+\Delta x$ 或 $y_1=$ etc.。

神秘的微分演算立刻把 $(x+\Delta x)$ 变为 $(x+dx)$,或者在牛顿那里变为

$x+\dot{x}$。这样一来，我们在右边即代数的一边也立刻得到了用 $x+dx$ 或 $x+\dot{x}$ 表达的二项式，而后它们就被当作普通的二项式来处理。把 $\Delta x$ 变为 $dx$ 或 $\dot{x}$ 的这种做法，数学上不加拒绝而先验地①假定下来，所以后来把二项式展开的一些项神秘地镇压掉就成为可能。

达兰贝尔从 $(x+dx)$ 出发，但是他把这表示式改正为 $(x+\Delta x)$ 或者 $(x+h)$；现在需要有一个使 $\Delta x$ 或 $h$ 变为 $dx$ 的演化，而这也就是实际所发生的全部演化。

不管是错误地从 $(x+dx)$ 出发，还是正确地从 $(x+h)$ 出发，把这个不定二项式代入 $x$ 的给定的代数函数中，它就变为一个具有一定幂次的二项式，犹如在 I a) 中代替 $x^3$ 出现了 $(x+h)^3$；并且实际上就变为这样一个二项式，它在一种情况下以 $dx$ 而在另一种情况下以 $h$ 作为它的后一项，因而在展开中也只是作为因子外在地依附于那些通过二项式导出的函数。

**因此我们在 I a) 中立刻现成地找到了 $x^3$ 的一阶导函数**，它就是 $3x^2$，就是级数第二项中附有 $h$ 的系数。从此以后，$3x^2=f'(x)$ 就不再改变。它本身决不是通过任何一种微分过程推导得来的，而是一开始就由二项式定理所提供的，并且确是因为我们一开始就把增长了的 $x$ 表示为二项式

$$x+\Delta x=x+h,$$

即表示为增长了 $h$ 的 $x$。现在的整个课题，就在于使这个似乎不是处于发育之中而是完全现成的 $f'(x)$，从它的因子 $h$ 及其他各附带项那里摆脱出来。

在 II a) 中则与此相反，增长了的 $x_1$ 进入代数函数在形式上完全象 $x$ 原先进入其中一样；$x^3$ 变为 $x_1^3$。通过相继 2 次而且两次性质完全不同的微分运算，才能最后得到导函数 $f'(x)$。

在等式 I b) 中，差值 $f(x+h)-f(x)$ 或 $y_1-y$ 确实为符号微系数的形成做好了准备；对于实在微系数来说，这个差值除了使它从级数的第二位移到第一位，因而使它得以从 $h$ 那里解放出来而外，没有任何改变。

在 II b) 中，我们两边都得到差值表示式；在代数的一边，对差值作这样的演化，使得 $(x_1-x)$ 作为因子出现在一个用 $x$ 和 $x_1$ 表达的导函数的旁边，而这个导函数是用 $x_1-x$ 去除 $x_1^3-x^3$ 得到的。只有差值 $x_1^3-x^3$ 的存在才使得有可能把它分解成两个因子。由于

$$x_1-x=h,$$

所以 $x_1^3-x^3$ 分解成的两个因子也可以写成 $h(x_1^2+x_1x+x^2)$。这表明了与

① “先验地”一词，原文为 a priori。

Ⅰb)的一个新的差别。$h$ 本身作为**预先导函数的因子**，通过差值 $x_1^3 - x^3$ 演化成两个因子的乘积才被推导出来的；而 $h$ 作为"导函数"的因子，如在Ⅰa)中那样，则是在任何一个差值形成以前就已现成地存在着的。至于从 $x$ 变到 $x_1$ 的不定增长采取与 $x$ 并列的因子 $h$ 这种分离的形式，在Ⅰ)中一开始就被假定了下来，在Ⅱ)中则是(由于 $x_1 - x = h$)通过推导而得到证明的。$h$ 在Ⅰ)中虽然一方面是不确定的，但另一方面，$x$ 的不定增长已显现为一个**固有的**量，$x$ 就增长了**这样**一个量，因而它能和 $x$ 并列地出现，就这一点而论，$h$ 却又是确定的。

在Ⅰc)中，现在要把 $f'(x)$ 从它的因子 $h$ 那里解放出来；这样，我们在左边得到 $\frac{y_1 - y}{h}$ 或 $\frac{f(x+h) - f(x)}{h}$，也就是得到微系数的一个仍然有限的表示式。但另一方面，当我们把 $h = 0$ 代入 $\frac{f(x+h) - f(x)}{h}$ 从而使它变为 $\frac{0}{0} = \frac{dy}{dx}$ 时，我们在Ⅰd)中一边得到符号微系数，另一边得到 $f'(x)$，它在Ⅰa)中已现成地存在，而现在则摆脱了其余各项而单独地出现在右边。

积极的演化只发生在左边，因为符号微系数是在这里产生出来的。在右边，这个演化仅仅在于使Ⅰa)中已通过二项式找到的 $f'(x) = 3x^2$，从它原先的附属物那里解放出来。$h$ 之变为 0 或 $x_1 - x = 0$ 在右边只具有这种消极的意义。

在Ⅱc)中则与此相反，通过 $x_1 - x(= h)$ 除两边，首先得到一个**预先导函数**。

最后在Ⅱd)中，由于肯定地令 $x_1 = x$，就得到**最终导函数**。然而这个 $x_1 = x$ 同时意味着令 $x_1 - x = 0$，因而使左边有限的比值 $\frac{y_1 - y}{x_1 - x}$ 变为 $\frac{0}{0}$ 或 $\frac{dy}{dx}$。

在Ⅰ)中象在神秘的微分法中一样，"导函数"很少是通过令 $x_1 - x = 0$ 或 $h = 0$ 而找到的。在这两种情况下都是为一开始就现成地出现的 $f'(x)$ 扫除掉附带的各项。不过，现在数学上是正确的，而在那里则采取了一次政变。

## [达兰贝尔方法分析的又一例]

现在我们根据达兰贝尔方法来进行演化：

a) $f(u)$ 或 $y=3u^2$；

b) $f(x)$ 或 $u=x^3+ax^2$。

$$y=3u^2, \tag{1}$$

$$f(u)=3u^2。 \tag{1a}$$

$$f(u+h)=3(u+h)^2,$$

$$f(u+h)-f(u)=3(u+h)^2-3u^2=3u^2+6uh+3h^2-3u^2=6uh+3h^2 \tag{2}$$

**(通过二项式定理，导函数在这里已经现成地作为 $h$ 的系数)，**

$$\frac{f(u+h)-f(u)}{h}=6u+3h。$$

通过除法，就使在(2)中已现成地给定的 $f'(u)=6u$ 从它的因子 $h$ 那里解放了出来。

$$\frac{f(u+0)-f(u)}{0}=6u,$$

$$\frac{y_1-y}{u_1-u} \text{ 或者 } \frac{0}{0}=\frac{dy}{du}=6u。$$

将等式 b)中 $u$ 的值代入这个式子，就得到

$$\frac{dy}{du}=6(x^3+ax^2)。$$

由于在 a)中 $y$ 要对 $u$ 微分，所以

$$(u_1-u)=h,\text{或 } h=(u_1-u),$$

因为 $u$ 是自变量。

于是，

$$\frac{dy}{du}=6(x^3+ax^2)。$$

(这是从 $f(u)$ 或 $y=3u^2$ 得到的。)

b) $f(x)$ 或 $u=x^3+ax^2$,

$$f(x+h)=(x+h)^3+a(x+h)^2,$$

$$
\begin{aligned}
f(x+h)-f(x) &= (x+h)^3+a(x+h)^2-x^3-ax^2= \\
&= \left.\begin{array}{l} x^3+3x^2h+3xh^2+h^3 \\ +ax^2+2axh+ah^2 \end{array}\right\{ \begin{array}{l} -x^3+ \\ -ax^2= \end{array} \\
&= (3x^2+2ax)h+(3x+a)h^2+h^3,
\end{aligned}
$$

$$\frac{f(x+h)-f(x)}{h}=3x^2+2ax+(3x+a)h+h^2。$$

如果我们现在令前面的 $h=0$,那末在右边:

$$\frac{0}{0}或\frac{du}{dx}=3x^2+2ax。$$

但是导函数 $3x^2+2ax$ 已经现成地包含在

$$f(x+h)=(x+h)^3+a(x+h)^2$$

中,因为这个式子给出了

$$x^3+3x^2h+3xh^2+h^3+ax^2+2axh+ah^2。$$

于是

$$x^3+ax^2+(3x^2+2ax)h+(3x+a)h^2+h^3。$$

导函数已经显现为 $h$ 的现成的系数。所以这导函数不是通过取差值,而是通过将 $f(x)$增长到 $f(x+h)$,也就是通过将 $x^3+ax^2$ 增长到$(x+h)^3+a(x+h)^2$而得到的。

它简直是这样得到的,即当 $x$ 变为 $x+h$ 时,我们在另一边得到一个具有一定幂次的 $x+h$ 的二项式,其第二项,即带有 $h$ 的一项就完全现成地包含了 $u$ 的导函数或 $f'(u)$。

所有后面的手续,只是为了使一开始就这样给定的 $f'(x)$从它自己的系数 $h$ 和所有其他各项中解放出来。

等式

$$\frac{f(x+h)-f(x)}{h}=\text{etc.}$$

提供了两点:首先,作为 $f(x)$的差值,最初作为 $=\Delta f(x)$,得到左边的分子;但在右边它只具有代数上的好处,那就是从$(x+h)^3+a(x+h)^2$的产物中去掉用 $x$ 表达的原函数 $x^3+ax^2$,等等。

---

我们还是再继续下去。对于 a)我们已获得了

$$\frac{dy}{du}=6(x^3+ax^2),$$

而对于 b)已获得了

$$\frac{du}{dx}=3x^2+2ax。$$

如果我们用$\frac{du}{dx}$去乘$\frac{dy}{du}$,那末

$$\frac{dy}{du}\cdot\frac{du}{dx}=\frac{dy}{dx},$$

这正是所要找的东西。若我们把找到的$\frac{dy}{du}$和$\frac{du}{dx}$的值代入其中,则得

$$\frac{dy}{dx}=6(x^3+ax^2)(3x^2+2ax),$$

因而,一般地说,如果我们有

$$y=f(u);\frac{dy}{du}=\frac{df(u)}{du},\ u=f(x);\frac{du}{dx}=\frac{df(x)}{dx},$$

那末

$$\frac{dy}{du}\cdot\frac{du}{dx}\text{ 或 }\frac{dy}{dx}=\frac{df(u)}{du}\cdot\frac{df(x)}{dx}。$$

如果现在我们在等式 a)中令 $h=u_1-u$, 在等式 b)中令 $h=x_1-x$, 那末事情就会是这样:

$$y\text{ 或 }f(u)=3u^2,$$

$$f(u+(u_1-u))=3(u+(u_1-u))^2=3u^2+6u(u_1-u)+3(u_1-u)^2,$$

$$f(u+(u_1-u))-f(u)=3u^2+6u(u_1-u)+3(u_1-u)(u_1-u)-3u^2,$$

因此

$$f(u+(u_1-u))-f(u)=6u(u_1-u)+3(u_1-u)^2,$$

$$\frac{f(u+(u_1-u))-f(u)}{u_1-u}=6u+3(u_1-u)。$$

因而如果 $u_1-u$ 在第一项中 $=0$,那末

$$\frac{dy}{du}=6u+0=6u。$$

---

这表明,如果一开始 $f(u)$就变为 $f(u+(u_1-u))$,以致它的增量显现为在右边确定的二项式的肯定的第二项,那末通过二项式定理得到的附有$(u_1-u)$或 $h$ 的第二项,立刻就是所要找的系数。如果这第二项有好几项,象

$$x^3+ax^2,\text{它变为}(x+h)^3+a(x+h)^2,$$

或

$$(x+(x_1-x))^3+a(x+(x_1-x))^2,$$

那末只要把附有 $(x_1-x)$ 一次幂，或者说 $h$ 一次幂的那些项相加起来作为 $h$ 或 $x_1-x$ 的系数，我们就现成地得到了所要找的系数。

这结果表明：

1）如果在达兰贝尔的演化中把 $x_1-x=h$ 反过来置为 $h=x_1-x$，而这样做对方法本身绝对没有什么改变，只是更清楚地突出了这个方法，那末通过 $f(x+h)$ 或 $f(x+(x_1-x))$，对于另一边的代数表示式就立即得到二项式而不是得到原函数，例如在当前情况下就不是得到 $3u^2$。

人们这样求得的附有 $h$ 或 $x_1-x$ 的第二项，就是现成的一阶导函数。现在的任务就在于把它从 $h$ 或 $x_1-x$ 那里解放出来，而这是容易做到的。导函数已经现成地在那里；所以它不是通过 $x_1-x=0$ 找到的，而是从它的因子 $(x_1-x)$ 和附加物那里解放出来的。正象它是用简单的乘法（二项式展开）作为附有因子 $x_1-x$ 的第二项找到的那样，最后用 $x_1-x$ 去除两边就使它从这个因子那里解放了出来。

中间过程却在于展开等式

$$f(x+h)-f(x) \text{ 或 } f(x+(x_1-x))-f(x)=$$

这等式在这里除了使起始函数能在右边消失以外，并没有其他目的，因为 $f(x+h)$ 的展开必然包含 $f(x)$ 及其由二项式展开得到的增量。因此起始函数就将从右边被去掉。

所以，例如在

$$(x+h)^3+a(x+h)^2-x^3-ax^2$$

中所发生的，便是把最初两项 $x^3$ 和 $ax^2$ 从二项式 $(x+h)^3+a(x+h)^2$ 中去掉；这样，我们就得到附有 $h$ 或 $x_1-x$ 的、早已现成的导函数作为这等式的第一项。

右边的第一次微分无非是简单地把原函数从其增长了的表示式中减去，所以它给予我们的是原函数所增长的增量，其附有 $h$ 的第一项已经是现成的导函数。其余各项则除 $h^2$ 等等或 $(x_1-x)^2$ 等等的系数外，不可能含有其他东西；通过第一次的用 $x_1-x$ 去除两边，这些项都将降低一个幂次，而在第一项中就没有 $h$ 出现。

2）同 $f(x_1)-f(x)=\text{etc.}$ 这方法的区别在于，例如当

$$f(x) \text{ 或 } u=x^3+ax^2$$

时我们得到

$$f(x_1) \text{ 或 } u_1 = x_1^3 + ax_1^2,$$

变量 $x$ 的最初增长并没有给我们一开始就提供现成造就的 $f'(x)$。

$$f(x_1) - f(x) \text{ 或 } u_1 - u = x_1^3 + ax_1^2 - (x^3 + ax^2)。$$

这里问题并不在于把原函数再去掉，因为 $x_1^3 + ax_1^2$ 没有以任何形式包含 $x^3$ 和 $ax^2$。相反，这最初的差值等式给我们提供了一个演化的机会，就是说把这原先两项的每一项都变为 $x_1$ 和 $x$ 的差值，即

$$= (x_1^3 - x^3) + a(x_1^2 - x^2)。$$

现在很清楚，如果我们把这两项的每一项再分解出因子 $x_1 - x$，那末我们就得到用 $x_1$ 和 $x$ 表达的函数作为 $x_1 - x$ 的系数，即

$$f(x_1) - f(x) \text{ 或 } u_1 - u = (x_1 - x)(x_1^2 + x_1x + x^2) + a(x_1 - x)(x_1 + x)。$$

如果我们用 $x_1 - x$ 去除，因而也去除其左边，那末

$$\frac{f(x_1) - f(x)}{x_1 - x} \text{ 或 } \frac{u_1 - u}{x_1 - x} = (x_1^2 + x_1x + x^2) + a(x_1 + x)。$$

通过这个除法，我们就得到了预先导函数。它的每一部分都含有用 $x_1$ 表达的项。

因此，只有当我们令 $x_1 = x$，因而 $x_1 - x = 0$ 时，才能最后得到所要找的用 $x$ 表达的一阶导函数。因为这时

$$x_1^2 = x^2, \qquad x_1x = x^2,$$

所以

$$(x_1^2 + x_1x + x^2) = 3x^2 \text{ 而 } x_1 + x = x + x = 2x,$$

从而

$$a(2x) = 2xa。$$

另一边的结果是

$$\frac{df(x)}{dx} = \frac{du}{dx} = \frac{0}{0}。$$

所以在这里只有通过令 $x_1 = x$ 因而 $x_1 - x = 0$，才会得到导函数。$x_1 = x$ 提供了用 $x$ 的实际函数表达的、最终的积极的结果。

但是 $x_1 = x$ 也提供了 $x_1 - x = 0$，因而除了这一积极的结果外，同时还为另一边提供了符号化的$\frac{0}{0}$或$\frac{dy}{dx}$。

我们也可以一开始就这样说：我们最终必须得到一个用 $x_1$ 和 $x$ 表达的**导函数**。一旦令 $x_1 = x$，它就能变为用 $x$ 表达的导函数。然而令 $x_1 = x$ 无异

于令 $x_1-x=0$，这个化为零是用公式 $x_1=x$ 肯定地表示的，这个公式是为把用 $x_1$ 和 $x$ 表达的导函数变为用 $x$ 表达的函数所必要的，而其否定形式 $x_1-x=0$ 则必须为我们提供那个符号。

3）对 $x$ 不引进和它并列的一个独立增量如 $x_1-x=\Delta x$ 或 $h$ 这样一种处理方法，即使很可能已经为人们所知道，而且我在博物馆里翻阅了约·兰登的东西以后将会确信这一点，然而它本质上的区别可能还没有被人们所理解。

但是这方法与拉格朗日方法的区别在于，这方法是要实际加以微分的，因而微分表示式也就出现在符号的一边，而在拉格朗日那里，求导不是用代数方法来表示差值，而是用代数方法直接从二项式中导出函数，它们的微分形式只是为了“对称”起见才被采用，因为大家已由微分演算知道，一阶导函数 $=\frac{dy}{dx}$，二阶导函数 $=\frac{d^2y}{dx^2}$，等等。

---

## [关于二阶微分]

$$x^3,\quad 3x^2=f'(x),\quad 6x=f''(x)。$$

I) $3x^2$ 或 $f'(x)=\dfrac{dy}{dx}$。

如果现在从 $f'(x)$ 作为 $\varphi(x)$ 出发，即令 $\varphi(x)=3x^2$，那末用前面的方法我们得到：

II) $6x=\varphi'(x)=\dfrac{d\varphi}{dx}$。这里之所以使用 $\varphi$，只是为了把它与一阶的 $f'(x)$ 区别开来。

$$d(3x^2)=d(f'(x))=d\left(\frac{dy}{dx}\right),$$

$$dy=f'(x)dx,\qquad d(dy)\text{或 }d^2y=d(f'(x)dx)。$$

如果 $dx$ 是常量，那末 $=d(f'(x))dx$。

$$\therefore\quad \frac{d^2y}{dx}=d(f'(x))。$$

但 $d(f'(x))=f''(x)dx$，$\therefore\quad \dfrac{d^2y}{dx}=f''(x)dx$ 或 $\dfrac{d^2y}{dx^2}=f''(x)=6x$。

a) $3x^2=f'(x)=\dfrac{dy}{dx}$。所以 $dy=f'(x)dx$。因此：

$$d^2y=d(f'(x)dx)=f''(x)dx^2。$$

b) $6x=f''(x)$。

假定 $dx$ 为常量，则：

$$d^2y=d(f'(x))dx,$$

$$\therefore\quad \frac{d^2y}{dx^2}=\frac{d(f'(x))}{dx}，\text{但 } d(f'(x))=f''(x)dx。$$

因此

$$\frac{d^2y}{dx^2}=f''(x)。$$

在这方法里，公式 $\dfrac{d^2y}{dx^2}=f''(x)$ 仅仅是由于假设

$$d(f'(x)dx)=d(f'(x))dx,$$

也就是假定 $dx$ 为常量才得来的。只有当 $dy=f'(x)dx$ 这结果中(这一次 $dx$ 不作为假定 $=d\ (f(x))$ 的 $dx$，而作为量 $x$ 的微分元)的 $dx$ 为常量这个假定不是任意的时候，这个证明才算是一个证明。如果在一阶微分那里它作为常量出现,那末在二阶微分那里它同样也是个常量。

# [拉格朗日方法分析]

## 逐次微分的预备知识

我们知道，如果我们有 $y=f(x)$，那末 $df(x)=f'(x)dx$。因此：

1）$\dfrac{df(x)}{dx}=f'(x)$；如果对 $f'(x)$ 本身进行微分，那末 $df'(x)=f''(x)dx$；因此：

2）$\dfrac{df'(x)}{dx}=f''(x)$；进而：$df''(x)=f'''(x)dx$；因此：

3）$\dfrac{df''(x)}{dx}=f'''(x)$；进而：$df'''(x)=f^{\mathrm{IV}}(x)dx$；因此：

4）$\dfrac{df'''(x)}{dx}=f^{\mathrm{IV}}(x)$ 等等。

A. 这些"导"函数本身可加以处理而不与原函数发生任何联系，另一方面则又可以把它们表示为前一个原函数的"导函数"。在这情况下这些函数显现为：

1）$y$ 或 $f(x)$ 及其导函数 $f'(x)$，因此象以前一样：$\dfrac{df(x)}{dx}=f'(x)$；这个函数独立看待时 $=\varphi(x)$；

2）$y$ 或 $\varphi(x)$ 和它的导函数：$d\varphi(x)=\varphi'(x)dx$ 以及 $\dfrac{d\varphi(x)}{dx}=\varphi'(x)$；这个函数独立看待时又当作 $F(x)$；

3）$y$ 或 $F(x)$ 及导函数：$dF(x)=F'(x)dx$ 以及 $\dfrac{dF(x)}{dx}=F'(x)$。

因此我们有：

1）$f'(x)=\dfrac{df(x)}{dx}$；

2）$f''(x)=\dfrac{df'(x)}{dx}$；若把 $f'(x)$ 的值代入其中，则将变为

$$\frac{d\left(\dfrac{df(x)}{dx}\right)}{dx}=\frac{d(df(x))}{dx^2},$$

因此

$$f''(x)=\frac{d(df(x))}{dx^2},$$

这是一个其分子的意义我们到此还未予以确定的表示式。

---

我们已经看到，这些不同的公式是怎样作为**导函数的符号表示式**，也就是作为**已实行的运算的符号**而得到的，并且从前面的演化中现在已不言而喻，它们将反过来变为**符号化的运算公式**，变为表明有待去**实行的运算**的，也就是去找与它们相应的实在等价物或导函数的一些公式。但是对这些公式本身，我们还必须进一步加以考察。

1）首先我们来考察 $f'(x)$，$f''(x)$，$f'''(x)$，$f^{\mathrm{IV}}(x)$ 这些符号微系数的分子，即考察 $df(x)$，$d(df(x))$，$d(d(df(x)))$，$d(d(d(df(x))))$ 等等。

这些表示式之所以会产生，只是因为不同的函数在它们的推导链列中是从原函数中和从前一个函数中得来的，因此它们显现为**逐阶导函数**，其中每一个总是从其前面的一个产生出来的。但是在另外的等式中它们本身也作为原函数出现，因而彼此之间以及与原先的原函数之间并没有任何联系，而这个原先的原函数又可以带有另一个原函数的胎记，表明这另一个原函数的导函数原来就是它。

这些函数就是这样没有联系地出现的：例如，$y$ 或 $f(x)=5x^4$（这个原函数本身一眼就可以看出是 $x^5$ **的一阶导函数**）。因此

$$f'(x)=\frac{df(x)}{dx}\text{或}\frac{dy}{dx}=5\cdot4\cdot x^3。$$

这个导函数独立出现时就称它 $=\varphi(x)$。

可是事先还应指出这样一点：当 $\frac{y_1-y}{x_1-x}$ 通过令 $x_1=x$ 即 $x_1-x=0$ 而化为 $\frac{dy}{dx}$ 之后，

$$f'(x)=\frac{df(x)}{dx}\text{ 或 }\frac{dy}{dx}=\frac{y_1-y}{x_1-x}。$$

为了把这一点固定下来，我们要在 $\frac{y_1-y}{x_1-x}$ 一旦经历了这种变形以后把它写成 $\left(\frac{y_1-y}{x_1-x}\right)$，所以这里的括号表明 $\frac{y_1-y}{x_1-x}$ 已变成了 $\frac{dy}{dx}$。

这样，我们就得到

$$f'(x)=\frac{df(x)}{dx}\text{或}\frac{dy}{dx}=\left(\frac{y_1-y}{x_1-x}\right)。$$

而且在这里还必须对**函数**一词做一个说明。

---

# 关于“函数”一词

“函数”一词，原先是在处理方程个数少于其中所含的未知量个数的所谓不定方程时引进代数中来的。在这里，例如 $y$ 的值要随着人们对例如 $x$ 所假定的数值 3,4,5 等等而变化。$y$ 在这里称为 $x$ 的函数，因为它必须服从 $x$ 的命令，正如每一个官员，甚至伟大的威廉一世也得依从某人一样。

在微分演算中，我们相应地在这意义下把“函数”一词转移给了因变量，例如 $y$。

因此 $y$ 或 $f(x)$，即我们例子中的 $y$ 或 $5x^4$，表明它是 $x$ **的**函数，而且确是用确定的表示式 $5x^4$ 表示的 $x$ 的函数，因为 $y$ 的值是随着 $x$ 在其固有的表示式 $5x^4$ 中变化所产生的值而变化的。

但是，一当拉格朗日引进了“导函数”的定义，并因此引进了它们所据以导出的原函数的定义的时候，就出现了一种至今还一直存在着的混乱。拉格朗日意义下的函数在处理现代的一切演算中已普遍使用，但“函数”一词的传统含义在那里还同时存在着：所以例如

当 $y=5x^4$ 时，我们就有：

$y$ 或 $f(x)$，或者更详细地说，$y$ **或** $f(x)$**或**函数 $5x^4=f(x)$ **或** $5x^4$。

这种混乱只有这样来消除，如果我们读成：

$y$，这 $x$ **的**函数，也就是在每一特殊情况下这个依赖于 $x$ 的东西，或**用确定的表示式** $5x^4$ **表达的** $x$ **的函数**，是 = **用** $x$ **表达的**原函数：$5x^4$；对于导函数也同样如此，$y$ 总是 $x$ **的**函数，而导函数则相反是**用** $x$ **表达的**函数。后一种意义下的“函数”一词表明：对于**原函数**来说它是 $x$ 原先在其中出现的代数组合，例如 $5x^4$，而对于**导函数**来说它是由于 $x$ 的变化和与之相应的微分而取代 $5x^4$ 的新值。

