TR/61 April 1976

THE ACCURATE NUMERICAL INVERSION OF LAPLACE TRANSFORMS

by

A. Talbot

The accurate numerical inversion of Laplace transforms

A. TALBOT
Mathematics Department, Brunei University, Uxbridge

Abstract

Inversion of almost arbitrary Laplace transforms is effected by trapezoidal integration along a special contour. The number n of points to be used is one of several parameters, in most cases yielding absolute errors of order 10^{-7} for n = 10, 10^{-11} for n = 20, 10^{-23} for n = 40 (with double precision working), and so on, for all values of the argument from 0+ up to some large maximum.

The extreme accuracy of which the method is capable means that it has many possible applications of various kinds, and some of these are indicated.

1. Introduction

The inversion of Laplace transforms is a topic of fundamental importance in many areas of applied mathematics, as would be evident by a glance at, for example, Carslaw and Jaeger (1948). In the more standard applications the inversion can he 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 et very large number of methods for numerical inversion have been published: see for example the partial bibliography in Piessens and Branders (1971) or the fuller one in Piessens (1974) . A useful critical survey of earlier work was given by Weeks (1966).

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 gaining 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 claimed have probably been those obtained by Piessens and 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. No preliminary computational work, is required. The method is almost universal in its application. 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, being typically of order 10^{-4} or 10^{-5} for n = 6, 10^{-7} for n = 10, 10^{-11} for n = 20, 10^{-23} for n = 40 (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 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+ to some maximum value, usually ranging between 20 and 100 or more, and depending on the accuracy required and the positions of the singularities of F(s). The computer execution time is roughly proportional to n. Using a CDC 7600 the average time per inversion when n = 20 (giving errors nearly all of order 10^{-11} or less) is about 1 ms.

In essence the method is contained in an unpublished Ph.D. thesis (J.S. Green, 1955) which was supervised by the present author. However the potentialities as regards accuracy attainable only became apparent much later, and turn on the correct choice of the various parameters.

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

(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

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

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 21 or more decimal places for $t \le 100$. Triple precision would raise the number of places to about 35, and so on.

2. Description of the method

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 $\mathrm{Re}\, s > \gamma_0$ but diverges in $\mathrm{Re}\, s < \gamma_0$. Our starting point for numerical inversion of F(s) is the standard inversion formula

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

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 $\mathrm{Im}\, s \to \pm\infty$. 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 $\mathrm{Re}\, s \to -\infty$ 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 $\mathrm{Re}\, s \le \gamma_0$ 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 by a given F(s), but for the time being we shall assume that it does hold.

To choose one particular path out of all possible equivalent paths, we write the integral in (2) as

$$\int_B e^{w(s)} ds, \quad w = u + iv \tag{5}$$

Now if $\hat{s}$ is a saddle-point of the integrand, i.e. a zero of dw/ds, then in general, as is well-known, there is a pair of steepest-descent paths through $\hat{s}$ on which, if we write $w(s) = \hat{u} + i\hat{v}$,

(a) $v = \text{const.} = \hat{v}$, so that e^w does not oscillate,

and

(b) u decreases steadily from $\hat{u}$ at $\hat{s}$ to $-\infty$ at both ends.

It is obvious that this pair of paths forms a contour L which, if equivalent to B, is likely to be very suitable for numerical integration of (5).

For later reference we note that if we write $\mu^2 = \hat{u} - u$, then $\mu^2 \geq 0$ on L, and we may suppose μ to increase steadily from $-\infty$ to $+\infty$ along L. Furthermore, on L we have

$$\mu^2 = \hat{w} - w = -\tfrac{1}{2}\hat{W}''.(s - \hat{s})^2 + \ldots, \tag{6}$$

from which, retaining only the first term of the Taylor expansion, we derive the Saddle-Point approximation formula

$$\frac{1}{2\pi i}\int_B e^w\,ds = \frac{e^{\hat{w}}}{2\pi i}\int_{-\infty}^{\infty} e^{-\mu^2}\frac{ds}{d\mu}\,d\mu \tag{7}$$

$$\sim\sim \frac{e^{\hat{w}}}{\sqrt{(2\pi\hat{w}'')}}\,. \tag{8}$$

Now referring to (7) we recall that for real integrals of the form $\int_{-\infty}^{\infty} e^{-x2}\,\phi(x)\,dx$ trapezoidal quadrature has long been known to yield abnormally accurate results. Loosely speaking, this can be explained by reference to the derivative form of the Euler-Maclaurin formula, on noting that the integrand, and all its derivatives vanish at $\pm\infty$. (See for example Hartree (1952), sec. 6.54.) An analysis by oodwin (1949) using contour integration provides a strict explanation of the phenomenon.

We may therefore expect that trapezoidal quadrature applied on a steepest-descent (S.D.) contour will be exceptionally accurate. Wow it would be quite impracticable to calculate the S.D. contour for each function F(s) to be inverted. A little thought however shows that this is unnecessary, for by the discussion above the S.D. contour for <u>any</u> F(s) is likely to produce good results for <u>all</u> F, and this does indeed turn out to be the case. Moreover, it is clear that this will continue to hold if u in (7) is replaced by any

other suitable parameter, for the integrand will still have the same type of "behaviour at the end-points.

For our method we take the simplest possible F, viz. $F(s) = 1/s$, and for simplicity take $t = 1$. This gives

$$w = s - \ell n s, \qquad \hat{s} = 1, \qquad \hat{v} = 0,$$

and the S.D. contour is, taking $\theta = \arg s$ as a parameter,

$$L : s = s_c(\theta) = \alpha + i\theta, \quad \alpha = \theta \cot \theta ; \ -\pi < \theta < \pi , \tag{9}$$

or

$$r = r_c(\theta) = \theta \operatorname{cosec} \theta .$$

Hence the suffix c denotes 'critical'. The reason for its use will become apparent later. The contour is shown in Fig.1.

If with L as in (9) condition (3) is satisfied by a given F(s), we can apply (2) with L in place of B. If however the condition is not satisfied by F(s), it will in general be satisfied by the modified function $F(\lambda s + \sigma)$ for suitable choice of the positive scaling parameter λ and the shift parameter a, for if s_j is a singularity of F(s), $F(\lambda s + \sigma)$ has the corresponding singularity

$$s_j^* \ \frac{s_j - \sigma}{\lambda} . \tag{10}$$

We may then replace (2) by

$$f(t) = \frac{\lambda}{2\pi i} \int_L e^{(\lambda s + \sigma)t} F(\lambda s + \sigma)\, ds , \tag{11}$$

and if F(s) is a real function it is now easy to derive the trapezoidal approximation

$$\tilde{f}(t) = \frac{\lambda e^{\sigma t}}{n} \sum_{k=0}^{n-1} \operatorname{Re}\left[(1 + i\beta) e^{s_c \tau} F(\lambda s_c + \sigma)\right]_{\theta = \theta_k} \tag{12}$$

where $\tau = \lambda t$,

$$\theta_k = k\pi/n , \quad k = 0, 1, \ \ldots \ , n-1, \tag{13}$$

and

$$\beta = -\frac{d\alpha}{d\theta} = \theta + \alpha(\alpha - 1)/\theta \tag{14}$$

The term for $k = n$, i.e. $\theta = \pi$, is omitted since it is zero. If

$$F(\lambda s_c + \sigma) = G + iH \tag{15}$$

where G and H are real, (12) takes the real form

$$\tilde{f}(t) = \frac{\lambda e^{\sigma t}}{n} \sum_{k=0}^{n-1} [e^{\alpha\tau} \{(G - \beta H)\cos \theta\tau - (H + \beta G)\sin \theta\tau\}]_{\theta = \theta_k} . \quad (16)$$

Equations (12) and (16) are the basic formulae of the method. By suitable choice of the parameters n, λ and a they are in general capable of yielding extreme accuracy. The principles governing the choice are simple, and depend on the error analysis given in the next section.

3. Error Analysis

Consider the conformal transformation

$$s = S(z) = \frac{z}{1 - e^{-z}} = \frac{z}{2}(1 + \coth \frac{z}{2}) . \quad (17)$$

This has branch-points at $z = \pm 2r\pi i$, $r = 1,2, \ldots$, and maps the imaginary z-plane interval M between $-2\pi i$ and $2\pi i$ on to the curve L in the s-plane. Some idea of the nature of the mapping in relation to L is given by Fig.2, where correspondences between z-plane and s-plane regions are indicated by shading, or its absence, and correspondences between points by their labels. In particular the region enclosed by L is mapped 1 - 1 from a z-plane region R bounded on the right by M, above by a curve CFH between $-\infty + \pi i$ and $2\pi i$, and below by the conjugate curve. We shall call points z in this region "principal inverses" of the corresponding points s, and shall denote them by the notation

$$z = Z(s). \quad (18)$$

Writing $z = x + iy$, and

$$Q(z) = \lambda e^{(\lambda S + \sigma)t} F(\lambda S + \sigma) \frac{dS}{dz} , \quad (19)$$

(11) becomes

$$f(t) = \frac{1}{2\pi i} \int_M Q dz = \frac{1}{2\pi} \int_{-2\pi}^{2\pi} Q(iy) dy , \quad (20)$$

and the trapezoidal approximation to f(t) is

$$\tilde{f}(t) = \frac{1}{n} {\sum_{k=0}^{n-1}}' 2 \, \mathrm{Re}\, Q(z_k) , \; z_k = 2k\pi i / n , \quad (21)$$

which is seen to agree with (12) on noting that $z_k = 2i\theta_k$ and $s(z_k) = s_c(\theta_k)$.

