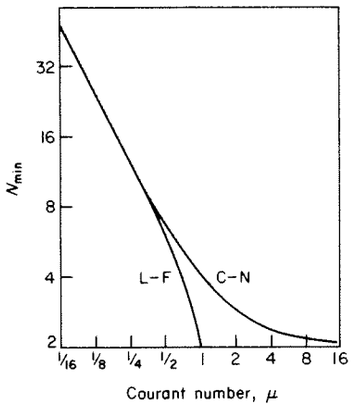

FIG. 4. Variation of the cut-off sampling rate with Courant number in the Leap-frog and Crank–Nicolson schemes.

5. Conclusion

The existence of the type of damping described in Section 4 has been noted before in the application of the Leap-frog scheme to the one-dimensional advection equation (Roache, 1972). To the author's knowledge the effect has not been explained quantitatively hitherto. The analysis given herein represents such an explanation. Damping is shown to occur in both explicit and implicit solutions of the shallow water equations.

The analysis technique outlined in Sections 3 and 4 consists of recasting the usual von Neumann stability analysis in terms of a dispersion relation between the frequency and wave-number of the Fourier solution. The dispersion relation can then be studied from two angles:

(a) Consider the wave-number to be real. This is equivalent to the von Neumann analysis and is useful in studying how the initial conditions develop in time.
(b) Consider the frequency to be real. This determines how boundary information propagates spatially.

Although only the example of the one-dimensional wave equation has been considered here, the approach can be applied equally well to finite difference representations of other linear differential equations.

The work described in this paper has been carried out as part of the research program of the Hydraulics Research Station, United Kingdom, and is published with the permission of the Director of Hydraulics Research.

REFERENCES

RICHTMEYER, R. D. & MORTON, K. W. 1967 *Difference Methods for Initial-Value Problems.* New York: Interscience Publishers.

ROACHE, P. J. 1972 *Computational Fluid Dynamics.* Albuquerque, N.M.: Hermosa Publishers.

J. Inst. Maths Applics (1979) **23,** 97–120

The Accurate Numerical Inversion of Laplace Transforms

A. TALBOT

Department of Mathematics, Brunel University, Uxbridge

[Received 28 May 1976 and in revised form 16 June 1978]

Numerical inversion of almost arbitrary Laplace transforms, for any value of t, to any prescribed accuracy up to at least three-quarters of the computer precision, is effected by trapezoidal integration along a special contour. The required number of points depends on t, the accuracy, and the transform singularity positions, and for moderate t is typically 11 for errors of order 10^{-6}, 18 for order 10^{-10}, 35 for order 10^{-20} (with double precision working).

1. Introduction

THE INVERSION of Laplace transforms is a topic of fundamental importance in many areas of applied mathematics. In the more standard applications the inversion can be accomplished by the use of a dictionary of transforms, or in the case of rational function transforms by partial fraction decomposition. Where these methods are of no avail recourse may be had to the inversion integral formula, which is likely to lead to an intractable integral, or to an infinite series, often with terms involving the roots of some transcendental function. It is clear that in all but the simplest cases considerable effort is needed to obtain an accurate numerical value of the inverse for a specified value of the argument.

It is therefore natural that attention has been paid by mathematicians, engineers, physicists and others to alternative ways of evaluating the inverse. Early methods (e.g. Widder, 1935; Tricomi, 1935; Shohat, 1940) involved expansion of the inverse in series of Laguerre functions. Salzer (1955) evaluated the inversion integral by Gaussian quadrature using an appropriate system of orthogonal polynomials. Since 1955 a very large number of methods for numerical inversion have been published, see for example the very full bibliography in Piessens (1975).

Many of the methods use either orthogonal series expansions, or weighted sums of values of the transform at a set of points, usually complex points. In either case considerable preliminary work must be carried out. In the second type this may be done in advance once and for all for each selected set of points, and the points and weights stored in the computer. However, if more points are desired for the sake of getting increased accuracy, much further computational effort must be expended first.

In general the methods hitherto published have been intended for use with transforms of particular types, e.g. rational functions in the transform variable s, functions of $\sqrt{s}$, functions representable by polynomials in $1/s$, and so on. The accuracies attainable have depended very much on the particular transform $F(s)$ to be inverted, as well as on the argument, t, of the inverse $f(t)$. The highest accuracies so far

0020-2932/79/010097+24 $02.00/0

claimed have probably been those obtained by Piessens & Branders (1971), Piessens (1971, 1972), and Levin (1975), who in particular cases obtained errors of orders 10^{-12} to 10^{-15}.

The method to be described in the present paper is of the second type, but is unlike any previously published method. The number n of points to be used is one of several arbitrary parameters. Little if any preliminary computational work is required. The method is almost universal in its application, the only transforms to which it is not applicable being those possessing an infinite number of singularities with imaginary parts extending to infinity. The theoretical error is expressible in closed form by means of contour integrals, and for a given t decreases roughly exponentially with increase of n. Typically, for t ranging up to about 10 or 20, the error, relative or absolute depending on whether the transform has singularities in the right half-plane or not, is of order 10^{-6} for $n = 11$, 10^{-11} for $n = 20$, 10^{-20} for $n = 35$ (with double precision working), and so on. The actual decrease of error is of course limited by the precision of the computer, but the "round-off" error is usually very easily estimated from the value of one single term. In practice the orders of error quoted are always attainable, by proper choice of the other parameters, for all values of t from $0+$ upwards, though of course the required value of n usually increases with t, beyond the range of t already mentioned. (This is not the case however if the transform has only real singularities. For example the Bessel function $I_0(t)$, whose transform is $1/\sqrt{(s^2-1)}$, can be evaluated to 20 or more significant figures taking $n = 35$ for all t up to 200 and beyond.) Further, the required value of n, for a specified accuracy, depends considerably on the positions of the singularities of $F(s)$, or rather on the positions of the "dominant" singularities, and so does the choice of the other parameters, as will be seen below in Section 5. The computer execution time is roughly proportional to n. Using a CDC 7600 the average time for inversion when $n = 20$ (giving errors of order 10^{-11} or less) is about 0·8 ms.

In essence, but in a much less general form, the method is contained in an unpublished Ph.D. thesis (J. S. Green, 1955) which was supervised by the present author. This gave good results for "moderate" t but not for small or large t. The improvement for small or fairly large t is here effected by fuller exploitation of the parameters λ and σ already introduced by Green, and for larger t by the introduction of a new parameter, v.

Possible applications of the method are numerous, and many have already been tested. These include the following.

(a) The direct one-step solution, for any specified value of the independent variable, of any linear constant-coefficient differential equation with arbitrary right-hand side possessing a Laplace transform calculable as a function of the complex transform variable s.

(b) The time-domain solution of any linear network or system (e.g. control system) using either standard network or system analysis or the solution of simultaneous algebraic linear equations.

(c) In particular, the solution of a system governed by a state-matrix A:

$$\frac{du}{dt} = Au + v(t),$$

by combining the inversion process with Fadeev's method for evaluation of $(sI-A)^{-1}$. That is to say, given any vector $v(t)$ and any initial conditions, the equation can be solved for the vector $u(t)$ for a given t *in one step* to almost any desired degree of accuracy.

(d) The direct one-step solution for any specified x and t of the parabolic equation

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}, \quad a \leqslant x \leqslant b, \quad t > 0$$

with arbitrary initial condition

$$u(x, 0) = \phi(x), \quad a < x < b$$

and a variety of (or perhaps arbitrary) end-conditions on u or $\partial u/\partial x$.

(e) The evaluation of some difficult integrals to great accuracy, by inversion of their transforms taken with respect to a pre-existing or artificially introduced parameter.

(f) The direct evaluation of transcendental functions, by inversion of their transforms, to many more decimal places than are available at present, provided a computer of sufficient precision in its arithmetic and in its exponential and sine-cosine subroutines is at hand. For example with a CDC 7600 in double precision $J_0(t)$ can be found to 20 or more decimal places for $t \leqslant 200$ (and probably for unlimited t), using values of n ranging from 35 for $t = 1$ to 170 for $t = 200$. Triple precision would raise the number of places to about 34, and so on.

2. A General Formula for Numerical Inversion

Let $f(t)$, defined for $t > 0$, have the Laplace transform

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt \tag{1}$$

with abscissa of convergence γ_0, so that the integral converges in the half-plane $\operatorname{Re} s > \gamma_0$ but diverges in $\operatorname{Re} s < \gamma_0$. Our starting point for numerical inversion of $F(s)$ is the standard inversion formula

$$f(t) = \mathscr{L}^{-1}F(s) = \frac{1}{2\pi i}\int_B e^{st}F(s)\, ds, \quad t > 0 \tag{2}$$

where B is the "Bromwich contour" from $\gamma - i\infty$ to $\gamma + i\infty$, where $\gamma > \gamma_0$, so that B is to the right of all singularities of $F(s)$. Direct numerical integration along B is impractical on account of the oscillations of e^{st} as $\operatorname{Im} s \to \pm\infty$. The difficulties were to some extent overcome by Filon (1929) and others since, and probably Levin (1975) has gone as far as anybody in this direction by use of his remarkable convergence-acceleration algorithms. But his method would require considerably more effort to improve on the orders of error 10^{-12} to 10^{-15} which he has been able to achieve for some functions $F(s)$ and some values of t.

