THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 61, No. 8, October 1982
Printed in U.S.A.

An Inversion Technique for the Laplace Transform

By D. L. JAGERMAN

(Manuscript received February 8, 1982)

In this paper we use an approximation sequence defined by the Widder Laplace transform inversion formula to provide a practical method for inverting the Laplace transform. The approximation sequence converges uniformly and retains essential structural characteristics of the original function, e.g., nonnegativity, monotonicity, and convexity. Thus, we approximate a distribution function by distribution functions and use enhancement techniques to increase the speed of convergence and to capture the quality of exponential decay. Also, we present a practical computational method illustrated by examples.

I. INTRODUCTION

The purpose of this paper is to summarize the techniques presented in the paper "An Inversion Technique for the Laplace Transform"[1] and to make available a useful reference of properties of the 'approximation sequence,' and a new numerical method developed since the publication of Ref. 1. The inversion, or approximation sequence, retains the essential structural characteristics of the original function, e.g., nonnegativity, monotonicity, and convexity. Thus, we approximate a distribution function by distribution functions. For application to queueing theory, this may be considered quite important. The basic inversion sequence, together with error estimates, is discussed in Section II; also, two enhancement procedures are given—namely, the construction of a sequence that is more rapidly convergent than the approximation sequence and which was not given in Ref. 1, and a method of accurate approximation to functions that decay exponentially. Section III discusses the new numerical method, and Section IV presents two examples of numerical inversion along with controls. Except for the new material whose derivations are given here, all proofs can be found in Ref. 1.

II. INVERSION SEQUENCE

Consider a function $f(t)$ for which the Laplace transform, $\bar{f}(s)$, defined by

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

exists for $s > c$ and $f(t) = 0(e^{ct})(t \to \infty)$; then a sequence of functionals, $(L_n)_0^\infty$, is defined by

$$L_n f(t) = f_n(t) = \frac{(-1)^n}{n!} s^{n+1} \left. \frac{d^n \bar{f}(s)}{ds^n} \right|_{s = \frac{n+1}{t}}. \tag{2}$$

One has

$$\lim_{n \to \infty} f_n(t) = f(t) \tag{3}$$

uniformly in every finite closed interval throughout which $f(t)$ is continuous. Equations (2) and (3) constitute a variant of Widder's theorem.[2] The sequence $[f_n(t)]_0^\infty$ is called the "approximation sequence" of $f(t)$.

The following lists some important properties of $f_n(t)$ valid for $t \geq 0$, $n \geq 0$; a dot indicates differentiation. Assumptions on $f(t)$ are to hold in $(0, \infty)$.

(*i*) $a \leq f(t) \leq b \Leftrightarrow a \leq f_n(t) \leq b$,

$$f_n(0+) = f(0+),$$

$$\dot{f}_n(0+) = \dot{f}(0+),$$

$$f_n(\infty) = f(\infty),$$

$$\dot{f}_n(\infty) = \dot{f}(\infty),$$

(*ii*) $f(t)$ monotone $\Rightarrow f_n(t)$ monotone in the same sense. $f(t)$ completely monotone, absolutely monotone, convex, log-convex implies the same property, respectively, for $f_n(t)$.

(*iii*) $f(t) \geq 0 \Rightarrow \dfrac{f_n(t)}{n+1}$ monotone decreasing in n, $f(t)$ convex $\Leftrightarrow$ $f(t) \leq f_n(t), \ddot{f}(t) \geq 0 \Rightarrow f_n(t)$ monotone decreasing in n.

(*iv*) If $f(t) * g(t) = \displaystyle\int_0^\infty f\left(\frac{t}{u}\right) g(u) \frac{du}{u}$ (Mellin convolution), then $f_n(t)$ is the approximation sequence of $f(t) \Rightarrow f_n(t) * g(t)$ is the approximation sequence of $f(t) * g(t)$.

