ASYMPTOTICS, SUPERASYMPTOTICS, HYPERASYMPTOTICS...

MICHAEL BERRY

H.H.Wills Physics Laboratory
Tyndall Avenue
Bristol BS8 1TL
United Kingdom

'Divergent series are the invention of the devil, and it is shameful to base on them any demonstration whatever' (Abel, 1828)

'2+2=5, for sufficiently large values of 2' (Princeton asymptotic graffito)

1. INTRODUCTION

My purpose is to describe several recent developments in our understanding of divergent series and the accurate calculation of the functions they represent. All the work has been[1,2] or is being published[3], so this will be an informal account, emphasising the new concepts and illustrating them with pictures.

It is useful to introduce some terminology. Typically, an asymptotic series for a function depending on a large parameter k and several variables $X=(X_1,X_2,...)$ has the form

$$y(k;X)=M(k;X)\exp\{k\phi(X)\}\sum_{r=0}^{\infty}Y_r(k,X), \quad \text{where } Y_0=1 \quad \text{and } Y_r \propto k^{-r} \tag{1}$$

(Often k - which will not always be written explicitly - serves simply as a book-keeping parameter, to order the terms in the series.) In the cases we are interested in, the series diverges and so is meaningless when interpreted conventionally. The usual 'asymptotics' is the study of the series truncated at fixed order $r=N$: according to Poincaré's definition[4], the series is asymptotic if the error is of order $k^{-(N+1)}$. However, as was known to Stokes [5] nearly half a century before Poincaré, much more accurate approximations can be obtained by truncating not at fixed order but at the least term, which typically increases with k. It is common to achieve errors of order $\exp(-k)$ with such optimal truncation, which therefore constitutes 'asymptotics beyond all orders' or, as I will call it, 'superasymptotics'. (After introducing this term I felt at first shamed by Barbara Levi's gentle satire[6] on physicists' predilection for terminological 'super'iority, but was later made unashamed by reading Lord Kelvin's memorial appreciation[7], in which he described Stokes' early work on divergent series as 'mathematical supersubtlety'.) We shall also require a term for systematic improvements to the exponentially small remainder of a optimally truncated series . I will call these 'hyperasymptotics'. Thus hypersaymptotics goes 'beyond asymptotics beyond all orders.'

Asymptotics beyond All Orders, Edited by H. Segur *et al.*
Plenum Press, New York, 1991

Underlying the recent work are two ideas. First, that an asymptotic series such as (1) is a compact encoding of a function, and its divergence should be regarded not as a deficiency but as a source of information about the function. In particular, divergence usually indicates the presence of exponentially small terms which the bare asymptotic series, uninterpreted, cannot capture. This is why superasymptotics can yield exponential accuracy. A consequence is that the late terms of the asymptotic series associated with one exponential are frequently related by 'resurgence' to the early terms of the series associated with another exponential. Second, that the divergences of the series obtained by a variety of methods, and representing a variety of functions, follow a common pattern: factorial divided by a power. Recognition of this universality and its cause leads to powerful resummation techniques enabling the asymptotics to be decoded to yield precise (hyperasymptotic) numerical information. These principles were systematically explored and exploited by Dingle in the 1950s, and summarised in his 1973 book[8], but are only now becoming widely known.

In the new results I describe here, Dingle's work is extended in two ways. The first concerns Stokes' phenomenon[9], namely rapid jumps, as the variables X are changed, in the multipliers M of a small (subdominant) exponential whilst hidden behind a big (dominant) one. In a sense this is the very heart of asymptotics, because such changes in form necessarily accompany the divergence of the asymptotic series associated with each exponential, reflecting its inability to describe the other exponentials. In my opinion, the persistent failure to understand Stokes' phenomenon (still evident in the literature) is in large measure responsible for what Littlewood[10] called the 'aroma of paradox and audacity' that has hung about the whole subject of divergent series, connection formulae in WKB theory, etc. By appropriate magnification and resummation, however, a precise description of the change in the subdominant multiplier can be obtained, in terms of a universal scaling function. This refinement and demystification of Stokes' phenomenon can be regarded as the consequence of just the first step into hyperasymptotics.

The second result goes much further. By systematically exploiting resurgence, the remainder in an optimally truncated expansion can itself be expressed as an asymptotic series, which has its own remainder,..... Iteration of this hyperasymptotic process leads to an intricate sequence of hyperseries in which the original asymptotic coefficients (the Y_r in (1)) are renormalised by certain universal functions. At the end, after hyperasymptotics has come to a natural halt, the error is reduced, not to zero but to less than the square of the superasymptotic error. For integrals of exponentials with several saddles, there is a remarkable resurgence identity connecting the expansions about the different saddles; this can be employed to refine the method of steepest descent into an exact technique, whose hyperasymptotics can be accomplished without resumming divergent series.

I will illustrate these general ideas with the Airy function:

$$\mathrm{Ai}(z) \equiv \frac{1}{2\pi}\int_C du \exp\left\{i\left(\frac{u^3}{3}+zu\right)\right\} \qquad (2)$$

Here the infinite contour C runs from $\infty\exp(5\pi i/6)$ to $\infty\exp(\pi i/6)$, so that the integral converges for all complex z. For real z, C can be deformed to the real axis, and Ai(z) is real. For any z, Ai depends on two real quantities: the modulus and phase of z. In terms of the general theory for the series (1), $|z|$ (or rather $|z|^{3/2}$) will be the large parameter k and arg z will be the variable X. In § 2 we show the ordinary asymptotics and superasymptotics of Ai, with and without the Stokes jump, which is represented as a discontinuity. In §3 the universal smoothing is described. §4 contains an account of resurgence and hyperasymptotics, again with Ai as an example.

Having these new techniques, I would like to hear from anybody who needs the Airy function to twenty decimals, but am not expecting an early call. Probably no application requires such accuracy. This being so, it is important to reveal the motivation for this renewed interest in the oldest and simplest problems of asymptotics. This I leave to the concluding §5.

