

## *études*

---

### *la résolution numérique des équations différentielles : un thème abordable par des lycéens*

*par Daniel Reisz*

L'objet de cet article est de présenter quelques méthodes classiques de résolution approchée d'équations différentielles du premier ordre, méthodes qui sont *dans leur principe* compréhensibles par des élèves de Première et de Terminales scientifiques et dont l'exécution est possible sur des calculatrices programmables ou des micro-ordinateurs. Avec une idée didactique plus fondamentale : ouvrir, par là, un champ de problèmes intéressants et significatifs souvent d'origine extra-mathématique (évolution de systèmes économiques, biologiques, démographiques, problèmes d'origine physique,...).

Mais que mon propos soit clair : il ne s'agit en aucun cas de pousser à un alourdissement de ce qui se fait dans nos classes. Il y a une différence fondamentale entre un éventuel thème d'activités et l'étude, *in extenso*, d'une question, entre une approche encore très expérimentale et intuitive et son étude théorique. Et je réponds par avance à une critique que l'on entend souvent dans la bouche de certains collègues des classes préparatoires ou de l'enseignement supérieur : "Tout effleurer, sans jamais rien approfondir, est très mauvais, très démobilisateur !". A cette critique

souvent émise dans le contexte général des actuels programmes, je ferai une réponse de Normand : oui et non. Oui, si on entend par "effleurer" un bavardage papillonnant qui ne s'appuie sur aucun acquis, sur aucune activité significative ; non, si on sous-entend que l'appropriation des concepts mathématiques se fait, ipso facto, par leur formulation théorique définitive et générale. Je crois qu'il faut du temps, des activités très diverses, des échecs, des questions, des périodes de maturation, pour aller d'une simple prise de conscience à l'étude théorique "définitive" en passant par des stades d'approfondissement successifs et diversifiés. Toute l'histoire des mathématiques plaide en faveur d'un tel processus. Rappelons-nous toujours que Bourbaki est postérieur à Archimède, Newton, Euler, Lagrange, ...

C'est dans ce contexte que j'ai voulu aborder le thème de cet article. Je n'ai aucun mérite quant à son contenu : j'ai pillé, sans vergogne, les idées de l'excellent livre "*Differential equations and their applications*" de Martin BRAUN (Springer, 1978) dont il existe d'ailleurs, chez le même éditeur, une traduction allemande. C'est aussi à cet ouvrage qu'on peut se reporter pour les justifications théoriques qui ne figurent pas dans ce texte. Ceci est signalé, dans la suite de l'article par [BRAUN].

## 1. Le problème

Soit une équation différentielle du premier ordre, supposée réduite à la forme

$$\frac{dy}{dt} = f(t, y) \quad (1)$$

Les solutions de l'équation différentielle (1) sont toutes les fonctions  $t \rightarrow y(t)$  qui vérifient (1).

Certains types d'équations différentielles se résolvent explicitement, c'est-à-dire qu'on est capable d'écrire explicitement les fonctions  $y(t)$ . Il en est ainsi par exemple d'équations de la forme

$$\frac{dy}{dt} = g(t) \quad \text{lorsqu'on connaît une primitive de } g$$

$$\frac{dy}{dt} = a(t)y + b(t) \quad (\text{équation linéaire})$$

$$\frac{dy}{dt} = \frac{g(t)}{f(y)} \quad (\text{équation à variables séparées}) \text{ lorsqu'on connaît une primitive de } f \text{ et de } g$$

Dans l'annexe nous montrerons qu'en réalité il est *tout à fait exceptionnel* de pouvoir envisager une résolution explicite d'une équation différentielle, d'où l'intérêt de pouvoir les résoudre de façon approchée. Cet intérêt est encore accentué par le fait que même lorsqu'une résolution explicite est théoriquement possible, celle-ci peut être d'un accès techni-

que difficile ou peu utilisable, alors qu'une solution approchée peut éventuellement donner les renseignements nécessaires pour le problème posé.