If the singularities of $F(\lambda s + \sigma)$ are all inside L, those of Q are all in the region R, and if M_1 is any path from $-2\pi i$ to $2\pi i$ passing

to the right of M, and M'_2 any such path to the left of M but close enough to it to exclude the singularities of Q, then by (21) and the Residue Theorem

$$\tilde{f}(t) = \frac{1}{2\pi i} \int_{M_1 - M_2} \frac{Q\,dz}{1 - e^{-nz}} . \tag{22}$$

(The integrand in (22) is regular at $z = \pm 2\pi i$ since by assumption F and hence Q satisfies condition (4).)
Paths M_1, M'_2 and their maps L_1, L'_2 are shown in Fig. 3.

It now follows by (20) and (22) that the theoretical approximation error is

$$E(t) = \tilde{f}(t) - f(t) = E_1 + E'_2$$

where

$$E_1 = \frac{1}{2\pi i} \int_{M_1} \frac{Q\,dz}{e^{nz} - 1} , \tag{23}$$

$$E_2 = \frac{1}{2\pi i} \int_{M'_2} \frac{Q\,dz}{e^{-nz} - 1} . \tag{23}$$

Now if L_2 can be found such that L_2, L'_2 enclose the poles but no other singularities of $F(\lambda s + \sigma)$, then M_2, the principal inverse of L_2, and M'_2 enclose the poles hut no other singularities of Q, and we may write

$$E'_2 = E_2 + E_Q,$$

where

$$E_2 = \frac{1}{2\pi i} \int_{M_2} \frac{Q\,dz}{e^{-nz} - 1} , \tag{24}$$

and, after simplification,

$$E_0 = \sum_j \frac{e^{s_j t}\, \operatorname{res} F(s_j)}{e^{-nz^*_j} - 1} . \tag{25}$$

Here the summation is over poles s_j of F(s), and

$$z^*_j = Z(s^*_j), \quad s^*_j = (s_j - \sigma)/\lambda . \tag{26}$$

Thus the theoretical error is

$$\tilde{E} = E_0 + E_1 + E_2. \tag{27}$$

Since $-z^*_j$ in (25), z in (23) and -z in (24) all have positive real parts, it is obvious that the components E_0, E_1 and E_2 all decrease exponentially to zero as n increases. However, there

9.

is still a round-off error component to consider. Now it is clear that the largest or near-largest term in (12) is the first, viz.

$$T_0 = \frac{\lambda}{2n} e^{(\lambda+\sigma)t} \quad F(\lambda+\sigma) \ . \tag{28}$$

Moreover, because of cancellations, $|\tilde{f}| << |T_0|$. Thus if the computer evaluates T_0 correct to c significant figures, the round-off error in $\tilde{f}$ is

$$E_r = 0(10^{-C} T_0), \tag{29}$$

all other round-off errors in the evaluation of $\tilde{f}$ being negligible by comparison. Finally, we may state that the actual error in $\tilde{f}$ is

$$E = E_0 + E_1 + E_2 + E_r \quad , \tag{30}$$

the four components being given by (25), (23), (24), (29) respectively. We shall now consider these components one by one, and obtain estimates of their orders of magnitude.

Component E_0. By (25) we may indicate the dominant exponential factor in E_0 by writing

$$E_0 \sim e^{-A_0} \ , \quad A_0 = \min_j \ (nu_j^* - p_j t), \tag{31}$$

where $u_j^* = -\mathrm{Re} z_j^* > 0$, but $p_j = \mathrm{Re}\ s_j$ may have either sign.

If s_j is known and X and 0 fixed, s_j^* is found as in (26). A rough idea of the value of u_j^* can then be obtained from Fig.4, which shows s-plane loci u = constant, $s = S(z)$, $z = -u + iy$, for various values of u. More accurate values of u_j^* may be found from Table 1. This gives values of $u = -\mathrm{Re} z$, where $S(z) = s = re^{i\theta}$, for various values of s inside L. Table 1 was computed by applying Newton's method to the equation

$$S(1 - e^{-Z}) - z = 0 \ . \tag{32}$$

Except when θ is near to π, a suitable starting value, ensuring convergence, is $z_0 = (\theta - 3)/18 + 2i\theta$, and it is convenient to take as independent variables in the Table θ and κ, where

$$\kappa = r_c(\theta)/r. \tag{33}$$

For θ near to π we take r and θ as the variables, and $z_0 = -4 + i\pi$. It will be seen from Table 1 that values of u increase rapidly as r approaches zero, and thus A_0 increases rapidly as λ increases.

We note also that poles $s_j^* = 0$ (or $s_j = 0$ if $\sigma = 0$) make no contribution to E_0, since $u \to \infty$ as $s \to 0$.

We may remark that in cases where E_0 is the dominant component, (25) is an accurate formula for E, and thus could be used as a correction term to convert $\tilde{f}(t)$ to $f(t)$. However in such cases it is much simpler and more efficient merely to increase the value of n.

Component E_1. Except near the end points of M_1, the order of magnitude of the integrand in (23) is mainly determined by the factor $e^{\tau S - nz}$. Since $S(z) \sim z$ when $\mathrm{Re}\, z > 2$, say, and M_1 may be deformed arbitrarily far to the right, the modulus of this factor is of order $e^{(t-n)\mathrm{Re}\, z}$ over much of L_1. It follows at once that a necessary condition for E_1 and therefore E(t) to be small is that

$$n > \tau = \lambda t. \tag{34}$$

Thus for large t coupled with large λ necessitated by the presence of remote singularities of F(s), the use of large n is essential for good results.(Large n may not suffice, however, because of the E_r term, as we shall see).

A close estimate of E_1 may be obtained by applying the saddle-point formula (8) to the integral (23). Details of the calculation are given in Appendix 1, but it is immediately obvious from (8) that if $z_1 = X_1 + iy_1$ is a saddle-point in (23), and $s_1 = S(z_1) = p_1 + iq_1$, then

$$E_1 \sim e^{-A_1}, \quad A_1 = nb_1 - \sigma t, \tag{35}$$

where $b_1 = x_1 - \rho p_1$, $\rho = \tau / n$.

Now it has been found empirically (but not so far explained) that for a great variety of functions F(s) and values of τ and n, b_1 is practically a function of ρ only, even though x_1 and p_1 may vary appreciably. Thus the estimation of E_1 and hence of E is greatly facilitated by Table 2 which shows values of $b_1(\rho)$ for a range of $\rho < 1$. (By (34) E_1 will be large if $\rho > 1$.) It will be seen that for optimum E_1, ρ should be between 0.3 and 0.5.

Component E_2. Similar results hold in general for E2, though there is a complication here in that unlike M_1, M_2 cannot be deformed arbitrarily far from M since it must remain to the right of non-polar singularities of $F(\lambda S + \sigma)$. Nevertheless, if $z_2 = -u_2 + iy_2$ is a saddle-point in (24) and $s_2 = S(z_2) = P_2 + iq_2$, then in a large number of cases

$$E_2 \sim e^{-A2}, \qquad A_2 = nb_2 - \sigma t, \tag{36}$$

where $b_2 = u_2 - \rho p_2$ is practically a function of ρ only, and is also given in Table 2.

It is clear that except for very small ρ, which are unlikely to be used, $b_2 > b_1$. Thus in practice we can usually neglect E_2 in comparison with E_1, and this we shall do henceforth unless otherwise indicated.

However, exceptional cases can arise when E_2 is by no means negligible, and in fact dominates E. To see this, suppose that F(s) has a branch point s_0, and that $s_0^* = (s_0 - \sigma)\lambda$ is inside L but "near" to it, i.e. such that if $z_0 = Z(s_0^*)$, $u_0 = -\mathrm{Re} z_0$ is small and positive. Then any modification M_2 of the path M'_2 in (23'), enclosing with M'_2 only poles of $F(\lambda s + \sigma)$, must remain to the right of z_0 and thus must contain an arc on which Re z is small. Thus E_2 will be abnormally large, although it will still decrease exponentially as n increases. An example of this situation occurs when $F(s) = 1/\sqrt{(s^2 + 1)}$ and t is large. Here so = i, and if $\sigma = 0$ the 'critical', i.e. minimum value of λ to keep s_0^* inside L is $\lambda_c = 2/\Pi$. But as is clear from (28) and (29) an increase of t tends to increase E_r unless X is reduced. Thus to make E_r acceptably small when t is large it may be necessary to reduce λ nearly to the value λ_c, and then E_2 may be unacceptably large unless n is very large. More details of this case, and of the improvement effected by judicious choice of σ, will be given later.

Much more work remains to be done in investigating this phenomenon. Experiments with $F(s) = 1/\sqrt{(s^2 + 1)}$ indicate that $d2 \doteq \frac{1}{2} n u_0$ in this case, at any rate when $d = d2 < d_{1r}$. (See (40).) To what extent this generalises to other functions is not at present known.

Component E_r . In the approximations (32), (35), (36) we have included only the exponential factors, and ignored factors of the form λF and other factors. We shall do the same for E_r, but shall include a contribution representing the denominator factor 2n in (28), which may be appreciable. Thus we write

$$E_r \sim e^{-Ar}, \quad A_r = 2.3(c + \log_{10} 2n) - \tau - \sigma t. \tag{37}$$

In practice it is always found that for fixed t, λ and σ, E decreases exponentially as n increases until dominated by E_r (which is practically independent of n), and then remains approximately steady.