Here we overcome the difficulty by avoiding it. We replace B by an equivalent contour L starting and ending in the left half-plane, so that $\operatorname{Re} s \to -\infty$ at each end.

This replacement is permissible, i.e. L is equivalent to B, if

(i) L encloses all singularities of $F(s)$, (3)

and

(ii) $|F(s)| \to 0$ uniformly in $\operatorname{Re} s \leqslant \gamma_0$ as $|s| \to \infty$. (4)

Condition (ii) holds for almost all functions likely to be encountered, and we shall assume it satisfied by all $F(s)$ considered here. Condition (i) may well not be satisfied with a particular L by a given $F(s)$, but can generally be made to hold for the modified function $F(\lambda s+\sigma)$ by suitable choice of the scaling parameter λ and shift parameter σ, for if $F(s)$ has a singularity s_0, $F(\lambda s+\sigma)$ has the corresponding singularity

$$s_0^* = (s_0-\sigma)/\lambda. \tag{5}$$

Then (2) may be replaced by

$$f(t) = \frac{\lambda e^{\sigma t}}{2\pi i}\int_L e^{\lambda s t}F(\lambda s+\sigma)\,ds, \quad t>0. \tag{6}$$

In his thesis, Green (1955) used a standard contour L which was a steepest-descent contour for the integral (2) in the special case where $F(s) = 1/s$ and $t = 1$, and then used trapezoidal quadrature, based on a parametric description of the curve, to evaluate the integral. It is not difficult to see intuitively that this is likely to give good results for any $F(s)$ (some details are given in Talbot, 1976), but since the present paper uses a generalized contour L whose scope is far wider than that of Green's contour (for example it enables much greater values of t to be dealt with) but which is in general definitely *not* a steepest-descent contour for any $F(s)$, we shall introduce this contour in quite a different way.

The method to be described was suggested by the error analysis in Green (1955), but is much more general. The treatment is modelled closely on that of Green, except for the rotation of an auxiliary complex plane through a right angle for greater convenience.

Let $z = x+iy$ be a complex variable, M the imaginary interval from $z = -2\pi i$ to $+2\pi i$, and $s = S(z)$ a real uniform analytic function of z which

(a) has simple poles at $\pm 2\pi i$, and residues there with imaginary parts respectively positive and negative,
(b) has no singularities in the strip $|y| < 2\pi$,
(c) maps $M1-1$ onto a contour L traversed upwards in the s-plane, which encloses all singularities of $F(\lambda s+\sigma)$ for some λ and σ,
(d) maps the half-strip $H : x > 0$, $|y| < 2\pi$ into the exterior of L.

Then (6) holds and may be re-written

$$f(t) = \frac{1}{2\pi i}\int_M Q(z)\,dz = \frac{1}{2\pi}\int_{-2\pi}^{2\pi} Q(iy)\,dy, \tag{7}$$

where

$$Q(z) = \lambda\, e^{(\lambda S+\sigma)t}F(\lambda S+\sigma)S'(z). \tag{8}$$

We note that by assumptions (a) and (c), $\operatorname{Re} s \to -\infty$ on L as $z \to \pm 2\pi i$ on M, and

$Q(\pm 2\pi i) = 0$. We note further that condition (c) depends on F as well as S, and cannot always be satisfied if $F(s)$ has an infinite number of singularities with imaginary parts extending to infinity. Such transforms constitute the only inadmissible class for the method of this paper, and we shall assume henceforth that F does not belong to this class. A trapezoidal approximation to $f(t)$ in (7) is now

$$\bar{f}(t) = \frac{2}{n}\sum_{k=0}^{n-1}{}' \operatorname{Re} Q(z_k), \quad z_k = 2k\pi i/n. \tag{9}$$

This is the general inversion formula considered in this paper.

Now let M_1 be any path from $-2\pi i$ to $+2\pi i$ in the half-strip H to the right of M, and M_2 a path from $-2\pi i$ to $+2\pi i$ to the left of M and close enough to it to exclude from the region between them any singularities of $F(\lambda S(z)+\sigma)$ and hence of $Q(z)$. Then by the Residue Theorem

$$\bar{f}(t) = \frac{1}{2\pi i}\int_{M_1-M_2}\frac{Q(z)\,dz}{1-e^{-nz}}, \tag{10}$$

the integrand being regular at $\pm 2\pi i$.

By assumptions (c) and (d), M in (7) may be replaced by the equivalent path M_1. If then we combine (7) and (10) we obtain

$$\bar{E}(t) = E_1(t)+E_2(t), \tag{11}$$

where $\bar{E}(t)$ is the theoretical error (which depends of course on S, λ, σ and n as well as on t) given by

$$\bar{E}(t) = \bar{f}(t)-f(t), \tag{12}$$

and

$$E_1(t) = \frac{1}{2\pi i}\int_{M_1}\frac{Q\,dz}{e^{nz}-1}, \tag{13}$$

$$E_2 = \frac{1}{2\pi i}\int_{M_2}\frac{Q\,dz}{1-e^{-nz}}. \tag{14}$$

Since $\operatorname{Re} z > 0$ on M_1 it is obvious intuitively that $E_1 \to 0$ as $n \to \infty$. A formal proof of this is given in Appendix 1, where it is shown that for large enough n the integrand in (13) has a bound of order

$$O(n^2 \exp(h\tau - b\sqrt{\tau n}+\sigma t)), \tag{15}$$

where h and b are constants and

$$\tau = \lambda t,$$

in a region U consisting of the union of a conjugate pair of triangles with vertices at $\pm 2\pi i$. It is clear therefore that if M_1 is taken to lie inside U, $E_1(t) \to 0$ as $n \to \infty$.

Since $\operatorname{Re} z < 0$ on M_2, it is again intuitively obvious that $E_2(t) \to 0$ as $n \to \infty$. A similar formal proof applies, though the region U is now restricted by having to be to the right of all singularities of $F(\lambda S+\sigma)$, which may have the effect of increasing h and reducing b in (15) and hence increasing E_2, if $F(\lambda s+\sigma)$ has singularities near to L.

Thus for fixed t, λ and σ,

$$\bar{E} \to 0 \quad \text{as } n \to \infty. \tag{16}$$

It will be clear from (15) that both the rate of convergence and the magnitude of $\bar{E}$ depend heavily on τ. We shall discuss this dependence later.

In Green (1955) and Talbot (1976), in which the mapping function used, namely

$$s = S(z) = \frac{z}{1-e^{-z}}, \tag{17}$$

satisfies not only the conditions (a) to (d) above, but also the additional condition (e), $S(z)$ maps a region R bounded on the right by $M1$–1 onto the interior of L, the error component E was split into two parts by replacing M_2 in (14) by another path M_2' which together with M encloses some or all of the poles (if any) of $F(\lambda S+\sigma)$ but no non-polar singularities. We may then write, as is easy to see,

$$E_2 = E_0 + E_2', \tag{18}$$

where

$$E_0 = \sum \frac{e^{s_l t} \operatorname{res} F(s_l)}{e^{-nz_l^*}-1} \tag{19}$$

(or a more complicated expression if the poles are multiple),

$$E_2' = \frac{1}{2\pi i}\int_{M_2} \frac{Q\,dz}{e^{-nz}-1}, \tag{20}$$

and in (19) summation is over the poles s_l of $F(s)$ whose "transforms"

$$s_l^* = (s_l-\sigma)/\lambda$$

under shift σ and scaling λ (see (5)) are enclosed by M and M_2, and z_l^* is the inverse image of s_l^* in the region R. However the more general maps used in this paper (see Section 4) may not satisfy condition (e), and a detailed study of the Riemann surface corresponding to the inverse function $z = S^{-1}(s)$ would be required in order to decompose E_2 as in (18). In the computational strategy to be presented here we shall use this decomposition only in cases where the simpler mapping function (17) is used. (Detailed examples of its use are given in Talbot (1976) where, however, the notations E_2, M_2, L_2 and E_2', M_2', L_2' are interchanged). We shall denote the unique inverse function in such cases by $z = Z(s)$.

