Numerical inversion of Laplace transforms: an efficient improvement to Dubner and Abate's method

F. Durbin

Commissariat a l'Energie Atomique Centre U—Service Electronique, B.P. 61, 92120—Montrouge, France

An accurate method is presented for the numerical inversion of Laplace transform, which is a natural continuation to Dubner and Abate's method. (Dubner and Abate, 1968). The advantages of this modified procedure are twofold: first, the error bound on the inverse $f(t)$ becomes independent of t, instead of being exponential in t; second, and consequently, the trigonometric series obtained for $f(t)$ in terms of $F(s)$ is valid on the whole period $2T$ of the series. As it is proved, this error bound can be set arbitrarily small, and it is always possible to get good results, even in rather difficult cases. Particular implementations and numerical examples are presented.

(Received June 1973)

1. Introduction

Let $f(t)$ be a real function of t, with $f(t) = 0$ for $t < 0$; the Laplace transform and its inversion formula are defined as follows:

$$F(s) = \mathscr{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)\, dt \quad s = a + i\omega \tag{1}$$

$$f(t) = \mathscr{L}^{-1}\{f(t)\} = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{st} F(s)\, ds \tag{2}$$

$a > 0$ is arbitrary, but is greater than the real parts of all the singularities of $F(s)$.

In case of singularities of $F(s)$ to the right of the origin, a suitable translation of the imaginary axis can always reject those singularities to the left of the origin.

In all this discussion we therefore assume that (1) and (2) exist for $\mathrm{Re}\,(s) \geqslant a > 0$.

2. Development into trigonometric integrals

Let us expand (1):

$$F(s) = \int_0^\infty e^{-at} f(t) \cos \omega t\, dt - i \int_0^\infty e^{-at} f(t) \sin \omega t\, dt \tag{3}$$

$$F(s) = \mathrm{Re}\,\{F(a + i\omega)\} + i\, \mathrm{Im}\,\{F(a + i\omega)\} \tag{4}$$

We now expand (2), with $ds = i\, d\omega$

$$f(t) = \frac{1}{2\pi i}\int_{-\infty}^{+\infty} e^{at} (\cos \omega t + i \sin \omega t) (\mathrm{Re}\{F(s)\} + i\, \mathrm{Im}\,\{F(s)\})\, i\, d\omega \tag{5}$$

$$f(t) = \frac{e^{at}}{2\pi}\left[\int_{-\infty}^{+\infty} (\mathrm{Re}\,\{F(s) \cos \omega t - \mathrm{Im}\,\{F(s)\} \sin \omega t)\, d\omega + i \int_{-\infty}^{+\infty} (\mathrm{Im}\,\{F(s)\} \cos \omega t + \mathrm{Re}\,\{F(s)\} \sin \omega t)\, d\omega\right] \tag{6}$$

The imaginary part in (6) cancels out, because of the parity of $\mathrm{Re}\,\{F(s)\}$ and $\mathrm{Im}\,\{F(s)\}$; using this parity again, we have:

$$f(t) = \frac{e^{at}}{\pi}\left[\int_0^\infty (\mathrm{Re}\,\{F(s)\} \cos \omega t - \mathrm{Im}\,\{F(s)\} \sin \omega t\right] d\omega \tag{7}$$

For $t < 0$, $f(t) = 0$, which means that:

$$\int_0^\infty (\mathrm{Re}\,\{F(s)\} \cos \omega t + \mathrm{Im}\,\{F(s)\} \sin \omega t)\, d\omega = 0 \tag{8}$$

Consequently, we obtain 3 formulas for the Laplace inverse $f(t)$ corresponding to $F(s)$:

$$f(t) = \frac{2e^{at}}{\pi}\int_0^\infty \mathrm{Re}\,\{F(s)\} \cos \omega t\, d\omega \tag{9}$$

$$f(t) = \frac{-2e^{at}}{\pi}\int_0^\infty \mathrm{Im}\,\{F(s)\} \sin \omega t\, d\omega \tag{10}$$

