Acta Applicandae Mathematicae **56:** 1–98, 1999.

The Devil's Invention: Asymptotic, Superasymptotic and Hyperasymptotic Series ⋆

JOHN P. BOYD
University of Michigan, Ann Arbor, MI 48109, U.S.A.

(Received: 30 July 1996; in revised form: 11 September 1998)

Abstract. Singular perturbation methods, such as the method of multiple scales and the method of matched asymptotic expansions, give series in a small parameter ε which are asymptotic but (usually) divergent. In this survey, we use a plethora of examples to illustrate the cause of the divergence, and explain how this knowledge can be exploited to generate a 'hyperasymptotic' approximation. This adds a second asymptotic expansion, with different scaling assumptions about the size of various terms in the problem, to achieve a minimum error much smaller than the best possible with the original asymptotic series. (This rescale-and-add process can be repeated further.) Weakly nonlocal solitary waves are used as an illustration.

Mathematics Subject Classifications (1991): 34E05, 40G99, 41A60, 65B10.

Key words: perturbation methods, asymptotic, hyperasymptotic, exponential smallness.

> "Divergent series are the invention of the devil, and it is shameful to base on them any demonstration whatsoever."
>
> Niels Hendrik Abel, 1828

⋆ This work was supported by the National Science Foundation through grant OCE9119459 and by the Department of Energy through KC070101.

1. Introduction

Divergent asymptotic series are important in almost all branches of physical science and engineering. Feynman diagrams (particle physics), Rayleigh–Schrödinger perturbation series (quantum chemistry), boundary layer theory and the derivation of soliton equations (fluid mechanics) and even numerical algorithms like the 'Nonlinear Galerkin' method [66, 196] are examples. Unfortunately, classic texts like van Dyke [297], Nayfeh [229] and Bender and Orszag [19], which are very good on the *mechanics* of divergent series, largely ignore two important questions. First, why do some series diverge for all nonzero ε where ε is the perturbation parameter? And how can one break the 'Error Barrier' when the error of an optimally-truncated series is too large to be useful?

This review offers answers. The roots of hyperasymptotic theory go back a century, and the particular example of the Stieltjes function has been well understood for many decades as described in the books of Olver [249] and Dingle [118]. Unfortunately, these ideas have percolated only slowly into the community of derivers and users of asymptotic series.

I myself am a sinner. I have happily applied the method of multiple scales for twenty years [67]. Nevertheless, I no more understood the reason why some series diverge than why my son is lefthanded.

In this review, we shall concentrate on teaching by examples. To make the arguments accessible to a wide readership, we shall omit proofs. Instead, we will discuss the key ideas using the same tools of elementary calculus which are sufficient to derive divergent series.

In the next section, we begin with a brief catalogue of physics, chemistry and engineering problems where key parts of the answer lie 'beyond all orders' in the standard asymptotic expansion because these features are *exponentially small* in $1/\varepsilon$ where $\varepsilon \ll 1$ is the perturbation parameter. The emerging field of 'exponential asymptotics' is not a branch of pure mathematics in pursuit of beauty (though some of the ideas *are* aesthetically charming) but a matter of bloody and unyielding engineering necessity.

In Section 3, we review some concepts that are already scattered in the textbooks: Poincaré's definition of asymptoticity, optimal truncation and minimum error, Carrier's Rule, and four heuristics for predicting divergence: the Exponential Reciprocal Rule, Van Dyke's Principle of Multiple Scales, Dyson's Change-of-

Table I. Nonsoliton exponentially small phenomena

Phenomena	Field	References
Dendritic crystal growth	Condensed matter	Kessler, Koplik and Levine [163] Kruskal and Segur [171, 172] Byatt-Smith [86]
Viscous fingering (Saffman–Taylor problem)	Fluid dynamics	Shraiman [276] Combescot *et al.* [103] Hong and Langer [146] Tanveer [288, 289]
Diffusion and merger of fronts on an exponentially long time scale	Reaction-diffusion systems	Carr [92], Hale [137], Carr and Pego [93] Fusco and Hale [130] Laforgue and O'Malley [173 – 176]
Superoscillations in Fourier integrals, quantum billiards, Gaussian beams	Applied mathematics, quantum mechanics, electromagnetic waves	Berry [31, 32]
Rapidly-forced pendulum	Classical physics	Chang [94] Scheurle *et al.* [275]
Resonant sloshing in a tank	Fluid mechanics	Byatt-Smith and Davie [88, 89]
Laminar flow in a porous pipe	Fluid mechanics, Space plasmas	Berman [23], Robinson [272], Terrill [290, 291], Terrill and Thomas [292], Grundy and Allen [135]
Jeffrey–Hamel flow stagnation points	Fluid mechanics, Boundary layer	Bulakh [85]
Shocks in nozzle	Fluid mechanics	Adamson and Richey [2]
Slow viscous flow past circle, sphere	Fluid mechanics (log and power series)	Proudman and Pearson [264], Chester and Breach [98] Skinner [283] Kropinski, Ward and Keller [170]
Log-and-power series	Fluids, electrostatic	Ward, Henshaw and Keller [308]

Table I. Nonsoliton exponentially small phenomena (continued)

Phenomena	Field	References
Log-and-power series	Elliptic PDE on domains with small holes	Lange and Weinitschke [179]
Equatorial Kelvin wave instability	Meteorology, oceanography	Boyd and Christidis [74, 75] Boyd and Natarov [76]
Error: Midpoint rule	Numerical analysis	Hildebrand [143]
Radiation leakage from a fiber optics waveguide	Nonlinear optics	Kath and Kriegsmann [162], Paris and Wood [258] Liu and Wood [183]
Particle channeling in crystals	Condensed matter physics	Dumas [119, 120]
Island-trapped water waves	Oceanography	Lozano and Meyer [185], Meyer [210]
Chaos onset: Hamiltonian systems	Physics	Holmes, Marsden and Scheurle [145]
Separation of separatrices	Dynamical systems	Hakim and Mallick [136]
Slow manifold in geophysical fluids	Meteorology	Lorenz and Krishnamurthy [184],
	Oceanography	Boyd [65, 66]
Nonlinear oscillators	Physics	Hu [149]
ODE resonances	Various	Ackerberg and O'Malley [1] Grasman and Matkowsky [133] MacGillivray [191]
French ducks ('canards')	Various	MacGillivray, Liu and Kazarinoff [192]

Sign Argument, and the Principle of Nonuniform Smallness. In later sections, we illustrate hyperasymptotic perturbation theory, which allows us to partially overcome the evils of divergence, through three examples: the Stieltjes function (Sections 4 and 5), a linear inhomogeneous differentiation equation (Section 6), and a weakly nonlocal solitary wave (Section 7).

Lastly, in Section 8 we present an overview of hyperasymptotic methods in general. We use the Pokrovskii–Khalatnikov–Kruskal–Segur (PKKS) method for

Table II. Selected examples of exponentially small quantum phenomena

Phenomena	References
Energy of a quantum double well (H_2^+, etc.)	Fröman [128] Čížek *et al.* [100] Harrell [140 – 142]
Imaginary part of eigenvalue of a metastable quantum species: Stark effect (external electric field)	Oppenheimer [255], Reinhardt [269], Hinton and Shaw [144], Benassi *et al.* [18]
Im(E): Cubic anharmonicity Im(E): Quadratic Zeeman effect (external magnetic field)	Alvarez [6] Čížek and Vrscay [101]
Transition probability, two-state quantum system (exponentially small in speed of variations)	Berry and Lim [42]
Width of stability bands for Hill's equation	Weinstein and Keller [313, 314]
Above-the-barrier scattering	Pokrovskii and Khalatnikov [262] Hu and Kruskal [150 – 152]
Anosov-perturbed cat map: semiclassical asymptotics	Boasman and Keating [46]

Table III. Weakly nonlocal solitary waves

Species	Field	References
Capillary-gravity water waves	Oceanography, marine engineering	Pomeau *et al.* [263] Hunter and Scheurle [153] Boyd [62] Benilov, Grimshaw and Kuznetsova [22] Grimshaw and Joshi [134] Dias *et al.* [114]
ϕ^4 Breather	Particle physics	Segur and Kruskal [278] Boyd [58]

Table III. Weakly nonlocal solitary waves (continued)

Species	Field	References
Fluxons, DNA helix modons in magnetic shear	Physics plasma physics	Malomed [195] Meiss and Horton [201]
Klein–Gordon envelope solitons	Electrical engineering	Boyd [67] Kivshar and Malomed [167]
Various	Review article	Kivshar and Malomed [168]
Higher latitudinal mode Rossby waves	Oceanography	Boyd [56, 57]
Higher vertical mode internal gravity waves	Oceanography, marine engineering	Akylas and Grimshaw [4]
Perturbed sine–Gordon	Physics	Malomed [194]
Nonlinear Schrödinger eq., cubic dispersion	Nonlinear optics	Wai, Chen and Lee [307]
Self-induced transparency eqs.: envelope solitons	Nonlinear optics	Branis, Martin and Birman [84] Martin and Branis [197]
Internal waves: stratified layer between 2 constant density layers	Oceanography, marine engineering	Vanden-Broeck and Turner [299]
Lee waves	Oceanography	Yang and Akylas [325]
Pseudospectra of matrices	Applied math., fluid mechanics	Reddy, Schmid and Henningson [267] Reichel and Trefethen [268]

'above-the-barrier' quantum scattering (Section 14) and 'resurgence' for the analysis of Stokes' phenomenon (Section 12) to give the flavor of these new ideas. (We warn the reader: 'beyond all orders' perturbation theory has become sufficiently developed that it is impossible, short of a book, to even summarize all the useful strategies.) The final section is a summary with pointers to further reading.

2. The Necessity of Computing Exponentially Small Terms

> Even the best toolmaker cannot wring five-figure accuracy out of the machining tolerances... This is how I come to find nearly all computations to more than three significant figures embarrassing. It's not a criticism of computer science because there is a direct analogy in asymptotic expansions. I find them plain embarrassing as a failure of realistic judgment.
>
> I was led to contemplate a heretical question: are higher approximations than the first justifiable? My experience indicates yes, but rarely. All differential equations are imperfect models and I would be embarrassed to publish a second approximation without convincing justification that the quality of the model validates it.
>
> Solutions as an end in themselves are pure mathematics; do we really need to know them to eight significant decimals?
>
> Richard E. Meyer (1992) [218]

Meyers' tart comment illuminates a fundamental limitation of hyperasymptotic perturbation theory: for many engineering and physics applications, a single term of an asymptotic series is sufficient. When more than one is needed, this usually means that the small parameter ε is not really small. Hyperasymptotic methods depend, as much as conventional perturbation theory, on the true and genuine smallness of ε and so cannot help. Numerical algorithms are usually necessary when $\varepsilon \sim O(1)$, either numerical or analytic [63].

And so, the first question of any adventure in hyperasymptotics is a question that patriotric Americans were supposed to ask themselves during wartime gas-rationing: 'Is this trip necessary?' The point of this review is that there is an amazing variety of problems where the trip *is* necessary.

Table I is a collection of miscellaneous problems from a variety of fields, especially fluid mechanics, where exponential smallness is crucial. Tables II and III are restricted selections limited to two areas where 'beyond all orders' calculations have been especially common: quantum mechanics and the weakly nonlocal solitary waves. The common thread is that for all these problems, some aspect of the physics is *exponentially small* in $1/\varepsilon$ where ε is the perturbation parameter. Since $\exp(-q/\varepsilon)$ where q is a constant cannot be approximated as a power series in ε – all its derivatives are zero at $\varepsilon = 0$ – such exponentially small effects are invisible to an ε power series. Such 'beyond all orders' features are like mathematical stealth aircraft, flying unseen by the radar of conventional asymptotics.

There are several reasons why such apparently tiny and insignificant features are important. In quantum chemistry and physics, for example, perturbations such as an external electric field may destabilize molecules. Mathematically, the eigenvalue E of the Schrödinger equation acquires an imaginary part which is typically exponentially small in $1/\varepsilon$. Nevertheless, this tiny $\Im(E)$ is important because it completely controls the lifetime of the molecule. J. R. Oppenheimer [255] showed that in the presence of an external electric field of strength ε, hydrogen atoms disassociated on a timescale which is inversely proportional to $\Im(E) = (4/3\varepsilon)\exp(-2/(3\varepsilon))$

and that electrons can be similarly sprung from metals. (This observation was the basis for the development of the scanning tunneling microscope by Binnig and Rohrer half a century later.) Only a few months after Oppenheimer's 1928 article, G. Gamow and Condon and Gurney showed that this 'tunnelling' explained the radioactive decay of unstable nuclei and particles, again on a timescale exponentially small in the reciprocal of the perturbation parameter.

Similarly, weakly nonlocal solitary waves do not decay to zero as $|x| \rightarrow \infty$ but to small, quasi-sinusoidal oscillations that fill all of space. For the species listed in Table III, the amplitude of the 'radiation coefficient' α is proportional to $\exp(-q/\varepsilon)$ for some q. When the appropriate wave equations are given a spatially localized initial condition, the resulting coherent structure slowly decays by radiation on a timescale inversely proportional to α.

For other problems, exponential smallness may hold the key to the very existence of solutions. For example, the melt interface between a solid and liquid is unstable, breaking up into dendritic fingers. Ivantsev (1947) develped a theory that successfully explained the parabolic shape of the fingers. However, experiments showed that the fingers also had a definite width. Attempts to predict this width by a power series in the surface tension ε failed miserably, even when carried to high order. Eventually, it was realized that the instability is controlled by factors that lay beyond all orders in ε. Kruskal and Segur [171, 172] showed that the complex-plane matched asymptotics method of Pokrovskii and Khalatnikov [262] could be applied to a simple model of crystal growth. In so doing, they not only resolved a forty-year old conundrum, but also furnished one of the (multiple) triggers for the resurgence in exponential asymptotics.

Even earlier, the flow of laminar fluid through a pipe or channel with porous walls had been shown to depend on exponential smallness. This nonlinear flow is not unique; rather there are *two* solutions which differ only through terms which are exponentially small in the Reynolds number R, which is the reciprocal of the perturbation parameter ε. As early as 1969, Terrill [292, 291] had diagnosed the illness and analytically determined the exponentially-small, mode-splitting terms [272, 135]

Similarly, the interactions between the electrostatic fields of atoms cause splitting of molecular spectra. The prototype is the quantum mechanical 'double well', such as the H_2^+ ion. The eigenvalues of the Schrödinger equation come in pairs, each pair close to the energy of an orbital of the hydrogen atom. The difference between each pair is exponentially small in the internuclear separation.

Lastly, Stokes' phenomenon in asymptotic expansions, which requires one exponential times a power series in ε in regions of the complex ε-plane, but *two* exponentials in other sectors, can only be smoothed and fully understood by looking at exponentially small terms.

In the physical sciences, smallness is relative. We can no more automatically assume an effect is negligible because it is proportional to $\exp(-q/\varepsilon)$ than a mother can regard her baby as insignificant because it is only sixty centimeters long.

3. Definitions and Heuristics

DEFINITION 1 (Asymptoticity). A power series is *asymptotic* to a function $f(\varepsilon)$ if, for *fixed* N and *sufficiently small* ε [19]

$$\left| f(\varepsilon) - \sum_{j=0}^{N} a_j\, \varepsilon^j \right| \sim \mathrm{O}\big(\varepsilon^{N+1}\big), \tag{1}$$

where O( ) is the usual 'Landau gauge' symbol that denotes that the quantity to the left of the asymptotic equality is bounded in absolute value by a constant times the function inside the parentheses on the right.

This formal definition, due to Poincaré, tells us what happens in the limit that ε tends to 0 for fixed N. Unfortunately, the more interesting limit is ε fixed, $N \to \infty$. A series may be asymptotic, and yet diverge in the sense that for sufficiently large j, the terms increase with increasing j.

However, convergence may be over-rated as expressed by the following amusing heuristic.

PROPOSITION 1 (Carrier's Rule). *Divergent series converge faster than convergent series because they don't have to converge.*

What George F. Carrier meant by this bit of apparent jabberwocky is that the leading term in a divergent series is often a very good approximation even when the 'small' parameter ε is not particularly small. This is illustrated through many numerical comparisons in [19]. In contrast, it is quite unusual for an ordinary convergent power series to be accurate when $\varepsilon \sim \mathrm{O}(1)$.

The vice of divergence is that for fixed ε, the error in a divergent series will reach, as more terms are added, an ε-dependent minimum. The error then increases without bound as the number of terms tends to infinity. The standard empirical strategy for achieving this minimum error is the following.

DEFINITION 2 (Optimal Truncation Rule). For a given ε, the minimum error in an asymptotic series is *usually* achieved by truncating the series so as to retain the *smallest* term in the series, discarding all terms of higher degree.

The imprecise adjective 'usually' indicates that this rule is empirical, not something that has been rigorously proved to apply to all asymptotic series. (Indeed, it is easy to contrive counter-examples.) Nevertheless, the Optimal Truncation Rule is very useful in practice. It can be rigorously justified for some classes of asymptotic series [158, 241, 169, 106, 107, 285].

To replace the lengthy, jaw-breaking phrase 'optimally-truncated asymptotic series', Berry and Howls coined a neologism [35, 30] which is rapidly gaining popularity: 'superasymptotic'. A more compelling reason for new jargon is that the standard definition of asymptoticity (Definition 1 above) is a statement about *powers* of ε, but the error in an optimally-truncated divergent series is usually an *exponential* function of the reciprocal of ε.

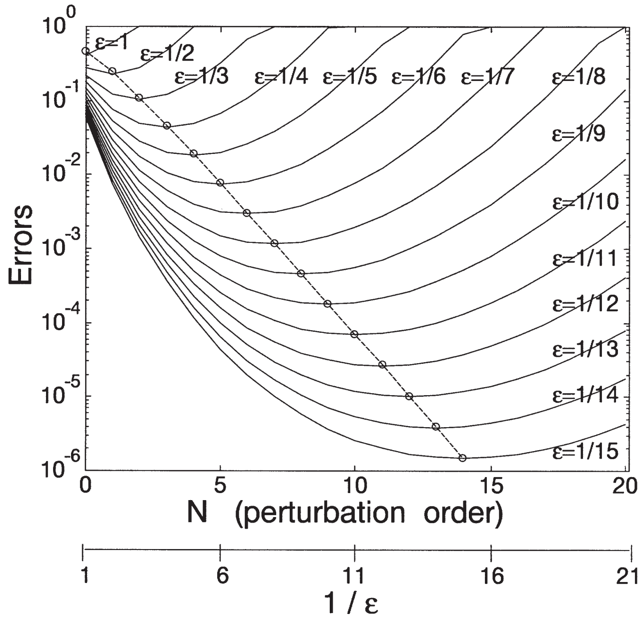

Figure 1. Solid curves: absolute error in the approximation of the Stieltjes function up to and including the Nth term. Dashed-and-circles: theoretical error in the optimally-truncated or 'superasymptotic' approximation: $E_{N_{\text{optimum}}}(\varepsilon) \approx (\pi/(2\varepsilon))^{1/2} \exp(-1/\varepsilon)$ versus $1/\varepsilon$. The horizontal axis is perturbative order N for the actual errors and $1/\varepsilon$ for the theoretical error.

DEFINITION 3 (Superasymptotic). An *optimally-truncated* asymptotic series is a '*superasymptotic*' approximation. The error is *typically* $O(\exp(-q/\varepsilon))$ where $q > 0$ is a constant and ε is the small parameter of the asymptotic series. The degree N of the highest term retained in the optimal truncation is proportional to $1/\varepsilon$.

Figure 1 illustrates the errors in the asymptotic series for the Stieltjes function (defined in the next section) as a function of N for fifteen different values of ε. For each ε, the error dips to a minimum at $N \approx 1/\varepsilon$ as the perturbation order N increases. The minimum error for each N is the 'superasymptotic' error.

Also shown is the theoretical prediction that the minimum error for a given ε is $(\pi/(2\varepsilon))^{1/2} \exp(-1/\varepsilon)$ where $N_{\text{optimum}}(\varepsilon) \sim 1/\varepsilon - 1$. For this example, both the exponential factor and the proportionality constant will be derived in Section 5.

The definition of 'superasymptotic' makes a claim about the exponential dependence of the error which is easily falsified. Merely by redefining the perturbation parameter, we could, for example, make the minimum error be proportional to the exponential of $1/\varepsilon^{\gamma}$ where γ is arbitrary. Modulo such trivial rescalings, however, the superasymptotic error is indeed exponential in $1/\varepsilon$ for a wide range of divergent series [30, 72].

The emerging art of 'exponential asymptotics' or 'beyond-all-orders' perturbation theory has made it possible to improve upon optimal truncation of an asymptotic series, and calculate quantities 'below the radar screen', so to speak, of the superasymptotic approximation. It will not do to describe these algorithms as the calculation of exponentially small quantities since the superasymptotic approximation, too, has an accuracy which is O(exp$(-q/\varepsilon)$ for some constant q. Consequently, Berry and Howls coined another term to label schemes that are better than mere truncation of a power series in ε:

DEFINITION 4. A *hyperasymptotic* approximation is one that achieves higher accuracy than a superasymptotic approximation by adding one or more terms of a *second* asymptotic series, with different scaling assumptions, to the optimal truncation of the original asymptotic expansion [30]. (With another rescaling, this process can be iterated by adding terms of a third asymptotic series, and so on.)

All of the methods described below are 'hyperasymptotic' in this sense although in the process of understanding them, we shall acquire a deeper understanding of the mathematical crimes and genius that underlie asymptotic expansions and the superasymptotic approximation.

But when does a series diverge? Since all derivatives of $\exp(-1/\varepsilon)$ vanish at the origin, this function has only the trivial and useless power series expansion whose coefficients are *all zeros*:

$$\exp(-q/\varepsilon) \sim 0 + 0\varepsilon + 0\varepsilon^2 + \cdots \tag{2}$$

for any positive constant q. This observation implies the first of our four heuristics about the nonconvergence of an ε-power series.

PROPOSITION 2 (Exponential Reciprocal Rule). *If a function $f(\varepsilon)$ contains a term which is an* exponential *function of the* reciprocal *of ε, then a power series in ε will not converge to $f(\varepsilon)$.*

We must use phrase 'not converge to' rather than the stronger 'diverge' because of the possibility of a function like

$$h(\varepsilon) \equiv \sqrt{1+\varepsilon} + \exp(-1/\varepsilon). \tag{3}$$

The power series of $h(\varepsilon)$ will *converge* for all $|\varepsilon| < 1$, but it converges to a number *different* from the true value of $h(\varepsilon)$ for all ε except $\varepsilon = 0$.

Fortunately, this situation – a convergent series for a function that contains a term exponentially small in $1/\varepsilon$, and therefore *invisible* to the power series – seems to be rare in applications. (The author would be interested in learning of exceptions.)

Milton van Dyke, a fluid dynamicist, offered another useful heuristic in his slim book on perturbation methods [297]:

PROPOSITION 3 (Principle of Multiple Scales). *Divergence should be expected when the solution depends on two independent length scales.*

We shall illustrate this rule later.

The physicist Freeman Dyson [122] published a note which has been widely invoked in both quantum field theory and quantum mechanics for more than forty years [164 – 166, 43 – 45]. However, with appropriate changes of jargon, the argument applies outside the realm of the quantum, too. Terminological note: a 'bound state' is a spatially localized eigenfunction associated with a discrete, negative eigenvalue of the stationary Schrödinger equation and the 'coupling constant' is the perturbation parameter which multiplies the potential energy perturbation.

PROPOSITION 4 (Dyson Change-of-Sign Argument). *If there are no bound states for* negative *values of the coupling constant ε, then a perturbation series for the bound states will diverge even for $\varepsilon > 0$.*

A simple example is the one-dimensional anharmonic quantum oscillator, whose bound states are the eigenfunctions of the stationary Schrödinger equation:

$$\psi_{xx} + \{E - x^2 - \varepsilon x^4\}\psi = 0. \tag{4}$$

When $\varepsilon \geqslant 0$, Equation (4) has a countable infinity of bound states with positive eigenvalues E (the energy); each eigenfunction decays exponentially with increasing $|x|$. However, the quartic perturbation will grow faster with $|x|$ than the unperturbed potential energy term, which is quadratic in x. It follows that when ε is negative, the perturbation will reverse the sign of the potential energy at $x = \pm 1/(-\varepsilon)^{1/2}$. Because of this, the wave equation has no bound states for $\varepsilon < 0$, that is, no eigenfunctions which decay exponentially with $|x|$ for all sufficiently large $|x|$.

Consequently, the perturbation series cannot converge to a bound state for negative ε, be it ever so small in magnitude, because there is no bound state to converge to. If this nonconvergence is divergence (as opposed to convergence to an unphysical answer), then the divergence must occur for all nonzero positive ε, too, since the domain of convergence of a power series is always $|\varepsilon| < \rho$ for some positive ρ as reviewed in elementary calculus texts.

This argument is not completely rigorous because the perturbation series could in principle converge for negative ε to *something* other than a bound state. Nevertheless, the Change-of-Sign Argument has been reliable in quantum mechanics [164].

Implicit in the very notion of a 'small perturbation' is the idea that the term proportional to ε is indeed small compared to the rest of the equation. For the anharmonic oscillator, however, this assumption always breaks down for $|x| > 1/|\varepsilon|^{1/2}$. Similarly, in high Reynolds number fluid flows, the viscosity is a small perturbation everywhere except in thin layers next to boundaries, where it brings the velocity to zero ('no slip' boundary condition) at the wall. This and other examples suggests our fourth heuristic:

PROPOSITION 5 (Principle of Nonuniform Smallness). *Divergence should be expected when the perturbation is not small, even for arbitrarily small ε, in some regions of space.*

When the perturbation is not small *anywhere*, of course, it is impossible to apply perturbation theory. When the perturbation is small *uniformly* in space, the ε power series usually has a finite radius of convergence. Asymptotic-but-divergent is the usual spoor of a problem where the perturbation is small-but-not-everywhere.

We warn that these heuristics are just that, and not theorems. Counterexamples to some are known, and probably can be constructed for all. In practice, though, these empirical predictors of divergence are quite useful.

Pure mathematics is the art of the provable, but applied mathematics is the description of what happens. These heuristics illustrate the gulf between these realms. The domain of a theorem is bounded by extremes, even if unlikely. Heuristics are descriptions of what is probable, not the full range of what is possible.

For example, the simplex method of linear programming can converge very slowly because (it can be proven) the algorithm could visit every one of the millions and millions of vertices that bound the feasible region for a large problem. The reason that Dantzig's algorithm has been widely used for half a century is that in practice, the simplex method finds an acceptable solution after visiting only a tiny fraction of the vertices.

Similarly, Hotellier proved in 1944 that (in the worst case) the roundoff error in Gaussian elimination could be 4^N times machine epsilon where N is the size of the matrix, implying that a matrix of dimension larger than 50 is insoluble on a machine with sixteen decimal places of precision. What happens in practice is that the matrices generated by applications can usually be solved even when $N > 1000$ [294]. The exceptions arise mostly because the underlying problem is genuinely singular, and not because of the perversities of roundoff error.

In a similar spirit, we offer not theorems but experience.

4. Optimal Truncation and Superasymptotics for the Stieltjes Function

The first illustration is the Stieltjes function, which, with a change of variable, is the 'exponential integral' which is important in radiative transfer and other branches of science and engineering. This integral-depending-on-a-parameter is defined by

$$S(\varepsilon) = \int_0^\infty \frac{\exp(-t)}{1+\varepsilon t}\,\mathrm{d}t. \tag{5}$$

The geometric series identity, valid for arbitrary integer N,

$$\frac{1}{1+\varepsilon t} = \sum_{j=0}^{N} (-\varepsilon t)^j + \frac{(-\varepsilon t)^{N+1}}{1+\varepsilon t} \tag{6}$$

allows an exact alternative definition of the Stieltjes function, valid for any finite N:

$$S(\varepsilon) = \sum_{j=0}^{N} (-\varepsilon)^j \int_0^\infty \exp(-t) t^j \, \mathrm{d}t + E_N(\varepsilon), \tag{7}$$

where

$$E_N(\varepsilon) \equiv \int_0^\infty \frac{\exp(-t)(-\varepsilon t)^{N+1}}{1+\varepsilon t} \, \mathrm{d}t. \tag{8}$$

The integrals in (3) are special cases of the integral definition of the Γ-function and so can be performed explicitly to give

$$S(\varepsilon) = \sum_{j=0}^{N} (-1)^j j! \varepsilon^j + E_N(\varepsilon). \tag{9}$$

Equations (5)–(9) are *exact*. If the integral $E_N(\varepsilon)$ is neglected, then the summation is the first $(N+1)$ terms of an asymptotic series. Both Van Dyke's principle and Dyson's argument forecast that this series is divergent.

The exponential $\exp(-t)$ varies on a length scale of $\mathrm{O}(1)$ where $\mathrm{O}(\,)$ is the usual 'Landau gauge' or 'order-of-magnitude' symbol. In contrast, the denominator depends on t only as εt, that is, varies on a 'slow' length scale which is $\mathrm{O}(1/\varepsilon)$. Dependence on two independent scales, i.e., t and (εt), is van Dyke's 'Mark of Divergence'.

When ε is negative, the integrand of the Stieltjes function is *singular* on the integration interval because of the simple pole at $t = -1/\varepsilon$. This strongly (and correctly) suggests that $S(\varepsilon)$ is not analytic at $\varepsilon = 0$ as analyzed in detail in [19]. Just as for Dyson's quantum problems, the radius of convergence of the ε power series must be zero.

A deeper reason for the divergence of the ε-series is that Taylor-expanding $1/(1+\varepsilon t)$ in the integrand of the Stieltjes function is an act of inspired stupidity. The inspiration is that an integral which cannot be evaluated in simple closed form is converted to a power series with explicit, analytic coefficients. The stupidity is that the domain of convergence of the geometric series is

$$|t| < 1/\varepsilon \tag{10}$$

because of the simple pole of $1/(1+\varepsilon t)$ at $t = -1/\varepsilon$. Unfortunately, the domain of integration is semi-infinite. It follows that the Taylor expansion is used *beyond* its interval of validity. The price for this mathematical crime is divergence.

The reason that the asymptotic series is useful anyway is because the integrand is *exponentially small* in the region where the expansion of $1/(1+\varepsilon t)$ is divergent. Split the integral into two parts, one on the interval where the denominator expansion is convergent, the other where it is not, as

$$S(\varepsilon) = S_{\mathrm{con}}(\varepsilon) + S_{\mathrm{div}}(\varepsilon), \tag{11}$$

where

$$S_{\text{con}}(\varepsilon) \equiv \int_0^{1/\varepsilon} \frac{\exp(-t)}{1+\varepsilon t}\,\mathrm{d}t, \qquad S_{\text{div}}(\varepsilon) \equiv \int_{1/\varepsilon}^{\infty} \frac{\exp(-t)}{1+\varepsilon t}\,\mathrm{d}t. \tag{12}$$

Since $\exp(-t)/(1+\varepsilon t)$ is bounded from above by $\exp(-t)/2$ for all $t \geqslant 1/\varepsilon$, it follows that

$$S_{\text{div}}(\varepsilon) \leqslant \frac{\exp(-1/\varepsilon)}{2}. \tag{13}$$

Thus, one can approximate the Stieltjes function as

$$S(\varepsilon) \approx S_{\text{con}}(\varepsilon) + \mathrm{O}(\exp(-1/\varepsilon)). \tag{14}$$

The magnitude of that part of the Stieltjes function which is inaccesible to a convergent expansion of $1/(1+\varepsilon t)$ is proportional to $\exp(-1/\varepsilon)$. This suggests that the best one can hope to wring from the asymptotic series is an error no smaller than the order-of-magnitude of $S_{\text{div}}(\varepsilon)$, that is, $\mathrm{O}(\exp(-1/\varepsilon))$.

5. Hyperasymptotics for the Stieltjes Function

It is possible to break the superasymptotic constraint to obtain a more accurate ‘hyperasymptotic’ approximation by inspecting the error integrals $E_N(\varepsilon)$, which are illustrated in Figure 2 for a particular value of ε. The crucial point is that the *maximum* of the *integrand* shifts to *larger* and *larger* t as N increases. When $N \leqslant 2$, the peak (for $\varepsilon = 1/3$) is still within the convergence disk of the geometric series. For larger N, however, the maximum of the integrand occurs for $T > 1$, that is, for $t > 1/\varepsilon$. (Ignoring the slowly varying denominator $1/(1+\varepsilon t)$, one can show by differentiating $\exp(-t)t^{N+1}$ that the maximum occurs at $t = 1/(N+1)$.) When $(N+1) \geqslant 1/\varepsilon$, the geometric series diverges in the very region where the integrand of E_N has most of its amplitude. Continuing the asymptotic expansion to larger N will merely accumulate further error.

The key to a hyperasymptotic approximation is to use the information that the error integral is peaked at $t = 1/\varepsilon$. Just as asymptotic series can be derived by several different methods, similarly ‘hyperasymptotics’ is not a single algorithm, but rather a family of siblings. Their common theme is to append a *second* asymptotic series, based on different scaling assumptions, to the ‘superasymptotic’ approximation.

