iment with different values of the step size; if the choice of $h$ worries you, that is all to the good since you will be in a better frame of mind for what is to come.

The improved Euler's method goes well on a hand calculator. It is useful if you realize its limitations.

## 3.3 RUNGE-KUTTA FORMULAS

The general explicit Runge-Kutta step can be written as follows. For

$$\frac{dy}{dx} = f(x,y); \quad y(x_0) = y_0.$$

Let

$$\begin{aligned} f_0 &= f(x_0,y_0), \\ f_1 &= f(x_0 + \alpha_1 h,\ y_0 + h\beta_{10}f_0) \\ f_2 &= f(x_0 + \alpha_2 h,\ y_0 + h\{\beta_{20}f_0 + \beta_{21}f_1\}), \\ &\ \vdots \\ f_k &= f(x_0 + \alpha_k h,\ y_0 + h\{\beta_{k0}f_0 + \beta_{k1}f_1 + \ldots + \beta_{k,k-1}f_{k-1}\}). \end{aligned} \tag{3.3.1}$$

Then

$$y(x_0 + h) \simeq y_1 = y_0 + h(c_0f_0 + c_1f_1 + \ldots + c_kf_k).$$

To illustrate some of the nuts and bolts of a formula, let's look at the case $k = 2$:

$$\begin{aligned} f_0 &= f(x_0,y_0), \\ f_1 &= f(x_0 + \alpha h,\ y_0 + h\beta f_0). \\ y(x_0 + h) \simeq y_1 &= y_0 + h(c_0f_0 + c_1f_1). \end{aligned} \tag{3.3.2}$$

The $\alpha$'s, $\beta$'s and $c$'s are constants. We need to find them such that the error in $y_1$ is as small as we can make it. Well, we don't need to really, because there are heroes of a very special kind who have done it for us. But just relax and watch it happen; you will be all the better for it.

Both sides of the last equation of (3.3.2) must be expanded in powers of $h$. For a start,

$$y(x_0 + h) = y(x_0) + hy'(x_0) + \tfrac{1}{2}h^2y''(x_0) + \tfrac{1}{6}h^3y'''(x_0) + \ldots \tag{3.3.3}$$

Now

$$y(x_0) = y_0, \text{ and } y'(x_0) = f(x_0,y_0) = f_0.$$

Further,

$$y'' = \frac{d}{dx}(y') = \frac{d}{dx}f(x,y) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = f_x + ff_y,$$

and

$$\begin{aligned} y''' &= \frac{d}{dx}(y'') = \frac{\partial y''}{\partial x} + \frac{\partial y''}{\partial y}\frac{dy}{dx} \\ &= f_{xx} + f_x f_y + 2ff_{xy} + ff_y^2 + f^2 f_{yy}. \end{aligned} \tag{3.3.4}$$

Then

$$\begin{aligned} y(x_0 + h) = y_0 + hf_0 &+ \tfrac{1}{2}h^2[f_x + ff_y]_0 \\ &+ \tfrac{1}{6}h^3[f_{xx} + f_x f_y + 2ff_{xy} + ff_y^2 + f^2 f_{yy}]_0 + \ldots \end{aligned} \tag{3.3.5}$$

All the derivatives are evaluated at the point $(x_0,y_0)$.

Next we take note of the start of a Taylor series of a function of two variables:

$$\begin{aligned} f(x + a, y + b) &= f(x,y) + af_x(x,y) + bf_y(x,y) \\ &+ \tfrac{1}{2}[a^2 f_{xx}(x,y) + 2abf_{xy}(x,y) + b^2 f_{yy}(x,y] + \ldots \end{aligned}$$

Therefore,

$$\begin{aligned} y_1 &= y_0 + hc_0 f_0 + hc_1 f_1 \\ &= y_0 + hc_0 f_0 + hc_1 f(x_0 + \alpha h, y_0 + h\beta f_0) \\ &= y_0 + hc_0 f_0 + hc_1[f(x_0,y_0) + \alpha h f_x(x_0,y_0) + h\beta f_0 f_y(x_0,y_0) \\ &\qquad + \tfrac{1}{2}(\alpha h)^2 f_{xx}(x_0,y_0) + (\alpha h)(h\beta f_0) f_{xy}(x_0,y_0) \\ &\qquad + \tfrac{1}{2}(h\beta f_0)^2 f_{yy}(x_0,y_0)] + \ldots \\ &= y_0 + hc_0 f_0 + hc_1 f_0 + h^2 c_1(\alpha f_x + \beta f_0 f_y)_0 \\ &\quad + h^3 c_1(\tfrac{1}{2}\alpha^2 f_{xx} + \alpha\beta f_0 f_{xy} + \tfrac{1}{2}\beta^2 f_0^2 f_{yy})_0 + \ldots \end{aligned} \tag{3.3.6}$$

Combining (3.3.5) and (3.3.6), we find that

$$\begin{aligned} y(x_0 + h) - y_1 = hf_0\,[1 - c_0 - c_1] & \\ + h^2[f_x(\tfrac{1}{2} - c_1\alpha) + ff_y(\tfrac{1}{2} - c_1\beta)] & \\ + h^3[f_{xx}(\tfrac{1}{6} - \tfrac{1}{2}c_1\alpha^2) + ff_{xy}(\tfrac{1}{3} - c_1\alpha\beta) & \\ + f_{yy}(\tfrac{1}{6} - \tfrac{1}{2}c_1\beta^2)f^2 + \tfrac{1}{6}f_x f_y + \tfrac{1}{6}ff_y^2] & \\ + \text{terms with higher powers of } h. & \end{aligned} \tag{3.3.7}$$

This quantity is to be made as small as possible for any possible function $f(x,y)$.