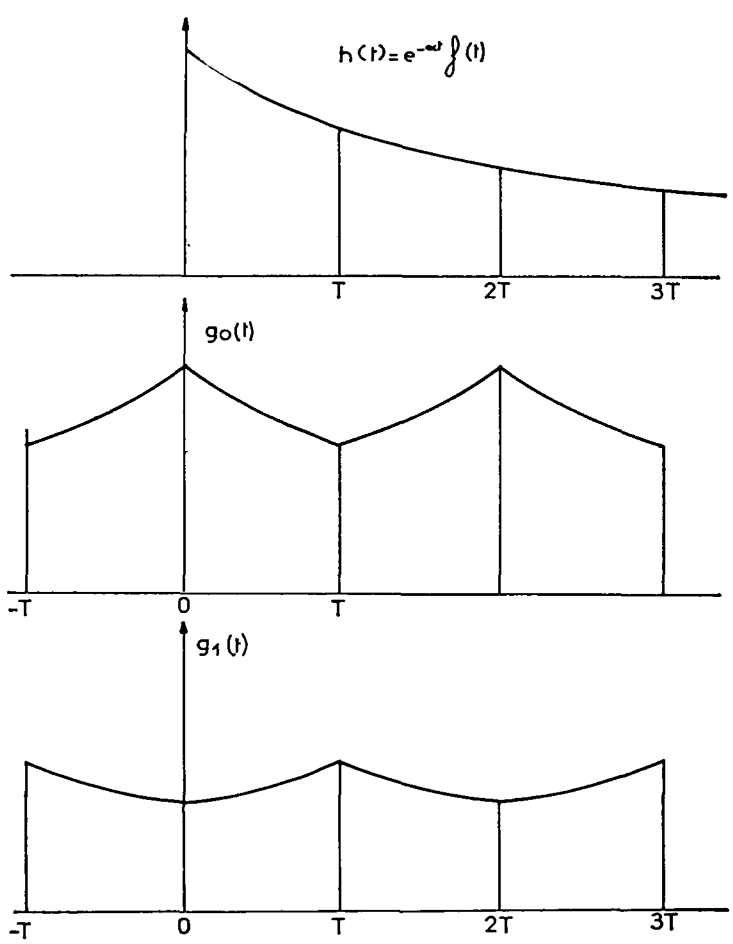

Fig. 1

$$f(t) = \frac{e^{at}}{\pi}\int_0^\infty (\mathrm{Re}\,\{F(s)\} \cos \omega t - \mathrm{Im}\,\{F(s)\} \sin \omega t)\, d\omega \tag{11}$$

3. Method

We first summarise Dubner and Abat's method; this **will** enable us to expose easily its natural continuation. Let $h(t)$ **be** a real function of t, with $h(t) = 0$ for $t < 0$;

(*a*) Consider sections of $h(t)$ in intervals like $(nT, (n + I)T)$, construct an infinite set of $2T$-periodic functions $g_n(t)$:

$$n = 0, 2, 4, \ldots\; g_n(t) = \begin{cases} h(nT - t) & -T \leqslant t \leqslant 0 \quad (12)\\ h(nT + t) & 0 \leqslant t \leqslant T \quad (13)\\ h((n+2)T - t) & T \leqslant t \leqslant 2T \quad (14)\end{cases}$$

$$n = 1, 3, 5, \dots\ g_n(t) = \begin{cases} h((n+1)T + t) & -T \leqslant t \leqslant 0 \quad (15) \\ h((n+1)T - t) & 0 \leqslant t \leqslant T \quad (16) \\ h((n-1)T + t) & T \leqslant t \leqslant 2T \quad (17) \end{cases}$$

(*b*) Develop each $g_n(t)$ into cosine Fourier series:

$$g_n(t) = \frac{A_{n,0}}{2} + \sum_{k=1}^{\infty} A_{n,k} \cos \Omega_k t;\ \Omega_k = k\frac{\pi}{T} \tag{18}$$

(*c*) Evaluate:

$$A_{n,k} = \frac{2}{T}\int_{nT}^{(n+1)T} h(t) \cos \Omega_k t\, dt \tag{19}$$

(*d*) Since it is always possible to write

$$h(t) = e^{-at} f(t) \tag{20a}$$

or

$$f(t) = e^{at} h(t) \tag{20b};$$

we have:

$$\sum_{n=0}^{\infty} A_{n,k} = \frac{2}{T} \int_0^{\infty} e^{-at} f(t) \cos \Omega_k t\, dt = \frac{2}{T} \operatorname{Re}\{F(a + i\Omega_k)\} \tag{21}$$

$$\sum_{n=0}^{\infty} e^{at} g_n(t) = \frac{2e^{at}}{T} \left[\tfrac{1}{2}\operatorname{Re}\{F(a)\} + \sum_{k=1}^{\infty} \operatorname{Re}\left\{F\left(a + i\frac{k\pi}{T}\right)\right\} \cos \frac{k\pi}{T} t\right] \tag{22}$$

(*e*) Use relations no. (13, 14, 16, 17, 20a, 20b) to obtain:

$$\sum_{n=0}^{\infty} e^{at} g_n(t) = f(t) + \sum_{k=1}^{\infty} e^{-2akT} (f(2kT + t) + e^{2at} f(2kT - t)) \tag{23}$$

$$= f(t) + \mathrm{ERROR1}\,(a, t, T) \tag{24}$$

In conclusion, for any $0 \leqslant t \leqslant 2T$, we can write:

$$f(t) + \mathrm{ERROR1}\,(a, t, T) = \frac{2e^{at}}{T} \left[\tfrac{1}{2}\operatorname{Re}\{F(a)\} + \sum_{k=1}^{\infty} \operatorname{Re}\left\{F\left(a + i\frac{k\pi}{T}\right)\right\} \cos \frac{k\pi}{T} t\right] \tag{25}$$

This is Dubner and Abate's formula; ERROR1 is a function of a, t, T; clearly the factor $\sum_{k=1}^{\infty} e^{-2akT} f(2kT - t)\, e^{2at}$ is the most disturbing one since it increases exponentially with t.

Numerically (25) is valid only for $t \leqslant T/2$.

4. The natural continuation of the method

Just as in Section 3, we consider $h(t)$ in the interval $(nT, (n+1)T)$, but this time we construct an infinite set of odd $2T$-periodic function $k_n(t)$. See **Fig. 2**.