As an illustration of the results of this section, consider the case

$$F(s) = \frac{s^3}{s^4 + 4}, \quad f(t) = \cos t \cosh t,$$

with t = 10, λ = 1, σ = 1. Using the CDC 7600 in double precision, for which c is about 27 or 28 for this function F, the following values of E (to 3 figures) were obtained for various n:

n	20	30	40	60	80	100	120	150
E	-2.67D-2	3.88D-5	-5.03D-8	-4.09D-14	1.3D-8	-4.96D-22	2.17D-22	4.91D-22

First, $T_0 = (9.7D7)/n$, so for $n \geq 100$, (29) gives $E_r = 0(10^{-21})$ or $0(10^{-22})$, agreeing with the above values. Next, F(s) has poles $1 \pm i$, $-1 \pm i$, and we find $z^*_j = -0.6465 \pm 2.7961_i$, $-0.5572 \pm 4.6273i$ respectively, and $\mathrm{res}F(s_j) = 1/4$ for each pole. Then by (31)

$$A_0 = \min\ (O.6465\, n - 10\ ,\ 0.5572\, n + 10)\ ,$$

and clearly the first pole-pair is dominant in E_0 if $n < 220$. Considering now E_1, we have $\tau = 10$ and $b_1 > 0.7$ for n between 20 and 200, by Table 2. Thus $E_0 >> E_1$ for all n under consideration, and E_0 is dominant in E up to n = 90, when E_r begins to dominate. One does indeed find that the terms in (25) for the pole-pair $1 \pm i$ do accurately give the above values of E for $n \leq 80$. (A slight discrepancy, of about 2D-21, for n = 80 is almost certainly due to the effect of E_r .)

4. Strategy, and some results

The order of E in (30) is determined by that of its largest component. Thus, writing $d_0 = A_0/2.3$, $d_1 = A_1/2.3$, etc., we have

$$E \sim 10^{-d}, \quad d = \min(d_0, d_1, d_2, d_r). \tag{38}$$

For a given F(s), the relative sizes of the d's vary as we vary t, n, λ, σ, c. Thus the best strategy (i.e. choice of the parameters n, λ, σ) will depend on the value of t, the nature and position of the singularities of F(s), the computer precision c, and the accuracy desired. Fortunately however, except when there are remote singularities or when t is large, a simple general strategy applies in all cases.

Noting that, as pointed out earlier, in general $d_2 > d_1$, we may write

$$d = \min(d_0, d_{1r}) \tag{39}$$

where

$$d_{1r} = \min(d_1, d_r) - \delta_{1r} - \sigma t/2.3, \tag{40}$$
$$\delta_{1r} = \min(nb_1/2.3, c + 2 - \tau/2.3).$$

Table 3 gives values (rounded below) of δ_{1r} when c = 14 or 27, for various values of n and τ, using Table 1 for b_1. In each column d_r is dominant at the bottom entry, and remains so below this. Whenever $d_0 > d_1$ or d_r, $\delta_{1r} - \sigma t/2.3$ gives a safe estimate of the order d of error (to within one or two units), and pairs (n, τ) can be readily selected to produce any attainable accuracy. If $d_r \leq d_1$ larger n yields practically no increase in d_{1r}, but may be needed to ensure that $d_0 > d_r$, i.e. that the value $d = d_{1r}$ is actually attained. This will certainly be the case if F(s) has no poles except possibly at the origin (so that $E_0 = 0$), provided $\lambda = \tau/t$ is large enough to bring all singularities inside L. In the general case when F(s) has poles away from s = 0, d_0 can be found with the help of Table 1. In fact, referring to Fig. 1, let $S_0 = (p_0, q_0)$ be any singularity, polar or non-polar, of F(s), and for fixed σ let $s_0 - \sigma = r_0 e^{i\theta}$. The radius to $s_0 - \sigma$ meets L at the 'critical' point

$$s_c = r_c e^{i\theta} = (p_c, q_c), \tag{41}$$
$$r_c = \theta \operatorname{cosec} \theta, \quad q_c = \theta.$$

The value of A used must be greater than the critical value λ_c bringing $s_0 - \sigma$ to s_c, say

$$\lambda = \lambda_c , \quad \lambda_c = q_0/\theta , \quad > 1 , \tag{42}$$

and then

$$\frac{s_0 - \sigma}{\lambda} = s^* = \frac{s_c}{\kappa} , \tag{43}$$

giving

$$\lambda = q_0 /\theta , \sigma = p_0 - q_0 \cot \theta . \tag{44}$$

If in particular s_0 is a pole s_j, then the corresponding u_j^* in (31) can he found by entering Table 1 with θ and , or θ and r, $r = |s^*|$.

In general, little is gained by taking σ non-zero, and in most of our results $\sigma = 0$. However, in cases where d_r and hence d is small because τ is large (see (40)), non—zero σ can be used to advantage. In fact, we have

$$d_r \text{ -- } c + 2 - (\lambda + \sigma) t / 2.3, \tag{45}$$

and to increase d_r we must make $\lambda + \sigma$ small. Now with a singularity s_0 as above and λ, σ as in (44), $\lambda + \sigma$ is minimised if

$$= \theta^2 \operatorname{cosec}^2\theta = \quad (\theta) \tag{46}$$

Then

$$\lambda + \sigma = p_0 + q_0 \gamma (\theta), \quad \gamma (\theta) = \theta \operatorname{cosec}^2\theta - \cot\theta . \tag{47}$$

Table 4 gives values of (θ) and γ (θ), and also of u(θ) corresponding to θ and (θ), for a number of values of θ. Clearly a compromise is needed between large θ, giving large u (and thus large d_0 If s_0 is a pole) and also large γ (hence small d_r), and small θ, giving small d_0 and large d_r , though it should be noted that one can always compensate for small u by taking n large.
As an illustration of this strategy consider

$$F(s) = \frac{1}{\sqrt{(s^2 + 1)}} , \quad f(t) = J_0 (t) ,$$

and suppose we require f(t) for t = 50. F(s) has branch-points at $S_0 = \pm i$. First, if $\sigma = 0$, the minimum λ is $\lambda_c = 2/\Pi = .637$, and corresponding τ_C -- 32.

The table below records values of d obtained with various n and τ and the corresponding d_r values. (In all cases $d_{lr} = d_r$.)

n	60	60	60	60	60	80	100	120	120	120
τ	35	40	45	50	55	50	50	40	50	60
d	5	10	10	7	6	7	8	12	7	3
d_r	14	12	10	7	5	7	8	12	8	3

It will be seen that in all except the first two cases $d \geq d_r$. With the (n, τ) pair (60, 35), $= \tau / \tau_C = 1.10$ and u = 0.14. The approximate empirical formula $d_2 \geq nu/2$ mentioned earlier gives $d_2 \geq 4$: cf.d = 5. With (60,40), = 1.26 and u = 0.33, giving $d_2 \geq 10 = d$. On the other hand with (120, 40) u is the same but $d_2 \geq 20$, so that now $d = d_r$, and does not increase with further increase of n.

If however we use (44), (47) and Table 4 we note that by (45) d_r decreases from 21 to 12 as θ increases from 0.5 to 1, and any such d_r value is attainable with sufficiently large n, viz. $n > n_2 = 2d_r/u$ if we assume $d_2 \geq nu/2$. For θ = 0.5, 0.6, ... 1 the values of n_2 are about 260, 180, 120, 80, 60, 40 corresponding to d_r = 21, 20, 18, 16, 14, 12. Thus if we aim for d = 20, then θ = 0.6 and $n \geq 80$ are indicated. In fact, n = 180 gives an error 1D—19, while n = 200 gives 1D-20, thus confirming the general strategy.

For larger t the d_r values and hence attainable d values would be smaller. For example, if t = 100 they would be 13, 11, 7, 3, ... for θ = 0.5, ... 1. To obtain d_r = 20 we would need θ = 0.3, giving n_2 = 670. If this is thought excessively high, recourse may be had to a modified contour, viz. our contour L expanded vertically by a factor v. With this it is possible to achieve d = 21 with n = 250 for t = 100, using a similar strategy. Some details of the use of this new parameter v are given in Appendix 2, but much work remains to be done in investigating its effect on the error.

We now consider various types of function F(s), and in a number

of selected examples compare actual d values obtained using various (n,τ) pairs with those obtained by other authors. For all results quoted a CDC 7600 was used, with double precision working unless otherwise stated. The execution time per answer is roughly proportional to the number of points n, i.e. the number of transform function evaluations, and is about 3ms. when n = 20. If single precision is used, then as Table 3 shows d -values of 11 are readily achieved with n = 20, and the average execution time is then 1 ms. If $n < 20$, single precision gives the same d - values as double precision.

In the examples below, each line of results starts with an (n, τ) pair.

I. Singularities only at s = 0.

In these cases $E_0 = 0$, and Table 3, adjusted for the appropriate value of c, may be used with confidence (to within one or two units) for d.

(a) $F(s) = e^{-1/s/\sqrt{s}}$, $f(t) = \cos(2\sqrt{t})/\sqrt{(\pi t)}$.
(10,4) : d = 5 - 8, $t \leq 20$
(20,8.5) : d = 11 - 14, $t \leq 50$
(40,10.5) : d = 23-24, $t \leq 50$.

