

L'Intégration Numérique en Calcul d'Orbites

Cours de l'Ecole GRGS 2002

Pierre EXERTIER^a, David COULOT^{a,b}

^aObservatoire de la Côte d'Azur (CERGA/URA6527), av. Copernic, F-06130 Grasse

^bInstitut Géographique National (LAREG, IGN/ENSG), 6-8 Av. B. Pascal, F-77455 Marne-la-Vallée

5 février 2003

Résumé

Les méthodes d'intégration numérique utilisées en calcul d'orbites sont basées sur le principe de la discrétisation. On cherche des valeurs approchées de la solution exacte $E_j(t)$ (le vecteur des 6 paramètres orbitaux, par exemple), sur un ensemble de points équidistants. L'algorithme consiste alors à faire correspondre, à tous les points d'abscisse t_n (où $t_n = t_0 + nh$, h étant le pas d'intégration), des nombres $(E_j)_n$ qui approchent les valeurs $E_j(t_n)$ de la solution exacte du système différentiel considéré aux points t_n .

Les algorithmes généralement employés aujourd'hui sont dits "à pas liés", c'est-à-dire qu'ils utilisent, à un instant t_n quelconque, les valeurs aux instants t_{n+1-k} déjà calculées. L'avantage de ces méthodes est d'avoir accès à tout moment aux termes $f((E_j)_{n+1-k}, t_{n+1-k})$, tableau qui permet aisément l'interpolation à une date donnée (date de mesure, par exemple). L'inconvénient, par rapport aux méthodes dites "à pas séparés", se situe au démarrage, puisqu'il faut bien constituer le premier tableau. Ceci se réalise alors souvent par itérations successives.

On peut noter par ailleurs que l'utilisation de différentes formules d'interpolation polynomiale, dans la mise en oeuvre d'un algorithme d'intégration, débouche sur diverses méthodes d'intégration. Cowell utilise l'interpolation de Stirling, les méthodes d'Adams utilisent les méthodes de Lagrange, Newton, etc.

Table des matières

1	Introduction	4
2	Vocabulaire Utile	6
3	Méthodes d'intégration numérique en calcul d'orbite	7
3.1	Les limites théoriques	7
3.2	Présentation des méthodes	7
3.3	Intégrateurs linéaires du premier degré	9
3.3.1	Aspects théoriques	9
3.3.2	Aspects pratiques	11
3.3.3	Formulation générale	13
3.3.4	Principes de stabilisation	14
3.4	Interpolation de Lagrange et différences divisées	17
3.4.1	Polynôme interpolateur de Lagrange	17
3.4.2	Méthode des différences divisées	17
3.5	Formulation des intégrateurs dits "à pas liés"	18
3.5.1	Méthodes d'Adams	19
3.5.2	Méthode de Bulirsch & Stoer	23
3.6	Intégrateurs linéaires du second degré	25
3.6.1	Méthode de (Adams-) Störmer	25
3.6.2	Méthode d'Adams-Moulton-Cowell (AMC)	26
3.6.3	Méthode de Cowell	28
4	Méthode d'Encke	32
4.1	Introduction	32
4.2	Mouvement de référence	32
4.3	Equations différentielles	33

4.3.1	Cas des éléments elliptiques usuels	33
4.3.2	Cas des coordonnées rectangulaires	34
5	La circularisation	37
5.1	Présentation générale de la méthode appliquée à l'algorithme de Cowell . . .	37
5.1.1	Position du problème	38
5.1.2	Méthode récurrente de calcul des coefficients modifiés	39
5.2	Initialisation de la récurrence : méthodes d'ordre 2 et 4	44
6	Discussion	46
	Références	47

1 Introduction

Les méthodes numériques d'intégration des systèmes différentiels ont connu des progrès très importants depuis le début des années 60. L'utilisation de calculateurs de plus en plus performants a permis, depuis 30 ans environ, d'employer ces méthodes en Mécanique Céleste et en Géodésie Spatiale, voire de les généraliser ([9]).

Le mouvement du satellite artificiel est décrit par les lois de Newton, qui font appel à l'ensemble des forces (le modèle) subies par le satellite :

$$\ddot{\mathbf{r}} = \frac{\mu\mathbf{r}}{r^3} + \mathbf{f}_G + \mathbf{f}_{NG} \quad (1)$$

où \mathbf{r} est le vecteur géocentrique Terre-satellite, $\ddot{\mathbf{r}}$ est l'accélération et μ la constante de la gravitation.

Quelle que soit la forme sous laquelle on présente l'équation fondamentale de la dynamique (1), il s'agit d'équations différentielles que l'on veut intégrer numériquement. Suivant les cas, c'est-à-dire suivant la paramétrisation, en coordonnées rectangulaires ou en éléments képlériens, on aura à faire avec des équations du second degré - au nombre de trois - ou du premier degré - six -, respectivement.

Les différents types de mesures spatiales (techniques spatiales), par leur interprétation en restitution de trajectoires, ont donné des informations très précises sur la connaissance des forces \mathbf{f} agissant sur les satellites artificiels de la Terre. Aujourd'hui, de la qualité de l'intégration numérique dépend la faculté que l'on a de savoir si les modèles de forces employés, voire aussi certaines données d'accélérométrie très fines, sont proches de la réalité ou non.

Les forces sont, de façon conventionnelle, séparées en deux catégories : les forces d'origine gravitationnelle \mathbf{f}_G , qui ne dépendent que des positions et sont par nature conservatives, et les forces d'origine non-gravitationnelle \mathbf{f}_{NG} (par unité de masse) qui peuvent dépendre des vitesses et sont, par nature également, non-conservatives voir dissipatives. Schématiquement, nous avons :

1. gravitationnel :

- le champ de gravité statique (décrit en harmoniques sphériques),
- le champ de gravité variable (marées, ...),
- l'attraction gravitationnelle du Soleil, de la Lune et des planètes (effets directs et indirects),
- les effets de la relativité

2. non-gravitationnel :

- le freinage atmosphérique dû aux effets de la densité,
- la pression de radiation solaire directe (incluant les effets d'ombre),
- la pression de radiation rediffusée par la Terre (effets de réémission et d'émissivité),
- les poussées thermiques dues au satellite,
- autres (effet de charge électrique, ...).

Ces équations du mouvement sont typiquement exprimées dans le système $J2000$, dont la réalisation effective dans les calculs d'orbite peut être soit un repère inertiel, souvent défini

par l'équateur et l'équinoxe moyen en 2000, soit un repère quasi-inertiel, généralement le repère céleste vrai de la date. Dans les deux cas, ceci implique de définir et d'utiliser des transformations de coordonnées appropriées pour pouvoir exprimer toutes les forces dans le même système. Par exemple, le champ de gravité statique est communément décrit dans un système terrestre moyen tournant, ce qui implique une transformation de la force vers le système d'intégration en utilisant la précession, la nutation, le mouvement du pôle et la rotation de la Terre (UT1). La variable indépendante, dans la terminologie actuelle, est le Temps Terrestre (TT) qui est équivalent au précédent Temps Dynamique Terrestre (TDT).

Un résumé récent des modèles de forces, d'orientation de la Terre et d'échelle de temps est donné dans les Conventions de l'International Earth Rotation Service (IERS) [12], notamment pour les satellites géodésiques.

Les méthodes d'intégration, dites dynamiques, apportent la meilleure prédiction possible du mouvement d'un objet spatial à partir d'une série de mesures de poursuite et d'un modèle d'évolution/restitution d'orbite. Pour cela, des solutions approchées d'équations approchées sont recherchées, sur la base de systèmes dynamiques dont l'évolution temporelle est décrite par un nombre fini d'équations déterministes : ici le système des équations du mouvement d'un satellite artificiel de la Terre (voir [16], par exemple).

Dans ce cours, nous verrons tout d'abord la formulation générale des intégrateurs linéaires, les grandes caractéristiques des algorithmes, le vocabulaire (notes de F. Faubert, GRGS, Grasse, 1986). Ensuite, nous décrirons plus finement le principe puis le calcul des coefficients des algorithmes, dans plusieurs cas. Enfin, nous décrirons plus spécifiquement deux types de méthodes, la circularisation et la méthode d'Encke, qui permettent de très bien stabiliser les calculs d'intégration numérique lorsque l'on cherche une très faible erreur relative (projets GRACE et GOCE, notamment).

2 Vocabulaire Utile

Le vocabulaire utilisé couramment en géodésie spatiale est très spécifique, pas toujours clair, redondant, parfois un peu obscur pour le lecteur non averti, comme dans de nombreuses spécialités. Le lecteur trouvera les expressions suivantes :

géodynamique	Etude des sources de variations du champ de gravité de la Terre (les déplacements de masses)
dynamique	Représentation des mouvements célestes ou spatiaux (dynamique orbitale ou spatiale) par intégration des équations de la dynamique Les systèmes dynamiques sont des systèmes d'équations différentielles non-linéaires
solution d'orbite	Ephéméride d'un satellite artificiel généralement obtenue par intégration numérique d'un modèle de forces
long terme	Les arcs longs représentent plusieurs milliers de révolutions orbitales (1 à 20 ans)
court terme	Les arcs courts représentent quelques centaines de révolutions (1 à 10 jours)
méthode classique	Méthode d'intégration numérique des équations du mouvement et des équations de sensibilité des variables à des paramètres dynamiques choisis
sol. géométrique	Calcul des positions/vitesses d'un satellite artificiel de la Terre, sans modèle dynamique, grâce à la redondance des données de poursuite
sol. analytique d'ordre i	Expressions formelles et littérales des variables du mouvement en fonction des coefficients des forces perturbatives
sol. semi-analytique	Calcul d'une éphéméride par une transformation analytique des équations du mouvement (filtrage ou moyenne) et intégration numérique des nouvelles équations
sensibilité	Dérivée partielle des variables décrivant le mouvement par rapport à des paramètres dynamiques choisis
précision des solutions	Ecart-type des différences entre quantités observées et mesurées (en référence aux estimateurs liés au moindres carrés) Différences entre solutions de différente nature
erreur d'orbite	Partie dépendant du temps : normalement décomposable en séries de Fourier de la période orbitale et de ses sous-multiples. Le signal peut néanmoins présenter des sauts de phase importants au fur et à mesure du temps Partie constante : erreur d'orbite toujours réitérée en un point géographique fixe sur la Terre (erreur géographiquement corrélée)
propagation d'erreur	Calcul de l'influence de l'erreur d'orbite (part constante essentiellement) dans le calcul d'un sous-produit : coordonnées de stations, profil altimétrique, ...
système spatial	Ensemble formé par un satellite (ou une constellation), son type d'orbite (altitude, ...), la technique de géodésie spatiale utilisée pour l'acquisition de mesures de poursuite, et le réseau des stations terrestres correspondantes
étalonnage (ou calibration)	Opération de contrôle d'un instrument visant à établir son systématisme de mesure par rapport à un autre instrument, plus exact et plus précis

3 Méthodes d'intégration numérique en calcul d'orbite

3.1 Les limites théoriques

La recherche de la meilleure représentation possible, dans notre approche de la détermination du mouvement orbital, est très exigeante. Ceci est dû à la proximité des trajectoires par rapport à la Terre ([8]), rendant plus complexe le volume des interactions dynamiques satellite-milieu spatial dont il faut rendre compte pour intégrer les équations du mouvement (1). L'exigence tient aussi et surtout à la très grande précision recherchée sur les positions et les vitesses fournies par les solutions d'orbite (les éphémérides) à court terme comme à long terme.

Les limites de prédiction, en précision et stabilité à long terme, ne sont pas inhérentes aux propriétés du système dynamique, le mouvement du satellite étant stable et régulier au sens de la Mécanique Céleste sur les durées considérées. Elles viennent en grande partie de nos capacités actuelles en ce qui concerne [15] :

- les conditions initiales du mouvement (position et vitesse du satellite à un instant donné). Elles ne sont connues qu'approximativement, étant issues soit des observations, soit de calculs approchés ;
- le modèle d'évolution (l'expression des forces). Son estimation est délicate et celle de l'exactitude de toutes les constantes numériques qui le décrivent reste difficile. Ajoutons que l'aspect chaotique de la rotation du satellite sur lui-même peut introduire par couplages "spin-orbite" des effets, certes faibles, mais dont il est très difficile de rendre compte dans la modélisation ;
- les équations différentielles du mouvement. Du fait que celles-ci ne sont pas intégrables, le développement de techniques d'intégration spécifiques comme l'intégration numérique se révèle inévitable, générant des erreurs sur le résultat final ;
- les techniques de mesure de poursuite. Il faut pouvoir transcrire, interpréter et exploiter correctement des observations de précisions parfois hétérogènes. L'ajustement des conditions initiales du mouvement ainsi que d'éventuels termes du modèle en dépend.

Compte tenu de ces considérations, il apparaît qu'un long effort est encore nécessaire pour optimiser la précision des solutions et assurer leur stabilité à très long terme ([9]).

3.2 Présentation des méthodes

Avant tout, il faut noter ici les quatre critères fondamentaux nécessaires à l'établissement de toute théorie d'intégration numérique (voir par exemple [11]) :

- un ensemble de valeurs permettant de calculer l'approximation suivante,
- un formalisme estimant les erreurs de troncature par pas,
- des critères permettant d'accepter un résultat en fonction de la précision recherchée (cas des méthodes avec prédiction-correction notamment),
- parfois aussi, un critère permettant de déterminer à quel instant on doit changer le pas (sur quelle partie de la trajectoire), et lequel choisir afin d'optimiser en même temps la précision et le temps de calcul (travaux de Kulikov, notamment).

On distingue les méthodes d'intégration numérique "à pas liés", "à pas séparés", à pas fixe ou variable, et aussi des méthodes qui ont un système de prédiction-correction.

De plus, ces méthodes sont souvent fondées sur l'approximation du second membre des équations par un polynôme ou, parfois, par une fraction rationnelle.

Nous allons étudier ici les trois méthodes les plus fréquemment employées dans les programmes de restitution d'orbite de satellite artificiel. Chacune d'entre elles représente, par sa spécificité une classe particulière de méthodes :

- les méthodes de type Adams-Moulton et Adams-Bashforth, utilisées dans le programme CODIOR (restitution d'orbite moyenne de satellite) du CERGA, caractérisées par :
 - un pas lié
 - une interpolation polynomiale
 - des équations différentielles du premier degré
- les méthodes de type Cowell, utilisées dans les programmes GINS et ZOOM du CNES, caractérisées par :
 - un pas lié
 - une interpolation par développement de Taylor des fonctions du second membre
 - des équations différentielles du second degré
- la méthode de Bulirsh & Stoer, utilisée dans le programme d'extrapolation d'orbite LAGRAN, caractérisée par :
 - un pas variable (éventuellement), le concept de pas ayant toutefois un sens spécifique à cette méthode
 - un système de prédiction-correction
 - une interpolation par fraction rationnelle
 - des équations différentielles du premier degré

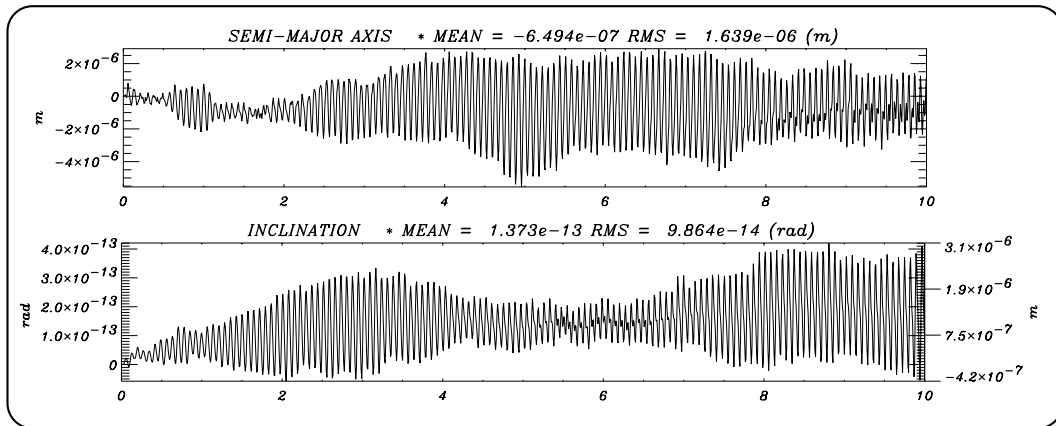


FIG. 1 – Evaluation de l'intégration numérique sur STARLETTE
 Différence sur 10 jours entre une orbite intégrée en éléments orbitaux et une orbite intégrée en coordonnées rectangulaires. Le potentiel gravitationnel perturbateur est développé jusqu'aux degré et ordre 20. Le pas d'intégration est de 30 secondes.

