# FAST NUMERICAL SIMULATION OF BIOCHEMICAL SYSTEMS

E.M.CHANCE and A.R.CURTIS
*Department of Biochemistry, University College London, London W.C.1.*
*and Mathematics Branch, A.E.R.E., Harwell, Berks., England*

Received 10 December 1969
Received version received 26 January 1970

Following the FEBS Summer School [1] at Edinburgh in 1968, we set out to apply to biochemical systems a new method designed by Gear [2, 3] for rapid integration of stiff sets of differential equations. This attempt has been, in our judgement, successful, and we report details below.

As described by Cooper [4], we suppose the differential equations to be linearised, in the form

$$\frac{dy_i}{dt} = \sum_{j=1}^{N} A_{ij}(t_0)(y_j - \alpha_j) \qquad (1)$$

The property of stiffness arises in biochemical systems largely because interacting chemical species (e.g. an enzyme and its substrate) are present in concentrations differing by large factors. It manifests itself by eigenvalues of the matrix $A_{ij}$ (which are negative or, if complex, have negative real parts) being numerically large in relation to the rates of change of the solution (these eigenvalues have the dimension of reciprocal time). With most integration methods, gross instability results if the product of the step length and the largest eigenvalue exceeds some definite number of order unity, and the techniques usually adopted for automatic control of step length prevent this happening; we explain this point below. Thus, in stiff systems, a very large number of steps must be taken, requiring much computer time. Gear's method avoids this problem, since it is stable whatever the product of step length with eigenvalue of $A_{ij}$ (provided that the eigenvalue has negative real part and, if complex, is not too close to the imaginary axis).

Until now, predictor-corrector methods have (rightly, in the absence of Gear's method) been rejected for biochemical systems in favour of Euler's method, with step length controlled by estimating the "truncation error" $\frac{1}{2}h^2 y''$; in practice, the difference of two successive values of $y'_i$ is taken as an estimate of $hy''_i$. We now proceed to analyse this method. Consider the component of y in the direction of the eigenvector of $A_{ij}$ corresponding to a negative eigenvalue, $-\lambda$ say. Writing z for this component, (1) leads to

$$\frac{dz}{dt} = -\lambda(z - \beta) \qquad (2)$$

If an error $\zeta(t)$ occurs in z at time t (e.g. as a rounding error), we find by subtracting (2) from the corresponding equation for $(z + \zeta)$ (as long as $\zeta$ is small enough to justify linearisation) that we should have

$$\zeta(t + h) = \zeta(t)e^{-\lambda h} \qquad (3)$$

In Euler's method with step h we obtain instead of (3)

$$\begin{aligned} z(t + h) + \zeta(t + h) &= z(t) + \zeta(t) - \lambda h[z(t) + \zeta(t) - \beta] \\ &= z(t + h) - \tfrac{1}{2}h^2 z''(t) + (1 - \lambda h)\zeta(t) \end{aligned} \qquad (4)$$

It follows that the actual error at time $(t + h)$ is

$$\zeta(t + h) = (1 - \lambda h)\zeta(t) - \tfrac{1}{2}h^2 z''(t) \qquad (5)$$

where we suppose $h^3 z'''$ and higher order terms negligible, i.e. that $z''$ changes slowly in relation to the step-length h, and can be treated as a constant over several steps.

If $\lambda h > 2$, $1 - \lambda h < -1$ and the error will grow numerically (with alternating sign at each step). There will therefore be instability unless $\lambda h \leqslant 2$ for each $\lambda$, and

the most unstable mode will correspond to the largest $\lambda$. The equations are stiff if $z''$ for this most unstable mode is small enough that $\lambda h >> 2$ when h satisfies $|h^2 z''| = 2E$, E being a preset error tolerance. In our experience, this condition is satisfied in biochemical problems, at least after a biochemically uninteresting initial transient.