2. DOMINANT AND SUBDOMINANT SERIES; STOKES' JUMP

We shall display several approximations to Ai(z) for large $|z|$, the aim being to understand the asymptotics in the upper half-plane (fig. 1), that is as θ=arg z varies from 0 to π. For large $|z|$ it is appropriate to approximate (2) by the saddle-point method[11]. There are two saddles, at $u=\pm iz^{1/2}$, at which the integrand is

$$\exp\left\{\pm\frac{1}{2}F\right\}\equiv\exp\left\{\mp\frac{2}{3}z^{3/2}\right\} \tag{3}$$

(note the signs- see fig. 1). Following Dingle[8], we have introduced the 'singulant' F, namely the difference between the two exponents. The full significance of this quantity will emerge later. In our first numerical illustrations we shall take the 'large parameter' as $|F|=3$, i.e. $|z|=1.7171$.

Study of the topology of the phase in (2) shows that when $\theta<120°$ the contour C can be deformed into a steepest path passing through only one of the saddles, yielding the lowest approximation

$$\mathrm{Ai}(z)\approx\mathrm{Ai1}(z)=\frac{1}{2z^{1/4}\sqrt{\pi}}\exp\left\{\frac{F}{2}\right\} \tag{4}$$

Fig. 2 is an Argand plot comparing Ai1 with the exact Ai (computed to high precision, e.g. by the convergent series) in the upper half-plane. Agreement is reasonable for small θ, but rapidly deteriorates, becoming worst on the negative real axis, where Ai1 is complex whereas Ai is real.

The natural next step is to include higher-order corrections to Ai1, giving the approximation[8]

$$\mathrm{Ai}(z)\approx\mathrm{Ai2}(z,N)=\frac{1}{2z^{1/4}\sqrt{\pi}}\exp\left\{\frac{F}{2}\right\}\sum_{r=1}^{N}Y_r\quad\text{where}\quad Y_r=\frac{\Gamma\left(r+\frac{1}{6}\right)\Gamma\left(r+\frac{5}{6}\right)}{2\pi F^r\,\Gamma(r+1)} \tag{5}$$

For large orders ('the asymptotics of the asymptotics'), the coefficients are

$$Y_r\xrightarrow[r\to\infty]{}\frac{(r-1)!}{2\pi F^r} \tag{6}$$

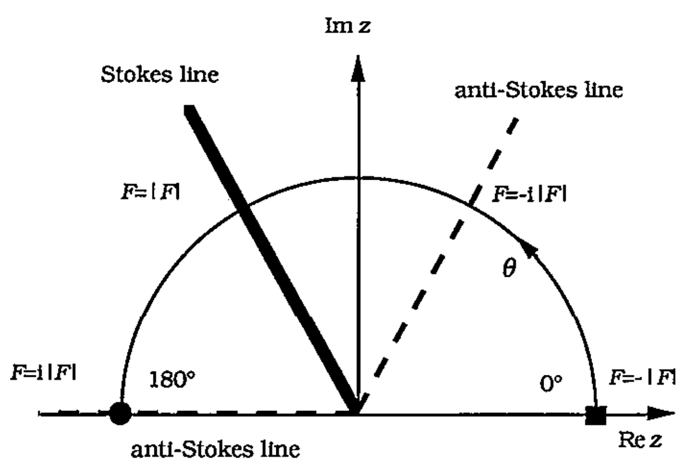

Fig. 1 Upper half-plane of argument z of the Airy function Ai(z).

Thus the smallest term, corresponding to optimal truncation, i.e. superasymptotics, is near $N=r^*=\mathrm{Int}|F|$. The superasymptotic Ai2 is shown in fig.2. The agreement is much better for small θ, but no better for $\theta=180°$.

The reason for the poor agreement near the negative real axis is the neglect of *Stokes' phenomenon*: for $\theta>120°$ the steepest-descent deformation of C passes through both saddles, so that the contribution of the second exponential in (3) should also be included. $\theta=120°$ is the *Stokes line* for Ai, defined as the locus of greatest disparity between the two exponentials, where F is positive real and the terms Y_r in the dominant series (5) all have the same sign, so that the divergence of the series is most severe. Stokes[9,12] argued that the extra exponential should be regarded as being born on this line, where it is smallest. Incorporating it into the lowest-order approximation gives

$$\mathrm{Ai}(z) \approx \mathrm{Ai3}(z) = \frac{1}{2z^{1/4}\sqrt{\pi}}\left(\exp\left\{+\frac{F}{2}\right\} + i\exp\left\{-\frac{F}{2}\right\}\mathrm{H}(\theta-120°)\right)$$
$$= \frac{1}{(-z)^{1/4}\sqrt{\pi}}\sin\left\{\tfrac{2}{3}(-z)^{3/2}+\tfrac{1}{4}\pi\right\} \quad \text{if} \quad \theta>120° \tag{7}$$

where H denotes the unit step. Note the factor i in the new exponential. This birth 'in quadrature' not only makes the jump most unobtrusive but also ensures that Ai3 is real on the negative axis, which is an anti-Stokes line for Ai, that is the locus of equal magnitude of the two exponentials. Fig. 3 shows the considerable improvement that this produces: the overall agreement is much better, and the discontinuity is indeed unobtrusive.

