Stokes's phenomenon for superfactorial asymptotic series†

By M. V. Berry

H. H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, U.K.

Superfactorial series depending on a parameter z are those whose terms $a(n,z)$ grow faster than any power of $n!$. If the terms get smaller before they increase, the function $F(z)$ represented by $\Sigma_0^\infty a(n,z)$ will exhibit a Stokes phenomenon similar to that occurring in asymptotic series whose divergence is merely factorial: across 'Stokes lines' in the z plane, where the late terms all have the same phase, a small exponential switches on in the remainder when the series is truncated near its least term. The jump is smooth and described by an error function whose argument has a slightly more general form than in the factorial case. This result is obtained by a method which is heuristic but applies to superfactorial series where Borel summation fails. Several examples are given, including an analytical interpretation of the sum, and a numerical test of the error-function formula, for the function represented by

$$F(z) = \sum_0^\infty \exp\{n^2/A - 2nz\}, \quad \text{where } A \gg 1.$$

1. Introduction

Many functions $F(z)$ occurring in mathematics and physics are represented formally by divergent asymptotic series, that is

$$F(z) = \sum_{n=0}^{\infty} a(n,z) \tag{1}$$

in which the terms $|a(n,z)|$ get smaller and then increase. Truncations of the series, optimally at the least term, often provide very accurate approximations to $F(z)$. For reasons explained by Dingle (1973), a common situation is for the late terms to grow as

$$a(n,z) \sim (\alpha n+\beta)!/z^n \quad (n \gg 1), \tag{2}$$

where α and β are constants; examples of functions in this universality class are solutions of linear second-order differential equations and integrals over exponentials containing a large parameter. In such cases the remainder when the series is truncated near its least term is a small exponential whose coefficient jumps as z crosses a 'Stokes line', characterized by the late terms all having the same phase. This jump is the phenomenon discovered by Stokes (1864). For (2) the Stokes line is the positive real z axis. Recently the jump was examined more closely by careful Borel summation of the tail of (1), and found to be not discontinuous but smoothed

† This paper was accepted as a rapid communication.

according to an error function whose argument has a universal form (Berry 1989). This result has been confirmed for several particular cases by techniques not involving summation of a divergent series (Olver 1990; Boyd 1990; Berry 1990*a*, *b*; Berry & Howls 1991; Paris & Wood 1991).

My purpose here is threefold. First, to show (§2) how the error function persists beyond the universality class (2), and describes the switching on of small exponentials in remainders of series which diverge much faster. Examples (§3) are $a \sim \exp(n^m)$ $(m > 1)$ and $a \sim \exp(\exp(n))$ (which should be compared with (2), for which $a \sim \exp(\alpha n \ln n)$). Such superfactorial divergences arise, for example, in probability generating functions on lattices. It should be noted that superfactorial growth is faster than that in the Gevrey series, where in (2) $(\alpha n+\beta)!$ is replaced by $[(\alpha n+\beta)!]^k$ (Ramis 1984; Ramis & Martinet 1990; Thomann 1990) and where the growth is essentially factorial, with $a \sim \exp(k\alpha n \ln n)$. We shall find a smoothing formula closely resembling that for factorially divergent series, but the argument in the error function has a more general form.

Second, to draw attention to the very simple – albeit non-rigorous – method used to obtain the Stokes smoothing. As well as being applicable to superfactorial series, where Borel summation is impotent, this has the merit of demystifying the error function by making its origin obvious. Moreover there are intriguing connections with the 'subtractions' employed to remove infinities in quantum field theory. That such a method might work was suggested to me (for the factorial case) by Martin Kruskal.

Third, to stimulate mathematicians to go beyond the study of Gevrey series and investigate superfactorial asymptotics. For example, a rigorous basis should be established for determining when divergent series (1) can be 'interpreted' to define unique functions $F(z)$. And the heuristic (and simplistic) manipulations I shall use to obtain the Stokes smoothing should be supplemented by properly justified procedures with strict error bounds. That such procedures exist is strongly suggested by the example in §4, where a particular superfactorial series (a generalized theta function) is summed (by a non-Borel method) and the error-function smoothing is shown to be numerically accurate.