In addition to the theoretical error (11) we must of course consider the computational round-off error. Now it is clear from (8) and (9) that because of the exponential factor in Q, the first term in (9), viz.

$$T_0 = \frac{\lambda}{n}\exp\left((\lambda S(0)+\sigma)t\right)F(\lambda S(0)+\sigma)S'(0), \tag{21}$$

is normally the largest or near-largest.† Moreover, it is always found that because of heavy cancellations in the summation, $|\bar{f}| \ll |T_0|$. Thus if the computer evaluates T_0 and its neighbours correct to c significant figures, the rounding error E_r in $\bar{f}$ is roughly

† An exceptional case is discussed in Section 6(d).

given by

$$E_r = O(10^{-c}T_0), \tag{22}$$

all other round-off errors in the evaluation of $\bar{f}$ being negligible by comparison. Finally the *actual* error in $\bar{f}$ is

$$E = E_1 + E_2 + E_r, \tag{23}$$

and by (17) $E = O(10^{-c}T_0)$ for sufficiently large n. This gives an asymptotic order of magnitude of error which cannot be improved on, given a particular choice of λ and σ and of contour L. The strategy to be described in Section 5 involves finding a near-optimum choice of L from the family of contours discussed in the next section, and corresponding values of λ and σ.

3. A Special Family of Contours

The conditions (a) to (d) on the mapping $S(z)$ are not very restrictive, and allow of many possible mapping functions. The simplest are of the form

$$S(z) = az - \frac{b}{z^2+4\pi^2} + c,$$

and indeed such functions can give good results, though their potentialities have not yet been explored. What we shall consider here is the family of mappings

$$s = S_\nu(z) = \frac{z}{2}\left(\coth\frac{z}{2}+\nu\right) = \frac{z}{1-e^{-z}} + az, \tag{24}$$

where ν is an arbitrary positive parameter and $a = (\nu-1)/2$.

The singularities of $S_\nu(z)$ are simple poles, at $\pm(2, 4, 6, \ldots)\pi i$, and those at $\pm 2\pi i$ have residues $\pm 2\pi i$. $S_\nu(z)$ maps the interval $M(-2\pi i, 2\pi i)$ onto a contour

$$L_\nu\colon s = s_\nu(\theta) = \alpha + \nu i\theta, \quad -\pi < \theta < \pi, \tag{25}$$

where we have taken $z = 2i\theta$ on M, and

$$\alpha = \alpha(\theta) = \theta\cot\theta. \tag{26}$$

In the special case $\nu = 1$ ($a = 0$) we obtain the curve L which was used both by Green (1955) and Talbot (1976), and was derived by Green as a steepest-descent curve when $F = 1/s$. When $\nu \neq 1$, L^ν consists of L expanded "vertically" by a factor ν (see Fig. 1), and can be shown not to be a steepest-descent curve for any F. For most normal purposes the use of L provides any desired accuracy (subject to the computer precision) with moderate values of n. In certain circumstances, however, namely when $F(s)$ has complex singularities and, in relation to their position, t is "large" (this will be clarified below), it is found that by using L^ν with $\nu > 1$ one can escape from the limitation on accuracy imposed by E_r.

It is clear that $S_\nu(z)$ satisfies the conditions (a) to (c) assumed for $S(z)$ in Section 2, and it is easy to verify that (d) is also satisfied. When $\nu = 1$ or $\nu \simeq 1$, the extra condition (e) also holds, but (e) ceases to hold as ν increases above 1.

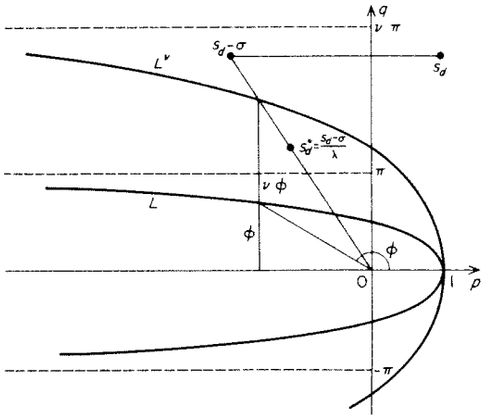

FIG. 1. Contours L and L^ν ($\nu = 2{\cdot}5$).

Now on L^ν we have

$$S'_\nu(z) = s'_\nu(\theta)/2i = (\nu + i\beta)/2,$$

where

$$\beta = \beta(\theta) = -d\alpha/d\theta = \theta + \alpha(\alpha-1)/\theta. \tag{27}$$

Thus for a given $F(s)$ and t, and choice of the parameters λ, σ, ν and n, the approximation (9) to the inverse $f(t)$ becomes

$$\bar{f}(t) = \frac{\lambda e^{\sigma t}}{n} \sum_{k=0}^{n-1}{}' \operatorname{Re}\left[(\nu + i\beta)\, e^{s_\nu \tau} F(\lambda s_\nu + \sigma)\right]_{\theta=\theta_k} \tag{28}$$

where

$$\theta_k = k\pi/n, \quad k = 0, 1, \ldots, n-1,$$

and

$$\tau = \lambda t. \tag{29}$$

If

$$F(\lambda s_\nu + \sigma) = G + iH, \tag{30}$$

where G and H are real, (28) takes the real form

$$\bar{f}(t) = \frac{\lambda e^{\sigma t}}{n} \sum_{k=0}^{n-1}{}' \left[e^{\alpha\tau}\{(\nu G - \beta H)\cos \nu\theta\tau - (\nu H + \beta G)\sin \nu\theta\tau\}\right]_{\theta=\theta_k}. \tag{31}$$

Before discussing the choice of parameters it is necessary to analyse the error components E_r, E_0, E_1 and E_2 in further detail, and this we do in the next section.

4. Error Analysis

We consider the error components in turn.

(i) E_r. When $S = S_\nu$, (21) gives

$$T_0 = \frac{\nu\lambda}{2n} e^{(\lambda+\sigma)t} F(\lambda + \sigma). \tag{32}$$

Thus by (22), using the symbol $\sim$ for convenience to denote order of magnitude,

$$E_r \sim e^{-A_r} = 10^{-d_r} \tag{33}$$

where

$$A_r = 2{\cdot}3c' - (\tau + \sigma t), \qquad d_r = A_r/2{\cdot}3, \tag{34}$$

$$c' = c + \log_{10}(2n/\nu\lambda F(\lambda + \sigma)). \tag{35}$$

The difference $c' - c$ can vary considerably in different cases, and may occasionally be abnormally high if $F(s)$ tends rapidly to zero as $s \to +\infty$ and t is small, leading to large λ for a selected τ.

(ii) E_0. As already stated, we shall consider E_0 only when $\nu = 1$. As we shall see this choice of ν is made for all t when the singularities of $F(s)$ are all real, and for small or "moderate" t in other cases. Then by (19) we may write

$$E_0 \sim e^{-A_0} = 10^{-d_0} \tag{36}$$

where

$$A_0 = \min\,(nu_l^* - p_l t), \qquad d_0 = A_0/2{\cdot}3, \tag{37}$$

$$u_l^* = -\operatorname{Re} z_l^* > 0, \qquad p_l = \operatorname{Re} s_l, \tag{38}$$

and

$$z_l^* = Z(s_l^*).$$

In order to use this we must be able to calculate $\operatorname{Re} Z(s)$ for any s. This can be done for example by applying Newton's method to the equation

$$w(z) \equiv s(1 - e^{-z}) - z = 0. \tag{39}$$

If z is any trial solution, the Newton correction is

$$\delta z = \frac{w}{w + z + 1 - s}. \tag{40}$$

For a given $s = (p, q) = re^{i\theta}$ a suitable starting value is $z = (-u, y)$, where

$$y = 2\pi - \frac{13}{5 - 2p - q - 0{\cdot}45\, e^p}, \tag{41}$$

and

$$u = \ln \frac{y - q}{r \sin(y - \theta)}. \tag{42}$$

(41) was derived empirically from inspection of curves $y =$ constant, (42) is satisfied by any solution $z = (-u, y)$ of (39). With these starting values convergence is found to be rapid or fairly rapid for any $s = s_l^*$ that will be encountered. However, the following alternative method which both avoids complex arithmetic and seems more efficient may be preferred. For any $z = (-u, y)$ let $w(z) = g + ih$. If u and y are connected as in (42), $h = 0$, while $g = g(u, y) = \hat{g}(y)$, say. Thus to solve $w(z) = 0$ for z we need only solve $\hat{g}(y) = 0$ for y, and this can be done by using a real Newton process. On carrying out the differentiations involved one finds that the Newton

correction to y is

$$\delta y = \frac{(q-y)g}{1-2re^{u}\cos(y-\theta)+r^2e^{2u}}. \tag{43}$$

A suitable starting value for y is again given by (41).