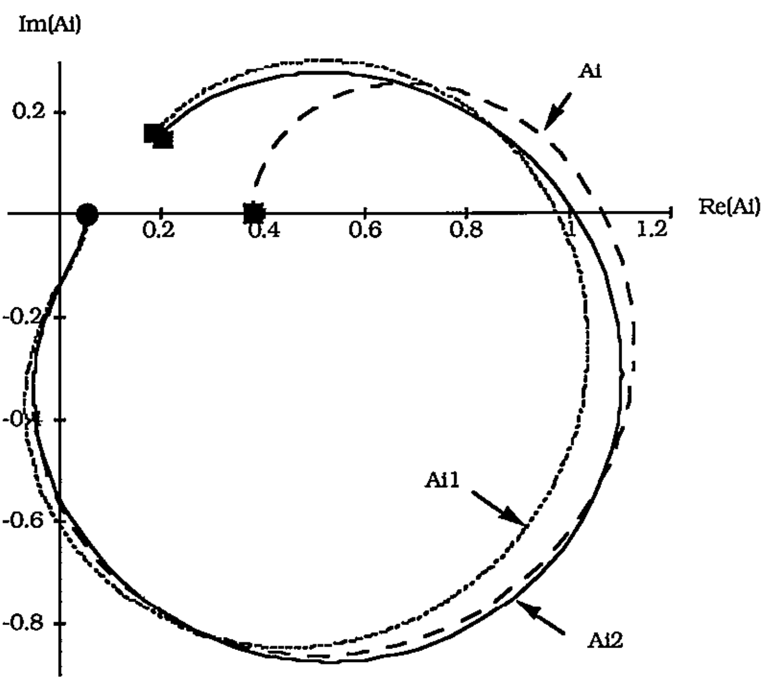

Fig. 2. Argand plot of exact Airy function Ai (dashed line) in the upper half-plane, along the semicircular path shown in fig.1, for $|F|=3$ (i.e. $|z|=1.7171$), compared with lowest-order (dominant exponential) asymptotics Ai1 (dotted line) and superasymptotics Ai2 with $N=3$ (dominant exponential x optimally truncated series) (full line). l marks $\theta=0$, n marks $\theta=180°$.

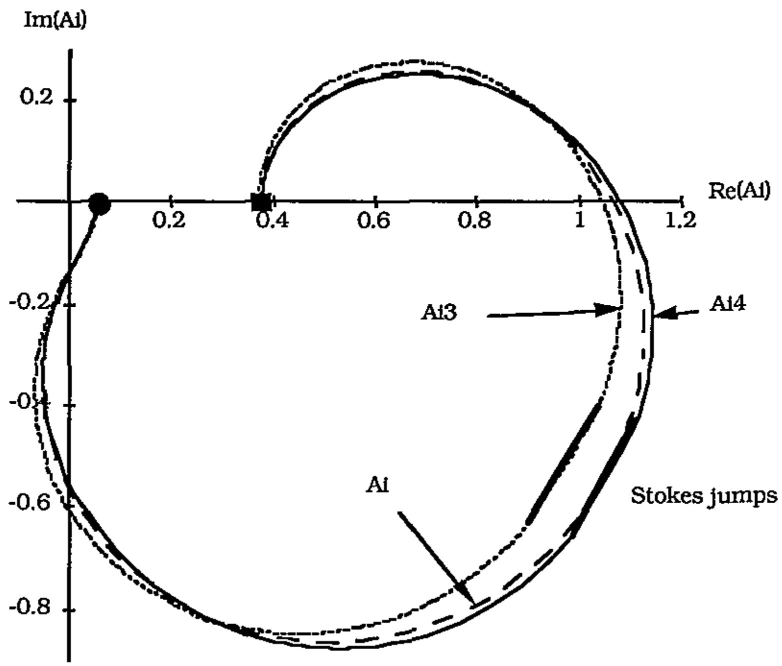

Fig. 3. As fig. 2, but comparing the exact Ai (dashed line) with Ai3 (dotted line) (Stokes jump included to lowest order) and Ai4 with N=3 (full line)(Stokes jump included superasymptotically). The jumps are shown as bold lines.

Stokes' analysis indicated that the best approximation is obtained by including the jump at the superasymptotic level, i.e. in the optimally truncated series. This gives

$$\mathrm{Ai}(z) \approx \mathrm{Ai4}(z,N) = \frac{1}{2z^{1/4}\sqrt{\pi}}\left[\exp\left\{+\frac{F}{2}\right\}\sum_{r=0}^{N} Y_r + \mathrm{i}\exp\left\{-\frac{F}{2}\right\}\sum_{r=1}^{N}(-1)^r Y_r\ \mathrm{H}(\theta-120°)\right] \tag{8}$$

As fig. 3 shows, Ai4 indeed gives a dramatic improvement, even here where $|F|=3$ and the new exponential appears with relative magnitude exp(-3)=0.0498, which is hardly small.

Figs. 2 and 3 extend a numerical experiment of Stokes[9] demonstrating the reality of his phenomenon. Since then, many people(e.g.[13, 14]) have rediscovered the increased accuracy achieved by correctly including small exponentials. For this to be a consistent procedure, it is essential to go to the superasymptotic level, where the first neglected term is exponentially small. Ordinary Poincaré asymptotics is inadequate because with this, as has often been remarked, the small exponentials $\exp(-|F|)$ are lost in the truncation errors $|F|^{-N}$.

3. SMOOTHING STOKES' DISCONTINUITY

The exact Airy function changes smoothly, so that any discontinuity at the Stokes line, where the subdominant exponential appears suddenly, must be an artefact of poor resolution. To get an approximation without discontinuity, it is necessary to go beyond superasymptotics. The first step into hyperasymptotics is sufficient to resolve the structure near the Stokes line. This can be accomplished by taking seriously what superasymptotics neglects, namely the tail of the series beyond the optimal truncation limit $N=r^*$. An important observation is that the late terms formula (6) (factorial divided by a power) is not restricted to Ai but has a very wide range of validity when modified as follows:

$$Y_r \xrightarrow[r\to\infty]{} \frac{M_-}{M}\frac{(r-1)!}{2\pi F^r} \tag{9}$$

Here F denotes the difference between the exponent $k\phi$ in (1) and the leading subdominant exponent, which we call $k\phi_-$, and M_- the multiplier appropriate to the subdominant series (for Ai, $M_-=M$). The simplicity of (9) is remarkable. Anybody who has computed asymptotic corrections in realistic (i.e. not textbook) applications knows that the algebra gets very heavy and generates enormous formulae (see e.g. pp 119-121 of Dingle's book[8]).