2. Generalized smoothing formula

In (1) we assume that when $z = x$, where $x \equiv \operatorname{Re} z$ is positive, all terms $a(n, x)$ are real and positive. Thus in the z plane the Stokes line is the positive real axis, and we seek to understand the behaviour of $F(z)$ across this line. Moreover we make the stipulation, whose significance will emerge later, that when z is real the function $F(x)$ represented by the infinite series is real. Let $N(x)$ be the minimum of $a(n, x)$; sometimes we shall simply write N instead of $N(x)$. We assume that we are in a régime where asymptotics is relevant, that is $N \gg 1$, and define the optimally truncated sum $F_{\mathrm{T}}(z)$ and remainder $R(z)$ by

$$F(z) = \sum_{n=0}^{\mathrm{Int}[N(x)-1]} a(n,z) + \sum_{n=\mathrm{Int}[N(x)]}^{\infty} a(n,z)$$
$$\equiv F_{\mathrm{T}}(z) + R(z). \tag{3}$$

Thus, defining

$$a(n,z) \equiv \exp\{b(n,z)\}, \tag{4}$$

we seek to interpret the remainder

$$R(z) = \sum_{\mathrm{Int}[N]}^{\infty} \exp\{b(n,z)\} \tag{5}$$

and find its lowest-order approximation across the Stokes line.

The first step is to replace the sum by an integral:

$$R(z) \approx \int_N^\infty \mathrm{d}n \exp\{b(n,z)\}. \tag{6}$$

Next, the exponent is approximated near its saddle, at $N(z)$, satisfying

$$\partial_n b(n,z) = 0 \quad \text{if} \quad n = N(z). \tag{7}$$

For complex z this is a complex number, to be distinguished from the real limit $N = N(x)$ of summation or integration. It is convenient to define

$$B(z) \equiv b(N(z),z); \quad B_2(z) \equiv \partial_{nn} b(N(z),z). \tag{8}$$

Thus

$$R(z) \approx \exp\{B(z)\} \int_N^\infty \mathrm{d}n \exp\{\tfrac{1}{2}[N(z)-n]^2 B_2(z)\}. \tag{9}$$

An obvious change of variable now gives

$$R(z) \approx \mathrm{i}\sqrt{\left(\frac{2}{B_2(z)}\right)} \exp\{B(z)\} \int_{\mathrm{i}\infty}^{\mathrm{i}[N-N(z)]\sqrt{(B_2(z)/2)}} \mathrm{d}t \exp\{-t^2\}. \tag{10}$$

The upper limit can be simplified by expanding to lowest order in the distance y from the Stokes line; using (7) we find

$$\mathrm{i}[N-N(z)]\sqrt{(\tfrac{1}{2}B_2(z))} \approx -y\,\partial_{nx} b(N(x),x)/\sqrt{(2B_2(x))}. \tag{11}$$

Thus if we define the 'Stokes multiplier' by

$$R(z) \equiv \mathrm{i}\sqrt{(2\pi/B_2(x))} \exp\{B(z)\} S(z), \tag{12}$$

we obtain the approximation

$$S(z) \approx \frac{1}{\sqrt{\pi}} \int_{\mathrm{i}\infty}^{\sigma(x,y)} \mathrm{d}t \exp\{-t^2\}, \tag{13}$$

where

$$\sigma(x,y) \equiv -y\,\partial_{nx} b(N(x),x)/\sqrt{(2B_2(x))}. \tag{14}$$

(Simple analysis based on (11) shows that the same result follows if $|N(z)|$ or the real part of the minimum of $b(n,z)$, rather than $N(x)$, is used as truncation limit in (3).)

Note that because of the lower limit this integral is divergent, reflecting the divergence of the original series and the fact that we are integrating through a minimum, rather than the customary maximum. Therefore the Stokes multiplier is still unspecified. This is not a defect of our summation method, but reflects an inevitable ambiguity in the interpretation of the series (1). To see this, observe that in (12) the exponential is small, corresponding to the smallest term of the sequence $a(n,z)$ (alternatively stated, $B(z) \ll b(0,z)$). Now, Stokes's phenomenon does not provide an absolute specification of the coefficient at any given z, because this information is not encoded in the bare series (1) but must be supplied from outside. Rather, the phenomenon describes the *change* in the coefficient of the small exponential across the Stokes line.