Cf. Nakhla et al (1973), where 22 points (per value of t) yield d = 3 up to t = 50. The method of Piessens (1972) for this function yields d = 12 to 14 for t between 1 and 10, using 31 points.

(b) $F(s) = e^{-1/\sqrt{s}}/\sqrt{s}$, $f(t) = \left(\int_0^\infty ue^{-u^2/4t} J_0(2\sqrt{u})du\right)/2t\sqrt{(\pi t)}$.
(20,6) : $d \geq 11$, $t < 10$; $d=14$, $10 \leq t \leq 100$
(30,13.5) : $d \geq 17$, $t < 4$; $d = 19$, $4 \leq t \leq 100$
(40,12) : d = 20+ (probably about 25), $t \leq 100$. (Values inferred from 20-figure results using 10 different (n, τ) pairs and comparison with Table 3.)

Cf. Piessens and Branders (1971), example 6, where $d = 9$, $10 \leq t \leq 100$, with 51 points; Piessens (1971), where d = 15, 14, 12 for t = 1, 10, 100, with 12 points; and Levin (1975) where Gaussian quadrature together with Levin's rational transformation of order 14/14 yields d = 12, 15, 12 for t = 1, 10, 100.

(c) $F(s) = (\sqrt{s} + 0.5)/(s + \sqrt{s} + 0.5)$. (For f(t), see Piessens and Branders (1971), example 3.)
(20, 6) : d = 11, t = 0.001; 13, $0.1 \leq t \leq 10$; 14, $10 < t \leq 100$
(30,13.5) : d = 17, t = 0.001: 18, $0.1 \leq t \leq 4$; 19, $6 \leq t \leq 100$
(40,12) : d = 20+ , $t \leq 100$.

Cf. Piessens and Branders (1971), where d = 7 is claimed for t = 2, 4, 6, 10 and d = 5 for t = 14, 20. (in fact the "exact" values quoted for t = 4 to 20 are in error, and d = 7 or 8 is achieved throughout.) Such functions F(s) are important in connection with electric networks containing mixed lumped and distributed elements.

Similar values of d have been obtained by our method for

$F(s) = 1/\sqrt{s}\,(f(t) = 1/\sqrt{(\pi t)}, e^{-\sqrt{s}}\ (e^{-1/4t}/2t\sqrt{(\pi t)})$, and other such cases.

II <u>Poles (if any) at s = 0, other singularities elsewhere</u>
As already indicated, the presence of branch points has a depressing effect on d_2, which can be countered by increasing λ (and so d_r) or n. In these cases d may be less than d_{1r}.
(d) $F(s) = 1/\sqrt{(s^2 + 1)}$, $f(t) = J_0(t)$.
(10,6) : d = 7, $t \leq 1$; 5, t = 5
(20,10) : d = 13, $t \leq 5$; 7, t = 10
(40,18) : d = 20, $t \leq 10$; 13, t = 20
(50,10) : d = 25, $t \leq 6$; 16, t = 10
(60, max(20,t)) : d = 19 or 20, $t \leq 20$; 13, t = 40; 8, t = 50. (Taking τ = max (20,t) ensures that $\lambda \geq 1$, i.e. $\kappa \geq 1.57$ and $u^* \geq 0.65$). Of. Piessens and Branders (1971), ex. 4, where 251 points yield d = 12 up to t = 10, and 11 at t = 20 decreasing to 3 at t = 100. An alternative method applicable only to special classes of functions yields d = 14 up to t = 20 but poor results thereafter.

As already discussed, we can obtain even better results by using $\sigma < 0$. For example, with σ = -1, n = 160 and τ = max(50, 1.5t) we obtain d= 12,14,14,18, 14, 12, 8 for t = 10, 20, 40, 50, 60, 80, 100. (Here we are not using the special strategy described earlier.)
Good results can likewise be obtained for $F(s) = 1/\sqrt{(s^2 - 1)}$, $f(t) = I_0(t)$, though the singularity positions are quite different. Taking n = 60 and τ = max (7, 2t) for example gives $d \geq 20$, $t \leq 5$; 19, t = 10; 9, t = 20.

(e) $F(s) = e^{-s\sqrt{(s+1)}}/s$.

This transform arises in pulse-propagation problems (see Longman, 1973), and its inverse f(t) is not known in explicit form. Levin (1975) gives 10- to 13- decimal place values of f(t) for t = 0.5(.5)2.5. The pairs $(n, \tau) = (40, 12)$ and (40,18), with $d_{1r} = 23$ and 21 respectively, give values of f(t) which are identical to at least 20 decimal places for t = 0.1 to 100, and may be presumed correct to 20 d.p. They confirm Levin's figures to 12 d.p., but show errors of 3×10^{-13} in his figures for t = 2 and t = 2.5.

Similar results are obtained with $F(s) = {}^1/_s \ell n(1 + s)$, f(t) = E, (t); $F(s) = \tan^{-1}(1/s)$ (with logarithmic branch—points at $s = \pm i$), f(t) = (sin t)/t; and so on.

III Rational functions

Since the method is based on an S.D. path for the inversion integral (2) when F(s) = 1/s it is not surprising that it should give good results with rational functions. Now however with poles present other than at s = 0, $E_0 \neq 0$ and

$$d = d_{01r} = \min(d_0, d_{1r}),$$

assuming as before that E_2 is negligible. Here $d_0 = A_0/2.3$ is given by (31). If $p_j < 0$, large t actually helps to increase d_0 and d, but if $p_j > 0$ the opposite occurs. d_0 can be made as large as desired by increasing n, but increase of λ (and use of σ) to increase u^*_j will permit of smaller n, though d_r may also decrease.

Thus compromise is needed to achieve a specified d, but there is usually no difficulty if d is not too close to the computer precision constant c.

The only properties of F(s) relevant to the choice of n, λ and σ are the distances and polar angles of its most distant poles, though only a rough indication is needed, and even without this a succession of choices with increasing n (or, up to a point, τ) will pinpoint f(t) with increasing accuracy. We give two examples.

(f) $F(s) = (s^4+4s^3+4s^2+4s+8)/(s + 1)^5$, $f(t) = e^{-t}(1-t^2+2t^3/3+5t^4/24)$.

The pole $s_j = -1$ being negative real, arbitrary λ may be used, but we note that u* increases with λ. If say n = 20 and $\tau = 9$, $d_{1r} = 11$ by Table 3b. Now $s^*_j = p/\lambda$, $p = -1$, $|s^*_j| = |p|/\lambda = r^*$, say, and

$$A_0 = nu^* - pt = n(u^* + pr^*),$$

which for varying t, i.e. varying r^*, is minimum when $du^*/dr^* = -\rho = -0.45$. Inspection of Table 1 shows that for $\theta = 180°$, this occurs roughly when $r^* = 1.7$, i.e. $\lambda = 0.6$, $t = 15$, $u^* = 1.06$, $A_0 = 36$, $d_0 = 16 > d_{lr}$. It follows that $d = d_{lr}$ for all t. In fact we find that d ranges between 12 and 14 for $t \leq 100$, and E(t) is indeed largest for t about 15. (Note that the multiple pole is not a problem, as in other methods.) Similarly $n = 30$, $\tau = 13.5$ gives $d = 19$ or 20 for $t \leq 100$, while $n = 40$, $\tau = 12$ gives d between 22 and 25, mostly 24. In Piessens and Branders (1971), ex.1, d is 5-7 up to $t = 16$.

(g) $F(s) = 999/(s+1)(s+1000) = 1/(s+1) - 1/(s+1000)$, $f(t) = e^{-t} - e^{-1000t}$.

Such cases, having a large ratio of time-constants, are often described as presenting difficulties for numerical inversion, but they are no problem with our method. For example, in Nakhla et al. (1973), 22 points yield d between 5 and 7 for $t \leq 50$, whereas here the (n, τ) pair (20,6) gives d between 13 and 17 for $t \leq 100$, (30,13.5) gives d between 19 and 22, and so on (up to d_r.)

IV General F(s).

No new principles are involved. We give two examples.

(h) $F(s) = s(s^2+1)\sqrt{(s+1)}$, $f(t) = \int_0^t \cos(t-u)\, e^{-u}\, du/\sqrt{(\pi u)}$.

Piessens and Branders (1971) obtain $d = 5$ or 6 up to $t = 14$, $d = 4$ at $t = 20$. With the pair (40,24) we obtain $d = 19$ or 20 up to $t = 10$, 17 up to $t = 20$.

(i) $F(s) = s\ell ns/(s^2+1)$, $f(t) = -\sin t\, Si(t) - \cos t\, Ci(t)$.

Levin (1975) obtained $d = 5$, $t = 0.1$; 12-14, $t = 1$-4. Results in Piessens and Branders (1971) and others quoted there are poor. This is stated elsewhere by Piessens to be due to the logarithmic singularity. However this does not affect the present method. For example, $n = 40$, $\tau = \max(10.5, 1.8t)$ gives d between 21+ and 11 for $t \leq 20$.