One strategy is to expand the denominator of the error integral $E_{N_{\text{optimum}}}(\varepsilon)$ in powers of $(t - 1/\varepsilon)$ instead t. In other words, expand the integrand about the point where it is peaked (when $N = N_{\text{optimum}}(\varepsilon) \approx 1/\varepsilon - 1$). The key identity is

$$\begin{aligned} \frac{1}{1+\varepsilon t} &= \frac{1}{2\{1+\frac{1}{2}(\varepsilon t - 1)\}} \\ &= \frac{1}{2}\sum_{k=0}^{M}\left(-\frac{1}{2}\right)^k (\varepsilon t - 1)^k. \end{aligned} \tag{15}$$

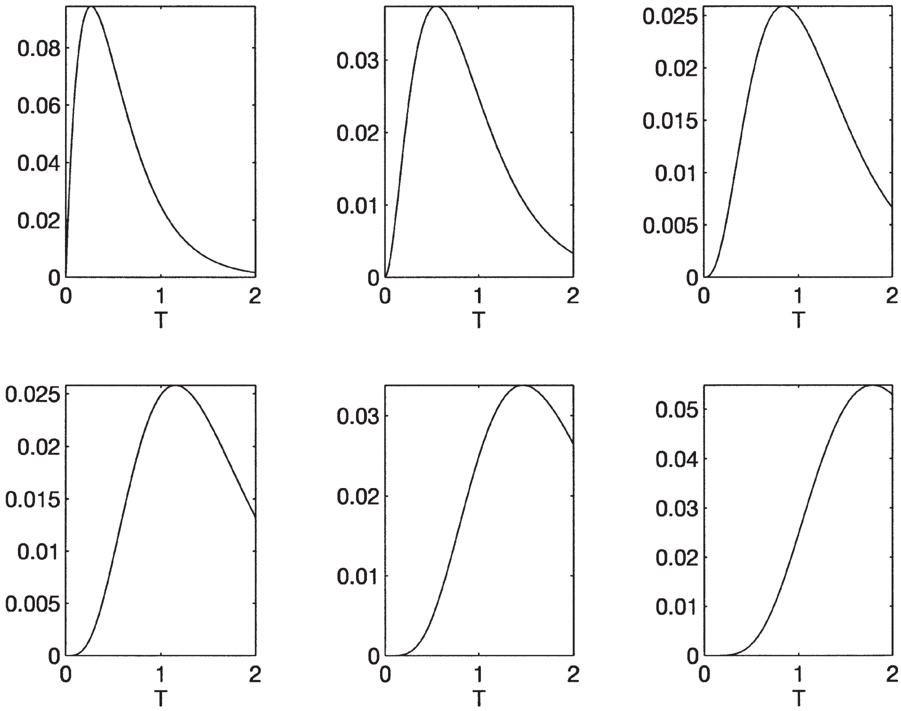

Figure 2. The integrands of the first six error integrals for the Stieltjes function, $E_0, E_1, \ldots, E_5$ for $\varepsilon = 1/3$, plotted as functions of the 'slow' variable $T \equiv \varepsilon t$.

$$S(\varepsilon) = \sum_{j=0}^{N} (-1)^j j! \varepsilon^j + \\ + \frac{1}{2} \sum_{k=0}^{M} \int_0^\infty \exp(-t)(-\varepsilon t)^{N+1} \left(\frac{1 - \varepsilon t}{2} \right)^k \mathrm{d}t + H_{NM}(\varepsilon), \tag{16}$$

where the hyperasymptotic error integral is

$$H_{NM}(\varepsilon) \equiv \frac{1}{2} \int_0^\infty \frac{\exp(-t)}{1 + \varepsilon t} (-\varepsilon t)^{N+1} \left(-\frac{1}{2} \right)^{M+1} (\varepsilon t - 1)^{M+1} \, \mathrm{d}t. \tag{17}$$

A crucial point is that the integrand of each term in the hyperasymptotic summation is $\exp(-t)$ multiplied by a polynomial in t. This means that the (NM)th hyperasympotic expansion is just a *weighted sum* of the first $(N + M + 1)$ terms of the original divergent series. The change of variable made by switching from (εt) to $(\varepsilon t - 1)$ is equivalent to the 'Euler sum-acceleration' method, an ancient and well-understood method for improving the convergence of slowly convergent or divergent series.

Let

$$a_j \equiv (-\varepsilon)^j j!, \tag{18}$$

$$S_N^{\text{Superasymptotic}} \equiv \sum_{0}^{[1/\varepsilon-1]} a_j, \tag{19}$$

where $[m]$ denotes the integer nearest m for any quantity m and where the upper limit on the sum is

$$N_{\text{optimum}}(\varepsilon) = 1/\varepsilon - 1. \tag{20}$$

Then the Euler acceleration theory [318, 70] shows

$$\begin{aligned} S_0^{\text{Hyperasymptotic}} &\equiv S_N^{\text{Superasymptotic}} + \frac{1}{2}a_{N+1}, \\ S_1^{\text{Hyperasymptotic}} &\equiv S_N^{\text{Superasymptotic}} + \frac{3}{4}a_{N+1} + \frac{1}{4}a_{N+2}, \\ S_2^{\text{Hyperasymptotic}} &\equiv S_N^{\text{Superasymptotic}} + \frac{7}{8}a_{N+1} + \frac{1}{2}a_{N+2} + \frac{1}{8}a_{N+3}. \end{aligned} \tag{21}$$

The lowest order hyperasymptotic approximation estimates the error in the superasymptotic approximation as roughly one-half a_{N+1} or explicitly

$$\begin{aligned} E_N &\sim (1/2)(-1)^{N+1}(N+1)!\varepsilon^{N+1} \quad [\varepsilon \approx 1/(N+1)] \\ &\sim \sqrt{\frac{\pi}{2\varepsilon}}\exp\left(-\frac{1}{\varepsilon}\right) \quad [\varepsilon = 1/(N+1)]. \end{aligned} \tag{22}$$

This confirms the claim, made earlier, that the superasymptotic error is an exponential function of $1/\varepsilon$.

Figure 3 illustrates the improvement possible by using the Euler transform. A minimum error still exists; Euler acceleration does not eliminate the divergence. However, the minimum error is roughly squared, that is, twice as many digits of accuracy can be achieved for a given ε [273, 274, 249, 77].

However, a hyperasymptotic series can also be generated by a completely different rationale. Figure 4 shows how the integrand of the error integral E_N changes with ε when $N = N_{\text{optimum}}(\varepsilon)$: the integrand becomes *narrower* and *narrower*. This narrowness can be exploited by Taylor-expanding the denominator of the integrand in powers of $1 - \varepsilon t$, which is equivalent to the Euler acceleration of the regular asymptotic series as already noted. However, the narrowness of the integrand also implies that one may make approximations in the *numerator*, too.

Qualitatively, the numerator resembles a Gaussian centered on $t = 1/\varepsilon$. The heart of the 'steepest descent' method for evaluating integrals is to (i) rewrite the rapidly varying part of the integral as an exponential (ii) make a change of variable so that this exponential is equal to the Gaussian function $\exp(-z^2/\varepsilon)$ and expand $\mathrm{d}t/\mathrm{d}z$, multiplied by the slowly varying part of the integral (here $1/(1 + \varepsilon t(z)$, in powers of z. Since this method is described in Section 11 below, the details will be omitted here. The lowest order is identical with the lowest order Euler approximation.

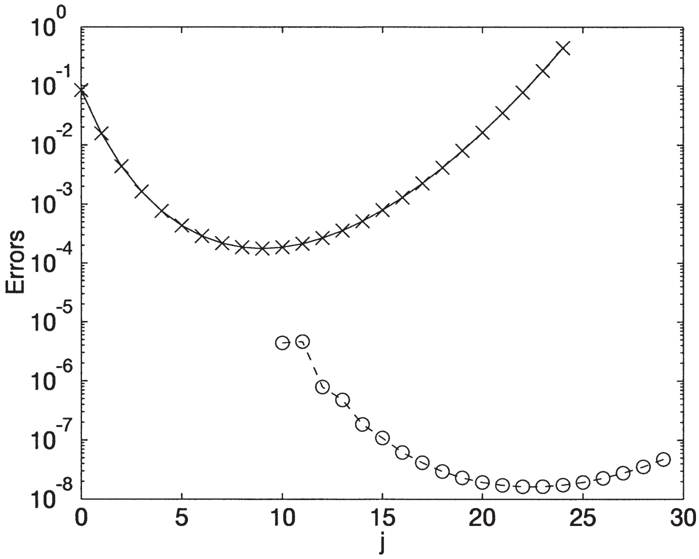

Figure 3. Stieltjes function with $\varepsilon = 1/10$. Solid-with-x's: Absolute value of the absolute error in the partial sum of the asymptotic series, up to and including a_j where j is the abscissa. Dashed-with-circles: The result of Euler acceleration. The terms up to and including the optimum order, here $N_{\text{opt}}(\varepsilon) = 9$, are unweighted. Terms of degree $j > N_{\text{opt}}$ are multiplied by the appropriate Euler weight factors as described in the text. The circle above $j = 15$ is thus the sum of nine unweighted and six Euler-weighted terms.

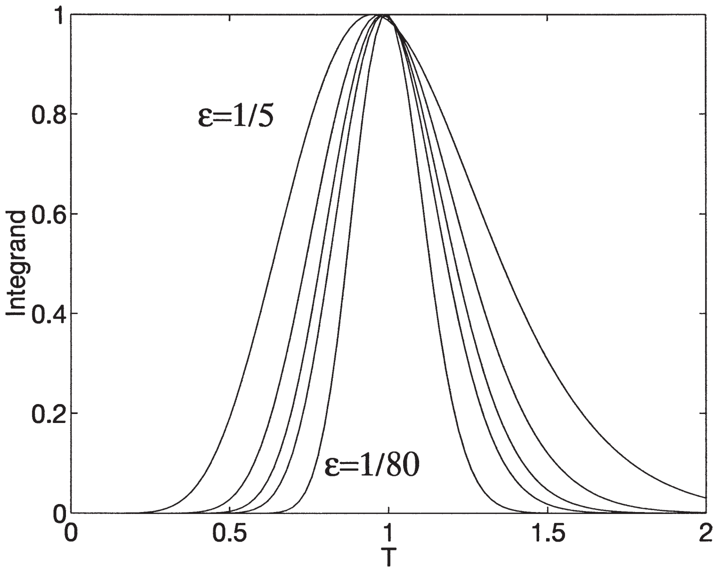

Figure 4. Integrand of the integral $E_{N_{\text{optimum}}}(\varepsilon)$, which is the error in the regular asymptotic series truncated at the Nth term, as a function of $T \equiv \varepsilon t$ for $\varepsilon = 1/5, 1/10, 1/20, 1/40, 1/80$ in order of increasing narrowness.

W. G. C. Boyd (no relation) has developed systematic methods for integrals that are Stieltjes functions, a class that includes the Stieltjes function as a special case [77 – 80]. The simpler treatment described here is based on Olver's monograph [249] and forty-year old articles by Rosser [273, 274].

6. A Linear Differential Equation

Our second example is the linear problem

$$\varepsilon^2 u_{xx} - u = -f(x) \tag{23}$$

on the infinite interval $x \in [-\infty, \infty]$ subject to the conditions that both $|u(x)|$, $|f(x)| \to 0$ as $|x| \to \infty$ where the subscripts denote second differentiation with respect to x, $f(x)$ is a known forcing function, and $u(x)$ is the unknown. This problem is a prototype for boundary layers in the sense that the term multiplying the highest derivative formally vanishes in the limit $\varepsilon \to 0$, but it has been simplified further by omitting boundaries. The divergence, however, is *not* eliminated when the boundaries are.

At first, this linear boundary value problem seems very different from the Stieltjes integral. However, Equation (23) is solved without approximation by the Fourier integral

$$u(x) = \int_{-\infty}^{\infty} \frac{F(k)}{1+\varepsilon^2 k^2} \exp(ikx)\,\mathrm{d}k, \tag{24}$$

where $F(k)$ is the Fourier transform of the forcing function:

$$F(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \exp(-ikx)\,\mathrm{d}x. \tag{25}$$

The Fourier integral (24) is very similar in form to the Stieltjes function. To be sure, the range of integration is now infinite rather than semi-infinite and the exponential has a complex argument. The similarity is crucial, however: for both the Stieljes integral and the Fourier integral, expanding the denominator of the integrand in powers of ε generates an asymptotic series. In both cases, the series is divergent because the expansion of the denominator has only a finite radius of convergence whereas the range of integration is unbounded.

The asymptotic solution to (23) may be derived by either of two routes. One is to expand $1/(1+\varepsilon^2 k^2)$ as a series in ε and then recall that the product of $F(k)$ with $(-k^2)$ is the transform of the second derivative of $f(x)$ for any $f(x)$. The second route is to use the method of multiple scales. If we assume that the solution $u(x)$ varies only on the same 'slow' O(1) length scale as $f(x)$, and not on the 'fast' $O(1/\varepsilon)$ scale of the homogeneous solutions of the differential equation, then the second derivative may be neglected to lowest order to give the solution $u(x) \approx f(x)$. This is called the 'outer' solution in the language of matched asymptotic expansions. Expanding $u(x)$ as a series of even powers of ε and continuing this reasoning to higher order gives

$$u(x) \sim \sum_{j=0}^{\infty} \varepsilon^2 \frac{\mathrm{d}^{2j} f}{\mathrm{d}x^{2j}}. \tag{26}$$

This differential equation seems to have little connection to our previous example, but this is a mirage. For the special case

$$f(x) = \frac{4}{1+x^2} \tag{27}$$

the Fourier transform $F(k) = 2\exp(-|k|)$. Using the partial fraction expansion $1/(1+\varepsilon^2 k^2) = (1/2)\{1/(1-i\varepsilon k) + 1/(1+i\varepsilon k)\}$, one can show that the solution to (23) is

$$\begin{aligned} u(x;\varepsilon) \;=\; & \frac{1}{1+ix}\left\{S\left(-\frac{i\varepsilon}{1+ix}\right) + S\left(\frac{i\varepsilon}{1+ix}\right)\right\} + \\ & + \frac{1}{1-ix}\left\{S\left(-\frac{i\varepsilon}{1-ix}\right) + S\left(\frac{i\varepsilon}{1-ix}\right)\right\}, \end{aligned} \tag{28}$$

where $S(\varepsilon)$ is the Stieltjes function. At $x = 0$, the solution simplifies to $u(0) = 2\{S(i\varepsilon) + S(-i\varepsilon)\}$. The odd powers of ε cancel, but the even powers reinforce to give

$$u(0) \sim 4\sum_{j=0}^{\infty}(2j)!(-1)^j \varepsilon^{2j}. \tag{29}$$

There is nothing special about the Lorentzian function (or $x = 0$), however. As explained at greater length in [61] and [69], the exponential decay of a Fourier transform with wavenumber k is generic if $f(x)$ is free of singularities for real x. The factorial growth of the power series coefficients with j, explicit in (29), is typical of the general multiple scale series (26) for all x for most forcing functions $f(x)$.

To obtain the optimal truncation, apply the identity $1/(1+z) = \sum_{j=0}^{N}(-z)^j + (-z)^{N+1}/(1+z)$ for all z and any positive integer N to the integral (24) with $z = \varepsilon^2 k^2$ to obtain, without approximation,

$$u = \sum_{j=0}^{N} \varepsilon^2 \frac{\mathrm{d}^{2j} f}{\mathrm{d}x^{2j}} + (-1)^{N+1}\varepsilon^{2(N+1)} \int_{-\infty}^{\infty} \frac{k^{2(N+1)} F(k)}{1+\varepsilon^2 k^2} \exp(ikx)\,\mathrm{d}k. \tag{30}$$

The Nth order asymptotic approximation is to neglect the integral. For large N, the error integral in Equation (30) can be approximatedly evaluated by steepest descent (Section 11 below). The optimal truncation is obtained by choosing N so as to minimize this error integral for a given ε. It is not possible to proceed further without specific information about the transform $F(k)$. If, however, one knows that

$$F(k) \sim \mathcal{A} \exp(-\mu|k|) \quad \text{as } |k| \to \infty \tag{31}$$

for some positive constant μ where $\mathcal{A}$ denotes factors that vary *algebraically* rather than exponentially with wavenumber, then independent of $\mathcal{A}$ (to lowest order), the optimal truncation as estimated by steepest descent is

$$N_{\text{opt}}(\varepsilon) \sim \frac{\mu}{2\varepsilon} - 1, \quad \varepsilon \ll 1, \tag{32}$$

and the error in the 'superasymptotic' approximation is

$$\left| u(x;\varepsilon) - \sum_{j=0}^{N_{\text{opt}}} \varepsilon^2 \frac{\mathrm{d}^{2j} f}{\mathrm{d}x^{2j}} \right| \leqslant \mathcal{A}' \exp\left(-\frac{\mu}{\varepsilon} \right), \quad \varepsilon \ll 1, \tag{33}$$

where $\mathcal{A}'$ denotes factors that vary algebraically with ε, i.e., slowly compared to the exponential, in the limit of small ε.

In textbooks on perturbation theory, the differential equation (23) is most commonly used to illustrate the method of matched asymptotic expansions. The multiple scales series (26) is the interior or 'outer' solution. To satisfy the boundary conditions

$$u(-1) = u(1) = 0 \tag{34}$$

it is necessary to add 'inner' solutions which are functions of the 'fast' variable $X = x/\varepsilon$. For (23), the exact solution is

$$u(x;\varepsilon) = u_p(x;\varepsilon) + a\exp(-[x+1]/\varepsilon) + b\exp([x-1]/\varepsilon), \tag{35}$$

where $u_p(x;\varepsilon)$, the particular solution, is the solution to the same problem on the infinite interval, already described above, and

$$\begin{aligned} a &= \frac{-u_p(-1;\varepsilon) + e^{-2/\varepsilon} u_p(1;\varepsilon)}{1 - \exp(-4/\varepsilon)}, \\ b &= \frac{-u_p(1;\varepsilon) + e^{-2/\varepsilon} u_p(-1;\varepsilon)}{1 - \exp(-4/\varepsilon)}. \end{aligned} \tag{36}$$

The 'inner' expansion is just the perturbative approximation to the exponentials in (35). The matched asymptotics solution is completed by matching the inner and outer expansions together, term-by-term.

It is important to note that for the finite domain $x \in [-1, 1]$, it is perfectly reasonable to choose a function like $g(x) = x^4/(1+x^2)$, which is unbounded as $|x| \to \infty$ and therefore lacks a well-behaved Fourier transform. However, the hyperasymptotic method can be extended to such cases by defining the function f in the Fourier integral to be

$$f(x) \equiv g(x)\frac{1}{2}\left\{\text{erf}(\lambda[x-2]) - \text{erf}(\lambda[x+2])\right\}. \tag{37}$$

If the constant λ is large, the multiplier of g differs from 1 by an exponentially small amount on the interval $x \in [-1, 1]$ so that $f \approx g$ on the finite domain. The modified function f, unlike g, decays exponentially with $|x|$ as $|x| \to \infty$ so that it has a well-behaved Fourier transform. We can therefore proceed exactly as before with f used to generate the 'outer' approximation in the form of a Fourier transform. For example, for the particular case $g = x^4/(1+x^2)$, the poles at $x = \pm i$ imply that $F(k)$ decays as $\exp(-|k|)$ so that the optimal truncation and error bound are the same as for the Lorentzian forcing, $f = 4/(1+x^2)$.

Since asymptotic matching is needed only because of the boundaries (and boundary layers), it is natural to assume that the inner expansion is the villain, responsible for the divergence of the matched asymptotic expansions. This is only half-true. In the perturbative scheme,

$$a \sim -u_p(-1; \varepsilon); \qquad b \sim -u_p(1; \varepsilon) \tag{38}$$

to all orders in ε with an error which is $O(\exp(-2/\varepsilon))$. The boundary layers have indeed enforced a minimum error below which the ordinary perturbative scheme cannot go, but it depends on the separation between the boundaries. Here, the boundary-layer-induced error is only the *square* of the minimum error in the power series for $u_p(x; \varepsilon)$ when $f(x) = 4/(1+x^2)$.

The outer solution is a greater villain. Even without boundaries, the multiple scales series is divergent.

7. Weakly Nonlocal Solitary Waves

> In general, the divergence of series in perturbation theory (while a good approximation is given by a few initial terms) is usually related to the fact that we are looking for an object which does not exist. If we try to fit a phenomenon to a scheme which actually contradicts the essential features of the phenomenon, then it is not surprising that our series diverge.
>
> V. I. Arnold (1937–) [7, p. 395]

Solitary waves, which are spatially localized nonlinear disturbances that propagate without change in shape or form, have been important in a wide range of science and engineering disciplines. Such diverse phenomena as the Great Red Spot of Jupiter, Gulf Stream rings in the ocean, neural impulses, vibrations in polymer lattices, and perhaps even the elementary particles of physics have been

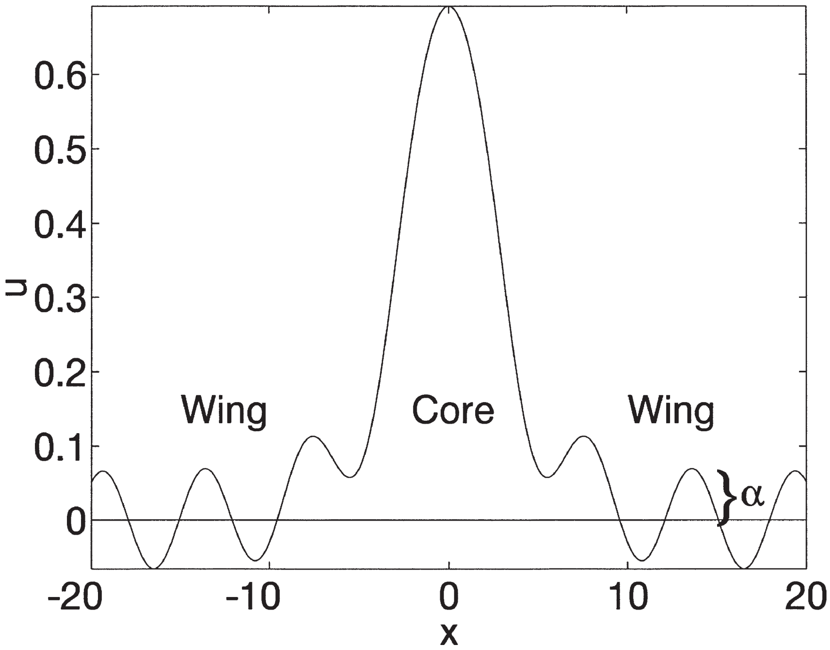

Figure 5. Schematic of a weakly nonlocal solitary wave or a forced wave of similar shape. The amplitude of the 'wings' is the 'radiation coefficient' α, which is exponentially small in $1/\varepsilon$ compared to the amplitude of the'core'.

identified, at least tentatively, as solitary waves; in ten years, most of our phone and data communications may be through exchange of envelope solitary waves in fiber optics.

Classic examples of solitary waves decay exponentially fast away from the peak of the disturbance. In the last few years, as reviewed in the author's book [72] and also [56], it has become clear that solitary waves which flunk the decay condition are equally important. Such 'weakly nonlocal' solitary waves decay not to zero, but to an oscillation of amplitude α, the 'radiation coefficient' (Figure 5). The amplitude of these oscillations is important because it determines the radiative lifetime of the disturbance.

The complication is that for many nonlocal solitary waves, the radiation coefficient α is an exponential function of $1/\varepsilon$ where ε is a small parameter proportional to the amplitude of the maximum of the solitary wave. This implies that an ordinary asymptotic series in powers of ε:

– must fail to converge to the solution,
– must tell us nothing about whether the solitary waves are classical or weakly nonlocal,
– must be useless for computing α.

However, it *is* possible to compute the radiation coefficient through a *hyperasymptotic* approximation [68, 72].

A full treatment of a weakly nonlocal soliton is too complicated for an introduction to hyperasymptotics, but it is possible to give the flavor of the subject through the closely-related inhomogeneous ordinary differential equation studied by Akylas and Yang [5]

$$\varepsilon^2 u_{xx} + u - \varepsilon^2 u^2 = \operatorname{sech}^2(x). \tag{39}$$

To lowest order in ε, the second derivative is negligible compared to u, just as in our previous example, and the quadratic term is also small so that

$$u(x) \sim \operatorname{sech}^2(x). \tag{40}$$

By assuming $u(x)$ may be expanded as a power series in even powers of ε, substituting the result into the differential equation and matching powers one finds

$$u(x) \sim \sum_{j=0}^{\infty} \varepsilon^{2j} u_j, \qquad u_j \equiv \sum_{m=1}^{j+1} a_{jm} \operatorname{sech}^{2m}. \tag{41}$$

When this series is truncated to finite order, $j \leqslant N$, all terms in the truncation decay exponentially with $|x|$ and therefore so does the approximation u^N. In reality, the exact solution decays to an oscillation, just as in Figure 5. The 'wings' are invisible to the multiple scales/amplitude expansion because the amplitude α of the wings is an exponential function of $1/\varepsilon$.

Boyd shows [68] [with notational differences from this review] that the residual equation which must be solved at each order is

$$u_{N+1} = r(u^N), \tag{42}$$

where $r(u^N) \equiv -\{\varepsilon^2 u^N_{xx} + u^N - \varepsilon^2 (u^N)^2 - \operatorname{sech}^2(x)\}$ is the 'residual function' of the solution up to and including Nth order. When the order $N = N_{\text{optimum}} \sim -1/2 + \pi/(4\varepsilon)$, the Fourier transform of the residual is peaked at wavenumber $k = 1/\varepsilon$. In other words, when the series is truncated at optimal order, the neglected second derivative is just as important as u_{N+1} in consistently computing the correction at next order. The hyperasymptotic approximation is to replace Equation (42) by

$$\varepsilon^2 u_{N+1,xx} + u_{N+1} = r(u^N) \tag{43}$$

for all $N > N_{\text{optimum}}$.

The good news is that the *nonlinear* term in the original forced-KdV equation is still negligible on the left-hand side of the perturbation equations at each order (though it appears in the residual on the right-hand side). The bad news is that the equation we must solve to compute the hyperasymptotic corrections, although linear, does not admit a closed form solution except in the form of an integral which cannot generally be evaluated analytically:

$$\varepsilon^2 u^{N+1}(x) = \int_{-\infty}^{\infty} \frac{R_N(k)}{1 - \varepsilon^2 k^2} \exp(ikx)\, \mathrm{d}k, \tag{44}$$

where $R_N(k)$ is the Fourier transform of the residual of the Nth order perturbative approximation.

The Euler expansion cannot help; a weighted sum of the terms of the original asymptotic series must decay exponentially with $|x|$ and therefore will miss the oscillatory wings. The integrand in Equation (44) is now *singular on the integration interval*, rather than off it as for the Stieltjes function. Indeed, when $N \approx N_{\text{optimum}}(\varepsilon)$, the numerator of the integrand is largest at $|k| = 1/\varepsilon$, precisely where the denominator is singular! No simple change in the center of the Taylor expansion of the denominator factor $1/(1 - \varepsilon^2 k^2)$ will help here.

Fortunately, it is possible to *partially* solve Equation (43) in the sense that we can analytically determine the amplitude of the radiation coefficient α. Boyd [68] shows that α is just the Fourier transform of the residual at the points of singularity. The result is an approximation to $\alpha(\varepsilon)$ with relative error $O(\varepsilon^2)$. This can be extrapolated to the limit $\varepsilon \to 0$ to obtain

$$\alpha(\varepsilon) = \frac{1.558823 + O(\varepsilon^2)}{\varepsilon^2} \exp\left(-\frac{\pi}{2\varepsilon}\right), \quad \varepsilon \ll 1. \tag{45}$$

As for the Stieltjes integral, several different hyperasymptotic methods are available for weakly nonlocal solitary waves and related problems. The most widely used is to match asymptotic expansions near the singularities of the solitary wave on the imaginary axis. Originally developed by Pokrovskii and Khalatnikov [262] for ‘above-the-barrier’ quantum scattering (WKB theory in the absence of a turning point), it was first applied to nonlinear problems by Kruskal and Segur [278, 172]. The book by Boyd [72] reviews a wide number of applications and improvements to the PKKS method.

Akylas and Yang [5, 323 – 325, 327] apply multiple scales perturbation theory in wavenumber space after a Fourier transformation. Chapman, King and Adams [96], Costin [104, 105] and Costin and Kruskal [106, 107], Écalle [123] have all shown that related but distinct methods can also be applied to nonlinear differential equations.

8. Overview of Hyperasymptotic Methods

Hyperasymptotic methods include the following:

(1) (Second) Asymptotic Approximation of Error Integral or Residual Equation for Superasymptotic Approximation
(2) Isolation Strategies, or Rewriting the Problem so the Exponentially Small Thing is the Only Thing
(3) Resurgence Schemes or Resummation of Late Terms
(4) Complex-Plane Matching of Asymptotic Expansions
(5) Special Numerical Algorithms, especially Spectral Methods
(6) Sequence Acceleration including Padé and Hermite–Padé Approximants
(7) Hybrid Numerical/Analytical Perturbative Schemes

The labels are suggestive rather than mutually exclusive. As shown amusingly in Nayfeh [229], the same asymptotic approximation can often be generated by any of half a dozen different methods with seemingly very dissimilar strategies. Thus, the Euler summation gives the exact same sequence of approximations, when applied to the Stieltjes function, as making a power series expansion in the error integral for the superasymptotic approximation.

In the next few sections,we shall briefly discuss each of these general strategies in turn.

9. Isolation of Exponential Smallness

Long before the present surge of interest in exploring the world of the exponentially small, some important problems were successfully solved without benefit of any of the strategies of modern hyperasymptotics. The key idea is *isolation*: in the region of interest (perhaps after a transformation or rearrangement of the problem), the exponentially small quantity is the only quantity so that it is not swamped by other terms proportional to powers of ε.

A quantum mechanical example is the 'WKB', 'phase-integral' or 'Liouville–Green' calculation of 'Below-the-Barrier Wave Transmission'. The goal is to solve the stationary Schrödinger equation

$$\psi_{xx} + \{k^2 - V(\varepsilon x)\}\psi = 0 \tag{46}$$

subject to the boundary conditions of (i) an incoming wave from the left of unit amplitude and (ii) zero wave incoming from the right:

$$\begin{aligned} &\psi \sim \exp(ikx) + \alpha \exp(-ikx), \quad x \to -\infty; \\ &\psi \sim \beta \exp(ikx), \quad x \to \infty. \end{aligned} \tag{47}$$

The goal is to compute the amplitudes of the reflected and transmitted waves, α and β, respectively. If $k^2 < \max(V(\varepsilon x))$, however, β is exponentially small in $1/\varepsilon$ for fixed k, and α differs from unity by an exponentially small amount also. Nevertheless, this problem was solved in the 1920's as reviewed in Nayfeh [229] and Bender and Orszag [19].

The crucial point is that on the right side of the potential barrier, the exponentially small transmitted wave is the entire wavefunction. There is no ambiguity: far to the right, the WKB approximation must approximate a transmitted, rightgoing wave and nothing else. This, in an analysis too widely published to be repeated here, allows the analytical determination of β through standard WKB or matched asymptotics expansions.

In contrast, standard WKB is quite impotent for determining the difference between the amplitude of the reflected wave and one because the large reflected wave swamps the exponentially small correction. However, α is easily found *indirectly* by combining the known values of the incoming and transmitted waves with conservation of energy. Similarly, WKB gives a good approximation to the bound

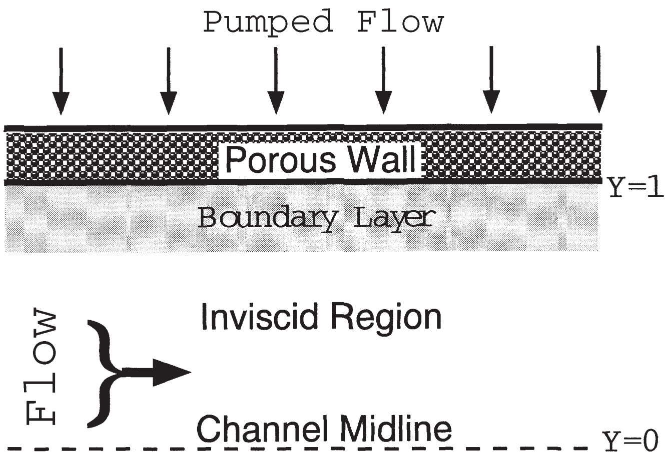

Figure 6. Schematic of the Berman–Terrill–Robinson problem. Fluid in the channel flows to the right, driven partly by fluid pumped in through the porous wall. Only half of the channel is shown because the flow is symmetric with respect to the midline of channel (dashed).

states and eigenvalues of a potential well: where the wavefunction is exponentially small (for large $|x|$), there is no competition from terms that are larger.

A nonlinear example is the 'Berman–Terrill–Robinson' or 'BTR' problem, which is interesting in both fluid mechanics and plasma physics [135, 154, 108, 193, 186, 109]. In its mechanical engineering application, the goal is to calculate the steady flow in a pipe or channel with porous walls through which fluid is sucked or pumped at a constant uniform velocity V. Berman [23] showed that for both the pipe and channel, the problem could be reduced to a nondimensional, ordinary differential equation which in the channel case is

$$\varepsilon f_{YYY} + f_Y^2 - f f_{YY} = \alpha^2, \tag{48}$$

where α is the eigenparameter which must be computed along with $f(Y)$. The boundary conditions are

$$f(1) = 1, \qquad f_Y(1) = 0, \qquad f(0) = 0, \qquad f_{YY}(0) = 0. \tag{49}$$

The small parameter is $\varepsilon = 1/R$ where R is the usual hydrodynamics 'Reynolds number' (very large in most applications). Symmetry with respect to the midline of the channel (at $Y = 0$) is assumed.

By matching asymptotic expansions, boundary layer to inviscid interior (Figure 6), one can easily compute a solution in powers of ε. Unfortunately, the numerical work of Terrill and Thomas [292] showed that there are actually *two* solutions for the circular pipe for all Reynolds numbers for which solutions exist. Terrill correctly deduced that the two modes differed by terms *exponentially small* in the Reynolds number (or equivalently, in $1/\varepsilon$) and analytically derived them in 1973 [291], quite independently of all other work on hyperasymptotics.

The early numerical work on the porous channel was even more confusing [265], finding one or two solutions where there are actually three. Robinson resolved these uncertainties in a 1976 article that combined careful numerical work with the analytical calculation of the exponentially small terms which are the sole difference between the two physically interesting solutions.

The reason that the exponential terms could be calculated without radical new technology is that the solution in the inviscid region ('outer' solution) is linear in Y plus terms exponentially small in ε:

$$f(Y) \sim \alpha(\varepsilon)Y + \gamma(\varepsilon)\left\{-3\frac{\varepsilon}{\alpha} + Y^3\right\} + \cdots, \tag{50}$$

$$\begin{aligned}\gamma(\varepsilon) \;=\; &\pm\frac{1}{6}\left(\frac{2}{\pi\varepsilon^7}\right)^{1/4}\exp\left(-\frac{1}{4}\right)\exp\left(-\frac{1}{4\varepsilon}\right)\times\\ &\times\left\{1-\frac{5}{4}\varepsilon-\frac{253}{32}\varepsilon^2+\mathrm{O}(\varepsilon^3)\right\}.\end{aligned} \tag{51}$$

(Note that because of the $\pm$ sign, there are *two* solutions for γ, reflecting the exponentially small splitting of a single solution (in a pure power series expansion) into the dual modes found numerically.) It follows that by making the almost trivial change-of-variable

$$g \equiv f - \alpha Y \tag{52}$$

we can recast the problem so that the 'outer' approximation is proportional to $\exp(-1/(4\varepsilon))$. Systematic matching of the 'inner' (boundary layer) and 'outer' flows gives the exponentially small corrections in the boundary layer, too, even though there are nonexponential terms in this region.

Other fluid mechanics cases are discussed in Notes 10 and 11 of the 1975 edition of Van Dyke's book [298]. Bulakh [85] as early as 1964 included exponentially small terms in the boundary-layer solution to converging flow between plane walls and showed that such terms will also arise at higher order in flows with stagnation points. Adamson and Richey [2] found that for transonic flow with shock waves through a nozzle, exponentially small terms are as essential as for the BTR problem.

Happily, there is a widely-applicable strategy for isolating exponential smallness which is the theme of the next section. The key idea is that the optimal truncation of the ε power series is always available to rewrite the problem in terms of a new unknown which is the *difference* between the original $u(x;\varepsilon)$ and the optimally-truncated series. Because this difference $\delta(x;\varepsilon)$ is exponentially small in $1/\varepsilon$, we can determine it without fear of being swamped by larger terms.

