Many Cheerful Facts: Divergent Series

Theo Johnson-Freyd

October 18, 2007

Abstract

Mathematicians through the ages have varied from terrified of divergent sums to only mildly scared of them: Euler, most famously, made great use of divergent series, whereas Abel called them "the invention of the devil". In this talk, I will survey the most important methods of summing divergent series, and make general vague remarks about them. I will quote many results, but will studiously avoid proving anything.

I should begin this talk with a disclaimer: (i) Everything I say in this talk is at least 50 years old. Indeed, this talk is more or less a book report: (almost) everything I say in this talk can be found in G.H. Hardy's last book *Divergent Series* (1949). [Hardy himself died Dec. 1, 1947; he had finished the galleys, but the final editing of the book was performed by Eggleston, Rogosinsky, Edmonds, and Bosanquet.] (ii) As such, I will understand this as a talk consisting of various cheerful facts, and I will try to avoid proving anything. (iii) Moreover, I am not in analysis: I find this material interesting, but I would not claim to have any expertise in the matter.

On the other hand, I am teaching Calculus 1B right now, and in that class we're studying *convergent* series. It's generally accepted that if the sequence of partial sums has a limit, then it's meaningful to talk about the sum of the infinite series. Although it was Cauchy who first correctly defined the limit of a sequence, even as far back as Archimedes mathematicians have worked with divergent series (which Abel called "the invention of the devil"), all the while knowing intuitively whether a series was convergent or divergent. In 1B, we declare that only those series whose sequences of partial sums converge are meaningful; in this talk I'd like to argue that some divergent series nonetheless have meaningful summations.

1 Preliminary remarks

Thus, let me begin with the favorite divergent series: the geometric series with ratio -1:

$$1 - 1 + 1 - 1 + 1 - 1 + \dots$$

It is well known that divergent infinite sums are not commutative: indeed, given even a *conditionally convergent* series, there is a rearrangement that sums to any value you choose. But this series is not even associative:

$$\begin{aligned} 1 &= 1+0+0+0+\dots \\ &= 1+(-1+1)+(-1+1)+(-1+1)+\dots \\ &= (1-1)+(1-1)+(1-1)+\dots \\ &= 0+0+0+\dots \end{aligned}$$

Convergent series are associative in the sense that if a non-parenthesized series has a limit, then so does any parenthesization, and to the same value. But the converse is not true: this example shows that a parenthesization can converge even when the original series diverges.

Nonetheless, reasonable formal manipulation (that Cauchy justifies on any convergent geometric series) yields a reasonable answer:

$$\begin{array}{rcl} S &=& (1-1+1-1+1-1+\dots) \\ +\ S &=& 0+(1-1+1-1+1-\dots) \\ \hline 2S &=& 1 \\ S &=& 1/2 \end{array}$$

This argument relied two properties that "summation" ought to have:

- Linearity: the sum of a series should be linear in the sequence of summands, or equivalently in the sequence of partial sums.
- Shift invariance: the series $a_0 + a_1 + a_2 + a_3 + \dots$ should equal the series $0 + a_0 + a_1 + a_2 \dots$.

I will often move back and forth between discussing a series as a sequence of summands and discussing a series as a sequence of partial sums. This is justified because there's an invertible (lower-triangular) matrix that shifts between them:

$$\sigma = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 0 & 0 & \\ 1 & 1 & 1 & 0 & \\ 1 & 1 & 1 & 1 & \\ \vdots & & & & \ddots \end{bmatrix} \qquad \sigma^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ -1 & 1 & 0 & 0 & \\ 0 & -1 & 1 & 0 & \\ 0 & 0 & -1 & 1 & \\ \vdots & & & \ddots & \ddots \end{bmatrix}$$

In any case, Cesàro provides a method through which to justify the sum $1/2 = 1-1+1-1+\dots$: replace the sequence of partial sums $1, 0, 1, 0, \dots$ by its corresponding sequence of averages $1, 1/2, 2/3, 1/2, \dots$. More generally, we can multiply a sequence of partial sums by the matrix

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \\ \vdots & & & & \ddots \end{bmatrix}$$

If the original sequence converges, then so does the new sequence, and to the same limit; if the new sequence converges, we can call its limit the "Cesàro limit" of the original sequence.

I'm writing in terms of matrices acting on column vectors to suggest that this is a good way to look for methods of summing divergent series: act on the sequence of partial sums by a matrix, and try to take the limit of that. It it natural to ask whether a given matrix M is "regular": given a convergent sequence s_n, is the new sequence $Ms = \sum_{j=0}^{\infty} M_n^j s_j$ necessarily convergent, and to the same limit?