3.3 Intégrateurs linéaires du premier degré

3.3.1 Aspects théoriques

Le but est de calculer numériquement la solution d'une équation différentielle du type de celle du problème de Cauchy :

$$\begin{cases} \frac{dy}{dt} = f(t, y(t)), & t \geq 0 \\ y(t_0) = y_0 \end{cases} \quad (2)$$

où f est une fonction suffisamment régulière des deux variables t et y . Ceci est équivalent à la recherche d'une fonction $\varphi \in \mathcal{C}^1[0, T]^1$ telle que :

$$\forall t \in [0, T], \varphi(t) = y_0 + \int_0^t f(s, \varphi(s)) ds$$

où $\forall h > 0$, h étant un pas d'évolution de la variable t :

$$\begin{cases} \varphi(h) = y_0 + \int_0^h f(s, \varphi(s)) ds \\ = y_0 + hf(0, y_0) + h\varepsilon(h) \end{cases}$$

D'où l'on peut déduire une récurrence explicite définissant y_{n+1} , selon la méthode d'Euler :

$$\begin{cases} y_{n+1} = y_n + hf(t_n, y_n) & t_n = nh \text{ et } n \geq 0 \\ y_0 = y(t_0) \end{cases} \quad (3)$$

Si nous pouvons estimer une erreur "locale" en $\Theta(h^2)$, pour passer de y_n à y_{n+1} , il existe, pour définir l'erreur "globale", un processus de cumul des erreurs locales. Car à l'instant t_n , au lieu de partir de y_n , on part en fait d'une valeur calculée \tilde{y}_n .

Convergence :

Une méthode converge si :

$$\max_{0 \leq n \leq N} |\tilde{y}_n - \varphi(t_n)| \rightarrow 0$$

quand $y(t_0) \rightarrow \varphi(t_0)$ et $h \rightarrow 0$.

Ceci revient à dire que, pour $t = nh$, $\lim_{n \rightarrow \infty} \tilde{y}_n(t) = \varphi(t)$. Dans ce cas, on a alors $h \rightarrow 0$.

La quantité $\max_{0 \leq n \leq N} |\tilde{y}_n - \varphi(t_n)|$ est l'erreur globale de la méthode.

¹ φ est la solution exacte de (2).

Pour obtenir une approximation de $\varphi(t)$ par la méthode retenue, il faudrait ainsi faire d'autant plus de calculs que h est petit ; le cumul des erreurs locales se posera alors de façon accrue. Il y a convergence si, à la limite, ce cumul s'efface devant l'augmentation de précision globale de la méthode.

Consistance :

On définit l'erreur de consistance comme étant la quantité :

$$e_n = \varphi(t_{n+1}) - y_{n+1}$$

la valeur y_{n+1} étant issue de (3), avec $y_n = \varphi(t_n)$.

Une méthode est dite consistante lorsque $\sum_{0 \leq n \leq N} |e_n|$ tend vers 0, lorsque h tend vers 0.

De plus, une méthode est consistante d'ordre α s'il existe une constante $C \geq 0$ telle que :

$$\forall n \in \mathbb{N}, 0 \leq n < N, |e_n| \leq Ch^{\alpha+1}$$

Cette notion affirme la "précision locale" d'une méthode (d'un schéma) : $\Theta(h^{\alpha+1})$.

On peut déduire de ceci la règle suivante : une consistance d'ordre $\Theta(h^{p+1})$ implique une convergence à $\Theta(h^p)$. Le schéma d'Euler converge : s'il est consistant d'ordre $\Theta(h^2)$, la convergence a lieu en $\Theta(h)$. Cependant, le cumul des erreurs conserve un caractère linéaire : les erreurs locales s'ajoutent.

Stabilité :

On dit que la méthode est stable s'il existe une constante $S \geq 0$ telle que, pour toutes suites (y_n) et (z_n) définies respectivement par :

$$\begin{aligned} y_{n+1} &= y_n + hf(t_n, y_n) \\ z_{n+1} &= z_n + hf(t_n, z_n) + \varepsilon_n \end{aligned}$$

on ait (où ε_n désigne une erreur d'arrondi) :

$$\max_{0 \leq n \leq N} |y_n - z_n| \leq S(|y_0 - z_0| + \sum_{0 \leq n \leq N} |\varepsilon_n|)$$

Ainsi, de petites erreurs dans le calcul génèrent finalement une erreur contrôlable.

On conçoit que la convergence du schéma exige la stabilité, mais celle-ci n'est pas forcément satisfaite sur tous les schémas linéaires (L).

La convergence exige la stabilité de (L), qui a lieu sous certaines conditions. En règle générale : "consistance + stabilité \Rightarrow convergence".

Exemple de schémas en $\Theta(h^3)$

Euler amélioré ("à pas séparés")

$$\begin{cases} y_{n+1}^* &= \tilde{y}_n + hf(\tilde{y}_n, nh) \\ \tilde{y}_{n+1} &= \tilde{y}_n + \frac{h}{2}[f(\tilde{y}_n, nh) + f(y_{n+1}^*, (n+1)h)] \end{cases} \quad (4)$$

Adams-Bashforth (AB) ("à pas liés")

$$\tilde{y}_{n+1} = \tilde{y}_n + \frac{h}{2}[3f(\tilde{y}_n, nh) - f(\tilde{y}_{n-1}, (n-1)h)] \quad (5)$$

Le premier requiert, à chaque calcul, deux appels au second membre. Le second, AB, n'en requiert qu'un, puisque f_{n-1} est déjà en mémoire. AB utilise, pour \tilde{y}_{n+1} , une extrapolation linéaire à partir de \dot{y}_{n-1} et \dot{y}_n ; la consistance est bien d'ordre 3, mais il faut faire attention aussi à la stabilité si l'intervalle $[0, T]$ est trop grand.

3.3.2 Aspects pratiques

On peut exprimer les choses de façon totalement numérique. Ainsi, le problème de Cauchy se traduit par :

$$y(t_0 + T) = y(t_0) + \sum_{i=0}^n \int_{t_0+ih}^{t_0+(i+1)h} g(t, y(t)) dt \quad (6)$$

avec : $g(t, y(t)) \simeq f(t, y(t))$.

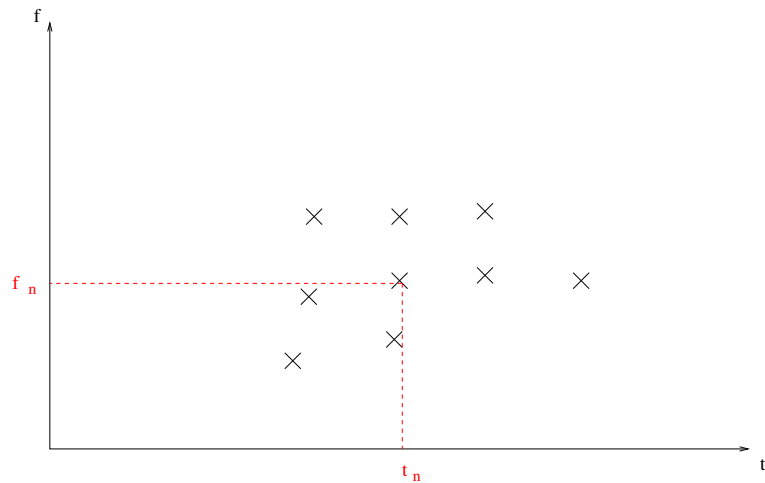
Les méthodes d'intégration numérique se distinguent par le type de construction de la fonction d'approximation g . Le pas h d'intégration peut être d'autant plus grand que g approche mieux la fonction f , mais alors le calcul de g est d'autant plus compliqué.

Rappelons que les méthodes sont à un pas (dites "à pas séparés") ou bien multi-pas (dites "à pas liés"). Voici un exemple d'une méthode "à pas séparés" de degré un et d'ordre N , où N correspond au nombre de fonctions f_p utilisées dans le schéma (Runge-Kutta, RK- N) :

$$\begin{aligned} y_{n+1} - y_n &= h \sum_{p=1}^N b_p f_p \\ f_p &= f\left[y_n + \sum_{j=1}^N a_{p,j} f_j, t_n + c_p h\right] \end{aligned} \quad (7)$$

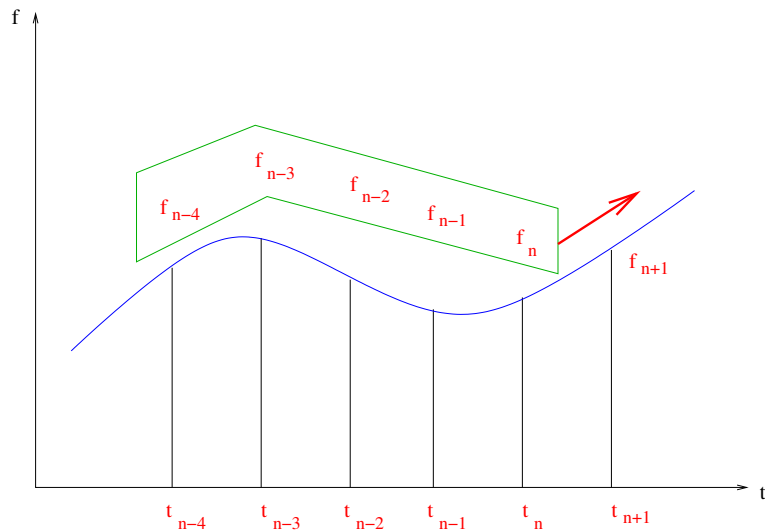
Avec les versions suivantes : RK-4 (Kutta), RK-6 (Butcher, 1964), RK-8 (Schanks, 1966).

⇒ A pas séparés (par exemple : RK) : les f_p sont des évaluations de la fonction ($t \mapsto f(t, y(t))$), faites aux pas précédents :



Il existe une grande richesse de possibilités de construction de réseau, pour une meilleure évaluation du second membre.

⇒ A pas liés (par exemple : Cowell) : à chaque pas, on cherche f_p en un réseau (t_p, y_p) : (i) voisin du pas courant (t_n, y_n) , et (ii) indépendant des autres pas.



Ici, le réseau de construction est imposé par les précédents pas d'intégration.

Différences entre les algorithmes "à pas séparés" et "à pas liés"

Type d'algorithme	"à pas séparés"	"à pas liés"
Volume de calcul à ordre égal	supérieur donc : temps de calcul et erreur d'arrondi en hausse	réduit au minimum
Nbe d'éval. du second membre par pas	N	1
Mémoires des pas précédents	sans	N pas disponibles
Procédure de démarrage	aucune	élaborée
Changement de pas d'intégration	à chaque instant	quasi-impossible
Construction du réseau d'évaluation	grande richesse	imposée par les pas précédents
	algorithme adaptable	on peut changer la classe des polynômes
Erreur de troncature	plus faible	

3.3.3 Formulation générale

Afin d'obtenir une consistance d'ordre élevé, on a intérêt à privilégier une évaluation de la fonction f (second membre), ainsi que de sa dérivée, par moyenne pondérée ; mais il faut alors définir les coefficients. D'où l'approximation inhérente à la classe des méthodes linéaires multi-pas (linear multistep (LM) methods) :

$$y_{n+1} + \sum_{p=p_0}^P a_p y_{n+1-p} = h \sum_{p=n_0}^N b_p f_p \quad (8)$$

Le premier membre est l'évaluation de la dérivée \dot{y} par moyenne, pondérée (a_p) de différences de y au voisinage de y_{n+1} , et le second l'évaluation de $f(t, y)$ par moyenne pondérée (b_p) d'évaluations f_p de la fonction autour de y_{n+1} . L'ordre de l'intégrateur est bien N , c'est-à-dire le nombre de fonctions f_p entrant dans la moyenne.

Un algorithme est donc déterminé par :

- le choix des lieux f_p d'évaluation de la fonction $f(t, y)$
- le choix des coefficients (a_p, b_p)

La stratégie de formation des f_p étant choisie, on peut, en choisissant les coefficients (b_p) convenablement, faire en sorte que l'équation de progression soit vérifiée pour quelques fonctions particulières $f_q(t, y)$. On privilégie alors un jeu de fonctions que l'intégrateur intégrera exactement, car l'équation de progression (par exemple : $y_{n+1} - y_n = h \sum_p b_p f_p$) est exactement vérifiée pour ces dernières.

→ Le jeu des fonctions $f_q(t, y)$ indépendantes, intégrées exactement, contient au plus N fonctions (ordre) ; on n'a donc de liberté que pour le choix de N constantes (b_p).

Exemple : dans l'algorithme de Cowell original, le jeu des N fonctions choisies correspond à N polynômes de degrés allant de 1 à N . La méthode de Cowell intègre ainsi

exactement toute fonction polynomiale de degré \leq à l'ordre d'intégration.

→ Il est possible de privilégier des fonctions autres que des polynômes, comme par exemple des fonctions circulaires, sinus ou cosinus, à la période du mouvement considéré.

Erreur de troncature

→ L'intégrateur passe exactement à travers une classe de fonctions privilégiées f_q . Il passe donc du pas n au pas $n + 1$ en "supposant" que la solution $y(t)$ est combinaison linéaire des fonctions qu'il intègre exactement.

En pratique, un tel algorithme de décomposition ne sera acceptable que lorsque le développement de Taylor de la fonction pourra être tronqué à l'ordre N avec une erreur négligeable, souvent évaluée à l'aide du terme de rang $(n + 1)$, lui-même négligé.

Si la solution $y(t)$ ne se décompose pas exactement sur la base des N fonctions f_q privilégiées, il y a troncature dans la décomposition.

L'évaluation de l'erreur de troncature se fait en pratique avec deux intégrations à h et $h/2$; la différence entre les résultats donne alors une idée de sa valeur.

→ *Facteurs dégradant l'erreur de troncature :*

- ordre d'intégration faible (d'où une mauvaise représentation de la solution)
- pas d'intégration h trop grand
- durée de l'intégration importante (l'erreur de troncature est bien évidemment fonction croissante du pas d'intégration).

Erreur d'arrondi

Plusieurs auteurs ont étudié, souvent de façon empirique, l'erreur d'arrondi et sa propagation dans les mouvements orbitaux intégrés numériquement sur le très long terme.

→ L'erreur d'arrondi se cumule avec le nombre de pas au carré : N^2 . Elle est proportionnelle à $\varepsilon = 2^{-b}$, où b est le nombre de bits de la mantisse réservés pour la représentation des nombres réels en machine ;

→ si la valeur moyenne de l'erreur d'arrondi était zéro, l'effet en longitude serait proportionnelle à $N^{3/2}$;

→ l'utilisation de la méthode d'Encke (voir plus loin) permet de réduire cet effet d'un facteur 0.006 (Nobili et al.).

3.3.4 Principes de stabilisation

Divers auteurs ont recherché des algorithmes qui intègrent exactement des fonctions non polynomiales, notamment en calcul d'orbites (fonctions trigonométriques, harmoniques, etc). Ceci peut se faire de deux façons :

1. en modifiant les coefficients (b_p) d'un algorithme tel que celui décrit par la formulation générale précédente (8),
2. en recherchant un changement de variable indépendante (ou descriptive), qui "régularise" un certain algorithme pour une famille donnée d'orbites.

