



怎样让发散数列收敛?

丁 玖

看到本文标题后,读者或许会像读到“怎样让厌食症患者喜欢吃饭?”一样的美好感觉。是的,纯粹数学如同伦理道德,总是“善”字当头。在人生中,我们希望自己积点德而离圣贤更近一点,在数学中,我们也想多做好事化腐朽为神奇。比方说,如果有种办法让一部分发散数列收敛起来,这便做了一件好事,它的一个数学功能就是让某些发散级数可以“广义”求和,它的一个物理效果便是让统计力学变为可能。

当然,那些学过初等微积分中关于无穷数列收敛理论但又比较“信奉教条”的读者,可能会问:“数列既然发散,岂有收敛之理?”不错,他们的质疑精神可嘉,应该大力提倡。任何标准的微积分教科书中都严格定义了何时称一个数列“收敛”、什么是收敛数列的“极限”,以及何时说一个数列是“发散”的。由于无穷级数的收敛理论本质上就是无穷数列的收敛理论对级数部分和数列的直接应用,反过来任何数列的收敛性可以转化为级数的收敛性,所以数列和级数几乎是一回事。我们在下面主要把级数拿来作为范例,向读者兜售“像老母鸡变鸭似的将发散数列变成收敛数列”的魔术技巧。

给定一个无穷级数 $\sum_{n=1}^{\infty} a_n$, 其中 a_n 是实数或复数无穷数列,称为该级数的通项。在本文的数列记号中我们不外加大括号“{ }”或圆括号“()”,以示符号简洁,故 a_n 既表示数列的第 n 项,又表示该数列本身,就好像常见的 $f(x)$ 既表示函数在 x 的值又表示函数本身一样。定义所给级数的部分和数列 $s_n = \sum_{k=1}^n a_k$, 即 s_n 是级数的前 n 项的和。因为项数是有限的,所以每一个部分和 s_n 都是可以

计算出的确定的数。如果部分和数列 s_n 当 n 趋向于无穷大时收敛到一个数 s ，即极限 $\lim_{n \rightarrow \infty} s_n = s$ ，则称级数 $\sum_{n=1}^{\infty} a_n$ 是收敛的并且收敛到 s 。这时 s 叫做该级数的和，写成 $s = \sum_{n=1}^{\infty} a_n$ 。在此情形下，这堆符号的组合 $\sum_{n=1}^{\infty} a_n$ 具有无歧义的数学意义，它代表了一个叫做“级数和”的数。

反之，如果部分和数列 s_n 当 n 趋向于无穷大时不收敛到一个数（即该数列发散），则所给级数 $\sum_{n=1}^{\infty} a_n$ 被说成是发散的，这时，也仅仅是在这时， $\sum_{n=1}^{\infty} a_n$ 只是一堆数学符号的混合体而已，不代表任何数，没有任何数学意义。

将可列无穷个数加起来以获得一个合理的“无穷级数和”这个数，如果没有高等数学中的极限概念和手段，应该是任何人算了一辈子也完成不了的任务。假设一个计算快手一秒钟能加一个数，他一个小时也只能加上 3600 个数，离加完无穷个数远得何止十万八千里！我国古代的哲学家庄子早就哀叹了完成这类费时无穷大工作量的不可能性：“一尺之棰，日取其半，万世不竭”。这是伟大的极限思想，将不可能变为可能，使得无穷多个数的求和变成现实。据说冯·诺伊曼就用级数求和法而不是初等方法，求解了别人考他智商的一道算术题：在匀速相向而行的两人之间，有一只苍蝇不间断地以匀速直线运动来回飞，那么在两人碰面时，苍蝇一共飞了多少距离？

正因为极限的概念是微积分大厦的基础，有必要在这里重温数列极限的严格定义。说一个数列 x_n 有极限 L ，意思是任给正数 ε ，存在自然数 N ，使得对所有的自然数 $n > N$ ，不等式 $|x_n - L| < \varepsilon$ 都成立。这个事实用符号表示就是 $\lim_{n \rightarrow \infty} x_n = L$ 。有极限的数列称为是收敛的，否则被叫做是发散的。