The pointwise error, $\epsilon_n(t; f)$, is defined by

$$\epsilon_n(t; f) = f_n(t) - f(t). \tag{4}$$

One has

$$\epsilon_n(t; f) \sim \sum_{l=1}^{\infty} \frac{1}{l!} G_l(-n-1, -n-1) t^l f^{(l)}(t), \qquad n \to \infty, \tag{5}$$

in which $G_l(-n-1, -n-1)$ are Poisson-Charlier polynomials.[3] Set $\beta_l = G_l(-n-1, -n-1)$; then the following recursion is satisfied:

$$\beta_{l+1} = \frac{l}{n+1}[\beta_l + \beta_{l-1}], \qquad \beta_0 = 1, \qquad \beta_1 = 0. \tag{6}$$

The terms of eq. (5) through order n^{-2} are

$$\epsilon_n(t; f) \sim \frac{1}{2}\frac{1}{n+1} t^2 f^{(2)}(t) + \frac{1}{(n+1)^2}\left[\frac{1}{3} t^3 f^{(3)}(t) + \frac{1}{8}\frac{n+3}{n+1} t^4 f^{(4)}(t)\right], \qquad n \to \infty. \tag{7}$$

Some bounds for $\epsilon_n(t; f)$ are

$$|\epsilon_n(t; f)| \le \frac{t^2}{2n+2} \sup_{x>0} |\ddot{f}(x)|, \tag{8}$$

for pointwise error, and

$$|\epsilon_n(t; f)| \le \alpha_n \sup_{t>0} |t\dot{f}(t)| \tag{9}$$

uniform over all t. The α_n satisfies

$$\alpha_n \sim \sqrt{\frac{2}{\pi n}}, \qquad n \to \infty. \tag{10}$$

It appears, therefore, that the convergence of $f_n(t)$ to $f(t)$ is not rapid. In fact, from eq. (7) it is $0(1/n)$ and is not improved by assuming a higher degree of smoothness for $f(t)$. However, one may decrease the error, $\epsilon_n(t; f)$, for a given n in at least two ways by: (*i*) modifying the sequence $f_n(t)$ to obtain a higher convergence rate, and (*ii*) modifying $f_0(t)$ to improve its approximation to $f(t)$. The first way may destroy the desirable properties of the approximation sequence as previously listed because the transformation from $f(t)$ to the members of the new sequence may no longer be positive; the second way can be implemented to preserve all the desirable properties with respect to a function different from the original $f(t)$.

The following procedure improves the convergence rate but does not preserve positivity. Define the functionals, σ_n, by

$$\sigma_n f = \sigma_n(t) = \left(2 + \frac{1}{n}\right) f_{2n}(t) - \left(1 + \frac{1}{n}\right) f_n(t); \tag{11}$$

then, from eq. (7)

$$\sigma_n(t) - f(t) \sim -\frac{1}{3}\frac{1}{(n+1)(2n+1)}t^3 f^{(3)}(t) - \frac{1}{8}\frac{2n^2+9n+5}{(n+1)^2(2n+1)^2}t^4 f^{(4)}(t), \qquad n \to \infty, \quad (12)$$

and hence

$$\sigma_n(t) - f(t) = 0(1/n^2). \quad (13)$$

Often, $f(t)$ decreases exponentially fast for $t \to \infty$; the inversion method may not capture this effect well for large t if the decrease is very rapid. This problem is largely overcome by the following technique, which constitutes an example of the second method of improving the approximation.

If $\tilde{f}(s)$ converges for $s \geq -\beta (\beta > 0)$, and $0 \leq \alpha \leq \beta$, then $\tilde{f}(s - \alpha)$ exists and is the transform of $g(t) = e^{\alpha t} f(t)$. If $g_n(t)$ is the approximation sequence for $e^{\alpha t} f(t)$ and

$$f_{n,\alpha}(t) = e^{-\alpha t} g_n(t), \quad (14)$$

then

$$\lim_{n \to \infty} f_{n,\alpha}(t) = f(t). \quad (15)$$

The error $\epsilon_{n,\alpha}(t; f) = f_{n,\alpha}(t) - f(t)$ can be much smaller than $\epsilon_n(t; f)$, corresponding to $\alpha = 0$, so that a real improvement in accuracy can result; of course,

$$\epsilon_{n,\alpha}(t; f) = e^{-\alpha t} \epsilon_n(t; g). \quad (16)$$

The approximation $f_{n,\alpha}(t)$ imitates the rapid exponential decrease of $f(t)$ with increasing accuracy as α approaches β. Since $g_n(t)$ is an approximation sequence, it possesses all of the properties given earlier, many of which carry over to $f(t)$.