Ces procédés s'appliquent dans le cadre de la stabilisation des méthodes aux différences finies (de nombreux calculs, présentés ici, sont issus de notes de G. Balmino, CNES/GRGS, Toulouse).

Régularisation

Le temps n'est pas une variable d'intégration bien adaptée pour les orbites très excentriques ($e > 0.1$ en géodésie spatiale, par exemple). La répartition des points sur l'orbite est insuffisante au périégée où les perturbations sont multiples et de fortes amplitudes.

→ L'intégration d'une équation différentielle de plus donne le temps t (ou anomalie moyenne) en fonction de la variable d'intégration. Dans le cas général (\vec{r} a pour coordonnées x, y, z), on a (idem pour y et z) :

$$\begin{cases} \frac{dt}{ds} &= f(r, ..) \\ \frac{dx}{ds} &= f(r, ..) \frac{dx}{dt} \\ \frac{d^2x}{ds^2} &= g(r, ..) [\vec{r} \cdot \vec{V}] \frac{dx}{dt} + [f(r, ..)]^2 \frac{d^2x}{dt^2} \end{cases}$$

La régularisation permet de ramener le calcul d'une orbite quelconque (en particulier avec une excentricité e importante), au calcul d'un mouvement circulaire plan, dans la mesure où la densité des points sur l'orbite est la même au périégée et à l'apogée.

Ceci est un moyen détourné, certes très efficace, qui ne remet pas en question la technique d'intégration donc l'erreur de troncature.

– La variable E (anomalie excentrique) est un bon compromis :

$$\begin{aligned} M &= E - e \sin E \\ \frac{dM}{dE} &= (1 - e \cos E) \rightarrow \frac{dM}{dE} = \frac{r}{a} \end{aligned}$$

Ici : $dt/ds = r/a$. On compte le temps à partir du début de l'intégration, d'autre part, on préfère intégrer numériquement des fonctions bornées, on pose donc : $t = \theta_0 + s + \sigma(s)$, d'où : $d\sigma/ds = r/a - 1$. La partie périodique du temps, prise nulle au périégée, est σ . Ce qui implique :

$$n \theta_0 = E_{\text{origine}} - M_{\text{origine}}$$

et, grâce à l'équation de Kepler :

$$\theta_0 = \frac{1}{n} e \sin E_{\text{origine}} = \frac{a}{\mu} \vec{r} \cdot \vec{V} (\text{origine})$$

On passe de la variable s à la variable temps, avant chaque appel à la fonction second membre, et on revient de la variable temps à la variable s ensuite. Les formules de

changement de variables sont les suivantes (idem pour y et z) :

$$\left\{ \begin{array}{l} \frac{dt}{ds} = \frac{r}{a} \\ \frac{dx}{ds} = \frac{r}{a} \frac{dx}{dt} \\ \frac{d^2x}{ds^2} = \frac{1}{a^2} [\bar{r} \cdot \bar{V}] \frac{dx}{dt} + \left[\frac{r}{a} \right]^2 \frac{d^2x}{dt^2} \end{array} \right.$$

– La variable v (anomalie vraie) est bien adaptée mais les calculs sont plus complexes :

$$\begin{aligned} r^2 \dot{v} &= n a^2 \sqrt{1 - e^2} = C_0 \\ &\rightarrow \frac{dM}{dv} = \frac{r^2}{a} \frac{1}{\sqrt{1 - e^2}} \end{aligned}$$

Le calcul de θ_0 ne peut se simplifier comme dans le cas de l'anomalie excentrique, mais il n'est fait qu'une fois pour toute l'intégration. Les formules de changement de variables sont les suivantes (idem pour y et z) :

$$\left\{ \begin{array}{l} \frac{dt}{ds} = \frac{r^2}{C_0} \\ \frac{dx}{ds} = \frac{r^2}{C_0} \frac{dx}{dt} \\ \frac{d^2x}{ds^2} = 2 \frac{r^2}{C_0^2} [\bar{r} \cdot \bar{V}] \frac{dx}{dt} + \left[\frac{r^2}{C_0} \right]^2 \frac{d^2x}{dt^2} \end{array} \right.$$

Circularisation

La circularisation permet d'intégrer exactement les mouvements plans circulaires uniformes, dans la mesure où les coefficients (b_p) de l'intégrateur correspondent, cette fois, au choix d'une fonction harmonique $\exp(j(i - 1)\omega t)$, de fréquence $(2\pi/T)$ donnée. Elle peut être utilisée indépendamment de la régularisation, mais il n'est cependant pas souhaitable de l'employer sans elle pour le calcul d'orbites excentriques.

→ Cette modification de l'algorithme peut se traduire par l'adjonction de termes correctifs dans les équations de l'intégrateur. Les formules corrigées tiennent compte, dans les coefficients, du rapport (pas/période).

Mais il faut alors connaître la période avec beaucoup de précision. Une façon efficace de réduire l'erreur de troncature (sur l'intégration de l'équation du temps, par exemple) est de tenir compte du moyen mouvement képlérien et des effets séculaires dus à J_2 , J_4 , J_6 .

Ainsi, le choix de l'option "circularisation" permet-il, dans une certaine mesure, de réduire ce type d'erreur de troncature. La fonction second membre devrait être mieux représentée que dans le cas des seuls polynômes.

La section 5 est consacrée à ce formalisme ; elle veut donner une base afin de montrer comment calculer, dans le cadre de la méthode de Cowell, les coefficients modifiés de l'algorithme "classique" par l'adjonction de fonctions harmoniques à plusieurs fréquences.

3.4 Interpolation de Lagrange et différences divisées

Sont rappelées brièvement dans cette partie la méthode d'interpolation de Lagrange et les principales propriétés des différences divisées.

On considère une fonction $f : [a, b] \rightarrow \mathbb{R}$ continue et une subdivision de $n + 1$ points, deux à deux distincts, t_0, t_1, \dots, t_n du segment $[a, b]$. On se place dans \mathcal{P}_n , \mathbb{R} espace vectoriel de dimension $n + 1$ des fonctions polynômes de degré inférieur ou égal à n .

3.4.1 Polynôme interpolateur de Lagrange

Le polynôme interpolateur de f sur $[a, b]$ selon la subdivision t_0, t_1, \dots, t_n est l'unique polynôme p_n de \mathcal{P}_n tel que :

$$\forall i \in \{0, 1, \dots, n\}, p_n(t_i) = f(t_i) \quad (9)$$

Définissant les fonctions élémentaires $l_i, i \in \{0, 1, \dots, n\}$, par :

$$\forall x \in [a, b], l_i(t) = \prod_{j=0, j \neq i}^n \frac{t - t_j}{t_i - t_j} \quad (10)$$

on a :

$$\forall t \in [a, b], p_n(t) = \sum_{i=0}^n f(t_i) l_i(t) \quad (11)$$

Ce polynôme peut également être calculé à partir des différences divisées de la fonction f .

3.4.2 Méthode des différences divisées

La méthode des différences divisées est une méthode simple de calcul du polynôme interpolateur de Lagrange.

Définition et propriétés. La différence divisée d'ordre k de la fonction f , notée $f[t_0, t_1, \dots, t_k]$, est le coefficient de plus haut degré (ou encore coefficient directeur) du polynôme interpolateur p_k de f . Reprenant le précédent polynôme interpolateur p_n , on peut montrer :

$$\forall t \in [a, b], p_n(t) = f(t_0) + \sum_{k=1}^n f[t_0, t_1, \dots, t_k](t - t_0)(t - t_1)\dots(t - t_{k-1}) \quad (12)$$

Les différences divisées de f vérifient la formule de récurrence [7] :

$$\forall k \in \mathbb{N}^*, f[t_0, t_1, \dots, t_k] = \frac{f[t_1, \dots, t_k] - f[t_0, \dots, t_{k-1}]}{t_k - t_0} \quad (13)$$

A l'aide de cette formule de récurrence et de l'algorithme de Hörner, on peut facilement calculer le polynôme interpolateur de Lagrange p_n [7].

Reprenant les deux expressions (11) et (12), on constate aisément, par identification du terme de plus haut degré des polynômes p_k , que :

$$\forall k \in \{0, 1, \dots, n\}, f[t_0, t_1, \dots, t_k] = \sum_{i=0}^k f(t_i) \cdot \Pi_i \quad (14)$$

avec :

$$\Pi_i = \prod_{j=0, j \neq i}^k \frac{1}{t_i - t_j}$$

Enfin, terminons avec deux propriétés utiles.

→ La $(k - 1)$ ème différence divisée d'un polynôme de degré $(k - 2)$ est nulle. *En effet, le polynôme de degré $(k - 2)$ est égal à son polynôme interpolateur de Lagrange. La $(k - 1)$ ème différence divisée de ce polynôme serait le coefficient directeur de son polynôme interpolateur si celui-ci était de degré $(k - 1)$; elle est donc nulle.* On peut ainsi écrire (les coefficients c_j étant arbitraires) :

$$\sum_{i=1}^k \frac{c_1 t_i + c_2 t_i^2 + \dots + c_{k-1} t_i^{k-1}}{t_i} \cdot \Pi_i = 0 \quad (15)$$

→ La $(k - 1)$ ème différence divisée d'un polynôme de degré $(k - 1)$ est égale au coefficient du terme de plus haut degré de ce polynôme. *En effet, le polynôme de degré $(k - 1)$ est égal à son polynôme interpolateur de Lagrange. La $(k - 1)$ ème différence divisée de ce polynôme est le coefficient directeur de son polynôme interpolateur donc son propre coefficient directeur.* On peut ainsi écrire (le coefficient c_k étant arbitraire) :

$$\sum_{i=1}^k \frac{c_k t_i^k}{t_i} \cdot \Pi_i = c_k \quad (16)$$

3.5 Formulation des intégrateurs dits "à pas liés"

Ces méthodes sont basées sur l'idée que plus d'information peut être utilisée à un moment donné dans le processus d'intégration, en prenant en compte non seulement y_n mais aussi y_{n-1} , y_{n-2} , etc.

Ces méthodes demandent en général moins de calculs que les méthodes à un pas, pour la même précision. En revanche, elles sont complexes à mettre en oeuvre (processus de démarrage) et, dans certains cas, dangereuses du point de vue des éventuelles instabilités numériques (voir par exemple [1]).

La plupart des méthodes "à pas liés" conventionnelles sont basées sur l'interpolation polynomiale. On privilégie ainsi, dans l'équation (6), l'intégration de classes de fonctions puissances de la seule variable t :

$$g_j(t) = t^{j-1} \quad (j = 1, N + 1) \quad (17)$$

qui sont des polynômes de degré $\leq N + 1$.

3.5.1 Méthodes d'Adams

Ces méthodes présentent les qualités de simplicité et d'efficacité qui en font celles les plus souvent choisies dans les programmes d'intégration numérique. Elles se rapprochent beaucoup de la méthode "élémentaire" de Runge-Kutta et semblent être les plus indiquées pour une première approche de l'intégration numérique. On progresse pas par pas avec une forme générale du type :

$$y_{n+1} - y_n = h \sum_{p=1}^N b_p f(y_{n+1-p}, t_{n+1-p}) \quad (18)$$

Les méthodes d'Adams nécessitent la connaissance d'un nombre fini de valeurs particulières de la fonction aux points t_n choisis. Elles résolvent des équations linéaires du premier degré.

Ces méthodes approchent la fonction f par son polynôme de Lagrange construit à partir des formules de Newton sur les différences successives d'une fonction aux points t_n . L'écriture (voir section précédente) montre bien la nécessité de connaître plusieurs valeurs de la fonction aux instants (t_0, t_1, \dots, t_n) . Les méthodes d'Adams s'inscrivent donc bien dans le cadre des méthodes dites "à pas liés".

Si l'on choisit :

$$\begin{cases} f(t, y(t)) = \sum_{j=1}^{N+1} c_j g_j(t) + k(t) \\ \text{avec} \quad : \quad g_j(t) = t^{j-1} \end{cases}$$

où $k(t)$ est le résidu de la décomposition de f sur la base des fonctions g (c'est-à-dire l'erreur de troncature), on obtient :

$$f_{n+1-p} = \sum_{j=1}^{N+1} c_j g_j(t_{n+1-p}) \quad p \in \{1, N\}$$

d'où :

$$\begin{bmatrix} f_n \\ \vdots \\ f_{n-N} \end{bmatrix} = \begin{bmatrix} g_{1,1} & \cdots & & \\ \vdots & g_{p,j} & & \vdots \\ & \cdots & g_{N,N+1} & \end{bmatrix} \times \begin{bmatrix} c_1 \\ \vdots \\ c_{N+1} \end{bmatrix} \quad \text{avec : } g_{p,j} = t_{n+1-p}^{j-1}$$

puis :

$$c_j = \sum_{p=1}^{N+1} G_{p,j}^{-1} f_{n+1-p} \quad (j = 1, \dots, N+1)$$

La fonction f prend les valeurs f_{n+1-p} aux points correspondants, pour $p \in \{1, N+1\}$; le polynôme passant par les points (f_{n+1-p}, t_{n+1-p}) de degré N est unique. Les valeurs t_{n+1-p} doivent toutes être distinctes, sinon le déterminant de la matrice $G_{p,j}$ (qui doit être régulière), du type Vandermonde, est nul.

Les coefficients (c_j) sont les coefficients du polynôme d'interpolation de Lagrange ; on peut alors utiliser la formule de Newton, faisant intervenir les différences descendantes de la fonction f :

$$\begin{cases} f(t, y(t)) \simeq P(t) = \sum_{j=0}^N (-1)^j C_j^{-s} \Delta_-^j f_n \\ t \in [t_0, t_n] \quad ; \quad s = (t - t_n)/h \end{cases}$$

avec :

$$\Delta_-^j f_n = \sum_{k=0}^j (-1)^k C_k^j f_{n-k}$$

Exemple pour $N = 2$

$$y_{n+1} - y_n = \int_{t_n}^{t_{n+1}} (c_1 + c_2 t) dt = c_1 h + \frac{c_2}{2} [t^2]_{t_n}^{t_{n+1}}$$

avec :

$$f_{n+1-p} = c_1 + c_2 t_{n+1-p} + c_3 t_{n+1-p}^2 + \dots$$

ce qui fournit le système suivant :

$$\begin{bmatrix} f_n \\ f_{n-1} \\ f_{n-2} \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & t_n & t_n^2 & \cdots \\ 1 & t_{n-1} & t_{n-1}^2 & \\ 1 & t_{n-2} & t_{n-2}^2 & \\ \vdots & & & \vdots \end{bmatrix} \times \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix}$$

d'où :

$$c_1 = \frac{f_{n-1} t_n - f_n t_{n-1}}{t_n - t_{n-1}} \quad \text{et : } c_2 = \frac{f_n - f_{n-1}}{t_n - t_{n-1}}$$

et :

$$y_{n+1} - y_n = h \left[\frac{3}{2} f_n - \frac{1}{2} f_{n-1} \right] \quad (19)$$

On retrouve donc bien la formule (5).

Les méthodes d'Adams sont construites sur ce schéma. Cependant, deux cas sont à envisager : le schéma explicite, dit "prédicteur", et le schéma implicite, dit "correcteur".