By definition, we have:

$$n = 0, 1, 2, \dots\ k_n(t) = \begin{cases} h(t) & nT \leqslant t \leqslant (n+1)T \\ -h(2nT - t) & (n-1)T \leqslant t \leqslant nT \end{cases}$$

Similarly, on the intervals $(-T, +T)$, $(0, T)$, $(T, 2T)$, we can write

$$n = 0, 2, 4, \dots\ k_n(t) = \begin{cases} -h(nT - t) & -T \leqslant t \leqslant 0 \quad (26a) \\ h(nT + t) & 0 \leqslant t \leqslant T \quad (26b) \\ -h((n+2)T - t) & T \leqslant t \leqslant 2T \quad (26c) \end{cases}$$

$$n = 1, 3, 5, \dots\ k_n(t) = \begin{cases} h((n+1)T + t) & -T \leqslant t \leqslant 0 \quad (27a) \\ -h((n+1)T - t) & 0 \leqslant t \leqslant T \quad (27b) \\ h((n-1)T + t) & T \leqslant t \leqslant 2T \quad (27c) \end{cases}$$

The Fourier representation for each odd function $k_n(t)$ is:

$$k_n(t) = \sum_{k=0}^{\infty} B_{n,k} \sin k\frac{\pi}{T} t = \sum_{k=0}^{\infty} B_{n,k} \sin \Omega_k t \tag{28}$$

Just as for $A_{n,k}$, we find:

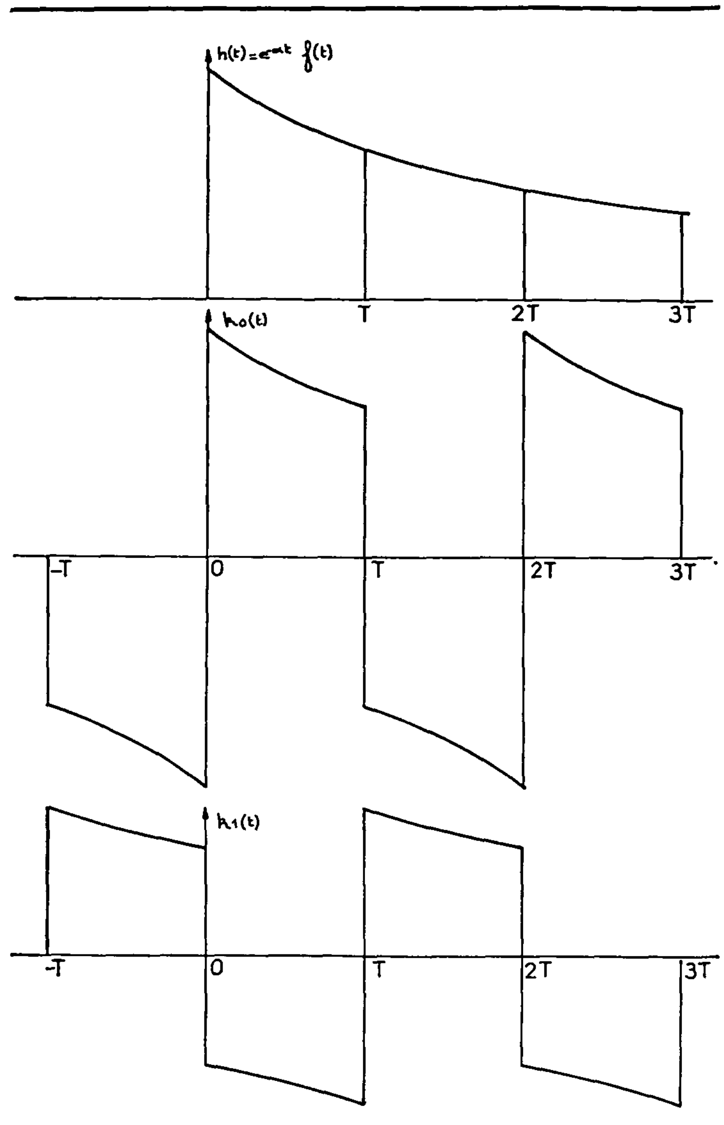

Fig. 2

$$B_{n,k} = \int_{nT}^{(n+1)T} e^{-at} f(t) \sin \Omega kt\, dt \tag{29}$$

Summing (29) over n and comparing it with (3):

$$\sum_{n=0}^{\infty} B_{n,k} = \frac{2}{T}\int_0^{\infty} e^{-at} f(t) \sin \Omega kt\, dt = -\frac{2}{T} \operatorname{Im}\left\{F\left(a + ik\frac{\pi}{T}\right)\right\}$$

Summing (28) over n and multiplying both sides by e^{at}, we obtain a relation similar to (22):

$$\sum_{n=0}^{\infty} e^{at} k_n(t) = -\frac{2e^{at}}{T}\left[\operatorname{Im}\left\{F\left(a + ik\frac{\pi}{T}\right)\right\} \sin k\frac{\pi}{T} t\right] \tag{30}$$