Dans la pratique on est souvent amené à résoudre le problème suivant : quelles sont, parmi les solutions de l'équation (1) celles qui vérifient les *conditions initiales*

$$y(t_0) = y_0 \quad (t_0 \text{ et } y_0 \text{ donnés})$$

problème que nous regrouperons sous la forme

$$\frac{dy}{dt} = f(t, y) \quad , \quad y(t_0) = y_0 \quad (3)$$

Il se pose alors trois questions :

1) Le problème admet-il des solutions ? (alors que nous ne pouvons pas les expliciter).

2) Le problème admet-il une seule solution ? (problème pratique essentiel).

3) Pourquoi se donner la peine de répondre aux deux questions précédentes si, de toute façon, nous ne pouvons pas expliciter la ou les solutions ?

La réponse à la dernière question réside dans le fait que dans la pratique, il suffit le plus souvent d'avoir pour  $y(t)$ , un tableau de valeurs numériques à une précision suffisante sur un intervalle utile. La réponse aux deux premières questions nécessite la mise en place et l'étude de la procédure d'itération de PICARD et n'est pas du ressort de cet article.

Disons, de façon simpliste, que si  $f$  et  $\frac{\partial f}{\partial y}$  sont continues sur des intervalles convenables, alors (3) admet une solution et une seule (voir par exemple [BRAUN]).

## 2. Une première méthode d'approximation : la méthode d'EULER

Soit une équation différentielle du premier ordre, soumise à des conditions initiales :

$$\frac{dy}{dt} = f(t, y) \quad , \quad y(t_0) = y_0 \quad (3)$$

Une calculatrice ou un micro-ordinateur ne peut pas (pour l'instant) fournir explicitement la solution  $y(t)$  sur un intervalle  $[a, b]$ . Il peut tout au plus fournir pour des valeurs  $t_0 = a, t_1, t_2, \dots, t_n = b$  des valeurs approchées  $y_0, y_1, \dots, y_n$  des quantités  $y(t_0), y(t_1), \dots, y(t_n)$ . Pour simplifier nous supposons la subdivision  $t_0, t_1, \dots, t_n$  de  $[a, b]$  régulière, c'est-à-dire que

$$t_k = t_0 + k h \quad \text{avec} \quad h = \frac{b-a}{n} \quad (\text{pas de la subdivision})$$

$$\text{et} \quad k = 0, 1, \dots, n$$

On connaît déjà le couple initial  $(t_0, y_0)$ . A partir de cette information on voudrait une valeur approchée  $y_1$  de  $y(t_1)$ , puis, à partir de cette valeur approchée  $y_1$ , une valeur approchée  $y_2$  de  $y(t_2)$ , etc.

Pour cela on peut utiliser :

— en classe de Première, l'approximation affine d'une fonction en un point :

$$y(t_{k+1}) = y(t_k + h) = y(t_k) + h \frac{dy}{dt}(t_k) + h\varepsilon(h)$$

— en classe de Terminale, le développement de Taylor jusqu'au second ordre, que l'on établit sans peine par intégrations successives, en partant de l'encadrement

$$m \leq y''(t) \leq M \text{ pour } t_k \leq t \leq t_{k+1}$$

$$y(t_{k+1}) = y(t_k + h) = y(t_k) + h \frac{dy}{dt}(t_k) + \frac{h^2}{2} \frac{d^2y}{dt^2}(t_k) + R_2(h)$$

$$\text{avec } |R_2(h)| < M \frac{h^3}{6}$$

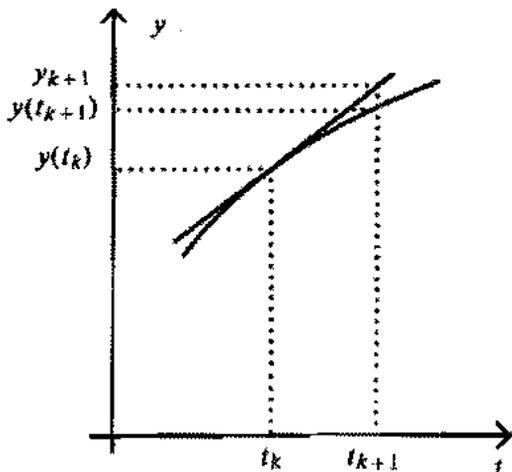
La méthode d'Euler repose sur le fait que l'on peut donc déterminer une approximation de  $y(t_{k+1})$  à partir de  $y(t_k)$ , de proche en proche, à partir de la donnée initiale  $y(t_0) = y_0$  et que l'équation initiale (3) donne immédiatement :

$$\frac{dy}{dt}(t_k) = f(t_k, y(t_k))$$

d'où l'idée toute naturelle d'utiliser l'approximation affine de  $y(t)$  en  $t_k$  :

$$y(t_{k+1}) = y(t_k) + h \frac{dy}{dt}(t_k) + h\varepsilon(h)$$

en négligeant le terme  $h\varepsilon(h)$ .



