The author acknowledges with thanks the aid of DOLORES UFFORD, who assisted in the calculations.

YUDELL L. LUKE

Midwest Research Institute
Kansas City 2, Missouri

[1] W. E. MILNE, "The remainder in linear methods of approximation," NBS, *Jn. of Research*, v. 43, 1949, p. 501–511.

[2] W. E. MILNE, *Numerical Calculus*. p. 108–116.

[3] M. BATES, *On the Development of Some New Formulas for Numerical Integration*. Stanford University, June, 1929.

[4] M. E. YOUNGBERG, *Formulas for Mechanical Quadrature of Irrational Functions*. Oregon State College, June, 1937. (The author is indebted to the referee for references 3 and 4.)

[5] E. L. KAPLAN, "Numerical integration near a singularity," *Jn. Math. Phys.*, v. 26, April, 1952, p. 1–28.

On the Numerical Solution of Equations Involving Differential Operators with Constant Coefficients

1. The General Linear Differential Operator. Consider the differential equation of order n

$$Ly + F(y, x) = 0, \tag{1}$$

where the operator L is defined by

$$Ly = \sum_{k=0}^{n} P_k(x) \frac{d^k y}{dx^k},$$

and the functions $P_k(x)$ and $F(y, x)$ are such that a solution y and its first n derivatives exist in $0 \leq x \leq X$. In the special case when (1) is linear the solution can be completely determined by the well known method of variation of parameters when n independent solutions of the associated homogeneous equations are known. Thus for the case when $F(y, x)$ is independent of y, the solution of the non-homogeneous equation can be obtained by mere quadratures, rather than by laborious stepwise integrations. It does not seem to have been observed, however, that even when $F(y, x)$ involves the dependent variable y, the numerical integrations can be so arranged that the contributions to the integral from the upper limit at each step of the integration, at the time when y is still unknown at the upper limit, drop out. Thus again the computation can be made to involve merely quadratures.

It is not often that the solution of the homogeneous equation can be simply determined, and it is perhaps for this reason that attention has not been given heretofore to the possibility of simplifying the numerical evaluation of the solution by making use of the solutions to the homogeneous equation. However, in the case when the functions $P_k(x)$ in L are constants, the solution of the homogeneous equation is easy to determine. This is particularly true when the order of the differential equation is fairly low. In the instance when the operator L is of second order, with constant coefficients, the method of using the integral equation often has decided advantages over the usual methods employed for solving differential equations.

For this reason attention will now be centered on a second order operator with constant coefficients.

2. Linear Differential Operators of Second Order with Constant Coefficients. Let Ly now be specialized as follows:

$$Ly = \frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy, \tag{2}$$

where b, c are real constants. Let $y(0) = \rho_0$, and $y'(0) = \rho_1$ be assigned. The differential equation (1) can be replaced by the integral equation

$$y = \alpha_1 e^{m_1 x} + \alpha_2 e^{m_2 x} - G_1(x) + G_2(x), \tag{3}$$

where

$$\alpha_1 + \alpha_2 = \rho_0; \quad \alpha_1 m_1 + \alpha_2 m_2 = \rho_1,$$

$$G_k(x) = \left[e^{m_k x}\int_0^x e^{-m_k t}F(y, t)dt\right]\Big/ (b^2 - 4c)^{\frac{1}{2}}; \quad k = 1, 2,$$

and m_1, m_2 are the roots of $m^2 + bm + c = 0$, provided $b^2 - 4c \neq 0$. In the special case when $b^2 - 4c = 0$, (3) becomes

$$y = e^{-\frac{1}{2}bx}\left[\alpha_1 + \alpha_2 x + \int_0^x (t - x)e^{\frac{1}{2}bt}F(y, t)dt\right], \tag{3a}$$

$$\alpha_1 = \rho_0, \quad \alpha_2 = \rho_1 + \tfrac{1}{2}b\rho_0.$$

It should be observed that when $b^2 - 4c$ is negative, m_1 and m_2 are conjugate complex numbers. When b, c, and the initial values are real, the imaginary component of (3) will drop out, and the discussion which is to follow will apply to this case and to (3a) as well. For the sake of simplicity, therefore, we shall now assume that m_1 and m_2 are real and distinct. Let

$$x_r = x_0 + rh, \quad x = x_0 + sh, \quad t = x_0 + ph, \quad y_r = y(x_r).$$

With the above, $G_k(x)$ takes the form

$$G_k(x) = e^{m_k(x - x_0)}G_k(x_0) + C_k(x), \tag{4}$$

where

$$C_k(x) = e^{m_k sh}h(b^2 - 4c)^{-\frac{1}{2}}\int_0^s e^{-m_k ph}F(y, x_0 + ph)dp. \tag{5}$$