Likewise, on the interval $(0, 2T)$, using (20a), (26b), (26c), (27b), (27c), we find:

$$\sum_{n=0}^{\infty} e^{at} k_n(t) = f(t) + \sum_{k=1}^{\infty} e^{-2akT} [f(2kT + t) - e^{2at} f(2kT - t)] \tag{31}$$

Another representation for $f(t)$ is therefore:

$$f(t) + \mathrm{ERROR2}(a, t, T) = -\frac{2e^{at}}{T}\left[\sum_{k=0}^{\infty} \operatorname{Im}\left\{F\left(a + ik\frac{\pi}{T}\right)\right\} \sin k\frac{\pi}{T} t\right] \tag{32}$$

5. Error analysis

Let us write down the two similar expansions (25) and (32):

$$f(t) + \sum_{k=1}^{\infty} e^{-2kaT} [f(2kT + t) + e^{2at} f(2kT - t)] =$$

$$\frac{2e^{at}}{T}\left[\tfrac{1}{2}\mathrm{Re}\,\{F(a)\} + \sum_{k=1}^{\infty}\mathrm{Re}\left\{F\left(a + ik\frac{\pi}{T}\right)\right\}\cos k\frac{\pi}{T}t\right] \quad (33)$$

$$f(t) + \sum_{k=1}^{\infty} e^{-2kaT}[f(2kT + t) - e^{2at} f(2kT - t)] =$$

$$-\frac{2e^{at}}{T}\left[\sum_{k=0}^{\infty}\mathrm{Im}\left\{F\left(a + ik\frac{\pi}{T}\right)\right\}\sin k\frac{\pi}{T}t\right] \quad (34)$$

Clearly, any one of these formulas does not show any specific advantage: both error terms contain a factor which is exponentially increasing with t; however these factors have opposite signs.

Since we know that $F(s)$ has no singularities for $\mathrm{Re}\,F(s) > 0$, then $|f(t)|$ is bounded at infinity by some function of the form Ct^m, where C is a constant and m a nonnegative integer. We consider first the important case of all physical functions such as $|f(t)| < C$; then ERROR1 (a, t, T) and ERROR2 (a, t, T) have the same bound, which is:

$$\left.\begin{array}{l}|\mathrm{ERROR1}\,(a, t, T)| \\ |\mathrm{ERROR2}\,(a, t, T)|\end{array}\right\} \leqslant \sum_{k=1}^{\infty} C\,e^{-2kaT}(1 + e^{2at}) =$$

$$C\frac{e^{-2aT}}{1 - e^{-2aT}}(1 + e^{2at})$$

$$\leqslant C\exp(-aT + at)\frac{\cosh at}{\sinh aT} \quad (35)$$

We are going to reduce considerably this bound as follows:
Let us sum half of both sides of (33) and (34):

$$f(t) + \sum_{k=1}^{\infty} e^{-2akT} f(2kT + t) = f(t) + \mathrm{ERROR3}(a, t, T) =$$

$$= \frac{e^{at}}{T}\left[\tfrac{1}{2}\mathrm{Re}\,\{F(a)\} + \sum_{k=1}^{\infty}\mathrm{Re}\left\{F\left(a + ik\frac{\pi}{T}\right)\right\}\cos k\frac{\pi}{T}t\right.$$

$$\left. - \sum_{k=0}^{\infty}\mathrm{Im}\left\{F\left(a + ik\frac{\pi}{T}\right)\right\}\sin k\frac{\pi}{T}t\right] \quad (36)$$

This time, if $|f(t)| < C$, the bound for ERROR3 (a, t, T) is:

$$|\mathrm{ERROR3}\,(a, t, T)| \leqslant \sum_{k=1}^{\infty} C\,e^{-2kaT} = \frac{C}{e^{2aT} - 1}. \quad (37)$$

The interest of this result is twofold:

1. ERROR3 (a, t, T) is now bounded by a fixed quantity; this allows us to use our representation of $f(t)$ on the interval $(0, 2T)$ instead of only $(0, T/2)$;
2. This fixed bound depends only on the product aT. Once the precision $Q = \mathrm{MAX}\,\{\mathrm{ERROR3}\,(a, t, T)\}$ is chosen, a is determined. For example, with $aT = 10$, we find $Q = C\,.\,2\,10^{-9}$ whereas the original method gave only $Q = C\,.\,10^{-5}, 0 \leqslant t \leqslant T/2$.

We now consider the case $|f(t)| < C\,.\,t^m$:

$$|\mathrm{ERROR3}\,(a, t, T)| \leqslant \sum_{k=1}^{\infty} e^{-2kaT} C(t + 2kT)^m < C(2T)^m$$

$$\sum_{k=1}^{\infty} e^{-2kaT}(k + 1)^m \quad (38)$$

Each term of the series of positive terms $u(k) = e^{-2kaT}(k + 1)^m$ is decreasing uniformly to zero for $k > k_1$; therefore $\sum_{k=k_1}^{\infty} u(k)$ and

$$\int_{k_1+1}^{\infty} e^{-2xaT}(x + 1)^m\,dx$$

are of the same nature. Clearly the integral is convergent; consequently, α being some positive constant such that:

$$\sum_{k=1}^{m} u(k) = \alpha\int_{1}^{\infty} e^{-2xaT}(x + 1)^m\,dx$$

we obtain the bound:

$$\mathrm{ERROR3}\,(a, t, T) \leqslant \alpha C(2T)^m \int_{1}^{\infty} e^{-2xaT}(x + 1)^m\,dx \quad (39)$$