III. NUMERICAL EVALUATION

A type of generating function may be constructed for $[f_n(t)]_0^\infty$. Let

$$G(z, t) = \frac{1}{t}\tilde{f}\left(\frac{1-z}{t}\right) \quad (17)$$

and define $[a_n(t)]_0^\infty$ by

$$G(z, t) = \sum_{n=0}^{\infty} a_n(t) z^n; \quad (18)$$

then

$$f_n(t) = a_n\left(\frac{t}{n+1}\right). \quad (19)$$

This permits the use of numerical procedures for obtaining coefficients in a power series. The following is such a method. The notation

$$e_q(x) = e^{i\frac{2\pi x}{q}} \tag{20}$$

will be used. Let $0 < r < 1$, q prime and $q > n$, and define the functional S by

$$Sf = \frac{n+1}{tqr^n} \sum_{j=1}^{q} e_q(-nj)\tilde{f}\left\{\frac{n+1}{t}[1 - re_q(j)]\right\}, \tag{21}$$

then, in terms of $f_n(t)$, one has

$$Sf = f_n(t) + \sum_{l=1}^{\infty} f_{n+lq}\left(\frac{n+lq+1}{n+1}t\right) r^{lq}. \tag{22}$$

For the purpose of computing $f_n(t)$, Sf in the form of eq. (21) will be used.

We will now assume that $f(t)$ is bounded so one may take $|f(t)| \leq A$. Clearly, from eq. (22) and property 1,

$$|Sf - f_n(t)| \leq A\frac{r^q}{1 - r^q} \tag{23}$$

and

$$\lim_{r \to 0+} Sf = f_n(t). \tag{24}$$

Thus, Sf is an accurate approximation to $f_n(t)$ when r is small; however, if the round-off error of each term of eq. (21) is bounded by ϵ, and the total round-off error by η, then

$$\eta \leq \epsilon \frac{n+1}{t} r^{-n}. \tag{25}$$

It follows that a choice of r may be made to ensure the round-off error, η, is not too large, in accordance with eq. (25). The parameter q may now be chosen to render the truncation error given by eq. (23) comparably small.

IV. EXAMPLES

A Fortran program for the evaluation of Sf, as shown in Eq. (21), was written by Bharat Doshi. Double precision arithmetic was used with a round-off error of 10^{-15}. As a rough estimate, it was assumed that the computations required to form each term of the summation in eq. (21) resulted in a round-off error of 10^{-11}; thus the choice $\epsilon = 10^{-11}$ was made. The choices

$$r = 0.83, \qquad q = 127; \qquad n = 50,$$

$$r = 0.91, \qquad q = 251; \qquad n = 100 \tag{26}$$

lead to about five units error in the sixth decimal; the values of q given result in negligible truncation errors. These errors in computing f_{50} and f_{100} were considered acceptably small.

The function

$$f(t) = 1/2e^{-t} + 1/2e^{-3t} \tag{27}$$

is chosen for the first example. It represents a complementary distribution function whose values may be accurately computed as a control, and which shows the enhancement of accuracy obtainable by use of eqs. (11) and (14). Tables I and II illustrate these results. Since

$$\tilde{f}(s) = \frac{1}{2}\frac{1}{s+1} + \frac{1}{2}\frac{1}{s+3}, \tag{28}$$

the choice $\alpha = 1$ is the appropriate value to use.

It may be observed, from Table I, that the error is halved in going from f_{50} to f_{100} as expected from the $0(1/n)$ behavior of eq. (7), while the decrease of error from f_{100} to σ_{50} is better than 1/50 as seen from the $0(1/n^2)$ rate of eq. (13). Table II shows that the effect of α is greater for larger t since the exponential term that is being tracked becomes dominant.

