Research into limits of divergent series has revitalised many areas of physics – as well as studies of asymptotic expansions themselves

Infinity interpreted

MICHAEL BERRY AND CHRISTOPHER HOWLS

Divergent series are the invention of the devil, and it is shameful to base on them any demonstration whatsoever – **Abel, 1828**

DESPITE the denunciations of the mathematician Abel, if the devil did invent divergent series it was because his creator counterpart chose to build our physical universe so that they are among the more useful ways to describe its finite properties.

It is common to consider physical situations as perturbations of an idealised problem that can be solved exactly, and get the solution of the real problem as a series in powers of a small parameter δ describing the strength of the perturbation:

$$S = \sum_{n=0}^{\infty} a_n \delta^n.$$

Sometimes these series converge (i.e. beyond a certain order n the magnitude of successive terms approaches zero and the sum approaches a finite value or asymptote); there is a prejudice in favour of such solutions. But the limit $\delta \rightarrow 0$ often turns out to be singular and then difficulties can arise: finite physical quantities in well founded theories can be represented by series that diverge, with successive terms beyond a certain order increasing in magnitude so that the sum tends to infinity.

Physicists have been uneasy about these series, but have tolerated them because truncations after the first few terms give very good approximations to observed quantities. Only recently has it been appreciated that the tails of divergent series have universal properties. These have allowed us to tame the tails and reveal a rich structure of exponentially small terms responsible for a variety of physical phenomena, and moreover answer questions in mathematical asymptotics – the study of limits – that have endured since the nineteenth century.

Reduction of theories

Limits arise in physics at a fundamental level because $\delta \rightarrow 0$ can describe the reduction of a general theory to a more restricted (usually earlier) theory. A case where the limit is unproblematic is special relativity, which reduces smoothly to Newtonian mechanics for objects with speed v small compared with the speed of light c. To see this we set $\delta = v/c$. Wherever a Lorentz factor $\gamma = 1/\sqrt{(1-\delta^2)}$ appears in an equation of special relativity it may be expanded as a convergent Taylor series in δ to give the corresponding low-speed formulae, reducing to Newtonian mechanics when $\delta = 0$.

Contrast this with the geometrical (ray) limit of physical optics, for example in the rainbow (figure 1 and front cover); in this limit, for each colour the wavelength is small compared with the radius a of a raindrop (we could choose $\delta = \lambda/a$), so that diffraction can be ignored. Imagine a bow at a single wavelength. The region under the bow is bright because two rays emerge in each direction. No rays reach the region above the bow – the dark side. The boundary is a caustic, where rays coalesce and the bow is most intense.

The illumination on the bright side was calculated by Young in 1804, who moved beyond the geometric optics limit by using an approximate theory in which wave energy travels along the rays. He was able to explain the observation that, beneath the bow, there is a series of "supernumerary" bows that arise from interference oscillations. (These effects are often washed out by waves of different colours and from droplets of different sizes, but supernumerary bows can be discerned beneath the main bow in figure 1.)

The oscillations can be represented as the usual two-wave

1 Supernumerary rainbows (faint fringes below the main arc) above Newton's birthplace (Woolsthorpe Manor, Lincolnshire). The irony is that Newton could not have understood these fringes with his corpuscular theory of light, because they are the result of wave interference. (Photograph courtesy of Professor Roy Bishop, Acadia University, Nova Scotia, Canada)

Airy's rainbow integral

The Airy integral is defined by

$$\mathrm{Ai}(z) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}t \exp\left\{\mathrm{i}\left(\frac{1}{3}t^3 + zt\right)\right\} \qquad (1)$$

when z is real, and by a similar integral with a modified contour when z is complex. Stokes found that when z is large and positive a formally exact representation is the series

$$\mathrm{Ai}(z) = \frac{\exp\{-\frac{2}{3}z^{3/2}\}}{2z^{1/4}\sqrt{\pi}}\sum_{r=0}^{\infty}(-1)^r T_r, \qquad (2)$$

with coefficients

$$T_r = \frac{1}{(36z^{3/2})^r}\frac{(3r-\frac{1}{2})!}{r!(r-\frac{1}{2})!} \qquad (3)$$