The computation of the integral is straightforward: (Gradshteyn and Riszhik, 1965)

$$\int_{1}^{\infty} e^{-2xaT}(x + 1)^m\,dx = e^{2aT}\int_{2}^{\infty} e^{-2uaT}u^m\,du$$

$$= \frac{e^{-2aT}}{2aT}\left(2^m + \sum_{k=1}^{m}\frac{(m(m-1)\ldots(m-k+1)}{(2aT)^k}2^{m-k}\right)$$

In conclusion:

$$|\mathrm{ERROR3}\,(a, t, T)| \leqslant$$

$$K(2T)^m e^{-2aT}\left(\frac{\alpha_1}{2aT} + \frac{\alpha_2}{(2aT)^2} + \cdots\frac{\alpha_{m+1}}{(2aT)^{m+1}}\right) \quad (40)$$

$$K, \alpha_1, \alpha_2, \ldots, \alpha_{m+1} = \text{Constants}\ .$$

Again, we see that the error term decreases very quickly with aT, but this time depends also upon T.

Comparison of equations (36) and (11) shows that our approximation is formally equivalent to the application of the trapezoidal rule to (11), the integration step being π/T. But the error bound we obtained, proportional to $\exp(-2aT)$, is much tighter than the bound directly associated with the trapezoidal rule, which decreases like $1/T^2$.

On the other hand, by applying directly the trapezoidal rule to (9), (10), or (11), therefore using implicitly a fundamental result (de Balbine and Frank, 1966), according to which this rule is as good as any other rule of quadrature for infinite range Fourier integrals, one could not have seen the influence of the parameter aT.

But above all, the possibility of cancellation for 2 exponentially increasing opposite error factors would not have been in a conspicuous position.

6. Numerical implementation

Since we are going to compare Dubner and Abate's method with the modified one, over the interval $(0, 2T)$, we change T into $T/2$ in (25) and (36).

Also, the infinite series involved can only be summed up to a number NSUM of terms; therefore truncation error Et and roundoff error Er must be accounted for:

$$f(t) + \mathrm{ERROR1}\,(a, t, T) = \frac{4e^{at}}{T}$$

$$\left[-\tfrac{1}{2}\mathrm{Re}\,\{F(a)\} + \sum_{k=0}^{NSUM}\mathrm{Re}\left\{F\left(a + ik\frac{2\pi}{T}\right)\right\}\cos\frac{2\pi}{T}t\right] \quad (41)$$

$$f(t) + \mathrm{ERROR3}\,(a, t, T) = \frac{2e^{at}}{T}\left[-\tfrac{1}{2}\mathrm{Re}\,\{F(a)\} + \right.$$

$$\sum_{k=0}^{NSUM}\left(\mathrm{Re}\left\{F\left(a + ik\frac{2\pi}{T}\right)\right\}\right.$$

$$\left.\left.\cos k\frac{2\pi}{T}t - \mathrm{Im}\left\{F\left(a + ik\frac{2\pi}{T}\right)\right\}\sin k\frac{2\pi}{T}t\right)\right] \quad (42)$$

We have proved in Section 5 that both ERROR1 (a, t, T) and ERROR3 (a, t, T) decreased with $\exp(-aT)$; but practically, for each t, Er and Et are amplified by the factor $\exp(at)/T$; too large a value of aT would require too large a value of *NSUM* for a given accuracy.

We also have tried various convergence acceleration methods, e.g. epsilon algorithm, Euler method and others (D. Shanks, 1955), but all these procedures are efficient only when the terms of the original series decrease monotonically in modulus.

$F(s)$ being a Laplace Transform, we know that

$$\mathrm{Re}\left\{F\left(a + ik\frac{2\pi}{T}\right)\right\}$$

and

$$\mathrm{Im}\left\{F\left(a + ik\frac{2\pi}{T}\right)\right\}$$

tend to 0 when k tends to infinity; to apply efficiently one of the above mentioned procedures, one would have first to find, in each case, the value of k after which

$$\left|F\left(a + ik\frac{2\pi}{T}\right)\right|$$

decreases monotonically to 0. This is virtually impossible for the very complicated $F(s)$ we had to invert, as shown in Section 8.

An economical, and, up to now, successful way of doing such summations is the following one: the real and imaginary part of $F(s)$ are evaluated together through a complex, single precision arithmetic subroutine, but are converted into double precision constants for the summation up to *NSUM*; the results are then turned back to single precision expressions.

Thus one avoids time and storage consuming systematic double precision computation; NSUM can be determined by the convergence criterion:

$$\left|\mathrm{Re}\left\{F\left(a + i\,\mathrm{NSUM}\frac{2\pi}{T}\right)\right\}\right| \text{ and }$$

$$\left|\mathrm{Im}\left\{F\left(a + i\,\mathrm{NSUM}\frac{2\pi}{T}\right)\right\}\right| \leqslant \frac{\varepsilon T}{\exp(aT)} \quad (43)$$

$$\varepsilon = 10^{-6} \text{ to } 10^{-10}.$$

We found that $aT = 5$ to 10 gave good results for NSUM ranging from 50 to 5000.