If the integrand of (5) is approximated by an $(s + 1)$-point polynomial, then

$$\int_0^s e^{-m_k ph}Fdp = \sum_{r=0}^{s} a_r e^{-m_k rh}F(y_r, x_r) + R_k, \tag{6}$$

where the coefficients a_r result from the integration of the polynomial and the truncation term, R_k, can be represented by

$$R_k = \int_0^s \phi(t)[t, x_0, x_1, \cdots, x_s]dt. \tag{7}$$

In (7) $[t, x_0, x_1, \cdots, x_s]$ is the divided difference[1] of order $(s + 1)$ of the function $e^{-mt}F(y, t)$ and $\phi(t)$ is the polynomial approximation of $e^{-mt}F(y, t)$.

Thus

$$(8)\quad -G_1(x)+G_2(x) = -e^{m_1(x-x_0)}G_1(x_0)+e^{m_2(x-x_0)}G_2(x_0) + h(b^2-4c)^{-\frac{1}{2}}\sum_{r=0}^{s-1} aF(y_r,x_r)\{e^{m_2h(s-r)}-e^{m_1h(s-r)}\}-R_1+R_2.$$

Note that the term involving $F(y_s, x_s)$ dropped out, since the coefficient of this term appeared with the same sign in G_1 and G_2. Although this fact is well known in the theory of integral equations, its importance from the viewpoint of the numerical evaluation of the solution needs emphasis. Thus the evaluation of $-G_1+G_2$ *does not depend on the value of the function at the end of the interval.* We therefore do not need a "predictor" formula (using MILNE's terminology[2]). The steps of integration at any stage can therefore be carried out as follows:

1) Evaluate $-G_1(x)+G_2(x)$ by (8).
2) Compute $\alpha_1 \exp(m_1x)+\alpha_2\exp(m_2x)$; hence knowing $(-G_1+G_2)$, we now know y_s, from (3).
3) Knowing y_s, we can use (6) to evaluate the integral $G_1(x)$ by a mere quadrature. $G_2(x)$ is now also known, since $G_1(x)$ and $-G_1(x)+G_2(x)$ are known.

3. The Truncation Term. If m_1 and m_2 are negative, $e^{-m_1t}F$ may require a higher order approximation formula than $F(t)$ itself (although by no means necessarily so). In any case, h must be small enough so that R_k in (7) is indeed negligible for the accuracy aimed at. In some cases it is actually possible to take a *larger* step h when the integral equation is used, than the one that can be taken when the differential form (2) is operated with. Moreover, the process of solving the integral equation may be more stable than the corresponding solution of the differential equation by stepwise integration. This is especially true when $(b^2-4c)^{\frac{1}{2}}$ is large numerically. The example following illustrates the case.

4. Example. Consider the differential equation

$$(9)\qquad u''+vu'+e^{-x}+f(u)=0,$$

where

$$f(u)=\exp\left(A-\frac{B}{u+d}\right).$$

This differential equation occurred in connection with certain steady state solutions, and among the various parameters which were used, one set was the following:

$$A=74.997736;\quad B=257.42325;\quad d=2.19885;\quad u(0)=1.294;$$
$$u'(0)=-M=-10;\quad v=24.38.$$

Suppose we attempted the evaluation of (9) by the usual method of stepwise integration, first computing an integral of u'', using (9), then integrating u' to obtain u. An analysis of the manner in which an error in u' is propagated shows that the interval h would have to be taken no larger than $1/(5v)$ over the entire range of the integration. Thus for $v=25$, an interval as small as 0.008 would be needed.

A study of the behavior of $f(u)$, which depends not only on v but also on M and $u(0)$, shows that for very small values of x, it is sometimes necessary to take h even smaller than $1/(5v)$. This is true whether we solve the equation by integrating (9) or by using (3). However, when $M = 10$, for example, and the integral equation is used, it is possible to use a step h as large as 0.1, for x larger than 0.4, and yet we would have to maintain an interval of about 0.008 if the differential equation were used. It can be verified that it is possible to lose all significance in u' within a relatively short range of x when (9) is used, unless the interval h is kept small enough. When (3) is used, on the other hand, the solution remains very stable, and the size of the interval can be increased just as quickly as the behavior of $f(u)$ permits. It turns out that $e^x f(u)$ behaves no worse than $f(u)$ itself. The successive derivatives of $f(u)$ with respect to x are numerically very large near the origin, but they go down in magnitude to reasonable levels at $x = 0.4$, where $M = 10$. For this problem, therefore, the form (3) is far superior to the usual method of solving the differential equation.