Adams-Bashforth : schéma explicite, prédicteur On a :

$$y_{n+1} - y_n = h \sum_{j=0}^N \xi_j \Delta_-^j f_n \quad (20)$$

avec :

$$\xi_j = (-1)^j \int_0^1 \mathcal{C}_j^{-s} ds \quad (21)$$

d'où le schéma explicite (AB) :

$$y_{n+1} - y_n = h \sum_{j=0}^N \sigma_{N,j} f_{n-j} \quad (22)$$

Les coefficients $(\xi_j, \sigma_{N,j})$ sont indépendants de f et sont calculés une fois pour toutes. On utilise pour cela une fonction génératrice $G(t)$ dont le développement de Mac-Laurin contient les coefficients (ξ_j) :

$$\begin{aligned} G(t) &= \sum_{j=0}^{\infty} \xi_j t^j = \sum_{j=0}^{\infty} (-t)^j \int_0^1 \mathcal{C}_j^{-s} ds \\ &= \int_0^1 (1-t)^{-s} ds = \int_0^1 e^{-s \ln(1-t)} ds \\ &= \frac{-t}{\ln(1-t)(1-t)} \end{aligned} \quad (23)$$

En identifiant les coefficients de même puissance en t , dans les équations (23) ci-dessus, on aboutit à :

$$\left| \begin{array}{l} \xi_0 = 1 \\ \xi_1 = \frac{1}{2} \\ \xi_2 = \frac{5}{12} \end{array} \right| \quad \left| \begin{array}{l} \xi_3 = \frac{3}{8} \\ \xi_4 = \frac{251}{720} \\ \xi_5 = \frac{95}{288} \end{array} \right| \quad \left| \begin{array}{l} \xi_6 = \frac{19087}{60480} \\ \xi_7 = \frac{36799}{120960} \\ \xi_8 = \dots \end{array} \right|$$

La formule (22), qui donne explicitement $(y_{n+1} - y_n)$ en fonction des différences $\Delta_-^j f_n$ est efficace dès lors que l'on peut faire des chargements dans la procédure.

Notons que si la différence du plus grand ordre (en N) est importante, la précision peut s'avérer insuffisante et le nombre N de termes pris en compte doit être augmenté.

Enfin, comme le nombre N est souvent fixé une fois pour toutes (pour des raisons de programmation informatique), il n'y a pas de raison particulière pour que les différences soient explicitées. On peut ainsi redonner une nouvelle formulation à la méthode, en exprimant les différences en terme d'ordonnées. En rassemblant les coefficients de même ordonnée, la formule d'AB apparaît ainsi sous la forme :

$$\sigma_{N,j} = (-1)^j \left[\mathcal{C}_j^j \xi_j + \mathcal{C}_j^{j+1} \xi_{j+1} + \dots + \mathcal{C}_j^N \xi_N \right] \quad (24)$$

Notons que la grandeur de ces coefficients ainsi que l'alternance des signes dans cette série sont des désavantages de cette méthode.

Adams-Moulton : schéma implicite, correcteur

La méthode précédente utilise les différences de la fonction f construites sur les points d'abscisses t_{n+1} et t_n ; elle n'est donc pas optimale. La méthode AM utilise une interpolation :

$$y_n - y_{n-1} = h \sum_{j=0}^N \xi_j^* \Delta_j^- f_n \quad (25)$$

avec :

$$\xi_j^* = (-1)^j \int_{-1}^0 C_j^{-s} ds \quad (26)$$

d'où le schéma implicite (AM) :

$$y_n - y_{n-1} = h \sum_{j=0}^N \sigma_{N,j}^* f_{n-j} \quad (27)$$

De même que précédemment, on utilise une fonction génératrice $G^*(t)$ donnée par :

$$G^*(t) = \sum_{j=0}^{\infty} \xi_j^* t^j = \frac{-t}{\ln(1-t)} \quad (28)$$

En identifiant les coefficients de même puissance en t , on trouve :

$$\left| \begin{array}{l} \xi_0^* = 1 \\ \xi_1^* = \frac{-1}{2} \\ \xi_2^* = \frac{-1}{12} \end{array} \right| \left| \begin{array}{l} \xi_3 = \frac{-1}{24} \\ \xi_3 = \frac{-19}{720} \\ \xi_4 = \frac{-3}{160} \end{array} \right| \left| \begin{array}{l} \xi_6 = \frac{-863}{60480} \\ \xi_7 = \dots \\ \xi_8 = \dots \end{array} \right|$$

En rassemblant les coefficients de même ordonnée, la formule AM apparaît sous la forme :

$$\sigma_{N,j}^* = (-1)^j \left[C_j^j \xi_j^* + C_j^{j+1} \xi_{j+1}^* + \dots + C_j^N \xi_N^* \right] \quad (29)$$

La différence de cette méthode avec AB réside dans le fait que seules les valeurs y_{p-1}, \dots sont connues, et que l'on détermine y_p au lieu de y_{p+1} .

D'autre part, la forme particulière de l'équation suggère une procédure itérative, car y_p apparaît aussi dans le membre de droite, comme argument de f_p . La solution est très rapide si h est suffisamment petit et si le prédicteur est suffisamment précis.

Méthodes de Prédiction-Correction

Ces méthodes profitent des avantages des méthodes fermées type Adams-Moulton (de très bonne précision), en réduisant les temps de calcul élevés : il suffit de trouver une valeur estimée $y_n^{(0)}$ proche de la valeur finale y_n . Pour cela, $y_n^{(0)}$ (appelé prédiction) est obtenue par

une formule ouverte type Adams-Bashforth de même ordre. C'est cette valeur $y_n^{(0)}$ qui est introduite dans le membre de droite de la formule fermée.

Ces méthodes contiennent en particulier les méthodes AB et AM décrites ci-dessus. Le processus de prédiction-correction (PC) s'arrête lorsque $[f_n^{(\nu)} - f_n^{(\nu-1)}]$ est négligeable (test établi par rapport à un critère préalablement choisi). La dernière valeur $f_n^{(\nu)}$ est prise comme valeur finale f_n .

Puisque la valeur y_n peut être entachée d'erreurs d'arrondi, suite à la sommation des séries, il faut être très prudent et, si possible, réévaluer le schéma AM une dernière fois.

Les constantes de stabilité ($\alpha_N = \sum_j |\sigma_{N,j}|$ et $\alpha_N^* = \sum_j |\sigma_{N,j}^*|$) sont beaucoup plus grandes pour les méthodes AB que pour les méthodes AM, surtout lorsque N croît. En fait, le domaine de stabilité absolu est faible pour une méthode (LM) explicite. C'est une des raisons pour lesquelles on utilise de préférence AM, malgré les complications apportées par le schéma implicite. Les méthodes AM sont plus stables et plus précises.

L'erreur d'interpolation ($f(t) - p_n(t)$) tend-elle vers zéro lorsque $n \rightarrow \infty$?

→ pour un support t_j convenablement choisi, la réponse est certes positive (il y a convergence), à condition toutefois que $f(t)$ soit "raisonnable". C'est en particulier le cas pour les fonctions développables en série entière autour d'un point quelconque de l'intervalle choisi pour définir $f(t)$.

→ dans certains cas, en particulier celui des points t_j régulièrement espacés, une divergence peut apparaître si n est trop grand : ce sont des fortes oscillations du polynôme entre les points t_j .

⇒ d'où la nécessité de ne pas utiliser un ordre trop élevé.

3.5.2 Méthode de Bulirsch & Stoer

C'est une méthode à pas liés qui intègre des équations différentielles du premier degré du type : $\dot{y} = f(t, y(t))$, avec $y(0) = y_0$. Son intérêt réside dans son algorithme de prédiction-correction très efficace et rapide, ainsi que dans une stabilité très importante.

C'est une méthode dite d'extrapolation : l'idée consiste à évaluer $y(t_0 + H)$ sur un certain nombre de sous-pas (h_s) de plus en plus petits par rapport au pas "global" H , afin d'extrapoler à $h = 0$, une meilleure valeur de y : $\tilde{y}(t_0 + H)$. L'inconvénient de cette méthode est la valeur élevée du pas global H qui diminue fortement la résolution temporelle de la solution.

Le prédicteur est à pas multiples, c'est-à-dire qu'il fournit plusieurs valeurs possibles $P(h_s, t)$ - vecteur de \mathbb{R}^n - de $y(t_0 + H)$ suivant le sous-pas h_s utilisé. Le correcteur donne la meilleure estimation, \tilde{y} , de $y(t_0 + H)$ d'après les différentes valeurs $P(h_s, t)$, en extrapolant ces valeurs pour $h = 0$ à l'aide d'une fraction rationnelle.

Algorithme de prédiction

La méthode numérique la plus simple (symétrique et explicite) est la méthode du point milieu, méthode de base telle que l'erreur soit une fonction paire de h (car on obtient une précision

double pour un même travail) :

$$\begin{aligned} y(t_0) &= y_0 \\ y_1 &= y_0 + h_s f(t_0, y_0) \\ y_{n+1} - y_{n-1} &= 2h f(t_n, y_n) \end{aligned}$$

Gragg (1965) a montré que cette méthode possède les propriétés de symétrie nécessaires. Malheureusement, la méthode du point milieu a un domaine de stabilité absolue vide, même sur un petit nombre de pas. D'où l'adoption d'une méthode modifiée : on prend un pas global H et un nombre $S + 1$ de sous-pas h_s avec : $N_s = [1, 2, 3, 4, 6, 8, 12, 16, \dots]$ pour $s = [0, 1, 2, 3, \dots, S]$ et $h_s = H/N_s$.

Cette suite de sous-pas est en effet celle qui rend la sensibilité de la méthode minimale aux erreurs d'arrondi. Avec cette suite, tendant vers zéro, les $(S + 1)$ déterminations $P(h_s)$ de $\tilde{y}(t_0 + H)$ sont telles que :

$$\tilde{y}(t_0 + H, h_s) = \frac{1}{4}[\tilde{y}_{N_s+1} + 2\tilde{y}_{N_s} + \tilde{y}_{N_s-1}] \quad (30)$$

ce qui correspond bien à une méthode d'ordre 2.

Il est possible de former une combinaison linéaire des valeurs $\tilde{y}(t_0 + H, h_s)$ pour approcher $\tilde{y}(t_0 + H)$; ceci est équivalent à l'extrapolation à $h = 0$ d'un polynôme passant par les $(S + 1)$ valeurs calculées. Le processus est efficace si l'on peut trouver une méthode numérique pour laquelle le développement asymptotique se mettra sous la forme :

$$\tilde{P}(h, t) = \tilde{y}(t) + h^2 \tau_1(t) + h^4 \tau_2(t) + \dots$$

Cette extrapolation est cependant meilleure lorsque l'on choisit une fonction rationnelle [6].

Algorithme de correction

Soit la fraction rationnelle $\mathcal{C}_S^{(i)} : \mathbb{R} \rightarrow \mathbb{R}^n$. La valeur estimée de $y(t)$ est : $\tilde{y} = \mathcal{C}_S^{(i)}(0) \simeq \mathcal{C}_S^{(i)}$; ces valeurs se calculent par récurrence (triangle de Romberg) :

$$\begin{aligned} \mathcal{C}_{-1}^{(i)} &= 0 \\ \mathcal{C}_0^{(i)} &= P(h_i, x) \\ \mathcal{C}_k^{(i)} &= \mathcal{C}_{k-1}^{(i+1)} + \frac{\mathcal{C}_{k-1}^{(i+1)} - \mathcal{C}_{k-1}^{(i)}}{\mathcal{A}} \quad (k \geq 1) \end{aligned} \quad (31)$$

avec :

$$\mathcal{A} = \left(\frac{h_i}{h_{i+k}} \right)^2 \left[1 - \frac{\mathcal{C}_{k-1}^{i+1} - \mathcal{C}_{k-1}^i}{\mathcal{C}_{k-1}^{i+1} - \mathcal{C}_{k-2}^{i+1}} \right]$$

Pour éviter la formation de différences répétées, on peut établir un calcul par récurrence de ces différences. Ces formules sont calculées successivement pour $s = 0, 1, 2, \dots$, l'indexage sur k est choisi afin d'indiquer la séquence de calculs en cours.

Un test de précision est effectué afin d'arrêter le processus de recherche de la solution par les sous-pas. Avec un facteur de précision de l'ordre de 10^{-14} , environ 10 étapes d'extrapolation suffisent en général.

3.6 Intégrateurs linéaires du second degré

Il s'agit d'intégrer ici une équation du type : $\ddot{y} = f(t, y(t))$, donc de passer directement de la dérivée seconde à la fonction inconnue. Cette formulation est donc beaucoup utilisée en géodésie spatiale (cf. équation de la dynamique (1)). L'algorithme PC général est le suivant :

Prédiction

$$y_{n+1}^{(1)} + \sum_{p=1}^P \alpha_p y_{n+1-p} = h^2 \sum_{p=1}^N \beta_p f_{n+1-p} \quad (32)$$

Correction

$$y_{n+1} + \sum_{p=1}^M \alpha_p^* y_{n+1-p} = h^2 \left[\sum_{p=1}^R \beta_p^* f_{n+1-p} + \beta_0^* f(t_{n+1}, y_{n+1}^{(1)}) \right] \quad (33)$$

La première formule prédit une valeur $y_{n+1}^{(1)}$, valeur qui est ensuite corrigée par la seconde formule, étape généralement itérée. Cet algorithme PC est écrit en "forme lagrangienne", c'est-à-dire en fonction des valeurs descendantes f_{n+1-p} et non en fonction des "différences rétrogrades". De nombreux algorithmes sont souvent représentés sous forme de différences.

3.6.1 Méthode de (Adams-) Störmer

Parmi les méthodes les plus répandues, on a l'algorithme PC de Störmer de schéma général :

Prédiction (explicite) :

$$y_{n+1} - 2y_n + y_{n-1} = h^2 \sum_{p=0}^K b_{-p} f_{n-p}$$

Correction (implicite) :

$$y_{n+1} - 2y_n + y_{n-1} = h^2 \sum_{p=-1}^K b_{-p} f_{n-p}$$

Le mode opératoire est le suivant. D'après le développement de Taylor en t_n de y , avec un reste intégrale :

$$y(t_{n+1}) - 2y(t_n) + y(t_{n-1}) = h^2 \int_{t_{n-1}}^{t_{n+1}} G(t) \ddot{y}(t) dt$$

où l'on remplace \ddot{y} par son polynôme d'interpolation de Newton, aux points (t_n, \dots, t_{n-K}) :

$$\ddot{y}(t) = \ddot{y}(t_n) + \frac{\Delta_- \ddot{y}(t_n)}{h} (t - t_n) + \dots + \frac{\Delta_-^K \ddot{y}(t_n)}{h^K K!} (t - t_n) \dots (t - t_{n-K+1})$$

Après intégration :

$$y(t_{n+1}) - 2y(t_n) + y(t_{n-1}) = h^2 \sum_{p=0}^K \beta_p \Delta_-^p \ddot{y}(t_n) + r_n^K$$

avec :

$$\beta_p = \frac{1}{p!} \int_0^1 (1-t)t \dots (t+p-1) dt + \frac{1}{p!} \int_0^{-1} (-1-t)t \dots (t+p-1) dt$$

Exemple : $\beta_p = 1, 0, 1/12, 1/12, 19/240, 3/40, \dots$

D'où :

$$y_{n+1}^* - 2y_n + y_{n-1} = h^2 \sum_{p=0}^K \beta_p \Delta_-^p f_n \quad (\text{prédiction})$$

Et si l'on remplace \ddot{y} par son polynôme d'interpolation aux points $(t_{n+1}, \dots, t_{n-K+1})$, on obtient :

$$y_{n+1} - 2y_n + y_{n-1} = h^2 \sum_{p=0}^K \gamma_p \Delta_-^p f_{n+1}^* \quad (\text{correction})$$

avec :

$$\gamma_p = \frac{1}{p!} \int_0^1 (1-t)(t-1)t \dots (t+p-2) dt + \frac{1}{p!} \int_0^{-1} (-1-t)(t-1)t \dots (t+p-2) dt$$

Exemple : $\gamma_p = 1, -1, 1/12, 0, -1/240, -1/240, \dots$

La méthode est parfois appelée méthode d'Adams-Störmer.

3.6.2 Méthode d'Adams-Moulton-Cowell (AMC)

Cette méthode a été programmée il y a plusieurs années par G. Balmino et son équipe, en plus de la méthode de Cowell, afin d'intégrer les équations du satellite artificiel. Nous l'avons reprise, et reprogrammée (travaux de P.Schaeffer, 1993), afin d'effectuer des tests.