在拉格朗日那里却相反，表示式 $f(x)$ 一旦处于用 $x$ 表达的代数表示式的左边，那末相对于**特殊的**表示式而言它只具有**一般的**、因而是**不定的表示式**的含义；而 $f(x+h)$ 相对于它展开了的展开级数表示式而言具有一般的、未展开的表示式的意义，正象例如在代数中 $(x+a)^m$ 是一般的、未展开的表示式，而在展开级数的右边则是 $x^m+\text{etc.}$。

对于某些目的来说，这是完全足够的并且也是适当的；虽然如此，但 $x$ **的**函数和**用** $x$ **表达的**函数之间的区别依然不可缺少，因为只有这种区别才说明

$x$ **的**函数能够具有一种具体的、与**用** $x$ **表达的**函数不同的存在，正象例如当 $x$ 为横坐标时有纵坐标存在一样，等等。

---

如果我们让不同的**导函数**本身立刻又作为原函数出现，也就是说如果我们不是把它们看作相互有联系的，那末：

若我们有

1) $y=f(x)=x^5$，则 $y_1=f(x_1)=x_1^5$。

如果一旦由于 $x_1-x=0$ 而 $y_1-y$ 变为 $dy$，我们就用 $(y_1-y)$ 来标记它，并且当 $x_1-x=0$ 时，同样用 $(x_1-x)$ 来标记 $x_1-x$，因此也用 $\frac{(f(x_1)-f(x))}{(x_1-x)}$ 来标记 $\frac{f(x_1)-f(x)}{x_1-x}$，那末

$$\frac{(y_1-y)}{(x_1-x)}\text{或}\frac{(f(x_1)-f(x))}{(x_1-x)\text{或}h}=f'(x)=5x^4,$$

以及

$$\begin{aligned}&dy\ \text{或}(y_1-y)\text{或}(f(x_1)-f(x))=\\&\quad=f'(x)(x_1-x)\text{或}f'(x)dx=5x^4(x_1-x)\text{或}5x^4dx。\end{aligned}$$

而在另一方法中：

$$\left(\frac{y_1-y}{x_1-x}\right)\text{或}\frac{(f(x+h)-f(x))}{(x_1-x)}=f'(x)=5x^4,$$

并同样得到

$$\begin{aligned}&dy\ \text{或}(y_1-y)\text{或}(f(x+h)-f(x))=\\&\quad=f'(x)(x_1-x)=5x^4(x_1-x)=5x^4dx。\end{aligned}$$

---

我们在这里看到，

1) $dy$ 或 $(y_1-y)=(f(x_1)-f(x))$，

2) $dy$ 或 $(y_1-y)=(\varphi(x_1)-\varphi(x))$，

3) $dy$ 或 $(y_1-y)=(F(x_1)-F(x))$

等等，

或者用另一种表示方式：

1) $dy$ 或 $(y_1-y)=(f(x+h)-f(x))$，

2) $dy$ 或 $(y_1-y)=(\varphi(x+h)-\varphi(x))$，

3) $dy$ 或 $(y_1-y)=(F(x+h)-F(x))$

等等，

因而 $dy$ 或 $(y_1-y)$ 表示 $x$ **的不同函数**的扬弃了的差值，而且确实在 1)，2)，3) 等等中每一次都表示**对相应的用 $x$ 表达的原函数通过微分**而获得的**不同函数**

$$f(x)=x^5,\quad \varphi(x)=5x^4,\quad F(x)=5\cdot 4x^3$$

的扬弃了的差值。

因此在 1)，2)，3) 中 $dy$ 或 $(y_1-y)$，正象它们也可以用**微分**来表示那样有三个完全不同的值：

1) $dy$ 或 $(y_1-y)=f'(x)(x_1-x)$，

2) $dy$ 或 $(y_1-y)=\varphi'(x)(x_1-x)$，

3) $dy$ 或 $(y_1-y)=F'(x)(x_1-x)$。

但在**符号微系数**的表示式

$$1)\ \frac{(f(x_1)-f(x))}{(x_1-x)},\quad 2)\ \frac{(\varphi(x_1)-\varphi(x))}{(x_1-x)},\quad 3)\ \frac{(F(x_1)-F(x))}{(x_1-x)},\cdots$$

中，$x_1-x$ 始终保持**相同**，因此在**微分**中也一样，在那里它在所有三个等式中都显现为同一个因子。然而这是不言而喻的。如果我们用那些以举例方式给出的具体表示式来取代函数 $f'(x)$，$\varphi'(x)$，$F'(x)$，那末：

$$1)\ f'(x)=5x^4,\quad 2)\ \varphi'(x)=5\cdot 4x^3,\quad 3)\ F'(x)=5\cdot 4\cdot 3x^2。$$

尽管所有这三个都是用 $x$ 表达的不同**函数**，但有一点它们是共同的，即它们都是**用同一个变量 $x$ 表达的函数**。它们都是通过**微分**，也就是通过令 $x_1=x$ 亦即令 $x_1-x=0$，$(x_1-x)=dx$ 而变成的。

---

如果我们现在从运算方法出发，在这种方法中把实际符号化的、从**微系数产生的微分**用作**出发点**，以便反过来求微系数，那末例如

$$1)\ dy \text{ 或 } (y_1-y)=f'(x)(x_1-x);$$

所以

$$\frac{dy}{dx}\text{或}\frac{(y_1-y)}{(x_1-x)}=f'(x)。$$

同样，

$$2)\ dy \text{ 或 } (y_1-y)=\varphi'(x)(x_1-x);$$

因此

$$\frac{(y_1-y)}{(x_1-x)}=\varphi'(x),\cdots$$

在微分中，$(x_1-x)$ 或 $dx$ 显现为用 $x$ 表达的一阶导函数的因子，即分别

显现为 $f'(x)$，$\varphi'(x)$，$F'(x)$ 等等的因子。因此只有当通过除法使后者从它们的因子那里解放出来时，换句话说，用 $(x_1-x)$ 或 $dx$ 去除 $(y_1-y)=dy$ 时，才能得到它们。

---

如果我们现在把这个和那些（11 页上①）通过**置换而得来的表示式**相比较，那末当我们将那里和这里 $f'(x)$，$f''(x)$，$f'''(x)$的，或者在我们情况下 $5x^4$，$5\cdot4x^3$，$5\cdot4\cdot3x^2$ 的不同表示式等同起来时，就有：

1) $\frac{df(x)}{dx}=f'(x)$（或者 $5x^4$）$=\frac{dy}{dx}$，

2) $\frac{d(df(x))}{dx^2}=f''(x)$（或者 $5\cdot4x^3$）$=\frac{dy'}{dx}$，

3) $\frac{d(d(df(x)))}{dx^3}=f'''(x)$（或者 $5\cdot4\cdot3x^2$）$=\frac{dy''}{dx}$。

因此，左边的表示式无非是说，在 1) 中原先的函数**第一次被微分**，在 2) 中它第二次被微分，在 3) 中它第三次被微分；而且由于我们用 $df(x)$ 表示这个一阶微分，所以我们可以用 $d^2f(x)$ 来表示二阶的 $d(df(x))$，并用 $d^3f(x)$ 来表示 $d(d(df(x)))$，其中 $d$，$d^2$，$d^3$ 无非表示：$f(x)$ 第一次被微分，以及如此得到的结果再次被微分，等等。

因此，我们得到：

1) $\frac{df(x)}{dx}=\frac{dy}{dx}$，　2) $\frac{d^2f(x)}{dx^2}=\frac{dy'}{dx}$，　3) $\frac{d^3f(x)}{dx^3}=\frac{dy''}{dx}$，…

由于 $f(x)=y$，所以我们在左边可以到处把 $f(x)$ 写成 $y$，于是我们得到：

1) $\frac{dy}{dx}=f'(x)=\frac{dy}{dx}$，2) $\frac{d^2y}{dx^2}=f''(x)=\frac{dy'}{dx}$，3) $\frac{d^3y}{dx^3}=f'''(x)=\frac{dy''}{dx}$。

这两种表示式之间的差别仅仅是表面的。

在左边，所有的导函数都表示成从原函数 $f(x)$ 导来的；

在右边，不仅表示成从原函数导来的，而且每一个导函数也都表示成从它前面的一个导来的。

1) 然而 $dy=(y_1-y)=(f(x_1)-f(x))=df(x)$，

2) 同样，$dy'=(y'_1-y')=(f'(x_1)-f'(x))$。

因此，如果在 1) 中表示**原函数** $f(x)$ **和** $f(x_1)$ **之间扬弃了的差值**，那末在 2) 中就表示一阶导函数 $f'(x)$ 和由于 $x$ 增长到 $x_1$ 而改变了的函数 $f'(x_1)$ 之间的

---

① 指手稿中的页码，相当于本文第 90 页。

扬弃了的差值。

但是

$$d^2f(x)=d(d(f(x)))$$

并不表示什么别的东西；因为 $df(x)$ 是原函数 $f(x)$ 的微分，等于 $f'(x)dx$，因而 $d(df(x))$是一阶微分 $f'(x)dx$ 的微分。

$d(f(x))$是原函数 $f(x)$的一阶微分。

$d(df(x))$ 或 $d^2f(x)$是对原函数的二阶微分，但也是对一阶微分 $f'(x)dx$ 的一阶微分。

**对于 $d(d(df(x)))$等等同样如此。**

$d(y'dx)$是对 $y$ 的 $d^2$，而且同样地 $d(y''dx^2)$是对 $dy$ 的 $d^2$，因此是对 $y$ 的 $d^3$。

这样，对于计算最有用的形式就此已经得到：若 $f(x)=y$，则

$$f'(x)=\frac{dy}{dx},\quad f''(x)=\frac{d^2y}{dx^2},\qquad f'''(x)=\frac{d^3y}{dx^3},\cdots$$

作为还从牛顿-莱布尼茨方法的年代以来就已存在着的混乱的表示式，在现代数学家那里，他们不仅……

---

# 泰勒定理

让我们来看

$$1)\ (x+h)^m = x^m + mx^{m-1}\frac{h}{1} + m(m-1)x^{m-2}\frac{h^2}{1\cdot 2} + \\ + m(m-1)(m-2)x^{m-3}\frac{h^3}{1\cdot 2\cdot 3} + \\ + m(m-1)(m-2)(m-3)x^{m-4}\frac{h^4}{1\cdot 2\cdot 3\cdot 4} + \cdots$$

这是**二项式定理**。

如果我们现在假定$(x+h)^m$中的$x$是变量而$h$是它的增量，那末当$h=0$时，$(x+h)^m=(x+0)^m=x^m$。

因此在$x$增长之前，用$x$表达的原函数是$x^m$或

a) $x^m = f(x) = y$,

b) $(x+h)^m = f(x+h) = y_1$。

因此前面的等式就变为

$$2)\ f(x+h) \text{ 或 } y_1 = x^m + mx^{m-1}\frac{h}{1} + m(m-1)x^{m-2}\frac{h^2}{1\cdot 2} + \\ + m(m-1)(m-2)x^{m-3}\frac{h^3}{1\cdot 2\cdot 3} + \\ + m(m-1)(m-2)(m-3)x^{m-4}\frac{h^4}{1\cdot 2\cdot 3\cdot 4} + \cdots$$

如果我们在右边令$h=0$，那末它在左边也将等于0，于是我们重新得到$y=x^m$或$=f(x)$(参见a))。因此$y_1$或$f(x+h)$的展开级数的首项必须$=f(x)=y$。

因此等式2)就变为

$$3)\ f(x+h) \text{ 或 } y_1 = y + mx^{m-1}\frac{h}{1} + m(m-1)x^{m-2}\frac{h^2}{1\cdot 2} + \\ + m(m-1)(m-2)x^{m-3}\frac{h^3}{1\cdot 2\cdot 3} + \\ + m(m-1)(m-2)(m-3)x^{m-4}\frac{h^4}{1\cdot 2\cdot 3\cdot 4} + \cdots$$

就$h$的系数而言，我们知道，

$x^m$的导函数 $= mx^{m-1}$,

$mx^{m-1}$ 的导函数 $\quad = m(m-1)x^{m-2}$,

$m(m-1)x^{m-2}$ 的导函数 $\quad = m(m-1)(m-2)x^{m-3}$,

$m(m-1)(m-2)x^{m-3}$ 的导函数 $\quad = m(m-1)(m-2)(m-3)x^{m-4}$

等等

所以等式 3)变为

4) $f(x+h)$ 或 $y_1 = y$ 或 $f(x) + f'(x)\frac{h}{1} + f''(x)\frac{h^2}{1\cdot 2} + f'''(x)\frac{h^3}{1\cdot 2\cdot 3} + + f^{\mathrm{IV}}(x)\frac{h^4}{1\cdot 2\cdot 3\cdot 4} + \cdots$

然而由于

$$f'(x) = \frac{dy}{dx},\quad f''(x) = \frac{d^2y}{dx^2},\quad f'''(x) = \frac{d^3y}{dx^3},\quad f^{\mathrm{IV}}(x) = \frac{d^4y}{dx^4},\cdots$$

并且由于我们可以把各导函数用它们的符号等价物来代替,所以:

5) $f(x+h)$或 $y_1 = f(x)$或 $y + \frac{dy}{dx}\cdot\frac{h}{1} + \frac{d^2y}{dx^2}\cdot\frac{h^2}{1\cdot 2} + \frac{d^3y}{dx^3}\cdot\frac{h^3}{1\cdot 2\cdot 3} + + \frac{d^4y}{dx^4}\cdot\frac{h^4}{1\cdot 2\cdot 3\cdot 4} + \cdots$,

这就是**泰勒定理**,也就是在增长一个正或负的增量 $h$ 时对**每个** $f(x)$进行微分的一般运算公式。这时我们只须把 $y$ 表示为 $x$ 的给定函数,并在我们为它算出了它的$\frac{dy}{dx}$,$\frac{d^2y}{dx^2}$等等以后,就把这样找到的$\frac{dy}{dx}$,$\frac{d^2y}{dx^2}$等等的值代入上面的公式中去,其中它们的数值因子这时还要通过与 $h$,$\frac{h^2}{1\cdot 2}$,$\frac{h^3}{1\cdot 2\cdot 3}$等等相乘而加以修改。

从我们所应用的代数方法的观点来看,只要人们按照表征这个做法的方式行事,那末这个定理是不能应用的,尽管如上所述,在由这个做法提供的事实基础上它是可以直接从二项式定理推导出来的。所以从这个做法的观点来看,对于微分演算的一切运算等式中这个最一般、最包罗万象的运算等式只须指出:

a) 为了使它**根本可以应用**,它要求(参见等式 4)用 $x$ 表达的原函数不仅可以展开成确定的,从而是有限的 $x$ 的函数的级数,而且此外还可以展开成附有正整数升幂的 $h$ 因子的这样一些函数的级数。我们以后将回到这一点上来。

b) 从等式 5)得出,

$$f(x+h) - f(x) \text{ 或 } y_1 - y = \frac{dy}{dx}h + \frac{d^2y}{dx^2}\cdot\frac{h^2}{1\cdot 2} + \frac{d^3y}{dx^3}\cdot\frac{h^3}{1\cdot 2\cdot 3} + + \frac{d^4y}{dx^4}\cdot\frac{h^4}{1\cdot 2\cdot 3\cdot 4} + \cdots$$

然而 $y_1-y$ 是一个有限差值，$=\Delta y$，因为它只不过是处于原始状态中的 $f(x)$ 和处于增长了的状态中的同一个 $f(x)$ 之间的差值。这个差值没有化为 $dy$。因此得出，有限差值

$$y_1-y \text{ 或 } f(x+h)-f(x)$$

可以用**微系数**乘以 $h$ 的(正整数)升幂而成的**和式**来表示。

这个和式是因变量 $y$ 当 $x$ 增长 $h$ 时所得到的增量。

然而从等式 4)得出，这无非是说 $f(x+h)-f(x)$(或 $y_1-y$)被表示为附有增量 $h$ 的上升幂次的从 $f(x)$导来的各函数等等之和。

但是现在我们知道，

$$\left.\begin{array}{l}
f'(x)=\dfrac{dy}{dx}=\dfrac{y_1-y}{x_1-x}\text{，其中已令 } x_1=x, x_1-x=0\text{，因而也已令 } y_1-y=0\text{；}\\[2ex]
\text{在令 } x_1-x=0 \text{ 之后，} f''(x)=\dfrac{(y')-y'}{(x_1-x)}=\dfrac{d^2y}{dx^2}\text{；}\\[2ex]
\text{在令 } x_1-x=0 \text{ 之后，} f'''(x)=\dfrac{(y'')-y''}{(x_1-x)}=\dfrac{d^3y}{dx^3}\text{，}\\[1ex]
\text{等等。}
\end{array}\right\}\text{(A)}$$

如果我们把这些不同的导函数看作是从 $f(x)$ 这个原函数逐次地导来的，情况看来就是这样。

但是，如果我们把这些导函数中的每一个只与同**它自己直接相关的原函数**联系起来看，也就是说，只与它本身所由直接产生的那个 $f(x)$ 联系起来看，那末我们只得到一系列相互没有任何联系的 $f(x)$和 $f'(x)$，因而对这些微系数中的每一个也只得到微分表示式$\dfrac{dy}{dx}$。

首先，例如：

$$\left.\begin{array}{l}
1)\ f(x)=x^3,\\[1ex]
\qquad f'(x)\text{ 或 } 3x^2=\dfrac{y_1-y}{x_1-x}\text{，其中} x_1-x=0\text{，从而}=\dfrac{dy}{dx}\text{；}\\[2ex]
2)\ f(x)=3x^2,\\[1ex]
\qquad f'(x)=6x=\dfrac{y_1-y}{x_1-x}\text{，其中 } x_1-x=0\text{，从而}=\dfrac{dy}{dx}\text{；}\\[2ex]
3)\ f(x)=6x,\\[1ex]
\qquad f'(x)=6=\dfrac{y_1-y}{x_1-x}\text{，其中 } x_1-x=0\text{，从而}=\dfrac{dy}{dx}\text{。}
\end{array}\right\}\text{(B)}$$

如果我们固守**于处理方式**(B)，那末我们每次也只得到

$$f(x+h) \text{ 或 } y_1 \approx f(x)+f'(x)h。$$

因此 $y_1-y \approx f'(x)h=\frac{dy}{dx}h$。只有这个 $f'(x)$，因而还有它的符号等价物 $\frac{dy}{dx}$，从而 $f'(x)h=\frac{dy}{dx}h$ 每次都在 $(B)_{1)}$，$(B)_{2)}$ 和 $(B)_{3)}$ 的这一个或那一个中。那个 $\approx f'(x)h=\frac{dy}{dx}h$ 的 $y_1-y$，在所有这三个等式中都保持**一样的形式**，但具有完全**不同的值**，这些值在这里彼此很少有联系，正象例如原函数 $x^3$ 和它所由导出的任何一个前面的原函数很少有联系一样；例如，假定我们已经得到了原函数 $x^4$，那末

$$(x+h)^4=x^4+4x^3h+\cdots$$

这里的一阶导函数 $f'(x)$ 是 $4x^3$，并且我们得到

$$4(x+h)^3=4x^3+4\cdot 3x^2h+4\cdot 3xh^2+4h^3。$$

如果我们用 4 去除两边，那末

$$(x+h)^3=x^3+3x^2h+3xh^2+h^3,$$

其中 $x^3$——这个我们在 b) 中所由出发的**原函数**——在这里本身就表现为**导函数**。象对 $(B)_{1)}$ 中的 $x^3$ 来说这一点很少妨碍我们那样，对 $(B)_{2)}$ 中的 $3x^2$，或 $(B)_{3)}$ 中的 $6x$，它也很少妨碍我们。“导函数”只是相对于作为原函数出发的那种函数而被导出的。

因此，在方法(B)中不把 $y_1-y$ 表示为导函数之和，在这里是不中用的。

另一方面，如果我们转到(A)，导函数在这里被表示为**逐次**从原函数导出的一些函数，因而被表示为一个导函数链列，那末

$$f(x+h)-f(x) \text{ 或 } y_1-y=\frac{dy}{dx}h+\frac{d^2y}{dx^2}\cdot\frac{h^2}{1\cdot 2}+\cdots$$

同样表示 $y_1-y$ 的一个链列。

就 $y_1-y$ 而言，公式

$$y_1-y \text{ 或 } f(x+h)-f(x)=\cdots$$

在这里只是说，通过对 $f(x)$ 的，因而也是对 $f(x+h)$ 或 $y_1$ 的逐次微分，我们就得到了那个级数；因此通过微分

$$\text{对 } y_1-y \text{ 得到 } f'(x)h \text{ 或 } \frac{dy}{dx}h+\cdots,$$

$$\text{对 } (y')_1-y' \text{ 得到 } f''(x)h \text{ 或 } \frac{d^2y}{dx^2}h+\cdots,$$

$$对\ (y'')_1 - y''\ 得到\ f'''(x)h\ 或\ \frac{d^3y}{dx^3}h + \cdots$$

因此，我们把 $x$ 的原函数看作较为潜在地包含着它的全部“导函数”，看作它本身是独立存在着的，因而在 $f(x+h)-f(x)$ 或 $y_1-y$ 之中也不但较为潜在地包含着 $y_1-y$，而且还包含着 $(y')_1-y'$，$(y'')_1-y''$ 等等。

在这里我们把用 $x$ 表达的原函数，例如 $x^m=y$，看作较为潜在地在它自身内包含着一切可以由它导出的函数，因而也把 $f(x+h)-f(x)$ 或 $y_1-y$ 看作可以用所有这些导函数来表示，而这些导函数以后将为与它们等价的微分符号，也就是将为与它们相对应的符号微系数所替代。

## [泰勒定理增补一]

我们把下面的

$$\text{a)}\ (x+h)^{m+1}\ \text{或}\ y=x^{m+1}+(m+1)x^{m}\frac{h}{1}+$$
$$+(m+1)mx^{m-1}\frac{h^2}{1\cdot2}+(m+1)m(m-1)x^{m-2}\frac{h^3}{1\cdot2\cdot3}+$$
$$+(m+1)m(m-1)(m-2)x^{m-3}\frac{h^4}{1\cdot2\cdot3\cdot4}+\cdots$$

假定为已知，其中$(x+h)^1$是一个通常的二项式，因而不把$x$当做变量看待。

我们的方法能对这个等式进行微分，也就是把$x$假定为变量，把$h$看作常量而不是看作$x$的增量，因为对我们来说，$x_1-x$只存在于这个差值形式之中，而不是作为$x_1-x=h$，因而也不是作为和式$x_1=x+h$。

---

但是这个二项式$(x+h)^m$（在它未展开的形式中）以及在它的展开级数中我们都有$(m+1)$倍①。这个因子$(m+1)$是一根脐带，表明它是从$(x+h)^{m+1}$导来的……，如果我们从$(x+h)^{m+2}$导出了它，那末作为脐带将会得到因子$(m+2)$。

因此，如果我们在A)中划掉这个因子$(m+1)$，那末二项式$(x+h)^m$就象它的出发等式$(x+h)^{m+1}$一样也会独立出现，因而我们得到：

$$\text{B)}\ (x+h)^m=x^m+mx^{m-1}h+m(m-1)x^{m-2}\frac{h^2}{1\cdot2}+$$
$$+m(m-1)(m-2)x^{m-3}\frac{h^3}{1\cdot2\cdot3}+$$
$$+m(m-1)(m-2)(m-3)x^{m-4}\frac{h^4}{1\cdot2\cdot3\cdot4}+\cdots$$

---

这个$m$次的等式比它所由导出的那个$(m+1)$次的等式要低一次。虽然如此，我们仍能把$y$或$f(x)$变为$y_1$或$f(x_1)$而不改变等式的代数现状一根毫毛。为此我们只须

① 在这句话的前面，手稿中写有等式

$$\text{A)}\ (m+1)\left(x^m+mx^{m-1}\frac{h}{1}+m(m-1)x^{m-2}\frac{h^2}{1\cdot2}+\cdots+\right)=(m+1)(x+h)^m。$$

1) 令 $x^m = f(x)$ 或 $= y$，而且我们在这里更加有理由这样做，因为我们一旦从 $x^{m+1}$ 导出了 $x^m$ 之后，所有随后的函数就可以先直接从 $x^m$，然后逐次从一个又一个中推导出来，因而所有随后的函数都可作为从 $x^m$ **逐次**导出的函数来表示；

2) 把 $h$ 这个在我们等式的推导中曾经象代数学里 $(x+a)^m$ 中的 $a$ 一样的普通常量，看作 $x$ 的（正或负的）**增量**。我们这样做也是有理由的，因为 $x_1 - x = \Delta x$ 以及这个 $\Delta x$ 本身，不象在我们方法中那样用来作为 $x$ 的差值即 $x_1 - x$ 的纯粹符号或纯粹记号，而也可以当作差值 $x_1 - x$ 的量值来对待，这个量值象 $x_1 - x$ 一样是不确定的，并且将随着它的变化而变化。因此 $x_1 - x = \Delta x$ 或 $=$ 不确定的量 $h$。由此得出 $x_1 = x + h$，而 $f(x_1)$ 或 $y_1$ 变为 $f(x+h)$。

因此我们得到：$f(x) = x^m$，

$$\text{A)}\quad f(x+h) \text{或} y_1 = (x+h)^m = x^m + mx^{m-1}h + m(m-1)x^{m-2}\frac{h^2}{1\cdot 2} + \cdots$$

如果我们现在考察这方程的两边，那末左边给我们指明，由于 $x$ 增长了 $h$，$x^m$ 或 $f(x)$ 就变成了 $(x+h)^m$ 或 $f(x_1) = f(x+h)$；二项式 $(x+h)^m$ 就是这样从单项式 $x^m$ 变来的，但是它现在作为 $x^m$ 的变化的表示式，而不是象在通常的二项式 $(x+a)^m$ 中那样作为两个常量之和的乘幂的表示式出现的。关于这个**一般的、未展开的表示式** $(x+h)^m$ 或 $f(x+h) = y_1$ 就讲到这里。