5. General remarks

It will be clear from the examples that the method is almost universal in its scope, except that there may be difficulties for large t, depending on the positions of singularities. Where the inversion problem arises from the solution of linear constant-coefficient differential equations the difficulty for large t can be overcome by using two or three steps to reach t instead of

only one, and making terminal values (including derivatives if necessary) of one step serve as initial values for the next. This would be particularly simple for state-matrix, i.e. first-order problems since no derivatives would need to be found. Alternatively, for all types of problem, the difficulty can often be overcome as we have seen by careful use of the shift parameter σ, and of the new expansion parameter ν. It may be remarked that if the difficulty is due to the existence of a remote pole s_j whose location is accurately known, then there is no need to choose λ so large as to bring the pole inside L : it can be left outside L and its effect taken into account by adding the residue term

$$e^{s_j t} \text{ res } F(s_j) \tag{48}$$

to $\tilde{f}(t)$.

Problems involving delay would seem at first sight to be failing cases. For example, if

$$F(s) = e^{-as} G(s), \tag{49}$$

where $a > 0$ and G satisfies condition (4), then F will in general not satisfy the condition, and the method would be inapplicable to F. However, this is a trivial failure, for we know that the inverse g(t) of G(s) can be found, and

$$\begin{aligned} f(t) &= 0, \quad t < a, \\ &= g(t - a), \quad t > a. \end{aligned}$$

Indeed, the integrand in (2) may be written $e^{s(t-a)} G(s)$, so that for $t > a$ the method may in fact be applied directly to F.

We note in passing that in this method, unlike others, there is no "Gibbs phenomenon" for t close to a. However, it can be shown that for $G(s) = 1/s$, i.e. f(t) a delayed unit step, $\tilde{f}(a) = 1 - 1/2n$, while for $t > a$ $\tilde{f}(t)$ is a function of n and τ only (if $\sigma = 0$), not of t or λ separately. For fixed n and λ, $\tilde{f}(t) \to \tilde{f}(a)$ as $t \to a+$. For fixed τ, $\tilde{f}(t) \to 1$ as n increases.

In general, if F(s) has an infinite number of complex singularities the method will fail, for no value of A can bring them all inside L. If however as a special case

$$F(s) = G(s)/(1 - e^{-as}), \tag{50}$$

where the inverse g(t) of G(s) is a pulse between t = 0 and t = a, then the method will give g(t) in (0,a), and f(t) is a periodic repetition of g(t).

6. Applications

Some indication of the variety of possible applications of the inversion method has already been given in the Introduction. Here we will mention a few of the results so far obtained, leaving a full description of the various processes involved to later papers.

(i) State matrix problems. The solution-vector u(t) of the system

$$\frac{du}{dt} = Au + v(t) \qquad (51)$$

is normally obtained either by using some Runge-Kutta process, or by inverting the transform

$$U(s) = (sI - A)^{-1} W(s), \quad W = V(s) + u(o) \qquad (52)$$

by partial-fraction expansion using the eigenvalues of A, assuming W(s) is rational. In the first case the accuracy attainable is very limited, and in the second case great care has to be taken to avoid serious loss of accuracy through errors in eigenvalues and residues. To find u(t) by numerical inversion of U(s), one must be able to evaluate $(sI-A)^{-1}$ for arbitrary complex s, and the Fadeev algorithm enables this to be done very efficiently and accurately. As an example of results obtainable, in a control problem concerning a boiler system of order 8, the vector output was obtained correct to 11 or more d.p. for 38 values of $t \leq 20$ using n = 20, τ =8 and single precision, with execution time of about 3 ms. per component per value of t. With double precision and n = 40, τ = 14, 22 or more d.p. were obtained.

Clearly the Fadeev algorithm has contributed very little error to these answers, and it is unlikely that the accuracy would be appreciably reduced in the case of much larger systems, but this has yet to be tested.

(ii) Diffusion equation. For the solution of

$$\frac{\partial 2u}{\partial x^2} = \frac{\partial u}{\partial t}, \quad a \leq x \leq b \qquad (53)$$

with initial condition

$$u(x,0) = \phi(x), \quad a<x<b \tag{54}$$

and various end-conditions, two transform variables and successive inversions are required. A new feature here is that the result of the first inversion, which is involved in the second inversion, is a non-real function of the second transform variable, and our formulae (12) and (16) have to be modified accordingly.

As an example of results obtainable, if a = 0, b = 1, u(0,t)=u(l,t) = 0, and

$$ø(x) = 1 - |2x- 1| ,$$

then with (n, τ, σ) = (30, 25, 0) in the first inversion, and (10, 2.5, - λ/2) in the second, and using single precision, d $\geq$ 7 for all x for t $\geq$ 0.1. Here, unlike the normal situation, and in accordance with the theoretical analysis, results are less good for small t: for t = 0.01, d reduces from 6 at x = 0.1 to 4 at x = 0.9. The execution time was about 20ms. per value of u. Similar results are obtained when ø(x) = 1, with a discontinuity in ø at x = 0 and 1 instead of the previous discontinuity in dø/dx at x = 0.5. There is no reason to think that either of these discontinuities worsens the results.(Later work with ø(x) =1, still using single precision, has given d $\geq$ 12 for t $\geq$ 0.3, but this work is not yet completed.)

It may prove possible to tackle Laplace's equation by similar means, but this remains to be investigated.

(iii) Miscellaneous quadratures. To illustrate the use of our method for numerical quadrature, we consider the integrals which were the subject of Burnett and Soroka (1972), namely

$$\begin{matrix} C \\ S \end{matrix}(t,R) = \int_c^d \sqrt{(1-R/x^2)} \cdot \begin{matrix} \cos \\ \sin \end{matrix} tx\,dx , \qquad c=\sqrt{R},\ d=\sqrt{(R+1/R)}.$$

In the paper a complicated approximation procedure was described by which C and S were evaluated correct to 7 d.p. for a range of values of R between 1 and 32 and between 1/2 and 1/32, and for t between 0.1 and 100. By transforming C and S with respect to t, and inverting the transforms, using the expansion parameter ν and the special strategy embodied in Table 4, one can easily obtain 20 d.p. over the whole range of R and of t, using double precision.

(iv) Evaluation of mathematical functions. We have already seen that we can evaluate $J_0(t)$ for $t \leq 100$ correct to 21 or more d.p. by using the CDC 7600, in double precision, to invert its transform $1/\sqrt{(s^2 + 1)}$. The only preparatory work needed is the choice of the parameters n, λ, σ and ν, or of n and θ if Table 4 is used. It would be a simple matter to obtain similar results with other standard mathematical functions.

Now the accuracy obtainable is limited by d_r and therefore by the precision constant c. If a computer were available giving higher precision, not only in its arithmetic working but also in its exponential and sine-cosine subroutines, then the attainable accuracy would be correspondingly increased, for it is always possible to obtain $d = d_r$ by using sufficiently large n. For example, the CDC in triple precision would have $c = 41$, and without any additional work we would be able to obtain $J_0(t)$ correct to 35 d.p.

7. Conclusion.

An almost universal method of numerical inversion has been described which is applicable over very wide ranges of t. It gives errors which can be made smaller than those obtained in any other hitherto published method, and uses only modest amounts of computer time. The method is simple, but requires the evaluation of the transform F(s) at sets of complex points.

So great is the accuracy attainable that it is suggested that transform inversion be considered as a method of solving any mathematical problem the solution of which can be regarded as a value or set of values of a function f(t) whose transform F(s) can be found as a function of s.

Acknowledgment

It has already been mentioned in the Introduction that the method described here was essentially contained in an unpublished Ph.D. thesis, Green (1955). It seems appropriate therefore to indicate which parts of the present paper are in essence to be found in the thesis. They are the following: most of section 2, but with $\lambda = 1$

and $\sigma = 0$ in (12) and (16); the introduction of shift σ and scaling factor λ (though not in combination) for the purpose of bringing singularities inside L; the use of a fixed λ (e.g. $\lambda = 2.$) for reducing errors for certain values of t (but not the use of a 'fixed' τ and hence varying $\lambda = \tau/t$, which is the key to the success of the present method); the error analysis of section 3 leading to the components E_0, E_1, E_2 (but not the component E_r, without which no proper understanding of the behaviour of the error or rational choice of parameters is possible); the condition (34) (with $\lambda = 1$) and the saddle-point formulae (35), (36) for E_1, E2 (but not the property that b_1 and b_2 are practically functions of p only); Figs. 2 - 4 and Appendix 1. Needless to say, the thesis was a constant source of reference in the earlier stages of the work leading to the present paper, and the author is glad to make this acknowledgment to his former student.

REFERENCES

BURNETT, D.S. and SOROKA, W.W. 1972 J.Inst.Maths Applies 10, 325-332,

CARSLAW, H.S. and JAEGER, J.C. I948 Operational methods in applied Mathematics. Oxford : University Press.

FILON, L.N.G. 1928 Proc.Roy.Soc.Edin.49 38-47.

GOODWIN, E.T. 1949 Proc.Camb.Phil.Soc.45, 241-5.

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

HARTREE, D.R. 1952 Numerical analysis. Oxford : University Press.

LEVIN, D. 1975 J.Comp.Appl.Maths. 1, 247 - 250.

LONGMAN, I.M. 1973 SIAM J. Appl .Math. 24, 429-440.

NAKHLA, M., SINGHAL, K. and VLACH, J. 1973 Proc. 16th Midwest Symposium on Circuit Theory 2, XIV. 5-1-9.

PIESSENS, R. 1971 J. Eng.Math.5, 1-9.

PIESSENS, R. 1972 J.Inst. Maths. Applies 10, 185-192.

