PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Neville's and Romberg's Processes: A Fresh Appraisal with Extensions

J. C. P. Miller

Phil. Trans. R. Soc. Lond. A 1968 **263**, 525-562
doi: 10.1098/rsta.1968.0032

NEVILLE'S AND ROMBERG'S PROCESSES: A FRESH APPRAISAL WITH EXTENSIONS

By J. C. P. Miller

University Mathematical Laboratory, Cambridge

(*Communicated by M. V. Wilkes, F.R.S.—Received* 9 *October* 1967)

CONTENTS

In this paper Neville's process for the repetitive linear combination of numerical estimates is re-examined and exhibited as a process for term-by-term elimination of error, expressed as a power series; this point of view immediately suggests a wide range of applications—other than interpolation, for which the process was originally developed, and which is barely mentioned in this paper—for example, to the evaluation of finite or infinite integrals in one or more variables, to the evaluation of sums, etc. A matrix formulation is also developed, suggesting further extensions, for example, to the evaluation of limits, derivatives, sums of series with alternating signs, and so on.

It is seen also that Neville's process may be readily applied in Romberg Integration; each suggests extensions of the other.

Several numerical examples exhibit various applications, and are accompanied by comments on the behaviour of truncation and rounding errors as exhibited in each Neville tableau, to show how these provide evidence of progress in the improvement of the approximation, and internal numerical evidence of the nature of the truncation error.

A fuller and more connected account of the behaviour of truncation errors and rounding errors is given in a later section, and suggestions are also made for choosing suitable specific original estimates, i.e. for choosing suitable tabular arguments in the elimination variable, in order to produce results as precise and accurate as possible.

1. Introduction

The idea of interpolation by repetitive* linear interpolation of estimates seems to have originated with A. C. Aitken (1932, 1938), with effective further development by Neville (1934). The resulting processes are highly suitable for automatic computation, but have been rather slow in gaining ground, partly because of the relative obscurity of the journals in which these early papers appeared, and partly because they seem, at first sight, to be restricted to interpolation, now a relatively less widespread requirement than in the past.

* The inappropriate adjective 'iterative' is commonly used; it is, however, more useful to reserve 'iterative' for *self-correcting* repetitive processes—Neville's process is *not* one of these.

In this paper Neville's process is re-examined and exhibited mainly as a process for the successive elimination of terms in an error expressed as a power series; this immediately suggests a wide range of applications to the evaluation of integrals, sums, etc. A matrix formulation is also developed, suggesting further extensions, for example, to the evaluation of limits, derivatives, and sums of series with alternating signs.

Applications to interpolation are barely mentioned in this paper; the accounts of Aitken and Neville should be consulted for details and numerical examples of such applications, and for Aitken's extension to quadratic extrapolation and for Neville's extension allowing use of derivatives. It may, however, be useful to mention that the methods of § 11, involving the determination and use of Lagrangian multipliers, extend readily to incorporate these developments, and also the osculatory interpolation formulae of Hermite (1878) and Fort (1948).

It is also seen that Neville's process may be applied to Romberg integration, and that each suggests extensions of the other. In particular, it is shown that error terms may be eliminated from estimates for integrals with singularities of the integrand at the end-points of the range, after determining the form of these terms as in Fox (1967).

A number of applications on the lines of the more direct parts of this paper can be found in recent literature. For example, occasional applications that may be regarded as cases of the 'deferred approach to the limit' of Richardson & Gaunt (1927) are foreshadowed in Henrici (1964), or Gragg (1965), to give but two examples. However, this paper is designed largely as an expository account of a number of extensions of Neville's and Romberg's processes, mainly new, and no attempt has been made to search the literature exhaustively.

Several numerical examples are included, exhibiting a variety of applications. In order to separate the effects of convergence (or the successive elimination of terms in the truncation error series) and of rounding errors (which accumulate and grow, in general), it has seemed desirable to retain a substantial number of figures in the calculations. This is not, of course, always necessary or even possible in practice, but serves to avoid confusion due to overlap of effects. Comments are added in nearly every case on truncation and rounding-error effects exhibited in the various Neville tableaux.

A more connected discussion of error effects is given later, in § 12, making some use of Lagrange type multipliers, developed in § 11, and useful also in obtaining isolated results in the Neville tableau. It is shown how such discussion of error may help in estimating the precision of results obtained, by study of a tableau alone, and may also help, more generally, in the choice of suitable arguments for the individual estimates for satisfactory reduction by Neville's process (§ 13).

2. Description of the process

2·1. *The basic step*

Neville's process for interpolation is a repetitive procedure of a type very convenient for inclusion in programs for automatic computers. A brief description of this process follows.

The basic step is implied in the inductive definition

$$f(x;\, x_0, x_1, \ldots, x_n) = \begin{vmatrix} \lambda(x-x_0) & f(x;\, x_0, x_1, \ldots, x_{n-1}) \\ \lambda(x-x_n) & f(x;\, x_1, x_2, \ldots, x_n) \end{vmatrix} \div \lambda(x_n - x_0) \qquad (2\cdot1)$$

with $f(x; x_r) = f(x_r)$, where $f(x)$ is a function of x tabulated at the points $x = x_r$, $(r = 0, 1, 2, \ldots, n)$. This is known as a *linear cross-mean*; the quantities in the first column of the matrix are known as *parts*, the divisor is the difference between the parts. The scaling factor λ should be chosen to suit convenience; it is here assumed that $\lambda = 1$ in theoretical discussion.

Clearly $f(x; x_0, x_1, \ldots, x_n)$ is a polynomial in x of degree not greater than n. Also $f(x_s; x_0, x_1, \ldots, x_n) = f(x_s)$ for $s = 0(1)n$. These facts follow by induction from $f(x; x_r) = f(x_r)$, a constant, $r = 0(1)n$.

Thus $f(x; x_0, x_1, \ldots, x_n)$ is the unique polynomial of degree not exceeding n through the $n+1$ points $(x_r, f(x_r))$, $r = 0(1)n$; that is, it is identical in all but form with the Lagrange polynomial, or with Newton's divided-difference polynomial, through these points. We may therefore quote the most usual form of error term in representing $f(x)$ by this polynomial $P_n(x)$, namely

$$\left.\begin{aligned} f(x) &= P_n(x) + R_n(x), \\ R_n(x) &= (x-x_0)\,(x-x_1)\ldots(x-x_n)\,\frac{f^{(n+1)}(\eta)}{(n+1)\,!} \end{aligned}\right\} \qquad (2\cdot2)$$

where

with η a point in the least closed interval including $x, x_0, x_1, \ldots, x_n$, provided that $f(x)$ is suitably well behaved in this interval. We note also a consequence of the uniqueness of $P_n(x)$; it means that the order in which $x_0, x_1, \ldots, x_n$ are used in the development by linear cross-means is immaterial, i.e. that

$$f(x; x_0, x_1, \ldots, x_n) = f(x; x_0', x_1', \ldots, x_n')$$

in which $\{x_0', x_1', \ldots, x_n'\}$ is any re-arrangement of the set $\{x_0, x_1, \ldots, x_n\}$.

It is sometimes useful to distinguish between two ways of using the process. One way is as a process for developing interpolating polynomials $P_n(x)$ as *functions* approximating a given function $f(x)$. The other way is to view it as a procedure for obtaining *numerical* approximations to a particular value $f(X)$, for a specified $x = X$. We use this convention of distinction by the argument x or X below, particularly as it is helpful to transfer the origin by writing $x = X + \xi$ and considering interpolation to $\xi = 0$.

There are two ways of developing the numerical interpolation that are most commonly used: Aitken's process and Neville's process.

2·2. *Aitken's process*

This method of using the basic step in computing $f(X)$ is illustrated in table 2·2. The quantities computed and the layout of results for desk computations are shown. In each column a *pivotal element* is chosen, and combined with each of the other elements in that column in turn. The elements of the 2×2 matrix of the basic step, see (2·1), are in each case at the corners of a rectangle.

TABLE 2·2

x	parts	$f(x)$		
x_0	$X-x_0$	$\underline{f(x_0)}$		
x_1	$X-x_1$	$f(x_1)$	$\underline{f(X; x_0, x_1)}$	
x_2	$X-x_2$	$f(x_2)$	$f(X; x_0, x_2)$	$\underline{f(X; x_0, x_1, x_2)}$
x_3	$X-x_3$	$f(x_3)$	$f(X; x_0, x_3)$	$f(X; x_0, x_1, x_3)$
x_4	$X-x_4$	$f(x_4)$	$f(X; x_0, x_4)$	$f(X; x_0, x_1, x_4)$

New values $f(x_r)$ may be appended as needed, if convergence to the desired interpolate is insufficient, but only at the bottom of the table; order is immaterial, so that $x_0, x_1, x_2, \ldots$ may be chosen in numerical order, or in order of distance from X, or in any other order.

2·3. *Neville's process*

In this method, only consecutive elements in any column are combined as in table 2·3. Results are recorded on a level mid-way between lines in the previous column, and the parts in the first column of the 2×2 matrix of the basic step (2·1) are found at the end of lines sloping diagonally backwards from the two elements used in the second column, up for the upper element, and down for the lower element. New values may be added at either end of the table, thus making it easier to preserve numerical order of arguments and regular progression of results, and yet append new arguments on both sides of x. This regular progression helps with checking.

TABLE 2·3

x	parts	$f(x)$		
x_0	$X-x_0$	$f(x_0)$		
			$f(X;\, x_0, x_1)$	
x_1	$X-x_1$	$f(x_1)$		$f(X;\, x_0, x_1, x_2)$
			$f(X;\, x_1, x_2)$	
x_2	$X-x_2$	$f(x_2)$		$f(X;\, x_1, x_2, x_3)$
			$f(X;\, x_2, x_3)$	
x_3	$X-x_3$	$f(x_3)$		$f(X;\, x_2, x_3, x_4)$
			$f(X;\, x_3, x_4)$	
x_4	$X-x_4$	$f(x_4)$		

2·4.

Aitken's process has a very effective extension involving quadratic parts, which is less easily applied with Neville's process. On the other hand, Neville's process may be easily adapted to use derivatives; these are much more difficult to use with Aitken's process, though it is of interest to note that Aitken (1932, p. 75) did suggest the use of one such derivative with a quadratic process. We are not concerned further with these particular possibilities in this paper.

In the rest of this paper we shall be concerned almost entirely with 'extrapolation to zero', in particular, with tables having parts of one sign only. In fact, the signs have no particular effect on the calculation, but it is worthwhile to remark that any use of Aitken's or Neville's process can be regarded as 'extrapolation or interpolation to zero', if the *column of parts* is regarded as the primary variable; the examples that follow, however, usually have a variable which vanishes naturally at the argument for which we desire to know the function value—this should not be allowed to obscure the fact that the processes are equally valid in the more general case.

3. NEVILLE'S PROCESS AS AN ELIMINATION PROCESS

New light is thrown on Neville's process, and extensions of its range of applications become clear, if we express it as a form of elimination.

3·1.

Suppose, as before, that $f(x)$ is known at $x = x_r$, $r = 0(1)n$, and that $f(X)$ is required. If we expand as a Taylor series at $x = X$, we obtain, writing $\xi = x - X$,

$$\left.\begin{aligned} f(x) &= f(X) + (x-X)\,f'(X) + (x-X)^2 f''(X)/2! + \ldots + (x-X)^s f^{(s)}(X)/s! + \ldots \\ &= f(X) + \xi f'(X) + \xi^2 f''(X)/2! + \ldots + \xi^s f^{(s)}(X)/s! + \ldots. \end{aligned}\right\} \quad (3{\cdot}11)$$

Then
$$\left.\begin{aligned} f(X;x_0) &= f(x_0) = f+\xi_0 f'+\xi_0^2 f''/2!+\xi_0^3 f'''/3!+\ldots+\xi_0^s f^{(s)}/s!+\ldots \\ f(X;x_1) &= f(x_1) = f+\xi_1 f'+\xi_1^2 f''/2!+\xi_1^3 f'''/3!+\ldots+\xi_1^s f^{(s)}/s!+\ldots \end{aligned}\right\} \tag{3·12}$$
where $\xi_r = x_r - X$ and the arguments (X) are omitted.

The linear cross-mean is then
$$f(X;x_0,x_1) = f+0-\xi_0\xi_1 f''/2!-\xi_0\xi_1(\xi_0+\xi_1)f'''/3!-\ldots-\xi_0\xi_1 H_{s-2}(\xi_0,\xi_1)f^{(s)}/s!\ldots \tag{3·13}$$
in which $H_p(\xi_0,\xi_1)$ is the sum of the homogeneous products of degree p in the two variables ξ_0,ξ_1, i.e.
$$H_p(\xi_0,\xi_1) = \sum_{s=0}^{p}\xi_0^s\xi_1^{p-s}.$$

We note that the term in $f'(X)$ has been eliminated.

3·2.

If we now take the linear cross-mean of (3·13) and of
$$f(X;x_1,x_2) = f+0-\xi_1\xi_2 f''/2!-\xi_1\xi_2(\xi_1+\xi_2)f'''/3!-\ldots-\xi_1\xi_2 H_{s-2}(\xi_1,\xi_2)f^{(s)}/s!-\ldots \tag{3·21}$$
we obtain similarly
$$f(X;x_0,x_1,x_2) = f+0+0+\xi_0\xi_1\xi_2 f'''/3!+\ldots+\xi_0\xi_1\xi_2 H_{s-3}(\xi_0,\xi_1,\xi_2)f^{(s)}/s!+\ldots \tag{3·22}$$
where $H_p(\xi_0,\xi_1,\xi_2)$ is the homogeneous product sum of degree p in the three variables ξ_0,ξ_1,ξ_2. We extend this notation similarly to any number of variables.

This time the term in $f''(X)$ has been eliminated.

3·3.

We have used above, for $r = 2, 3$, the relation
$$\left.\begin{aligned} &\begin{vmatrix} \xi_0 & H_p(\xi_0,\xi_1,\ldots,\xi_{r-1})\times\xi_0\xi_1\ldots\xi_{r-1} \\ \xi_r & H_p(\xi_1,\xi_2,\ldots,\xi_r)\times\xi_1\xi_2\ldots\xi_r \end{vmatrix} \div(\xi_0-\xi_r) \\ &\qquad = \xi_0\xi_1\ldots\xi_r \begin{vmatrix} 1 & H_p(\xi_0,\xi_1,\ldots,\xi_{r-1}) \\ 1 & H_p(\xi_1,\xi_2,\ldots,\xi_r) \end{vmatrix} \div(\xi_0-\xi_r) \\ &\qquad = -\xi_0\xi_1\ldots\xi_r H_{p-1}(\xi_0,\xi_1,\ldots,\xi_r). \end{aligned}\right\} \tag{3·31}$$
The last line follows because
$$\left.\begin{aligned} &\frac{1}{\xi_0-\xi_r}\{H_p(\xi_r,\xi_1,\xi_2,\ldots,\xi_{r-1})-H_p(\xi_0,\xi_1,\xi_2,\ldots,\xi_{r-1})\} \\ &\qquad = \frac{1}{\xi_0-\xi_r}\sum_{s=0}^{p}(\xi_r^s-\xi_0^s)\,H_{p-s}(\xi_1,\xi_2,\ldots,\xi_{r-1}) \\ &\qquad = -\sum_{s=0}^{p}H_{s-1}(\xi_0,\xi_r)\,H_{p-s}(\xi_1,\xi_2,\ldots,\xi_{r-1}) \\ &\qquad = -H_{p-1}(\xi_0,\xi_1,\ldots,\xi_r). \end{aligned}\right\} \tag{3·32}$$

3·4.

This elimination process may clearly be continued to yield generally
$$\begin{aligned} &f(X;x_0,x_1,\ldots,x_n) \\ &= f(X)+(-1)^n\{\xi_0\xi_1\ldots\xi_n f^{(n+1)}(X)/(n+1)!+\xi_0\xi_1\ldots\xi_n(\xi_0+\xi_1+\ldots+\xi_n)f^{(n+2)}(X)/(n+2)! \\ &\qquad +\ldots+\xi_0\xi_1\ldots\xi_n H_s(\xi_0,\xi_1,\ldots,\xi_n)f^{(n+s+1)}(X)/(n+s+1)!+\ldots\}. \end{aligned} \tag{3·41}$$
This gives a *remainder series* for the Lagrange polynomial through the $n+1$ points $(x_r, f(x_r))$, $r = 0(1)n$. This may be compared with the equivalent and more usual remainder term (2·2).