10. Darboux's Principle and Resurgence

> Evidently, the determination of the remainder [beyond the superasymptotic approximation] entails the evaluation of several transcendental functions. In other words, the calculation of the correction can be more formidable than that of the original asymptotic expansion. One is reminded of the dictum, sometimes asserted in physics, that getting an extra decimal place demands 100 times the effort expended on the previous one. Fortunately, the multiplying factor is not so huge in our case but it is perforce appreciable.
>
> D. S. Jones (1990) [155, p. 261]

Jones' mildly pessimistic remarks are still true: hyperasymptotics is more work than superasymptotics and one does have to evaluate additional transcendentals. However, Dingle showed in a series of articles in the late fifties and early sixties, collected in his 1973 book, that there is a suprising universality to hyperasymptotics: a quartet of generic transcendentals suffices to cover almost all cases. The key to his thinking, refined and developed by Berry and Howls, Olver and many others, is the following.

DEFINITION 5. (Darboux's Principle). One may derive an asymptotic expansion in degree j for the coefficients a_j of a series solely from knowledge of the *singularities* of the function $f(z)$ that the series represents. This principle applies to power series [110, 111, 123, 82, 83], Fourier, Legendre and Chebyshev series [55], and divergent power series [118].

'Singularity' is a collective terms for poles, branch points and other points where a complex function $f(z)$ ceases to be an analytic function of z. If $f(z)$ is singular, on the same Riemann sheet as the origin, at the set of points $\{z_j\}$, then the radius of convergence of the power series for $f(z)$ is $\rho = \min |z_j|$, as proven in most introductory calculus courses. Darboux showed that if the convergence-limiting singularity was such that $f(z) = (z - z_c)^r g(z)$ where $g(z)$ is nonsingular at the convergence-limiting singularity, then the power series coefficients are asymptotically (if $j \neq$ integer)

$$a_j \sim j^{-1-r} z_c^{-j} \{1 + \mathrm{O}(1/j)\}. \tag{53}$$

Asymptotics-from-singularities can be extended to logarithms and other singularities, too. As reviewed in [55], one can derive similar asymptotic approximations to the coefficients of Fourier, Chebyshev, Legendre and other orthogonal expansions from knowledge of the singularities of $f(z)$.

Dingle [116, 117] realized in the late 50's that Darboux's Principle applies to divergent series, too. If one makes an asymptotic expansion by performing a power series expansion inside an integral and then integrating term-by-term, the coefficients of the divergent expansion will be simply those of the power series in the integration variable multiplied by the effect – usually a factorial – of the term-by-term integration. For example, consider the class of functions

$$f(\varepsilon) \equiv \int_0^\infty \exp(-t)\Phi(\varepsilon t)\,\mathrm{d}t, \tag{54}$$

where $\Phi(z)$ has the power series

$$\Phi(z) = \sum_{j=0}^{\infty} b_j z^j \tag{55}$$

then

$$f(\varepsilon) \sim \sum_{j=0}^{\infty} a_j \varepsilon^j; \quad a_j = j! b_j. \tag{56}$$

Because the coefficients of the divergent series $\{a_j\}$ are merely those of the power series of Φ, multiplied by $j!$, it follows that the asymptotic behavior of the coefficients of the divergent series must be controlled by the singularities of $\Phi(z)$ as surely as those of the power series of Φ itself. In particular, the *singularity of the integrand which is closest to* $t = 0$ *must determine the leading order of the coefficients of the divergent expansion*. This implies that all $f(\varepsilon)$ that have a function $\Phi(z)$ with a convergence-limiting singularity of a given type (pole, square root, etc.) and a given strength (the constant multiplying the singularity) at a given point z_c will have coefficients that asymptote to a common form, even if the functions in this class are wildly different otherwise.

EXAMPLE. The 'double Stieltjes' function

$$SD(\varepsilon) \equiv S(\varepsilon) + S(\varepsilon/2), \tag{57}$$

where $S(\varepsilon)$ is the Stieltjes function described earlier. The asymptotic series is

$$SD(\varepsilon) \sim \sum_{j=0}^{\infty} a_j \varepsilon^j; \quad a_j = (-1)^j \; j! \left\{ 1 + \frac{1}{2^j} \right\}. \tag{58}$$

The integrand of $S(\varepsilon)$ is singular at $t = -1/\varepsilon$ while that of $S(\varepsilon/2)$ is singular twice as far away at $t = -2/\varepsilon$. In the braces in Equations (58), the first and nearer singularity contributes the one while the rapidly decaying factor $1/2^j$ comes from the more distant pole of the integrand, that of $S(\varepsilon/2)$. The crucial point is that in the limit $j \to \infty$, the coefficients of the divergent series for the double Stieltjes function asymptote to those of the ordinary Stieltjes function.

As explained above, the optimal truncation of the ε power series for the Stieltjes function is to stop at $N = [1/\varepsilon]$, that is, at the integer closest to the reciprocal of ε; the error in the resulting 'superasymptotic' approximation is proportional to

$\exp(-1/\varepsilon)$. The dominance of the asymptotic coefficients of the double Stieltjes function by the pole at $t = -1/\varepsilon$ implies that all these conclusions should apply to the optimal truncation of the divergent expansion for $SD(\varepsilon)$, too:

$$SD(\varepsilon) \sim \sum_{j=0}^{N_{\text{opt}}} a_j \varepsilon^j + \mathrm{O}\Big(\sqrt{\pi/(2\varepsilon)} \exp(-1/\varepsilon)\Big); \quad N_{\text{opt}}(\varepsilon) = [1/\varepsilon], \tag{59}$$

where the factor in front of the exponential is justified in [19]. More important, if we add the error integral for Stieljes function to the $N_{\text{opt}}(\varepsilon)$-term truncation of the series for the double Stieltjes function, we should obtain an improved approximation. Since the first neglected term in the series for $SD(\varepsilon)$ differs from that included in the Stieltjes error integral by a relative error of $\mathrm{O}(\varepsilon^{2N})$, the best we can hope for is to improve upon the superasymptotic approximation by a factor of 2^N, which, because $N_{\text{opt}} \approx 1/\varepsilon$, can be rewritten as $\exp(-\log(2)/\varepsilon)$. Thus,

$$\begin{aligned} &SD(\varepsilon) \sim \sum_{j=0}^{N} a_j \varepsilon^j + E_N(\varepsilon) + \mathrm{O}\big(\exp(-\{1 + \log(2)\}/\varepsilon)\big); \\ &N(\varepsilon) = [1/\varepsilon], \end{aligned} \tag{60}$$

where $E_N(\varepsilon)$ is the error integral for the Stieltjes function defined by Equation (8). Figure 7 shows that the error estimate in Equation (60) is accurate.

If the location of the second-worst singularity is known – that is, the pole or branch point of the integrand which is closer to $t = 0$ than all others except the one which asymptotically dominates – one can do better. Since the second pole of $SD(\varepsilon)$ is at twice the distance of the first, if we add the next N contributions of the *second singularity only* only to the approximation of Equation (60), the result should be as accurate as the optimal truncation of a series derived from the second singularity (i.e., $S(\varepsilon/2)$ for this example), that is, have an error proportional to $\exp(-2/\varepsilon)$:

$$SD(\varepsilon) \sim \sum_{j=0}^{N} a_j \varepsilon^j + E_N(\varepsilon) + \sum_{j=N+1}^{2N} (-1)^j \frac{j!}{2^j} \varepsilon^j + \mathrm{O}\left\{\exp\left(-\frac{2}{\varepsilon}\right)\right\}. \tag{61}$$

Figure 7 confirms this. (Howls [147] and Olde Daalhuis [241] have developed improved hyperasymptotic schemes with smaller errors, but for expository purposes, we have described the simplest approach.)

A key ingredient in Dingle's strategy is Borel summation. Under certain conditions [318], a divergent series can be summed by the integral of $\exp(-t)$ multipied by a function $\Phi(\varepsilon t)$ which is defined to be that function whose power series has the coefficients of the divergent series divided by $j!$. That is to say, the integral in (54) is the Borel sum of the ε power series for the function $f(\varepsilon)$ on the left in the same equation. (We are again reminded of the interplay between different strategies in hyperasymptotics; a series acceleration method, which is a hyperasymptotic method in its own right when combined with Padé approximation of $\Phi(\varepsilon t)$ ['Padé–Borel' method [315, 316]], is also a key justification for a different and sometimes

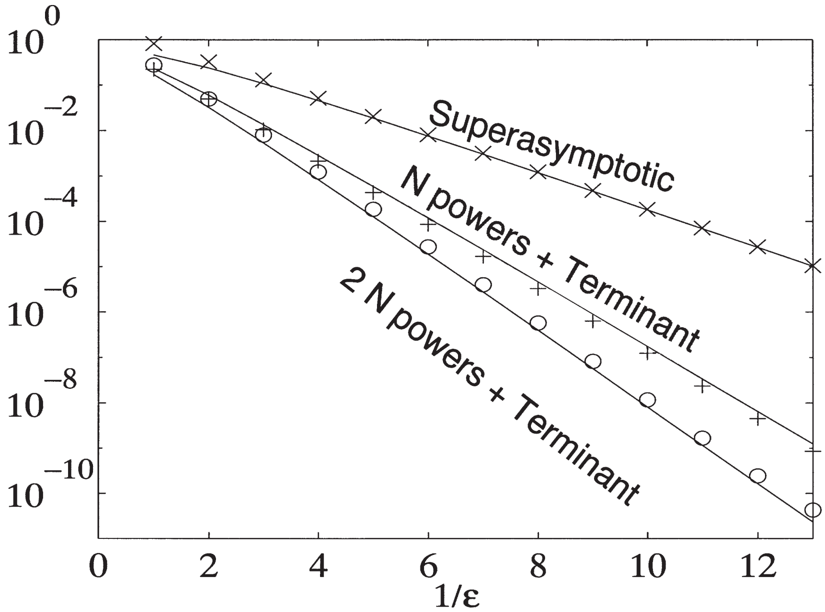

Figure 7. Double Stieltjes function: errors in three approximations. x's: Errors in optimally-truncated asymptotic series (the 'superasymptotic' approximation. Pluses: Superasymptotic approximant plus the 'terminant'. Circles: Approximation defined by Equation (55). Solid curves: Predicted errors, which are respectively the following – (top) $q \exp(-1/\varepsilon)$, (middle) $q \exp(-1.693/\varepsilon)$, (bottom) $q \exp(-2/\varepsilon)$ where $q(\varepsilon) \equiv (\pi/(2\varepsilon))^{1/2}$.

more powerful hyperasymptotic scheme.) Dingle's twist is that he applies Borel summation only to the *late* terms in the asymptotic series. The first few terms in the sum for $SD(\varepsilon)$ are very different from those of the Stieltjes function; the only way to obtain the right answer is to sum these leading terms directly without tricks. Dingle's key observation is that the *late* terms, meaning those neglected in the optimal truncation, are essentially the same as those for the ordinary Stieltjes function. Thus, the error integral $E_N(\varepsilon)$ for *one* function, $S(\varepsilon)$, provides a hyperasymptotic approximation to an entire *class* of functions, namely all those of the form of Equation (54) for which the convergence-limiting singularity of $\Phi(z)$ is a simple pole at $z = -1$.

It might seem as if we would have to repeat the analysis for each different species of singularity – one family of error integrals when the singularity is a simple pole, another when the dominant singularity of Φ is a logarithm and so on. In reality, Dingle shows that for a very wide range of asymptotic expansions, both from integral representations, the WKB method, and so on, the coefficients are asymptotically of the form

$$a_j \sim q(-1)^j \frac{\Gamma(j+1-\beta)}{\rho^{j+1-\beta}} \tag{62}$$

for some constants q, ρ and β. The error integral for the Stieltjes function is almost the theory of everything.

In the next three sections, we describe how Dingle's theory has been extended to the method of steepest descent and the mystery of Stokes phenomenon. A couple of historical, semantic, and notational grace notes are needed first, however.

The first is that the work of Dingle and others is couched not in terms of the error integrals $E_N(\varepsilon)$ but rather in terms of the following:

DEFINITION 6 (Terminant). A function $T_N(\varepsilon)$ is a '*terminant*' if it is used to weight the Nth term in an asymptotic series so as to approximate the exact sum.

The reason for working with terminants instead of errors is mostly historical. Stieltjes [286] showed that for an alternating series, one could considerably improve accuracy for both convergent and divergent series merely by multiplying the last retained term by a weight factor of 1/2. Airey developed an early (1937) hyperasymptotic method, restricted to alternating series for which the general term is known, which computed an improved, N-dependent replacement for Stieltjes' 1/2 [3]. Later studies have generally followed this convention. However, terminants are sometimes more convenient than error integrals as in the smoothing of Stokes phenomenon.

The second comment is that Dingle found it helpful to define four canonical (approximate) terminants instead of one. One reason is that the Stieltjes error integral, and the equivalent terminant, have poles on the negative real axis away from the integration interval, which is the positive real axis. Stokes phenomenon happens when the poles coincide with integration interval, which makes it convenient to define a second terminant. Dingle's fundamental pair are

$$\Lambda_m(1/\varepsilon) \equiv \frac{1}{\Gamma(m+1)} \int_0^\infty \mathrm{d}t \frac{\exp(-t)t^m}{1+\varepsilon t}, \tag{63}$$

$$\overline{\Lambda}_m(1/\varepsilon) \equiv \frac{1}{\Gamma(m+1)} P \int_0^\infty \mathrm{d}t \frac{\exp(-t)t^m}{1-\varepsilon t}, \tag{64}$$

where P denotes the Cauchy Principal Value of the integral. These two fission into two more because many expansions proceed in powers of ε^2 rather than ε itself, which makes it convenient to define terminants for even powers of ε, his Π_m and $\overline{\Pi_m}$.

Furthermore, newer classes of problems have required additional terminants, as illustrated in Delabaere and Pham [113]. When the hyperasymptotic process is iterated so as to add additional terms, with different scalings, one needs generalizations of the Dingle terminants called 'hyperterminants'. Olde Daalhuis [240, 242] has given algorithms for the numerical computation of terminants and hyperterminants. The need for these generalizations, however, should not obscure the fundamental unity of the idea of adding error integrals or terminants that match the *dominant* singularity to convert a superasymptotic approximation into a hyperasymptotic approximation.

There is a close parallel between Dingle's universal terminants for asymptotic series and the universal error envelopes for Chebyshev and Fourier spectral methods which were derived by Boyd [55, 59]. For example, Boyd found that the error envelope was always a linear combination of the same two meromorphic functions (the 'Lorentzian' and 'serpentine' functions, defined in [59]), regardless of whether the function being interpolated was entire, meromorphic, or had logarithmic singularities. Even when $f(x)$ is nonanalytic but infinitely differentiable at a point on the expansion interval, and thus has only a divergent power series about that point, the error envelope is the sum of these two functions. The reason for the similarity is that Darboux's Principle applies to Fourier and Chebyshev series, too. Asymptotically, functions that are very dissimilar in their first few terms resemble each other more and more closely in the late terms. One or two terminants can encapsulate the error for very different classes of functions, even ones whose late coefficients are decaying, because of the magic of Taylor expansions with respect to degree.

11. Steepest Descents

> The resultant series is asymptotic, rather than convergent, because the range of integration extends beyond the circle of convergence of [the power series of the metric factor], the radius of this circle being fixed by the zero of $\mathrm{d}\phi/\mathrm{d}t$ in the complex w-plane lying closest to the origin.
>
> R. B. Dingle [118, p. 111], with translation of notation into the symbols used in the section below.

The method of steepest descent is commonly applied to evaluate the integral

$$I(z) \equiv \int \exp(z\phi(t))\,\mathrm{d}t \tag{65}$$

in the limit $|z| \to \infty$. As described in standard texts [19], the 'saddle points' or 'stationary points' $\{t_s\}$ play a crucial role where these are defined as the roots of the first derivative of the 'phase function' $\phi(t)$:

$$\frac{\mathrm{d}\phi}{\mathrm{d}t}(t_s) = 0. \tag{66}$$

The path of integration is deformed so as to pass through one or more saddle points. The next step is to identify the *dominant* saddle point, which is the one *on the deformed contour of integration* for which $\Re(\phi(t_s))$ is largest. Restricting t_s to the dominant saddle point, one then makes the exact change-of-variable

$$w \equiv \sqrt{\phi(t_s) - \phi(t)} \tag{67}$$

so that the integral becomes

$$I(z) = \int \exp(-zw^2)\frac{\mathrm{d}t}{\mathrm{d}w}(w)\,\mathrm{d}w. \tag{68}$$

The final steps are (i) extend the integration interval to the entire real w-axis and (ii) expand the 'metric factor' $\mathrm{d}t/\mathrm{d}w$ in powers of w and integrate term-by-term to obtain an exponential factor multiplied by an inverse power series in the large parameter z. (By setting $\varepsilon = 1/z$, this series is similar – and similarly divergent – to the ε power series explored earlier.) We omit details and generalizations because the mechanics are so widely described in the literature [19, 319].

Unfortunately the standard texts hide the fact that the asymptotic expansion is based on the same mathematical atrocity as the divergent series for the Stieltjes function: employing a power series in the integration variable with a *finite* radius of convergence under integration over an *infinite* interval. Hyperasymptotics is greatly simplified by the following.

THEOREM 1 (Singularities of the Steepest Descent Metric Function). *If an integral of the form of Equation* (65) *is transformed by the mapping Equation* (67) *into the integral over w, Equation* (68), *then the metric factor* $\mathrm{d}t/\mathrm{d}w$ *has branch points of the form*

$$\frac{\mathrm{d}t}{\mathrm{d}w} = \frac{g(w)}{\sqrt{w - w_s}} + h(w), \tag{69}$$

where $g(w)$ and $h(w)$ are analytic at $w = w_s$. All such points w_s are the images of the saddle points t_s under the mapping $w(t)$; conversely, the metric factor is singular at all points w_s which are images of saddle points except for $w = 0$. The metric factor may also be singular with singularities of more complicated type at points w which are images of points where the 'phase factor' $\phi(t)$ is singular.

Proof. The first step is to differentiate the definition of the mapping Equation (67) to obtain

$$\frac{\mathrm{d}w}{\mathrm{d}t} = -\frac{1}{2\sqrt{\phi(t_s) - \phi(t)}}\frac{\mathrm{d}\phi}{\mathrm{d}t} \quad \Longleftrightarrow \quad \frac{\mathrm{d}t}{\mathrm{d}w} = -2\frac{w}{\mathrm{d}\phi/\mathrm{d}t}. \tag{70}$$

This shows that the metric factor can be singular only at the w-images of those points t in the original integration variable where (i) $\phi(t)$ is singular or (ii) saddle points where by the very definition of a saddle point, $\mathrm{d}\phi/\mathrm{d}t = 0$ and the denominator of the right-hand side of Equation (70) is zero. This is really just a restatement of the implicit function of elementary calculus, which states that if $\mathrm{d}w/\mathrm{d}t$ is non-zero at a point, then the inverse function $t(w)$ exists and is analytic at that point and its derivative $\mathrm{d}t/\mathrm{d}w = 1/(\mathrm{d}w/\mathrm{d}t)$. The point $w = 0$ is exceptional because the numerator of the right-hand side (w) cancels the zero in the denominator.

To obtain an expression for $\mathrm{d}t/\mathrm{d}w$ in the neighborhood of a saddle point, we expand $w(t)$ about the saddle point $t = t_s$. The constant w_s can be moved to the left side of the equation and the linear term is zero because $\mathrm{d}w/\mathrm{d}t$ is zero at the saddle point. Taking the square root gives

$$\sqrt{w - w_s} = (t - t_s)\sqrt{\frac{1}{2}\frac{\mathrm{d}^2 w}{\mathrm{d}t^2}(t_s)}\left\{1 + \mathrm{O}(t - t_s)\right\}. \tag{71}$$

It follows that $\mathrm{d}t/\mathrm{d}w$ is proportional to $1/\sqrt{w - w_s}$ near the saddle point, which demonstrates the theorem.

Denote the image-of-a-saddle-point of smallest absolute value by w_{min}. The coefficients b_j of the power series of the integrand will then asymptote, for sufficiently high degree j, to those of a constant times $1/\sqrt{w - w_{\mathrm{min}}}$; the contributions of the singularities that are more remote in the complex w-plane will decrease exponentially fast with j compared to the contribution of the square root branch point at $w = w_{\mathrm{min}}$. Applying the binomial theorem to compute the power series coefficients of the square root singularity and then integrating term-by-term shows that the coefficients a_j of the asymptotic series for the integral itself will asymptote for large j to

$$a_j \sim q\frac{\Gamma(j + 1/2)}{|w_{\mathrm{min}}|^{j+1/2}}, \tag{72}$$

where the constant q is proportional to $g(w_{\mathrm{min}})$ in the theorem. Dingle [118, p. 457], gives the basic terminant (with some changes in notation)

$$T_N \sim q_0\Lambda_{N-1}(-F) + q_2\Lambda_{N-2}(-F) + q_4\Lambda_{N-3}(-F) + \cdots, \tag{73}$$

where the q_{2j} are functions of Dingle's 'chief singulant' F, which in our notation is

$$F \equiv z w_{\mathrm{min}}^2 \tag{74}$$

and $q_{2j} \sim \mathrm{O}(F^j)$. This situation is more complicated than for the double Stieltjes function in that we have a *series* of terminants, rather than a single terminant. (Each term of the expansion of $\mathrm{d}t/\mathrm{d}w$ in half-integral powers of $w - w_{\mathrm{min}}$ will generate its own contribution to the terminant series.) The underlying ideas remain simple even though the algebraic complexity rapidly leaves one muttering: 'Thank heavens for Maple! [and similar symbolic manipulation languages like Mathematica, Reduce and so on].' □

12. Stokes Phenomenon

> about the present title [*Divergent Series*], now colourless, there hung an aroma of paradox and audacity.
>
> Sir John E. Littlewood (1885–1977) [139]

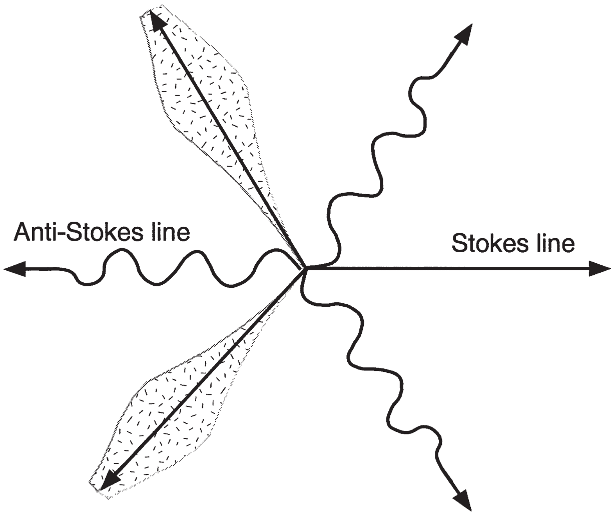

Figure 8. Stokes Lines (Monotonic Growth/Decay) and anti-Stokes Lines (Pure Oscillation) for the Airy Functions Ai and Bi. The shaded regions show the transition zone for the Stokes' multiplier of Ai, that is, the regions where it varies from 1 to 0 as an error function. The positive real axis is a Stokes Line for Bi but not Ai. The shaded regions narrow for large $|z|$ because for the Airy function, the width of the transition zone, expressed in terms of the angle $\theta \equiv \arg(z)$, decays as $|z|^{-3/4}$.

Stokes phenomenon has contributed much to the 'aroma of paradox and audacity' of asymptotic series. It is easiest to explain by example.

The Airy function $\mathrm{Ai}(z)$ asymptotes for large positive z to the product of a decaying exponential with a series in inverse powers of $z^{3/2}$. For negative real z, the Airy function is real and oscillatory; approximated by the product of a cosine with a inverse power series plus a sine with a different inverse power series. However, the multiplier of the leading inverse power, the cosine, is the sum of *two* exponentials. If we track the asymptotic approximation for fixed $|z|$ as $\theta = \arg(z)$ varies from 0 to π, one exponential must somehow metamorphosize into two.

The classical analysis hinges on two species of curves in the complex z-plane: 'Stokes lines', where the exponentials grow or decay without oscillations, and 'anti-Stokes' lines where the exponentials oscillate without change in amplitude.* (Figure 8.) Stokes' own interpretation is that the coefficient of the 'recessive' (decaying) exponential jumps discontinuously on the Stokes line (for $\mathrm{Ai}(z)$, at $\arg(z) = \pm 2\pi/3$), that is, where this exponential is smallest relative to the 'dominant' exponential that grows as $|z|$ increases along the Stokes line. As the negative real axis (an anti-Stokes line) is approached, the two exponentials become more and more similar until finally both are purely oscillatory with coefficients of equal magnitude on the anti-Stokes line itself.

---

* We employ the convention of Heading, Dingle, Olver, and Berry, but other authors such as Bender and Orszag reverse the meaning of 'Stokes' and 'anti-Stokes'.

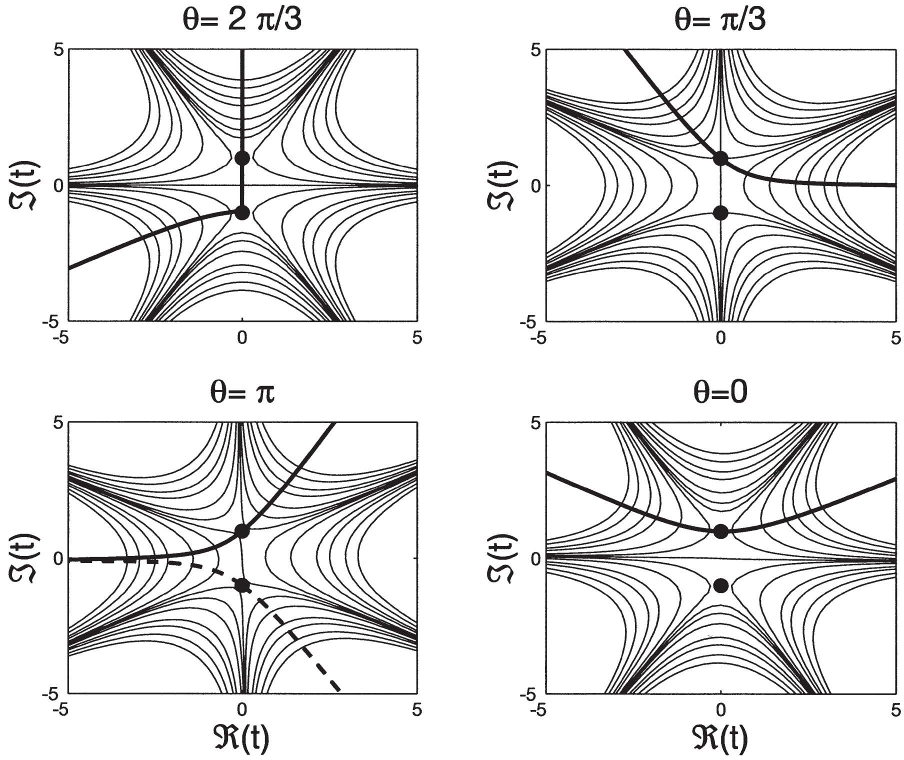

Figure 9. Steepest descent paths of integration in the complex plane of the original integration variable t for four different values of z. The two saddle points are marked by black discs. The contours of $\log(|\exp(z^{3/2}i\{t+t^3/3\})|)$ are also shown. For $\theta = \pi$ [negative real z-axis, lower left panel], the integration contour comes from large t in the upper right quadrant, returns to infinity along the negative t-axis, and then returns to pass through the right saddle point and depart to infinity via the lower right t-quadrant.

The annoying and unsatisfactory part of this discontinuous jump is that the Airy function itself is an entire function, completely free of all jumps, infinities and pathologies of all kinds except at $|z| = \infty$. Sir Michael Berry has recently smoothed this 'Victorian discontinuity', to quote from one of his papers, by combining Dingle's ideas with the standard and long-known asymptotic approximation to an integral when the saddle point and a pole nearly coincide. To understand Berry's jump-free hyperasymptotics, we need some preliminaries.

First, let us represent the solution by an integral which can be approximated by the method of steepest descent for large z. (Berry's smoothing is equally applicable to WKB approximations to differential equations and a wide variety of other asymptotics, but steepest descent is the most convenient for explaining the concepts.)

The integral representation for the Airy function is

$$\mathrm{Ai}(z) \equiv \frac{z^{1/2}}{2\pi} \int_{\mathcal{C}} \exp\bigl(z^{3/2}\bigl\{i(t + t^3/3)\bigr\}\bigr)\,\mathrm{d}t, \tag{75}$$

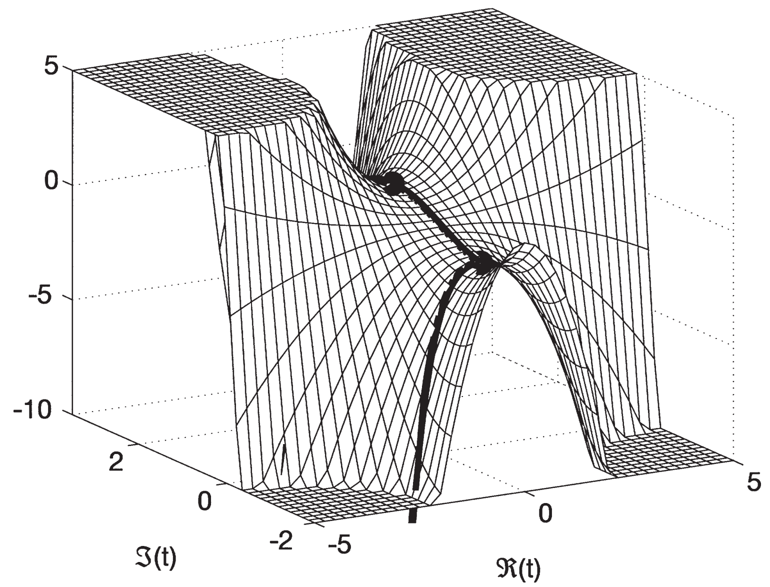

Figure 10. A surface plot of $\log(|\exp(z^{3/2}\phi)|)$ for the Airy integral for $\arg(z) = 2\pi/3$, that is, on the Stokes line. The steepest descent path is marked by the heavy solid line; the disks denote the two saddle points at $t_s = \pm i$. The surface has been truncated at the vertical axis limits for graphical clarity.

where $\mathcal{C}$ is a contour that originates at infinity at an angle $\arg(t) = (5/6)\pi - (1/2)\arg(z)$ and returns to infinity at $\arg(t) = (1/6)\pi - (1/2)\arg(z)$.

Figure 9 shows the steepest descent paths of integration for the Airy integral representation. As explained in the preceding section, the easiest way to generate the coefficients of the asymptotic series is to begin with a change-of-coordinate to a new integration variable w. To illustrate the key topological ideas, however, it is perhaps more illuminating to illustrate the steepest descent path in the t-plane as we have done in the figure. In either plane, the path of integration is deformed so as to pass through a saddle point, and then curve so that at each point t on the path, $\Im(z\phi(t_s)) = \Im(z\phi(t_s))$. This condition that the phase of the integrand matches that of the saddle point ensures that the magnitude of the integrand decreases as steeply as possible from its local maximum, i.e., that the curve is really is a path of 'steepest descent'.

As a student, I was much puzzled because my texts and teachers expended a lot of energy on determining the exact shape of the steepest descent contour even though it does not appear *explicitly* in the answer, even at higher order! The steepest descent path is actually important only for *topological* reasons: it is essential to know *which* saddle points lie on the path, but the shape of the contour is otherwise irrelevant.

For the Airy function, for example, there are two saddle points for all z but only one is on the contour for large positive z. As the argument of z varies, however, the steepest descent paths in the t (or w) planes must vary also. For some z, the steepest descent path through one saddle point must collide with the other; this happens precisely on the Stokes lines.

As shown in Figure 9, the Stokes lines are a change in the *topology* of steepest descent paths: a single saddle point on the contour on one side of the Stokes line, two saddle points on the other side of the Stokes line and on the Stokes line itself. Thus, Berry's title for one of his articles, 'Smoothing a Victorian discontinuity', is a bit misleading since the discontinuity is not removed in a *topological* sense. The jump is, however, smoothed *numerically*.

Parenthetically, note that at the Stokes line itself ($\arg(z) = 2\pi/3$ for $\mathrm{Ai}(z)$), the steepest descent path descends from one saddle point directly to another saddle point, then makes a right angle turn and then continues to descend monotonically from the second saddle point (Figure 10). For $\arg(z) > 2\pi/3$, the steepest descent contour from one saddle point does not runs off to infinity parallel to the negative imaginary axis. To be continuous and still terminate at $\infty \exp(i\pi/6)$, however, the contour must return and pass through the second saddle point. At $\arg(z) = \pi$, the contributions of both saddle points are equal.

The properties of Stokes lines may be summarized as follows:

(1) There are TWO saddle points on the steepest descent integration path in the t-plane.
(2) $\Im\{z(\phi(t_+) - \phi(t_-)\} = 0$ where t_+ and t_- are the two saddle points on the steepest descent contour and where $\phi(t)$ is the steepest descent phase function defined by Equation (65).
(3) The terminants for the series each have a simple pole *on* the real w-axis, which is the integration interval after the usual steepest descent change of variable, the poles being at the saddle point which contributes the 'recessive' saddle point.
(4) The terms b_j of the asymptotic inverse power series are, for sufficiently large degree j, all of the same sign.

When there is a discontinuity in asymptotic form, the first three properties are each equivalent definitions of a 'Stokes line'.

The proofs of these assertions and also generalizations of Stokes phenomenon to solutions of nonlinear differential equations and so on are given by the theory of 'resurgence'. Écalle [123] invented 'resurgence' [123] and the formalism of the 'alien calculus' and 'multisummability'. This has been extended by a group that includes Voros, Pham, Sternin, Shatalov, Delabaere, and others too numerous to list. The monograph by Sternin and Shatalov [285] and the collection of articles edited by Braaksma [83] are good summaries. (Berry, who was visiting Pham when he developed his smoothing scheme, was strongly influenced by Écalle's three-volume book and the follow-up work of the 'French school'.) The alien calculus and multisummability theory are very general but accordingly also very abstract. Berry and Howls, Olde Daalhuis and Olver, Costin, Kruskal, Hu and others have developed simplified variants of resurgence and applied them to concrete problems in special functions and physics.

As shown by the sheer length of the Table IV, which is a selected bibliography of works on resurgence and Stokes phenomenon, it is quite unfeasible to summarize this powerful theory here. (Prof. Écalle's pioneering treatise is in three volumes!) Still, one can give a little of the flavor of resurgence.