(iii) E_1. Except near the end-points of M_1 the order of magnitude of the integrand in (13) is determined by its dominant factor

$$\exp(\tau S_\nu - nz + \sigma t) = \exp(\tau S_1 - n'z + \sigma t), \tag{44}$$

where S_1 is S_ν with $\nu = 1$ ($a = 0$), and

$$n' = n - a\tau. \tag{45}$$

Since by (25) $S_1 \sim z$ when $\mathrm{Re}\, z > 3$, say, and M_1 may be deformed arbitrarily far to the right, a necessary condition for E_1 to be small is seen at once to be

$$\rho < 1, \tag{46}$$

where

$$\rho = \tau/n'. \tag{47}$$

In Talbot (1976) the special case $a = 0$, $n' = n$ was investigated by obtaining saddle-point estimates of E_1, and it was found empirically after examining a large number of results that, neglecting a factor involving the value of $F(s)$ at the saddle-point and other minor factors, we can write

$$E_1 \sim e^{-A_1} = 10^{-d_1}, \tag{48}$$

$$A_1 = nb_1 - \sigma t, \qquad d_1 = A_1/2{\cdot}3, \tag{49}$$

where b_1 is practically a function of $\rho = \tau/n$ alone. Since the only difference between (44) and the earlier case is in having n' instead of n, (48) and (49) still hold but with n' for n, i.e.,

$$A_1 = n'b_1(\rho) - \sigma t, \tag{50}$$

with ρ given by (47).

If we examine the derivation of (A.15) in Appendix 1, we see that in the case $S = S_1 + az$ which we are now considering it would be possible, just as in (44), to account for the term az in S by replacing n by n'. If this is done we may expect from (A.15) that A_1 is of the form

$$A_1 \simeq b\sqrt{\tau n'} - h\tau - \sigma t. \tag{51}$$

Comparing (50) with (51) gives

$$b_1 \simeq b\sqrt{\rho} - h\rho, \tag{52}$$

which is indeed a function of ρ only, if h and b are constant. Actual values of $b_1(\rho)$ were given in Table 2 in Talbot (1976), and it is interesting, as largely confirming the analysis in Appendix 1, that (52) agrees with those values in two important respects, namely:

(1) b_1 is approximately proportional to $\sqrt{\rho}$ for small ρ;
(2) b_1 rises from zero at $\rho = 0$ to a peak between 0 and 1 and returns to zero or below for $\rho \geqslant 1$.

In the sequel we shall need an explicit formula for $b_1(\rho)$, and the following, based or (52), provides a very close fit, with maximum error about 0·004 in the range $0 < \rho < 0{\cdot}8$:

$$b_1(\rho) = 3{\cdot}46\sqrt{\rho} - 1{\cdot}8\rho(\rho + 0{\cdot}8) + 0{\cdot}03. \tag{53}$$

We note that (50) may be written alternatively

$$A_1 = \tau c_1 - \sigma t \tag{54}$$

where

$$c_1 = c_1(\rho) = b_1(\rho)/\rho. \tag{55}$$

Here c_1 is a monotonic decreasing function of ρ.

For the strategy of Section 5 it will be necessary to invert the relationship in (55) i.e. to find ρ for a given c_1. This is readily done using Newton's method, but it is convenient to have explicit approximate formulae for the purpose. Writing

$$e = c_1 + 1, \tag{56}$$

we have (with relative error ranges indicated in brackets)

$$\rho = \begin{cases} (24{\cdot}8 - 2{\cdot}5e)/(16 + 4{\cdot}3e), & 1{\cdot}6 < e \leqslant 4{\cdot}4 \quad (-0{\cdot}8\% \text{ to } +0{\cdot}4\%), \\ (129/e - 4)/(50 + 3e), & 4{\cdot}4 < e \leqslant 10 \quad (-0{\cdot}8\% \text{ to } +0{\cdot}4\%), \\ (256/e + 0{\cdot}4)/(44 + 19e), & 10 < e < 20 \quad (-0{\cdot}2\% \text{ to } -0{\cdot}7\%). \end{cases} \tag{57}$$

Although (53) and hence (55) gives increasing errors as ρ increases beyond 0·8, i.e. a e decreases below 2·0, nevertheless in the application we shall make of (57) its firs formula may be safely used for all $e \leqslant 4{\cdot}4$, and its third formula for all $e > 10$.

(iv) E_2. The estimation of E_2 (or of $E_2' = E_2 - E_0$) is much more difficult than tha of E_1, because of the presence of the "transformed" singularities $s_j^* = (s_j - \sigma)/\lambda$ in th interior of L^ν. In Talbot (1976) it was found that, with $\nu = 1$, if the non-polar s_j^*, i any, are far enough away from L, then $E_2' < E_1$, and thus E_2' can be neglected. I many cases however this does not hold, even when $\nu = 1$, and indeed E_2 becomes th dominant error component, apart from E_r, as t increases. The way in which E_2 i taken account of will be explained in Section 5.

5. A Universal Strategy

The strategy to be described is based on an intensive study of a large number c experimental results using many different types of function $F(s)$. It is in two parts:

(I) choice of the "geometrical" parameters λ, σ, ν for given $F(s)$, t, and compute precision c (as defined for (22));
(II) choice of n for prescribed accuracy.

We consider these parts in order.

(I) Our strategy for λ, σ, ν depends heavily on the geometry of the singularities c $F(s)$, and especially on the position of what we shall call the "dominant" singularity

Before defining this however we must note the following:

(1) if $F(s)$ has no singularities in the half-plane Re $s > 0$, then the inverse $f(t)$ as given by (2) may be expected to be of order about $O(1)$, since B may be taken as the imaginary axis, indented if necessary, and $|F(s)| \to 0$ at both ends of B, by (4);

(2) if $F(s)$ has singularities in the half-plane Re $s > 0$, and their maximum real-part is $\hat{p}$, then $f(t) = e^{\hat{p}t}\,\mathscr{L}^{-1}F(s+\hat{p})$, where $F(s+\hat{p})$ is of the type of transform discussed in (1) so that $f(t)$ is of exponential order $O(e^{\hat{p}t})$, $\hat{p} > 0$.

It is clear that for functions of type (1) absolute and relative error are similar in magnitude and it is sensible to speak of either, while for functions of type (2) it is only sensible to speak of relative error or equivalently significant figures, and that moreover the number of correct significant figures in an approximation $\bar{f}(t)$ to $f(t)$ is roughly equal to the decimal position of the absolute error in the corresponding approximation to $\mathscr{L}^{-1}F(s+\hat{p})$.

Now the formulae for the error components E_0, E_1, E_2, E_r given in Sections 2 and 4 all refer to absolute, not relative, error. This is unavoidable because it is impossible to relate values of the E's to values of $\bar{f}$, which are given by the summation formula (28) in which there is always heavy cancellation. In view of the remarks above we can give a unified strategy for functions $F(s)$ of types (1) and (2) in the following way.

Let the singularities of $F(s)$ be at $s_j = p_j + iq_j$, and as above let

$$\hat{p} = \max_j p_j. \tag{58}$$

Now write

$$\sigma_0 = \max(0, \hat{p}), \tag{59}$$

and apply an *initial shift* σ_0 to $F(s)$. Then the resulting function

$$F_0(s) = F(s+\sigma_0) \tag{60}$$

is always of type (1), and our strategy will aim at producing an absolute error in $\bar{f}_0(t)$ of order 10^{-D}, where D is specified (or in other words $D-1$ correct decimal places in $\bar{f}_0(t)$), and accordingly a similar error in $\bar{f}(t)$ if F is of type (1), or D correct significant figures in $\bar{f}(t)$ if F is of type (2). We note that the singularities of $F_0(s)$ are at

$$s_j' = s_j - \sigma_0 = p_j' + iq_j, \qquad p_j' = p_j - \sigma_0 \leqslant 0. \tag{61}$$

The strategy as applied to $F_0(s)$ will involve a further shift, say σ' (which may be zero), making a total shift σ, so that

$$\sigma' = \sigma - \sigma_0, \tag{62}$$

and it is important to notice that in the formulae (34) and (50) for A_r and A_1 we must replace σ by σ', and in the formula (37) for A_0 we must replace p_l by p_l'.

After applying the initial shift σ_0 the next step is to find the "dominant" singularity of $F(s)$, provided there are some complex singularities. If s_j is one of these, with $q_j > 0$, then the radius from the origin to the corresponding s_j' meets L at a point where the ordinate is $\theta_j = \arg s_j'$. Thus s_j' is situated (q_j/θ_j) times as far out as that point, and we define as the most significant or dominant singularity s_d that for which

this ratio is greatest, i.e. $s_d = p_d + iq_d$ is such that

$$\frac{q_d}{\theta_d} = \max_{q_j > 0} \frac{q_j}{\theta_j}, \tag{63}$$

where

$$\theta_j = \arg s_j'. \tag{64}$$

If on the other hand the singularities are all real, they do not affect our choice of λ, σ, ν, and there is no need to find a dominant singularity for this purpose. It is however convenient to write $q_d = 0$ and $\theta_d = \pi$ in this case.