右边级数展开的展开表示式中，第一项 $x^m$ 已不再象二项式定理中那样是二项式 $(x+h)^m$ 的幂次最高的第一项；它是 $f(x)$，因为 $y = x^m$，而所有其他各项则仅仅表示 $f(x)$ 或 $y$ 或 $x^m$ 由于 $x$ 增长了 $h$ 而得到的增量。

---

所以等式 A) 也可以写成：

$$\text{B)}\quad f(x+h) = y_1 = y + mx^{m-1}h + m(m-1)x^{m-2}\frac{h^2}{1\cdot 2} + m(m-1)(m-2)x^{m-3}\frac{h^3}{1\cdot 2\cdot 3} + \cdots$$

所有这些冗长的准备工作为我们所必需，因为在我们的方法中，$y_1$ 不被表示为 $f(x)$ + 它的导来项这种和式，而反过来一般地通过 $y_1 = A(x_1^m - x^m)$ 表示为 $f(x_1)$ 和 $f(x)$ 的差值（其中 $A$ 可以表示任意一个常量）。

与此相反，本来的微分方法是从 $x_1 - x = h$（即 $= \Delta x$）出发的；所以 $x_1 = x + h$；因而 $x_1$ 一开始就显现为 $x + h$，即显现为一次幂的二项式，因为 $x + h = (x+h)^1$。因此，它的微分表示式也是 $(x + dx)$。所以除了**用 $x$ 表达的**一次函数以外，对于所有其他**用 $x$ 表达的**函数，一旦 $x$ 增长 $h$，就要去算二次幂

以上的**二项式**,而其展开本身在于二项式定理的应用。

因此在这里不言而喻，展开级数的也就是二项式展开的第一项 $=f(x)$ 或 $y$，所有其余各项 = 这个函数由于 $x$ 变为 $x+h$ 而获得的增量，因而二项式一般公式 $(x+h)^m$ 巧妙的表示式就立刻被表示**成等式** B)**的形式**。

---

整个事情这样开始较好：

## I

假定已给：

$$f(x)=x^6,\qquad f(x_1)=x_1^6。$$

我们以前(见第一份手稿)已经指出：如果 $f(x)=x^m, f(x_1)=x_1^m$,那末

$$\begin{aligned}y_1-y=f(x_1)-f(x)&=x_1^m-x^m=\\&=(x_1-x)(x_1^{m-1}+x_1^{m-2}x+x_1^{m-3}x^2+x_1^{m-4}x+\\&\quad+\text{一直到第 } m \text{ 项 } x_1^{m-m}x^{m-1})。\end{aligned}$$

如果我们现在用 $(x_1-x)$ 去除它,那末

$$\frac{y_1-y}{x_1-x}\text{或}\frac{f(x_1)-f(x)}{x_1-x}=(x_1^{m-1}+\cdots+x_1^{m-m}x^{m-1})。$$

如果现在 $x_1$ 变为 $=x$,因而 $x_1-x=0$,那末

$$\frac{dy}{dx}\text{或}\frac{(f(x_1)-f(x))}{(x_1-x)}=mx^{m-1}。$$

凡是对于 $x^m$ 的一阶**导函数**所做的，对于所有其后的导函数都同样适用。它们都将用**同一个方法**所找到,而这个方法则基于这样一个**代数**定理：如果我们有一个形如 $x^m-a^m$ 的差值，那末它总可以用 $x-a$ 除尽，也就是总可以表示成 $(x-a)P$。

---

我们从代数学知道,在二项式 $(x+h)^m$ (例如, $=(a+c)^m$, 其中 $m$ 表示正整数幂) 中那些用 $x$ 表达的表示式以二项式的后一项即这里的 $h$ 的升幂作为因子,

$$(x+h)^6 \text{ 是 } =(x+h)(x+h)(x+h)(x+h)(x+h)(x+h)。$$

这里无论 $x$ 现在是变量、或者是代数中的未知常量、或者甚至象 $(a+h)^6$ 中的 $a$ 那样是已知量,这后一项，即这里的常量项 $h$，总将以正整数升幂 $h^0(=1)$, $h^1, h^2$ 等等 (在一定的假定下) 构成逐次表示式的因子，这些逐次表示式是 $x$ 作为代数中的常量时通过逐次自乘而得到的，而变量 $x$ 的这些函数则通过对

中间“导函数”的微分而得到的。

---

**必须注意**：$h$ 在这里不但开始时是作为一个普通**常量**引进的，象例如代数中的 $(x+a)^6$，$(x+c)^6$，$(a+0)^6$，$(a+h)^6$ 那样，而且必须永远保持其为一个**常量**，并在这里所选定的代数的微分方法的基础上，它既不能自己是个变量，又不能成为变量的增量，因为在这方法中自变量的差值 $x_1-x$（以及相应的因变量的差值 $y_1-y$ 或 $f(x_1)-f(x)$）始终保持其这种原来的形式，因此决不能令 $x_1-x$ 和 $y_1-y$ 等于某一个差值，$=\Delta x$ 或 $h$，因而也决不能象 $x_1-x=\Delta x$ 或 $h$ 那样使 $x_1$ 变为 $=x+\Delta x$ 或 $=x+h$，也不能使 $y_1-y$ 变为 $=\Delta y$ 或 $k$。如果我们把**预先导函数**的等价物写成 $\dfrac{y_1-y}{x_1-x}$ 或 $\dfrac{\Delta y}{\Delta x}$，那末对我们来说，它在这里只是用以表明不确定的 $x_1-x$ 和 $y_1-y$ 的记号，而不是**差值的值**，所以不是把 $x_1-x=\Delta x$ 作为不定差值的值，也不是把 $y_1-y=\Delta y$ 作为这样的值。在 $f(x)$ 本身的求导中，增长了的 $x$ 总是表现为 $x_1$，而另方面，增长了的 $y$ 总是表现为 $y_1$，所以不能同时表现为 $x+\Delta x$，$y+\Delta y$。

---

当 I）中的 $y_1=x^m+\text{etc.}$① 独立出现时，我们一见到它就能把它看作是形式最简单的二项式 $(x+h)^m=x^m+\text{etc.}$ 的普通代数表示式，而且确是一种正整数幂的二项式的普通代数表示式，因为这是从一开头就对幂次——指数 $m$——这样假定了的。我们首先在它这种形式下把它看作转向微分方法的过渡点，在这方法中我们令 $x_1-x=h$［因而也令 $y_1-y=k$，只要它用得到］②，因而也令 $x_1=x+h$。它给了我们一个机会从迄今所用的方法直接［过渡到微分方法中去］。在前一方法中，$x_1-x$，$y_1-y$ 只以它们作为差值的这种普适形式出现，并且在这个方法里等式右边的 $x$ 一旦增长，它只以 $x_1$ 出现，因而与原先的 $x$ 一样处于不确定的形式之中，而决不会变为 $x_1=x+\Delta x$ 或 $x+h$。但是，假如有人认为可以反过来进行，在再令 $h$ 等于 $x_1-x$ 之下就能从微分方法过渡到我们的方法中来，那就错了。

---

① 这里所指的是等式

$$\text{I) } y_1 \text{ 或 } f(x+h)=(x+h)^m=x^m+mx^{m-1}\frac{h}{1}+m(m-1)x^{m-2}\frac{h^2}{1\cdot2}+$$
$$+m(m-1)(m-2)x^{m-3}\frac{h^3}{1\cdot2\cdot3}+\cdots+x^{m-m}h^m.$$

② 这对括号是在马克思的手稿里原来就有的。

$$\begin{cases} \text{I)} & (x+h)^5 = x^5 + 5x^4 h + 5\cdot 4x^3 \dfrac{h^2}{1\cdot 2} + 5\cdot 4\cdot 3\, x^2 \dfrac{h^3}{1\cdot 2\cdot 3} + \\ \text{a)} & \qquad + 5\cdot 4\cdot 3\cdot 2\, x \dfrac{h^4}{1\cdot 2\cdot 3\cdot 4} + h^5\text{。} \end{cases}$$

这是一个普通代数等式；一方面，在左边是两个常量 $x$ 和 $h$ 的二项式未展开的一般表示式，即 $(x+h)^5$；另一方面，在右边是由二项式定理所提供的展开级数。如果我们在一般表示式 $(x+h)^5$ 中令 $h=0$，那末代替 $(x+h)^5$ 我们将得到 $(x+0)^5=x^5$。如果我们现在把 $x$ 看作变量，那末 $x^5$ 是一个**用 $x$ 表达的**确定函数，也就是 $=f(x)$，但 $x^5$ 也是一个 $x$ **的**函数，因为 $f(x)$ 的值将随着 $x$ 的变化而变化；只要我们现在把这样的函数看作是依赖于 $x$ 的函数，我们就称它为 $y$；而只要认为 $f(x)$ 处于这种依赖关系之下，它就和 $y$ 意义相同。另一方面由于 $x^5=f(x)$，$(x+h)^5$ 将变为 $f(x+h)$ 或 $y_1$。

因此，代替 I)a) 我们得到：

$$y \text{ 或 } f(x) = x^5,$$

$$y_1 \text{ 或 } f(x+h) = (x+h)^5 =$$

$$= x^5 + 5x^4 h + 5\cdot 4\, x^3 \frac{h^2}{1\cdot 2} + 5\cdot 4\cdot 3\, x^2 \frac{h^3}{1\cdot 2\cdot 3} + \cdots + h^5\text{。}$$

如果我们再令 $h=0$，那末这个级数将化为 $x^5$，而在左边 $f(x+h)$ 将化为 $f(x)$ 或者 $y_1$ 将化为 $y$。所以我们得 $y$ 或 $f(x)=x^5$。如果我们把 $x^5$ 的这个值代入等式 I)a)，那末它就变为等式 b)〔见下面〕。通常两个常量的二项式展开级数其第一项仍将作为第一项留在它的位置上，但它的特性已经完全改变。它不再是通常二项式 $(x+h)^5$ 的第一项，而是变量 $x$ 的原先的函数——即 $x^5$——在变为 $f(x+h)$，也就是在 $x^5$ 变为 $(x+h)^5$ 之前的自身。

因此，这等式现在变为：

$$\text{b)}\ y_1 \text{ 或 } f(x+h) = y \text{ 或 } f(x) + \left(5x^4 h + 5\cdot 4x^3 \frac{h^2}{1\cdot 2} + \right.$$

$$\left. + 5\cdot 4\cdot 3\, x^2 \frac{h^3}{1\cdot 2\cdot 3} + 5\cdot 4\cdot 3\cdot 2\, x \frac{h^4}{1\cdot 2\cdot 3\cdot 4} + h^5\right)\text{。}$$

另一方面从 b) 得出：

$$\text{c)}\ y_1 - y \text{ 或 } f(x+h) - f(x) = 5\, x^4 h + 5\cdot 4\, x^3 \frac{h^2}{1\cdot 2} +$$

$$+ 5\cdot 4\cdot 3\, x^2 \frac{h^3}{1\cdot 2\cdot 3} + 5\cdot 4\cdot 3\cdot 2\, x \frac{h^4}{1\cdot 2\cdot 3\cdot 4} + h^5\text{。}$$

因此，随着使第一项 $x^5$ 转变为 $f(x)$ 的这第一次突击，二项式的整个其余级数就从第二项起立刻变为一个用 $x$ 表达的导函数的级数。

作为因子附在这些“导”函数上的 $h, h^2, h^3$ 等等，现在由于 $x_1-x=h$ 或 $x_1=x+h$，就从通常二项式 $(x+h)^5$ 中的第二项变为两变量 $x_1$ 和 $x$ 或 $(x+h)$ 和 $x$ 的差值的各种不同乘幂，因为 $((x+h)-x)=h$ 是因变量 $x$ 增长到 $x_1$ 或 $x+h$ 而产生的差值。所以**整个级数**现在就是**用 $x$ 表达的原函数与 $x$ 增长到 $(x+h)$ 时的函数的差值**。这个差值也可以肯定地表示为用 $x$ 表达的原函数因 $x$ 增长到 $x_1$ 或 $x+h$ 而获得的增量。

因此，通过一个非常简单的技巧就使 a) 中 $(x+h)^5$ **的整个展开级数**从两个常量的二项式表示式变成了这样一个级数，其第一项是变量 $x$ 的原函数，而所有其他各项则是这原函数在它从 $f(x)$ 变为 $f(x+h)$ 时所获得的增量之和。可见这些增量都由它的运动所产生，而不是作为一个普通代数二项式的展开项得到的。因此得出了（见上面）等式 c）。

然而令 $h=0$ 的做法不但对我们起着这样一种形态变化的作用，使两个常量的通常二项式变为变量 $x$ 的函数当这个变量增长时的展开式。它还一开始就给我们指出了**那个**一方面为把用 $x$ 表达的已完全现成地包含着的那些逐阶导函数从它们依次存在的展开式中解脱出来，另一方面为引出与它们相应的符号微系数所必须应用的**方法**。

A）级数第一项 $x^5$ 以前之所以可以置为 $=f(x)$，1）因为它本身有 $h^0=1$ 作为因子，所以从开始起就不依赖于因子 $h$；2）因为通过令 $h=0$，一切其他的项预先都被清除掉了，因而整个级数就化为 $x^5$。所以，为了使导函数 $5x^4$，$5\cdot4x^3$ 等等能处于类似 $x^5$ 那样的状态，就必须 1）把它们逐个地从因子 $h, h_2$ 等等那里解放出来，这只有逐次用 $h$ 去除才能做到；2）一个“导函数”一旦这样地摆脱了 $h$，也就是作为“导”函数被解脱了出来之后，就必须象在第一项那里一样令 $h=0$，也就是把它的一切附带项预先清除掉，而且象以前在第一项那里一样，把整个级数化为这个得到了解放的“导函数”。

B）在对级数第一项进行的运算中，令 $h=0$ 就使它的左边从 $f(x+h)$ 或 $y_1$ 变为 $f(x+0)$ 或 $y$，也就是变为 $f(x)$。但由于只有通过用 $h$ 去除，才能使那些**导函数**摆脱掉它们用 $h$ 表达的因子，所以现在这个除法在左边就产生了

$$\frac{y_1-y}{h}\left(=\frac{y_1-y}{x_1-x}\right)\text{或}\frac{f(x+h)-f(x)}{h}。$$

因此，为了把整个级数化为得到解放的“导函数”而 $h$ 在右边一旦变为 0，那末左边就必须变为

$$\frac{y_1-y}{0}=\frac{0}{0}\text{或}\frac{f(x+0)-f(x)}{0}=\frac{0}{0}=\frac{dy}{dx}。$$

所以令 $h=0$ 就在左边产生了**用 $x$ 表达的**"导"函数的符号微系数。

因此，现在已经**证明**，为获得级数第一项 $=f(x)$ 或 $y$ 所用的那个最初运算一箭双雕地提供了：

1）通常的二项式 $(x+h)^5=x^5+\text{etc.}$ 转变为 $f(x+h)=f(x)+$ 从 $f(x)$ 导来的一些函数的一个级数，这些导函数附有因子 $h, h^2$ 等等，也就是附有这样一种增量的乘幂，这种增量是自变量 $x$ 当从 $x$ 变为 $x_1$，因而把 $h=x_1-x$ 作为它的增量时得到的。

2）一种**方法**，它把用 $x$ 表达的早已完全现成地存在着的（事实上借助于两个常量 $x$ 和 $h$ 的普通二项式的展开而产生的）那些"导"函数本身**解放了出来**，而同时把它们的**符号微分表示式**放在与它们对立的一边。

因此，现在我们可以重新回到本题上来。

$$\begin{array}{l} \mathrm{I)} \\ \mathrm{a)} \end{array}\left\{\begin{array}{l} (x+h)^5=x^5+5x^4h+5\cdot4\,x^3\dfrac{h^2}{1\cdot2}+5\cdot4\cdot3x^2\dfrac{h^3}{1\cdot2\cdot3}+ \\ \qquad\qquad\qquad\qquad\qquad\qquad +5\cdot4\cdot3\cdot2\,x\dfrac{h^4}{1\cdot2\cdot3\cdot4}+h^5。\end{array}\right.$$

# [泰勒定理增补二]

我们曾经通过**微分**从等式

$$(x+h)^{m+1}=x^{m+1}+(m+1)x^{m}h+\cdots$$

导出了起始等式

$$1)\ (x+h)^{m}=x^{m}+mx^{m-1}h+\cdots$$

本身。

但在这样做时，我们总是把**二项式定理**看作已给定的出发点；为了通过**微分**去找二项式$(x+h)^m$，因而为了得到一个作为通过微分给定的出发点，我们只把二项式$(x+h)$的$(m+1)$幂次作为出发点。

**代数二项式**根本只考察**确定幂次**的二项式，从我们只与常量打交道的代数方面来看，$x+h$ 本身就是一次的二项式 $=(x+h)^1$。我们可以象在等式 3)和 4)① 中所发生的那样，把从$(a+h)$得到的级数任意加以延伸，并把这种延伸用 + etc. 来表示，因为 $m$ 是一个数值**不定**的代数量；然而这样做并不妨碍级数的有限性。我们也可以把它有头有尾地写成：

$$x^{m}+mx^{m-1}h+\cdots+mxh^{m-1}+h^{m}。$$

我们在微分演算中所应用的那个 $y_1$ 或 $f(x+h)$，在这里只是并且永远只是$(x+h)$的一定的、即使是任意的幂次的二项式的一种符号。

因此，我们只是**形式上**用 + etc. 来截断而得到了一个无穷级数，事实上它象被截断的那样也可用头和尾以及居于中间的间隔 + etc. 来表示。但这还不是所有的一切。各个函数的系数 $h^0$（或 1），$h^1$，$h^2$，$h^3$ 等等表明，我们已经把$(x+h)^m$ 表示为正整数升幂；因此，我们不但以代数的、因而必然处于任一已给幂次的二项式为基础，而且甚至以二项式定理的单方面形式为基础。此外还不应忘记，为了得到 $h$ 作为用 $x$ 表达的函数的纯粹因子，我们选择了形式$(x+a)^m$（因为 $h$ 在代数二项式中象 $a$ 一样是一个常量），其中 $x$ 为前一项、$h$ 为后一项，而不是选择两项位置相反的形式$(a+x)^m$。

当然，我们也可以从一个具有负整数幂或分数幂 $-m$ 或$\frac{m}{n}$的二项式出发，从而得到一个无穷级数。

但是除了以后行将阐明的其他一些限制外，事实上我们还可再以具有一

---

① 见本文第 96—97 页。

个确定幂次的二项式作为基础，因为 $-m$ 或 $\frac{m}{n}$ 象 $m$ 一样都是确定的幂次。无限的展开级数本身在这里是具有确定幂次的一般表示式(如$(x+h)^{-m}$，$(x+h)^{\frac{m}{n}}$)的展开，即使是负整数幂或分数幂。

我们所获得的 $y_1$ 或 $f(x+h)$，在这里永远仅仅是任何一个正整数幂、负整数幂或分数幂的二项式的符号；因此与 $y_1$ 或 $f(x+h)$ 相应的展开级数，始终不过是一个具有一定幂次的二项式的一般表示式，而且事实上只不过是具有一定代数形式的**某个二项式的一般表示的例子**而已。

这种缺陷也许可以避免。如果我们求助于代数的待定系数法，我们或许能够摆脱代数二项式的局限性。首先我们把等式 2) 中的表示式重新带回到它原先的代数形式，也就是从

$$2)\ f(x+h)\ 或\ y_1=x^m+mx^{m-1}\frac{h}{1}+m(m-1)x^{m-2}\frac{h^2}{1\cdot2}+$$
$$+m(m-1)(m-2)x^{m-3}\frac{h^3}{1\cdot2\cdot3}+\cdots$$

带回到

$$2a)\ f(x+h)\ 或\ y_1=x^m+\left(\frac{m}{1}x^{m-1}\right)h+\left(\frac{m(m-1)}{1\cdot2}x^{m-2}\right)h^2+$$
$$+\left(\frac{m(m-1)(m-2)}{1\cdot2\cdot3}x^{m-3}\right)h^3+\cdots$$

这些函数在这里不再显现为**整个函数**，而象它们原先在**二项式展开**中所表示的那样显现为

$$x^m+mx^{m-1}h+\frac{1}{2}m(m-1)x^{m-2}h^2+$$
$$+\frac{1}{1\cdot2\cdot3}\left(或\frac{1}{6}\right)m(m-1)(m-2)x^{m-3}h^3+\cdots$$

并且这是用括号来标明的。

在这种形式上的修正中，$h,h^2,h^3$ 等等从它们的分母那里被解放了出来，而那些用 $x$ 表达的函数则把自己表示成象在它们原先的代数推导中的那样，其中只有第二项 $mx^{m-1}$ 以整个函数出现，第三项以整个函数的 1/2 出现，等等。在这样一种修正之后，现在我们用不定系数 $A,B,C$ 等等来替换 $x$ 的这些函数，从而在 3)中(因此也对于 4))我们得到

$$f(x+h)或\ y_1=f(x)或\ y+Ah+Bh^2+Ch^3+Dh^4+Eh^5+\cdots。$$

$A,B,C,D,E$ 等等在这里显现为 $x$ 的函数(请与 2a)比较)，而这些函数恰是还有待于去寻找的。此外，这级数实际上可以任意延长，而且现在已不能

再被化为一个其末项象首项一样可以被确定的级数，因为我们继续可以任意地写下 $Fh^6$，$Gh^7$ 等等，等等，这是由于那些数值上的推导叛逆者，即系数 $\frac{1}{2}$，$\frac{1}{1\cdot2\cdot3}$，$\frac{1}{1\cdot2\cdot3\cdot4}$ 等等，在不定系数 $A,B,C,D,E,F$ 等等中已经和导函数本身一起都看不见，所以通过右边的这一技巧，使左边的 $f(x+h)$ 或 $y_1$ 从一个任一幂次的代数二项式的符号变为这个无穷级数的一般的未展开的表示式，这个级数虽然包含着每一个幂次，但它并不表明任何一个幂次。

$f(x)$ 是变量 $x$ 的一个可以无限制加以展开的函数；$f(x+h)$ 是 $x$ 的这一函数当 $x$ 变为 $x+h$ 时的一般的未展开表示式。

此外，就等式

3) $y_1$ 或 $f(x+h)=y+f'(x)h+f''(x)\frac{h^2}{1\cdot2}+\cdots$

而论，不言而喻，对于函数 $f'(x)$，$f''(x)$ 等等不能再象在等式 4) 中所做的那样，简单地用它们的符号等价物来替换，而这些符号等价物是先要通过微分来寻找的东西。

## [泰勒定理增补三]

泰勒的起始等式是：

I) $y_1$ 或 $f(x+h)=y+Ah+Bh^2+Ch^3+Dh^4+Eh^5+\cdots$

为了对这个等式可以有所着手，就必须应用一个本身仍由二项式所提供的技巧，而且确是在于把二项式定理应用到多项的表示式上。

如果我们作为例子有$(x+a)^2$，并用$(x+a+b)^2$来代替它，那末我们可以把后者当作$((x+a)+b)^2$或者当作$(x+(a+b))^2$来加以展开：

在第一种情形下，我们得到：

$$(x+a+b)^2=(x+a)^2+2(x+a)b+b^2;$$

在第二种情形下：

$$(x+a+b)^2=x^2+2x(a+b)+(a+b)^2。$$

在这例子中，问题仅仅在于在这两个右边之间存在着一种形式上的，即使明显地只是一种形式上的差异；而这两个右边的恒等一开始就由两个左边的恒等得到了证明。

---

现在我们要问：用什么方法能使我们把形式最简单的代数二项式$(x+h)^m$用不定系数隐藏起来的表示式 I)变为 IV)① 的呢？

是不是通过这种伎俩，即我们先把 I)对 $h$ 微分，然后对 $x$ 微分，从而为 $y_1$ 得到了两个微分等式，其一般表示式$\frac{dy_1}{dx}$和$\frac{dy_1}{dh}$恒等，因而其相应的展开级数也等价，所以它们单个的项就必须可以等同起来呢？

决不是这样。给我们指明**两等式中 $x$ 的哪些函数可以等同起来**的那个准则，是由因子 $h^0(=1), h^1, h^2, h^3, h^4$ 等等提供给我们的。只有那些以同一幂次的 $h$ 为因子的函数才可以**等同起来**；然而支配整个过程的这些因子 $h^0, h^1, h^2, h^3$ 等等本身，是由什么东西提供给我们的呢？

起始等式 I)：

$$\text{I)}\quad y_1=yh^0+Ah^1+Bh^2+Ch^3+Dh^4+Eh^5+\cdots$$

---

① 这里是指表示式

$$\text{IV)}\quad y_1=y+\frac{dy}{dx}h+\frac{d^2y}{dx^2}\,\frac{h^2}{1\cdot 2}+\frac{d^3y}{dx^3}\,\frac{h^3}{1\cdot 2\cdot 3}+\frac{d^4y}{dx^4}\,\frac{h^4}{1\cdot 2\cdot 3\cdot 4}+\cdots$$

我们虽然把从二项式导来的那些函数 $f'(x),f''(x),f'''(x)$等等用**不定系数** $A,B,C,D,E$ 等等伪装了起来，可是我们还记得，这些不定系数压根儿都是 $x$ 的函数，否则我们就既不能把它们对 $x$ 微分，又不能把它们对 $h$ 微分；但是我们把因子 $h^0$、$h^1$、$h^2$、$h^3$ 等等以其由二项式定理所提供的那种原始形式带到了我们的起始等式中去，而如果不把它们带去，那末即使有了能把等式 I)变为微分形式的两种等价而形式不同的翻译这种伎俩，也是无济于事的。

---

这在后来想给 I)中的微分推导以这样一种形式的那些尝试中更加明显地暴露出来了。在这种形式中不但 $x$ 的函数，**就是 $h$ 的各个幂次也似乎都是通过微分**找到的。因为起始等式在这里被写成：

$$\text{Ia)}\ y_1 = y + Ah^\alpha + Bh^\beta + Ch^\gamma + Dh^\delta + \cdots$$

于是我们得到：

$$1)\ \frac{dy_1}{dh} = \alpha Ah^{\alpha-1} + \beta Bh^{\beta-1} + \gamma Ch^{\gamma-1} + \delta Dh^{\delta-1} + \cdots,$$

$$2)\ \frac{dy_1}{dx} = \frac{dy}{dx} + \frac{dA}{dx}h^\alpha + \frac{dB}{dx}h^\beta + \frac{dC}{dx}h^\gamma + \frac{dD}{dx}h^\delta + \cdots$$