For a fair comparison between Dubner and Abate's method and the modified method, we took NSUM terms for the cosine series, but we used NSUM/2 sine terms and NSUM/2 cosine terms for the modified method.

Clearly, the running time will be less for the 2nd method, since the subroutine which anyway computes

$$\mathrm{Re}\left\{F\left(a + ik\frac{2\pi}{T}\right)\right\}$$

and

$$\mathrm{Im}\left\{F\left(a + ik\frac{2\pi}{T}\right)\right\}$$

using complex arithmetic has to be called NSUM/2 times instead of NSUM times.

We tested the following examples:

function 1: $F(s) = s(s+1)^{-2}$; $f(t) = (t/2)\sin(t)$

function 2: $F(s) = s^{-1}\exp(-10s)$; $f(t) = U(t-10)$

U = Heaviside's step function.

We took $aT = 5$, $T = 20$, NSUM = 2000.

METHOD1 = Dubner and Abate's method

METHOD2 = Modified method.

The computer was an IBM 370/155, run through the time sharing option (TSO). See **Tables 1** and **2** for the results.

One can see that METHOD2 gives accurate results on $(0, 2T)$, whereas METHOD1 breaks down for $t \geqslant T/4$. It is interesting to notice that for function 2, which possesses a discontinuity at $t = 10$, METHOD2 gives a numerical value, namely 0·506790, very close to its theoretical one,

$$\tfrac{1}{2}[f(10+0) + f(10-0)] = 0{\cdot}5\,;$$

METHOD1 does not follow this discontinuity.

7. Final implementation

The 'Fast Laplace Inverse Transform'. This implementation will be called 'FLIT' in the sequel.

Table 1

Test function 1 $F(t) = (t/2)\sin t = \mathscr{L}^{-1}\{s(s^2+1)^{-2}\}$

t	Method1	Method2	FLIT	Exact $F(t)$
0·0	0·12437E + 00	0·62186E − 01	0·62186E − 01	0·0
1·0	0·49700E + 00	0·42566E + 00	0·42565E + 00	0·42073E + 00
2·0	0·78468E + 00	0·90776E + 00	0·90775E + 00	0·90929E + 00
3·0	−0·10387E + 00	0·14530E + 00	0·14530E + 00	0·21168E + 00
4·0	−0·17074E + 01	−0·15867E + 01	−0·15867E + 01	−0·15136E + 01
5·0	−0·20114E + 01	−0·24075E + 01	+0·24075E + 01	−0·23973E + 01
6·0	0·17604E + 00	−0·77050E + 00	−0·77050E + 00	−0·83824E + 00
7·0	0·30208E + 01	0·23865E + 01	0·23865E + 01	0·22994E + 01
8·0	0·28200E + 01	0·39821E + 01	0·39821E + 01	0·39574E + 01
9·0	−0·15723E + 01	0·17886E + 01	0·17886E + 01	0·18543E + 01
10·0	−0·56405E + 01	−0·28202E + 01	−0·28202E + 01	−0·27201E + 01
11·0	−0·25923E + 01	−0·55413E + 01	−0·55413E + 01	−0·54999E + 01
12·0	0·76657E + 01	−0·31588E + 01	−0·31588E + 01	−0·32194E + 01
13·0	0·13538E + 02	0·28427E + 01	0·28426E + 01	0·27310E + 01
14·0	0·13008E + 01	0·69941E + 01	0·69941E + 01	0·69342E + 01
15·0	−0·25504E + 02	0·48254E + 01	0·48254E + 01	0·48771E + 01
16·0	−0·34295E + 02	−0·24241E + 01	−0·24241E + 01	−0·23032E + 01
17·0	−0·34397E + 01	−0·82514E + 01	−0·82514E + 01	−0·81718E + 01
18·0	0·42842E + 02	−0·67196E + 01	−0·67196E + 01	−0·67589E + 01
19·0	0·44738E + 02	0·15511E + 01	0·15511E + 01	0·14236E + 01
	1·510″	1·210″	1·110″	Running time ← (seconds)