The high orders have the form

$$T_r \approx \frac{(r-1)!}{2\pi(\frac{4}{3}z^{3/2})^r}, \quad (r \gg 1). \qquad (4)$$

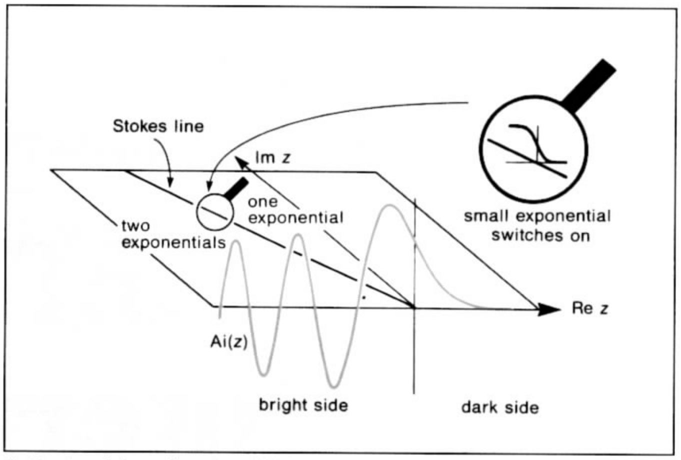
When z is large the terms decrease at first and then increase, so that the series (2) diverges. The least term is near $r = r^* = 4z^{3/2}/3$. Truncating at this order and summing the divergent tail (using Borel summation) generates a finite representation valid throughout the upper z half-plane and incorporating the second exponential and the Stokes phenomenon:

$$\mathrm{Ai}(z) = \frac{\exp\{-\frac{2}{3}z^{3/2}\}}{2z^{1/4}\sqrt{\pi}}\sum_{r=0}^{r^*-1}(-1)^r T_r + \mathrm{i}S(z)\frac{\exp\{+\frac{2}{3}z^{3/2}\}}{2z^{1/4}\sqrt{\pi}}\sum_{r=0}^{r^*-1} T_r. \qquad (5)$$

Here $S(z)$ switches rapidly but smoothly from 0 to 1 across the Stokes line $z = |z|\exp\{2\mathrm{i}\pi/3\}$ (figure 2), thereby generating the second exponential, which is subdominant there, and, for negative real z, the supernumerary rainbow oscillations. A good approximation to $S(z)$ involves the error function. □

fringe pattern with intensity $I = \sin^2(2\pi z/\lambda)$ where z is a rainbow-crossing coordinate. This is mathematically singular as $\lambda \to 0$, because the sine oscillates infinitely fast, and a smooth description, reproducing geometrical optics, can only be achieved by averaging. On the dark side, there is a different singularity, typical where waves reach places that rays do not, namely exponential decay: $I = \exp(-Az/\lambda)$ where $A > 0$. Across the caustic $z = 0$, something peculiar happens: one real exponential (on the dark side) has transformed into two complex ones (comprising the sine on the bright side). How has the second exponential appeared?

Sudden appearances of exponential terms are common in physical asymptotics. For example in radioactive decay the above formulae give semiclassical approximations to the quantum tunnelling of a particle from a potential well (inside the nucleus) into a barrier. The character of the wave changes from oscillatory to decaying at the classical turning point (where the particle momentum is zero).



2 Airy's rainbow function Ai(z), with small exponential switching on across Stokes line in the complex plane

The above approximations based on exponentials are supposed to be valid in the asymptotic regime where δ can be regarded as small (i.e. the wavelength or, in the quantum case, Planck's constant, are small compared with characteristic parameters of the problem). But they fail at caustics and turning points, obscuring the transitions to geometrical optics (i.e. from waves to rays) and classical mechanics.

