分析中的严密性

19 世纪初越来越多的数学家认识到分析（导数、积分、级数等）的概念以及证明中的不严密性问题。导数概念中的无穷小量、无穷大量始终没有合理的解释；积分无论是作为“越来越多次求和”还是作为求导数的反运算也都说不清楚；级数在没有收敛性的保证下出现了悖论（早在 18 世纪初大伯努利就由发散级数出发得到许多矛盾的结果，它称之为“悖论”）。尽管如此，人们所得到的大量结果却仍然是正确的。这种状况导致了一些数学家决心从含混的分析学中整理出一套逻辑上清晰的理论，这就是由布尔扎诺 (Bolzano)(1781-1848) 、哥西 (Cauchy)(1789-1857) 、阿贝尔 (Abel)(1802-1829) 和狄里赫勒 (Dirichlet)(1806-1859) 开始的、外尔斯特拉斯 (Weierstrass)(1815-1897) 发展的所谓“批判运动”，其核心是建立“极限”的理论。这个运动的开始和非欧几何的创立大体上年代相同。由于欧几里德几何（第 5 公设）的合理性已被怀疑，所以严密的分析学不能建立在几何的基础之上，因此只能建立在算术的概念之上（不能建立在代数（字母的运算）基础上的原因是：当时的分析学就是当时的代数的推广。顺便说一句，后来的代数学确实可以用来定义极限）。

1. **函数**. “函数”的概念清晰化是建立严密的分析学的第一步（导数及积分都是关于函数而言的）。关于“函数”，尽管早在 1714 年莱布尼兹已被定义为“表示依赖于变量的量”，但是实际上 18 世纪的大多数数学家的直观意识中都觉得函数应当在处处有统一的有限的解析表达式，包括早期的高斯 (Gauss)(1777-1855) 在内。欧拉和拉格朗日 (Lagrange)(1736-1813) 允许函数在不同的区域内有不同的（有限）表达式。晚年的拉格朗日把（无穷）幂级数也看成是一种函数。傅立叶 (Fourier)(1768-1830) 摆脱了“函数是代数函数及其推广”（例如幂级数）的信念，他实际上讨论的函数允许在任一有限区间内有有限多个间断点（这样的函数在有限区间内可以展开为傅立叶级数）。哥西从“变量”出发（他把变量定义为“依次取许多互不相同的值的量”）定义了自变量和函数。他说：“当变量之间互相联系，即给定这些变量之一的值就可以决定所有其他变量的值，人们通常想象这些量是用其中的一个来表达的，这时这个量就取名为自变量，而由这个自变量表示的其他的量就叫作自变量的函数。”他不要求函数一定要有解析表达式。现今的（单值）函数的定义是由狄里赫勒在 1837 年给出的：如果对于给定区间上的每一个 x 的值有唯一的一个 y 的值与它对应，则称 y 是 x 的函数。至于在整个区间上 y 是否按照一种或多种规律依赖于 x ，或者 y 对于 x 的依赖关系是否可以用数学运算来表达，都是无关紧要的。事实上，早在 1829 年他给出了一个函数的例子：在有理点处取值为 c ，而在无理点处取值为 d 。

函数概念的推广使得对于代数函数以及初等超越函数而言不成问题的“连续”、“可微”、“可积”等性质都必须重新考虑。

2. **连续与极限**. “连续”的恰当的定义是由布尔扎诺给出的。他是波希米亚的一个神父、哲学家和数学家。他研究连续性的动因是试图给出代数基本定理的一个纯算术的证明以代替高斯在 1799 年给出的基于几何思想的证明。他在 1821 年的《纯粹分析的证明》一书中说：如果在一个区间内的 x 处，只要 ω（的绝对值）充分小，就能使得差 $f(x+\omega)-f(x)$（的绝对值）任意小，就称 $f(x)$ 在 x 处是连续的。从这里我们已经可以看到极限概念的正确想法。他证明了多项式函数的连续性。他对于微积分的建立（除了实数理论之外）有正确的概念。不过他的工作在半个世纪内没有被引起注意。

哥西抓住了极限、无穷小量、无穷大量和连续性的概念实质。 1821 年的《代数分析教程》中说：“当一个变量逐次所取的值无限趋近一个定值，最终使得变量的值与该定值之差