L'algorithme provient des travaux de Cohen et Hubbard, lui-même re-écrit dans l'article d'Oesterwinter & Cohen duquel nous nous sommes inspirés ([13]). Sa particularité est de mélanger les coefficients de formules d'Adams, avec des coefficients de Cowell; on a donc, en quelques sortes, deux intégrateurs en un : l'un pour passer de l'accélération à la position, l'autre pour passer de l'accélération à la vitesse.

En tant que méthode à pas liés, l'algorithme utilise le tableau d'initialisation suivant (avec les valeurs typiques : $c = 14$, qui est l'ordre de démarrage, $a = b = 7$, et $m = 12$ l'ordre de

progression) :

$$\begin{array}{c}
 | \quad - \quad c \\
 | \quad - \quad (c-1) \\
 \vdots \\
 | \quad - \quad (a+1) \\
 | \quad - \quad a \\
 \vdots \\
 | \quad - \quad 1 \\
 | \quad \quad 0 \\
 | \quad \quad 1 \\
 \vdots \\
 | \quad - \quad (b-1) \\
 | \quad - \quad b \\
 | \quad - \quad (b+1) \\
 \vdots \\
 | \quad \quad (c-1) \\
 | \quad \quad c \\
 \vdots
 \end{array}$$

La phase de démarrage est la suivante (h est le pas) :

$$\begin{aligned}
 \ddot{r}_n &= \ddot{r}_0 && \text{pour : } -a \leq n \leq b \\
 \ddot{r}_{-1} &= \ddot{r}_0 - h \dot{r}_0 + h^2 \sum_{i=0}^{a+b} \gamma_i \ddot{r}_{b-i}
 \end{aligned}$$

Les accélérations :

$$\begin{aligned}
 \ddot{r}_{-n} &= \sum_{j=0}^{a+b} b_j \ddot{r}_{-n+1+j} && \text{pour : } n = a+1, a+2, \dots, c \\
 \ddot{r}_n &= \sum_{j=0}^{a+b} b_j \ddot{r}_{n-1-j} && \text{pour : } n = b+1, b+2, \dots, c
 \end{aligned}$$

Les positions et vitesses, uniquement sur $[a, b]$:

$$\begin{aligned}
 \bar{r}_n &= 2\bar{r}_{n-1} - \bar{r}_{n-2} + h^2 \sum_{j=0}^c \alpha_j^* \ddot{r}_{n-j} && \text{pour : } n = 1, 2, \dots, b \\
 \bar{r}_{-n} &= 2\bar{r}_{-n+1} - \bar{r}_{-n+2} + h^2 \sum_{j=0}^c \alpha_j^* \ddot{r}_{-n+j} && \text{pour : } n = 2, 3, \dots, a
 \end{aligned}$$

et :

$$\begin{aligned}
 \dot{r}_n &= \dot{r}_{n-1} + h \sum_{j=0}^c \beta_j^* \ddot{r}_{n-j} && \text{pour : } n = 1, 2, \dots, b \\
 \dot{r}_{-n} &= \dot{r}_{-n+1} + h \sum_{j=0}^c \beta_j^* \ddot{r}_{-n+j} && \text{pour : } n = 1, 2, \dots, a
 \end{aligned}$$

A ce stade, le programme doit contrôler le nombre de cycles nécessaires à la convergence (en général 4 sont suffisants); pour cela, il est recommandé de tester la différence absolue dans les accélérations successives par rapport à une valeur a priori.

La phase de progression est double (puisque nous sommes dans un schéma de prédiction et correction); soit, la prédiction :

$$\begin{aligned}\bar{r}_n &= 2\bar{r}_{n-1} - \bar{r}_{n-2} + h^2 \sum_{j=0}^m \alpha'_j \ddot{r}_{n-1-j} \\ \dot{\bar{r}}_n &= \dot{\bar{r}}_{n-1} + h \sum_{j=0}^m \beta'_j \ddot{r}_{n-1-j}\end{aligned}$$

et la correction, où de nouvelles accélérations seront calculées, soit :

$$\begin{aligned}\bar{r}_n &= 2\bar{r}_{n-1} - \bar{r}_{n-2} + h^2 \sum_{j=0}^m \alpha_j \ddot{r}_{n-j} \\ \dot{\bar{r}}_n &= \dot{\bar{r}}_{n-1} + h \sum_{j=0}^m \beta_j \ddot{r}_{n-j}\end{aligned}$$

De même qu'en phase de démarrage, il est important de tester la convergence de ce schéma. Nous avons cependant modifié légèrement l'algorithme en choisissant de tester non plus les accélérations mais les vitesses successives. Nous avons pu contrôler que cela nous apporte une plus grande stabilité numérique des solutions (en référence à une méthode a priori "plus stable", celle de Bulirsch & Stoer).

3.6.3 Méthode de Cowell

Cette méthode est une des rares méthodes qui permettent d'intégrer directement des équations différentielles du second degré. Elle a été appliquée par Cowell et Crommelin pour mettre en évidence l'existence du 8^{ème} satellite de Jupiter, ainsi que pour décrire très correctement le retour de la comète de Halley en 1910.

C'est une méthode à pas liés qui nécessite la connaissance d'une position et d'une vitesse à l'instant t_0 (conditions initiales du mouvement). Elle utilise l'interpolation polynomiale de Stirling afin d'exprimer les dérivées en termes de différences. L'algorithme de Cowell à l'ordre $2p$ intègre exactement toute solution polynomiale de degré $2p$.

On peut noter aussi que cette méthode possède une extension, créée par Kulikov, qui permet de changer le pas en cours d'intégration.

Soit h le pas d'intégration. On considère deux valeurs de l'argument t : $t = t_k - h$ et $t = t_k + h$. On développe ensuite la valeur de la "position" \mathbf{x} en fonction de ses dérivées, à l'aide d'une série de Taylor :

$$\mathbf{x}(t_k - h) = \mathbf{x}_{k-1} = \mathbf{x}_k + \sum_{n=1}^{\infty} (-1)^n \frac{h^n}{n!} \left(\frac{d^n \mathbf{x}}{dt^n} \right)_k \quad (34)$$

et :

$$\mathbf{x}(t_k + h) = \mathbf{x}_{k+1} = \mathbf{x}_k + \sum_{n=1}^{\infty} \frac{h^n}{n!} \left(\frac{d^n \mathbf{x}}{dt^n} \right)_k \quad (35)$$

On introduit la différence du premier ordre Δ^1 comme étant :

$$\begin{aligned} \Delta_{k-1/2}^1 &= \Delta^1 \mathbf{x}_{k-1/2} = \mathbf{x}_k - \mathbf{x}_{k-1} \\ \Delta_{k+1/2}^1 &= \Delta^1 \mathbf{x}_{k+1/2} = \mathbf{x}_{k+1} - \mathbf{x}_k \end{aligned}$$

Dès lors, la différence du second ordre est :

$$\Delta_k^2 = \Delta_{k+1/2}^1 - \Delta_{k-1/2}^1 = \mathbf{x}_{k+1} - 2\mathbf{x}_k + \mathbf{x}_{k-1}$$

En utilisant les deux premières relations (34) et (35), on obtient

$$\Delta^2 \mathbf{x}_k = 2 \sum_{n=1}^{\infty} \frac{h^{2n}}{2n!} \left(\frac{d^{2n} \mathbf{x}}{dt^{2n}} \right)_k$$

A partir de l'équation du mouvement $d^2 \mathbf{x}/dt^2 = R(\mathbf{x}, t)$, et avec $f = h^2 \ddot{\mathbf{x}} = h^2 R$, on introduit la valeur de \mathbf{x} et l'on obtient :

$$\Delta^2 \mathbf{x}_k = f_k + 2 \sum_0^{\infty} \frac{h^{2n}}{(2n+2)!} \left(\frac{d^{2n} f}{dt^{2n}} \right)_k \quad (36)$$

Les expressions des dérivées $(d^n f/dt^n)_k$, en terme de différences, sont données par des formules d'interpolation polynomiale. Pour Cowell, on utilise la formule de Stirling :

$$f(t_k + zh) = f_k + z f_k^1 + \frac{z^2}{2!} f_k^2 + \frac{z(z^2 - 1)}{3!} f_k^3 + \frac{z^2(z^2 - 1)}{4!} f_k^4 + \dots \quad (37)$$

\Rightarrow A titre d'exemple, la formule de Bessel permet d'exprimer les dérivées de f en t_k en terme de différences, pour les lignes $n = k + 1/2$, et est exploitée dans une variante de la méthode de Cowell :

$$\begin{aligned} f(t_k + zh) &= f_k + z f_{k+1/2}^1 + \frac{z(z-1)}{2!} f_{k+1/2}^2 + \frac{z(z-1)(z-2)}{3!} f_{k+1/2}^3 \\ &+ \frac{(z+1)z(z-1)(z-2)}{4!} f_{k+1/2}^4 + \dots \end{aligned}$$

A partir de la formule de Stirling (37), les dérivées sont obtenues en dérivant $f(t_k + zh)$ par rapport à z . On a : $df/dz = (df/dt).(dt/dz)$ et, comme $t = t_k + zh$, $dt/dz = h$ et $df/dz = h(df/dt)_k$, ainsi :

$$\frac{df}{dz} = f_k^1 + \frac{2z}{2!} f_k^2 + \frac{3z^2 - 1}{3!} f_k^3 + \frac{4z^3 - 2z}{4!} f_k^4 + \dots$$

De même :

$$\frac{d^2 f}{dz^2} = \frac{2}{2!} f_k^2 + \frac{6z}{3!} f_k^3 + \frac{12z^2 - 2}{4!} f_k^4 + \dots$$

soit :

$$h^2 \left(\frac{d^2 f}{dt^2} \right)_k = f_k^2 - \frac{1}{12} f_k^4 + \dots \quad (38)$$

On trouve ainsi la formule de base de la méthode de Cowell :

$$\Delta^2 \mathbf{x}_k = f_k + \frac{1}{12} f_k^2 - \frac{1}{240} f_k^4 + \frac{31}{60480} f_k^6 + \dots \quad (39)$$

La formule de Cowell donnant directement \mathbf{x}_k est obtenue par deux sommations de la série précédente, d'abord sur : $k' = j, (0, i)$ et ensuite sur $i, (0, R - 1)$. On obtient finalement :

$$\mathbf{x}_k = f_k^{-2} + \frac{1}{12} f_k - \frac{1}{240} f_k^2 + \frac{31}{60480} f_k^4 - \frac{289}{3628800} f_k^6 + \frac{317}{22809600} f_k^8 + \dots \quad (40)$$

avec :

$$f_k^{-2} = f_0^{-2} + \sum_{i=0}^{k-1} f_{i+1/2}^{-1} \quad (41)$$

$$f_{i+1/2}^{-1} = f_{-1/2}^{-1} + \sum_{j=0}^i f_j$$

Les valeurs initiales des sommes f_0^{-2} et $f_{-1/2}^{-1}$ sont arbitraires ; elles sont choisies grâce à la formule (38), à condition que les données de départ soient définies en terme de position et de vitesse à l'instant t_0 :

$$\begin{aligned} f_0^{-2} &= \mathbf{x}_0 - \frac{1}{12} f_0 + \frac{1}{240} f_0^2 - \frac{31}{60480} f_0^4 + \dots \\ f_{-1/2}^{-1} &= h \dot{\mathbf{x}}_0 - \frac{1}{2} f_0 + \frac{1}{12} f_0^1 - \frac{11}{720} f_0^3 + \dots \end{aligned} \quad (42)$$

Les relations précédentes représentent le formalisme mathématique de Cowell. Pour débiter le procédé d'intégration, il faut calculer la table suivante :

$$f_0^{-2}, f_{-1/2}^{-1}, f_{-2}, f_{-1}, f_0, f_1, f_2$$

De plus, lorsque l'on calcule les valeurs initiales des deux premières sommes, équations (42), on doit faire attention à la précision, notamment pour la seconde, car l'erreur sur ce terme est proportionnelle au temps.

Ensuite, le second membre des équations de la dynamique doit être calculé pour différents instants autour de t_0 :

$$t_0 - 3h, t_0 - 2h, t_0 - h, t_0, t_0 + h, t_0 + 2h, t_0 + 3h$$

En partant du fait que les différences du sixième ordre sont constantes (hypothèse principale de la méthode), et en tenant compte de la relation (38) et de la formule donnant les ordonnées f_k (due à Tokmalayena), les formules utilisées dans les programmes sont les suivantes :

$$\begin{aligned}
x_{-3} &= f_0^{-2} - 3f_{-1/2}^{-1} + \frac{237671}{3628800}f_{-3} + \frac{645569}{604800}f_{-2} + \frac{91415}{48384}f_{-1} + \frac{18937}{181440}f_0 \\
&- \frac{14513}{241920}f_1 + \frac{11729}{604800}f_2 - \frac{9829}{3628800}f_3
\end{aligned} \tag{43}$$

$$x_{-2} = \dots$$

⋮

le vecteur initial \mathbf{x}_0 étant déjà disponible. Les valeurs initiales des sommes (42) peuvent être calculées par :

$$\begin{aligned}
f_0^{-2} &= x_0 + \frac{289}{3628800}f_{-3} - \frac{599}{604800}f_{-2} + \frac{1793}{241920}f_{-1} - \frac{2497}{25920}f_0 \\
&+ \frac{1793}{241920}f_1 - \frac{599}{604800}f_2 + \frac{289}{3628800}f_3
\end{aligned} \tag{44}$$

et :

$$\begin{aligned}
f_{-1/2}^{-1} &= h\dot{x}_0 - \frac{191}{120960}f_{-3} + \frac{211}{15120}f_{-2} - \frac{7843}{120960}f_{-1} - \frac{1}{2}f_0 \\
&+ \frac{7843}{120960}f_1 - \frac{211}{15120}f_2 + \frac{191}{120960}f_3
\end{aligned} \tag{45}$$

le vecteur initial $\dot{\mathbf{x}}_0$ étant également disponible.

Une procédure itérative est à mettre en place. On considère le second membre des équations du mouvement comme constant pendant l'intervalle de temps $[t_0 - 3h, t_0 + 3h]$ et égal à f_0 , qui est calculée avec \mathbf{x}_0 . Les sommes (42) et (43) sont calculées, puis les valeurs de f pour tous les temps en utilisant (41) et $f = h^2\ddot{\mathbf{x}} = h^2R$. Dès lors, les sommes (42) et (43) sont recalculées n fois avec les nouvelles valeurs f_i . Le procédé d'approximations successives est poursuivi jusqu'à ce que les différences entre deux valeurs consécutives des modules de f_0^{-2} , $f_{-1/2}^{-1}$, et f soient inférieures au facteur de convergence choisi au départ.

4 Méthode d'Encke

4.1 Introduction

Le principe de cette méthode est d'intégrer les équations du mouvement d'un corps S par rapport au mouvement de référence d'un corps fictif \tilde{S} ; ce dernier étant prédéterminé en fonction des conditions initiales et de l'environnement (modèle physique). Ainsi, les grandeurs caractéristiques du mouvement différentiel conservent des amplitudes faibles, ce qui permet de gagner en précision. Dans le cadre des calculs d'orbites, l'application de cette méthode consiste ainsi à intégrer le mouvement du satellite par rapport à une orbite "moyenne" comme l'illustre la figure (2).