For a start, we can make sure that

$$\left.\begin{aligned} 1 - c_0 - c_1 &= 0, \\ \frac{1}{2} - c_1\alpha &= 0, \\ \frac{1}{2} - c_1\beta &= 0. \end{aligned}\right] \tag{3.3.8}$$

Equations (3.3.8) are called *equations of condition*. In this case they can be satisfied with one degree of freedom, and this freedom can be used with various strategies to try to minimize the $h^3$ term. Taking advantage of (3.3.8), the $h^3$ term of (3.3.7) can be written as

$$h^3\left(f_{xx}\left(\frac{1}{6} - \frac{1}{4}\alpha\right) + ff_{xy}\left(\frac{1}{3} - \frac{1}{2}\alpha\right) + f^2 f_{yy}\left(\frac{1}{6} - \frac{1}{4}\alpha\right) + \frac{1}{6} f_x f_y + \frac{1}{6} ff_y^2\right). \tag{3.3.9}$$

This is the principal term in the error of the method; it confirms that the order of the method is two. Notice that both the modified and improved Euler's methods are special cases of (3.3.9). If the leading term in the error is written as $Ah^3$, notice the complexity of the formula for $A$ when compared with a similar formula for Euler's method.

## 3.4 STEP SIZE CONTROL—SOME TACTICS TO AVOID

The remainder term (3.3.9) is unwieldy, and such terms for higher order Runge-Kutta methods become much more involved, so they cannot in practical terms be made the basis for estimation of local truncation error or for step size control. We shall see in a moment how these formulas can be used implicitly; but first I want to warn against an approach that has often been used in the past in Runge-Kutta calculations and should never be used again. This approach involves shifting tactics between halving and doubling the interval.

Let us say we are using the algorithm (3.1.15) and have just completed a step with step size $h$. But you feel uneasy and develop cold feet about $h$. Was $h$ too large? Have you accumulated a ruinous·truncation error? One way to deal with this alarm is to go back to the start of the step just taken and then take *two* steps, each with step size $h/2$. If the difference between the end results exceeds some preset tolerance, then your alarm was justified: throw out $h$ and use $h/2$

as the step size. BUT IS THIS REDUCED STEP SIZE TOO LARGE? Calm down; if it will make you feel better, you can halve the interval again and repeat the test. But remember that if you go on doing this, then the difference between the end results could arise from round-off error, and after that you will continue to halve the interval forever, or at least until the men in white coats take you away.

Suppose on the other hand that you passed the halving test and feel sure that the step size $h$ was all right. Now you start to be afraid that you are doing too much work and taking too many steps. That is, is your step size too small? Well, we have just found that the one step using $h$ was all right. So take another step also using $h$. If this is also all right, go back to the start of the first step and take a step with step size $2h$. Then compare the end results. If the difference is small enough, then the larger step size can be used. Or do you want to try doubling it again?

Consider a model like that for project 4.37. A comet moves in an uneventful orbit, then has a hectic encounter with Jupiter, and then resumes a relatively tranquil path. During the encounter with Jupiter, the step size will need to be drastically reduced; afterward it can be increased again. The technique of halving and doubling would be time-consuming and extravagant, since the main expense of the calculation is likely to be evaluating the functions on the right-hand sides of the equations. It will also very probably become unstable; so would you.

## 3.5 FEHLBERG'S METHOD

E. Fehlberg has published Runge-Kutta methods of several orders in which the operator of a program chooses an upper bound for the local truncation error, and the program finds a "safe" step size consistent with this bound; it also checks at the end of the step to see whether it really was safe. To introduce Fehlberg's method we shall illustrate it with an example of low order. This example, and the principal method to be used in this text, appeared in NASA Technical Report R315 in 1969.

You can verify from (3.3.2) and (3.3.8) that the following method is of second order:

$$\left.\begin{aligned} &\frac{dy}{dx} = f(x,y); \quad y(x_0) = y_0. \\ &f_0 = f(x_0,y_0), \\ &f_1 = f(x_0 + h,\, y_0 + hf_0), \\ &y_1 = y_0 + h\left(\frac{1}{2}f_0 + \frac{1}{2}f_1\right). \end{aligned}\right] \qquad (3.5.1)$$

So

$$y(x_0 + h) - y_1 = Ah^3 + Bh^4 + \ldots \qquad (3.5.2)$$