Il en résulte le processus d'approximations successives suivant :

$$\begin{aligned} y_0 &= y(t_0) \text{ donné} \\ y_{k+1} &= y_k + hf(t_k, y_k) \end{aligned} \quad (5)$$

Disons un mot de l'erreur commise au  $k^{\text{ième}}$  pas. On démontre, mais cela est hors de portée d'un lycéen, que

$$|y_k - y(t_k)| \leq A h$$

où  $A$  est une constante dont la valeur importe peu dans la pratique présente. En effet, ce simple résultat théorique justifie la règle de conduite suivante : en remplaçant  $h$  par  $\frac{h}{2}$  l'erreur est, grosso modo, réduite de moitié.

Il faut aussi noter que des approximations itératives, à partir d'une valeur exacte  $y_0$ , accumulent les erreurs et, plus on s'éloigne de  $t_0$ , moins l'approximation  $y_k$  de  $y(t_k)$  sera bonne. Or, dans la pratique, il est bien évident que les valeurs de  $y(t_k)$  dont on a besoin ne correspondent précisément pas à des valeurs de  $t$  "voisines" de  $t_0$ .

En réalité, cette méthode d'Euler est surtout instructive sur le plan des idées, mais trop peu performante dans la pratique pour être souvent utilisée.

### 3. Une amélioration de la méthode d'EULER

Une première amélioration, que l'on pourrait envisager serait de prendre en compte, dans le développement de Taylor, un terme de plus :

$$y_{k+1} = y_k + h \frac{dy}{dt}(t_k) + \frac{h^2}{2} \frac{d^2y}{dt^2}(t_k)$$

Cette idée a un inconvénient à la fois scientifique et didactique. En effet on n'a pas directement

$$\frac{d^2y}{dt^2}(t_k)$$

et il faut alors utiliser le fait que

$$\frac{d^2y}{dt^2} = \frac{\partial f}{\partial t} + f \frac{\partial f}{\partial y}$$

d'où

$$\frac{d^2y}{dt^2}(t_k) = \left[ \frac{\partial f}{\partial t} + f \frac{\partial f}{\partial y} \right](t_k, y(t_k))$$

Ceci est à la fois lourd à calculer et en dehors du programme d'analyse des lycées.

Une autre idée, d'efficacité comparable à la précédente, mais plus accessible aux élèves et aux calculatrices, repose sur la constatation suivante : l'équation différentielle

$$y' = f(t, y)$$

intégrée entre  $t_k$  et  $t_{k+1}$ , donne

$$y(t_{k+1}) - y(t_k) = \int_{t_k}^{t_{k+1}} f(t, y(t)) dt$$

soit

$$y(t_{k+1}) = y(t_k) + \int_{t_k}^{t_{k+1}} f(t, y(t)) dt \quad (6)$$

Si nous comparons (6) à la formule itérative de la méthode d'Euler

$$y_{k+1} = y_k + h f(t_k, y_k)$$

et si nous assimilons  $y_{k+1}$  à  $y(t_{k+1})$ ,  $y_k$  à  $y(t_k)$  nous voyons (fig. 2) que l'approximation réalisée en utilisant la méthode d'Euler consiste, en gros, à remplacer l'aire dans la courbe  $Y = Y(t) = f(t, y(t))$  par l'aire  $y_k$  du rectangle hachuré.