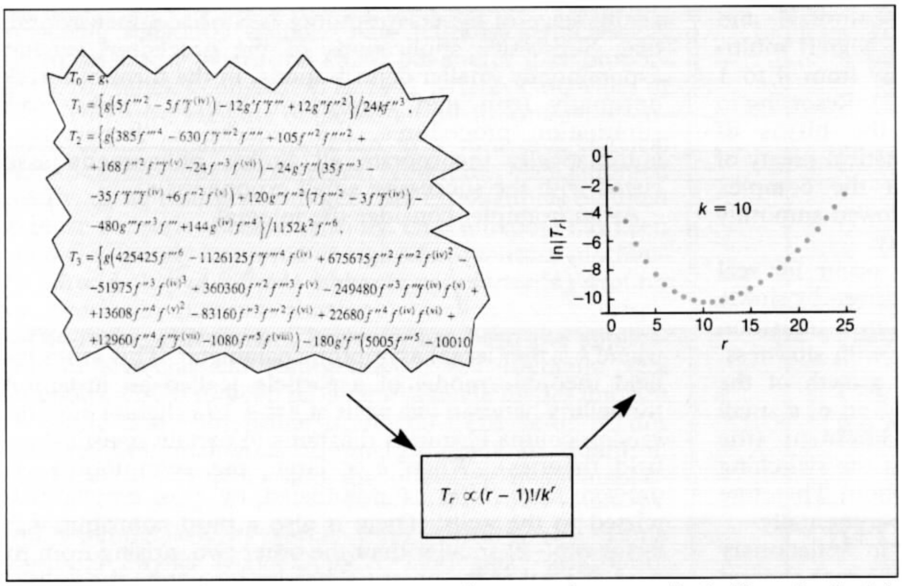
Among crossovers between theories in physics, smooth transitions – as with special relativity – are exceptional; usually the limits are singular. Other examples are the transitions from statistical mechanics to thermodynamics, where the singularity describes critical phenomena, and from viscous to inviscid flow, where it involves turbulence.

This is not to say that the theories are necessarily incompatible – each is self-consistent and agrees well with observation in its own domain. Singularities mean that on the borderline between the theories there is a wall impenetrable to elementary mathematical interpretation. Only now are nonelementary interpretations being devised, making some of these walls transparent.

Stokes and the rainbow

In 1838, Airy devised a diffraction integral $\mathrm{Ai}(z)$ (see box) representing, for each colour, the variation of the wave across a rainbow (the intensity is $\mathrm{Ai}^2(z)$). He was, however, unable to compute $\mathrm{Ai}(z)$ far on the bright side (z large and negative) and so compare his theory with measurements of the angles of supernumerary bows (and make the connection with geometrical optics). This was because, although he had found a convergent series, the number of terms needed to get a satisfactory approximation increased with $|z|$, and the integral itself converged very slowly. This situation is similar to that in quantum chromodynamics today, where technical difficulties inhibit the calculation of observed quantities.

In 1847, Stokes devised approximations for $\mathrm{Ai}(z)$ in the desired asymptotic regions $|z| \gg 1$ (that is, far from the caustic at $z = 0$), and showed that the function behaves as a damped exponential on the dark side (positive z) and sinusoidally on the bright side. The accuracy increased with $|z|$, and was sufficient to enable comparison with measured supernumerary bows. Stokes attacked the problem posed by the different numbers of exponentials



3 Dingle's universal "factorial divided by power" emerges from complicated formulae for late terms of an asymptotic series (in this case for an integral with a saddle)

on the two sides. He realised that the key lies in regarding z as a complex variable (figure 2). The second exponential appears across certain lines (now called Stokes lines) on a path in the z plane from the dark to bright sides (positive to negative real axes) avoiding the caustic singularity at $z=0$. On Stokes lines, the nascent exponential is smallest (subdominant) relative to the other. The birth of the subdominant exponential is called the Stokes phenomenon. (On the bright side, the second exponential becomes the second ray in the interfering pair.)

A major feature of Stokes' analysis was his appreciation that the pure exponentials in his approximation should each be multiplied by a series of corrections (see box) that are powers of $1/z$. He obtained these by solving the differential equation that Ai(z) satisfies, anticipating the WKB method of wave mechanics. The resulting asymptotic series was expected to give better approximations for larger $|z|$. As it turns out, successive terms do get smaller initially, and the accuracy is indeed improved by including them, but then the terms begin to increase, and the series diverges. For numerical purposes the optimal procedure is to truncate the series at its least term.

However, something has gone wrong, because since the series diverges its numerical sum would be infinity rather than Ai(z). The mathematical description of the physics is formally exact – there has been no approximation in the calculation of the correction terms – but we need to interpret the infinite result to rescue something sensible. Stokes appreciated that the divergence of series representing physics in asymptotic regimes is not peculiar to Airy's rainbow integral but is a general phenomenon.

