PC4 : Numerical Integration of Ordinary Differential Equations (Part II)

1 Introduction

The Petite Classe is divided into three parts. We first investigate the stability properties of the RK methods as well as some other explicit methods available in the community, such as the Dormand and Price class of methods and draw some stability diagrams in order to properly understand the influence of stability function / diagram on the performances of the method. The second part is devoted to the integration of the Belousov-Zhabotinsky oscillating reaction. The objective of this second part is two-fold. First, we want to characterize the stiffness of the system of equations and investigate how a standard Runge-Kutta method is dealing with such a stiffness, and what is the impact on the global error. Second, we will use an adaptative embedded Runge-Kutta method in order to experiment the time step adaptation and clarify the impact of such a numerical strategy at two levels : 1- the size of the time steps and the correspondance with the local stiffness of the dynamics of the system of equations, 2- what is the local error estimate, its link to the effective local error, and finally the resulting global error. At this level it is essential to have a clear understanding of the various errors as well as a clear picture of the relevance of the local error estimate. Finally, we switch to one of the most difficult and renown test case: the van der Pol oscillator[1]. The purpose is here to get a precise idea about the influence of the concepts of stability / accuracy / order, and their response to the presence of stiffness for high order Runge-Kutta methods, whereas we have investigated such a behavior in the contexte of Euler methods in PC 3.

2 Stability diagrams: graphical representations and their use

The notebook proposed with this PC contains a program in order to plot the stability diagram studied during the course.

2.1 Compare the four diagram for the four Runge-Kutta methods of order 1, 2, 3 and 4. Comment on the evolution of the stability conditions with the order of the method. What is the meaning of the part of the stability diagram for high order methods for which the real part of z is positive?

2.2 Relying on the notebook of the previous PC on the Curtiss and Hirschfelder model, what happens when the time step is chosen so that the method is still in its stability domain, but close to the boundary? Does stability imply an accurate integration of the system?

2.3 Propose a representation of the stability domain of the DOPRI methods. How does it compare to the ones of the RK methods. What is going to be the impact on the resolution of the system?

[1] For those interested in the history of this model, we refer to the nice article [2].

3 High order RK methods for BZ reaction dynamics integration

The dynamics of the oscillating reaction discovered by Belousov and Zhabotinsky [4, 1], can be modeled through the so-called Brusselator model [7] depending on two parameters:

$$\begin{cases} \mathrm{d}_t y_1 = 1 - (b+1)y_1 + a\, y_1^2 y_2 \\ \mathrm{d}_t y_2 = b\, y_1 - a\, y_1^2 y_2 \\ y_1(0) = y_1^0 \\ y_2(0) = y_2^0 \end{cases} \tag{1}$$

For the purpose of illustrating the concepts introduced in the course, we use $a = 1$, $b = 3$, $y_1^0 = 1.5$ and $y_2^0 = 3$. For that set of parameters, the dynamics admits a limit cycle as ω-limit set for large times and after a short transition period of time, the dynamics of the system approaches a periodic behavior.

3.1 Stiffness?

3.1.1 Using the proposed parameters, integrate the system in the interval $[0, 20]$ and explain why we reach what is called a quasi-exact solution. Propose a representation of the eigenvalues of the Jacobian matrix of the non-linear system (1). Explain in what sense the model is considered to have a stiff dynamics. Propose an order of magnitude in terms of variation of the dynamics.

3.1.2 Using a change of the initial condition, comment on the impact of the stiffness of the system. Is the stiffness the same on the limit cycle and in the initial time dynamics leading the dynamics on the limit cycle?

3.2 Integration of the system using high order RK methods with fixed time steps

3.2.1 In the spirit of what has been done in PC 3, explain how we can reach the stability limits and what happens when the time step for various RK methods are too large for the stability to be guaranteed. Since the notion of stability is only valid for eigenvalues with negative real part, clarify in what part of the dynamics, the stability criterion is going to be valid.

3.2.2 In connection with the previous exercise on stability diagrams and previous questions on stiffness, explain what will be the limiting factor in terms of stability for the time step of the RK methods (Heuristics proposed in the notes and in Class). What will be the influence of the initial conditions?

3.2.3 For a given discretization, what is the impact of switching form first order to higher orders? Provide and answer in terms of 1- stability, 2- accuracy, 3- computational cost (number of evaluation of the function).