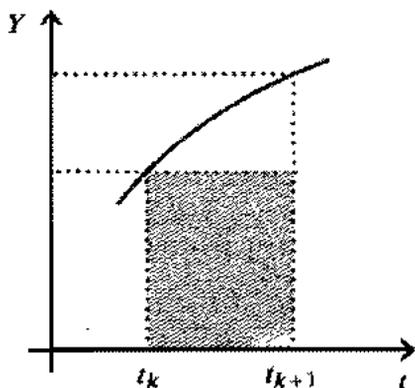


Figure 2

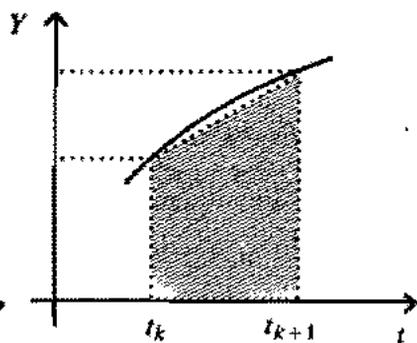


Figure 3

Il est alors naturel de préférer l'approximation trapézoïdale suggérée par la figure 3, c'est-à-dire remplacer la relation récurrentielle de la méthode d'Euler :

$$y_{k+1} = y_k + h f(t_k, y_k)$$

par

$$y_{k+1} = y_k + \frac{h}{2} [f(t_k, y(t_k)) + f(t_{k+1}, y(t_{k+1}))]$$

ou, numériquement,

$$y_{k+1} = y_k + \frac{h}{2} [f(t_k, y_k) + f(t_{k+1}, y_{k+1})]$$

On objectera évidemment à cela que dans le second membre figure  $y_{k+1}$  qu'on est précisément entrain de calculer. Toute l'astuce consiste alors,

dans ce second membre, à utiliser pour  $y_{k+1}$  la valeur approchée que donne précisément la méthode d'Euler, soit

$$y_{k+1} = y_k + h f(t_k, y_k)$$

d'où

$$y_{k+1} = y_k + \frac{h}{2} [f(t_k, y_k) + f(t_k, y_k + h f(t_k, y_k))]$$

On peut montrer qu'avec cette méthode (voir [BRAUN]), hors de la portée d'un lycéen)

$$|y(t_k) - y_k| < B h^2$$

où  $B$  est une constante. D'où la règle de conduite suivante : le remplacement de  $h$  par  $\frac{h}{2}$  divise, grosso modo, l'erreur par 4.

#### 4. La méthode de Runge-Kutta

Nous allons, en dernier lieu, proposer, sans aucune justification théorique, une méthode mise au point vers 1900 par les mathématiciens RUNGE et KUTTA. Cette méthode est suffisamment performante pour rester encore aujourd'hui l'une des plus utilisées. On démontre en effet que l'erreur y est majorée par

$$|y(t_k) - y_k| < C h^4 \text{ où } C \text{ est une constante.}$$

Le remplacement de  $h$  par  $\frac{h}{2}$  divise donc, grosso modo, l'erreur par 16.

Cette méthode consiste à utiliser la formule

$$\begin{cases} y_{k+1} = y_k + \frac{h}{6} [L_{k,1} + 2 L_{k,2} + 2 L_{k,3} + L_{k,4}] \\ y_0 \text{ donné} \end{cases}$$

où

$$L_{k,1} = f(t_k, y_k)$$

$$L_{k,2} = f\left(t_k + \frac{h}{2}, y_k + \frac{h}{2} L_{k,1}\right)$$

$$L_{k,3} = f\left(t_k + \frac{h}{2}, y_k + \frac{h}{2} L_{k,2}\right)$$

$$L_{k,4} = f(t_k + h, y_k + h L_{k,3})$$

On peut, pour avoir une petite idée des raisons d'être de cette formule, voir dans la quantité

$$\frac{1}{6} [L_{k,1} + 2 L_{k,2} + 2 L_{k,3} + L_{k,4}]$$

une moyenne pondérée de différentes valeurs de  $f(t, y)$  entre  $(t_k, y_k)$  et  $(t_{k+1}, y_{k+1})$ , c'est-à-dire que le trapèze décrit à la figure 3 est remplacé par un trapèze du genre de celui de la figure 4.