3·5. *An alternative development of the elimination*

It is sometimes convenient to start with a Taylor expansion about an origin that is not at $x = X$, where $f(X)$ is the desired interpolate. We choose, without loss of generality, the origin $x = 0$ for this purpose. Then

$$\left.\begin{aligned} f(x) &= f_0+xf'_0+x^2f''_0/2!+\ldots+x^rf_0^{(r)}/r!+\ldots \\ \text{and}\quad f(X) &= f_0+Xf'_0+X^2f''_0/2!+\ldots+X^rf_0^{(r)}/r!+\ldots,\end{aligned}\right\} \tag{3·51}$$

$$\text{so that}\quad \left.\begin{aligned} f(x) &= f(X)+(x-X)f'_0+(x^2-X^2)f''_0/2!+\ldots+(x^r-X^r)f_0^{(r)}/r!+\ldots \\ &= f(X)+\xi f'_0+\xi H_1(x,X)f''_0/2!+\ldots+\xi H_{r-1}(x,X)f_0^{(r)}/r!+\ldots\end{aligned}\right\} \tag{3·52}$$

with $\xi = x-X$ and H_p defined as in § 3·2.

We now take the linear cross-means of $f(X;\,x_0) = f(x_0), f(X,x_1) = f(x_1)$, expressed in the form (3·52). This yields

$$f(X;\,x_0,x_1) = f(X)+0-\xi_0\xi_1f''_0/2!-\ldots-\xi_0\xi_1H_{r-2}(x_0,x_1,X)f_0^{(r)}/r!-\ldots \tag{3·53}$$

$$\text{since}\quad \left.\begin{aligned} \begin{vmatrix}\xi_0 & \xi_0H_p(x_0,X)\\ \xi_1 & \xi_1H_p(x_1,X)\end{vmatrix}\div(\xi_0-\xi_1) &= \xi_0\xi_1\begin{vmatrix}1 & H_p(x_0,X)\\ 1 & H_p(x_1,X)\end{vmatrix}\div(x_0-x_1) \\ &= -\xi_0\xi_1H_{p-1}(x_0,x_1,X).\end{aligned}\right\} \tag{3·54}$$

The elimination proceeds as before, with the difference that the homogeneous products have an extra argument X, and the derivatives $f^{(r)}(0)$ are at a fixed point independent of X, giving the remainder series in a different form.

3·6.

As an elimination procedure, Neville's process is immediately seen to be relevant (as has been remarked by Henrici and others) to processes such as Richardson's 'deferred approach to the limit' or 'h^2-extrapolation', or to the summation of series, by elimination of an appropriate series of remainder terms. In each case there is a remainder term which may usually be expressed as a continuous function, for example, of the interval in a formula for quadrature using equally spaced intervals, or of $1/n$, where n is the number of terms in the partial sums of a series. In the following paragraphs we consider several examples of such applications in turn, designed to bring out various points that may arise.

4. Applications to quadrature and summation

The first examples are concerned with the evaluation of an integral $\int_a^b f(x)\,\mathrm{d}x$. We use throughout the *trapezoidal rule*—the simplest possible formula—after dividing the (finite) range into intervals $h = (b-a)/n$. The Euler–Maclaurin expansion indicates that the trapezoidal sum has the form given by

$$T_h = h(\tfrac{1}{2}f_0+f_1+\ldots+f_{n-1}+\tfrac{1}{2}f_n) = \int_a^b f(x)\,\mathrm{d}x+Ah^2+Bh^4+\ldots+Kh^{2n}+R_n \tag{4·1}$$

where $f_r = f(x_r)$, $x_r = a+rh$, and R_n is a remainder term of order h^{2n+2} in general, but which often includes another part decreasing, for example, like $e^{-\lambda/h}$, and usually quite negligible in practice except when h is relatively large (which, however, we like it to be), see § 12 for further discussion. Thus we may use Neville's process with elimination variable h^2, and extrapolate for $h^2 = 0$, with caution when a result for large h is used.

In § 4·5 we illustrate an application to summation, closely related to quadrature.

4·1. *Quadrature* (i)

Table 4·1 illustrates the applications of the process, by extrapolation to $h^2 = 0$, i.e. as $n \to \infty$, to

$$I = \int_0^1 \frac{dx}{1+x^2} = \tfrac{1}{4}\pi \doteqdot 0{\cdot}78539\,81634\,0$$

using estimates

$$T_h = h(\tfrac{1}{2}f_0 + f_1 + f_2 + \ldots + f_{n-1} + \tfrac{1}{2}f_n) \quad h = 1/n \quad f_r = 1/(n + r^2h^2).$$

Table 4·1

h^2	T_h						
			0·785...	0·7853...	0·78539...		
1	0·75						
		0·78333 33333					
$\frac{1}{4}$	0·775		64 10257				
		0·78538 46154		8 59080			
$\frac{1}{9}$	0·78076 92308		40 18529		81285		
		0·78539 75435		9 76397		82267	
$\frac{1}{16}$	0·78279 41176		39 83138		82240		81606
		80365		9 81591		81613	
$\frac{1}{25}$	0·78373 15284		81978		81638		81635
		81261		9 81634		81635	
$\frac{1}{36}$	0·78424 07666		81689		81635		
		81582		9 81635			
$\frac{1}{100}$	0·78498 14972		81644				
		81628					
$\frac{1}{144}$	0·78510 88117						

Comments on table 4·1

Not unexpectedly, the 10th decimal is not quite accurate when results have settled down. For 10-decimal accuracy it is clear that one of the 10 and 12 interval formulae is needed; probably $h = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{12}$ would suffice. However, rounding-errors would make it necessary to use further guard figures.

We may note also that here, as in many subsequent examples, it is advantageous to rewrite the basic step (2·1) in terms of n ($n = 1/h$); this may lead to an improvement in rounding-error effects when n is integral.

4·2.

The process may be applied equally well to multi-dimensional integrals, provided that the *ratio* of mesh-lengths in all dimensions is kept the same in all approximations, so that the remainder may be a function of one variable only. Consider, for example,

$$I = \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} \cos x \cos y \, dx\, dy, \quad \text{with} \quad h = 2/n \quad (n = 1, 2, 3, \ldots)$$

at interval the same in both variables. For comparison, the true value to 9 decimals is $I = 0{\cdot}70807\,3418$. Extrapolation with h^2 is appropriate if we use the trapezoidal rule in both variables.

Table 4·2

h	$\frac{1}{4}h^2$	T_h				
				0·708...	0·70807...	
2	1	0·29192 6582				
			0·69353 4870			
1	$\frac{1}{4}$	0·59313 2798		32 7546		
			0·70668 3915		1210	
$\frac{2}{3}$	$\frac{1}{9}$	0·65621 6752		08 7231		3423
			0·70773 6402		3362	
$\frac{1}{2}$	$\frac{1}{16}$	0·67875 6599		07 4903		
			0·70799 0278			
$\frac{1}{3}$	$\frac{1}{36}$	0·69499 7532				

4·3.

As a tour de force, with a double integral over a square domain, we integrate a harmonic function, for which it is possible to use h^4 as elimination variable (see Miller 1960). We evaluate

$$I = \frac{1}{4}\int_{-2}^{2}\int_{-2}^{2} \cos x \cosh y \,\mathrm{d}x\,\mathrm{d}y \doteqdot 3{\cdot}29789\ 48363\ 11236$$

with the same interval $h = 4/n$, $n = 1(1)5$, 10 in both variables and apply the trapezoidal rule in each.

TABLE 4·3

h	$h^4/4^4$	T_h		3·297…	3·29789…		
4	1	−6·26250 33412 29730					
			3·38238 47765 84110				
2	$\frac{1}{16}$	+2·78042 30192 20745		72 24537 73570			
			3·29877 87787 46540		49469 44356		
$\frac{4}{3}$	$\frac{1}{81}$	+3·19638 75176 05642		89 42731 42908		48362 82156	
			3·29794 95547 43135		48364 59216		48363 11237
1	$\frac{1}{256}$	+3·26581 46914 30100		89 48220 38319		48363 11234	
			3·29790 19153 96863		48363 11471		
$\frac{4}{5}$	$\frac{1}{625}$	+3·28475 89884 60077		89 48361 95858			
			3·29789 50174 23404				
$\frac{2}{5}$	$\frac{1}{10000}$	+3·29707 40156 13196					

Comments on table 4·3

It is notable here that, crude though the original approximation with $h = 4$ may be, it makes a definite improvement in accuracy every time it is used, in that the first item in any column is always an improvement on the second item in the previous column. This improvement on using the crudest approximation obtained is very often present, though not invariably so. In table 4·1, for instance, the first line does not show this improvement, but indicates, in fact, the presence of an element in the error term R_n that is difficult or impossible to eliminate since it does not fit into the extended power series in h^2, but which is important for large h (only). This element is akin to similar elements involving, e.g. $\mathrm{e}^{-\lambda x}$, that may be ignored when considering asymptotic series in powers of $1/x$ in Poincaré's sense. The top line in the table thus provides either an improvement, or information about the type of error that can be present.

4·4.

We may also evaluate infinite integrals in (a, ∞) or $(-\infty, \infty)$ in a similar way. We need the range of integration to depend on h; since approximations must involve finite sums or integrals, we take limits $\pm R/h$ instead of $\pm\infty$, with a suitable constant R. A note of caution is needed in these cases, for the error may not be wholly expressible as a series in ascending powers in h; a singularity at $h = 0$ can upset things, and this quite frequently happens without obvious warning with infinite integrals and sums.

Consider

$$I = \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2}$$

and use the trapezoidal rule with $h = 1/n$, and with $R = 2$, i.e. use limits of integration $\pm 2n = \pm 2n^2 h$; all working in table 4·4 was with 7 or more decimals.

Comments on table 4·4

Convergence is not very rapid here. We shall see later that it can be improved when we have demonstrated a feasible method (see §§ 8·1, 10·1) of using the fact that the error series consists only of odd powers of h, while Neville's process perforce eliminates all integral powers of the variable used, starting from the first, a wasteful feature in the present example. The main difficulty is, however, the presence of a fairly substantial part of the error (for large h) that cannot be expressed as a power series in h, and so is not well eliminated, if at all, by Neville's process. A similar error was mentioned in § 4·3, and such errors are discussed in § 12; here it renders the first line rather unhelpful, and prohibits the easy use of intervals $h = 2$ (one strip) and $h = \frac{2}{3}$ (3 strips), because this particular part of the error in this example depends on the position of the origin within the strip containing it; this is readily appreciated by considering trapezoidal approximations with $x = 0$ at mid-strip or at the junction of two strips. For the infinite integral, as $h \to \infty$, the former gives $T_h \to 0$, the latter $T_h \to \infty$. We have used in table 4·4 only approximations with $x = 0$ at a junction of intervals.

Table 4·4

n	h	T_h					
1	1	2·2					
			3·101014				
2	$\frac{1}{2}$	2·650507		3·148072			
			3·132386		3·142058		
3	$\frac{1}{3}$	2·811133		3·143561		3·141921	
			3·137974		3·141944		3·141524
4	$\frac{1}{4}$	2·892843		3·142483		3·141574	
			3·140228		3·141666		
6	$\frac{1}{6}$	2·975305		3·141972			
			3·141100				
8	$\frac{1}{8}$	3·016754					

With $n = (1, 2, 4, 5, 10, 20)$ the result 3·14158 84 is obtained and, with $n = (1, 2, 3, 4, 6, 8, 12, 24)$ or $n \mid 24$ (n divides 24), we get 3·14159 272.

4·5. *Summation of series* (i)

Many infinite series are known, or may be shown or assumed, to have a remainder term R_n, where $S = S_n + R_n$, expressible as a power series in $1/n$. It is clear that Neville's process may be applied to a set of values of S_n to eliminate R_n from $S_n = S - R_n$, as we have done above with $T_h = I - R_h$ for an integral.

It is equally simple to obtain a finite sum S_N for some (large) N by extrapolating the variable $1/n$ to $1/N$ rather than to zero.

We illustrate summation with $\Sigma(1/n^2) = \frac{1}{6}\pi^2 \doteqdot 1{\cdot}64493\ 40668$.

Table 4·5

$1/n$	S_n		1·6...	1·64...	1·6449...		1·64493 4...	
1	1							
		1·5						
$\frac{1}{2}$	1·25		25					
		1·58333 33333		351 85185				
$\frac{1}{3}$	1·36111 11111		3888 88889		6 29632			
		1·61111 11111		472 22224		4 52547		
$\frac{1}{4}$	1·42361 11111		4277 77779		4 74683		5003	
		1·62694 44445		489 11568		3 55757		0174
$\frac{1}{6}$	1·49138 88889		4409 86397		3 79542		0577	
		1·63552 15421		492 39150		3 43107		
$\frac{1}{8}$	1·52742 20522		4459 38049		3 52216			
		1·63915 04472		493 14527				
$\frac{1}{10}$	1·54976 77312		4476 26288					
		1·64102 11744						
$\frac{1}{12}$	1·56497 66384							

Comment on table 4·5

In this case enough figures have been retained to see that the effect of rounding-errors is showing strongly in the last two or three figures, as a sudden slowing down of convergence.

4·6.

We consider next a double series

$$\sum_{m,\,n}' \frac{1}{(m^2+n^2)^{\frac{3}{2}}}$$

summed over all integers m, n, positive or negative, except $m = n = 0$, as indicated by the prime. If we take finite sums over diamonds, i.e. if we write

$$S = S_r + R_r = \sum_{\substack{m,\,n \\ 0<|m|+|n|\leqslant r}}' \frac{1}{(m^2+n^2)^{\frac{3}{2}}} + R_r$$

the main part of R_r can be written as a power series in $1/r$. The table below (computed on the Cambridge Titan) shows that there is again an awkward part to this error term, noticeable for small r, but dying away very rapidly and not representable by the power series.

Table 4·6

$1/r$	S_r			9·03...			
1	4						
$\frac{1}{2}$	5·91421 35624	7·82842 71247					
$\frac{1}{3}$	6·77790 34633	8·50528 32651	8·84371 13351				
$\frac{1}{4}$	7·27016 23715	8·74693 90961	8·98859 49271	688 94577			
$\frac{1}{5}$	7·58697 37667	8·85421 93475	9·01513 97249	283 62569	182 29567		
$\frac{1}{6}$	7·80765 67105	8·91107 14289	9·02477 55915	441 14581	519 90589	587 42793	
$\frac{1}{7}$	7·97009 05453	8·94469 35539	9·02874 88665	404 65665	377 28975	320 24331	275 71255
$\frac{1}{8}$	8·09460 29345	8·96618 96579	9·03067 79703	389 31437	373 97209	371 98147	389 22751
$\frac{1}{9}$	8·19306 38619	8·98075 12807	9·03171 69611	379 49427	367 21913	361 81679	356 73441
$\frac{1}{10}$	8·27286 40343	8·99106 55865	9·03232 28097	373 64567	364 87281	362 52647	362 99961
$\frac{1}{11}$	8·33884 32309	8·99863 51977	9·03269 84475	370 01483	363 66085	362 20649	361 93989
$\frac{1}{12}$	8·39430 23929	9·00435 31743	9·03294 30579	367 68889	363 03703	362 16365	362 12083
$\frac{1}{13}$	8·44156 96809	9·00877 71365	9·03310 89289	366 18317	362 79531	362 40855	362 69425

Comments on table 4·6

The awkward rapidly decaying error (here alternating in sign) is evident in the first two lines after three applications of the Neville step, and extends to four lines after the fifth application. In the last column, rounding errors are affecting at least the four or more final digits near the bottom of the column. A general examination of the results suggest no better than $S = 9{\cdot}03362$, which is correct to 5 decimals.

We could also sum over squares $|m|$, $|n| \leqslant r$, but this involves using more points—the error term might however be better behaved, it has not been examined.

5. A matrix formulation of Neville's process

In order to extend Neville's process further, it is convenient to express it first in matrix terms.

5·1.