The second example concerns the complementary busy-period distribution for an $M/M/1$.[4] The arrival rate is ρ and the service rate is $\mu = 1$. For this case

$$f(t) = 1 - \rho^{-1/2} \int_0^t e^{-(1+\rho)x} I_1(2x\sqrt{\rho}) \frac{dx}{x}, \tag{29}$$

and

$$\tilde{f}(s) = \frac{1}{2\rho s}\left[\rho - 1 - s + \sqrt{(\rho + 1 + s)^2 - 4\rho}\right], \tag{30}$$

Table I—$\alpha = 0$

t	f	f_{50}	f_{100}	σ_{50}
2	0.06891	0.07203	0.07048	0.06890
4	0.00916	0.01064	0.00990	0.00915
8	0.00017	0.00030	0.00023	0.00016

Table II—$\alpha = 1$

t	f	$f_{50,1}$	$f_{100,1}$	$\sigma_{50,1}$
2	0.06891	0.06911	0.06901	0.06891
4	0.00916	0.00916	0.00916	0.00916
8	0.00017	0.00017	0.00017	0.00017

Table III—$\rho = 0.9,\ \alpha = 0$

t	f	f_{50}	f_{100}	σ_{50}
3	0.2919	0.2942	0.2931	0.2919
6	0.1950	0.1967	0.1959	0.1950
9	0.1514	0.1528	0.1521	0.1514

Table IV—$\rho = 0.5,\ \alpha = 0$

t	f	f_{50}	f_{100}	σ_{50}
3	0.1803	0.1834	0.1819	0.1803
6	0.0764	0.0785	0.0775	0.0764
9	0.0399	0.0415	0.0407	0.0399

Table V—$\rho = 0.1,\ \alpha = 0$

t	f	f_{50}	f_{100}	σ_{50}
3	0.0732	0.0774	0.0753	0.0732
6	0.0088	0.0103	0.0096	0.0088
9	0.0013	0.0018	0.0016	0.0013

Table VI—$\rho = 0.1,\ \alpha = 0.4675$

t	f	$f_{50,\alpha}$	$f_{100,\alpha}$	$\sigma_{50,\alpha}$
3	0.0732	0.0742	0.0737	0.0732
6	0.0088	0.0090	0.0089	0.0088
9	0.0013	0.0014	0.0014	0.0013

in which $I_1(2x\sqrt{\rho})$ is a modified Bessel function. The evaluation of the control was accomplished by using a quadrature procedure. The values $\rho = 0.9, 0.5$, and 0.1 were used. The best choice of α is, from Section II, $(1 - \sqrt{\rho})^2$; however, when α is small, one could simply set $\alpha = 0$, since there is little improvement in using the best α. For the case $\rho = 0.1$, one has $\alpha = 0.4675$; accordingly, a comparison was made with $\alpha = 0$. Tables III through VI sample the results obtained.

V. CONCLUSION

It may be observed that the best α for $\rho = 0.1$ noticeably improved the approximation. The performance of the σ_n functionals in all cases created a marked improvement in accuracy; however, one should remember that the desirable properties of the L_n functionals are not possessed by the σ_n. Nonetheless, there is a property of the σ_n, important in numerical work, which follows from eq. (11). Namely, if δ is the round-off error in the computation of $f_n(t)$ and $f_{2n}(t)$, then $(3 + 2/n)\delta$, $n \geq 1$ bounds the induced round-off error in σ_n (i.e., round-off errors are not significantly magnified).

An interesting use of the inversion technique follows. A function $f(t)$ is given in the form

$$f(t) = \int_0^t g(t, x)dx, \tag{31}$$

with $g(t, x)$ known; it is required to evaluate $f(t)$ over a wide range of values of t. Quadrature methods are accurate, for t small, or can be designed, for t large, but do not apply equally well over all values of t, including the important transition region from small to large. If the transform $\tilde{f}(s)$ can be obtained, then the inversion method described here can be used to obtain sufficiently accurate values of $f(t)$ over the entire range of t, since eq. (9) shows that the errors, $\epsilon_n(t; f)$, are uniformly bounded.

REFERENCES

1. D. L. Jagerman, "An Inversion Technique for the Laplace Transform with Application to Approximation," B.S.T.J., *57,* No. 3 (March 1978), pp. 669–710.
2. I. I. Hirschman and D. V. Widder, *The Convolution Transform*, Princeton, NJ: Princeton University Press, 1955.
3. D. L. Jagerman, "Some Properties of the Erlang Loss Function," B.S.T.J., *53,* No. 3 (March 1974), pp. 525–51.
4. R. B. Cooper, *Introduction to Queueing Theory*, New York: Macmillan Co., 1972.