Il est intéressant de comparer sur une base aussi rigoureuse que possible les différentes méthodes vues jusqu'à maintenant. On a constaté que plus l'ordre d'une méthode est élevé, plus cette méthode est précise. Par contre, plus l'ordre de la méthode est élevé, plus elle est coûteuse en temps de calcul. Par exemple, la méthode d'Euler (d'ordre 1) ne nécessite qu'une seule évaluation de la fonction $f(t, y)$ à chaque pas de temps, alors que la méthode d'Euler modifiée (d'ordre 2) en demande 2 et que la méthode de Runge-Kutta d'ordre 4 exige 4 évaluations de la même fonction. En d'autres termes, la méthode de Runge-Kutta d'ordre 4 demande à peu près deux fois plus de calculs que la méthode d'Euler modifiée et quatre fois plus que la méthode d'Euler.

Il est raisonnable de se demander s'il n'est pas préférable d'utiliser la méthode d'Euler avec un pas de temps 4 fois plus petit ou la méthode d'Euler modifiée d'ordre 2 avec un pas de temps 2 fois plus petit, plutôt que de se servir de la méthode de Runge-Kutta d'ordre 4. L'exemple qui suit permet de comparer les différentes méthodes sur une base plus équitable.

Exemple 7.9

On considère l'équation différentielle habituelle :

$$y'(t) = -y(t) + t + 1 \quad (y(0) = 1)$$

On recourt à 3 méthodes de résolution : la méthode d'Euler avec un pas $h = 0{,}025$, la méthode d'Euler modifiée avec $h = 0{,}05$ et la méthode de Runge-Kutta d'ordre 4 avec $h = 0{,}1$. Ces valeurs de h permettent de comparer ces 3 méthodes sur la base de coûts de calculs à peu près équivalents. Le tableau suivant présente les résultats obtenus en $t = 1$ pour ces différents choix. La valeur exacte de la solution est $y(1) = 1{,}367\,879\,4412$.

Comparaison des différentes méthodes : $y'(t) = -y(t) + t + 1$				
Méthode	h	Nombre de pas	Résultat	Erreur
Euler	0,025	40	1,363 232 374 17	$0{,}464 \times 10^{-2}$
Euler modifiée	0,05	20	1,368 038 621 67	$0{,}159 \times 10^{-3}$
Runge-Kutta	0,1	10	1,367 879 774 41	$0{,}333 \times 10^{-6}$

Les résultats sont éloquents. Même en prenant un pas de temps quatre fois plus petit, la méthode d'Euler reste très imprécise par rapport à celle de Runge-Kutta d'ordre 4. On peut porter le même jugement sur la méthode d'Euler modifiée. *Il est donc généralement préférable d'utiliser des méthodes d'ordre aussi élevé que possible.* ∎

7.4.3 Contrôle de l'erreur

Nous n'avons considéré jusqu'à maintenant que des méthodes de résolution à pas de temps h fixé dès le départ. Ces méthodes sont certes très utiles, mais souffrent de 2 inconvénients majeurs :

– il faut déterminer le pas de temps h un peu au hasard ;

– il n'y a aucun contrôle sur la précision des résultats.

Il est cependant important dans certaines situations de bien contrôler la précision des résultats numériques. Ce sera le cas lorsque la solution d'un problème présente de très brusques variations dans certaines régions. Une valeur donnée du pas de temps h peut être tout à fait adéquate dans certaines régions et trop grande (ou trop petite) ailleurs.

L'idée est alors de se servir du pas de temps h pour contrôler la précision. Si la solution présente des variations brusques, on prendra un pas de temps plus petit et on l'augmentera éventuellement là où la solution varie plus lentement.

La famille des méthodes de Runge-Kutta se prête parfaitement à cet exercice. Tout comme pour les méthodes de Runge-Kutta d'ordre 2 (section 7.4.1), on sait qu'il existe toute une famille de méthodes de Runge-Kutta d'ordre 4, 5 et même plus. Nous en avons vu une variante d'ordre 4 à la section 7.4.2. Il existe cependant une combinaison particulièrement utile due à Fehlberg [14] et que nous allons maintenant décrire. On définit en premier lieu les 6 constantes suivantes :

$$
\begin{aligned}
k_1 &= hf(t_i, y_i) \\
k_2 &= hf\left(t_i + \frac{1}{4}h, y_i + \frac{1}{4}k_1\right) \\
k_3 &= hf\left(t_i + \frac{3}{8}h, y_i + \frac{3}{32}k_1 + \frac{9}{32}k_2\right) \\
k_4 &= hf\left(t_i + \frac{12}{13}h, y_i + \frac{1932}{2197}k_1 - \frac{7200}{2197}k_2 + \frac{7296}{2197}k_3\right) \\
k_5 &= hf\left(t_i + h, y_i + \frac{439}{216}k_1 - 8k_2 + \frac{3680}{513}k_3 - \frac{845}{4104}k_4\right) \\
k_6 &= hf\left(t_i + \frac{1}{2}h, y_i - \frac{8}{27}k_1 + 2k_2 - \frac{3544}{2565}k_3 + \frac{1859}{4104}k_4 - \frac{11}{40}k_5\right)
\end{aligned}
$$