I will state without proof the answer. For a matrix to be regular, the following conditions are necessary and sufficient: (a) the sum of each row $m_n = \sum_{j=0}^{\infty} M_n^j$ is absolutely convergent; (b) the absolute sums $m_n^* = \sum_{j=0}^{\infty} |M_n^j|$ are bounded by some K not depending on n; (c) the limit of the sums m_n tends to 1 as $n \to \infty$; and (d) the columns M_n^j (for fixed j) tend to 0 as $n \to \infty$.

There are, of course, methods of assigning limits to infinite sequences that do not arise as matrix methods. For instance, any real-valued sequence $s : \mathbb{N} \to \mathbb{R}$ extends uniquely to a function $\tilde{s} : \beta\mathbb{N} \to \overline{\mathbb{R}}$, where $\beta\mathbb{N}$ is the Stone-Cěch compactification of $\mathbb{N}$ and $\overline{\mathbb{R}}$ is the standard two-point compactification of $\mathbb{R}$. Then by picking an infinite point $\mathcal{N} \in \beta\mathbb{N} \setminus \mathbb{N}$ and evaluating $\tilde{s}(\mathcal{N})$, we get either a finite value or one of $+\infty$ or $-\infty$; if $s_n \to l$ as $n \to \infty$, then $\tilde{s}(\mathcal{N}) = l$, so in general $\tilde{s}(\mathcal{N})$ deserves to be called the $\mathcal{N}$-limit of s. But $\tilde{s}(\mathcal{N})$ is always finite if s is bounded, whereas it is a theorem that any matrix which sends all bounded sequences to convergent sequences cannot respect the limits of convergent sequences. (A matrix M that sends bounded sequences to convergent sequences necessarily has $\lim_{n\to\infty} \sum_{j=0}^{\infty} M_n^j = \sum_{j=0}^{\infty} \lim_{n\to\infty} M_n^j$.)

For the rest of this talk I will restrict my discussion to summation methods representable as matrices. Indeed, I will be primarily interested in methods representable as lower-triangular matrices.

2 Nörlund Means and power series

The Cesàro method, by which we averaged the first n terms of the sequence of partial sums, is a good one, and immediately suggests a generalization: let's allow different weights in our averages. Specifically, Nörlund (well, Voronoi before him) suggests, let's pick a sequence $w_0, w_1, w_2, \ldots$ of non-negative (and such that $w_0 > 0$) "weights", and turn any sequence of partial sums $s_0, s_1, \dots$ into the sequence $t_0, t_1, \dots$, where $t_n = (s_0 w_n + s_1 w_{n-1} + \dots + s_n w_0)/(w_0 + w_1 + \dots + w_n)$. As a matrix, we multiply:

$$\begin{bmatrix} t_0 \\ t_1 \\ t_2 \\ t_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} \frac{1}{w_0} & 0 & 0 & 0 & \cdots \\ 0 & \frac{1}{w_0+w_1} & 0 & 0 & \\ 0 & 0 & \frac{1}{w_0+w_1+w_2} & 0 & \\ 0 & 0 & 0 & \frac{1}{w_0+w_1+w_2+w_3} & \\ \vdots & & & & \ddots \end{bmatrix} \begin{bmatrix} w_0 & 0 & 0 & 0 & \cdots \\ w_1 & w_0 & 0 & 0 & \\ w_2 & w_1 & w_0 & 0 & \\ w_3 & w_2 & w_1 & w_0 & \\ \vdots & & & & \ddots \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ \vdots \end{bmatrix}$$

An easy calculation shows that such a method is regular if an only if $w_n/(w_0 + w_1 + \dots + w_n) \to 0$, and that any two regular Nörlund methods are consistent, in the sense that if they both assign a "limit" to the same sequence, then they assign the same limit.

Indeed, if a regular Nörlund matrix makes a series $a_0 + a_1 + a_2 + \dots$ sum to s, then the corresponding power series $a(x) = a_0 + a_1 x^1 + a_2 x^2 + \dots$ (i) has positive radius of convergence, (ii) this function of x analytically extendable in an open neighborhood of the half-open interval $[0, 1)$, and (iii) as $x \to 1$ through real values, $a(x) \to s$. Thus, Nörlund means fail to sum series that diverge faster than any exponential.