We start from the Taylor expansion (3·12) for $f(x_r)$, with remainder term starting in $f^{(k)}(X)$. With $\xi_r = x_r - X$, and dropping (X) as before, it is

$$f(x_r) = f + \xi_r f' + \ldots + \xi_r^s f^{(s)}/s! + \ldots + \xi_r^{k-1} f^{(k-1)}/(k-1)! + R_r^{(k)} \quad r = 0(1)\,n \qquad (5{\cdot}11)$$

which we may write as

$$\begin{pmatrix} 1 & \xi_0 & \xi_0^2 & \dots & \xi_0^{k-1} \\ 1 & \xi_1 & \xi_1^2 & \dots & \xi_1^{k-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \xi_j & \xi_j^2 & \dots & \xi_j^{k-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \xi_n & \xi_n^2 & \dots & \xi_n^{k-1} \end{pmatrix} \begin{pmatrix} f(X) \\ f'(X) \\ \dots \\ f^{(i)}(X)/i! \\ \dots \\ f^{(k-1)}(X)/(k-1)! \end{pmatrix} + \begin{pmatrix} R_0^{(k)} \\ R_1^{(k)} \\ \dots \\ R_j^{(k)} \\ \dots \\ R_n^{(k)} \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \dots \\ f(x_j) \\ \dots \\ f(x_n) \end{pmatrix} \tag{5·12}$$

or

$$\mathbf{A}_{n,k}\,\mathbf{t}_k + \mathbf{r}_n^{(k)} = \mathbf{f}_n \tag{5·13}$$

in which $\mathbf{A}_{n,k}$ is a Van der Monde matrix.

Neville's process is applied to the vector $\mathbf{f}_n$, to convert it step-by-step into the vector of leading diagonal elements of table 2·3 or as exhibited in the tableaux of § 4. Intermediate stages give vectors of elements along the diagonal for several columns, and then down the rest of the column thus reached. By retaining in this way the top elements on the diagonal at each stage, we conveniently retain a vector with $n+1$ elements throughout.

5·2.

The elimination process of § 3 outlines the effect of Neville's process on the matrix $\mathbf{A}_{n,k} = (\mathbf{u}_n|\mathbf{A}^*_{n,k})$, where $\mathbf{u}_n = (1, 1, \dots, 1)^T$ is the first column of $\mathbf{A}_{n,k}$, and $\mathbf{A}^*_{n,k}$ is the remaining matrix of non-zero powers. In fact, *Neville's process is a simple repetitive process for zeroizing sub-diagonal elements of the matrix* $\mathbf{A}^*_{n,k}$, and the steps for the first column of reductions are collectively obtained by premultiplying by the $(n+1)\times(n+1)$ matrix

$$\mathbf{N}_n^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ \frac{-\xi_1}{\xi_0-\xi_1} & \frac{\xi_0}{\xi_0-\xi_1} & 0 & 0 & \dots \\ 0 & \frac{-\xi_2}{\xi_1-\xi_2} & \frac{\xi_1}{\xi_1-\xi_2} & 0 & \dots \\ 0 & 0 & \frac{-\xi_3}{\xi_2-\xi_3} & \frac{\xi_2}{\xi_2-\xi_3} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

Likewise the next column results from a further premultiplication by

$$\mathbf{N}_n^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & \frac{-\xi_2}{\xi_0-\xi_2} & \frac{\xi_0}{\xi_0-\xi_2} & 0 & \dots \\ 0 & 0 & \frac{-\xi_3}{\xi_1-\xi_3} & \frac{\xi_1}{\xi_1-\xi_3} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

and so on for subsequent columns.

The elements of these matrices are two-point Lagrange multipliers. We shall write, for the $(k+1)$-point Lagrange multipliers, of degree k,

$$L_k(x; x'_r; x'_0, x'_1, \dots, x'_k) = \frac{\pi(x)}{(x-x'_r)\,\pi'(x'_r)}$$

in which $\pi(x) = (x-x_0')\,(x-x_1')\dots(x-x_k')$, the product of $(x-x_s')$ for all x_s' mentioned after the second semicolon in L_k, and $\pi'(x)$ is its derivative. The x_s' are some set of $k+1$ chosen from the x_r. In the matrices below, in which $x = X$ for every L_k, we shall further abbreviate by omitting X, and retaining only the suffices in the x_r used. Thus $L_2(X;\, x_1;\, x_1, x_2, x_3)$ becomes $L_2(1;\, 1(1)3) \equiv L_k(1;\, 1, 2, 3)$ and

$$\mathbf{N}_n^{(1)} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & \dots \\ L_1(0;\,0,1) & L_1(1;\,0,1) & 0 & \dots & 0 & 0 & \dots \\ 0 & L_1(1;\,1,2) & L_1(2;\,1,2) & \dots & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & L_1(r;\,r,r+1) & L_1(r+1;\,r,r+1) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

$$\mathbf{N}_n^{(2)} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots \\ 0 & L_1(0;\,0,2) & L_1(2;\,0,2) & \dots & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & L_1(r-1;\,r-1,r+1) & L_1(r+1;\,r-1,r+1) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

and the product is clearly

$$\mathbf{L}_n^{(2)} = \mathbf{N}_n^{(2)}\mathbf{N}_n^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ L_1(0;\,0,1) & L_1(1;\,0,1) & 0 & 0 & 0 & \dots \\ L_2(0;\,0(1)2) & L_2(1;\,0(1)2) & L_2(2;\,0(1)2) & 0 & 0 & \dots \\ 0 & L_2(1;\,1(1)3) & L_2(2;\,1(1)3) & L_2(3;\,1(1)3) & 0 & \dots \\ 0 & 0 & L_2(2;\,2(1)4) & L_2(3;\,2(1)4) & L_2(4;\,2(1)4) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

since the result of application to $\mathbf{f}_n$ consists, apart from the first two elements, of a set of quadratic (Lagrange) interpolation polynomials at $x = X$.

5·3.

It seems worth while to exhibit the corresponding matrices for Aitken's process. They are

$$\mathbf{N}_n^{(1)*} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \dots \\ L_1(0;\,0,1) & L_1(1;\,0,1) & 0 & \dots & 0 & \dots \\ L_1(0;\,0,2) & 0 & L_1(2;\,0,2) & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ L_1(0;\,0,r) & 0 & 0 & \dots & L_1(r;\,0,r) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

$$\mathbf{N}_n^{(2)*} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \dots \\ 0 & 1 & 0 & \dots & 0 & \dots \\ 0 & L_1(1;\,1,2) & L_1(2;\,1,2) & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & L_1(1;\,1,r) & 0 & \dots & L_1(r;\,1,r) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

with product

$$\mathbf{L}_n^{(2)*} = \mathbf{N}_n^{(2)*}\,\mathbf{N}_n^{(1)*} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \dots \\ L_1(0;\,0,\,1) & L_1(1;\,0,\,1) & 0 & \dots & 0 & \dots \\ L_1(0;\,0,\,1,\,2) & L_2(1;\,0,\,1,\,2) & L_2(2;\,0,\,1,\,2) & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ L_2(0;\,0,\,1,\,r) & L_2(1;\,0,\,1,\,r) & 0 & \dots & L_2(r;\,0,\,1,\,r) & \dots \end{pmatrix}$$

5·4.

The application of the Neville process for k columns ($k \leqslant n$) then yields

$$\mathbf{L}_n^{(k)}\,\mathbf{A}_{n,\,k}\,\mathbf{t}_k + \mathbf{L}_n^{(k)}\,\mathbf{r}_n^{(k)} = \mathbf{L}_n^{(k)}\,\mathbf{f}_n. \tag{5·41}$$

The right-hand side consists, as mentioned above, of the first $k+1$ diagonal elements of the Neville tableau, together with the remainder of the $(k+1)$st column; these are the results of computation on the estimates. The left-hand side gives a vector of expressions equivalent to the results on the right. Consider the terms of (5·41) in turn. We can write

$$\mathbf{A}_{n,\,k}\,\mathbf{t}_k = (\mathbf{u}_n|\mathbf{A}^*_{n,\,k})\,\mathbf{t}_k = \mathbf{u}_n f(X) + (0|\mathbf{A}^*_{n,\,k})\,\mathbf{t}_k \tag{5·42}$$

and consider premultiplication by $\mathbf{L}_n^{(k)}$. Clearly

$$\mathbf{L}_n^{(k)}\,\mathbf{u}_n f(x) = \mathbf{u}_n f(x)$$

since this is just interpolation of a constant and yields a vector of identical terms. We have also seen in § 5·2 that $\mathbf{L}_n^{(k)}\,\mathbf{A}^*_{n,k}$ results in a matrix consisting entirely of zeros below the main diagonal.

We shall now consider the two terms involving $(0|\mathbf{A}^*_{n,\,k})\,\mathbf{t}_k$ and $\mathbf{r}_n^{(k)}$. These provide a vector of remainders, each expressed as a series, as mentioned in § 3·4, ending with (or, when $r > k$, consisting of) a remainder term of the form given in (2·2). For each row we can replace the series by a *single equivalent remainder term* of the form in (2·2), the usual Lagrange remainder term, as remarked in § 3·4. This remainder term, for the rth row, will contain r factors $(x-x_i)$ when $r \leqslant k$ and will be denoted by $R^{*(r-1)}_{r-1}$; when $r > k$ there are $k+1$ factors $(x-x_i)$ and the term will then be denoted by $R^{*(k)}_{r-1}$. We thus obtain the remainder vector

$$\mathbf{r}_n^{*(k)} = (R_0^{*(0)},\, R_1^{*(1)},\, \dots,\, R_k^{*(k)},\, \dots,\, R_s^{*(k)},\, \dots,\, R_n^{*(k)})^T \tag{5·43}$$

in which the upper affix in each remainder is one less than the degree of the product-coefficient of the derivative

$$f^{(r)}(\eta_{r-1})/r!,\ r \leqslant k \quad \text{or} \quad f^{(k+1)}(\eta_{r-1})/(k+1)!,\ r > k.$$

This term is highly ineffective as a remainder for early approximations on the Neville tableau diagonal, but improves as k increases, until it can be satisfactorily neglected. Thus finally (5·41) becomes

$$\mathbf{u}_n f(X) + \mathbf{r}_n^{*(k)} = \mathbf{L}_n^{(k)}\,\mathbf{f}_n \tag{5·44}$$

and only the right-hand side has to be computed, until satisfactory agreement (see § 12) indicates that the $R_n^{*(k)}$ reached are negligible. These remarks, by suitable choice of k, clearly apply to any approximation in the Neville tableau; intercomparison of items may be used to give reasonable estimates of the error in most of them, and to suggest which gives the best result, and how good this result may be. Full discussion of error analysis is beyond the scope of the present paper, but is carried somewhat further in § 12.

6. Steps in the elimination process

Before suggesting extensions we again list below the precise steps to be taken in thinking about, and in applying, Neville's process. We want to make the actual numerical process as simple and effortless as possible, increasing the work to be done only if necessary to achieve wider application or higher precision.

I. We first isolate the quantity it is desired to compute, as a single term in each equation. This corresponds to the detachment of $\mathbf{u}_n f(X)$ from $\mathbf{A}_{n,k}\,\mathbf{t}_k$ leaving $(0|\mathbf{A}^*_{n,k})\,\mathbf{t}_k$.

II. We next apply a premultiplying matrix $\mathbf{L}_n^{(k)}$ so that $\mathbf{L}_n^{(k)}(0|\mathbf{A}^*_{n,k})$ is reduced to zeros on and below the main diagonal.

III. The terms $\mathbf{L}_n^{(k)}(0|\mathbf{A}^*_{n,k})\,\mathbf{t}_k+\mathbf{L}_n^{(k)}\mathbf{r}_n^{(k)}$ are then combined into a single vector $\mathbf{r}_n^{*(k)}$ of error terms, in a tidying up process. The sole use made of these terms is to search for negligible ones.

IV. The only numerical work to be done is to evaluate $\mathbf{L}_n^{(k)}\mathbf{f}_n$, to obtain estimates of the result. This depends on the fact that $\mathbf{L}_n^{(k)}\mathbf{u}_n = \mathbf{u}_n$.

V. We are particularly anxious to choose a method for developing $\mathbf{L}_n^{(k)}$ in its application to $\mathbf{f}_n$ that is easy to apply. This is exactly why the Neville process is so attractive; it works only on a matrix $\mathbf{A}^*_{n,k}$ in which each row is formed of powers ξ_i^r, $r = 1(1)k$, i.e. $\xi_i, \xi_i^2, \xi_i^3, \ldots$; *all must be assumed present, starting from the first.* Any values ξ_i may be used.

We discuss in §§ 8, 9, 10 ways in which these steps may be usefully varied.

7. Romberg integration

Romberg (1955) introduced a method of integration making use of the trapezoidal rule, with successive halving of the interval. Successive results are then used (as in Richardson's 'Deferred Approach to the Limit') to eliminate in turn errors of the form $\alpha h^2, \beta h^4, \gamma h^6, \ldots$, which normally occur when the function is 'well-behaved', and h not 'too large'. Neville's process does, in fact, apply in this case, and can be used with *exactly* the same multipliers in every step in each column, though differing from column to column. The formula is, in fact

$$f\left(0;\,h_i^2,\frac{h_i^2}{2^2},\ldots,\frac{h_i^2}{2^{2j}}\right) = \begin{vmatrix} h_i^2 & f(0;\,h_i^2,\ldots,h_i^2/2^{2j-2}) \\ h_i^2/2^{2j} & f(0;\,h_i^2/2^2,\ldots,h_i^2/2^{2j}) \end{vmatrix} \div \left(h_i^2-\frac{h_i^2}{2^{2j}}\right)$$

$$= \begin{vmatrix} 2^{2j} & f(0;\,h_i^2,\ldots,h_{i+j-1}^2) \\ 1 & f(0;\,h_{i+1}^2,\ldots,h_{i+j}^2) \end{vmatrix} \div (2^{2j}-1)$$

which corresponds to the elimination of h^{2j}.

This method thus restricts values of h to be in geometrical progression. A result of this, used by Fox (1967), is that *the powers eliminated need not be in geometrical progression*, as they must be for Neville's process to apply; nor need they start with the first power—they can be quite arbitrary, and can be eliminated in arbitrary order.

In fact, if the error is of the form $\alpha_1 h^{\beta_1}+\alpha_2 h^{\beta_2}+\ldots$ we eliminate $\alpha_j h^{\beta_j}$ by evaluating

$$f(0;\,h_i^2,h_{i+1}^2,\ldots,h_{i+k}^2) = \begin{vmatrix} 2^{\beta_j} & f(0;\,h_i^2,\ldots,h_{i+k-1}^2) \\ 1 & f(0;\,h_{i+1}^2,\ldots,h_{i+k}^2) \end{vmatrix} \div (2^{\beta_j}-1)$$

and so on. It is, however, essential that $h_{i+1} = \lambda h_i$, with the same λ for all i (here $\lambda = 2$), so that the column vector $h_i^{\beta_r}$ has *all altered elements multiplied by the same constant.*

We can thus include a *Romberg process* in our battery of methods for elimination of errors. Fox (1967) gives some examples (with a slightly less systematic method of use). We note that the Romberg process involves an amount of work comparable with Neville's process in similar circumstances.

8. Extensions of the processes

8·1. *Premultiplication by a diagonal matrix*

The error expansion in powers of h sometimes involves odd powers only, or powers in arithmetic progression with indices other than the complete series of positive integers. We may be able to convert the matrix $\mathbf{A}^*_{n,k}$ into standard form (and so avoid having to eliminate 'vacant' columns corresponding to absent terms) by premultiplication by a diagonal matrix.

For example

$$\mathbf{D}(x_0, x_1, \ldots, x_n)\begin{pmatrix} x_0 & x_0^3 & x_0^5 & \ldots \\ x_1 & x_1^3 & x_1^5 & \ldots \\ x_2 & x_2^3 & x_2^5 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{pmatrix} = \begin{pmatrix} x_0^2 & x_0^4 & x_0^6 & \ldots \\ x_1^2 & x_1^4 & x_1^6 & \ldots \\ x_2^2 & x_2^4 & x_2^6 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{pmatrix}$$

the last matrix being in standard form for Neville's process with variable x^2. Here $\mathbf{D}(x_0, x_1, \ldots, x_n)$ represents a diagonal matrix with diagonal as indicated; more briefly it will be written as $\mathbf{D} = \mathbf{D}(\mathbf{x})$, with $\mathbf{x} = (x_0, x_1, \ldots, x_n)^T = (x_r)$ as diagonal.

Equation (5·41) is then replaced by

$$\mathbf{L}_n^{(k)}\mathbf{D}\mathbf{A}_{n,k}\,\mathbf{t}_k + \mathbf{L}_n^{(k)}\mathbf{D}\,\mathbf{r}_n^{(k)} = \mathbf{L}_n^{(k)}\mathbf{D}\mathbf{f}_n \tag{8·1}$$

where $\mathbf{L}_n^{(k)}$ is the appropriate matrix for variable x^2. Equation (5·44) then becomes

$$\mathbf{L}_n^{(k)}\mathbf{v}_n f(X) + \bar{\mathbf{r}}_n^{(k)} = \mathbf{L}_n^{(k)}\,\mathbf{D}\,\mathbf{f}_n \tag{8·2}$$

in which $\mathbf{v}_n = \mathbf{D}\mathbf{u}_n$ and $\bar{\mathbf{r}}_n^{(k)}$ is still neglected (except perhaps for study concerning which approximations give it small values). We now have to compute $\mathbf{D}\mathbf{f}_n$ (a simple matter) before applying $\mathbf{L}_n^{(k)}$, and *and we must also compute* $\mathbf{L}_n^{(k)}\mathbf{v}_n$; this is an extra evaluation of a complete Neville tableau. The approximations to $f(X)$ are now given by *ratios* of corresponding terms in $\mathbf{L}_n^{(k)}\,\mathbf{D}\mathbf{f}_n$ and $\mathbf{L}_n^{(k)}\mathbf{v}_n$. The amount of work is thus approximately doubled, but with the gain of superior convergence in many cases, and of application to cases not otherwise accessible to Neville's process.