It is just this change which *is* uniquely given by (13), because the infinity cancels on subtracting from $S(z)$ its value at some reference point z_0, leaving

$$S(z)-S(z_0) \approx \tfrac{1}{2}\operatorname{erf}\{\sigma(x,y)\}-\tfrac{1}{2}\operatorname{erf}\{\sigma(x_0,y_0)\}. \tag{15}$$

There is an analogy here with quantum field theory (Itzykson & Zuber 1980), where similar subtractions of 'identical infinities' are used to produce finite and consistent values for physical quantities.

Now recall that for the series (1) we are providing the supplementary information that the function $F(z)$ as real on the Stokes line $y = 0$. Because of the factor i in (12), this suffices to disambiguate the multiplier, and we obtain the main result

$$S(z) \approx \tfrac{1}{2}\operatorname{erf}\{\sigma(x,y)\}. \tag{16}$$

(Of course, other definitions are possible, and occur in applications: a common situation is that the multiplier is zero far 'below' the Stokes line, i.e. at $y_0 = -\infty$.)

With this method of derivation the origin of the universal error function is clearly exposed as the integration of a gaussian (in equation (10)) from a limit of integration near to, but not coinciding with, its minimum.

One way to express the smoothing is as an asymptotic limiting form for the rescaled remainder from the optimally truncated sum:

$$-\mathrm{i}\sqrt{(B_2(x)/2\pi)}\exp\{-B(z)\}[F(z)-F_{\mathrm{T}}(z)] \approx \tfrac{1}{2}\operatorname{erf}\{\sigma(x,y)\}, \quad \text{if} \quad N(x) \gg 1. \tag{17}$$

Note that the approximation to the remainder given by (12), (14) and (16) is imaginary near the Stokes line and vanishes on it. Of course the exact remainder includes real terms, whose magnitude is comparable with $a(N(x,)\,x)$. However, these terms will be small compared with the error-function contribution if $|B_2(x)| \ll 1$, and we shall confirm that this is the case in the examples which follow.

3. Examples

(*a*) *Factorial divergence*

We choose here the simplest Gevrey series, a special case of (2) which arises naturally in, for example, the method of steepest descent for integrals with saddles (Berry & Howls 1991), where

$$a(0,z) = 1, \quad a(n,z) \underset{n\to\infty}{\longrightarrow} (n-1)!/2\pi z^n. \tag{18}$$

(In such applications z is called the 'singulant'; it could denote, for example, the difference in heights of two saddles, or the difference between the actions of two classical paths.) From Stirling's formula and (4) we obtain

$$b(n,z) = (n-\tfrac{1}{2})\ln n - n - n\ln z - \ln\sqrt{(2\pi)} \tag{19}$$

and thence, from (7) and (8)

$$\left.\begin{aligned} N(z) = z, \quad B(z) &= -\tfrac{1}{2}\ln z - z - \ln\sqrt{(2\pi)}, \\ B_2(z) &= -\partial_{nz} b(N(z),z) = 1/z. \end{aligned}\right\} \tag{20}$$

The asymptotic condition $N \gg 1$ is satisfied if $|z| \gg 1$, and then the condition $|B_2(x)| \ll 1$, necessary for the error function to dominate the remainder, is satisfied too.

The remainder now follows from (12)–(14) and (16) as

$$\begin{aligned} R(z) &\approx \mathrm{i}\exp(-z)\times\tfrac{1}{2}\operatorname{erf}(y/\surd(2x)) \\ &= \mathrm{i}\exp(-z)\times\tfrac{1}{2}\operatorname{erf}(\operatorname{Im} z/\surd(2\operatorname{Re} z)), \end{aligned} \tag{21}$$

where the second form of writing is included to emphasize the relation between the exponent and the argument of the error function. In this case the simple general procedure has reproduced the result already obtained (Berry 1989) by Borel summation. Simplicity is the only virtue here, because the method fails to suggest the many refinements, going far beyond the error-function approximation, that Borel summation can provide.