要多小就有多小，这个定值就叫做所有其他值的极限。”在该书的的序言中他说：“当一个变量的数值无限减小以致收敛到极限 0，则人们就称该变量成为无穷小；当变量的数值无限增大以致收敛到极限 ∞，则该变量就称为无穷大。”这就澄清了莱布尼兹的无穷小的概念（莱布尼兹的无穷小量 $\mathrm{d}x$ 是一个在运算中有时必须保留有时又必须舍去的量，这是长时间内的一种普遍想法。哥西则认为是一种变量。）在无穷小量的基础上哥西定义了连续性：如果变量的无穷小增量总产生函数自身的一个无穷小增量，则称函数对于自变量保持连续。用现代的观点衡量，哥西著作中的严密性是不够的，因为他用了“无限趋近”、“想要多小就多小”等直观却是含糊的说法作为定义的基础，但是对于理解极限的严格定义是有益的。

外尔斯特拉斯改进了布尔扎诺和哥西的工作。他力求避免直观而把分析的基础奠定在纯算术之上。他是在 1841-1856 年作中学教师时做这些工作的，因此直到 1859 年到柏林担任教授之前他的工作并不为人所知。他攻击“一个变量趋于一个极限”的说法，因为这使得人们联想到时间与运动。他把变量简单地解释为一个字母，该字母代表它可以取值的集合中的任何一个数。这样，运动就消除了。为了消除布尔扎诺和哥西的定义中的不明确性，他给出了连续性的 ε-δ 的定义，即：如果给定任一正数 ε，都存在一个正数 δ，使得对于区间 $|x-x_0|<\delta$ 中的所有 x，都有 $|f(x)-f(x_0)|<\varepsilon$，则称 $f(x)$ 在 $x=x_0$ 处连续。如果在这个说法中用 L 代替 $f(x_0)$，则称 $f(x)$ 在 $x=x_0$ 处有极限 L。海涅 (Heine)(1821-1881) 在 1870 年定义了一元和多元函数的一致连续性。

回顾极限定义的历史，就不难理解我们今天掌握这个概念为何如此困难。在这个定义中，实际上用到了数理逻辑中的两个重要的谓词：“对于所有的”和“存在”，将一个无限的过程转变为一个有限的陈述。

在人们接受了连续性的思想之后，许多原来直观上明显的事实都需要加以证明。例如连续函数的中值定理、有界无穷序列必有收敛子序列（即存在聚点）、闭区域上连续函数的最大最小值定理、有限覆盖定理等。此时正是建立实数理论的时期，在实数理论建立之前，这些定理的证明总有含糊之处（如布尔扎诺对于中值定理的证明），而在实数理论建立之后的证明就是严格的了（例如伯莱尔对于有限覆盖定理的证明）。

3. **导数**. 布尔扎诺第一次（1817 年）把 $f(x)$ 的导数定义为当 Δx 经由负值和正值趋于 0 时，比值 $\frac{f(x+\Delta x)-f(x)}{\Delta x}$ 无限接近趋向的量 $f'(x)$。他强调 $f'(x)$ 不是两个 0 的商，也不是两个消失了的量之比，而是上述的比趋近的一个数。

哥西沿用了布尔扎诺关于导数的定义。他进一步明确指出导数与微分的关系：$\mathrm{d}y=f'(x)\mathrm{d}x$，这样他把牛顿的导数与莱布尼兹的微分统一了起来。他又证明了增量 Δy 与 Δx 之间的关系：$\Delta y=f'(x+\theta\Delta x)\Delta x$（他用到了 $f'(x)$ 的连续性）。这个结果是以前拉格朗日已经知道的。

关于连续性和可微性之间的关系在很长时间内弄不清楚。哥西时代的几乎所有数学家都相信并且“证明”连续函数一定可微。这种状况持续了半个世纪。布尔扎诺是个例外，他在没有完成也没有发表的《函数论》（1834 年）中给出了处处不可微的连续函数的例子（一条没有解析表达式的曲线）。真正讲清楚连续与可微的差别的是黎曼 (Riemann)(1826-1866) 在 1854 年为了取得格丁根的无薪大学教师资格所写的论文中给出的例子：用 (x) 表示 x 与他最接近的整数的差（如果 x 是半整数，则 (x) 规定为 0）；定义

$$f(x)=\frac{(x)}{1}+\frac{(2x)}{4}+\frac{(3x)}{9}+\cdots.$$

此级数对于所有的 x 收敛。它的全部间断点为 $x = \frac{m}{2n}$ ，其中 $(m, 2n) = 1$（所以在任一小区间内都有无穷多个间断点），间断点处函数值跳跃 $\frac{\pi^2}{8n^2}$ 。 $f(x)$ 可积，其原函数 $F(x) = \int f(x)\mathrm{d}x$ 就处处连续但在 $f(x)$ 的间断点处不可微。这个例子直到 14 年后才发表，也并没有引起多大的注意。1860 年塞莱里耶 (Cellérer)(1818-1889) 给出了处处连续处处不可微的函数的例子