The underlying reason for the simplification, well explained by Dingle[8], is that all asymptotic methods (saddle-point and end-point integration, WKB solution of differential equations, etc.) are generated by *local expansions*. Thus, successive terms in the expansions involve successive derivatives, and late terms correspond to high derivatives. But by Darboux's theorem[8] the high derivatives of a function $f(t)$ at, say, $t=0$ are dominated by the nearest singularity, at $t=t^*$, say. Typically this will be a pole or branch point, and then the high derivatives will indeed have the form factorial/power. A common case is the simple pole

$$f(t) \xrightarrow[t\to t^*]{} \frac{A}{(t-t^*)} \tag{10}$$

Successive differentiation swells the range of validity of this formula from the neighbourhood of t^* to $t=0$, so that

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r} f(0) \xrightarrow[r\to\infty]{} -A\frac{r!}{(t^*)^{r+1}} \tag{11}$$

For example[8], if

$$f(t) = \frac{1}{1+\log(1+t)} \tag{12}$$

Darboux's principle gives, from the pole at $t^*=\mathrm{e}^{-1}-1$,

$$r!\frac{\mathrm{d}^r}{\mathrm{d}t^r} f(0) \xrightarrow[r\to\infty]{} \frac{(-1)^r}{\mathrm{e}\left(1-\mathrm{e}^{-1}\right)^{r+1}} \tag{13}$$

If $r=0$,this formula gives 0.58, in poor agreement with the exact value 1, but for the 'late' term $r=8$ it gives 22.8300, close to the exact value 38371/1680=22.8399.

Formally, we can write the expansion (1) as the optimally truncated series plus the divergent tail with its terms approximated by (9), in the form

$$y \approx M\exp\{k\phi\}\sum_{r=0}^{r^*} Y_r + iM_- S\exp\{k\phi_-\}, \quad \text{where}$$
$$S(F) \equiv \frac{-i}{2\pi}\exp(F)\sum_{r^*+1}^{\infty}\frac{(r-1)!}{F^r} \tag{14}$$

In coded form, $S(F)$ is the Stokes multiplier, describing the appearance of the subdominant exponential across the Stokes line F positive real. To decode it, we employ *Borel summation*, that is[4,8] writing the factorial in the familiar integral representation and then evaluating the sum. This replaces (14) by a convergent integral, which must be approximated for large $|F|$ (this is 'the asymptotics of the asymptotics of the asymptotics'). I have done this elsewhere[1], and do not repeat the details here. The important point is that the evaluation of the Borel integral is greatly simplified by optimal truncation, because then (and only then) a pole and saddle in its integrandcoincide.

In the case (e.g. Ai) where there is no subdominant exponential 'before' the Stokes line, i.e. for Im F<<0, the multiplier takes the very simple form

$$S(F) \approx \frac{1}{2}\left[1+\mathrm{Erf}\left\{\frac{\mathrm{Im}F}{\sqrt{2\,\mathrm{Re}\,F}}\right\}\right] \tag{15}$$

where Erf denotes the familiar error function[15]. As the argument of Erf increases from $-\infty$ to ∞ (i.e. between the two anti-Stokes lines adjacent to the Stokes line Im F=0), $S(F)$ increases from -1 to 1. Recalling that F is proportional to the large parameter k, we see that the 'width' of the Stokes line, that is the range in the space of variables X over which the subdominant exponential enters, scales as $k^{-1/2}$.

I wish to make four remarks about the error-function smoothing formula. The first concerns its generality. Although it provides a refined description of the Stokes phenomenon in Airy, Bessel, hypergeometric, and Mathieu functions (and even the error function itself), it is not restricted to these special functions, nor to the solutions of certain differential equations, nor to integrals with a large parameter. Its range of applicability is all functions whose asymptotic series diverge as (factorial/power).

The second remark concerns the extent to which the derivation of the error-function smoothing requires the resummation of divergent series. Until now, resummation provides the most direct and context-free route to the formula. It does not however seem to be popular amongst mathematicians - certainly not those who have taken up the important question of providing a rigorous justification for the smoothing, with error bounds. In particular cases where this has been possible[16,18,19], it was achieved by using special methods, appropriate to particular classes of integrals, where the remainder can be expressed in closed form rather than as a divergent series (see §4 for a wide generalisation of such cases). Other problems for which the smoothing can be obtained without resummation, albeit sill non-rigorously, are certain second-order differential equations[20], or equivalent first-order systems[21].

The third remark concerns the importance of optimal truncation. Without this, the Stokes multiplier is still defined as in (14), but with a different summation limit. This change seems innocuous but actually makes a big difference[20-22]. For non-optimal truncation (which means that N lies outside the range $|F|-\sqrt{|F|}$ to $|F|+\sqrt{|F|}$), $S(F)$ still increases from 0 to 1, but with exponentially large oscillations and over an X-range bigger than $k^{-1/2}$.

The fourth remark is that the smoothing has applications in physics. In wave theory, small exponentials represent complex evanescent rays, so that Stokes' phenomenon describes the gentle birth of rays[12]- in contrast to caustics, which represent the violent transformation of real into complex rays. Mathematically, rays correspond to saddles of diffraction integrals, Stokes' phenomenon to two saddles having the same (imaginary part of) height, and caustics to two saddles colliding. For integrals more complicated than that describing Ai, Stokes' phenomenon can occur on surfaces in the space of real parameters X, and can take interesting forms[23]. One direct application of the smoothing formula is to the generation of exponentially weak reflections[20], for example above a potential barrier in quantum mechanics. Another is to the history of a quantal transition between two states, induced by a slowly changing field; in this case, optimal truncation corresponds to a particular choice of basis states, and suggests new experiments[21,22] to detect the Stokes phenomenon.

Applied to Ai, the smoothing (15) spreads the Stokes jump over the region between the anti-Stokes lines at θ=60° and θ=180°, and gives the approximation

$$\mathrm{Ai}(z) \approx \mathrm{Ai5}(z,N)$$

$$= \frac{1}{2z^{1/4}\sqrt{\pi}} \left(\exp\left\{+\frac{F}{2}\right\} \sum_{r=0}^{N} Y_r + \mathrm{i}\exp\left\{-\frac{F}{2}\right\} \frac{\left[1+\mathrm{Erf}\left\{-\sin\left(\frac{3}{2}\theta\right)\sqrt{\frac{-|F|}{2\cos\left(\frac{3}{2}\theta\right)}}\right\}\right]}{2} \mathrm{H}(\theta-60°) \right) \quad (16)$$

(despite the step on the antiStokes line, there is no discontinuity). In fig. 4 this is compared with the exact Ai. Evidently the agreement is again much improved: the curves can hardly be distinguished over the whole range of θ. (Actually the approximation Ai5 is defective in that it is not real when θ=180°, but the cure, which is to include in the subdominant contribution the first N terms of its asymptotic series (cf. (8)) - a procedure for which there is theoretical justification[2] - leads to a curve which cannot be distinguished from Ai5 in fig. 4.)