PIESSENS, R. 1974 A bibliography on numerical inversion of the Laplace transform. Report TW20, Applied Mathematics and Programming Division, Katholicke Universiteit, Leuven, Belgium.

PIESSENS, R. and BRANDERS, M. 1971 Proc. IEE 118, 1517-1522.

SALZER, H.E. 1955 M.T.A.C. 2, 1614-177.

SHOHAT, J. 1940 Duke Math. J. 6 615-626.

TRICOMI, P. 1935 RC Accad.Naz.Iincei, Cl.Sci.Fis. 13, 232-239, 420-426.
WEEKS, W.T. 1966 J.ACM 13, 419-426.
WIDDER, D.V. 1935 Duke Math.J. 1, 126-136.

Appendix 1 : S.P. estimation of E_1 and E2.

The integral for E_1 in (23) is of the form $\int e^{w(z)}dz$, with

$$w(z) = \tau S(z) + \ell nS' + \ell nF(\phi) - \ell n(e^{nz} - 1), \qquad (A1)$$

where $\phi = \phi(z) = \lambda s + \sigma$. Then

$$w'(z) = (\tau+\lambda P)S' + \frac{S''}{S'} - \frac{ne^{nz}}{e^{nz}-1} \qquad (A2)$$

$$w'' = (\tau+\lambda P)S'' + \left(\frac{S''}{S'}\right)' + \lambda^2 P'(\phi)S'^2 + \frac{n^2 e^{nz}}{(e^{nz}-1)^2}, \qquad (A3)$$

where $P = P(\phi) = F'(\phi)/F$.

With S(z) given by (17), we find, for efficient computation of S' etc.,

$$S' = \frac{S}{z}(1 + z - S),$$

$$\frac{S''}{S'} = 1 - \frac{2S}{z} + \frac{1}{1+z-S}, \qquad (A4)$$

$$\left(\frac{S''}{S'}\right)' = \left(\frac{S}{z}-1\right)\left(\frac{2S}{z} + \frac{1-S}{(1+z-S)^2}\right).$$

The saddle-point Z_1 is a root of $w'(z)$, and may be found by a (complex) Newton process. A suitable starting value z_{l0}, usually leading to rapid convergence of the iteration, is obtained by considering the special case when $F(s) = 1/s$, and is given by

$$Z_{10} = 2\pi i + (1-i)\sqrt{(\pi\rho)}. \qquad (A5)$$

Taking into account the conjugate saddle-point $\bar{z}_1$ (assuming F(s) is real) we find by (8) the approximate S.P. formula

$$E_1 \doteq E_{1s} = \mathrm{Re}\left\{\frac{e^{\tau S+\sigma t}\ F(\lambda S+\sigma)\ S'(z)}{(e^{nz}-1)\sqrt{(\pi w''(z)/2)}}\right\}_{z_1}. \qquad (A6)$$

It will be noticed that the value of $w''(Z_1)$ required for E_{ls} will already have been found for the Newton algorithm.
The evaluation of (24) for E_2 is almost identical: the only change needed in the computer program is the replacement of n by -n, and the corresponding replacement of $\sqrt{(\pi\rho)}$ in (A5) by $\sqrt{(\pi\rho)}/i$, since $\rho = \tau/n$. Thus a starting value for the saddle-point z_2 is

$$z_{20} = 2\pi i - (1+i)\sqrt{(\pi\rho)}, \qquad (A7)$$

and

$$E_2 \;=\; E_{2s} \;=\; \mathrm{Re}\left\{\frac{e^{\tau S+\sigma t}\;F(\lambda S+\sigma)\;S'(z)}{(e^{-nz}-1)\;\sqrt{\;(\pi w''(z)/2)}}\right\}_{z_2} \qquad \text{(A8)}$$

Numerous checks have shown that (A6) gives an approximation to E_1 which, is accurate to within one or two percent; similarly for (A8) and E_2, except where a branch-point of F(s) abnormally increases E_2.

Appendix 2 : Generalized contours.

If the error analysis leading to (27) is examined, it will be seen that it does not rest on the particular nature of the function S(z): any function which maps the z-interval M $(-2\pi i,\ 2\pi i)$ onto an s-curve similar in appearance to L would probably do equally well, and would give errors E_0, E_1, E_2 tending exponentially to zero as n increases.

As a particular case, consider the family of mappings

$$s \;=\; \frac{z}{2}\,(1+\cot\frac{z}{2}) \;+\; az\ . \qquad \text{(A9)}$$

Taking $z = 2i\theta$, $-\Pi < \theta < \Pi$, on M gives the s-curve

$$L_\nu : \quad s \;=\; s_\nu \;=\theta\cot\theta \;+\nu\, i\theta\ ,\quad \nu \;=\; 2a+1\ . \qquad \text{(A 10)}$$

In the special case $\nu = 1$ $(a = 0)$ we obtain the curve L. For $\nu > 1$ L_ν consists of L expanded 'vertically' by a factor ν. It is immediately obvious that by using an appreciable value of ν, one can reduce the value of λ and hence τ required to bring a singularity of F(s) inside L, and thus enable d_r and so d to be increased, though naturally at the cost of an increase in n.

The use of this new parameter ν entails slight changes to previous formulae. We replace s_c by s_ν in (12) and (15), 1 by ν in (12). The angles $\theta\nu$ and the terms G,H are multiplied by ν in (16). The strategy described in Section *h* involving the use of (44), (46) and (47) can still be used, the only change being that q_0 is replaced by q_0/ν. This strategy was used in obtaining $J_0(100)$ correct to 21 d.p.

Many features of the use of the parameter ν have still to be investigated, as has the possibility of the use of quite different mapping functions and contours.

KAPPA =	1.05	1.10	1.15	1.20	1.23	1.30	1.35	1.43	1.45	1.50	1.55	1.63	1.70	1.80	1.90	2.00	2.10	2.20	2.40	2.60	2.80	3.0
THETA(DEGREES)																						
0	.10	.19	.27	.35	.43	.50	.57	.64	.70	.76	.82	.88	.98	1.08	1.17	1.26	1.34	1.41	1.55	1.68	1.80	1.9
5	.10	.19	.27	.35	.43	.50	.57	.64	.70	.76	.82	.88	.98	1.08	1.17	1.26	1.34	1.41	1.55	1.68	1.80	1.9
10	.10	.19	.27	.35	.43	.50	.57	.64	.70	.76	.82	.87	.98	1.08	1.17	1.25	1.33	1.41	1.55	1.68	1.79	1.9
15	.10	.19	.27	.35	.43	.50	.57	.63	.70	.76	.82	.87	.98	1.07	1.16	1.25	1.33	1.41	1.55	1.67	1.79	1.8
20	.10	.19	.27	.35	.43	.50	.57	.63	.69	.75	.81	.87	.97	1.07	1.16	1.24	1.32	1.40	1.54	1.67	1.78	1.8
25	.09	.18	.27	.35	.42	.49	.56	.63	.69	.75	.81	.86	.97	1.06	1.15	1.24	1.32	1.39	1.53	1.66	1.77	1.8
30	.09	.18	.27	.34	.42	.49	.56	.6?	.68	.74	.80	.86	.96	1.05	1.14	1.23	1.31	1.38	1.52	1.65	1.76	1.8
35	.09	.18	.26	.34	.41	.49	.55	.62	.68	.74	.79	.85	.95	1.05	1.13	1.22	1.30	1.37	1.51	1.64	1.73	1.8
40	.09	.18	.26	.34	.41	.48	.55	.61	.67	.73	.78	.84	.94	1.03	1.12	1.21	1.29	1.36	1.50	1.62	1.74	1.8
45	.09	.18	.26	.33	.40	.47	.54	.60	.66	.72	.77	.83	.93	1.02	1.11	1.19	1.27	1.35	1.48	1.61	1.7?	1.8
50	.09	.17	.25	.33	.40	.47	.53	.59	.65	.71	.76	.82	.92	1.01	1.10	1.18	1.26	1.33	1.46	1.39	1.70	1.8
55	.09	.17	.25	.32	.39	.46	.52	.58	.64	.70	.75	.80	.90	.99	1.08	1.16	1.24	1.31	1.43	1.57	1.68	1.7
60	.09	.17	.24	.31	.38	.45	.51	.57	.63	.68	.74	.79	.89	.98	1.06	1.14	1.22	1.29	1.42	1.54	1.66	1.7
65	.08	.16	.24	.31	.37	.44	.50	.56	.62	.67	.72	.77	.87	.96	1.04	1.12	1.20	1.27	1.40	1.32	1.63	1.7
70	.08	.16	.23	.30	.36	.43	.49	.55	.60	.66	.71	.76	.85	.94	1.02	1.10	1.17	1.24	1.37	1.49	1.60	1.7
75	.08	.15	.22	.29	.35	.42	.48	.33	.59	.64	.69	.74	.83	.92	1.00	1.08	1.15	1.22	1.35	1.46	1.57	1.6
80	.08	.15	.22	.28	.34	.40	.46	.52	.57	.62	.67	.72	.81	.89	.97	1.05	1.12	1.19	1.32	1.43	1.54	1.6
85	.07	.14	.21	.27	.33	.39	.45	.50	.55	.60	.63	.70	.79	.87	.93	1.02	1.09	1.16	1.28	1.40	1.50	1.6
90	.07	.14	.20	.26	.32	.38	.43	.48	.53	.58	.63	.67	.76	.84	.92	.99	1.06	1.13	1.25	1.36	1.46	1.5
95	.07	.13	.19	.25	.31	.36	.41	.46	.51	.56	.60	.65	.73	.81	.89	.96	1.03	1.09	1.21	1.32	1.42	1.5
00	.06	.12	.18	.24	.29	.34	.39	.44	.49	.53	.58	.62	.70	.78	.85	.92	.99	1.05	1.17	1.28	1.38	1.4
05	.06	.12	.17	.22	.28	.33	.37	.42	.46	.51	.55	.59	.67	.75	.82	.88	.95	1.01	1.12	1.23	1.33	1.4
10	.06	.11	.16	.21	.26	.31	.35	.40	.44	.48	.52	.56	.64	.71	.78	.84	.91	.96	1.08	1.18	1.27	1.3
15	.05	.10	.15	.20	.24	.29	.33	.37	.41	.45	.49	.53	.60	.67	.74	.80	.86	.92	1.02	1.13	1.22	1.3
20	.05	.09	.14	.18	.22	.27	.31	.35	.38	.42	.46	.49	.56	.63	.69	.75	.81	.87	.97	1.07	1.16	1.2
25	.04	.08	.12	.16	.20	.24	.28	.32	.35	.39	.42	.46	.32	.58	.64	.70	.76	.81	.91	1.00	1.09	1.1
30	.04	.06	.11	.15	.18	.22	.25	.29	.32	.35	.39	.42	.48	.54	.59	.65	.70	.75	.85	.94	1.02	1.1
35	.03	.07	.10	.13	.16	.19	.23	.26	.29	.32	.35	.38	.43	.49	.54	.59	.64	.69	.78	.86	.94	1.0
40	.03	.36	.08	.11	.14	.17	.20	.22	.25	.28	.30	.33	.38	.43	.48	.53	.57	.62	.70	.78	.86	.9
45	.02	.35	.07	.09	.12	.14	.17	.19	.21	.24	.26	.28	.33	.38	.42	.46	.50	.54	.62	.70	.77	.8
50	.02	.04	.06	.08	.10	.12	.14	.16	.18	.20	.22	.24	.28	.31	.35	.39	.43	.46	.53	.60	.67	.7
55	.01	.03	.04	.06	.07	.09	.11	.12	.14	.15	.17	.19	.22	.25	.28	.32	.35	.38	.44	.50	.56	.6
60	.01	.02	.03	.04	.05	.06	.08	.09	.10	.11	.12	.14	.16	.19	.21	.24	.26	.29	.34	.39	.44	.4
65	.01	.31	.02	.03	.03	.04	.05	.06	.06	.07	.08	.09	.11	.12	.14	.16	.18	.20	.24	.27	.31	.3