(*b*) *Exponential-power divergence*

Hence the function $b(n,z)$ in (4) is chosen as

$$b(n,z) = Kn^m - nz, \tag{22}$$

where K and m are positive constants. The series (1) is superfactorial when $m > 1$. From (7) and (8) we find

$$\left.\begin{aligned} N(z) &= (z/mK)^{1/(m-1)}, \quad B(z) = -(m-1)K[N(z)]^m, \\ \partial_{nz} b(n,z) &= -1, \quad B_2(z) = m(m-1)K[N(z)]^{m-2}. \end{aligned}\right\} \tag{23}$$

The asymptotic and error-function-domination conditions $N \gg 1$ and $|B_2(x)| \ll 1$ are fulfilled if $K \ll 1$. The argument of the error function is, from (14),

$$\sigma(x,y) = y/\surd\{2m(m-1)K[N(x)]^{m-2}\}. \tag{24}$$

A short calculation now shows that the remainder can be written

$$R(z) \approx \mathrm{i}\sqrt{\left(\frac{2\pi}{Km(m-1)}\right)}\left(\frac{mK}{x}\right)^{\frac{m-2}{m-1}} \exp\{-[-B(x)]\}\times\tfrac{1}{2}\operatorname{erf}\left\{\frac{\operatorname{Im}[-B(z)]}{\surd\{2m\operatorname{Re}[-B(z)]\}}\right\}. \tag{25}$$

Thus apart from the factor $1/\surd m$ the argument of the error function bears the same relationship to the exponent as in the familiar case of factorial divergence. Because of the $1/\surd m$, the effect of superfactorial divergence is to make the Stokes smoothing wider than in the factorial case.

(*c*) *Exponential-exponential divergence*

Here the function $b(n,z)$ in (4) is chosen as

$$b(n,z) = A\exp(Cn) - nz, \tag{26}$$

where A and C are positive constants. The quantities entering into the smoothing formula are

$$\left.\begin{aligned} N(z) = \frac{1}{C}\ln\left(\frac{z}{AC}\right), \quad B(z) = -\frac{z}{C}\left[\ln\left(\frac{z}{AC}\right)-1\right], \\ \partial_{nz} b(n,z) = -1, \quad B_2(z) = Cz. \end{aligned}\right\} \tag{27}$$

The asymptotic and error-function-domination conditions $N \gg 1$ and $|B_2(x)| \ll 1$ are fulfilled if $C \ll 1$.

Thus the remainder is

$$R(z) \approx \mathrm{i}\sqrt{\left(\frac{2\pi}{Cz}\right)}\exp\left\{-\frac{z}{C}\left[\ln\left(\frac{z}{AC}\right)-1\right]\right\}\times\tfrac{1}{2}\operatorname{erf}\left\{\frac{y}{\surd(2Cx)}\right\}. \tag{28}$$

For comparison with the form for factorial divergence we write this as

$$R(z) \approx \mathrm{i}\sqrt{(2\pi/Cx)}\exp\{-[-B(z)]\} \times \tfrac{1}{2}\operatorname{erf}\{\operatorname{Im}[-B(z)]/\sqrt{\{2\ln(x/AC)\operatorname{Re}[-B(z)]\}}\}. \quad (29)$$

Again the form is very similar to the factorial case, the difference here being the logarithmic factor in the denominator. This further increases the width of the Stokes smoothing in comparison with the exponential-power case, presumably because the superfactorial divergence is stronger.

4. Numerical illustration

For detailed study we choose the function defined by interpreting

$$F(z) \equiv \sum_{n=0}^{\infty} \exp\left\{\frac{n^2}{A} - 2nz\right\}, \quad (30)$$

where A is positive real and $x = \operatorname{Re} z > 0$. This is the special case $m = 2$ of the series considered in §3*b*. If A were negative, $F(z)+F(-z)$ gives the series for one of the Jacobi theta functions. For the minimum of the summand we find

$$N(z) = Az \quad (31)$$

and from §3*b* the remainder is

$$R(z) \approx \mathrm{i}\sqrt{(\pi A)}\exp\{-Az^2\} \times \tfrac{1}{2}\operatorname{erf}\{y\sqrt{A}\}. \quad (32)$$

We can expect this to be a good approximation when $A \gg 1$.