Table 2

Test Function 2 $F(t) = U(t-10) = \mathscr{L}^{-1}\{s^{-1}\exp(-10s)\}$

t	Method1	Method2	FLIT	Exact $F(t)$
0·0	0·15367E − 01	0·67834E − 02	0·67836E − 02	0·0
1·0	0·17966E − 01	0·67859E − 02	0·67842E − 02	0·0
2·0	0·25222E − 01	0·67897E − 02	0·67859E − 02	0·0
3·0	0·37186E − 01	0·67967E − 02	0·67905E − 02	0·0
4·0	0·56909E − 01	0·68076E − 02	0·67965E − 02	0·0
5·0	0·89423E − 01	0·68250E − 02	0·68244E − 02	0·0
6·0	0·14304E + 00	0·68566E − 02	0·68248E − 02	0·0
7·0	0·23142E + 00	0·69171E − 02	0·68520E − 02	0·0
8·0	0·37715E + 00	0·70523E − 02	0·69223E + 02	0·0
9·0	0·61742E + 00	0·74942E − 02	0·71546E − 02	0·0
10·0	0·10147E + 01	0·50679E + 00	0·50738E + 00	0·5
11·0	0·10179E + 01	0·10056E + 01	0·10061E + 01	1·0
12·0	0·10252E + 01	0·10060E + 01	0·10063E + 01	1·0
13·0	0·10371E + 01	0·10061E + 01	0·10064E + 01	1·0
14·0	0·10569E + 01	0·10062E + 01	0·10065E + 01	1·0
15·0	0·10568E + 01	0·10062E + 01	0·10065E + 01	1·0
16·0	0·11430E + 01	0·10063E + 01	0·10065E + 01	1·0
17·0	0·12314E + 01	0·10063E + 01	0·10065E + 01	1·0
18·0	0·13771E + 01	0·10064E + 01	0·10065E + 01	1·0
19·0	0·16173E + 01	0·10065E + 01	0·10066E + 01	1·0
	2·020″	1·520″	1·415″	Running time ← (seconds)

As Dubner and Abate did, we are going to use the Fast Fourier Transform (FFT) to speed up the computation and to increase accuracy; but this time, FFT will be more efficiently applied: there will be one real and one imaginary argument entered into the FFT subroutine, instead of one real argument only.

If we require $f(t)$ for N equidistant points $t_j = j\Delta t = jT/N$ $j = 0, 1, 2, \ldots N-1$, (42) can be written:

with

$$s_k = a + ik\frac{2\pi}{T}$$

and

$$F(s_k) = \mathrm{Re}\,\{F(s_k)\} + i\,\mathrm{Im}\,\{F(s_k)\}$$

$$f(t_j) + \mathrm{ERROR3}\,(a, t, T) + E_t + E_n = \frac{2e^{aj\Delta t}}{T}$$

$$\left[-\tfrac{1}{2}\,\mathrm{Re}\,\{F(a)\} + \mathrm{Re}\left\{\sum_{k=0}^{NSUM} F(s_k)\left(\cos kj\frac{2\pi}{N} + i\sin kj\frac{2\pi}{N}\right)\right\}\right] \quad (44)$$

Putting

$$C(j) = e^{aj\Delta t};\ W = \cos\frac{2\pi}{N} + i\sin\frac{2\pi}{N} = \exp\left(i\frac{2\pi}{N}\right)$$

and since $W^{jk} = W^{j(k+lN)}$, $l = 1, 2, 3, \ldots,$ we can group terms like $\mathrm{Re}\,\{F(a + i(k + lN))\}$ and $\mathrm{Im}\,\{F(a + i(k + lN))\}$, and write:

$$f(t_j) + \mathrm{ERROR3}\,(a, t, T) + Er + Et =$$

$$C(j)\left[-\tfrac{1}{2}\,\mathrm{Re}\,\{F(a)\} + \mathrm{Re}\left\{\sum_{k=0}^{N-1}(A(k) + iB(k))\,W^{jk}\right\}\right]$$

$$A(k) = \sum_{l=0}^{L} \mathrm{Re}\left\{F\left(a + i(k + lN)\frac{2\pi}{T}\right)\right\}.$$

$$B(k) = \sum_{l=0}^{L} \mathrm{Im}\left\{F\left(a + i(k + lN)\frac{2\pi}{T}\right)\right\}.$$

Table 3

$f(t) = 2\sum_{k=0}^{\infty}(-1)^k\,U(t-2k) = \mathscr{L}^{-1}\{2/s(1+\exp(-2s))\}$

t	FLIT	Exact $F(t)$
0·0	1·006615	1·0
1·0	2·012790	2·0
2·0	1·006321	1·0
3·0	0·000511	0·0
4·0	1·007322	1·0
5·0	2·012833	2·0
6·0	1·005719	1·0
7·0	0·000549	0·0
8·0	1·008010	1·0
9·0	2·012913	2·0
10·0	1·005268	1·0
11·0	0·000820	0·0
12·0	1·009086	1·0
13·0	2·013551	2·0
14·0	1·005602	1·0
15·0	0·002814	0·0
16·0	1·013075	1·0
17·0	2·019597	2·0
18·0	1·017248	1·0
19·0	0·039362	0·0
	4·047″	Running time ← (seconds)

(Cooley and Tukey, 1965; Gentleman and Sande, 1966; Cooley, Lewis, and Welch, 1967).

To be able to use this formulation, we must take NSUM $= L \times N$, but this is not a limitation. The input arrays for the FFT are $A(k)$ and $B(k)$; the output arrays are

$AX(j)$ and $BX(j)$, with $AX(j) = f(t_j)$.

We tested FLIT again with our two previous test functions; we took $L = 20$ in order to have NSUM = 2000 for METHOD1, METHOD2, and FLIT. Tables 1 and 2 show the improvement from the right to the left. Again $N = 100$; only 20 points are printed.