A more discriminating test is shown in fig. 5. Here the error-function smoothing (15) is compared with the exact multiplier defined by (14) and (5), namely

$$S(F) = -2\mathrm{i}z^{1/4}\sqrt{\pi}\exp\left(\frac{F}{2}\right)\left[\mathrm{Ai}(z) - \mathrm{Ai2}(z,r^*)\right] \quad (17)$$

This multiplier, predicted to be of order unity, is the difference between two quantities which near the Stokes line are both exponentially large: the exact Ai and its superasymptotic approximation. Even under this magnification, the agreement is excellent, and, as expected, is better for the larger singulant $|F|=10$ (refinement of the general theory[1] shows that the error in the error-function smoothing is of order $F^{-1/2}$ for Im S, and of order F^{-1} for Re S).

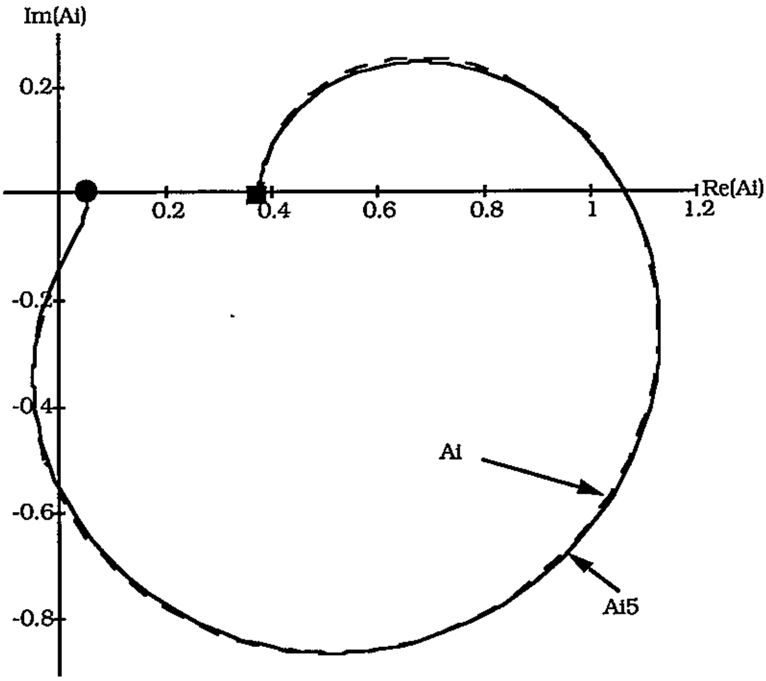

Fig. 4. As fig. 2, but comparing the exact Ai (dashed line)with Ai5 with N=3 (full line) (superasymptotics + Stokes smoothing of subdominant exponential).

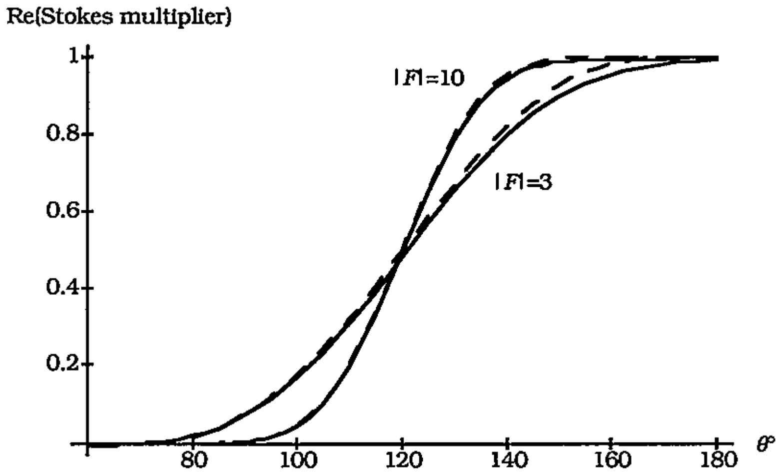

Fig.5. Real part of Stokes multiplier for the Airy function, across the Stokes line $\theta=120°$. The full lines are the exact $S(F)$ (equation 17); the dashed lines are the theoretical smoothing (15).

4. RESURGENCE

Deeper penetration of the asymptotic series (1) requires more accurate asymptotics of the asymptotics' than the leading term (9). Resurgence is a principle that greatly assists the determination of such higher-order approximations to the late terms. The idea is that if (1) is regarded as a complete asymptotic expansion which can represent the function $y(k;X)$ exactly, resummation of the late terms must yield not only the leading-order subdominant exponential but also the corrections terms in *its* asymptotic series. This must hold for all component asymptotic series, so each must contain, encoded in its late terms, all the terms of all the other series. Systematic exploration of this requirement of mutual consistency is still not completed. Dingle[8] gave several examples of resurgence; Écalle[24] described it at length (as well as inventing the term); Voros[25] applied it to differential equations (calling it 'analytic bootstrap'); and Flagolet and Odlyzko[26] examined applications to generating functions with exotic singularities.

Before illustrating resurgence, I should point out that for some simple functions it occurs only in rudimentary form. One such class (which includes the integrals Ei and Erf) is where the form (factorial/power) holds for all the terms Y_r, not just as $r\to\infty$; then a single resummation terminates the series exactly. Another class (which includes log $\Gamma(z)$[32]) is where the Y_r are given by an infinite convergent series of (factorial/power) terms, each of which can be exactly terminated by a single resummation. (A curiosity is that the superasymptotics of $n!$ requires summing to the least term $r\approx\pi n$, which involves $[\text{Int}(2\pi n)]!$ - a case of runaway self-reference, if not resurgence.)