À l'aide de ces constantes k_i, on construit une méthode de Runge-Kutta d'ordre 4 ($\tau_{n+1}(h) = O(h^4)$) :

$$
y_{n+1} = y_n + \left(\frac{25}{216}k_1 + \frac{1408}{2565}k_3 + \frac{2197}{4104}k_4 - \frac{1}{5}k_5\right) = y_n + h\phi(t_n, y_n)
$$

et une autre d'ordre 5 ($\tilde{\tau}_{n+1}(h) = O(h^5)$) :

$$\tilde{y}_{n+1} = y_n + \left(\frac{16}{135}k_1 + \frac{6656}{12\,825}k_3 + \frac{28\,561}{56\,430}k_4 - \frac{9}{50}k_5 + \frac{2}{55}k_6\right) = y_n + h\tilde{\phi}(t_n, y_n)$$

qui sont 2 approximations d'ordres différents de la même quantité. L'intérêt de cette combinaison de deux méthodes est qu'elle ne requiert le calcul que de 6 constantes k_i au total, et ce, pour les deux approximations.

La méthode de base sera celle d'ordre 4 et l'on se servira de la méthode d'ordre 5 pour contrôler la valeur du pas de temps h et par le fait même l'erreur. Nous supposerons dans ce qui suit que l'algorithme débute à partir d'une valeur y_n exacte, c'est-à-dire $y_n = y(t_n)$. Cette hypothèse est importante, car elle suppose qu'aucune erreur n'a été accumulée jusqu'au pas de temps t_n. Puisque nous avons supposé que $y_n = y(t_n)$, on a :

$$\begin{aligned} y(t_{n+1}) - y_{n+1} &= y(t_{n+1}) - y_n - h\phi(t_n, y_n) \\ &= y(t_{n+1}) - y(t_n) - h\phi(t_n, y(t_n)) \\ &= h\tau_{n+1}(h) = O(h^5) \end{aligned}$$

De même, on montre que :

$$y(t_{n+1}) - \tilde{y}_{n+1} = h\tilde{\tau}_{n+1}(h) = O(h^6)$$

ce qui entraîne que :

$$y(t_{n+1}) - y_{n+1} = (y(t_{n+1}) - \tilde{y}_{n+1}) + (\tilde{y}_{n+1} - y_{n+1})$$

et donc que :

$$h\tau_{n+1}(h) = h\tilde{\tau}_{n+1}(h) + (\tilde{y}_{n+1} - y_{n+1})$$

En regardant de chaque côté de l'égalité, on constate que :

$$O(h^5) = O(h^6) + (\tilde{y}_{n+1} - y_{n+1})$$

Pour que l'égalité puisse avoir lieu, il faut que :

$$y(t_{n+1}) - y_{n+1} \simeq (\tilde{y}_{n+1} - y_{n+1}) = O(h^5)$$

et l'on en conclut que la partie la plus importante de l'erreur commise fait tout simplement intervenir la différence entre les approximations d'ordre 4 et 5. On en conclut également que la quantité E définie par :

$$E = (\tilde{y}_{n+1} - y_{n+1}) = \left(\frac{1}{360}k_1 - \frac{128}{4275}k_3 - \frac{2197}{75\,240}k_4 + \frac{1}{50}k_5 + \frac{2}{55}k_6\right) \tag{7.13}$$

est déterminante pour contrôler l'erreur. En effet, pour une précision *tol* spécifiée à l'avance, on calcule l'approximation 7.13 et l'on modifie h de manière à rencontrer le critère de précision. Cela signifie que la valeur de h sera augmentée si l'erreur constatée est très petite et sera diminuée si l'erreur est jugée trop grande. Pour y arriver, rappelons que :

$$y(t_{n+1}) - y_{n+1} = C(h)^5 \simeq E$$

On remplace alors h par βh de sorte que :

$$y(t_{n+1}) - y_{n+1} = C(\beta h)^5 = tol$$

En faisant le quotient des deux dernières relations, on a :

$$\beta^5 \simeq \frac{tol}{E} \text{ ou encore } \beta \simeq \left(\frac{tol}{E}\right)^{1/5}$$

ce qui nous donne le facteur de réduction (ou d'augmentation si $\beta > 1$) souhaité du pas de temps h. Pour plus de sécurité, on peut introduire un facteur 2 supplémentaire de sorte que :

$$\beta \simeq \left(\frac{tol}{2E}\right)^{1/5} \simeq 0{,}8\left(\frac{tol}{E}\right)^{1/5} \tag{7.14}$$

ce qui donne une meilleure assurance d'atteindre la précision désirée. On remplace ainsi la valeur actuelle de h par une nouvelle quantité βh. En pratique, on évitera de diminuer (d'augmenter) trop fortement la valeur de h en lui fixant une valeur maximale. On peut prendre par exemple :

$$h_{max} = \frac{(t_f - t_0)}{16}$$

où t_f est le temps final que l'on souhaite atteindre. Le facteur 16 est arbitraire, mais semble donner de bons résultats en pratique. De même, on peut prendre :

$$h = \frac{h_{max}}{8}$$

comme valeur initiale du pas de temps. Ainsi, l'algorithme résultant n'exige pas de donner cette valeur initiale puisqu'on la détermine automatiquement.

Exemple 7.10

Un parachutiste de masse $m = 70$ kg tombe en chute libre à partir d'une vitesse initiale nulle. L'air exerce une friction proportionnelle à sa vitesse avec un coefficient de proportionnalité c_f estimé à 13 kg/s. Au bout de 10 secondes, il ouvre son parachute et, heureusement pour lui, le coefficient