A good approximation to what? The answer requires an interpretation of the divergent series (30) in conventional terms. Borel summation, that is inserting the identity

$$1 = \frac{1}{n!}\int_0^{\infty} \mathrm{d}t\, t^n \exp(-t) \quad (33)$$

into (30) and exchanging summation and integration, fails because the resulting series still diverges. However, the stronger identity

$$1 = \exp\left\{\frac{n^2}{A}\right\}\sqrt{\left(\frac{A}{\pi}\right)}\int_{-\infty}^{\infty} \mathrm{d}t \exp(-At^2+2nt) \quad (34)$$

yields a formally summable geometric series and leads to the following convergent representation as a convergent principal-value integral

$$F(z) = \sqrt{\left(\frac{A}{\pi}\right)}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\exp\{-A(t+z)^2\}}{1-\exp\{2t\}}. \quad (35)$$

With this interpretation, we can find an exact expression to the remainder by splitting off the first N terms of the series (30), using the identity

$$\frac{1}{1-\exp\{2t\}} = \sum_{n=0}^{N-1} \exp\{2nt\} + \frac{\exp\{2Nt\}}{1-\exp\{2t\}}. \quad (36)$$

Elementary manipulations give

$$R(z) = \sqrt{\left(\frac{A}{\pi}\right)}\exp\{-Az^2\}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\exp\{-At^2+2(N-Az)t\}}{1-\exp\{2t\}}. \quad (37)$$

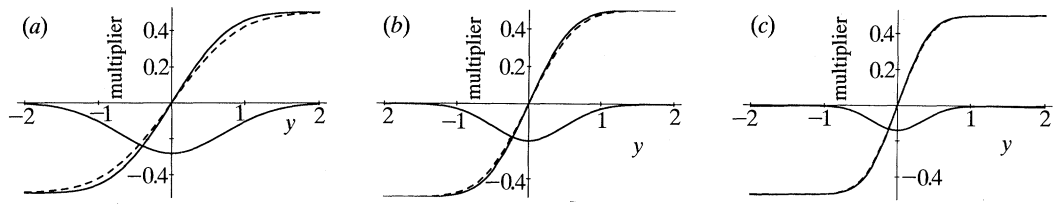

Figure 1. Real and imaginary parts of the Stokes multiplier $S(1+\mathrm{i}y)$ across the Stokes line $y=0$, computed from the integral representation (left-hand side of (38)) as explained in Appendix A, for (*a*) $A=1$, (*b*) $A=2$, (*c*) $A=4$. The odd curves show $\operatorname{Re} S$ and the even curves show $\operatorname{Im} S$; the dashed curves show the error function approximation (right-hand side of (38)).

The Stokes multiplier defined by (12) can now be isolated, and our main result gives the following prediction for its asymptotic limit:

$$S(z) = -\frac{\mathrm{i}}{\pi}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\exp\{-At^2+2(N-Az)\,t\}}{1-\exp\{2t\}} \underset{A\to\infty}{\longrightarrow} \tfrac{1}{2}\operatorname{erf}\{y\surd A\}. \tag{38}$$

An analytical demonstration of this smoothing formula, independent of the heuristic arguments of §2, is based on the fact that for optimal truncation ($N \approx |Az|$) the saddle of the integrand in (38) lies close to the pole at $t=0$. The argument is very similar to that for the Borel integral in the factorial case (Berry 1989).

Instead of repeating it, I give in figure 1 a numerical illustration, based on computing the real and imaginary parts of the Stokes multiplier $S(z)$ using the integral representation on the left-hand side of (38). The method used to compute $S(z)$ is explained in appendix A. Computations were carried out for $z=1+\mathrm{i}y$ and $A=1,2$ and 4. The agreement between $\operatorname{Re} S(z)$ and $\frac{1}{2}\operatorname{erf}(y\surd A)$ is very good – even for $A=1$, when the asymptotic series contains just the single term $n=0$ – and as expected it improves with A.