然后我们就这样来论证：1)和 2)两个的左边相等，所以它们的展开级数等价，而它们相应的同类项可以等同起来；因此第一步 $\alpha Ah^{\alpha-1} = \frac{dy}{dx}$。但由于 $\frac{dy}{dx}$以 $1(=h^0)$ 作为因子，所以必须是 $h^{\alpha-1}=h^0$，因而 $\alpha-1=0$，即 $\alpha=1$。

这样就一举三得：1)我们现在知道，由于 $\alpha=1$，所以起始等式中的 $Ah^\alpha = Ah^1 = Ah$；2)起始等式中也已**较为潜在地**把 $h^\beta$，$h^\gamma$，$h^\delta$ 等等中不定幂次 $\beta$，$\gamma$，$\delta$ 等等的值给了我们；3)由于 $\alpha Ah^{\alpha-1} = \frac{dy}{dx}$，亦即 $\alpha Ah^{\alpha-1} = 1\cdot A\cdot h^{1-1} = A$，所以也得到 $A = \frac{dy}{dx}$。

第二个等式是 $\beta Bh^{\beta-1} = \frac{dA}{dx}h^\alpha$。因此 $h^{\beta-1}=h^\alpha$；也就是 $\beta-1=\alpha$；但由于 $\alpha=1$，所以 $\beta-1=1$，由此得出 $\beta=2$，而 $\beta Bh^{\beta-1}$ 也就 $=2Bh^{2-1}=2Bh$。

现在我们知道，在起始等式 Ia)中我们第一可以用 $Bh^2$ 去代替 $Bh^\beta$；因此，这是**推导出来**的，而不是一开始就假定的。

但其次是 $\beta Bh^{\beta-1} = \frac{dA}{dx}h^\alpha$ 变为 $2Bh = \frac{dA}{dx}h$(因为 $h^\alpha = h^1$)，因此 $2B = \frac{dA}{dx}$，

也就是 $B=\frac{1}{2}\frac{dA}{dx}$；如果我们把 $A$ 的值代入其中，那末

$$B=\frac{1}{2}\cdot\frac{d}{dx}\left(\frac{dy}{dx}\right)=\frac{1}{2}\cdot\frac{d^2y}{dx^2}。$$

没有必要再继续进行下去。

---

泰勒展开的这种改善了的、偏于矫揉造作的版本，归结到一点就是如下：整个关键在于把两个等式中的第一项等同起来：$\alpha Ah^{\alpha-1}=\frac{dy}{dx}$；由于$\frac{dy}{dx}$以 $h^0$ 为其因子，所以必须 $h^{\alpha-1}=h^0$，因此 $\alpha-1=0$，也就是 $\alpha=1$ 以及 $\alpha Ah^{\alpha-1}=Ah^{1-1}=A$。

但是如果假定起始等式

$$y_1=yh^0+Ah^\alpha[+\cdots]$$

中的 $\alpha$ 为负（而且象二项式定理在其最一般的展开中所指出的那样，这不定的 $\alpha$ 可以是一切可能的数），那末 $Ah^\alpha$ 变为 $Ah^{-\alpha}$，而 $\alpha Ah^{\alpha-1}$ 变为 $-\alpha Ah^{-\alpha-1}$。如果我们现在象前面展开中得出的那样令 $\alpha=1$，那末 $-\alpha Ah^{-\alpha-1}=-Ah^{-1-1}=-Ah^{-2}$。按照我们以前的论证，就必须作这样的结论：由于

$$\frac{dy}{dx}=\frac{dy}{dx}\times1(\text{或}h^0),$$

所以必须 $h^{-2}=h^0$，因而 $-2=0$，或者，如果我们愿意用另一种方法来表述的话，由于 $\frac{dy}{dx}\cdot h^0=-\alpha Ah^{-1-1}$，所以必须 $h^{-1-1}=h^0$，因而 $-1-1=0$，它给我提供了 $-1=+1$ 这样一个结果。可是这个乏味的结果只是证明了：我们暗中所假定的 $h^\alpha$ 不过是 $h^1$ 的一个化了装的表示式，也就是暗中假定了 $\alpha=+1$；并且还证明了：我们关于因子 $h^0$，$h^1$，$h^2$，$h^3$ 等等的展开级数不是简单地由假定得来、而它们相应的 $h$ 的数字幂次 1,2,3 等等是通过微分过程从不定幂次 $\alpha,\beta,\gamma$ 等等推导得来的那个论点，一开头就尽是欺骗。

再者，$Ah^\alpha$ 中的 $\alpha$ 也可能小于 1,……

因此问题仍然是这样，即起始等式

$$y_1=y+Ah^\alpha+Bh^\beta+Ch^\gamma+Dh^\delta+\cdots$$

**或者只是**从二项式带来的 $h^1$，$h^2$，$h^3$ 等等的**一种装腔作势的伪装**，或者，当我们把譬如 $\alpha$ 真的看作幂次的**纯粹一般符号**，也就是看作 $\alpha$ 的所有可能形式的纯粹一般的符号时，两个用微分导出的等式的各项是不能等同起来的，因而

是根本一文不值的。

然而我们知道，确有这样一些 $x$ 的函数，它们在 $x$ 一旦增长 $h$ 时就给出 $h$ 的负数和分数乘幂。

**结果**因而是：那些从二项式得来的、用 $x$ 表达的不同阶的导函数的［因子］，也就是 $h^0, h^1, h^2$ 等等，仍然是必须加以**假定的**，由二项式所提供而不是由微分演化提供给我们的；因而所由出发的那个二项式的形式，是一个完全确定的形式，也就是因子 $h$ 在其中被展开成**正整数升幂**的形式。

---

但是我们还没有结束。对于那些伪装在不定系数 $A, B, C, D$ 等等中的用 $x$ 表达的导函数(参见 3)) $f'(x), f''(x)$ 等等，事情也不完全干净利落。

在那里我们已经导出了 $f'(x)=\frac{dy}{dx}$，然而在那里我们也已经有了 $f(x)=x^m$，所以 $f'(x)=mx^{m-1}$，也就是说，象所有后来通过二项式导出的用 $x$ 表达的函数一样，$f'(x)$ 曾是一个**用 $x$ 表达的确定而有限的表示式**；但必须注意，那些利用二项式加以展开的原先的函数，在它们的**一般形式**中总有其确定的幂次，例如当 $f(x)=x^m$ 时就是如此，当它 $=\frac{a^m}{x^m-a^m}$ 时也同样是如此。

这些导函数可以表示成一个有限的或无限的级数，它的每一项都是一个用 $x$ 表达的确定的、因而是**有限的**表示式。例如在 $x^4$(或 $x^m$)的情况下，$4x^3$(或 $mx^{m-1}$)是一个确定而有限的表示式，这决不影响它作为用变量 $x$ 表达的函数的可展开性。$4x^3$ 由于 $x$ 的重新变为 $x+h$ 而变为 $4(x+h)^3$，尽管 $4x^3$ 是一个**确定的**因而是**有限的**表示式。

同样：

$$\frac{a}{a-x}=1+\frac{x}{a}+\frac{x^2}{a^2}+\frac{x^3}{a^3}+\cdots$$

这级数的每一表示式 $\frac{x}{a}$，$\frac{x^2}{a^2}$，$\frac{x^3}{a^3}$ 等等都是 $x$ 的一个**因确定而为有限的**函数，而且无论我通过用 $a-x$ 去除 $a$，或者还是通过对 $\frac{a}{a-x}$ 进行逐次微分这个捷径找到了展开级数，对此完全无关紧要。

但是一旦我们在 $f'(x)$，$f''(x)$ 的地方代入了不定系数 $A$，$B$ 等等，那末我们又来到了进退两难的境地，就象在 $h^0, h^1, h^2$ 等等这些完全来自二项式某一特殊形式而在泰勒起始等式中保持为导函数的因子那里所遇到的一样。

要末 $A, B, C$ 等等——这些不定系数——仅仅是从二项式借来的**用 $x$ 表**

**达的确定的、因而是"有限的"导函数**的别名，要末 $A,B,C$ 等等作为**用 $x$ 表达的**导函数的一般符号，不但当这些导函数确定而有限时必须是它们的符号，而且当它们 $=0$，$=\infty$ 或 $=-\infty$ 时也必须是它们的符号。而我们知道，这种情况事实上是出现的。

---

所以，在泰勒定理中如果 1)从二项式定理的一个假定为 $(x+h)^m$ 的**特殊形式**那里把 $m$ 为**正整数幂**接受过来，因而用 $h$ 表达的因子也就 $=h^0,h^1,h^2,h^3$ 等等，即 $h$ 以整的，上升的，正的幂次出现；那末 2)象在**一般**形式的代数二项式定理中那样，**$x$ 的导函数是用 $x$ 表达的**确定的、因而是**有限的函数**。但是还要加进第三种情况。当变数 $x$ 取**特殊值**例如 $x=a$ 时，$x$ 的导函数只能变为 $=0$，$=+\infty$，$=-\infty$。同样，$h$ 只能 $=h^{-1}$ 或 $h^{m/n}$(例如 $h^{1/2}$)。

概括起来：只有当 1)自变量 $x$ 保持一般的**不确定形式** $x$，2)用 $x$ 表达的原函数本身通过取差值可以展开成**用 $x$ 表达的、确定的**，因而是**有限的导函数**的级数，并附有相应的正整数升幂的因子 $h$，即附有 $h^1$，$h^2$，$h^3$ 等等的时候，**泰勒定理**一般才能应用于展开那些用 $x$ 表达的函数，而在这展开中 $x$ 变成了 $x+h$ 或者说从 $x$ 增长到了 $x_1$。

但是所有这些条件都是下面这一事实的另一种表述，即：这个定理不过是翻译成微分演算语言的、带有**正整数幂次**的二项式定理。

凡是这些条件不满足的地方，也就是说，凡是**泰勒定理不能应用**的地方，就会出现微分演算中的所谓这定理的"**失误**"。

但是泰勒定理的最大**失误**，并不是这些特殊的应用上的失误，而是**一般的失误**，它把任何幂次的二项式的仅仅符号的表示式

$$y=f(x),\qquad y_1=f(x+h)$$

变成了这样一个表示式，其中 $f(x)$ 是包含所有幂次、因而本身没有任何幂次的 $x$ 的函数，以致 $y_1=f(x+h)$ 同样包含所有幂次而本身没有幂次，或者不如说 $f(x)$ 是当 $x$ 增长时变量 $x$ 的任一函数的**未展开的一般表示式**。因此这个没有幂次的 $f(x+h)$ 所展开成的展开级数，即 $y+Ah+Bh^2+Ch^3+\text{etc.}$，也包含所有幂次而本身不属于任何一个幂次。

从**普通代数**到**变数代数**的、而且确是**借助于普通代数**的这个跳跃，被认为是一个**既成事实**，它没有得到证明，并且从表面看来是和普通代数的**一切法则相矛盾**的；在普通代数中，$y=f(x)$，$y_1=f(x+h)$ 决不可能具有这种意义。

换句话说：起始等式

$$y_1 \text{ 或 } f(x+h)=y \text{ 或 } f(x)+Ah+Bh^2+Ch^3+Dh^4+Eh^5+\text{etc.}$$

不仅**没有得到证明**，而且还自觉或不自觉地假定了**常量**可以用**变量**来代替——因为代数学，从而代数二项式也只容许有常量，而且只容许有**已知的**和**未知的**两类常量——这种代替就与代数的一切法则相矛盾。因此这等式的来自代数的推导似乎是建立在一个骗局上的。

如果**泰勒定理**——它的失误在应用中几乎无足轻重，因为它们实际上只限于那些用微分方法得不出什么结果、因而根本不能用微分演算来处理的 $x$ 的函数——在实践中现在确实还证明它是整个演算的最概括、最一般和最富有成果的**运算公式**，那末这只是为他① 所从属的牛顿学派和微分演算的牛顿-莱布尼茨整个发展时期所建立起来的大厦加了一顶花冠，这个微分演算就在其开头的一些假设中从错误的前提得出了正确的结果。

泰勒定理的代数证明现在由**拉格朗日**所提供，而且这个证明根本就构成了**他的**微分演算的代数方法基础。关于这件事本身我将在本手稿的可能要写的历史部分中详加叙述。

这里作为历史的漫游只是提一下，拉格朗日并没有追溯到泰勒所未曾意识到的基础即二项式定理，而且是最初等形式的二项式定理上去；在这形式中它只由两个量 $(x+a)$ 或者在这里只由 $(x+h)$ 所组成，并且有一个正的指数。

他更加没有追溯下去而问自己，为什么翻译成微分形式的，并同时通过暴力从它的代数条件中解放出来的牛顿二项式定理竟会显现为他所奠定的微分演算的概括的一般运算公式呢？回答很简单；因为牛顿一开始就令 $x_1-x=dx$，从而 $x_1=x+dx$。所以**差值**的展开立即变为**和式**的展开，变为二项式 $(x+dx)$ 的展开（这里我们完全不管他本来应该令 $x_1-x=\Delta x$ 或 $h$）（从而 $x_1=x+\Delta x$ 或 $x+h$）。泰勒只是把这体系的基础发展到了它最一般和最概括的形式，但只有在微分演算的所有基本运算已经发现时，根本才有可能做到这一点。因为在人们不是对所有的 $x$ 的主要函数已能演化出它们相应的 $\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$ 等等的时候，他的 $\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$ 等等又有什么意义？

**拉格朗日则相反，他直接追随泰勒定理**，当然从这样一个立足点出发：一方面，牛顿-莱布尼茨时代的继承者已为他提供了 $x_1-x=dx$ 的修正了的版本，因而也提供了 $y_1-y=f(x+h)-f(x)$，另一方面，他恰恰在把泰勒公式代数化时建立了他自己的“**导**”**函数**理论。〔就这样，费希特追随康德，谢林追随费希特，黑格尔追随谢林，而无论费希特，谢林，黑格尔也许都根本没有研究过

① 这里的“他”是指泰勒。

康德的一般基础即唯心主义；否则他们就不会把它继续发展。]①

---

如果我们取泰勒的起始公式：

$$y = f(x),$$

$$\text{I)}\ y_1 \text{ 或 } f(x+h) = f(x) + Ah + Bh^2 + Ch^3 + Dh^4 + Eh^5 + \cdots,$$

那末可以把它写成：

$$= f(x) + h(A + Bh + Ch^2 + Dh^3 + Eh^4 + Fh^5 + \cdots)。$$

如果我们把括号里的整个表示式称作 $P$，那末：

$$1)\ y_1 \text{ 或 } f(x+h) = f(x) + Ph。$$

拉格朗日说：这就是一旦变量 $x$ 变为 $x+h$，从而 $f(x)$ 变为 $f(x+h)$ 时整个展开级数所必须化到的那个表示式；因为如果我们令 $h=0$，那末得到 $f(x+h)=f(x)$，也就是说，$f(x+h)$ 又化到了它原来的表示式。这就给我们证明了，$f(x+h)$ 的展开级数的第一项必须置为 $=f(x)$ 或 $y$。

现在，我们来详细研究 $Ph$，它是：

$$Ph = h(A + Bh + Ch^2 + Dh^3 + Eh^4 + \cdots);$$

因而

$$P = A + (B + Ch + Dh^2 + Eh^3 + \cdots)h。$$

如果我们令这个 $A=p$，并令括号中的表示式 $=Q$，那末 $P = p + Qh$。

如果我们现在把 $P = p + Qh$ 的值代入

$$1)\ f(x+h) = f(x) + Ph,$$

那末我们得到：

$$f(x+h) = f(x) + (p + Qh)h = f(x) + ph + Qh^2;$$

也就是：

$$2)\ f(x+h) = f(x) + ph + Qh^2。$$

拉格朗日起初只考察第二项。由于 $p$ 只在其**外部**附有作为因子的 $h$（不象 $Q$ 本身在**其内部**还包含着 $h$ 的函数），并且由于除 $x$ 和 $h$ 外我们根本没有级数的任何其他构成元素，所以 $p$ 必须是 $x$ 的一个函数，而且确是**它的一阶导函数**这个 $f(x)$ 的最小展开。然后他证明，$p$ 既不能 $=0$ 或 $\infty$，又不能附有具负数或分数指数的 $h$。在这证明中，巧妙的一点在于，因为 $f(x)$ 中的变量 $x$ 是不定的、一般的，因而决不会取一个 $=a$ 等等的特殊值，而且总是能够作任意增长，所以 $f(x+h)$ 在它的级数展开式中和在包括后者的一般展开定

---

① 这对括号是在马克思的手稿里原来就有的。

律中恰恰排除了在泰勒那里作为失误而出现的一切特殊情况。[巧妙的另一点是，那个同 $f(x+h)$ 的展开为级数交织在一起的导函数理论，又可以立即应用来更加详细地确定这个级数的各项。但是这里我不想对这点深入下去。]①

所以到目前为止我们有

$$f(x+h)=f(x)+ph\ (\text{或}\ f'(x)h)+\cdots$$

如果我们现在比较详细地来考察 $Qh^2$,那末它是

$$=(B+Ch+Dh^2+Eh^3+Fh^4+\cdots)h^2=$$
$$=Bh^2+h^3(C+Dh+Eh^2+\cdots)。$$

如果我们称 $B=q$，并把括号中的表示式称为 $R$，那末 $Qh^2=(q+Rh)h^2$；因此：$Q=q+Rh$。

如果我们把这表示式代入 2)，那末得到：

$$3)\ f(x+h)=f(x)+ph+(q+Rh)h^2,$$
$$f(x+h)=f(x)+ph+qh^2+Rh^3,$$

照这样继续演化下去，我们就得到：

$$R=r+Sh,\qquad S=s+Th,\cdots,$$
$$f(x+h)=y\ (\text{或}\ f(x))+ph\ (\text{或}\ f'(x)h)+qh^2+rh^3+Sh^4+\cdots$$

这个级数不会完结，因为每次总会重新获得一个新的表示式，象上面获得 $S=s+Th$ 那样，会获得 $T=t+Uh$，其中 $U$ 在自身中又包含着 $x$ 和 $h$ 的函数。因此这种展开方式就排除了展开级数的任何一个最后的终结。

① 这对括号是在马克思的手稿里原来就有的。

# 麦克劳林定理

I）泰勒定理所展开的是 $f(x+h)$ 或 $f(x_1)=y_1$；而麦克劳林则是$f(x)$或 $y$；也就是说，$x$ 的函数本身不是要用代数方法而是要用微分方法来展开；因此，事实上它无非就是要通过微分把 $x$〔各幂次〕的**常系数**找出来。

II）如果我们有一个二项式，例如$(x+c)^4$，那末根据我把 $x$ 或 $c$ 取作前一项或后一项的不同，能够把它展开成两种形式：

$$\text{a)}\ \ (x+c)^4=x^4+4x^3c+6x^2c^2+4xc^3+c^4,$$

$$\text{b)}\ \ (c+x)^4=c^4+4c^3x+6c^2x^2+4cx^3+x^4。$$

在 a)中，用 $x$ 表达的函数以被展开的形式出现而具有正整数升幂的 $c$，也就是 $c^0,c^1,c^2$ 等等作为纯粹的因子。在 b)中则相反，用 $c$ 表达的函数以被展开的形式出现，而具有正整数升幂的 $x$ 作为因子。

等式 b)显然是 a)的**倒逆等式**，因为如果我倒过来念 a)，那末就是

$$c^4+4c^3x+6c^2x^2+4cx^3+x^4。$$

因此，如果泰勒的起始等式是：

$$\alpha)\ \ y_1=y+Ah+Bh^2+Ch^3+Dh^4+\cdots,$$

那末，当我们用 $A$ 表示 $c^4$，用 $B$ 表示 $4c^3$，用 $C$ 表示 $6c^2$，用 $D$ 表示 $4c$ 时：

$$\beta)\ \ y=A+Bx+Cx^2+Dx^3+\cdots$$

在 $\alpha$)中，不定系数 $A,B,C$ 等等都是 $x$ 的函数，而在 $\beta$)中，它们则是常量 $c$ 的函数。

III）A）$y$ 或 $f(x)$　或 $(c+x)^m=$

$$=c^m+mc^{m-1}x+\frac{m(m-1)}{1\cdot2}c^{m-2}x^2+\frac{m(m-1)(m-2)}{1\cdot2\cdot3}c^{m-3}x^3+$$

$$+\frac{m(m-1)(m-2)(m-3)}{1\cdot2\cdot3\cdot4}c^{m-4}x^4+\cdots$$

由于只有**变量**才可以微分，所以通过微分来演化 $x$ 的函数的常系数似乎是一种附加语的矛盾①。

但是，我们将按照在泰勒定理中所做的那样来进行。如果我们令 $x=0$，那末

---

① “附加语的矛盾”一词，原文为 contradictio in adjecto。

$$y \text{ 或 } f(x)=(c+x)^m$$

变为

$$y \text{ 或 } f(0) \text{ 或 } (c+x)^m=(c+0)^m=c^m。$$

如果我们就这样通过令 $(c+x)^m$ 中 $x=0$ 而获得了 $c$ 的第 $m$ 次幂即 $c^m$，那末用同样的方法就可以找到 $c^m$ 的一切导函数，为此我们只要先通过微分运算求出$(c+x)^m$ 的导函数，然后再令 $x=0$；因此，由于

$$y \text{ 或 } f(x)=(c+x)^m,$$

所以

$$\frac{dy}{dx}=m(c+x)^{m-1};$$

如果我们在这里令 $x=0$，那末 $m(c+x)^{m-1}$ 就变为

$$m(c+0)^{m-1}=mc^{m-1}=f'(0)。$$

我们就这样陆续得到级数的各项。

因此我们得到：

$$f(x)=f(0)+f'(0)\,x+f''(0)\,\frac{x^2}{1\cdot 2}+f'''(0)\,\frac{x^3}{1\cdot 2\cdot 3}+\cdots=$$

$$=(y)_0+\left(\frac{dy}{dx}\right)_0 x+\frac{1}{1\cdot 2}\left(\frac{d^2y}{dx^2}\right)_0 x^2+$$

$$+\frac{1}{1\cdot 2\cdot 3}\left(\frac{d^3y}{dx^3}\right)_0 x^3+\cdots+\frac{1}{1\cdot 2\cdot 3\cdots m}\left(\frac{d^my}{dx^m}\right)_0 x^m,$$

式中括号$(\ \ )_0$表明，这些符号微系数相应于其中已令 $x=0$ 的那些“导”函数。

## [泰勒定理增补四]

I) $y_1$ 或 $f(x+h)$或$(x+h)^m=$

$$=x^m+mx^{m-1}\frac{h}{1}+m(m-1)x^{m-2}\frac{h^2}{1\cdot2}+m(m-1)\ (m-2)x^{m-3}\ \frac{h^3}{1\cdot2\cdot3}+$$

$$+m(m-1)(m-2)(m-3)x^{m-4}\frac{h^4}{1\cdot2\cdot3\cdot4}+\cdots$$

如果我们令 $h=0$,那末得到

$$f(x)=x^m。$$

秘密就此识破。$(x+h)^m$ 的第一项，也就是 $x^m$，不被看作二项式的第一项,而被看作变量 $x$ 的已知函数，在这里是 $x^m$。因此，I)中右边的整个表示式**减去** $x^m$ 本身以后似乎是这样产生的,即 $x^m$ 中的变量 $x$ 增长并变为 $x+h$，而这个 $x^m$ 就变为$(x+h)^m$，或者变为象我们[上面]所写的那样。

在把 $x_1$ 看作 $x+h$ 的系统中，就是这种**用以把已经包含在级数中的那些逐阶"导"函数**找出来的方法,也已经从我们用以找到 $x^m=f(x)$ 的第一次运算中得出来了。$x^m$ 只有 $h^0$(或 1) 作为因子；所以一旦令 $h=0$,其他附有 $h$ 的项就此消失,只留下 $x^m$ 作为 $y_1$ 的等价物,而 $y_1$ 则重新变为 $y$。

因此，其他一些 $x$ 的函数都必须从它们各自的 $h$ 那里逐个地解放出来，然后令 $h=0$。如同在第一次运算中通过令 $h$ 为零使 $f(x+h)$ 变为 $f(x+0)$，变为 $f(x)$ 那样，借助于有关的运算，左边将得到它的相应形式作为符号微系数。

由于

$$y_1\ (或\ f(x+h))=y\ (或\ f(x))+mx^{m-1}h+m\ (m-1)\ x^{m-2}\frac{h^2}{1\cdot2}+\cdots$$

所以

$$y_1-y\ (或\ f(x+h)-f(x))=mx^{m-1}h+m(m-1)\ x^{m-2}\frac{h^2}{1\cdot2}+\cdots$$

如果我们用 $h$ 去除两边,那末

$$\frac{y_1-y}{h}\left(或\frac{f(x+h)-f(x)}{h}\right)=$$

$$=mx^{m-1}+m(m-1)x^{m-2}\ \frac{h}{1\cdot2}+m(m-1)\ (m-2)x^{m-3}\ \frac{h^2}{1\cdot2\cdot3}+\cdots$$

$mx^{m-1}$ 在这里占有以前 $x^m$ 所占有的同样地位……

---

如果

$$(x+h)^m = x^m + mx^{m-1}h + \frac{m(m-1)}{1\cdot 2}x^{m-2}h^2 + \cdots,$$

那末微分演算与泰勒无关地已经证明了：若 $f(x)=x^m$，则

$$\frac{dy}{dx} \text{ 或 } f'(x) = mx^{m-1},$$

$$\frac{d^2y}{dx^2} \text{ 或 } f''(x) = m(m-1)x^{m-2},$$

等等。

但是如何证明的呢？

作为例子，让我们看变量的乘积 $xz$。

如果它们变化，那末这乘积就变为$(x+dx)(z+dz)$，这是一个**二次的二项式**。它与 $(x+a)(x+a)$或$(x+a)^2$ 形式上的差异仅仅在于我们把第一个$(x+a)$用$(x+dx)$来代替，而**把第二个**$(x+a)$用$(z+dz)$来代替，所以我们得到的不是 $x^2+ax+ax+aa$ 而是 $xz+zdx+xdz+dxdz$；如果我们从这里减去$xz$，那末得到：$zdx+xdz+dxdz$；如果我们删去末项，那末就得：

$$zdx+xdz。$$

在这样借助于二项式演化出 $d(xz)=zdx+xdz$ 以后，这个式子就可以应用到任意个变量的乘积上去，例如

$$\frac{d(xzuv)}{xzuv} = \frac{zuvdx}{xzuv} + \frac{xuvdz}{xzuv} + \frac{uzxdv}{xzuv} + \frac{xzvdu}{xzuv}。$$