3.2.4 When you increase the number of time steps in the proposed interval and thus reduce the time step, what is the influence on the global error? In what part of the dynamics is the global error the most important? Relate the answer to the stiffness of the system and explain if this is related to the problem of stability of the method.

3.2.5 Choosing a time step of $1/60$ and a first order RK method, does the global error build up? That is, to what extent do the errors issued from the stiff zone influence the dynamics afterwards? Try to explain why.

3.2.6 If we switch to a first order implicit method for this system of equations, the method is supposed to be A-stable. However, there is a maximum time step beyond which we can not retrieve the periodic dynamics. Can you propose an equivalent theory as the one of the stability, but for equations with eigenvalue with a positive real part?

3.3 Integration of the system using adaptive time stepping

In this part, we will rely on the embedded Runge-Kutta described in class, based on the 3/8 rule fourth-order Runge-Kutta method. The purpose is here to investigate how an adaptive time stepping strategy can lead to a given error within a limited amount of numerical work. An appendix is provided in order to describe such a strategy.

3.3.1 Explain and illustrate how the adaptive time-stepping is related to the stiffness of the system. Compared to the 4th order RK method with fixed time step, evaluate the number of time steps as a function of the global error.

3.3.2 Is the number of time steps the only key issue? How can we evaluate the actuel cost of the methods?

3.3.3 Plotting on the same graph, the local error estimate, the exact local error (describe how it is calculated), as well as the global error, comment on how the embedded method is efficient for the proposed problem. Is the proposed error estimate relevant? Illustrate your answer and consider various tolerances.

3.3.4 What is the evolution of the number of steps as well as of the computational cost when we decrease the tolerance. Compare to the fixed time step methods.

3.4 Integration of the system using Dormand and Price method

In this part, we will rely on the Dormand and Price [5] method with time step adaptation. The purpose is here to compare the efficiency of the method to the one of the previous method.

3.4.1 For various levels of tolerance, evaluate the number of time steps needed for both methods as well as the precision of the results.

3.4.2 What is the impact of using a higher order and more precise method?

3.4.3 In terms of computational cost (based here on the number of accepted time step), what conclusion can be drawn from the calculations conducted with various levels of tolerance?

4 The van der Pol oscillator: a discriminating test-case

In the third part of the Petite Classe, we will tackle the van der Pol oscillator [11, 12, 2, 6], which is a very nice and discriminating test-case for numerical methods. It is an autonomous system of ordinary differential equation (ODE) written as :

$$\begin{cases} \mathrm{d}_t y_1 = y_2 \\ \mathrm{d}_t y_2 = \varepsilon \left(1 - y_1^2\right) y_2 - y_1 \quad \text{avec } \epsilon > 0 \end{cases} \tag{2}$$

where the stiffness can be tuned through the ε parameter.

4.1 Stiffness?

4.1.1 For a range of ε ranging from 1 to 20, how much time does the solution take in order to reach the limit cycle? Is the initial transient stiffer than the rest of the dynamics? Where does the stiffness comes from (summarize what has been explained in Class and presented in the second appendix)?

4.1.2 Describe the evolution of the stiffness of the system through time. How does it compare to the one encountered in the Brusselator case for a range of ε ranging from 1 to 20. How does it evolve through one period on the limit cycle?

4.1.3 The system is integrated using the Dormand and Price solver with various tolerances. How many time steps does it take to integrate the system as a function of tolerance on the one side, and ε on the other side? Propose a synthetic view on the results and explain.

4.1.4 Describe the work conducted by the method for the integration using Dormand and Price in the case $\varepsilon = 20$ and compare to what would be needed with a RK4 method with fixed time stepping leading to the same level of error.

4.1.5 Compare this to the DOPRI853 method for the same level of tolerance.

4.1.6 Conclude on the advantages and limits of the proposed schemes, making the link with the stiffness of the equation.

Appendix A: Adaptive time stepping strategy

The idea is to adapt the time step to the local dynamics in order to provide an efficient integrations strategy. The user should provide a tolerance and the adaptation has to rely on an error estimate and should produce a time step so that the local error estimate is below the given tolerance. The idea, in order to provide such an error estimate is to combine two methods with different orders such that the difference between the two is a conservative error estimate. However, building up a lower order method from a given one, should not result in an important increase of the computational effort. Thus the idea of embedded methods in order to minimize the number of function evaluations, and consequently the computational effort.