The imaginary part of the multiplier is smaller than the real part, and although appreciable for $A=1$ it gets smaller as A increases. Its magnitude is largest on the Stokes line, where (38) can, fortuitously, be evaluated exactly when Ax is an integer, confirming the decrease and exhibiting its precise form:

$$S(\text{Integer}/A) = -\frac{\mathrm{i}}{\pi}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\exp\{-At^2\}}{1-\exp\{2t\}} = -\frac{\mathrm{i}}{2\surd(\pi A)}. \tag{39}$$

A slight generalization is that for large A the multiplier anywhere on the Stokes line is

$$S((\text{Integer}+\delta)/A) \approx -[\mathrm{i}/2\surd(\pi A)]\,(1+\delta) \quad (0 \leqslant \delta < 1). \tag{40}$$

Calculation based on (12) confirms that the corresponding remainder is of the same order as the first term omitted in the optimal truncation of (30).

Appendix A

This is the expression of the principal-value integral (38) in a form suited to computation. Splitting the integral at its singularity gives

$$S(z) = -\frac{\mathrm{i}}{\pi}\int_{\epsilon}^{\infty} \mathrm{d}t\,\exp\{-At^2\}\left[\frac{\exp\{-2(N-Az)\,t\}}{1-\exp\{-2t\}} + \frac{\exp\{2(N-Az)\,t\}}{1-\exp\{2t\}}\right], \tag{A 1}$$

where ϵ is a positive infinitesimal. Expanding the denominators in powers of $\exp(-2t)$ and interchanging summation and integration gives

$$S(z) = -\frac{\mathrm{i}}{2\sqrt{(\pi A)}} \sum_{n=0}^{\infty} [\mathrm{erfc}\{Q_n^+\} \exp\{(Q_n^+)^2\} - \mathrm{erfc}\{Q_n^-\} \exp\{(Q_n^-)^2\}],$$

where

$$Q_n^+ \equiv (n+N-Az)/\sqrt{A} \quad \text{and} \quad Q_n^- \equiv (n+1-N+Az)/\sqrt{A}. \tag{A 2}$$

This series of complementary error functions (not related to the error function in the smoothing formula (38)) is convergent, but only slowly: for $n \gg 1$ the terms in the sum decrease as n^{-2}. Convergence can be substantially hastened by approximating the sum of the tail by its integral, and this trick was used in computing figure 1.

I thank Martin Kruskal for suggesting that the error function might emerge from integrating a divergent series near its minimum, and Christopher Howls and William Boyd for a careful reading of the typescript.

References

Berry, M. V. 1989 *Proc. R. Soc. Lond.* A **422**, 7–21.

Berry, M. V. 1990*a* *Proc. R. Soc. Lond.* A **427**, 265–280.

Berry, M. V. 1990*b* *Proc. R. Soc. Lond.* A **429**, 61–72.

Berry, M. V. & Howls, C J. 1991 *Proc. R. Soc. Lond.* A **434**, 657–675.

Boyd, W. G. C. 1990 *Proc. R. Soc. Lond.* A **429**, 227–246.

Dingle, R. B. 1973 *Asymptotic expansions: their derivation and interpretation.* New York and London: Academic Press.

Itzykson, C. & Zuber, J. B. 1980 *Quantum field theory.* Singapore: McGraw-Hill.

Olver, F. W. J. 1990 In *Proc. Int. Symp. on Asymptotic and Computational Analysis* (ed. R. Wong), pp. 329–355. New York: Marcel Dekker.

Paris, R. B. & Wood, A. D. 1991 *J. Comp. appl. Math.* (In the press.)

Ramis, J. P. 1984 *Mem. Am. math. Soc.* no. 296.

Ramis, J. P. & Martinet, J. 1990 In *Computer algebra and differential equations* (ed. E. Tournier), pp. 117–214. New York and London: Academic Press.

Stokes, G. G. 1864 *Trans. Camb. phil. Soc.* **10**, 106–128. (Reprinted in *Mathematical and physical papers by the late Sir George Gabriel Stokes* vol. IV, pp. 77–109. Cambridge University Press 1904.

Thomann, J. 1990 *Numer. Math.* **58**, 503–535.

Received 8 July 1991; accepted 16 August 1991