R =	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.2	1.4	1.6	1.8	2.0	3.0	4.0	5.0	6.0	8.0	10.0	12.0
70	3.98	3.19	2.74	2.42	2.17	1.98	1.82	1.68	1.56	1.45	1.28	1.14	1.02	.92	.83	.53.	36	.26	.19.	10	.06	.03
75	4.00	3.21	2.76	2.44	2.20	2.01	1.85	1.71	1.60	1.49	1.32	1.18	1.06	.96	.88	.59	.42	.31	.23	.15	.10	.02
80	4.02	3.23	2.78	2.47	2.23	2.04	1.88	1.75	1.63	1.53	1.36	1.22	1.11	1.01	.92	.63	.47	.36	.28	.19	.13	.11

TABLE 1. VALUES OF U=-RE(Z(S)) FOR ARG(S)=THETA, ABS(S)=R1: KAPPA = R(CRIT.)/R.

ρ	.025	.05	.075	.10	.15	.20	.25	.30	.35	.40	.45	.50
b_1	.530	.726	.862	.966	1.118	1.221	1.290	1.332	1.352	1.353	1.338	1.306
b_2	.535	.743	.896	1.021	1.223	1.385	1.523	1.644	1.751	1.848	1.937	2.019

ρ	.55	.60	.65	.70	.80	.90
b_1	1.260	1.199	1.124	1.033	.803	.491
b_2	2.094	2.165	2.232	2.295	2.410	2.515

Table 2.

τ / n	4	6	8	10	12	14	16	18	20	22	24	26	28	30	33
10	5	5	3												
15	8	8	8	7	5										
20	10	11	11	11	10	9	7	4							
25	12	13	12	12	11	10	9	8	7	6					
30	14										5	4	3		
35														3	
40															2

Table 3a. Values of δ_{1r} when c = 14

τ / n	4	6	8	10	12	14	16	18	20	22	24	26	28	30	33	36	39	42	45	48	51	54	57
10	5	5	3																				
15	8	8	8	7	5																		
20	10	11	11	10	9	7	4																
25	12	13	14	14	14	13	12	10	8	6													
30	14	15	17	17	17	17	16	15	14	12	10	8	4										
35	15	17	19	20	20	20	20	19	18	17	16	14	12	9	3								
40	16	19	21	22	23	23	22	21	20	19	18	17	16	16	12	8							
45	18	20	22	24	24										15	13	12	6					
50	19	22	24	25														10	9	5			
55	20	23	25																	8	6		
60	21	25																					5 4
65	22	26																					
100	27																						

Table 3b. Values of δ_{1r} when c = 27.

θ	o.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
κ(θ)	1.055	1.088	1.129	1.181	1.244	1.320	1.412	1.523	1.658	1.820	2.018	2.261
γ(θ)	0.272	0.345	0.420	0.499	0.583	0.673	0.770	0.876	0.993	1.123	1.269	1.437
u(θ)	0.104	0.161	0.228	0.307	0.394	0.489	0.592	0.701	0.816	0.936	1.062	1.191

Table 4.