Now referring to (34), with σ' for σ as explained above, it is clear that the round-off error is controlled by the quantity

$$\omega = (\lambda + \sigma')t, \tag{65}$$

and experimental results have shown that the correct choice of ω is vital for an efficient strategy. They have also shown that if we write

$$v = q_d t, \tag{66}$$

the optimum strategy depends on the value of v. In fact there are two distinct cases to consider.

Case 1. $$v \leqslant \omega\theta_d/1{\cdot}8. \tag{67}$$

We then use only the initial shift σ_0 and a scaling factor λ, and do not expand L: we take

$$\begin{aligned} \lambda &= \omega/t \qquad (\tau = \omega), \\ \sigma &= \sigma_0 \qquad (\sigma' = 0), \\ \nu &= 1 \qquad (a = 0), \end{aligned} \tag{68}$$

and note that with this choice (65) is satisfied. Case 1 always occurs if the singularities of $F(s)$ are all real.

Case 2. $$v > \omega\theta_d/1{\cdot}8. \tag{69}$$

In this case we use the expanded contour L^v, as shown in Fig. 1. It is clear that with ϕ defined as in Fig. 1,

$$p_d - \sigma = \frac{q_d}{v}\cot\phi. \tag{70}$$

Now if λ_v is the value of λ which would bring $s_d - \sigma$ just onto L^v,

$$\lambda_v = q_d/v\phi. \tag{71}$$

We now define κ as the ratio λ/λ_v, which is a measure of the extent to which s_d^* is inside L^v, i.e. away from the contour and approaching the origin. Then

$$\kappa = v\lambda\phi/q_d. \tag{72}$$

Finally, if we regard ω, ϕ, κ as three new parameters we can solve (65), (70) and (72)

for λ, σ, v and obtain

$$\lambda = \kappa q_d/v\phi, \quad \sigma = p_d - \frac{q_d}{v}\cot\phi, \quad v = q_d\left(\frac{\kappa}{\phi} - \cot\phi\right)\Big/\left(\frac{\omega}{t} + \sigma_0 - p_d\right).$$

These can be made to give good results, but even better results are obtained if we replace p_d by $\hat{p}$ (thus ensuring that $\sigma' > 0$). If then we write

$$\mu = \left(\frac{\omega}{t} + \sigma_0 - \hat{p}\right)\Big/\left(\frac{\kappa}{\phi} - \cot\phi\right), \tag{73}$$

our formulae in Case 2 becomes

$$\begin{aligned} \lambda &= \kappa\mu/\phi, \\ \sigma &= \hat{p} - \mu\cot\phi, \\ v &= q_d/\mu. \end{aligned} \tag{74}$$

We still need to fix the values of ω (in Cases 1 and 2) and of κ and ϕ in Case 2. Tests with numerous functions $F(s)$ and values of t have shown that for optimum results ω, κ and ϕ should be taken as functions of $v = q_d t$. The precise values are not critical and the optimum choice will vary with F and t, but for κ and ϕ the following choices have been found to be very satisfactory in all cases tried:

$$\begin{aligned} \kappa &= 1{\cdot}6 + 12/(v+25), \\ \phi &= 1{\cdot}05 + 1050/\max\,(553{,}800 - v). \end{aligned} \tag{75}$$

For ω the choice must depend on the precision of the computer. For the CDC 7600, for which c is normally about 14 in single-precision (S.P.), 27 in double-precision (D.P), the following have been used successfully:

$$\begin{aligned} c = 14: &\quad \omega = \min\,(6 + v/2, 10) \\ c = 27: &\quad \omega = \min\,(11 + v/2, 19). \end{aligned} \tag{76}$$

We may assume that as a general formula, suitable for any value of c, we may take

$$\omega = \min\,(0{\cdot}4(c+1) + v/2, 2(c+1)/3) \tag{77}$$

Summarizing, the choice of the parameters λ, σ, v for a given $F(s)$ and t is made as follows: with $\hat{p}, \sigma_0, s_d = p_d + iq_d, \theta_d, v$ and ω defined as in (58), (59), (63), (64), (66), and (77) we test for Case 1 by (67). If this is satisfied, λ, σ and v are given by (68). If (67) is not satisfied, we find κ and ϕ by (75) and hence λ, μ, v by (74).

(II) It remains now to determine a suitable value of n, to achieve a desired D significant figures (if $\hat{p} > 0$), or error in Dth decimal place (if $\hat{p} \leqslant 0$). We proceed as follows, in general terms. We shall find an n_0 such that the error E_0, with σ replaced by σ', is small enough, and similarly an n_1 and n_2. Then we take

$$n = \max\,(n_0, n_1, n_2). \tag{78}$$

We discuss n_0, n_1, n_2 in turn.

(i) n_0. Both for n_0 and n_2 we use a slightly modified value of D to take account of multiple poles. To be quite general, for any singularity s_j and any desired value of D we define

$$D_j = D + \min\,(2m_j - 2, 2) + [m_j/4], \tag{79}$$

where $m_j = 0$ if s_j is non-polar, and otherwise m_j is the multiplicity of s_j.

For n_0 there are several possibilities to consider.

(a) If we have Case 2, we take $n_0 = 0$, since n_0 is always found to be less than n_1 and n_2.

(b) In Case 1, if the dominant (complex) singularity s_d is not a pole, we again take $n_0 = 0$.

(c) If, however, in Case 1, s_d is a pole, we apply the algorithm (41)–(43) to $s = s_d^* = (s_d - \sigma_0)/\lambda$. This rapidly gives the value of $u_d^* = -\mathrm{Re}\,Z(s_d^*)$. (Note: for the purpose in hand, n_0 need only be known to within about 1%. Thus the algorithm can be stopped when $|\delta y| < 10^{-4}$, say.) Then by (37), with d_0 replaced by D_d,

$$n_0 \geqslant n_{0d} = (2{\cdot}3D_d + p_d' t)/u_d^* \tag{80}$$

where $p_d' = p_d - \sigma_0$. Strictly speaking one should also find similar values n_{0l} corresponding to other poles, s_l, and take $n_0 = \max\,[n_{0l}] + 1$, but with our choice of s_d it is very unlikely that this will be necessary. Thus we take

$$n_0 = [n_{0d}] + 1. \tag{81}$$

(d) A special problem arises if the singularities of $F(s)$ are all real, for no dominant singularity has been defined in this case. Clearly we can take $n_0 = 0$ if there are no poles, and otherwise

$$n_0 = \max\,[n_{0l}] + 1. \tag{82}$$

In fact however this need rarely be calculated, i.e. we can normally take $n_0 = 0$. For we may write

$$n_{0l} = \frac{2{\cdot}3D_l + p_l' t}{u_l^*} = \tau R_l, \tag{83}$$

where

$$p_l' = p_l - \sigma_0 = \lambda p_l^* \leqslant 0, \qquad R_l = \frac{h_l + p_l^*}{u_l^*}, \qquad h_l = \frac{2{\cdot}3D_l}{\tau}, \qquad u_l^* = -\mathrm{Re}\,Z(p_l^*),$$

and (since we necessarily have Case 1) $\tau = \omega$.

Now a study of R_l as a function of p_l^* shows that, for fixed h_l, R_l is a maximum when $p_l^* \simeq -0{\cdot}43h_l$, i.e. $p_l' \simeq -D_l/t$, and then $R_l \simeq \hat{R}(h_l)$, where

$$\hat{R}(h) = 0{\cdot}09 + 0{\cdot}11h + 0{\cdot}085h^2 + 0{\cdot}005h^3. \tag{84}$$

Using this it is readily shown that when $c = 14$ or 27 and correspondingly $\omega = 6$ or 11, $n_{0l} < n_1$ if $D_l \leqslant c - 2$, which presumably holds for any c. Thus in calculating n_0 by (82) we need only consider poles p_l for which $D_l \geqslant c - 1$, and usually there will be none such. We note also that poles for which $p_l' = 0$ can always be ignored, since they give $u_l^* = \infty$.