Usually, though, we can expect resurgence to arise in all its glory, which will be illustrated now with a brief description of a new result obtained with Howls[3]. Consider the integral

$$I_j(k)=\int_{C_j(k)} dz\,G(z)\exp\{-k\phi(z)\} \qquad (18)$$

where $G(z)$ and $\phi(z)$ are nonsingular and $\phi(z)$ has saddles at a number of points z_j. $C_j(k)$ is one of the two infinite oriented steepest-descent contours through z_j. It is convenient, although not necessary, to think of k as complex, with $|k|$ as the large parameter and arg k as the variable X. Standard steepest-descents[8] yields the following series of the form (1), in which for convenience the prefactor and the coefficients Y_r have been amalgamated:

$$I_j(k) = \exp\{-k\phi_j\} \sum_{r=0}^{\infty} T_{jr}(k) \tag{19}$$

The terms are

$$T_{jr}(k) = \frac{\left(r-\frac{1}{2}\right)!}{2\pi i} \oint_j dz \frac{G(z)}{\left[F_j(z)\right]^{r+1/2}}, \text{ where}$$

$$F_j(z) \equiv k\left[\phi(z) - \phi_j\right] \tag{20}$$

Here the subscript j means 'evaluated at z_j', and the contour is a positive circuit of z_j. The integrals can all be evaluated explicitly in terms of derivatives of ϕ and G at z_j. For example,

$$T_{j0}(k) = \left(\frac{2\pi}{k\phi_j''}\right)^{1/2} G_j \tag{21}$$

(primes denote z derivatives).

The series (19) is a local expansion about the saddle j, and diverges because of the other saddles l. An explicit and exact expression, whose derivation[3], (involving P2C2E[27] here) has been obtained for this resurgence, showing how the integrals through certain other saddles give the remainder of the truncated series for a given saddle:

$$I_j(k) = \exp\{-k\phi_j\} \sum_{r=0}^{N-1} T_{jr}(k) +$$

$$+ \frac{\exp\{-k\phi_j\}}{2\pi i} \sum_l \frac{(-1)^{\gamma_{jl}}}{\left(F_{jl}\right)^{N-1/2}} \int_0^{\infty} dv \frac{v^{N-1/2}\exp(-v)}{F_{jl} - v} \left\{\exp\left(\frac{kv}{F_{jl}}\right) I_l\left(\frac{kv}{F_{jl}}\right)\right\} \tag{22}$$

Here

$$F_{jl} \equiv k\left(\phi_l - \phi_j\right) \tag{23}$$

denotes the singulant giving the exponent difference beteween the two saddles. The contributing other saddles l are determined by a topological rule: They are the saddles reached by lines of constant phase of $F_j(z)$ issuing from z_j. The sign of each contribution is determined by γ_{jl}, which is zero if the expanded loop contour through z_j has the same sense at z_l as $C_l(kv/F_{jl})$, and unity otherwise. Being exact, (22) contains the Stokes phenomenon (contribution from the pole at $v=F_{jl}$), and the higher terms of the original expansion (from the expression of the v integral as a factorial for large N).

Substitution of the series (19) converts (22) into a formally exact resurgence relation between the coefficients for the saddle j and the contributing other saddles l:

$$T_{jr} = \frac{1}{2\pi i} \sum_l (-1)^{\gamma_{jl}} \sum_{s=0}^{\infty} \frac{(r-s-1)!}{\left(F_{jl}\right)^{r-s}} T_{ls} \tag{24}$$

For large r, the leading contribution is the term s=0 from the saddle l for which $|F_{jl}|$ is smallest. Denoting this singulant by F, we obtain

$$T_{jr} \xrightarrow[r\to\infty]{} \frac{(-1)^{\gamma_{jl}}}{2\pi}\left(\frac{2\pi}{-k\phi_l''}\right)^{1/2} G_l \frac{(r-1)!}{F^r} \tag{25}$$

With appropriate identifications, this is the same as the previous late-terms approximation (9). Now, however, we have the complete expression, giving all the corrections, from all the contributing saddles. Of course it is only formal, because in the terms $s>r-1$ the factorials diverge; but it can be made to converge by resummation, which reproduces (22).

Hyperasymptotics consists of iterating (22), with optimal truncations, and substituting at each stage the truncated asymptotic series for the I_l, a procedure that can be regarded as multiple scattering among the saddles. Each iteration produces an exponential improvement. The result is, in the general case[3], an intricate sequence of truncated asymptotic series, eventually involving all the saddles (i.e. not just those which contribute to (22) at a given stage). These series involve the original asymptotic coefficients T_{jr}, which of course depend on the particular function being approximated, and certain 'generalised terminants', in the form of multiple integrals, which are universal functions of the singulants F_{jl}. The advantage of the integral (22) over the 'pure asymptotic' resurgence (24) is that at every stage of hyperasymptotics there is an explicit expression for the remainder, which will, we anticipate, be indispensable in constructing rigorous error bounds.

Hyperasymptotics comes to a natural halt, because each hyperseries is shorter than its predecessor, and eventually contains only one term. The decreasing length is a consequence of the 'live now, pay later' philosophy, natural in asymptotics, that the terms must continue to decrease, not only within each hyperseries but from each hyperseries to the next. A typical result[2] is that hyperasymptotics reduces the ultimate error from $\exp\{-|F|\}$ (superasymptotics) to $\exp\{-(1+2\log 2)|F|\}\approx\exp\{-2.386|F|\}$; thus the error is reduced to less than its square.