如果我化约两边，那末得到：

$$\frac{d(xzuv)}{xzuv} = \frac{dx}{x} + \frac{dz}{z} + \frac{dv}{v} + \frac{du}{u},$$

$$\frac{d(x^4)}{x^4} = \frac{dx}{x} + \frac{dx}{x} + \frac{dx}{x} + \frac{dx}{x} = \frac{4dx}{x},$$

$$d(x^4) = 4\frac{dx}{x}\cdot x^4 = 4x^3dx,\quad \frac{d(x^4)}{dx} = \frac{dy}{dx} = 4x^3。$$①

假定说，这是由 $m$ 个可变因子而不是由 4 个因子组成的乘积 $xzuvty$ 等等，那末

$$\frac{d(x^m)}{x^m} = \frac{mdx}{x};$$

---

① 后两行是马克思用铅笔补写上去的。

因此

$$d(x^m)=mx^m\frac{dx}{x},\qquad d(x^m)=mx^{m-1}dx$$

或者

$$\frac{d(x^m)}{dx}\left(=\frac{dy}{dx}\right)=mx^{m-1};\quad 而\frac{d(mx^{m-1})}{dx}=m(m-1)x^{m-2}\text{ 等等。}$$

所以这个式子只是由于应用了作为运算公式的

$$d(xy)=ydx+xdy$$

才得到的，而这个运算公式则是应用二项式得到的。

---

关于等式 IV)：

$$y_1=y\ (或\ f(x))+f'(x)\frac{h}{1}+f''(x)\frac{h^2}{1\cdot2}+f'''(x)\frac{h^3}{1\cdot2\cdot3}+$$
$$+f^{\mathrm{IV}}(x)\frac{h^4}{1\cdot2\cdot3\cdot4}+f^{\mathrm{V}}(x)\frac{h^5}{1\cdot2\cdot3\cdot4\cdot5}+\cdots$$

现在也许可以这样说，虽然$f'(x)=mx^{m-1}$，$f''(x)=m(m-1)x^{m-2}$ 等等也是为**微分演算**所证实了的；但数字因子$\frac{1}{1\cdot2}$，$\frac{1}{1\cdot2\cdot3}$，$\frac{1}{1\cdot2\cdot3\cdot4}$等等却是**直接从二项式定理接受过来的**。对于我们的目的来说，问题恰恰在这里，但它与从第二项开始上升的函数的系数 $h,h^2,h^3,h^4,h^5$ 等等总是直接“从二项式定理”导来相比，并不更足以为奇。

至于一般的二项式系数$=\frac{m(m-1)\cdots(m-(r-1))}{1\cdot2\cdots r}$(这是**与** $m$ **个元素无重复地取** $r$ **个的组合形式数目**相同的**表示式**)，象在我们的情况下，若 $r=m$ ($r$ 决不能比 $m$ 大)，它就变为

$$=\frac{m(m-1)\cdots(m-(m-1))}{1\cdot2\cdots m}=\frac{m(m-1)(m-2)\cdots2\cdot1}{1\cdot2\cdots m},$$

这是通过组合理论得到证明的，而具有正整数乘幂的二项式定理本身不过是组合理论的一个特殊应用(**组合形式的数目用排列形式的数目去除**)而已。

但对于 $x_1=x+h$，$f(x_1)=f(x+h)$等等的微分方法来说重要的是，凡通过二项式定理给出的，就表现为由微分演算本身导来的。

但是这种推导本身，象我们在用微分从 $x^m$ 导出 $mx^{m-1}$ 等等时所看到的那样，又只能在二项式定理的基础上进行。

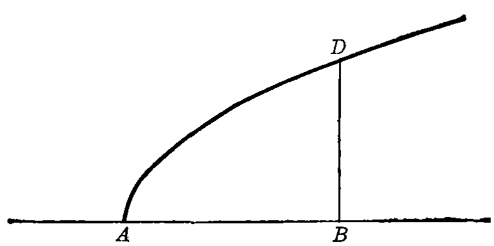
---

# [七十年代的几份手稿]

## 《曲边形求积》

(摘自 1669 年**牛顿**向伦敦皇家协会会长提出的论著《**运用无限多项方程的分析**》。)

所讨论的是形如 $y=ax^{m/n}$ 的(简单)曲线(其中例如对于抛物线 $a=p^{1/2}$ 而 $x^{m/n}=x^{1/2}$)。




$AB=x, BD=y, a, b, c$ 为给定的量,$m, n$ 为整数。

如果 $y=ax^{m/n}$,那末

$$\text{面积 } ABD=\frac{an}{m+n}x^{\frac{m+n}{n}}。$$

(这是不加证明地提出的。)

作为例子,他给出:

1) $x^2=1\cdot x^{2/1}=y$; $a=1, m=2, n=1$; 因此

$$\text{面积 } ABD=\frac{1}{3}x^3。$$

2) 如果 $4\sqrt{x}=4x^{1/2}=y$ $(a=4, m=1, n=2)$,那末

$$\text{面积 } ABD=\frac{8}{3}x^{3/2}。$$

他在这样列举时,无论对于一般的定理,抑或对于例子的任何说明,都没

有给以证明。

如果我们看**第一个例子**：$y=x^2$；那末面积元 $=ydx=x^2dx$。

因此

$$面积=\int x^2dx=\frac{x^{2+1}}{2+1}=\frac{1}{3}x^3。$$

**第二个例子**：$y=4x^{1/2}$；面积元 $=ydx=4x^{1/2}dx$。因此

$$面积=\int 4x^{1/2}dx=\frac{4x^{\frac{1}{2}+1}}{3/2}=\frac{8}{3}x^{3/2}。$$

一般而论：如果

$$y=ax^{m/n},$$

那末

$$面积元=ydx=ax^{m/n}dx;$$

因此

$$面积=\int ax^{m/n}dx;$$

$$\int ax^{m/n}dx=\frac{ax^{\frac{m}{n}+1}}{\frac{m}{n}+1}=\frac{ax^{\frac{m+n}{n}}}{\frac{m+n}{n}}=\frac{an}{m+n}x^{\frac{m+n}{n}}。$$

牛顿从解析几何已经知道，抛物线等等的面积元 $=ydx$，即等于所要求的曲边形面积的微分；由于 $y$ 根据曲线方程 $=ax^{m/n}$，所以这个微分 $=ax^{m/n}dx$，即等于 $y$ 的横坐标表示式乘以 $dx$。再者，他从同一个来源知道，面积可以看作这些面积元的无限累加，即看作

$$\int ax^{m/n}dx。$$

他还知道，对于 $x^m dx$（其中 $m$ 可以是任何一个$\frac{m}{n}$，只要 $m$ 和 $n$ 都是整数；例如 $n=1$，则 $x^{m/n}=x^m$），其结果为 $\frac{x^{m+1}}{m+1}$，这就给出

$$\int ax^m dx=a\frac{x^{m+1}}{m+1}=\frac{1}{m+1}ax^{m+1}。$$

此外，例如对 $x^m$ 的微分运算曾告诉他，其微分为 $mx^{m-1}dx$，也就是二项式定理的第二项，而它将重又变为第一项，也就是按下面的公式积分得

$$\frac{mx^{m-1+1}dx}{(m-1+1)dx}=x^m。$$

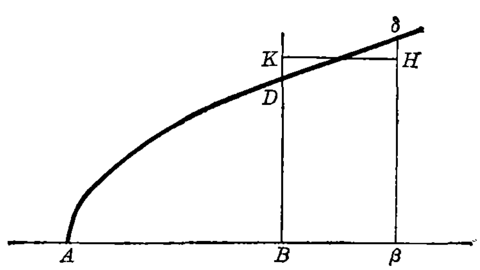
如果他知道了这个公式，那末他就会从解析几何知道，它是 $ydx$ 的积分，也就

是曲线方程中被微分过的 $x$ 的函数的积分。但是他完全没有能够把微积分应用到解析几何上去，他对一般定理的下述证明就说明了这一点。在这个证明中：1)辅助▭，或者不如说辅助梯形，不是由 $dx$ 和 $y+dy$ 所构成，而是由 $dx$ 和某一个高所构成，这个高不是纵坐标，因而当 $dx$ 消失时，它也不会消失；2)他不是从方程 $y=$ etc.，而是在假定面积为**已知的**情况下用几何方法来作曲线；3)利用对 $x$ 的已知函数的微分找出高度 $y$，然后反过来推断：如果现在 $y$ 这个高度 = 如此这般，那末当 $y$ 用这种表示式给定时，面积就反过来必须是这样那样。但是他却回避去真正进行积分，或者去指出如何用演算来实现这个逆过程；4)否则他也就会发现，这个公式并非对于所有简单的曲线都已够用，而还要 $+C$，这个常量在许多情况下并不是 0，而是必须加以确定的。我们将在下一页如实地给出他的所谓证明。

必须注意：在我们写作 $dx, dx^2$ 等等的地方，他都写作 $o, o^2$ 等等。

1）**证明的准备**。所要计算的面积 $ABD=z$；$AB=x$；$BD=y$；$B\beta=dx$；$BK=v$；▭$B\beta KH(vdx)=$ 面积 $B\beta\delta D$

$$A\beta=x+dx;\qquad A\delta\beta=z+vdx。$$




从 $x$ 和 $z$ 之间任意假定的关系出发**我按下列方式求** $y$：

A）设 $z=\frac{2}{3}x^{3/2}$ 或 $z^2=\frac{4}{9}x^3$。然后用 $x+dx$ 代替 $x$，$z+vdx$ 代替 $z$，这样就得到：

B）$\frac{4}{9}(x+dx)^3=(z+vdx)^2$；因此：

$$\frac{4}{9}(x^3+3x^2dx+3xdx^2+dx^3)=z^2+2zvdx+v^2dx^2。$$

$\frac{4}{9}x^3 = z^2$；消去这两项并把剩下的用 $dx$ 去除，就得到

$$\frac{4}{9}(3x^2 + 3xdx + dx^2) = 2zv + v^2dx。$$

如果我们现在假定 $B\beta = dx$ 无限地减少并最后消失或者实际上变为 0（在这之前他早就称它为 $o$），那末与 $o$ 相乘的那些项就此消失，而我们得到：

C) $\frac{4}{9} \cdot 3x^2 = 2zv$，或 $\frac{4}{3}x^2 = 2zv$；因而

$$\frac{2}{3}x^2 = zv。$$

但现在 $v$（为什么？$v$ 又不是纵坐标，$dx$ 又不是 $v$ 的函数？）当 $dx$ 变为 0 时就变为 $= y$，因此

$$\frac{2}{3}x^2 = zv = zy = \frac{2}{3}x^{3/2}y。$$

但若 $\frac{2}{3}x^2 = \frac{2}{3}x^{3/2}y$，则

$$y = \frac{\frac{2}{3}x^2}{\frac{2}{3}x^{3/2}} = \frac{x^2}{x^{3/2}} = x^{2-\frac{3}{2}} = x^{1/2}。$$

［所以当我们 1）不是从作图得出而是令微分矩形 $= ydx$，然后通过对 $\frac{4}{9}x^3$ 的微分从所假定的面积表示式找到了 $vz = \frac{2}{3}x^2$ 的时候，我们就找到了 $y$。］①

现在**牛顿**补充说——这一点应该证明他已经懂得**积分**：

“**因此**（因何？）如果反过来 $x^{1/2} = y$（就是说，如果我们有了曲线方程），那末面积 $z = \frac{2}{3}x^{3/2}$”。

所以这个过程是这样：如果给定了面积方程，其中面积是用横坐标表示的，那末我利用微分就会找到**曲线方程**，例如 $y = x^{1/2}$（因而 $y = 1 \cdot x^{1/2}$，如果我把 1 换成 $a$ 或者其他给定的数，并把 $\frac{1}{2}$ 换成 $\frac{m}{n}$，那末 $y = ax^{m/n}$），也就是会找到用横坐标 $x$ 来确定纵坐标 $y$。然后我推论出，反过来也必定可以从**曲线方程**——也就是从被积函数——求出这个面积，而且这个面积必须就是我由之求出

$$y = ax^{m/n}$$

① 这对括号是在马克思的手稿里原来就有的。

的那个面积。这样做时，我还根本没有进行过**积分**，就象原先在导出面积 $ABD$ 的微分时根本没有直接对曲线方程——实际上它是我反过来作为结果求得的——进行过微分。此外，要是牛顿认为这种他称为“证明的准备”的做法本身就已经是证明的话，那他就不会跟着再来一个**证明**了。但是这个所谓的**证明**仅仅在于用**一般代数的**形式来代替**确定的**数，即用 $z=\frac{n}{m+n}ax^{\frac{m+n}{n}}$ 来代替 $z=\frac{2}{3}x^{3/2}$ 而把同一件事重复一遍罢了。

就请看下面的**证明**吧。

**证明**。如果现在**一般地** $z=\frac{n}{m+n}ax^{\frac{m+n}{n}}$（以代替 $=\frac{2}{3}x^{3/2}$；这是代替数字的一般代数表示式）或者（它仅仅是一种代数的简化，用比较简单的表示式来置换，而以后可以再用它们的值代回去），**或者** $\frac{an}{m+n}=c$ 以及 $m+n=p$，如果（通过这个置换，它并不提供新的证明论据，而只是把**假定**用另一种较为简单的符号**重复一遍**）$z=cx^{p/n}$ 或 $z^n=c^nx^p$，那末当（完全象以前在准备中那样）用 $x+dx$ [象上面那样他写做 $x+o$]① 代替 $x$，并用 $z+vdx$（象上面那样他写做 $z+vo$）或（以后会证明是同样的）$z+ydx$ 代替 $z$ 时[这里的改进在于，不用啰嗦就直接把面积元 ▭ 作为 $ydx$ 确定了下来]②，

$$c^n(x^p+px^{p-1}dx+\cdots)=z^n+nz^{n-1}ydx+\cdots。$$

$c^nx^p$ 与 $z^n$ 相抵销，而其余的用 $dx$ 去除以后，就得

$$c^npx^{p-1}=nz^{n-1}y=\frac{ynz^n}{z}=\frac{ync^nx^p}{cx^{p/n}};$$

或除以 $c^nx^p$：

$$\frac{c^npx^{p-1}}{c^nx^p}=\frac{ync^nx^p}{cx^{p/n}c^nx^p},$$

$px^{-1}=\frac{yn}{cx^{p/n}}$；由此得出

$$ny=px^{-1}cx^{p/n}=pcx^{\frac{p}{n}-1}=pcx^{\frac{p-n}{n}}。$$

如果再把 $p$ 和 $c$ 的值代入，那末最后得到：$y=ax^{m/n}$。（即这时：

$$ny=(m+n)\frac{an}{m+n}x^{\frac{m+n-n}{n}},\ ny=anx^{m/n},\ y=ax^{m/n}。)$$

因此，如果**反过来**令 $y=ax^{m/n}$（也就是说曲线方程已给定），那末就得 $z$ 的值 $\frac{m}{m+n}ax^{\frac{m+n}{n}}$，证毕。这就是说，并没有证明出什么来，而只是用一般的代数形式把“证明的准备”**重复**了一遍而已。

---

①② 这对括号是在马克思的手稿里原来就有的。

## [用$\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$，$\frac{d^3y}{dx^3}$等等符号代替$\frac{0}{0}$]

比值

$$\frac{f(x+h)-f(x)}{h}\text{或}\frac{f(x+h)-f(x)}{x_1-x}\text{或}\frac{y_1-y}{x_1-x}\text{或}\frac{\Delta y}{\Delta x}$$

表明：$f(x)$原来的量与其增长了的量$f(x+h)$之间的差值，或者$x$的函数即$f(x)$的增长率与$f$就是其函数的变量$x$的增长率之间的关系。

这是$x$的函数的差值与变量$x$本身的差值之间的比值。

在**分子**中，我们有$x$**的函数**之间的**差值**；在**分母**中，有变量$x$本身原来的量与增长后的量之间的**差值**；分母中是$x$变化的量度，分子中是其函数变化的量度。

**$\Delta y$ 是 $y$ 的一阶差值，而 $\Delta x$ 是 $x$ 的一阶差值。**

如果$\Delta x$变为$=0$，那末$\Delta y$也变为$=0$，因为$y$只是由于$x$变成了$=x+\Delta x$而变为$=y_1$的。

我们在第一个和第二个表示式

$$\frac{f(x+h)-f(x)}{h}=\frac{f(x+h)-f(x)}{x_1-x}=\frac{y_1-y}{x_1-x}=\frac{\Delta y}{\Delta x}$$

中看到：如果$h$变为0，我们就有

$$\frac{f(x+0)-f(x)}{0}=\frac{f(x)-f(x)}{0}=\frac{0}{0}。$$

至于一旦$\Delta x$或$h$变为0，$y_1-y$或$\Delta y$也要变为$=0$这一点，那是显而易见的。

因为

$$x_1-x=(x+\Delta x)-x=\Delta x,$$

所以一旦$\Delta x$变为0，$\Delta y$就要变为0。

因此很明显，$\Delta y$或$y_1-y$在这里不但要变为0，而且只是由于$\Delta x$的变为0，或者$x_1$和$x$的变为相等即$x_1=x$才变为0的；由于$x_1-x=\Delta x$，即$(x+\Delta x)-x=\Delta x$，所以当$\Delta x$变为0时，左边只能变为0，或者$x+\Delta x$变为$=x$。

因此，即使在$\Delta y$的消失中函数$y$还保持着对其自变量$x$的依赖关系，而且$\Delta y$的最后转变为0，它的最后消失本身就是$\Delta x$这个变量$x$的增量消失的

结果；函数 $y$ 对变量 $x$ 的依赖关系一直保持到化为零。但是在$\frac{0}{0}$这个表示式中，函数 $y$ 与 $y$ 就是其函数的变量 $x$ 之间的这个**质的关系**也同样**消失了**。分子和分母之间，变量的函数和变量本身之间的**质的**差异，其任何痕迹都已在表示式$\frac{0}{0}$中一笔勾销。

因此，为了表明$\frac{0}{0}$的起源和意义，我们用 $dx$ 来代替消失着的 $\Delta x$，这样一来，消失着的 $\Delta y$ 就自行变为 $dy$。

所以，$\frac{dy}{dx}$不仅是$\frac{0}{0}$的一个符号，同时也是在原始等式的确定的已知条件下$\frac{0}{0}$所由产生的这一**过程**的符号；而且它表明：$\Delta y$ 的变为 0 产生于函数 $y$ 对变量 $x$ 的质的关系，因此 $\Delta y$ 的转变为 $dy$ 是 $\Delta x$ 转变为 $dx$ 的结果，而这是$\frac{0}{0}$所无法表明的。所以这个**质的关系**就这样在否定中被固定了下来，而转变就是这样的否定。相反，在$\frac{0}{0}$中表示**不出什么东西消失了**；所表示的只是**量的**方面，即分子消失了，分母同样消失了，因而比值本身也消失了；但是**所存在着的一个质的关系**，即分子中的 0 只是分母中的 0 的一个**结果**，因而本身是函数对其变量的依赖关系的一种表示这个**质的关系**，并没有被表达出来。完全正确，$\frac{0}{0}$可以表示任何一个量，但是 $x$ 同样也可以表示任何一个量；无论是$\frac{0}{0}$的或是 $x$ 的特殊意义，任何时候都依赖于确定的条件，或者依赖于$\frac{0}{0}$或 $x$ 在其中出现的那个函数，以及依赖于无论是$\frac{0}{0}$所由产生的或者是 $x$ 所由变化的那些确定的条件。

但是，我们之所以会被引导到用符号$\frac{dy}{dx}$来表示$\frac{\Delta y}{\Delta x}$的转变为$\frac{0}{0}$，不但由于研究了$\frac{0}{0}$的演化过程，而且还由于从原始等式所得出的那个结果。这个结果就是$\frac{0}{0}=f'(x)$，而不是$\frac{0}{0}=0$或者任何其他的一个任意实在值。

这就是

1) $\frac{0}{0}(1)=f'(x)$，　　2) $\frac{0}{0}(2)=\frac{1}{2}f''(x)$，

3) $\frac{0}{0}(3)=\frac{1}{2\cdot 3}f'''(x)$，　　4) $\frac{0}{0}(4)=\frac{1}{2\cdot 3\cdot 4}f^{\mathrm{IV}}(x)$

等等，等等。

因此我们看到，$\frac{0}{0}$的第一个实在内容等于 $x$ **的一阶导函数，或**$f'(x)$，而$\frac{0}{0}$的其他内容或实在值，则都由变量 $x$ 的确定的、从 $x$ 的原函数导出的、依次按一定规律一个接着一个产生出来的那些函数所组成。

［我们一开始就能说出什么时候$\frac{0}{0}$会变为 $=0$，从而无论是微系数或者是极限都将消失。在过程中变量本身一旦消失，或者变为等于一个常量，那就是这种情况。

例如，如果我们有

$$y=f.x,$$
$$y_1=f.(x+h),$$

那末

$$y_1=f.x+h;$$

因为当 $h$ 变为 0 时，$y_1$ 在这里就变为 $=f.x=y$。

因此

$$y_1-y \text{ 或 } \Delta y=h; \qquad \frac{y_1-y}{x_1-x}=\frac{f.(x+h)-f.x}{x_1-x}=\frac{\Delta y}{\Delta x}=1,$$

即 $f.\ \frac{dy}{dx}$。

因此$\frac{0}{0}=1$，变量 $x$ 在这里已不见。

如果我们用 $\frac{dy}{dx}$ 代替 $\frac{0}{0}$，那末 $\frac{dy}{dx}=1$；因此 $dy=dx$；从而 $1=1$（因为 $\frac{dy}{dx}=\frac{a}{a}=1$）。

$\frac{0}{0}$在这里变成了等于一个常量……

假定我们对$\frac{dy}{dx}$进行第二次微分；由于常量的增量 $=0$，所以$\frac{d}{dx}\left(\frac{dy}{dx}\right)=\frac{d^2y}{dx^2}=0$ 或 $\frac{0}{0}=0$。］①

因此，只有当人们用$\frac{dy}{dx}$把$\frac{0}{0}$中的质的关系固定下来（符号化）的时候，人们才有可能去固定那些和第一个$\frac{0}{0}$或$\frac{0}{0}$(1)不同的、但与之相联系的、并从它

① 这对括号是在马克思的手稿里原来就有的。

相继产生的$\frac{0}{0}(2)$，$\frac{0}{0}(3)$，$\frac{0}{0}(4)$；使它们成为彼此有规律地联系着的产生它们的过程的符号，并且由它们自己通过其中出现它们的不同实在值的右边说出它们之间的这个联系，而这些实在值彼此处于确定的关系之中，并且都或近或远地来源于 $x$ 的原函数以及这原函数尚存在其中的第一个等式 A)。

---

A) **原始等式**：

$$f(x+h) \text{ 或 } y_1 = f(x) \text{ (或 } y) + f'(x)\,h + \frac{1}{2}f''(x)\,h^2 + $$
$$+ \frac{1}{2\cdot 3}f'''(x)\,h^3 + \frac{1}{2\cdot 3\cdot 4}f^{\mathrm{IV}}(x)h^4 + \cdots$$

首先发生的是，我们从 $f(x+h)$减去 $f(x)$ 或者从 $y_1$ 减去 $y$。于是只要 $h>0$，我们就得到

$$\Delta y = f'(x)h + \frac{1}{2}f''(x)\,h^2 + \frac{1}{2\cdot 3}f'''(x)\,h^3 + \frac{1}{2\cdot 3\cdot 4}f^{\mathrm{IV}}(x)h^4 + \cdots$$

然后为了要找出 $f(x+h)$ 与 $f(x)$ 之间的差值对于变量 $x$ 的增量的比值(由于 $x_1 - x = h$)，两边都用 $h$ 去除，我们就得到了

$$\frac{\Delta y}{\Delta x} = f'(x) + \frac{1}{2}f''(x)\,h + \frac{1}{2\cdot 3}f'''(x)\,h^2 + \frac{1}{2\cdot 3\cdot 4}f^{\mathrm{IV}}(x)h^3 + \cdots$$

为获得比值$\frac{\Delta y}{\Delta x}$所必需的这个用 $h$ 去除，把**原始等式的第二项**和这新等式的第一项，实际上就是只有以 $h$ 的一次幂作为因子的那一项，从 $h$ 那里解放了出来，但同时也把原始等式的所有其他的项作了修改，使它们每一个都**附有一个降低了 1 次幂的** $h$；所以，例如第三项就附有 $h^{3-1}$ 或 $h^2$，而不是附有 $h^3$，等等。然而重要的是要牢记，把原始等式的第二项从因子 $h$ 那里解放出来的那个过程，也把所有其他的项**修改了**。

其次，我们从出现 $\Delta y$ 的(也就是由于 $\Delta y = f(x+h) - f(x)$而**还有** $x$ **的原函数参与其中的**)等式看到，$h$ 的量越是减少，那末随后各项的每一项与它前面一项相比就显得越小，以致 $f'(x)$ 仅附有 $h$ 一次幂的那第一项 $f'(x)h$ 表示 $y_1$ 与 $y$ 之间的**最大差值**，而且 $h$ 变得越小，这第一项就越加变为其总体由 $A$ 或 $\Delta y$ 的级数所表示的**那些部分差值的总和的极限**，或者说 $\Delta y$ 比$f'(x)h >$得越少。

---

# [代数中表示式$\frac{0}{0}$和微分演算中表示式$\frac{dy}{dx}$之间质的差异]

$\frac{0}{0}$，无论是$\left(\frac{0}{0}\right)_1$，$\left(\frac{0}{0}\right)_2$等等，亦即无论是我们在第一次微分时得到的$\left(\frac{0}{0}\right)$，以及作为逐次微分的结果出现的$\left(\frac{0}{0}\right)_1$等等，凡是在$x$表现为自变量和$y$表现为因变量的地方，因而在它从$\frac{\Delta y}{\Delta x}$中产生出来并化为$\frac{dy}{dx}$的地方，在质上不同于象在普通代数中那样$x$是一个常量时我们所得到的$\frac{0}{0}$。