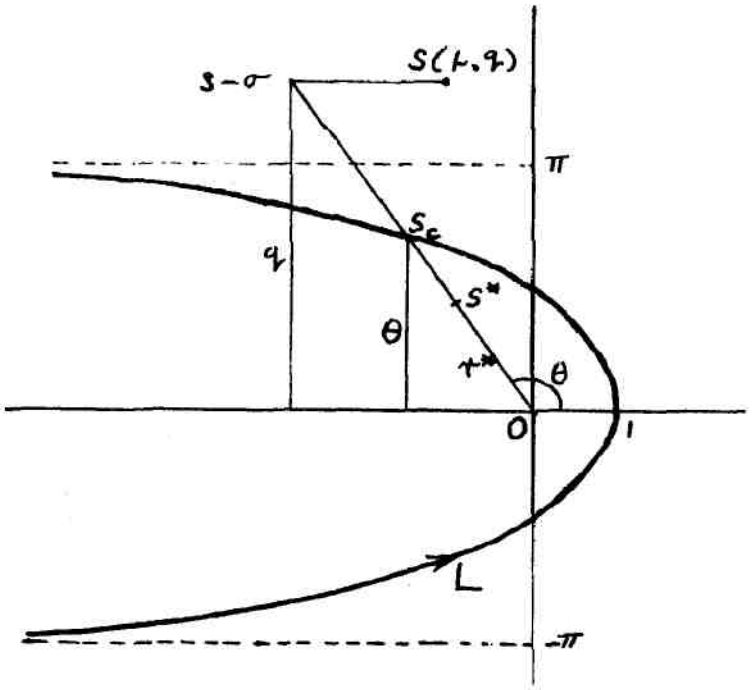
s-σ
S(r,ϑ)
π
S_c
q
S^*
θ
τ^*
θ
0
1
L
-π

Fig. 1

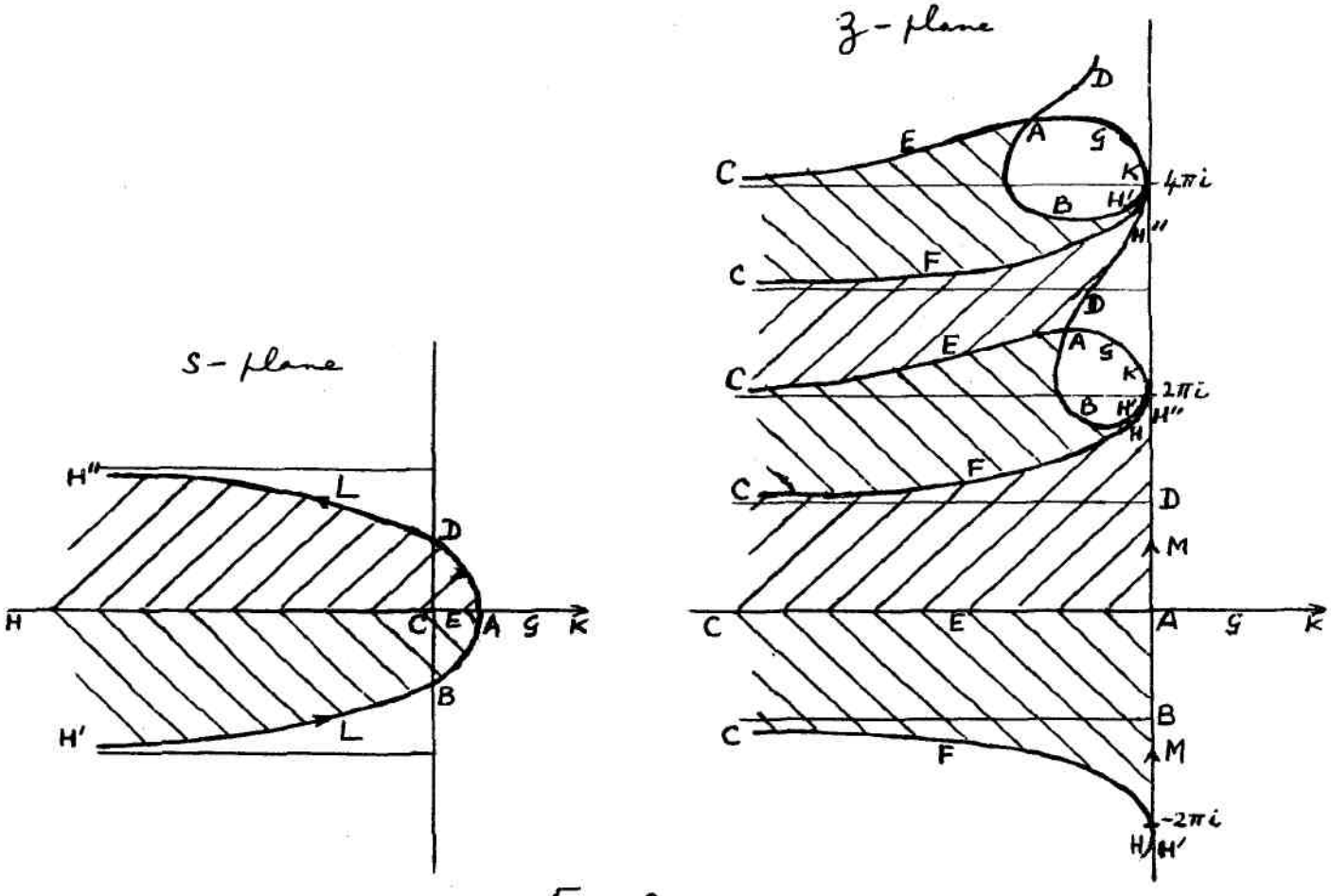
s-plane
H''
L
D
H
C
E
A
G
K
B
H'
L
z-plane
D
C
E
A
G
K
4πi
B
H'
H''
H
F
C
D
E
A
G
K
2πi
B
H'
H''
H
F
C
D
M
C
E
A
G
K
B
C
F
M
-2πi
H
H'

Fig. 2

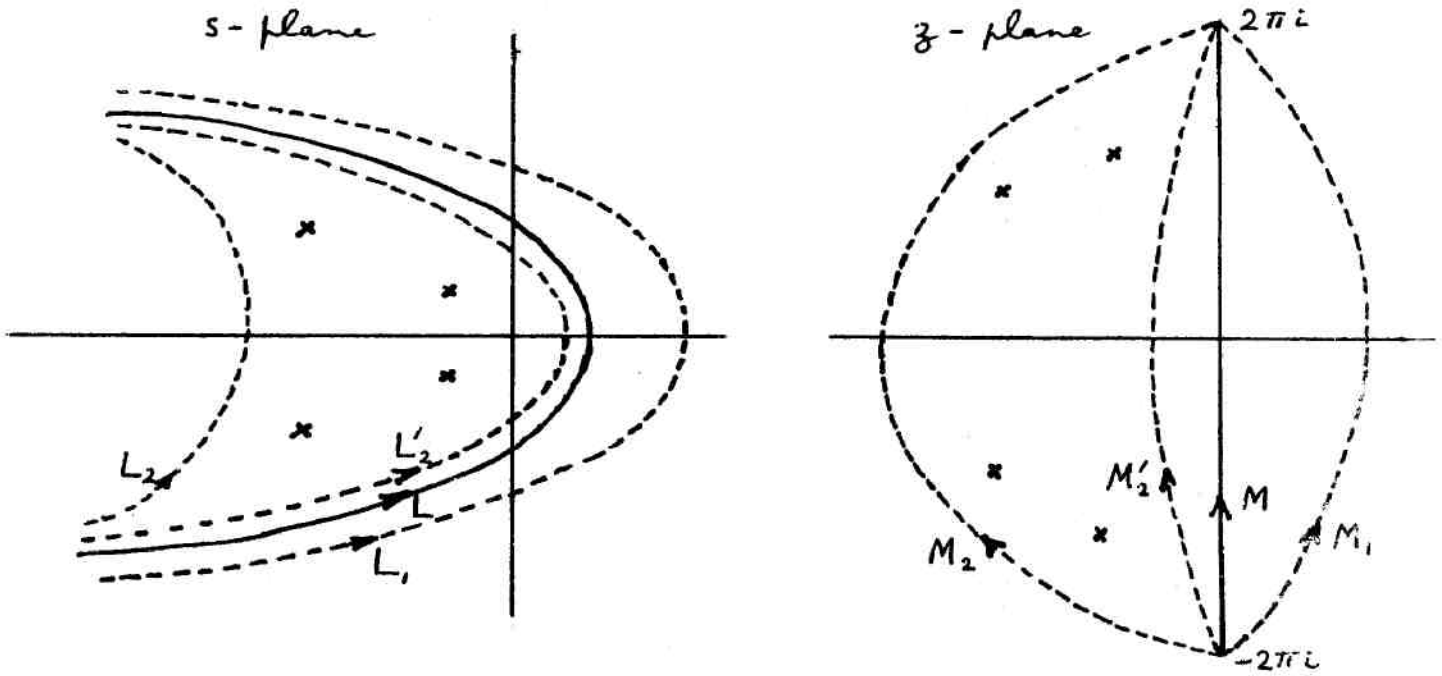

Fig. 3

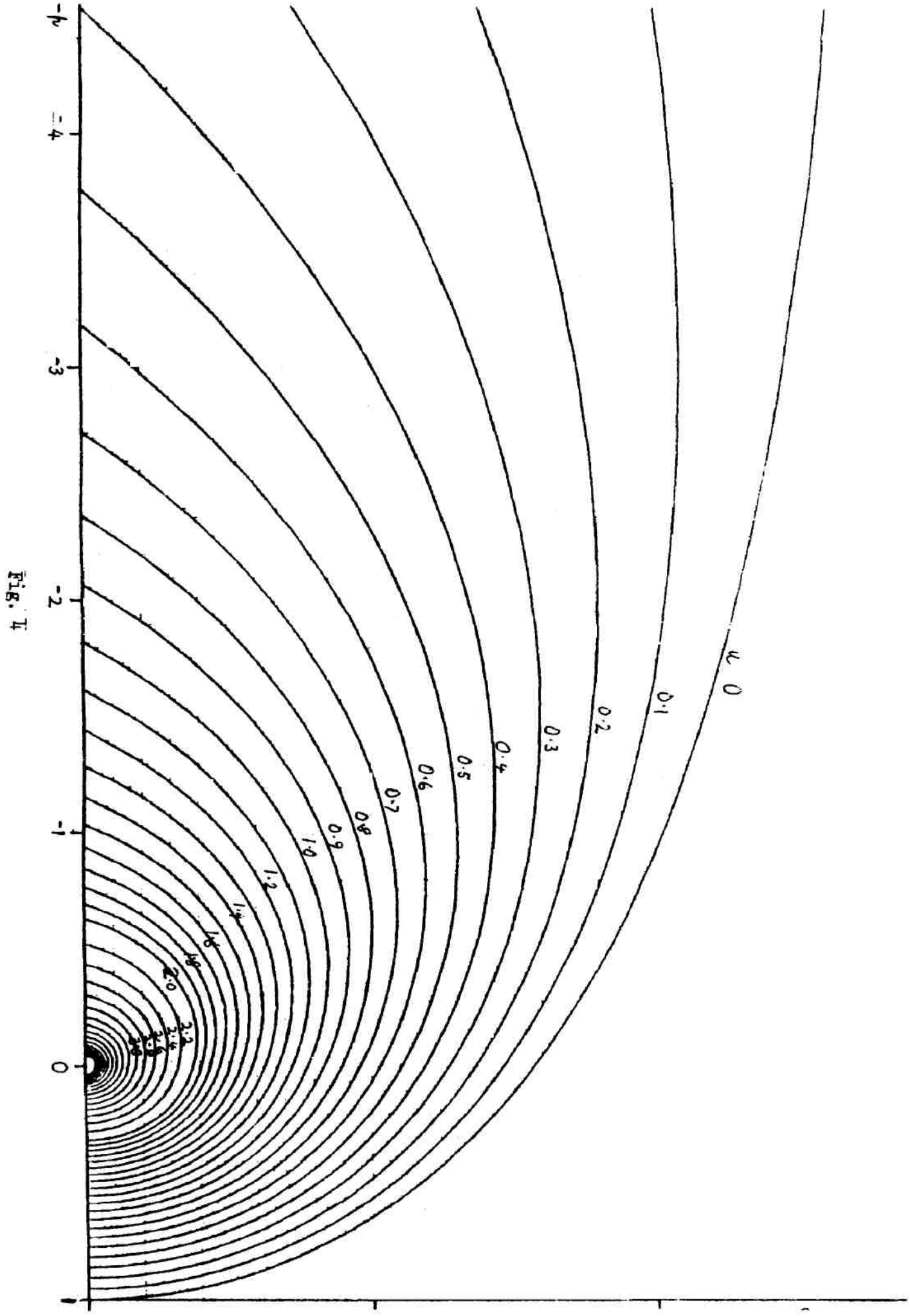

Fig. 1