To illustrate hyperasymptotics, we again employ Ai, defined by (2) This example is special because there are only two saddles (and therefore only one singulant $F_{jl}= -F_{lj}\equiv F$). and hyperasymptotic multiple scattering is simply back and forth between them. Moreover, the pre-exponential factor is $G=1$. The terms in the dominant and subdominant series, including the prefactors M in (1), are the same (apart from signs) and the general resurgence (24) reduces in the notation of (1) and (5) to

$$Y_r = \frac{1}{2\pi F^r}\sum_{s=0}^{\infty}(r-s-1)!(-F)^s Y_s \tag{26}$$

This relation holds not only for integrals with two saddles but also - as was discovered by Dingle[8] - for the coefficients in the asymptotic expansion of solutions of second-order linear ordinary differential equations with a single transition point[2,17]. Its compactified form - a special case of (22) - is equivalent to the self-Stieltjes transform relation for Airy functions, as noticed and exploited by Boyd[18]. In this case, each hyperseries is half the length of its predecessor, so hyperasymptotics stops after $\mathrm{Int}\log_2|F|$ stages, with a total of $2\mathrm{Int}|F|$ terms (as opposed to $\mathrm{Int}|F|$ terms in the zeroth stage of hyperasymptotics, namely superasymptotics).

We express the numerical results in terms of $Y(z)$, defined (cf.5) by

$$\mathrm{Ai}(z) \equiv \frac{1}{2z^{1/4}\sqrt{\pi}}\exp\left\{\frac{F}{2}\right\}Y(F) \tag{27}$$

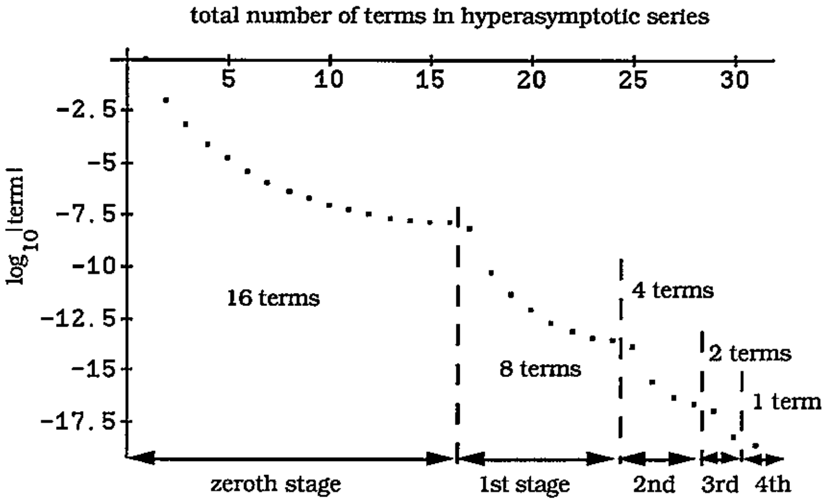

Fig. 6. Decrease of the terms in the five hyperseries constituting hyperasymptotics for $Y(-16)$ for the Airy function.

For F=-3 (i.e. z= 1.7171), the exact value is $Y(-3)$=0.96419.... Thus the lowest approximation$Y$~1 (i.e. Ai1 - cf. (4)) is in error by 0.036. Superasymptotics, i.e.Ai2 (equation 5) with N=4, gives Y≈0.95895..., an error of -0.00524, about ten times better. With a single stage of hyperasymptotics we do much better: Y≈0.96410..., an error of -0.00009, about 60 times better than superasymptotics

The improvements are even more dramatic for larger $|F|$. Fig. 6 (taken from [2]) shows the magnitudes of the terms in the five hyperseries for Ai when F=-16, that is z=+5.2414827884177932413..., corresponding to θ=0 (fig. 1). The exact value of $Y(-16)$ is 0.99183679918826259891....Thus the lowest approximation Y~1 is in error by 8.163×10^{-3}. Superasymptotics, i.e. Ai2 with 16 terms, gives Y≈0.99183679351132345911..., an error of -5.677×10^{-9}. With hyperasymptotics, taken to its natural halt, we obtain Y≈0.99183679918826260060...., an error of 1.151×10^{-18} (close to that predicted theoretically). Similar accuracy is obtained on the Stokes line θ=120° (fig.1) and on the anti-Stokes line θ=180°.(where, for example, hyperasymptotics can be employed to solve the 'eigenvalue problem' of determing the zeros of Ai).

5. OUTLOOK

There are several scientific reasons, as well as purely mathematical ones, for seeking such detailed understanding of the relatively simple asymptotic problems I have been discussing. One is that asymptotics is often deeply involved in the conections between physical theories[28]. It is common for a more general theory to 'reduce' to a less general theory when some parameter vanishes. For example, special relativity reduces to Newtonian mechanics as the particle speed vanishes; wave optics reduces to ray optics as the wavelength vanishes; quantum mechanics reduces to classical mechanics as Planck's constant vanishes; Navier-Stokes fluid motion reduces to Eulerian flow as the viscosity vanishes; statistical mechanics reduces to thermodynamics as the reciprocal number of particles vanishes, etc., etc. Only in the first of these examples is the limit regular and the expansion in the small parameter convergent. In all the other cases, the limits are singular and lead to divergent series. Associated with the singularities are important phenomena such as ray caustics, turbulence and critical behaviour. This connection between asymptotics and theory reduction, not appreciated by philosophers (at least in my experience), is sufficient reason to try to understand divergent series as deeply as possible.

A more concrete reason is that in practice we often encounter asymptotics which is intrinsically more complicated than what I have been describing here, and there is no hope of understanding the complicated problems unless we penetrate the simpler ones first. One important extension would be to the case of *many exponentials*. In Stokes' phenomenon and its smoothing (§§3 and 4) only two exponentials are essentially involved: the dominant and leading subdominant ones. With the resurgence relation (22) we make the first step towards the consistent treatment of many exponentials . However, the integral (18) only involves a single variable, and the topology involved in the derivation[3] of (22) suggests that the results might not be the same when many exponentials arise in the multiple integrals of diffraction theory, or the infinite-dimensional functional integrals of quantum mechanics, statistical mechanics and field theory.