$$f(x) = \sum_{n=1}^{\infty} a^{-n} \sin a^n x,$$

其中 a 为一个大的整数。这个例子直到 1890 年才发表。最引人注目的例子是外尔斯特拉斯在 1872 年给出的：

$$f(x) = \sum_{n=1}^{\infty} b^n \cos(a^n \pi x),$$

其中 a 是一个奇数， b 是一个小于 1 的常数满足 $ab > 1 + \frac{3\pi}{2}$ 。此级数一致收敛，因而定义了一个连续函数。

连续性与可微性的差别的发现使得数学家更加不敢信赖直观的或几何的思考了。

3. **积分**. 自从微积分诞生以来，人们更倾向于牛顿的“积分是微分的反运算”的说法，而忽略了莱布尼兹的积分是微元的“和”的思想。但是在处理间断函数的积分时牛顿的方法就不如莱布尼兹的了。而在哥西时代函数的傅立叶展开已经发展起来，其中处理的积分（求展开系数）经常是对于间断函数进行的。所以必须对于莱布尼兹的积分做严格的定义。

哥西在 1823 年的《无穷小分析教程概论》一书中对于定积分作了最系统的开创性工作。他首先对于连续函数给出定积分作为“和的极限”的确切定义，即：如果 $f(x)$ 在 $[x_0, X]$ 上连续，区间 $[x_0, X]$ 被 $x_1, x_2, \cdots, x_n = X$ 所分割，则定积分

$$\int_{x_0}^{X} f(x)\mathrm{d}x = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)(x_i - x_{i-1}),$$

其中 $x_i - x_{i-1}$ 的最大值趋于 0 。哥西证明了无论怎样选取 x_i ，积分总是存在的并且相同（由于没有（闭区间上连续函数必然）一致连续的概念，它的证明不是严格的）。接着他考虑变上限积分

$$F(x) = \int_{x_0}^{x} f(x)\mathrm{d}x. \qquad (x \in [x_0, X])$$

他证明了 $F(x)$ 在 $[x_0, X]$ 上连续，并且有

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\int_{x}^{x+h} f(x)\mathrm{d}x.$$

利用积分中值定理，他证明了

$$F'(x) = f(x).$$

这样，哥西给出了微积分基本定理的第一个证明。他引入了不定积分的标准记法，即

$$F(x)=\int f(x)\mathrm{d}x=\int_{x_0}^{x} f(x)\mathrm{d}x+C.$$

他又指出，如果 $f'(x)$ 连续，则 $f'(x)$ 的定积分等于 $f(x)$ 的函数值的差，即

$$\int_a^b f'(x)\mathrm{d}x=f(b)-f(a).$$

之后他讨论了奇异积分（函数无界或积分区域无限的情形）。

他用积分 定义 了面积、曲线的长度、曲面围成的区域的体积和表面积，例如由 $y=f(x)$ 定义的曲线在 $x=a,b$ 之间的弧长就定义为 $\int_a^b \sqrt{1+(y')^2}\mathrm{d}x$ 。他不自觉地给函数加了必要的条件（例如 $f(x)$ 可微）。哥西对于二重积分指出：如果被积函数不连续，则积分的次序会影响到最后的结果。

哥西研究的函数都是比较规则的。出于数学的考虑应当研究更一般的函数。现在我们所知道的黎曼可积函数要比连续函数或分段连续函数广泛的多。黎曼在 1854 年给出了可积的充分必要条件：小区间上函数的振幅与该区间的长度的乘积的和（在小区间的最大长度趋于 0 时）的极限等于 0 。达布 (Darboux)(1842-1917) 在 1875 年将此条件该述为上和与下和相等。

积分的下一个最有意义的推广是勒贝格积分。这个概念以及一系列相关的结果是勒贝格 (Lebesque) (1875-1941) 在 1903 年给出的。这种积分在傅立叶级数理论中特别有用。

4. **级数**. 18 世纪末，由于不加限制地应用无穷级数而导致一些可疑的或荒谬的结果，使得人们追究无穷级数运算的合理性。 1810 年前后，傅立叶、高斯、布尔扎诺等人开始正确地处理无穷级数。