One key concept is what one might call 'saddle point democracy'. Instead of focusing in quickly on one or two dominant saddle points (on the steepest descent path), resurgence treats all saddle points on an equal footing. One may define an integral passing through an arbitrary saddle point; the coefficients of the steepest descent expansion about that point encodes the expansions about all the other saddle points. Furthermore, the late terms in the asymptotic expansion about a dominant saddle point can be expressed in terms of the early terms of a subdominant series, and vice-versa. The reason is that the late terms in the expansion about the dominant saddle point are controlled, via Darboux's Principle, by the singularities created by the other saddle points.

13. Smoothing Stokes Phenomenon: Asymptotics of the Terminant

> Having these new techniques [hyperasymptotics], I would like to hear from anybody who needs the Airy function to twenty decimals, but am not expecting an early call.
>
> Berry (1991) [30, p. 2]

Berry's amusing comment is a frank admission that the smoothing of the discontinuity along a Stokes line is not a matter of great arithmurgical significance. The term that changes dramatically in the neighborhood of the Stokes line is exponentially small compared to the sum of the asymptotic series. However, the smoothing does provide deep insights into the interlocking systems of caverns – interlocking systems of expansions about different saddle points and branch points – that lie beneath the surface of asymptotic approximations.

The numerical smoothing of the discontinuity along a Stokes lines is based on the following ideas which will be explained below:

(1) The exponentially small Stokes multiplier $\mathcal{M}$ can be *isolated* by subtracting the optimal truncation of the standard asymptotic series for $f(z)$ from it so that the multiplier is no smaller than the other terms left after the subtraction.
(2) The subdominant saddle point, the one whose Stokes multiplier is to change, lies directly *on* the steepest descent path leading down from the dominant saddle point when z or ε is on the Stokes line.
(3) When the asymptotic approximation for $f(z)$ is optimally truncated, the saddle point of the integral representation of Dingle's terminant will coincide with the subdominant saddle point and therefore with the pole of the integrand.
(4) The method of steepest descent, applied to the integrand of the *terminant*, replaces the integrand's sharp peak at its saddle point with a Gaussian function,

Table IV. Theory of Stokes phenomenon and resurgence

Description	Special functions	References
Fundamental theory		Écalle [123]
Quantum eigenproblems	Anharmonic oscillator generalized zeta funcs.	Voros [301 – 303, 305, 306]
Critical phenomena		Zinn-Justin (monograph) [328]
Erfc smoothing of Stokes phenomenon	Dawson's integral, Bi Airy function Ai Various integrals	Berry (1989a) [24] Berry (1989b) [25] Jones [155 – 157] Olver [250], McLeod [200]
Hyperasymptotics		Berry and Howls [35]
Diffraction catastrophes,		Berry and Howls [36]
Waves near Stokes lines		Berry [26]
Adiabatic quantum transitions		Berry [27]
$\Im$ (eigenvalue) exponentially small	Airy function	Wood and Paris [322, 321, 259]
2d order ODEs		Hanson [138]
Hyperasymptotics with saddles		Berry and Howls [37]
Infinitely many Stokes smoothings	Gamma function	Berry [28]
Superfactorial series		Berry [29]
Uniform hyperasymptotics with error bounds	Generalized exponential integral	Olver [251]
Uniform exponentially-improved asymptotics with error bounds	Confluent Hypergeometric functions	Olver [252]
Transcendentally small reflection	2d order ODEs	Gingold and Hu [132]
Multisummability	 Confluent hypergeometric	Martinet and Ramis [198] Olde Daalhuis [237], Olver [253]

Table IV. Theory of Stokes phenomenon and resurgence (continued)

Description	Special functions	References
Stokes phenomenon: Mellin–Barnes integral and high-order ODEs		Paris [256, 257]
Exponential asymptotics Smoothing Stokes discontinuities	Gamma function	Paris and Wood [259]
Coalescing saddles		Berry and Howls [38]
Brief (4 pg.) review		Berry and Howls [39]
Superadiabatic renormalization		Berry and Lim [42]
ODEs	Fifth-order KdV Eq.	Tovbis [293]
Steepest descent: Error bounds		W. Boyd [78]
Stokes phenomenon and hyperasymptotics		Olde Daalhuis [238]
Écalle ‘alien calculus’	REVIEW (in French)	Candelpergher *et al.* [90]
Overlapping Stokes smoothings		Berry and Howls [40]
Quantum billiards		Berry and Howls [41]
Écalle theory	REVIEW	Delabaere [112]
Weyl expansion		[148]
Reduction of theories	Philosophy of science	Berry [33]
Stokes phenomenon and Stieltjes transforms		W. Boyd [77]
Coefficients of ODEs		Olver [254]
ODEs: irregular singularities		Olde Daalhuis and Olver [244, 245, 247]
Steepest descent	Gamma function	W. Boyd [79]
Higher order ODEs		Olde Daalhuis [239, 241] Murphy and Wood [228]

Table IV. Theory of Stokes phenomenon and resurgence (continued)

Description	Special functions	References
Matched asymptotics and Stokes phenomenon		Olde Daalhuis *et al.* [243]
Stokes multipliers: Linear ODEs		Olde Daalhuis and Olver (1995b) [246]
Multisummability		Balser [12 – 17, 81]
Quantum resurgence		Voros [304]
Riemann–Siegel expansion	Zeta function	Berry [34]
ODEs		Dunster [121]
Brief reviews		Paris and Wood [260, 320]
Multidimensional integrals		Howls [147]
Steepest descent	ODEs	W. Boyd [80]
Multisummability; Gevrey separation		Ramis and Schafke [266]
Quantum eigenproblem	Quartic oscillator	Delabaere and Pham [113]
Re-expansion of remainders	Integrals	Byatt-Smith [87]

thereby reducing the asymptotics of the terminant to that of a Gaussian divided by a simple pole at the origin.

(5) If we allow the small parameter ε or the equivalent large parameter, $z = 1/\varepsilon$, to move a little way δ *off* the Stokes line, the terminant integral becomes the Fourier transform of a Gaussian divided by a pole at (or very near) the maximum of the Gaussian.

(6) The Fourier transform of a Gaussian divided by a pole is that of the integral of the Fourier transform of the Gaussian, which is the error function erf.

To illustrate these ideas, define the ‘singulant’ F via

$$F \equiv \Im\{z(\phi(t_+) - \phi(t_-))\}, \tag{76}$$

where t_+ and t_- are the two saddle points on the steepest descent contour and where the a_j are the coefficients of the inverse power series. (The real part of the difference between $z\phi(t)$ at the two points is zero along a Stokes line, and this can be used to define a Stokes line.) The singulant is proportional to some positive power of the large parameter z so that the inverse power series in z can be expressed as inverse powers of F.

The Stokes multiplier $\mathcal{M}$ may then be *defined* by

$$f(z) \sim \exp(z\phi(t_+))\left\{\sigma_+(z)\sum_{j=0}^{N_{\mathrm{opt}}(z)} a_j F^{-j+\beta-1} + i\mathcal{M}\sigma_-(z)\exp(-F)\right\}, \tag{77}$$

where N_{opt} denotes the optimal truncation of the asymptotic series for a given z, $\sigma_\pm(z)$ are slowly varying factors of z (usually proportional to a *power* of z rather than an exponential), β depends on the class of asymptotic approximation, and the coefficients have been scaled so that $a_0 = 1$ by absorbing factors into $\sigma_\pm$ if necessary. (For steepest descent as discussed here, $\beta = 1$, but other values do occur when the integral involves a contribution from an endpoint of integration interval or certain other classes of asymptotics [25].) This definition is equivalent to

$$\mathcal{M} \equiv -i\frac{\exp(F)}{\sigma_-(z)}\left\{f(z)\exp(-z\phi(t_+)) - \sigma_+(z)\sum_{j=0}^{N_{\mathrm{opt}}(F)} a_j F^{-j+\beta-1}\right\}. \tag{78}$$

Replacing $f(z)\exp(-z\phi(t_+))$ by the infinite asymptotic series and subtracting

$$\mathcal{M} \sim -i\exp(F)\frac{\sigma_+(z)}{\sigma_-(z)}\left\{\sum_{j=N_{\mathrm{opt}}(F)+1}^{\infty} a_j F^{-j+\beta}\right\}. \tag{79}$$

Note F is real and positive on the Stokes line.

The next step is to sum the series for the Stokes multiplier via Borel summation. The follow-up is crucial: instead of employing the exact power series coefficients a_j in the Borel sum, we use the asymptotic approximation to them as $j \to \infty$. This is legitimate since only *late* terms, i.e., those for $j > N_{\mathrm{opt}}(F)$, appear in the sum. This approximates the Stokes multiplier in terms of Dingle's singular terminant $\overline{\Lambda_N}(F)$.

To illustrate this general strategy, we shall return to the specific example of the Airy function, which has the asymptotic approximation

$$Ai(z) \sim z^{-1/4}\frac{1}{2\sqrt{\pi}}\{E_- + i\mathcal{M}E_+\}, \tag{80}$$

where

$$E_\pm \equiv \frac{\exp(\pm(2/3)z^{3/2})}{\Gamma(5/6)\Gamma(1/6)}\sum_{n=0}^{\infty}\frac{\Gamma(n+5/6)\Gamma(n+1/6)}{\Gamma(n+1)(\mp F)^n}. \tag{81}$$

The Stokes' multiplier $\mathcal{M}$ is zero when $\arg(z) = 0$ and is unity when $\arg(z) = \pi$.

Dingle's singulant is defined by

$$F \equiv -\frac{4}{3}z^{3/2} \tag{82}$$

which as always is the difference between the arguments of the two exponentials in the asymptotic approximation. Note the sign convention: F is negative when z and ζ are real and positive.

If the coefficients of the asymptotic series for E_-, which is the dominant exponinant exponential near the Stokes line at $\arg(z) = (2/3)\pi$, are denoted by a_j, then the argument given above implies that

$$\mathcal{M} \sim -i\exp(F)\left\{\sum_{j=N_{\text{opt}}(F)+1}^{\infty} a_j F^{-j}\right\} \tag{83}$$

$$\sim -i\exp(F)\left\{\sum_{j=N_{\text{opt}}(F)+1}^{\infty} \frac{1}{2\pi}(j-1)!F^{-j}\right\} \tag{84}$$

$$\sim -\frac{i}{2\pi}\int_0^{\infty} \exp(F(1-t))t^{N_{\text{opt}}}\frac{1}{1-t}\,\mathrm{d}t. \tag{85}$$

In the second line, we have replaced the a_j by their asymptotic approximation as $j \to \infty$ [derived through the large degree asymptotics of the gamma functions plus the identity $\Gamma(1/6)\Gamma(5/6) = 2\pi$]. The third line was derived from the second by taking the Borel sum of the series, which happens to be an integral with an integrand that can be written down explicitly. We can check that the integral is correct by expanding the integrand about $t = 0$ and then integrating term-by-term. The integral is, with a change in integration variable, proportional to Dingle's singular terminant.

The integral is not completely specified until one makes a choice about how to deal with the pole on the path of the integration. Since we know that for the Airy function, Stokes' multiplier must increase from 0 for real, positive z to 1 for real, negative z, the proper choice is to indent the path of integration *above* the pole.

The integral is also not fully determined until the optimal truncation N_{opt} has been identified. However, the coefficients asymptotic series for E_-, which is the multiplier of the exponential which is dominant near the Stokes line $\arg(z) = 2\pi/3$, are asympotically factorials, just the same as for the Stieltjes function (Equation (81)). This implies that our earlier analysis for $S(\varepsilon)$ applies here, too, to suggest

$$N_{\text{opt}} = |F|. \tag{86}$$

When F is real and positive, that is, when z is on the Stokes line, the factor

$$\chi \equiv \exp(F(1-t))t^{N_{\text{opt}}} = \exp\{F(1-t) + N_{\text{opt}}\log(t)\} \tag{87}$$

has its maximum at $t = 1$, which coincides with the singularity of the factor $1/(1-t)$ which is rest of integrand for the Stokes multiplier. This coincidence of the saddle point with the pole requires only a slight modification of standard descent to approximate $\mathcal{M}$ near the Stokes line. Wong [319, p. 356–360] gives a good discussion, attributing the original analysis to van der Waerden [296].

The key idea is to expand the factor χ as a power series about $t = 1$, rather than the saddle point, which is slightly shifted away from $t = 1$ when $\Im(F) \neq 0$. Let $T \equiv t - 1$ and $F \equiv F_{\mathrm{r}} + i F_{\mathrm{im}}$. Furthermore, since the integral is strongly peaked about $T = 0$, the lower limit of integration has been extended from $T = -1$ to $-\infty$. The Stokes multiplier is approximately

$$\mathcal{M} = -\frac{i}{2\pi} \int_{-\infty}^{\infty} \exp\left(- i F_{\mathrm{im}} T - -\frac{1}{2} F_{\mathrm{r}}^2 T^2 \right) \frac{1}{T} \,\mathrm{d}T, \tag{88}$$

where terms of $\mathrm{O}(T^3)$ in the exponential have been neglected.

This approximation is just the Fourier transform of a Gaussian function $(\exp(-(1/2) F_{\mathrm{r}}^2 T)$, divided by iT. The identity

$$\lim_{\delta \to 0} \frac{-i}{2\pi} \int_{-\infty}^{\infty} \exp(-ikx) \exp(-a^2 k^2) \frac{1}{(k + i\delta)} \,\mathrm{d}k = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{x}{2a}\right) \tag{89}$$

shows that the Stokes multiplier is

$$\mathcal{M} = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{F_{\mathrm{im}}}{\sqrt{2F_{\mathrm{r}}}}\right). \tag{90}$$

Figure 11 shows that this approximation is very accurate.

The error function does not cover all cases; Chapman [95] has shown that other smoothing functions are needed in some circumstances. However, the complementary error function does remove the ‘Victorian discontinuity’ of Stokes for a remarkably wide class of functions.

14. Matched Asymptotic Expansions in the Complex Plane: The PKKS Method

In ‘above-the-barrier’ quantum scattering, there are no turning points where the coefficient of the undifferentiated term in the Schrödinger equation is zero except at complex values of the spatial coordinate. When there are real-valued turning points, it was discovered in the 1920s that the scattering – including the exponentially small transmission through the barrier – can be computed by means of the so-called turning point connection formulas. (The transmission coefficient can be calculated without heroics because the exponentially small transmitted wave is the *whole* solution on the far side of the barrier, isolating it from terms proportional to powers of ε as noted earlier.) Later, it was shown that the connection formulas are really just a special case of the method of matched asymptotic expansions [229, 20, 21]. The solution in the neighborhood of the turning point can be expressed (to lowest order) in terms of the Airy function Ai. This is matched to standard WKB approximations which describe the solution everywhere else.

For ‘above-the-barrier’ scattering, however, what is one to do? Pokrovskii and Khalatnikov [262] had a flash of insight: actually, there *are* turning points, but only

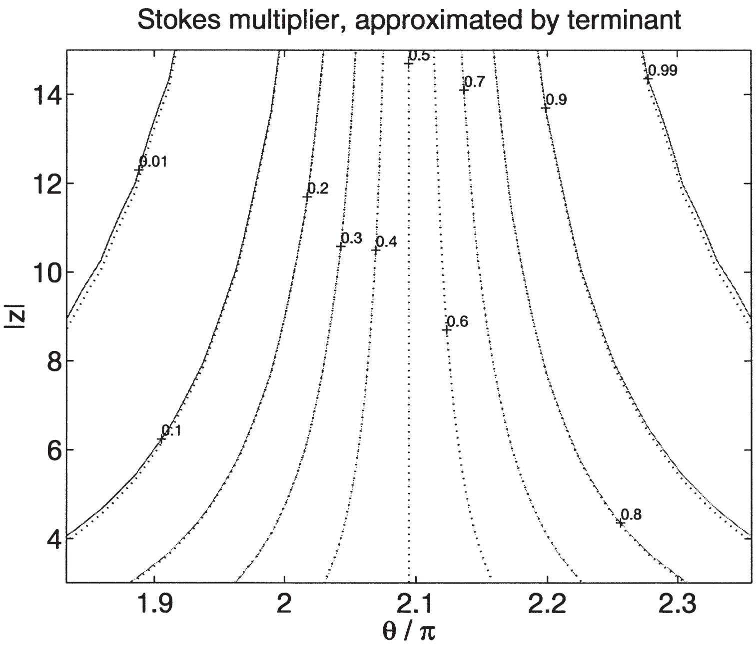

Figure 11. Solid: contours of the integral approximating the Stokes multiplier for the Airy function. Dashed: contours of the error function approximation to this integral. The solid and dashed contours are almost indistinguishable, which is a graphical demonstration that the steepest descent approximation to the integral is very accurate.

for *complex* x. In the vicinity of these off-the-real-axis turning points, the reflected wave is *not* small, so the usual connection formulas apply with only minor modifications. The amplitude of the reflected wave decays exponentially as $\Im(x) \to 0$ so that, on the real x-axis, the reflection coefficient is exponentially small in $1/\varepsilon$, the inverse width of the barrier.*

Kruskal and Segur [171, 172, 278] showed that matching expansions at off-the-real-axis critical points was a powerful method for nonlinear problems, too. Their first application resolved a forty year old conundrum in the formation of multi-branched fingers ('dendrites') on a solid-liquid interface. The unique length scale observed in the laboratory is imposed by surface tension. However, the scale-selecting effect lies 'beyond all orders' in a power series expansion in the surface tension parameter. Their method, which we shall henceforth call the 'PKKS' [Pokrovskii–Khalatnikov–Kruskal–Segur] scheme for short, has been widely used for weakly nonlocal solitary waves (Table III).

To illustrate the PKKS method, we shall apply it to the linear problem:

$$u_{xx} + u = f(\varepsilon x), \tag{91}$$

* Pokrovskii relates an amusing story: when he presented his work to the Nobel laureate, Lev Landau, the great man thought he and Khalatnikov were crazy! He eventually changed his mind [261].

where $f(x)$ will be restricted to functions that (i) decay exponentially as $|x| \to \infty$ on and near the real axis and (ii) have a complex conjugate pair of double poles at $x = \pm i$ as the singularities nearest the real axis and (iii) are symmetric with respect to $x = 0$, that is, $f(x) = f(-x)$. This seems like a rather special and restrictive problem. However, as Dingle observed long ago, every simple example is a master-key to an entire class of problems, as we shall show. This linear problem is identical to that solved earlier, Equation (26), except for the sign of the undifferentiated term in u.

We shall impose the boundary condition that

$$u \sim \alpha \sin(|x|) \quad \text{as } |x| \to \infty \tag{92}$$

for some constant α which will be determined as part of the solution. This excludes the homogeneous solutions $\sin(x)$ and $\cos(x)$ so as to yield a unique solution. (Note the absolute value bars inside the argument of the sine function in the boundary condition.)

The PKKS method has the following steps:

(1) Identify the singularities or critical points which are nearest the real x-axis.
(2) Define an ‘inner’ problem, that is, a perturbative scheme which is valid in the neighborhood of one of these critical points, using a complex coordinate y whose origin is at the critical point.
(3) Asymptotically solve the ‘inner’ problem as $|y| \to \infty$, that is, compute the ‘outer limit of the inner solution’.
(4) Sum the divergent outer limit of the inner problem by Borel summation or otherwise determine the connection formula, that is, the magnitude and phase of the discontinuity along the Stokes line radiating from the critical point to the real x-axis.
(5) Match the outer limit of the inner solution to the inner limit of the outer expansion.
(6) Continue the matched outer expansion back to the real x-axis to compute the (exponentially small) magnitude of the Stokes jump for real x.

The domains of the ‘inner’ and ’outer’ regions are illustrated in Figure 12.

Step one has already been accomplished by the specification of the problem: the relevant critical points are the double poles of $f(\varepsilon x)$ at $x = \pm i/\varepsilon$ where the change of variable from x to εx has reduced the residues to $1/\varepsilon^2$. The shifted coordinate (for matching in the upper half-plane) is

$$y \equiv x - i/\varepsilon. \tag{93}$$

Step two pivots on the observation that in the vicinity of its double pole, it is a legitimate approximation to replace $f(\varepsilon x)$ by the singular term only, even though this is a poor approximation everywhere except near the pole. The inner problem is then

$$U_{yy} + U = 1/y^2; \quad U \equiv \varepsilon^2 u. \tag{94}$$

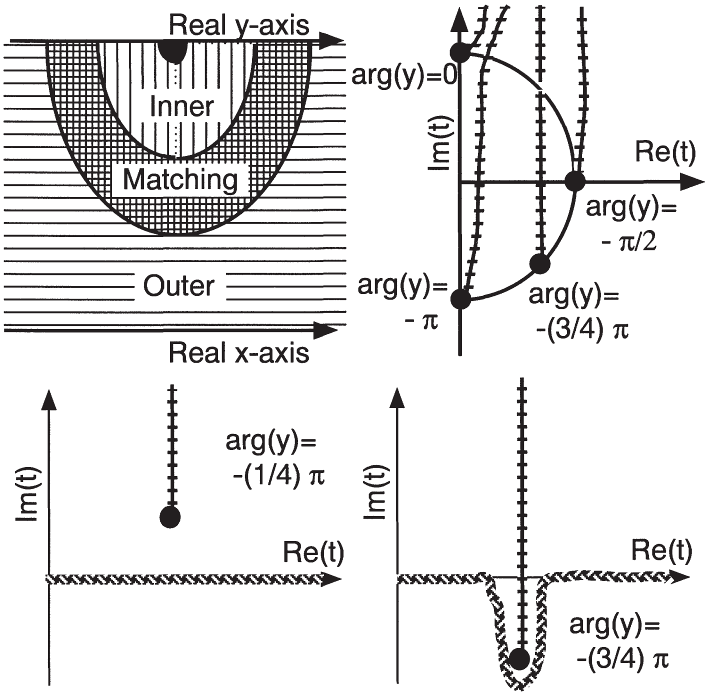

Figure 12. (a) [Upper left corner] Schematic of complex y-plane where y is the shifted coordinate. (The real axis in the original coordinate x is the arrow at the bottom.) The location of the double pole (at $y = 0$) is the large solid dot at top. The ‘matching’ region, shaped like a half annulus, is where both the inner and outer solutions are valid, allowing them to be matched. (b) [Upper right corner] The complex t-plane where t is the integration variable for the Borel-logarithm function, $\mathrm{Bo}(y)$. The four large black discs show the location of the logarithmic singularity of the integrand for four different values of $\arg(y)$. The branch cut (cross-hatched lines) goes to $i\infty$ for all cases. As $\arg(y)$ increases, the location of the branch cut rotates clockwise. For $\arg(y) < -\pi/2$, the branch cut crosses the real t-axis as shown in the lower right half diagram. (c) [Bottom half of the figure]. Both left and right panels illustrate the path of integration in the complex t-plane (heavy, patterned curves) and the branch cuts for the logarithm of the integrand (cross-hatched lines). The left diagram shows the situation when $\arg(y) = -\pi/4$, or any other point such that the branch point is in the upper half of the t-plane: the branch cut does not cross the real axis. When $\arg(y) < -\pi/2$ [right, bottom diagram], the integration path must be deformed below the real t-axis to avoid crossing the branch cut. The integration around the branch cut adds an additional contribution.

Step three, computing an outer expansion for the inner problem, is obtained by an inverse power series in y:

$$U(y) \sim \sum_{j=1}^{\infty} \frac{(-1)^{j+1}(2j-1)!}{y^{2j}}. \tag{95}$$

For the inner problem to be sensible, $|y| \ll 1/\varepsilon$. For the inverse power series to be an accurate approximation to the inner solution, we must have $|y| \gg 1$. It follows that the inverse power series is a good approximation only in the annulus

$$1 \ll |y| \ll 1/\varepsilon. \tag{96}$$

It is fair to dub this annulus the 'matching region' because it will turn out that the inner limit of the outer expansion will also be legitimate in this annulus. However, 'annulus' is a slightly misleading label because Equation (96) ignores Stokes phenomenon, which will limit the validity of Equation (95) to a sector of the annulus.

To sort out Stokes phenomenon, it is helpful to sum the divergent series by Borel summation. For this simple case, the Borel transform can be written in closed form to give, without approximation,

$$U(y) = (1/2)\mathrm{Bo}(y); \qquad \mathrm{Bo}(y) \equiv \int_0^{\infty} \exp(-t)\log\{1 + t^2/y^2\}\,\mathrm{d}t. \tag{97}$$

The integrand is logarithmically singular at $t = \pm iy$. As the argument of y varies from 0 to $-\pi$, that is, through a semicircle in the lower half of the y-plane, the singularity initially in the upper half of the t-plane rotates clockwise through a semicircle in the right half of the t-plane to exchange places with the other branch point. As $\arg(y)$ passes through $-\pi/2$, that is, through the negative imaginary y-axis, the branch points of the 'Borel-logarithm' function $\mathrm{Bo}(y)$ are forced to cross the real t-axis. To avoid discontinuously redefining the branch points of the logarithm in the integrand, the path of integration must be deformed to pass below the real t-axis (in the right half t-plane). This gives an extra contribution which is the Stokes jump for this function with the negative imaginary y-axis as the Stokes line. One finds

$$\mathrm{Bo}(y) - \mathrm{Bo}(-y) = 2\pi i \exp(-iy). \tag{98}$$

The positive and negative real y-axis are the anti-Stokes lines for $\mathrm{Bo}(y)$.

The outer expansion is the same as the multiple scales series for Equation (26) except for alternating signs. The final result on the real x-axis is

$$u(x) \sim \begin{cases} -\dfrac{\pi}{\varepsilon^2}\exp\left(-\dfrac{x_s}{\varepsilon}\right) + \displaystyle\sum_{j=0}^{\infty}(-1)^j \varepsilon^{2j}\frac{\mathrm{d}^{2j} f}{\mathrm{d}X^{2j}}, & x \geqslant 0, \\ \dfrac{\pi}{\varepsilon^2}\exp\left(-\dfrac{x_s}{\varepsilon}\right) + \displaystyle\sum_{j=0}^{\infty}(-1)^j \varepsilon^{2j}\frac{\mathrm{d}^{2j} f}{\mathrm{d}X^{2j}}, & x < 0, \end{cases} \tag{99}$$

where the outer expansion has been written in terms of derivatives of $f(\varepsilon x) \equiv f(X)$ with respect to the 'slow' variable $X \equiv \varepsilon x$ to explicitly, rather than implicitly, display the dependence of the jth term on ε^{2j}.

Extrapolating back to the real axis reduces the magnitude of the jump by $\exp(-x_s/\varepsilon) = \exp(-1/\varepsilon)$ where x_s/ε is simply the distance of the singularity from the real x-axis. Note that the exponential dependence on ε is controlled *entirely* by x_s/ε; the strength of the residue and the type of singularity (simple pole, double pole or logarithm) only alters factors that vary as powers of ε or slower.

We chose this particular example because the theory of Pomeau, Ramani and Grammaticos [263] for the Fifth-Order Korteweg–deVries equation, later extended to higher order by Grimshaw and Joshi [134], is very similar. In particular, the dominant singularities – in their case, of the lowest order approximation to the solitary wave – are also double poles on the imaginary axis. It is also true that the outer limit of the inner solution is the Borel-logarithm function, $\mathrm{Bo}(y)$ [to lowest order]. Consequently, the lowest order theory for this *nonlinear eigenvalue* problem is almost identical to that for this *linear, inhomogeneous* problem. The major difference is that the nonlinearity multiplies $\mathrm{Bo}(y)$ by a constant which can only be determined numerically by extrapolating the recurrence relation. The early terms of the series in inverse powers of y in the matching region is strongly affected by the nonlinearity, but the coefficients *asymptote* to those of $\mathrm{Bo}(y)$, another triumph of Dingle's maxim: Always look at the late terms where a whole class of problems asymptote to the same, common form.

As noted by a reviewer, the integral for Bo can be integrated by parts to express it as the sum of two Dingle terminants, and the connection formulae can then be evaluated through residues. This alternative derivation of the same answer emphasizes the remarkable universality of hyperasymptotics; again and again, one keeps falling over the same small set of terminants.

Table III records many successes for the PKKS method, but it is a curious success. It is a *general* truth that the *exponential* dependence on ε is controlled entirely by x_s, the distance from the relevant singularities or critical points to the real axis. This is usually almost trivial to determine. Roughly 90% of the work of the PKKS method is in determining the 'prefactor', that is, the product of a constant times *algebraic* factors of ε, such as logarithms and powers, which multiplies the exponential. Not only is the determination of the prefactor (comparatively) arduous, but the final step of determining the overall multiplicative constant must always be done numerically. Pomeau, Ramani and Grammaticos and later workers such as Akylas and Yang [5] and Boyd [68] have simplified the numerical bit to extrapolating a sequence derived from a recurrence, a task much easier than directly solving a differential equation. However, the fact that the PKKS method is an analytical method that is not entirely analytic gives much ground to alternatives such as spectral methods which are discussed later.

15. Snares and Worries: Remote but Dominant Saddle Points, Ghosts, Interval-Extension and Sensitivity

There are, so to speak, in the mathematical country, precipices and pit-shafts down which it would be possible to fall, but that need not deter us from walking about.

Lewis F. Richardson (1925)

More subtle perils in deriving even the lowest order correctly also lurk. Balian, Parisi and Voros [11] describe an integral where the convergence is controlled by a saddle point at $t = 2$, but the error is dominated by the exponentially larger contributions of another saddle point at $t = 3$. Their function is

$$I(z) \equiv \int_{-\infty}^{\infty} \exp\left(-z\left\{36t^2 - 20t^3 + 3t^4\right\}\right) \mathrm{d}t. \tag{100}$$

For large z, the integrand is steeply peaked about the dominant saddle point at $t = 0$ (Figure 13). The contributions of the other two saddle points will be proportional to the integrand evaluated at these saddle points:

$$\exp(-z\phi(t = 2)) = \exp(-32z), \qquad \exp(-z\phi(t = 3)) = \exp(-27z). \tag{101}$$

Because the saddle point at $t = 2$ controls convergence, the smallest term in the asymptotic series for a given z will be $O(\exp(-32z))$, so we would expect this to be the magnitude of the error in the optimally-truncated series in inverse powers of z. In reality, the superasymptotic error is dominated by the contribution of the saddle point at $t = 3$, which is $O(\exp(-27z))$ and therefore larger than the smallest term in the optimally-truncated series by $O(\exp(5z))$.

One of the charms of resurgence theory is that during the early stages, all saddle points are treated equally. This 'saddle point democracy' is valuable in detecting such pathologies, and correctly retaining the contributions of all the important saddle points. Still, if the asymptotic series is derived not from an integral but directly from a differential equation so that no information is available but the coefficients of the series, it would be easy to be fooled, and assume that the magnitude of the smallest retained term was a genuine estimate of the superasymptotic error.

Fortunately, it appears that this is rare in practice. The applied mathematical landscape is littered with deep sinkholes which fortunately have an area of measure zero. The Balian–Parisi–Voros example was contrived by its authors rather than derived from a real application. However, related difficulties are not contrived.

For the so-called ϕ^4 breather problem [278, 58], the convergence of the divergent series is controlled by the constant in the Fourier series with an expected minimum error of $O(\exp(-\sqrt{2}\pi/(2\varepsilon)))$. However, the far field radiation has a magnitude α which has been shown to be $O(\exp(-\sqrt{6}\pi/(2\varepsilon)))$. Thus, after the ε-power series has been truncated at optimal order and subtracted from the solution, the correction is still exponentially large in ε relative to the weakly nonlocal radiation. Complex-plane matched asymptotics is not inconvenienced [278], but the hyperasymptotic method of Boyd [67] would likely fail.

Another danger is illustrated by the function

$$f(\varepsilon) \equiv S(\varepsilon) + \exp(-(1/2)/\varepsilon), \tag{102}$$

where $S(\varepsilon)$ is the Stieltjes function. The asymptotic expansion for this function is the *same* as for the Stieljes function; because $\exp(-(1/2)/\varepsilon)$ and all its derivatives vanish as $\varepsilon \to \infty$, this function makes no contribution to the divergent power

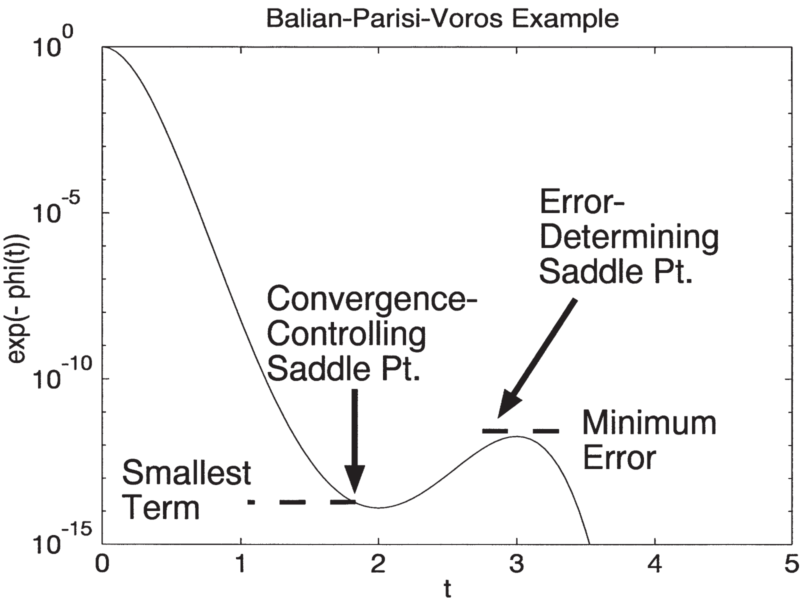

Figure 13. The integrand of the example of Balian, Parisi and Voros: $\exp(-\phi)$. The dominant saddle point is at $t = 0$. The secondary peak (saddle point) at $t = 2$ controls the asymptotic form of the coefficients of the asymptotic series; because of it, the series for dt/dw converges only for $|w| \leqslant |w(t = 2)|$. However, the contribution of the more distant saddle point at $t = 3$ dominates the error. The terms of the power series in $1/z$ reach a minimum at roughly $\exp(-z\phi(t = 2))$, but the error in the optimally-truncated series is exponentially large compared to this minimum term, being roughly $\exp(-z\phi(t = 3))$.

series of $f(\varepsilon)$. It follows that if we manipulate the power series in the usual way, we arrive at a superasymptotic approximation which, from the size of the smallest term, has an error *apparently* of $O(\exp(-1/\varepsilon)$. Adding a Dingle terminant gives a hyperasymptotic approximation of even smaller error – more fool we! Because we are approximating $f(\varepsilon)$ (at best!) by the Stieltjes function $S(\varepsilon)$, the error is actually the magnitude of the second term – exponentially larger than $\exp(-1/\varepsilon)$. Quick to defend the honor of hyperasymptotics, a reviewer argued that this is merely a problem of definition. An engineer's answer: a wrong answer is never just a matter of definition, but rather a good reason to lie awake at night, and retain a lawyer.