(ii) n_1. By (50), with σ replaced by σ', we must make

$$\tau c_1 = n_1' b_1 = 2{\cdot}3D + \sigma' t, \tag{85}$$

where $n_1' = n_1 - a\tau$. Then

$$e = c_1 + 1 = (2{\cdot}3D + \omega)/\tau. \tag{86}$$

The solution ρ is given by (57), and we now take

$$n_1 = [\tau(a + 1/\rho)] + 1. \tag{87}$$

(iii) n_2. As already stated, E_2, unlike E_0 and E_1, cannot be estimated readily by formulae applying in all cases, and our value of n_2 is based entirely on a careful and extensive examination of empirical results. Writing

$$\gamma = \sigma'/\lambda, \tag{88}$$

and

$$D' = \max D_j \text{ if all singularities } s_j \text{ are real,}$$

$$D' = D_d \text{ otherwise,}$$

we take

$$n_2 = [\eta v(2{\cdot}3D' + \omega)/(3 + 4\gamma + e^{-\gamma})] + 1, \tag{89}$$

where η is a factor $\leqslant 2$. It is safe to take $\eta = 2$, but for large t, when n_2 is the largest of n_0, n_1, n_2, computational economy requires η to be reduced as much as possible, and we take

$$\eta = (1{\cdot}09 - 0{\cdot}92y + 0{\cdot}8y^2) \min (1{\cdot}78, 1{\cdot}236 + 0{\cdot}0064(1{\cdot}78)^{D'}) \tag{90}$$

where

$$y = 10^{-3}v.$$

(Note: if the s_j are all real, n_2 is independent of t.)

We conclude this section by remarking that on the basis of a great deal of computational experience it is believed that with n_0, n_1, n_2 calculated as explained, n as given by (78) is not far from minimal in nearly all cases.

6. General Remarks and Practical Considerations

(a) Since c' in (35) varies only slightly with n, when λ, σ, v have been found in Part I of the strategy, d_r can be closely estimated from (34), with σ' for σ, by assuming a rough value of n. There is then clearly no point in aiming for a desired accuracy D which exceeds d_r or perhaps even $d_r - 1$ (although frequently the value of D achieved in practice is approximately equal to d_r, and sometimes exceeds it).

(b) The factors $\cos v\theta\tau$ and $\sin v\theta\tau$ in (31) can be an appreciable source of error. In Case 2 the angle $v\theta\tau = \kappa\theta v/\phi$, and for large t, i.e. large v, $\kappa/\phi \simeq 0{\cdot}6$. Then $v\theta\tau$ may easily be of order 10^2, and thus subject to absolute computational error of order $10^{-(c-2)}$, leading to similar orders of error in the cosine and sine, quite apart from the subroutine errors which are again of similar order due to the large angle. Thus the effective value of c may be reduced by one or two units on this account. Such reduction may be avoided by the following procedure.

The basic formula (28) may be written

$$\bar{f}(t) = \frac{\lambda e^{\sigma t}}{n} \operatorname{Re} \sum_0^{n-1}{}' a_k e^{ki\psi}, \tag{91}$$

where

$$a_k = [e^{\sigma\tau}(v + i\beta)F(\lambda s_v + \sigma)]_{\theta=\theta_k}, \quad \psi = \tau v\pi/n. \tag{92}$$

Now the factors $e^{ki\psi}$ in (91) satisfy the same recurrence relation as the Chebyshev polynomials $T_k(\cos\psi)$:

$$e^{(k+1)i\psi} + e^{(k-1)i\psi} = 2\cos\psi \,.\, e^{ki\psi}. \tag{93}$$

It follows that the sum $\sum'$ can be evaluated by an algorithm almost identical to Clenshaw's for Chebyshev sums, viz.:

$$\begin{aligned} &b_{n+1} = b_n = 0, \\ &b_k = a_k + ub_{k+1} - b_{k+2} \quad (u = 2\cos\psi), \quad k = n-1, \ldots, 1, \\ &\textstyle\sum' = \tfrac{1}{2}(a_0 + ub_1) - b_2 + ib_1 \sin\psi. \end{aligned} \tag{94}$$

Moreover, since the only trigonometrical evaluations required are $\cos\psi$ and $\sin\psi$ (where ψ is about 1 or 2 radians), it is clear that, compared with (31), (91) offers a considerable saving in execution time, in addition to the increase in accuracy. In fact the total execution time is reduced by about 20%.

(c) Another possible source of error would seem to be in the calculation of β in (27) for small θ, for then $1 - \alpha$ is small and in fact of order θ^2. Two alternative remedies based on series expansions are available: one using the expansion of $\alpha(\theta)$, with Bernoulli numbers as coefficients, the other simpler, though involving more operations, and obtained by expressing $(1 - \alpha)/\theta^2$ as a ratio of two series. Both have been tried, up to $\theta = 0{\cdot}35$ and $\theta = 0{\cdot}6$ respectively, retaining appropriate numbers of terms in the series, but no appreciable improvement in performance has been discerned, so no further details are given here.

(d) If $F(s) = e^{-a/s}/\sqrt{s}$, with an essential singularity at $s = 0$, we have Case 1, with $\lambda = \omega/t$, $\tau = \omega$, $\sigma = 0$, $v = 1$, and $\omega = 6$ or 11. Thus λ is small when t is large, and $F(\lambda s)$ is subject to the abnormal behaviour of an analytic function near an essential singularity. Detailed analysis in this case shows that when $a > 0$ and $at > \tau^2/2$, T_0 ceases to be the largest term in (9), and as t increases the largest term increases rapidly (almost exponentially). Thus, for large t, E_r is much larger than is indicated in (22), and so thwarts our general strategy by considerably reducing d_r.

The remedy is simple: one must set a lower bound to λ. With $\omega = 5 + at/30$ ($\lambda = 5/t + a/30$) in single-precision, or $\omega = 10 + at/30$ ($\lambda = 10/t + a/30$) in double-precision, normal results are obtained, and are to be expected likewise for other transforms having $e^{-a/s}$ as a factor.

If $a < 0$ the problem does not arise, and the normal choice of ω may be made.

(e) In Case 1, if t is small, $\lambda = \omega/t$ is large and s_d^* is near the origin. The algorithm (41)–(43) used in finding n_{0d} may then fail through the argument of the logarithm in (42) becoming negative. One should then abandon the algorithm and simply take $n_{0d} = 0$, for with s_d^* near to zero, u_d^* will be large and n_{0d} small, and certainly smaller than n_1.

(f) In the strategy of Section 5 an essential part is played by the dominant singularity s_d, but it should be noted that the position of s_d need not be known accurately: there is nothing critical in the choice of λ, σ, ν. Now the larger the imaginary part q_d, the larger does $v = q_d t$ become for a given t, and the larger will be the value of n required to achieve a given D. If however the position of s_d is known exactly, and s_d is some distance above all other singularities more precisely (see (63)): $q_d/\theta_d \gg$ other q_j/θ_j, then the required value of n can be significantly reduced by setting aside s_d, using the next dominant singularity $s_{d'}$, in finding λ, σ, ν and n, and taking account of s_d by simply adding the residue term

$$e^{s_d t} \operatorname{res} F(s_d) \tag{95}$$

to $\bar{f}(t)$. This assumes however that with the chosen parameters, s_d^* is outside L^ν, the condition for which is (cf. (72) and Fig. 1)

$$\tan^{-1} \frac{q_d}{\nu(p_d - \sigma)} < \frac{q_d}{\lambda \nu}. \tag{96}$$

(g) Suppose condition (ii), in (4), is not satisfied by $F(s)$ but is satisfied by $G(s) = e^{as}F(s)$, $a > 0$. Since

$$e^{st}F(s) = e^{s(t-a)}G(s),$$

we may continue to use F in (28) to obtain a good approximation $\bar{f}(t)$ to $f(t)$ in (2), provided $t > a$, and provided we replace t by $t - a$ in the strategy of Section 5 (though not in the formula (28)). For $t < a$ however the method would give values of $\bar{f}(t)$ quite different from the correct value of $f(t)$, namely zero.

7. Results

The method and strategy described in previous sections have been applied to a large number of miscellaneous transforms, collectively exhibiting most if not all of the features likely to be encountered in practice. The results to be quoted were obtained partly on the CDC 7600 of the University of London Computing Centre, partly on the CDC 6600 of the Center for Numerical Analysis at Austin, Texas.

Twenty transforms were selected for special attention. For each the absolute error $\bar{f}(t) - f(t)$, or relative error $(\bar{f}(t) - f(t))/f(t)$, as appropriate, was computed for a range of values of t up to 200 or 500, and of D up to 12 for S.P. ($c = 14$), 24 for D.P. ($c = 27$).

In discussing the results we divide the transforms into two groups: those with all their singularities real, and those with some complex singularities. The functions in the first group are shown in Table 1. Since Case 1 occurs for all t, the strategy leads to the same λ, $\sigma'(=0)$, and $\nu(=1)$ for all t (except in the case of F_8: see Section 6(d)), and to n depending slightly on t and on the multiplicities m_j but mainly on D. The approximate relationship is set out below in Table 2. In every case the desired D was attained. It should be noted that since execution time in D.P. is about three times that in S.P., S.P. should be used for $D \leqslant 12$, for transforms with all their singularities real.