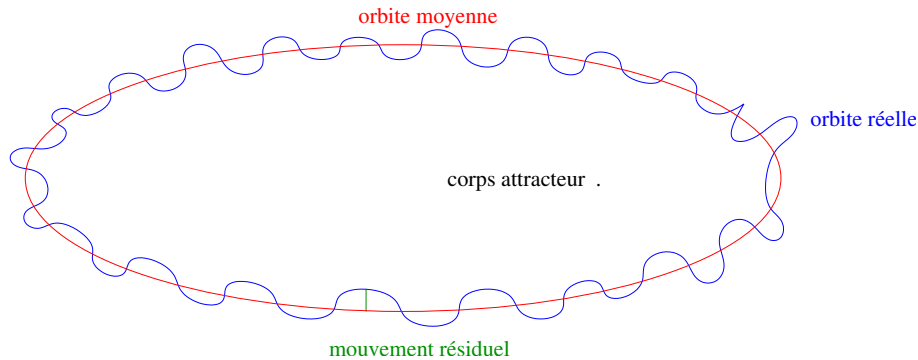


FIG. 2 – Méthode d'Encke pour le calcul d'orbites.

Pour les calculs d'orbites, on peut utiliser des variables cartésiennes (coordonnées rectangulaires) pour lesquelles les équations du mouvement constituent un système différentiel de degré 2 et de dimension 3 ou des éléments elliptiques pour lesquels les équations du mouvement sont les équations de Lagrange et/ou les équations de Gauss [5].

Les mouvements de S et de \tilde{S} doivent être décrits suivant le même type de variables, même si, dans la pratique, on part toujours des éléments elliptiques usuels (moyens) pour décrire la trajectoire de \tilde{S} .

4.2 Mouvement de référence

Dans le cas de satellites artificiels terrestres, on sait que les orbites s'écartent peu d'ellipses keplériennes en précession. En éléments elliptiques, ceci se traduit, pour cette ellipse moyenne de référence, par les relations :

$$\begin{aligned}\tilde{a} &= \tilde{a}_0 \\ \tilde{e} &= \tilde{e}_0 \\ \tilde{I} &= \tilde{I}_0 \\ \tilde{\Omega} &= \tilde{\Omega}_0 + \dot{\tilde{\Omega}}_0(t - t_0)\end{aligned}$$

$$\begin{aligned}\tilde{\omega} &= \tilde{\omega}_0 + \dot{\tilde{\omega}}_0(t - t_0) \\ \tilde{M} &= \tilde{M}_0 + \dot{\tilde{M}}_0(t - t_0)\end{aligned}$$

où les éléments initiaux $(\tilde{a}_0, \tilde{e}_0, \tilde{I}_0, \tilde{\Omega}_0, \tilde{\omega}_0, \tilde{M}_0)$ ainsi que les termes de variations séculaires $(\dot{\tilde{\Omega}}_0, \dot{\tilde{\omega}}_0, \dot{\tilde{M}}_0 = \dot{\tilde{M}}_0)$ dépendent des caractéristiques moyennes du mouvement ainsi que des forces considérées.

Pour déterminer cette ellipse précessante "moyenne", on peut effectuer une intégration numérique avec un modèle de forces simplifié pour déterminer les éléments elliptiques à partir desquels on peut, par moindres carrés, déterminer les éléments moyens. On peut également utiliser des théories analytiques telles que [4], fournissant les éléments elliptiques osculateurs en fonction des éléments moyens.

Si l'on utilise des coordonnées rectangulaires, il existe des relations "classiques" permettant de déterminer ces coordonnées à partir des éléments elliptiques (voir [5], entre autres).

4.3 Equations différentielles

4.3.1 Cas des éléments elliptiques usuels

Dans le cas des éléments elliptiques "usuels", on part des équations de Gauss-Lagrange de la forme :

$$\begin{aligned}\dot{a} &= f_a \\ \dot{e} &= f_e \\ \dot{I} &= f_I \\ \dot{\Omega} &= f_{\Omega} \\ \dot{\omega} &= f_{\omega} \\ \dot{M} &= f_M\end{aligned}$$

et on utilise les décompositions $a = \tilde{a}_0 + \Delta a, \dots, \Omega = \tilde{\Omega}_0 + \dot{\tilde{\Omega}}_0(t - t_0) + \Delta\Omega, \dots$ pour finalement aboutir au système différentiel :

$$\begin{aligned}\frac{d\Delta a}{dt} &= f_a \\ \frac{d\Delta e}{dt} &= f_e \\ \frac{d\Delta I}{dt} &= f_I \\ \frac{d\Delta\Omega}{dt} &= f_{\Omega} - \dot{\tilde{\Omega}}_0 \\ \frac{d\Delta\omega}{dt} &= f_{\omega} - \dot{\tilde{\omega}}_0\end{aligned}$$

$$\frac{d\Delta M}{dt} = f_M - \bar{n}_0$$

que l'on résout à l'aide d'une méthode d'intégration numérique à partir des conditions initiales : $\Delta a_0 = a_0 - \tilde{a}_0, \dots, \Delta M_0 = M_0 - \tilde{M}_0$.

A chaque pas de calcul de l'intégrateur numérique, on détermine ainsi $a = \tilde{a} + \Delta a, \dots, M = \tilde{M} + \Delta M$ et on peut également avoir accès aux coordonnées rectangulaires à l'aide des relations de passage précédemment évoquées.

4.3.2 Cas des coordonnées rectangulaires

Mouvement de référence. Nous partons du mouvement de référence vu dans le repère mobile défini par :

- le vecteur \bar{P} unitaire passant par le périastre de l'ellipse en précession,
- le vecteur \bar{Q} unitaire, orthogonal à \bar{P} dans le plan de cette ellipse.

Dans ce paragraphe, nous appellerons X, Y, Z les coordonnées du système d'axes de l'intégration numérique ; le repère formé par (O, P, Q) définissant un système de coordonnées (x, y) , avec :

$$\begin{aligned}\bar{r} &= x\bar{P} + y\bar{Q} \\ \dot{\bar{r}} &= \dot{x}\bar{P} + \dot{y}\bar{Q} + x\dot{\bar{P}} + y\dot{\bar{Q}} \\ \ddot{\bar{r}} &= \ddot{x}\bar{P} + \ddot{y}\bar{Q} + 2\dot{x}\dot{\bar{P}} + 2\dot{y}\dot{\bar{Q}} + x\ddot{\bar{P}} + y\ddot{\bar{Q}}\end{aligned}\quad (46)$$

On utilise également les formules suivantes (où E est l'anomalie excentrique) :

$$\begin{aligned}r &= a(1 - e \cos E) \\ M &= E - e \sin E = \bar{n}(t - t_0) \\ x &= a(\cos E - e) = r \cos v \\ y &= a\sqrt{1 - e^2} \sin E = r \sin v \\ \dot{x} &= -a \sin E \dot{E} \\ \dot{y} &= a\sqrt{1 - e^2} \cos E \dot{E} \\ \ddot{x} &= -a(\sin E \ddot{E} - \cos E \dot{E}^2) = a\dot{E}^2 \frac{e - \cos E}{1 - e \cos E} = -a\dot{E}^2 \cos v \\ \ddot{y} &= a\sqrt{1 - e^2}(\cos E \ddot{E} - \sin E \dot{E}^2) = -a\dot{E}^2 \frac{\sqrt{1 - e^2} \sin E}{1 - e \cos E} = -a\dot{E}^2 \sin v\end{aligned}$$

On a donc :

$$\ddot{\bar{r}} = -\frac{a^3}{r^2} \bar{n}^2 \cos v \bar{P} - \frac{a^3}{r^2} \bar{n}^2 \sin v \bar{Q} + \bar{\rho} \quad (47)$$

avec :

$$\bar{\rho} = 2\dot{x} \dot{\bar{P}} + 2\dot{y} \dot{\bar{Q}} + x\ddot{\bar{P}} + y\ddot{\bar{Q}} \quad (48)$$

Attention, alors que l'on a bien $\tilde{n}_0^2 \tilde{a}_0^3 = \mu$, on n'a pas $\bar{n}^2 a^3 = \mu$. On peut transformer l'équation précédente en remplaçant $\cos v$ par x/r et $\sin v$ par y/r ; d'où :

$$\ddot{\bar{r}} = -\bar{n}^2 \frac{a^3}{r^3} (x\bar{P} + y\bar{Q}) + \bar{\rho}$$

c'est-à-dire :

$$\ddot{\vec{r}} = -\bar{n}^2 \frac{a^3}{r^3} \vec{r} + \bar{\rho} \quad (49)$$

Le calcul du vecteur d'accélération apparente $\bar{\rho}$ se fait comme suit :

$$\bar{P} = \begin{pmatrix} \cos \Omega \cos \omega & - \cos I \sin \omega \sin \Omega \\ \sin \Omega \cos \omega & + \cos I \sin \omega \cos \Omega \\ \sin I \sin \omega & \end{pmatrix} = \begin{pmatrix} P_X \\ P_Y \\ P_Z \end{pmatrix} = -\frac{\partial \bar{Q}}{\partial \omega}$$

$$\bar{Q} = \begin{pmatrix} -\cos \Omega \sin \omega & - \cos I \cos \omega \sin \Omega \\ -\sin \Omega \sin \omega & + \cos I \cos \omega \cos \Omega \\ \sin I \cos \omega & \end{pmatrix} = \begin{pmatrix} Q_X \\ Q_Y \\ Q_Z \end{pmatrix} = \frac{\partial \bar{P}}{\partial \omega}$$

D'où (I étant constant pour le mouvement de référence, ainsi que $\dot{\Omega}$ et $\dot{\omega}$) :

$$\begin{aligned} \dot{\bar{P}} &= \frac{\partial \bar{P}}{\partial \Omega} \dot{\Omega} + \frac{\partial \bar{P}}{\partial \omega} \dot{\omega} \\ \dot{\bar{Q}} &= \frac{\partial \bar{Q}}{\partial \Omega} \dot{\Omega} + \frac{\partial \bar{Q}}{\partial \omega} \dot{\omega} \\ \ddot{\bar{P}} &= \frac{\partial^2 \bar{P}}{\partial \Omega^2} \dot{\Omega}^2 + 2 \frac{\partial^2 \bar{P}}{\partial \Omega \partial \omega} \dot{\Omega} \dot{\omega} + \frac{\partial^2 \bar{P}}{\partial \omega^2} \dot{\omega}^2 \\ \ddot{\bar{Q}} &= \frac{\partial^2 \bar{Q}}{\partial \Omega^2} \dot{\Omega}^2 + 2 \frac{\partial^2 \bar{Q}}{\partial \Omega \partial \omega} \dot{\Omega} \dot{\omega} + \frac{\partial^2 \bar{Q}}{\partial \omega^2} \dot{\omega}^2 \end{aligned}$$

avec :

$$\begin{aligned} \frac{\partial \bar{P}}{\partial \Omega} &= \begin{pmatrix} -P_Y \\ P_X \\ 0 \end{pmatrix}, & \frac{\partial \bar{Q}}{\partial \Omega} &= \begin{pmatrix} -Q_Y \\ Q_X \\ 0 \end{pmatrix} \\ \frac{\partial \bar{P}}{\partial \omega} &= \bar{Q}, & \frac{\partial \bar{Q}}{\partial \omega} &= -\bar{P} \\ \frac{\partial^2 \bar{P}}{\partial \Omega^2} &= \begin{pmatrix} -P_X \\ -P_Y \\ 0 \end{pmatrix}, & \frac{\partial^2 \bar{P}}{\partial \Omega \partial \omega} &= \frac{\partial \bar{Q}}{\partial \Omega}, & \frac{\partial^2 \bar{P}}{\partial \omega^2} &= -\bar{P} \\ \frac{\partial^2 \bar{Q}}{\partial \Omega^2} &= \begin{pmatrix} -Q_X \\ -Q_Y \\ 0 \end{pmatrix}, & \frac{\partial^2 \bar{Q}}{\partial \Omega \partial \omega} &= -\frac{\partial \bar{P}}{\partial \Omega}, & \frac{\partial^2 \bar{Q}}{\partial \omega^2} &= -\bar{Q} \end{aligned}$$

On utilise enfin, dans la formule (48), les expressions de x, y, \dot{x}, \dot{y} pour le mouvement de référence.

Mouvement différentiel. Le système des équations du mouvement en coordonnées rectangulaires est re-écrit en mettant en évidence le terme principal du potentiel du corps central, et les autres forces perturbatives $\Delta \bar{F}$:

$$\ddot{\vec{r}} = -\mu \frac{\vec{r}}{r^3} + \Delta \bar{F} \quad (50)$$

et le système pour le mouvement de référence est, d'après ce qui précède :

$$\ddot{\bar{r}}_{\text{ref}} = -\bar{n}^2 \frac{a_{\text{ref}}^3}{r_{\text{ref}}^3} \bar{r}_{\text{ref}} + \bar{\rho}_{\text{ref}} \quad (51)$$

en écrivant explicitement "ref" pour plus de clarté dans les variables (et pour éviter les confusions entre variables \tilde{a} , les vecteurs \bar{a} , et les vecteurs de référence qu'il aurait fallu écrire $\tilde{\tilde{a}}$. Ici, \bar{n} est toujours le "moyen" moyen mouvement de référence (identique à \bar{n}_0 des équations xx).

Des équations (50) et (51), on obtient :

$$\ddot{\bar{r}} - \ddot{\bar{r}}_{\text{ref}} = \Delta\ddot{\bar{r}} = -\mu \left(\frac{\bar{r}}{r^3} - \frac{\bar{n}^2 a_{\text{ref}}^3}{\mu r_{\text{ref}}^3} \bar{r}_{\text{ref}} \right) - \bar{\rho}_{\text{ref}} + \Delta\bar{F} \quad (52)$$

On introduit, afin d'améliorer l'expression entre parenthèses ci-dessus, le moyen mouvement képlérien de référence : $n_{\text{ref}} (= \bar{n}_0) = \sqrt{\mu/a_{\text{ref}}^3} (= \sqrt{\mu/\tilde{a}_0^3})$, et donc :

$$\Delta\ddot{\bar{r}} = -\mu \left[\frac{\bar{r}}{r^3} - \left(\frac{\bar{n}}{n_{\text{ref}}} \right)^2 \frac{\bar{r}_{\text{ref}}}{r_{\text{ref}}^3} \right] - \bar{\rho}_{\text{ref}} + \Delta\bar{F} \quad (53)$$

Sous cette forme, (53) pose un problème numérique car les deux termes du crochet sont très voisins (si l'ellipse de référence est bien choisie). On peut utiliser la double précision (en 128 bits) afin d'effectuer cette soustraction (en ayant bien calculé tous les termes de référence en double précision, avant). Ou bien, on peut aussi transformer cette différence suivant :

Posons : $(\bar{n}/n_{\text{ref}})^2 = 1 + \delta$, et $\bar{\varepsilon} = \Delta\bar{r} = \bar{r} - \bar{r}_{\text{ref}}$ (où δ est proche de zéro dans la pratique, et $\|\bar{\varepsilon}\|$ reste petit) ; on écrit alors :

$$\begin{aligned} \frac{\bar{r}}{r^3} - (1 + \delta) \frac{\bar{r}_{\text{ref}}}{r_{\text{ref}}^3} &= \frac{1}{r_{\text{ref}}^3} \left[\bar{r} \frac{r_{\text{ref}}^3}{r^3} - (1 + \delta) \bar{r}_{\text{ref}} \right] \\ &= \frac{1}{r_{\text{ref}}^3} \left[\bar{r} \left(\left(\frac{r_{\text{ref}}}{r} \right)^3 - 1 \right) + \bar{\varepsilon} - \delta \bar{r}_{\text{ref}} \right] \end{aligned}$$

On a donc résumé le problème à une différence de scalaires, proches de l'unité, au lieu d'une différence de vecteurs. Soit alors $\beta = (r_{\text{ref}}/r)^3 - 1$, cette différence. Nous allons transformer cette opération en une somme de petites quantités, en posant : $(r_{\text{ref}}/r)^2 = 1 + \alpha$, et $\beta = (1 + \alpha)^{3/2} - 1$; d'où :

$$\begin{aligned} \Delta\ddot{\bar{r}} &= -\mu\bar{D} + \Delta\bar{F} - \bar{\rho}_{\text{ref}} \\ \bar{D} &= [\beta(\Delta\bar{r} + \bar{r}_{\text{ref}}) + \Delta\bar{r} - \delta \bar{r}_{\text{ref}}] \end{aligned} \quad (54)$$

avec les conditions initiales suivantes : $\Delta\bar{r}_0 = \bar{r}_0 - \bar{r}_{\text{ref}}(t_0)$ et $\Delta\dot{\bar{r}}_0 = \dot{\bar{r}}_0 - \dot{\bar{r}}_{\text{ref}}(t_0)$

5 La circularisation

Les méthodes classiques d'intégration numérique (méthode de Cowell, méthodes d'Adams-Bashforth, d'Adams-Moulton, ..., voir sections précédentes) sont basées sur des approximations polynomiales et intègrent donc exactement des fonctions polynomiales. Lorsque ces méthodes sont appliquées aux équations de la mécanique céleste et, plus particulièrement, aux équations du mouvement d'un satellite d'un corps central (à dominantes périodiques), elles peuvent théoriquement être sources d'instabilités numériques (par la propagation des erreurs).