例如，假使我们有$\frac{x^2-a^2}{x-a}$——因而它也可以写成$=\frac{(x-a)(x+a)}{x-a}$；如果令$x=a$，那末$x^2-a^2=a^2-a^2$以及$x-a=a-a$；$\therefore\frac{x^2-a^2}{x-a}=\frac{0}{0}$；但它不是因为分母$=0$而变为$\frac{0}{0}$，而是因为一当我们把$x$和$x^2$用它们的值$a$和$a^2$代入时，分子和分母将同时变为0……

与此相反，在

$$\frac{f(x+h)-f(x)}{x_1-x}=\frac{y_1-y}{x_1-x}$$

中分子的值则决定于并且依赖于分母的值。虽然我可以说：如果分子中的$h$变为等于0，那末

$$f(x+h)-f(x)=f(x)-f(x),$$

因而$=0$；但是我在这里只有当$x_1-x=h=0$，亦即只有当$x_1-x$变为$=x-x=0$时才能令$h=0$。

分子只有当分母改变其量时，才改变它的量……

在$f(x+h)-f(x)$中则不同，在$x_1-x$没有事先变为$=0$或$x_1=x$之前，我就不能令$h=0$。

所以在分子不是变量$x$的函数而分子和分母中的$x$都是同一个常量，即使它是未知的和还未确定的但总是一个常量的地方，$dx$和$dy$之间这种质的

关系是不存在的。

再者，$\frac{x^2-a^2}{x-a}=\frac{(x-a)(x+a)}{x-a}=x+a$。所以当 $x=a$ 时，$\frac{x^2-a^2}{x-a}=2a$。

这个 $2a$ 不是 $\frac{x^2-a^2}{x-a}$ 象 $\frac{dy}{dx}=m$ 那样意义下的极限，$2a$ 是分式 $\frac{x^2-a^2}{x-a}$ 单纯通过除法而得到的实际值。它仅仅是这个意义下的极限，正象任何一个比数的实际值是它的极限一样。

所以，$\frac{6}{3}=2$。$\frac{6}{3}$ 既不 $>2$，也不 $<2$，在这个意义下每一个等式都表示某个极限，而且甚至对于每一个常量如 3 等等，它的极限和它的存在是同时发生的。3 不是 2，不是 4，也不是 2 与 4 之间的一个分数，而是 $=2+1$ 或 $4-1$。

$\frac{1}{3}$ 本身是它自己的极限。如果我把它表示成级数，那末：

$$\begin{array}{r|l} 1 & \underline{3} \\ 10 & 0.33 \\ \underline{-\ 9} & \\ 10 & \end{array}\text{，所以}\frac{1}{3}=\frac{3}{10}+\frac{3}{100}+\frac{3}{1000}+\frac{3}{10000}+\cdots$$

在这种情况下，$\frac{1}{3}$ 变成它的无穷级数的极限。

---

$$y=x^3, \qquad y_1=x^3+3x^2h+3xh^2+h^3,$$

$$y_1=(x+h)^3, \qquad y_1-y=3x^2h+3xh^2+h^3,$$

$$\therefore \quad \frac{y_1-y}{h}=3x^2+3xh+h^2;$$

如果 $h$ 减少到 0，那末 $\frac{y_1-y}{h}=3x^2$；所以 $3x^2$ 是 $\frac{y_1-y}{h}$ 当 $h$ 逐渐减少时所趋向的极限；但这时就会 $\frac{y_1-y}{h}=\frac{0}{0}$，或

$$\frac{\text{函数 } y \text{ 的增量}}{\text{变量 } x \text{ 的增量}}=\frac{0}{0};$$

因此 $\frac{0}{0}=3x^2$。

这与普通代数还符合，因为 $\frac{0}{0}$ 可以 $=$ 任何一个量。但由于在 $\frac{0}{0}$ 中无论函数的或变量 $x$ 的任何痕迹都已不见，所以用表示式 $\frac{dy}{dx}$ 代替 $\frac{0}{0}$，就使我们回忆

起函数是 $y$ 而变量是 $x$；$dy$ 和 $dx$ 是消失的量；$\frac{dy}{dx}=3x^2$；$\frac{dy}{dx}$或者不如说它的值 $3x^2$ 是函数 $y$ 的微系数。

---

引进消失的量以及它们的比值$\frac{dy}{dx}$来代替$\frac{0}{0}$，已经不属于代数的范围；而更坏的是：虽然 $\frac{dy}{dx}$“是代表极限 $3x^2$ 的符号”，因而“$dx$ 本来应该总是写在 $dy$ 的下面”①，但是“为了便于进行代数运算”，我们把$\frac{dy}{dx}$当作一个普通分式，把$\frac{dy}{dx}=3x^2$ 作为一个普通等式来处理；这样，在消除了这等式的分母之后得到的 $dy=3x^2dx$ 这个——以相当模棱两可的方式得到的——表示式，就称为函数 $y$ 的微分。

---

① 这句话出自布夏拉的微积分教科书。

# [两种不同的微分方法]

## [第一种方法：极限方法]

1）在基础或出发点，也就是在**求微系数**方面，我们最初有这样的方法：

$$\text{I)}\ y=f(x);$$

$$\text{II)}\ y_1=f(x+h)。$$

例如当

$$y=f(x)=ax^2$$

时，那末

$$y_1=f(x+h)=a(x+h)^2;$$

由此

$$y_1=ax^2+2ahx+ah^2$$

以及

$$\text{III)}\ y_1-y=2ahx+ah^2。$$

用 $h$ 除两边，我们得到：

$$\text{IV)}\ \frac{y_1-y}{h}=2ax+ah。$$

$y_1-y$ 等于 $y_1$ 的差值（$y_1$ 对 $y$ 的增量），所以 $=\Delta y$；

$x+h=x_1$；$h$ 等于 $x_1$ 的差值，等于 $\Delta x$（即 $=x_1$ 对 $x$ 的超额部分）；因此我们可以不写成 $\frac{y_1-y}{h}=2ax+ah$ 而写成

$$\text{V)}\ \frac{\Delta y}{\Delta x}=2ax+ah。$$

这等式的左边表示**函数 $y$ 的有限差值对自变量 $x$ 的有限差值的比值**。如果我们现在令 $h=0$，那末就得到：

$$\text{VI)}\ \frac{0}{0}=2ax。$$

这还在普通代数的范围之内。如果我们例如有 $\frac{x^2-a^2}{x-a}$，那末它 $=\frac{(x+a)(x-a)}{x-a}$。因此 $\frac{x^2-a^2}{x-a}=x+a$，因为在右边

$$(x+a)\cdot\frac{(x-a)}{(x-a)}=(x+a)\cdot 1。$$

如果我们两边都令 $x=a$，那末我们得到$\frac{0}{0}=2a$。由于$\Delta x=h$，所以如果 $h=0$，那末 $\Delta x=0$；并且由于 $y$ 只是因为 $x$ 增长了 $h$ 而变为 $y_1$ 的，所以当 $h=0$ 即 $x+h=x$ 时 $y_1=y$；因此 $\Delta x=0$，$\Delta y=0$。在这种形式下对此等式是无所事事的，它既不含有函数的任何痕迹，也不含有主要的变量，$\frac{0}{0}$表明两个差值 $\Delta y$ 和 $\Delta x$都已消失，但是我们要把已经消失了的因子的特性固定下来；我们要把它们作为消失量(在所要否定的那个东西的特性的否定中)固定下来，因此用 $\frac{dy}{dx}$ 来代替$\frac{0}{0}$；我们不用 $\Delta$ 而用 $d$，不用差值而用它的最小值即微分。所以：

$$\text{VII)}\ \frac{dy}{dx}=2ax。$$

这首先表明，组成比值的这些项是消失量，而一旦我们有了 $2ax$，则表明它们确实已经消失或者已经变为$\frac{0}{0}$。

所以 $2ax$ 是它们变化的极限。

这个微系数因此有两个表示方式，一个表示运动 $\frac{dy}{dx}$，另一个表示它的值即它的**极限**。

运算完成之后所消失的是$\left\{\frac{dy}{dx}\right.=0$；所以如果不把它们去掉的话，那末就将是计算中的一个错误。

所以唯一的困难是把消失量之间比值加以固定的这种辩证观念，而当它**完成了自己的任务的**时候，**比值** $\left(\frac{0}{0}\right)$**也就在演算的结果中消失。**

---

## [第二种方法：拉格朗日方法]

这种方法的重要是为了以后的解析运算，而不是为了基本运算。因为在我们有了

$$u_1=f(x+h)=u+Ah+Bh^2+Ch^3+\cdots$$

之后，那末当令 $h=0$ 的时候，运算方才变为$\frac{u_1-u}{h}=\frac{du}{dx}=A$；如果我们不令它 $=0$，那末为了丢掉 $Bh^2+Ch^3+\text{etc.}$ 就须假定这些项与 $A$ 相比小得消失，而

在此地这是完全多余的想象。但重要的是，当给定 $u=f(x)$ 并且 $x$ 变为 $x+h$，$u$ 变为 $u_1$ 时，就有

$$u_1-u=f(x+h)-f(x),$$

而这个**必须依赖于 $h$ 即 $x$ 的增量的、$x+h$ 和 $x$ 的函数之间的差值**即 $u_1-u$，可以用形如 $Ah+Bh^2+Ch^3+\text{etc.}$ 的级数来表示，因而

$$u_1=u+Ah+Bh^2+Ch^3+\cdots,$$

或

$$u_1=f(x)+Ah+Bh^2+Ch^3+\cdots,$$

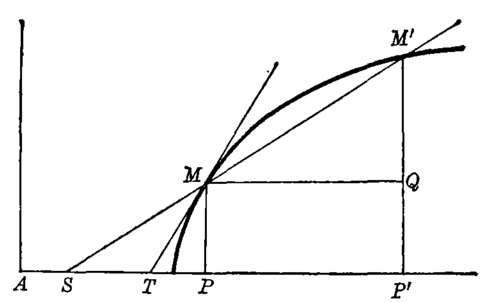
其中 $h$ **的幂次是上升的**；再则**这些幂次只是正的指数**(而且是整数，不是分数)；第一项必须是 $u=f(x)$；$A$ **这个 $h$ 一次幂的系数是一阶微系数**，**而** $Ah$ $(=Adx)$**是** $u_1$ **和** $u$ **之间**或 $f(x+h)$ 和 $f(x)$**之间的差值的第一项。**

微分演算的主要目的**在于求系数** $A,B,C$ **等等的值**；所以**微分**是 $f(x+h)$ **展开的第二项。**

---

## ［切线问题：两种不同的解法］

［这第二种应用了泰勒定理的方法，比第一种更加复杂而且更加令人厌烦；它似乎回避了这样一个困难，那就是要使弧 $MM'$ 与弦 $MM'$ 重合起来，因而用一个一边事实上是一段弧的三角形来代替；然而很清楚，当横坐标的




增量减小时，$M'$ 就趋近于 $M$，直到纵坐标 $M'P'$ 与 $MP$ 重合为止，因而割线 $MM'$ **也不过是切线** $TM$ **的延伸**，所以 $SP$ 也就与次切线 $PT$ 相重合。另一

方面，这种回避仅仅是表面的；因为整个奥妙只是在于两个三角形的相似，而由于辅助三角形的两边由 $dx$ 和 $dy$ 构成，它们比点还小，所以在这种情况下也就可以毫无顾虑地把弦与弧，或者反过来把弧与弦重合起来。此外，即使在第一种方法中也只是把两条直角边相互作了比较，而对于斜边的特性可以让幻想来臆测。]①

---

① 这对括号是在马克思的手稿里原来就有的。

# 关于函数概念

A）如果一个问题根本可以确定的话，那末为了确定它就需要有同所要确定的**未知量**个数**一样多的等式**。因此所有的问题，只要其中给出的等式同未知量个数一样多，那就都属于**确定分析**的范围。

相反，如果一个问题没有提供同未知量个数一样多的等式，那末其中必定有几个仍然是不确定的，并且［在即将提到的界限内］①**可任意地**由我们来确定；因此这类问题称为**不确定的**，而且构成代数学一个特殊部分即**不定分析**的对象。

由于在这样的情况下，一个或几个未知量可以取任意的值，所以它们就容许有不同的解。另一方面，这类问题的提出通常是附有条件的，即所要找的数是**正整数**或者至少是**有理数**；这样一来，可能的解的数目就往往减到只有很少几个；有时它的数目可以无限，但难以得到；有时则不可能有解。

a）**要找其和 $=10$ 的两个正整数。**

这样我们就有如下一个问题：

$$1)\ x+y=10, \qquad 2)\ x=10-y,$$

其中 $y$ 只确定到这个程度，它必须是一个**正整数**。

若我们令 $y=10$，则我们得到 $x=10-10=0$；但 $x=0$ 要除外，因为 $x$ 同样应当是一个正整数。于是就排除了 $y=10$ 这一确定；我们用试验方式可以取的 $y$ 的数值，因而只能是 1—9 。这就是 $y$ 可能的量的界限，它是由问题本身所给定的。

另一方面，从最初的假定：$y=10$，我们已经看到，$x$ 将因此而变为 0，这是要除外的。

同样，若我们取 $y=11$，则

$$x=10-11=-1,$$

就违反了数必须是**正的**这个条件。

但是这两个由于所提的问题而被除外的假定已经表明，**$x$ 的值依赖于 $y$ 的值并且随之而变化**；因为当 $y=10$ 时，势将 $x=0$，当 $y=11$ 时，势将

① 这对括号是在马克思的手稿里原来就有的。

$x=-1$。而且对等式的进一步处理，也给我们表明了这一点。

$y$ **可能的值**$=1,2,3,4,5,6,7,8,9$。而这时 $x$ 相应的值$=9,8,7,6,5,4,3,2,1$，因为 $x=10-y$，所以如果 $y=1$，那末 $x=10-1=9$ 等等。

**因此在这里，未知量 $x$ 的值依赖于未知量 $y$ 的值，并且随着 $y$ 所取的任意的、但在由问题所确定的界限 1—9 内的值而变化**。正是根据不定方程中的这个关系，使得一个**这样的未知量**，如这里的 $x$，最初称为 $y$〔另一个不依赖于 $x$ 而相继取不同值的未知量〕① **的函数**。这是第一个在普通代数学范围内所提出的把**一个未知量表征为另一个未知量的函数**的动机。这时立即可以一开始就把一些象 $a,b,c$ 那样的量，例如上述等式 $x+y=10$ 中的 10 置之不顾，而只把 $x$ 确定为 $y$ 的函数，即它所依赖的**那个未知量的函数**，因为 $a$ 或 10 都**已经确定**而且在等式的每一个可能的解中保持**同一个值**。

$f(y)$ 或 $x$ 依赖于**未知量** $y$ **的变化**而改变其值，$f(y)$ 或 $x$ 是 $y$ 的函数。

但是 $y$ **的变化**本身只在于此，即在一定的界限内可以**任意**赋予它不同的数值，例如上面 9 个不同的值。若我们赋予它数 9，则 $x=10-9=1$；如果 8，那末 $x=10-8=2$，等等。$y$ 的每一个这样的任意数值从 1—9，都能解这个方程，因为它给 $x$ 产生一个与 $y$ 所取的值相对应的值；但 $y$ 始终是一个**常量**，不管我们赋予它的是值 1，或 8，或 3；在它自身中并不发生任何使它从 1 到 2 或者从 8 到 9 的变化；因此，它并**不是变量**，虽然它的值可以——在一定界限内——为我们任意改变。在表示式$\frac{m}{m-1}$中也是这样，其中 $m$ **不象** $y$ **那样是一个未知量**，$m$ 的值可以由我们任意**改变**，但 $m$ 决不因此而变成一个**变量**；它只不过是一个**未定的常量**，正是由于这个原因它可以取任意的而且任何一个任意的数值。如果我们令 $m=1$，那末$\frac{m}{m-1}=\frac{1}{0}=\infty$；我们若给 $m$ 以数值 3，那末 $\frac{m}{m-1}=\frac{3}{2}$等等。同样，这里等式 $x=10-y$ 中的**未知量** $y$ 不是因为 10 是常量、$y$ 是变量而与 10 有所区别，而是因为 10 是一个**确定的**常量值，它在等式的任何一个可能的解中总保持为确定的，而 $y$ 是一个**不确定的**、但每次同样也是一个**常量值**，因而在等式的解中可以交替地确定为 1，2 等等。

因此这里的**独立未知量** $y$，象表达式 $\frac{m}{m-1}$ 中的 $m$ 一样，也不是**变量**；它是**不确定的**，犹如 $m$ 与算术上的数值相比是不确定的一样；它和 $m$ 的区别，在于代数表示式中的 $m$：1)不象 $y$ 那样是一个**未知量**；2)不是有另一个**未**

---

① 这对括号是在马克思的手稿里原来就有的。

知量 $x$ 的数值的确定依赖于 $m$ 的数值的确定。相反，如果我们有等式

$$x=\frac{m}{m-1},$$

那末 $x$ 的值势将依赖于我们赋予 $m$ 的不同数值。

这样，人们看到，**函数的概念**就象它原先在**不定分析**中所产生的那样，还只有一种非常局限的、仅仅适用于某几种等式的意义。

如果我们现在回到等式 $x=10-y$ 的解上来，那末

$$y=1,2,3,4,5,6,7,8,9,$$

$$x=9,8,7,6,5,4,3,2,1。$$

$y$ 的最后四个值 $=6,7,8,9$ 为我们给 $x$ 提供了它的相应的最初四个值：4,3,2,1。

所以便有等式

$$6+4=10,\quad 7+3=10,\quad 8+2=10,\quad 9+1=10。$$

但是通过 $y=1,2,3,4$ 和 $x=9,8,7,6$ 也可以为我们提供同样这四个等式。

因此，这个问题事实上只容许五组**不同的解**：

$$y=1,2,3,4,5,\quad x=9,8,7,6,5。$$

譬如若我们有 $x=\frac{y}{y-1}$，则其解不外乎就是 $\frac{m}{m-1}$ 的解；至于 $x$ 必须是正整数这个条件，也许会给问题增加困难，但不会改变等式的性质。

b）如果我们有 2 **个等式而不是**象上面那样只有**一个等式**，那末仅仅在这 2 个等式中包含多于 2 个**未知量**的时候，问题才是**不确定的**。**这类问题**［此地假定等式是一次的］① 在普通的**初等算术课本**中就已出现并且是用所谓**虚假位置规则**来解的。

例如，男人、女人和小孩共 30 人，在饮食店里一起化了 50 个先令；男的每人化 3 个先令，女的每人化 2 个先令，小孩每人化 1 个先令；**问每种各有几人**？

**设男人数** $=p$，**女人数** $=q$，**小孩数** $=r$；那末我们有

1）$p+q+r=30$；　　2）$3p+2q+r=50$；

从这里要在**正整数**范围内找 $p,q,r$。

等式 1)给出 $r=30-p-q$；所以

$$p+q<30;$$

---

① 这对括号是在马克思的手稿里原来就有的。

我们若把 $r$ 的值代入第 2 个等式，则

$$3p+2q+30-p-q=50;$$

从而

$$2p+q+30=50;$$

因此

$$q=20-2p, \qquad p+q=20-p<30。$$

从等式 $q=20-2p$ 得出，当 $p=10$ 时，$q=20-20=0$；所以 $p$ 的值若取得比 10 大，则 $q$ 将变为负；例如，当 $p=11$ 时，$p+q=20-p$ 就变为 $11+q=20-11$，或者 $q=20-22=-2$。这是要除外的。因此对于 $p$ 我们可以取不 $>10$ 的一切数。于是，考虑到 $p+q<30$ 和 $q=20-2p$ [**因而若** $p=0$，**则** $q=20$]①，我们就得到 11 个答案：

$$\begin{aligned}
p&=\ 0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\\
q&=20,\ 18,\ 16,\ 14,\ 12,\ 10,\ 8,\ 6,\ 4,\ 2,\ 0,\\
r&=10,\ 11,\ 12,\ 13,\ 14,\ 15,\ 16,\ 17,\ 18,\ 19,\ 20。
\end{aligned}$$

如果我们把

$$\text{I)}\quad p=0, \qquad q=20, \qquad r=10,$$

和

$$\text{II)}\quad p=10, \qquad q=0, \qquad r=20$$

除外，亦即把第一组和最后一组解（其中 I) $p=0$，而 II) $q=0$）除外，那末就剩下 9 组解。

---

B)1) 因此，**函数**概念原先只限于**不定等式**中的（即在等式个数少于其中出现的未知量个数的那种等式中的）**未知量**，其**值依赖于并因此**随着其他未知量所取的**不同的**[或者是完全任意的，或者是由问题本身所规定的**界限内任意的**]② **值而变化**。

例如

$$y=ax^2+bx+c;$$

这里 $y$ 是 $x$ 的函数；在 $y=axz+bx^2+cz^2$ 中，$y$ 是 $x$ 和 $z$ 的函数。在等式

$$\text{a)}\ x^3+y^3=axy, \qquad \text{b)}\ x^3+y^3+z^3=axz+byz+cxy$$

中，1）就 b）来说，$x$，$y$，$z$ 彼此互为函数；2）就 $y$ 来说，它是 $x$ 的或 $x$ 和 $z$ 的**显函数**，因为当 $x$ 和 $z$ 确定时，它的值也就给定，但在 b）中，它是 $x$ 和 $z$ 的**隐函**

①② 这对括号是在马克思的手稿里原来就有的。

**数**，因为即使它们为已知，就 $y$ 来说也还需要解一个代数方程。

因此，从不定等式中得来的**函数**概念是这样的：如果人们要表示一个**量**，它在没有**给其他量以确定的值**之前**无从确定**，而这些量可能得到的值在同一问题中又有不确定的数目，那末为了表明这种**依赖性**，人们就使用**函数**这个词。

2）**函数**的概念后来**一般化了**；它扩充到**每一个代数未知量**，它与它所依赖的未定量之间的关系可用一个**代数方程**来表示。这种代数未知量就叫做**代数函数**。当代数函数用它们固有的形式表示的时候，它们总是只包含一定数目的项；但是**既约分式**，如我们看到过的那样，只能展开成无穷级数；**函数**概念也推广到它身上，从而开辟了一条通往**超越函数**的道路，如**对数**，就只能通过无限个平方根来表示；**正弦和余弦**，如用它们的弧来表示，也是如此。

C）**进一步的一般化**，就是在几个量之间**不需要等式**而使其中的一个成为其他的**隐函数**；只要它的值依赖于其他量的值就行。例如在圆中，正弦是弧的隐函数，尽管没有一个代数等式能把它表示出来，因为两者之中一个当着另一个确定时也就确定，反之亦然。（这里没有计及半径，因为它并不涉及某条确定的弧。）

D）由于笛卡儿把代数应用到几何上去，也就是通过解析几何或高等几何，**函数**概念获得了新的发展和重要性。未知量 $x$，$y$ 等等变成了**变量**，而已知量变成了**常量**。

一个变量的函数是另一个变量，它的值随着第一个的值的变化而变化，也就是说它**依赖于**第一个变量。它和不定等式中的函数具有这样一个共同点，即当给变量一个特定值时，它的函数就得到一个**确定的相应的值**。

# [拉格朗日方法分析初稿]

## I. 拉格朗日在代数的基础上对泰勒定理的演化(有某些修正)

在同微分演算的结果相比较时，我们会发现，$f'(x)$ 是$\frac{dy}{dx}$的实在等价物，$f''(x)$ 是 $\frac{d^2y}{dx^2}$的实在等价物等等，或者反过来，$\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$等等是实在微系数的微分表示式，即拉格朗日那里 $x$ 的导函数的微分表示式。

---

拉格朗日自己说，以 $dx$ 代替 $h$,$dx^2$ 代替 $h^2$ 等等只是为了记号一致而采取的。

表示式$\frac{dy}{dx}$在这里变为运算符号，通过这种运算就会获得 $f(x+h)$ 展开中 $h$ 的系数；$\frac{d^2y}{dx^2}$，$\frac{d^3y}{dx^3}$等等表明，重复同一过程就会给出 $h$ 其他幂次的系数。因此对于每一个函数，人们只要通过**代数规则**，就能知道$\frac{dy}{dx}$等等该是什么。例如，对于 $x^m$，$\frac{dy}{dx}$ 是什么呢？我们必须把 $(x+h)^m$ 按二项式定理加以展开，这就给出 $x^m+mx^{m-1}h+\text{etc.}$；由于 $\frac{dy}{dx}$表示这展开中 $h$ **一次幂**的系数，所以$\frac{dy}{dx}=mx^{m-1}$ 等等。

因此，整个问题就归结为通过分析的过程来展开代数所能提供的各种不同函数。所以这方法以**所有这些函数的代数展开**为前提，并且这时就会自行得出相应于它们的微分表示式。微分演算中的微分形式——即它们所指明的运算——则反过来用于按自己的捷径去找这些函数。

[拉格朗日也用 $y'$，$y''$ 等等代替$\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$等等来表示导函数。]①

[虽然在拉格朗日的方法中**微分演算的原理**是在与对**极限**、**无限小和消失**

---

① 这对括号是在马克思的手稿里原来就有的。

量的一切考虑无关的情况下证明的，但是一旦他转到**应用**上来，例如确定体积、表面积、曲线的长度，求次切线的表示式等等，他自己却不得不老是逃到这些概念那里去。他的方法要求熟悉他所提出的、把 $x+h$ 的各种函数展开为 $h$ 的正整数升幂的那个解析方法，然而这种展开往往非常困难。况且，泰勒定理和麦克劳林定理一旦确立以后，它们就为展开许多函数提供了十分容易的方法，而用普通代数去展开则要特别费力周旋。]①

---

[我们刚才已经看到，当例如要对 $ax^m$ 展开 $f(x+h)$ 时，只要按二项式定理展开 $a(x+h)^m$ 就已足够：

$$mx^{m-1}ah,\ m(m-1)x^{m-2}\cdot\frac{ah^2}{1\cdot 2},\cdots$$

于是我们知道，$h$ 的系数 $=\frac{dy}{dx}$，$\frac{h^2}{1\cdot 2}$的系数 $=\frac{d^2y}{dx^2}$ 等等。

与此相反，一旦泰勒定理确立以后，**二项式定理**就可以反过来**从它演化出来**，而且更加简单地可以从初等微分运算演化出来。

在例如先证明了

$$d(xy)=xdy+ydx$$

之后，一般就有

$$d(xyztu)=xyztdu+yztudx+ztuxdy+tuxydz+xyzudt。$$

如果我现在用 $xyztu$ 去除两边，那末

$$\frac{d(xyztu)}{xyztu}=\frac{xyztdu}{xyztu}+\frac{yztudx}{yztux}+\frac{ztuxdy}{ztuxy}+\frac{tuxydz}{tuxyz}+\frac{xyzudt}{xyzut}。$$