所以，如果部分和数列 s_n 不收敛，给出的级数 $\sum_{n=1}^{\infty} a_n$ 就发散，就是一堆无意义的符号，遑论求和了。然而，无穷级数的求和工作来自于人类对“求和”的合理定义，这时，经典的定义不能对级数求和，我们寻找的是能不能补充定义，对它“广义求和”。这篇文章的主题实际上是怎样广义求和发散级数，以及更一般地对发散数列求收敛。这是个有趣也有用的问题，因为不仅在数学上许多数列或级数不幸是发散的，而且在物理上也存在非常重要的发散级数以及不可能收敛的运动粒子。正如力学家孙博华教授最近告诉我的，对描绘非平衡态热力学系统之统计行为的玻尔兹曼方程的“查普曼 - 恩斯科格展开 (Chapman-Enskog expansion)”进行求解，就面临着发散级数的棘手问题。

既然收敛级数的“狭义求和”定义合理，发散级数的“广义求和”也需要一个合理的定义。这里的合理性自然包括要满足两个基本要求。一个是，如果级数本身在通常的意义下已经收敛，由广义求和法得到的“和”就应该等于级数在原先意义上的和。这个要求说明“广义求和”具有“狭义求和”的“遗传性”。另一个要求是基于传统求和法的线性性质。我们知道，微积分中的许多运算如求极限、求导数、求积分等都具有线性特征，例如求导代数法则

$$[af(x) + bg(x)]' = af'(x) + bg'(x).$$

对于级数而言，也有类似的断言：如果 $\sum_{n=1}^{\infty} a_n$ 和 $\sum_{n=1}^{\infty} b_n$ 均为收敛级数且 c 和 d 为常数，则级数 $\sum_{n=1}^{\infty} (ca_n + db_n)$ 也是收敛的，而且有等式

$$\sum_{n=1}^{\infty} (ca_n + db_n) = c \sum_{n=1}^{\infty} a_n + d \sum_{n=1}^{\infty} b_n.$$

我们希望发散级数的广义求和法也保持这个性质。

怎样定义满足如上两个合理条件的发散级数广义求和法呢？一个好的思路是“平均化处理”，或用更时髦的专业术语：“切萨罗算术平均法”。事实上，这个法子本身就是用来对付不收敛数列的，而级数的收敛性或发散性，根据教科书里的标准定义，实际上是关于给定级数的部分和数列而言的。所以我们来考虑怎样让一个不收敛的数列转变为一个“收敛数列”。先举个简单例子。

考虑数列 $a_n = (-1)^{n-1}$ 。它是 1 和 -1 交替出现的无穷数列，当然不收敛。然而如果我们取这个数列的前 n 项的算术平均数，得到的

$$A_n = \frac{1}{n} \sum_{k=1}^n a_k = \frac{1}{2n} [1 - (-1)^n]$$

称为原数列 a_n 的切萨罗算术平均数列，它的各项写出来就是

$$0, \frac{1}{3}, 0, \frac{1}{5}, 0, \dots, \frac{1}{2k-1}, 0, \dots,$$

所以当 n 趋向于无穷大时 A_n 趋向于 0。这样，对于这个发散的数列，通过平均化处理，我们获得了一个收敛的数列。

一般地，对于一个数列 a_n ，如果它对应的切萨罗算术平均数列 $A_n = \frac{1}{n} \sum_{k=1}^n a_k$ 收敛并收敛到极限 L ，则称原数列 a_n 在切萨罗算术平均的意义下收敛并收敛到 L 。平均化思想不仅在数学上对数列的收敛性有巨大帮助，而且它也让统计物理成长为一门受人尊敬的学科。甚至对于人类社会的福祉和安定，现代国家在税收上实行的“富人多缴税，穷人拿福利”政策，体现的多半也是仁慈的平均化思想。



切萨罗 (1859–1906)