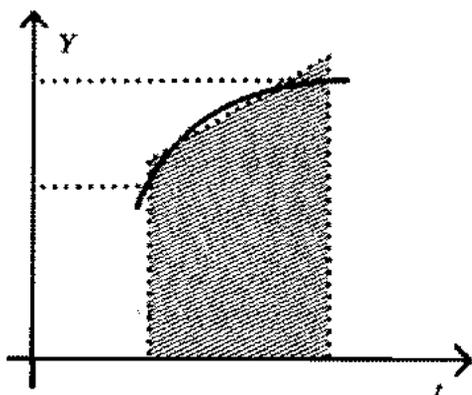


Fig.4

### 5. Quand s'arrêter ? Quelle méthode choisir ?

Quand s'arrêter ? Ici, pour ce type d'approximations, la pratique (que l'on peut justifier) est la suivante :

on remplace  $h$  par  $\frac{h}{2}$  et on compare les résultats. Si ces résultats ne sont pas stables, modulo la précision désirée, on recommence. Sinon on recommence encore *une* fois et si la stabilité se maintient toujours on peut être assuré que  $y_k$  est une approximation de  $y(t_k)$  à la précision désirée.

On mettra en garde les élèves sur le fait que de telles pratiques ne sont pas généralisables à n'importe quel type de convergence (pratique assez répandue chez les élèves...).

Quelle méthode choisir ? Une réponse hâtive de mathématicien superficiel reposerait sur l'analyse illustrée par l'exemple simpliste suivant : on veut que

$$|y_k - y(t_k)| < 10^{-8}$$

et on sait que sur l'intervalle  $[0;1]$ , où on est censé se trouver, on a

$$|y_k - y(t_k)| < Ah \quad \text{avec Euler}$$

$$|y_k - y(t_k)| < Ah^2 \quad \text{avec Euler amélioré}$$

$$|y_k - y(t_k)| < Ah^4 \quad \text{avec Runge-Kutta}$$

$$\text{avec } h = \frac{1}{n}$$

Il faut donc s'assurer que

$$Ah < 10^{-8} \implies n > A \cdot 10^8 \quad \text{(Euler)}$$

$$Ah^2 < 10^{-8} \implies n > \sqrt{A} \cdot 10^4 \quad \text{(Euler amélioré)}$$

$$Ah^4 < 10^{-8} \implies n > A^{1/4} \cdot 10^2 \quad \text{(Runge-Kutta)}$$

c'est-à-dire qu'on passe de l'ordre de  $10^6$  itérations avec la méthode d'Euler à l'ordre de  $10^2$  itérations avec la méthode de Runge-Kutta.

Attention, ce serait confondre vitesse de convergence d'une suite et performance d'un algorithme. En effet la méthode de Runge-Kutta nécessite, à chaque itération, le calcul de quatre valeurs de  $f(t, y)$ , alors que la méthode d'Euler n'en nécessite qu'un seul. Mais l'efficacité de la méthode de Runge-Kutta est telle que le mathématicien superficiel aura fait le bon choix !

## 6. Quelques exemples pour investir

1) Comparer, sur l'intervalle  $[0; 1]$  les trois méthodes et comparer ces résultats à ceux donnés par la solution explicite, pour

$$y' = 1 + (y-t)^2 ; \quad y(0) = \frac{1}{2}$$

$$\text{(solution explicite : } y(t) = t + \frac{1}{2-t}\text{)}$$

2) Connaissant  $y(0)$ , donner des approximations de  $y(1)$  avec les trois méthodes présentées et comparer ces approximations aux résultats donnés par la solution explicite (on commence avec un pas  $h = 0,1$ ):

(a)  $y' = 2ty ; \quad y(0) = 2$

$$\text{(solution explicite } y = 2e^{t^2}\text{)}$$

(b)  $y' = -1 + 2t + \frac{y^2}{(1+t^2)^2} ; \quad y(0) = 1 \quad (y = 1+t^2)$

(c)  $y' = te^{-y} + \frac{t}{1+t^2} ; \quad y(0) = 0 \quad (y = \ln(1+t^2))$

3) Donner avec 4 décimales exactes, la valeur approchée de  $y(1)$ , sachant que

(a)  $y' = y(1+e^{-y}) + e^t ; \quad y(0) = 0$

(b)  $y' = ty^3 - y ; \quad y(0) = 1$

(c)  $y' = \frac{t^2 + y^2}{1+t+y^2} ; \quad y(0) = 0$

(d)  $y' = \sin(y+t) + e\sqrt{t} ; \quad y(0) = 0,7$

4) On sait que  $y(t)$  vérifie

$$y' + ty = \varphi(t) ; \quad y(0) = 1$$

où  $\varphi(t)$  et  $\varphi'(t)$  sont connus par le tableau de valeurs suivant :