To be sure of FLIT's efficiency, we tested the difficult case of a function $f(t)$ with an infinite number of discontinuities:

$f(t) = 2 \sum_{k=0}^{\infty} (-1)^k U(t - 2k)$, whose value is 2 for

$$2k \leqslant t \leqslant 2k + 1$$

and 0 for $2k + 1 \leqslant t \leqslant 2(k + 1)$. Here

$$F(s) = 2/s(1 + \exp(-2s)) \ .$$

We ran this test with $T = 20$, $aT = 5$, $L = 50$, NSUM = 5000, $N = 100$. The results are displayed in **Table 3**.

8. Conclusion

We wish to mention here the specific problem which brought us to develop FLIT; it might interest some electrical and electronical engineers.

We had to find the influence of various parameters upon the response of a circuit containing a coaxial cable. The Laplace expression for the voltage across the impedance loading this coaxial cable was:

$$V(s) = \left[\frac{C_o V_o Z_{(o,s)}}{1 + C_o s(L_o s + R_o + Z_{(o,s)})}\right] \times \left[\frac{Z}{Z_c \sinh \gamma l + Z \cosh \gamma l}\right]$$

with:

$$Z_{(o,s)} = Z_c \left(\frac{Z + Z_c \tanh \gamma l}{Z_c + Z \tanh \gamma l}\right)$$

$$\gamma = \sqrt{(R\sqrt{s} + L_s)(G + C_s)} \quad Z_c = \sqrt{\frac{R\sqrt{s} + L_s}{G + C_s}}$$

R, L, G, C: cable constants.
l: length of the cable.
Z: load impedance.
V_0, L_0, R_0, C_0: electrical parameters of the circuit; their influence upon the circuit response is investigated.

Whenever we decided to compare theory and experience, the computed voltage was found to be identical to what was observed on the scope.

A FORTRAN listing is available on request; the author would welcome the submission of any difficult case he has not thought of.

Acknowledgement

I wish to thank Mr. Jacques Chartier for many fruitful discussions about this work.

References

COOLEY, J. W., and TUKEY, J. W. (1965). An algorithm for the machine calculation of complex Fourier series, *Math. Comp.*, Vol. 19, pp. 297-301.

COOLEY, J. W., LEWIS, P. A. W., and WELCH, P. D. (1967). Application of the Fast Fourier Transform to the computation of Fourier integrals, Fourier series, and convolution integrals. *IEEE Transactions*, Vol. AE-15, pp. 79-85.

DE BALBINE, G., and FRANKLIN, J. (1966). The calculation of Fourier integrals, *Maths. Comp.*, Vol. 20, pp. 570-589.

DUBNER, R., and ABATE, J. (1968). Numerical inversion of Laplace transforms by relating them to the finite Fourier Cosine transform, *JACM*, Vol. 15, No. 1, pp. 115-123.

GENTLEMAN, W. M., and SANDE, G. (1966). Fast Fourier Transforms. *Procs. AFIPS. Joint Computer Conference*, Vol. 29, pp. 563-578.

GRADSHTEYN, I. S., and RIZHIK, I. M. (1965). *Table of Integrals, Series and Products*, New York and London, Academic Press.

SHANKS, D. (1955). Non-linear transformations of divergent and slowly convergent series, *J. Math. and Phys.*, Vol. 34, pp. 1-42.

Book review

Functional Analysis of Information Processing, by Grayce M. Booth, 1974; 269 pages. (*John Wiley*, £7·70.)

This book has to be judged in the context of the claims made for it in the preface and introductory chapter.

Its aim is to provide an aid to the information systems analyst, designer, or programmer in the analysis of complex computer systems. For this purpose a new approach is put forward—the approach of the structured, functional analysis of information processing. The functions referred to are all related to the processing machine, i.e. the computer, hardware and software. The approach is 'really a method of logically structuring the systems analysis and design process. It will also furnish (the designer) with a complete set of hardware and software functions which he can evaluate when designing an information processing system.'

In practice the author offers a six level scheme of hierarchically classifying a computer system, ranging from level I—the network level (two components: information processing, and network processing) to level VI—the level of device techniques.

Like most classification schemes it is often arbitrary and sometimes idiosyncratic. For example, the category 'simulation' (level VI) appearing in the level V category of 'other languages' puts simulation of one computer on another in the same class as simulation languages, and it is the only place in which emulation is described.

The rigid structure imposed by the six level classification system prohibits analysis where more than six levels may be appropriate. Thus the title operating systems much used in the text cannot be found a place in the classification, all operating system functions being separately defined under the level III—classification, 'software functions'.

More seriously, many functions important to the designer are not classified or may be missing altogether. No reference is made to different methods of file access organisation, such as index sequential, random algorithmic, lists, or inverted files. The level V entry—printers has no lower level components although a designer could well be concerned with further entries such as line printers, character printers, impact printers, non-impact printers and sub-classes of these.

The analysis provides descriptions of class components in various levels of detail but little in the way of quantitive information which could help the systems designer. Hence, it fails in its major objective.

It is to some extent redeemed by the clarity of the writing independent of the system of classification. Some of the descriptive pieces, as for example those relating to data management, and its component data description language and data manipulation language, are well written but not detailed enough for anything except a first appraisal. The main use of the book may be as a check-list of systems components for information systems designers.

F. F. LAND (London)