切萨罗 (Ernesto Cesàro) 是意大利的微分几何学家。尽管他写过一本关于内蕴几何的书，其中描绘了现被称为“切萨罗曲线”的一类分形，以及几本微

积分教材，但他提出让可能发散的数列收敛的平均化途径，或许对后世的影响最大，数学上受惠最大的当推遍历理论，缺了它统计力学则登不上科学的殿堂。

读者自然会问，如果数列本来就收敛，那么它的切萨罗算术平均数列也收敛并收敛到同一极限吗？答案为“是的”。这是数列极限理论中的一个简单命题，在这里我们不妨把它证出来，顺便复习一下上面所述的“ ε - N ”数列极限语言。假设 $a_n \rightarrow L$ 。任给正数 ε ，存在自然数 M ，使得对于所有的自然数 $n > M$ ，不等式 $|a_n - L| < \varepsilon/2$ 成立。现在，对所有的自然数 $n > M$ ，有

$$\begin{aligned} |A_n - L| &= \frac{1}{n} \left| \sum_{k=1}^n (a_k - L) \right| \leq \frac{1}{n} \left| \sum_{k=1}^n |a_k - L| \right| \\ &= \frac{1}{n} \left(\sum_{k=1}^M |a_k - L| + \sum_{k=M+1}^n |a_k - L| \right) \\ &< \frac{1}{n} \sum_{k=1}^M |a_k - L| + \frac{n-M}{n} \cdot \frac{\varepsilon}{2} \leq \frac{1}{n} \sum_{k=1}^M |a_k - L| + \frac{\varepsilon}{2}. \end{aligned}$$

因为 M 已被固定，上面不等式右端的第一项当 n 趋向于无穷大时趋向于 0，故存在自然数 $N > M$ 使得当 $n > N$ 时，这一项小于 $\varepsilon/2$ ，因此 $|A_n - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ 。证毕。

另外显见，切萨罗平均运算是线性的，即数列 $ca_n + db_n$ 的切萨罗平均等于 c 乘以 a_n 的切萨罗平均加上 d 乘以 b_n 的切萨罗平均。再取极限可见，切萨罗平均的极限运算满足线性性质。这样，如果把用于发散数列的切萨罗算术平均法移植到对于发散级数的广义求和，这个方法满足前面提出的遗传性和线性两个基本要求。

综上所述，我们有了发散级数的切萨罗广义求和算术平均法：对于给定的发散级数 $\sum_{n=1}^{\infty} a_n$ ，如果它的部分和数列 $s_n = \sum_{k=1}^n a_k$ 在切萨罗算术平均的意义下收敛到极限 s ，则称原级数 $\sum_{n=1}^{\infty} a_n$ 在切萨罗算术平均的意义下有广义和 s 。由前面我们知道，如果给定的级数 $\sum_{n=1}^{\infty} a_n$ 本身已经收敛到和 s ，那么它也在切萨罗算术平均的意义下收敛到广义和 s 。此外，切萨罗广义求和算术平均法是线性的广义求和法。

举一个历史上既简单又著名的发散级数例子： $\sum_{n=1}^{\infty} (-1)^{n-1} = 1 - 1 + 1 - 1 + \cdots$ 。它的部分和数列是 $s_n = [1 - (-1)^n]/2$ 。对所有的自然数 k ： $s_{2k-1} = 1$ 和 $s_{2k} = 0$ ，故数列 s_n 不收敛，因此级数发散。另一方面，部分和数列 s_n 的切萨罗算术平均数列为：

$$A_{2k-1} = \frac{k}{2k-1} \text{ 和 } A_{2k} = \frac{1}{2}; \quad k = 1, 2, 3, \dots$$

从而 $A_n \rightarrow \frac{1}{2}$ 。或言之，所给级数在切萨罗算术平均意义下的广义和为 $\frac{1}{2}$ 。

之所以要举出上面这个例子，是因为十八世纪的瑞士人欧拉对这个级数也曾经给出和为 $\frac{1}{2}$ 的结论。然而这位历史上最多产的大数学家玩起无穷级数来，有时玩得太过自由了，因为他偶尔会自作主张地给幂级数在不收敛的点处赋上一值，上面欧拉的答案就是这样得出的：他知道著名的幂级数求和公式（这是公比为 r 的几何级数 $\sum_{n=0}^{\infty} r^n = 1/(1-r)$ 等式的直接结果，其中 $|r| < 1$ ）