Weakly nonlocal solitary waves are a nontrivial example of phenomena with exponentially small 'ghosts': the solitons can be expanded in nontrivial power series in ε, but the amplitude α of the sinusoidal oscillations of the soliton for large $|x|$ is proportional to $\exp(-\mu/\varepsilon)$ for some constant μ. The terms in the power series are, in the simplest cases, powers of $\text{sech}(\varepsilon x)$ and therefore each term individually decays exponentially fast as $|x| \to \infty$. It follows that standard acceleration methods must fail because reweighting the terms of the power series still gives nothing at infinity, and thus misses the far field oscillations completely. The Dingle terminants method, which is based on the asymptotics of the power series coefficients, has never been successfully applied to this sort of problem either.

Fortunately, the PKKS method, spectral algorithms, the spectral space asymptotics of Akylas and Yang [5, 325, 326] and the hyperasymptotic scheme of Boyd [68] all work well for nonlocal solitary waves. Nevertheless, the failure of some alternative schemes for this class of problems because the quantity of interest is invisible to the power series are vivid reminders of the truism: Fear and caution are healthy character traits in an applied mathematician!

Another pitfall is extending the interval of integration to infinity. For example, the Bessel function $I_0(z)$ has a representation that is an integral over a *finite* interval:

$$I_0(z) \equiv \frac{1}{2\pi} \exp(z) \int_{-\pi}^{\pi} \exp\big(z[\cos(t) - 1]\big)\, \mathrm{d}t. \tag{103}$$

Without approximation, we can make the change of variable $2w^2 = 1 - \cos(t)$ to obtain

$$I_0(z) = \frac{1}{\pi} \exp(z) \int_{-1}^{1} \exp(-2zw^2) \frac{1}{(1 - w^2)^{1/2}}\, \mathrm{d}w. \tag{104}$$

The usual procedure is to expand the $1/\sqrt{1 - w^2}$ as a power series, extend the interval of integration to $w \in [-\infty, \infty]$, and integrate term-by-term to obtain the expansion given in most references:

$$I_0(z) \sim \exp(z) \left(\frac{1}{2\pi z}\right)^{1/2} \left\{1 + \frac{1}{8} z^{-1} + \frac{9}{128} z^{-2} + \cdots \right\}. \tag{105}$$

The sole reason for the divergence of this series is the extension of the interval. If we expand and integrate term-by-term on the original interval $w \in [-1, 1]$, the result is a series that *converges* – albeit rather slowly because the radius of convergence of the power series under the integrand is just equal to the limit of integration so that the terms of the integrated series decrease only as $O(1/j^{3/2})$ for the coefficient of $1/z^j$ in the series. Why, then, is interval-extension so ubiquitous in asymptotics?

The answer is two-fold. First, the terms of the series are greatly simplified at the price of divergence. The one-term approximation is simplified from an error function to a Gaussian, for example, and the higher order terms of the convergent series become more and more complicated; to the same order as above, the convergent series is

$$\begin{aligned} I_0(z) \;=\; & \frac{1}{\pi} \exp(z) \left\{ \sqrt{\frac{\pi}{2\,z}}\, \mathrm{erf}(\sqrt{2z}) + \right. \\ & + \sqrt{\frac{\pi}{2}} \frac{\mathrm{erf}(\sqrt{2z})}{8 z^{3/2}} - \frac{1}{4} \frac{\exp(-2z)}{z} + \end{aligned}$$

$$+\frac{3}{256z^{5/2}}\Big(-12\sqrt{z}\exp(-2z)-16z^{3/2}\exp(-2z)+ \\ +3\sqrt{2\pi}\,\mathrm{erf}(\sqrt{2z})\Big)\Bigg\}. \tag{106}$$

The second reason is that because the convergent series is only slowly convergent, it is far from obvious that the error can be reduced much below the superasymptotic limit unless one uses a very large number of terms in the convergent series. It is more practical to restrict z to such large values that the superasymptotic approximation is acceptably accurate, and use the ordinary Taylor series for smaller z.

Another snare is that exponentially small quantities, when paired with a nontrivial powers, are often extremely sensitive to small changes in parameters. For example, the solution of the differential equation equation

$$u_{xx}+u=\mathrm{sech}(\varepsilon x)+\varepsilon^{2n-1}d(\varepsilon)(2n-1)!\,\mathrm{sech}^{2n}(\varepsilon x) \tag{107}$$

asymptotes to a sinusoidal oscillation with an amplitude $\alpha(\varepsilon)$, which is an exponential function of $1/\varepsilon$. One can choose an O(1) function $d(\varepsilon)$ such that $\alpha(\varepsilon)$ is zero. And yet if $\varepsilon = 1/10$ and $n = 3$, the second term in the forcing is more than a thousand times smaller than the first!

The moral is that in physical applications, it is not sufficient merely to calculate the exponentially small effects. One must also look at how small perturbations of the idealized problem might drastically change the exponentially small contributions.

16. Asymptotics as Hyperasymptotics for Chebyshev, Fourier and Other Spectral Methods

One important but neglected area of asymptotics is numerical analysis, specifically, approximations to the error as a function of the grid spacing h (or other discretization parameters). For the familiar numerical integration scheme known as the trapezoidal rule, for example, which is defined by

$$I\equiv\int_0^1 g(y)\,\mathrm{d}y\sim T\equiv\frac{1}{M}\left\{\frac{1}{2}g(0)+\frac{1}{2}g(1)+\sum_{j=1}^{M-1}g\left(\frac{j}{M}\right)\right\}, \tag{108}$$

where the grid spacing h is

$$h\equiv 1/M, \tag{109}$$

the Euler–Maclaurin sum formula gives the following asymptotic series for the error [143]

$$I-T\sim-\sum_{j=1}^{N}h^{2j}\frac{B_{2j}}{(2j)!}\left\{g^{(2j-1)}(1)-g^{(2j-1)}(0)\right\}+E_N, \tag{110}$$

where the B_{2j} are the Bernoulli numbers and the error is

$$E_N \equiv -h^{2N+2}\frac{B_{2N+2}}{(2N+2)!}g^{(2N+2)}(\xi) \tag{111}$$

with ξ a point somewhere on the interval [0, 1].

Elementary numerical analysis courses usually stop with the observation that since the leading term is proportional to h^2, the trapezoidal rule is a 'second order' method. However, the series in powers of h contains hidden depths. First, if the rule is applied twice with different grid spacings h and $2h$, one can take a weighted difference of $T(h)$ and $T(2h)$ such that the quadratic terms cancel in (110):

$$\frac{4T(h) - T(2h)}{3} \sim I + \mathrm{O}(h^4). \tag{112}$$

This improved integration scheme is 'Simpson's Rule'. This strategy can be iterated to obtain methods of progressively higher accuracy and complexity not only for quadrature schemes, but for finite difference approximations, too. The generalization to arbitrary order is known as 'Richardson extrapolation' [270, 271] or to use his own term, 'deferred approach to the limit'.

Second, if the integrand $g(x)$ is a *periodic* function, then $g^{2j-1}(1) = g^{2j-1}(0)$ to all orders j, and the error expansion reduces to the trivial one:

$$I \sim T + 0h^2 + 0h^4 + 0h^6 + \cdots. \tag{113}$$

Within the Poincaré asymptotic framework, this suggests the trapezoidal rule is exact for all periodic functions – Wrong!

The integral of $g(x)$ is proportional to the constant in the Fourier expansion of $g(x)$; the usual formula for the error in computing Fourier coefficients through the trapezoidal rule [55] gives

$$I - T = -\sum_{j=1}^{\infty} a_{jN}, \tag{114}$$

where the $\{a_j\}$ are the exact Fourier cosine coefficients of $g(x)$, i.e.,

$$a_j \equiv 2\int_0^1 g(x)\cos\big(2\pi j[x - 1/2]\big)\,\mathrm{d}x. \tag{115}$$

(Note that in (114), the degree of the Fourier coefficient is the product of the sum variable j with N so that only every Nth coefficient appears in the error series.) For a *periodic* function, it is known that

$$a_j \sim \nu(j)\exp(-2\pi\mu j), \tag{116}$$

where μ is the absolute value of that singularity in the complex x-plane which is closest to the real axis and where the prefactor ν is an *algebraic* rather than exponential function of j which depends on the type of singularity (simple pole, logarithm, etc.). Inserting this into Equation (114) gives the correct conclusion that for periodic functions, free of singularity for real x, the error in the trapezoidal rule is

$$I - T \sim \nu \exp(-2\pi\mu/h). \tag{117}$$

In other words, the error lies beyond all orders in powers of the grid spacing h.

Presumably such exponential dependence on $1/h$ lurks in the Euler–Maclaurin series even for nonperiodic functions. We can confirm this suspicion by considering a particular case: a function $g(x)$ whose singularity nearest the interval $x \in [0, 1]$ is a simple pole of unit residue at $x = -\sigma$. In the limit that the order $j \to \infty$, the $(2j - 1)$st derivative will be more and more dominated by this singularity. (Recall that the convergence of the power series in x about the origin has a radius of convergence σ controlled by this nearest singularity, and that the coefficients of the power series are the derivatives of $g(x)$ at the origin.) Using the known asymptotics of the Bernoulli numbers, $B_{2j} \sim (-1)^{j-1}(2j)!/(2^{2j-1}\pi^{2j})$, we find

$$h^{2j}\frac{B_{2j}}{(2j)!}\left\{g^{(2j-1)}(1) - g^{(2j-1)}(0)\right\} \sim 2(-1)^j \frac{h^{2j}}{2^{2j}\pi^{2j}\sigma^{2j}}\, j!, \quad j \to \infty. \tag{118}$$

This implies one can obtain an improved trapezoidal rule by subtracting the leading term of the hyperasymptotic approximation to the error in the ordinary trapezoidal rule. This is obtained by optimally-truncating the Euler–Maclaurin series and approximating the error by a Dingle terminant:

$$E_N \sim \Lambda_N\left(\frac{h^2}{4\pi^2\sigma^2}\right). \tag{119}$$

One important implication of the factorial divergence of the Euler–Maclaurin series is that it shows that Richardson extrapolation will diverge, too, if applied to arbitrarily high order for *fixed* h. Romberg integration, which is a popular and robust algorithm for numerical integration, does in fact employ Richardson extrapolation of arbitrary order. However, at each stage, the grid spacing h is halved. In the limit that $1/h$ and the order of Richardson extrapolation *simultaneously* tend to infinity, the quadrature scheme converges.

Lyness and Ninham [190] noted this exponential dependence on $1/h$ in quadrature errors nearly thirty years ago. Lyness has emphasized that the Euler–Maclaurin series for the trapezoidal rule is closely related to a general formula for the coefficients of a Fourier series. His ‘FCAE’ [Fourier Coefficient Asymptotic Expansion] takes the form [188, 189]

$$\begin{aligned} a_n = {} & \sum_{k=1}^{M}(-1)^{n+k-1}\frac{g^{2k-1}(1) - g^{2k-1}(0)}{(2\pi n)^{2k}} + \\ & + \frac{(-1)^{M+1}}{(2\pi n)^{2M+1}}\int_0^1 g^{(2M+1)}(x)\, \sin(2\pi n[x - 1/2])\, \mathrm{d}x \end{aligned} \tag{120}$$

plus a similar series for the sine coefficients. It is derived by integration-by-parts. As noted by Lyness, it is usually divergent.

As explained in the books by Boyd [55] and Canuto *et al.* [91], spectral methods usually employ a basis set so that the error is exponentially small in $1/h$ or equivalently, in the number of degrees of freedom M. Fourier series, for example, are restricted to periodic functions. Chebyshev polynomials give exponential convergence even for nonperiodic functions, provided only that $f(x)$ is free of singularities on $x \in [-1, 1]$. These polynomials are defined by $T_n(\cos[\theta]) \equiv \cos(n\theta)$ so that the expansion of $f(x)$ as a Chebyshev series is identical, under this change of variable, with the Fourier expansion of $f(\cos(\theta))$. The transformed function is *always* periodic in θ, so the error in truncating a Chebyshev series after M terms decreases exponentially fast with M. For Chebyshev and Fourier spectral methods, the power series in h is always the trivial one with zero coefficients.

It follows that all asymptotic approximations to Fourier, Chebyshev, and other spectral coefficients for large M are *implicitly* hyperasymptotic (Table V). One might object that this catalogue of asymptotics for orthogonal series is out of place here because it is 'beyond all orders' [in h] only because the coefficients of all powers of h are zero. The main tools for the work of Elliott, Miller, Luke, Weideman and Boyd were steepest descents and the calculus of residues – no explicit use of hyperasymptotic thinking at all.

Nevertheless, Table V makes several important points. First, much of hyperasympotics *is* steepest descent and the calculus of residues. In discussing the Stieltjes function earlier, for example, we noted that one could go beyond the superasympotic approximation by applying steepest descent to the error integral of the optimally-truncated ε power series. Similarly, the heart of the PKKS technique of matched asymptotic expansions in the complex plane is the notion that the singularities or other critical points closest to the real axis control the hyperasymptotic behavior. The asymptotic behavior of the coefficients of orthogonal expansions is likewise controlled by complex singularities. In both cases, the constant q inside the exponential factor, $\exp(-q/\varepsilon)$ or $\exp(-q/h)$, is simply the distance from the dominant singularity to the real x-axis.⋆

Second, an important hyperasymptotic strategy is to isolate the exponentially small contributions. For the large degree behavior of Chebyshev and Fourier coefficients, this isolation is a free gift, the result of choosing the sensible spectral basis – Chebyshev or Legendre polynomials for nonperiodic problems, spherical harmonics for problems on the surface of a sphere and so on. For less trivial problems, the key to isolation is to subtract an *optimally truncated* asymptotic expansion. This is the justification for applying steepest descent to the error integral for Stieltjes function, for applying Euler's method or other sum acceleration scheme, for Dingle's universal terminants and for Berry's smoothing of Stokes

⋆ For Chebyshev series, 'distance from the real axis' means distance in the transformed coordinate θ, the argument of the equivalent Fourier series, rather than the argument of the Chebyshev polynomials, $x = \cos(\theta)$.

Table V. Asymptotics of Fourier, Chebyshev, Hermite and other spectral methods.

Functions	Comments	References
Entire functions, Meromorphic functions, Branch points on $[-1, 1]$		Elliott [124, 125]
Entire functions, $f(x)$ when Laplace transform known	Uniform as well as large n asymptotics	Elliott and Szekeres [126]
Whittaker Exponential integral, Error integral, Confluent hypergeometric		Miller [227] [230, 231] Németh [232 – 235] Németh (1992) [236]
Whittaker	Asymptotics for Meier G-function [exact Chebyshev functions]	Wimp [317]
Many (monograph)	Many (complicated) exact coefficients and asymptotics	Luke [187]
		Tuan and Elliott [295]
$\exp(-A/x)$ [Laguerre] Contour integrals for arbitrary $f(x)$	Misleading title; 'Fourier' coefficients are Jacobi, Laguerre and Hermite coeffs.	Elliott and Tuan [127]
Stieltjes functions	Upper bound on β	Boyd [49]
Stieltjes functions	Lower bound: $\beta \geqslant 1 - r/2$	Boyd [49]
General	Hermite functions	Boyd [48, 52]
General	Rational Chebyshev	Boyd [50, 53, 54]
General	Fourier and Chebyshev error envelopes	Boyd (1990c) [59]
Entire functions	Chebyshev	Ciasullo and Cochran [99]
General	Mapped Fourier for infinite interval	Cloot and Weideman [102], Weideman and Cloot [312]
Entire functions: Exp (Gaussian)	Fourier coeffs.	Boyd [64]
Error function	Rational Chebyshev series	Weideman [309 – 311]

Note: β is the spectral 'exponential index of convergence' such that $a_n \sim \exp(-\text{constant}\, n^{\beta})$. r is defined by $r = \limsup_{n\to\infty} \log|b_n|/(n \log n)$

phenomenon. The difference between a quantity and its superasymptotic approximation is as trivially isolated as the spectral coefficients of a Fourier series, for which the superasymptotic approximation is zero.

Third, the asymptotics of spectral series and other numerical processes is a relatively under-cultivated area. Can the ideas reviewed here lead to optimal order Richardson extrapolation for numerical quadrature, various classes of differential equations, and so on?

Fourth, there have been some limited but important excursions beyond conventional asymptotics in the analysis of the convergence of spectral series. For example, Boyd's 1982 article on the optimization of Chebyshev methods on an unbounded domain notes that there are two different species of contributions to the asymptotic spectral coefficients: (i) saddle point contributions that depend on how rapidly the function $f(x)$ being expanded decays as $|x| \to \infty$ and (ii) contributions from the poles and other singularities of $f(x)$. In the limit that the degree n of the rational Chebyshev coefficient tends to infinity for fixed value of the 'map parameter' L, one type of contribution will be exponentially large compared to the other, and it is inconsistent (in the Poincaré sense of asymptotics) to retain the other. Boyd points out that to optimize numerical efficiency, one should allow L to vary with the truncation N of the Chebyshev series. Convergence is optimized when N and L *simultaneously* tend to infinity in a certain way so that both the pole and saddle point contributions are of equal order. This sort of analysis does not explicitly use exponentially-improved asymptotics of the Dingle–Berry–Olver sort. Nevertheless, hyperasymptotic thinking – considering the role of terms that at first glance are exponentially small compared to the dominant terms – is absolutely essential to this kind of numerical optimization.

A few other interesting studies of the role of exponential smallness in numerical analysis have already been made. For example, the usual second order differential equation for the nonlinear pendulum, $q_{tt} + \sin(q) = 0$, can be written as the equivalent system

$$p_t = \sin(q), \qquad q_t = p, \tag{121}$$

where q is the angle of deflection of the pendulum with $q = 0$ when the pendulum is standing (unstably) on its head and p is the momentum. Hakim and Mallick [136] show that when the first equation is discretized by a forward difference and the second equation by a backward difference, the system (121) becomes what in dynamical systems theory, with a slight change in notation, is called the 'standard mapping':

$$p_{n+1} = p_n + \tau \sin(q_n), \qquad q_{n+1} = q_n + \tau\, p_{n+1}. \tag{122}$$

The usual numerical analysis description begins and ends with the statement that this algorithm is second order accurate, that is, has an error which is proportional to τ^2. ⋆ Hakim and Mallick point out that there are also changes which are exponen-

⋆ The one-sided differences are only first order, but by eliminating p, one can show that the system is equivalent to applying centered, 2d order differences to the second order differential equation.

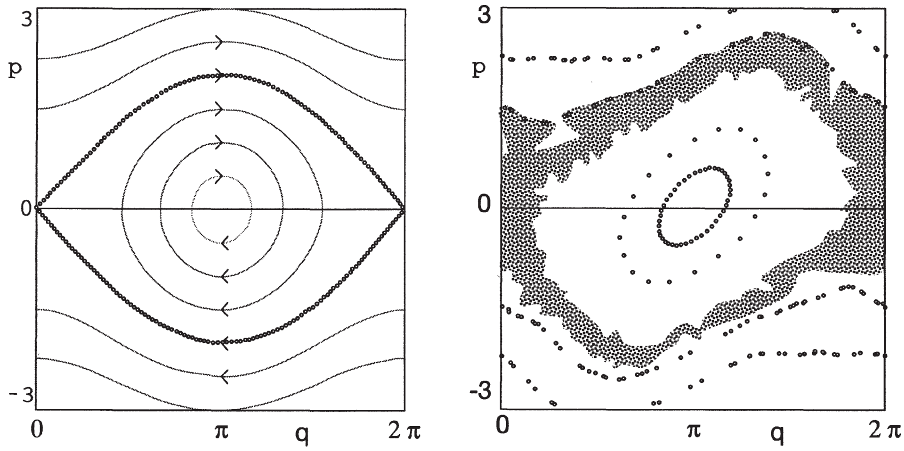

Figure 14. Phase plane for the nonlinear pendulum. Left panel: trajectories for the exact solution. Right panel: trajectories when the differential equation is integrated by a finite difference scheme, i.e., the 'standard mapping', with $\tau = 1$. The cross-hatched area is the region spanned by a single chaotic trajectory. To avoid printer overload, the individual iterates were erased and replaced by a uniform texture.

tially small in $1/\tau$ – and qualitatively different from the effects of finite differencing which are proportional to powers of the timestep.

These changes are easiest to explain by examining the trajectories of the differential equation and the difference system in the phase plane (Figure 14). The nonlinear pendulum is an exactly integrable system, and all trajectories are periodic. The closed curves in the phase plane represent side-to-side, small amplitude oscillations of the pendulum. The open curves at top and bottom show trajectories in which the pendulum swings through complete loops like a propeller. These two species of trajectories are divided by the 'separatrix', which is a trajectory that passes through $q = 0$, the unstable equilibrium, with zero momentum p. The separatrix and trajectories near it are super-sensitive to perturbations because only a tiny additional amount of momentum will suffice to push a large amplitude oscillation over the top and thereby convert it into a propeller-like motion.

The difference system, alias 'standard mapping', is not integrable and has chaotic solutions. When the time step τ is very small, however, one would expect that in some sense the difference and differential systems would be close to one another. Indeed, trajectories away from the separatrix are not drastically altered by the discretization; finite differences give a good approximation to these trajectories of the nonlinear pendulum. The neighborhood of the separatrix, however, dissolves into chaotic motion. The incoming and outgoing separatrices from the equilibrium split with a splitting angle at the point where the separatrices cross of approximately [136, 180]

$$\phi \sim \frac{3514.9}{\tau^3} \exp\left(-\frac{\pi^2}{\tau}\right). \tag{123}$$

The width of the region of chaos around the separatrices has a similar exponential dependence.

The bland statement 'the method is second order' or even a formal expansion of the errors in powers of the timestep τ completely misses the spawning of this region of chaos. Such exponentially small qualitative changes are perhaps less important in the real world, where the original differential equation probably has regions of chaotic motion anyway, than in the idealized world of exactly integrable systems such as the nonlinear pendulum, the Korteweg–deVries equation and so on. Still, it reiterates the theme that hyperasymptotics is important to numerical analysis.

Hakim and Mallick observe that the discretized system can be interpreted in two ways: (i) a second order accurate approximation to the pendulum system or (ii) a fourth order accurate approximation to a nonlinear differential equation which is obtained by modifying the pendulum equation by the addition of a higher derivative with a τ-dependent coefficient. Although the second interpretation seems rather artificial, it is also illuminating. Weakly nonlocal solitary waves and hydrodynamic boundary layers arise in this same way through addition of a higher derivative with a coefficient proportional to the small parameter. The result of such a singular perturbation is that the power series in the small parameter is divergent, and there are effects which depend on the exponential of the reciprocal of the small parameter.

17. Numerical Methods for Exponential Smallness or: Poltergeist-Hunting by the Numbers, I: Chebyshev and Fourier Spectral Methods

Because of the messiness of hyperasymptotic methods even for classical special functions and ordinary differential equations, numerical algorithms are important both as checks and as alternatives to hyperasymptotics. The exponential dependence on $1/\varepsilon$ for $\varepsilon \ll 1$ cries out for numerical schemes whose error also falls exponentially with the number of degrees of freedom M. Fortunately, Chebyshev and Fourier spectral methods [55, 91] and also Padé approximants [9, 19] have this property. In this section, we shall discuss spectral methods while Padé algorithms are described in the following section.

However, when the unknown function $f(\varepsilon)$ has only a divergent power series and also has contributions that lie beyond all orders in ε, both the rate of convergence and (sometimes) the methodology are altered. For example, when a function $f(x)$ which is analytic on $x \in [-1, 1]$ is expanded in a Chebyshev series, the error decreases *geometrically* with M, the truncation of the Chebyshev series. In other words, the error $E_M = O(\exp(-\mu M))$ as $M \to \infty$.

When $f(\varepsilon)$ has only a divergent power series about $\varepsilon = 0$, the Chebyshev or other spectral series on an interval that includes this point will lack geometric convergence. However, as long as the function is infinitely differentiable (with bounded derivatives), it is easy to prove by integration-by-parts that the error must decrease faster than any finite power of M. 'C^{∞}' singularities do not defeat spectral methods, but merely slow them down.

By using the method of steepest descents, one can often show that the spectral coefficients for functions which are infinitely-differentiable-but-nonananalytic on the expansion are 'subgeometric with exponential index of convergence β':

$$a_n \sim \nu(n)\exp(-\mu n^{\beta}), \tag{124}$$

where ν is a prefactor that varies algebraically rather than exponentially with n. Elliott [124, 125] pioneered this method, but Miller [227] first applied it to estimate Chebyshev coefficients for functions with divergent power series about one endpoint of the expansion interval.

For example, denoting the coefficients of the corresponding divergent power series by b_n and the Chebyshev expansion interval by $\varepsilon \in [0, \gamma]$, Miller found for $S(\varepsilon)$ and $\Re(S(-\varepsilon))$, respectively,

$$b_n = (-1)^n n! \to a_n \sim (-1)^n \sqrt{\frac{16\pi}{3\gamma}} \exp\left(\frac{1}{3\gamma}\right) \exp\left(-3\frac{n^{2/3}}{\gamma^{1/3}}\right), \tag{125}$$

$$\begin{aligned} b_n &= n! \to a_n \sim i(-1)^n \sqrt{\frac{16\pi}{3\gamma}}\ \exp\left(-\frac{1}{3\gamma}\right) \times \\ &\times \exp\left(-i\frac{3^{3/2}}{2}\frac{n^{2/3}}{\gamma^{1/3}}\right) \exp\left(-\frac{3}{2}\frac{n^{2/3}}{\gamma^{1/3}}\right). \end{aligned} \tag{126}$$

(One can show that the error in truncating an exponentially convergent series after the Nth term is proportional to a_N as explained in [55, 73].) Note that even when the asymptotic series is *monotonic*, corresponding to a principal value integral with a simple pole at $t = -1/\varepsilon$ on the path of integration, the Chebyshev series still happily converges. However, roughly twice as many terms are needed to achieve the same accuracy for $\Re(S(-\varepsilon))$ as for $S(\varepsilon)$ (Figure 15). Asymptotic approximations to the subgeometrically decreasing Chebyshev coefficients for many other special functions are given by Miller [227] and in the books by Luke [187] and Németh [236].*

It easy to derive asymptotic approximations for Chebyshev or other spectral coefficients for specific functions, but few general results are known. One such theorem applies to the class of Stieltjes functions, which is defined to be the set of all functions that can be written in the form

$$f(\varepsilon) \equiv \int_0^\infty \frac{\rho(t)}{1+\varepsilon t}\,\mathrm{d}t \tag{127}$$

for some positive semi-definite weight function $\rho(t)$ such that the moment integrals

$$b_n \equiv \int_0^\infty t^n \rho(t)\,\mathrm{d}t \tag{128}$$

* It appears from his English-language monograph that Néméth independently derived many asymptotic approximations in Hungarian-language articles in the mid-60's.

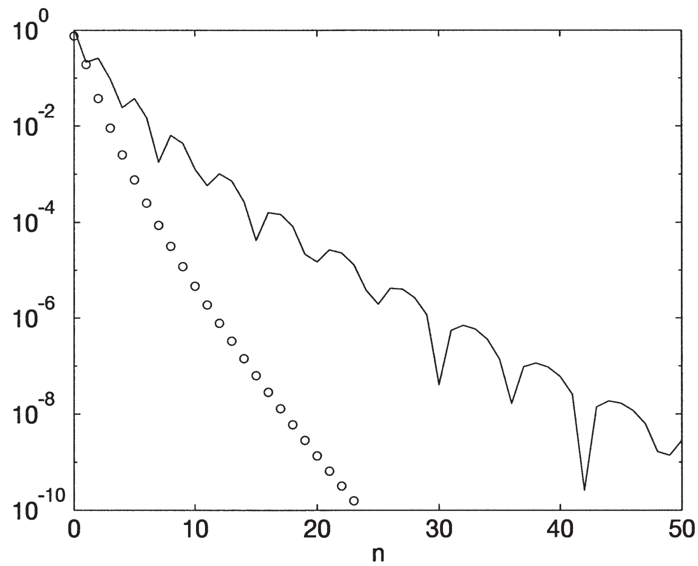

Figure 15. Chebyshev coefficients for the expansion on $\varepsilon \in [0, 1]$ of the Stieltjes function, $S(\varepsilon)$ [lower curve, circles] and $\Re[S(-\varepsilon)]$ (upper curve, solid).

exist for all nonnegative integers n. (These moments are also the coefficients of the power series expansion of $f(\varepsilon)$ about $\varepsilon = 0$.) W. G. C. Boyd has developed a general theory for hyperasymptotics for Stieltjes functions.

J. P. Boyd [49, 51] showed that if the power series coefficients diverge as $(n!)^r$ or $(rn)!$, i.e.,

$$\limsup_{n\to\infty} \frac{\log|b_n|}{n \log n} = r \tag{129}$$

then the Chebyshev coefficients satisfy the inequalities

$$\frac{2}{r+2} \geqslant \limsup_{n\to\infty} \frac{\log|(\log|a_n|)|}{\log n} \geqslant 1 - \frac{r}{2}. \tag{130}$$

Less precisely, the theorem implies that if the Chebyshev coefficients are decreasing like $O(\exp(-\mu n^{\beta}))$ for some constants μ and β, then β must be smaller $2/(r+2)$, implying subgeometric convergence for all $r > 0$, i.e., all factorial divergence. However, the exponential index of convergence β cannot be smaller than $1 - r/2$.

The integration-by-parts theorem shows that even for $r > 2$, the convergence of Chebyshev series is beyond all orders in $1/N$. However, it would be highly desirable to extend Boyd's theorem to more general classes of asymptotic functions, and perhaps sharpen it, too.

Berry [29] has shown that his error-function smoothing of Stokes phenomenon applies even to several broad classes of functions whose coefficients diverge faster than any factorial, so-called 'superfactorial asymptotics'. His numerical illustration

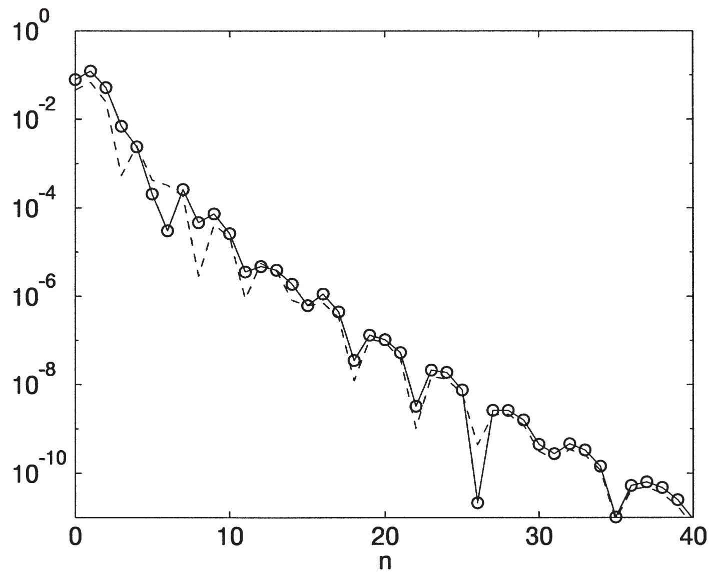

Figure 16. Solid with circles: The absolute values of the Chebyshev coefficients a_n for the expansion of Berry's superfactorial function, $\mathrm{BeS}(\varepsilon; A = 4)$ on $\varepsilon \in [0, 1]$. The thin dashed curve, closely tracking the solid curve, illustrates the coefficients of $\exp(-2/\varepsilon)$, which are known to decay as $\exp(-3\ 2^{1/3}\ n^{2/3})$.

is the function

$$\begin{aligned}\mathrm{BeS}(\varepsilon; A) &\sim \sum_{n=0}^{\infty} \exp\left(\frac{n^2}{A}\right)\varepsilon^n \\ &\equiv \sqrt{\frac{A}{\pi}}\mathrm{PV}\int_{-\infty}^{\infty} \frac{\exp\{-A(t-(1/2)\log(\varepsilon))^2\}}{1-\exp(2t)}\,\mathrm{d}t, \end{aligned} \tag{131}$$

where we have replaced his z by $-(1/2)\log(\varepsilon)$ so that the asymptotic series is a conventional power series. Figure 16 shows that Chebyshev polynomials have no problem in providing a highly accurate approximation even though the power series coefficients are blowing up like Gaussians of n. Unfortunately, there are a hundred papers on the asymptotics and hyperasymptotics of the confluent hypergeometric function for every one on the asymptotics of spectral series.

When exponentially small effects are present, there are often algorithmic challenges present, too. Numerical checking of the prefactors in front of $\exp(-q/\varepsilon)$ as obtained by the PKKS matched asymptotics (or whatever) may be impossible in single precision because the exponentially small quantity may fall below the single precision threshold for moderate ε, making it impossible to determine whether numerical differences are due to errors in the asymptotics, or the neglected effects of higher order terms at not-so-small ε. When α was smaller by a factor of 10^{-48} than the core of the solitary wave, Boyd [71] numerically computed the radiation coefficient α to a relative precision of six decimal places by using 70 decimal place arithmetic in Maple.

These calculations would seem to be very expensive since (i) spectral methods generate full matrices instead of the sparse matrices produced by finite differences and (ii) multiple precision arithmetic is excruciatingly slow in comparison to single precision operations, which are executed directly in silicon. However, most spectral solutions to boundary value and eigenvalue problems are performed with preconditioning. That is to say, the bulk of the code for solving a differential equation in multiple precision with spectral accuracy is to write a *low order* finite difference or finite element code to solve the inhomogeneous version of the problem in *single precision*. By repeatedly calling this finite difference solver, evaluating the residual in *multiple precision* with *spectral* methods at the end of each iteration, one can obtain a multiple precision, spectrally accurate solution without ever factoring (or even computing) the full, dense spectral matrix. By use of the Fast Fourier Transform, the spectral evaluation of the residual of an ordinary differential equation can be performed at a cost that grows as $O(N \log_2 N)$ operations where N is the number of degrees of freedom.

Another special difficulty that arises mostly in exponentially small phenomena is that of solutions on the infinite interval which do not decay to zero for $|x|$, but rather to sinusoidal oscillations. Two good strategies have been developed.

The first is to approximate the infinite interval by a large but spatially periodic interval, and then expand the solution as a Fourier series. The drawback is that the radiation coefficient α is sensitive to the spatial period P (modulo the wavelength of the far field oscillations, W). However, the periodic solutions themselves are often interesting. (In the atmosphere, for example, the solutions are *always* periodic in latitude and longitude.) In addition, the parameter P/W is actually a manifestation of a genuine degree of freedom, the 'phase factor' Φ, of the infinite interval. Consequently, it is possible to trace the entire parameter space for the unbounded domain by using the device of a large but periodic computational interval.