Other diagonal matrices of use are $\mathbf{D}(h_0^{-2}, h_1^{-2}, h_2^{-2}, \ldots)$ which reduces the Simpson rule error $ah^4 + bh^6 + ch^8 + \ldots$ to standard form; $\mathbf{D}(x_0^{-1}, x_1^{-1}, x_2^{-1}, \ldots)$ which reduces the x_i column (cofactor of $f'(X)$) to $\mathbf{u}_n$, see § 10·5, later; $\mathbf{D}(h_0^{\frac{3}{2}}, h_1^{\frac{3}{2}}, h_2^{\frac{3}{2}}, \ldots)$ which reduces to standard form the error term for the Trapezoidal rule applied to $\int_0^1 \surd\{x(1-x)\}\,\mathrm{d}x$.

8·2. *Replacement of the unit vector*

We have seen in § 8·1 that premultiplication by $\mathbf{D}$ to render $\mathbf{A}^*_{n,k}$ more amenable to Neville's process replaces $\mathbf{u}_n$ by $\mathbf{D}\mathbf{u}_n = \mathbf{v}_n$, which is no longer a unit vector, and must be processed. In fact, a non-unit vector $\mathbf{v}_n$ may arise in other ways.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

An example occurs in the summation of series of terms of alternating sign. If we require

$$S = \sum_{0}^{\infty} (-1)^r a_r$$

we may find that S_r is expressible in the form

$$S_r = S + (-1)^r R_r$$

in which $$R_r = \sum_{j=1}^{\infty} \alpha_j r^{-j}.$$

If we now write

$$(-1)^r S + \frac{\alpha_1}{r} + \frac{\alpha_2}{r^2} + \ldots = (-1)^r S_r \quad (r = 0, 1, 2, \ldots)$$

we have the matrix equation, involving several partial sums S_r,

$$\mathbf{v}_n S + (0|\mathbf{A}^*_{n,k})\,\mathbf{d}_k + \mathbf{r}^{(k)}_n = \mathbf{s}_n$$

where
$$\mathbf{s}_n = (S_0, -S_1, S_2, -S_3, \ldots)^T$$
$$\mathbf{d}_k = (S, \alpha_1, \alpha_2, \ldots, \alpha_k)^T$$
and
$$\mathbf{v}_n = (1, -1, 1, -1, \ldots)^T$$
with $n+1$ elements, or more generally, with $r = n_0, n_1, n_2, \ldots$

$$\mathbf{v}_n = ((-1)^{n_0}, (-1)^{n_1}, (-1)^{n_2}, \ldots)^T = ((-1)^{n_r})$$
and
$$\mathbf{s}_n = ((-1)^{n_0} S_{n_0}, (-1)^{n_1} S_{n_1}, \ldots)^T = ((-1)^{n_r} S_{n_r}).$$

As in § 8·1, reduction involves using the Neville process on both $\mathbf{s}_n$ and $\mathbf{v}_n$; this can be successfully and readily done. We may also combine the processes of §§ 8·1 and 8·2, as in the example of § 10·3, with negligible extra work.

In the tables in § 11 below we shall drop the suffix n and use $\mathbf{u}$ for the vector $(1, 1, 1, \ldots)^T$ also denoted by (1), assuming the appropriate number of elements. Similarly†, for arguments $h_r = 1/n_r$, we use $\mathbf{v}^{(1)}$ for the vector $((-1)^{n_r})$, $\mathbf{v}^{(2)}$ for (n_r), $\mathbf{v}^{(3)}$ for $(1/n_r)$; these are the only ones for which results are tabulated, except in discussion of rounding-error, where $\mathbf{v}^{(4)} = (1, -1, 1, -1, \ldots)^T$, with *strict* alternation, is used.

9. Initial isolation of two terms

9·1. *Derivatives*

In § 5·4, we detached from $\mathbf{A}_{n,k}\mathbf{t}_k$ the initial column $\mathbf{u}_n f(X)$. We could also detach the second column, writing

$$\mathbf{A}_{n,k}\mathbf{t}_k = (\mathbf{u}_n|\mathbf{w}^{(1)}_n|\mathbf{A}^{**}_{n,k})\,\mathbf{t}_k$$
$$= \mathbf{u}_n f(X) + \mathbf{w}^{(1)}_n f'(X) + (0|0|\mathbf{A}^{**}_{n,k})\,\mathbf{t}_k$$

in which $$\mathbf{w}^{(s)}_n = (\xi^s_0, \xi^s_1, \xi^s_2, \xi^s_3, \ldots)^T, \quad \text{with } s = 1.$$

For the application of Neville's process we must also restore $\mathbf{A}^{**}_{n,k}$ to standard form by premultiplication by $\mathbf{D}(\xi_1^{-1}, \xi_2^{-1}, \ldots) = \mathbf{D}(\mathbf{w}^{(-1)}_n)$. Equation (5·41) thus becomes

$$\mathbf{L}^{(k)}_n \mathbf{w}^{-1}_n f(X) + \mathbf{u}_n f'(X) + \{\mathbf{L}^{(k)}_n (0|0|\mathbf{A}^*_{n,k-1})\,\mathbf{t}_k + \mathbf{L}^{(k)}_n \mathbf{D}\mathbf{r}^{(k)}_n\} = \mathbf{L}^{(k)}_n \mathbf{D}\mathbf{f}_n$$

† Note that a *single symbol* for a vector always stands for a *column* vector unless ^{T} is affixed. Thus the column vector with typical element u_r is written (u_r) so that

$$(u_r)^T = (u_1, u_2 \ldots, u_n) \quad \text{or} \quad (u_r) = (u_1, u_2, \ldots, u_n)^T.$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

since $$\mathbf{D}\mathbf{A}^{**}_{n,k} = \mathbf{A}^{*}_{n,k-1}, \quad \mathbf{D}\mathbf{w}^{(1)}_{n} = \mathbf{u}_{n}, \quad \mathbf{D}\mathbf{u}_{n} = \mathbf{w}^{(-1)}_{n}.$$

The remainder (comprising the terms in { }), as usual, will be treated as negligible, with careful examination to decide how far this is so. We now pick out *two* accurate equations, for terms where both errors are judged quite negligible, and solve these simultaneously for $f(X)$ and $f'(X)$. In this way a derivative may be calculated and, by use of other equations or pairs, checked.

It is clear that this method is capable of extension to higher derivatives.

9·2. *Removal of a 'rogue' error term*

The same principle may be applied if the error of the initial estimates consists of a standard series of powers with an extra term, for example, if

$$R(\xi) = a\xi + b\xi^2 + c\xi^3 + \ldots + \alpha\phi(\xi).$$

In place of (5·13) we then have

$$(\phi(\xi_r)|\mathbf{A}_{n,k})\ (\alpha|\mathbf{t}^T_k)^T + \mathbf{r}^{(k)}_n = \mathbf{f}_n$$

whence, in place of (5·44)

$$\alpha\mathbf{L}^{(k)}_n(\phi(\xi_r)) + \mathbf{u}_n f(X) + \mathbf{r}^{*(k)}_n = \mathbf{L}^{(k)}_n \mathbf{f}_n.$$

As before, $\mathbf{r}^{*(k)}_n$ is neglected, and we use *two* good estimates to determine (or eliminate) α and so find $f(X)$.

10. Further applications and examples

10·1. *Quadrature* (ii)

In the example of § 4·4, namely

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2},$$

the error in T_h is, as $h \to \infty$, of form $\alpha h + \beta h^3 + \gamma h^5 + \ldots$. We therefore multiply by $\mathbf{D}(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8})$ and eliminate $\alpha h^2 + \beta h^4 + \gamma h^6 + \ldots$. The vectors $\mathbf{v}^{(3)} = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8})^T$ and $\mathbf{D}\mathbf{f}_r = (T_1, \frac{1}{2}T_{\frac{1}{2}}, \frac{1}{3}T_{\frac{1}{3}}, \frac{1}{4}T_{\frac{1}{4}}, \frac{1}{6}T_{\frac{1}{6}}, \frac{1}{8}T_{\frac{1}{8}})^T$ are listed one below the other, and Neville's process applied successively to the pairs. We work to eight decimals (to allow some accumulation of rounding error), and retain rational form when working on $\mathbf{v}^{(3)}$. Ratios are given to seven decimals.

Table 10·1

n	h^2	hT_h h					
1	1	2·2 1					
			1·03367 120 $\frac{1}{3}$				
2	$\frac{1}{4}$	1·32525 340 $\frac{1}{2}$		0·57557 789 $\frac{11}{60}$			
			0·62647 715 $\frac{1}{5}$		0·37643 726 $\frac{151}{1260}$		
3	$\frac{1}{9}$	0·93704 437 $\frac{1}{3}$		0·38888 355 $\frac{13}{105}$		0·25082 894 $\frac{503}{6300}$	
			0·44828 195 $\frac{1}{7}$		0·25431 806 $\frac{17}{210}$		0·17626 928 $\frac{244963}{4365900}$
4	$\frac{1}{16}$	0·72321 081 $\frac{1}{4}$		0·26926 978 $\frac{3}{35}$		0·17743 427 $\frac{1957}{34650}$	
			0·31402 282 $\frac{1}{10}$		0·18223 951 $\frac{67}{1155}$		
6	$\frac{1}{36}$	0·49588 415 $\frac{1}{6}$		0·19447 814 $\frac{13}{210}$			
			0·22436 431 $\frac{1}{14}$				
8	$\frac{1}{64}$	0·37709 422 $\frac{1}{8}$					

TABLE 10·1 (*cont.*)

Ratios

2·2	3·10101 36	3·13951 58	3·14113 21	3·14159 51	3·14159 30
2·65050 68	3·13238 58	3·14098 25	3·14157 60	3·14159 30	
2·81113 31	3·13797 36	3·14148 08	3·14159 16		
2·89284 32	3·14022 82	3·14157 00			
2·97530 49	3·14110 03				
3·01675 38					

Comment on table 10·1

The improvement in the result over that of § 4·4 is noticeable. Using $n = 1, 2, 4, 5, 10, 20$ gives 3·14159 2665, with 10-figure working.

10·2.

For the integral

$$I = \int_0^1 \sqrt{\{x(1-x)\}}\, dx = \tfrac{1}{8}\pi \doteqdot 0{\cdot}39269\ 90817,$$

Fox (1967) has shown that the trapezoidal rule gives an error of form $\alpha h^{\frac{3}{2}} + \beta h^{\frac{5}{2}} + \gamma h^{\frac{7}{2}} + \ldots$. We shall therefore reduce trapezoidal sums T_h by using Neville's process on $T_h/\sqrt{h}$ and on $1/\sqrt{h}$ with variable h.

TABLE 10·2

h	$T_h/\sqrt{h}$					
1	0	0·70710 67812	1·11426 04239	1·47707 03089	2·00585 90217	2·77069 38212
$\frac{1}{2}$	0·35355 33906	1·01247 20132	1·40450 83319	1·95298 01502	2·73245 20812	
$\frac{1}{4}$	0·68301 27019	1·24769 38044	1·84328 57867	2·65450 48881		
$\frac{1}{5}$	0·79594 89224	1·60504 89938	2·49226 10678			
$\frac{1}{10}$	1·20049 89581	2·27045 80493				
$\frac{1}{20}$	1·73547 85037					

h	$1/\sqrt{h}$					
1	1	1·82842 71248	2·83823 95419	3·76132 61822	5·10787 65478	7·05551 38590
$\frac{1}{2}$	1·41421 35624	2·58578 64376	3·57670 88541	4·97322 15112	6·95813 19934	
$\frac{1}{4}$	2	3·18033 98875	4·69391 89798	6·75964 09452		
$\frac{1}{5}$	2·23606 79775	4·08848 73429	6·34649 65521			
$\frac{1}{10}$	3·16227 76602	5·78199 42498				
$\frac{1}{20}$	4·47213 59550					

Ratios

0	0·38672 95402	0·39258 85773	0·39269 93399	0·39269 91974	0·39269 90828
0·25	0·39155 28361	0·39268 17611	0·39269 92084	0·39269 90870	
0·34150 63509	0·39231 46106	0·39269 65494	0·39269 90959		
0·35595 91794	0·39257 77089	0·39269 87193			
0·37963 11036	0·39267 73274				
0·38806 47908					

Comment on table 10·2

Nearly eight correct decimals have been achieved with 21 points, and results are still improving.

This example, where no integer powers of h appear in the error expansion, suggests strongly that the form of error to be eliminated depends mainly, if not entirely, on the ends of the range.

10·3. *Summation of series* (ii)

As another example in which the vector $\mathbf{u}_n$, multiplying the desired end-product, is replaced by a more general vector $\mathbf{v}_n$, consider the series

$$S = \sum_1^\infty (-1)^{n-1}/n.$$

For comparison, we note that $S = \ln 2 = 0{\cdot}69314\ 71806$. Here we take

$$S = \sum_1^{n-1} (-1)^{r-1}/r + (-1)^{n-1}/2n + R_n = S_n + R_n$$

where it can be shown that

$$(-1)^{n-1}R_n = \frac{\alpha}{n^2}+\frac{\beta}{n^4}+\frac{\gamma}{n^6}+\ldots+E_n = R_n^*$$

in which E_n is supposed to die out very rapidly indeed, and to be negligible except perhaps for a very few small values of n. We write the relation in the form

$$(-1)^n S_n = (-1)^n S - R_n^*$$

and eliminate R_n^*, using $n = n_0, n_1, n_2, \ldots$, so that $\mathbf{v}_n = ((-1)^{n_0}, (-1)^{n_1}, \ldots)^T$. Results are exhibited in table 10·3, in which $1/n^2$ is used as elimination variable.

TABLE 10·3

n	$(-1)^n S_n$						
1	−0·5						
2	+0·75	+1·16666 66667					
3	−0·66666 66667	−1·8	−2·17083 33334				
4	+0·70833 33333	+2·47619 04762	+3·90158 73016	+4·30641 53440			
5	−0·68333 33333	−3·15740 74074	−6·32630 62170	−8·27447 64109	−8·79868 02341		
6	+0·7	+3·84393 93938	+9·44501 68347	+14·70212 45187	+17·57419 96349	+18·32771 04883	
7	−0·68809 52381	−4·53205 12821	−13·25704 15693	−24·26410 01895	−33·03150 07489	−37·52978 52271	−38·69348 30545
8	+0·69702 38095	+5·22174 60318	+17·76234 25783	+37·64656 31857	+58·28345 09768	+73·22589 76216	+80·60960 98120
10	0·69563 49206	0·69316 57847	−3·65782 30802	−15·70666 62637	−33·49107 60805	−50·97193 83773	−63·25524 08392
12	0·69487 73448	0·69315 55816	+0·69314 74191	+2·93733 22029	+9·15199 83585	+18·11062 74423	+26·74594 81697
14	0·69441 94693	0·69315 15064	0·69314 72614	0·69314 71849	−0·05491 44878*	−2·12646 98782	−5·08510 98373
16	0·69412 18503	0·69314 96282	0·69314 72134	0·69314 71826	+0·69314 71818	+0·87022 45819	+1·36059 27663

n	$(-1)^n$						
1	−1						
2	+1	$+\frac{5}{3}$					
3	−1	$-\frac{13}{5}$	−3·13333 33333				
4	+1	$+\frac{25}{7}$	+5·62857 14286	+6·21269 84127			
5	−1	$-\frac{41}{9}$	−9·12698 41269	−11·93756 61375	−12·69382 71605		
6	+1	$+\frac{61}{11}$	+13·62626 26261	+21·21067 82105	+25·35420 87543	26·44129 54947	
7	−1	$-\frac{85}{13}$	−19·12587 41259	−35·00569 80057	−47·65438 26545	−54·14403 52245	−55·82289 62808
8	+1	$+\frac{113}{15}$	+25·62564 10261	+54·31250 97143	+84·08524 56216	+105·64263 93395	+116·29508 43104
10	1	1	−5·27712 41830	−22·65992 96137	−48·31740 93897	−73·53696 27256	−91·25802 22696
12	1	1	+1	+4·23767 45786	+13·20354 26427	+26·12811 23973	+38·58624 67877
14	1	1	1	1	−0·07922 48595*	−3·06784 75475	−7·33626 27443
16	1	1	1	1	+1	+1·25546 86866	+1·96292 04340

Ratios

n			0·6931...			
2	0·7					
3	0·69230 76923	0·69281 91490				
4	0·69333 33333	0·69317 54089	6 34304			
5	0·69308 94309	0·69314 31160	4 60162	4 63713		
6	0·69316 93989	0·69314 80190	4 73088	4 72327	4 72209	
7	0·69313 72549	0·69314 69632	4 71609	4 71757	4 71781	4 71785
8	0·69315 21281	0·69314 72489	4 71844	4 71811	4 71807	4 71806
10	0·69316 57847	0·69314 70538	4 71779	4 71804	4 71805	4 71806
12	0·69315 55816	0·69314 74191	4 71845	4 71807	4 71806	4 71806
14	0·69315 15064	0·69314 72614	4 71849	4 71781*	4 71806	4 71806
16	0·69314 96282	0·69314 72134	4 71826	4 71818	4 71818	4 71815

Comments on table 10·3

This rather lengthy example exhibits several points.