Abel suggested that the only divergent series he was willing to consider were those whose corresponding power series had radius of convergence equal to 1, in which case he would take the limit as $x \to 1$ of $a_0 + a_1 x^1 + a_2 x^2 + \dots$. This corresponds to using the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \cdots \\ 1 & \frac{2}{3} & \frac{4}{9} & \frac{8}{27} & \frac{16}{81} & \cdots \\ 1 & \frac{3}{4} & \frac{9}{16} & \frac{27}{64} & \frac{81}{256} & \cdots \\ 1 & \frac{4}{5} & \frac{16}{25} & \frac{64}{125} & \frac{256}{625} & \cdots \\ \vdots & & & & & \ddots \end{bmatrix} \sigma^{-1}$$

(The σ^{-1} converts sequences of partial sums to sequences of summands.) As is easily checked, this is a regular method. However, care must be taken when following Abel's philosophy: rather than using $a_0 + a_1 x^1 + a_2 x^2 + \dots$, why not $a_0 x^{d_0} + a_1 x^{d_1} + a_2 x^{d_2} + \dots$ for some other choice of a (strictly increasing) sequence of exponents d_n? For instance, if $d_n = 1, 2, 4, 5, 7, 8, \dots$, skipping every multiple of 3, then we would expect $1 - 1 + 1 - 1 + \dots = x - x^2 + x^4 - x^5 + \dots \to 1/3$. Indeed, we can get any fraction m/n by expanding $(1 - x^m)/(1 - x^n) = 1 - x^m + x^n - x^{n+m} + x^{2n} - x^{2n+m} + \dots$, and any sequence of d_n gives a regular method. What's going on here, of course, is that the series is getting spaced out. Thus, the failure of associativity remarked upon earlier is even stronger than imagined:

$$1 - 1 + 1 - 1 + 1 - 1 + \dots \neq 1 - 1 \quad + \quad 1 - 1 \quad + \quad 1 - 1 \quad + \quad \dots$$

I think it would be interesting to study associativity in this light: where the distance between operations determines how much you do one before the other, and full-out parenthesizing corresponds to introducing an infinite distance.

One weak consistency theorem does hold: any methods with exponents $d_n = n^k$ for fixed k are consistent. Any such method will sum $1-1+1-1+\ldots$, whereas $d_n = 2^n$ will not. In general, the more powerful a method is, in the sense of summing more violently divergent series, the less subtle it is, in that it will not sum slowly divergent series.

3 Cesàro means and multiplication

Cesàro's mean

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \\ \vdots & & & & \ddots \end{bmatrix}$$

will sum $1-1+1-1+\ldots$ to $1/2$, but fails to sum $1-2+3-4+\ldots$. On the other hand, if we take averages twice, we do get a converging sequence, tending to $1/4$. In general, we could try to sum various series by repeatedly applying C.

However, C^k, for $k > 1$ is no longer a Nörlund mean, and we might worry that it would be inconsistent. $C = \nu\sigma$, where ν is the diagonal matrix $1/n$ and σ is as before:

$$\nu = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & \frac{1}{2} & 0 & 0 & \\ 0 & 0 & \frac{1}{3} & 0 & \\ 0 & 0 & 0 & \frac{1}{4} & \\ \vdots & & & & \ddots \end{bmatrix} \qquad \sigma = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 0 & 0 & \\ 1 & 1 & 1 & 0 & \\ 1 & 1 & 1 & 1 & \\ \vdots & & & & \ddots \end{bmatrix}$$

Then $C^k = \nu\sigma\nu\sigma\ldots\nu\sigma$, alternately summing and dividing. We could instead do all our summing and then a single division: $C_k = \mu_k\sigma^k$, where μ_k is the diagonal matrix so that $\mu_k\sigma^k$ has the sequence $1, 1, 1, 1, \ldots$ as an eigensequence: the nth term in μ_k is $k!/(n+1)^k$. C_k, the "kth Cesàro mean", is clearly a Nörlund mean. Moreover, it is a theorem that C_k and C^k are equivalent: if one sums a series, then so does the other, and to the same value. The C_k and C^k methods are weaker than Abel's method: if $C_k \cdot \sigma a$ converges, for a sequence of summands a_n, then $a_0 + a_1x + a_2x^2 + \ldots$ has radius of convergence at least 1, and the limit of the C_k method is the limit as $x \to 1$ of $a_0 + a_1x + a_2x^2 + \ldots$.

The Cesàro methods are important, though, for a few reasons. In particular, they have a good theory of multiplication.

Of course, the linearity of all our methods assures us that we can multiply series by finite sums. But given $\sum a_n = A$ and $\sum b_m = B$, we would like a prescription through which we can write down a series $\sum c_l$ that sums to AB. If $\sum a_n$ and $\sum b_m$ are absolutely convergent (so that even commutativity is preserved), then the double sum $\sum\sum a_n b_m$ can be rearranged arbitrarily and still converge to the correct limit.

Cauchy provides the right prescription. He says that $(\sum a_n)(\sum b_m) = \sum c_l$, with $c_l = a_0 b_l + a_1 b_{l-1} + \cdots + a_l b_0$, suggested by formally multiplying the corresponding power series (which converge absolutely within their radii of convergence, and so the product is well-defined) and combining like terms. This is an effective method because if the three sums $\sum a_n = A$, $\sum b_m = B$, and $\sum c_l = C$ each converge, then $AB = C$.