The second group of transforms is shown in Table 3. For those there is a

TABLE 1

Transforms with all singularities real

i	$F_i(s)$	$f_i(t)$
1	$1/\sqrt{(s^2-1)}$	$I_0(t)$
2	$\exp(-s/\sqrt{(s+1)})/s$	—
3	$\exp(-\sqrt{s})$	$\exp(-1/4t)/2t\sqrt{(\pi t)}$
4	$1/(s+1)^5$	$t^4 e^{-t}/24$
5	$999/(s+1)(s+1000)$	$e^{-t} - e^{-1000t}$
6	$-(\gamma + \ln s)/s$	$\ln t$
7	$2/(\sqrt{s}+\sqrt{(s+1)})$	$(1-e^{-t})/t\sqrt{(\pi t)}$
8	$e^{-1/s}/\sqrt{s}$	$\cos 2\sqrt{t}/\sqrt{(\pi t)}$

TABLE 2

Values of n for F_1–F_8

D	6	8	10	11	12	12	16	20	22	24
n (S.P.)	11	14	18	20	22					
n (D.P.)						22	28	35	39	43

TABLE 3

Transforms with complex singularities

i	$F_i(s)$	$f_i(t)$
9	$\tan^{-1}(1/s)$	$\sin t/t$
10	$1/\sqrt{(s^2+1)}$	$J_0(t)$
11	$1/(s+\sqrt{(s^2+1)})^{1/2}$	$\sin t/t\sqrt{(2\pi t)}$
12	$1/\sqrt{(s^2+1)}.(s+\sqrt{(s^2+1)})^{1/2}$	$\sin t/\sqrt{(\pi t/2)}$
13	$1/\sqrt{(s^2+1)}.(s+\sqrt{(s^2+1)})^{3/2}$	$(\sin t - t\cos t)/t\sqrt{(\pi t/2)}$
14	$\exp(s-\sqrt{(s^2+1)})/\sqrt{(s^2+1)}.(s+\sqrt{(s^2+1)})^{1/2}$	$\sqrt{(2/\pi(t+2))}.\sin\sqrt{(t(t+2))}$
15	$s/(s^2+1)\sqrt{(s+1)}$	—
16	$s^3/(s^4+4)$	$\cos t \cosh t$
17	$1/(s^4-1)$	$(\sinh t - \sin t)/2$
18	$1/(s^2+1)^2$	$(\sin t - t\cos t)/2$
19	$s^2/(s^3+8)$	$(e^{-2t} - 2\cos\sqrt{3}t.e^t)/3$
20	$\ln((s^2+1)/(s^2+4))$	$2(\cos 2t - \cos t)/t$

transition from Case 1 to Case 2 at a value t_d of t which depends on the position of the dominant singularity s_d, and is found by solving the equation

$$\omega = 1{\cdot}8v/\theta_d \tag{97}$$

(see (67)) for v. For the functions F_9–F_{16} and F_{18}, $q_d = 1$, $\theta_d = \pi/2$, and $t_d = 8{\cdot}7$ (S.P.) or 16·6 (D.P.). For F_{17} these values are $1, 3\pi/4$, 13 or 24·9; for F_{19}, $\sqrt{3}$, $\pi/2$, 5 or 9·6; for F_{20}, 2, $\pi/2$, 4·4 or 8·3.

In general the required values of n for prescribed D differ only slightly from those shown in Table 2 when $t < t_d$, but rise rather rapidly with t when $t > t_d$. In Tables 4(a)–4(e) we show these values for the functions F_9–F_{20} for $t = 5, 10, 20, 50, 100, 200$.

TABLE 4(a)

Values of n for F_9–F_{14}

t	$D=$ 6	8	10	11	12	12	16	20	22	24
5	12	15	18	19	21	23	28	34	38	42*
10	17	22	27	29	32	25	30	35	38	41*
20	21	26	31	34	37	34	44	55	61	67*
50	32	38	52	56	60	47	57	69	77*	84*
100	50	62	91	99	106	71	87	102	110*	118*
200	86	100	147	159	170	114	139	164	177*	190*

TABLE 4(b)

Values of n for F_{15}–F_{17}

t	$D=$ 6	8	10	11	12	12	16	20	22	24
5	13	15	18	19	21	23	28	34	38	42*
10	17 (14)	22 (16)	27 (19)	29 (21)	32 (24)	25	30	36	39	43*
20	21	26	31	34	37	34 (28)	44 (32)	55 (37)	61* (44*)	67* (50*)
50	35	52	60	64	69	47	57	69	77*	84*
100	62	91	106	113	121	79	94	110	118	125*
200	100	147	170	182	194†	127	152	177	189	202*

Bracketed entries refer to F_{17}.

TABLE 4(c)

Values of n for F_{18}

t	$D=$ 6	8	10	11	12	12	16	20	22	24
5	14	17	20	22	24	24	28	34	38	42
10	17	22	27	29	32	27	34	41	44	48
20	23	27	31	34	37	34	44	55	61	67*
50	52	60	69	73	77	51	60	69	77	84*
100	91	106	121	128†	135†	91	106	122	129	137*
200	147	170	194†	206†	218†	146	170	195	208*	220*

TABLE 4(d)

Values of n for F_{19}

t	$D=$ 6	8	10	11	12	12	16	20	22	24
5	17	22	28	30	33*	25	29	35	38	41*
10	20	25	30	33†	36†	33	42	53	59	66*
20	27	38	44	47	50	40	50	62	69	76*
50	55	82	95	101	108	71	84	98	105*	112*
100	92	135	157	167	178	116	139	162	174†	185*
200	146	216	251	268†	286†	187	224	267	291	316*

TABLE 4(e)

Values of n for F_{20}

t	$D=$ 6	8	10	11	12	12	16	20	22	24
5	17	22	27	29	32	25	30	35	38	41*
10	21	26	31	34	37	34	44	55	61	67*
20	29	34	43	46	50	43	53	65	71*	79*
50	50	62	91	99	106	71	87	102	110*	118*
100	86	100	147	159	170	114	139	164	177*	190*
200	156	174	244	263	283	202	239	282	306*	331*

The values of n shown are for S.P. working for $D \leqslant 12$, D.P. for $D \geqslant 12$, and ar the values as calculated by the strategy of Section 5.

The desired values of D were attained almost everywhere for $D \leqslant 10$ (S.P.), $D \leqslant 2$ (D.P.), the only exception being at $t = 200$, $D = 10$ for F_{18}, which could in fact hav been avoided by a slight modification of (79), leading to a slight increase in n in case of double poles, such as F_{18}. For larger D an asterisk * against a value of n denote that D was not attained for some or all of the transforms involved because c rounding error (i.e. because $d_r < D$), while a dagger † denotes failure for some othe reason, not yet fully diagnosed.

Remark 1. For the functions F_i, $i = 1, 2, 10, 15$ the inverses $f_i(t)$ cannot be readil found from explicit formulae involving standard computer subroutines. In these case the values of $f_i(t)$ needed for finding the errors $\bar{f}_i(t)-f_i(t)$ at the various values of t wer obtained to at least 24 figures or places by choosing λ, σ, v so that $d_r > 24$ and the increasing n until successive values of f agreed to 24 figures or places.

Remark 2. Even where explicit formula for $f_i(t)$ involving elementary subroutines ar available, there may be a loss of accuracy in evaluating $f_i(t)$ for particular values of which may contribute appreciably to the error $\bar{f}_i(t)-f_i(t)$. This certainly occurs fo example for $f_{19}(t)$ when $t = 10$, and $|\tan\sqrt{3}t| \simeq 25$, for then an error ε in th computer evaluation of $\sqrt{3}$ produces a relative error of 250ε in $\cos\sqrt{3}t$. Whether thi accounts fully for the failures † for $t = 10$ in Table 4(d) has not yet been investigate Similarly there is an error in the evaluation of $f_{18}(t)$ for large t which may account fo the failures † in Table 4(c).

8. Conclusions

A method has been presented for inverting all Laplace transforms except those with an infinite number of singularities whose imaginary parts extend to infinity. The method yields absolute or relative errors less than 10^{1-D}, and usually of this order, for any prescribed values of D up to about three-quarters of the computer precision c, and in many cases up to five-sixths of c. For inversion of a transform at a particular t the method requires the evaluation of the transform at complex points, the number and positions of which are determined by t, D, and the position (or approximate position) of the "dominant" singularity of the transform.

Applications are numerous, and it is suggested that transform inversion be considered as a possible method of solving any mathematical problem of which the solution can be regarded as a value or set of values of a function $f(t)$ whose transform $F(s)$ can be found as a calculable analytic function of s.