It turns out that, far from being a hindrance, the divergence of the asymptotic series, suitably interpreted, actually explains the Stokes phenomenon. In his prescient but crude analysis Stokes thought that the birth of the second exponential is a sudden event; in fact it occurs smoothly, but it was more than a century before the details were discovered.

Universal divergences

Mathematicians responded to Stokes' work by ignoring it. The theory of asymptotic series developed along different lines, initiated by Poincaré in the 1880s. For each function possessing an asymptotic series, only a fixed number of terms is retained, and the divergent tail discarded and replaced by an exact remainder. The accuracy of the truncation can then be gauged by an estimate of the remainder, which has to be calculated separately for each function. In this work, divergent and convergent series are treated on the same basis, so the main issue of divergence is side-stepped.

In particular, Poincaré asymptotics fails to achieve the accuracy of truncation at the least term. For this reason, we call Stokes' technique of optimal truncation "superasymptotics" (reflecting Kelvin's description of it as "mathematical supersubtlety"). But even superasymptotics cannot provide a detailed description of exponential effects such as the Stokes phenomenon. For this, we need what Martin Kruskal of Rutgers University calls "asymptotics beyond all orders".

In the 1950s, R B Dingle, a physicist at St Andrews, returned to the original ideas of Stokes and confronted the divergences head-on, introducing ideas which have now led to a great simplification in their interpretation. He regarded the divergent tail of a formal expansion not as indicating lack of precision but rather as a source of information to be decoded to reveal the remainder. Dingle traced the divergences of asymptotic series to singularities (for example the caustic at $z=0$ in the rainbow). This led him to the result that whole classes of singularities arising in physics determine the form of the high-order (late) terms in a universal and simple way. This is astonishing because the early terms are not universal and are usually very complicated (figure 3). If k (proportional to $1/\delta$) is the large asymptotic parameter in such a series then Dingle's common form for the late terms T_r is

$$T_r \propto (r-1)!/k^r, \qquad r \gg 1.$$

An example is given in the box (there, $k=4z^{3/2}/3$).

For large k the terms rapidly diminish at first, but the factorial in the numerator eventually dominates the power in the denominator, and the terms increase unboundedly. The least term has $r \approx |k|$ and size $\exp(-|k|)$, i.e. exponentially small. This is a common pattern of divergence, not restricted to the Airy and other mathematical functions but also describing – as was later discovered – the divergences of series with much more complicated structure, such as those in quantum field theory (where, for example, δ could be the fine-structure constant).

Dingle used his late term approximations in conjunction with a procedure called Borel summation, in which the factorial is replaced by a well known integral representation. Thus he interpreted the infinity and obtained a *finite* representation of the divergent tail directly, in the form of readily calculable integrals called terminants. These extend beyond superasymptotics, and constitute a decoding of the information contained in the divergent tail. The universality of the late terms is mirrored in the universality of the terminants, which depend only on k and the order of the term where the series is truncated.

Terminants can also be adapted to give a refined description of the Stokes phenomenon. In 1988 we showed that if an asymptotic series is optimally truncated, i.e. near its least term $r \approx |k|$, the terminant reproduces the subdominant exponential switching on across the

Stokes line. The form of the switching is universal, and consists of an error function (probability integral) multiplying the exponential, and rising smoothly from 0 to 1 over a range proportional to $1/\sqrt{|k|}$ (figure 2). Resorting to paths in imaginary space to explain the births of exponentials might appear to be a mathematical nicety of no relevance to the physical world, but the complex approach enables the asymptotics to be followed smoothly around a seemingly impenetrable singularity.

Moreover, the Stokes phenomenon can occur for real variables. For example, a quantum system driven by slowly changing forces can jump between states, in a transition whose probability diminishes exponentially with slowness. The history of the transition, that is the growth of the occupancy of the new state, is the switching-on of a small exponential. Using an appropriate measurement (the physical counterpart of optimal truncation) the switching would conform to the universal error function. Therefore the Stokes smoothing could be studied experimentally.