We will rely on the a Runge-Kutta of order 4 with 4 stages, known as the 3/8 rule [5] :

$$\begin{array}{c|cccc} 0 & & & & \\ 1/3 & 1/3 & & & \\ 2/3 & -1/3 & 1 & & \\ 1 & 1 & -1 & 1 & \\ \hline & 1/8 & 3/8 & 3/8 & 1/8 \end{array} \tag{3}$$

and will construct a 3rd order embedded method. Starting from the k_i values, $i = 1, 2, 3, 4$, obtained through the previous Runge-Kutta method, we will build a $s+1$ stages method of order 3

$$\widehat{y}_1 = y_0 + \Delta t\,(\widehat{b}_1 k_1 + \ldots, \widehat{b}_s k_s + \widehat{b}_{s+1}\, f(t_1, y_1)),$$

where the last point has to be evaluated anyway and $\widehat{b}_{s+1}$ provides more flexibility. The order conditions are obtained using the usual way, except that we have another stage here ($a_{s+1,i} = b_i$, $i = 1, \ldots, s$) and yield four equations:

$$\begin{aligned} \widehat{b}_1 + \widehat{b}_2 + \widehat{b}_3 + \widehat{b}_4 + \widehat{b}_5 &= 1 \\ \widehat{b}_2 c_2 + \widehat{b}_3 c_3 + \widehat{b}_4 + \widehat{b}_5 &= 1/2 \\ \widehat{b}_2 c_2^2 + \widehat{b}_3 c_3^2 + \widehat{b}_4 + \widehat{b}_5 &= 1/3 \\ \widehat{b}_3 a_{32} c_2 + \widehat{b}_4 (a_{42} c_2 + a_{43} c_3) + \widehat{b}_5/2 &= 1/6. \end{aligned} \tag{4}$$

We have five unknowns and four equations. We choose $\widehat{b}_5 = 1/6$ and obtain:

$$\widehat{b}_1 = 2b_1 - 1/6, \quad \widehat{b}_2 = 2(1 - c_2)b_2, \quad \widehat{b}_3 = 2(1 - c_3)b_3, \quad \widehat{b}_4 = 0,$$

Thus, using a time step Δt, we obtain:

$$y_1 - \widehat{y}_1 = y_1 - y(t_0 + \Delta t) + y(t_0 + \Delta t) - \widehat{y}_1 = O((\Delta t)^{p+1}) + O((\Delta t)^{\widehat{p}+1}) \approx C\,(\Delta t)^{\widehat{p}+1}. \tag{5}$$

The optimal time step Δt_{opt} is given by the fact that

$$Tol \approx C\,(\Delta t_{\text{opt}})^{\widehat{p}+1}$$

so that by eliminating the constant C between the last two equations we get

$$\Delta t_{\text{opt}} = 0.9\,\Delta t \sqrt[\widehat{p}+1]{\frac{Tol}{||y_1 - \widehat{y}_1||}}, \tag{6}$$

where the 0.9 factor is called a security factor.

For practical purposes and robustness of the method, it is standard to replace the last evaluation of the time step by

$$\Delta t_{\text{opt}} = \Delta t \, \min\left(5, \max\left(0.2, 0.9 \sqrt[\widehat{p}+1]{\frac{Tol}{||y_1 - \widehat{y}_1||}}\right)\right), \tag{7}$$

and the norm is taken as a mix of the relative and absolute l^2 norm:

$$||y_1 - \widehat{y}_1|| = \sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(\frac{y_{j1} - \widehat{y}_{j1}}{1 + max(|y_{j0}|, |y_{j1}|)}\right)^2} \tag{8}$$

In order to be clear on what has been coded in the notebook, here is the algorithm, which starts from an initial condition y_0, a tolerance Tol and a given time step Δt_1 at $n = 1$:

Algorithm 1 Automatic selection of the adaptive time step

A) With the current time step Δt_n and from y_{n-1} evaluate y_n, $\widehat{y}_n$ and $err = ||y_n - \widehat{y}_n||$ as well as $\Delta t_{\text{opt},n}$ using the definitions above
B) Advance in time or adapt the time step
if $err \leq Tol$ **then**
  (the time step is accepted)
  $t_{n+1} := t_n + \Delta t_n$
  $\Delta t_{n+1} = \min(\Delta t_{\text{opt},n}, t_{end} - \Delta t_n)$
  the new state of the system if taken as y_n, $n := n + 1$
else
  (the time step is rejected)
  $\Delta t_n = \Delta t_{\text{opt},n}$
end if
C) If the current time is $t_n = t_{end}$ the simulation is over, else we start again at A)

Appendix B: Van der Pol's equation

The first examples have been introduced by Rayleigh in 1883 [10] and then by Baltasar Van der Pol (1920-1926) in a series of papers on nonlinear oscillations [11, 12]. In general, given α, the equation of the oscillator:

$$\mathrm{d}_t^2 y + \alpha \mathrm{d}_t y + y = 0, \tag{9}$$

admits a damped solution for $\alpha > 0$ and an unstable solution for $\alpha < 0$. The idea here is to make α depend on the solution so that it is positive for large enough y and negative for small enough y. One possibility is to take $\alpha(y) = \epsilon(y^2 - 1)$, $\epsilon > 0$, that is:

$$\mathrm{d}_t^2 y + \epsilon(y^2 - 1)\mathrm{d}_t y + y = 0, \tag{10}$$

or switching to a first order system of equations:

$$\begin{cases} \mathrm{d}_t y_1 = y_2, \\ \mathrm{d}_t y_2 = \varepsilon\left(1 - y_1^2\right) y_2 - y_1, \end{cases} \tag{11}$$

which is the system we will study. One expects the existence of a limit cycle and the convergence of the dynamics toward a unique periodic solution in time, whatever the initial condition, or at least in a neighborhood of the limit cycle. The existence of a limit cycle can be studied using the notion of Poincaré map [9] and uniqueness was first proved by Liénard in 1928 [8, 3]. Such a result can also be obtained by Poincaré-Bendixon Theorem, which we will study in the remaining of the course.

For ϵ sufficiently large, we can look for quasi-stationary states through a singular perturbation analysis. For that purpose, we change the time scale and introduce $\tau = t/\epsilon$. We then take $z_1 = y_1$, $z_2 = \epsilon y_2$ and introduce $\mu = 1/\epsilon^2$:

$$\begin{cases} \mathrm{d}_\tau z_1 = z_2, \\ \mu \mathrm{d}_\tau z_2 = (1 - z_1^2)\, z_2 - z_1. \end{cases} \tag{12}$$

As μ tends to 0 with $\epsilon \to +\infty$, the dynamics of the systems lies on the slow manifold:

$$\bar{z}_2 = \frac{\bar{z}_1}{1 - \bar{z}_1^2}, \qquad \bar{y}_2 = \frac{\bar{y}_1}{\epsilon(1 - \bar{y}_1^2)}, \tag{13}$$

and injecting this relation in the first equation, the slow dynamics of the system reads:

$$\mathrm{d}_t \bar{y}_1 = \frac{\bar{y}_1}{\epsilon(1 - \bar{y}_1^2)}, \tag{14}$$

for which we have a solution:

$$\log(\bar{y}_1) - \frac{\bar{y}_1}{2} = \frac{t - t_0}{\epsilon} + cst. \tag{15}$$

It can be shown that the solution lives on such quasi-stationary states for long periods of time before switching abruptly to another branch, all the more abruptly as ϵ is small.

References

[1] I.R. Epstein and J.A. Pojman. *An Introduction to Nonlinear Chemical Dynamics.* Oxford University Press, 1998. Oscillations, Waves, Patterns and Chaos.

[2] J.-M. Ginoux and C. Letellier. Van der pol and the history of relaxation oscillations: Toward the emergence of a concept. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(2):023120, 2012.

[3] Jean-Marc Ginoux. Self-Excited Oscillations : from Poincare to Andronov. *Nieuw Archief voor Wiskunde*, 13(3):170–177, 2012.

[4] P. Gray and S.K. Scott. *Chemical Oscillations and Instabilities.* Oxford Univ. Press, 1994.

[5] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. I - Nonstiff problems.* Springer-Verlag, Berlin, 1987.

[6] E. Hairer and G. Wanner. *Solving ordinary differential equations. II - Stiff and differential-algebraic problems.* Springer-Verlag, Berlin, 1991.