The usual technique for step-length control is to take the divided difference of two successive calculated first derivatives as an estimate of $z''$, choosing the next step h in an attempt to make $|\tfrac{1}{2}h^2 z''| = E$. Suppose that $h = h_r$ for the rth step, and that the exact solution and the error after this step are $z_r$, $\zeta_r$. Then this divided difference is

$$\begin{aligned} D_r &= -\frac{\lambda}{h_r}[(z_r + \zeta_r - \beta) - (z_{r-1} + \zeta_{r-1} - \beta)] \\ &= \frac{1}{h_r}[z'_r - z'_{r-1} - \lambda(\zeta_r - \zeta_{r-1})] \\ &= (1 + \tfrac{1}{2}\lambda h_r)\, z'' + \lambda^2 \zeta_{r-1} \end{aligned} \tag{6}$$

by (5). Then we choose

$$h_{r+1} = [2E/|D_r|]^{1/2} \tag{7}$$

This step-length control scheme is successful in keeping $|\zeta_r/E|$ bounded by a number of order unity, because in a stiff system $|\zeta|$ grows through instability until the last term in (6) dominates, so that (7) becomes approximately

$$\lambda h_{r+1} = [2E/|\zeta_{r-1}|]^{1/2} \tag{8}$$

If $|\zeta_{r-1}|$ is significantly greater than E, a succession of steps with $\lambda h$ of order unity or less is taken, thus reducing $|\zeta|$; if in the other hand $|\zeta_{r-1}|$ is smaller than $\tfrac{1}{2}E$, one or more steps with $\lambda h > 2$ are taken, increasing $|\zeta|$. It can be shown that

$$\lambda h_{r+1} = \lambda h_r / |\lambda h_{r-1} - 1|^{1/2} \tag{9}$$

approximately. This non-linear recurrence relation admits a steady-state solution.

$$\lambda h_r = 2, \zeta_r = \pm \tfrac{1}{2}E \tag{10}$$

and we have observed this state to be reached, at least when additional precautions were taken, such as imposing a bound on ratio $h_{r+1}/h_r$. However, it does not seem possible to prove that this state will always be reached, and probably $\lambda h_r$ may oscillate finitely.

Assuming that (10) is obeyed at least on average a good estimate of the largest eigenvalue is twice the number of steps taken to advance the integration by one second. We have found in the models we have used that this estimate gives values of the order $10^5$ $sec^{-1}$, in good agreement with estimates by independent methods; this implies a shortest relaxation time constant of order 10 $\mu$sec.

By contrast, Gear's method controls the step-length purely in relation to the time-scale of change of the solution, and we found that on these same problems the step size it uses can grow to one second or longer; naturally, each step is more complicated than in Euler's method. Gear uses a predictor-corrector technique, and two or three derivative evaluations per step are needed, corresponding to two or three corrector iterations. Moreover, in order to make these iterations converge with large values of $\lambda h$, it forms and uses an estimate of the matrix $A_{ij}$. Writing the original equations in the form

$$\frac{dy_i}{dt} = f_i(y_1, y_2, ..., y_N) \tag{11}$$

we have

$$A_{ij} = \partial f_i / \partial y_j \tag{12}$$

and we have found it satisfactory in practice to estimate these partial derivatives by finite differences. To estimate $A_{ij}$ thus requires N additional evaluations of the derivative vector, one for each column of the matrix; this is done at every step, but only when failure of corrector iterations to converge indicates the necessity.

At each corrector iteration, a set of N linear equations (whose matrix of coefficients is derived from $A_{ij}$) has to be solved. We have taken advantage of the sparseness of the matrix $A_{ij}$ to do this solution economically and we then find that the large increase in step-length gives a large increase in computing speed. However, because of storage requirements for $A_{ij}$ and for other quantities used in the method, the computer program is rather large.