The second strategy is add one or more additional basis functions which are chosen to mimic the required asymptotic behavior of the 'wings' of the weakly nonlocal solitary wave (or whatever). When the width of the core structure is inversely proportional to ε and the wavenumber of the wing oscillations is k_f, an effective radiation basis function is

$$\begin{aligned}\phi_{\text{rad}}(x) \;\equiv\; & H(x+\Phi;\varepsilon)\sin(k_f x+\Phi) \\ & + H(-x+\Phi;\varepsilon)\sin(k_f - x+\Phi),\end{aligned} \tag{132}$$

where a smoothed approximation to the step function is defined by

$$H(x;\varepsilon) \equiv (1/2)\{1+\tanh(\varepsilon x)\}. \tag{133}$$

Boyd [60] has successfully applied a mixed basis of the rational Chebyshev polynomials [53] plus a single 'radiation function' to compute quantum scattering in one dimension. Boyd [62] shows that one can construct a basis function that depends *nonlinearly* on its coefficient – which is an approximation to α – so as to apply this method even when the oscillations for large $|x|$ are allowed to weakly self-interact, as is inevitable for nonlinear differential equations.

Boyd [54] used rational Chebyshev functions on the semi-infinite interval to approximate the J_0 Bessel function. This asymptotes to a sinusoidal oscillation rather than decaying, so a naive expansion in basis functions appropriate to an unbounded interval will fail disastrously. Nevertheless, he wrote, in a form mimicking the large x asymptotics,

$$J_0(x) \approx \frac{1}{\sqrt{(1+x)}} \big\{ P(x)\cos(x - \pi/4) + Q(x)\sin(x - \pi/4) \big\}. \tag{134}$$

Using a total of only 17 coefficients for $P(x)$ and $Q(x)$ combined gave a maximum absolute error of less than 2×10^{-7} *uniformly* on $x \in [0, \infty]$.

Numerical methods to replace divergent power series do demand some technology which is not otherwise widely used. It is encouraging that now this technology mostly exists and has been tested in applications.

18. Numerical Methods, II: Sequence Acceleration and Padé and Hermite–Padé Approximants

Sequence acceleration or 'summability' methods have a long history [139, 318, 315, 316]. The Euler sum acceleration is an elderly but still potent algorithm as already shown for the Stieltjes function. It is, however, but one of many schemes in the literature. We must refer to specialized reviews [139, 318, 315, 316] for an in depth discussion, but it is important to note one principle and one algorithm.

The principle is that acceleration methods are slaves to the oscillations of the jth term in the series with respect to j. For alternating series, that is, those for which the sign of the $(j + 1)$st term is opposite the sign of the jth term, and for nearly alternating series, acceleration methods are very effective. For monotonic series, that is, expansions whose terms are all of the same sign, some different but effective acceleration schemes are also known [315, 316]. However, when the series is slowly oscillating in degree j but not strictly monotonic, sequence acceleration algorithms tend to perform very badly [70].

The $[p/q]$ Padé approximant to a function $f(\varepsilon)$ is a polynomial of degree p divided by a polynomial of degree q which is chosen so that the leading terms of the power series of the approximant match the first $(p + q + 1)$ terms of the power series of $f(\varepsilon)$. The good news is that the Padé approximation usually *converges* even when the power series from whence it came *diverges*. For example, it has been rigorously proved that the $[N/N]$ approximant to the Stieltjes function $S(\varepsilon)$ converges with an error that decreases proportional to $\exp(-4N^{1/2}/\varepsilon^{1/2})$ – in other words, exponential but subgeometric convergence, similar to the Chebyshev series for this function [19].

Unfortunately, the Padé approximant fails along the branch cut for the Stieltjes function, which is the negative real ε-axis. Because the integral that defines the

Stieltjes function has a pole on the integration path when ε is real and negative, the function is not completely specified until we choose how the pole is treated. The Stieltjes function is real-valued if the integral is interpreted as a Principal Value integral, but has an imaginary part which is exponentially small in $1/\varepsilon$ if the path of integration is indented above or below the pole in the t-plane. That is,

$$S(-\varepsilon) = PV \int_0^{\infty} \exp(-t) \frac{1}{1-\varepsilon t}\, \mathrm{d}t \pm i\pi \frac{1}{\varepsilon} \exp(-1/\varepsilon), \tag{135}$$

where the sign of the imaginary part depends on whether the contour is idented above or below the real t-axis. Since the terms in the Stieltjes power series, $S(\varepsilon) \sim \sum_{j=0}^{\infty} (-1)^j j!$, are all real-valued, one can prove that the coefficients of the numerator and denominator polynomials in the Padé approximant are real, too. Even if the approximants converged, they would inevitably miss the imaginary part of $S(-|\varepsilon|)$.

The same difficulty arises in quantum mechanics. For the quartic oscillator, for example, the eigenvalue E of the stationary Schrödinger equation is complex-valued when the coupling constant ε is negative; physically, the imaginary part is the inverse of the lifetime of a metastable bound state, which eventually radiates away from the local minimum of the potential energy. The exact $\Im(E)$ is not analytically known, but it does decrease exponentially fast with $1/|\varepsilon|$. Because $\Im(E)$ gives the lifetime of the state, and therefore the rate of radiation, it is a very important quantity even when small. Ordinary Padé approximants fail when ε is real and negative, however, just as for the Stieltjes function. (In fact, it has been proved that both the eigenvalue of the quartic oscillator and $S(\varepsilon)$ belong to a class of functions called Stieltjes fuctions, and thus are close cousins.)

Shafer [282] developed a generalized Padé approximant which has been used successfully by several groups to calculate exponentially small imaginary parts of quantum eigenvalues [300, 131, 280, 287, 281]. The approximant $f[K/L/M]$ is defined to be the solution of the quadratic equation

$$P(f[K/L/M])^2 + Qf[K/L/M] + R = 0, \tag{136}$$

where the polynomials P, Q and R are of degrees K, L and M, respectively. These polynomials are chosen so that the power series expansion of $f[K/L/M]$ agrees with that of f through the first $N = K + L + M + 1$ terms. (The constants in P and Q can be set equal to one without loss of generality since these choices do not alter the root of the equation, so the total number of degrees of freedom is as indicated.) As for ordinary Padé approximants, the coefficients of the polynomials can be computed by solving a matrix equation and the most accurate approximations are obtained by choosing the polynomials to be of equal degree, so-called 'diagonal' approximants.

Figure 17 shows the diagonal approximant of the Stieltjes function for negative real ε; the polynomials for the [4/4/4] approximation are

$$P = 1 - (160/3)\varepsilon + (3680/3)\varepsilon^2 + (3680/3)\varepsilon^3 + 7240\varepsilon^4,$$

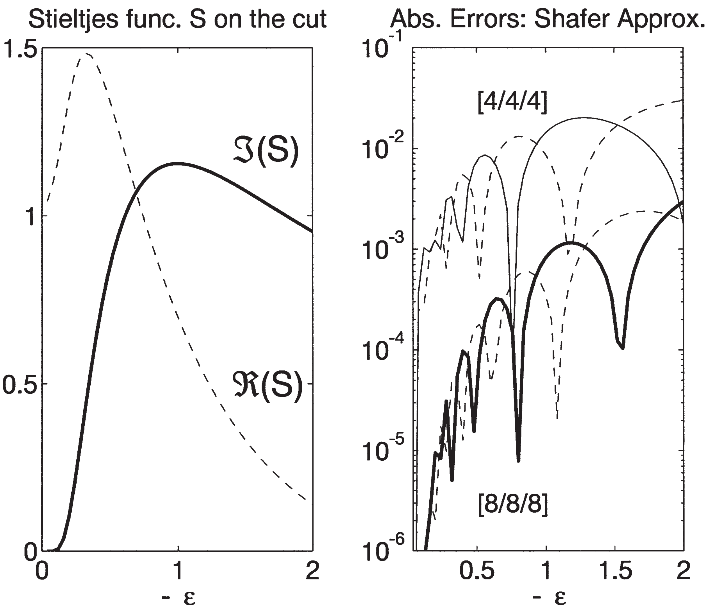

Figure 17. Left panel: the real and imaginary parts of the Stieltjes function on the negative real ε-axis. Right: the absolute value of the absolute error for the real part (thin dashed) and imaginary part (thick solid) part in the [4/4/4] Shafer approximant (top pair of curves) and the [8/8/8] approximant (bottom solid and dashed curves).

$$\begin{aligned} Q &= 1 + (2239/9)\varepsilon - (2698/9)\varepsilon^2 + (43658/3)\varepsilon^3 + 5056\varepsilon^4, \\ R &= -2 - (1732/9)\varepsilon - (7126/9)\varepsilon^2 - (41504/3)\varepsilon^3 - (2452/9)\varepsilon^4. \end{aligned} \quad (137)$$

The two roots of the quadratic equation give us the result of choosing to indent the contour either above or below the path of integration; discarding the imaginary part gives the Principal Value of the integral. The [10/10/10] approximant gives a maximum relative error for both parts of less than 5×10^{-6}.

Shafer's idea can be generalized to polynomials of higher degree in the approximation. The result is usually called a 'Hermite–Padé' approximant. The quadratic or 'Shafer' approximants seem to be quite successful for most quantum problems [300, 131, 280, 287]. However, Sergeev and Goodson describe fast algorithms for computing approximants of higher degree and also solve some problems where such higher degree approximants, capable of representing cube roots and higher branch points, are very useful [281].

Padé approximants have been generalized in several other directions, too [129]. Reinhardt [269] has developed a procedure using *double* Padé approximants which works well even for monotonic, factorially-diverging series, including the computation of exponentially small $\Im(E)$, although it has been largely displaced by Shafer approximants. Another generalization is to approximate f by the solution of an ordinary differential equation ([10] and earlier references cited there) or a partial differential equation ([97] and earlier articles therein) where again the coefficients of the differential equation are polynomials in ε chosen so the approximant has a power series matching f up to some finite order.

Padé methods have limitations; they seem to be quite useless for calculating the exponentially small radiation coefficient of a weakly nonlocal solitary wave, for example. Nonetheless, many problems in quantum mechanics have fallen to Hermite–Padé approximants.

19. High-Order Hyperasymptotics versus Chebyshev and Hermite–Padé Approximations

> I wrote an analytical solution to a sixth order differential equation as a hypergeometric integral, derived asymptotic approximations, matched the boundary conditions, and finally went to a computer to make graphs. The machine took about a minute. Then I solved the whole problem numerically, and the same machine took about two seconds. That was the last analytical work I ever did!
>
> R. E. Dickinson [115]

As illustrated by Dickinson's amusing experience, very complicated analytical answers are really just another form of numerical solution. High order asymptotic and hyperasymptotic solutions are usually in this category because a long string of terms adds little insight to the lowest term, only greater numerical accuracy. Consequently, a proverb in perturbation theory is: One term: insight; several terms: numerics.

If many terms, as opposed to three or four terms, are available, it is possible to deduce some non-numerical information from the series such as the approximate location of the convergence-limiting singularities in the complex plane (another use of Darboux's Principle!) However, this analytic information is mostly used only to transform the series to improve the rate of convergence as described in van Dyke's book [298], for example. Fluid dynamicists are not too interested in branch points at complex values of the spatial coordinates!

In the rest of this section, we shall focus on answering the question suggested by these considerations: How useful are high order hyperasymptotics *numerically* in comparison to other numerical methods?

Although many books and articles on beyond-all-orders methods offer numerical tables, head-to-head comparisons between hyperasymptotics and other numerical algorithms are rare. Most of the work catalogued in Table IV has a decidely pre-computer spirit.

One reason for this lack of efficiency contests is that most software libraries already have quite efficient routines for computing special functions. Typically, the algorithms for computing Bessel and Airy functions, the Error Integral, and so on employ *two* expansions for each function. Through experimentation and theory, a breakpoint ζ is chosen for each function. A convergent power series is used for $|z| \leqslant \zeta$ whereas a divergent expansion in form of an exponential or trigonometric pre-factor multiplied by a series of inverse powers of z is employed for $|z| > \zeta$.

(The divergent series may actually be different expansions in different parts of the complex plane because of Stokes' phenomenon.) Hyperasymptotics might seem like a natural way to add a few decimal places of additional accuracy.

Unfortunately, the extension beyond the superasymptotic approximation is a series of terminants like Equation (73). Olde Daalhuis [240, 242] has developed good numerical methods for evaluating terminants, but the approximations are series of hypergeometric functions which can only be evaluated by recurrence. This implies that each terminant is itself as expensive to evaluate as the special function it helps to approximate. Thus, terminant series are numerically inefficient. It is probably more sensible to simply increase the 'break point' where the subroutine switches from the power series to the inverse power series and also increase the number of terms in each series.

In practice, both series are replaced by the equivalent Chebyshev series, which converge faster and also are much more resistant to roundoff error. It is useful to illustrate how easily these expansions can be derived to replace the standard divergent asymptotic series.

For example, the Stieltjes function $S(\varepsilon)$ satisfies the ordinary differential equation

$$\varepsilon^2 \frac{\mathrm{d}S}{\mathrm{d}\varepsilon} + (1+\varepsilon)S = 1. \tag{138}$$

For paper-and-pencil or symbolic language calculation, the simplest method is the 'Lanczos τ-method'. He observed [177, 284] that if we perturb the right-hand side of Equation (138) by a polynomial of degree M, multiplied by an unknown constant τ, we can then solve this perturbed equation *exactly* by a polynomial of degree M. (Instead of the usual strategy of *approximately* solving the exact differential equation, the τ-method *exactly* solves an *approximate* differential equation.) If the perturbing polynomial is ε^M, then the τ-method yields the first M terms of the usual divergent power series in ε.

However, this is actually a rather stupid choice if the goal is uniform accuracy on some interval $\varepsilon \in [0, z]$ where z is a complex number. The power function ε^M is extremely nonuniform – very small near the origin, but increasing very rapidly away from it. The polynomial of degree M (and leading coefficient of one) which is most uniform on $[0, z]$ is the shifted Chebyshev polynomial $T_M^*(\varepsilon/z)$ [177, 178, 55].

Table VI is a short code in the symbolic manipulation language Maple to solve the ordinary differential equation

$$\varepsilon^2 \frac{\mathrm{d}S}{\mathrm{d}\varepsilon} + (1+\varepsilon)S = 1 + \tau T_M^*(\varepsilon/z) \tag{139}$$

through a Chebyshev τ-method. The most important feature of the table is simply its brevity: all the necessary algebra is performed in exact, rational arithmetic under the control of only nine lines of code! The choice of Maple is arbitrary; the same calculation could be performed with equal brevity in any of the other

Table VI. A Maple program to compute the Chebyshev-τ approximation for the Stieltjes function $S(x)$.

```
# For brevity, epsilon is replaced by x
M := 8; M1 := M + 1; # degree of Chebyshev polynomial
# Next, compute the shifted Chebyshev polynomial TN
T[0] := 1; T[1] := 2*y - 1;
for j from 2 by 1 to M do T[j] := 2*(2*y - 1)*T[j - 1] - T[j - 2]; od;
y := x/z; TM := simplify(T[M]);
S := a0; for j from 1 by 1 to N do S := S + a.j*x**j; od;
resid := x*x*diff(S, x) + (1 + x)*S - 1 - tau*TM; resid := collect(resid, x);
for j from 0 by 1 to M1 do eq.j := coeff(resid, x, j); od;
eqset := eq.(0..M1); varset := tau, a.(0..M); asol := solve(eqset, varset); assign(asol);
x := z; Srational := simplify(S);
```

widely used symbolic algebra languages including Mathematica, MACSYMA and Reduce. Boyd [55, 63] gives many examples of problem-solving via spectral methods in algebraic manipulation languages. Note that because all operations involve polynomials, not transcendentals, the code also executes very speedily.

The Chebyshev-τ result is rather messy: polynomial in ε, rational in z. However, the Chebyshev (or Chebyshev-like) expansion will obviously converge most rapidly when the expansion interval $[0, z]$ is small as possible for a given ε. This implies that for best results, one should choose $z = \varepsilon$. Making this substitution not only optimizes accuracy for a given ε, but also simplifies the result to a *rational* function of ε alone. The $M = 8$ approximation, which is a polynomial of degree 7 over a polynomial of degree 8, is

$$\begin{aligned} S_8 \equiv\ & 16\{4 + 124\varepsilon + 1336\varepsilon^2 + 6168\varepsilon^3 + 12173\varepsilon^4 + \\ & +8955\varepsilon^5 + 1737\varepsilon^6 + 33\varepsilon^7\}/ \\ & \{256 + 8192\varepsilon + 93184\varepsilon^2 + 473088\varepsilon^3 + 1108800\varepsilon^4 + \\ & +1128960\varepsilon^5 + 423360\varepsilon^6 + 40320\varepsilon^7 + 315\varepsilon^8\}. \end{aligned} \tag{140}$$

Figure 18 shows that for small positive ε, this simple rational approximation is on the whole a lot more useful than either the superasymptotic or hyperasymptotic series.

Chebyshev polynomial approximations are usually polynomials rather than rational functions and are optimized for a particular line segment in the complex plane. By computing symbolically, we have obtained an approximation that is more complicated (because it is rational rather than polynomial) but has the great virtue of being as accurate, for a given ε, as the standard Chebyshev approximation of degree M along the segment $[0, \varepsilon]$ even when ε is complex-valued.

In the previous section, we have already given the ordinary asymptotics of the Chebyshev coefficients of the Stieltjes function. However, the comparison of convergent Chebyshev series with divergent power series is not completely straightforward. The asymptotic series uses a number of terms N which is inversely proportional to ε. What happens if we compare the N-term Chebyshev series on the interval $\varepsilon \in [0, 1/N]$ with the N-term optimally truncated power series for $\varepsilon = 1/N$?

Through an elementary steepest descent analysis of the usual inner product integrals for the coefficients of an orthogonal series, one finds that the Nth Chebyshev coefficient for the series on $[0, 1/N]$ is (previously unpublished)

$$\begin{aligned} a_N &\sim 2.98\sqrt{N}\exp(-2.723N) \\ &\sim 2.98\varepsilon^{-1/2} \quad \exp(-2.723/\varepsilon). \end{aligned} \tag{141}$$

(One can show that the error in truncating the Chebyshev series after N terms is proportional to a_N [55, 73].) Intriguingly, the errors for Padé approximants and for hyperasymptotics are of this same form:

$$|f - f_N| \leqslant \exp(-q/\varepsilon), \quad N \sim \mathrm{O}(1/\varepsilon), \tag{142}$$

where the constant $q > 0$ depends on the precise Chebyshev, Padé, or hyperasymptotic scheme used. There are likely deep connections between these different families of approximations which are now only dimly understood [51].

One can make more entertaining approximations by using other spectral basis sets. For example, the rational Chebyshev functions are a good basis set for the semi-infinite interval, $x \in [0, \infty]$. Boyd [54] gives three examples in which the usual *pair* of series – divergent series in $1/x$ for large x and convergent power series for small x – can be replaced by a *single* expansion over the entire semi-infinite range. The examples range from the K_1 Bessel function, which has a pole at the origin, to the J_0 Bessel function, in which separate series multiply the sine and cosine in the uniform approximations, to the ground state eigenvalue of the quantum quartic oscillator as a function of the coupling constant ε, which is a Stieltjes function with a factorially divergent power series about $\varepsilon = 0$ [47, 19]. These uniform approximations are much complicated and converge more slowly than the pair of Chebyshev series they replace, but have the advantage of avoiding a conditional statement, which is needed in the traditional approach to switch between large and small x approximations.

Lastly, one must not overlook non-series alternatives. Schulten, Anderson and Gordon [277] have developed an efficient subroutine to evaluate the Airy functions at arbitrary points in the complex plane. Instead of using an asymptotic approximation for large $|z|$, they use a clever optimized Gaussian quadrature to directly evaluate the integral representations for Ai and Bi, even on Stokes' lines. Their double precision code, which is accurate to at least 11 decimal places for all $|z|$ (with use of the power series about $z = 0$ near the origin) employs a maximum of just *six* quadrature points!

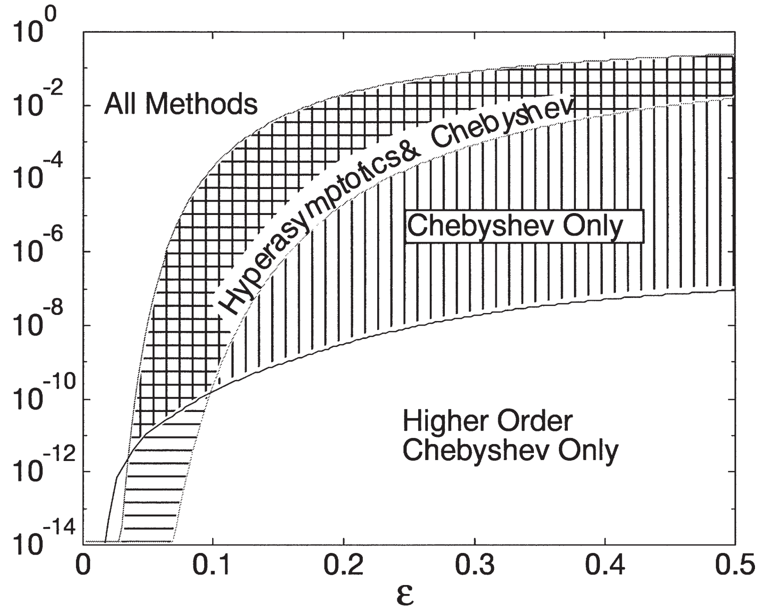

Figure 18. A comparison of the rational τ-Chebyshev approximation S_8 versus the superasymptotic and hyperasympotic approximations for the Stieltjes function, $S(\varepsilon)$. The three solid curves plot the errors for each of the three methods versus ε. If acceptable accuracy for a given ε is a point in the unshaded region in the upper left corner, all three methods are satisfactory. In the unshaded lower right region, none of the three approximations is sufficiently good (although such tiny errors $< 1 \times 10E^{-12}$ can be achieved by simply using a Chebyshev τ approximation of higher order). The vertical shading – most of the graph – shows where the Chebyshev approximation S_8, a polynomial of degree 7 divided by degree 8, is successful, but the asymptotic series fail because their minimum error is larger than the required tolerance. In the vertically-and-horizontally shaded area, both hyperasympotics and $S_8(\varepsilon)$ are successful. Finally, there is a tiny region of horizontal shading where only hyperasymptotics is successful (though a Chebyshev approximation of higher order than $S_8(\varepsilon)$ would succeed). The hyperasymptotic errors were calculated using the Berry–Howls scheme (where the errors are $O(\exp(-2.386/\varepsilon))$, but employing the more accurate hyperasymptotic methods of later authors such as Olde Daalhuis would not change the theme of the graph.

Detailed comparisons between high order hyperasymptotics and other methods of numerical approximation have not yet been carried out. Still, the examples and illustrations above show that the comparison, except perhaps for special cases, is likely to be unfavorable to high order asymptotics.

Although hyperasymptotics look comparable to Chebyshev and Padé schemes when $N \sim O(\varepsilon)$, the Chebyshev and Padé have the profound advantage of *converging* as $N \to \infty$ for fixed ε. Furthermore, these approximations are built from ordinary polynomials whereas hyperasymptotic approximations are series of hyperterminants, which in turn are approximated by series of hypergeometric functions.

There may be a few exceptions: problems where no alternatives are available. In quantum chaology, Gutzwiller's divergent series for the quantum spectrum has been summed using resurgence [182, 39]. The practical result has been greatly improved energies for quantum mechanics odd-shaped billiard tables – idealized but popular for testing theories [182, 39]. Another application is computing the zeros of the Riemann zeta function, where resurgence has proved to be much better than the best available competitor, the (un-resurgent) Riemann–Siegel formula [34].

Berry's amusing quote,'I am not expecting an early call', is a frank admission that most extensions of exponential asymptotics beyond the lowest nontrivial term are arithmurgically useless. The proper use of exponential asymptotics is to give insight. A sensible application is to compute a small term that is also the leading term to approximate some crucial feature of a problem, perhaps the lifetime of a quantum bound state or a nonlocal solitary wave.

20. Hybridizing Asymptotics with Numerics

The hyperasymptotic scheme of Boyd [68] and the PKKS method [171, 172] are both blends of analysis and numerics in the sense that the final step, the determination of the proportionality constant which multiplies the exponential of $1/\varepsilon$, requires a computation. However, the prior analysis has reduced the problem to a very small calculation that returns an answer as the product of a number with an analytical factor. This is far different than a brute force calculation that requires a hundred times as much computer time to return only a number.

The flow past a sphere or cylinder at small Reynolds number Re [264, 159, 161, 160, 98, 283] has frustrated fluid dynamicists for over forty years, but there has been, very recently, a partial breakthrough by means of a hybrid numerical-asymptotic method [170]. The source of pain is that these expansions are double series in powers of Re and $1/\log(\mathrm{Re})$ or, defining $\varepsilon = 1/\log(\mathrm{Re})$, in powers of $\exp(-1/\varepsilon)$ and ε. Formally, one should include an infinite number of logarithmic corrections to the drag coefficient before computing the first correction proportional to Re. For the flow past the circle, however, $\mathrm{Re} > (1/\log(3.70/\mathrm{Re}))^4$ for all $\mathrm{Re} > 1/12000$. (Real fluid flows are typically at much larger Re.) A systematic scheme for the transcendentally small terms is still an open problem. For the sphere, which is probably the easier of the two, Chester and Breach conclude sadly 'the expansion is of practical value only in the limited range $\mathrm{Re} < 1/2$ and that in this range there is little point in continuing the expansion further'.

Kropinski, Ward and Keller [170] made the crucial observation that if the outer ('Oseen') problem is solved numerically, the numerical solution will implicitly incorporate an infinite number of logarithmic corrections. Better yet, the outer solution is independent of whether the body is a cylinder, an elliptic cylinder, or some other smooth shape: a single numerical solution provides a good answer to a whole spectrum of body shapes. The inner solution differs from shape to shape, but is easy to calculate analytically. They have successfully applied this same idea, outer-numerical/inner-analytical, to other problems with logarithmic corrections [308]. The end product neglects higher powers of ε and the necessary numerics is the full solution to a PDE, but still, their work is real progress after two decades of no advance at all.

It seems likely that such hybrid numerical-asympotic methods will flourish in the next few years for following reasons:

- Analytical perturbation theory has enjoyed at least a century of development, and it is hard to grow good ideas in such old soil.
- Mathematics departments, even in the noncomputational areas, are becoming more computer-friendly.
- Hyrid algorithms have successfully attacked a number of problems already.
- There are broad areas where hybrids have not yet been tried.

21. History

Improving upon the minimum error of an asymptotic series has a long history; Stieltjes himself discussed the possibility in his 1888 doctoral thesis. Oppenheimer's calculation of the exponentially large decay time in the quantum Stark effect and the independent discovery of quantum tunnelling by Gamow and by Condon and Gurney all happened in 1928. The Euler acceleration of the Stieltjes function series was first analyzed by Rosser [273] in 1951.

One can distinguish several parallel lines of development. The first is the calculation of 'converging factors' or terminants for the asymptotic expansions of special functions, beginning with Airey in 1937 [3] and reaching a high degree of sophistication in the books of Dingle (1973) and Olver (1974) [118, 249], who also give good histories of earlier work.

Another was quantum mechanics, beginning with discovery of tunnelling in 1928, continuing with the Pokrovskii–Khalatnikov solution for 'above-the-barrier' quantum scattering, and continued to the present with studies of high order perturbation theory. The books written by Arteca, Fernandez and Castro [8] and edited by LeGuillou and Zinn-Justin [181] and Braaksma [83] are good testaments, as is a special issue of *International Journal of Quantum Chemistry* [269].

A third area is KAM theory and dynamical systems theory in general. Under perturbations, integrable dynamical systems become chaotic, but the chaos is confined to exponentially thin regions around the separatrices [136] for small ε. Through 'Arnold diffusion', dynamical systems can move great distances in phase space (on exponentially long time scales) even when the perturbation is very weak.

A fourth area is 'weakly nonlocal solitary waves', that is, nonlinear coherent structures that would be immortal were it not for weak radiation away from the core of the structure. These seem to be as ubiquitous as classical, decaying-to-zero solitary waves. Nonlocal solitary waves arise in fiber optics, hydrodynamics, plasmas and a wide variety of other applications. Meiss and Horton (1983) [201] seem to have done the earliest explicit calculations. However, the existence of slowly radiating solitary waves in particle physics (‘ϕ^4 breathers’) and oceanography (Gulf Stream Rings) was known from observations and initial value computations a decade earlier. The subsequent eruption of activity is catalogued in the book by Boyd [72].

A fifth area is crystal formation and solidification. The 1985 work of Kruskal and Segur [171, 172] resolved a long-standing roadblock in the theory of dendritic fingers on melt interfaces, and touched off a great plume of activity. There was rapid cross-fertilization with nonlocal solitary waves because Segur and Kruskal applied their new PKKS method to the ‘ϕ^4 breather’ of particle physics, contributing to the rapid growth of exponential asymptotics for nonlinear waves.

A sixth area is fluid mechanics. The Berman–Terrill–Robinson problem [135] in flows with suction, the radiative decay of free oscillations bound to islands [185] and Kelvin wave instability in oceanography and atmospheric dynamics [74, 75] were all examples in which exponential smallness had been calculated in the seventies or early eighties. Somehow, these problems remained isolated. However, boundary layer theory always involves divergent power series and exponential smallness as showed by example above. Fluids is an area where hyperasymptotic technology is likely to have a vigorous future.

A seventh line of research is that pursued by Richard E. Meyer and his students. This began with studies of adiabatic invariants [202 – 204, 219, 205, 207]. He also devised an independent solution to ‘above the barrier’ quantum scattering: recasting the problem as an integral equation so that the reflection coefficient appears as the dominant contribution instead of as an exponentially small correction [206, 208]. This led to further studies of exponential smallness in water waves trapped around an island [185, 209, 220], connection across WKB turning points and wave dynamics and quantum tunnelling [221, 211, 222 – 224, 212, 214 – 216, 225, 218, 226]. Meyer has also written four reviews [210, 213, 217, 218].

An eighth line of development is the abstract theory of resurgence and multisummability. This began with Écalle [123] and continued with important contributions from Pham, Ramis, Delabaere, Braaksma and others too numerous to mention as reviewed in [285, 83, 15].

Lastly, a ninth area is the development of resurgence and Stokes phenomenon by physicists and applied mathematicians. This grew out of the abstract theory of Écalle, which Berry learned during a visit to France, but took resurgence in a direction that was less rigorous but much more pragmatic and applied. The trigger was Berry’s 1989 realization that the discontinuity in the numerical value of an asymptotic expansion at a Stokes’ line could be smoothed. (The change in *topology* of the steepest descent path at a Stokes lines is unavoidable, however.) Building on the books of Dingle, Olver and Écalle, Berry, Howls, Olver, Olde Daalhuis, Paris, Wood, W. G. C. Boyd and others have developed smoothed, high order hyperasymptotic approximations for many species of special functions, for the WKB method and for other schemes for differential equations. A selection is given in Table IV.

Dingle's ideas of generic forms for the late terms in asymptotic series and universal terminants now seem as important to the rise of exponential asymptotics as the comet-crash(?) that put an end to the dinosaurs was in biology. Only a year later, Olver's book developed similar ideas, with error bounds, for ordinary differential equations. And yet, though these books were widely bought and read, their net effect at the time was as quiet as a sandcastle washed away by a rising tide. Bothered by long-term illness, Dingle never published again.

In recent years, however, the analysis of exponentially small terms has exploded. A special program of study at the Newton Institute at Cambridge has brought together researchers from a wide range of fields for a workshop lasting the whole first half of 1995. The books by Boyd [72] and Segur, Tanveer and Levine (eds) [279] are good introductions to the vigour and diversity of this interest.

Why was this revolution in asymptotics so slow, so long delayed? Perhaps the most important factor is that alterations in scientific world-view, like atom bombs, require assembling a critical mass. Part of this critical mass was provided by the parallel threads of slow development outlined above; when ideas began to cross disciplinary boundaries, exponential asymptotics exponentiated. Another trigger was the popularization of algebraic manipulation languages, which made it easier to compute many terms of an asymptotic series. Lastly, applied mathematics is subject to fads and enthusiasms.

I myself read both the Dingle and Olver books when they first appeared while I was still in graduate school, but was unimpressed. First, my ε was not very small. Second, a string of messy hyperasymptotic corrections seemed a poor alternative to numerical algorithms, which were fast and efficient even a quarter century ago. Modern exponential asymptotics still shares these limitations, but there is now a cadre of enthusiasts who are unbothered as there was not in Dingle's time.*

Still, with the emergence of exponential asymptotics as a subfield of its own with ideas shared widely from physics to fluids to nonlinear optics, hyperasymptotics has been very useful, at least as the lowest hyperasymptotic order, in a wide variety of practical applications. When the parallel threads ceased to be parallel and converged, the ancient topic of asymptotics suddenly became very interesting again.

22. Books and Review Articles

The theme of extending asymptotic series through Borel summation and other methods of re-expanding remainder integrals is treated in the classic books of Dingle [118] and Olver [249]. Jones' 1997 book is very short (160 pages), a primer of steepest descent and hyperasymptotics that is perhaps closest in style and spirit to

* In a language of Papua New Guinea, the word 'mokita' is used to denote 'things we all know but agree not to talk about'.

the Dingle and Olver books, but at a somewhat more elementary level. It includes a short appendix on nonstandard analysis as well as exercises at the end of each chapter.

Écalle's 1981 three-volume treatise greatly extended and generalized earlier ideas on hyperasymptotics. Unfortunately, his work has not been translated from French. However, Sternin and Shatalov is a recent presentation of the abstract theory of resurgence [285]. The collection of articles edited by Braaksma [83] gives a broader but less coherent state of the abstract resurgence work. Balser [15] is only one hundred pages long, but is very readable, based on a course taught by the author.

Kowalenko *et al.* [169] is a short monograph devoted entirely to the hyperasymptotics of a fairly narrow class of integrals. Maslov [199] is a broad treatment of the WKB method.

Segur, Tanveer and Levine [279] is a collection of articles from a NATO Workshop that displays the remarkable breadth of application of beyond-all-orders asymptotics that existed even in 1991. Arteca, Fernandez and Castro [8] and LeGuillou and Zinn-Justin [181] describe the calculation of exponentially small terms in quantum mechanics through large order perturbation theory and summation methods. Boyd [72] is focused particularly on nonlocal solitary waves, but it includes a chapter on general applications of hyperasymptotics and several chapters on numerical methods.

Curiously, review articles seem rarer than books. Berry and Howls [39], Paris and Wood [260] and Wood [320] have written short, semi-popular reviews. Delabaere [112] has written (in English) an introduction to Écalle alient calculus. Olde Daalhuis and Olver [248] describe hyperasymptotics (and numerical methods) for linear differential equations. Byatt-Smith [87], based on an unpublished but widely circulated manuscript of seven years earlier, is not technically a review, but it nonetheless is one of the most readable treatments of re-expansion of remainder integrals and the error function smoothing of Stokes phenomenon.

This profusion of books and reviews is helpful, but there are still some large gaps. This present article was written to fill in some of these holes and point the reader to other summaries of progress.