傅立叶在 1811 年给出了级数收敛的正确的思想：级数的部分和随项数的增加而趋近于一个固定的值。他还强调收敛的必要条件是单项趋于 0 。

高斯在 1812 年研究超几何级数

$$F(\alpha,\beta,\gamma,x)=1+\sum_{n=1}^{\infty}\frac{\alpha(\alpha+1)\cdots(\alpha+n-1)\beta(\beta+1)\cdots(\beta+n-1)}{\gamma(\gamma+1)\cdots(\gamma+n-1)}x^n$$

($\alpha,\beta,\gamma\in\mathbb{C},\quad \gamma\neq 0,-1,-2\cdots$ 。但是高斯当时研究的是 α,β,γ 为实数的情形) 时给出了这个级数的收敛性的判定准则： $|x|<1$ 时收敛， $|x|>1$ 时发散， $x=1$ 时当且仅当 $\alpha+\beta<\gamma$ 时收敛，而 $x=-1$ 时当且仅当 $\alpha+\beta<\gamma+1$ 时收敛。他的文章中的异乎寻常的严密性使得当时的数学家丧失了兴趣。他没有讨论级数收敛的一般原则。

布尔扎诺在 1817 年就已经对于序列的收敛有了正确的概念，但是他的工作没有被广泛了解。他给出的收敛性条件被后人归功于哥西。

哥西在《代数分析教程》 (1821 年) 中用部分和的极限给出了级数收敛与发散的定义。接着他叙述了哥西收敛判别准则，即序列 $\{S_n\}$ 收敛当且仅当 $|S_{n+r}-S_n|$ 对于一切 r 和充分大的 n 都小于任何指定的量。他证明了这个条件的必要性。由于缺乏实数的知识，他没能证明充分性。

在同一本书中哥西给出了正项级数 $u_1+u_2+\cdots$ 收敛或发散的一些条件，其中包括：若 $\lim\limits_{n\to\infty}\sqrt[n]{u_n}<1$ 则级数收敛，若 $\lim\limits_{n\to\infty}\sqrt[n]{u_n}>1$ 则级数发散，若 $\lim\limits_{n\to\infty}\frac{u_{n+1}}{u_n}<1$ 则级数收敛，若 $\lim\limits_{n\to\infty}\frac{u_{n+1}}{u_n}>1$ 则级数发散，以及比较判别法和积分判别法。他还证明了收敛级数取极限与加、乘法可交换次序。对于非正项级数证明了绝对收敛必收敛以及交错级数的莱布尼兹判别法。

哥西也研究了函数项级数。他指出了：余项收敛于 0 的泰勒级数收敛到产生该级数的函数。他指出了拉格朗日的《函数论》一书中的错误，拉格朗日说：如果函数 $f(x)$ 在 $x=x_0$ 处有各阶导数，则 $f(x)$ 可以表为在 x_0 附近的收敛到 $f(x)$ 的泰勒级数 (哥西给出了反例：$f(x)=\mathrm{e}^{-\frac{1}{x^2}}\ (x\neq 0)$ ， $f(0)=0$ 。

哥西在研究一般的函数项级数 $F(x)=\sum_{n=1}^{\infty}u_n(x)$ 时出现了错误。他说：如果级数收敛并且每个 $u_n(x)$ 都连续，则 $F(x)$ 连续，并且可以对于级数可以逐项积分，即 $\int_a^b F(x)\mathrm{d}x=\sum_{n=1}^{\infty}\int_a^b u_n(x)\mathrm{d}x$ 。在这里他忽视了一致收敛的要求（当时没有一致收敛的概念）。

阿贝尔很赞赏哥西的严密性的工作， 1826 年他在给他的老师的一封信中称赞哥西“是当今懂得怎样对待数学的人”。当时他对于分析的状况很不满意。在给另一位教授的信中说：“人们在分析中确实发现了惊人的含糊不清之处。这样一个完全没有计划和体系的分析，竟有那么多人能研究过，真是奇怪。” 阿贝尔发现了哥西的错误，他给出了例子

$$\sin x-\frac{\sin 2x}{2}+\frac{\sin 3x}{3}-\cdots,$$

此和函数在 $x=(2n+1)\pi$ 处是不连续的（发散的），因此每项都连续的函数项级数不一定连续。然后他在一致收敛的条件下证明了和函数的连续性，但是他并没有提出一致收敛的概念。

斯托克斯 (Stokes)(1819-1903) 在 1848 年清楚认识到了一致收敛的概念。他是第一流的数学物理学家。他并没有给出一致收敛的明确定义，而是给出了一致收敛的一种直观解释。

