

LVII. *On the Numerical Integration of Differential Equations.*
 By H. T. H. PIAGGIO, M.A., D.Sc., University College,
 Nottingham*.

I. Introduction and Summary.

MANY physical problems lead to differential equations which cannot be integrated in finite terms by any of the usual devices. In such cases some form of approximation must be used. The earliest method of approximating was by an infinite series, but this is often very tedious when accurate numerical values are required. Runge (*Math. Ann.* vol. xlvi. 1895) has given a formula for calculating approximately the increment of y corresponding to a small increment of x , when x and y are connected by the differential equation

$$\frac{dy}{dx} = f(x, y),$$

and it is given that $y=b$ when $x=a$.

Other approximate formulæ have been given by Heun and Kutta (*Zeitschrift f. Math. u. Physik*, vols. xlv. and xlvi.).

In applying any of these methods to actual examples, it is important to know how far wrong the result may be. Runge assumes that when two steps of his approximations come fairly close together, the error in the final result will be of about the same order of magnitude as one-third of their difference. But he does not give any definite upper limit for this error.

The object of the present paper is to supply this omission. Four simple formulæ are found which give four numbers, between the greatest and least of which the required increment of y must lie. A new approximate formula is derived from these.

As an illustration this new formula is applied to the example given by Runge. The result is more accurate than those given by the methods of Runge, Heun, or Kutta.

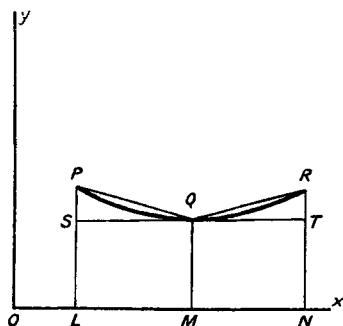
II. Limits between which the value of a definite integral lies.

Let $F(x)$ be a function which, together with its first and second differential coefficients, is continuous (and therefore finite) between $x=a$, and $x=a+h$. Let $F''(x)$ be of constant sign in the interval. In the figure this sign is taken as

* Communicated by the Author.

positive, making the curve concave upwards. LP, MQ, NR are parallel to the axis of y , M is the middle point of LN, and SQT is the tangent at Q. OL = a , LN = h .

Fig. 1.



Then the area PLNR lies between that of the trapezium SLNT and the sum of the areas of the trapezia PLMQ, QMNR.

That is, $\int_a^{a+h} F(x) dx$ lies between

$$hF(a + \frac{1}{2}h) = A \text{ say,}$$

and $\frac{1}{4}h\{F(a) + 2F(a + \frac{1}{2}h) + F(a + h)\} = B \text{ say.}$

In the figure $F''(x)$ is positive and A is the lower limit, B the upper. If $F''(x)$ were negative, A would be the upper limit and B the lower.

As an approximation to the value of the integral it is best to take, not the arithmetic mean of A and B, but $\frac{2}{3}B + \frac{1}{3}A$, which is exact when PQR is an arc of a parabola with its axis parallel to the axis of x . It is also exact for the more general case when $F(x) = a + bx + cx^2 + ex^3$, as is proved in most treatises on the Calculus in their discussion of Simpson's Rule.

III. Extension of preceding results to functions defined by differential equations.

Consider the function defined by

$$\frac{dy}{dx} = f(x, y); \quad y = b \text{ when } x = a,$$

where $f(x, y)$ is subject to the following limitations in the

range of values a to $a+h$ for x and b to $b+h$ for y . It will be seen from what follows below that the increment of y is less than h , so that all values of y will fall in the above range. The limitations are :

(1) $f(x, y)$ is finite and continuous, as are also its first and second partial differential coefficients.

(2) It never numerically exceeds unity. If this condition is not satisfied, we can get a new equation in which it is satisfied by taking y instead of x as the independent variable.

(3) Neither d^3y/dx^3 nor $\partial f/\partial y$ changes sign.

Let m and M be any two numbers, such that

$$m < j < M \leq 1.$$

Then if the values of y when x is $a + \frac{1}{2}h$ and $a+h$ are denoted by $b+j$ and $b+k$ respectively,

$$\frac{1}{2}mh < j < \frac{1}{2}Mh \leq \frac{1}{2}h, \quad (1)$$

and

$$mh < k < Mh \leq h. \quad (2)$$

We shall now apply the formulæ of the last section, taking y to be the same function as that defined by

$$y = b + \int_a^{a+x} F(x) dx,$$

so that

$$k = \int_a^{a+h} F(x) dx.$$

We have to express the formulæ in terms of f instead of F .