[7] R. Lefever and G. Nicolis. Chemical instabilities and sustained oscillations. *J. Theoretical Biology*, 30(2):267–284, 1971.

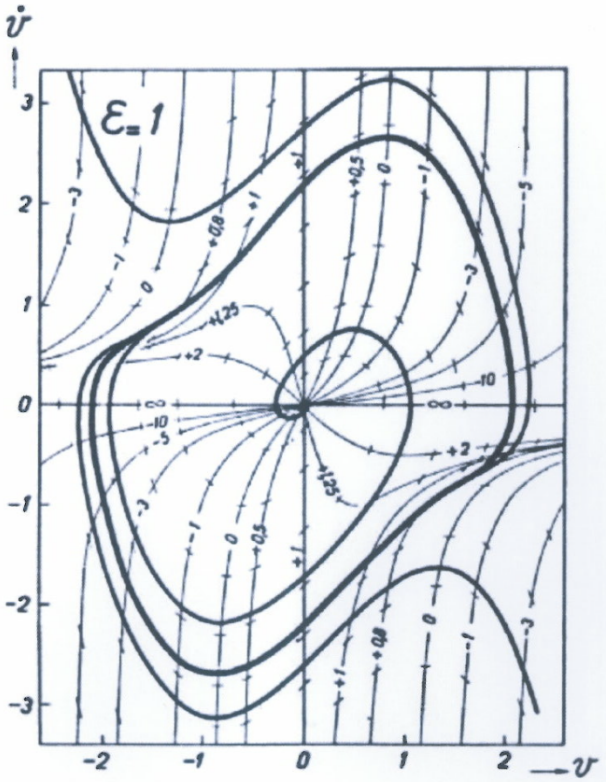
[8] A. Liénard. Etude des oscillations entretenues. *Revue générale de l'électricité*, 23:901–912 and 946–954, 1928.

[9] Henri Poincaré. Mémoire sur les courbes définies par une équation différentielle (2nde partie). *Journal de mathématiques pures et appliquées*, 8:251–296, 1882.

[10] Lord Rayleigh. Etude des oscillations entretenues. *Phil. Mag.*, 15:229–235, 1883.

[11] B. van der Pol. On "relaxation-oscillations". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 7:978–992, 1926.

[12] B. van der Pol. *Selected scientific papers.* 2 vols. Edited by H. Bremmer and C. J. Bouwkamp, with an introduction bt H. B. G. Gasimir. North-Holland Publishing Co., Amsterdam, 1960.



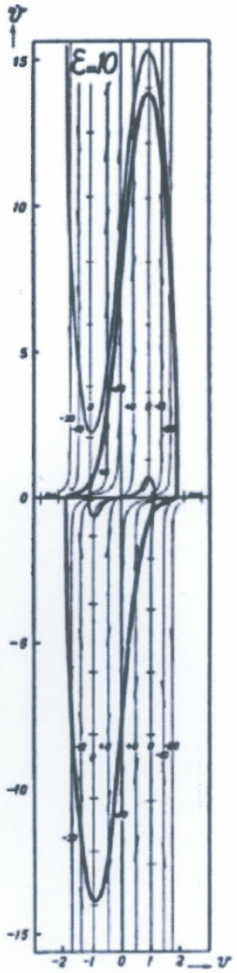


Fig. 16.1. Illustrations from Van der Pol (1926) (with permission)

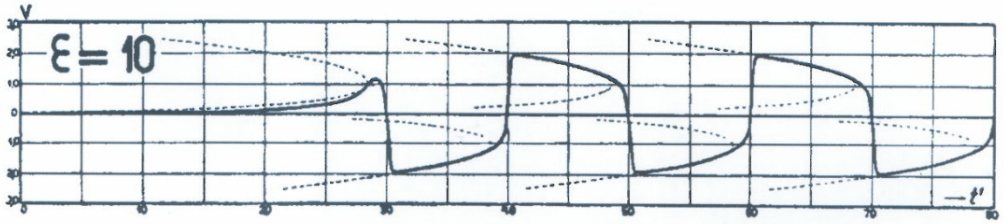


Fig. 16.3. Solution of Van der Pol's equation for $\varepsilon = 10$ compared with steady state approximations

Figure 1: Phase portrait and solution of the Van der Pol equation reproduced from the original work in [5].