因此：

$$\frac{d(xyztu)}{xyztu}=\frac{du}{u}+\frac{dx}{x}+\frac{dy}{y}+\frac{dz}{z}+\frac{dt}{t}。$$

如果现在令 $x,y,z,t,u$ 彼此相等，也就是使它们例如都 $=x$，且若它们的数目 $=m$，那末

$$\frac{d(x^m)}{x^m}=\frac{mdx}{x};$$

于是

$$d(x^m)=\frac{mx^m dx}{x}=mx^{m-1}dx。$$

---

① 这对括号是在马克思的手稿里原来就有的。

因此$\frac{d(x^m)}{dx}$或$\frac{dy}{dx}=mx^{m-1}$，并按同样方式得$\frac{d^2y}{dx^2}=m(m-1)x^{m-2}$等等，这样就演化出了 $x^m$ 的全部导函数。]①

---

拉格朗日 1）**用代数方法**证明了泰勒假定的那个东西：只要$x$保持不确定，$f(x+h)$总可以表示成无穷级数$=f(x)+ph+qh^2+\text{etc.}$；他为微分演算奠定了**代数基础**，但就这点而论只能作为出发点来使用，因为微分演算通过它自己的方法可以简单得多地提供的东西，如用代数方法来烦琐地加以演化，那是完全多余的。

2）他一开始就证明了，当$x$为不确定时，$f(x+h)$的一般展开级数把所有那些作为泰勒定理的**失误**而出现的特殊情况都排除了出去。

3）他通过**变量的导函数**这个概念，给微分演算一个完全新的支持，它消除了许多无所裨益的困难。

如果把泰勒定理写成：

$$f(x+h)=f(x)+\frac{df(x)}{dx}\cdot\frac{h}{1}+\frac{d^2f(x)}{dx^2}\cdot\frac{h^2}{1\cdot2}+\frac{d^3f(x)}{dx^3}\cdot\frac{h^3}{1\cdot2\cdot3}+\cdots,$$

那末在这公式中还并没有包含导函数$f'(x)$，$f''(x)$等等概念，而只是说**对同一个原先的** $f(x)$已经实行过**逐次的微分运算**；在这公式中逐阶微系数**不是作为$x$的逐阶导函数**来表述的。

另一方面，凡是在拉格朗日写成

$$f(x+h)=f(x)+ph+q\cdot\frac{h^2}{1\cdot2}+r\cdot\frac{h^3}{1\cdot2\cdot3}+\cdots$$

的地方，这里就假定了

$$f(x+h)=f(x)+ph+qh^2+\cdots$$

中的系数 $p,q$ 等等已经化为它们的值$=f'(x)$，$f''(x)$等等。

---

## II．泰勒定理的基础，从二项式定理的代数语言翻译成微分表示方式

A）**布夏拉**在其《微积分学》的第二个注记（附录）中指出："除了从三角公

---

① 这对括号是在马克思的手稿里原来就有的。

式容易导出的圆函数的微分外，所有其他**单项式**如 $x^m$，$a^x$，$\log x$ 等等的微分，都是单从**二项式定理**演化出来的；在**确定指数公式中的常量** $A$ 时则应用到了**麦克劳林**定理，但也可以不必用到它等等[以后将回到这后一点上来]①。由此可见，**微分的全部原理**是单**以二项式定理为基础的**。"

另一方面，**泰勒**是在这样的一个时期建立他的**定理**的(它连同**麦克劳林**定理——后者本身又可表示为泰勒定理的特殊情况——是微分学运算的最重要的定理)，那时一方面不仅二项式定理已经知道，而且也已知道通过微分演算本身的途径所提供的 $x$ **的函数**的演化，以及一般而论微分演算的所谓要素的演化。

$f(x+h)$相应于二项式定理总是处于右边，即展开级数的一边，并附有因子 $h^0(=1)$，$h$，$\frac{h^2}{1\cdot 2}$，$\frac{h^3}{1\cdot 2\cdot 3}$等等(**附有 $h$ 的正整数升幂**，而不去管负数幂，分数幂和对数幂，根据按拉格朗日方法所展开的，这种幂在这里我们可以暂时不予考虑)；$h$ 相继幂次的**未定系数**，即不同的 $x$ **的逐阶导函数**或微系数自然依据所要展开的原先的 $f(x)$，例如 $f(x)$ 是否 $=x^m$ 或 $a^x$ 或 $\log x$ 或 $\sin x$ 或其他函数等等而具有不同的形式。但在泰勒定理那里显然是以二项式定理即 $f(x)=x^m$ 最简单的应用为基础的。

因此，如果我们根据二项式定理展开 $f(x+h)=(x+h)^m$。

那末，例如

$$(x+h)^m=x^m+mx^{m-1}h+\frac{m\,(m-1)}{1\cdot 2}x^{m-2}h^2+\cdots=$$

$$=x^m+mx^{m-1}h+\frac{1}{2}m(m-1)x^{m-2}h^2+\cdots$$

在第三项$\frac{1}{2}\,m(m-1)x^{m-2}h^2$ (即拉格朗日在他上面的展开式中写成$\frac{1}{2}f''(x)h^2$ 的那一项)中，是直接按 $x$ 从 $mx^{m-1}$ 导来的函数 $m(m-1)x^{m-2}$；因此，为了不是得到半个函数，而是得到整个函数，这里和以后都必须写做 $m(m-1)x^{m-2}\frac{h^2}{1\cdot 2}$，也就是把数字除数放在 $h^2$，$h^3$ 等等的下面。

于是我们

---

① 这对括号是在马克思的手稿里原来就有的。

| | 根据二项式定理得到 | 把 $x$ 看作为变量时得到 | |
|---|---|---|---|
| $x^m$ | $=x^m$ | $=f(x)$ | |
| $mx^{m-1}h$ | $=mx^{m-1}h$ | $=f'(x)h$ | $=\frac{dy}{dx}h$ |
| $m(m-1)x^{m-2}\frac{h^2}{2}$ | $=\frac{m(m-1)}{2}x^{m-2}h^2$ | $=f''(x)\frac{h^2}{2}$ | $=\frac{d^2y}{dx^2}\cdot\frac{h^2}{2}$ |
| $m(m-1)(m-2)x^{m-3}\frac{h^3}{2\cdot3}$ | $=\frac{m(m-1)(m-2)}{2\cdot3}x^{m-3}h^3$ | $=f'''(x)\frac{h^3}{2\cdot3}$ | $=\frac{d^3y}{dx^3}\cdot\frac{h^3}{2\cdot3}$ |
| $m(m-1)(m-2)(m-3)x^{m-4}\frac{h^4}{2\cdot3\cdot4}$ | $=\frac{m(m-1)(m-2)(m-3)}{2\cdot3\cdot4}x^{m-4}h^4$ | $=f^{\mathrm{IV}}(x)\frac{h^4}{2\cdot3\cdot4}$ | $=\frac{d^4y}{dx^4}\cdot\frac{h^4}{2\cdot3\cdot4}$ |
| 等等 | 等等 | 等等 | 等等 |

此外泰勒已经知道如何用微分演算的方法来求 $d(x^m)=mx^{m-1}dx$，也就是$\frac{d(x^m)}{dx}$或$\frac{dy}{dx}=mx^{m-1}$，以及同样来求

$$\frac{d^2y}{dx^2}=m(m-1)x^{m-2}$$

等等；换句话说，由二项式定理所提供的 $x$ 的导函数和作为逐阶微系数而出现的那些导函数是等同的；他也同样知道，在用微分演算方法求这些函数时，无论是 $h$ 或者是它的数字系数$\frac{1}{1\cdot2}$，$\frac{1}{1\cdot2\cdot3}$ 等等都会消失。我们得到函数 $mx^{m-1}$，$m(m-1)x^{m-2}$ 等等作为用 $\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$ 等等表示的微分运算的结果。

另一方面，二项式定理指出，

$$f(x+h)\text{这里是}(x+h)^m=f(x)+f'(x)h+f''(x)\frac{h^2}{1\cdot2}+f'''(x)\frac{h^3}{1\cdot2\cdot3}+\cdots$$

因此，为了得到 $f(x+h)$ 的实际展开式，就必须用$\frac{h}{1}$乘第二项，用$\frac{h^2}{1\cdot2}$乘第三项等等；也就是说必须把微分过程中已消失的 $h$，$h^2$ 等等和它们的数字因子恢复过来。

例如，当 $x^m=x^3$ 时，就有

$$f(x+h)=(x+h)^3=x^3+3x^2h+3xh^2+h^3。$$

因此通过二项式定理人们就得到 $x$ 的导函数，以及借助于微分而得到同

样的这些函数：

$$3x^2 = 3x^2, \qquad 3\cdot 2x = 6x, \qquad 6x^0 = 6。$$

如果把那些在微分中已消失的$h$，$\frac{h^2}{1\cdot 2}$，$\frac{h^3}{1\cdot 2\cdot 3}$等等恢复过来，那末：

$$mx^{m-1}\cdots\text{代替 } 3x^2 \text{ 得到}\cdots\cdots\cdots 3x^2h$$

$$m(m-1)x^{m-2}\cdots\text{代替 } 6x \text{ 得到}\cdots\cdots\cdots 6x\frac{h^2}{2} = 3xh^2,$$

代替$6x^0$ [即 $m(m-1)(m-2)x^{m-3}$，$(x+h)^3$ 的三阶导函数 =

$$3(3-1)(3-2)x^{3-3} = 3\cdot 2\cdot 1x^0 = 6 \text{ 得到}\cdots\cdots\frac{6h^3}{1\cdot 2\cdot 3} = h^3]①。$$

因此

$$f(x+h)\text{或}(x+h)^3 = x^3 + 3x^2h + 3xh^2 + h^3,$$

这就是从二项式定理已经知道的结果。

再者，如上所述，泰勒已经知道，在微分演算中由二项式定理所表明的那个级数，从原函数到它的导函数是这样表示的：$x^m$ 作为原函数 $f(x)$或$y$；二项式定理中的第二个函数 $mx^{m-1}$ 作为一阶微系数$\frac{dy}{dx}$的实在值，第三个函数作为二阶微系数$\frac{d^2y}{dx^2}$的实在值等等。如果我现在不用

$$(x+h)^m = x^m + mx^{m-1}h + m(m-1)x^{m-2}\frac{h^2}{1\cdot 2} +$$

$$+ m(m-1)(m-2)x^{m-3}\frac{h^3}{1\cdot 2\cdot 3} + \cdots$$

而把这些函数的微分表示式放在它们的位置上，并用不确定的 $f(x+h)$ 来代替$(x+h)^m$，那末我就得到

$$f(x+h) = f(x)\ (\text{或 } y) + \frac{dy}{dx}h + \frac{d^2y}{dx^2}\cdot\frac{h^2}{1\cdot 2} + \frac{d^3y}{dx^3}\cdot\frac{h^3}{1\cdot 2\cdot 3} + \cdots,$$

这就是泰勒定理。

[还必须指出，如果作为$(x+h)^m$ 的四阶导函数我们得到

$$m(m-1)(m-2)(m-3)x^{m-4}\frac{h^4}{1\cdot 2\cdot 3\cdot 4},$$

它的微分表示式 = $\frac{d^4y}{dx^4}$——那末对于$(x+h)^3$ 来说它将

$$= 3(3-1)(3-2)(3-3)x^{3-4} = 3\cdot 2\cdot 1\cdot 0\cdot x^{-1} = 0,$$

① 这对括号是在马克思的手稿里原来就有的。

乘上$\frac{h^4}{1\cdot2\cdot3\cdot4}$之后还是$=0$,所以在这情况下$\frac{d^4y}{dx^4}$也$=0$。因此,**二项式定理**也在这里再一次指出,微分演算中的变数$x$一旦从导函数中被消去,也就是后者变为**常量**以后,那末和它相应的$\frac{dy}{dx}$也就变为$=0$;这就是说,推导$x$的新的函数因而亦即新的微分就到此为止。]①

虽然在这里$f(x+h)$的泰勒公式现在只是从二项式定理的最初等应用,即在$x^m$中用$x+h$来代替$x$,从而展开$(x+h)^m$才得到的。但是这对于**结果的普遍性**绝对没有什么改变,因为1)无论什么样的$f(x)$,具有正整数升幂的$h$因子[如果人们乐意的话,可以从$h^0=1$作为$f(x)$的展开级数第一项的因子开始]②总管保持一样;2)系数$\frac{dy}{dx}$,$\frac{d^2y}{dx^2}$等等,这些事实上只是有待实行的微分运算的符号,自然将随原函数$f(x)$的特殊性质而提供不同的结果。例如当$f(x)=a^x$时,就和当它是$ax^n$等等时所提供的结果不同。在所有情况下,它们只提供$f'(x)$,$f''(x)$等等$f(x)$的导函数,况且象拉格朗日曾经用事实证明过的那样,这些导函数也都可用代数方法、而且主要还是在二项式定理的基础上演化出来的。

$f(x+h)$虽然是不确定的,不具有确定的幂次,因而当它展开时也就展开成一个无穷级数。但是只要$m$还没有取得一个确定值,$f(x+h)[=(x+h)^m]$同样也是完全不确定的并且也只能展开成无穷级数;因此在翻译到微分语言时,正象这种情况所要求的那样,它也提供一个无穷级数。

首先把证明真正加以推广的是拉格朗日,但在泰勒那里则相反,如我们即将看到的那样,这个推广只是一种**推测性的**假定,况且从一开始起他就一次也没有对这**假设**所包含的那些条件作过探讨。

[从外表上可以看出,如果

$$f(x+h)=f(x)+ph+q\frac{h^2}{2}+r\frac{h^3}{2\cdot3}+\cdots=$$

$$=f(x)+f'(x)h+f''(x)\frac{h^2}{1\cdot2}+f'''(x)\frac{h^3}{1\cdot2\cdot3}+\cdots=$$

$$=f(x)+\frac{dy}{dx}h+\frac{d^2y}{dx^2}\frac{h^2}{1\cdot2}+\frac{d^3y}{dx^3}\frac{h^3}{1\cdot2\cdot3}+\cdots,$$

那末差值$f(x+h)-f(x)$或$y_1-y$就等于$x$的导函数或微系数的无穷累加。右边——即左边的一般表示式$f(x+h)$展开成无穷级数的那一边——用$+$号与$f(x)$连起来的所有各项,共同构成了$f(x)$和$f(x+h)$之间的差值。就导

①② 这对括号是在马克思的手稿里原来就有的。

函数或微系数的**无穷级数**而言，绝大多数的函数事实上只能表示成这样的级数。

**指数函数、对数函数和三角函数**按其本性不可能表示成**具有有限项的代数表示式**。在本来意义下的**代数**函数中也有大量的只能表示成无穷级数，如 $\frac{a}{a-x}$ 等等。只是在某些代数函数如 $(x+h)^4$ 那里，才有确定数目的导函数；这种函数最后变为**常量**（$x$ 已消去），因此其导函数 $=0$，正象在恒等式中那样这也是不言而喻的。此外，$f'(x)=0$（这里 $f'$ 代表所有以后的 $f''$，$f'''$ 等等）并不意味着 $x$ 已变为 $=0$，即已被消去，而意味着 $f'(x)=0$ **是一个具有确定次数的等式**：因为每一个等式如把它的两边都写到一边去时就 $=0$，而这时恰恰通过把微分表示式归到一边而它的实在值归到另一边，从 $f'(x)=0$，就把 $x$ 演化了出来。由于这个缘故，$f'(x)=0$ 在极大极小理论中起着如此重要的作用。]①

B）代替

$$f(x+h)=f(x)\ (\text{或}\ y)+\frac{dy}{dx}h+\frac{d^2y}{dx^2}\cdot\frac{h^2}{1\cdot 2}+\frac{d^3y}{dx^3}\cdot\frac{h^3}{1\cdot 2\cdot 3}+\cdots$$

这定理也可写成：

$$f(x+h)=f(x)+\frac{df(x)}{dx}\cdot\frac{h}{1}+\frac{d^2f(x)}{dx^2}\cdot\frac{h^2}{1\cdot 2}+\frac{d^3f(x)}{dx^3}\cdot\frac{h^3}{1\cdot 2\cdot 3}+\cdots$$

完全按照拉格朗日：

$$f(x\pm h)=f(x)\pm ph+\frac{qh^2}{1\cdot 2}\pm\frac{rh^3}{1\cdot 2\cdot 3}+\cdots$$

或

$$f(x\pm h)=f(x)\pm f'(x)h+f''(x)\frac{h^2}{1\cdot 2}\pm f'''(x)\frac{h^3}{1\cdot 2\cdot 3}+\cdots$$

C）泰勒不是把他的定理表示为译成微分语言的二项式定理，而是通过一个假设使它看来好象**一般的**证明。

1）**假定**函数 $f(x+h)$ 被展开成 $h$ 的正整数升幂，那末

$$y_1\ \text{或}\ f(x+h)=y\ (\text{或}\ f(x))+Ah+Bh^2+Ch^3+\cdots,\qquad\text{(I)}$$

其中 $A,B,C$ 等等未定系数是 $x$ 的未知函数。

事实上，通过二项式定理并结合微分演算的已知结果人们已经找到了等式（15 页②最后一行）

① 这对括号是在马克思的手稿里原来就有的。

② 见本文第 150 页。

$$f(x+h)=f(x)+\frac{dy}{dx}h+\cdots,$$

所以就不难想到，再用未定的系数——而这种未定系数法在代数中，例如在对数的展开中，是经常用到的——如 $A$，$B$ 等等来代替$\frac{dy}{dx}$等等，也就是来代替微系数或导函数，以便现在**反过来从这些系数**借助于微分演算本身演化出微系数并从而给它们一个**一般的**推导。泰勒**不依赖**于微分演算所做到的恰恰是 $f(x+h)$可以展开成级数 $=f(x)+Ah+\text{etc.}$；但在他那里这只是一个假设；证明是由拉格朗日首先提供的。

如果人们可以这样假定：他私下曾以我们在 IIA) 中所表述的那种方式发现了他的定理，那末，以后用 $A$，$B$，$C$ 等等来代替 $x$ 的导函数或其微分表示式——作为将来有待实行的微分运算本身的出发点——肯定不是什么不可思议的事。

2）要从等式

$$y_1=y+Ah+Bh^2+Ch^3+\cdots$$

中确定系数 $A,B,C$ 等等，显然只有一个方法：把这一个等式作成二个，它们的左边，即函数的未展开的表示式大家**相同**，而它们的右边，即**展开成一系列项的函数**，则具有不同的形式。由于这两个等式左边恒等，所以右边也必须恒等，因此那些附有 $h$ 相同幂次的项($y$ 附有因子 $h^0=1$)可以相等起来。

如果我们在 $x$ 是常量、$h$ 是变量的假定下进行微分，那末 $y$ 就此消失，因为它是 $x$ 的函数，不含 $h$。同时我们得到没有 $h$(附有 $h^0=1$)的 $A$，但其他的项则给出数字系数，因为 $h^1$，$h^2$ 等等都带有数字指数。如果我们在 $h$ 是常量、$x$ 是变量的假定下进行微分，那末我们得到向上升的$\frac{dy_1}{dx}$如$\frac{dy}{dx}+\frac{dA}{dx}h+\text{etc.}$，这方法的关键**在于对第一项的微分**

$$\frac{dy_1}{dh}=A+\cdots,\qquad \frac{dy_1}{dx}=\frac{dy}{dx}+\cdots,$$

所以如果一旦找到了

$$A=\frac{dy}{dx},$$

那末就会自然而然地得出$\frac{d^2y}{dx^2}$等等其他系数。

## III. 麦克劳林定理也纯粹是从二项式定理的代数语言翻译成微分语言

A) **通过泰勒定理**给出了一个公式，它把每一个用 $x$ 表达的函数(在前已阐明的条件下）当 $x$ 增长一个正的或负的增量 $h$，从而变为 $f(x\pm h)$ 时，表示成一个级数，其中 $f(x)$ 是第一项，而随后的、附有升幂的 $h$ 作为因子的各项 $\frac{dy}{dx}$ 等等是**微系数**，或者不如说是一些表明如何用微分方法来导出逐次的 $x$ 的函数的符号，而这些函数连同各自的因子 $h, h^2$ 等等的总和就等于 $f(x+h)$ 与 $f(x)$ 的差值。

**麦克劳林**定理则相反，它用于把用 $x$ 表达的**函数本身**，例如

$$y=\frac{1}{a+x},\quad y=(a^2+bx)^{1/2},\quad y=(a+x)^m,\cdots$$

展开成级数(而且确是按 $x$ **的升幂**展开)。$x$ **的函数** $=(a+x)^{-1}$ 或 $=(a^2+bx)^{1/2}$，或 $=(a+x)^m$ 等等。由于 $f(x)$ 应按 $x$ 的升幂加以展开，所以 $x$ 在这里起着和泰勒定理中增量 $h$ 一样的作用。它是二项式的第二项，因此只显现为升幂的因子，正象在那里是 $h^0, h^1, h^2$，在这里则是 $x^0, x^1, x^2$ 等等。例如在泰勒定理中真正被展开的是第一项：**变量** $x$ 的导函数，而增量 $h$ 这个第二项，只表现为从 $h^0=1$ 开始的升幂的因子。在麦克劳林定理中**变量** $x$ 则正好反了过来；因而通过这定理所要做到的是展开第一项，此项在这里是一个**常量**，而巧妙之处正在于通过微分演算把包含在 $f(x)$ 中的常系数的代数推导……

---

## IV. 关于泰勒定理的其他事项

这种代数的演化是我自己从麦克劳林定理的发明者——即从柯林·麦克劳林本人，第六版(《代数论著三编等等》，伦敦，1796 年)那里得来的。出版者**安妮·麦克劳林**是麦克劳林的寡妇。这是一部遗著。在序言中说到："由于**伊萨克·牛顿爵士**的一些关于高次方程求解及其根的性质的**规则**，在他的《**算术大全**》中**大都是不加任何证明而给出的**，所以麦克劳林先生曾经打算把这本论著充当牛顿那部著作的注释。因为我们看到，伊萨克爵士书中的所有那些长期以来使学代数的人感到困惑的难懂段落，都已清楚地加以解释和证明。"

至于如何从高次方程导出低次方程的方法，人们不晓得究竟牛顿是通过微分演算发现了它呢，还是反过来——作为二项式定理的发明者——是通过

二项式定理发现了他的全部微分理论。

**麦克劳林**(**柯林**)1698年生于苏格兰，死于1746年。1720年(即22岁)他发表了甚至使牛顿感到惊奇的关于曲线的论著。

**泰勒**(J.布鲁克)1685年生于埃特蒙顿(米德尔塞克斯)，死于1731年(46岁)。他发表了《**正反增量方法**》，伦敦，1715—1717年，他的定理可以说是该书的概述。此外，还发表了各种数学的以及一些形而上学的著作。

---

1) **本来意义下的代数学**最伟大的发现是**二项式定理**(同时也可以把多项式包括在内)。只有通过这个定理才不但有可能象前面那样求解一定幂次的**方程**，而且还有可能建立**方程的一般理论**。

---

2) 但二项式定理不是单单用于发展方程(也包括含有几个未知量的方程)的一般理论，诸如发展三角函数、指数函数等等的组合理论；它是**微分演算的一般基础**，因此有理由提出这样一个问题：是牛顿这个既是二项式定理又是微分演算的发现者，或者至少是他的学生**泰勒**和**麦克劳林**(前者在年代上比后者早)这两个以他们的普遍化公式而使微分演算在技术上的应用大为简化的人，纵然在**暗中**直接从二项式定理的应用取得了他们的结果。

---

如果泰勒级数不能给出$(x+h)$的函数的展开，那末就发生它的(所谓)失误；当函数的**特殊值不能用** $h$ **的正整数幂与有限系数的相结合来表示**时，就是这种情况。

---

# 泰勒定理，麦克劳林定理和拉格朗日的导函数理论

## I

牛顿发现二项式定理（在它的应用中也包括多项式），使整个代数学来了个革命，因为这个发现第一次使得有可能建立**方程的一般理论**。

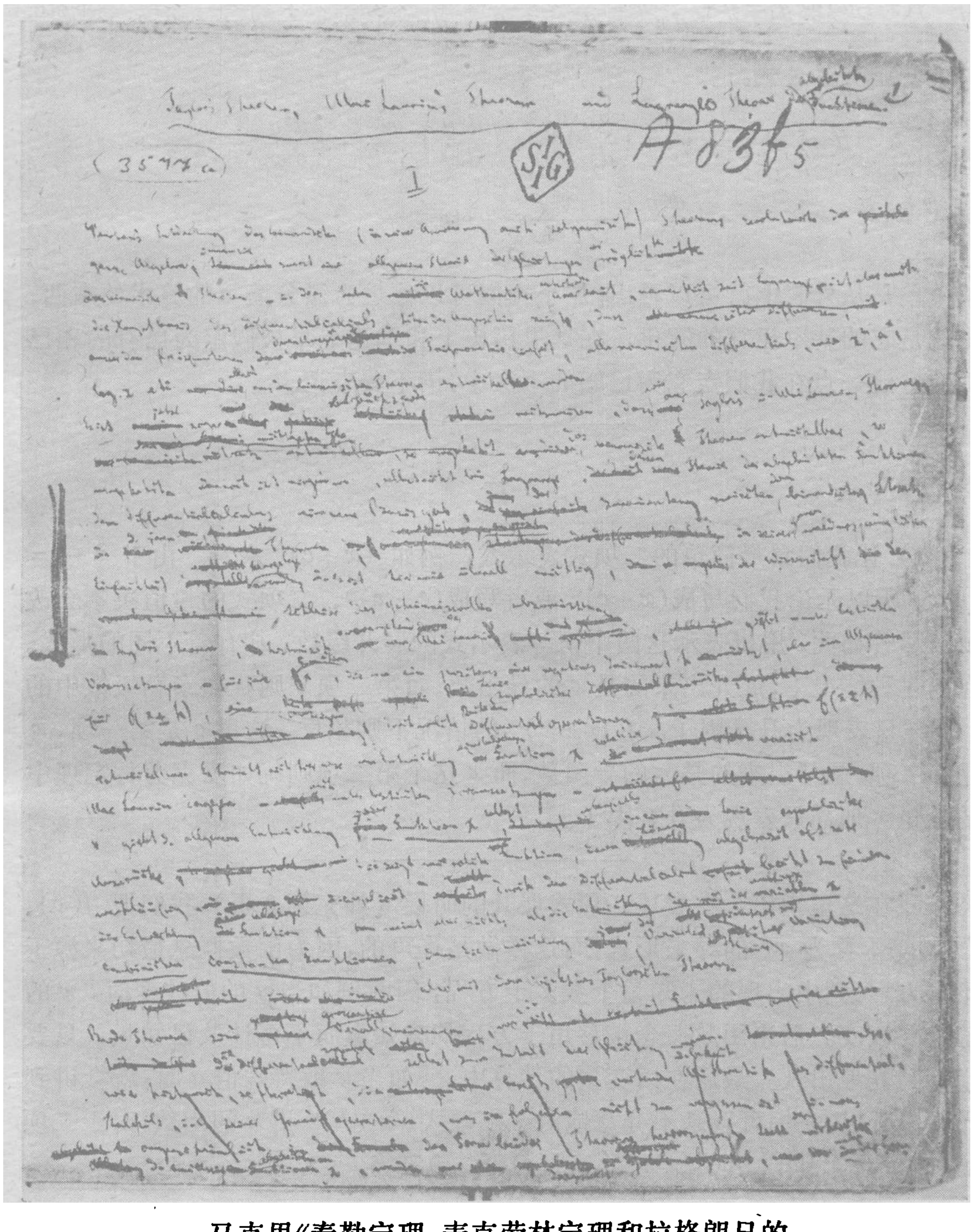
但是二项式定理——这是数学家们，尤其从拉格朗日以来所绝对公认的——也是微分演算的主要基础。从外表上已经可以看出，除那些发源于三角学的圆函数以外，所有如 $x^m$，$a^x$，$\log x$ 等等单项式的微分，光是从二项式定理就能演化出来。