The Stokes phenomenon also occurs in inflationary models of cosmology. E Calzetta (Buenos Aires) showed that in a "mixmaster" model of the early Universe the smooth creation of exponentially small numbers of particles is represented by universal error functions. A multitude of such smooth creations stokes the inflation of the Universe.

Resurgence and hyperasymptotics

The simple "factorial divided by power" formula for T_r is only a first approximation to the late terms. Dingle also found an asymptotic series for the corrections. A remarkable and beautiful feature is that Dingle's series connects the terms in the different asymptotic series representing a given function. For example, the late terms of the asymptotic series multiplying a dominant exponential can be expressed systematically by the early terms of a subdominant series, and vice versa. For functions with many exponentials (such as integrals in diffraction theorv or statistical mechanics, whose saddles represent rays or equilibrium phases), all associated asymptotic series are similarly related. These formulae are called resurgence relations, after a related idea developed in the 1980s by the mathematician Jean Ecalle of Orsay.

By the sequential use of resurgence relations and Borel summation, it is now possible to get very accurate reconstructions of functions from their divergent series. This is a systematic extension of superasymptotics, developed since 1990; we call it "hyperasymptotics". The zeroth stage of hyperasymptotics is Stokes' superasymptotics. Successive applications of the procedure generate exponentially smaller contributions, in the form of further optimally truncated asymptotic series. Unlike other summation procedures, the successive "hyperseries" automatically incorporate all Stokes phenomena associated with the successive small exponentials.

As an example, consider the integral

$$N(k) = \sqrt{k}\int_{-\infty}^{+\infty} \mathrm{d}z \exp\left(-k(z^2-1)^2\right), \quad k > 0$$

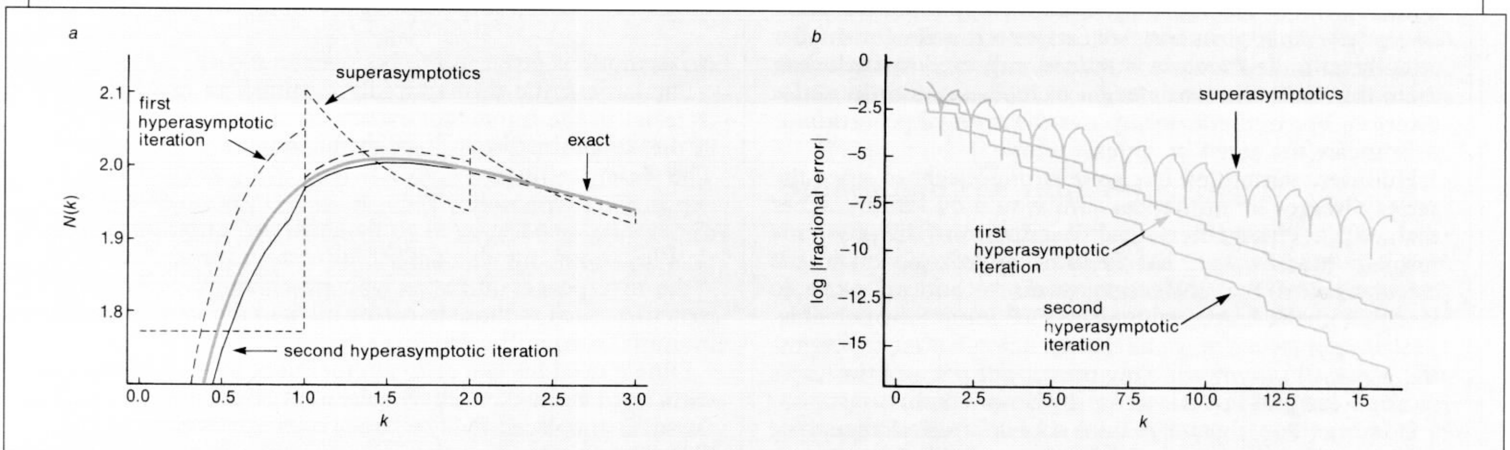
where k is the (large) asymptotic parameter. This arises in a field theoretic model of a particle (called an instanton) tunnelling between two wells at $z = \pm 1$ (a slightly modified version counts Feynman diagrams in certain statistical and field theories). When k is large, the asymptotic series derived from $N(k)$ is dominated by two exponentials, related to the wells. There is also a third contribution, a factor $\exp(-k)$ smaller than the other two, arising from the peak at $z = 0$ of the potential barrier separating the wells. It is this small exponential that can be physically interpreted as the tunnelling of an instanton through the barrier.