(i) We can select the sums S_n to be included in the table. Here we have chosen $n = 1(1)8(2)16$; this gives $\mathbf{v}_n$ with units of alternating sign up to the eighth term, with a tail consisting of units as in $\mathbf{u}_n$. This allows comparison to some extent of results using either $\mathbf{v}_n$ or $\mathbf{u}_n$.

(ii) The part with alternating signs involves extra work, noticeable with desk machines, less so with an automatic computer.

(iii) Accuracy is better where signs alternate because of the larger divisors used in obtaining S; the rounding-error effects depend on the basic Neville steps, and not on what happens to $\mathbf{u}_n$ or $\mathbf{v}_n$. These errors are noticeable in the lower lines to the right of the table. See also § 12·4.

(iv) The possible presence of small divisors is well exhibited in items marked with an asterisk *. The increased rounding-error is apparent, but ephemeral; it does not affect later columns.

(v) Rounding-error effects increase towards the bottom of the table and *tend to persist* in upward sloping diagonals; not unexpectedly. Comparison of results in the same vertical column provides the most useful information on rounding errors. See also § 12·7 later. Values were computed on the Cambridge Titan with nearly 11-figure accuracy.

10·4. *Evaluation of a limit*

As an example of the type involving two unknowns, we shall not obtain an actual derivative, but use the example of § 10·2 to obtain the value of α. In other words, we shall evaluate

$$-\alpha = \lim_{h\to 0} h^{-\frac{3}{2}} \left\{\int_0^1 \sqrt{\{x(1-x)\}}\,\mathrm{d}x - T_h\right\} = \lim_{h\to 0} (I-T_h)/h^{\frac{3}{2}}$$

in which T_h is the trapezoidal estimate to the integral

$$T_h = \sum_{r=1}^{n-1} \sqrt{\{rh(1-rh)\}} \qquad (nh=1).$$

To obtain this we write
$$Ih^{-\frac{3}{2}} + \alpha + \beta h + \gamma h^2 + \ldots = T_h/h^{\frac{3}{2}}$$
giving
$$\mathbf{v}_n I + \mathbf{u}_n \alpha + \mathbf{A}^*_{n,\,k-1}(\beta, \gamma, \delta, \ldots)^T = (T_{h_1} h_1^{-\frac{3}{2}}, T_{h_2} h_2^{-\frac{3}{2}}, \ldots)^T$$
with
$$\mathbf{v}_n = (h_1^{-\frac{3}{2}}, h_2^{-\frac{3}{2}}, \ldots)^T.$$

We thus have to apply the Neville process to $\mathbf{v}_n$ and $(T_{h_r} h_r^{-\frac{3}{2}})$. If in the Neville tableaux we denote the sth elements in the rth columns $(r = 0(1)k-1)$ by $a_{r,s}$ and $b_{r,s}$ respectively, we then proceed as follows. From two consecutive elements in any column we have

$$a_{r,s} I + \alpha = b_{r,s}$$
$$a_{r,s+1} I + \alpha = b_{r,s+1}$$

whence
$$I = \frac{b_{r,s+1} - b_{r,s}}{a_{r,s+1} - a_{r,s}}$$
$$\alpha = b_{r,s} - a_{r,s} I$$
with
$$b_{r,s+1} = \alpha + a_{r,s+1} I$$
as a check in hand calculation. Results are in table 10·4.

TABLE 10·4

h	$T_h/h^{\frac{3}{2}}$					
1	0					
$\frac{1}{2}$	0·70710 67812	1·41421 35624				
$\frac{1}{4}$	2·73205 08076	4·75699 48340	5·87125 52579			
$\frac{1}{5}$	3·97974 46120	8·97051 98296	11·77953 64933	13·25660 68022		
$\frac{1}{10}$	12·00498 95810	20·03023 45500	27·40337 76969	31·30933 79978	33·31519 70195	
$\frac{1}{20}$	34·70957 00740	57·41415 05670	69·87545 59060	80·49347 54583	85·95837 96206	88·72907 34417

h	$1/h^{\frac{3}{2}}$					
1	1					
$\frac{1}{2}$	2·82842 71248	4·65685 42496				
$\frac{1}{4}$	8	13·17157 28752	16·00981 24171			
$\frac{1}{5}$	11·18033 98875	23·90169 94375	31·05511 71457	34·81644 33278		
$\frac{1}{10}$	31·62277 66020	52·06521 33165	70·84088 92358	80·78733 22583	85·89520 88061	
$\frac{1}{20}$	89·44271 91000	147·26266 15980	178·99514 43585	206·03370 81392	219·94997 21260	227·00548 59849

Values of I and α

0·38672 95402				
−0·38672 95403				
0·39155 28361	0·39258 85773			
−0·40037 18812	−0·41401 42223			
0·39231 46106	0·39268 17611	0·39269 93399		
−0·40646 60769	−0·41524 15995	−0·41578 75100		
0·39257 77089	0·39269 65494	0·39269 92084	0·39269 91976	
−0·40940 76055	−0·41559 50653	−0·41578 34252	−0·41578 25556	
0·39267 73274	0·39269 87193	0·39269 90959	0·39269 90870	0·39269 90828
−0·41255 78213	−0·41570 80374	−0·41577 54518	−0·41577 36225	−0·41577 26958

Comments on table 10·4

The values of I agree with those in § 10·2; this is to be expected, since they result from elimination between the *same sets of values* though the terms eliminated are removed in a different order and agreement is rather better than might have been expected. For I, as seen in § 10·2, the final estimate is the best, with error about 10^{-9}. The final value is also best for α; using the known value $\frac{1}{8}\pi$ for I and subtracting this from T_h, we can then divide $T_h - \frac{1}{8}\pi$ by $h^{\frac{3}{2}}$ and use the results in an ordinary Neville reduction; by using nine values of h, $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{10}, \frac{1}{12}, \frac{1}{20})$ the best estimate obtained is $\alpha = 0{\cdot}41577\,2448+$. An estimate may also be obtained from approximations given in Fox (1967), but it seems difficult to obtain great precision in this way.

The single elements in the final columns are not needed; they are simply recombinations of the two in the previous column. In other words, the pairs $(a_{r,s}, b_{r,s})$ and $(a_{r,s+1}, b_{r,s+1})$ give results identical with those from $(a_{r,s+1}, b_{r,s+1})$ and $(a_{r+1,s}, b_{r+1,s})$.

It may be possible to reduce rounding-error—at the expense of increased truncation error—by taking the bottom element in a column and combining it, not with the element next above it, but with another higher in the same column. This needs investigating.

10·5. *Derivatives*

A first derivative is obtained in a manner precisely similar to the process in the previous § 10·4. The only difference is that $\mathbf{v}_n$ is now $(1/\xi_1, 1/\xi_2, 1/\xi_3, \ldots)^T$ and the final vector is (f_i/ξ_i). It does not seem necessary to give an example.

We may, however, remark that rounding errors are known to cause difficulty with derivatives, and it seems likely that separation of the final elements used and judicious choice of the ξ_i will help.

10·6. *Integration* (iii)

A final example in which the Romberg process is most effective is now given.

The mid-point rule approximations to the integral

$$I = \int_0^{0\cdot 8} \frac{\cosh x}{\sqrt{x}}\,\mathrm{d}x$$

have a series of correction terms of the form exhibited,

$$M_h = I + \alpha h^{\frac{1}{2}} + \beta h^2 + \gamma h^{\frac{5}{2}} + \delta h^4 + \epsilon h^{\frac{9}{2}} + \dots$$

the half-integer powers coming from the lower limit and the even integer powers from the upper limit of integration. We use $h = 0\cdot 8$, $0\cdot 4$, $0\cdot 2$, $0\cdot 1$, $0\cdot 05$, and $0\cdot 025$ and, to eliminate h^ρ for $\rho = \frac{1}{2}, 2, \frac{5}{2}, 4, \dots$ in turn from elements $M_h^{(r)}$ in column r (where $r = 0$ gives M_h) we use

$$M_h^{(r+1)} = \frac{2^\rho M_{\frac{1}{2}h}^{(r)} - M_h^{(r)}}{2^\rho - 1}.$$

The results are exhibited in table 10·6. The hyperbolic cosines were computed to 9 decimals; the value of the integral is 1·90678 3801 to this accuracy.

Table 10·6

h	M_h					
0·8	1·36746 04048					
		1·90378 77578				
0·4	1·52454 70496		1·90689 96471			
		1·90612 16748		1·90686 67943		
0·2	1·63630 76698		1·90687 26019		1·90678 65098	
		1·90668 48701		1·90679 15276		1·90678 37475
0·1	1·71549 93183		1·90680 58596		1·90678 38696	
		1·90677 56122		1·90678 43482		
0·05	1·77152 28477		1·90678 81509			
		1·90678 50162				
0·025	1·81114 02196					
2^ρ	$\sqrt{2}$	4	$4\sqrt{2}$	16	$16\sqrt{2}$	

Comment on table 10·6

Convergence is rapid, though a little erratic, indicating the presence of the now familiar truncation error, noticeable for large h, not representable by a power series of the type chosen.† The last result is correct to nearly 7 decimals.

With Romberg integration it needs as many integrand evaluations as have already been made to add another estimate. With Neville's process, one would have to eliminate all powers of $h^{\frac{1}{2}}$ to the desired limit. Eventually this would be possible with fewer integrands than for the Romberg process, but only for a fairly large number of points.

11. Lagrange multipliers

11·1.

It is possible to obtain individual results in the general elimination tableau by means of Lagrange multipliers or weights applied to the first column, as we have seen in § 5·2 for interpolation. These weights may be conveniently obtained by use of Neville's process

† If enough is known about this error, it might be possible to treat it as a 'rogue', as indicated in § 9·2.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

itself, applied to a set of initial vectors each consisting of zeros except for a single unit in each position in turn, i.e. we write $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_1 + \ldots + \mathbf{u}_n$ in which $\mathbf{u}_r = (0, 0, \ldots, 0, 1, 0, \ldots 0)^T$ with a 1 in the $(r+1)$th position and zeros elsewhere.

Each vector gives a triangular tableau, and results are obtained by taking one element from each tableau (all in corresponding positions) and multiplying by the appropriate value in the vector $\mathbf{f}$ of initial estimates, and then taking the sum.

It is convenient to collect into a single matrix the vectors of multipliers by which the vector $\mathbf{f}$ of initial estimates must be multiplied to give all the results in the Neville tableau as a single vector. For convenience, we take the transpose $\mathbf{f}^T$, a row vector of initial estimates, and postmultiply by a matrix of columns, which we shall denote by $\mathbf{L}^T$, comprising all columns of multipliers—excepting any that consists of a column of a unit matrix, or that is just a copy of a previous column—of the partitioned matrix $(\mathbf{L}_n^{(2)T}|\mathbf{L}_n^{(3)T}|\mathbf{L}_n^{(4)T}|\ldots)$. We go a little further than this in order to produce rational coefficients, by using integer numerators and a common denominator for each column in $\mathbf{L}^T$; if the row vector of denominators is denoted by $\mathbf{d}^T = (d_{r,s})^T$, then the matrix of numerators is given by $\mathbf{L}^T\mathbf{D}(\mathbf{d}) = (\mathbf{D}(\mathbf{d})\,\mathbf{L})^T$, which is tabulated, together with $\mathbf{d}^T$, and certain other vectors useful for error analysis. Thus results are given as a row vector by

$$\mathbf{f}^T(\mathbf{D}(\mathbf{d})\,\mathbf{L})^T\,\mathbf{D}^{-1}(\mathbf{d}),$$

or as a column vector by $$\mathbf{D}^{-1}(\mathbf{d})\ (\mathbf{D}(\mathbf{d})\,\mathbf{L}\,\mathbf{f})$$

and we note that $$\mathbf{DLu} = \mathbf{Du} = \mathbf{d}.$$

We observe also that the final vector of results contains none of the initial estimates (these can be appended at the start, if desired), and no repetition of estimates.

The matrix $\mathbf{L}^T$ thus consists of rows (transposed to columns) taken from the matrices $\mathbf{L}_n^{(r)}$ described in § 5·2, but omitting in the tables the unit matrix $\mathbf{L}_n^{(0)}$ which would give the first column $\mathbf{f}$ of initial estimates in the Neville tableau, and any repetitions. The tables are partitioned to indicate the separation of successive columns in the final Neville tableau; there are also extra columns in one or two cases, to allow for estimates, not appearing in the tableau, that may be useful on occasion, or which might appear in a smaller tableau by omitting some intermediate argument, e.g. with tables 11·3, 11·4, where $h = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{12}, \frac{1}{24})$, the result based on $h = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{6})$ would not appear in the Neville tableau, nor would the corresponding multipliers appear in the tables unless specially added. The result for $h = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{12})$ also fails to appear in the tableau, but the *multipliers* are the same as for $h = (\frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{12}, \frac{1}{24})$ in which each h is halved; these multipliers do appear in the tables.

11·2.

We illustrate with $h = (1, \frac{1}{2}, \frac{1}{5}, \frac{1}{10})$, i.e. $h = 1/n$, $n = 1, 2, 5, 10$ or $n|10$. There are four triangular tableaux.

h	$\mathbf{u}_0$			
1	1			
		-1		
$\frac{1}{2}$	0		$\frac{1}{4}$	
		0		$-\frac{1}{36}$
$\frac{1}{5}$	0		0	
		0		
$\frac{1}{10}$	0			

h	$\mathbf{u}_1$			
1	0			
		2		
$\frac{1}{2}$	1		$-\frac{4}{3}$	
		$-\frac{2}{3}$		$\frac{1}{3}$
$\frac{1}{5}$	0		$\frac{1}{6}$	
		0		
$\frac{1}{10}$	0			

h	$\mathbf{u}_2$			
1	0			
		0		
$\frac{1}{2}$	0		$\frac{25}{12}$	
		$\frac{5}{3}$		$-\frac{25}{12}$
$\frac{1}{5}$	1		$-\frac{5}{3}$	
		-1		
$\frac{1}{10}$	0			

h	$\mathbf{u}_3$			
1	0			
		0		
$\frac{1}{2}$	0		0	
		0		$\frac{25}{9}$
$\frac{1}{5}$	0		$\frac{5}{2}$	
		2		
$\frac{1}{10}$	1			

These are combined to give the matrix $(\mathbf{DL})^T$; in the illustrative table below the initial unit matrix is not omitted, $\mathbf{v}^{(1)}$ is the vector $((-1)^{n_r})$, $\mathbf{v}^{(2)}$ is (n_r), i.e. (1, 2, 5, 10).

	h										
$(\mathbf{DL})^T$	1	1	0	0	0	−1	0	0	3	0	−1
	$\frac{1}{2}$	0	1	0	0	2	−2	0	−16	1	12
	$\frac{1}{5}$	0	0	1	0	0	5	−1	25	−10	−75
	$\frac{1}{10}$	0	0	0	1	0	0	2	0	15	100
$(\mathbf{DLu})^T = \mathbf{d}^T$		1	1	1	1	1	3	1	12	6	36
$(\mathbf{DLv}^{(1)})^T = (\mathbf{d}^{(1)*})^T$		−1	1	−1	1	3	−7	3	−44	26	188
$(\mathbf{DLv}^{(2)})^T = (\mathbf{d}^{(2)*})^T$		1	2	5	10	3	21	15	96	102	648
$(\mathbf{D}\epsilon)^T$		1	1	1	1	3	7	3	44	26	188
$(\mathbf{D}\sigma)^T$		1	1	1	1	2·2	5·4	2·2	30	18	126

The last two lines give respectively the sums of the absolute values, and the square root of the sum of squares of the elements of each column of $(\mathbf{DL})^T$, for use in determining maximum and least-square rounding errors. We note that $(\mathbf{D}\epsilon) = (d_{r,s}\epsilon_{r,s}) = (d_{r,s}|e|_{rs})$ where $\mathbf{e} = (\mathbf{Lv}^{(4)})$ in which $\mathbf{v}^{(4)} = (1, -1, 1, -1)$ with alternating signs; in this table $\mathbf{v}^{(4)} = -\mathbf{v}^{(1)}$. In general, if $(\mathbf{Lv}^{(r)}) = \mathbf{d}^{(r)}$, we write $(\mathbf{DLv}^{(r)}) = (\mathbf{Dd}^{(r)}) = \mathbf{d}^{(r)*}$; $\mathbf{D} = \mathbf{D}(\mathbf{d})$ always.