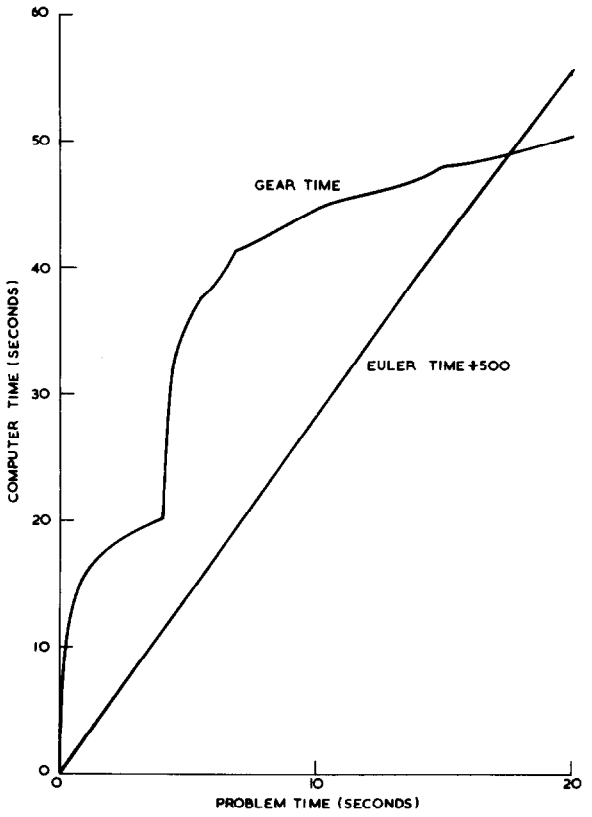


Fig. 1. Timing comparison on the Garfinkel-Hess model.

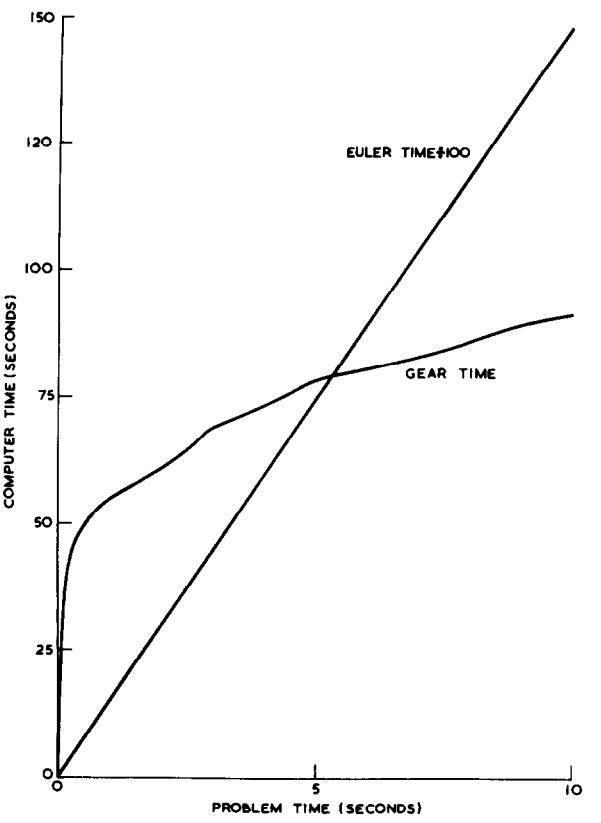


Fig. 2. Timing comparison on the Chance-Hess model.

We have carried out two kinds of test on each of two different biochemical systems. The first system was intended to be the Garfinkel and Hess [5] model for the glycolytic pathway in ascites cells; owing to a misunderstanding, the model used was not exactly that of Garfinkel and Hess, but the discrepancy in no way invalidates the tests described below, and we are clear that the exact model would give the same timings. The second system is one due to Chance and Hess [6] for glycolysis in yeast; it is similar in complexity to the first system, but evolves over a shorter time-scale. Each system involves about 65 chemical species and about 90 unidirectional reactions.