23. Summary

> What they [engineers] want from applied mathematics ... is information that illuminates.
>
> Richard E. Meyer (1992) [218, p. 43]

Key concepts:

– Divergence is a disease caused by a perturbative approximation which is true for only part of the interval of integration or part of the Fourier spectrum.

- A power series is asymptotic when the perturbative assumption is bad only for a part of the spectrum or integrand that makes an exponentially small contribution.
- When a factorially divergent series is truncated at its smallest term, this 'optimal truncation' gives an error which is typically an *exponential* function of $1/\varepsilon$. The usual Poincaré definition of asymptoticity, which refers only to *powers* of ε, is therefore rather misleading. The neologism 'superasymptotic' was therefore coined by Berry and Howls to describe the error in an optimally-truncated asymptotic series.
- By appending one or more terms of a second asymptotic series (with a different rationale) to the optimal truncation of a divergent series, one can reduce the error below that of the superasymptotic approximation to obtain a 'hyperasymptotic' approximation. This, too, is divergent, but with a minimum error far smaller than the best 'superasymptotic' approximation. (This rescale-and-add-another-series step can be repeated for further error reduction.)
- There are many different species of hyperasymptotic methods including:

  (1) Sequence acceleration schemes such as the Euler, Padé and Hermite–Padé (Shafer) approximations.
  (2) Complex-plane matched asymptotics (the Pokrovskii–Khalatnikov–Krusal–Segur method).
  (3) Resurgence schemes.
  (4) Isolation of exponential smallness.
  (5) Special numerical algorithms, usually employing Chebyshev or Fourier spectral methods or Gaussian quadrature.

- The history of exponential asymptotics stretches back at least a century with several parallel lines of slow development that reached a critical mass only within the last six years, culminating in an explosion of both applications and theory that will touch almost every field of science and engineering as well as mathematics.

The list of open problems is large. One is a rigorous numerical test of many-term, high order hyperasymptotic expansions versus competing methods, such as Chebyshev series, for special function software. (The arguments presented above suggest that the results are likely to be unfavorable to hyperasymptotics.)

Another is to create an expanded theory for the connection between the rate of growth of power series coefficients or other properties of functions with divergent power series and the rate of convergence of Chebyshev series and Padé approximants. Some theorems exist for the special class of Stieljtes functions (Chebyshev [49, 51] and Padé [19]), but little else.

An important issue is whether the Dingle terminant formalism can be extended to weakly nonlocal solitary waves. The radiation coefficient α, which is proportional to the function $\exp(-\mu/\varepsilon)$ for some constant μ, has only the trivial power series $0 + 0 \cdot \varepsilon + 0 \cdot \varepsilon^2 + \cdots$. Does α somehow influence the coefficients of the ε power series subtly so that terminants can be applied, or is the radiation condition truly a ghost, forever invisible to methods that look only at the asymptotic form of the power series coefficients?

A fourth domain of future study is to apply exponential asymptotics to new realms. We have shown above that the theory of numerical algorithms contains hidden beyond-all-orders terms, but this aspect of numerical analysis is largely *terra incognita*.

Although applications and fundamental research on exponentially small terms will doubtless continue for many years, we have tried to show that the underlying principles are neither complicated nor obscure.

Acknowledgements

This work was supported by the National Science Foundation through grant OCE9119459 and by the Department of Energy through contract KC070101. I thank Richard Meyer, Michael Ward and Robert O'Malley for helpful correspondence or conversations, and others too numerous to mention for supplying reprints and references. I am grateful to the three referees for their extremely careful reading of this long paper.

References

1. Ackerberg, R. C. and O'Malley, R. E., Jr.: Boundary layer problems exhibiting resonance, *Stud. Appl. Math.* **49** (1970), 277–295. Classical paper illustrating the failure of standard matched asymptotics; this can be resolved by incorporating exponentially small terms in the analysis (MacGillivray, 1997).
2. Adamson, T. C., Jr. and Richey, G. K.: Unsteady transonic flows with shock waves in two-dimensional channels, *J. Fluid Mech.* **60** (1973), 363–382. Shows the key role of exponentially small terms.
3. Airey, J. R.: The "converging factor" in asymptotic series and the calculation of Bessel, Laguerre and other functions, *Philos. Magazine* **24** (1937), 521–552. Hyperasymptotic approximation to some special functions for large $|x|$.
4. Akylas, T. R. and Grimshaw, R. H. J.: Solitary internal waves with oscillatory tails, *J. Fluid Mech.* **242** (1992), 279–298. Theory agrees with observations of Farmer and Smith (1980).
5. Akylas, T. R. and Yang, T.-S.: On short-scale oscillatory tails of long-wave disturbances, *Stud. Appl. Math.* **94** (1995), 1–20. Nonlocal solitary waves; perturbation theory in Fourier space.
6. Alvarez, G.: Coupling-constant behavior of the resonances of the cubic anharmonic oscillator, *Phys. Rev. A* **37** (1988), 4079–4083. Beyond-all-orders perturbation theory in quantum mechanics.
7. Arnold, V. I.: *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1978. Quote about why series diverge: p. 395.
8. Arteca, G. A., Fernandez, F. M. and Castro, E. A.: *Large Order Perturbation Theory and Summation Methods in Quantum Mechanics*, Springer-Verlag, New York, 1990, p. 642; beyond-all-orders perturbation theory.
9. Baker, G. A., Jr. and Graves-Morris, P.: *Pade Approximants*, Cambridge University Press, New York, 1996.

10. Baker, G. A., Jr., Oitmaa, J. and Velgakis, M. J.: Series analysis of multivalued functions, *Phys. Rev. A* **38** (1988), 5316–5331. Generalization of Padé approximants; the power series for a function $u(z)$, known only through its series, is used to define the polynomial coefficients of a differential equation, whose solution is then used as an approximation to u.
11. Balian, R., Parisi, G. and Voros, A.: Quartic oscillator, in: S. Albeverio, P. Combe, R. Hoegh-Krohn, G. Rideau, M. Siruge-Collin, M. Sirugue and R. Stora (eds), *Feynman Path Integrals*, Lecture Notes in Phys. 106, Springer-Verlag, New York, 1979, pp. 337–360.
12. Balser, W.: A different characterization of multisummable power series, *Analysis* **12** (1992), 57–65.
13. Balser, W.: Summation of formal power series through iterated Laplace integrals, *Math. Scand.* **70** (1992), 161–171.
14. Balser, W.: Addendum to my paper: A different characterization of multisummable power series, *Analysis* **13** (1993), 317–319.
15. Balser, W.: *From Divergent Power Series to Analytic Functions*, Lecture Notes in Math. 1582, Springer-Verlag, New York, 1994, p. 100; good presentation of Gevrey order and asymptotics and multisummability.
16. Balser, W., Braaksma, B. L. J., Ramis, J.-P. and Sibuya, Y.: Multisummability of formal power series of linear ordinary differential equations, *Asymptotic Anal.* **5** (1991), 27–45.
17. Balser, W. and Tovbis, A.: Multisummability of iterated integrals, *Asymptotic Anal.* **7** (1992), 121–127.
18. Benassi, L., Grecchi, V., Harrell, E. and Simon, B.: Bender–Wu formula and the Stark effect in hydrogen, *Phys. Rev. Lett.* **42** (1979), 704–707. Exponentially small corrections in quantum mechanics.
19. Bender, C. M. and Orszag, S. A.: *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York, 1978, p. 594.
20. Bender, C. M. and Wu, T. T.: Anharmonic oscillator, *Phys. Rev.* **184** (1969), 1231–1260.
21. Bender, C. M. and Wu, T. T.: Anharmonic oscillator. II. A study of perturbation theory in large order, *Phys. Rev. D* **7** (1973), 1620–1636.
22. Benilov, E., Grimshaw, R. H. and Kuznetsova, E.: The generation of radiating waves in a singularly perturbed Korteweg–de Vries equation, *Physica D* **69** (1993), 270–276.
23. Berman, A. S.: Laminar flow in channel with porous walls, *J. Appl. Phys.* **24** (1953), 1232–1235. Earliest paper on an ODE (Berman–Robinson–Terrill problem) where exponentially small corrections are important.
24. Berry, M. V.: Stokes' phenomenon; smoothing a Victorian discontinuity, *Publ. Math. IHES* **68** (1989), 211–221.
25. Berry, M. V.: Uniform asymptotic smoothing of Stokes's discontinuities, *Proc. Roy. Soc. London A* **422** (1989), 7–21.
26. Berry, M. V.: Waves near Stokes lines, *Proc. Roy. Soc. London A* **427** (1990), 265–280.
27. Berry, M. V.: Histories of adiabatic quantum transitions, *Proc. Roy. Soc. London A* **429** (1990), 61–72.
28. Berry, M. V.: Infinitely many Stokes smoothings in the Gamma function, *Proc. Roy. Soc. London A* **434** (1991), 465–472.
29. Berry, M. V.: Stokes phenomenon for superfactorial asymptotic series, *Proc. Roy. Soc. London A* **435** (1991), 437–444.
30. Berry, M. V.: Asymptotics, superasymptotics, hyperasymptotics, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 1–14.
31. Berry, M. V.: Faster than Fourier, in: J. S. Auandan and J. L. Safko (eds), *Celebration of the 60th Birthday of Yakir Aharonov*, World Scientific, Singapore, 1994.
32. Berry, M. V.: Evanescent and real waves in quantum billiards and Gaussian beams, *J. Phys. A* **27** (1994), L391–L398.

33. Berry, M. V.: Asymptotics, singularities and the reduction of theories, in: D. Prawitz, B. Skyrms and D. Westerstahl (eds), *Logic, Methodology and Philosophy of Science IX*, Elsevier, Amsterdam, 1994, pp. 597–607.
34. Berry, M. V.: Riemann–Siegel expansion for the zeta function: High orders and remainders, *Proc. Roy. Soc. London A* **450** (1995), 439–462. Beyond all orders asymptotics.
35. Berry, M. V. and Howls, C. J.: Hyperasymptotics, *Proc. Roy. Soc. London A* **430** (1990), 653–668.
36. Berry, M. V. and Howls, C. J.: Stokes surfaces of diffraction catastrophes with codimension three, *Nonlinearity* **3** (1990), 281–291.
37. Berry, M. V. and Howls, C. J.: Hyperasymptotics for integrals with saddles, *Proc. Roy. Soc. London A* **434** (1991), 657–675.
38. Berry, M. V. and Howls, C. J.: Unfolding the high orders of asymptotic expansions with coalescing saddles: Singularity theory, crossover and duality, *Proc. Roy. Soc. London A* **443** (1993), 107–126.
39. Berry, M. V. and Howls, C. J.: Infinity interpreted, *Physics World* (1993), 35–39.
40. Berry, M. V. and Howls, C. J.: Overlapping Stokes smoothings: Survival of the error function and canonical catastrophe integrals, *Proc. Roy. Soc. London A* **444** (1994), 201–216.
41. Berry, M. V. and Howls, C. J.: High orders of the Weyl expansion for quantum billiards: Resurgence of periodic orbits, and the Stokes phenomenon, *Proc. Roy. Soc. London A* **447** (1994), 527–555.
42. Berry, M. V. and Lim, R.: Universal transition prefactors derived by superadiabatic renormalization, *J. Phys. A* **26** (1993), 4737–4747.
43. Bhattacharyya, K.: Notes on polynomial perturbation problems, *Chem. Phys. Lett.* **80** (1981), 257–261.
44. Bhattacharyya, K. and Bhattacharyya, S. P.: The sign-change argument revisited, *Chem. Phys. Lett.* **76** (1980), 117–119. Criterion for divergence of asymptotic series.
45. Bhattacharyya, K. and Bhattacharyya, S. P.: Reply to "another attack on the sign-change argument", *Chem. Phys. Lett.* **80** (1981), 604–605.
46. Boasman, P. A. and Keating, J. P.: Semiclassical asymptotics of perturbed cat maps, *Proc. Roy. Soc. London A* **449** (1995), 625–653. Shows that the optimal truncation of the semiclassical expansion is accurate to within an error which is an exponential function of $1/\hbar$. Stokes phenomenon, Borel resummation, and a universal approximation to the late terms are used beyond the superasymptotic approximation.
47. Boyd, J. P.: A Chebyshev polynomial method for computing analytic solutions to eigenvalue problems with application to the anharmonic oscillator, *J. Math. Phys.* **19** (1978), 1445–1456.
48. Boyd, J. P.: The nonlinear equatorial Kelvin wave, *J. Phys. Oceanogr.* **10** (1980), 1–11.
49. Boyd, J. P.: The rate of convergence of Chebyshev polynomials for functions which have asymptotic power series about one endpoint, *Math. Comput.* **37** (1981), 189–196.
50. Boyd, J. P.: The optimization of covergence for Chebyshev polynomial methods in an unbounded domain, *J. Comput. Phys.* **45** (1982), 43–79. Infinite and semi-infinite intervals; guidelines for choosing the map parameter or domain size L.
51. Boyd, J. P.: A Chebyshev polynomial rate-of-covergence theorem for Stieltjes functions, *Math. Comput.* **39** (1982), 201–206. Typo: In (24), the rightmost expression should be $1-(r+\varepsilon)/2$.
52. Boyd, J. P.: The asymptotic coefficients of Hermite series, *J. Comput. Phys.* **54** (1984), 382–410.
53. Boyd, J. P.: Spectral methods using rational basis functions on an infinite interval, *J. Comput. Phys.* **69** (1987), 112–142.
54. Boyd, J. P.: Orthogonal rational functions on a semi-infinite interval, *J. Comput. Phys.* **70** (1987), 63–88.

55. Boyd, J. P.: *Chebyshev and Fourier Spectral Methods*, Springer-Verlag, New York, 1989, p. 792.
56. Boyd, J. P.: New directions in solitons and nonlinear periodic waves: Polycnoidal waves, imbricated solitons, weakly non-local solitary waves and numerical boundary value algorithms, in: T.-Y. Wu and J. W. Hutchinson (eds), *Advances in Applied Mechanics*, Adv. in Appl. Mech. 27, Academic Press, New York, 1989, pp. 1–82.
57. Boyd, J. P.: Non-local equatorial solitary waves, in: J. C. J. Nihoul and B. M. Jamart (eds), *Mesoscale/Synoptic Coherent Structures in Geophysical Turbulence: Proc. 20th Liége Coll. on Hydrodynamics*, Elsevier, Amsterdam, 1989, pp. 103–112. Typo: In (4.1b), 0.8266 should be 1.6532.
58. Boyd, J. P.: A numerical calculation of a weakly non-local solitary wave: the ψ^4 breather, *Nonlinearity* **3** (1990), pp. 177–195. The eigenfunction calculation (5.15, etc.) has some typographical errors corrected in Chapter 12 of Boyd (1998).
59. Boyd, J. P.: The envelope of the error for Chebyshev and Fourier interpolation, *J. Sci. Comput.* **5** (1990), 311–363.
60. Boyd, J. P.: A Chebyshev/radiation function pseudospectral method for wave scattering, *Computers in Physics* **4** (1990), 83–85. Numerical calculation of exponentially small reflection.
61. Boyd, J. P.: A comparison of numerical and analytical methods for the reduced wave equation with multiple spatial scales, *Appl. Numer. Math.* **7** (1991), 453–479. Study of $u_{xx} \pm u_x = f(\varepsilon x)$. Typo: ε^{2n} factor should be omitted from Equation (4.3).
62. Boyd, J. P.: Weakly non-local solitons for capillary-gravity waves: Fifth-degree Korteweg–de Vries equation, *Physica D* **48** (1991), 129–146. Typo: at the beginning of Section 5, 'Newton–Kantorovich (5.1)' should read 'Newton–Kantorovich (3.2)'. Also, in the caption to Figure 12, '500,000' should be '70,000'.
63. Boyd, J. P.: Chebyshev and Legendre spectral methods in algebraic manipulation languages, *J. Symbolic Comput.* **16** (1993), 377–399.
64. Boyd, J. P.: The rate of convergence of Fourier coefficients for entire functions of infinite order with application to the Weideman–Cloot sinh-mapping for pseudospectral computations on an infinite interval, *J. Comput. Phys.* **110** (1994), 360–372.
65. Boyd, J. P.: The slow manifold of a five mode model, *J. Atmos. Sci.* **51** (1994), 1057–1064.
66. Boyd, J. P.: Time-marching on the slow manifold: The relationship between the nonlinear Galerkin method and implicit timestepping algorithms, *Appl. Math. Lett.* **7** (1994), 95–99.
67. Boyd, J. P.: Weakly nonlocal envelope solitary waves: Numerical calculations for the Klein–Gordon (ψ^4) equation, *Wave Motion* **21** (1995), 311–330.
68. Boyd, J. P.: A hyperasymptotic perturbative method for computing the radiation coefficient for weakly nonlocal solitary waves, *J. Comput. Phys.* **120** (1995), 15–32.
69. Boyd, J. P.: Eight definitions of the slow manifold: Seiches, pseudoseiches and exponential smallness, *Dyn. Atmos. Oceans* **22** (1995), 49–75.
70. Boyd, J. P.: A lag-averaged generalization of Euler's method for accelerating series, *Appl. Math. Comput.* **72** (1995), 146–166.
71. Boyd, J. P.: Multiple precision pseudospectral computations of the radiation coefficient for weakly nonlocal solitary waves: Fifth-order Korteweg–de Vries equation, *Computers in Physics* **9** (1995), 324–334.
72. Boyd, J. P.: *Weakly Nonlocal Solitary Waves and Beyond-All-Orders Asymptotics: Generalized Solitons and Hyperasymptotic Perturbation Theory*, Math. Appl. 442, Kluwer Acad. Publ., Dordrecht, 1998, p. 608.
73. Boyd, J. P.: *Chebyshev and Fourier Spectral Methods*, Dover, New York, 1999. Second edition of Boyd (1989a), in press.

74. Boyd, J. P. and Christidis, Z. D.: Low wavenumber instability on the equatorial beta-plane, *Geophys. Res. Lett.* **9** (1982), 769–772. Growth rate is exponentially in $1/\varepsilon$ where ε is the shear strength.
75. Boyd, J. P. and Christidis, Z. D.: Instability on the equatorial beta-plane, in: J. Nihoul (ed.), *Hydrodynamics of the Equatorial Ocean*, Elsevier, Amsterdam, 1983, pp. 339–351.
76. Boyd, J. P. and Natarov, A.: A Sturm–Liouville eigenproblem of the fourth kind: A critical latitude with equatorial trapping, *Stud. Appl. Math.* **101** (1998), 433–455.
77. Boyd, W. G. C.: Stieltjes transforms and the Stokes phenomenon, *Proc. Roy. Soc. London A* **429** (1990), 227–246.
78. Boyd, W. G. C.: Error bounds for the method of steepest descents, *Proc. Roy. Soc. London A* **440** (1993), 493–516.
79. Boyd, W. G. C.: Gamma function asymptotics by an extension of the method of steepest descents, *Proc. Roy. Soc. London A* **447** (1993), 609–630.
80. Boyd, W. G. C.: Steepest-descent integral representations for dominant solutions of linear second-order differential equations, *Methods Appl. Anal.* **3** (1996), 174–202.
81. Braaksma, B. L. J.: Multisummability and Stokes multipliers of linear meromorphic differential equations, *Ann. Inst. Fourier (Grenoble)* **92** (1991), 45–75.
82. Braaksma, B. L. J.: Multisummability of formal power-series solutions of nonlinear meromorphic differential equations, *Ann. Inst. Fourier (Grenoble)* **42** (1992), 517–540. Proves a theorem of Écalle that formal power series of nonlinear meromorphic differential equations are multisummable.
83. Braaksma, B. L. J. (ed.): *The Stokes Phenomenon and Hilbert's Sixteenth Problem: Groningen, The Netherlands, 31 May–3 June 1995*, World Scientific, Singapore, 1996.
84. Branis, S. V., Martin, O. and Birman, J. L.: Self-induced transparency selects discrete velocities for solitary-wave solutions, *Phys. Rev. A* **43** (1991), 1549–1563. Nonlocal envelope solitons.
85. Bulakh, B. M.: On higher approximations in the boundary-layer theory, *J. Appl. Math. Mech.* **28** (1964), 675–681.
86. Byatt-Smith, J. G. B.: On solutions of ordinary differential equations arising from a model of crystal growth, *Stud. Appl. Math.* **89** (1993), 167–187.
87. Byatt-Smith, J. G. B.: Formulation and summation of hyperasymptotic expansions obtained from integrals, *European J. Appl. Math.* **9** (1998), 159–185.
88. Byatt-Smith, J. G. B. and Davie, A. M.: Exponentially small oscillations in the solution of an ordinary differential equation, *Proc. Roy. Soc. Edinburgh A* **114** (1990), 243.
89. Byatt-Smith, J. G. B. and Davie, A. M.: Exponentially small oscillations in the solution of an ordinary differential equation, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 223–240.
90. Candelpergher, B., Nosmas, J. C. and Pham, F.: Introduction to Écalle alien calculus, *Ann. Inst. Fourier (Grenoble)* **43** (1993), 201–224. Review. The "alien calculus" is a systematic theory for resurgence and Borel summability to generate hyperasymptotic approximation. Written in French.
91. Canuto, C., Hussaini, M. Y., Quarteroni, A. and Zang, T. A.: *Spectral Methods for Fluid Dynamics*, Springer-Verlag, New York, 1987.
92. Carr, J.: Slowly varying solutions of a nonlinear partial differential equation, in: D. S. Broomhead and A. Iserles (eds), *The Dynamics of Numerics and the Numerics of Dynamics*, Oxford University Press, Oxford, 1992, pp. 23–30.
93. Carr, J. and Pego, R. L.: Metastable patterns in solutions of $u_t = \varepsilon^2 u_{xx} - f(u)$, *Comm. Pure Appl. Math.* **42** (1989), 523–576. Merger of fronts on an exponentially slow time scale.
94. Chang, Y.-H.: Proof of an asymptotic symmetry of the rapidly forced pendulum, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 213–221.

95. Chapman, S. J.: On the non-universality of the error function in the smoothing of Stokes discontinuities, *Proc. Roy. Soc. London A* **452** (1996), 2225–2230.
96. Chapman, S. J., King, J. R. and Adams, K. L.: Exponential asymptotics and Stokes lines in nonlinear ordinary differential equations, *Proc. Roy. Soc. London A* (1998). To appear.
97. Chen, J., Fisher, M. E. and Nickel, B. G.: Unbiased estimation of corrections to scaling by partial differential approximants, *Phys. Rev. Lett.* **48** (1982), 630–634. Generalization of Padé approximants.
98. Chester, W. and Breach, D. R.: On the flow past a sphere at low Reynolds number, *J. Fluid Mech.* **37** (1969), 751–760. Log-and-power series.
99. Ciasullo, L. M. and Cochran, J. A.: Accelerating the convergence of Chebyshev series, in: R. Wong (ed.), *Asymptotic and Computational Analysis*, Marcel Dekker, New York, 1990, pp. 95–136.
100. Cizek, J., Damburg, R. J., Graffi, S., Grecchi, V., II, E. M. H., Harris, J. G., Nakai, S., Paldus, J., Propin, R. K. and Silverstone, H. J.: $1/R$ expansion for H_2^+: Calculation of exponentially small terms and asymptotics, *Phys. Rev. A* **33** (1986), 12–54.
101. Cizek, J. and Vrscay, E. R.: Large order perturbation theory in the context of atomic and molecular physics – interdisciplinary aspects, *Int. J. Quantum Chem.* **21** (1982), 27–68.
102. Cloot, A. and Weideman, J. A. C.: An adaptive algorithm for spectral computations on unbounded domains, *J. Comput. Phys.* **102** (1992), 398–406.
103. Combescot, R., Dombe, T., Hakim, V. and Pomeau, Y.: Shape selection of Saffman–Taylor fingers, *Phys. Rev. Lett.* **56** (1986), 2036–2039.
104. Costin, O.: Exponential asymptotics, trans-series and generalized Borel summation for analytic nonlinear rank one systems of ODE's, *Internat. Math. Res. Notices* **8** (1995), 377–418.
105. Costin, O.: On Borel summation and Stokes phenomenon for nonlinear rank one systems of ODE's, *Duke Math. J.* **93** (1998), 289–344. Connections with Berry smoothing and Écalle resurgence.
106. Costin, O. and Kruskal, M. D.: Optimal uniform estimates and rigorous asymptotics beyond all orders for a class of ordinary differential equations, *Proc. Roy. Soc. London A* **452** (1996), 1057–1085.
107. Costin, O. and Kruskal, M. On optimal truncation of divergent series solutions of nonlinear differential systems; Berry smoothing, *Proc. Roy. Soc. London A* **452** (1998). Submitted. Rigorous proofs of some assertions and conclusions in the smoothing of discontinuities in Stokes phenomenon.
108. Cox, S. M.: Two-dimensional flow of a viscous fluid in a channel with porous walls, *J. Fluid Mech.* **227** (1991), 1–33. Multiple solutions differing by exponentially small terms.
109. Cox, S. M. and King, A. C.: On the asymptotic solution of a high-order nonlinear ordinary differential equation, *Proc. Roy. Soc. London A* **453** (1997), 711–728. Berman–Terrill–Robinson problem with good review of earlier work.
110. Darboux, M. G.: Mémoire sur l'approximation des fonctions de très-grands nombres, et sur une classe étendue de développements en série, *J. Math. Pures Appl.* **4** (1878), 5–56.
111. Darboux, M. G.: Mèmoire sur l'approximation des fonctions de trés-grands nombres, et sur une classe étendue de développements en série, *J. Math. Pures Appl.* **4** (1878), 377–416.
112. Delabaere, E.: Introduction to the Écalle theory, in: E. Tournier (ed.), *Computer Algebra and Differential Equations*, London Math. Soc. Lecture Notes Ser. 193, Cambridge University Press, Cambridge, 1994, pp. 59–102.
113. Delabaere, E. and Pham, F.: Unfolding the quartic oscillator, *Ann. Phys.* **261** (1997), 180–218. Resurgence and "exact WKB methods"; confirm the branch structure found by Bender and Wu.
114. Dias, F., Menasce, D. and Vanden-Broeck, J.-M.: Numerical study of capillary-gravity solitary waves, *Eur. J. Mech. B Fluids* **15** (1996), 17–36.

115. Dickinson, R. E.: Numerical versus analytical methods for a sixth order hypergeometric equation arising in a diffusion-wave theory of the quasi-biennial oscillation QBO. Seminar, 1980.
116. Dingle, R. B.: Asymptotic expansions and converging factors I. General theory and basic converging factors, *Proc. Roy. Soc. London A* **244** (1958), 456–475.
117. Dingle, R. B.: Asymptotic expansions and converging factors IV. Confluent hypergeometric, parabolic cylinder, modified Bessel and ordinary Bessel functions, *Proc. Roy. Soc. London A* **249** (1958), 270–283.
118. Dingle, R. B.: *Asymptotic Expansions: Their Derivation and Interpretation*, Academic Press, New York, 1973. Beyond all orders asymptotics.
119. Dumas, H. S.: Existence and stability of particle channeling in crystals on timescales beyond all orders, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 267–273.
120. Dumas, H. S.: A Nekhoroshev-like theory of classical particle channeling in perfect crystals, *Dynamics Reported* **2** (1993), 69–115. Beyond all orders perturbation theory in crystal physics.
121. Dunster, T. M.: Error bounds for exponentially improved asymptotic solutions of ordinary differential equations having irregular singularities of rank one, *Methods Appl. Anal.* **3** (1996), 109–134.
122. Dyson, F. J.: Divergence of perturbation theory in quantum electrodynamics, *Phys. Rev.* **85** (1952), 631–632.
123. Écalle, J.: *Les fonctions résurgentes*, Université de Paris-Sud, Paris, 1981. Three volumes, Earliest systematic development of resurgence theory.
124. Elliott, D.: The evaluation and estimation of the coefficients in the Chebyshev series expansion of a function, *Math. Comput.* **18** (1964), 274–284. This and the next two papers are classic contributions to the asymptotic theory of Chebyshev coefficients.
125. Elliott, D.: Truncation errors in two Chebyshev series approximations, *Math. Comput.* **19** (1965), 234–248. Errors in Lagrangian interpolation with a general contour integral representation and an exact analytical formula for $1/(a + x)$.
126. Elliott, D. and Szekeres, G.: Some estimates of the coefficients in the Chebyshev expansion of a function, *Math. Comput.* **19** (1965), 25–32. The Chebyshev coefficients are exponentially small in the degree n.
127. Elliott, D. and Tuan, P. D.: Asymptotic coefficients of Fourier coefficients, *SIAM J. Math. Anal.* **5** (1974), 1–10.
128. Fröman, N.: The energy levels of double-well potentials, *Ark. Fysik* **32** (1966), 79–96. WKB method for exponentially small splitting of eigenvalue degeneracy.
129. Frost, P. A. and Harper, E. Y.: An extended Padé procedure for constructing global approximations from asymptotic expansions: an explication with examples, *SIAM Rev.* **18** (1976), 62–91.
130. Fusco, G. and Hale, J. K.: Slow motion manifolds, dormant instability and singular perturbations, *J. Dynamics Differential Equations* **1** (1989), 75–94. Exponentially slow frontal motion.
131. Germann, T. C. and Kais, S.: Large order dimensional perturbation theory for complex energy eigenvalues, *J. Chem. Phys.* **99** (1993), 7739–7747. Quadratic Shafer–Padé approximants, applied to compute imaginary part of eigenvalue.
132. Gingold, H. and Hu, J.: Transcendentally small reflection of waves for problems with/without turning points near infinity: A new uniform approach, *J. Math. Phys.* **32** (1991), 3278–3284. Generalized WBK (Liouville–Green) for above-the-barrier scattering.
133. Grasman, J. and Matkowsky, B. J.: A variational approach to singularly perturbed boundary value problems for ordinary and partial differential equations with turning points, *SIAM J. Appl. Math.* **32** (1976), 588–597. Resolve the failure of standard matched asymptotics for

the problem of Ackerberg and O'Malley (1970) by applying a non-perturbative variational principle; MacGillivray (1997) solves the same problem by incorporating exponentially small terms into matched asymptotics.

134. Grimshaw, R. H. J. and Joshi, N.: Weakly non-local solitary waves in a singularly-perturbed Korteweg–de Vries equation, *SIAM J. Appl. Math.* **55** (1995), 124–135.
135. Grundy, R. E. and Allen, H. R.: The asymptotic solution of a family of boundary value problems involving exponentially small terms, *IMA J. Appl. Math.* **53** (1994), 151–168.
136. Hakim, V. and Mallick, K.: Exponentially small splitting of separatrices, matching in the complex plane and Borel summation, *Nonlinearity* **6** (1993), 57–70. Very readable analysis.
137. Hale, J. K.: Dynamics and numerics, in: D. S. Broomhead and A. Iserles (eds), *The Dynamics of Numerics and the Numerics of Dynamics*, Oxford University Press, Oxford, 1992, pp. 243–254.
138. Hanson, F. B.: Singular point and exponential analysis, in: R. Wong (ed.), *Asymptotic and Computational Analysis*, Marcel Dekker, New York, 1990, pp. 211–240.
139. Hardy, G. H.: *Divergent Series*, Oxford University Press, New York, 1949.
140. Harrell, E. and Simon, B.: The mathematical theory of resonances whose widths are exponentially small, *Duke Math. J.* **47** (1980), 845.
141. Harrell, E. M.: On the asymptotic rate of eigenvalue degeneracy, *Comm. Math. Phys.* **60** (1978), 73–95.
142. Harrell, E. M.: Double wells, *Comm. Math. Phys.* **75** (1980), 239–261. Exponentially small splitting of eigenvalues.
143. Hildebrand, F. H.: *Introduction to Numerical Analysis*, Dover, New York, 1974. Numerical; asymptotic-but-divergent series in h for errors.
144. Hinton, D. B. and Shaw, J. K.: Absolutely continuous spectra of 2d order differential operators with short and long range potentials, *Quart. J. Math.* **36** (1985), 183–213. Exponential smallness in eigenvalues.
145. Holmes, P., Marsden, J. and Scheurle, J.: Exponentially small splitting of separatrices with applications to KAM theory and degenerate bifurcations, *Contemp. Math.* **81** (1988), 214–244.
146. Hong, D. C. and Langer, J. S.: Analytic theory of the selection mechanism in the Saffman–Taylor problem, *Phys. Rev. Lett.* **56** (1986), 2032–2035.
147. Howls, C. J.: Hyperasymptotics for multidimensional integrals, exact remainder terms and the global connection problem, *Proc. Roy. Soc. London A* **453** (1997), 2271–2294.
148. Howls, C. J. and Trasler, S. A.: Weyl's wedges, *J. Phys. A: Math. Gen.* **31** (1998), 1911–1928. Hyperasymptotics for quantum billiards with nonsmooth boundary.
149. Hu, J.: Asymptotics beyond all orders for a certain type of nonlinear oscillator, *Stud. Appl. Math.* **96** (1996), 85–109.
150. Hu, J. and Kruskal, M.: Reflection coefficient beyond all orders for singular problems, 1, separated critical-points on the nearest critical-level line, *J. Math. Phys.* **32** (1991), 2400–2405.
151. Hu, J. and Kruskal, M.: Reflection coefficient beyond all orders for singular problems, 2, close-spaced critical-points on the nearest critical-level line, *J. Math. Phys.* **32** (1991), 2676–2678.
152. Hu, J. and Kruskal, M. D.: Reflection coefficient beyond all orders for singular problems, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 247–253.
153. Hunter, J. K. and Scheurle, J.: Existence of perturbed solitary wave solutions to a model equation for water waves, *Physica D* **32** (1988), 253–268. FKdV nonlocal solitons.
154. Jardine, M., Allen, H. R., Grundy, R. E. and Priest, E. R.: A family of two-dimensional nonlinear solutions for magnetic field annihilation, *J. Geophys. Res. – Space Physics* **97** (1992), 4199–4207.