11·3.

There are six tables of coefficients, tables 11·1 to 11·6, covering three typical sets of arguments h, for the elmination of powers of h, and of powers of h^2. They are given in order to help in the evaluation and checking of individual eliminations, and also in order to help with the discussion of rounding-errors, see § 12·4.

The tables are all concerned with extrapolation to zero for the variable h. Tables of ordinary Lagrange interpolation coefficients may be used in a similar way, but are too familiar to need inclusion here. The table in Lyness & McHugh (1963) gives the first column in each block of table 11·2, and is extended to $n = 10$; this paper and table gave the initial impulse for the investigation given in the present work—the usefulness of including the crudest estimate was doubted, a way was sought to avoid its use (Neville's process provided the means) and, to the writer's surprise, the usefulness of the crude estimate largely confirmed.

In all tables

$$\mathbf{u} = (1, 1, 1, \ldots)^T = (1)$$

$$\mathbf{v}^{(1)} = ((-1)^{n_0}, (-1)^{n_1}, \ldots)^T = ((-1)^{n_r}),\ h_r = 1/n_r$$

$$\mathbf{v}^{(2)} = (n_0, n_1, \ldots)^T = (n_r)$$

$$\mathbf{v}^{(3)} = (1/n_0, 1/n_1, \ldots)^T = (1/n_r)$$

$$\mathbf{v}^{(4)} = (1, -1, 1, -1, \ldots)^T = ((-1)^r), \text{ strictly alternating}$$

$$\mathbf{DLu} = \mathbf{d},\quad \mathbf{Lv}^{(r)} = \mathbf{d}^{(r)},\quad \mathbf{DLv}^{(r)} = \mathbf{Dd}^{(r)} = \mathbf{d}^{(r)*}\quad (r = 1, 2, 3)$$

$$\mathbf{Lv}^{(4)} = \mathbf{e},\quad \mathbf{DLv}^{(4)} = \mathbf{De},\quad \epsilon = (\epsilon_{r,s}) = (|e_{r,s}|).$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

TABLE 11·1

$h = 1/n$, $n = 1(1)8$. Powers of h eliminated.

	n																		
$(\mathbf{DL})^T$	1	−1	.	.	.	.	.	.	+1	.	.	.	.	.	−1	.	.	.	.
	2	+2	−2	.	.	.	.	.	−8	+2	.	.	.	.	+24	−8	.	.	.
	3	.	+3	−3	.	.	.	.	+9	−9	+9	.	.	.	−81	+81	−9	.	.
	4	.	.	+4	−4	.	.	.	.	+8	−32	+8	.	.	+64	−192	+64	−64	.
	5	.	.	.	+5	−5	.	.	.	.	+25	−25	+25	.	.	+125	−125	+375	−125
	6	.	.	.	.	+6	−6	.	.	.	.	+18	−72	+18	.	.	+72	−648	+648
	7	.	.	.	.	.	+7	−7	.	.	.	.	+49	−49	.	.	.	+343	−1029
	8	.	.	.	.	.	.	+8	.	.	.	.	.	+32	.	.	.	.	+512
$\mathbf{d}^T$		1	1	1	1	1	1	1	2	1	2	1	2	1	6	6	2	6	6
$(\mathbf{d}^{(1)*})^T$		+3	−5	+7	−9	+11	−13	+15	−18	+19	−66	+51	−146	+99	+170	−406	+270	−1430	+2314
$(\mathbf{d}^{(2)*})^T$		3	5	7	9	11	13	15	12	9	24	15	36	21	60	84	36	132	156
$\boldsymbol{\epsilon}^T$		3	5	7	9	11	13	15	9	19	33	51	73	99	28	68	135	240	390
$\boldsymbol{\sigma}^T$		2·2	3·6	5·0	6·4	7·8	9·2	11	6·0	12	21	32	45	61	18	40	79	140	220

	n										
$(\mathbf{DL})^T$	1	+1	.	.	.	−1	.	.	+1	.	−1
	2	−64	+4	.	.	+160	−32	.	−384	+32	+896
	3	+486	−81	+81	.	−2430	+1215	−243	+10935	−2187	−45927
	4	−1024	+384	−1024	+64	+10240	−10240	+5120	−81920	+30720	+5 73440
	5	+625	−625	+3750	−625	−15625	+31250	−31250	+2 34375	−1 56250	−27 34375
	6	.	+324	−5184	+1944	+7776	−38880	+77760	−2 79936	+3 49920	+58 78656
	7	.	.	+2401	−2401	.	+16807	−84035	+1 17649	−3 52947	−57 64801
	8	.	.	.	+1024	.	.	+32768	.	+1 31072	+20 97152
$\mathbf{d}^T$		24	6	24	6	120	120	120	720	360	5040
$(\mathbf{d}^{(1)*})^T$		−2200	+1418	−12440	+6058	+36232	−98424	+231176	+7 25200	+10 23128	+170 95248
$(\mathbf{d}^{(2)*})^T$		360	120	600	180	2520	3240	3960	20160	12600	1 81440
$\boldsymbol{\epsilon}^T$		92	240	520	1010	300	820	1930	1010	2800	3400
$\boldsymbol{\sigma}^T$		54	134	290	550	170	450	1030	550	1500	1800

TABLE 11·2

$h = 1/n$, $n = 1(1)8$. Powers of h^2 eliminated.

n														
$(DL)^T$	1	−1	.	.	.	.	.	.	+5	.	.	.	.	.
	2	+4	−4	.	.	.	.	.	−128	+28	.	.	.	.
	3	.	+9	−9	.	.	.	.	+243	−243	+729	.	.	.
	4	.	.	+16	−16	.	.	.	.	+320	−4096	+704	.	.
	5	.	.	.	+25	−25	.	.	.	.	+4375	−3125	+8125	.
	6	.	.	.	.	+36	−36	.	.	.	.	+2916	−31104	+4860
	7	.	.	.	.	.	+49	−49	.	.	.	.	+26411	−16807
	8	.	.	.	.	.	.	+64	.	.	.	.	.	+13312
$\mathbf{d}^T$		3	5	7	9	11	13	15	120	105	1008	495	3432	1365
$(\mathbf{d}^{(1)*})^T$		+5	−13	+25	−41	+61	−85	+113	−376	+591	−9200	+6745	−65640	+34979
$(\mathbf{d}^{(2)*})^T$		7	19	37	61	91	127	169	478	607	7678	4687	38878	18007
$(\mathbf{d}^{(3)*})^T$		1	1	1	1	1	1	1	22	13	94	37	214	73
ϵ^T		1·7	2·6	3·6	4·6	5·5	6·5	7·5	3·1	5·6	9·1	13·6	19	26
σ^T		1·4	2·0	2·6	3·3	4·0	4·7	5·4	2·3	3·8	6·0	8·8	12	16

n										
$(DL)^T$	1	−7	.	.	.	.	+42	.	.	.
	2	+896	−768	.	.	.	−24576	+1056	.	.
	3	−6561	+19683	−13365	.	.	+5 31441	−72171	+9 38223	.
	4	+8192	−81920	+1 80224	−1 06496	.	−20 97152	+7 20896	−272 62976	+7 45472
	5	.	+78125	−5 46875	+10 15625	−5 46875	+19 53125	−19 53125	+1777 34375	−136 71875
	6	.	.	+4 35456	−25 19424	+41 99040	.	+14 69664	−3762 33984	+680 24448
	7	.	.	.	+17 64735	−90 58973	.	.	+2421 21642	−1210 60821
	8	.	.	.	.	+57 67168	.	.	.	+692 06016
$\mathbf{d}^T$		2520	15120	55440	1 54440	3 60360	3 62880	1 66320	172 97280	32 43240
$(\mathbf{d}^{(1)*})^T$		+15656	−180496	+11 75920	−54 06280	+195 72056	−46 06336	+42 16912	−8242 91200	+2727 08632
$(\mathbf{d}^{(2)*})^T$		14870	120458	5 59162	18 88742	51 84398	29 22230	17 21542	2198 82230	489 91582
$(\mathbf{d}^{(3)*})^T$		302	1322	3802	8702	17222	31238	11014	9 27014	1 45750
ϵ^T		6·2	11·9	21	35	54	12·7	25	48	84
σ^T		4·2	7·6	13	21	32	8·0	15	28	48

n							
$(DL)^T$	1	−66	.	.	+429	.	−715
	2	+1 68960	−1 46432	.	−46 85824	+1 46432	+328 00768
	3	−97 43085	+253 32021	−23 02911	+6839 64567	−621 78597	−1 17517 54833
	4	+922 74688	−5452 59520	+1363 14880	−1 39586 43712	+27262 97600	+48 85525 29920
	5	−2441 40625	+31738 28125	−17089 84375	+7 93457 03125	−3 05175 78125	−534 05761 71875
	6	+1813 98528	−63489 48480	+72559 41120	−15 67283 28192	+12 24440 06400	+2115 83243 05920
	7	.	+39546 53486	−118639 60458	+9 68890 10407	−19 37780 20814	−3323 29305 69601
	8	.	.	+64424 50944	.	+10 30792 15104	+1759 21860 44416
$\mathbf{d}^T$		199 58400	2594 59200	2594 59200	62270 20800	38918 88000	65 38371 84000
$(\mathbf{d}^{(1)*})^T$		+5277 25952	−140481 68064	−274099 54688	−34 76103 36256	+45 26074 43072	+7782 43541 78048
$(\mathbf{d}^{(2)*})^T$		2078 95394	33526 89266	40209 59978	8 07895 66956	6 09826 70382	1027 30598 00268
$(\mathbf{d}^{(3)*})^T$		13 10354	136 13954	112 92602	3250 24572	1670 67678	2 79711 76092
ϵ^T		26	54	106	56	116	119
σ^T		16	31	59	32	65	66

TABLE 11·3

$h = 1/n$, $n|24$, i.e. $n = 1, 2, 3, 4, 6, 8, 12, 24$. Powers of h eliminated.

	n							
$(\mathbf{DL})^T$	1	−1	.	.	.	.	.	.
	2	+2	−2	.	.	.	.	.
	3	.	+3	−3	.	.	.	.
	4	.	.	+4	−2	.	.	.
	6	.	.	.	+3	−3	.	.
	8	.	.	.	.	+4	−2	.
	12	.	.	.	.	.	+3	−1
	24	.	.	.	.	.	.	+2
$\mathbf{d}^T$		1	1	1	1	1	1	1
$(\mathbf{d}^{(1)*})^T$		+3	−5	+7	+1	+1	+1	+1
$(\mathbf{d}^{(2)*})^T$		3	5	7	10	14	20	36
$\boldsymbol{\epsilon}^T$		3	5	7	5	7	5	3
$\boldsymbol{\sigma}^T$		2·2	3·6	5·0	3·6	5·0	3·6	2·2

	n						
$(\mathbf{DL})^T$	1	+1	.	.	.	.	.
	2	−8	+2	.	.	.	.
	3	+9	−9	+3	.	.	.
	4	.	+8	−8	+2	.	.
	6	.	.	+6	−9	+3	.
	8	.	.	.	+8	−8	+1
	12	.	.	.	.	+6	−3
	24	.	.	.	.	.	+3
$\mathbf{d}^T$		2	1	1	1	1	1
$(\mathbf{d}^{(1)*})^T$		−18	+19	−5	+1	+1	+1
$(\mathbf{d}^{(2)*})^T$		12	9	13	18	26	44
$\boldsymbol{\epsilon}^T$		9	19	17	19	17	7
$\boldsymbol{\sigma}^T$		6·0	12	10	12	10	4·4

	n					
$(\mathbf{DL})^T$	1	−1	.	.	.	.
	2	+24	−1	.	.	.
	3	−81	+9	−9	.	.
	4	+64	−16	+40	−1	.
	6	.	+9	−90	+9	−1
	8	.	.	+64	−16	+4
	12	.	.	.	+9	−6
	24	.	.	.	.	+4
$\mathbf{d}^T$		6	1	5	1	1
$(\mathbf{d}^{(1)*})^T$		+170	−17	+23	+1	+1
$(\mathbf{d}^{(2)*})^T$		60	15	105	30	50
$\boldsymbol{\epsilon}^T$		28	35	41	35	15
$\boldsymbol{\sigma}^T$		18	20	24	20	8

	n				
$(\mathbf{DL})^T$	1	+1	.	.	.
	2	−60	+5	.	.
	3	+405	−81	+3	.
	4	−640	+240	−20	+1
	6	+324	−405	+90	−15
	8	.	+256	−128	+40
	12	.	.	+60	−45
	24	.	.	.	+24
$\mathbf{d}^T$		30	15	5	5
$(\mathbf{d}^{(1)*})^T$		−782	+177	−1	+5
$(\mathbf{d}^{(2)*})^T$		480	345	165	270
$\boldsymbol{\epsilon}^T$		48	66	60	25
$\boldsymbol{\sigma}^T$		28	36	34	13

	n								
$(\mathbf{DL})^T$	1	−1	.	.	+1	.	−1	−1	−1
	2	+140	−1	.	−308	+7	+644	+20	+132
	3	−1701	+27	−3	+6237	−297	−20493	−45	−1485
	4	+4480	−120	+28	−24640	+1848	+1 13344	.	+3520
	6	−6804	+405	−210	+74844	−10395	−5 73804	+36	−3564
	8	+4096	−512	+448	−90112	+19712	+10 36288	.	.
	12	.	+216	−420	+36288	−16632	−8 34624	.	+1728
	24	.	.	+192	.	+6912	+3 31776	.	.
$\mathbf{d}^T$		210	15	35	2310	1155	53130	10	330
$(\mathbf{d}^{(1)*})^T$		+3614	−39	+41	−10166	+1749	+94118	+102	+3302
$(\mathbf{d}^{(2)*})^T$		5040	525	1995	83160	68145	31 87800	120	9240
$\boldsymbol{\epsilon}^T$		82	85	37	101	48	55	10·2	32
$\boldsymbol{\sigma}^T$		44	47	19	54	25	28	6·1	17

TABLE 11·4

$h = 1/n$, $n|24$, i.e. $n = 1, 2, 3, 4, 6, 8, 12, 24$. Powers of h^2 eliminated.