References

Filon, L. N. G. 1928 *Proc. R. Soc. Edinb.*, **49,** 39–47.

Green, J. S. 1955 The calculation of the time-responses of linear systems. Ph.D. Thesis, University of London.

Levin, D. 1975 *J. Comp. Appl. Math.* **1,** 247–250.

Piessens, R. 1971 *J. Engng Math.* **5,** 1–9.

Piessens, R. 1972 *J. Inst. Maths Applics* **10,** 185–192.

Piessens, R. 1975 *J. Comp. Appl. Math.* **1,** 115–128. Also supplement in **2** (1976), 225–228.

Piessens, R. & Branders, M. 1971 *Proc. IEE* **118,** 1517—1522.

Salzer, H. E. 1955 *M.T.A.C.* **9,** 167–177.

Shohat, J. 1940 *Duke Math. J.* **6,** 615–626.

Talbot, A. 1976 The accurate numerical inversion of Laplace transforms. *Report TR/61.* Brunel University.

Tricomi, F. 1935 *RC Accad. Naz. Lincei, Cl. Sci. Fis.* **13,** 232–239, 420–426.

Widder, D. V. 1935 *Duke Math. J.* **1,** 126–136.

Appendix

The object here is to prove that $E_1(t) \to 0$ as $n \to \infty$. Let $S(z)$ satisfy the conditions (a)–(d) of Section 2. Let the residue at the pole $-2\pi i$ be k. Then

$$\arg k = \frac{3\pi}{2} + \delta, \qquad |\delta| < \frac{\pi}{2}. \tag{A.1}$$

Take angles α and β between $|\delta|$ and $\pi/2$, so that

$$0 \leqslant |\delta| < \alpha < \beta < \frac{\pi}{2}, \tag{A.2}$$

and define T as the triangle bounded by rays issuing from $z = -2\pi i$ with inclinations α and β, and the real axis $\operatorname{Re} z = 0$, as shown in Fig. 2.

Taking

$$\xi = z + 2\pi i = re^{i\theta}, \tag{A.3}$$

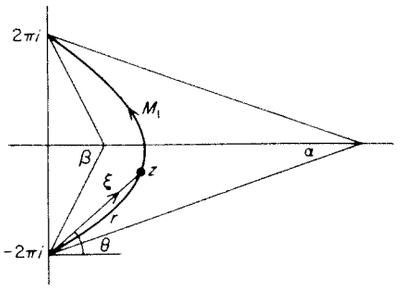

Fig. 2.

where

$$\alpha \leqslant \theta \leqslant \beta \text{ in } T, \tag{A.4}$$

we may write

$$S(z) = \frac{k}{\xi} + A' + iA'' + \xi B(\xi), \tag{A.5}$$

where A' and A'' are real constants and by condition (b) $B(\xi)$ is regular and bounded in T, say $|B(\xi)| \leqslant L$ in T. Similarly

$$S'(z) = \frac{-k}{\xi^2} + C(\xi), \tag{A.6}$$

where $C(\xi) = B(\xi) + \xi B'(\xi)$ is regular and bounded in T, say $|C(\xi)| \leqslant M$ in T. Then for z in T we have

$$\operatorname{Re} S(z) \leqslant -\frac{|k|}{r} \sin(\theta - \delta) + A' + rL, \tag{A.7}$$

$$|S'(z)| \leqslant \frac{|k|}{r^2} + M. \tag{A.8}$$

Next we obtain lower bounds for the denominator $e^{nz} - 1$ in (13). Now if g is fixed in the range $0 < g < 1$, and $nr \leqslant g$,

$$\begin{aligned} |e^{nz} - 1| = |e^{n\xi} - 1| &> nr(1 - g/2 - g^2/6 \ldots) \\ &> cnr, \quad c = 2(1-g)/(2-g). \end{aligned} \tag{A.9}$$

Again, if $nr > g$ and $\theta \leqslant \beta$,

$$|e^{nz} - 1| > e^{nr\cos\theta} - 1 > de^{nr\cos\theta}, \quad d = 1 - e^{-g\cos\beta}. \tag{A.10}$$

Now let $P = S'(z)\, e^{\tau S}/(e^{nz} - 1)$, and take $z \in T$. Then

(i) $$\text{if } 0 < nr \leqslant g, \qquad |P| < f(r)/n,$$

where

$$f(r) = \frac{1}{cr}\left(\frac{|k|}{r^2} + M\right) \exp \tau(-|k| \sin(\theta - \delta)/r + A' + rL) \tag{A.11}$$

$$\to 0 \quad \text{as} \quad r \to 0,$$

and $|P| < \varepsilon/n$ if $r < r_0(\varepsilon)$,

(ii) if $nr > g$, $$|P| < \frac{1}{d}\left(\frac{|k|}{r^2}+M\right)\times \exp\{\tau(-|k|\sin(\theta-\delta)/r + A' + rL) - nr\cos\theta\}$$
$$< h(n) = \frac{1}{d}\left(\frac{n^2|k|}{g^2}+M\right)\exp(h\tau - b'\sqrt{\tau n}, \tag{A.12}$$

where

$$h = A' + 2\pi L/\sin\alpha, \tag{A.13}$$

and

$$b'^2 = 4|k|\sin(\theta-\delta)\cos\theta$$
$$\geqslant b^2 = 4|k|\min(\sin(\alpha-\delta)\cos\alpha, \sin(\beta-\delta)\cos\beta) > 0. \tag{A.14}$$

Thus $h(n) \to 0$ as $n \to \infty$, and $|P| < \varepsilon$ if $n > n_0(\varepsilon)$.

Take $n > n_1(\varepsilon) = \max(1 + [g/r_0], n_0)$. Then we either have case (i) above, with $r < r_0$, or case (ii) with $n > n_0$, and in either case $|P| < \varepsilon$. The same obviously holds if $z \in \bar{T}$, the triangle conjugate to T. Thus $|P| \to 0$ as $n \to \infty$, uniformly in the union $U = T \cup \bar{T}$. Since the integrand in (13) is $\lambda e^{\sigma t} PF(\lambda S + \sigma)$, and F is bounded in U, it follows that $E_1(t) \to 0$ as $n \to \infty$ if M_1 lies in U.

Further, since (A.12) holds except for points close to $\pm 2\pi i$ for which $|z \mp 2\pi i| \leqslant g/n$, where $g < 1$, it is clear we can write

$$|E_1(t)| < \lambda C\, e^{\sigma t}\|F\| n^2 \exp(h\tau - b\sqrt{\tau n}), \tag{A.15}$$

where $\|F\| = \max|F(\lambda S(z) + \sigma)|$, taken over U, and C, h and b depend only on the mapping function $S(z)$ and the angles α and β.

J. Inst. Maths Applics (1979) **23**, 121–130

The Equivalence of Bernoulli's Equation and a Set of Integral Relations for Periodic Waves

J. G. B. Byatt-Smith
Department of Mathematics, University of Edinburgh

[Received 21 April 1978]

In a recent paper Longuet-Higgins (1978) discovered some new relations between Stokes' coefficients in the theory of periodic gravity waves. These were shown to give rise to a set of integral relations. In this paper we show that this set is complete in the sense that it is equivalent to Bernoulli's equation. We also show that a suitably redefined set exists in the theory of the solitary wave.

1. Integral Relations for Waves of Infinite Depth

By Bernoulli's theorem, the condition that the pressure is constant on the free surface may be written

$$gy + \tfrac{1}{2}q^2 = \text{constant} \tag{1.1}$$

where g denotes gravity, q is the particle speed and the y-coordinate is taken to be vertically upwards. By a unique choice of origin of y, the constant on the right-hand side may be made to vanish. If also for convenience we take units so that

$$g = 1, \qquad \lambda = 2\pi, \tag{1.2}$$

where λ denotes the wavelength, we have

$$2y + q^2 = 0. \tag{1.3}$$

Now we consider a wave progressing to the left with velocity $c > 0$. In the moving frame of reference in which the wave is stationary, the flow tends to a uniform flow to the right as $y \to -\infty$.

Therefore

$$\begin{aligned} x &\to \phi/c \\ y &\to \psi/c + \tfrac{1}{2}H_0 \end{aligned} \tag{1.4}$$

where H_0 is a constant, determined by our choice of Bernoulli constant in (1.3).

We now consider the following sequence of conformal mappings

$$z(\equiv x+iy) \to \xi(\equiv \phi + i\psi) \to \zeta = e^{-i\xi/c},$$

as illustrated in Fig. 1.

The fluid region occupying one wavelength in the z-plane transforms to the unit circle $|\zeta| \leqslant 1$ in the ζ-plane. (1.4) then implies that the transformation $z(\zeta)$ has a

0020-2932/79/020121+10 $02.00/0