Now, $F(a)$ = the value of dy/dx when $x=a$,

so

$$F(a) = f(a, b).$$

Similarly, $F(a + \frac{1}{2}h) = f(a + \frac{1}{2}h, b + j)$,

and

$$F(a+h) = f(a+h, b+k).$$

Now, if $\partial f/\partial y$ is positive, so that f increases with y , the inequalities (1) and (2) lead to

$$f(a + \frac{1}{2}h, b + \frac{1}{2}mh) < f(a + \frac{1}{2}h, b + j) < f(a + \frac{1}{2}h, b + \frac{1}{2}Mh), \quad (3)$$

$$\text{and } f(a+h, b+mh) < f(a+h, b+k) < f(a+h, b+Mh); \quad (4)$$

while if $\partial f/\partial y$ is negative,

$$f(a + \frac{1}{2}h, b + \frac{1}{2}mh) > f(a + \frac{1}{2}h, b + j) > f(a + \frac{1}{2}h, b + \frac{1}{2}Mh), \quad (5)$$

$$\text{and } f(a+h, b+mh) > f(a+h, b+k) > f(a+h, b+Mh). \quad (6)$$

Thus if $F''(x) = dy/dx$ is positive and $\partial f/\partial y$ is also positive, the result of section II.,

$$A < k < B,$$

may be replaced by

$$p < k < Q, \dots \dots \dots (7)$$

where $p = hf(a + \frac{1}{2}h, b + \frac{1}{2}mh)$

and $Q = \frac{1}{4}h\{f(a, b) + 2f(a + \frac{1}{2}h, b + \frac{1}{2}Mh) + f(a + h, b + Mh)\}$;

while if $F''(x)$ is positive, and $\partial f/\partial y$ is negative,

$$P < k < q, \dots \dots \dots (8)$$

where $P = hf(a + \frac{1}{2}h, b + \frac{1}{2}Mh)$

and $q = \frac{1}{4}h\{f(a, b) + 2f(a + \frac{1}{2}h, b + \frac{1}{2}mh) + f(a + h, b + mh)\}$.

Similarly, if $F''(x)$ and $\partial f/\partial y$ are both negative,

$$p > k > Q, \dots \dots \dots (9)$$

while if $F''(x)$ is negative and $\partial f/\partial y$ positive,

$$P > k > q, \dots \dots \dots (10)$$

These results may be summed up by saying that *in every case* (subject to the limitations on f stated at the beginning of this section) k lies between the greatest and least of the four numbers $p, P, q,$ and Q .

As an approximate formula we use $k \doteq \frac{2}{3}B + \frac{1}{3}A$, replacing B by Q or q , and A by p or P .

IV. Application to a numerical example.

Consider the example selected by Runge and Kutta to illustrate their methods

$$\frac{dy}{dx} = \frac{y-x}{y+x}; \quad y=1 \text{ when } x=0.$$

It is required to find the increment k of y when x increases by 0.2 . Here $f(x, y) = (y-x)/(y+x)$. This function satisfies the conditions laid down in the last section.

We take $M=1, m=(1-0.2)/(1.2+0.2)=4/7$.

Then

$$p = 0.1654321,$$

$$P = 0.1666667,$$

$$q = 0.1674987,$$

$$Q = 0.1690476.$$

Thus k lies between p and Q .	Errors.
$\frac{2}{3}Q + \frac{1}{3}p = 0.1678424,$	0.0000007
Kutta's value 0.1678449,	0.0000032
Runge's value 0.1678487,	0.0000070
Heun's value 0.1680250,	0.0001833

The second, third, and fourth of these were calculated by Kutta. They are also given in Bateman's 'Differential Equations,' p. 229. Now this particular example admits of integration in finite terms, giving

$$\log(x^2 + y^2) - 2 \tan^{-1}(x/y) = 0.$$

Hence we may find the accurate value of k .

$$\text{Accurate value} = 0.1678417.$$

Thus in this example our result is the nearest to the accurate value, the errors being as stated above.

We may also test the method by taking a larger interval $h=1$. Of course a more accurate way of obtaining the result would be to take several steps, say $h=0.2, 0.3,$ and finally $0.5,$ as Runge does.

Still, it is interesting to see how far wrong the results come for the larger interval.

$$\text{We take } M=1, m=(1-1)/(2+1)=0.$$

Then	$\frac{2}{3}Q + \frac{1}{3}p = 0.50000.$
True value = 0.49828,	Errors.
Kutta's value = 0.49914,	0.00086
Our value = 0.50000,	0.00172
Heun's value = 0.51613,	0.01785
Runge's value = 0.52381,	0.02553

This time Kutta's value is the nearest, and ours is second.

University College, Nottingham,
February 22nd, 1919.