A l'époque où les ordinateurs n'étaient pas des calculateurs aussi performants qu'aujourd'hui (fin des années 60, début des années 70), une méthode permettant de modifier les coefficients caractéristiques des méthodes d'intégration pour intégrer exactement des fonctions circulaires (d'où le nom de circularisation) a été développée par Bettis et Stiefel, [14], [2] et [3]. Le but de ce chapitre est de présenter cette méthode et son utilisation dans le cadre de la dynamique orbitale.

5.1 Présentation générale de la méthode appliquée à l'algorithme de Cowell

On s'intéresse aux deux types d'équations différentielles $\ddot{x}(t) = f(x, t)$ et $\dot{x}(t) = f(x, t)$. On résout ces équations par intégration numérique à l'aide des différences de la fonction f . On se donne un pas d'intégration h et on définit ces différences par la relation de récurrence :

$$\Delta^m f_n = \Delta^{m-1} f_{n+\frac{1}{2}} - \Delta^{m-1} f_{n-\frac{1}{2}} \quad (55)$$

sachant que : $\Delta^0 f_n = f_n = f(x(nh), nh)$.

On résout ainsi les deux équations précédentes, respectivement, à l'aide des deux relations suivantes qui, dans la pratique, sont tronquées à un indice m donné, ordre de la méthode utilisée :

$$\Delta^2 x_0 = h^2 \sum_{k=0}^{\infty} \alpha_k \Delta^k f_{p-\frac{k}{2}} \quad (56)$$

$$x_1 - x_0 = h \sum_{k=0}^{\infty} \alpha_k \Delta^k f_{p-\frac{k}{2}} \quad (57)$$

Rappel :

- si $p = 1$ dans (56), on utilise la méthode de Cowell,
- si $p = 0$, c'est la méthode de Störmer,
- si $p = 1$ dans (57), on utilise la méthode d'Adams-Moulton,
- si $p = 0$, c'est la méthode d'Adams-Bashforth.

Dans toute la suite de cette partie, on ne s'intéresse plus qu'à la seule méthode de Cowell [14]. On se place dans le contexte de la résolution de l'équation différentielle $\ddot{x}(t) = f(x, t)$ avec cette méthode, d'ordre n ,

Rappel :

$$\Delta^2 x_0 = h^2 \sum_{k=0}^n \alpha_k \Delta^k f_{1-\frac{k}{2}}$$

Les coefficients de la méthode classique peuvent être obtenus en utilisant la fonction spécifique $f(x, t) = z^{t/h}$. En résolvant l'équation différentielle avec cette fonction, on aboutit à :

$$x(t) = \frac{h^2}{(\ln z)^2} z^{\frac{t}{h}} \quad \text{et :} \quad \Delta^2 x_0 = \frac{zh^2}{(\ln z)^2} \left[1 - \frac{1}{z}\right]^2$$

On peut montrer que :

$$\Delta^k f_{1-\frac{k}{2}} = z \left[1 - \frac{1}{z}\right]^k$$

En posant $\varepsilon = 1 - z^{-1}$, on obtient ainsi la relation :

$$\left(\frac{\varepsilon}{\ln(1-\varepsilon)}\right)^2 = P(\varepsilon)$$

avec $P(\varepsilon) = \sum_{k=0}^n \alpha_k \text{varepsilonpsilon}^k$. Effectuant alors le développement en série entière de la fonction génératrice $\varepsilon \mapsto P(\varepsilon)$, on obtient les coefficients classiques α_k par identification. A l'ordre 6, par exemple, on a [14] :

$$\left(\frac{\varepsilon}{\ln(1-\varepsilon)}\right)^2 \simeq 1 - \varepsilon + \frac{1}{12}\varepsilon^2 - \frac{1}{240}\varepsilon^4 - \frac{1}{240}\varepsilon^5 - \frac{221}{60480}\varepsilon^6 \quad (58)$$

On désire adapter cette méthode de Cowell ("classique") pour intégrer exactement des fonctions circulaires.

5.1.1 Position du problème

Supposons tout d'abord que l'on souhaite intégrer exactement les deux fonctions ($t \mapsto \cos \omega t$) et ($t \mapsto \sin \omega t$). La fonction spécifique devient donc : $f(x, t) = e^{i\omega t}$. Afin de se ramener à Cowell (où $f(x, t) = z^{\frac{t}{h}}$, il faut donc poser :

$$z = e^{2i\sigma} \quad \text{et :} \quad \sigma = \omega h/2$$

Pour que la méthode de Cowell intègre exactement les deux fonctions circulaires précédentes, il faut donc que la relation définissant $P(\varepsilon)$ soit vérifiée pour :

$$\varepsilon = 1 - e^{-2i\sigma} \quad \text{et :} \quad \varepsilon = 1 - e^{2i\sigma}$$

Ceci revient à modifier deux coefficients du polynôme P , les deux coefficients de plus haut degré, par exemple ; les autres coefficients demeurent ceux de la méthode classique.

Le nouveau jeu de coefficients α_k permet alors d'intégrer exactement toute fonction de la forme $(t \mapsto Q_{n-2}(t) + a \cos \omega t + b \sin \omega t)$, Q_{n-2} désignant une fonction polynôme de degré inférieur ou égal à $n - 2$ [2].

On suppose maintenant que l'ordre de la méthode est pair ($n = 2\nu$) et on souhaite intégrer exactement toute fonction de la forme :

$$t \mapsto a_0 + \sum_{j=1}^{\nu} (a_j \cos \omega_j t + b_j \sin \omega_j t)$$

donc possédant ν fréquences. Pour ce faire, le polynôme P doit vérifier les $2\nu + 1$ relations :

$$\left(\frac{\varepsilon_j}{\ln(1 - \varepsilon_j)} \right)^2 = P(\varepsilon_j), j \in \{-\nu, \dots, \nu\} \quad (59)$$

avec $\forall j \in \{-\nu, \dots, \nu\}$, $\varepsilon_j = 1 - e^{2i\sigma_j}$, $\sigma_{-j} = -\sigma_j$ et donc $\sigma_0 = 0$.

Notant :

$$L(\sigma) = \left(\frac{\varepsilon}{\ln(1 - \varepsilon)} \right)^2 = \left(\frac{\sin \sigma}{\sigma} \right)^2 e^{-2i\sigma} \quad (60)$$

le problème est donc de déterminer les $2\nu + 1$ coefficients modifiés α_k , $k \in \{0, 1, \dots, 2\nu\}$ à partir des $2\nu + 1$ relations $L(\sigma_j) = P(\varepsilon_j)$, $j \in \{-\nu, \dots, \nu\}$.

5.1.2 Méthode récurrente de calcul des coefficients modifiés

Le polynôme P cherché n'est autre que le polynôme interpolateur de la fonction L aux $2\nu + 1$ points ε_j . Toute formule d'interpolation classique pourrait être utilisée ; c'est celle de Lagrange qui est retenue. Les deux coefficients de plus haut degré de P peuvent ainsi être déterminés à l'aide de l'interpolation de Lagrange. Les autres pourraient l'être également mais ce serait fastidieux. On utilise donc une méthode récurrente.

Différences divisées de la fonction $(x \mapsto 1/x)$. Les différences divisées de la fonction $(x \mapsto \frac{1}{x})$ interviennent directement dans les calculs de la méthode de circularisation.

Considérons donc la fonction $g : x \mapsto 1/x$. La différence divisée d'ordre 0 de g (on utilise la notation simplifiée $[x_0]$ pour $g[x_0]$) est $[x_0] = 1/x_0$. La formule de récurrence (13) donne alors facilement $[x_0, x_1] = -1/x_0x_1$. Il semble ainsi que :

$$[x_0, \dots, x_k] = (-1)^k \prod_{i=0}^k \frac{1}{x_i} \quad (61)$$

En l'admettant pour les ordres inférieurs ou égaux à k , (13) donne :

$$[x_0, \dots, x_{k+1}] = (-1)^{k+1} \prod_{i=0}^{k+1} \frac{1}{x_i}$$

(61) est donc vérifiée, et (14) permet alors d'écrire :

$$\prod_{i=1}^k \frac{1}{x_i} = (-1)^{k-1} \sum_{i=1}^k \frac{1}{x_i} \cdot \Pi_i$$

avec $\Pi_i = \prod_{j=1, j \neq i}^k \frac{1}{x_i - x_j}$.

Les deux relations (15) et (16) donnent alors (les coefficients c_j sont arbitraires) :

$$\prod_{i=1}^k \frac{1}{x_i} = -c_k + \sum_{i=1}^k \frac{(-1)^{k-1} (1 + c_1 x_i + c_2 x_i^2 + \dots + c_{k-1} x_i^{k-1}) + c_k x_i^k}{x_i} \cdot \Pi_i \quad (62)$$

Calcul des deux coefficients de plus haut degré. Le coefficient $\alpha_{2\nu}$ est le coefficient de plus haut degré du polynôme interpolateur de Lagrange de la fonction L . Par définition des différences divisées, ce coefficient est donc la différence divisée d'ordre 2ν $L[\sigma_{-\nu}, \dots, \sigma_\nu]$. Or, $\alpha_{2\nu} \in \mathbb{R}$ et la fonction L a une formulation complexe. Il est donc nécessaire de calculer ce coefficient d'une autre façon. P ayant pour expression :

$$P(\varepsilon) = \sum_{j=-\nu}^{\nu} L(\sigma_j) \text{prod}_{m=-\nu, m \neq j}^{\nu} \frac{\varepsilon - \varepsilon_m}{\varepsilon_j - \varepsilon_m} \quad (63)$$

le coefficient de plus haut degré de P est donc :

$$\alpha_n = \sum_{j=-\nu}^{\nu} L(\sigma_j) \text{prod}_{m=-\nu, m \neq j}^{\nu} \frac{1}{\varepsilon_j - \varepsilon_m} \quad (64)$$

soit encore :

$$\begin{aligned} \alpha_n &= L(\sigma_0) \prod_{m=-\nu, m \neq 0}^{\nu} \frac{1}{\varepsilon_0 - \varepsilon_m} + \sum_{j=-\nu}^{-1} L(\sigma_j) \prod_{m=-\nu, m \neq j}^{\nu} \frac{1}{\varepsilon_j - \varepsilon_m} \\ &+ \sum_{j=1}^{\nu} L(\sigma_j) \prod_{m=-\nu, m \neq j}^{\nu} \frac{1}{\varepsilon_j - \varepsilon_m} \end{aligned} \quad (65)$$

avec $\varepsilon_0 = 0$ et :

$$\prod_{m=-\nu, m \neq 0}^{\nu} \frac{-1}{\varepsilon_m} = \prod_{m=1}^{\nu} \frac{1}{\varepsilon_{-m} \varepsilon_m} = \prod_{m=1}^{\nu} \frac{1}{u_m}$$

où $u_m = 4 \sin^2 \sigma_m$.

En opérant un changement d'indice, on aboutit finalement à l'expression :

$$\begin{aligned} \alpha_n &= L(\sigma_0) \prod_{k=1}^{\nu} \frac{1}{u_k} \\ &+ \sum_{k=1}^{\nu} \frac{L(\sigma_k) \varepsilon_{-k} (\varepsilon_{-k} - 1)^{\nu-1} - L(\sigma_{-k}) \varepsilon_k (\varepsilon_k - 1)^{\nu-1}}{(\varepsilon_k - \varepsilon_{-k}) u_k} \cdot \Pi_k \end{aligned} \quad (66)$$

où :

$$\Pi_k = \prod_{m=1, m \neq k}^{\nu} \frac{1}{u_k - u_m}$$

Utilisant alors la suite polynomiale S_n dont les propriétés sont exposées plus loin, on aboutit à :

$$\begin{aligned} \alpha_n &= L(\sigma_0) \prod_{k=1}^{\nu} \frac{1}{u_k} \\ &+ \sum_{k=1}^{\nu} \left[\frac{S_{2\nu-4}(u_k)}{u_k^{\nu-1}} + \frac{S_{2\nu-2}(u_k)}{u_k^{\nu}} \right] \cdot \frac{\Pi_k}{4\sigma_k^2} \end{aligned} \quad (67)$$

Or :

$$\frac{S_{2\nu-4}(u_k)}{u_k^{\nu-1}} + \frac{S_{2\nu-2}(u_k)}{u_k^{\nu}} = (-1)^{\nu} [1 + Q_{\nu-2}^*(u_k)]$$

où Q_p^* désigne une fonction polynôme de degré p dont le terme constant est nul.

D'après le calcul des différences divisées de la fonction ($x \mapsto 1/x$) et les propriétés de ces différences (les coefficients c_j sont arbitraires) :

$$\prod_{k=1}^{\nu} \frac{1}{u_k} = -c_{\nu} + \sum_{k=1}^{\nu} \frac{(-1)^{\nu-1} (1 + c_1 u_k + c_2 u_k^2 + \dots + c_{\nu-1} u_k^{\nu-1}) + c_{\nu} u_k^{\nu}}{u_k} \cdot \Pi_k$$

Prenant $c_{\nu} = c_{\nu-1} = 0$, on obtient alors :

$$\alpha_n = (-1)^{\nu-1} \sum_{k=1}^{\nu} \left[\frac{1 + c_1 u_k + \dots + c_{\nu-2} u_k^{\nu-2}}{u_k} - \frac{1 + Q_{\nu-2}^*(u_k)}{4\sigma_k^2} \right] \cdot \Pi_k \quad (68)$$

Les coefficients c_j , $j \in \{1, \dots, \nu-2\}$, étant arbitraires, on peut les choisir de façon à vérifier :

$$1 + c_1 u_k + \dots + c_{\nu-2} u_k^{\nu-2} = 1 + Q_{\nu-2}^*(u_k)$$

ce qui donne finalement :

$$\alpha_n = - \sum_{k=1}^{\nu} \frac{S_{2\nu-3}(u_k)}{u_k^{\nu-1}} \cdot \left[\frac{1}{u_k} - \frac{1}{4\sigma_k^2} \right] \cdot \Pi_k \quad (69)$$

où α_n est donc la $(\nu-1)$ ^{ème} différence divisée de la fonction :

$$u \mapsto \frac{S_{2\nu-3}(u)}{u^{\nu-1}} \cdot \left[\frac{1}{u} - \frac{1}{4\sigma^2} \right]$$

aux points u_1, \dots, u_{ν} .