155. Jones, D. S.: Uniform asymptotic remainders, in: R. Wong (ed.), *Asymptotic Comput. Anal.*, Marcel Dekker, New York, 1990, pp. 241–264.
156. Jones, D. S.: Asymptotic series and remainders, in: B. D. Sleeman and R. J. Jarvis (eds), *Ordinary and Partial Differential Equations*, Volume IV, Longman, London, 1993, p. 12.
157. Jones, D. S.: Asymptotic remainders, *SIAM J. Math. Anal.* **25** (1994), 474–490. Hyperasymptotics; shows that the remainders in a variety of asymptotic series can be uniformly approximated by the same integral.
158. Jones, D. S.: *Introduction to Asymptotics: A Treatment Using Nonstandard Analysis*, World Scientific, Singapore, 1997, 160 pages; includes a chapter on hyperasymptotics.
159. Kaplun, S.: Low Reynolds number flow past a circular cylinder, *J. Math. Mech.* **6** (1957), 595–603. Log-plus-power expansions.
160. Kaplun, S.: *Fluid Mechanics and Singular Perturbations*, Academic Press, New York, 1967. Ed. by P. A. Lagerstrom, L. N. Howard and C. S. Liu; Analyzed difficulties of log-plus-power expansions.
161. Kaplun, S. and Lagerstrom, P. A.: Asymptotic expansions of Navier–Stokes solutions for small Reynolds number, *J. Math. Mech.* **6** (1957), 585–593.
162. Kath, W. L and Kriegsmann, G. A.: Optical tunnelling: Radiation losses in bent fibre-optic waveguides, *IMA J. Appl. Math.* **41** (1988), 85–103. Radiation loss is exponentially small in the small parameter, so “beyond all orders” perturbation theory is developed here.
163. Kessler, D. A., Koplik, J. and Levine, H.: Pattern selection in fingered growth phenomena, *Adv. Phys.* **37** (1988), 255–339.
164. Killingbeck, J.: Quantum-mechanical perturbation theory, *Reports on Progress in Theoretical Physics* **40** (1977), 977–1031. Divergence of asymptotic series.
165. Killingbeck, J.: A polynomial perturbation problem, *Phys. Lett. A* **67** (1978), 13–15.
166. Killingbeck, J.: Another attack on the sign-change argument, *Chem. Phys. Lett.* **80** (1981), 601–603.
167. Kivshar, Y. S. and Malomed, B. A.: Comment on ‘nonexistence of small amplitude breather solutions in ψ^4 theory’, *Phys. Rev. Lett.* **60** (1988), 164–164. Exponentially small radiation from perturbed from perturbed sine-Gordon solitons and other species.
168. Kivshar, Y. S. and Malomed, B. A.: Dynamics of solitons in nearly integrable systems, *Revs. Mod. Phys.* **61** (1989), 763–915. Exponentially small radiation from perturbed sine-Gordon solitons and other species.
169. Kowalenko, V., Glasser, M. L., Taucher, T. and Frankel, N. E.: *Generalised Euler–Jacobi Inversion Formula and Asymptotics Beyond All Orders*, London Math. Soc. Lecture Note Ser. 214, Cambridge University Press, Cambridge, 1995.
170. Kropinski, M. C. A., Ward, M. J. and Keller, J. B.: A hybrid asymptotic-numerical method for low Reynolds number flows past a cylindrical body, *SIAM J. Appl. Math.* **55** (1995), 1484–1510. Log-and-power series in Re.
171. Kruskal, M. D. and Segur, H.: Asymptotics beyond all orders in a model of crystal growth, Tech. Rep. 85-25, Aeronautical Research Associates of Princeton, 1985.
172. Kruskal, M. D. and Segur, H.: Asymptotics beyond all orders in a model of crystal growth, *Stud. Appl. Math.* **85** (1991), 129–181.
173. Laforgue, J. G. L. and O’Malley, R. E., Jr.: Supersensitive boundary value problems, in: H. G. Kaper and M. Garbey (eds), *Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters*, Kluwer Acad. Publ., Dordrecht, 1993, pp. 215–223.
174. Laforgue, J. G. L. and O’Malley, R. E., Jr.: On the motion of viscous shocks and the supersensitivity of their steady-state limits, *Methods Appl. Anal.* **1** (1994), 465–487. Exponential smallness in shock movement.
175. Laforgue, J. G. L. and O’Malley, R. E., Jr.: Shock layer movement for Burgers’ equation, *SIAM J. Appl. Math.* **55** (1995), 332–347.

176. Laforgue, J. G. L. and O'Malley, R. E., Jr.: Viscous shock motion for advection-diffusion equation, *Stud. Appl. Math.* **95** (1995), 147–170.
177. Lanczos, C.: Trigonometric interpolation of empirical and analytical functions, *J. Math. Phys.* **17** (1938), 123–199. The origin of both the pseudospectral method and the tau method. Lanczos is to spectral methods what Newton was to calculus.
178. Lanczos, C.: *Discourse on Fourier Series*, Oliver and Boyd, Edinburgh, 1966.
179. Lange, C. G. and Weinitschke, H. J.: Singular perturbations of elliptic problems on domains with small holes, *Stud. Appl. Math.* **92** (1994), 55–93. Log-and-power series for eigenvalues with comparisons with numerical solutions; demonstrates the surprisingly large sensitivity of eigenvalues to small holes in the membrane.
180. Lazutkin, V. F., Schachmannski, I. G. and Tabanov, M. B.: Splitting of separatrices for standard and semistandard mappings, *Physica D* **40** (1989), 235–248.
181. Le Guillou, J. C. and Zinn-Justin, J. (eds): *Large-Order Behaviour of Perturbation Theory*, North-Holland, Amsterdam, 1990. Exponential corrections to power series, mostly in quantum mechanics.
182. Lim, R. and Berry, M. V.: Superadiabatic tracking for quantum evolution, *J. Phys. A* **24** (1991), 3255–3264.
183. Liu, J. and Wood, A.: Matched asymptotics for a generalisation of a model equation for optical tunnelling, *Europ. J. Appl. Math.* **2** (1991), 223–231. Compute the exponentially small imaginary part of the eigenvalue λ, $\Im(\lambda) \sim \exp(-1/\varepsilon^{1/n})$, for the problem $u_{xx} + (\lambda + \varepsilon x^n)u = 0$ with outward radiating waves on the semi-infinite interval.
184. Lorenz, E. N. and Krishnamurthy, V.: On the nonexistence of a slow manifold, *J. Atmos. Sci.* **44** (1987), 2940–2950. Weakly non-local in time.
185. Lozano, C. and Meyer, R. E.: Leakage and response of waves trapped by round islands, *Phys. Fluids* **19** (1976), 1075–1088. Leakage is exponentially small in the perturbation parameter.
186. Lu, C., MacGillivray, A. D. and Hastings, S. P.: Asymptotic behavior of solutions of a similarity equation for laminar flows in channels with porous walls, *IMA J. Appl. Math.* **49** (1992), 139–162. Beyond-all-orders matched asymptotics.
187. Luke, Y. L.: *The Special Functions and Their Approximations*, Vols I and II, Academic Press, New York, 1969.
188. Lyness, J. N.: Adjusted forms of the Fourier coefficient asymptotic expansion, *Math. Comput.* **25** (1971), 87–104.
189. Lyness, J. N.: The calculation of trigonometric Fourier coefficients, *J. Comput. Phys.* **54** (1984), 57–73. A good review article which discusses the integration-by-parts series for the asymptotic Fourier coefficients.
190. Lyness, J. N. and Ninham, B. W.: Numerical quadrature and asymptotic expansions, *Math. Comp.* **21** (1967), 162. Shows that the error is a power series in the grid spacing h plus an integral which is transcendentally small in $1/h$.
191. MacGillivray, A. D.: A method for incorporating transcendentally small terms into the method of matched asymptotic expansions, *Stud. Appl. Math.* **99** (1997), 285–310. Linear example has a general solution which is the sum of an antisymmetric function $A(x)$ which is well-approximated by a second-derivative-dropping outer expansion plus a symmetric part which is exponentially small except at the boundaries, and can be approximated only by a WKB method (with second derivative retained). His nonlinear example is the Carrier–Pearson problem whose exact solution is a KdV cnoidal wave, but required to satisfy Dirichlet boundary conditions. Matching fails because each soliton peak can be translated with only an exponentially small error; MacGillivray shows that the peaks, however many are fit between the boundaries, must be evenly spaced.
192. MacGillivray, A. D., Liu, B. and Kazarinoff, N. D.: Asymptotic analysis of the peeling-off point of a French duck, *Methods Appl. Anal.* **1** (1994), 488–509. Beyond-all-orders theory.

193. MacGillivray, A. D. and Lu, C.: Asymptotic solution of a laminar flow in a porous channel with large suction: A nonlinear turning point problem, *Methods Appl. Anal.* **1** (1994), 229–248. Incorporation of exponentially small terms into matched asymptotics.
194. Malomed, B. A.: Emission from, quasi-classical quantization, and stochastic decay of sine-Gordon solitons in external fields, *Physica D* **27** (1987), 113–157. Explicit calculations of exponentially small radiation.
195. Malomed, B. A.: Perturbation-induced radiative decay of solitons, *Phys. Lett. A* **123** (1987), 459–468. Explicit calculations of exponentially small radiation for perturbed sine-Gordon solitons, fluxons, kinks and coupled double-sine-Gordon equations.
196. Marion, M. and Témam, R.: Nonlinear Galerkin methods, *SIAM J. Numer. Anal.* **26** (19), 1139–1157.
197. Martin, O. and Branis, S. V.: Solitary waves in self-induced transparency, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 327–336.
198. Martinet, J. and Ramis, J.-P.: Elementary acceleration and multisummability-I, *Ann. Inst. H. Poincaré-Phys. Theor.* **54** (1991), 331–401.
199. Maslov, V. P.: *The Complex WKB Method for Nonlinear Equations I: Linear Theory*, Birkhäuser, Boston, 1994. Calculation of exponentially small terms.
200. McLeod, J. B.: Smoothing of Stokes discontinuities, *Proc. Roy. Soc. London A* **437** (1992), 343–354.
201. Meiss, J. D. and Horton, W.: Solitary drift waves in the presence of magnetic shear, *Phys. Fluids* **26** (1983), 990–997. Show that plasma modons leak radiation for large $|x|$, and therefore are nanopterons.
202. Meyer, R. E.: Adiabatic variation. Part I: Exponential property for the simple oscillator, *J. Appl. Math. Phys. ZAMP* **24** (1973), 517–524.
203. Meyer, R. E.: Adiabatic variation. Part II: Action change for simple oscillator, *J. Appl. Math. Phys. ZAMP* (1973).
204. Meyer, R. E.: Exponential action of a pendulum, *Bull. Amer. Math. Soc.* **80** (1974), 164–168.
205. Meyer, R. E.: Adiabatic variation. Part IV: Action change of a pendulum for general frequency, *J. Appl. Math. Phys. ZAMP* **25** (1974), 651–654.
206. Meyer, R. E.: Gradual reflection of short waves, *SIAM J. Appl. Math.* **29** (1975), 481–492.
207. Meyer, R. E.: Adiabatic variation. Part V: Nonlinear near-periodic oscillator, *J. Appl. Math. Phys. ZAMP* **27** (1976), 181–195.
208. Meyer, R. E.: Quasiclassical scattering above barriers in one dimension, *J. Math. Phys.* **17** (1976), 1039–1041.
209. Meyer, R. E.: Surface wave reflection by underwater ridges, *J. Phys. Oceangr.* **9** (1979), 150–157.
210. Meyer, R. E.: Exponential asymptotics, *SIAM Rev.* **22** (1980), 213–224.
211. Meyer, R. E.: Wave reflection and quasiresonance, in: *Theory and Application of Singular Perturbation*, Lecture Notes in Math. 942, Springer-Verlag, New York, 1982, pp. 84–112.
212. Meyer, R. E.: Quasiresonance of long life, *J. Math. Phys.* **27** (1986), 238–248.
213. Meyer, R. E.: A simple explanation of Stokes phenomenon, *SIAM Rev.* **31** (1989), 435–444.
214. Meyer, R. E.: Observable tunneling in several dimensions, in: R. Wong (ed.), *Asymptotic and Computational Analysis*, Marcel Dekker, New York, 1990, pp. 299–328.
215. Meyer, R. E.: On exponential asymptotics for nonseparable wave equations I: Complex geometrical optics and connection, *SIAM J. Appl. Math.* **51** (1991), 1585–1601.
216. Meyer, R. E.: On exponential asymptotics for nonseparable wave equations I: EBK quantization, *SIAM J. Appl. Math.* **51** (1991), 1602–1615.
217. Meyer, R. E.: Exponential asymptotics for partial differential equations, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Oreders*, Plenum, Amsterdam, 1991, pp. 29–36.

218. Meyer, R. E.: Approximation and asymptotics, in: D. A. Martin and G. R. Wickham (eds), *Wave Asymptotics*, Cambridge University Press, New York, 1992, pp. 43–53. Blunt and perceptive review.
219. Meyer, R. E. and Guay, E. J.: Adiabatic variation. Part III: A deep mirror model, *J. Appl. Math. Phys. ZAMP* **25** (1974), 643–650.
220. Meyer, R. E. and Painter, J. F.: Wave trapping with shore absorption, *J. Engg. Math.* **13** (1979), 33–45.
221. Meyer, R. E. and Painter, J. F.: New connection method across more general turning points, *Bull. Amer. Math. Soc.* **4** (1981), 335–338.
222. Meyer, R. E. and Painter, J. F.: Irregular points of modulation, *Adv. Appl. Math.* **4** (1982), 145–174.
223. Meyer, R. E. and Painter, J. F.: Connection for wave modulation, *SIAM J. Math. Anal.* **14** (1983), 450–462.
224. Meyer, R. E. and Painter, J. F.: On the Schrödinger connection, *Bull. Amer. Math. Soc.* **8** (1983), 73–76.
225. Meyer, R. E. and Shen, M. C.: On Floquet's theorem for nonseparable partial differential equations, in: B. D. Sleeman (ed.), *Eleventh Dundee Conference in Ordinary and Partial Differential Equations*, Pitman Adv. Math. Res. Notes, Longman–Wiley, New York, 1991, pp. 146–167.
226. Meyer, R. E. and Shen, M. C.: On exponential asymptotics for nonseparable wave equations III: Approximate spectral bands of periodic potentials on strips, *SIAM J. Appl. Math.* **52** (1992), 730–745.
227. Miller, G. F.: On the convergence of Chebyshev series for functions possessing a singularity in the range of representation, *SIAM J. Numer. Anal.* **3** (1966), 390–409.
228. Murphy, B. T. M. and Wood, A. D.: Exponentially improved asymptotic solutions of second order ordinary differential equations of arbitrary rank, *Methods Appl. Anal.* (1998).
229. Nayfeh, A. H.: *Perturbation Methods*, Wiley, New York, 1973. Good reference on the method of multiple scales.
230. Németh, G.: Polynomial approximation to the function $\psi(a, c, x)$, Technical Report, Central Institute for Physics, Budapest, 1965.
231. Németh, G.: Chebyshev expansion of Gauss' hypergeometric function, Technical Report, Central Institute for Physics, Budapest, 1965.
232. Németh, G.: Chebyshev expansions of the Bessel function. I, *Proceedings of the KFKI* **14** (1966), 157.
233. Németh, G.: Chebyshev expansions of the Bessel functions. II, *Proceedings of the KFKI* **14** (1966), 299–309.
234. Németh, G.: Note on the zeros of the Bessel functions, Technical Report, Central Institute for Physics, Budapest, 1969.
235. Németh, G.: Chebyshev series for special functions, Technical Report 74-13, Central Institute for Physics, Budapest, 1974.
236. Németh, G.: *Mathematical Approximation of Special Functions: Ten Papers on Chebyshev Expansions*, Nova Science Publishers, New York, 1992, p. 200. [Tables, with some theory and coefficient asymptotics, for Bessel functions, Airy functions, zeros of Bessel functions, and generalized hypergeometric functions.]
237. Olde Daalhuis, A. B.: Hyperasymptotic expansions of confluent hypergeometric functions, *IMA J. Appl. Math.* **49** (1992), 203–216.
238. Olde Daalhuis, A. B.: Hyperasymptotics and the Stokes phenomenon, *Proc. Roy. Math. Soc. Edinburgh A* **123** (1993), 731–743.
239. Olde Daalhuis, A. B.: Hyperasymptotic solutions of second-order linear differential equations II, *Methods Appl. Anal.* **2** (1995), 198–211.

240. Olde Daalhuis, A. B.: Hyperterminants I, *J. Comput. Appl. Math.* **76** (1996), 255–264. Convergent series for the generalized Stieltjes functions that appear in hyperasymptotic expansions.
241. Olde Daalhuis, A. B.: Hyperasymptotic solutions of higher order linear differential equations with a singularity of rank one, *Proc. Roy. Soc. London A* **454** (1997), 1–29. Borel–Laplace transform; new method to compute Stokes multipliers.
242. Olde Daalhuis, A. B.: Hyperterminants, II, *J. Comput. Appl. Math.* **89** (1998), 87–95. Convergent and computable expansions for hyperterminants so that these can be easily evaluated for use with hyperasymptotic perturbation theories. The expansions involve hypergeometric ($_2F_1$) functions, but these can be computed by recurrence.
243. Olde Daalhuis, A. B., Chapman, S. J., King, J. R., Ockendon, J. R. and Tew, R. H.: Stokes phenomenon and matched asymptotic expansions, *SIAM J. Appl. Math.* **6** (1995), 1469–1483.
244. Olde Daalhuis, A. B. and Olver, F. W. J.: Exponentially improved asymptotic solutions of ordinary differential equations. II. Irregular singularities of rank one, *Proc. Roy. Soc. London A* **445** (1994), 39–56.
245. Olde Daalhuis, A. B. and Olver, F. W. J.: Hyperasymptotic solutions of second-order linear differential equations. I, *Methods Appl. Anal.* **2** (1995), 173–197.
246. Olde Daalhuis, A. B. and Olver, F. W. J.: On the calculation of Stokes multipliers for linear second-order differential equations, *Methods Appl. Anal.* **2** (1995), 348–367.
247. Olde Daalhuis, A. B. and Olver, F. W. J.: Exponentially-improved asymptotic solutions of ordinary differential equations. II: Irregular singularities of rank one, *Proc. Roy. Soc. London A* **2** (1995), 39–56.
248. Olde Daalhuis, A. B. and Olver, F. W.: On the asymptotic and numerical solution of ordinary differential equations, *SIAM Rev.* **40** (1998). In press.
249. Olver, F. W. J.: *Asymptotics and Spectral Functions*, Academic Press, New York, 1974.
250. Olver, F. W. J.: On Stokes' phenomenon and converging factors, in: R. Wong (ed.), *Asymptotic and Computational Analysis*, Marcel Dekker, New York, 1990, pp. 329–356.
251. Olver, F. W. J.: Uniform, exponentially-improved asymptotic expansions for the generalized exponential integral, *SIAM J. Math. Anal.* **22** (1991), 1460–1474.
252. Olver, F. W. J.: Uniform, exponentially-improved asymptotic expansions for the confluent hypergeometric function and other integral transforms, *SIAM J. Math. Anal.* **22** (1991), 1475–1489.
253. Olver, F. W. J.: Exponentially-improved asymptotic solutions of ordinary differential equations I: The confluent hypergeometric function, *SIAM J. Math. Anal.* **24** (1993), 756–767.
254. Olver, F. W. J.: Asymptotic expansions of the coefficient in asymptotic series solutions of linear differential equations, *Methods Appl. Anal.* **1** (1994), 1–13.
255. Oppenheimer, J. R.: Three notes on the quantum theory of aperiodic effects, *Phys. Rev.* **31** (1928), 66–81. Shows that Zeeman effect generates an imaginary part to the energy which is exponentially small in the reciprocal of the perturbation parameter.
256. Paris, R. B.: Smoothing of the Stokes phenomenon for high-order differential equations, *Proc. Roy. Soc. London A* **436** (1992), 165–186.
257. Paris, R. B.: Smoothing of the Stokes phenomenon using Mellin–Barnes integrals, *J. Comput. Appl. Math.* **41** (1992), 117–133.
258. Paris, R. B. and Wood, A. D.: A model for optical tunneling, *IMA J. Appl. Math.* **43** (1989), 273–284. Exponentially small leakage from the fiber.
259. Paris, R. B. and Wood, A. D.: Exponentially-improved asymptotics for the gamma function, *J. Comput. Appl. Math.* **41** (1992), 135–143.
260. Paris, R. B. and Wood, A. D.: Stokes phenomenon demystified, *IMA Bulletin* **31** (1995), 21–28. Short review of hyperasymptotics.
261. Pokrovskii, V. L.: Science and life: conversations with Dau, in: I. M. Khalatnikov (ed.), *Landau, the Physicist and the Man: Recollections of L. D. Landau*, Pergamon Press, Oxford,

1989. Relates the amusing story that the Nobel Laureate Lev Landau believed the Pokrovskii–Khalatnikov (1961) “beyond-all-orders” method was wrong. The correct answer (but with incorrect derivation) is given in the Landau–Lifschitz textbooks. Eventually, Landau realized his mistake and apologized.

262. Pokrovskii, V. L. and Khalatnikov, I. M.: On the problem of above-barrier reflection of high-energy particles, *Soviet Phys. JETP* **13** (1961), 1207–1210. Applies matched asymptotics in the complex plane to compute the exponentially small reflection which is missed by WKB.

263. Pomeau, Y., Ramani, A. and Grammaticos, G.: A Structural stability of the Korteweg–de Vries solitons under a singular perturbation, *Physica D* **21** (1988), 127–134. Weakly nonlocal solitons of the FKdV equation; complex-plane matched asymptotics.

264. Proudman, I. and Pearson, J. R.: Expansions at small Reynolds numbers for the flow past a sphere and a circular cylinder, *J. Fluid Mech.* **2** (1957), 237–262. Log-and-powers expansion.

265. Raithby, G.: Laminar heat transfer in the thermal entrance region of circular tubes and two-dimensional rectangular ducts with wall suction and injection, *Internat. J. Heat Mass Transfer* **14** (1971), 223–243.

266. Ramis, J. P. and Schafke, R.: Gevrey separation of fast and slow variables, *Nonlinearity* **9** (1996), 353–384. Iterated averaging transformations of perturbed one phase Hamiltonian systems, not necessarily conservative.

267. Reddy, S. C., Schmid, P. J. and Henningson, D. S.: Pseudospectra of the Orr–Sommerfeld equation, *SIAM J. Appl. Math.* **53** (19), 15–47. Exponentially sensitive eigenvalues.

268. Reichel, L. and Trefethen, N. L.: The eigenvalues and pseudo-eigenvalues of Toeplitz matrices, *Linear Algebra Appl.* **162** (1992), 153–185.

269. Reinhardt, W. P.: Padé summation for the real and imaginary parts of atomic Stark eigenvalues, *Int. J. Quantum Chem.* **21** (1982), 133–146. Two successive Padé transformations are used to compute the exponentially small imaginary part of the eigenvalue.

270. Richardson, L. F.: The deferred approach to the limit. Part I. – Single lattice, *Philos. Trans. Roy. Soc.* **226** (1927), 299–349. Invention of Richardson extrapolation, which is an asymptotic but divergent procedure because of beyond-all-orders terms in the grid spacing h. Reprinted in Richardson’s Collected Papers, ed. by O. M. Ashford *et al.*

271. Richardson, L. F.: The deferred approach to the limit. Part I.–Single lattice, in: O. M. Ashford, H. Charnock, P. G. Drazin, J. C. R. Hunt, P. Smoker and I. Sutherland (eds), *Collected Papers of Lewis Fry Richardson*, Cambridge University Press, New York, 1993, pp. 625–678.

272. Robinson, W. A.: The existence of multiple solutions for the laminar flow in a uniformly porous channel with suction at both walls, *J. Engg. Math.* **10** (1976), 23–40. Exponentially small difference between two distinct nonlinear solutions.

273. Rosser, J. B.: Transformations to speed the convergence of series, *J. Res. Nat. Bureau of Standards* **46** (1951), 56–64. Convergence factors; improvements to asymptotic series.

274. Rosser, J. B.: Explicit remainder terms for some asymptotic series, *J. Rat. Mech. Anal.* **4** (1955), 595–626.

275. Scheurle, J., Marsden, J. E. and Holmes, P.: Exponentially small estimate for separatrix splitting, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 187–196. Show that the splitting is proportional to $\nu(\varepsilon)\exp(-\pi/(2\varepsilon))$ where $\nu(\varepsilon)$ has as essential singularity at $\varepsilon = 0$ and must be represented as a Laurent series rather than a power series. No examples of essentially-singular $\nu(\varepsilon)$ for nonlocal solitons are as yet known.

276. Schraiman, B. I.: On velocity selection and the Saffman–Taylor problem, *Phys. Rev. Lett.* **56** (1986), 2028–2031.

277. Schulten, Z., Anderson, D. G. M. and Gordon, R. G.: An algorithm for the evaluation of the complex Airy functions, *J. Comput. Phys.* **31** (1979), 60–75. An alternative to hyperasymptotics – a very efficient one.

278. Segur, H. and Kruskal, D.: On the nonexistence of small amplitude breather solutions in ϕ^4 theory, *Phys. Rev. Lett.* **58** (1987), 747–750. Title not withstanding, the ϕ^4 breather does exist, but is nonlocal.
279. Segur, H., Tanveer, S. and Levine, H. (eds): *Asymptotics Beyond All Orders*, Plenum, New York, 1991, 389 pages.
280. Sergeev, A. V.: Summation of the eigenvalue perturbation series by multivalued Padé approximants – application to resonance problems and double wells, *J. Phys. A: Math. Gen.* **28** (1995), 4157–4162. Shows that Shafer's generalization of Padé approximants, when the approximant is the solution of a quadratic equation with polynomial coefficients, converge to the lowest eigenvalue of the quantum quartic oscillator even when the perturbation parameter ε ("coupling constant") is real and negative and thus lies on the branch cut of the eigenvalue. (Ordinary Padé approximants fail on the branch cut.)
281. Sergeev, A. V. and Goodson, D. Z.: Summation of asymptotic expansions of multiple-valued functions using algebraic approximants: Application to anharmonic oscillators, *J. Phys. A: Math. Gen.* **31** (1998), 4301–4317. Show that Shafer's (1974) generalization of Padé approximants can successfully sum the exponentially small imaginary part of some functions with divergent power series, as illustrated through the quantum quartic oscillation with negative coupling constant.
282. Shafer, R. E.: On quadratic approximation, *SIAM J. Numer. Anal.* **11** (1974), 447–460. Generalization of Padé approximants. A function $u(z)$, known only through its power series, is approximated by the root of a quadratic equation. The coefficients of the quadratic are polynomials which are chosen so that the power series of the root of the quadratic equation will match the power series of u to a given order.
283. Skinner, L. A.: Generalized expansions for slow flow past a cylinder, *Quart. J. Mech. Appl. Math.* **28** (1975), 333–340. Log-and-series in Re.
284. Snyder, M. A.: *Chebyshev Methods in Numerical Approximation*, Prentice-Hall, Englewood Cliffs, New Jersey, 1966, p. 150.
285. Sternin, B. Y. and Shatalov, V. E.: *Borel–Laplace Transform and Asymptotic Theory: Introduction to Resurgent Analysis*, CRC Press, New York, 1996.
286. Stieltjes, T. J.: Recherches sur quelques séries semi-convergentes, *Ann. Sci. École Norm. Sup.* **3** (1886), 201–258. Hyperasymptotic extensions to asymptotic series.
287. Suvernev, A. A. and Goodson, D. Z.: Perturbation theory for coupled anharmonic oscillators, *J. Chem. Phys.* **106** (1997), 2681–2684. Computation of complex-valued eigenvalues through quadratic Shafer–Padé approximants; the imaginary parts are exponentially small in the reciprocal of the perturbation parameter.
288. Tanveer, S.: Analytic theory for the selection of Saffman–Taylor finger in the presence of thin-film effects, *Proc. Roy. Soc. London A* **428** (1990), 511.
289. Tanveer, S.: Viscous displacement in a Hele-Shaw cell, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 131–154.
290. Terrill, R. M.: Laminar flow in a uniformly porous channel with large injection, *Aeronautical Quarterly* **16** (1965), 323–332.
291. Terrill, R. M.: On some exponentially small terms arising in flow through a porous pipe, *Quart. J. Mech. Appl. Math.* **26** (1973), 347–354.
292. Terrill, R. M. and Thomas, P. W.: Laminar flow in a uniformly porous pipe, *Appl. Sci. Res.* **21** (1969), 37–67.
293. Tovbis, A.: On exponentially small terms of solutions to nonlinear ordinary differential equations, *Methods Appl. Anal.* **1** (1994), 57–74.
294. Trefethen, L. N. and Bau, D., III: *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
295. Tuan, P. D. and Elliott, D.: Coefficients in series expansions for certain classes of functions, *Math. Comp.* **26** (1972), 213–232.

296. Van der Waerden, B. L.: On the method of saddle points, *Appl. Sci. Res. B2* (1951), 33–45. Steepest descent for integral with nearly coincident saddle point and pole.
297. Van Dyke, M.: *Perturbation Methods in Fluid Mechanics*, 1st edn, Academic Press, Boston, 1964.
298. Van Dyke, M.: *Perturbation Methods in Fluid Mechanics*, 2nd edn, Parabolic Press, Stanford, California, 1975.
299. Vanden-Broeck, J.-M. and Turner, R. E. L.: Long periodic internal waves, *Phys. Fluids A* **4** (1992), 1929–1935.
300. Vaĭnberg, V. M., Mur, V. D., Popov, V. S. and Sergeev, A. V.: Strong-field Stark effect, *JETP Lett.* **44** (1986), 9–13. Shafer–Padé approximants are used to compute the complex-valued eigenvalues of the hydrogen atom in an electric field. The imaginary part is exponentially in the reciprocal of the perturbation parameter.
301. Voros, A.: Semi-classical correspondence and exact results: The case of the spectra of homogeneous Schrödinger operators, *J. Physique-Lett.* **43** (1982), L1–L4.
302. Voros, A.: The return of the quartic oscillator: the complex WKB method, *Ann. Inst. H. Poincaré, Physique Théorique* **39** (1983), 211–338.
303. Voros, A.: Schrödinger equation from $O(\hbar)$ to $o(\hbar^{\infty})$, in: M. C. Gutzwiller, A. Inomata, J. R. Klauder and L. Streit (eds), *Path Integrals from meV to MeV*, No. 7 in Bielefeld Encounters in Physics and Mathematics, Bielefeld Center for Interdisciplinary Research, World Scientific, Singapore, 1986, pp. 173–195. Review.
304. Voros, A.: Quantum resurgence, *Ann. Inst. Fourier* **43** (1993), 1509–1534. In French.
305. Voros, A.: Aspects of semiclassical theory in the presence of classical chaos, *Prog. Theor. Phys.* **116** (1994), 17–44.
306. Voros, A.: Exact quantization condition for anharmonic oscillators (in one dimension), *J. Phys. A: Math. Gen.* **27** (1994), 4653–4661.
307. Wai, P. K. A., Chen, H. H. and Lee, Y. C.: Radiation by "solitons" at the zero group-dispersion wavelength of single-mode optimal fibers, *Phys. Rev. A* **41**(19), 426–439. Nonlocal envelope solitons of the TNLS Eq. Their (2.1) contains a typo and should be $\bar{q}^{(2)} = -(39/2)A^2\bar{q}^{(0)} + 21|\bar{q}^{(0)}|^2\bar{q}^{(0)}$.
308. Ward, M. J., Henshaw, W. D. and Keller, J. B.: Summing logarithmic expansions for singularly perturbed eigenvalue problems, *SIAM J. Appl. Math.* **53** (1993), 799–828.
309. Weideman, J. A. C.: Computation of the complex error function, *SIAM J. Numer. Anal.* **31** (1994), 1497–1518. [Errata: 1995, **32**, 330–331.] These series of rational functions are useful for complex-valued z.
310. Weideman, J. A. C.: Computing integrals of the complex error function, *Proceedings of Symposia in Applied Mathematics* **48** (1994), 403–407. Short version of Weideman (1994a).
311. Weideman, J. A. C.: Errata: computation of the complex error function, *SIAM J. Numer. Anal.* **32** (1995), 330–331.
312. Weideman, J. A. C. and Cloot, A.: Spectral methods and mappings for evolution equations on the infinite line, *Comput. Meth. Appl. Mech. Engr.* **80** (1990), 467–481. Numerical.
313. Weinstein, M. I. and Keller, J. B.: Hill's equation with a large potential, *SIAM J. Appl. Math.* **45** (1985), 200–214.
314. Weinstein, M. I. and Keller, J. B.: Asymptotic behavior of stability regions for Hill's equation, *SIAM J. Appl. Math.* **47** (1987), 941–958.
315. Weniger, E. J.: Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series, *Comput. Phys. Reports* **10** (1989), 189–371.
316. Weniger, E. J.: On the derivation of iterated sequence transformations for the acceleration of convergence and the summation of divergent series, *Comput. Phys. Comm.* **64** (1991), 19–45.
317. Wimp, J.: The asymptotic representation of a class of G-functions for large parameter, *Math. Comp.* **21** (1967), 639–646.

318. Wimp, J.: *Sequence Transformations and Their Applications*, Academic Press, New York, 1981.
319. Wong, R.: *Asymptotic Approximation of Integrals*, Academic Press, New York, 1989.
320. Wood, A.: Stokes phenomenon for high order differential equations, *Z. Angew. Math. Mech.* **76** (1996), 45–48. Brief review.
321. Wood, A. D.: Exponential asymptotics and spectral theory for curved optical waveguides, in: H. Segur, S. Tanveer and H. Levine (eds), *Asymptotics Beyond All Orders*, Plenum, Amsterdam, 1991, pp. 317–326.
322. Wood, A. D. and Paris, R. B.: On eigenvalues with exponentially small imaginary part, in: R. Wong (ed.), *Asymptotic and Computational Analysis*, Marcel Dekker, New York, 1990, pp. 741–749.
323. Yang, T.-S.: On traveling-wave solutions of the Kuramoto–Sivashinsky equation, *Physica D* **110** (1998), 25–42. Shocks with oscillations, exponentially small in $1/\varepsilon$, which grow slowly in space, and thus are (very!) nonlocal. Applies the Akylas–Yang beyond-all-orders perturbation method in wavenumber space to compute the far field for oscillatory shocks. These are then matched to the nonlocal regular shocks to create solitary waves (that asymptote to the same constant as $x \to \pm\infty$); these are confirmed by numerical solutions.
324. Yang, T.-S. and Akylas, T. R.: Radiating solitary waves of a model evolution equation in fluids of finite depth, *Physica D* **82** (1995), 418–425. Solve the Intermediate-Long Wave (ILW) equation for water waves with an extra third derivative term, which makes the solitary waves weakly nonlocal. The Yang–Akylas matched asymptotics in wavenumber is used to calculate the exponentially small amplitude of the far field oscillations.
325. Yang, T.-S. and Akylas, T. R.: Weakly nonlocal gravity-capillary solitary waves, *Phys. Fluids* **8** (1996), 1506–1514.
326. Yang, T.-S. and Akylas, T. R.: Finite-amplitude effects on steady lee-wave patterns in subcritical stratified flow over topography, *J. Fluid Mech.* **308** (1996), 147–170.
327. Yang, T.-S. and Akylas, T. R.: On asymmetric gravity-capillary solitary waves, *J. Fluid Mech.* **330** (1997), 215–232. Asymptotic analysis of classical solitons of the FKdV equation; demonstrates the coalescence of classical bions.
328. Zinn-Justin, J.: *Quantum Field Theory and Critical Phenomena*, Oxford University Press, Oxford, 1989.