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots,$$

于是他轻率地在等式两边代入 $x = 1$ ，得到等式

$$\frac{1}{2} = 1 - 1 + 1 - 1 + \cdots.$$

然而，这离真理还差了一步。今日，每一个学过初等级数理论的理工科大学生都知道上述幂级数的收敛半径为 1，且收敛区域仅仅是开区间 $(-1, 1)$ 。所以欧拉用了错误的幂级数赋值法所得到的是发散级数的广义和。其实，如果他将 -1 分别乘以如上幂级数展开的两端，再将级数奇偶项彼此对调，得到一个非幂级数形式的函数项级数

$$-\frac{1}{1+x} = x - 1 + x^3 - x^2 + \cdots,$$

然后再如法炮制地代入 $x = 1$ ，便有同一常数项级数的另一个“和”

$$-\frac{1}{2} = 1 - 1 + 1 - 1 + \cdots,$$

这是多么荒唐的“数学”啊！

不过，如果欧拉不用直接赋值法，而是在上面对满足 $0 < x < 1$ 的 x 都收敛的幂级数等式左端的函数 $1/(1+x)$ 取 $x \rightarrow 1^-$ 时的极限（即 x 趋向于 1 的左极限），就得到与切萨罗广义求和算术平均法结果相同的另一种意义下的广义和

$$\frac{1}{2} = 1 - 1 + 1 - 1 + \cdots.$$

将这个法子一般化，我们就得到了发散级数的第二种经典的广义求和法：对于给定的发散级数 $\sum_{n=1}^{\infty} a_n$ ，形式地写出对应的幂级数 $\sum_{n=1}^{\infty} a_n x^n$ 。假如这个级数关于满足不等式 $0 < x < 1$ 的每一个数 x 都收敛（换句话说，此幂级数的收敛半径不小于 1），并且它的和函数 $f(x)$ 当 $x \rightarrow 1^-$ 时有极限 s ： $\lim_{x \rightarrow 1^-} f(x) = s$ ，则此值 s 称为给定级数在泊松 - 阿贝尔幂级数意义下的广义和。这个方法曾由法国数学家泊松用于三角级数这一特殊的情形，但其思想根植于下一段中定理的主人、挪威伟大却早逝的数学家阿贝尔，所以让他们共享这个命名荣誉是合适的。

“基于幂级数和函数极限意义下的广义和”的定义陈述，让我们想起了微积分中关于幂级数在收敛区间端点性质的阿贝尔定理：假设幂级数 $\sum_{n=1}^{\infty} a_n x^n$ 的收敛半径为 $R > 0$ 。若 $\sum_{n=1}^{\infty} a_n R^n$ 收敛，则该幂级数在闭区间 $[0, R]$ 上是一致收敛的。“函数项级数 $\sum_{n=1}^{\infty} f_n(x)$ 在点集 A 上一致收敛到和函数 $f(x)$ ”这个性质大大强于它在 A 上逐点收敛到 $f(x)$ ，后者的定义是： $\sum_{n=1}^{\infty} f_n(x)$ 在 A 中一点 x 收敛到数 $f(x)$ ，是指对于任意的正数 ε ，存在自然数 $N = N(x, \varepsilon)$ ，使得当 $n > N$ 时，

$$|\sum_{k=1}^n f_k(x) - f(x)| < \varepsilon.$$

由此可见，逐点收敛定义中的自然数 N 不仅依赖于 ε ，也依赖于 x 。至于一致收敛，在它的定义中，自然数 N 取值不依赖于 A 中的点 x ： $\sum_{n=1}^{\infty} f_n(x)$ 在 A 上一致收敛到 $f(x)$ ，若对任意的正数 ε ，存在自然数 $N = N(\varepsilon)$ ，使得当 $n > N$ 时，对 A 中所有的 x ，都有