Another complication is that when many exponentials appear they need not always be ordered in a dominance hierarchy: the singulants can all be imaginary. An important class of such problems occurs in quantum chaology[28], that is the semiclassical asymptotics of quantum systems whose classical counterparts have chaotic trajectories. There, the exponentials appearing in the asymptotic expansion of (for example) the density of energy levels are associated with classical periodic orbits. Even in lowest order, where each exponential is included bare - that is, without its correction terms - the proliferation of periodic orbits makes the sum diverge, in ways that are still mysterious and probably related to the Riemann hypothesis of arithmetic. There are strong hints[28,29] that at this higher level resurgence may again prove to be an important guiding principle, this time relating the exponentials associated with long and short orbits rather than the late and early terms of the series associated with individual orbits.

There are many directions for further research. One is to break through the $\exp\{-2.386|F|\}$ barrier (ultraasymptotics?). Another is to extend hyperasymptotics to multiple integrals, and to the Schrödinger equation wiith many transition points (the work of Balian and Bloch[30], Knoll and Schaeffer[31], and Voros[25] could be helpful here). Yet another is to provide rigorous mathematical underpinning for the results already obtained, hopefully going beyond proofs for particular functions[16,18,19] to genericity theorems which would delineate the universality class for which the results (for example the smoothing (15)) are valid. Going further, it would be helpful to know something about the late terms of the asymptotic series corresponding to each of the classical closed orbits in quantum chaology (for example whether the divergences of these individual series are related to the divergence of the sum over lowest-order exponentials for all the orbits).

REFERENCES

1. M.V. Berry, Uniform asymptotic smoothing of Stokes's discontinuities, *Proc. Roy. Soc. Lond*, A422: 7-21(1989).
2. M.V. Berry and C.J.Howls, Hyperasymptotics,*Proc. Roy. Soc. Lond*, A430: 653-668 (1990).
3. M.V. Berry and C.J. Howls, Hyperasymptotics for integrals with saddles, submitted to *Proc. Roy. Soc. Lond.* (1991).
4. G.H. Hardy, "Divergent Series", Clarendon Press, Oxford (1949).
5. G.G.Stokes, On the numerical calculation of a class of definite integrals and infinite series, *Trans. Camb. Phil. Soc.* 9: 379-407 (1847).
6. B.G. Levi, A super time to renormalize, *Physics Today*, April 1988: 25.
7. Lord Kelvin, The scientific work of Sir George Stokes, *Nature* 67: 337-338 (1903).
8. R.B.Dingle, "Asymptotic Expansions: their Derivation and Interpretation", Academic Press, New York and London (1973).
9. G.G.Stokes, On the discontinuity of arbitrary constants which appear in divergent developments, *Trans. Camb. Phil. Soc.* 10: 106-128 (1864).
10. J.E. Littlewood , preface to ref.4.
11. N.G. de Bruijn , "Asymptotic methods in analysis", North-Holland, Amsterdam (1958).
12. M.V. Berry, Stokes' phenomenon; smoothing a Victorian discontinuity, *Publ. Math.of the Institut des Hautes Études scientifique*, 68: 211-221 (1989).
13. R.Balian, G.Parisi, and A.Voros, Discrepancies from asymptotic series and their relation to complex classical trajectories, *Phys. Rev, Lett.* 41: 141-1144(1978).
14. F.W.J. Olver, "Asymptotics and special functions" Academic Press, New York and London (1974).

15. M.Abramowitz and I.A.Stegun,1964, "Handbook of mathematical functions", National Bureau of Standards, Washington (1964).
16. F.W.J. Olver, On Stokes' phenomenon and converging factors, in "Proceedings of International Symposium on Asymptotic and Computational Analysis (Manitoba, Winnipeg 1989)", R.Wong, ed., Marcel Dekker, New York (1990), pp329-355.
17. M.J. Rakovic, and E.A.Solov'ev, Higher orders of semiclassical expansion for the one-dimensional Schrödinger equation, *Phys. Rev.* A40: 6692-6694 (1989).
18. W.G.C. Boyd , Stieltjes transforms and the Stokes phenomenon, *Proc. Roy. Soc.* Lond., A429: 227-246 (1990).
19. D.S.Jones, Uniform asymptotic remainders, in "Proceedings of International Symposium on Asymptotic and Computational Analysis (Manitoba, Winnipeg 1989)" R. Wong, ed., Marcel Dekker, New York (1990), pp241-264.
20. M.V. Berry, Waves near Stokes lines, *Proc. Roy. Soc. Lond*, A427: 265-280 (1990).
21. M.V. Berry, Histories of adiabatic quantum transitions, *Proc. Roy. Soc. Lond*, A429: 61-72 (1990).
22. R. Lim and M.V. Berry, Submitted to *J.Phys.A.* (1991).
23. M.V. Berry and C.J. Howls, Stokes surfaces of diffraction catastrophes with codimension three, *Nonlinearity*, 3:281-291 (1990).
24. J.Écalle, "Les fonctions résurgentes " (3 volumes) Publ. Math. Université de Paris-Sud (1981), and "Cinq applications des fonctions résurgentes ", Preprint 84T62, Orsay (1984).
25. A.Voros, The return of the quartic oscillator. The complex WKB method, *Ann. Inst H. Poincaré*, 39:211-338 (1983).
26. P. Flagolet and A.M. Odlyzko, Singularity analysis of generating functions, *SIAM J. Discrete Math.*, 3 (2): 216-240 (1990).
27. P2C2E=processes too complicated to explain, see S. Rushdie, "Haroun and the sea of stories", Granta, London (1990).
28. M.V. Berry , Some quantum-to-classical asymptotics, in "Chaos and quantum physics",Les Houches Lecture Series 52, M J Giannoni and A Voros, eds. ,North-Holland, Amsterdam (1991).
29. M.V. Berry and J.P. Keating, A rule for quantizing chaos?, *J. Phys. A.*, 23: 4839-4849 (1990).
30. R. Balian and C. Bloch, Solution of the Schrödinger equation in terms of classical paths, *Ann. Phys.* (*N.Y.*) 85: 514-545 (1974).
31. J. Knoll and R. Schaeffer, Semiclassical scattering theory with complex trajectories. 1. Elastic waves, *Ann. Phys.* (*N.Y.*), 97: 307 (1976).
32. M.V. Berry, Infinitely many Stokes smoothings in the Gamma Function, submitted to *Proc. Roy. Soc. Lond*, (1990).