n																							
$(DL)^T$	1	−1	.	.	.	.	.	.	+5	.	.	.	.	.	−7	.	.	.	.	+1	.	.	.
	2	+4	−4	.	.	.	.	.	−128	+28	.	.	.	.	+896	−7	.	.	.	−560	+77	.	.
	3	.	+9	−9	.	.	.	.	+243	−243	+15	.	.	.	−6561	+162	−81	.	.	+10935	−4374	+27	.
	4	.	.	+16	−4	.	.	.	.	+320	−64	+28	.	.	+8192	−512	+704	−7	.	−32768	+28160	−440	+1
	6	.	.	.	+9	−9	.	.	.	.	+84	−243	+15	.	.	+567	−3564	+162	−1	+34992	−1 20285	+5940	−54
	8	.	.	.	.	+16	−4	.	.	.	.	+320	−64	+1	.	.	+4096	−512	+8	.	+1 31072	−16384	+320
	12	.	.	.	.	.	+9	−1	.	.	.	.	+84	−6	.	.	.	+567	−28	.	.	+16632	−945
	24	.	.	.	.	.	.	+4	.	.	.	.	.	+15	.	.	.	.	+56	.	.	.	+1728
d^T		3	5	7	5	7	5	3	120	105	35	105	35	10	2520	210	1155	210	35	12600	34650	5775	1050
$(d^{(1)*})^T$		+5	−13	+25	+5	+7	+5	+3	−376	+591	+5	+105	+35	+10	+15656	−114	+1317	+210	+35	−9272	+43398	+5721	+1050
$(d^{(2)*})^T$		7	19	37	38	74	76	84	478	607	293	1214	586	296	14870	1826	13957	3652	1066	110566	4 26538	102473	32372
$(d^{(3)*})^T$		1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{12}$	22	13	3	$\frac{13}{2}$	$\frac{3}{2}$	$\frac{1}{4}$	302	17	67	$\frac{17}{2}$	$\frac{5}{6}$	1006	1957	227	$\frac{49}{2}$
ϵ^T		1·7	2·6	3·6	2·6	3·6	2·6	1·7	3·1	5·6	4·7	5·6	4·7	2·2	6·2	5·9	7·3	5·9	2·7	6·3	8·2	6·8	2·9
σ^T		1·4	2·0	2·6	2·0	2·6	2·0	1·4	2·3	3·8	3·0	3·8	3·0	1·6	4·2	3·7	4·7	3·7	1·8	3·9	5·2	4·2	1·9

n										
$(DL)^T$	1	−11	.	.	+5	.	−1	−1	−1
	2	+25872	−11	.	−48048	+7	+38640	+112	+2288
	3	−12 40029	+1458	−3	+53 73459	−2106	−98 08695	−567	−1 04247
	4	+75 69408	−17600	+88	−615 01440	+45760	+2020 76160	.	+5 85728
	6	−311 77872	+2 00475	−2772	+6755 20560	−12 16215	−51789 90960	+1296	−16 67952
	8	+335 54432	−5 24288	+14336	−17448 30464	+59 63776	+2 50819 37920	.	.
	12	.	+5 13216	−38808	+16930 52928	−155 67552	−6 49003 62240	.	+29 85984
	24	.	.	+67584	.	+265 42080	+11 00753 14176	.	.
d^T		87 31800	1 73250	40425	5675 67000	157 65750	6 52702 05000	840	18 01800
$(d^{(1)*})^T$		+112 11880	+1 70334	+40431	+5568 20072	+157 69962	+6 52898 22392	+1976	+20 10296
$(d^{(2)*})^T$		1079 77502	31 01090	12 54719	1 01811 33310	4907 88950	203 33637 06778	6298	278 58842
$(d^{(3)*})^T$		4 89926	6725	933	219 62470	3 62135	14974 53778	82	83666
ϵ^T		8·4	7·3	3·1	7·4	3·1	3·1	2·4	3·0
σ^T		5·3	4·4	2·0	4·4	2·0	2·0	1·7	1·9

TABLE 11·5

$h = 2^{-r}$, $r = 0(1)8$; Romberg. Powers of h eliminated.

	2^r									
$(\mathbf{DL})^T$	1	1	−1	+1	−1	+1	−1	+1	−1	+1
	2	.	+2	−6	+14	−30	+62	−126	+254	−510
	4	.	.	+8	−56	+280	−1240	+5208	−21336	+86360
	8	.	.	.	+64	−960	+9920	−89280	+7 55904	−62 17920
	16	.	.	.	.	+1024	−31744	+6 66624	−120 94464	+2056 05888
	32	.	.	.	.	.	+32768	−20 64384	+873 92256	−31835 75040
	64	.	.	.	.	.	.	+20 97152	−2663 38304	+2 26387 55840
	128	.	.	.	.	.	.	.	+2684 35456	−6 84510 41280
	256	.	.	.	.	.	.	.	.	+6 87194 76736
$\mathbf{d}^T$		1	1	3	21	315	9765	6 15195	781 29765	1 99230 90075
$(\mathbf{d}^{(2)*})^T$		1	3	21	315	9765	6 15195	781 29765	199230 90075	1018 06990 28325
ϵ^T		1	3	5	6·4	7·3	7·8	8·0	8·1	8·2
σ^T		1	2·2	3·3	4·1	4·5	4·8	4·9	5·0	5·0

TABLE 11·6

$h = 2^{-r}$, $r = 0(1)6$; Romberg. Powers of h^2 eliminated.

	2^r							
$(\mathbf{DL})^T$	1	1	−1	+1	−1	+1	−1	+1
	2	.	+4	−20	+84	−340	+1364	−5460
	4	.	.	+64	−1344	+22848	−3 71008	+59 57952
	8	.	.	.	+4096	−3 48160	+237 44512	−15433 93280
	16	.	.	.	.	+10 48576	−3575 64416	+9 76150 85568
	32	.	.	.	.	.	+10737 41824	−146 56575 89760
	64	.	.	.	.	.	.	+439 80465 11104
$\mathbf{d}^T$		1	3	45	2835	7 22925	7395 52275	302 84665 66125
$(\mathbf{d}^{(2)*})^T$		1	7	217	27559	140 82649	2 88271 82503	23612 34518 82073
$(\mathbf{d}^{(3)*})^T$		1	1	7	217	27559	140 82649	2 88271 82503
ϵ^T		1	1·7	1·9	1·9	2·0	2·0	2·0
σ^T		1	1·4	1·5	1·5	1·5	1·5	1·5

12. ERRORS

Comments on errors have been made on several of the numerical examples and, in fact, the whole process has been described in terms of error elimination. Some connected general remarks now seem helpful.

We note that, in general, there are four kinds of error with which we have to deal.

(i) *The eliminable error* that it is our aim to remove. The form of this error must be known or assumed. It is usually supposed to consist of a power series in some variable—ξ, h, $1/n$, $h^{\frac{1}{2}}$, etc. It may, however, contain more complicated terms, known or unknown. See, for example, Fox (1967) where terms of type $h \ln h$ or $h^{\frac{1}{2}} \ln h$ are shown to occur.

(ii) The remaining *uneliminated error*, which includes further terms of the series of eliminable error, but may also include those functions that cannot be expressed as a power series, and so cannot be eliminated by Neville's process. Such extra functions often affect, for example, approximations to an integral using large interval, h, or to sums using a very small number, n, of terms. They often decrease very rapidly as h or $1/n$ decreases, and show up as irregularities in the first few lines of the Neville tableau; they can then be avoided by use of sufficiently small h or $1/n$, but it must be noted that a major consideration is to extract all possible information from all estimates, particularly those that are easiest to obtain. We elaborate this discussion below in § 12·3.

(iii) *Rounding-errors*. These can affect end-figures quite severely, and merit study. This effect is seen in several of the examples, and is discussed below, where it will be seen that the Neville process itself provides a satisfactory method of study.

(iv) *Blunders*; to be found and eradicated.

We consider these types of error in turn.

12·1. *Eliminable error*

This is, as stated, normally supposed to be a power series in some convenient variable. We shall not discuss here exactly how this is to be determined; Fox (1967) gives a number of suggestions for quadrature, and experience has also to be called into play. Again, considerable further study is needed in order to determine *from the Neville reduction tableau itself*, whether an alternative reduction variable or method might not be more effective.

We repeat here the *Neville's process* results in the elimination of a power series in a single variable, term by term, using *all* integer powers starting from the first. A zero coefficient is no help! Arguments may be arbitrary.

Romberg's process, on the other hand, removes an *arbitrary set* of powers, but requires arguments in geometric progression.

Finally, we mention that straightforward *triangularization* of the matrix of eliminable terms can also be carried out quite generally, but with great increase in the labour involved, if this turns out to be the only way.

Progress may be determined fairly simply by careful consideration of agreement and disagreement of intermediate results, allowing for the effects of rounding-error, discussed in § 12·4 below.

12·2. *The remainder term*

This is an expression for the uneliminated part of the error, excluding rounding-error or blunder, in any item in the tableau.

For an interpolation process, the ordinary Lagrange remainder term (2·2) or the remainder series in (3·41) are applicable. The former, in particular, gives a neat one term form for $R_n(x)$, provided that corresponding derivatives exist.

For many of the later applications, however, there seems no obvious simple function $f(x)$ that can be regarded as an interpoland, so that conversion from the series (3·41) to the one-term form of (2·2) is less obviously possible. Nevertheless, we shall suppose that the *behaviour* of R_n, as exhibited when the form (2·2) is possible, is also applicable to the error term in many other cases. This suggests, for instance, that 'extrapolation to zero' from arguments all of one sign will, *in general*, produce results approaching from one side of the true value (corresponding to persistence of sign in $f^{(n+1)}(\eta)$ for successive arguments of a set—not invariable, but also not uncommon). It is this remark which suggests that, when oscillation in sign of the error occurs only for large h (in an integral) or for small n (in a sum), this is due to an uneliminable term (i.e. not included in the power series) with oscillating but rapidly diminishing effect; this can be demonstrated in some cases. The uneliminable error need not, of course, oscillate in sign.

We have also seen, in § 8·2, how to deal with an error that alternates in sign, with numerical value decreasing regularly. To do this we have converted the main error series into a monotonic form, by weighting the vector of initial elements suitably. The comments of the previous paragraph then apply.

Other cases of regular, though not monotonic, behaviour of R_n can occur; if this error is known well enough it may be possible to eliminate it by a more general and elaborate reduction of the matrix of error coefficients. This can be a very laborious and time-consuming process.

12·3. *Uneliminable error and blunders*

We have just referred to that part of the error not represented by the power series, and of unknown form. The main points to remember are that it often exists, that it shows up in approximations most 'remote' from the desired result (i.e. with large h or small n) and that we need recognize it only in order to avoid relying too much on the corresponding approximations. The last remark applies to blunders also; these, however, may occur anywhere in the tableau, and should be removed if they can be suitably disentangled from other errors for identification.

A specific numerical illustration may be helpful here. For the integral $\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2}$ used in §§ 4·4 and 10·1, the estimates T_h contain errors of which part is expressible as a power series in (odd) powers of h, and part not expressible in this way. In this case, if we evaluate the infinite sums for each fixed h, the part expressible as a power series in h is completely removed for each h, but the more difficult part of the error still remains. In fact, we find

$$T_h(\infty) = h \sum_{-\infty}^{\infty} \frac{1}{1+n^2h^2}$$

as given in the table below.

h	$T_h(\infty)$	$T_h(\infty)-\pi$
2	3·42537 71499	0·28378 44963
1	3·15334 80949	0·01175 54413
$\frac{1}{2}$	3·14161 45653	2 19117
$\frac{1}{3}$	3·14159 26945	409
$\frac{1}{4}$	3·14159 26536	0

It can, in fact, be shown that

$$h \sum_{-\infty}^{\infty} \frac{1}{1+n^2h^2} = \pi \frac{e^{2\pi/h}+1}{e^{2\pi/h}-1}$$

so that the error is $O(e^{-2\pi/h})$ and diminishes rapidly as $h \to 0$, much faster than any power of h. Thus estimates for large h have limited usefulness though they may nevertheless help to improve final results, so long as the uneliminable error does not dominate the part that can still be eliminated.

Thus, if we note any irregularity in the tableau, we treat it as a blunder and attempt to find and correct it; we may, however, expect a failure to correct the irregularity if it is on or near the top line of the tableau as set out in the examples.

As a rule it is useful to keep the top line, even if it is only to exhibit the irregularities and to detect when they cease.

It is perhaps appropriate to say here that considerable effort has been spent on making as sure as is reasonably possible that blunders have all been removed from the numerical examples in this paper. However, since correct rounding is not vital to the demonstration of the various properties that the examples are intended to exhibit, the final digits have not been exhaustively checked, and trivial departures from 'best' rounding have not necessarily been 'corrected'.

12·4. *Rounding-errors: derivation of bounds*

Rounding-errors have considerable effect on the various results in the Neville tableau, and a convenient means of study is provided by the Neville–Romberg process itself, applied to the vectors of a unit matrix. This yields the tables of Lagrange multipliers already considered in § 11; multipliers in any individual column may then be combined to yield maximum or mean-square errors as desired.

We shall assume at this stage that *each initial estimate is subject to an independent rounding-error* (see § 12·9 for a discussion of other possibilities) and, for the moment, that each is subject to the same maximum error ϵ, or mean least-square-error σ; we shall also suppose that, by use of guard figures, the further rounding-errors made during the Neville reduction are relatively negligible. The resulting maximum error multiple $|e_{r,s}|$ of ϵ at any stage of the Neville reduction is then given by the sum of the magnitudes of the corresponding Lagrange multipliers, $e_{r,s} = \sum\limits_t |l_{r,s,t}|$, where $l_{r,s,t}$ $(t = 0, 1, \ldots)$ are the multipliers yielding the sth element in the rth column to the right of the original estimates in the Neville tableau. The vector $\boldsymbol{\epsilon}^T = (|e_{r,s}|)^T$ is given in the tables of § 11. The vector $\boldsymbol{\sigma}^T = (\sigma_{r,s})^T$ is also given, where $\sigma^2_{r,s} = \sum\limits_t l^2_{r,s,t}$.

The *signed* multipliers $e_{r,s}$ can be developed collectively by an application of Neville's process to a single vector consisting of alternate positive and negative units. To see this, consider the general Lagrange multiplier for the point (x_s, f_s) from the set

$$(x_r, f_r), \quad r = 0(1)m.$$

It is, for the argument $x = X$,

$$L(s; 0(1)m) = \frac{(X-x_0)\ (X-x_1)\ldots(X-x_{s-1})\,.\,(X-x_{s+1})\ldots(X-x_m)}{(x_s-x_0)\ (x_s-x_1)\ldots(x_s-x_{s-1})\,.\,(x_s-x_{s+1})\ldots(x_s-x_m)}.$$

Now, this coefficient is symmetric in the arguments so that the order of use in derivation when the Neville process is applied is not relevant. We shall therefore consider only the arrangement in which $x_r < x_s$ when $r < s$. It is then clear that the sign of $L(s; 0(1)m)$ is $(-1)^s$ when $X = 0$, all $x_r > 0$, as is the case with the examples of this paper. The signs thus *alternate strictly* for $s = 0(1)m$. This applies to the non-zero elements of every column in $\mathbf{L}^T$ in the tables of § 11. If then, we start with the vector $\mathbf{v}^{(4)} = (1, -1, 1, -1, \ldots)$ with strictly alternating signs, and apply the Neville process for a set of positive arguments in ascending (or descending) order, we obtain a tableau of signed elements $e_{r,s}$ such that $\epsilon_{r,s} = |e_{r,s}|$ is the maximum possible multiplier of ϵ for the corresponding entry in a Neville tableau obtained from initial estimates having the same set of arguments, each estimate subject independently to the same error-bound ϵ.

It is further evident that the signs of $e_{r,s}$ do not depend in any way on the equality in magnitude of the error bounds of the initial elements, that is, on the sizes of the elements in $\mathbf{v}^{(4)}$, but only on the signs; we thus have the result

Maximum error bounds in a Neville tableau for 'extrapolation to zero' with original estimates having arguments of one sign arranged in numerical order, and with independent maximum error bounds, may be obtained by applying the Neville process to the vector of error bounds for the original estimates used, taken with strictly alternating signs, and taking the magnitudes of the final results.

It is also readily seen that:

If the arguments used are of both signs, still arranged in order (including sign), then the vector to which the Neville process must be applied to yield error bounds consists of the vector of individual bounds with signs alternating strictly except that the two bounds for the arguments nearest zero, one on each side, must have the same sign.

In both these statements we must remember that we have assumed that rounding-errors made during the Neville reduction are relatively negligible. It is, however, quite easy to allow for these by increasing the magnitude of each element in the error-bound tableau, as it is formed, by the amount of the corresponding rounding-error bound for that step in the reduction. We do not consider this possibility further.

No such simple method has been found to produce the mean square errors $\sigma_{r,s}$, but it is clear from the few examples given in the tables of § 11, that $\sigma_{r,s}$ is normally a moderate submultiple of $\epsilon_{r,s}$, not too difficult to estimate; nowhere in the tables does the ratio appear to exceed 2.

12·5. *Discussion of rounding-errors*

Table 11·1 shows that use of $h = 1/n$ with consecutive values of n yields large rounding-error effects when eliminating powers of h, particularly when n is large, and in later columns in the tableau, the maximum final error is about 3400 times an individual original error. The arguments of table 11·3 give better results, particularly in the final column where the error is only 55 times an original error; the maximum is about 100. The Romberg arrangement of table 11·5 is better again with an error factor of about 8 at most.