现在甚至在教科书中也已成为一种时尚，总要象二项式定理可以从泰勒定理和麦克劳林定理中推导出来那样去证明反过来也行。然而无论何处，都没有把二项式定理和这些定理之间的联系以其完全原始的简单性解释清楚，即使是拉格朗日也没有做到，虽然他的导函数理论给了微分演算一个新的基础。因而在这里象在其他各处一样，重要的是应该给科学撕掉这块神秘的面纱。

泰勒定理历史上在麦克劳林定理之前，它——在一些确定的假定下——对每个增长一个正的或负的增量 $h$ 的 $x$ 的函数，因而一般地对 $f(x\pm h)$ 给出了一系列符号表示式，这些表示式表明，用哪样的一系列微分运算可以把 $f(x\pm h)$ 展开。所以这里所涉及的是 $x$ 的任意**函数当它变化时**的展开。

麦克劳林则与此相反，他——也是在一些确定的假定下——同样用一系列符号表示式给出了**每个 $x$ 的函数本身**的一般展开，这些表示式表明，对于用代数方法求解往往非常烦琐而复杂的那种函数，如何能用微分演算很容易找到其解。但是 $x$ 的任意函数的展开，无非就是指那些**和自变量 $x$ 相结合的常函数的展开**，因为变量本身的展开等同于它的变化，因而和泰勒定理的目标是相一致的。

这两个定理都是了不起的推广，在这些推广中微分符号本身变成了等式的内容。这里所得到表示的，不是那些实际上逐次导出的 $x$ 的函数，而只是以它们的符号等价物形式出现的导函数，这些符号等价物表明有这么多有待实行的运算手续，而与函数 $f(x)$ 或函数 $f(x+h)$ 的形状无关。这样，就得到两个公式，它们在一定的限制下可以应用于 $x$ 或 $x+h$ 的所有特殊函数。



马克思《泰勒定理，麦克劳林定理和拉格朗日的导函数理论》手稿的第一页

**泰勒公式：**

$$f(x+h)\text{或}y_1=y+\frac{dy}{dx}h+\frac{d^2y}{dx^2}\frac{h^2}{1\cdot 2}+\frac{d^3h}{dx^3}\frac{h^3}{1\cdot 2\cdot 3}+\frac{d^4y}{dx^4}\frac{h^4}{1\cdot 2\cdot 3\cdot 4}+\text{etc.}$$

**麦克劳林公式：**

$$f(x)\text{ 或 }y=(y)+\left(\frac{dy}{dx}\right)\frac{x}{1}+\left(\frac{d^2y}{dx^2}\right)\frac{x^2}{1\cdot 2}+\left(\frac{d^3y}{dx^3}\right)\frac{x^3}{1\cdot 2\cdot 3}+$$
$$+\left(\frac{d^4y}{dx^4}\right)\frac{x^4}{1\cdot 2\cdot 3\cdot 4}+\text{etc.}$$

单是从外表上就可以看出，那个人们可以称之为**微分演算的算术**的东西，也就是它的基本运算的演化，在这里历史上和理论上早被假定为已有的和已知的。这一点在我假定它为已知的下文中，请不要忘记。

---

## II

1）让我们拿最简单的二项式表示式，例如 $(x+c)^m$ 来看。由于 $x+c=c+x$，所以无论我们写成$(x+c)^m$ 或者写成$(c+x)^m$，二项式的量值决不会改变。尽管如此，用以表示这两个恒等表示式的展开级数，却有着不同的形式。在前一种情况下，展开出来的是 $x$ 的导函数，而第二项 $c$ 则犹如泰勒级数中的 $h$ 那样只表现为升幂因子。在$(c+x)^m$ 的情况下则相反，这里 $c$ 是第一项，$x$ 是第二项，展开出来的是 $c$ 的导函数，而 $x$ 这个第二项则犹如麦克劳林定理中的变量 $x$ 那样只表现为升幂因子。

---

我在这里指出：在表述二项式展开的时候（补充第 3 页）我们把 $f(x)$、$f'(x)$等等称为 $x^m$ 的导来形式，而这样做是许可的，因为函数概念最初来源于不定方程，其中出现的未知量个数多于方程，因而例如当 $y$ 的值改变时，$x$ 的值也要改变。后来函数概念被移用于方程的**未知量**而不再计及已知量，只要它们独立地出现。最后在用变量的计算中，譬如当 $x$ 取特殊值 $a$ 时，就会讲到 $f(a)$。所以，我们可以毫不介意地对于二项式 $(x+h)^m$ 讲 $x$ 的函数，对于二项式$(c+x)^m$ 我们讲 $f(c)$和导函数 $f'(c)$、$f''(c)$等等。

---

再者，如果方程是**一般的**，那末我们必须用 $f(x)=y$ 来代替 $f(x)=x^m$，其中变量 $x$ 没有确定的幂次，但可以取任何一个幂次，因而 $(x+h)^m$ 也就得到一般形式 $f(x+h)$。

但凡是在方程左边发生的，也必须在右边重演，这就是说，我们必须划掉由二项式$(x+h)^m$的幂次 $m$ 所给出的后面一些项并用 + etc. etc. 来代替，用以表明一般的 $f(x+h)$其新的导函数将无穷无尽地排列下去。

于是我们得到

$$y_1 \text{ 或 } f(x+h)=f(x) \text{ 或 } yh^0+Ah+Bh^2+Ch^3+Dh^4+\cdots,$$

而这就是泰勒在表述他的定理时所由出发的基本方程。

因此：

$$y_1 \text{ 或 } f(x+h)=f(x)+\frac{dy}{dx}h+\frac{d^2y}{dx^2}\cdot\frac{h^2}{1\cdot 2}+\cdots$$

所以，为了得到$f(x+h)$的展开，我们只要对每一个给定的 $x$ 的函数，当 $x$ 发生变化也就是当它变为 $x+h$ 时，演化出微系数$\frac{dy}{dx}$、$\frac{d^2y}{dx^2}$等等这些 $x$ 的逐阶导函数，然后把那些在微分过程中已消失的因子 $h$，$\frac{h^2}{1\cdot 2}$等等重新恢复过来。

泰勒定理就这样显现为二项式定理从代数语言到微分语言的简单翻译。

---

现在就产生了一个问题：牛顿这个二项式定理的发现者，是否通过对这一定理的暗中应用已经找到了使微分演算的应用得到如此不寻常简化的泰勒定理以供他个人使用呢？这是应当无条件地予以否定的。如果是这样，那他一定会以这一收获而夸耀自己。如果他看到了这个简单的联系，那末他就不会留下任何东西让泰勒、麦克劳林甚至拉格朗日去发现，并且事实上他自己就会把微分演算弄到底。因为虽然泰勒的（相应地麦克劳林的）定理最初只处理那些具有**单独一个自变量**的函数，但它依然也是多变量函数展开的基础，不管这些函数是以隐的或显的形式给出的。

这同一问题在泰勒那里显得较为可疑，他一方面看到了早已提供的牛顿的代数学（《算术大全》），另一方面也看到了早已提供的牛顿和莱布尼茨的微分演算。

诚然，人们也许可以说，在代数中总是只涉及一定幂次的二项式，如在$(x+h)^m$ 中只涉及 $m$ 次幂的那样，而 $f(x+h)$则把每一个确定的幂次只作为一个要素包括在内，但绝对不把它作为界限。可是这个异议也许反过来可用于反对泰勒，因为二项式定理比他的定理范围要广泛得多。前者可以让 $h^{-1}$、$h^{1/m}$、$h^{m/n}$ 等等，简言之即带有一切可能指数的 $h$ 作为 $x$ 的函数的因子，而泰勒定理只有在这种情况下才可应用（也就是不失误）：如果 $x$ 的函数虽然不确定而可以任意改变，但每一个 $x$ 的函数如 $f(x)$，$f'(x)$等等都是有限表示式（这

决不损害它的可变性)，另一方面 $h$ 可以展开成正整数升幂而作为它们的逐次因子的话。但是泰勒没有试图去证明这个不定的、能任意加以展开的 $f(x+h)$ 可按二项式展开的规则来予以表示。他恰恰由于这一点而引起了人们的怀疑，象例如在 $(x+h)^m$ 的情况下，由于 $m$ 的不确定性可以使级数变为无穷那样，他满足于把 $(x+h)^m$ 写成 $f(x)+f'(x)h+\cdots+\text{etc.}$ 而这时却忘记了，虽然只要我们保持 $m$ 为不确定，$(x+h)^m$ 就具有无限性，但我们总是知道它的终结的，因为最后第二项只能含有 $x$，而最后一项只能是 $x^0h^m=h^m$。

但在我看来这一点绝对可靠，泰勒对于他的定理和二项式定理之间的简单联系丝毫也没有想象到。他完全活动在微分演算本身的地盘上，而没有追溯到它的根源。

---

在由二项式理论所提供的那个出发等式上，泰勒没有什么改变，只是使 $f(x+h)$ 变为无幂次的，因而使之适用于任何展开，因而也使右边用 $+\text{etc.}$ 成为没有终结。因此他应用二项式定理(或者同样可以说，由这定理所提供的一个未知量的一般等式的形式)只限于为他提供自己的出发等式，没有证明它在这里是可以应用的。而对于多项式本身，他就立刻从微分演算的观点来加以处理。

---

2）麦克劳林定理

象在前面演化泰勒的出发等式时一样，我们可以把这等式①加以推广，使之变为

$$f(x)=A+Bx+Cx^2+\cdots+\text{etc. etc.}\ (\text{没有终结}),$$

这就是麦克劳林的出发等式。

---

……人们将会摆脱这个仅适用于一定幂次的等式 $x^n$，它在级数倒过来之前曾构成其第一项，而现在则必须构成其最后一项。它将被 $+\text{etc. etc.}$ 所代替。这样一来，多项式就得到了一般形式，这是当我们用一般表示式 $f(x)$ 代替 $(c+x)^n$ 或 $x$ 的任何别的确定函数时所必需的形式，其中 $f(x)$ 没有什么幂次，但在它的展开中包含了所有的幂次。于是我们一般地得到：

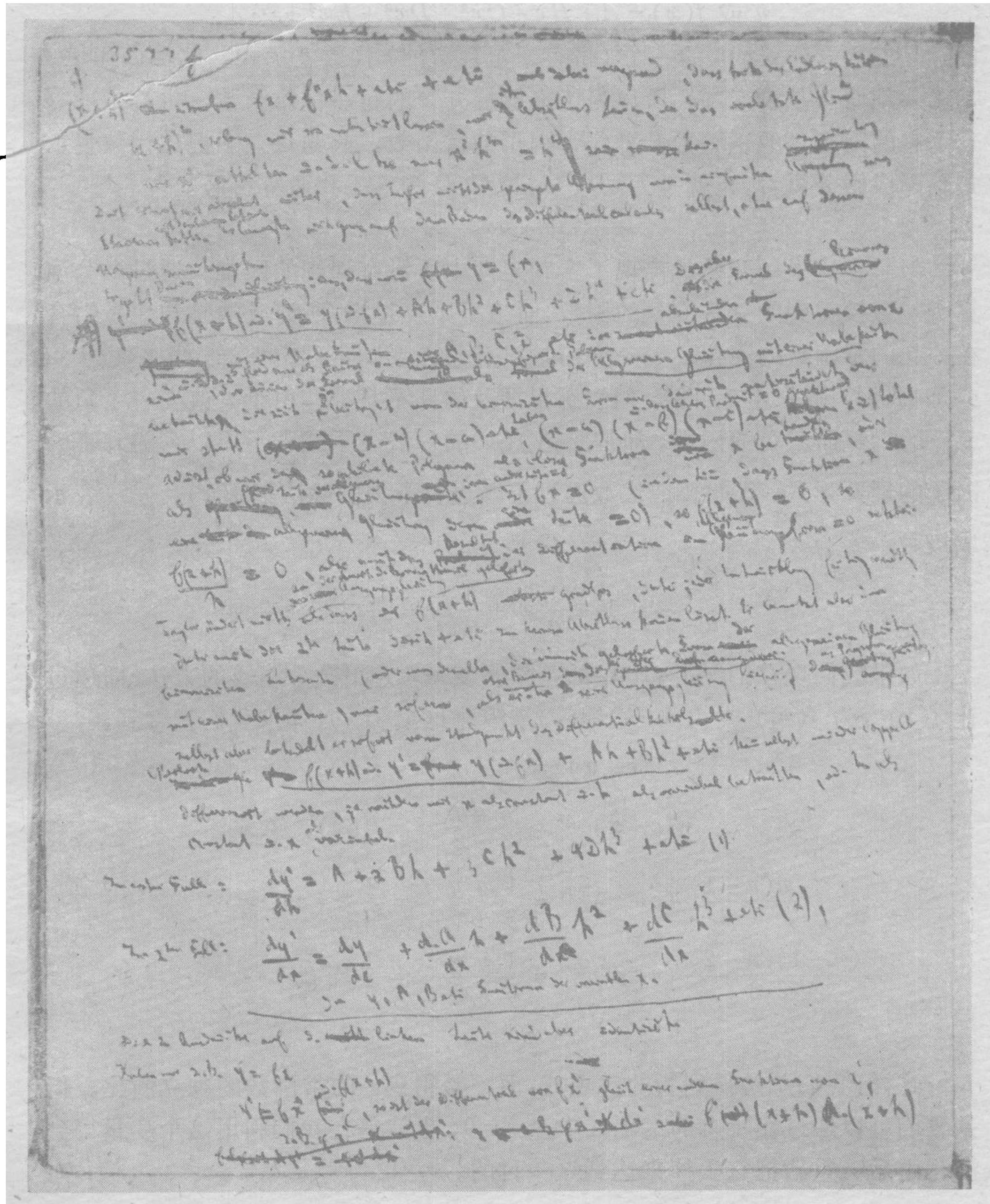
$$f(x)\ \text{或}\ y=A+Bx+Cx^2+Dx^3+Ex^4+\cdots$$

这是麦克劳林在演化他的定理时所由出发的基本等式。

---

① 指等式

$$f(x)=(c+x)^n=Ax^0+Bx+Cx^2+Dx^3+Ex^4+\cdots+ncx^{n-1}+x^n.$$



马克思《泰勒定理，麦克劳林定理和拉格朗日的导函数理论》手稿的第五页

因此，麦克劳林表述的出发点

$$y \text{ 或 } f(x)=A+Bx+Cx^2+Dx^3+Ex^4+\cdots$$

已经是二项式$(c+x)^n$的一个（带有未定系数的）代数表示式，此地已知量$c$是第一项而$x$是第二项。所以为了能够展开$x$的每一任意函数，也就是能够表示它的附有$x$的正整数升幂作为因子的常函数，问题只在于把这代数表示式翻译成微分语言，也就是为系数$A$、$B$、$C$、$D$等等去找微分符号。

---

但是在这里一开始就出现了泰勒定理中所没有遇到的困难。通过微分演算所能**直接**得到的只是变量的函数，而这里则反过来，所涉及的是与变量结合在一起的常函数的演化。另一方面，只有象在泰勒定理中那样，当$x$变为$x_1$或$x+h$时，微分才有可能，但在这里涉及的并不是那些由于$x$改变一个负的或正的增量而获得的函数，却是要把$f(x)$的一般表示式表示成以$x$的升幂为因子的展开形式，犹如普通代数通过不断进行下去的除法把$f(x)=\frac{a}{a-x}$表示成级数

$$\frac{a}{a-x}=1+\frac{1}{a}x+\frac{1}{a^2}x^2+\frac{1}{a^3}x^3+\cdots$$

如果我们现在把$(c+x)^n$中的$x$作为变量，并用微分演算的方法(即通过让$f(x)$变为$f(x_1)$或$f(x+h)$等等)来演化这个$f(x)$，那末我们就得到：

$$(c+x)^n=y=f(x),$$

$$n(c+x)^{n-1}=\frac{dy}{dx}=f'(x),$$

$$\frac{n(n-1)}{1\cdot 2}(c+x)^{n-2}=\frac{1}{2}\frac{d^2y}{dx^2}=\frac{1}{2}f''(x),$$

$$\frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}(c+x)^{n-3}=\frac{1}{6}\frac{d^3y}{dx^3}=\frac{1}{6}f'''(x)$$

等等

这样，我们总是得到新的二项式，而不是得到$c$的常函数的独立演化。但是我们可以通过在所有的二项式中令$x=0$，也就是在我们利用这个变量得到演化之后又重新把它去掉来达到这一目的。

---

麦克劳林定理可以看作泰勒定理的**特殊情况**。

在泰勒的情况下我们有：

$$y=f(x),$$

$$y_1=f(x+h)=f(x)\text{或}\, y+\frac{dy}{dx}h+\frac{1}{2}\frac{d^2y}{dx^2}h^2+\text{etc.}+$$
$$+\left[\frac{1}{1\cdot 2\cdot 3\cdots n}\right]\frac{d^ny}{dx^n}h^n+\text{etc.}$$

如果在 $f(x+h)$ 中，并同样在处于右边的 $y$ 或 $f(x)$ 及其在 $\frac{dy}{dx}$，$\frac{d^2y}{dx^2}$ 等等形式下的符号导函数中，我们令 $x=0$，以致这些函数除了 $x$ 的常元素的展开外不再包含更多的东西，那末

$$f(h)=(y)+\left(\frac{dy}{dx}\right)h+\left(\frac{d^2y}{dx^2}\right)\frac{h^2}{1\cdot 2}+\left(\frac{d^3y}{dx^3}\right)\frac{h^3}{1\cdot 2\cdot 3}+\text{etc.}$$

于是 $y_1=f(x+h)=f(0+h)$ 就变为 $h$ 的、与 $x$ 的函数 $y=f(x)$ 同样的一个函数，因为 $h$ 进入 $f(h)$ 中正象 $x$ 进入 $f(x)$ 中，以及 $(y)$ 进入 $\left(\frac{dy}{dx}\right)$ 等等中一样，$x$ 的任何痕迹都已消失。

因此，我们可以在两边用 $x$ 代替 $h$，于是得：

$$f(x)=(y)\text{ 或 }f(0)+\left(\frac{dy}{dx}\right)x+\left(\frac{d^2y}{dx^2}\right)\frac{x^2}{1\cdot 2}+\text{etc.}+$$
$$+\left(\frac{d^ny}{dx^n}\right)\frac{x^n}{1\cdot 2\cdot 3\cdots n}+\text{etc.}$$

或者象其他人所曾把它写成的那样：

$$f(x)=f(0)+f'(0)x+f''(0)\frac{x^2}{1\cdot 2}+f'''(0)\frac{x^3}{1\cdot 2\cdot 3}+\text{etc.}$$

例如在 $f(x)$ 或 $(c+x)^m$ 的展开中：

$$(c+0)^m=f(0)=c^m,$$
$$m(c+0)^{m-1}x=mc^{m-1}x=f'(0)x\ \text{etc.}$$

下面在我们转到拉格朗日的时候，我将不再考虑仅仅作为泰勒定理的特殊情况的麦克劳林定理。这里只是还要指出，它象泰勒定理一样也有其所谓“失误”的地方。这些失误，在前者都来源于常函数的无理性质，而在后者则来源于变函数的无理性质。

现在人们可以问：

牛顿是否只是把这些结果公之于世，象他在《**算术大全**》中对那些最困难的情况所做的那样，而不是为他自己个人的应用私下已经从他发现的二项式定理中导出了泰勒和麦克劳林定理呢？这是绝对确定地应予否定的：他不是一个让他的学生据有这种发明的人。事实上他那时还埋头于完成那些后来被泰勒和麦克劳林所假定为已有并且已知的微分运算本身的制定工作。而且，

象他关于微分演算的最初几个基本公式所表明的那样，牛顿显然首先是从力学的而不是从属于纯粹分析的出发点达到它们的。

另一方面，就泰勒和麦克劳林而论，他们一开始就在微分演算自己的地盘上工作和活动，所以没有任何一点动机要去寻找微分演算的尽可能简单的代数出发点，更不用说在牛顿学派和莱布尼茨学派之间正当围绕着将微分演算的确定的既成形式当作一种新发现的、完全特殊的、和普通代数有天壤之别的数学学科而相争论的时候了。

他们各自的**起始等式**同二项式定理的联系，对他们来说是不言而喻的，然而这个联系并不象例如在微分 $xy$ 或 $\frac{x}{y}$ 时认为这些是由普通代数所提供的表示式那样的不言而喻。

新旧事物之间的实际的因而是最简单的联系，总是只有在新事物本身获得了完善的形式以后才能发现，可以说，微分演算通过泰勒和麦克劳林定理获得了这个关系。因此把微分演算归结到严格的代数基础上去的责任首先落在拉格朗日身上。在这方面，18 世纪中叶的英国数学家**约翰·兰登**在他的《**差分分析**》一书中或许走在他前面。但是在我能对此作出判断之前，我必须先在博物馆里查阅一下这本书。

## III．拉格朗日的函数理论

拉格朗日从泰勒定理的代数论证出发，因而也就是从微分演算的一般公式出发。

关于泰勒的起始等式

$$y_1 \text{ 或 } f(x+h)=y \text{ 或 } f(x)+Ah+Bh^2+Ch^3+\text{etc.},$$

只须指出：

1）这个级数并没有得到证明；$f(x+h)$并不是一个**确定**幂次的二项式；倒不如说 $f(x+h)$是 $x$ 的任一函数当 $x$ 增长一个正的或负的增量 $h$ 时的不确定的一般表示式；所以 $f(x+h)$包括任何幂次的 $x$ 的函数，但同时排除了展开级数本身的任何一个确定的幂次。因此泰勒本人在级数的末尾加了“+etc.”。但是展开级数的这个适用于获得增量的 $x$ 的确定函数的定律——不管这些函数可用有限等式或用无穷级数表示——可以毫无问题地应用于不确定的一般的 $f(x)$，因而同样也可以应用于不确定的一般的 $f(x_1)$ 或 $f(x+h)$，则是尚需**证明的**。

2）通过对这等式的两次微分，即对 $y_1$ 一次把 $h$ 作为变量而 $x$ 作为常

量，然后把 $x$ 作为变量而 $h$ 作为常量，它就被翻译成了微分演算的语言。这样就产生了两个等式，它们的左边相同，而它们的右边则形式各异。但是为了能把这两个右边的都是 $x$ 的函数的不确定系数等同起来，就要求再假定各个系数 $A, B$ 等等**虽是不确定的，但却是有限的量**，并且同样要求假定伴随它们的因子 $h$ 是以**正整数的幂次**上升的。假定说(但实际并非如此)，在 $f(x)$ 中的 $x$ 保持**一般**的情况下泰勒已经对 $f(x+h)$ 证明了这一切，那末决不因此而当 $x$ 的函数采取确定的特殊值时，这也还适用。这些特殊值可能反过来和用他的级数

$$y_1 = y + \frac{dy}{dx}h + \frac{d^2y}{dx^2}h^2 + \text{etc.}$$

来处理是不相容的。

一句话：包含在泰勒的未经证明的起始等式中的条件或假定，自然也出现在从它导出的定理

$$y_1 = y + \frac{dy}{dx}h + \frac{d^2y}{dx^2}h^2 + \text{etc.}$$

中。所以它不能应用于与那些假定相矛盾的 $x$ 的某些函数。这定理的所谓“失误”就是由此而来的。

拉格朗日用代数方法为起始等式建立了基础，并同时用它的展开本身指明了哪些特殊情况是由于和它的**一般**性质相矛盾，也就是和 $x$ 的函数的一般的、不确定的性质相矛盾而自行排除了出去。

---

H) 1) 拉格朗日的巨大功绩，不但在于他用纯粹代数的分析为泰勒定理，从而根本为微分演算建立了基础，而且尤其在于他也引进了导函数的概念，这概念事实上或多或少都为所有他的后继者所使用，即使没有把它说出来。但是他并不以此为满足。他用 $h$ 上升的、正整数的幂次为所有可能的 $x+h$ 的函数给出了纯粹代数的展开，而后用微分演算中的名称来命名它们。微分演算(泰勒定理等)所具有的一切敏捷性和简便性都将因此而丧失，并且往往要用烦琐而复杂得多的代数运算来代替。

2) 就纯粹分析而论，拉格朗日确实摆脱了一切在他看来存在于牛顿的流数，莱布尼茨的不同阶的无限小量，消失量的极限理论以及把 $\frac{0}{0}\left(=\frac{dy}{dx}\right)$ 当作微系数的符号等等中的形而上学的先验性①。然而这并不妨碍他在应用其理论于曲线等等的时候，自己却常常要用到其中的这个或那个“形而上学”的观念。

---

① “先验性”一词，原文为 Transzendenz。