$t$	$\varphi(t)$	$\varphi'(t)$	$t$	$\varphi(t)$	$\varphi'(t)$
0,0	1	0	0,6	1,16412	0,49520
0,1	1,00499	0,99500	0,7	1,21579	0,53539
0,2	1,01980	0,19601	0,8	1,27059	0,55736
0,3	1,04399	0,28660	0,9	1,32660	0,55945
0,4	1,07683	0,36842	1,0	1,38177	0,54030
0,5	1,11730	0,43879			

Utiliser la méthode d'Euler pour avoir une approximation de  $y(t)$  pour  $t = 0,1; 0,2; \dots; 0,9; 1$ .

5) On sait que  $e^{-x^2}$  n'admet pas de primitive élémentaire.

a) Se servir du fait que  $x \mapsto y = e^{x^2} \int_0^x e^{-t^2} dt$  vérifie  $y' = 2xy + 1$  pour calculer, avec cinq décimales exactes

$$\int_0^{0,5} e^{-x^2} dx$$

b) Comparer cette méthode à une méthode d'intégration numérique (rectangle, trapèze, Simpson, ...).

## 7. Annexe

L'objet de cette annexe est de montrer de façon un peu plus précise pourquoi la possibilité de résoudre explicitement une équation différentielle du premier ordre est une situation exceptionnelle.

Soit donc l'équation

$$\frac{dy}{dt} = f(t, y) \quad (1)$$

Résoudre de façon explicite (1) nécessite de pouvoir mettre (1) sous la forme

$$\frac{d}{dt} \varphi(t, y) = 0 \quad (2)$$

qui mène, par intégration, à

$$\varphi(t, y) = c \quad (c \text{ étant une constante arbitraire})$$

forme qu'on peut espérer résoudre en  $y(t)$ .

A quelle condition peut-on mettre (1) sous la forme (2) ? Pour étudier cela, remarquons que

$$\frac{d}{dt} \varphi(t, y(t)) = \frac{\partial \varphi}{\partial t} + \frac{\partial \varphi}{\partial y} \cdot \frac{dy}{dt}$$

ce qui montre qu'une équation

$$M(t,y) + N(t,y) \cdot \frac{dy}{dt} = 0$$

ne peut s'écrire sous la forme  $\frac{d}{dt} \varphi(t,y) = 0$  que s'il existe une fonction  $\varphi(t,y)$  qui vérifie

$$M(t,y) = \frac{\partial \varphi}{\partial t} \quad \text{et} \quad N(t,y) = \frac{\partial \varphi}{\partial y}$$

La réponse à notre question est *presque toujours* négative, comme le montre le théorème classique suivant :

*Théorème : Soit  $M(t,y)$  et  $N(t,y)$  deux fonctions continues en  $t$  et  $y$ , aux dérivées partielles du premier ordre continues, pour  $(t,y)$  appartenant à l'ouvert  $]a,b[ \times ]c,d[$ . Il existe une fonction  $\varphi(t,y)$  telle que*

$$\frac{\partial \varphi}{\partial t} = M(t,y) \quad \text{et} \quad \frac{\partial \varphi}{\partial y} = N(t,y)$$

*si et seulement si*

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial t} \quad \text{sur l'ouvert } ]a,b[ \times ]c,d[$$

(Pour la démonstration de ce théorème on se reportera par exemple à [BRAUN]).

Cette contrainte  $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial t}$  n'étant *presque jamais* satisfaite, il en résulte que la résolution explicite de l'équation (1) est *presque toujours* impossible.

**Vous vous êtes promis d'assister  
aux Journées Nationales A.P.M.E.P...  
Remplissez sans tarder votre  
fiche d'inscription (page 329)**