This can be proven with the theory of analytic functions, but an elementary proof follows from the following theorem: if $\sum a_n$ can be summed via the C_k method, and $\sum b_m$ via the C_j method, then $\sum c_l$ can be summed via the C_{k+j+1} method to the product of the sums.

Thus, not only is it reasonable to multiply divergent series and manipulate the products to achieve "correct" summations, but, moreover, we should never worry about multiplying two convergent series: the product might not converge, but it will necessarily converge after taking averages just once.

4 Hausdorff means and concluding remarks

By toying with sequences, you learn that a good way to study a sequence is by "Taylor expansion": write down the first term of the sequence, and then the (negative of the) first term of the sequence of differences, the first "second derivative", the (negative of the) first "third derivative", and so on. As a matrix, this is the alternating Pascal's triangle:

$$\rho = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ 1 & -1 & 0 & 0 & \\ 1 & -2 & 1 & 0 & \\ 1 & -3 & 3 & -1 & \\ \vdots & & & & \ddots \end{bmatrix}$$

It is an involution: $\rho^2 = 1$. Moreover, it diagonalizes the Cesàro means: $C_k = \rho\lambda_k\rho$, where the nth term in the (diagonal matrix) λ_k is $\binom{n+k}{k}^{-1}$. In general, by a "Hausdorff matrix", I will mean any diagonal matrix conjugated by ρ.

Being conjugates of diagonal matrices, any two Hausdorff matrices commute, and thus their corresponding summation methods are consistent in the strongest sense. I should mention the conditions for a Hausdorff matrix to be regular. As always, I quote without proof: the matrix $\rho\lambda\rho$, where λ is the diagonal matrix given by the sequence l_n, is regular if and only if (i) the absolute sums $L_m = \sum_{n=0}^{m} \binom{m}{n} |(\sigma^{n-m} l)_n|$ are uniformly bounded (i.e. $L_m < K$ for some K not depending on m), (ii) the sequence $\rho l \to 0$, and (iii) $l_0 = 1$.

With this formalism, it's now immediate to define Euler's summation method, depending on the positive real parameter q: he used the matrix $\rho\lambda_q\rho$ where λ_q is the diagonal with

$(q+1)^{-n}$ in the nth spot. Euler discovered this method by expanding $\sum_{n=0}^{\infty} a_n \left(\frac{y}{1-qy}\right)^{n+1}$ in powers of y, and sending $y \to 1/(1+q)$; the latter can converge when $\lim_{x\to 1} \sum_{n=0}^{\infty} a_n x^{n+1}$ does not (because the radius of convergence is too small). For instance, Euler's method sums $1 + z + z^2 + z^3 + \dots$ inside the circle centered at $z = -q$ with radius $q + 1$. These are all consistent (and, indeed, commutative), and so, as $q \to \infty$, we can sum $\sum z^n$ in the entire half plane $\Re(z) < 1$.

Euler's method when applied to slowly convergent series tends to make them converge more quickly, and makes many divergent series converge. Even when it doesn't succeed at summing a divergent series, it often makes divergent series diverge less rapidly. For instance, for no q does Euler's method make $1 - 1! + 2! - 3! + \dots$ converge, but Euler nonetheless used it to approximate $1 - 1! + 2! - 3! + \dots = -e\left(\gamma - 1 + \frac{1}{2\cdot 2!} - \frac{1}{3\cdot 3!} + \dots\right) \approx .5963$. Euler defined $f(x) = x\sum_{n=0}^{\infty} n!(-x)^n$, which has an essential singularity at the origin, and observed that $xf'(x)+f(x) = x$; he found the unique solution so that $f(x)/x \to 1$ as $x \to 0$ along the positive reals, and used his q method to make the solution he got from this converge faster.

In any case, for different series, we need to use different methods to make them converge. No matrix method works for everything — indeed, methods that sum violently divergent series tend to fail to sum slowly divergent series. But we can find methods that are all consistent, and these even have a theory of convergence.

Is this the correct way to make series converge? Consistency, to some, is a proof, but why not, for instance, use the Stone-Cěch method mentioned at the beginning of the talk? The answer, I think, comes from asking precisely what structure we're interested in our generalized limits respecting. Inherent in the idea of a "sequence" is the topological structure: Stone-Cěch is the right way to make sequences converge, because it respects continuous functions. But inherent in a *series* is the arithmetic structure of $\mathbb{R}$. We do lose some arithmetic: divergent series are not associative or commutative, but they fail to be so in interesting and sensitive ways. We should look for methods that respect as much arithmetic as possible, and in fact there are some — the Cesàro matrices, and any matrices that are consistent with them — that even respect multiplication.