The first kind of test was of the accuracy of Gear's method. Using it, we produced exactly the same results, for each system, as we obtained with Euler's method. In the case of the Garfinkel and Hess system, we carried out the accuracy tests on a modified model, in which the large eigenvalues were reduced by a factor 100 by reducing enzyme turnover numbers (one of the expedients described by Garfinkel and Hess in their paper), in order to save computer time with Euler's method. The Chance and Hess system was treated in its original form, the Euler run taking 4 hr of computer time. These tests have satisfied us that both Euler's and Gear's methods produce correct results, if used with care (for example, with double precision arithmetic).

In the second kind of test we compared the speeds of the methods. For the Garfinkel and Hess model with Euler's method we compared the initial speed on the original model with that after we had changed the eigenvalues, and simply multiplied the computer time with the reduced eigenvalues by the observed initial speed ratio (which was of the order of 100, as expected). With Gear's method we did the timing runs on the ori-

ginal model; in fact, both the original and modified models ran at essentially the same speed. For the Chance and Hess system, the comparison was quite direct, since a full run had been done on the original model using Euler's method. IBM 360-65 computers at U.C.L. and A.E.R.E. were used for the test, but different Euler programs were used on the two models.

We show in figs. 1 and 2 the computer time needed by both methods for the simulation, as a function of problem time. Note that the computer times shown for Euler's method are on greatly reduced scales in order to show them on the same graph as the Gear times. Because Gear's method follows faithfully the initial transients (of which there are two in the Garfinkel and Hess problem), and takes time steps related to the behaviour of the solution, and also because of the need for occasional matrix estimations, the computer time is a far from linear function of the problem time. This function was tabulated by using a subroutine which reads the internal clock of the computer at regular intervals of problem time, and printing the results. Because Euler's method takes steps related to the eigenvalues, which change more slowly, the Euler graph for Garfinkel and Hess's model is almost linear; for Chance and Hess's model we had only the total Euler time, so we assumed linearity in drawing the graph.

The figures show that Gear's method is far faster than Euler's method, and tends to show up better the longer the total simulation time. Over the last halves of the ranges, speed ratios of 2400 and 550 are shown in the two figures, although over the full ranges these are reduced to 550 and 160. We emphasise that these times refer to solution of the complete set of differential equations, without any rate law or other approximations. We have since carried out runs on a modified Chance-Hess model to 2000 sec of problem time in only 110 sec of computer time.

Our present program can handle up to 70 chemical species, with up to 100 one-way reactions. It occupies 160 K bytes of System/360 memory. We have since integrated Gear's method with the well-known data input scheme maintained by one of us [7], but the program described here has its own rather less convenient data input scheme.

We should like to record our thanks to M.J.Hopper for invaluable assistance in running tests and correcting programs. The program described was developed at A.E.R.E., and is arranged to be adaptable to other applications involving stiff sets of differential equations. Furthermore, we are indebted to the computer centres at the A.E.R.E. and U.C.L. for use of their facilities. One of us (E.M.C.) wishes to acknowledge support from the U.S. Public Health Service (Grant No. AM.10435).

## References

[1] FEBS Letters 2 (1969) Suppl. 2.
[2] C.W.Gear, Numerical integration of stiff ordinary differential equations, Report No. COO-1469-0051, Dept. of Computer Science, Univ. of Illinois.
[3] C.W.Gear, The automatic integration of stiff ordinary differential equations, Proceedings of I.F.I.P. Congress, 1968.
[4] G.J.Cooper, The numerical solutions of stiff differential equations, FEBS Letters 2 (1969) 522–529.
[5] D.Garfinkel and B.Hess, A detailed computer model of the glycolytic pathway in ascites cells, J. Biol. Chem. 239 (1964) 971–983.
[6] E.M.Chance and B.Hess, Unpublished data.
[7] E.M.Chance and E.P.Shephard, Automatic techniques in enzyme simulation, in: Computers in biomedical research, Vol. 2 (1969) p. 321.