For elimination of h^2, tables 11·2, 11·4, 11·6 show similar relative error effects, but with smaller multipliers, respectively about 120, 8 and 2. For error effects with vectors $\mathbf{v}^{(i)}$, the elements of $(\mathbf{D}\boldsymbol{\epsilon})^T$ must be divided by the corresponding elements of the vector

$$(\mathbf{d}^{(i)*})^T = (\mathbf{DLv}^{(i)})^T;$$

this is readily done and is illustrated by examples below.

12·6.

In table 10·1, the maximum error-multiples in the tableau may be obtained from table 11·4, or by direct application to $\mathbf{v}^{(4)} = (1, -1, 1, -1, 1, -1)^T$ and suppression of signs. The result is the table of error-multiples on the left below, which apply to the elements in the reduction of hT_h in the upper tableau of table 10·1. After division by elements of the $\mathbf{v}$ tableau, the error-multiples (to the nearest integer) of the final ratios are given on the right (including those for T_h, corresponding to equal error maxima in hT_h).

h						
1	1					
		$\frac{5}{3}$				
$\frac{1}{2}$	1		$\frac{47}{15}$			
		$\frac{13}{5}$		$\frac{1957}{315}$		
$\frac{1}{3}$	1		$\frac{197}{35}$		$\frac{9907}{1575}$	
		$\frac{25}{7}$		$\frac{208}{35}$		$\frac{9195953}{1091475}$
$\frac{1}{4}$	1		$\frac{163}{35}$		$\frac{15776}{1925}$	
		$\frac{13}{5}$		$\frac{563}{77}$		
$\frac{1}{6}$	1		$\frac{197}{35}$			
		$\frac{25}{7}$				
$\frac{1}{8}$	1					

h						
1	1					
		5				
$\frac{1}{2}$	2		17			
		13		52		
$\frac{1}{3}$	3		45		79	
		25		73		150
$\frac{1}{4}$	4		54		145	
		26		126		
$\frac{1}{6}$	6		91			
		50				
$\frac{1}{8}$	8					

The final 3·14159 30 thus may have a maximum error of 150ϵ or 8 units in the 7th decimal, if $\epsilon = \frac{1}{2} \times 10^{-8}$; the truncation error should be small, so that about 6 decimal accuracy is attained, as indicated. In fact, errors are considerably smaller than maximum here, even if we disregard extra errors in computing the tableau.

12·7.

In table 10·3, we must compare maximum error multiples given by operation of $\mathbf{L}$ on the vector $\mathbf{v}^{(4)} = (1, -1, 1, -1, \ldots)^T$ with those given by the denominators $(\mathbf{L}\mathbf{v}^{(1)})$. The ratios, omitting signs are

n							
1	1						
		1					
2	1		1				
		1		1			
3	1		1		1		
		1		1		1	
4	1		1		1		1
		1		1		1	
5	1		1		1		1
		1		1		1·4	
6	1		1		1		1·3
		1		1		3	
7	1		1		1·5		3
		1		2		26	
8	1		3		4		14
		5		7		93	
10	1		14		711		76
		6		35			
12	1		19		84		
		7		54			
14	1		26				
		8					
16	1						

This shows the good effect on rounding-error of the alternation of signs in $\mathbf{v}^{(1)}$, when present, and the deterioration when it is abandoned. The effect of a very small divisor on error is apparent, and clearly very local. There is also the adverse effect of *truncation error* in the final ratios near the top of the table in § 10·3 when n is small, but this can occur independently of the signs in $\mathbf{v}^{(1)}$. Best results are thus obtained on lines, sloping *upward* near $n = 8$ or 10, in later columns; obvious convergence and small rounding error. However, rounding-errors tend to persist in sign and size along such upward-sloping lines; comparison between successive lines is needed to estimate rounding-error.

12·8.

Finally, we consider the example of § 10·4. The tableau for the vector $\boldsymbol{\epsilon}$ is given below, for $h = 1/n$, $n|20$, headed $\epsilon_{r,s}$

n	$\epsilon_{r,s}$					
1	1					
		3				
2	1		5			
		3		22·5		
4	1		17		22·5	
		9		18		$\frac{525}{38} = 13·9$
5	1		11		12	
		3		9		
10	1		5			
		3				
20	1					

n	$\epsilon(I)$				
1					
	1·09				
2		0·70			
	0·39		1·46		
4		1·12		0·88	
	0·63		0·70		0·26
5		0·43		0·22	
	0·10		0·15		
10		0·063			
	0·035				
20					

n	$\epsilon(\alpha)$				
1					
	0·4				
2		2·2			
	0·7		11		
4		6·5		21	
	2·1		15		17
5		7·4		14	
	0·8		8		
10		2·4			
	0·8				
20					

Now, from §10·4, we have

$$I = \frac{b_{r,s+1}-b_{r,s}}{a_{r,s+1}-a_{r,s}}, \quad \alpha = b_{r,s}-a_{r,s}I = \frac{b_{r,s}a_{r,s+1}-b_{r,s+1}a_{r,s}}{a_{r,s+1}-a_{r,s}}$$

whence maximum rounding errors are readily found to be

$$\epsilon_{r,s}(I) = \frac{\epsilon_{r,s+1}+\epsilon_{r,s}}{a_{r,s+1}-a_{r,s}}, \quad \epsilon_{r,s}(\alpha) = \frac{b_{r,s}\epsilon_{r,s+1}+b_{r,s+1}\epsilon_{r,s}}{a_{r,s+1}-a_{r,s}}$$

on noting that all quantities occurring, including denominators, are positive in this particular example (as is quite commonly the case) so that maximum errors in the numerator subtractions reinforce. The multipliers $\epsilon(I)$ and $\epsilon(\alpha)$ are also tabulated.

It is evident, in this example, that it is not sufficiently extended for rounding-errors to be significant.

12·9. *Error-bounds of original estimates*

Errors have been assumed independent, and discussion has been confined to cases where their bounds are also equal. There are obvious cases where this assumption is unrealistic, although it can always be attained by use of guard-figures in deriving the estimates and then rounding them to a lower precision.

In cases where we wish to make the most of our original material we must work with other assumptions. Cases that come to mind are

(*a*) Sums. Here S_n may have ϵ_n of order n for a single sum and of order n^2 for a double sum, while errors in sums are *not independent*.

(*b*) Integrals. Subdivisions in the range (a, b) of integration give submultiples of $b-a$ which means that each pair of trapezoidal estimates share ordinates to a greater or lesser extent. Errors in estimates are thus not independent. It may, for example, be deduced from table 11·6, that Romberg integration has a positive multiplier for *every* ordinate, whence a maximum error $(b-a)\,\epsilon$ (the maximum for a single ordinate being ϵ) for every result in the tableau.

Double integrals likewise need similar consideration.

(*c*) Weighted reduction. If we use a vector $\mathbf{v}$ instead of the unit vector $\mathbf{u}$, we modify the maximum errors in the original estimates, i.e. the first column, of the reduction tableau; it may well be worthwhile to allow for this in error estimation.

We do not pursue these considerations further, except to remark that we may clearly replace the initial error vector $\mathbf{v}^{(4)} = (1, -1, 1, -1, \ldots)^T$ by the more realistic $(\epsilon_1, -\epsilon_2, \epsilon_3, -\epsilon_4, \ldots)^T$ where the ϵ_r are maximum errors.

13. Choice of arguments

We are now in a position to recommend the choice of arguments most appropriate for the solution of various problems of elimination of error, or interpolation.

Points for consideration are

(i) The amount of labour involved in the evaluation of the initial estimates.
(ii) The effects of truncation error.
(iii) The effects of rounding error.
(iv) Extra work involved in the use of non-unit vectors $\mathbf{v}$.

13·1. *Labour*

Of the examples we have considered, the amount of labour in obtaining estimates is least for sums of series, assuming individual terms relatively easy to obtain; the same applies to direct interpolation, assuming a table of values already available. Finite integrals in one variable come next; each new estimate normally involves an increasing number of extra integrand evaluations (see § 13·5). Double sums probably come next, then double integrals, and infinite integrals in one variable, and so on to higher integrals and sums.

Precision required may also cause variations in labour involved; we remark here that, in view of various rounding-error effects, it is useful to keep several guarding figures if these can be retained at little cost.

13·2. *Truncation error*

This occasionally makes it unhelpful to use one or more of the crudest estimates available. This usually shows up clearly in the Neville tableau and can be remedied if necessary by appending further estimates.

13·3. *Rounding errors*

We have discussed this in some detail in § 12. In general, Romberg arguments give the smallest errors, with arguments $h = 1/n$ with equally spaced values of n giving the largest. Vectors $\mathbf{v}$ with alternating sign have a very good effect on rounding error.

13·4. *Non-unit vectors*

These involve about the same amount of extra work, per vector, as a full Neville reduction. The gain in extra or more speedy results is usually pronounced, due to more rapid elimination of truncation error, or to reduced rounding-error effects.

13·5. *Number of points in finite one-dimensional integrals*

Romberg integration uses $h = 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$, etc., so that use of 2^n+1 points (trapezoidal rule) gives $n+1$ distinct estimates, with integrands used again in all estimates subsequent to their first appearance. The use of the estimates proposed by Lyness & McHugh, with $h = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ involves intermittent re-use of most early function values, but fewer new ones, and gives cruder but easier estimates than the later Romberg ones. In this paper we suggest use of estimates for $h = 1/d$, where either (*a*) d is chosen to require as few new function values as possible at each stage, or (*b*) d is chosen to form a complete set of divisors of some suitable number N (say 12, 24, 20, 30) with a relatively large number of such divisors. The characteristics of these choices are intermediate between those of Romberg (many points for depth of eliminations, small rounding-error) and of Lyness & McHugh (fewer points, large rounding error) so far as errors are concerned (with (*b*) nearer Romberg, (*a*) nearer Lyness & McHugh); (*a*) is generally more economical than either in the number of points required; (*b*) is mentioned because of good rounding-error, though Lyness & McHugh may need fewer points.

The minimum number of points needed for elimination of depth $k-1$ in the various cases are tabulated; this needs k estimates.

Numbers of points needed

number of estimates k	4	5	6	7	8	9	10	11	12	13	14	15	16
Romberg	9	17	33	65	129	257	—	—	—	—	—	—	—
Lyness & McHugh	7	11	13	19	23	29	33	43	47	59	65	73	81
Miller (*a*)	7	9	13	17	21	25	31	37	43	49	57	65	73
Miller (*b*)	7	.	13	.	25	37	49	.	61	.	.	.	121

For the numbers of points (*a*) there may be more than one possible configuration; we list a few

$k = 4$ $(0, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, 1)$ or $(0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1)$,

$k = 5$ $(0, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, 1)$,

$k = 6$ as for $k = 5$ with either $(\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5})$ or $(\frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8})$ or $(\frac{1}{12}, \frac{5}{12}, \frac{7}{12}, \frac{11}{12})$,

$k = 7$ as for $k = 5$ with two of the extra sets in $k = 6$ or with $(\frac{1}{10}, \frac{1}{5}, \frac{3}{10}, \frac{2}{5}, \frac{3}{5}, \frac{7}{10}, \frac{4}{5}, \frac{9}{10})$,

$k = 8$ $(0, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, 1)$ and two of the sets $(\frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8})$, $(\frac{1}{10}, \frac{3}{10}, \frac{7}{10}, \frac{9}{10})$ and $(\frac{1}{12}, \frac{5}{12}, \frac{7}{12}, \frac{11}{12})$,

$k = 9$ 0, 1 and all fractions with denominators 2, 3, 4, 5, 6, 8, 10, 12,

$k = 10$ as for 9, with fractions having denominator either 7 or 9 appended.

A few useful denominators N for the choices (*b*) are also listed; these always need $N+1$ points, while the factor 5 may have other features of usefulness with a decimal system.

N	k	divisors
12	6	1, 2, 3, 4, 6, 12
20	6	1, 2, 4, 5, 10, 20
24	8	1, 2, 3, 4, 6, 8, 12, 24
30	8	1, 2, 3, 5, 6, 10, 15, 30
36	9	1, 2, 3, 4, 6, 9, 12, 18, 36
40	8	1, 2, 4, 5, 8, 10, 20, 40
48	10	1, 2, 3, 4, 6, 8, 12, 16, 24, 48
60	12	1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60
72	12	1, 2, 3, 4, 6, 8, 9, 12, 18, 24, 36, 72
120	16	1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 20, 24, 30, 40, 60, 120

Others are

$k = 10$ $N = 80, 112, \ldots$
$k = 12$ $N = 84, 90, 96, 108, 126, 132, 140, \ldots$
$k = 15$ $N = 144, \ldots$
$k = 16$ $N = 168, 210, 216, \ldots$
$k = 18$ $N = 180, 252, 288, 300, \ldots$
$k = 20$ $N = 240, \ldots$

13·6. *Practical suggestions*

We may summarize the suggestions made or implied as follows.

Sums of series of positive terms. If the terms are easy to obtain we should use Romberg estimates, or submultiples of some suitable N. If terms are less easy to get, or if high precision is possible, we may use all the early estimates of the required result, as in Lyness & McHugh, but *retain several guard figures*, if possible. If terms are difficult both in number and precision we must use early estimates, but it may possibly help a little to get further reduced values not given in the main tableau, for comparisons, by omitting one or two of the larger values of n, other than the largest.

Sums of series of terms with alternating sign. Use of successive estimates with errors of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

alternating sign is unambiguously indicated by the example in §10·3, unless extra precision is very easy, and the results of applying the Neville process to **v** are unavailable.

Integrals in one variable. Here the number of integrands should be kept down, subject to the possibility of adequate precision and control of rounding-error, and, to a lesser extent, the easy availability of certain function values. Possible special sets are indicated in §13·5.

Multiple sums and integrals. We must here make the best possible use of all estimates obtainable, which suggests the use of $n = 1, 2, 3, \ldots$ in succession. For double sums we may use sums over squares $|r|, |s| \leqslant n$, but it is worth while to consider diamonds, with $|r|+|s| \leqslant n$, as in §4·6, to give estimates for fewer terms with the same k.

For infinite integrals with a maximum at the origin, there is slight danger in using estimates that alternately use and do not use the function value at the origin, as an error of alternating signs may be present that is not easily eliminable in this case.

14. Further applications

More applications come fairly readily to mind. Whenever a deferred approach to the limit is appropriate, so are the methods outlined above. We may mention the solution of ordinary and partial differential equations in particular—these appear to merit the thorough study that is still needed, and which it is hoped may be provided later.

The results in this paper have accumulated over several years, and owe a considerable debt to several students taking the course for the Cambridge Diploma in Computer Science. I wish to mention in particular, Dr F. Rodriguez, from Brazil, who worked on the matrix representation and evaluation of derivatives. I am also grateful to Mr M. Atkinson, Mr S. C. P. Parry and Mr J. Dawson for a thorough checking and study of all the examples and theory.

References

Aitken, A. C. 1932 On interpolation by iteration of proportional parts without the use of differences. *Proc. Edinb. Math. Soc.* (2), **3**, 56–76.

Aitken, A. C. 1938 Studies in practical mathematics, III. The application of quadratic extrapolation to the evaluation of derivatives, and to inverse interpolation. *Proc. Roy. Soc. Edinb.* **58**, 161–175.

Fort, T. 1948 *Finite differences and difference equations in the real domain.* Oxford: Clarendon Press.

Fox, L. 1967 Romberg integration for a class of singular integrands. *Computer J.* **10**, 87–93.

Gragg, W. B. 1965 On extrapolation algorithms for ordinary initial value problems. *J. SIAM Numer. Anal.* (B) **2**, 384–403.

Henrici, P. 1964 *Elements of numerical analysis.* New York: Wiley.

Hermite, C. 1878 Sur la formule d'interpolation de Lagrange. *J. reine angew. Math.* (*Crelle's*) **84**, 70–79; *Oeuvres* **3**, 432–443.

Lyness, J. N. & McHugh, B. J. J. 1963 Integration over multidimensional hypercubes. I. A progressive procedure. *Computer J.* **6**, 264–270.

Miller, J. C. P. 1960 Numerical quadrature over a rectangular domain in two or more dimensions. Part 3. Quadrature of a harmonic integrand. *Maths. Comput.* **14**, 240–248.

Neville, E. H. 1934 Iterative interpolation. *J. Indian Math. Soc.* **20** (Jubilee Memorial Volume), 87–120.

Richardson, L. F. & Gaunt, J. A. 1927 The deferred approach to the limit. Part 1. Single lattice. By L. F. Richardson, pp. 299–349. Part 2. Interpenetrating lattices. By J. A. Gaunt, pp. 350–361. *Phil. Trans.* A **226**, 299–361.

Romberg, W. 1955 Vereinfachte numerische integration. *K. Norske Vid. Selsk. Forh.* (Trondheim) **28**, 30–36.