Considérons maintenant le coefficient α_{n-1} . Par identification à partir de l'expression du polynôme interpolateur P (63) :

$$\alpha_{n-1} = - \sum_{k=-\nu}^{\nu} L(\sigma_k) \frac{\sum_{m=-\nu, m \neq k}^{\nu} \varepsilon_m}{\prod_{m=-\nu, m \neq k}^{\nu} (\varepsilon_k - \varepsilon_m)} \quad (70)$$

En isolant le terme correspondant à $k = 0$, on aboutit à l'expression :

$$\begin{aligned} \alpha_{n-1} &= \left(1 - \sum_{m=1}^{\nu} u_m\right) \alpha_n - \frac{L(\sigma_0)}{\prod_{m=1}^{\nu} u_m} \\ &+ \sum_{k=-\nu, k \neq 0}^{\nu} L(\sigma_k) \frac{\varepsilon_k - 1}{\prod_{m=-\nu, m \neq k}^{\nu} (\varepsilon_k - \varepsilon_m)} \end{aligned} \quad (71)$$

Après un calcul similaire à celui mené pour α_n , on aboutit à :

$$\alpha_{n-1} - \left(1 - \sum_{m=1}^{\nu} u_m\right) \alpha_n = - \sum_{k=1}^{\nu} \frac{S_{2\nu-5}(u_k)}{u_k^{\nu-2}} \cdot \left[\frac{1}{u_k} - \frac{1}{4\sigma_k^2}\right] \cdot \Pi_k \quad (72)$$

où $\alpha_{n-1} - (1 - \sum_{m=1}^{\nu} u_m) \alpha_n$ est donc la $(\nu - 1)^{\text{ème}}$ différence divisée de la fonction :

$$u \mapsto \frac{S_{2\nu-5}(u)}{u^{\nu-2}} \cdot \left[\frac{1}{u} - \frac{1}{4\sigma^2}\right]$$

aux points u_1, \dots, u_{ν} .

Il reste désormais à déterminer les autres coefficients modifiés du polynôme P .

Calcul des autres coefficients. On veut calculer les coefficients modifiés restants α_m , $m \in \{1, \dots, n - 2\}$. Pour ce faire, on suppose que l'on connaît le jeu des coefficients modifiés pour la méthode d'ordre $(n - 2)$: $\alpha_1^*, \dots, \alpha_{n-2}^*$, et que l'on connaît également les deux coefficients de plus haut degré pour la méthode d'ordre n calculés précédemment : α_{n-1} et α_n .

On dispose de P^* , polynôme interpolateur de L , de degré $n - 2$, aux points $\varepsilon_0, \varepsilon_{\pm 1}, \dots, \varepsilon_{\pm(\nu-1)}$ et on s'intéresse au polynôme interpolateur P de L , de degré n , aux points $\varepsilon_0, \varepsilon_{\pm 1}, \dots, \varepsilon_{\pm \nu}$. Un polynôme ayant pour zéros les points $0, \varepsilon_{\pm 1}, \dots, \varepsilon_{\pm(\nu-1)}$ est :

$$\varepsilon \cdot (\varepsilon^2 - u_1 \varepsilon + u_1) \dots (\varepsilon^2 - u_{\nu-1} \varepsilon + u_{\nu-1})$$

On peut donc écrire :

$$\forall \varepsilon \in \mathbb{R}, P(\varepsilon) = P^*(\varepsilon) + \varepsilon \cdot (\varepsilon^2 - u_1 \varepsilon + u_1) \dots (\varepsilon^2 - u_{\nu-1} \varepsilon + u_{\nu-1}) \cdot (a\varepsilon + b) \quad (73)$$

Pour déterminer a et b , on identifie les deux termes de plus haut degré des deux membres de l'équation (73). On obtient finalement la relation liant les deux polynômes interpolateurs de deux méthodes d'ordres pairs successifs :

$$P(\varepsilon) = P^*(\varepsilon) + \varepsilon \left[\alpha_n \varepsilon + \alpha_{n-1} + \alpha_n \sum_{m=1}^{\nu-1} u_m \right] \prod_{m=1}^{\nu-1} (\varepsilon^2 - u_m \varepsilon + u_m) \quad (74)$$

Il suffit alors de procéder à l'identification des coefficients de même degré des polynômes de chaque membre de l'équation (74).

Tous les outils nécessaires à la mise en oeuvre de la circularisation (suites polynomiales, différences divisées et méthodes de calcul) sont définis de façon récurrente, ce qui en facilite la programmation informatique. Les coefficients modifiés des méthodes d'ordres 2 et 4 permettant d'initialiser la récurrence sont fournis plus loin.

Suites polynomiales. La méthode de circularisation se base, pour les calculs des nouveaux coefficients des méthodes d'intégration numérique considérées, sur l'utilisation de deux suites de fonctions polynômes R_n et S_n , [2] et [14].

Ces suites polynomiales sont fonctions de la variable $u = 4 \sin^2 \sigma$, σ étant liée au pas de la méthode d'intégration. Ces deux suites sont définies par les relations de récurrence suivantes :

$$R_{n+1}(u) = u(R_n(u) - R_{n-1}(u)), R_0(u) = 2, R_1(u) = u \quad (75)$$

$$S_{n+1}(u) = u(S_n(u) - S_{n-1}(u)), S_0(u) = 0, S_1(u) = u \quad (76)$$

Les premiers polynômes de ces deux suites sont donnés dans le tableau (1) suivant.

indice n	$R_n(u)$	$S_n(u)$
0	2	0
1	u	u
2	$u^2 - 2u$	u^2
3	$u^3 - 3u^2$	$u^3 - u^2$
4	$u^4 - 4u^3 + 2u^2$	$u^4 - 2u^3$
5	$u^5 - 5u^4 + 5u^3$	$u^5 - 3u^4 + u^3$

TAB. 1 – Premières valeurs des suites polynomiales R_n et S_n .

On peut démontrer les propriétés suivantes par récurrence, propriétés utiles pour les calculs liés à la circularisation :

$$\forall n \in \mathbb{N},$$

$$\begin{aligned} 2 \cos[2n\sigma](2 \sin \sigma)^{2n} &= (-1)^n R_{2n}(u) \\ 2 \sin[(2n+1)\sigma](2 \sin \sigma)^{2n+1} &= (-1)^n R_{2n+1}(u) \\ \frac{\sin[(2n)\sigma]}{\cos \sigma} (2 \sin \sigma)^{2n+1} &= (-1)^{n+1} S_{2n}(u) \\ \frac{\cos[(2n+1)\sigma]}{\cos \sigma} (2 \sin \sigma)^{2n+2} &= (-1)^{n+1} S_{2n+1}(u) \end{aligned} \quad (77)$$

$$\begin{aligned}
R_{2n}(u) &= (-1)^n u^n (2 + Q_n^*(u)) \\
R_{2n+1}(u) &= (-1)^n u^{n+1} (2m + 1 + Q_n^*(u)) \\
S_{2n}(u) &= (-1)^{n-1} u^{n+1} (n + Q_{n-1}^*(u)) \\
S_{2n+1}(u) &= (-1)^n u^{n+1} (1 + Q_n^*(u))
\end{aligned} \tag{78}$$

où Q_p^* désigne un polynôme de degré p dont le terme de degré 0 est nul.

Enfin, si $\epsilon = 1 - e^{-2i\sigma}$ et $\epsilon_- = 1 - e^{2i\sigma}$, on a également les deux relations utiles :

$$\forall n \in \mathbb{N},$$

$$\epsilon^n - \epsilon_-^n = \frac{2i \sin[2\sigma] S_n(u)}{u} \tag{79}$$

$$\epsilon^n + \epsilon_-^n = R_n(u) \tag{80}$$

5.2 Initialisation de la récurrence : méthodes d'ordre 2 et 4

Pour initialiser la méthode récurrente de calcul des coefficients modifiés par la circularisation, il faut calculer ces coefficients pour la méthode de Cowell d'ordre 4 de pas h .

A titre d'exemple, commençons par la méthode d'ordre 2, où :

$$\Delta^2 x_0 = h^2 \left[\alpha_0 f_1 + \alpha_1 \Delta^1 f_{\frac{1}{2}} + \alpha_2 \Delta^2 f_0 \right]$$

On cherche à déterminer les coefficients α_0 , α_1 et α_2 pour lesquels la méthode intègre exactement les deux fonctions circulaires ($t \mapsto \cos \omega t$) et ($t \mapsto \sin \omega t$). D'après ce qui précède, il faut que la relation² :

$$\alpha_0 + \alpha_1(1 - e^{-2i\sigma}) + \alpha_2(1 - e^{-2i\sigma})^2 = \left(\frac{\sin \sigma}{\sigma} \right)^2 . e^{-2i\sigma}$$

soit vérifiée pour tout σ réel. Avec $\sigma = 0$, on trouve :

$$\alpha_0 = 1$$

En développant les exponentielles complexes, on aboutit à l'équation :

$$\left(\frac{\sin \sigma}{\sigma} \right)^2 = [\cos 2\sigma(1 + \alpha_1) - 4\alpha_2 \sin^2 \sigma - \alpha_1] + i[\sin 2\sigma + \alpha_1 \sin 2\sigma]$$

qui donne, si $\sigma \neq k\pi/2$, k entier :

$$\alpha_1 = -1 \quad \text{et} \quad \alpha_2 = \frac{1}{4} \left(\frac{1}{\sin^2 \sigma} - \frac{1}{\sigma^2} \right)$$

²si elle est vérifiée pour tout σ , elle le sera d'autant plus pour $\frac{\omega h}{2}$ et pour $-\frac{\omega h}{2} \dots$

En utilisant la relation :

$$\Delta^2 x_0 = h^2 \left[f_0 + \frac{1}{4} \left(\frac{1}{\sin^2 \sigma} - \frac{1}{\sigma^2} \right) \Delta^2 f_0 \right] \quad (81)$$

avec $\sigma = \frac{\omega h}{2}$ si $\frac{\omega h}{2} \neq k\pi/2$, k entier, on intègre donc exactement toute fonction de la forme :

$$t \mapsto a_0 + a_1 \cos \omega t + b_1 \sin \omega t$$

La récurrence ne permet pas de calculer les coefficients modifiés pour l'ordre 4. Il faut alors utiliser une méthode de calcul alternative, présentée dans [14], reposant également sur l'utilisation des suites polynomiales précédentes. Les calculs aboutissent aux résultats suivants :

$$\left\{ \begin{array}{l} \alpha_0 = 1 \\ \alpha_1 = -1 + u_1 u_2 \frac{\lambda(u_1) - \lambda(u_2)}{u_1 - u_2} \\ \alpha_2 = \frac{\lambda(u_2)u_1 - \lambda(u_1)u_2}{u_1 - u_2} + u_1 u_2 \frac{\lambda(u_2) - \lambda(u_1)}{u_1 - u_2} \\ \alpha_3 = (u_1 + u_2) \frac{\lambda(u_1) - \lambda(u_2)}{u_1 - u_2} \\ \alpha_4 = \frac{\lambda(u_2) - \lambda(u_1)}{u_1 - u_2} \end{array} \right.$$

où $u_i = 4 \sin^2 \sigma_i$ et :

$$\lambda(u) = \frac{1}{4} \left[\frac{1}{\sin^2 \sigma} - \frac{1}{\sigma^2} \right]$$

6 Discussion

En géodésie spatiale, le problème de l'évaluation objective de la précision d'une solution d'orbite (une éphéméride) est délicat, en particulier lorsque l'on cherche des signaux de l'ordre de 1 mm.

L'intégration numérique, qui affecte les calculs d'erreurs de troncature et d'erreurs d'arrondi, erreurs qui se propagent avec le temps dans la solution, implique donc sa propre part d'inexactitude. Elle est difficilement identifiable, puisque (i) l'on ne peut toujours refaire une estimation de paramètres géodésiques en changeant de méthode d'intégration (lourdeur), (ii) les méthodes d'estimation des paramètres ont tendance à absorber les petites erreurs contenues dans les variables du mouvement (subtilité).

En conséquence, il faut évaluer un algorithme d'intégration numérique en-dehors des processus d'estimation ; ceci se fait généralement en extrapolant le mouvement dans plusieurs cas, et en comparant ensuite les solutions obtenues (voir par exemple [1]). En particulier, la technique d'aller et retour peut être utilisée ; elle est d'un apport intéressant.

De plus, les tests en extrapolation doivent s'effectuer dans des conditions réalistes. D'une part, les variables (coordonnées rectangulaires ou éléments képlériens, ou autres) doivent être bien adaptées (cas des orbites circulaires, ou très excentriques, etc), d'autre part le modèle (les perturbations) doit être le plus réaliste possible.

⇒ Par exemple, estimer la précision d'un intégrateur, éventuellement par rapport à un autre, en utilisant des éléments képlériens classiques (singuliers en excentricité) et une perturbation en J_2 en traitant l'orbite du satellite CHAMP (à 400 km d'altitude, sur une orbite quasi-circulaire) ne signifie rien.

Ainsi, à titre d'exemple, lorsque l'on traite l'orbite de Starlette ($e = 0.02$, alt.= 800 km) d'une part avec l'algorithme de Bulirsch & Stoer en éléments képlériens (BS-eK, pas $H=1200$ sec, puis pas $H=900$ sec), d'autre part avec l'algorithme d'Adams-Moulton-Cowell en coordonnées rectangulaires (AMC-cR, pas=30 sec), puis avec la méthode d'Adams-Basforth-Moulton (ABM-eK, pas=60 sec), on trouve les différences suivantes :

Cas	Durée (j)	Modèle	Err. Along-T	Err. Across-T
BS/1200-AMC	14	Pot. 10,10	0.9	0.001
BS/1200-ABM	id.	id.	33.	0.1
id.	id.	Pot. 20,20	3.4	0.004
id.	id.	Pot. 50,50	197.	0.3
BS/900-AMC	id.	id.	195.	0.3
id.	720	Pot. 15,15+LS	3430.	3.
BS/1200 (A.-R.)	id.	id.	1330.	1.5

TAB. 2 – Différences sur orbite de Starlette avec (BS-eK, pas=1200 sec, puis pas=900 sec), (AMC-cR, pas=30 sec), (ABM-eK, pas=60 sec), sur 14 jours, puis sur 2 ans ; unités en mm.

Références

- [1] Balmino G., Barriot J.P. Numerical integration techniques revisited. *Manusc. Geodaetica*, 15 :1-10, 1990.
- [2] Bettis D.G. Numerical integration of products of Fourier and ordinary polynomials. *Num. Math.*, 14 :421-434, 1970.
- [3] Bettis D. G. Stabilization of finite difference methods of numerical integration. *Celestial Mechanics*, 2 :282-295, 1970.
- [4] Brouwer D. Solution of the problem of artificial satellite theory without drag. *The Astronomical Journal*, 64 :378-397, 1959.
- [5] Brouwer D., G. M. Clemence. *Methods of Celestial Mechanics*, Academic Press, New York and London, 1961.
- [6] Bulirsch R., Stoer J. Numerical treatment of ordinary differential equations by extrapolation methods. *Numerische Math.*, 8 :93-104, 1966.
- [7] Demailly J.-P. *Analyse numérique et équations différentielles*. Presses Universitaires de Grenoble. B.P.47X-38040 Grenoble Cedex, 1996.
- [8] Kaula W. M. *Theory of Satellite Geodesy*. Blaisdell Publishing Company, Waltham, Massachusetts. Toronto. London, 1966.
- [9] Kovalevsky J. Applications of computers to Celestial Mechanics. *Celestial Mechanics*, 45 :11-17, 1989
- [10] Kustaanheimo P., E. Stiefel. Perturbation theory of Kepler motion based on spinor regularization. *J. Reine Angew. Math.*, 218 :204-219, 1965.
- [11] Laudembach F. *Calcul différentiel et intégral*. Editions de l'école polytechnique, Ellipses, 2000.
- [12] McCarthy D.D. IERS Conventions. IERS technical note 21. Observatoire de Paris, 1996.
- [13] Oesterwinter C., Cohen C.H. New orbital elements for Moon and Planets. *Celest. Mechanics*, 5 :317-395, 1972.
- [14] Stiefel E., D. G. Bettis. Stabilization of Cowell's method. *Num. Math*, 13 :154-175, 1969.
- [15] Szebehely V. Limits of predictability of gravitational systems. *Celest. Mechanics*, 43 :139-145, 1988.
- [16] Zarrouati O. *Trajectoires spatiales*. Centre National d'Etudes Spatiales, Cepadues-Editions, 1987.