The permissible magnitude of h near the origin, when (3) is employed, can be determined by usual methods, and the necessary starting values can be computed in a simple manner; hence it is not deemed worth while to give the specific details of the solution. We shall, however, re-write (3) to apply specifically to (9). Here

$$F = f(u) + e^{-x}; \quad m_1 = 0; \quad m_2 = -v; \quad \rho_0 = u(0); \quad \rho_1 = -M.$$

When the above are substituted into (3), the equation becomes

$$u = U_1 - \frac{1}{v}\int_0^x f\{u(t)\}\,dt + \frac{1}{v}e^{-vx}\int_0^x e^{vt} f\{u(t)\}\,dt,$$

where

$$U_1 = \frac{e^{-x}}{v-1} + \frac{e^{-vx}}{v}\left[M - \frac{1}{v-1}\right] + u_0 - \frac{M+1}{v}, \quad v(v-1) \neq 0.$$

Thus let

$$G(x) = \frac{1}{v}\int_0^x f\,dt; \quad S(x) = \frac{1}{v}e^{-vx}\int_0^x e^{vt} f\,dt - \frac{1}{v}\int_0^x f\,dt,$$

$$H(x) = \frac{1}{v}e^{-vx}\int_0^x e^{vt} f\,dt = S(x) + G(x).$$

Using the notation

$$\phi_s = \phi(sh)$$

for any function ϕ, we have, by the previous analysis,

$$u_{s+1} = U_{s+1} + \frac{1}{v}e^{-2vh}[S_{s-1} + G_{s-1}] - \frac{1}{v}G_{s-1} + \phi_{s+1},$$

where

$$\phi_{s+1} = \frac{1}{v}e^{-v(s+1)}\int_{(v-1)h}^{(v+1)h} e^{vt} f\,dt - \frac{1}{v}\int_{(v-1)h}^{(v+1)h} f\,dt.$$

If ϕ_{s+1} is evaluated by Simpson's rule, we have

$$\phi_{s+1} = \frac{h}{3v}[4(e^{-vh} - 1)f_n + (e^{-2vh} - 1)f_{s-1}] + R + \delta_{s+1},$$

where δ_{s+1} is the rounding error. The total error at x can be approximated, roughly, by

$$\epsilon_{s+1} = -\frac{1}{v}\int_0^{x_{s+1}} (\delta + R)\frac{\partial f}{\partial u}\,dt + \frac{1}{v}e^{-vx_{s+1}}\int_0^{x_{s+1}} e^{vt}(\delta + R)\frac{\partial f}{\partial u}\,dt.$$

Differences can be used in the usual manner to approximate the magnitude of derivatives. The recommendations for the special example are:

a) Compute ϕ_{s+1} by Simpson's rule or some other suitable integration formula.
b) Evaluate u_{s+1}.
c) Evaluate G_{s+1} by quadrature (since u_{n+1} is now known to the required accuracy).
d) Write $S_{s+1} = u_{s+1} - U_{s+1}$.

Return to (a) for the evaluation of ϕ_{s+2} in the succeeding step.

The process can be readily coded for high-speed machines.

One may inquire whether (3) is always the better form to use, compared with (1), or whether there are ranges of the parameters where one of the forms is better than the other. An examination of the way in which the constants enter into the solution shows that when $b^2 - 4c$ is large form (3) would always have advantages over (1), since the square root of this quantity enters into the denominator of some of the terms. In the case when this quantity is small, however, it is likely that (1) would be the better form to operate with.

5. The General Case. It may be worth while to remark that the method applies to the general case mentioned in Section I, where the coefficients $P_k(x)$ are functions of x. The solution can be *constructed formally* by using the usual method of variation of parameters for linear equations.[3] The complete solution will be of the form

$$y = \sum_{k=1}^{n} u_k V_k,$$

where the n functions u_k are the known solutions of the homogeneous equation, and the n integrals V_k must be generated numerically by using the Wronskian of the solutions u_k. In the numerical process, the value of y at some point x is obtained from the property that the contribution to $y(x)$ from the upper limit of integration drops out. After $y(x)$ is known, each individual $V_k(x)$, which is required for carrying forward the numerical process, is then evaluated by a mere quadrature.

Gertrude Blanch

National Bureau of Standards
Los Angeles 24

The preparation of this paper was sponsored (in part) by the Office of Naval Research, USN.

[1] See L. M. Milne-Thomson, *The Calculus of Finite Differences*, London, 1933.
[2] W. E. Milne, *Numerical Calculus*, Princeton, 1949.
[3] E. L. Ince, *Ordinary Differential Equations*, London, 1927, section 5.23.