实际上，外尔斯特拉斯早在 1842 年就有了一致收敛的概念，他利用这个概念给出了函数项级数逐项积分和在积分号下求微分的条件。通过外尔斯特拉斯的学生，人们了解到一致收敛的的重要性。海涅在 1870 年明确地提出了这个概念。

外尔斯特拉斯在做中学教师 (1841-1856) 时证明了“一致收敛的连续函数项级数的和函数连续”的逆命题的强形式：闭区间上的任一连续函数可以表为绝对一致收敛的多项式级数。这个定理对于多元函数也成立。在 19 世纪最后的 1/4 世纪中，这个定理被推广成多种形式，用于将复变函数表示为多项式级数或有理函数级数。

关于交换级数的项的次序， 1837 年狄里赫勒证明了绝对收敛级数的项可以任意交换，他还给出例子说明对于条件收敛级数这样做是不行的。 1854 年黎曼证明了条件收敛级数适当地调换项的次序可以收敛到任意指定的值

19 世纪下半叶许多第一流的数学家推导了无穷级数收敛的很多判别法则。

5. **三角级数**. 除了幂级数之外，另一类重要的级数是三角级数。最早研究三角级数的是欧拉。 1753 年他发表了在 1729 年得到的结果，通过一系列巧妙的但并不严格的计算他得到用三角级数给出的无穷插值公式 (构造 $f(x)$ 使得 $f(n)$ 为指定的值， n 取所有整

数），然后他给出了形式上与后来的傅立叶级数完全相同的级数。 1754 年达朗贝尔出于对于行星之间的距离的考虑也得到了同样的结果。 1777 年欧拉已经应用三角函数的正交性计算函数的三角级数展开式的系数，就象我们现今所作的一样。但是对于哪些函数可以表示为三角级数引起了长期的争论，欧拉、达朗贝尔、（丹尼尔）伯努利、拉普拉斯、拉盖尔都卷入这场争论。除了伯努利之外，其他人都不认为三角级数可以表示任意的解析函数，但是伯努利只是坚持一种信仰，他无法给出表达式。直到 1811 年傅立叶在研究热传导方程的过程中在广泛的考虑的基础上断言：任意函数都可以表示成三角级数，事实上他所说的“任意函数”是指在有限区间内有有限多个间断点的函数。由于当时还没有严格的积分理论，所以他无法严格地证明他的结论。

函数的幂级数展开（泰勒级数）是由函数在某一点附近的性质完全决定的，展开式通常只在该点的一个邻域内成立。而三角级数却是由函数在整个定义域上的取值所决定，展开式成立的范围也几乎是整个定义域，这个观点是傅立叶首先提出的。他还将非周期函数稍加处理变为周期函数来处理（允许函数在不同的区间有不同的表达式）。

在分析学的严密性的过程中，狄里赫勒在 1829 年给出了函数可以展成傅立叶级数的一组充分条件。黎曼在 1854 年给出了形如

$$a_0 + \sum_{n=1}^{\infty} a_n \sin nx + b_n \cos nx$$

的级数收敛到一个给定的周期为 2π 的函数的充分必要条件。康托 (Cantor)(1829-1920) 得到了关于傅立叶展式唯一性的充分必要条件。

分析的严密化并未结束于哥西时代。特别是实数理论这个基础还有待于建立。在 20 世纪这个严密化仍然继续着，并且发展成为一个学科：实变函数论。

严密化震撼了整个数学界。在一次会议上当哥西提出级数的收敛性的理论后，拉普拉斯急忙赶回家隐居起来，直到查完他在《天体力学》一书中所用到的级数。幸好他用的级数都是收敛的。若当 (Jordan)(1838-1922) 的《分析教程》前后三版 (第一版： 1822-1887，第二版： 1893-1886，第三版： 1909-1915) 不断在严密化方面进行改动。

关于这种严密化，大多数数学家并不表示赞赏。庞卡莱在 1899 年说：“半个世纪以来我们已经看到一大堆离奇古怪的函数。··· 诚然，从逻辑的观点看来，这些陌生的函数是最一般的；另一方面，最常见的函数以及遵从简单规律的函数却是一种特殊情形，这种情形只是函数中很小的一角。过去，人们为了一个实际的目的而创造一个新的函数；今天人们为了说明先辈在推理方面的不足而故意造出这些函数来，而从这些函数所能推出来的东西也就是仅此而已。” 厄米特在给斯蒂尔吉斯的一封信中说：“我怀着惊恐的心情对不可导函数的令人痛惜的祸害感到厌恶。” 对于发散级数的排斥，无论是阿贝尔还是哥西都曾感到不安。发散级数的意义我们在上一节已经讲过。