Hyperasymptotics can discriminate between the three contributions (figure 4) and generate accurate approximations even when the "large parameter" k is smaller than unity, where Poincaré asymptotics and superasymptotics certainly fail. Moreover this integral lies precisely on a Stokes line, where the instanton is created, which is a regime conventionally regarded as not amenable to Borel summation.

Recent applications

Small exponentials are common in nonlinear physics, and it is perhaps here that the new theories could be most fruitful. For example, Kruskal and others have recently used exponential asymptotics to elucidate the influence of surface tension on the selection and stability of viscous fingers when one fluid is forced into another, and to analyse a geometric model for dendritic growth of crystals. And in chaology Vincent Hakim of the Ecole Normale Supérieure in Paris and others have summed divergent perturbation series in accurate calculations of the exponentially small widths of the chaotic regions that open up when a non-chaotic system is perturbed.

A notable success of the new approach has emerged from quantum chaology, where a long-standing problem is to calculate the energies of high-lying states of microscopic



4 Superasymptotics and several levels of hyperasymptotics for the instanton integral systems whose classical limit is chaotic (e.g. atomic electrons in strong magnetic fields). This requires semiclassical approximations, where the small parameter δ is proportional to Planck's constant $\hbar$. In 1971 Martin Gutzwiller of IBM Yorktown Heights obtained an asymptotic representation of the quantum spectrum as a sum over classical periodic orbits, but this diverges. Now it has been summed exactly, using resurgence to generate exponentials required to make the series behave sensibly. One outcome has been to provide greatly improved energies for quantum mechanics on odd-shaped billiard tables (a natural testing-ground for theories in quantum chaology).

The main achievement, however, has been the application of semiclassical resurgence to the Riemann zeta function (related to deep problems of mathematics through its connection with prime numbers). The zeros of this function are analogous to energy levels in quantum chaology (see "Physics and the queen of mathematics" by Jonathan Keating *Physics World* April 1990 pp46–50, and "Semiclassicists saved from infinity" by Tania Monteiro *Physics World* October 1991 pp21–22). Until recently, the zeta function was computed by a technique – the Riemann–Siegel formula – which, although very accurate, involves discontinuities detrimental to the computation of zeros. However, the Riemann zeta analogue of the new quantum formula is readily calculable, involves no discontinuities, and gives, term for term, approximations orders of magnitude better than its Riemann–Siegel cousin.

Prospects

Asymptotics was conceived in Victorian times, when it provided the fastest and most accurate means of approximation. Now that workstations with powerful numerical routines (e.g. for integration) are standard desktop furniture, the need for some of the purely numerical applications of asymptotics has diminished. On the other hand, computer algebra now makes it practicable to evaluate many terms of asymptotic series which were previously impenetrably complicated, thereby opening new areas of application. Moreover, the computational role of asymptotics is being supplemented by their vital use as an analytical tool in understanding the behaviour of physical systems with small exponentials.

After a sudden emergence – reminiscent of chaology in the 1970s – asymptotics is now perceived as a common thread in many different fields. This has been reflected in at least three recent international conferences, and a six-month workshop being planned for 1995 at the Isaac Newton Institute in Cambridge.

Further reading

M V Berry and C J Howls 1991 Hyperasymptotics for integrals with saddles *Proc. R. Soc. Lond.* **A434** 657–75

M V Berry and J P Keating 1992 A new approximation for $\zeta(1/2+it)$ and quantum spectral determinant *Proc. R. Soc. Lond.* **A437** 151–73

R B Dingle 1973 *Asymptotic Expansions: Their Derivation and Interpretation* (Academic Press, New York and London)

R Lim and M V Berry 1991 Superadiabatic tracking for quantum evolution *J. Phys. A: Math. Gen.* **24** 3255–64

H Segur and S Tanveer (eds) 1991 *Asymptotics Beyond All Orders* (Plenum, New York)

Michael Berry and **Christopher Howls** are in the Department of Physics, H H Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK