



HAL
open science

LES CONTRAINTES OPERATIONNELLES DANS LES MODELES DE PREVISION NUMERIQUE DU TEMPS : IMPLICATIONS SCIENTIFIQUES

Pierre Benard

► **To cite this version:**

Pierre Benard. LES CONTRAINTES OPERATIONNELLES DANS LES MODELES DE PREVISION NUMERIQUE DU TEMPS : IMPLICATIONS SCIENTIFIQUES. Océan, Atmosphère. Université Paul Sabatier (Toulouse), 2021. tel-03420313

HAL Id: tel-03420313

<https://hal.science/tel-03420313>

Submitted on 9 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Mémoire pour l'obtention du
diplôme d'habilitation à diriger des recherches**

présenté à

l'Université Paul Sabatier

par

Pierre Bénard

**LES CONTRAINTES OPERATIONNELLES DANS
LES MODELES DE PREVISION NUMERIQUE DU TEMPS :
IMPLICATIONS SCIENTIFIQUES**

soutenue le 6 Janvier 2021

devant le jury composé de :

Jean-Pierre Chaboureau (LA, Toulouse, Président)
Thomas Dubos (LMD, Palaiseau, Rapporteur)
René Laprise (UQAM, Montréal, Rapporteur)
Nigel Wood (UKMO, Exeter, Rapporteur)
Serge Gratton (INP, Toulouse, Directeur)
Christine Lac (METEO-FRANCE, Toulouse, Examinatrice)

Table des matières

1	Introduction	11
1.1	Cadre général des travaux de recherche : la prévision numérique déterministe	11
1.2	Esprit et plan du mémoire	12
2	Généralités et formalisme symbolique	15
2.1	Modèles d’atmosphère, systèmes d’équations, et discrétisation	15
2.2	Système linéaire, système linéaire-tangent d’un système non-linéaire	18
2.3	Matrice jacobienne, modes propres	20
2.4	Matrice d’amplification d’un système linéaire discretisé en espace et temps	20
2.5	Systèmes faiblement non-linéaire	21
2.6	Méthode spectrale	22
3	Stabilité des schémas numériques en PNT	25
3.1	Stabilité des schémas discrets dans les modèles numériques	25
3.2	Diagnostic ou analyse	26
3.3	Systèmes raides et schémas implicites	27
3.4	Principe des schémas semi-implicites	28
3.4.1	Schémas semi-implicites traditionnels (SI)	28
3.4.2	Schémas implicites itératifs centrés (ICI)	30
3.5	Géométrie horizontale et stabilité	31
3.5.1	Contexte	31
3.5.2	Instabilité liée à la linéarisation du facteur d’échelle	32
3.5.3	Epilogue (2004)	33
3.6	Stabilité du système d’Equations d’Euler en semi-implicite	34
3.6.1	Contexte historique	34
3.6.2	Premières tentatives en Equations d’Euler	37
3.6.3	Premières analyses et formalisme symbolique	37
3.6.4	Stabilité en SI des Equations d’Euler sans relief en 2-TL	38
3.6.5	Forces et faiblesses des analyses proposées	39
3.6.6	Problème de l’orographie	40
3.7	Mise en place d’un modèle de PN en équations d’Euler	42
3.8	Conclusion	42
4	PNT sur des ellipsoïdes de révolution	45
4.1	Introduction	45
4.2	Géométrie ellipsoïdale en atmosphère profonde	47
4.3	Géométrie ellipsoïdale en atmosphère mince	50
4.3.1	Dissocier erreurs géométriques et gravitationnelles	51
4.3.2	Planètes ‘réelles’ et planètes ‘forgées’	51
4.3.3	Vitesse angulaire centrifuge et de Coriolis	52
4.3.4	Un quartet de représentations de planètes en rotations	53
4.3.5	Estimation quantitative des erreurs	54
4.3.6	Conclusion	54
5	PNT sur la sphère avec des algorithmes locaux	57

6 Ondes de Rossby sur la sphère	61
6.1 Ondes de Rossby sur la sphère dans le système en “eau peu profonde”	61
6.2 Stabilité des ondes de Rossby sur la sphère dans le système ‘BVE’	64
7 Conclusion et perspectives	67
7.1 Perspectives	68
8 Appendice 1 : Incompressible fluid between two fixed rigid surfaces	71
8.1 Case of a plane flow	72
8.1.1 Evolution of vorticity on the plane	73
8.1.2 Conservation of potential vorticity on the plane	73
8.1.3 Particular case of a stationary flow on the plane with $\eta = \eta(x)$ and $(\partial/\partial y) = 0$. .	74
8.1.4 Standalone formulation on the plane	74
8.1.5 Particular case of uniform depth $\eta = \eta_0$ on the non-rotating plane	75
8.1.6 A personal digression	77
8.2 Flow on the rotating sphere with $\eta = \eta_0$	79
8.2.1 Comparison with Charney’s equation	79
8.3 Flow on the rotating sphere with $\eta = \eta(\lambda, \varphi)$	80
9 Appendice 2 : Gravitational potential in simple cases	83
9.1 Motivation	83
9.2 Generalities	84
9.3 Spherical geopotential approximation	85
9.4 Geopotential by two punctual masses on the polar axis	86
9.5 Geopotential by a lineic density along an equatorial circle	89
9.6 Discussion	92

Curriculum Vitae

Pierre Bénard

Né le 12 avril 1961 à Orsay (S.-et-O.)

Diplômes

1978 : Baccalauréat Série C (Lycée de Foix)

1984 : Ingénieur diplômé de l'Ecole Polytechnique

1986 : Ingénieur de la Météorologie

1989 : Docteur de l'UPS (Toulouse)

Parcours universitaire

1978-1981 : CPGE, Lycée Louis-le-Grand, Paris

1981-1984 : Ecole Polytechnique, Palaiseau

1984-1986 : Ecole Nationale de la Météorologie, Toulouse

1986-1989 : Doctorat, Université Paul Sabatier/CNRM, Toulouse

Parcours professionnel

1986-1988 : Centre National de Recherches Météorologiques, Toulouse

Formation Complémentaire par la Recherche (FCPLR).

1988-1991 : Service Météorologique Inter-Régional Nord, Lille

Bureau d'Etudes et de Développement (BED).

1992-2019 : Centre National de Recherches Météorologiques, Toulouse

Groupe de Modélisation et d'Assimilation pour la Prévision (GMAP)

Liste des travaux et publications

Publications dans des revues scientifiques de rang A

- Bénard, P., J.-L. Redelsperger, and J.-Ph. Lafore, 1992 : Nonhydrostatic Simulation of Frontogenesis in a Moist Atmosphere. Part I : General Description and Narrow Rainbands. *J. Atmos. Sci.*, **49**, 2200-2217.
- Bénard, P., J.-Ph. Lafore, and J.-L. Redelsperger, 1992 : Nonhydrostatic Simulation of Frontogenesis in a Moist Atmosphere. Part II : Moist Potential Vorticity Budget and Wide Rainbands. *J. Atmos. Sci.*, **49**, 2218-2235.
- Bubnová, R., G. Hello, P. Bénard, and J.-F. Geleyn, 1995 : Integration of the Fully Elastic Equations Cast in the Hydrostatic Pressure Terrain-Following Coordinate in the Framework of the ARPEGE/Aladin NWP System. *Mon. Wea. Rev.*, **123**, 515-535.
- Yessad, K. and P. Bénard, 1996 : Introduction of a local mapping factor in the spectral part of the Météo-France global variable mesh numerical forecast model. *Q. J. R. Meteorol. Soc.*, **122**, 1701-1719.
- Prtenjak, M. T., A. Marki and P. Bénard, 1998 : Refinement of the vertical diffusion scheme in the ARPEGE/ALADIN model. *Geofizika*, **vol 15 No.1**, December 1998, 27-51.
- Bénard, P., A. Marki, P.N. Neytchev, and M.T. Prtenjak, 2000 : Stabilization of Nonlinear Vertical Diffusion Schemes in the Context of NWP Models. *Mon. Wea. Rev.*, **128**, 1937-1948.
- Bénard, P., and K. Yessad, 2001 : On the Discretization of the Continuity Equation in Semi-Lagrangian Models with a Pressure-Type Vertical Coordinate. *Mon. Wea. Rev.*, **129**, 1746-1749.
- Gospodinov, I.G., V.G. Spiridonov, P. Bénard and J.-F. Geleyn, 2002 : An alternative semi-Lagrangian vertical trajectory scheme applied to an hydrostatic atmospheric model. *Q. J. R. Meteorol. Soc.*, **128**, 323-336.
- Bénard, P., 2003 : Stability of Semi-Implicit and Iterative Centred-Implicit Time Discretisations for various Equation Systems Used in NWP. *Mon. Wea. Rev.*, **131**, 2479-2491.
- Bénard, P., R. Laprise, J. Vivoda, P. Smolíková, 2004 : Stability of Leap-Frog Constant-Coefficients Semi-Implicit Schemes for the Fully Elastic System of Euler Equations. Flat-Terrain Case. *Mon. Wea. Rev.*, **132**, 1306-1318.
- Bénard, P., 2004 : On the Use of a Wider Class of Linear Systems for the Design of Constant-Coefficients Semi-Implicit Time-Schemes in NWP. *Mon. Wea. Rev.*, **132**, 1319-1324.
- Geleyn, J.-F., P. Bénard and R. Fournier, 2005 : A general-purpose extension of the Malkmus band-model average equivalent width to the case of the Voigt line profile. *Q. J. R. Meteorol. Soc.*, **131**, 2757 - 2768.
- Bénard, P., J. Mašek, and J. Vivoda, 2005 : Stability of leap-frog constant-coefficients semi-implicit schemes for the fully elastic system of Euler equations. Case with orography. *Mon. Wea. Rev.*, **133**, 1065-1075.
- Catry, B., J.-F. Geleyn, M. Tudor, P. Bénard and A. Trojaková, 2007 : Flux-conservative thermodynamic equations in a mass-weighted framework. *Tellus*, **59A**, pp.71-79.
- Váňa F., P. Bénard, J.-F. Geleyn, A. Simon and Y. Seity, 2008 : Semi-Lagrangian advection scheme with controlled damping : An alternative to nonlinear horizontal diffusion in a numerical weather prediction model. *Q. J. R. Meteorol. Soc.*, **134**, 523-537.

- Voitus F., P. Termonia et P. Bénard , 2009 : Well-posed Lateral Boundary Conditions for Spectral Semi-implicit semi-Lagrangian schemes : Tests in a one-dimensional model. *Mon. Wea. Rev.*, **137**, 315-330.
- P. Bénard, J. Vivoda, J. Mašek, P. Smolíková, K. Yessad, Ch. Smith, R. Brožková and J.-F. Geleyn, 2010 : Dynamical kernel of the Aladin-NH spectral limited-area model : revised formulation and sensitivity experiments *Q. J. R. Meteorol. Soc.* , **136** , 155-169.
- Yano, J.-I., Bénard, P. , Couvreur, F. and A. Lahellec, 2010 : NAM-SCA : Nonhydrostatic Anelastic Model under Segmentally-Constant Approximation *Mon. Wea. Rev.* , **138** , 1957-1974.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, Ch. and V. Masson , 2011 : The AROME-France convective scale operational model *Mon. Wea. Rev.* , **139** , 976-991.
- Bénard, P., 2014 : An oblate-spheroid geopotential approximation for global meteorology. *Q. J. R. Meteorol. Soc.*, **140**, 170-184.
- Bénard, P., 2015 : An assessment of global forecast errors due to the spherical geopotential approximation in the shallow water case. *Q. J. R. Meteorol. Soc.*, **141**, 195-206.
- Caluwaerts, S., Degrauwe, D., Termonia, P., Voitus, F., Bénard, P. and Geleyn, J.-F. 2015 : Importance of temporal symmetry in spatial discretization for geostrophic adjustment in semi-implicit Z-grid schemes. *Q. J. R. Meteorol. Soc.*, **141**, 128-138.
- Colavolpe, Ch., Voitus, F., and Bénard, P., 2017 : RK-IMEX HEVI schemes for fully compressible atmospheric models with advection : analyses and numerical testing. *Q. J. R. Meteorol. Soc.*, **143**, 1336-1350.
- Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derkova, M., El Khatib, R., Hamdi, R., Masek, J., Pottier, P., Pristov, N., Seity, Y., Smolikova, P., Spaniel, O., Tudor, M., Wang, Y., Wittmann, C., and Joly, A., 2018 : The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1. *Geosci. Mod. Dev.*, **11**, 257-281.
- Bénard, P., 2019a : Numerical investigation of Rossby waves for nonlinear Shallow-Water equations on the sphere. *Q. J. R. Meteorol. Soc.* , **145**, 1461–1473.
- Bénard, P., and Ginton, M., 2019 : Circumventing the pole problem of reduced lat-lon grids with local schemes. Part I : analysis and model formulation. *Q. J. R. Meteorol. Soc.* , **145**, 1377–1391.
- Ginton, M., and Bénard, P., 2019 : Circumventing the pole problem of reduced lat-lon grids with local schemes. Part II : validation experiments. *Q. J. R. Meteorol. Soc.* , **145**, 1392–1405.
- Voitus, F., Bénard, P., Kühnlein, C. and Wedi, N., 2019 : A semi-implicit unified dynamical system cast in a mass-based coordinate : model formulation and numerical testing. *Q. J. R. Meteorol. Soc.* , **145**, 3387– 3408
- Bénard, P., 2020 : Stability of Rossby-Haurwitz waves. *Q. J. R. Meteorol. Soc.*, **146**, 613– 628

Conférences avec comité de lecture et proceedings (en premier auteur seulement)

(N.B. : à partir de 2010, seules les conférences invitées sont indiquées)

- Bénard, P., J.-L. Redelsperger, and Ph. Veyre, 1987 : Nonhydrostatic Numerical simulation of frontogenesis and frontal rainbands with a fine-scale convection model. Third Conference on Mesoscale Processes, AMS, 21-26 August 1987, Vancouver, B.C., Canada.
- Bénard , P., 1996 : Stabilization of ARPEGE/IFS semi-Lagrangian scheme for diabatic variable resolution purpose, and other pre-operational problems. ECMWF Workshop on semi-Lagrangian methods, 6-8 November 1995, Reading, UK.
- Bénard , P., and Bubnová, R., 1996 : Non-hydrostatic dynamics in Aladin model. First SRNWP Workshop on Non-Hydrostatic Modelling, 11-13 March 1996, Bad Orb, Germany.
- Bénard , P., and Thépaut, J.-N. , 1996 : Some aspects related with efficiency improvement in the ARPEGE operational NWP system. 11th AMS Conference on Numerical Weather Prediction, 19-23 August 1996, Norfolk, VA, USA.

- Bénard, P., 2001 : Study of the robustness of partially implicit temporal schemes in the Euler system. Third Workshop on the Solution of Partial Differential Equations on the Sphere, May 15-18 2001, Montréal, Québec, Canada.
- Bénard, P., P. Smolíková, J. Mašek and J. Vivoda, 2003 : A potentially dangerous albeit unnecessary restriction in classical semi-implicit schemes design strategy. Fifth international SRNWP workshop on non-hydrostatic modelling, 27-29 October 2003, Bad Orb, Germany.
- Bénard, P., 2004 : ALADIN-NH/AROME dynamical core : status and possible extension to IFS. Proceedings of ECWMF Seminar on "Recent developments in numerical methods for atmosphere and ocean modelling", 6-10 Sept 2004, Reading, U.K.
- Bénard, P., 2010 : Non-hydrostatic modelling with AROME and some properties of quasi-elastic systems, ECMWF Workshop on Non-hydrostatic modelling, 8-10 Nov 2010, Reading, U.K.
- P. Bénard, P., 2013 A non-hydrostatic SI dynamical core : current-state, limitations and perspectives. ECMWF Workshop on Scalability, 02-05 Sept. 2013, Reading, UK.
- Bénard, P., 2013 : Météorologie et Climatologie sur des planètes sphéroïdales en rotation : approche cohérente et évaluation de l'impact pour la Terre. XXVIe Entretiens Jacques Cartier, 27-28 Nov. 2013, Lyon, France.
- Bénard, P., 2013 : Earth's flattening : which impact for meteorology and climatology ?. Conference WWOSC2014, 16-21 August 2014, Montreal, Canada.
- Bénard, P., 2015 : High-order scalable numerical schemes on the sphere with unstaggered reduced Lat-Lon grids. *Workshop "Mathematics of the Weather"*, 21 – 24 September 2015, Erquy, France
- Bénard, P., 2017 : Work on Dynamical kernels at météo-France *Workshop "Mathematics of the Weather"*, 2 – 5 September 2017, Erquy, France
- Bénard, P., 2017 : Circumventing the pole problem of spherical coordinates with local schemes ; some applications *Workshop "Mathematics of the Weather"*, 2 – 5 September 2017, Erquy, France
- Bénard, P., 2017 : Pole problem of spherical coordinates *Workshop "Partial Differential Equations on the Sphere"*, 03 – 07 April 2017, Paris, France

Activités d'encadrement (thèses)

- Radmila Bubnová, 1993-1995, encadrée en tant que co-directeur de Thèse. PhD/thèse en co-tutelle Université Charles (Prague) / Université Paul Sabatier (Toulouse) Diplôme obtenu le 25 avril 1995.
- Ilian Gospodinov, 1997-2000, encadré en tant que co-directeur de Thèse. Thèse Université Paul Sabatier (Toulouse) Diplôme obtenu le 7 décembre 2000.
- Filip Váňá, 2000-2003, encadré en tant que co-directeur de Thèse. PhD/thèse en co-tutelle Université Charles (Prague) / Université Paul Sabatier (Toulouse) Diplôme obtenu le 30 septembre 2003.
- Fabrice Voitus, 2004-2008, encadré en tant que co-directeur de Thèse. Thèse Université Paul Sabatier (Toulouse) Diplôme obtenu le 3 décembre 2008.

Activités d'enseignement

Module (20 h) "Les bases de la Prévision Numérique" à l'Institut Hydrométéorologique de Formation et de Recherche (Oran, Algérie) de 2006 à 2009

Activités diverses

Associate Editor au "Quarterly Journal of the Royal Meteorological Society" de 2005 à 2010.

Contribution régulière aux comités de lecture de revues météorologiques, principalement :

- Monthly Weather Review,
- Quarterly Journal of the Royal Meteorological Society,
- Tellus (Series A).

Chapitre 1

Introduction

1.1 Cadre général des travaux de recherche : la prévision numérique déterministe

Le fil conducteur choisi pour ce mémoire de HDR portera sur les implications scientifiques sous-tendues par les contraintes liées à la Prévision Numérique du Temps (PNT) dans ses aspects numériques et dynamiques, implications sur lesquelles j'ai été amené à me pencher au fil de l'évolution de cette discipline, au cours des trois décennies passées.

Il sera essentiellement question ici de la PNT telle qu'elle est abordée en France, au sein de Météo-France, avec donc les spécificités liées aux choix stratégiques propres à cet Etablissement Public. D'autres Centres Nationaux de Météorologie (CNM) ont fait des choix stratégiques parfois très différents pour bâtir leurs outils de PNT. La PNT (Coiffier, 2011) est l'ensemble des activités permettant, à partir de données observées, de conduire à la production opérationnelle de prévisions météorologiques à court ou moyen terme (à moins de 10 jours d'échéance, pour fixer les idées). Au sens strict, la collecte des données et la diffusion des produits de prévision n'entrent pas dans le champ de la PNT. L'activité se divise ainsi en deux grandes branches : (i) d'une part l'assimilation de données, visant à l'élaboration d'une image du ou des états initiaux de l'atmosphère à un instant donné, propre à être exploitée numériquement ; (ii) d'autre part la prévision numérique en elle-même qui à partir de l'état initial et via une intégration temporelle des équations d'évolution reflétant d'hypothétiques "lois d'évolution" de l'atmosphère réelle, fournit une séquence des états futurs de l'atmosphère modélisée jusqu'à l'échéance souhaitée. Il sera uniquement question ici de travaux relatifs à cette seconde branche, l'intégration temporelle.

Il est à noter qu'une réflexion d'importance grandissante tente, au sein des CNM, de comprendre quelle attitude adopter face à l'aspect chaotique ou non-déterministe de l'évolution de l'atmosphère aux échelles fines et échéances longues visées par la prévision moderne. Plusieurs voies d'approche face à ce problème sont possibles, on les regroupe sous le nom de "prévision d'ensemble" (e.g. Bouttier et Buizza, 2018 et les références citées). Cette branche considère des perturbations de l'état initial et/ou du modèle de prévision pour effectuer de multiples prévisions voisines, le but étant d'estimer la dispersion de ces prévisions d'une part et des probabilités d'occurrence d'autre part. Aucun de ces aspects ne sont considérés dans ce mémoire, l'évolution de l'atmosphère étant donc supposée déterministe à partir du meilleur état initial et du meilleur modèle disponible. Ceci peut limiter l'échéance de validité, et l'on suppose donc implicitement que l'on se trouve à l'intérieur de cet intervalle.

Afin de lever une possible ambiguïté, précisons enfin que ce que les Services Météorologiques opérationnels appellent une "prévision" en tant que produit fini délivrable aux usagers finaux, est en fait la combinaison de prévisions "brutes" directement issues de modèles numériques, et d'une expertise humaine venant préciser, critiquer, adapter ou affiner cette prévision brute. Ce dernier aspect est nécessaire voire primordial, ne serait-ce que parce que les prévisions numériques brutes disponibles à un instant donné pour une date donnée peuvent être multiples (issues de plusieurs sources) et parfois divergentes sur certains aspects. L'expertise humaine, apportée par les prévisionnistes, consiste donc à prendre l'ensemble de décisions qui à partir des résultats numériques bruts, vont permettre d'aboutir à une représentation la plus pertinente possible de l'état prévu, et de traduire cette image en un langage adapté au public auquel elle est destinée. Etant donné le contexte spécifique de ce mémoire, par souci de concision, nous parlerons de "prévision" pour désigner en fait ce qu'il conviendrait d'appeler "prévision numérique brute" dans un

cadre plus général.

Comme mentionné au premier paragraphe, le cadre de l'exposé est principalement focalisé sur les aspects dynamiques et numériques de la prévision en elle-même. La référence aux "aspects dynamiques" signifie qu'on se penche surtout sur la formulation et la résolution des équations régissant l'évolution de l'atmosphère considérée comme un fluide parfait, c'est à dire en ignorant les phénomènes physiques tels que les changements de phase de l'eau atmosphérique, les processus visqueux ou turbulents, les processus radiatifs des ondes électromagnétiques, etc. La référence aux "aspects numériques" concerne le détail du choix des méthodes algébriques retenues pour résoudre numériquement les équations d'évolution de l'atmosphère. L'importance de ces "détails" est extrême : l'élaboration d'une prévision météorologique nécessite de répéter automatiquement des milliards d'opérations algébriques, de sorte que des différences minimales dans les algorithmes utilisés peuvent conduire à d'énormes différences dans la qualité de la prévision. Ces aspects numériques sont étudiés par une branche des mathématiques appliquées dénommée "analyse numérique".

Bien que n'apparaissant pas explicitement dans le vocable, une qualification indissociable du concept de "Prévision Numérique du Temps" est celle de son caractère "opérationnel". C'est ce qui fait la spécificité de la PNT par rapport à d'autres disciplines voisines telles que par exemple la modélisation climatique ou la modélisation des processus atmosphériques. En PNT, l'opérationnalité est au cœur du problème, et se décline à son tour en de multiples contraintes, dont les principales sont :

- disponibilité : le processus de prévision doit présenter de bonnes garanties d'être mené à son terme chaque jour. Ceci suppose donc des algorithmes stables tout au long de la chaîne de prévision.
- ponctualité : le processus doit délivrer chaque jour ses produits de prévision à l'heure pré-définie. Ceci suppose des algorithmes dont l'efficacité est indépendante de la situation météorologique du jour.
- qualité : la prévision doit être de la meilleure qualité possible pour le coût imparti. Ceci suppose notamment un processus d'estimation de la qualité (a posteriori).

Le dernier point de cette liste suppose également des travaux scientifiques indiquant de façon convaincante que l'utilisation d'autres méthodes de prévision conduirait à une moindre qualité. D'où l'imbrication très forte entre les enjeux opérationnels et les enjeux scientifiques. En fait, la justification des choix effectués et des outils utilisés en PNT est basée sur des études scientifiques en amont, qui permettent de définir par consensus dans chaque CNM la stratégie à adopter pour bâtir et faire évoluer son système de PNT. La diversité des techniques utilisées en PNT dans les différents CNM de par le monde indique toutefois qu'il est difficile de démontrer scientifiquement la supériorité irréfutable d'une stratégie donnée. De plus une stratégie 'optimale' à une époque donnée peut devenir inadaptée ultérieurement, du fait de l'évolution des technologies utilisées pour les machines de calcul scientifique. Tout ceci concourt à un balancement plus ou moins permanent entre les études scientifiques en amont et le développement des applications en aval, afin que l'outil de PNT soit en permanence adapté de façon optimale aux conditions présentes. Les travaux rapportés dans ce mémoire se rattachent en grande majorité à ces études scientifiques en amont. Il s'agit là clairement de recherche appliquée, et non pas de recherche fondamentale.

1.2 Esprit et plan du mémoire

J'ai renoncé à trouver un fil conducteur unique et un titre vraiment spécifique, car les travaux décrits ici s'articulent selon trois grands thèmes qui se relient difficilement entre eux : la stabilité des schémas numériques en PNT, les apports potentiels d'une géométrie non-sphérique pour la PNT globale, et enfin les problèmes liés aux pôles pour la PNT globale en coordonnées polaires (i.e. coordonnées sphériques ou sphéroïdales).

L'esprit adopté pour ce mémoire est plutôt celui d'une mise en perspective que celui d'un exposé scientifique. Dans la mise en perspective choisie ici, les travaux sont présentés sous l'angle de leur genèse historique et leur maturation dans un contexte international, afin de montrer comment les avancées engrangées ont été le fruit d'une démarche progressive et d'un échange constant avec les conceptions prévalant dans les voies de recherche voisines menées par d'autres équipes à l'étranger.

Le mémoire est articulé comme suit. Quelques concepts de base et notations symboliques sont introduites dans le chapitre 2. Le chapitre 3 expose en détail les travaux sur la stabilité des schémas numériques en PNT qui ont occupé la majeure partie de mon activité. Le chapitre 4 aborde une source d'amélioration possible de la qualité de la PNT dans le domaine de la dynamique, via la prise en compte des effets

non-sphérique dans la formulation des équations gouvernant l'atmosphère météorologique sur le globe. Le chapitre 5 traite d'une stratégie possible pour pouvoir utiliser des algorithmes locaux sur le globe en remplacement de la représentation spectrale en harmonique sphérique (non-locale) utilisée aujourd'hui. Ce recours à des algorithmes locaux est fortement dicté par l'évolution récente et à venir de l'architecture des machines de calcul scientifique. Le chapitre 6 aborde une étude numérique des ondes de Rossby-Haurwitz sur la sphère pour le système en eau peu profonde.

Chapitre 2

Généralités et formalisme symbolique

Dans ce chapitre sont introduits quelques concepts généraux et un formalisme symbolique permettant clarifier et simplifier l'exposé des chapitres suivants.

2.1 Modèles d'atmosphère, systèmes d'équations, et discrétisation

Comme dans beaucoup d'autres domaines de la physique, l'évolution atmosphérique est bien trop complexe pour être mise en équation dans son entière exactitude. Sous peine de s'exposer à un inévitable embourbement, même si on se limite d'emblée à la mécanique Classique Newtonienne dans tous les cas, on est bien obligé de négliger le détail de certains phénomènes qui se produisent dans la réalité, mais que l'on n'est pas en mesure de prendre en compte au moment où l'on conçoit et développe un système de prévision. Les exemples qui viennent à l'esprit immédiatement sont les phénomènes d'échelle insignifiante, sans effet macroscopique notable, bien qu'en toute rigueur ils influent sur l'évolution et devraient donc idéalement être représentés (le célèbre battement d'aile d'un papillon ...). Néanmoins, il n'est pas besoin de se limiter aux phénomènes microscopiques pour trouver des effets plus ou moins insignifiants, le domaine macroscopique possède lui aussi de tels effets, dont la prise en compte peut être sujette à discussion : doit-on tenir compte de la non-sphéricité de la Terre, de l'affaiblissement de l'attraction terrestre avec l'altitude, ou encore des effets de marées lunaires sur l'atmosphère ? Où faut-il s'arrêter ? Quant aux effets qui ne sont ni microscopiques ni macroscopiques, mais difficiles à décrire mathématiquement, comment doit-on les traiter ?

Finalement, la question qui se pose est celle de savoir où est la frontière entre ce qu'il faut décrire et ce qui peut être négligé, pour obtenir des résultats optimaux. Précisons tout de suite que la réponse n'est pas simple, et loin de faire l'unanimité. Le fruit de cette toute première étape dans la conception d'un système de prévision est le choix d'un "modèle d'atmosphère", c'est à dire finalement le choix de tout un ensemble d'hypothèses simplificatrices, si possible cohérentes entre elles, qui vont dans une étape ultérieure permettre la mathématisation du modèle. A ce stade, il n'est donc pas encore question d'équations à proprement parler.

Par exemple, pour l'élaboration de la première prévision numérique de l'histoire, en 1950, Charney et ses collaborateurs s'étaient proposés de considérer l'atmosphère comme une "couche infiniment mince de fluide parfait incompressible, sur une planète sphérique en rotation, l'écoulement de ce fluide étant de plus supposé non-divergent à chaque instant, et indépendant de l'altitude (écoulement bidimensionnel)". C'était là un "modèle d'atmosphère", arbitraire, sur lequel il était obligatoire de se mettre d'accord avant de continuer à avancer, mais à l'évidence ce choix purement "phénoménologique" ne débouche pas encore en soi sur un système d'équations concret à résoudre en pratique. L'histoire de la genèse de cette première prévision (p. ex. Lynch 2004, p. 36) montre que l'obtention d'un consensus sur ce point avait nécessité un véritable travail critique de la part des protagonistes, d'autres "modèles d'atmosphère" légèrement plus complexes ayant aussi été envisagés (en autorisant par exemple des écoulements horizontalement divergents), mais abandonnés dès avant le passage à l'étape ultérieure.

Une fois effectué le choix d'un "modèle d'atmosphère", il est ensuite nécessaire d'introduire les outils permettant la mathématisation –la mise en équation– du problème à traiter, c'est à dire ici essentiellement un système de coordonnées, et un ensemble de variables d'état. Ce sont ces variables d'état, champs

scalaires ou vectoriels fonctions de l'espace et du temps (exprimés dans un système de coordonnées), qui vont se prêter à une forme mathématisée des lois d'évolution, appelée "système d'équation continues". Il faut bien voir que la forme particulière que va prendre le système d'équation représentant le modèle d'atmosphère dépend beaucoup du choix des coordonnées et des variables pronostiques, même si pour un même modèle d'atmosphère, il y aura une parenté formelle plus ou moins grande entre les systèmes d'équations déclinés avec divers choix de coordonnées et/ou de variables d'état. Notamment, il est assez logique de penser que quelles que soient les variables pronostiques choisies, leur nombre devrait être le même, car ce nombre représente *in fine* le nombre de degrés de liberté autorisé par le modèle d'atmosphère retenu.

Pour illustrer cette étape de mathématisation et introduire des notations symboliques utiles pour la suite, nous supposerons que les lois que nous avons choisies pour régir l'évolution continue de l'atmosphère peuvent s'écrire, après choix d'un système de coordonnées et de variables, sous la forme générale :

$$\frac{\partial \mathcal{X}(\mathbf{r}, t)}{\partial t} = \mathcal{M}[\mathcal{X}(\mathbf{r}, t)]. \quad (2.1)$$

Ici, (\mathbf{r}, t) représentent les coordonnées d'espace et de temps et \mathcal{X} représente l'ensemble minimal des variables d'état de l'atmosphère qui sont nécessaires pour pouvoir calculer de façon univoque l'évolution selon le système choisi, c'est à dire les variables d'état nécessaires pour fermer le système. Il s'agit donc d'un ensemble fini de champs continus bi- ou tri-dimensionnels, dont (en météorologie) les plus usuels sont les variables thermodynamiques au sens large (tels que par exemple température absolue, pression, vent) mais qui peuvent être aussi plus nombreux pour des systèmes plus sophistiqués (énergie cinétique turbulente, température sous la surface du sol, contenu spécifique de divers espèces d'hydrométéores, de vapeur d'eau ou d'aérosols, etc.). Enfin, dans cette équation, \mathcal{M} représente un opérateur non-linéaire mais purement spatial, agissant sur l'état de l'atmosphère $\mathcal{X}(\mathbf{r}, t)$ à un instant donné t .

Il est à noter que certains modèles d'atmosphère dits 'filtrés', ne se prêtent pas aisément à une mathématisation directement sous cette forme, du fait d'une équation diagnostique liant de manière non triviale deux ou plusieurs champs spatiaux à un instant donné. Prenons par exemple le système du tourbillon barotrope sur la sphère. La manière la plus naturelle de mathématiser ce système est la suivante

$$\begin{aligned} \frac{\partial \zeta}{\partial t} &= f(\zeta, \psi) \\ \zeta &= \nabla^2 \psi \end{aligned} \quad (2.2)$$

Toutefois, on peut remarquer que sur la sphère, l'opérateur ∇^2 est inversible sous réserve de supposer une jauge (arbitraire) pour ψ (par exemple une intégrale nulle). Dans ce cas, on peut simplement écrire $\psi = \nabla^{-2}\zeta$, et le système mathématisé devient

$$\begin{aligned} \frac{\partial \zeta}{\partial t} &= f(\zeta, \nabla^{-2}\zeta) \\ &= g(\zeta) \end{aligned} \quad (2.3)$$

où g est un opérateur spatial agissant sur ζ uniquement, ce qui ramène bien à la forme supposée en (2.1). Cette forme est atteignable pour les systèmes filtrés (ce qui inclut tous les systèmes anélastiques) sous réserve que l'équation diagnostique possède une solution unique pour le champ qui sera éliminé grâce à cette équation, ce qui en principe est le cas puisque l'évolution sous-tendue par le système entier reste par hypothèse dans le cadre déterministe.

A partir de maintenant, on se place définitivement dans la forme canonique supposée ci-dessus en (2.1). Si, par exemple le système choisi est monodimensionnel en x , et peut être fermé avec les champs de vent u (selon x), température absolue T et pression p , alors \mathcal{X} se réduira à l'ensemble de ces trois champs tridimensionnels, $\mathcal{X} = (u, T, p)$, étant entendu que ces trois champs sont des champs scalaires continus, définis sur l'ensemble du domaine d'intégration spatio-temporel. Chacun de ces trois champs est appelé variable pronostique, tandis que le vecteur colonne \mathcal{X} qui les regroupe toutes pourrait être désigné sous le nom de "variable globale d'état" pour le système d'équations considéré. Dans notre exemple, d'un point de vue formel, la variable globale d'état est donc un vecteur à trois composantes (chacune d'elle étant une fonction de l'espace et du temps), et l'opérateur \mathcal{M} est un opérateur vectoriel que l'on peut décomposer

en $(\mathcal{M}_u, \mathcal{M}_T, \mathcal{M}_p)$. On pourrait dans notre cas, noter aussi $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3)$ et $(\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3)$, et, dans un formalisme légèrement plus général :

$$\begin{aligned} \frac{\partial \mathcal{X}_1(\mathbf{r}, t)}{\partial t} &= \mathcal{M}_1[\mathcal{X}_1(\mathbf{r}, t), \dots, \mathcal{X}_N(\mathbf{r}, t)] \\ \frac{\partial \mathcal{X}_2(\mathbf{r}, t)}{\partial t} &= \mathcal{M}_2[\mathcal{X}_1(\mathbf{r}, t), \dots, \mathcal{X}_N(\mathbf{r}, t)] \\ &\dots \\ \frac{\partial \mathcal{X}_N(\mathbf{r}, t)}{\partial t} &= \mathcal{M}_N[\mathcal{X}_1(\mathbf{r}, t), \dots, \mathcal{X}_N(\mathbf{r}, t)], \end{aligned} \quad (2.4)$$

avec dans notre exemple particulier, $N = 3$. Si un terme à droite dans la i -ème équation met en jeu une dépendance à $\mathcal{X}_j(\mathbf{r}, t)$ avec $i \neq j$, on dit alors qu'il s'agit d'un terme de couplage. Ce sont ces termes qui empêchent de résoudre séparément et indépendamment les N équations ci-dessus. Nous n'avons pas plusieurs équations à résoudre, mais plusieurs équations interdépendantes : il s'agit donc d'un système dynamique multivarié couplé.

Pour illustrer simplement la différence entre modèle d'atmosphère et système d'équation, remarquons que si on choisit la température potentielle θ à la place de la température absolue T comme variable pronostique (pour le même "modèle d'atmosphère"), le vecteur \mathcal{X} et l'opérateur \mathcal{M} ne seront plus définis mathématiquement sous la même forme. Nous aurons une variante légèrement différente du système, d'un point de vue formel, alors que la réalité décrite sera exactement la même.

Le but de la prévision météorologique est donc de calculer le vecteur \mathcal{X} à un instant $t > 0$, supposant connue sa valeur à l'instant initial $t = 0$, en résolvant le système (2.1). Néanmoins, le système continu défini par (2.1) n'est pas soluble algébriquement dans le cas d'un état initial quelconque. Il existe des états initiaux où le système peut être résolu algébriquement, mais ces états sont trop particuliers pour être d'un intérêt quelconque pour la prévision en pratique : il s'agit le plus souvent d'états pour lesquels l'évolution est nulle (états stationnaires), ou d'états infiniment proches de ces derniers. L'idée de la modélisation numérique est de transformer le système continu (2.1) en un système discrétisé en temps et en espace, c'est à dire que parmi l'infinité de degrés de liberté qui caractérise $\mathcal{X}(\mathbf{r}, t)$, on n'en retiendra qu'un nombre fini, choisis le plus judicieusement possible.

Puisque en prévision numérique on résout les équations d'évolution avec une machine de calcul et que celle-ci possède une capacité de mémoire finie ainsi qu'une vitesse de calcul finie, ce sont ces contraintes matérielles vont imposer le nombre maximal de degrés de liberté que nous pourrons manipuler dans le temps imparti pour la prévision. Il apparaît ainsi que le fait de représenter l'atmosphère terrestre et son évolution sur une échéance donnée par un nombre fini de degrés de libertés (c'est à dire de valeurs numériques) va nécessairement entraîner l'apparition d'une résolution spatiale et temporelle pour la description adoptée, c'est à dire des échelles d'espace et de temps au-dessous desquelles les processus et les phénomènes ne sont pas descriptibles explicitement. Le processus de discrétisation, en imposant une certaine résolution spatiale et temporelle va donc lui aussi influencer à son tour sur la frontière entre les phénomènes qui doivent être retenus dans le modèle d'atmosphère, ceux qui peuvent être traités de façon approchée, et ceux qui doivent être négligés. Tout le processus de conception, déjà fort complexe, se reboucle sur lui-même par ce biais, et il n'est pas surprenant qu'il n'existe pas deux systèmes de prévisions identiques dans le monde.

Lorsqu'on applique une discrétisation spatiale au système (2.1), chacune des fonctions continues qui correspond à l'une des variables pronostiques \mathcal{X}_i de \mathcal{X} se trouve représentée par un ensemble fini P_i de valeurs numériques, et la représentation spatiale devient discrète. En conséquence, ce qui permet de décrire l'atmosphère à un instant donné devient un "vecteur d'état" formé de $P = P_1 + \dots + P_N$ valeurs numériques. Les opérateurs continus qui constituaient \mathcal{M} doivent donc maintenant être représentés par des opérateurs agissant non plus sur des champs continus mais sur les valeurs numériques discrètes du vecteur d'état, cette transformation des opérateurs étant appelée 'discrétisation spatiale'.

$$\mathcal{M}[\mathcal{X}(t)] \xrightarrow{\text{sp. discr.}} \mathbf{M}[\mathbf{X}(t)]. \quad (2.5)$$

De même, lorsqu'on applique une discrétisation temporelle au système, la dimension temps, continue à l'origine, se retrouve représentée par une suite d'états successifs discrets t_i , en nombre finis pour un intervalle de temps donné. l'opérateur temporel $\partial/\partial t$, est représenté par un opérateur discret $(\delta/\delta t)$ (par

exemple une différence finie d'ordre 2 entre des état t_i voisins). Symboliquement, nous noterons donc le nouveau système discrétisé par :

$$\frac{\delta \mathbf{X}(t_i)}{\delta t} = \mathbf{M}[\mathbf{X}(t_i)]. \quad (2.6)$$

En pratique chacun des champs individuels (u, T, p) dans l'exemple ci-dessus, devient un vecteur colonne lui aussi (de dimension P_u, P_T et P_p respectivement), et c'est l'agrégation en colonne de tous ces vecteurs colonne individuels qui constitue le 'vecteur d'état' $\mathbf{X} = (X_1, \dots, X_P)$, comprenant la liste de tous les degrés de liberté spatiaux de toutes les variables, avec donc $P = P_u + P_T + P_p$ dans notre exemple. L'opérateur spatial discret est représenté symboliquement par \mathbf{M} . On peut donc écrire, de façon analogue à ce qui a été fait en (2.4) :

$$\begin{aligned} \frac{\delta X_1(t_i)}{\delta t} &= \mathbf{M}_1[X_1(t_i), \dots, X_P(t_i)] \\ &\dots \\ \frac{\delta X_P(t_i)}{\delta t} &= \mathbf{M}_P[X_1(t_i), \dots, X_P(t_i)]. \end{aligned} \quad (2.7)$$

Les symboles $\mathbf{M}_1, \dots, \mathbf{M}_P$ désignent des fonctions agissant sur des P -uplets et délivrant une valeur numérique. Evidemment, P est beaucoup plus grand que le N du système (2.4), mais en contrepartie chaque $X_j(t_i)$ est maintenant un simple nombre au lieu d'une fonction d'espace $\mathcal{X}_j(t)$. Dans notre exemple à trois variables pronostiques, si chaque champ est représenté par un million de valeurs numériques, nous avons donc une valeur de P égale à trois millions, avec donc autant d'équations scalaires à résoudre. Tout au long du processus de prévision, l'atmosphère est donc représentée par une suite de vecteurs d'état successifs, et on voit que l'ensemble de toutes ces données n'est finalement qu'un paquet indexé de nombres, qu'il faut manipuler en essayant de leur faire respecter au mieux les lois qui régissent l'atmosphère réelle. La contrepartie négative de l'approche numérique discrétisée est évidemment que la représentation n'est plus exacte mais approchée, puisqu'on a remplacé des champs et les opérateurs continus par des champs et des opérateurs discrétisés, desquels une partie de l'information initiale est absente.

En analyse numérique, les quatre combinaisons possibles (continu/discret ; espace/temps) ont un intérêt : (i) le système pleinement continu en temps et en espace représente la réalité idéalisée que nous choisissons de simuler, (ii) le système discret en espace mais continu en temps permet d'analyser et comprendre les effets de la discrétisation en espace, (iii) le système continu en espace mais discret en temps permet de comprendre les effets de la discrétisation temporelle, et (iv) le système pleinement discret reflète le comportement du modèle numérique utilisé pour les prévisions. Pour ce qui est du problème plus spécifique de la stabilité, qui nous occupera au chapitre 3, c'est plutôt la dimension temporelle qui intervient, et donc on s'intéresse plutôt aux propriétés comparées des systèmes (i), (iii) et (iv). Pour les problèmes de précision des opérateurs spatiaux près des poles, tels qu'examinés au chapitre 5, ce sont plutôt les propriétés comparées des systèmes (i), (ii) et (iv) qui seront pertinentes.

2.2 Système linéaire, système linéaire-tangent d'un système non-linéaire

Le système (2.1) régissant l'évolution de l'atmosphère est dit linéaire quand l'opérateur continu \mathcal{M} est linéaire par rapport aux variables $(\mathcal{X}_1, \dots, \mathcal{X}_N)$ de (2.4). L'opérateur \mathcal{M} peut alors être représenté par une "matrice d'opérateurs" \mathcal{M}_{ij} de sorte que :

$$\begin{pmatrix} \partial \mathcal{X}_1 / \partial t \\ \dots \\ \partial \mathcal{X}_N / \partial t \end{pmatrix} = \mathcal{M}_{ij} \cdot \begin{pmatrix} \mathcal{X}_1 \\ \dots \\ \mathcal{X}_N \end{pmatrix} \quad (2.8)$$

où chacun des éléments \mathcal{M}_{ij} est un opérateur linéaire agissant sur la variable \mathcal{X}_j . L'opérateur \mathcal{M} étant linéaire, l'opérateur discrétisé en espace \mathbf{M} sera linéaire lui aussi, en supposant l'étape de discrétisation cohérente sur ce point. L'opérateur \mathbf{M} possèdera alors une nature matricielle classique (agissant sur des nombres), ce qui n'était pas le cas pour un système non-linéaire. Pour un système linéaire, dans les deux

cas (continu ou discret), l'arsenal puissant de l'algèbre linéaire matricielle pourra alors être utilisé pour examiner les propriétés de l'évolution du système.

Toutefois, en météorologie, dans le cas concret qui nous intéresse, les opérateurs \mathcal{M} et \mathbf{M} sont non-linéaires et donc l'opérateur discrétisé \mathbf{M} ne peut pas être représenté par une matrice. Pour ces systèmes non-linéaires, l'étude des propriétés de l'évolution temporelle ne peut être menée que dans des cas particuliers où le problème se simplifie. Nous examinons ici le cas discrétisé en espace et continu en temps, mais la même discussion peut se transposer pour les autres cas. Dans le cas où l'opérateur \mathbf{M} est non-linéaire, l'évolution à un instant t_0 donné d'une petite perturbation infinitésimale \mathbf{x}' autour d'un vecteur d'état $\mathbf{X}(t_0) = \mathbf{X}_0$ sera donnée par :

$$\frac{\partial \mathbf{x}'}{\partial t} = \left. \frac{\partial \mathbf{M}}{\partial \mathbf{X}} \right|_{\mathbf{X}_0} \cdot \mathbf{x}', \quad (2.9)$$

où $\left. \frac{\partial \mathbf{M}}{\partial \mathbf{X}} \right|_{\mathbf{X}_0}$ est l'opérateur linéaire tangent à l'opérateur \mathbf{M} en \mathbf{X}_0 à l'instant t_0 . En effet l'opérateur linéaire tangent à \mathbf{M} en \mathbf{X}_0 est l'opérateur linéaire qui satisfait :

$$\forall \mathbf{X}_1, \quad \mathbf{M}(\mathbf{X}_1) - \mathbf{M}(\mathbf{X}_0) = \left. \frac{\partial \mathbf{M}}{\partial \mathbf{X}} \right|_{\mathbf{X}_0} \cdot (\mathbf{X}_1 - \mathbf{X}_0) + O(\|\mathbf{X}_1 - \mathbf{X}_0\|^2). \quad (2.10)$$

Or, puisque

$$\mathbf{M}(\mathbf{X}_1) - \mathbf{M}(\mathbf{X}_0) = \frac{\partial \mathbf{X}_1}{\partial t} - \frac{\partial \mathbf{X}_0}{\partial t}, \quad (2.11)$$

on retrouve bien l'équation (2.9), en substituant $\mathbf{x}' = \mathbf{X}_1 - \mathbf{X}_0$ et en négligeant le terme résiduel puisque \mathbf{x}' a été supposée infinitésimale. Si les opérateurs spatiaux mis en œuvre dans \mathbf{M} sont raisonnablement simples, l'opérateur linéaire tangent se calcule aisément, ce qui est en général le cas pour les opérateurs qui concernent la partie dynamique de l'évolution, beaucoup moins pour ceux qui concernent l'évolution physique à petite échelle (turbulence, changements de phase,...) et qui peuvent présenter des phénomènes de seuils ou de bifurcations difficiles à représenter de façon précise par un opérateur linéaire.

Il existe un cas particulier important pour l'étude des propriétés du système (2.6), c'est le cas des états stationnaires. Si \mathbf{X}_0 est un état tel que $\mathbf{M}(\mathbf{X}_0) = 0$, alors $\mathbf{X}(t) = \mathbf{X}_0$ est solution du système quelle que soit la valeur de t . On dit que \mathbf{X}_0 est un état stationnaire du système. L'évolution par le système complet (2.6) d'une perturbation infinitésimale quelconque \mathbf{x}' autour de cet état \mathbf{X}_0 est alors donnée par le système linéaire tangent, et donc :

$$\frac{\partial \mathbf{x}'(t)}{\partial t} = \left. \frac{\partial \mathbf{M}}{\partial \mathbf{X}} \right|_{\mathbf{X}_0} \cdot [\mathbf{x}'(t)]. \quad (2.12)$$

Cette équation donne donc accès à l'évolution réelle du système complet à relativement long terme, c'est à dire tant que $\mathbf{X}_0 + \mathbf{x}'$ reste relativement proche de \mathbf{X}_0 , de telle sorte que l'évolution donnée par le modèle linéaire tangent reflète celle du modèle complet dans cette situation particulière.

De cette discussion, il ressort que grâce à l'arsenal de l'algèbre linéaire, on peut aisément examiner les propriétés – et notamment la stabilité – des systèmes non-linéaires autour de leurs états stationnaires, puisque l'évolution du modèle va se réduire à un problème d'algèbre linéaire, et donc, essentiellement, de modes propres. La principale restriction est que par cette méthode on ne peut explorer qu'un faible échantillon de toutes les situations possibles dans la réalité. Le fait de considérer le système linéaire tangent et non pas le système non-linéaire complet signifie en pratique que les propriétés du schéma ne peuvent être étudiées que pour de petites perturbations autour de l'état choisi. La méthode décrite ne permet donc pas de connaître les propriétés du système pour de grandes perturbations autour de cet état. Il faut bien voir toutefois que cette dernière restriction est peu gênante pour l'étude de la stabilité, car la pratique montre que les instabilités démarrent le plus souvent avec une amplitude infinitésimale, liée par exemple au bruit numérique introduit par la troncature des nombres réels représentés dans une machine par des nombres décimaux avec une arithmétique à virgule flottante de précision finie.

Dans le cas général où l'état réel \mathbf{X}_0 n'est pas un état stationnaire, l'évolution à long terme d'une perturbation autour de \mathbf{X}_0 va consister en une superposition de l'évolution de \mathbf{X}_0 par \mathbf{M} et de l'évolution de \mathbf{x}' par son opérateur linéaire tangent. Il devient difficile dans ce cas d'étudier les propriétés du comportement du système d'une façon théorique.

Pour résumer, on peut dire que la stabilité des systèmes linéaires peut être étudiée pour tous les états et toutes les perturbations possibles, tandis que pour les systèmes non-linéaires, on ne peut étudier aisément que la stabilité du système pour les états stationnaires et des perturbations infinitésimales (ceci en restreignant l'étude au système linéaire tangent au lieu du système non-linéaire complet).

2.3 Matrice jacobienne, modes propres

La matrice jacobienne des opérateurs \mathcal{M} ou \mathbf{M} permet de calculer en pratique l'opérateur linéaire-tangent, utile pour l'étude de la stabilité des systèmes dynamiques considérés. Dans le cas continu en espace, si on écrit l'opérateur \mathcal{M} comme la fonction vectorielle définie en (2.4), la matrice jacobienne en \mathcal{X}_0 est la matrice de taille $N \times N$ formée des opérateurs suivants : $\mathcal{J}_{ij} = (\partial \mathcal{M}_i / \partial \mathcal{X}_j)$, évalués en \mathcal{X}_0 . Dans le cas discret en espace, en partant de la forme (2.7), la matrice jacobienne en un état \mathbf{X}_0 est la matrice numérique de taille $P \times P$ formée des éléments $\mathbf{J}_{ij} = (\partial \mathbf{M}_i / \partial \mathbf{X}_j)$ évalués pour l'état \mathbf{X}_0 donné, chacun des \mathbf{J}_{ij} étant donc une simple valeur numérique. La matrice jacobienne n'est autre qu'une représentation élément par élément de l'opérateur linéaire tangent ($\partial \mathcal{M} / \partial \mathcal{X}$) ou ($\partial \mathbf{M} / \partial \mathbf{X}$) en un état donné. Elle permet d'accéder à l'expression du système linéaire tangent en pratique, ou plus précisément à l'effet de l'opérateur linéaire tangent sur tout vecteur d'état représentant une perturbation par rapport à \mathbf{X}_0 . En effet si χ' est une petite perturbation superposée à \mathbf{X}_0 , on a

$$\frac{\partial \chi'(t)}{\partial t} = \left. \frac{\partial \mathcal{M}}{\partial \mathcal{X}} \right|_{\mathcal{X}_0} \cdot [\chi'(t)]. \quad (2.13)$$

Dans le cas continu, la matrice jacobienne \mathcal{J}_{ij} étant diagonalisable sur \mathbb{C} , toute perturbation autour de \mathcal{X}_0 se décompose en combinaison linéaire de modes propres χ' , et les valeurs propres associées λ sont les fréquences complexes continues de ces modes propres, puisque

$$\frac{\partial \chi'(t)}{\partial t} = \left. \frac{\partial \mathcal{M}}{\partial \mathcal{X}} \right|_{\mathcal{X}_0} \cdot [\chi'(t)] = \lambda \chi'(t), \quad (2.14)$$

ce qui entraîne :

$$\chi'(t) = \chi'(0) \exp(\lambda t). \quad (2.15)$$

Un raisonnement similaire s'applique aux cas discrétisés en espace et/ou en temps et conduit à des fréquences complexes discrètes mais cette fois-ci en nombre fini, λ_n .

2.4 Matrice d'amplification d'un système linéaire discretisé en espace et temps

Considérons maintenant la version discrétisée en espace et en temps du système dynamique basé sur (2.12), ou n'importe quel autre système linéaire de la forme suivante.

$$\frac{\partial \mathcal{X}}{\partial t} = \mathcal{L} \mathcal{X} \quad (2.16)$$

Le vecteur d'état (discrétisé en espace) à un instant donné est noté \mathbf{X} comme ci-dessus. Pour la version discrétisée en temps, on peut toujours définir un vecteur d'état *généralisé* \mathbf{Z} contenant exactement l'intégralité de l'information variable dans le temps qui est nécessaire pour calculer l'état futur à partir de l'état présent dans cette discrétisation temporelle. Ce vecteur d'état est encore une liste de nombre comme \mathbf{X} , mais possiblement plus grand que \mathbf{X} lui-même. Par exemple, dans le cas d'une discrétisation explicite Leap-Frog, on aura $\mathbf{Z}(t) = [\mathbf{X}(t), \mathbf{X}(t - \Delta t)]$ car pour calculer l'état futur, on a besoin à la fois de l'état présent et de l'état passé. Dans ce cas, \mathbf{Z} est donc exactement deux fois plus long que \mathbf{X} . Une fois cette définition de \mathbf{Z} soigneusement effectuée, le système dynamique de départ (discret en temps et espace) peut toujours être réorganisé afin de faire apparaître $\mathbf{Z}(t + \Delta t)$ (noté \mathbf{Z}^+) à gauche comme une fonction linéaire de $\mathbf{Z}(t)$ (noté \mathbf{Z}^0).

Puisque cette fonction est linéaire et numérique, c'est donc une matrice réelle que l'on peut noter \mathbf{A} , et appelée matrice d'amplification. Dans notre exemple d'un schéma Leap-Frog explicite, on aurait

$$\frac{\mathbf{X}^+ - \mathbf{X}^-}{2\Delta t} = \mathbf{L} \mathbf{X}^0, \quad (2.17)$$

où $\mathbf{Z}(t - \Delta t)$ est noté \mathbf{Z}^- . La réorganisation évoquée ci-dessus consiste donc simplement ici à écrire :

$$\mathbf{X}^+ = 2\Delta t \mathbf{L} \mathbf{X}^0 + \mathbf{I} \mathbf{X}^- \quad (2.18)$$

$$\mathbf{X}^0 = \mathbf{I} \mathbf{X}^0 + \mathbf{0} \mathbf{X}^- \quad (2.19)$$

où \mathbf{I} et $\mathbf{0}$ sont les matrices identité et nulle de taille adéquate. Nous avons donc simplement $\mathbf{Z}^+ = (\mathbf{X}^+, \mathbf{X}^0)$, $\mathbf{Z}^0 = (\mathbf{X}^0, \mathbf{X}^-)$, et

$$\mathbf{A} = \begin{bmatrix} 2\Delta t \mathbf{L} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}. \quad (2.20)$$

Evidemment, pour des schémas temporels plus sophistiqués, le vecteur généralisé \mathbf{Z} et la matrice \mathbf{A} peuvent prendre une forme plus complexe. La matrice \mathbf{A} est la matrice qui contient l'intégralité du processus qui fait passer d'un état généralisé 'courant' $\mathbf{Z}(t)$ à l'état généralisé 'futur' $\mathbf{Z}(t + \Delta t)$ pour la discrétisation spatiale et temporelle choisie :

$$\mathbf{Z}(t + \Delta t) = \mathbf{A} \mathbf{Z}(t). \quad (2.21)$$

Il est clair que cette matrice, via ses valeurs propres, est liée à la stabilité du système discrétisé en temps et espace. Toute autre matrice, plus simple, ne peut prétendre en elle-même donner accès à la stabilité du schéma de discrétisation de façon correcte.

2.5 Systèmes faiblement non-linéaire

Les systèmes utilisés en PN sont par nature non-linéaires, mais en un certain sens, ils sont parfois qualifiés de 'faiblement non-linéaires', car les termes qui sont responsables des variations les plus rapides sont presque linéaires. Cette définition est laissée ici sous une forme peu formalisée et assez floue, mais illustrée plus en détail maintenant sur un exemple concret, intéressant directement la météorologie : examinons le système des Equations d'Euler, celui qui décrit sans approximation l'évolution d'un gaz parfait non visqueux soumis à un champ de gravité \mathbf{g} dans un référentiel galiléen. Après choix des variables pronostiques (\mathbf{V}, T, p) , l'évolution peut s'écrire sous la forme :

$$\frac{\partial \mathbf{V}}{\partial t} = -\frac{RT}{p} \nabla p - (\mathbf{V} \cdot \nabla) \mathbf{V} + \mathbf{g} \quad (2.22)$$

$$\frac{\partial T}{\partial t} = -\frac{RT}{C_v} \nabla \cdot \mathbf{V} - \mathbf{V} \cdot \nabla T \quad (2.23)$$

$$\frac{\partial p}{\partial t} = -p \frac{C_p}{C_v} \nabla \cdot \mathbf{V} - \mathbf{V} \cdot \nabla p \quad (2.24)$$

où \mathbf{V} , T , p sont respectivement le vecteur tridimensionnel du champ de vent, la température absolue et la pression, R , la constante des gaz parfaits, C_v , C_p les chaleurs spécifiques à volume et pression constante du gaz considéré, et ∇ représente le symbole de l'opérateur différentiel du calcul vectoriel.

Dans l'atmosphère décrite par ce système appliqué à l'air sec, les termes d'évolution qui – de beaucoup – peuvent être les plus grands en valeur absolue, sont le premier terme de chaque membre de droite. Les termes d'évolution de nature advective (seconds termes de droite) sont d'amplitude plus faible, c'est à dire qu'ils donnent des évolution beaucoup moins rapides. Supposons que les fonctions thermodynamiques C_v et C_p sont constantes dans les domaines qui nous intéressent (gaz parfait de Laplace), le système est tout de même non-linéaire, puisque par exemple nous avons des termes comme $(RT \nabla p / p)$ qui ont une dépendance non-linéaire par rapport aux variables d'état. Cependant, on se rend compte que la non-linéarité des termes les plus rapides provient de dépendances en T , p qui sont certes variables, mais relativement peu. En un lieu donné, ces deux champs s'écartent peu d'une valeur moyenne constante dans le temps : typiquement la pression s'écarte d'environ $\pm 5\%$, et la température absolue d'environ $\pm 10\%$ autour de la valeur moyenne en ce lieu. On dit que ces termes sont faiblement non-linéaires. Finalement,

les termes qui sont le moins “faiblement non-linéaires” (dans le sens précédent) sont les termes advectifs, mais ils contribuent peu à la valeur totale de l’évolution puisqu’ils entraînent des tendances plus faibles que les termes rapides.

Puisque les termes potentiellement prépondérants pour l’évolution sont presque linéaires, on pourrait penser que les propriétés d’un algorithme de discrétisation pour le système complet devraient logiquement rester assez proches des propriétés du système linéaire tangent, ce qui tend donc à justifier les études sur le système linéaire tangent. Néanmoins, d’autres aspects (non décrits dans l’exemple ci-dessus) peuvent venir invalider ce raisonnement. Notamment, la présence du relief terrestre impose, d’une manière ou d’une autre l’apparition de termes non-linéaires importants. Dans une moindre mesure, les aspects liés aux processus microphysiques, radiatifs, turbulents, etc. peuvent créer des non-linéarités qui écartent notablement le système complet d’un système faiblement linéaire *stricto sensu*. Dans la réalité, le système que l’on résout pour faire de la prévision numérique, ne ressemble pas du tout à (2.22)-(2.24), mais quelque chose de bien plus compliqué, avec une multitude de termes fortement non-linéaires, mais globalement moins importants que les termes rapides.

Ceci montre qu’il est finalement assez illusoire de vouloir analyser de façon exacte les propriétés d’un algorithme de discrétisation du système d’équation complet que l’on a choisi pour décrire l’évolution de l’atmosphère : dès lors qu’un certain réalisme est atteint dans la représentation de l’atmosphère, les équations s’éloignent notablement d’un système simple et faiblement non-linéaire tel que celui décrit plus haut en (2.22)-(2.24) et la validité de l’analyse s’en trouve nécessairement réduite. En dernier ressort, l’analyse des propriétés d’une discrétisation d’un système donné ne vaut que dans des cadres restreints par rapport à l’ensemble des situations que sera amené à rencontrer le système dans la réalité. Le corollaire un peu désespérant de tout ce que nous venons de dire est donc le suivant : une analyse qui prédit un schéma instable dans certaines conditions se verra très probablement confirmée par une manifestation inévitable de cette instabilité en pratique, mais une analyse qui ne détecte pas d’instabilité ne garantira jamais que le système discrétisé sera stable en pratique, puisque l’analyse ne dit rien sur les situations qu’elle n’examine pas.

2.6 Méthode spectrale

La discrétisation spatiale des opérateurs spatiaux (notamment différentiels) peut être effectuée en suivant une multitude d’approches. La plus simple est celle des différences finies, dans laquelle toute les dérivée de long d’une direction donnée du système de coordonnées curvilignes est évaluée par différences des valeurs des points voisins, ce qui, au passage suppose une certaine structuration régulière de la grille de discrétisation. Une autre approche, la méthode spectrale, est brièvement introduite maintenant car elle est appliquée dans les modèles de PNT mis en œuvre à Météo-France.

Dans le cas du modèle global (représentant l’atmosphère sur l’ensemble du globe), les variations horizontales d’un champ continu sont représentées par une superposition finie d’harmoniques sphériques Y_n^m , avec une troncature triangulaire afin d’obtenir une résolution horizontale approximativement isotrope (e.g. Coiffier, 2011). Dans le cas d’un modèle en aire limitée (représentant l’atmosphère sur une région du globe limitée par un quadrilatère), les champs sont représentés par des fonctions de Fourier dans chacune des deux directions horizontales. Dans les deux cas, ces fonctions sont choisies parce qu’elles sont les modes propres de l’opérateur Laplacien, d’une importance centrale dans l’évolution des variables en jeu.

La transformation spectrale, qui fait passer des champs physiques aux amplitudes des fonctions de la représentation spectrale, est linéaire. Les harmoniques sphériques ainsi que les fonctions bi-Fourier évoquées ci-dessus forment une base orthogonale de l’espace vectoriel considéré. Néanmoins, le produit de deux champs représentables dans l’espace vectoriel généré par cette base n’est pas nécessairement représentable dans cette base (le produit quadratique d’une composante de Fourier par elle-même donne par exemple un champ de fréquence spatiale double qui peut ne pas être représenté avec la troncature initiale).

Comme les termes sources des modèles météorologiques contiennent de nombreux termes non-linéaires, on applique pour résoudre les équations temporellement une méthode dite de “transformation spectrale”, dans laquelle toutes les variables d’état (spectrales) sont recalculées dans l’espace physique (sur une grille dite de collocation) à chaque pas de temps en appliquant les formules de transformation spectrale, ceci afin de pouvoir les combiner en produits lorsque cela est nécessaire. Une fois ces produits effectués, les termes sources correspondants sont transformées à nouveau dans l’autre sens vers l’espace spectral afin de calculer l’évolution des variables d’état spectrales. Par rapport à une méthode complètement

spectrale, dans laquelle les produits seraient calculés via les coefficients d'interaction, cette méthode par transformation est computationnellement avantageuse, mais elle impose de bien maîtriser les éventuels problèmes d'aliasing, ce qui n'est pas toujours facile (e.g. Courtier et Naughton, 1994).

Chapitre 3

Stabilité des schémas numériques en PNT

3.1 Stabilité des schémas discrets dans les modèles numériques

L'objet n'est pas ici de disserte sur la nature et la définition théorique de la stabilité ou de l'instabilité de tel ou tel système dynamique dans un cadre pleinement continu (en temps et en espace), c'est à dire de l'instabilité hydrodynamique au sens large. Il est au contraire question dans ce chapitre d'un problème plus concret, qui est la stabilité intrinsèque des schémas de discrétisation (en temps et en espace), désignée par le terme 'stabilité numérique'.

Il n'est cependant pas aisé de définir de façon immédiatement intelligible ce que l'on entend par 'stabilité numérique' d'un schéma de discrétisation, du moins, d'un point de vue théorique, et ce point n'est pas non plus abordé ici. D'un point de vue pratique, la définition est un peu plus simple si on ne s'embarrasse pas trop des détails : un schéma numérique est instable si de temps en temps les champs prévus ont un aspect manifestement irréaliste (bruit) ou si la prévision ne peut être menée à son terme (explosion). L'instabilité numérique est donc un artefact par lequel l'évolution simulée devient rapidement non-physique. Au bout d'un certain temps d'intégration, les résultats perdent tout réalisme, soit par la présence de bruit numérique dans les résultats, masquant le signal physique cherché, soit tout simplement par l'émergence de valeurs numériques sortant du domaine de la représentation autorisée par la machine de calcul ("overflow") ou physiquement interdites (pression ou température absolue négative).

Généralement, le premier de ces deux symptômes (le bruit), n'est d'ailleurs que le signe annonciateur du second (l'explosion), qui advient un peu plus tard si l'on pousse plus avant l'intégration temporelle. Cependant, tel n'est pas toujours le cas : il peut s'établir un régime non-physique conforme à l'idée que l'on se fait de l'instabilité, mais qui finit par 'saturer' aux grandes amplitudes à cause de la survenue d'autres rétroactions négatives, étrangères à la cause première de l'instabilité (e.g. Kalnay and Kanamitsu, 1988; Girard and Delage, 1990; Bénard et al. , 2000). Dans ces cas plus contestables, certains parleront d'instabilité tandis que pour d'autres c'est plutôt la consistance qui est mise en défaut, ce qui montre que la frontière entre les deux causes (instabilité ou inconsistance) n'est pas toujours limpide pour des cas concrets.

Un dernier point mérite d'être mentionné : à haute résolution, certains champs atmosphériques ont eux-même un aspect bruité, notamment les contenus spécifiques en hydrométéores précipitants. Une prévision réaliste des ces champs se doit donc d'avoir une apparence bruitée, ce qui dans certains cas corse un peu plus la définition du caractère stable ou instable d'un schéma de discrétisation. Une piste pour départager le naturel de l'artificiel est que le bruit naturel est généralement peu structuré ; à l'inverse le bruit lié à un schéma de discrétisation numérique instable possède souvent des régularités suspectes (bandes parallèles, structure spatiale en damier, ou au contraire oscillations en chaque point possédant une période commensurable avec le pas de temps). Mais une telle structuration n'est pas toujours le cas, et il nous est parfois arrivé, en examinant des prévisions préparées avec deux options différentes et ne différant que par le niveau de 'bruit' dans les cartes, de nous demander avec une certaine perplexité si la meilleure prévision n'était pas en définitive celle qui contenait le plus de signal à petite échelle, c'est à dire le plus de ce qui pouvait ressembler à du bruit, avec là encore une certaine difficulté à conclure. Cependant, dans la grande majorité des cas, un algorithme numériquement instable conduit à un comportement facilement

identifiable comme tel. La suite de ce chapitre se concentre sur ce type d'instabilité, tel qu'il a été souvent rencontré au cours du développement de nos modèles opérationnels.

3.2 Diagnostic ou analyse

Dans l'immense majorité des cas, lorsque l'on est confronté à un système dynamique discrétisé présentant une évolution instable, il est impossible de comprendre d'où provient ce comportement instable sans une analyse poussée, s'insinuant jusqu'au cœur du schéma de discrétisation. Notamment, il est généralement impossible d'expliquer et comprendre l'instabilité rien qu'en observant le comportement spatio-temporel des champs issus de la prévision, même au moyen de diagnostics détaillés. Il n'y a jamais un "accident initial" identifiable sur tel ou tel champ prévu qui "déraillerait" tout d'un coup, déclenchant ainsi l'instabilité initiale, accident qui puisse être décelé par un simple examen diagnostic de l'état courant à un moment donné.

L'éclairage ci-dessus à l'aide de la matrice d'amplification et de ses valeurs propres est instructif à cet égard : une instabilité sera finalement la manifestation du fait que l'une des valeurs propres, λ_n de la matrice d'amplification du système discrétisé en temps possède une partie réelle non-nulle, alors que les matrices jacobiniennes originelles (\mathcal{J} et \mathbf{J} de la section 2.3) avait toutes leurs valeurs propres purement imaginaires.

Il se peut aussi que toutes les valeurs propres de \mathcal{J} (continu en espace) soient imaginaires, mais que \mathbf{J} (discret en espace) possède des valeurs propres avec une partie réelle non-nulle. C'est un cas plus subtil, conduisant également à de l'instabilité, mais alors indépendamment de la discrétisation temporelle. La discrétisation spatiale est alors déféctueuse et une instabilité est susceptible d'apparaître quel que soit le schéma temporel utilisé. Fabrice Voituz a lancé récemment quelques pistes de recherche dans ce dernier axe, en proposant d'examiner dans un contexte Hamiltonien la conservation de la pseudo-énergie du système linéarisé autour d'un état de base simple. ceci conduit à des contraintes de discrétisation spatiale qui permettent d'assurer que le schéma discret en espace est exempt de tels défauts.

Dans tous les cas évoqués ci-dessus, les modes propres instables sont réalistes dans leur structure spatiale, cette dernière n'ayant rien de non-physique. Un tel mode pourrait se manifester dans la réalité ou même croître en amplitude de façon légitime pendant un certain temps au cours de l'évolution. Seulement, même dans cette éventualité, il sera 'stable' dans le cas continu, c'est à dire que par le jeu des rétroactions négatives, ou simplement du fait des changements dans la structure des champs due à l'évolution, l'état courant du système sera significativement modifié, et alors, dans la matrice jacobienne autour de ce nouvel état courant, ce mode ne sera plus instable.

Dans le cas discrétisé en temps au contraire, pour des raisons qui se cachent au cœur de l'algorithme de discrétisation, ce mode sera instable au début comme son équivalent physique (continu) pouvait très bien l'être, mais contrairement à celui-ci, la rétroaction négative n'existe pas et la croissance de ce mode peut poursuivre indéfiniment, et ce, de façon exponentielle. C'est là le comportement typique d'une instabilité linéaire, telle qu'elle apparaît le plus souvent dans les schémas discrétisés. Les instabilités non-linéaires sont plus rares en pratique, et ne sont pas abordées ici.

C'est cette dissimulation de la cause première de l'instabilité tout au fond des algorithmes qui en rend la compréhension parfois difficile notamment pour les systèmes complexes, la seule solution pour comprendre l'instabilité étant de l'analyser mathématiquement, cela n'étant possible que dans des cas très particuliers où le système est débarrassé de nombre de processus, et où ce système simplifié est lui-même appliqué à des écoulements stationnaires et extrêmement simplifiés.

Nous arrivons avec cette dernière remarque à une sorte de paradoxe débouchant sur un point méthodologique important. Une infime partie des situations réelles \mathcal{X} est susceptible de se prêter à une analyse de stabilité pour les modèles non-linéaires (ce sont essentiellement des états stationnaires presque triviaux). Et pourtant, ce genre d'analyse de stabilité a bien souvent représenté au cours de l'histoire une démarche fructueuse pour démêler l'origine d'un comportement instable, et trouver des parades qui finalement permettaient de stabiliser le système entier, même pour des écoulements complexes, c'est à dire très éloignés de celui avec lequel avait été effectuée l'analyse.

La principale réponse, à mon sens, est que les instabilités étudiées ici sont des instabilités linéaires et que l'atmosphère réelle étant elle-même faiblement non-linéaire, la garantie d'une stabilité pour des écoulements stationnaires linéaires entraîne souvent le constat d'une stabilité pour des écoulements réels. Il est d'ailleurs remarquable que lorsque une instabilité se manifeste dans un modèle complexe, on parvient en général à réduire la complexité de l'écoulement tout en conservant le caractère instable du schéma de

discrétisation temporelle, et ce, jusqu'à obtenir un comportement instable dans un contexte accessible à l'analyse, notamment par le fait que l'écoulement est stationnaire et l'évolution réelle parfaitement linéaire. C'est la démarche qui a presque invariablement été adoptée pour résoudre les cas d'instabilité observés, non seulement au cours de l'histoire, mais aussi dans les articles personnels que je présente ci-dessous.

Face à un système dynamique instable, la seule démarche efficace et fructueuse n'est donc pas d'examiner le comportement par l'extérieur (les résultats), mais par l'intérieur (les algorithmes).

Selon la méthode discutée précédemment, face à une instabilité se manifestant dans des cas réels, il faut 'purifier' l'instabilité, c'est à dire parvenir à la faire se manifester de façon essentiellement inchangée dans des contextes de plus en plus épurés et simplifiés, jusqu'au point où l'écoulement est si possible stationnaire et l'évolution purement linéaire. Pour cela, on essaie donc de simplifier le système dynamique lui-même (en retirant des sources peu susceptibles d'être cruciales, comme par exemple la force de Coriolis en grille A, ou les termes de courbure sur la sphère), puis les conditions de l'écoulement, c'est à dire la complexité de la variable globale d'état \mathcal{X} ou du vecteur d'état \mathbf{X} , tout cela en vérifiant à chaque étape que l'on conserve les caractéristiques de l'instabilité initiale. Ainsi, on peut espérer distinguer quels sont les facteurs qui jouent un rôle actif dans l'instabilité et devront donc être conservés dans l'analyse, et ceux qui n'entrent pas en ligne de compte, qui pourront être laissés de côté dans l'analyse. Avant de pouvoir débiter le travail d'analyse en lui-même, il y a donc à effectuer auparavant un important travail de diagnostic sur le comportement du système, en le simplifiant par étapes, et en cherchant si l'instabilité est toujours présente dans des formes similaires à celle que l'on tente d'expliquer initialement.

3.3 Systèmes raides et schémas implicites

Dans le domaine des systèmes dynamiques (linéaires ou non), la notion de système raide se réfère à des systèmes pour lesquels les échelles de temps qui caractérisent les évolutions s'étalent sur un large spectre. Pour un système physique non-linéaire \mathcal{M} , dire que le système est "raide" signifie que, au moins pour certains états \mathcal{X} , la matrice jacobienne du système linéaire tangent autour de \mathcal{X} possède un spectre très large en excluant les éventuelles valeurs nulles, c'est à dire que certaines des valeurs propres non-nulles ont un module beaucoup plus grand que certaines autres.

Les systèmes dynamiques qui nous intéressent ici sont des systèmes couplés, pour lesquels la raideur se manifeste le plus souvent au travers des termes de couplage entre plusieurs variables. Il s'agit donc de termes d'ajustement représentant des interactions rapides entre ces variables, alors que d'autres termes sources génèrent des échelles caractéristiques en temps beaucoup plus longues.

Dans de nombreux domaines, les systèmes dynamiques décrivant le monde physique réel, dès qu'ils sont un tant soit peu complexes, sont caractérisés par une raideur importante, que ce soit pour la chimie, la physique de la matière, la combustion, etc. La mécanique des fluides et en particulier la prévision numérique n'échappent pas à cette règle, du fait de la coexistence dans l'atmosphère de phénomènes avec très courtes échelles de temps coexistant avec des phénomènes beaucoup plus lents. Parmi les phénomènes les plus rapides représentés dans les modèles actuels, on peut citer l'ajustement hydrostatique (responsable des ondes acoustiques), l'ajustement gravitationnel (responsable des ondes de gravité), et certaines composantes de l'ajustement turbulent ou microphysique. Les phénomènes plus lents sont par exemple l'advection de champs lisses, la force de Coriolis, les phénomènes radiatifs, etc.

Dans le cas pleinement continu, la raideur signifie donc que l'évolution permise par le système peut consister en une superposition d'évolutions très lentes et d'évolutions très rapides. Mais dans le cas de systèmes discrétisés en temps, les évolutions rapides, mal représentées à cause de la résolution temporelle limitée se transforment souvent en évolutions instables et ce sont donc elles qui sont le plus souvent incriminées dans les instabilités liées aux schémas de discrétisation. Idéalement, il devrait y avoir une certaine adéquation entre les mouvements les plus rapides autorisés par le système continu, et la résolution temporelle liée à la discrétisation, mais ce n'est pas toujours le cas, surtout pour les discrétisations dites hautement efficaces, qui mettent en jeu des schémas implicites.

Pour les systèmes raides, la mise en place d'une intégration temporelle hautement efficace, avec des pas de temps bien supérieurs aux périodes les plus courtes des phénomènes représentables entraîne donc nécessairement des erreurs pour ces phénomènes, aussi doit-elle résulter d'un choix raisonné. La description des phénomènes les plus rapides étant nécessairement erronée, il faut accepter le fait de commettre ces erreurs, et justifier que leur commission a un impact acceptable.

L'examen des systèmes d'équations météorologiques suggère que les perturbations les plus rapides,

responsables d'une grande partie de la raideur observée, ne charrient souvent qu'une faible part des échanges d'énergie mis en jeu dans l'atmosphère, et que les phénomènes qu'ils décrivent sont d'intérêt météorologique plus médiocre que celui des perturbations lentes. Il faut cependant reconnaître que cette proposition est mathématiquement assez floue et parfois controversée.

A titre d'exemple de cette controverse, on peut citer le déclenchement 'à distance' de la convection par des ondes de gravité elles-mêmes liées à la convection : lorsqu'une cellule convective violente survient, disons au-dessus d'une plaine surchauffée, cette cellule peut générer des ondes de gravité qui vont se propager rapidement autour d'elle et peuvent alors, au gré d'un soulèvement distant (par ces ondes) du fluide instable, entraîner peu après, le déclenchement à distance d'une seconde cellule convective, et ainsi de suite, dans une sorte de réaction en chaîne, par laquelle le périmètre de la zone affectée par la convection peut s'étendre bien plus rapidement que la vitesse du vent. De tels déclenchements à distance sont parfois spectaculairement visibles en imagerie satellitaire. Si les termes d'ajustement rapides liés aux ondes de gravité sont fortement distordus par la discrétisation temporelle, il y a un risque que la réaction en chaîne se fasse de manière peu réaliste, ou même pas ne se produise pas du tout. C'est ce genre de risque qu'il faut accepter a priori quand on utilise des grands pas de temps rendus possibles par des algorithmes 'hautement efficaces'.

L'acceptation *a priori* de ces risques (quitte à en évaluer les conséquences pratiques *a posteriori*) conduit donc à la justification d'une classe de schémas de discrétisation temporelle dans lesquels on accepte de distordre l'évolution liée aux termes les plus rapides, afin de réduire 'artificiellement' la raideur du système discrétisé en temps, et de s'affranchir ainsi de fortes contraintes sur la longueur du pas de temps et l'efficacité globale du système. Cette acceptation d'une certaine distorsion au prix de l'efficacité a prévalu depuis la naissance de la Prévision Numérique, conduisant à la grande popularité des algorithmes dits 'implicites'. Cependant, avec le développement de la météorologie d'échelle fine ayant l'ambition de prévoir individuellement chaque cellule convective, ces schémas pourraient être abandonnés dans le futur.

Les algorithmes implicites sont des algorithmes de discrétisation temporelle dans lesquels l'inconnue (i.e. le vecteur d'état futur) se retrouve non seulement dans le membre de gauche du système d'équation tel que (2.4) ou (2.7), mais aussi dans le membre de droite, du moins dans certains des termes de ce membre de droite. L'analyse numérique montre que les termes du membre de droite qui sont ainsi traités implicitement n'interviennent plus dans la contrainte sur le pas de temps, car ces algorithmes limitent par construction la version discrète du temps caractéristique des processus liés à ces termes (période discrète) à une valeur supérieure au pas de temps. En d'autres termes la fréquence discrète de ces phénomènes traités implicitement est ramenée à une valeur inférieure à l'inverse du pas de temps.

Le traitement implicite d'une partie des termes d'évolution peut donc augmenter la stabilité du système global, mais rend l'intégration temporelle plus compliquée : en effet l'inconnue n'est plus isolable algébriquement puisqu'elle se trouve maintenant en de multiples endroits du système d'équation de part et d'autre du signe d'égalité, soumises à des opérateurs couplés beaucoup plus complexes que le simple opérateur $\partial/\partial t$ ou $\delta/\delta t$ du membre de gauche. Par ailleurs, adopter un traitement implicite suppose, comme mentionné ci-dessus que les évolutions rapides liées aux termes responsables de la raideur sont jugées moins intéressantes que les évolutions lentes, car du fait du recours à ces algorithmes implicites, l'évolution de ces phénomènes rapides sera certes stabilisée, mais aussi largement modifiée et déformée (via l'altération de leur fréquence discrète).

3.4 Principe des schémas semi-implicites

3.4.1 Schémas semi-implicites traditionnels (SI)

Supposons un système dynamique discrétisé selon (2.6). Le principe du schéma semi-implicite est d'identifier les termes de \mathbf{M} qui sont responsables des phénomènes rapides, de proposer une version linéaire pour ces termes, ce qui aboutit donc à un opérateur linéaire noté \mathbf{L}^* , puis d'appliquer la discrétisation temporelle suivante :

$$\frac{\delta \mathbf{X}(t)}{\delta t} = (\mathbf{M} - \mathbf{L}_*) [\mathbf{X}(t)] + \mathbf{L}_* [\bar{\mathbf{X}}^t(t)] \quad (3.1)$$

où $\bar{\mathbf{X}}^t$ représente un opérateur discret de moyenne temporelle, incluant nécessairement le niveau temporel futur (afin que l'algorithme soit bien implicite). Pour que la précision du schéma soit au moins d'ordre 2

en temps, il faut que cet opérateur de moyenne soit centré. On voit donc que dans ce type de schémas, certains termes (rapides) linéaires sont discrétisés de façon implicite, tandis que le reste du système est discrétisé de façon explicite (d'où le nom *semi-implicite*). Dans ce contexte, les termes composant la partie traitée en explicite ($\mathbf{M} - \mathbf{L}_*$) sont appelés "résidus non-linéaires". Evidemment, le calcul de l'état futur est plus compliqué que dans le cas d'un schéma explicite, car l'inconnue se trouvant non seulement dans le membre de gauche mais aussi dans le dernier terme du membre de droite, il est nécessaire d'inverser un système linéaire pour accéder à sa détermination ce qui entraîne la résolution d'un problème en espace appelé 'problème implicite'.

A dessein, nous sommes restés vagues jusqu'ici sur ce qu'est concrètement \mathbf{L}_* . En effet, la forme que doit prendre cet opérateur pour une efficacité ou une robustesse optimale du schéma n'est pas connue, et la définition concrète d'un opérateur \mathbf{L}_* qui répond correctement aux exigences souhaitées relève parfois davantage du savoir-faire que de d'une méthode gravée dans le marbre qu'il suffirait de suivre. Jusqu'à une époque relativement tardive, un dogme solidement ancré a consisté à chercher \mathbf{L}_* comme le linéaire tangent à l'opérateur \mathbf{M} autour d'un état donné \mathbf{X}_* , très simple et stationnaire. J'ai toutefois montré, au moins dans deux cas précis et avec des conséquences concrètes en prévision numérique (Bénard, 2004) que ce n'est pas toujours la meilleure façon de procéder, et qu'une définition élargie du système linéaire par rapport à cette conception pouvait être largement bénéfique.

Les schémas SI, dans la mesure où ils sont stables, sont efficaces, car ils permettent de s'affranchir dans une large mesure des limitations sur la longueur du pas de temps liées aux évolutions rapides, telles qu'elles se présentent dans les schémas explicites. Pour illustrer ceci, on remarquera que si \mathbf{M} est lui-même un système linéaire, dont le module maximum des valeurs propres est noté λ_{\max} , alors une discrétisation explicite possèdera une limitation pour le pas de temps de l'ordre de

$$\delta t < O(1/\lambda_{\max}), \quad (3.2)$$

tandis qu'une discrétisation semi-implicite reviendra dans ce cas particulier à un schéma implicite centré (du fait de l'absence de résidus non-linéaires) et sera donc inconditionnellement stable (sous réserve d'une discrétisation spatiale pertinente).

Il est généralement reconnu que les schémas implicites, bien que plus complexes à mettre en œuvre, présentent en définitive un net avantage sur les schémas explicites, surtout quand les systèmes à résoudre sont raides et que la représentation en détail des phénomènes rapides n'est pas essentielle. Pour les systèmes météorologiques, le caractère raide ne fait pas vraiment question, car même dans le modèle d'atmosphère le plus simple possédant des termes d'ajustement (système en eau peu profonde), il y a déjà des ondes de gravité se déplaçant à une vitesse proche de celle du son, alors que les phénomènes météorologiques les plus intéressants se déplacent à une vitesse de l'ordre de quelques mètres par secondes, c'est à dire à la vitesse du fluide atmosphérique lui-même.

Une autre façon de voir les schémas semi-implicites est de réécrire (3.1) sous la forme :

$$\frac{\delta \mathbf{X}(t)}{\delta t} = \mathbf{M}[\mathbf{X}(t)] + \mathbf{L}_* [\bar{\mathbf{X}}^t(t) - \mathbf{X}(t)]. \quad (3.3)$$

Le dernier terme apparait alors comme une "correction semi-implicite" appliquée au modèle explicite, simplement représenté par tous les autres termes.

Les schémas semi-implicites ont été jusqu'ici d'une importance capitale en PN, par l'efficacité et la robustesse qu'ils ont permis d'atteindre, en comparaison avec les schémas explicites. Seulement, ces schémas sont difficiles à mettre en œuvre sans analyse de leur comportement. Les premiers travaux de prévision numérique (Charney et al., 1950) associent déjà l'analyse de stabilité à la Prévision Numérique, à tel point que le concept d'analyse de stabilité de Von Neumann, général à toute la physique, s'appuie en fait sur cet article fondateur, initialement appliqué à la météorologie.

Enfin, signalons que les analyses de stabilité de Von Neumann peuvent indifféremment être appliquées au système continu en espace ou discrétisé en espace. La première variante donnera des enseignements sur la discrétisation temporelle du système dans une plus grande généralité (puisque le choix arbitraire d'une discrétisation spatiale n'est pas encore fait), mais en revanche, la seconde variante donnera des enseignements sur la stabilité telle qu'elle sera observée en pratique dans le système correspondant. Pour illustrer cette distinction, disons que si une analyse de stabilité révèle qu'un système continu en espace est instable, il est illusoire de penser qu'une quelconque discrétisation spatiale de ce système pourra être stable, où alors cette stabilité sera induite et obtenue pour des raisons artificielle liées à la discrétisation (application d'une forte diffusion numérique, ou d'importants décentrages temporels...).

3.4.2 Schémas implicites itératifs centrés (ICI)

A ce point il est nécessaire d'introduire de nouveaux types de schémas, qui sont en quelque sorte une extension conceptuelle des schémas SI. Dans la section 3.4.1, nous avons mentionné que si le système complet M est linéaire et qu'on choisit comme système linéaire de référence L_* précisément le système M , alors le schéma SI est inconditionnellement stable sous réserve d'une discrétisation spatiale pertinente. De ce constat est née l'idée que pour un système général (non-linéaire) M une discrétisation temporelle de Crank-Nicolson (dite aussi "trapézoïdale") devrait être une sorte d'optimum de stabilité. Dans le formalisme symbolique adopté ci-dessus, ce schéma temporel s'écrit

$$\frac{\delta \mathbf{X}(t)}{\delta t} = M \left[\overline{\mathbf{X}}^t(t) \right], \quad (3.4)$$

ou bien, sous une forme proche et plus facile à mettre en œuvre :

$$\frac{\delta \mathbf{X}(t)}{\delta t} = \overline{M[\mathbf{X}(t)]}^t. \quad (3.5)$$

Pour un schéma à deux niveaux temporels, la discrétisation vise donc à définir \mathbf{X}^+ par

$$\frac{\mathbf{X}^+ - \mathbf{X}^0}{\Delta t} = \frac{M(\mathbf{X}^+) + M(\mathbf{X}^0)}{2}. \quad (3.6)$$

Le problème est qu'en général cette méthode n'est pas applicable directement car l'opérateur M du second membre est non-linéaire ce qui conduit donc à un problème implicite non-linéaire. L'idée des schémas implicites itératifs centrés (ICI) est donc d'essayer d'approcher ce schéma par une méthode itérative, détaillée dans Bénard (2003). Ces schémas sont donc une méthode de quasi-Newton pour approcher la solution cherchée en (3.5)–(3.6). Par rapport à un schéma explicite, la première itération conduit au schéma SI classique, tandis que la seconde itération conduit à ce qui est parfois appelé schéma "prédicteur/correcteur" (PC) dans la littérature (p. ex. Cullen, 2001).

On définit donc un indice d'itération noté (k) , et les états successifs valables à $t + \Delta t$ sont notés $\mathbf{X}^{+(k)}$. Après choix d'un système linéaire L_* , l'état après la première itération, noté $\mathbf{X}^{+(0)}$ sera l'état donné par le schéma SI, via (3.1), puis pour les itérations suivantes on prendra (pour un schéma à deux niveaux temporels) :

$$\frac{\mathbf{X}^{+(k)} - \mathbf{X}^0}{\Delta t} = \frac{M(\mathbf{X}^{+(k-1)}) + M(\mathbf{X}^0)}{2} + \frac{L_*(\mathbf{X}^{+(k)}) - L_*(\mathbf{X}^{+(k-1)})}{2} \quad (3.7)$$

On remarque que *si le processus itératif converge*, après un nombre infini d'itérations le dernier terme disparaît, donnant ainsi la réponse souhaitée, c'est-à-dire celle du schéma de Crank-Nicolson, avec donc la promesse espérée d'une très grande stabilité numérique. Mais si l'on recherche uniquement un gain quantitatif de stabilité par rapport au schéma SI, on peut envisager de se limiter à une ou deux itérations au-delà du SI (toujours en supposant que le processus soit convergent lorsque un grand nombre d'itérations est employé).

Néanmoins, si le processus ne converge pas, il y a un risque que le schéma obtenu soit encore plus instable que le schéma SI, et ceci, même avec seulement une ou deux itérations. L'intérêt de cette approche, qui n'est que potentiel et hypothétique (plus grande stabilité, pas de temps plus longs) se paie d'une faiblesse qui, elle, est tout à fait avérée (le coût élevé, dû aux multiples itérations nécessaires pour compléter un pas de temps). Partant, le jeu en vaut-il vraiment la chandelle ? Il y a une sorte d'attraction à reculons pour ces schémas, et ils sont toujours examinés en gardant ce point à l'esprit, c'est à dire que le bénéfice de chaque itération supplémentaire devra être dûment démontré, pour une application opérationnelle, et doit être révisé à chaque modification.

Dans les deux premières applications mettant en œuvre ce type de schémas, Côté and staniforth (1988), Bubnová et al. (1995), seul un seul terme de l'opérateur M (celui jugé le plus dangereux) était soumis au traitement itératif. Symboliquement, cela revient à dire que l'opérateur M était scindé en deux parties M_1 et M_2 , et que le premier terme de droite (3.7) était remplacé par

$$\frac{M(\mathbf{X}^{+(k-1)}) + M(\mathbf{X}^0)}{2} \rightarrow \frac{M_1(\mathbf{X}^{+(0)}) + M_1(\mathbf{X}^0)}{2} + \frac{M_2(\mathbf{X}^{+(k-1)}) + M_2(\mathbf{X}^0)}{2} \quad (3.8)$$

On voit que le terme M_1 est traité comme dans un simple schéma semi-implicite, tandis que le terme M_2 est itéré pour tenter de s'approcher du schéma de Crank-Nicolson pour cette partie des sources.

Cette approche partielle permettait de limiter le surcoût du processus itératif tout en espérant un gain substantiel de stabilité. Dans les deux papiers cités, le seul terme inclus dans le traitement itératif M_2 était le terme lié au forçage orographique. À l'autre extrême, on peut inclure tout le système M dans le processus itératif, et –proprement– c'est alors qu'on parle de schéma ICI. C'est cette dernière approche qui est appliquée de nos jours lorsqu'on se réfère à ce type de schéma (par exemple dans le modèle AROME). Dans les implémentations semi-lagrangiennes, des variantes sont possibles avec réévaluation ou non du point origine de la trajectoire des particules à chaque itération.

3.5 Géométrie horizontale et stabilité

3.5.1 Contexte

En se plaçant tout à fait en amont de la démarche, le météorologiste qui doit bâtir un système d'équations pour un modèle de prévision dispose au départ de deux corpus de connaissances générales très différents qu'il doit "mélanger" pour les adapter au problème spécifique de la météorologie. D'une part la mécanique des fluides, qui décrit l'évolution des grandeurs pertinentes sans hypothèse sur la géométrie du problème. Ceci débouche sur la formulation d'une forme invariante (vectorielle ou même Hamiltonienne) des équations, qui dépeignent les lois et principes naturels dans toute leur généralité. Le système des équations d'Euler est un exemple de système qui peut être présenté sous une forme totalement indépendante de la géométrie. D'autre part le cadre géométrique particulier de l'atmosphère terrestre que l'on étudie. Entrent notamment dans ce cadre l'existence d'une limite inférieure rigide (le sol), l'existence d'un champ de gravité à l'importance prépondérante, l'existence d'une rotation diurne, et enfin l'existence de symétries plus ou moins grossièrement vérifiées pour la forme de la planète, et pour son champ de gravité apparente.

Contrairement à d'autres disciplines comme la géodésie ou la cartographie, il a été jusqu'ici largement admis en météorologie que les écarts à une symétrie sphérique simple pouvaient être négligés. On admet donc en Prévision Numérique une planète sphérique, avec un champ de gravité apparente à symétrie sphérique purement radial et on admet que les surfaces horizontales sont des sphères concentriques. Ici on appelle 'surface horizontale' (ou isopotentielle) une surface en chaque point de laquelle la bulle d'un niveau à bulle au repos par rapport au sol apparaît centrée. J'ai montré dans Bénard 2014 qu'en toute rigueur il n'est pas possible que deux surfaces isopotentielles soient des sphères concentriques autour d'une planète en rotation. Néanmoins, dans ce même article, je précise que pour ce qui est de la mécanique des fluides autour de la planète, la vitesse de rotation Ω_{Cent} permettant de définir l'accélération centrifuge du champ de gravité apparent, et la vitesse de rotation Ω_{Cor} permettant de définir les forces de Coriolis peuvent être spécifiées indépendamment l'une de l'autre tout en garantissant tout de même une cohérence physique (voir également Appendice 2). Finalement donc, cette géométrie à 'géopotentiel sphérique' utilisée en Prévision Numérique est tout à fait appropriée¹.

Dans ce cadre à 'géopotentiel sphérique', il est naturel d'adopter le système des coordonnées sphériques avec sa métrique relativement simple pour décrire la position dans l'espace ainsi que les différents champs liés à la gravité. En utilisant cette coordonnée, on transforme chaque opérateur spatial initialement sous forme invariante (divergence, gradient, Laplacien...), en un opérateur spécifique exprimé dans la géométrie sphérique, et on applique cette forme spécifique des opérateurs spatiaux à chacune de leur occurrence dans le système. Le champ de gravité et le paramètre de Coriolis f prennent des formes très simples dans ce cadre sphérique. Ainsi, on obtient le système d'évolution choisi pour décrire l'évolution du fluide planétaire que constitue l'atmosphère, valable dans un système de coordonnées sphériques, exprimé mathématiquement par des équations d'évolutions à variables scalaires (réelles).

Ceci étant posé, un des problèmes de base qui s'est rapidement posé à la PNT a été de trouver des méthodes pour simuler à échelle fine l'atmosphère météorologique, au moins localement sur un domaine d'intérêt, le reste de l'atmosphère étant simulé à une échelle plus lâche, donc de façon plus économique. Deux approches ont été utilisées en PNT pour cela : les modèles en aire limitée, couplés sur leurs frontières latérale à un modèle global plus lâche ; et les modèles globaux à maille variable. Dans les modèles mis en œuvre à Météo-France, les deux approches sont utilisées : ARPEGE est un modèle global à maille variable, tandis qu'AROME est un modèle en aire limitée, couplé au modèle ARPEGE (ou IFS).

1. Une option moins restrictive, dans laquelle la planète et les surfaces horizontales sont des ellipsoïdes de révolution est examinée au chapitre 4

Dans les deux cas (maille variable ou aire limitée), on part des équations formulées en coordonnées sphériques, puis on applique une transformation géométrique non-isométrique agissant sur les coordonnées horizontales, afin d'aboutir à un système de coordonnées satisfaisant les propriétés voulues de résolution accrue sur la zone d'intérêt. Pour ARPEGE, il s'agit d'une transformation conforme de sphère à sphère en coordonnées sphériques (Schmidt, 1977), qui entraîne une résolution variant d'un facteur arbitraire entre deux points opposés de la sphère initiale. Pour AROME, il s'agit d'une transformation conforme de projection de la sphère vers un plan, décrit en coordonnée cartésienne. Plusieurs projections conformes sont disponibles : Mercator (naturel ou basculé), Lambert, et stéréographique polaire.

Cet ajout d'une transformation géométrique va poser des problèmes spécifiques pour la mise en place de schémas semi-implicites robustes et efficaces. Dans cette section on s'intéressera surtout aux problèmes liés à ces transformations géométrique horizontales, et la dimension verticale passe au second plan de la discussion. L'opérateur différentiel ∇ représente donc ici seulement sa partie horizontale (à coordonnée verticale constante).

Dans tous les cas, la transformation géométrique évoquée ci-dessus étant conforme, les opérateurs différentiels du calcul vectoriel (divergence et rotationnel de vecteurs, gradient et laplaciens de scalaires,...) dans la géométrie transformée se déduisent des opérateurs natifs (dans la géométrie sphérique réelle) via une simple division par le facteur d'échelle de la transformation m , variable dans l'espace :

$$\nabla = m\nabla' \tag{3.9}$$

Dans ce contexte particulier de transformation conforme appliquée à la sphère, on désigne ∇ sous le nom d'opérateur de dérivation natif, ou physique, dit "géographique", tandis que ∇' est l'opérateur de dérivation purement numérique dans la coordonnées transformée, appelé opérateur "réduit".

Pour fixer les idées, dans la configuration opérationnelle actuelle du modèle à maille variable ARPEGE, la valeur de l'étirement dans la transformation de Schmidt est $c = 2.4$ et le facteur d'échelle m varie donc de $1/2.4$ à 2.4 entre les antipodes et la France. La résolution spatiale sur la France est donc 5.76 fois meilleure qu'à l'antipode. Dans le modèle AROME, la valeur du facteur d'échelle m est exactement 1 au centre du domaine, et augmente avec la distance au centre, si bien que sa valeur maximale dépend de la taille du domaine. Pour une projection de Lambert avec un domaine de l'ordre de 3000 km d'extension (dans les directions N-S et E-O), la valeur maximale du facteur d'échelle est de l'ordre de 1.03. Les variations du facteur d'échelle dans des applications typiques d'AROME sont beaucoup plus faibles que pour la configuration d'ARPEGE opérationnelle.

3.5.2 Instabilité liée à la linéarisation du facteur d'échelle

La coordonnée utilisée dans le modèle pour le calcul des opérateurs différentiels spatiaux étant la coordonnée transformée, l'espace spectral dans lequel est inversé le problème implicite est nécessairement celui des modes propres des opérateurs réduits de dérivée spatiale ∇' ou ∇'^2 (ces modes propres sont, pour ARPEGE les harmoniques sphériques, et pour AROME les composantes de Fourier bidimensionnelles). Or, dans les équations complètes ce sont les opérateurs géographiques qui apparaissent. La dépendance entre les deux types d'opérateurs n'étant pas linéaire (puisque le facteur d'échelle varie localement), il est donc nécessaire de faire un choix pour linéariser les termes dérivés spatialement dans l'opérateur linéaire \mathbf{L}_* qui sert à définir le schéma semi-implicite. Il avait été supposé, et constaté empiriquement avec un modèle en eau peu profonde (Courtier and Geleyn, 1988 – cf. notamment section 3.c), qu'il était suffisant de linéariser ces termes en remplaçant le facteur d'échelle local par sa valeur maximale m_* sur le domaine, pour obtenir un schéma stable. Un terme qui par exemple s'écrivait $\nabla p = m\nabla'p$ dans le modèle complet \mathbf{M} , devenait $m_*\nabla'p$ dans le modèle linéaire \mathbf{L}_* du schéma SI.

Cette approche fut appliquée dans le modèle en équation primitives hydrostatiques (HPE selon l'abréviation anglaise consacrée) ARPEGE. En 1995 ARPEGE fonctionnait avec un schéma eulérien en troncature T119 et un étirement $c = 3.5$ à l'époque. Le pas de temps était modulable selon chaque situation, déterminé par essais-erreurs à chaque prévision, car la stabilité des schémas eulériens est directement liée à la valeur maximale du vent, c'est là leur principal inconvénient. Lorsqu'il fut décidé de faire passer en opérationnel une version semi-lagrangienne en troncature T149 avec le même étirement, pour s'affranchir de cette contrainte handicapante sur le pas de temps, de graves problèmes d'instabilité furent rencontrés.

Le syndrome principal de cette instabilité était une dérive inexorable des températures vers le (et même au-delà du) zéro absolu sur l'Antarctique. Comme bien souvent lors de l'instabilité d'un système dynamique, il était malaisé de comprendre l'origine du problème par un simple examen de l'évolution

temporelle des champs simulés. En l'absence d'analyse du phénomène, c'est à Karim Yessad que revient le mérite d'avoir proposé (sans preuve donc) une hypothèse correcte pour l'origine possible de cette instabilité : il s'agissait du trop fort déséquilibre de l'amplitude des corrections semi-implicites entre les zones de fortes résolution et les zones de faible résolution, pour les termes contenant des dérivées horizontales. En effet, si les corrections semi-implicites, données par le dernier terme de (3.3), avaient une amplitude correcte sur la zone d'intérêt (où l'on avait effectivement $m = m_*$), elles étaient plus de douze (3.5 au carré) fois trop fortes à l'antipode (où $m = 1/m_*$). Ceci n'est pas une preuve directe car appliquer une correction semi-implicite forte n'est pas nécessairement en soi un facteur d'instabilité. Néanmoins, si l'on considère le résidu non-linéaire (traité explicitement) constitué par le premier terme du membre de droite de (3.1), il devient également très grand dans cette approche, avec ici une explication plus crédible du risque d'instabilité.

La solution proposée consistait à abandonner l'idée que les termes linéaires contenant des dérivées horizontales pouvaient simplement être obtenus en linéarisant les termes complets autour d'une valeur uniforme m_* , comme cela était fait jusqu'ici. Avec cette nouvelle solution, un terme de la forme ∇^2 serait linéarisé non plus sous la forme $m_*^2 \nabla'^2$, mais sous la forme exacte $m^2 \nabla'^2$. Cette solution fut documentée et validée dans l'article Yessad et Bénard (1996). Cela était rendu possible par le fait que dans la transformation de Schmidt (1977), le carré du facteur d'échelle m^2 peut se décomposer sur les trois premiers éléments dit 'zonaux' de la base des harmoniques sphériques (soit : Y_0^0 , Y_1^0 et Y_2^0). Rappelons que les harmoniques sphériques sont par construction les modes propres de l'opérateur laplacien horizontal réduit sur la sphère, ∇'^2 . Par conséquent, d'après les formules donnant la décomposition spectral d'un produit de deux harmoniques sphériques (via les coefficients de Clebsch-Gordan, ou symboles dits $3j$ -Wigner) la décomposition spectrale du produit $m^2 \nabla'^2$ pour chaque mode propre s'exprime comme combinaison des modes propres possédant des nombres d'onde totaux n' voisins à plus ou moins deux unités près du nombre d'onde total n de l'harmonique sphérique initiale.

Techniquement, l'application de la méthode spectrale devenait donc légèrement plus compliquée, l'inversion dans l'espace spectral de l'opérateur laplacien horizontal ∇'^2 devenant un problème d'inversion de matrices penta-diagonales, au lieu d'un problème trivial d'inversion de matrices diagonales lorsque m était représenté par un simple nombre m_* . Néanmoins, ce petit surcroît de complexité était jugé acceptable, car il permettait de sauver l'idée de l'étirement en semi-implicite avec les longs pas de temps autorisés par le schéma semi-lagrangien. L'inversion du problème penta-diagonal ne se traduisait que par un surcoût très limité en nombre de calculs (relativement au coût total du modèle) car tout comme le problème diagonal, le calcul des opérateurs inverse peut-être effectué une bonne fois pour toute au début de l'intégration du modèle.

Le résultat fut sans appel : la stabilité du modèle fut aussitôt restaurée, et les avantages économiques du schéma semi-lagrangien (en terme de longueur du pas de temps) purent être pleinement exploités. Un an après ce succès, l'efficacité du schéma fut encore doublée avec la mise en œuvre d'un schéma dit "à deux niveaux temporels", en remplacement du traditionnel schéma "saute-mouton" à trois niveaux temporels (Temperton et Staniforth, 1987). Le facteur d'échelle ne posa pas de nouveau problème lors de ce passage.

3.5.3 Epilogue (2004)

Ce n'est qu'en 2004, lorsque notre équipe fut en possession d'une culture plus solide en matière d'analyse de stabilité pour nos algorithmes de calculs (cf. section suivante), que l'origine de cette instabilité put être analysée et confirmée de façon théorique, ainsi que la pertinence de la solution adoptée. Les résultats de ces études de stabilité sont exposés dans le rapport d'un stage effectué dans l'équipe par Fabrice Voitus en 2004 (Voitus, 2004). Dans ce rapport de stage, il est notamment montré que le modèle hydrostatique (en équations primitive) avec sa discrétisation temporelle, est instable en présence de relief, de grands facteurs d'étirement c , et de résidus thermiques non-linéaires. Ceci est donc entièrement compatible avec l'instabilité qui avait été rencontrée sur l'Antarctique dix ans plus tôt, phénomène qui était resté sans véritable explication théorique à l'époque. En l'absence de relief, le choix initial avec m_* dans la partie linéaire aurait fonctionné, mais ce n'est malheureusement pas le cas pour le continent Antarctique, d'altitude très élevée.

3.6 Stabilité du système d'Equations d'Euler en semi-implicite

3.6.1 Contexte historique

Au moment où les schémas semi-implicites font leur première apparition en Météorologie (Marchuk, 1965 en tant que précurseur, sans mise en œuvre pratique), il est connu depuis plusieurs années que les systèmes linéaires traités avec un schéma implicite centré sont inconditionnellement stables (pour les systèmes discrétisés en espace cette propriété requiert certaines conditions de consistance dans la discrétisation spatiale). Pour les systèmes non-linéaires, même traités avec des algorithmes implicites, le problème de la stabilité est plus complexe, du fait des non-linéarités (voir, p. ex. Richtmyer, 1957). Si l'on considère les systèmes d'équations météorologiques comme "faiblement non-linéaires" il était donc logique d'espérer obtenir une bonne stabilité en traitant une grande partie (linéaire) de leurs termes sources les plus rapides avec le schéma implicite centré, d'où certainement l'idée des schémas SI.

Les deux études pionnières pour l'utilisation des schémas SI en météorologie sont Kwizak and Robert, 1971; et Robert et al., 1972. Ce sont ces deux publications qui ont montré l'avantage que pouvaient représenter les algorithmes SI pour la Prévision Numérique, en terme de stabilité et d'efficacité. Dans Kwizak and Robert, 1971, le schéma SI est expérimenté pour un modèle en 'eau peu profonde' (SW, pour "shallow-water" en anglais) en aire limitée. Les auteurs décrivent le schéma utilisé et montrent que le modèle développé sur ces idées se révèle beaucoup plus efficace que les modèles explicites, pour des résultats presque identiques. C'est donc une victoire éclatante.

On trouve dans ce papier un embryon d'analyse de stabilité, pour tenter d'apporter une justification théorique à l'utilisation du schéma SI. Malheureusement, cette analyse est très simplifiée et avec le recul, elle se révèle presque inutile, et surtout inopérante pour identifier les problèmes qui surgiront dans les schémas SI au cours des trois décennies suivantes. C'est d'autant plus dommage que l'analyse qui aurait été pertinente était à peine plus compliquée à mener, pour ce système très simple.

L'analyse telle qu'elle apparaît dans Kwizak and Robert, 1971 possède deux principales limitations. Tout d'abord, afin de simplifier les calculs, l'analyse ne s'intéresse qu'à un système dynamique comprenant une seule variable pronostique, notée F alors que le système SW est un système comprenant au minimum deux variables pronostiques couplées (par exemple le vent u et le géopotentiel ϕ). Ensuite, au lieu de s'intéresser aux résidus non-linéaires dus à différence entre l'état réel ϕ et l'état de référence ϕ_* , l'analyse s'intéresse à une séparation des termes sources en termes sources en termes advectifs (lents, traités explicitement) et termes d'ajustement (rapides, traités implicitement). Si l'analyse contient bien deux phénomènes de fréquence distincte, elle néglige le rôle potentiellement déstabilisant des résidus de termes d'ajustement traités explicitement. Cette dernière caractéristique conduit à une interprétation très limitative des résultats et une vision trop optimiste de la stabilité réelle du système discrétisé. Les auteurs concluent que le schéma est inconditionnellement stable dès que le vent advectif u est inférieur à la vitesse d'ajustement $\sqrt{\phi_*}$. Comme cette condition est largement respectée pour l'atmosphère, le schéma est déclaré inconditionnellement stable.

Mais cet argument assez rudimentaire reflète mal ce qui peut poser réellement problème dans le système SW traité en SI, notamment lorsque l'état de référence ϕ_* diffère de l'état réel ϕ , si bien que l'analyse échoue à identifier tous les problèmes à venir liés à la discrétisation temporelle SI. Pourtant, l'analyse du comportement du système SW SI en présence de résidus $(\phi - \phi_*)$ traités explicitement n'était pas plus complexe à mener, et aurait pu alerter dès le début sur le fait que pour le système SW SI décrit, le schéma n'est inconditionnellement stable qu'à l'intérieur d'un certain domaine en $\alpha = (\phi - \phi_*)/\phi_*$. Et alerter également sur le fait qu'à l'extérieur de ce domaine, le schéma est inconditionnellement instable! L'embryon d'analyse proposée dans Kwizak and Robert, 1971 résonne donc, a posteriori comme une occasion ratée passant très près d'une analyse beaucoup plus pertinente. A la décharge de ce papier toutefois, les valeurs de ϕ_* qui seraient potentiellement dangereuses pour le schéma SI sont assez éloignées de la valeur moyenne du géopotentiel dans l'écoulement, et le choix qui est fait dans le papier (noté ϕ_0 au début de la section 5) est suffisamment raisonnable pour ne pas entraîner de risque d'instabilité. Le schéma mis en place dans le modèle fonctionne donc conformément aux attentes, mais sans réelle prise de conscience des limites de la méthode à ce stade.

Dans le second papier sur le sujet, Robert et al., 1972, l'idée du schéma SI est appliquée pour la mise en œuvre d'un modèle en equations primitives hydrostatiques (HPE ci-après), avec cinq niveaux équirépartis en pression sur la verticale (900, 700, 500, 300, 100 hPa). Le succès de l'algorithme est clair, mais là encore, aucune analyse de stabilité ne vient encadrer et fonder scientifiquement le schéma ainsi que

ses limites de validité. Il est simplement signalé que l'analyse présentée dans l'article précédent, Kwizak and Robert, 1971, ne peut s'appliquer dans ce cas plus complexe du modèle en équation primitives : le schéma est donc utilisé avec succès mais sans véritable garantie scientifique, ni compréhension réelle de la raison pour laquelle il fonctionne.

Finalement, dans cette période pionnière, les modèles mis en œuvre, relativement simples, semblent fonctionner sans problème quand on applique un schéma SI sans précaution particulière, et donc aucune analyse vraiment pertinente pour la stabilité de ce schéma n'est proposée. A la décharge des pionniers, reconnaissons toutefois qu'étant donnés les succès obtenus, le besoin d'analyses plus pertinentes ne se fait pas grandement sentir non plus...

Les premiers problèmes surviennent quelques années plus tard, et imposent, pour leur compréhension et pour l'élaboration d'une solution, le recours à une analyse plus rigoureuse de la stabilité du schéma. Simmons et al., 1978 (SHB78 ci-après), appliquant à leur tour la technique décrite par Robert et al., 1972, mais dans un modèle à 16 niveaux verticaux avec une atmosphère non-isotherme comprenant une troposphère et une stratosphère, se heurtent tout à coup à une sévère instabilité. Ils proposent pour la première fois une tentative d'analyse de stabilité qui décrit fidèlement les mécanismes à l'œuvre dans le modèle discrétisé en temps, avec le système d'équation couplées, dans une situation volontairement simplifiée, épurée.

Tout d'abord, afin de rendre l'analyse techniquement possible, ils réduisent l'opérateur non-linéaire noté \mathbf{M} dans l'équation (3.3) à un opérateur linéaire que nous noterons $\bar{\mathbf{L}}$. Cet opérateur est obtenu de la même manière que l'opérateur \mathbf{L}_* , mais en linéarisant autour d'un état $\bar{\mathbf{X}}$ différent² de l'état de référence \mathbf{X}_* utilisé pour définir \mathbf{L}_* . Le système à examiner n'est donc plus (3.3), mais devient

$$\frac{\delta \mathbf{X}(t)}{\delta t} = \bar{\mathbf{L}}[\mathbf{X}(t)] + \mathbf{L}_* [\bar{\mathbf{X}}^t(t) - \mathbf{X}(t)]. \quad (3.10)$$

Ensuite, afin de simplifier encore les calculs, les auteurs proposent d'analyser uniquement la situation simplifiée dans laquelle les profils de température de l'état $\bar{\mathbf{X}}$ et de l'état de référence \mathbf{X}_* ne diffèrent que par un facteur constant α :

$$T_* = \alpha \bar{T}. \quad (3.11)$$

Ce cadre d'étude n'est pas assez riche pour refléter fidèlement ce qui produit vraiment l'instabilité dans leur cas précis, mais ils parviennent tout de même à identifier des conditions sur α pour lesquelles le schéma devient inconditionnellement instable. Ceci constitue une sorte de coup de tonnerre : les schémas explicites sont en général conditionnellement stable, c'est à dire que dans le pire des cas, le choix d'un pas de temps suffisamment petit assure la stabilité du schéma. Ici, SHB78 montrent que les schémas SI peuvent certes être plus efficaces que les schémas explicites, mais qu'une formulation légèrement fautive entraîne une instabilité inconditionnelle telle qu'aucun pas de temps, aussi petit soit-il, ne permet de mener à bien la prévision ! Pour la première fois, un inconvénient sévère de la méthode est identifié.

Par ailleurs, au moyen d'expériences numériques conduites avec le modèle complet, les auteurs parviennent à identifier de façon correcte les conditions d'apparition de l'instabilité dans le modèle complet : le schéma SI utilisé devient instable lorsque la tropopause de l'état réel est située *au-dessous* de la tropopause de l'état de référence. Et dans ce cas, conformément aux analyses, réduire le pas de temps ne permet pas de résoudre le problème, et il se révèle effectivement impossible de mener à bien la prévision.

En réalité, comme nous le verrons ci-dessous la cause profonde de l'instabilité réside dans le fait que, lorsque cette différence d'altitude entre les deux tropopauses existe, la stabilité statique de l'état réel est beaucoup plus grande que celle de l'état de référence, dans la couche située entre les deux niveaux de tropopause. Il est clair qu'en imposant $T_* = \alpha \bar{T}$ à chaque niveau, l'analyse de SHB78 ne permettait pas de traiter la situation où les tropopauses des deux profils se trouvent à des niveaux différents. L'analyse de SHB78, trop simplifiée ne permettait donc pas de comprendre véritablement l'instabilité, et les auteurs doivent se contenter de conclusions basées sur les présomptions issues des expériences numériques.

Quoi qu'il en soit, Simmons et al, 1978 constatent que l'utilisation d'un état de référence \mathbf{X}_* isotherme chaud pour définir l'opérateur \mathbf{L}_* et formuler le schéma semi-implicite pour ce système d'équations HPE conduit à un schéma stable, et font donc cette préconisation. Ce résultat se comprend aisément avec le recul : avec un tel choix, les résidus non-linéaires sont certes plus importants qu'avec un état de référence correspondant à l'état moyen ou standard de l'atmosphère, mais le schéma résultant est stable pour tous

2. Noter la distinction entre l'état choisi ici $\bar{\mathbf{X}}$ et la notation utilisée pour l'opérateur de moyenne temporelle discrète $\bar{\mathbf{X}}^t$ appliqué à un état \mathbf{X}

les états réalistes de l'atmosphère réelle que l'on cherche à simuler car, du fait qu'un état isotherme chaud possède une très grande stabilité statique, la stabilité statique est alors partout plus grande dans l'état de référence \mathbf{X}_* que dans l'état réel \mathbf{X} , même quand la tropopause de l'état réel se trouve au-dessous de la tropopause de l'état de référence.

Côté et al., 1983 (CBS83 ci-après) proposent alors une approche un peu différente en vue de l'analyse de la stabilité du schéma SI pour les équations primitives : l'analyse est menée pour le système discrétisé verticalement. Ce faisant, la taille du problème à résoudre pour l'analyse se trouve augmentée, mais la nature du problème est simplifiée : il s'agit alors de résoudre un simple problème de valeurs propres pour des matrices purement numériques.

Dans ce papier CBS83, l'analyse de stabilité est étendue au cas où l'état réel et l'état de référence sont deux états stationnaires de profils thermiques différents quelconques. Les auteurs de CBS83 parviennent à montrer par des raisonnements compliqués d'algèbre linéaire, que la stabilité du schéma est liée aux profils de stabilité statique de l'état réel et de l'état de référence, donnant ainsi une explication satisfaisante de l'instabilité observée dans SHB78. Une explication plus simple sera donnée par Bénard, 2004 (section 2.a), avec une méthode plus proche de celle utilisée dans SHB78.

Ce type d'analyse (proposé par CBS83) permet d'estimer la stabilité du système pour des profils thermiques quelconques, mais la contrepartie négative de cette approche discrétisée verticalement, c'est que le diagnostic est purement numérique, sous forme d'une liste de taux de croissance constituant les valeurs propres d'une matrice numérique, la matrice d'amplification du système discret. Ainsi, lorsque le schéma est instable, cette méthodologie ne permet pas d'identifier réellement les causes profondes (i.e. algébriques) de l'instabilité : en améliorant la prédiction ce qui se passe concrètement dans le modèle discret, on perd donc la possibilité de comprendre l'essence du phénomène algébrique qui engendre l'instabilité. Le diagnostic numérique prend le pas sur l'algèbre analytique. L'analyse discrète devient alors plutôt une 'simulation simplifiée', produisant le reflet anticipé de ce que donnerait un modèle basé sur telle ou telle discrétisation, sans avoir à développer réellement le lourd arsenal d'un modèle basé sur la discrétisation en question.

Pour terminer ce tour d'horizon des études pionnières sur les schémas SI en Prévision Numérique, terminons par une digression. La littérature ne permet pas de comprendre réellement la raison pour laquelle Robert et al., 1972 ne se sont pas heurtés au grave problème d'instabilité rencontré six ans plus tard par Simmons et al. 1978, alors que le cadre et la stratégie étaient très similaires. Deux hypothèses pouvaient expliquer l'absence d'instabilité chez Robert et al. 1972 : soit le système d'équations discrètes utilisé, un peu différent de celui de Simmons et al., 1978, était par nature plus robuste et plus stable que ce dernier, soit la stabilité constatée du système était simplement la conséquence de l'absence de tropopause dans les expériences de Robert et al. 1972. Pour trancher, nous avons donc mené (lors d'un stage ENM en 2011 effectué par Louise Lescop), l'analyse discrétisée verticalement du système utilisé par Robert et al., 1972. La conclusion de cette étude a clairement révélé que la seconde hypothèse était la bonne.

En effet, Robert et al., 1972 avaient mené leurs expériences avec cinq niveaux équi-répartis en pression sur la verticale, et donc le niveau de température le plus élevé était situé à 100 hPa. Pour les profils thermiques, ils avaient très probablement choisi des profils standard de type "OACI", avec une décroissance uniforme de la température avec l'altitude, de 6.5 K/km (ce point n'est pas précisé explicitement dans leur papier). Le niveau le plus élevé de leur modèle se trouvant au voisinage de la tropopause, il n'y avait donc pas de rupture de stabilité statique, ni dans l'état de référence, ni dans l'état réel simulé (l'état du 21 février 1969 à 0000 GMT). La source de l'instabilité identifiée par Simmons et al., 1978, n'était tout simplement pas présente dans leur simulation, du fait de sa trop grande pauvreté en résolution verticale. Noter que Hoskins and Simmons, 1975 avaient mené des expériences similaires à Robert et al., 1972, également avec 5 niveaux sur la verticale, sans rencontrer non plus de problème de stabilité.

Ensuite, nous avons montré que l'utilisation du même système que celui utilisé par Robert et al., 1972, avec le même schéma numérique, mais avec davantage de niveaux verticaux conduit au même type d'instabilité que celui rencontré dans Hoskins et al., 1978, dès que la discrétisation verticale permet une bonne représentation de la tropopause et de la stratosphère. Il n'y avait donc pas une stabilité accrue inhérente aux système utilisé par Robert et al., 1972, et ces auteurs se seraient eux aussi heurtés à ce type d'instabilité dès 1972 s'ils avaient eu les moyens de simuler l'atmosphère avec plus de niveaux dans leur système dès 1972.

Pour clore cette digression, notons qu'en fait le système utilisé par Robert et al., 1972, bien que conceptuellement plus compliqué que celui de Simmons et al., 1978, possède absolument les mêmes caractéristiques de stabilité que ce dernier. L'instabilité apparaît dans les mêmes conditions, avec des taux de croissances tout à fait similaires et des structures de modes instables elles aussi tout à fait similaires.

3.6.2 Premières tentatives en Equations d'Euler

L'utilisation d'un schéma semi-implicite pour intégrer les équations d'Euler pleinement compressibles (avec ondes acoustiques) se révèle dans certains cas plus délicat qu'avec les équations primitives hydrostatiques. Le premier papier défrichant ce domaine de recherche est Tanguay et al., 1990 au Canada. Dans ce papier, un schéma SI bâti sur les mêmes idées que dans CBS83 semble fonctionner sans problème majeur, et quelques centres de recherche se lancent alors dans l'aventure du non-hydrostatique SI, dont notre propre groupe de recherche (le GMAP).

En 1993, nous abordons le développement d'un modèle en aire limitée non-hydrostatique (Aladin-NH) avec un schéma Leap-frog SI (3-TL SI en anglais), analogue à celui présenté dans Tanguay et al., 1990 (le reste du modèle est assez différent : coordonnée verticale, discrétisation spatiale...) L'extension non-hydrostatique du modèle hydrostatique Aladin est réalisée en 1993-1994 par Radmila Bubnová (Brožková depuis) et Gwenaëlle Hello, sous la houlette de Jean-François Geleyn et moi-même. Cependant, le succès initial apparaît rapidement comme assez mitigé : la mise en œuvre initiale avec un schéma Eulérien et des petits pas de temps semble fonctionner assez correctement (Bubnová et al., 1995), mais dès 1995, les tentatives d'extension à un schéma semi-Lagrangien plus efficace (grâce à des pas de temps plus longs) se heurte à des instabilités récalcitrantes qui interdisent généralement de mener les prévisions à leur terme. Cette impossibilité d'utiliser Aladin-NH dans une version viable, ainsi que la volonté politique de concentrer les efforts du CNRM sur le développement d'un autre modèle basé sur une alternative anélastique explicite (modèle Meso-NH) mettent alors transitoirement les travaux sur Aladin-NH plus ou moins en sommeil de 1996 à 2000.

Parallèlement à ces travaux sur les modèles en Equations d'Euler, les progrès sur le schéma semi-lagrangien se poursuivaient et suite aux recherches du CEPMMT, un schéma semi-lagrangien à deux niveau temporels (SL 2-TL en anglais) était développé dans le modèle en vue d'une utilisation opérationnelle (originellement pour le système global hydrostatique). Cette nouvelle version 2-TL du schéma semi-lagrangien était environ deux fois plus économique que la version initiale Leap-Frog, pour une qualité tout à fait équivalente (Simmons and Temperton, 1997 ; Temperton et al., 2001). Il apparut alors que dans notre version NH du modèle Aladin, ce nouveau schéma temporel SL 2-TL induisait une instabilité d'une telle ampleur qu'il était impossible de mener l'intégration sur plus de quelques pas de temps (14 en général) ! Paradoxalement, c'est la 'monumentalité' de cet échec qui va débloquent la situation dans mon esprit, psychologiquement d'abord, et techniquement ensuite.

Tout d'abord, l'ampleur du désastre m'a incliné à penser que si l'instabilité était aussi fulgurante, elle devait être relativement facile à expliquer, comparée à l'instabilité plus légère observée avec le SL 3-TL par exemple, qui pouvait donc sembler plus sournoise, plus profondément cachée au cœur des algorithmes. L'avenir prouvera que cette idée était fautive, que l'origine des deux types d'instabilité était en fait tout autant cachée au plus profond des algorithmes, et qu'il n'y avait pas de différence notable dans la difficulté d'analyser ces instabilités. Mais il n'en reste pas moins que l'étendue de l'échec du SL 2-TL avec le système EE a joué pour moi comme un aiguillon, qui a permis de faire sauter un verrou psychologique et de passer à une approche nouvelle pour résoudre ces problèmes. Il faut rappeler qu'à cette époque, la culture de notre groupe de recherche (GMAP) pour résoudre un problème d'instabilité de ce type était souvent "empirique" : soit on cachait le problème sous le tapis avec le filtre approprié, soit on essayait d'imaginer une solution un peu à l'aveuglette, on la testait, et si cela ne fonctionnait pas, on tentait autre chose. Avec les instabilités rencontrées ici, ce type de méthode "ancestrales" n'avait pratiquement aucune chance de succès, comme il avait fallu se résoudre à le reconnaître après plusieurs tentatives infructueuses.

3.6.3 Premières analyses et formalisme symbolique

Un problème cependant était qu'à l'époque personne n'avait entrepris de bâtir des analyses de stabilité de type SHB78 (continues en espace) ou CBS83 (discrètes en espace) pour ce nouveau système des Equations d'Euler, relativement complexe. Les canadiens n'en avaient pas éprouvé le besoin puisqu'ils n'avaient pas rencontré de problème particulier, et de notre côté, nous ne voyions pas comment aborder cette analyse compliquée, d'autant plus que ce genre d'analyse de stabilité n'entraînait pas dans la culture

de notre Groupe. Le problème d'instabilité évoqué ci-dessus demeura donc non résolu durant ces quatre années (1996-2000), jusqu'à ce que je prenne le problème à bras le corps après avoir subitement compris que si nous étions capables d'écrire l'enchaînement des combinaisons algébriques menant à l'équation de structure de ce système (ce qui était acquis depuis le début), nous devions *nécessairement* être capables de mener avec succès la généralisation de l'analyse de SHB78 ou CBS83 à notre nouveau système d'équation. La raison de cela est que la trame de l'analyse peut toujours être calquée sur la trame de la dérivation de l'équation de structure. Il s'agit dans tous les cas d'éliminer une à une les variables couplées pour aboutir à une équation algébrique unique, portant sur l'évolution d'une variable pronostique unique.

Bien que le détail de l'analyse de stabilité soit plus complexe que la dérivation de l'équation de structure, car mettant en jeu beaucoup plus de termes, son canevas, c'est-à-dire l'enchaînement des combinaisons algébriques à réaliser pour aboutir au résultat final, est le même dans les deux cas. Partant de l'analyse (alors bien maîtrisée) menant à l'équation de structure, il suffisait donc d'être confiant et d'enchaîner pas à pas les calculs (plus volumineux) menant à l'équation de stabilité.

Les premières tentatives d'analyses que j'ai menées vers 2000 (sur des feuilles A3 en mode "paysage" tout de même...) furent immédiatement éclairantes : notre formulation ne pouvait pas être stable. Dans un second temps, la complexité nouvelle de ces analyses fut peu à peu domptée : grâce à un formalisme synthétique bien choisi, elles purent être rationalisées et rendues de plus en plus limpides, concises, synthétiques, et partant, s'appliquent facilement à davantage de situations (changement de coordonnées, de variables pronostiques, de schémas temporels...).

La mise au point de ce formalisme symbolique à partir de 2000 a été importante pour pouvoir progresser et communiquer au sein de la communauté. C'est ce qui a motivé la publication de Bénard 2003, qui proposait d'utiliser ce formalisme pour étudier la stabilité des schémas implicites (SI ou ICI) utilisés en Prévision Numérique. A titre d'illustration, un manuscrit de M. Cullen m'avait été adressé à cette époque pour examen. Il proposait une analyse de stabilité du schéma PC (ICI avec une itération au-delà du SI) pour les équations hydrostatiques, analyse que j'avais aussi réalisé de mon côté peu de temps auparavant, pour me faire la main sur la stabilité des schémas itératifs (ICI). Les résultats de Cullen divergeaient complètement des miens alors qu'ils auraient dû être identiques. Grâce à la compacité du formalisme proposé dans Bénard 2003, il me fut facile non seulement d'identifier la faille dans son analyse mais surtout de pointer clairement et efficacement ce point faible au cours de nos discussions, ce qui lui permit de rectifier son erreur initiale (voir les remerciements dans cet article). L'erreur consistait à supposer que dans un schéma itératif (PC ou ICI) la structure spatiale des modes instables devait être la même à chaque itération intermédiaire (k). Cette hypothèse restrictive n'est pas du tout vérifiée dans la réalité, et une analyse basée sur cette hypothèse conduit à des résultats complètement erronés, y compris pour ce qui est des domaines de stabilité.

Le point important est que ces analyses de type SHB78 (continues en espace) pour les Equations d'Euler firent aussitôt apparaître où se trouvait le point faible du schéma 3-TL que nous avons développé en 1993-95 pour les equations d'Euler. Certains résidus non-linéaires thermiques apparaissaient dans la divergence tri-dimensionnelle et créaient un couplage néfaste entre les différents termes non-linéaires des équations, responsable de l'instabilité. Ce point est discuté en détail dans Bénard et al., 2004 (section 4). La mise en évidence du point faible du schéma sous la forme d'un 'accident algébrique' [voir le terme imaginaire dans les équations (63) et (64)] ouvrait directement la voie vers des solutions exemptes de cet accident, et donc, de cette instabilité [voir les équations (66) et (75)], car l'origine de l'accident pouvait être "tracée" à rebours jusqu'au choix initial des variables pronostiques. Un simple changement (minime) de variable pronostique suffisait à rendre analytiquement stable le schéma semi-implicite en SI 3-TL !

L'implémentation de cette modification, qui n'aurait probablement jamais pu être trouvée "à tâtons" montra que la stabilité du schéma était bien restaurée dans le modèle complet avec son schéma temporel saute-mouton (3-TL, i.e. leap-frog) sur des cas réels avec orographie modérée, que ce soit en eulérien ou en semi-lagrangien. Toutefois, le modèle demeurerait toujours extrêmement instable avec un schéma à deux niveaux temporels (2-TL). Un nouveau travail d'analyse fut donc nécessaire pour parvenir à franchir cette seconde étape.

3.6.4 Stabilité en SI des Equations d'Euler sans relief en 2-TL

L'instabilité observée en 2-TL était extrêmement violente. Un travail de diagnostic montra aisément que cette instabilité existait aussi dans les conditions similaires à celles de SHB78 (état au repos, isotherme). Il était donc clair que l'analyse du problème "à la SHB78" pouvait conduire à une compréhension du problème et peut-être à une solution.

Ce point est exposé dans Bénard 2004 (section 3). Le problème fondamental était que la température de référence T_* apparaissait au numérateur dans certaines équations du système non-hydrostatique, et au dénominateur pour certaines autres. Le signe des résidus thermiques non-linéaires était donc opposés pour ces termes, ce qui là aussi créait un accident algébrique, et quel que soit le choix de T_* , le domaine de stabilité était vide dans le cas d'un schéma 2-TL. Une solution proposée fut de définir deux températures de référence : une chaude pour les termes où T_* apparaissait au numérateur (notée SITR dans le code), et une beaucoup plus froide (dénotée T_{e*} , SITRA dans le code) pour les termes où T_* apparaissait au dénominateur. Une fois cette infime modification implémentée dans le modèle, le schéma 2-TL fut immédiatement stabilisé (toujours avec orographie modérée). A ce stade, une stabilité satisfaisante était donc acquise pour la plupart des classes de schémas temporels disponibles dans le modèle, (2-TL ou 3TL, SI ou PC ou ICI) du moins pour les résolutions utilisées en opérationnel (certaines combinaisons particulières avec les schémas ICI apparaissant toutefois comme instables).

3.6.5 Forces et faiblesses des analyses proposées

La publication de Bénard 2003 et Bénard et al. 2004 souleva deux questions intéressantes. La première sur les conditions aux limites inférieure et supérieure ; la seconde sur les long pas de temps utilisés dans les analyses.

Comme expliqué dans Bénard 2003, le surcroît de complication apporté dans les analyses continues en espace par la présence de conditions aux limites matérielles à la base et au sommet du domaine peut être éliminé en examinant un système voisin, sans conditions aux limites (un système dans lequel le domaine est non-borné verticalement, donc). L'équipe anglaise du UKMO fit alors remarquer que par cette opération, on pouvait aboutir à une estimation erronée de la stabilité. Les travaux sur ce sujet furent d'abord présentés par l'UKMO comme un "Comment on..." (de Bénard et al. 2004, je crois). Dans mon esprit, un "comment on..." était comme une annonce officielle que l'article incriminé était faux, et je bataillai ferme pour présenter mes arguments. Les taux de croissance étaient certes augmentés significativement si on prenait en compte les conditions aux limites (ce qui complexifiait l'analyse), mais le domaine de stabilité était le même, d'où une certaine utilité de l'analyse simplifiée tout de même. Finalement, devant mon opiniâtreté, ces remarques furent publiés sous une autre forme, dans un article indépendant, plus neutre et plus constructif (Staniforth et Wood, 2005) examinant le problème sous un angle plus vaste.

Avec le recul, je pense qu'on peut admettre que les deux parties avaient partiellement raison. Il est vrai que simplifier l'analyse en éliminant les conditions aux limites peut modifier les conclusions, mais d'un autre côté, si une analyse simplifiée permet de pointer un défaut de formulation, de suggérer une parade, et que la parade se révèle opérante dans le modèle réel (qui lui est borné verticalement), le bénéfice de l'analyse reste entier. C'est un argument qui avec le recul me paraît important. Nous analysons des systèmes dynamiques très complexes. Il est extrêmement difficile de démontrer *a priori* qu'une analyse de stabilité d'un système simplifié dérivé du système complexe originel reflètera la stabilité effective du modèle complet. Pire, il peut y avoir, dans le processus menant à l'analyse, des simplifications qui vont aussi bien dans le sens d'une surestimation de la stabilité que d'autres qui vont dans le sens d'une sous-estimation, de telle sorte qu'en définitive, on ne sait plus trop quelle est la portée réelle des résultats obtenus.

L'expérience venant, je suis défenseur d'une approche plus pragmatique. Une analyse vaut ce qu'elle vaut, c'est à dire pas grand chose dans l'absolu si on tient absolument à clamer qu'elle reflète le comportement du modèle complet, dans toute sa complexité. Concrètement, au cours du développement d'un modèle de prévision, ce type d'analyse peut avoir deux objectifs différents : D'une part, l'analyse peut-être vue comme une tentative de sonder la viabilité d'un schéma donné sans avoir à se lancer dans le lourd codage dans un vrai modèle. Un analyse aux résultats désastreux n'encourage pas à se lancer dans un tel codage. En effet, même si elle est pessimiste par certaines simplifications, il est probable que le modèle résultant soit peu attractif. L'analyse fonctionne ici comme une alerte. L'autre objectif possible de l'analyse est de remédier à un défaut de stabilité constaté dans un schéma déjà codé, en suggérant concrètement une modification, une parade. Dans ce cas, même si une analyse est imparfaite, si elle pointe avec succès un défaut possible dans la formulation codée et un remède potentiel exempt de ce défaut, cela vaut la peine de tenter la modification suggérée. Il n'y a donc pas de preuve formelle que le défaut sera éliminé du système complexe, mais seulement une indication que la modification va dans le bon sens. Le point crucial est que souvent dans ce cas, la parade n'aurait pas pu être imaginée sans l'analyse, toute imparfaite qu'elle soit. L'analyse fonctionne ici comme source d'inspiration.

Du coup, on peut se demander quelle est l'utilité de publier des analyses qui ne valent que pour ceux qui y croient, qui ne valent que dans un cadre effroyablement spécifique et restreint, et ensuite de se glorifier des résultats de ces analyses comme une justification scientifique des schémas mis en place. L'information publiée possède alors une chance infime (pour ne pas dire nulle) de servir telle quelle à un autre modélisateur quelque part ailleurs dans le monde. Dès lors, quelle utilité ? A mon sens, il faut voir la publication de telles analyses davantage comme une source d'inspiration que comme la présentation d'une vérité immanente. Un modélisateur confronté à un problème concret avec son propre modèle pourra s'approprier la démarche publiée et la modifier pour l'appliquer à son propre problème.

Paragraphe suivant incorrect, à reformuler.

C'est exactement ce qui s'est produit dans notre cas : l'analyse de SHB78 n'est pas aisément généralisable au cas d'une coordonnée hybride η , au système d'équations plus complexe des équations d'Euler, et aux schémas temporels plus complexes que les 3-TL SI. Les simplifications qui résultent de l'utilisation d'une coordonnée simple σ et des équations hydrostatiques font que l'analyse peut être menée dans un formalisme simple permettant finalement de raisonner en terme de comparaison des fréquences continues des ondes rapides dans l'état de référence et dans l'état de base. Pour les configurations plus compliquées, cela n'est plus possible, et un nouveau formalisme est nécessaire.

Néanmoins, l'analyse de SHB78, avec son principe audacieux et novateur (double linéarisation et considération du résidu explicite) a été une source d'inspiration exceptionnelle qui a permis de faire progresser de façon spectaculaire notre compréhension du schéma SI appliqué à nos Equations d'Euler. Le déclic permettant de dépasser le caractère simplifié de l'analyse de SHB78 dans notre cas a été le constat qu'il suffit de suivre la trame de l'élimination algébrique continue pour arriver à coup sûr à une équation de stabilité s'exprimant sous la forme d'un simple polynôme dans l'espace des nombres complexes.

Une seconde objection de l'équipe anglaise visait le fait d'examiner la stabilité dans la limite des grands pas de temps (cette équipe examinait plutôt la stabilité dans la limite des petits pas de temps jusqu'alors). Mon argument était de dire qu'on souhaitait utiliser ces schémas avec de longs pas de temps, donc il était légitime d'examiner la stabilité aux pas de temps infinis.

Je dois reconnaître avec le recul que l'argument, même s'il ne manquait pas de logique (il n'y eu pas de controverse durable sur la question), possédait ses propres faiblesses, ainsi que nous le comprendrions mieux par la suite. En effet, supposer des pas de temps infinis fait en quelque sorte dégénérer le système examiné, et malgré les apparences, cela peut donc aboutir à une surestimation de la stabilité. De nombreux termes sont divisés par le pas de temps dans l'équation régissant la stabilité. ces termes disparaissent donc aux pas de temps infinis. On peut imaginer que seuls subsistent dans l'analyse les termes relatifs à l'onde la plus rapide permise par le système. Pour les équations d'Euler, les ondes rapides sont des ondes gravito-acoustiques et acousto-gravité. C'est à dire qu'il n'y a pas d'onde de gravité pure ou acoustique pure, toutes les ondes rapides étant en fait mixtes. L'examen de la stabilité aux pas de temps infinis revient à considérer que les seules ondes rapides sont les ondes acoustiques, laissant ainsi de côté le rôle potentiellement déstabilisant des ondes de gravité et surtout de la mixité des ondes.

Fabrice Voitus a montré (non publié) pour certains systèmes qu'un schéma implicite peut être inconditionnellement stable (dans un certain domaine) aux pas de temps infinis, et inconditionnellement instable dans ce même domaine pour des pas de temps finis. Il faut bien admettre que la restriction à une classe de pas de temps particuliers dans l'analyse (infiniment petits ou infiniment grands) est une source d'erreur dans la stabilité estimée. Mais là encore, si l'analyse simplifiée permise par cette restriction réussit à mettre en lumière une faiblesse particulière d'un schéma donné, c'est un apport qu'il convient de ne pas négliger, même s'il ne s'agit pas d'une démonstration mathématique sans faille.

3.6.6 Problème de l'orographie

Les progrès précédents étant engrangés, il apparut rapidement en augmentant la résolution dans le modèle complet pour des études exploratoires, qu'en présence d'orographie avec des pentes fortes la stabilité était encore prise en défaut. Ce nouveau problème, plus complexe, nous mettait devant un défi important car ici, il n'y avait pas de cadre d'analyse immédiatement pertinent. Le cadre proposé par SHB78 ou CBS83, que nous avons enrichi, ne permettait pas de façon évidente de décrire des situations avec orographie.

Il fallut donc s'atteler à la mise au point d'un cadre théorique permettant d'analyser le comportement du modèle en présence d'une pente. Pour atteindre ce but, je proposai de se focaliser sur l'analyse d'un système avec pente constante s s'étirant sur une longueur infinie le long d'une direction x . Ce cadre théorique présente quelques problèmes de réalisme car la pression au sol, au lieu d'être à peu près constante et uniforme en moyenne comme sur la Terre, tend vers l'infini et vers zéro aux deux extrémités du domaine $x = \pm\infty$, ce qui n'est pas très physique. De même l'altitude du sol varie de $-\infty$ à $+\infty$ conjointement avec x , ce qui n'est pas très physique non plus. La Figure 3.1 dépeint cette situation.

Je dois dire que cette faiblesse apparente du cadre d'analyse me mettait un peu mal à l'aise à cette époque, car il prêtait le flanc à la critique du bien-fondé de l'analyse en elle-même. Avec le recul et l'expérience, je crois que cette pudeur était largement exagérée, et je serais même prêt à adopter aujourd'hui une position encore plus extrême consistant à dire qu'il peut parfaitement y avoir une utilité à examiner la stabilité d'un système dynamique en tant qu'objet mathématique même s'il n'est pas complètement réaliste en tant que description du monde physique. Une illustration simple et extrême est qu'on peut très bien se proposer d'examiner par exemple la stabilité des équations primitives hydrostatique en SI comme l'ont fait SHB78, lorsque la température absolue de l'état réel est négative. Rétrospectivement, cette remarque me paraît importante, car elle montre que parfois, en fixant des limites à sa propre réflexion, on prend le risque de passer à côté de voies de recherches fructueuses.

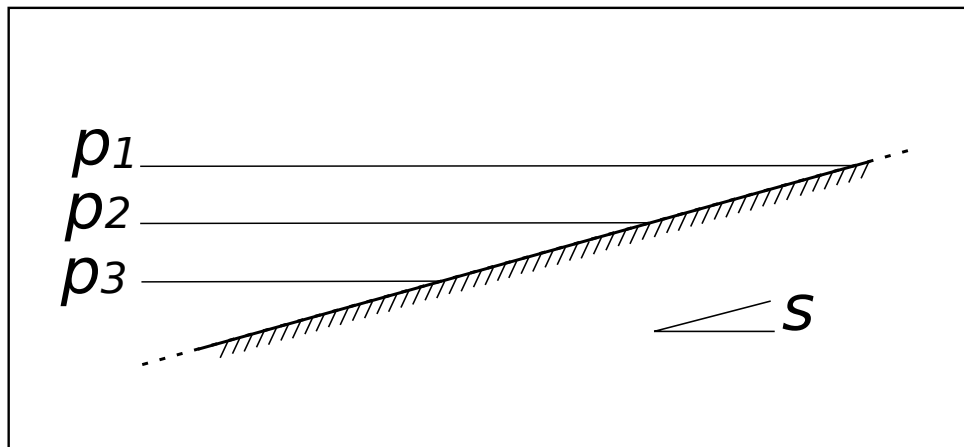


FIGURE 3.1 – Diagramme montrant la situation analysée dans Bénard et al., 2004. Le domaine horizontal selon x est infini, le terrain est incliné avec une pente constante s . La pression au sol tend vers des valeurs non-physiques (zéro et vers l'infini) aux bouts du domaine, qui sont situés à l'infini en x . Néanmoins la pression au sol n'intervient pas en elle-même dans l'analyse, seule la pente s intervient.

Une fois le cadre du système avec pente uniforme accepté malgré son étrangeté physique, on peut en dérouler mécaniquement les conséquences sans grande difficulté pour obtenir un système linéaire dont il est pertinent d'analyser la stabilité. Ce travail est exposé dans Bénard et al. (2005). La principale justification du cadre choisi est alors que même si l'orographie devient infinie, les termes sources liés à l'orographie ont en chaque point une amplitude normale, bornée, comme dans la nature. D'un point de vue mathématique, on a donc simplement ajouté quelques termes sources d'amplitude parfaitement réaliste dans le système linéaire. La nouveauté étant que dans certaines configurations, ces termes source ont des conséquences ravageuses pour la stabilité.

Ici encore, l'analyse montre le choix des variables pronostiques traitées implicitement peut avoir un impact majeur sur la stabilité du schéma SI dans les écoulements avec orographie. La raison, toujours identique, est que ce type de changement, bien qu'il ne change pas la structure du système linéaire en lui-même, entraîne une modification en profondeur de la structure (et de certains signes) des résidus qui sont traités explicitement. L'article montre par une analyse de type SHB78 qu'un passage de la variable d à la variable \mathbb{D} permet de stabiliser le schéma en présence d'une pente uniforme. Une expérience avec le modèle réel Aladin-NH sur un écoulement académique en plan vertical confirme cette stabilisation.

3.7 Mise en place d'un modèle de PN en équations d'Euler

En 2003, suite à de vifs débats internes, il fut décidé au niveau de la direction de Météo-France que le futur modèle de PN non-hydrostatique en aire limitée serait basé sur le noyau dynamique d'Aladin-NH, plutôt que sur celui de Méso-NH (modèle anélastique avec un schéma explicite). La date de fin 2008 fut alors retenue pour la mise en opération, le projet étant de coupler le noyau dynamique Aladin-NH avec une physique proche de celle de Méso-NH, jugée plus avancée et mieux adaptée à l'échelle fine que celle d'ARPEGE/Aladin utilisée jusqu'alors en opérationnel. Il fallait aussi mettre en place un système d'assimilation, le projet initial consistant sur ce point à étendre le 3D-VAR d'Aladin au nouveau modèle, en mettant l'accent sur l'utilisation de données radars, sur la fréquence du cycle d'assimilation, et l'utilisation de statistiques d'erreur d'ébauche adaptées aux nouvelles échelles résolues par le modèle. Le projet visant l'intégration de toutes ces composantes en un système capable de produire des Prévisions Numériques fut dénommé AROME (Application de la Recherche pour l'Opérationnel à Meso-Echelle), et ce nom fut ensuite donné au nouveau système de PN, ainsi qu'au modèle d'intégration temporelle en lui-même. La maille horizontale visée pour cette mise en opérationnel initiale était de 2.5 km.

Pour ce qui est du noyau dynamique, un candidat viable était disponible depuis peu en rassemblant les améliorations proposées dans les deux derniers articles (Bénard et al., 2004 ; et Bénard et al. 2005). Ces modifications permettaient de penser qu'il était possible de maintenir une stabilité et une qualité suffisante en présence des pentes rencontrées à ces résolutions et d'utiliser un simple schéma SI à deux pas de temps (2-TL) en semi-lagrangien. Vers 2004 ou 2005 un prototype du noyau dynamique exploitable sur cas réels en adiabatique (sans paramétrisations physique) était donc disponible. A cette époque j'ai donc effectué des tests intensifs pour cerner la robustesse réelle du noyau dynamique, en me concentrant sur des écoulements sévères (en adiabatique) au-dessus des Pyrénées, et en explorant les plages des paramètres conduisant aux meilleures caractéristiques de stabilité. A cette résolution (2.5 km), j'obtenais des intégrations toujours stables sur 24 heures pour des pas de temps de l'ordre de 300 secondes, c'est à dire des valeurs significativement plus grandes que celles qui étaient envisagées en pratique, pour le modèle complet avec physique (de l'ordre de 90 secondes). L'ensemble de ces validations a été publié dans Bénard et al., 2010. Le décalage important entre les expériences et cette date très tardive s'explique par ma réticence personnelle à présenter la publication de ces résultats si pour une raison quelconque le modèle échouait à passer l'étape de la mise en opérationnel avec succès, ceci afin d'éviter de prendre le risque d'un triomphalisme anticipé qui se serait trouvé ensuite injustifié à l'épreuve quotidienne des faits.

La direction du projet AROME en lui-même était assurée par François Bouttier. d'une part, l'évolution vers un outil opérationnel nécessitait l'inclusion des paramétrisations physiques de Méso-NH et l'écriture d'une interface physique/dynamique, ce qui fut réalisé par Yann Seity et Sylvie Malardel. D'autre part l'installation d'un système d'assimilation via l'adaptation du schéma d'assimilation 3D-VAR d'Aladin aux spécificités d'AROME fut l'œuvre de Pierre Brousseau principalement. Le système AROME délivra sa première prévision en opérationnel le 18 décembre 2008, soit 12 jours avant la date qui avait été planifiée plusieurs années auparavant, fait suffisamment rare pour être souligné, en ce qui concerne les projets internes à météo-France. Les caractéristiques du système pour cette mise en opérationnel ont été décrites dans Seity et al., 2011. Gwenaëlle Hello fut chargée des aspects plus institutionnels avec notamment le projet PARME pour assurer la sensibilisation et la formation des prévisionnistes à ce nouvel outil.

3.8 Conclusion

La stabilité des schémas semi-implicites avec solveurs directs dans les modèles spectraux a occupé un bon nombre d'années de ma carrière. Comme souvent en analyse numérique, c'est un domaine austère, mais gratifiant par le saut en efficacité qu'un succès peut amener, avec toutes les conséquences pratiques et même financières que cela implique. Ce qui est fascinant dans cette discipline, c'est qu'un obstacle apparemment infranchissable, conduisant à un modèle de prévision totalement instable, peut être levé par une modification impliquant seulement deux ou trois lignes de code. L'exemple évoqué dans la section 3.6.4, avec l'introduction de la température de référence T_{e*} alias SITRA pour rendre viable les schémas 2-TL en SI SL est une illustration frappante de cet aspect. Bien que l'article correspondant (Bénard, 2004) soit très court et relativement peu cité, je considère que c'est l'un des plus importants de ceux que j'ai écrits.

Le problème des schémas dits "à coefficient constants sur l'horizontale" abordés dans ce chapitre est que leur robustesse semble être mise en défaut au-delà d'une certaine pente de l'orographie. Ce problème

est réel et très concret, de sorte que lorsque l'on réduit la maille d'AROME aux alentours de 100 m, une proportion non-négligeable des prévisions de cas réels sur la France ne peuvent être menées jusqu'à leur terme (disons, jusqu'à 24 h d'échéance). L'origine véritable du problème, difficile à identifier n'est pas clairement certaine et pourrait être multiple (manque de représentation implicite de certains termes orographiques, faiblesse dans la discrétisation ou le choix des variables pronostiques, dans le choix de la coordonnée verticale "terrain-following",...). D'intéressantes perspectives semblent toutefois s'offrir avec les schémas implicites à coefficients non-constant sur l'horizontale, au travers des travaux de F. Voitus, L. Auger et Th. Burgot (voir le chapitre conclusions et perspectives).

Chapitre 4

PNT sur des ellipsoïdes de révolution

4.1 Introduction

Dès ses origines et jusqu'à maintenant, la PNT globale est basée sur des équations supposant que la terre (ainsi que son champ de gravité) est exactement sphérique, alors que dans la réalité, celle-ci est légèrement aplatie aux pôles. L'idée ou plutôt la suspicion que la qualité des prévisions numériques puisse pâtir de cette approximation de la réalité est apparue très tôt. White et al., 2008 (WSW08 ci-après), font remonter cette idée aux années 1970 (voir par exemple la référence Veronis, 1973 dans leur introduction). A l'époque, ces questionnements servaient surtout, au travers d'une estimation des erreurs via des analyses d'ordre de grandeur, à montrer que ces erreurs avaient un impact indétectable. Ces études initiales servaient donc uniquement à justifier l'hypothèse d'une Terre sphérique et ne débouchaient pas vers des solutions permettant de résoudre véritablement les équations sur autre chose que sur une sphère, car là n'était pas leur but. Pour cela il faut attendre les années 2000 (par exemple Gates, 2004, inspiré par Gill, 1982), en soulignant toutefois que les premières tentatives dans cette voie manquaient d'une compréhension suffisante pour approcher le problème de façon rigoureuse et pleinement rationnelle. Cette compréhension ne sera acquise que vers les années 2010, comme relaté dans ce chapitre.

Parallèlement à ces développements progressifs sur les équations à résoudre, on rencontre dans la littérature l'idée que les prévisions opérationnelles sont effectivement parvenues à un tel niveau de qualité que les erreurs liées à l'hypothèse sphérique pourraient très bientôt cesser d'atteindre un niveau indétectable dans les scores de prévision, d'autant plus qu'un aspect systématique est souvent attribué à ces erreurs. Le sujet devient progressivement plus brûlant donc, même s'il est pour l'instant difficile de dire quand (et si) l'abandon de l'hypothèse sphérique en PNT deviendra absolument incontournable pour continuer à progresser en qualité.

Ce chapitre traite de ces problématiques et de quelques autres questions connexes liées à la géométrie du domaine terrestre (valables aussi pour les planètes géantes du système solaire). Mes travaux sur le sujet sont publiés dans Bénard, 2014a (B14a ci-après) et Bénard 2014b (B14b ci-après). Ce chapitre se référera également à une série importante de papiers produit par l'équipe anglaise active dans cette thématique : outre WSW08, citons White et Wood, 2012 (WW12 ci-après) ; et finalement Staniforth et White, 2015a, 2015b.

L'erreur minimale commise en assimilant le niveau moyen de la mer à une sphère (environ ± 10 km) est beaucoup plus grande que celle commise en l'assimilant au meilleur ellipsoïde de révolution possible (environ ± 50 m). Nous admettrons dans tout ce chapitre que l'erreur de prévision liée à la géométrie sphérique adoptée jusqu'ici se trouve entièrement éliminée en se donnant la liberté de raisonner en termes d'ellipsoïdes au lieu de sphères (c'est-à-dire que les erreurs résiduelles liées à cet écart de ± 50 m seront négligées). Les premières tentatives pour exprimer les équations de la météorologie sur une planète ellipsoïdale faiblement aplatie aux pôles (pour laquelle nous employons le terme de sphéroïde) n'avaient pas bien saisi le problème dans son entier, alors que paradoxalement dans le champ de la géodésie et de l'astronomie gravitométrique, une compréhension correcte du problème était acquise depuis les travaux de Clairaut (1743) et Stokes (1849). Le problème du passage aux sphéroïdes comporte deux aspects d'égale importance pour la météorologie : la modification de la forme de la Terre, et la modification du champ de gravité. C'est ce second aspect qui était mal appréhendé dans les premières études au début des années 2000.

Précisons une bonne fois pour toute que nous parlerons ici du champ de gravité “apparente”, c’est à dire mesurée dans le référentiel lié à la sphère en rotation, et qui est donc la somme de la force d’attraction Newtonienne et de la force centrifuge liée à la rotation diurne planétaire :

$$\mathbf{g} = \mathbf{g}_N + \mathbf{g}_C. \quad (4.1)$$

Le champ de gravité apparente dérive d’un potentiel ϕ dénommé géopotential et on a donc avec des notations évidentes

$$\phi = \phi_N + \phi_C. \quad (4.2)$$

Les équations qui régissent les variations spatiales du géopotential dans l’espace situé à l’extérieur de la planète sont

$$\Delta\phi_N = 0, \quad (4.3)$$

$$\Delta\phi_C = 2\Omega^2, \quad (4.4)$$

où Ω est la vitesse angulaire de rotation diurne de la planète par rapport aux étoiles fixes.

Dans le cadre examiné ici, avec notamment la symétrie de révolution, la donnée du champ de gravité apparente se résume en fait à la définition d’une famille de courbes fermées symétriques et concentriques, représentant, dans un plan méridien, la famille des surface horizontales (toujours pour la gravité apparente), appelées “les isopotentielles” (ou “les géopotentielles” dans la littérature anglaise). Les lignes verticales sont alors les courbes qui sont orthogonales aux surfaces horizontales en chaque point, ce qui suffit à les définir entièrement et de façon univoque.

Afin de clarifier le débat, précisons ici qu’il existe deux conventions pour le géopotential et la valeur algébrique de la gravité notée g : une convention astronomique et une convention météorologique. Ici nous utilisons la convention météorologique qui veut que la valeur algébrique de la gravité g soit positive pour des attractions vers le centre O de la planète. Dans ce cas, si r est la longueur du rayon vecteur pris à partir de O , on peut définir conventionnellement les signes du géopotential et de la valeur algébrique de la gravité de telle sorte qu’au premier ordre :

$$\phi \approx +\Omega^2 \frac{\rho^2}{2} + \frac{C}{r}, \quad g = -\frac{\partial\phi}{\partial r}, \quad (4.5)$$

où ρ est la distance du point à l’axe des pôles et C une constante positive décrivant la gravité newtonienne. Dans le plan équatorial (où $\rho = r$), nous avons donc

$$\phi \approx +\Omega^2 \frac{r^2}{2} + \frac{C}{r}, \quad g = -\Omega^2 r + \frac{C}{r^2}. \quad (4.6)$$

La force centrifuge est bien répulsive, tandis que la force newtonienne est bien attractive. Il existe un rayon pour lequel l’attraction totale (i.e. la gravité apparente dans le référentiel tournant) s’annule ; c’est donc le rayon de l’orbite géostationnaire. Sur l’axe des pôles (où $\rho = 0$), on a simplement

$$\phi \approx +\frac{C}{r}, \quad g = +\frac{C}{r^2}. \quad (4.7)$$

Sur cet axe, la gravité apparente est toujours une attraction purement newtonienne, comme attendu. Cette convention ‘météorologique’ est adoptée dans tous les papiers de la série (2008-2015) citée ci-dessus, et ici même ; il est important de le préciser une bonne fois pour toute. On peut vérifier que cette convention est adoptée en se reportant à l’Eq. (4) de WSW08, l’Eq. (23) de B14a, et l’Eq. (2) de Staniforth and White, 2015b, en ne retenant dans chacune de ces équations la partie en r^{-1} et la partie en r^2 . Dans tous les article modernes traitant du sujet (dans la série entre 2008 et 2015), le paramètre astronomique Ω est remplacé par un facteur non-dimensionnel m défini par

$$m = \frac{\Omega^2 a^3}{\gamma M}, \quad (4.8)$$

où γ est la constante de la gravitation universelle, M est la masse de la planète et a son grand axe. Pour la Terre, les valeurs des paramètres astronomiques sont tels que $m \approx 0.0018$.

Dans l'approximation dite "approximation de géopotential sphérique" (SGA en anglais) la Terre est considérée comme une sphère et le champ de gravité apparente est supposé uniforme sur chaque surface horizontale. En conséquence, les surfaces isopotentielles sont donc des sphères concentriques. Dans l'approximation dite "approximation de géopotential ellipsoïdal" (EGA en anglais), on se donne la forme de la Terre comme un ellipsoïde, et toutes les isopotentielles sont aussi supposées elles aussi avoir une forme ellipsoïdale, ce qui n'est pas exactement le cas dans la nature, même si on se restreint au cas où la planète en rotation est elle-même exactement ellipsoïdale. A strictement parler ce cadre EGA constitue donc bien une approximation (mais bien meilleure que l'approximation SGA).

A ce stade de l'exposé, il est nécessaire d'introduire un théorème important, permettant de mesurer le degré de réalisme du champ de géopotential que l'on doit choisir pour une planète en rotation lorsque l'on souhaite utiliser l'approximation EGA.

Théorème de Clairaut (1743) : Considérons une planète 'réelle', nous entendons par là une planète constituée de couches de densité uniforme superposées les unes sur les autres, et telles que les surfaces délimitant ces couches sont toutes horizontales. Supposons que cette planète est animée d'une faible vitesse de rotation diurne, elle est donc faiblement aplatie et possède au premier ordre une ellipticité $\epsilon = (a - c)/a$, où a est son grand axe, et c son petit axe. Alors la gravité de surface au pôle g_{s_P} et la gravité de surface à l'équateur g_{s_E} vérifient (au premier ordre) :

$$\frac{g_{s_P}}{g_{s_E}} = 1 - \epsilon + 2\epsilon_H, \quad (4.9)$$

où ϵ_H est l'ellipticité qu'aurait le sphéroïde en rotation de même masse s'il avait une densité homogène, égale à la densité moyenne de la planète. Dans la littérature récente on utilise le paramètre m défini ci-dessus plutôt que cette ellipticité ϵ_H , les deux étant liés par $\epsilon_H = 5m/4$ et le Théorème s'exprime alors par :

$$\frac{g_{s_P}}{g_{s_E}} = 1 - \epsilon + (5/2)m. \quad (4.10)$$

Pour clore cette introduction, abordons une autre approximation, qui sera discutée dans ce chapitre. Cette approximation est dite 'd'atmosphère mince' (Shallow Atmosphere ou SA en anglais), par opposition à la représentation plus exacte d'une atmosphère épaisse (Deep Atmosphere ou DA en anglais). Pour rester simple à ce stade, disons que l'approximation SA consiste à considérer que les lignes verticales sont des droites et que la métrique du système de coordonnées curvilignes est indépendante de l'altitude (aucune de ces deux propriétés n'est en général valide dans une atmosphère épaisse). Généralement, on associe en un paquet unique cette approximation géométrique avec celle qui consiste en SA à supposer que la gravité g ne diminue pas avec l'altitude, tandis qu'en DA elle diminue conformément à la loi de Newton c'est à dire proportionnellement à l'inverse du carré de la distance au centre. L'écriture des équations d'évolution est relativement aisée dans le cadre de l'approximation SGA, que ce soit en atmosphère mince ou profonde. Mais lorsqu'il s'agit de s'affranchir de l'approximation SGA pour passer à l'approximation EGA la tâche devient, comparativement, beaucoup plus complexe pour une atmosphère profonde que pour une atmosphère mince. La plupart des modèles météorologiques actuels utilisent à la fois l'approximation SGA et l'approximation SA et se placent donc dans la cadre le plus simple possible pour faire de la prévision globale. Les modèles opérationnels anglais toutefois font exception, puisqu'ils utilisent depuis plusieurs décennies l'approximation SGA mais dans une atmosphère profonde (DA), avec une gravité g décroissant avec l'altitude (mais sans avoir jamais réussi à démontrer le bénéfice concret de cette approche).

4.2 Géométrie ellipsoïdale en atmosphère profonde

Dans cette section sont abordés les aspects liés au passage de l'approximation SGA à l'approximation EGA en atmosphère profonde (DA), c'est à dire la combinaison la plus complexe parmi toutes celles évoquées dans les paragraphes précédents. C'est paradoxalement cette combinaison –la plus complexe donc– qui a été examinée en premier dans la littérature météorologique dans les études pionnières (Gates,

2004), ce qui fut une source d'échec, alors que le passage en EGA dans une atmosphère mince aurait été plus simple. Pourquoi avoir voulu d'emblée traiter le cas le plus complexe en premier, surtout si c'est sans succès au début ? C'est essentiellement ce qui va être discuté dans cette section.

Mon interprétation tient aux doutes qu'il y avait l'époque quant à l'impact relatif de chacune des approximations en jeu, prise individuellement. Traiter la Terre comme une sphère, négliger les variations de la métrique avec l'altitude, et prendre g constant verticalement induisait trois sources d'erreurs, mais comme on ignorait laquelle était la plus importante, et si une prise en compte partielle était dynamiquement consistante, on avait sans doute décidé de les éliminer toutes les trois d'un seul coup, dans un effort certes important, mais du moins, unique, et forcément cohérent. Car, une fois encore, à cette époque, la nécessité d'une certaine cohérence était entrevue, mais on était à peu près incapable de statuer clairement sur celle-ci, et d'obtenir un consensus clair, faute de concepts théoriques adéquats.

Il y avait donc une grande incertitude sur la 'consistance dynamique' liée au fait d'appliquer une ou plusieurs approximations séparément. Au prix d'un long travail de consensus dans les années 1950-60, il avait été accepté dans la communauté météorologique que l'ensemble d'approximation de départ (SGA, SA, g constant) couplé avec le fait de négliger certains termes de Coriolis offrait un cadre 'consistant', mais reprendre cette tâche difficile dans le cadre encore plus compliqué de l'EGA devait alors paraître largement dissuasif.

Incidentement, il est intéressant de noter que dans les années 1950-60, l'approximation SGA faisait consensus sans trop de problème, ainsi que celle de g constant (en latitude et en altitude), mais les problèmes de consensus étaient liés à l'approximation SA couplée avec des interrogations sur l'expression des termes de Coriolis, aspects mal maîtrisés au début. C'est Phillips (1966) qui clôt le débat en proposant une approche rationnelle basée sur deux raisonnements un peu séparés : (i) l'existence d'invariants possédant une forme 'logique', et (ii) une dérivation rationnelle des équations originales (sous leur forme vector-invariant), vers leur forme exploitable, projetée dans un système de coordonnées curviligne *défini uniquement par sa métrique* et non plus exprimé 'artisanalement'. Cette démarche rationnelle, utilisant uniquement la métrique pour dériver des équations scalaires à partir des équations vectorielles, sera reprise par Dubos (2018) pour l'approximation EGA, dans un article non publié.

Rétrospectivement, il apparaît que les difficultés initiales rencontrées autour de ce couple de questions (SA / Coriolis) étaient liées au fait que fondamentalement, il n'existe pas, dans notre espace Euclidien, de système de coordonnées curvilignes possédant la métrique impliquée par l'approximation SA. Puisque les invariants contenus dans les équations de base (vectorielles) s'expriment eux-mêmes par des opérateurs de calcul vectoriel, il est logique qu'il soit fondamental d'exprimer correctement ces opérateurs à partir de la métrique adoptée pour que les invariants subsistent. C'est bien ce que souligne Phillips en dérivant son équation (10) qui pose la forme exacte de l'opérateur 'rot' dans cette coordonnée (d'un espace non-euclidien), ce qui conduit directement à l'équation (11), clôturant une bonne fois pour toutes, en une ligne dans un journal, ce long débat.

Pour en revenir, après cette digression, au sujet qui nous occupe ici c'est à dire l'abandon de l'approximation SGA, on pensait donc que relaxer les trois approximations initiales d'un seul coup (SGA, g constant, et SA/Coriolis) devait nécessairement offrir un cadre consistant puisque c'était un cadre sans approximation, entièrement conforme à la physique classique newtonienne dans un univers euclidien.

En revanche, on pensait que relaxer uniquement certaines approximations tout en maintenant certaines autres comportait un risque pour la consistance. Risque confus, assez mal défini, et difficile à évaluer à l'époque, car ce sont là des problèmes complexes sur lesquels les méthodes de réflexion n'étaient pas mûres, et sur lequel il était à peu près impossible d'obtenir un consensus et une vision claire. Ce n'est qu'avec le développement d'un mode de pensée basé sur les concepts de la mécanique hamiltonienne appliquée à l'hydrodynamique (e.g. Morisson , 1988 ; Salmon 1990) que la consistance de telle ou telle approximation météorologique aurait pu être clarifié et démontré. Mais ces travaux n'avaient pas encore significativement percolé en Météorologie.

Il s'avère cependant que ce concept d'un "effort unique" pour passer de l'approximation SGA-SA (et g indépendant de l'altitude) vers le cadre EGA-DA (et g variable), initialement privilégié, était trop difficile à manier, et les résultats obtenus alors étaient soit suspects soit faux, ce qui ne constituait donc pas une avancée. En effet remplacer les erreurs liées à un cadre sphérique mais consistant, par des erreurs mal comprises dans un cadre ellipsoïdal, pouvait être vu davantage comme un choix hasardeux que comme une avancée. En particulier, le système proposé par Gates, 2004 est non opérant, notamment par manque de réalisme, la gravité étant plus faible au pôle qu'à l'équateur. Face à ce constat d'échec, on assiste alors à une sorte de lassitude vers 2005-2010, personne n'y comprenant plus grand chose...

Ce sont les équipes anglaises qui vont débloquent la situation dans des articles comme White et al., 2005 (dans le cadre SGA) et White and Wood, 2012 (dans le cadre EGA), en proposant des combinaisons partielles et en justifiant plus solidement leur ‘consistance physique’, même si là encore quelques erreurs subsistaient.

Il est à peu près clair que des idées fausses circulèrent longtemps sur toutes ces thématiques, mais même sur ce point il n’y a pas consensus complet, et –pire encore– la signification même de la notion de ‘consistance dynamique’ est encore âprement débattue entre spécialistes ! On peut se reporter à Tort et Dubos, 2014a, 2014b, d’une part et Charron et Zadra, 2014, Charron et al. 2015 d’autre part, tous ces papiers étant difficiles à lire pour le profane. Le problème est donc épineux, et l’idée initiale de l’effort unique reposait sans doute sur le constat qu’il était plus simple de résoudre les équations exactes que d’utiliser un ensemble partiel d’approximations dont on ne savait pas d’une part s’il permettait d’éliminer vraiment les plus grosses erreurs dont on cherchait à s’affranchir, et d’autre part s’il avait un sens tout simplement. Le problème qui allait apparaître en relâchant toutes les approximations à la fois est que cela n’est pas si simple qu’initialement espéré.

Là-dessus vient se greffer le problème du choix d’un système de coordonnées pour décrire l’espace et exprimer concrètement les différents opérateurs de calcul vectoriel inhérents à la météorologie. Par simplicité on décide toujours de choisir des coordonnées curvilignes (analogues à longitude, latitude, et altitude) telles que leur iso-surfaces correspondent aux isopotentielles et aux verticales. Cela possède aussi l’avantage qu’alors la force de gravité se projette uniquement selon la direction verticale. Cet avantage, purement technique et contingent sur le papier, peut néanmoins devenir important lors de la discrétisation, et la plupart des auteurs s’accordent pour reconnaître qu’une gravité ayant une projection non nulle sur les composantes ‘horizontales’ du système de coordonnées n’est pas une bonne idée pour un modèle de prévision numérique.

Enfin, un dernier problème avec l’approximation EGA est que généralement on n’est pas capable de fournir une forme algébrique pour décrire les lignes verticales et, par conséquent, pour exprimer les formules de transformations entre les coordonnées euclidiennes naturelles et les coordonnées curvilignes de travail. Plusieurs solutions seront proposées pour pallier ce problème, et finalement dans le système proposé par Staniforth et White, 2015b, cet obstacle parvient à être levé en grande partie.

La Figure 4.1 illustre quelques exemples de familles d’isopotentielles pour une atmosphère épaisse. Dans l’approximation sphérique (Fig. a), on suppose que la terre et toutes les isopotentielles sont des sphères concentriques. Les verticales sont donc des lignes droites. Dans l’approximation sphéroïdale (Figs. b, c, d) on suppose que les isopotentielles dans le plan méridien sont des ellipses concentriques, et tout le problème est alors de choisir comment varie leur aplatissement en fonction de l’altitude, ce qui détermine le choix d’une approximation du champ de gravité. Dans la première tentative rapportée dans la littérature (Gates, 2004) les isopotentielles sont supposées être des ellipses confocales (Fig. b). La gravité *décroit* du pôle vers l’équateur, ce qui, au regard du théorème de Clairaut et des observations est irréaliste. On note également que la courbure des verticales est dirigée vers l’équateur, ce qui n’est pas réaliste non plus. White et al., 2008, proposent d’utiliser une famille de sphéroïdes similaires (d’ellipticité constante) pour définir le champ de géopotentiel (Fig. c). Cela représente un progrès indéniable, la courbure des verticales est orientée vers le pôle et la gravité de surface est plus grande au pôle (g_{sP}) qu’à l’équateur (g_{sE}). Ces deux caractéristiques sont toutes deux en accord avec la réalité. Néanmoins, le rapport $(g_{sP} - g_{sE})/g_{sE}$ n’a pas la valeur imposée par le théorème de Clairaut, le déficit étant de 30 % environ. Des solutions permettant d’atteindre plus de réalisme de ce point de vue ont été proposées par Bénard, 2104a, et Staniforth et White, 2015b, cette dernière proposition possédant l’avantage d’être plus algébrique et moins approchée. La Fig. d illustre la forme des isopotentielles dans ces deux cas (pour lesquels les différences graphiques sont minimales).

Dans cette thématique de l’approximation EGA en atmosphère épaisse, ma contribution a été de proposer une première solution réaliste (B14a), mais il faut reconnaître que la solution proposée peu après par Staniforth et White, 2015b est bien meilleure (moins approchée, plus analytique) et beaucoup plus élégante. Afin de garder la tête haute, on pourra donc dire que l’inélégance relative de B14a a pu au moins servir d’aiguillon pour susciter Staniforth et White, 2015b...

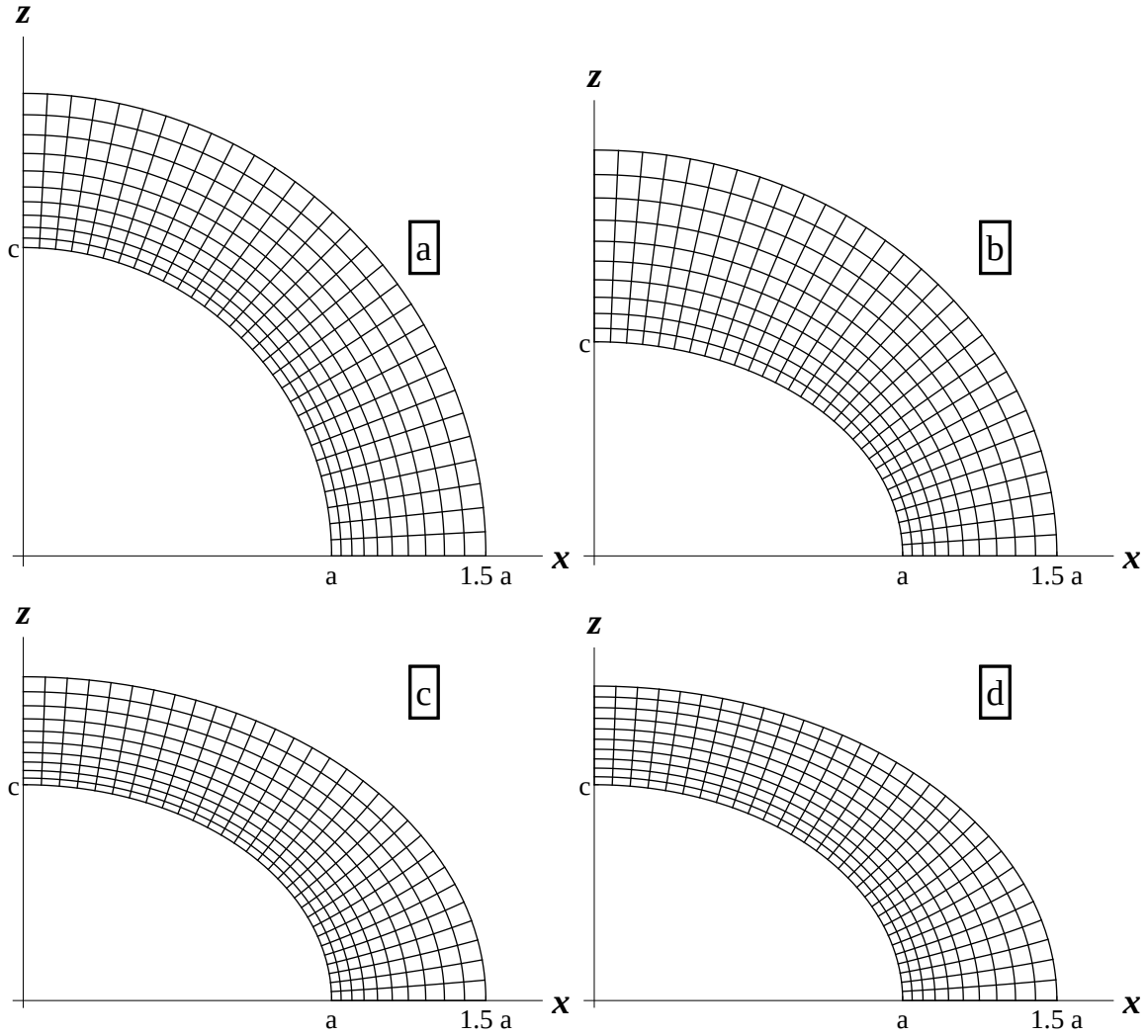


FIGURE 4.1 – Quatre possibilités de géométries pour la forme de la terre et la forme du champ de géopotentiel dans le cadre d'une atmosphère épaisse. La gravité en un lieu donné est inversement proportionnelle à la longueur du côté vertical du petit rectangle curviligne en ce lieu. Les quatre graphiques sont construits en stipulant que la gravité décroît avec la distance au centre selon la loi de Newton. Noter que l'aplatissement de la planète dans les Figs. b–d a été exagérée pour la clarté des figures. De même, le domaine vertical représenté (jusqu'à 1.5 fois le rayon terrestre) est exagéré par rapport à l'extension réelle de l'atmosphère, pour la clarté du propos.

Fig. a : géométrie sphérique : la Terre est les isopotentielle sont des cercles. La gravité au sol est uniforme du pôle à l'équateur. Fig. b : géométrie "Confocal oblate spheroids" comme dans Gates 2004. Les sphéroïdes sont confocaux. La gravité de surface est plus intense à l'équateur qu'au pôle, ce qui n'est pas réaliste. Fig. c : géométrie "Similar oblate spheroids" comme dans White et al. 2008. Les sphéroïdes sont d'ellipticité constante. La gravité de surface est plus intense au pôle qu'à l'équateur, ce qui est réaliste, mais le rapport entre ces deux valeurs de la gravité n'est pas réaliste. Fig. d : une géométrie réaliste. La gravité de surface est plus intense au pôle qu'à l'équateur, et le rapport entre ces deux valeurs de la gravité est conforme à la valeur théorique donnée par Clairaut, donc réaliste.

4.3 Géométrie ellipsoïdale en atmosphère mince

Le problème de l'approximation EGA en atmosphère mince est beaucoup plus simple qu'en atmosphère épaisse. Notamment le choix d'un système de coordonnées curvilignes ne pose alors pas de problème, et la métrique est simple, connue, et directement exprimable analytiquement. La dérivation mathématique des équations d'évolution dans ce système de coordonnées curviligne ne pose pas de problème particulier. De

plus, l'idée que l'approximation d'atmosphère mince (SA) produit moins d'erreurs que l'approximation SGA justifie la pertinence ce choix, comme discuté au tout début de la section 7 de B14b.

4.3.1 Dissocier erreurs géométriques et gravitationnelles

Dans B14b, je proposai d'évaluer pour la première fois de façon concrète les erreurs de prévision liées à l'approximation SGA, dans un modèle en eau peu profonde (shallow-water, SW en anglais). Il est très simple de développer un modèle SW qui puisse indifféremment être intégré avec l'approximation SGA ou l'approximation EGA, toutes choses égales par ailleurs, ce qui permet une estimation rigoureuse des erreurs liées spécifiquement à l'approximation SGA. Néanmoins, il y avait une subtilité que je mettais en avant dans ce papier. Dans l'introduction du présent chapitre ci-dessus, j'ai souligné que le passage de l'approximation SGA à l'EGA élimine deux types d'erreurs : l'erreur géométrique et l'erreur de gravité. Dans l'approximation SGA, d'une part le parallèle de Paris (par exemple) n'a pas la même circonférence que dans la réalité, ce qui crée une erreur de prévision de nature géométrique ; et d'autre part, la gravité de surface est uniforme, ce qui crée une autre erreur, de nature physique.

La question n'ayant pas été étudiée en détail, il était plus ou moins admis que ces deux erreurs n'étaient pas dissociables. Mon premier souhait était de montrer qu'il était tout à fait possible au contraire d'imaginer des expériences permettant de dissocier ces deux sources d'erreur individuellement, afin, notamment de pouvoir en estimer l'importance relative. Pour cela, une relecture soigneuse des travaux de Stokes, 1849 avec un travail d'interprétation de leurs conséquences était nécessaire.

4.3.2 Planètes 'réelles' et planètes 'forgées'

Rappelons auparavant qu'on appelle ici une planète 'réelle' une planète constituée de couches de densité uniforme superposées les unes sur les autres, et telles que les surfaces délimitant ces couches sont toutes horizontales, c'est-à-dire isopotentielle. Toutes les planètes du système solaire et même toutes les étoiles actives satisfont à cette propriété. La raison en est simple, tous ces corps célestes sont purement et simplement constitués de fluides de densité variable, superposées de telle sorte que les plus denses se retrouvent au centre, et que chaque couche est bien délimitée par deux isopotentielle. La partie solide des planètes telluriques est une couche si fine en flottaison sur le magma en fusion qu'elle n'intervient pour ainsi dire pas dans le problème.

Néanmoins on peut très bien imaginer, d'un point de vue purement conceptuel, qu'il en va tout autrement, et que les parties internes possédant une même densité ne sont pas délimitées par des isopotentielle. Par exemple, on peut imaginer une planète entièrement solidifiée, dans laquelle des grandes cavités seraient creusées, ou au contraire, contenant enfermées çà et là des poches de matière très dense. Une telle configuration ne peut pas vraiment survenir en pratique pour une planète de taille similaire à la Terre puisque la cohésion mécanique des roches solides connues n'est pas suffisante pour que cela puisse être possible en dehors d'accidents d'une taille très limitée comparée à celle de la planète (les contraintes mécaniques seraient énormes et briseraient la planète). De plus, l'immense majorité de la partie interne de notre planète est fluide, interdisant purement et simplement ce genre d'accidents de densité. Mais cela reste possible pour des expériences conceptuelles, et on peut donc s'autoriser à considérer une planète en rotation dont la forme de la surface, bien qu'isopotentielle, n'obéisse pas à la forme naturelle requise pour une planète 'réelle'. Nous parlerons alors de planète 'forgée' (ici le sens du mot 'forgé' doit être pris comme celui de l'expression 'forgé de toute pièce').

Pour illustrer cette idée, on peut très bien imaginer une planète en rotation, mais dont la surface, bien qu'isopotentielle, serait purement sphérique. Pour cela, il faudrait nécessairement qu'il y ait une accumulation de matière dense localisée le long de l'axe des pôles (ce qui ne peut pas être le cas pour une planète 'réelle'). A l'opposé, on peut très bien imaginer une planète qui ne tourne pas, mais dont la surface, bien qu'isopotentielle, soit un sphéroïde aplati aux pôles. Pour cela, il faudrait plutôt une accumulation de matière dense dans le plan équatorial. Bref, dès qu'on abandonne l'hypothèse d'une planète 'réelle', la forme de la surface (isopotentielle) n'est plus directement contrainte par la composition matérielle et le taux de rotation de la planète.

A titre d'illustration concrète, je présente ci-dessous dans l'Appendice 2 (en anglais) le cas simplifié de deux masses ponctuelles située sur l'axe polaire. Ce cas montre qu'avec une vitesse angulaire idoine, on peut obtenir une planète dont la surface est isopotentielle avec un rayon polaire c égal au rayon équatorial a , et une gravité de surface non-uniforme, dont l'intensité présente le même ratio (g_s^P/g_s^E) entre le pôle

et l'équateur que la gravité terrestre réelle. Symétriquement, je montre qu'avec une distribution de masse le long d'un cercle sur le plan équatorial, on peut avoir une planète dont la surface est isopotentielle d'aplatissement c/a égal à celui de la Terre, mais une gravité de surface telle que $(g_s^P/g_s^E) = 1$. Ces deux caractéristiques ne sont possibles qu'avec des distributions de masse non-naturelles.

4.3.3 Vitesse angulaire centrifuge et de Coriolis

Un autre fait notable doit être abordé à ce stade. Lorsqu'on écrit les équations de la dynamique dans un référentiel tournant avec la vitesse angulaire Ω , deux types d'accélération liées à la rotation du référentiel apparaissent : celle qui affecte un point immobile dans le référentiel tournant (c'est l'accélération centrifuge, toujours proportionnelle à Ω^2) et celle qui n'affecte qu'un point en mouvement dans le référentiel (c'est l'accélération de Coriolis, toujours proportionnelle à Ω). Si on écrit les équations du mouvement disons dans les coordonnées cartésiennes canoniques, les termes en Ω^2 et en Ω vont apparaître en des endroits différents des équations du mouvement, via l'apparition de termes correspondants à ces deux types d'accélération. Il doit pouvoir être démontré (mais nous ne le ferons pas ici) que dans tous les cas et toutes les géométries le système reste 'physiquement consistant' si on assigne deux valeurs différentes de Ω à ces deux types de termes.

Avec cette distinction, des termes en Ω_{CEN}^2 et des termes en Ω_{COR} apparaissent, les valeurs numériques de ces deux vitesses angulaires n'étant plus nécessairement égales. L'énergie totale, correspondant à un premier invariant, fera apparaître Ω_{CEN}^2 (dans l'énergie potentielle), tandis que le tourbillon potentiel, autre invariant, fera apparaître Ω_{COR} , présent dans l'expression du tourbillon absolu.

A titre d'illustration, considérons un modèle shallow-water (SW) en géométrie EGA, pour lequel les équations d'évolution sont écrites dans un système de coordonnées elliptiques de type longitude, latitude (λ, φ) , φ étant la latitude géodésique. On arrive donc au système (18)–(20) de B14b, retranscrit ici :

$$\begin{aligned} \frac{\partial u}{\partial t} = & -\frac{uR}{a^2 \cos \varphi} \frac{\partial u}{\partial \lambda} - \frac{vR^3}{a^2 c^2} \frac{\partial u}{\partial \varphi} + 2\Omega_{\text{COR}} \sin \varphi v \\ & - \frac{R}{a^2 \cos \varphi} g \frac{\partial H}{\partial \lambda} + uv \frac{R \tan \varphi}{a^2} \end{aligned} \quad (4.11)$$

$$\begin{aligned} \frac{\partial v}{\partial t} = & -\frac{uR}{a^2 \cos \varphi} \frac{\partial v}{\partial \lambda} - \frac{vR^3}{a^2 c^2} \frac{\partial v}{\partial \varphi} - 2\Omega_{\text{COR}} \sin \varphi u \\ & - \frac{R^3}{a^2 c^2} \frac{\partial (gH)}{\partial \varphi} - u^2 \frac{R \tan \varphi}{a^2} \end{aligned} \quad (4.12)$$

$$\begin{aligned} \frac{\partial H}{\partial t} = & -\frac{uR}{a^2 \cos \varphi} \frac{\partial H}{\partial \lambda} - \frac{vR^3}{a^2 c^2} \frac{\partial H}{\partial \varphi} \\ & - \frac{R}{a} H \left[\frac{1}{a \cos \varphi} \frac{\partial u}{\partial \lambda} + \frac{R^2}{a c^2} \frac{\partial v}{\partial \varphi} - \frac{v}{a} \tan \varphi \right], \end{aligned} \quad (4.13)$$

où $R(\varphi) = \sqrt{a^2 \cos^2 \varphi + c^2 \sin^2 \varphi}$.

Conformément à la tradition, on peut dénoter par $f = 2\Omega_{\text{COR}} \sin \varphi$ le facteur de Coriolis. Pour que le système ci-dessus soit fermé, il faut évidemment spécifier la variation de g en fonction de φ , c'est-à-dire (par exemple) les équations (21)–(23) de B14b, retranscrites ici :

$$g(\varphi) = \frac{g(0) R(\varphi) \sqrt{R(\varphi)^2 - 2\epsilon (a^2 \cos^2 \varphi)}}{\sqrt{1 - 2\epsilon a^2 \cos^2 \varphi + (1 + \epsilon - 5m/2) c^2 \sin^2 \varphi}}, \quad (4.14)$$

$$g(0) = \gamma M (1 + \epsilon - 3m/2) / a^2, \quad (4.15)$$

$$m = \Omega_{\text{CEN}}^2 a^3 / \gamma M, \quad (4.16)$$

où Ω_{CEN} n'apparaît que par son carré (via le paramètre m). Si, comme expliqué ci-dessus, on accepte d'affecter deux valeurs distinctes pour Ω_{CEN} et Ω_{COR} pour les forces gravitationnelles et les forces de Coriolis, on voit donc que ces forces s'exprimeront au moyen des paramètres suivants :

$$g = g(\varphi, \Omega_{\text{CEN}}^2), \quad (4.17)$$

$$f = f(\varphi, \Omega_{\text{COR}}). \quad (4.18)$$

On pressent donc que ces deux fonctions peuvent en fait être spécifiées séparément. Un argument plus précis, basé sur une expérience de pensée est développé dans l'Appendice 2, pour illustrer le problème, dans la géométrie SGA.

4.3.4 Un quartet de représentations de planètes en rotations

On considère ici des planètes en rotation, avec une vitesse angulaire Ω_{CEN} , la vitesse angulaire Ω_{COR} étant laissée libre. Les travaux de Stokes ont montré que le champ de géopotential (et donc de gravité) autour d'une planète en rotation (réelle ou forgée) est entièrement défini pour peu que l'on se donne une surface fermée entourant toute sa matière pesante, et la valeur du géopotential en chaque point de cette surface. En corollaire, le champ de géopotential autour d'une planète en rotation dont la surface est isopotentielle est entièrement déterminé dès que l'on se donne la forme de sa surface, et la valeur (uniforme) du géopotential sur cette surface. On notera cependant qu'alors les variations géographiques de la gravité à la surface ne constituent plus un degré de liberté : seule la valeur (uniforme) du géopotential sur la surface est libre (cette donnée se réduisant en fait à une condition de jauge peu intéressante en soi, physiquement). Rappelons que ϕ_N , ϕ_C et ϕ désignent respectivement le potentiel newtonien, centrifuge et total. Les équations (4.3)–(4.4) doivent être vérifiées à l'extérieur de la planète :

$$\Delta\phi_N = 0, \quad (4.19)$$

$$\Delta\phi = \Delta\phi_C = 2\Omega_{\text{CEN}}^2. \quad (4.20)$$

où Δ est l'opérateur laplacien. Nous en arrivons donc (voir plus de précisions dans B14a, B14b) à proposer quatre situations, qui seront pertinentes pour parvenir à dissocier quantitativement les erreurs géométriques et les erreurs de gravité liées à l'approximation SGA :

- (i) Une planète sphérique avec gravité de surface uniforme. C'est la situation postulée dans l'approximation SGA. Cela peut être obtenu en spécifiant $\epsilon = 0$ et $\Omega_{\text{CEN}} = 0$ c.-à-d. $m = 0$ (tout en conservant la valeur terrestre pour Ω_{COR})¹.
- (ii) Une planète sphérique avec une variation méridienne de la gravité de surface conforme aux observations terrestres. Ceci est obtenu pour $\epsilon = 0$ et $m = (2/5)(g_{sP}/g_{sE} - 1)$.
- (iii) Une planète sphéroïdale possédant l'ellipticité ϵ de la Terre mais une gravité de surface uniforme, ce qui est obtenu en prenant $m = (2/5)\epsilon$.
- (iv) Une planète sphéroïdale possédant l'ellipticité ϵ de la Terre et une variation méridienne de la gravité de surface conforme aux observations terrestres. Ceci est obtenu en prenant pour m la valeur terrestre observée. C'est donc la situation pleinement réaliste au regard de l'approximation EGA.

La comparaison des prévisions (ii) et (i) donne une estimation des erreurs de gravité résultant de l'approximation SGA, et la comparaison des prévisions (iii) et (i) donne une estimation des erreurs géométriques résultant de l'approximation SGA. Et finalement, la comparaison entre les prévisions (i) et (iv) donne l'estimation de l'erreur totale commise en appliquant l'approximation SGA.

Faire accepter cette idée d'une séparabilité des erreurs et le bien-fondé de ces quatre variantes (dans B14a, B14b) fut un combat difficile, les réticences provenant de l'équipe anglaise qui seule publiait sur le même sujet à cette époque, et qui donc révisait mes manuscrits. Guidés par certains de leurs travaux sur la 'consistance physique', les critiques face aux variantes (ii) et (iii) portaient sur l'idée (fausse) que ces variantes étaient supposées créer artificiellement des sources de tourbillon, ce qui n'est pas le cas. On trouve d'ailleurs une trace explicite de cette idée fausse dans le papier WW12, au quatrième paragraphe de la section 6, dont l'encre était encore fraîche lorsque je soumettais mes manuscrits. Inutile de dire que puisque ces deux variantes contredisaient frontalement les idées récemment exprimées dans WW12, elles furent l'objet de commentaires 'majeurs' pour mes manuscrits. Pour faire accepter la légitimité de ces variantes (ii) et (iii), il me fallut donc batailler âprement. Dans une longue réponse aux reviewers, il

1. Noter que lorsque $\Omega_{\text{CEN}} \neq 0$, un géopotential de la forme $\phi(r) = \Omega_{\text{CEN}}^2 r^2/3 + C r^{-1}$ avec $C > 0$ répond mathématiquement à la question, c'est-à-dire qu'il assure une gravité de surface g_S uniforme tout en satisfaisant (4.19)–(4.20). Mais la gravité correspondante, $g(r) = -2\Omega_{\text{CEN}}^2 (r - a)/3 + g_S (a/r)^2$, n'est pas physique puisqu'elle augmente avec le rayon r au lieu de tendre vers zéro pour r tendant vers l'infini, notamment au-dessus des pôles (ici a est le rayon terrestre). Cette solution non physique –aussitôt que g varie avec l'altitude– n'est pas retenue.

fut nécessaire pour étayer le propos, de revenir en détail sur les conséquences pratiques des travaux de Stokes (pas assez reconnus et compris des météorologues), les reformuler en un corollaire plus intelligible, tel qu'énoncé ci-dessus, puis soigneusement dépeindre l'expérience de pensée consistant à imaginer les planètes 'forgées' évoquées ci-dessus, pour clamer que si de telles planètes venaient à posséder une atmosphère, celle-ci devait être parfaitement stable, évidemment exempte de sources de tourbillon, et que les météorologues habitant ces planètes étaient donc tout à fait fondés à s'adonner à leur science favorite, c'est à dire la météorologie sur les planètes (ii) et (iii).

Donc, il est tout à fait légitime d'imaginer une planète en rotation exactement sphérique, dont le sol est partout exactement horizontal, mais dont la gravité de surface n'est pas uniforme. Bien qu'une telle planète ne puisse en aucun cas être 'réelle' (elle doit être 'forgée'), il est tout à fait légitime d'y faire de la météorologie. Cet argument très concret, acheva finalement de rallier les détracteurs qui s'étaient montrés sceptiques à plusieurs reprises durant le processus de révision des manuscrits. D'ailleurs, peu après, le bien-fondé de ce quartet sera pleinement reconnu, et discuté sous un éclairage un peu différent, dans Staniforth et White, 2015a (section 4.6). Dans cette discussion, les auteurs présentent la variante (ii) comme une forme d'approximation EGA dans laquelle une et une seule isopotentielle se trouve être exactement sphérique, et dans laquelle 'par chance' c'est cette isopotentielle-ci qui coïncide avec la surface de la planète en rotation. Cette vision est tout à fait exacte et peut aider didactiquement. Le but de cette section dans Staniforth et White, 2015a est donc de retravailler cette nouvelle donnée, peu naturelle à leurs yeux initialement, afin de l'assimiler, la comprendre, et l'intégrer dans leur modèle de pensée en tant que météorologues, avec aussi un peu de didactique.

Dans ce papier, les auteurs présentent chaque système en eau peu profonde comme la spécialisation d'un modèle tri-dimensionnel 'parent'. Avec raison, ils identifient le fait que les variantes (ii) à (iv) nécessitent un modèle parent utilisant l'approximation EGA, même si dans la variante (ii), la planète est sphérique. L'idée que la forme de la planète ne détermine pas le type d'approximation du géopotentiel arrive donc ici à maturité. C'est pour cette raison que peu à peu dans la littérature, les notions floues de 'spherical approximation' et de 'ellipsoidal approximation', rencontrées au début de cette lignée de travaux, ont laissé place aux appellations plus pertinentes : 'spherical geopotential approximation' (SGA) et 'ellipsoidal geopotential approximation' (EGA). En effet, dans toute cette affaire, c'est l'hypothèse que l'on fait sur la forme des isopotentielles qui importe, autant que celle que l'on peut faire sur la forme de la planète.

Il faut bien voir que ces "débats d'experts" sont plus importants qu'il ne peut le paraître à la lecture de cette section, car il s'agit là de la formulation des équations au plus haut niveau, avant bien d'autres spécialisations (notamment l'introduction de la prise en compte concrète de l'orographie). On ne peut en aucun cas bâtir une météorologie correcte sur une base géophysique incorrecte ou même simplement douteuse. Avant de spécialiser les équations, il est absolument nécessaire d'être certain que le cadre géophysique général est solide et pertinent. Toute absence de consensus doit donc être vue comme un point bloquant à ce stade, car cela signifie que d'un côté ou de l'autre, un point conceptuel reste problématique.

4.3.5 Estimation quantitative des erreurs

Dans B14b, j'ai développé un modèle de prévision simple capable de d'intégrer les équations en eau peu profonde dans les quatre variantes précédentes. Ce modèle est en fait un précurseur de celui qui sera décrit au chapitre 5, mais incluant ici les fonctionnalités nécessaires pour expérimenter les quatre variantes en question. Les résultats montrent que le gain de réalisme offert par l'approximation EGA, l'impact de la gravité variable est plus important que l'impact de la forme non-sphérique de la planète. Le rapport est environ d'un facteur 3.

L'article est illustré par trois types d'expériences : un cas d'ondes de Rossby communément employé ; un cas d'instabilité barotrope proposé par Galewski et al. (2004) et quelques cas réels anciens figurant dans la série de tests proposés par Williamson et al. (1992).

4.3.6 Conclusion

L'inclusion de l'approximation EGA pour PNT pourrait être effectuée dans un futur proche, en se restreignant à l'approximation de l'atmosphère mince. Dans ce cadre, l'adaptation n'est pas très compliquée. Cependant, il y a tout de même quelques obstacles et limitations. Tout d'abord, il semble difficile d'envisager le passage en EGA pour un modèle spectral passant par une représentation des champs en

harmoniques sphériques. En effet, en EGA, il faudrait représenter les champs par des harmoniques sphériques. On connaît l'expression de ces fonctions, mais en revanche on ne connaît pas de formules de transformation permettant de passer de la représentation spectrale vers la représentation en champ physiques (ni pour les transformations inverses). De plus le passage à une géométrie basculée et/ou étirée comme dans le modèle ARPEGE actuel, rend la solution du problème absolument hors de portée, si tant est qu'il soit soluble. En définitive, l'option EGA n'est pas du tout adaptée pour les modèles spectraux. Pour des modèles utilisant uniquement des algorithmes locaux, comme discuté au chapitre 5 ou comme le nouveau noyau dynamique FVM du CEPMMT, ou DYNAMICO du LMD, le passage à l'approximation EGA-SA ne pose pas de problème majeur.

Pour ce qui est d'utiliser l'approximation EGA en atmosphère épaisse, c'est une autre complication. Bien qu'ici, le système de coordonnées proposé par Staniforth and White 2015b semble répondre de manière satisfaisante au problème posé. Néanmoins, comme précisé ci-dessus, l'avantage concret de travailler avec une atmosphère épaisse n'est pas très clair pour la météorologie terrestre (il pourrait en aller différemment pour certaines planètes gazeuses géantes). Enfin, pour l'anecdote, signalons qu'une inclusion pleinement rigoureuse des données observées (notamment les occultations GPS) dans un système d'assimilation en EGA-SA peut facilement tourner au cauchemar, la notion d'altitude d'un point dans l'espace extérieur recouvrant tant de définitions différentes selon les disciplines (géodésie, altimétrie satellitaire, météorologie...), et même à l'intérieur de chaque discipline !

Chapitre 5

PNT sur la sphère avec des algorithmes locaux

La technique spectrale utilisée dans nos modèles de prévision numérique (ARPEGE, AROME) ne s'adapte pas bien à l'architecture massivement parallèle en mémoire distribuée des futures machines de calcul intensif. La nécessité de procéder à des transpositions globales de données entraîne des communications massives et très handicapantes entre tâches (communications dites "all-to-all" en anglais). La question se pose donc d'explorer des pistes permettant de s'affranchir de cette technique spectrale pour nos futurs outils de PNT.

Dans tous les Centres Nationaux (notamment ceux où la technique spectrale est utilisée), d'ambitieux projets ont donc été lancés pour trouver de nouveaux algorithmes permettant de mettre en œuvre des modèles sur la sphère (ou sur l'ellipsoïde) dont l'efficacité reste compétitive sur des machines de calcul massivement parallèles.

Météo-France mène une collaboration rapprochée avec le CEPMMT pour ce qui est des noyaux dynamiques : ceux d'IFS et d'ARPEGE sont basés sur un code développé en commun, IFS/ARPEGE. Les pistes explorées pour les futurs noyaux dynamiques dans ces deux centres opérationnels sont d'une part la technique des volumes finis sur grille déstructurée, au travers du module FVM développé au CEPMMT, et la technique des différences finies sur grille semi-structurée au travers du prototype en eau peu profonde GRASS développé à Météo-France. La grille semi-structurée de GRASS est très proche de la grille d'IFS/ARPEGE, c'est une grille lat-lon réduite : sur chaque cercle de latitude d'indice $j \in 1, 2, \dots, N_\varphi$ on répartit uniformément $N_\lambda(j)$ points de grille, de telle sorte que la résolution physique soit quasi-uniforme.

Les articles Bénard et Glinton, 2019 (BG19 ci-après) et Glinton et Bénard, 2019 (GB19 ci-après) décrivent les idées théoriques sur lesquelles le modèle GRASS est bâti (BG19), et quelques résultats sur des expériences idéalisées afin de valider le concept (GB19).

La motivation première poursuivie dans le développement du modèle GRASS était de répondre à la question suivante, posée par la hiérarchie de Météo-France : "serait-il possible de se passer de la méthode spectrale dans ARPEGE tout en gardant l'ensemble du reste inchangé?". En d'autres termes est-il possible de bâtir une version modifiée d'ARPEGE dans laquelle la seule modification serait que la méthode spectrale n'est plus utilisée. En particulier la structure de la grille (dite lat-lon réduite) et son agencement (dite grille-A, ou non-décalée) sont conservés.

La méthode spectrale sert à assurer trois fonctions principales dans ARPEGE : le calcul des dérivées horizontales, le calcul de la diffusion horizontale, et l'inversion du problème horizontal créé par le schéma temporel implicite. Ce troisième et dernier aspect est laissé de côté dans le modèle GRASS car il est un peu indépendant de la question principale : la stratégie scientifique de Météo-France prévoyait déjà de sortir de la méthode spectrale pour la résolution du problème implicite. La raison de ceci est que le solveur spectral actuel qui ne peut résoudre que des problèmes linéaires horizontalement homogènes, ne permet pas d'inclure les termes orographiques dans le problème implicite, et se révèle instable en présence de fortes pentes du relief. La stratégie pour contourner ce problème était donc de développer un solveur tridimensionnel en point de grille basé sur la méthode de Krylov, dans lequel les termes orographiques peuvent être pris en compte (aspect exploré par Thomas Burgot durant son travail de doctorant).

Le travail de recherche sur le modèle GRASS se propose donc d'explorer la possibilité de bâtir un schéma numérique efficace et précis pour simuler l'évolution d'un système en eau peu profonde explicite

en temps, sur la sphère ou l'ellipsoïde, au moyen d'algorithmes de calcul exclusivement 'locaux'. Le sens du mot 'local' utilisé ici est important et doit être précisé. Il signifie d'abord qu'il n'y a plus aucune communication de type "all-to-all". De plus, le volume des communication pour une tâche parallélisée implique un halo de moins d'une centaine de points autour de la région de calcul, et ce halo concerne un très petit nombre de tâches adjacentes (huit dans le pire des cas). Une région parallélisée augmentée de son halo est désignée par le terme de 'région étendue'. La Figure 5.1 illustre une configuration possible des halos dans GRASS près de l'équateur, dans les régions tempérées, et près des pôles. Près de l'équateur, le nombre de point sur chaque cercle de latitude $N_\varphi(j)$ est constant, et la grille est donc une simple grille lat-lon régulière. Dans les zones tempérées et près des pôles, ce nombre varie avec j , les points de grille ne sont plus alignés méridionalement, et les régions parallélisées ne sont donc plus à proprement parler délimitées latéralement par des segments de méridiens (la figure est schématique sur ce point).

Chaque région parallélisée comporte environ 50 ou 100 points, et la profondeur du halo en direction de l'extérieur de la région est d'environ 3 ou 4 points (pour un ordre de précision de 6 ou 8 respectivement). Ainsi chaque halo est lui-même constitué d'une cinquantaine de points au maximum, répartis sur huit régions adjacentes au maximum. C'est le sens du mot 'local' employé ci-dessus. Il est admis que les communications liées à ce découpage ne sont pas handicapantes pour des machines massivement parallèles, la discussion de ce dernier point dépassant largement le cadre de l'étude entreprise pour le modèle GRASS.

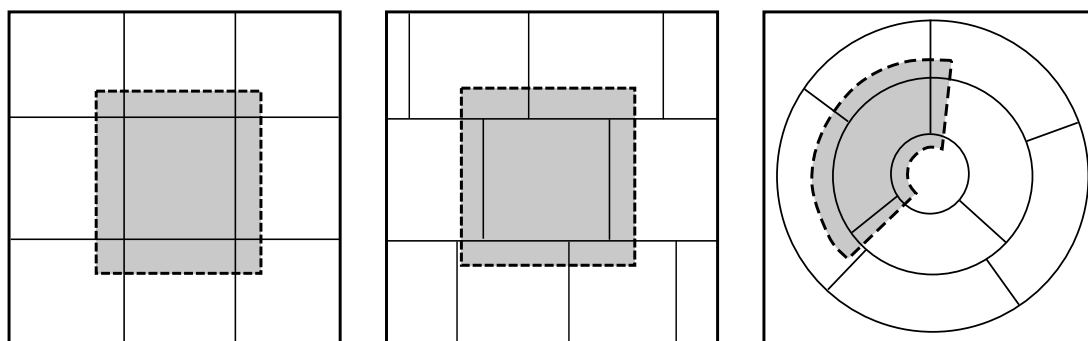


FIGURE 5.1 – Géométries possibles pour les régions parallélisées et les halos près de l'équateur (gauche), dans les zones tempérées (centre) et près du pôle (droite). Les lignes fines continues représentent les limites géographiques de régions correspondant à des tâches parallélisées, les lignes épaisses pointillées représentent les limites extérieures des halos. Les régions étendues (région + halo) sont colorées en gris. Les limites latérales des régions sont abusivement représentées par des segments de droites sur les figures du centre et de droite, alors qu'en réalité, il y a de petites sinuosités autour de ces segments.

L'article BG19 analyse les éventuels problèmes liés à l'adoption d'algorithmes locaux de type différences finies sur ce type de grille (lat-lon réduite). Les équations d'évolution sont formulées en composantes (zonales et méridiennes) pour les champs vectoriels. Les opérateurs de calcul vectoriel (gradient, divergence, rotationnel) sont aussi formulés en termes de composantes, ce qui implique des dérivées zonales et méridiennes de champs scalaires. Pour les dérivées zonales, puisque les points de grille ne sont pas alignés le long d'un méridien, des opérateurs d'interpolation zonale sont nécessaires (afin de calculer des données alignées méridionalement). Il apparaît que toute erreur de discrétisation des opérateurs de dérivée et d'interpolation zonales pour ces champs scalaires au voisinage immédiat des pôles crée une énorme erreur pour l'évaluation de l'opérateur vectoriel en question. En d'autres termes, une petite erreur pour la dérivée zonale des composantes d'un champ vectoriel crée une grosse erreur pour la divergence de ce champ. ce problème est connu depuis Shuman (1970), sous le nom de 'pole problem'. Plus la résolution est élevée, plus l'erreur sera grande près de pôles. Néanmoins, ce problème n'affecte que quelques rares membres de la famille des harmoniques sphériques qui peuvent être représentées par le champ discrétisé.

Cette analyse du problème amène donc à la conclusion que pour ces quelques harmoniques-là, le schéma de discrétisation doit être totalement exempt d'erreurs près des pôles pour les opérateurs zonaux. Une solution possible est donc d'utiliser la représentation de Fourier au voisinage des pôles pour estimer les dérivées zonales. L'analyse et l'expérience montrent que le recours à cette représentation n'est nécessaire que sur un petit nombre de cercles de latitude (environ 5), ce qui fait que le schéma reste local au sens défini ci-dessus. A chaque pôle, on peut définir une région parallélisée en forme de calotte polaire comportant 5 rangées de latitude, entièrement discrétisée dans l'espace de Fourier, en longitude. Cette

région comporte moins de 100 points et communique avec peu de région voisines, comme le montre la Figure 5.2, où le halo implique trois régions voisines seulement.

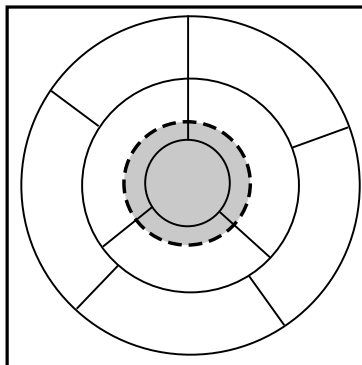


FIGURE 5.2 – Comme la Figure 5.1, mais pour la région polaire (la représentation de Fourier est utilisée dans la calotte polaire).

Un des avantages du schéma de discrétisation proposé est que, du fait de sa simplicité exploitant la semi-structure très forte de la grille, les schémas d'ordre de précision élevée sont très faciles à obtenir. Ainsi dans GRASS, des schémas d'ordre 2, 4 6, et 8 sont implémentés, permettant d'évaluer de façon tout à fait rigoureuse l'intérêt éventuel des ordres élevés pour la PNT.

L'article GB19 aborde plus complètement le volet validations du modèle GRASS. Les premiers tests abordent plusieurs écoulement stationnaires avec divers profils de vitesse au-dessus des pôles. Il est montré que – comme attendu – lorsque l'on élève l'ordre de précision, la réponse du modèle approche celle d'un modèle spectral. Un test très important est ensuite présenté : celui proposé par Scott et al. 2014. C'est le test le plus sévère à ce jour, pour un modèle global en eau peu profonde. Les tests montrent que le modèle GRASS est capable de satisfaire les critères (très stricts) permettant de valider ce cas idéalisé. Une étude sur l'impact de l'ordre de précision suggère qu'il pourrait être plus avantageux de travailler à ordre élevé, car pour obtenir une qualité de prévision égale à ordre peu élevé, il faut une résolution plus élevée, ce qui implique un coût significativement plus élevé.

La conclusion générale de ces deux papiers est qu'il n'est absolument pas impossible (moyennant un peu de soin aux pôles) d'utiliser des différences finies sur cette grille semi-structurée, et que cela permet un accès facile aux ordres élevés sur la sphère, ce qui est loin d'être le cas pour d'autres discrétisations (volumes finis, ...), surtout si l'on s'interdit de mettre à profit la possibilité d'une semi-structure. La véracité de l'idée que l'ordre 2 de précision en espace (d'ailleurs souvent même pas atteint sur la sphère en volume finis) est suffisant pour une PNT compétitive aux longues échéances n'a jamais été démontré, loin de là. Ma conviction personnelle est que l'ordre 4 est probablement bénéfique et d'un surcoût très raisonnable, le bénéfice réel des ordres ultérieurs restant lui, plus hypothétique, du fait du surcoût induit. Disposer d'un modèle permettant "au dernier moment" de choisir l'ordre de précision utilisé en opérationnel me semble un atout de poids, la réalité des scores à coût égal restant alors l'argument ultime, comme cela devrait toujours être en PNT. Utiliser le dogme d'une suprématie présumée de l'ordre 2 (voire 1.5 souvent) me semble dangereux. Un modèle souple sur ce point permet au moins de trancher le débat de façon claire.

Chapitre 6

Ondes de Rossby sur la sphère

6.1 Ondes de Rossby sur la sphère dans le système en “eau peu profonde”

L'équation du tourbillon barotrope (BVE en anglais, pour ‘Barotropic Vorticity Equation’) est, en sciences de l’atmosphère planétaire, un système qui décrit l’évolution d’un fluide homogène incompressible et inviscide confiné entre deux parois horizontales très proches supposées en outre séparées en chaque point par une distance constante verticalement, de sorte que les seuls mouvements possibles sont ceux parallèles aux deux parois (cf. Appendice 1). Par sa nature, ce système est purement cinématique, c’est-à-dire qu’il ne fait pas intervenir la masse ou la masse volumique du fluide, non plus que la gravité. Pour la science météorologique, on peut distinguer deux cas d’intérêt particulier en ce qui concerne la géométrie de la surface sous-jacente : surface plane (Fig. 6.1.a) ou sphérique (Fig. 6.1.b). La seule variable pronostique de ce système est le champ de vent \mathbf{v} . Mais comme ce vent est nécessairement horizontal et non-divergent (du fait de l’incompressibilité du fluide et du parallélisme des deux parois), le seul degré de liberté subsistant est la composante verticale du tourbillon ζ . Ce champ scalaire ζ est donc finalement la seule variable pronostique du système BVE. Une extension possible (mais pas utilisée en pratique) serait de considérer l’écoulement entre deux parois toujours fixes, mais dont la distance verticale n’est pas constante dans l’espace, en exprimant la conservation du tourbillon potentiel pour chaque particule matérielle.

Le système d’équation en eau peu profonde (SW en anglais pour Shallow-Water) décrit l’évolution du même fluide, parfait et incompressible, sauf que la paroi supérieure est absente, la surface supérieure du fluide étant libre. Il faut donc une force de gravité dirigée de haut en bas pour assujettir le fluide à la paroi inférieure, supposée isopotentielle ici. La pesanteur joue alors un rôle dans l’évolution, qui devient donc un phénomène dynamique. Toutefois, la masse volumique du fluide, constante du fait de l’incompressibilité supposée, n’intervient toujours pas dans l’évolution. Les variables pronostiques du système sont le champ vectoriel du vent \mathbf{v} qui maintenant est susceptible d’avoir une divergence non nulle, et le champ scalaire de l’épaisseur du fluide notée η . Les figures 6.1.c et 6.1.d illustrent la configuration du fluide pour la géométrie cartésienne et sphérique respectivement (dans le cas d’une gravité d’intensité uniforme).

Si l’on dispose d’un champ de vent valide pour un système BVE, c’est à dire non divergent, alors, en tenant compte des variations spatiales du paramètre de Coriolis f , il existe (à une constante près) un unique champ d’épaisseur η qui possède la propriété de maintenir nulle la tendance de divergence dans le système SW (c’est à dire qui ne génère pas d’ondes de gravité). Même dans le cas où le champ \mathbf{v} est analytique, ce champ η est difficile à exprimer analytiquement, mais dans certains cas simples, cela reste possible. Par exemple sur la sphère, Phillips 1959 propose une expression analytique de η lorsque ζ est une harmonique sphérique Y_n^m avec $(n - m) = 1$. Lorsque l’écart $n - m$ grandit, le problème devient de plus en plus compliqué, et dès lors que la valeur de $n - m$, est égale ou supérieure à 2, il n’existe plus d’expression analytique connue (ou plutôt disponible) pour η .

Au cours du développement du modèle GRASS, j’ai été amené à me poser des questions au sujet des modes de Rossby sur la sphère pour le système en eau peu profonde. Les ondes de Rossby ont été identifiées pour la première fois par Rossby (1939) dans le cadre d’un système BVE Cartésien mono-dimensionnel : Rossby avait montré que certaines ondes planes, inconnues jusqu’alors, pouvaient se manifester dans un simple fluide soumis à l’équation du tourbillon barotrope sur une surface plane (Fig. 6.1.a), si le paramètre

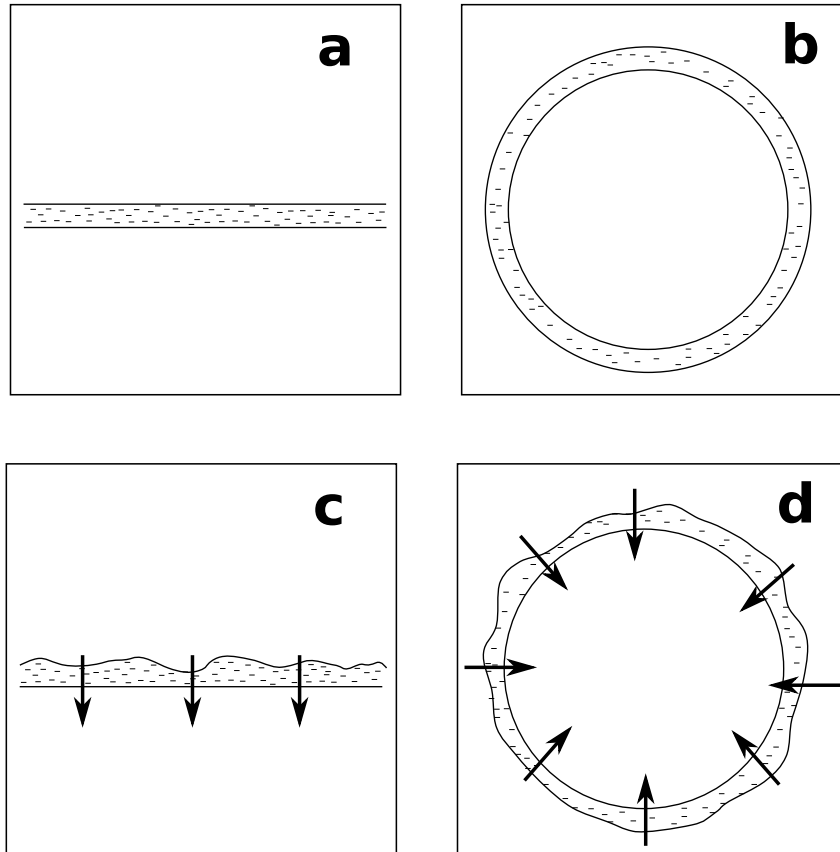


FIGURE 6.1 – Figures a et b : Coupes dans un fluide homogène et incompressible (tirets) confiné entre deux surfaces "parallèles" rapprochées, et dont les mouvements – parallèles aux surfaces – sont donc régis par l'équation du tourbillon barotrope. Fig. a : géométrie cartésienne ; Fig. b : géométrie sphérique. Figures c et d : Coupes dans un fluide homogène et incompressible (tirets) gisant sur une surface "inférieure" grâce à la gravité (supposée uniforme, orthogonale à la surface inférieure et représentée graphiquement par les flèches). La surface supérieure est une surface libre. Les mouvements de ce fluide sont régis par le système des équations en eau peu profonde. Fig. c : géométrie cartésienne ; Fig. d : géométrie sphérique.

de Coriolis présentait des variations spatiales régulières le long de cette surface (ce que l'on désigne sous le nom de β -plan). Haurwitz (1940) identifia ensuite l'ensemble des modes mûs par le même phénomène physique sur une sphère en rotation toujours pour le système BVE (Fig. 6.1.b), d'où le nom d'ondes de Rossby-Haurwitz sur la sphère.

Les modèles en eau peu profonde (SW) sont d'un degré de réalisme immédiatement supérieur au système BVE, tout en restant relativement simples. Comme nous venons de le voir, une des caractéristiques supplémentaires des systèmes SW par rapport aux systèmes BVE est que le mouvement horizontal du fluide peut être localement divergent (ou convergent), ce qui entraîne une diminution (ou augmentation) locale de la hauteur du fluide. Je n'ai pas trouvé trace dans la littérature d'une étude décrivant les modes de Rossby sur un β -plan pour le système SW, et j'ignore donc s'il existe une expression analytique de ces modes. Du reste, étant donné le peu de réalisme de cette géométrie pour la météorologie, ce n'est pas très important.

Un mode de Rossby du système SW sur la sphère serait donc un champ de vent et de hauteur qui, en vérifiant les équations d'évolution du système SW, se propagerait longitudinalement (à vitesse angulaire uniforme) sans aucun changement de forme au cours du temps. Le fait qui nous intéresse ici, c'est que la structure et même l'existence des modes de Rossby pour le système SW sur la sphère en rotation ne sont pas connus : on ne dispose pas d'expression analytique, car on ne sait tout simplement pas comment attaquer le problème de leur détermination. Devant ce constat d'impuissance, il me paraissait intéressant de pouvoir se faire au moins une idée expérimentale sur l'existence et la structure de ces ondes de Rossby,

en ayant recours à des moyens purement numériques.

Lorsqu'on initialise un modèle SW par l'expression du champ de vent d'un mode de Rossby valable pour le système BVE, et le champ de hauteur dont il a été question plus haut, on obtient une évolution qui est presque stationnaire, mais avec des oscillations rapides qui semblent avoir lieu autour d'un état moyen, lui-même plus ou moins proche d'une propagation zonale (Thuburn et Li, 2000, TL2000 ci-après). La méthode que je proposai pour atteindre les modes de Rossby du système SW était donc tout simplement de filtrer temporellement ces évolutions rapides pour tenter d'isoler une structure stationnaire. La méthode se révéla fonctionner correctement (avec quelques précautions) dans certaines plages des paramètres. Ces travaux sont présentés dans Bénard, 2019a, où je montre que si la structure est relativement facile à approcher, les modes obtenus sont instables sur le long terme. La plage d'instabilité est similaire à celle observée pour les modes en BVE. Le fait nouveau et marquant est que ces modes ont une divergence non-nulle, comme on pouvait s'y attendre.

Lors de la rédaction de cet article, le comportement observé dans la Figure 7 m'a beaucoup intrigué. Cette figure (reproduite ici en tant que Fig. 6.2) montre l'erreur de prévision sur l'évolution du vent en 25 jours, en fonction de l'amplitude de l'onde, pour une onde de géométrie donnée, précisément celle pour laquelle vorticité est proche de l'harmonique sphérique examinée par Phillips (1959), c'est à dire Y_5^4 . Nous nous intéressons ici seulement à la courbe la plus à droite, qui correspond à des ondes stationnaires dans le référentiel terrestre (les autres courbes correspondent à des ondes propagatives, soit vers l'est, soit vers l'ouest). Ce qui est appelé 'erreur' ici est tout simplement l'évolution constatée, puisque le champ recherché est supposé être purement stationnaire. On voit que lorsque l'amplitude de l'onde croît, l'évolution au bout d'un temps donné (i.e. l'erreur de prévision) est d'abord indétectable, puis, dans une certaine plage d'amplitude, autour de 40 m/s, se met apparemment à suivre une courbe ascendante.

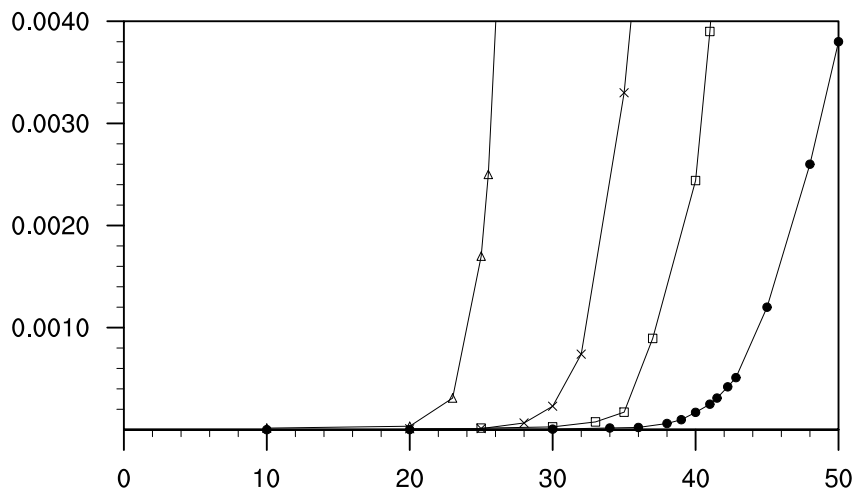


FIGURE 6.2 – Croissance en 25 jours de l'erreur sur le champ de vent méridien δv pour diverses ondes de Rossby du système SW sur la sphère en rotation, préparées numériquement. Ordonnées : erreurs δv , abscisses : amplitude de l'onde v_{\max} (valeur maximum du champ v). Seule la courbe la plus à droite est commentée ici, elle correspond au cas d'ondes stationnaires pour lesquelles la vorticité ζ est proche de l'harmonique sphérique Y_5^4 .

Une autre caractéristique qui n'apparaît pas sur cette figure, c'est que l'erreur possède une structure spatiale qui croît dans le temps, mais qui dans son évolution reste stationnaire relativement à l'onde elle-même. Cette caractéristique m'avait fait penser qu'il s'agissait d'une instabilité stationnaire, ce que j'avais écrit dans la première version du manuscrit. Un des deux Reviewers exprima des doutes sur ce sujet. Je dois dire que bien lui en prit, car après trois "navettes" du manuscrit peu à peu modifié, il fallut me rendre à la conclusion que ce n'était pas une instabilité stationnaire, mais plus prosaïquement l'effet résiduel de la faible diffusion horizontale appliquée dans le modèle de prévision. L'effet de cette diffusion croît très rapidement à partir du moment où le spectre de l'onde principale s'enrichit de composantes de petite échelle, ce qui se produit aux alentours de l'amplitude où v_{\max} est proche de 40 m/s. Face aux doutes réitérés du Reviewer, qu'il réussit à me communiquer, je résolus d'en passer par une analyse de

stabilité (pour le système BVE), bien que cela n'avait pas été expressément demandé lors de la révision de l'article.

6.2 Stabilité des ondes de Rossby sur la sphère dans le système 'BVE'

La méthode d'analyse la plus traditionnelle pour examiner ce problème ne me convenait pas. Cette méthode traditionnelle d'analyse, basée sur la 'wave interaction theory' et les 'coefficients d'interaction', met en jeu les symboles 3- j de Wigner, de calcul et d'emploi assez complexe, cf. p. ex. Hoskins (1973, H73 ci-après) et Baines (1976, B76 ci-après). Ces nombres ne sont pas grands, mais font intervenir des factorielles de grands nombres qui se neutralisent (au numérateur et au dénominateur). Il sont difficiles à calculer avec précision. C'est probablement cette complexité qui explique d'une part que ce type d'analyse n'ait pas connu beaucoup de développements ultérieurs en météorologie, et d'autre part que personne n'ait eu le courage de vérifier qui, de Baines ou Hoskins, avait raison puisque leurs résultats étaient contradictoires sur le point examiné ici, à savoir l'instabilité stationnaire de l'onde Y_5^4 .

J'ai donc mis au point une méthode alternative d'analyse de stabilité pour les ondes de Rossby sur la sphère dans le système BVE. Cette méthode, plus élémentaire et plus convaincante, est publiée dans Bénard, 2020. Cette analyse repose sur trois points spécifiques :

- une reformulation originale de l'équation de stabilité sous une forme factorisée qui simplifie la résolution du problème ;
- l'évaluation des termes non-linéaires dans l'espace physique (avec de simple différences finies), dans une stratégie proche de la 'transform method' utilisée dans les modèles spectraux de prévision météorologique (par exemple IFS ou ARPEGE). ceci évite le recours aux coefficients d'interaction, délicats à calculer.
- Une formulation rigoureuse du processus de troncature, nécessaire dans un analyse discrete.

L'ensemble du problème s'écrit alors sous forme matricielle, formulé dans l'espace propre de l'opérateur Laplacien discret, et l'analyse en elle-même s'exprime comme un problème aux valeurs propres classique. Le processus de troncature se traduit essentiellement par la définition d'une matrice de troncature (matrice rectangulaire). Cette méthode permet de faire aisément la distinction entre deux types d'instabilités différentes (appelées respectivement propagative ou p-instabilité, et stationnaire ou s-instabilité) selon que la perturbation instable se déplace zonalement ou reste stationnaire par rapport à l'onde de Rossby qui est examinée.

Un aspect important pour la crédibilité de ce type d'analyse, dite en 'modes normaux', est l'étendue de l'espace des perturbations examiné. Dans H73 et B76, cet espace était très limité, ce qui rendait les résultats sujets à caution. Comme la méthode reposait entièrement sur les simplifications qui résultent de cette étendue limitée (dénommée 'severe truncation assumption' en anglais), il était difficile de savoir comment les résultats seraient modifiés en relaxant cette hypothèse de troncature sévère.

La nature des perturbations examinées par H73 conduisent exclusivement à analyser la s-instabilité. De son côté, B76 a considéré deux types de perturbations, le premier type conduit exclusivement à évaluer la p-instabilité, tandis que le second est simplement la réitération de la perturbation utilisée dans l'analyse de H73 et conduit donc exclusivement à analyser la s-instabilité de l'onde. Un premier problème frustrant est que les résultats pour la s-instabilité ne sont pas en accord dans ces deux études, sans qu'on n'ait jamais pu départager qui avait raison ni les causes de cet écart : si l'on considère les ondes de Rossby pour lesquelles la vorticité est une harmonique sphérique de la forme Y_{m+1}^m , H73 trouve que les ondes pour $m \leq 5$ sont s-stables, tandis que B76 trouve que seules les ondes pour $m \leq 2$ sont s-stables. Outre la contradiction gênante, le statut des ondes pour $m = 3, 4, 5$ est donc inconnu en ce qui concerne leur s-stabilité. Puisque dans Bénard 2019a, j'examinais spécifiquement l'onde pour ($m = 4, n = 5$), il était impossible de conclure sur sa stabilité stationnaire à partir de la littérature existante.

Grâce à cette nouvelle analyse, le résultat était clair : l'onde P_5^4 ne peut pas présenter d'instabilité stationnaire pour le système BVE. Il est donc très peu probable que son homologue pour le système SW puisse être sujet à une instabilité stationnaire, et la croissance des erreurs observée sur la Fig. 7 devait trouver une autre explication. C'est donc suite à cette impossibilité probable d'un comportement instable stationnaire que j'ai dû battre en retraite sur l'origine véritable des erreurs de la figure 7 de cet article (Bénard, 2019a), au cours d u processus de revue, anormalement long, à cause de ce point qui restait en

suspens. D'autres causes possibles ayant été éliminées (aliasing,...), il ne restait plus que l'impact de la faible diffusion numérique appliquée dans le modèle comme candidat possible.

Malgré cette péripétie au cours du processus de revue de l'article 2019a, ce que j'apprécie dans ce papier, c'est qu'il parvient à isoler et décrire un écoulement stationnaire et divergent pour le système SW sur la sphère en rotation. C'est, dans la littérature, la première mention et description connue d'un écoulement possédant ces propriétés (stationnarité malgré la présence d'une divergence non-nulle), et c'est notamment la première visualisation du champ de divergence associé. L'inconvénient de la méthode utilisée reste néanmoins qu'elle ne fournit là qu'une description purement numérique, ce qui en d'autres termes signifie que l'expression analytique (i.e. exacte) de l'écoulement n'est toujours pas connue. Mais, du moins, la structure de l'écoulement est mise en évidence. On voit par exemple que pour les amplitudes faibles de l'onde, le champ de divergence possède une structure spatiale proche de celle du vent zonal, mais décalée de $\pi/2$ en longitude (Fig. 2 du papier). En revanche, dès que l'amplitude de l'onde croît, la structure du champ de divergence se déforme très rapidement (Figs. 2 et 4) et présente des structures de plus petite échelle que l'onde principale.

En ce qui concerne l'article 2020, excepté qu'il répond à une longue incertitude sur la s-stabilité des ondes de Rossby Y_4^3 , Y_5^4 et Y_6^5 , il n'apporte au fond pas grand chose de nouveau par rapport à B76. En effet, il confirme que toutes les ondes de Rossby sont instables pour la p-instabilité. Pour la majorité d'entre elles, cette p-instabilité se manifeste dès les faibles amplitudes. Incidemment, cette instabilité barotrope des ondes de Rossby (mais sur le β -plan périodique, et uniquement pour des ondes plane) est le mécanisme qui avait été mis en avant dans l'article de Lorenz (1972) pour illustrer concrètement le fameux "effet papillon" en météorologie.

La méthode d'analyse proposée dans Bénard, 2020 confirme également que pour une géométrie d'onde donnée Y_n^m l'instabilité propagative se manifeste bien avant l'instabilité stationnaire lorsque l'amplitude de l'onde croît, de telle sorte que les résultats de B76, bien qu'imparfaits, donnent une bien meilleure image du domaine de stabilité réelle des ondes de Rossby que les résultats de H73. Les ondes susceptibles de rester stables pour les plus grandes amplitudes sont celles pour lesquelles les valeurs de n et m sont petites. Les ondes les plus stables sont : Y_4^2 , Y_4^3 , Y_4^4 , Y_5^2 , Y_5^3 et Y_6^5 . Les autres ondes sont généralement instables à très faible amplitude et ne peuvent donc pas véritablement se développer dans l'atmosphère sans se briser, et par cela, alimenter d'autres ondes à leur tour. Ces résultats sont compatibles avec le fait que les ondes observées dans l'atmosphère possèdent un nombre d'onde zonal m préférentiellement compris entre 3 et 5, même si l'analyse ne dit rien de précis sur le processus qui permet la croissance de ces ondes.

Chapitre 7

Conclusion et perspectives

A l'orée des années 1990, la mise en place d'un cadre permettant d'effectuer à Météo-France des prévisions du temps avec un modèle pleinement compressible fut un défi excitant, qui nécessita de nombreuses investigations scientifiques. Avec Jean-François Geleyn, Radmila Bubnová (Brozková ensuite) et Gwenaëlle Hello, nous posâmes les premiers jalons de ce qui allait devenir le noyau dynamique du modèle opérationnel AROME. Avec, reconnaissons-le, une certaine naïveté dans l'évaluation de l'ampleur de la tâche, mais armée d'un optimisme sans limite, cette petite équipe initiale s'est lancée dans cette aventure qui constituait un rebond de perspective après les accomplissements qu'avaient constitués le passage à la méthode spectrale semi-implicite (années 1970-80) et la maîtrise complète de la méthode semi-lagrangienne semi-implicite (années 1980-1990). L'équipe s'est étoffée au fil des années avec notamment l'arrivée d'un bon nombre de scientifiques issus des services météorologiques des pays d'Europe centrale et orientale, comme en témoigne la liste des coauteurs de l'équipe initiale, dans les publications qui ont suivi la première publication BHBG95.

Au sein de Météo-France, un petit noyau de collaborateurs a également participé à l'aventure, bien souvent en tant que cheville ouvrière, permettant de soutenir le projet scientifique dans sa pérennité et/ou assurant un soutien précieux sur des aspects techniques particuliers. A ce titre il faut citer Y. Seity, P. Brousseau, K. Yessad, R. El Khatib et F. Bouttier, sans qui rien n'aurait été possible. Le soutien de la haute hiérarchie (Direction de la Recherche) a, quant à lui, connu des hauts et des bas (avec un net passage en creux vers 1996-1999), mais l'attitude généralement bienveillante sur le long terme a finalement permis la valorisation opérationnelle de l'outil si longuement mis au point. Aujourd'hui, l'outil 'historique' ARPEGE est régulièrement remis en question (car faisant en partie double emploi avec le modèle IFS du CEPMMT), tandis que la légitimité d'AROME comme fer de lance de notre savoir-faire reste entière, du moins pour l'instant.

Une grande partie des activités de recherche menées au cours de ma carrière a été consacrée à la conception et à l'évolution d'un noyau dynamique en Equations d'Euler stable et précis, c'est à dire bien adapté à la Prévision Numérique. Le franchissement de chaque étape délicate (schéma semi-Lagrangien, schéma d'intégration à deux niveaux temporels, haute résolution et fortes pentes...) a nécessité une bonne dose d'investigation, démontrant qu'avec ce système d'équations difficile à manier, tout ne coule pas de source, contrairement avec ce qui avait pu être observé en général avec les systèmes d'équations plus simples utilisés auparavant. Ce travail par étapes s'est traduit notamment par un nombre relativement important de publications, en comparaison de ce qui avait été produit par Météo-France avec les autres noyaux dynamiques (même s'il est difficile de comparer la valorisation à des époques différentes, avec des volontés politiques différentes). Les obstacles qui ont dû être surmontés (et donc les publications correspondantes) peuvent plutôt être comparés, en terme de nombres et d'importance, à ceux liés à l'introduction de la technique semi-Lagrangienne, une dizaine d'années auparavant. Cette technique, également délicate à mettre en œuvre, avait conduit à une masse impressionnante de publications au niveau international (Météo-France n'ayant pris sur ce sujet qu'une assez faible part à cet effort de publication, avec Coiffier et al. 1987 et Yessad et Bénard, 1996). Pour ce qui est de la galaxie ARPEGE/IFS, c'est véritablement le CEPMMT qui avait donc joué sur cet aspect 'semi-lagrangien' un rôle de leader en terme de recherches et de publications. Pour l'aspect 'non-hydrostatique' au contraire, la balance a penché intégralement dans l'autre sens, ce qui au reste est normal puisque ce sujet n'entraînait pas réellement dans les attributions du CEPMMT, jusqu'à date assez récente. Sur le chapitre 'non-hydrostatique pour la P.N.', notre équipe n'a donc pas démérité en terme de publication.

Les autres sujets abordés dans ce mémoire sont moins centraux pour la prévision opérationnelle dans l’immédiat. En ce qui concerne l’abandon de la géométrie sphérique, grâce aux résultats élaborés progressivement et de conserve avec l’équipe anglaise de l’UKMO (et une intervention un peu plus tardive également du LMD dans le débat), le chemin est solidement pavé, et nous possédons une idée assez claire de ce qu’il convient de faire ainsi que de la manière d’y parvenir. C’était loin d’être le cas avant la série de papiers (2012–2015) sur le sujet. Il est absolument clair notamment que l’abandon de la géométrie sphérique est incompatible avec une technique spectrale telle qu’actuellement utilisée pour la sphère. De plus, il ne fait pas de doute (du moins de ce côté-ci du Channel) que le passage à une atmosphère profonde entraînerait plus de désagréments que d’avantages en géométrie non-sphérique, et qu’il convient donc de se garder, pour plusieurs années encore, de s’engager dans cette voie prématurée.

Avant de passer aux perspectives, je voudrais terminer sur une note plus personnelle en revenant sur la tonalité générale de ma carrière de chercheur d’un point de vue plus humain.

Il ne fait aucun doute qu’une grande partie de mes travaux est intimement entrelacée avec ceux de l’équipe du UKMO (essentiellement Andy White, Andrew Staniforth, et Nigel Wood), bien souvent sur le mode de la controverse, et avec l’émulation résultant de ce qui transparaît parfois comme une certaine adversité. Avec le recul il me semble que toutes nos controverses se sont progressivement apaisées, et que nous partageons une identité de vue assez profonde sur l’ensemble des sujets abordés ici (le périmètre d’activité de l’équipe du UKMO est beaucoup plus vaste que ce qui a été abordé dans ce mémoire). Les glissements de positions ayant permis notre convergence de vue sur ces sujets ont été mutuels, si bien que je crois pouvoir dire qu’en nous aiguillonnant mais aussi en nous épaulant mutuellement, nous avons pu faire naître un consensus scientifique et progresser un peu ensemble dans la compréhension des mouvements de l’atmosphère et la manière de les simuler numériquement.

La relation avec les équipes canadiennes a été étroite aussi, mais là il s’agissait plutôt d’une sorte de filiation scientifique, la grande figure étant André Robert, qui disparut à peu près au moment où je débutai dans le domaine (... et il n’y avait pas grand chose à contester dans ses travaux !) En attribuant le rôle de ‘grand-père’ spirituel à André Robert, il est logique d’attribuer le rôle de ‘pères’ spirituels outre-Atlantique à Jean Côté, René Laprise, et encore une fois Andrew Staniforth dans la phase canadienne de sa carrière. Les qualités qui m’ont le plus impressionné chez tous les chercheurs listés ci-dessus sont l’audace, et l’opiniâtreté à poursuivre la quête du détail tant que tout n’est pas entièrement clarifié¹.

Côté français, les figures tutélaires sont Jean Coiffier et Jean-François Geleyn mais là c’est plus l’aspect humain qui ressort de notre cheminement commun durant toutes ces années.

7.1 Perspectives

Pour les aspects liés à la modélisation non-hydrostatique, de nombreuses idées nouvelles, ouvrant des perspectives attrayantes pour AROME sont portées par Fabrice Voitus. Citons entre autres :

- (i) d’une part, à court terme, l’utilisation d’une nouvelle variable (baptisée W) dans la partie ‘point de grille’ des calculs, dont la condition aux limites inférieure est homogène horizontalement, ce qui permet d’espérer une plus grande robustesse en présence de fortes pentes du terrain ;
- (ii) d’autre part, à plus long terme, une approche basée la conservation d’une pseudo-énergie au niveau discret, afin de garantir la stabilité du système en l’absence de résidus non-linéaires (ce qui n’est même pas le cas pour la formulation actuelle).

En ce qui concerne les aspects en géométrie non-sphériques pour la prévision météorologique, le point de basculement permettant d’envisager concrètement un transfert vers l’opérationnel n’est pas encore atteint. Cela tient au fait que la relaxation de l’hypothèse sphérique est à peu près incompatible avec la méthode spectrale actuellement utilisée, et qu’il est très difficile de prédire combien de temps encore la méthode spectrale conservera sa supériorité par rapport aux méthodes alternatives envisagées pour la remplacer. En l’état actuel des choses, ces travaux en géométrie ellipsoïdale doivent donc être regardés comme prospectifs et ressortissant de l’animation du débat scientifique (débat d’ailleurs essentiellement franco-britannique, à l’heure actuelle). Les travaux que j’ai menés ont permis de faire avancer le consensus sur deux idées principales : (i) l’introduction des géométries non-sphériques est infiniment plus aisée

1. La forme lumineuse des termes de Coriolis et de courbure dans l’équation (A-10)–(A.12) de WW12, est une source de rêverie : beauté et profondeur !

dans l'approximation dite en atmosphère mince qu'en atmosphère profonde ; (ii) l'approximation en atmosphère mince entraîne moins de dommages que l'hypothèse sphérique, pour la qualité des prévisions. En conséquence de cela, à terme (une fois les modèles spectraux tombés aux oubliettes), les modèles de prévision météorologiques devraient utiliser la géométrie ellipsoïdale avec l'approximation en atmosphère mince. Deux modèles actuellement en chantier (FVM au CEPMMT, et DYNAMICO au LMD) sont développés avec ce cadre en vue.

Pour ce qui est de l'abandon de la technique spectrale (et plus généralement des algorithmes non-locaux sur la sphère ou dans le plan), celle-ci paraît inéluctable à terme, mais sans que l'on sache véritablement à quelle échéance. Cet abandon concerne au premier chef le modèle global, à cause de la difficulté croissante à faire fonctionner correctement les transformées de Legendre quand la résolution augmente, puis dans un second temps, le modèle en aire limitée. Il s'agit donc de bâtir des noyaux dynamiques basés uniquement sur des algorithmes locaux. A Météo-France, nous proposons d'examiner des méthodes relativement simples, mais donnant un accès possible aux ordres de précision élevés, ce qui tranche avec l'approche du CEPMMT. Sur les aspects généraux (indépendants de la géométrie sphérique ou plane), des travaux préparatoires ont été menés dans le cadre des modèles ALADIN et AROME en vue de répondre aux deux principales interrogations suivantes.

- (i) L'utilisation de la grille 'A' (grille co-localisée) est-elle compatible avec les algorithmes locaux ? La réponse est affirmative : thèse de Steven Caluwaerts, et utilisation de plus en plus répandue de cette grille (initialement décrite) dans plusieurs modèles de PN de par le monde.
- (ii) Les solveurs non-spectraux de type Krylov sont-ils prometteurs pour les applications visées en PN à haute résolution ? Les travaux de thèse de Thomas Burgot semblent ici aussi permettre une réponse affirmative (et d'autres modèles utilisent cette technique aussi avec succès). La dernière question est de savoir si la mise en place d'opérateurs de diffusion appropriés (non-spectraux) est possible afin de contrôler la forme de la queue du spectre des champs prévus. La réponse affirmative à cette question ne semble pas, ici encore, soulever de problème particulier.

Toujours en vue d'abandonner la technique spectrale, pour les aspects spécifiquement liés à la sphère (si tant est que nous conservions en opération un modèle global propre à Météo-France), les travaux que j'ai menés avec GRASS indiquent qu'au moins pour la partie discrétisation horizontale des équations, il n'y a pas de problème majeur. Pour ce qui est de la possibilité d'utiliser un solveur spatial basé sur des opérateurs purement locaux (en vue du maintien de schémas semi-implicites) la réponse est un peu plus spéculative au stade actuel. L'objectif ici est d'utiliser un solveur de Krylov, avec un préconditionneur adapté. En l'absence de véritable maquette sur la sphère, nous ne disposons que d'éléments partiels permettant de penser que cette technique ne devrait pas poser de problème majeur : existence d'un solveur similaire sur la sphère dans le modèle russe (SL-AV model), même si la nature du solveur est un peu différente du fait de l'approche Fourier en longitude choisie pour ce modèle ; et comportement prometteur de la maquette en aire limitée sur le plan utilisant un solveur de Krylov similaire à celui qui serait utilisé sur la sphère.

Chapitre 8

Appendice 1 : Incompressible fluid between two fixed rigid surfaces

With the aim of establishing an analogy between the filtered geopotential equation (FGE) used by Charney et al. 1950, and the barotropic vorticity equation (BVE), the evolution of a perfect (inviscid) incompressible fluid between two smooth surfaces close together is examined here (implicitly the two surfaces are supposed rigid and in a mutual position independent of time).

By "two smooth surfaces close together" is understood the satisfaction of the following property : for at least one of the two surfaces, denoted S henceforth, any straight line emerging orthogonally to S intersects the other surface S' once and only once, and this, at a distance η that is significantly smaller than the curvature radius of S at the point of emergence. The surface S for which the property is true is termed "lower surface" henceforth, and (S' the "upper surface"). If the property is satisfied by the two surfaces, the lower surface may be chosen arbitrarily, as for instance the one for which the lower bound of the curvature radius is the largest (i.e. the smoothest), or any other criterion. At any point P of S the orthogonal line intersects S' once and only once, at a distance $\eta(P)$, thereby defining a straight segment the length of which is termed the "depth of the fluid at the point P ". The two surfaces are therefore roughly parallel, but the distance between the surface may vary from a place to another. The situation is depicted in 8.1.

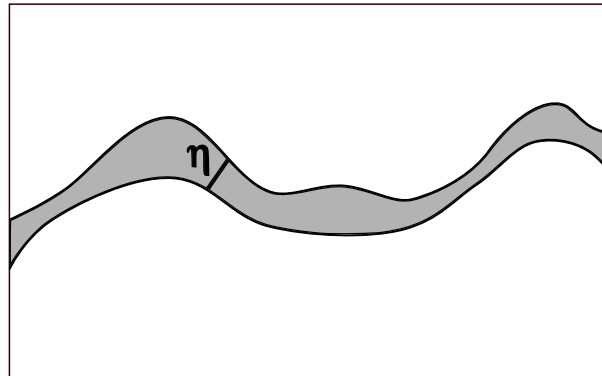


FIGURE 8.1 – Illustration (through a cross-section) of two surfaces that are "smooth, close together, but non-parallel". The fluid lays in the layer depicted in gray. The depth between the surfaces is defined by the length η of the straight segment orthogonal to S (here, the lower surface).

A system of curvilinear coordinates (s_1, s_2) is defined in the surface S . One may therefore write without any ambiguity $\eta = \eta(s_1, s_2)$. In the space located between the two surfaces, a third coordinate s_3 may be defined along the straight segment discussed above, so that on S , $s_3 = 0$, and on S' , $s_3 = 1$. Due to the above curvature radius condition, two segments emerging from two different points of S do not intersect themselves inside the fluid layer. Hence, conditions are fulfilled for defining a space transformation from the canonical Euclidean coordinate system (x, y, z) to the system (s_1, s_2, s_3) , with all associated metric

factors. Relative variations of metric factors along s_3 are assumed small enough to be neglected. All the above hypotheses and postulates are purely geometric and only involve the shape of the surfaces, not the flow. They describe what is traditionally termed as a "thin layer approximation" or, more specifically in meteorology, a "shallow atmosphere approximation".

Beside this geometric approximation, all physical variables describing the state of the fluid are assumed to be constant along the depth segment. This is the hypothesis of two-dimensionality of the flow. When the two approximations of "thin layer" and "two-dimensional flow" are combined, the governing system describing the evolution of the flow becomes purely two-dimensional, the coordinate s_3 completely disappears from the mathematical problem, so that finally s_3 may be ignored.

The goal of this chapter is therefore to study analogies and differences between the system discussed above (when the layer is filled with an inviscid incompressible fluid), and that considered by Charney et al. 1950. Unless specified, the two surfaces are assumed to be resting in the Galilean frame of fixed stars. The wind vector and the pressure inside the fluid are denoted $\mathbf{v} = \mathbf{v}(s_1, s_2)$ et $p = p(s_1, s_2)$ respectively. The governing equations of the system for such surfaces resting in the Galilean frame are simply

$$\frac{d\mathbf{v}}{dt} = -\frac{1}{\rho_0} \text{grad}(p) \quad (8.1)$$

$$\text{div}(\eta\mathbf{v}) = 0 \quad (8.2)$$

where grad and div are the gradient and divergence operators.

8.1 Case of a plane flow

In this section, to illustrate a conceptually simple case, the lower surface S is assumed to be a non-rotating plane, with an assumed Cartesian coordinate system (x, y) and the corresponding velocity components (u, v) . Governing equations of the system are thus simply

$$\frac{du}{dt} = -\frac{1}{\rho_0} \frac{\partial p}{\partial x} \quad (8.3)$$

$$\frac{dv}{dt} = -\frac{1}{\rho_0} \frac{\partial p}{\partial y} \quad (8.4)$$

$$\frac{\partial}{\partial x}(\eta u) + \frac{\partial}{\partial y}(\eta v) = 0 \quad (8.5)$$

where the material derivative is defined by :

$$\frac{d}{dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} \quad (8.6)$$

The depth η being a given data of the problem, there are three 2D state variables (u, v, p) and three 2D independent equations; the system is therefore closed, but the variable p has no prognostic equation. The state of the system is constrained at each time by the continuity equation (8.5). One may write

$$\frac{d\eta u}{dt} = -\frac{\eta}{\rho_0} \frac{\partial p}{\partial x} + u^2 \frac{\partial \eta}{\partial x} + uv \frac{\partial \eta}{\partial y} \quad (8.7)$$

$$\frac{d\eta v}{dt} = -\frac{\eta}{\rho_0} \frac{\partial p}{\partial y} + uv \frac{\partial \eta}{\partial x} + v^2 \frac{\partial \eta}{\partial y} \quad (8.8)$$

Besides :

$$\frac{d}{dt} \left(\frac{\partial \eta u}{\partial x} \right) = \frac{\partial}{\partial x} \left(\frac{d\eta u}{dt} \right) - \frac{\partial u}{\partial x} \frac{\partial \eta u}{\partial x} - \frac{\partial v}{\partial x} \frac{\partial \eta u}{\partial y} \quad (8.9)$$

$$\frac{d}{dt} \left(\frac{\partial \eta v}{\partial y} \right) = \frac{\partial}{\partial y} \left(\frac{d\eta v}{dt} \right) - \frac{\partial u}{\partial y} \frac{\partial \eta v}{\partial x} - \frac{\partial v}{\partial y} \frac{\partial \eta v}{\partial y} \quad (8.10)$$

Summation of the two equations leads to

$$\frac{1}{\rho_0} \left[\frac{\partial}{\partial x} \left(\eta \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(\eta \frac{\partial p}{\partial y} \right) \right] = \frac{\partial}{\partial x} \left(u^2 \frac{\partial \eta}{\partial x} + uv \frac{\partial \eta}{\partial y} \right) + \frac{\partial}{\partial y} \left(uv \frac{\partial \eta}{\partial x} + v^2 \frac{\partial \eta}{\partial y} \right) - \frac{\partial u}{\partial x} \frac{\partial \eta u}{\partial x} - \frac{\partial u}{\partial y} \frac{\partial \eta v}{\partial x} - \frac{\partial v}{\partial x} \frac{\partial \eta u}{\partial y} - \frac{\partial v}{\partial y} \frac{\partial \eta v}{\partial y} \quad (8.11)$$

The left-hand side implies a modified Laplacian operator Δ_1 in which the depth η plays a role similar to a metric factor :

$$\Delta_1(p) = \frac{1}{\eta} \left[\frac{\partial}{\partial x} \left(\eta \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(\eta \frac{\partial p}{\partial y} \right) \right]. \quad (8.12)$$

The diagnostic equation (8.11) gives the pressure field at each time from the wind and depth fields. The knowledge of p then gives access to the tendency of the wind field through momentum equations. Technically, this system may be integrated in time as an anelastic system : from the wind field, the pressure field is diagnosed by solving a problem in space, then, with this pressure field, the wind field may be computed at next time-step and so on.

8.1.1 Evolution of vorticity on the plane

For this Cartesian system, the vorticity ζ is

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \quad (8.13)$$

The material derivative of the vorticity is given by :

$$\begin{aligned} \frac{d\zeta}{dt} &= \frac{\partial}{\partial x} \left(\frac{dv}{dt} \right) - \frac{\partial}{\partial y} \left(\frac{du}{dt} \right) - \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \frac{\partial u}{\partial x} - \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial v}{\partial y} \frac{\partial u}{\partial y} \\ &= \frac{\partial}{\partial x} \left(\frac{dv}{dt} \right) - \frac{\partial}{\partial y} \left(\frac{du}{dt} \right) - \zeta D \end{aligned} \quad (8.14)$$

where D is the wind divergence :

$$D = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}. \quad (8.15)$$

Consequently, whatever are the specific forces acting on the fluid, if they derive from a potential, the vorticity equation reduces to :

$$\frac{d\zeta}{dt} = -\zeta D. \quad (8.16)$$

This equation means that the vorticity ζ of a fluid parcel may be modified only if this parcel is subjected to a divergent flow, which in turn may occur only if the parcel experiences a variable depth along its displacement. When the parcel travels toward a region of thinner depth, the divergence is positive, and the parcel must occupy a larger area, so that, (if non-zero) the magnitude of its vorticity decreases, exactly as for the famous ice-skater.

8.1.2 Conservation of potential vorticity on the plane

The time derivative of the potential vorticity $q = \zeta/\eta$ for a material parcel is

$$\frac{dq}{dt} = \frac{1}{\eta} \frac{d\zeta}{dt} - \frac{\zeta}{\eta^2} \frac{d\eta}{dt}. \quad (8.17)$$

Therefore, by inserting the rigid boundary condition $(\partial\eta/\partial t) = 0$:

$$\eta^2 \frac{dq}{dt} = \eta \frac{d\zeta}{dt} - \zeta \frac{d\eta}{dt} \quad (8.18)$$

$$= -\zeta \eta \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) - \zeta \left(u \frac{\partial \eta}{\partial x} + v \frac{\partial \eta}{\partial y} \right) \quad (8.19)$$

$$= -\zeta \left(\frac{\partial \eta u}{\partial x} + \frac{\partial \eta v}{\partial y} \right) \quad (8.20)$$

The kinematic constraint of the continuity equation then leads to the material conservation of the potential vorticity, valid as soon as the surfaces are rigid and the total force derives from a potential :

$$\frac{dq}{dt} = 0. \quad (8.21)$$

8.1.3 Particular case of a stationary flow on the plane with $\eta = \eta(x)$ and $(\partial/\partial y) = 0$

If $\eta = \eta(x)$ and $(\partial/\partial y) = 0$, then $u = u(x)$, $v = v(x)$. The flow being assumed stationary, the second momentum equation imposes that v is a constant, $v = v_0$. From the first momentum equation,

$$u \frac{\partial u}{\partial x} = -\frac{1}{\rho_0} \frac{\partial p}{\partial x}, \quad (8.22)$$

so that $(u^2/2 + p/\rho_0)$ is a constant, accordingly to Bernoulli's Theorem, valid for stationary one-dimensional flows. The pressure field is given by

$$p = p_0 + \frac{\rho_0 u_0^2}{2} \left(1 - \frac{\eta_0^2}{\eta^2} \right), \quad (8.23)$$

where u_0 , p_0 , η_0 are values at an arbitrary point (e.g. where the depth is minimum). To illustrate this, assume that the depth consists in a single 'narrowing' or 'bottleneck' parallel to y axis and localised near $x = x_0$. Whatever is the incident wind, the presence of the bottleneck has no effect on the field v , and only u is perturbed. Streamlines therefore locally undergo a bending, but only due to the acceleration of u when the flow passes the bottleneck, as illustrated in figure 8.2. Note also that since $\zeta = 0$ in the unperturbed area and ζ/η is conserved along parcel trajectories, ζ is zero even in the perturbed area where streamlines are bended. The bending of streamlines in the figure is only due to divergence, the vorticity is zero everywhere.

8.1.4 Standalone formulation on the plane

The pressure force is a 'dynamical' force applied to each parcel, but in turn, this force depends on the velocity only, so that the evolution of the velocity field only depends on the velocity field itself. If the pressure is appropriately eliminated from the system, the evolution becomes purely kinematic (i.e. gravity, pressure and density no longer play any role). The potential vorticity conservation equation, which also is purely kinematic, constitutes a bridge between the 'physical world' formulation with (u, v, p) and 'scalar world' with a standalone formulation *à la* BVE, involving a single prognostic equation for a single scalar quantity.

Concretely, since the non-divergent vector field is here $\eta \mathbf{v}$, a streamfunction may be associated to this field. For convenience, this streamfunction is denoted $\eta\psi$, so that ψ has the dimension of a usual velocity field streamfunction. Hence

$$\eta u = -\frac{\partial \eta \psi}{\partial y} \quad (8.24)$$

$$\eta v = \frac{\partial \eta \psi}{\partial x} \quad (8.25)$$

The potential vorticity writes

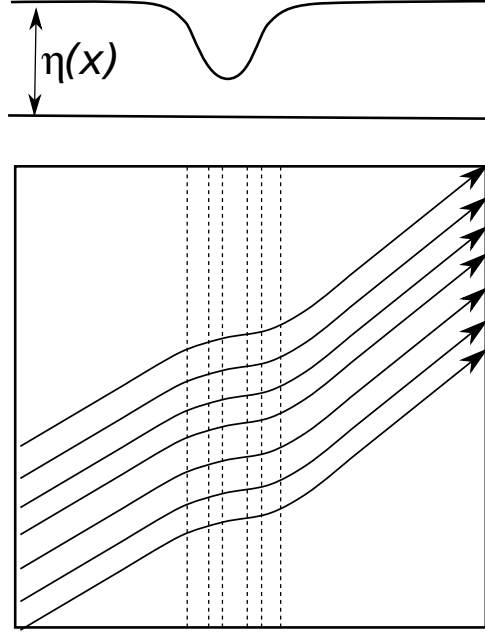


FIGURE 8.2 – Streamlines (arrows) and pressure field (dashed lines) for a stationary flow with $\eta = \eta(x)$ and $\partial/\partial y = 0$. The depth field $\eta = \eta(x)$ is depicted above the main panel. The pressure field shows a low at the same place, but narrower than η .

$$q = \frac{1}{\eta} \left[\frac{\partial}{\partial x} \left(\frac{1}{\eta} \frac{\partial \eta \psi}{\partial x} \right) + \frac{\partial}{\partial y} \left(\frac{1}{\eta} \frac{\partial \eta \psi}{\partial y} \right) \right] = \frac{1}{\eta} \Delta_2(\psi). \quad (8.26)$$

where Δ_2 denotes another modified Laplacian operator, in which η play the role of a metric factor. The conservation of q then implies

$$\frac{\partial \Delta_2 \psi}{\partial t} = \frac{\partial \eta \psi}{\partial y} \frac{\partial}{\partial x} \left(\frac{1}{\eta} \Delta_2 \psi \right) - \frac{\partial \eta \psi}{\partial x} \frac{\partial}{\partial y} \left(\frac{1}{\eta} \Delta_2 \psi \right) \quad (8.27)$$

This standalone scalar equation completely encloses the integrality of the evolution law for the system examined here, exactly as the single BVE equation under its traditional form.

8.1.5 Particular case of uniform depth $\eta = \eta_0$ on the non-rotating plane

In the case of the non-rotating plane with $\eta = \eta_0$, the above equation is identical to the standard BVE equation on the non-rotating plane :

$$\frac{\partial \Delta \psi}{\partial t} = \frac{\partial \psi}{\partial y} \frac{\partial}{\partial x} (\Delta \psi) - \frac{\partial \psi}{\partial x} \frac{\partial}{\partial y} (\Delta \psi). \quad (8.28)$$

The standard BVE system on the plane is therefore equivalent to that of a shallow incompressible fluid enclosed between two parallel rigid surfaces.

When $\eta = \eta_0$, the wind divergence identically vanishes, and the vorticity ζ , being then proportional to the potential vorticity q is conserved for any parcel. The diagnostic pressure equation is reduced to

$$\frac{1}{\rho_0} \Delta p = \left[- \left(\frac{\partial u}{\partial x} \right)^2 - \left(\frac{\partial v}{\partial y} \right)^2 - 2 \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} \right] = 2 \left(\frac{\partial u}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} \right).$$

\Rightarrow

$$\frac{1}{\rho_0} \Delta p = 2 J(u, v). \quad (8.29)$$

Even if the depth is constant, a space variation of the flow velocity induces the existence of space variations for the pressure field.

The physical meaning of (8.29) is not totally obvious at first sight, but more insight is accessed in simple cases. The streamfunction, assumed smooth, could be decomposed as a superposition of many isotropic Gaussian localised functions, with various amplitudes and scales, with their centres spreaded in the whole plane (and with possibly overlapping supports). A single member of this set (or ‘decomposition’) represents, in the physical world, a stationary circular vortex. Let us consider for instance the following element, centred at the origin O

$$\psi(r) = -(C/2) \exp(-r^2), \quad (8.30)$$

where $r = \sqrt{(x^2 + y^2)}$ is the length of the radius vector, and C is some algebraic constant. As for any purely circular vortex —i.e. function of the form $\psi = \psi(r)$ — the state examined here is stationary, due to the material conservation of ζ or q . For this flow,

$$\begin{aligned} u &= -C y e^{-r^2} & \frac{\partial u}{\partial x} &= 2C x y e^{-r^2} & \frac{\partial u}{\partial y} &= (2y^2 - 1)C e^{-r^2} \\ v &= xC e^{-r^2} & \frac{\partial v}{\partial x} &= (1 - 2x^2)C e^{-r^2} & \frac{\partial v}{\partial y} &= -2xyC e^{-r^2} \end{aligned}$$

The algebraic derivation of various physical fields leads to

$$\begin{aligned} V &= \sqrt{(u^2 + v^2)} = rC e^{-r^2} \\ \zeta &= 2C (1 - r^2) e^{-r^2} \\ J(u, v) &= (1 - 2r^2)C e^{-2r^2} \\ p &= p_0 - (\rho_0/4)C e^{-2r^2}. \end{aligned}$$

Variations of ψ , V , ζ , p and $J(u, v)$ along the (O, r) are depicted in figure 8.3. The shape of fields p and J is indeed independant of the sign of the vortex rotation (clockwise or anti-clockwise) since they depend quadratically on ψ or C . Jacobian J vanishes at the place where V is maximum. The pressure perturbation is centred at O and is also negative whatever is the sign of the vortex rotation.

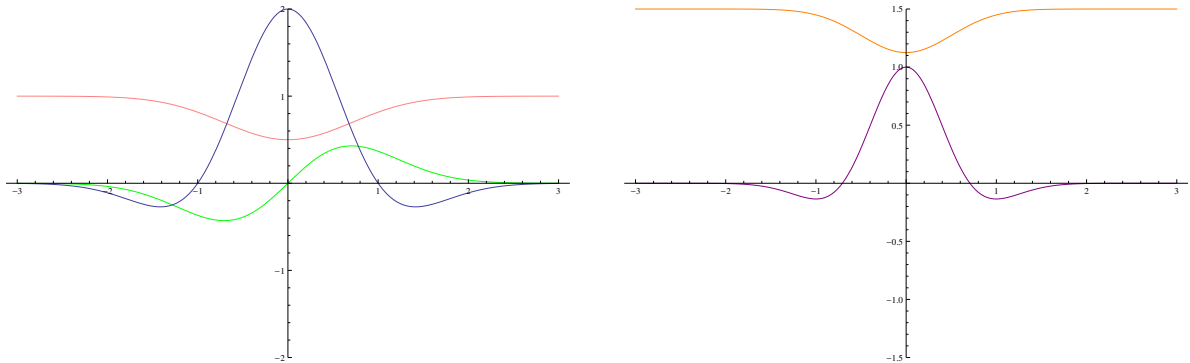


FIGURE 8.3 – Radial variations of various physical fields for the single vortex discussed in the text, for an incompressible fluid between two plane parallel fixed surfaces. Left panel : ψ (pink) ; v (green) ; ζ (blue). Right panel : p (orange) ; $J(u, v)$ (purple)

8.1.6 A personal digression

Here I would like to include some rather personal views, inspired by the latter derivations for the motions of an incompressible fluid between two parallel planes. This digression is totally speculative without any claim of a scientific value, and should be regarded more as my two cents added in the 'how-I-see-the-world' endless debate, on a more aesthetic than fully scientific side, say.

The 'universe' contemplated here (the fluid lying between the two surfaces) may be described in two equivalent ways : (i) a 'physical' representation, with a number of physical and observable fields possessing an immediate meaning to our common sense (u, v, p), with physical invariants also possessing an immediate meaning (potential vorticity, energy, etc.), and a number of evolution rules deriving from a number of physical conservation laws, also possessing an immediate meaning ; and (ii) a 'mathematical' representation where all the state of the universe is enclosed in a unique but non-observable scalar function ψ , and a standalone but quite obscure evolution equation as (8.27), the meaning of which –as one would easily admit– is not immediately accessible to our sense (the local evolution law being differential, must be supplemented by some gauge conditions for ψ). Of course, these two aspects are two facets of the same reality, but the fascinating feature is that a very rich and complex universe (with complicated motions, numerous invariants...) ultimately results from a single standalone equation involving a single scalar field. These two facets show a striking analogy with the mutually-opposed spirit of Newtonian and Quantum Mechanics descriptions of the world.

If one calls 'particle' a circular vortex of the form (8.30), there is something fascinating in the fact that observable concepts as mass, kinetic energy, velocity, momentum, or angular momentum of the particle are just the numerical result of integro-differential space operators applied to the non-observable function ψ defining this vortex. Of course there are weaknesses in this view, since an initially circular vortex might cease to be exactly circular when interacting with another vortex, thereby compromising the view of a vortex/particle as a necessarily stable and self-identical structure in time. Many other similar objections might be found, but none of them is really redhibitory to an analogy with Wave Mechanics (in a 2D universe here, though).

In the Bohr-Einstein debate, a key issue is the postulate (by Copenhagen School) that the governing law of the real world is fundamentally not deterministic, that the wave function only and inherently represents probabilities, and that one *must not* try to discover any determinism 'hidden' at some lower level : determinism fundamentally does not exist (i.e. evolution laws may only handle probabilities, and will *never* be able to handle *nothing more* than probabilities). This interpretation leads to various corollaries to which Einstein was highly allergic. The most famous one is that in the double slit experiment of Young, the interference pattern is possible 'because and only because' one does not know by which slit each particle has passed. This is discussed extensively in Chapter I of Feynman's book (1965) : as soon as one switches-on a apparatus that is able to 'look' which slit particles went through, the interference figure *must be* necessarily destroyed, because some determinism is introduced by this measurement, and the interference is a consequence of indeterminism (sections 1.5 and 1.6 therein).

Fascinating experiments with macroscopic 'dual wave-particle' systems (termed "dual walkers", Couder and Fort 2006, see also the website dualwalkers.com or Bush et al., 2018, for many other references) suggest that this argument by Feynman (as a follower of Bohr's dogma) might be totally flawed (thereby flawing in turn Copenhagen dogma). These dual walkers experiments show that it is perfectly possible to build a deterministic system for which interference pattern occurs even when the actual path of the particles is known (and the pattern being of course destroyed if one slit is obturated). When individual particles are passing one of the two slits, their trajectories on the screen build-up an interference pattern, this pattern being of course only observable by statistics on a large number of particles, exactly as in the double slit photon experiment, when photons are emitted one by one.

The weakness of Feynman's argument would then originate from assuming that the deterministic evolution laws for a simple or 'isolated' system must be simple (for instance he implicitly postulates straight trajectories for particles, past the slit, in a restrictive view astonishingly reminiscent of Newtonian Mechanics simplifications). In dual walker experiments the observable system is very simple (one single particle moving along some perfectly observable path with some perfectly observable velocity), but the underlying mechanics is of an appalling complexity, because the evolution (the path and velocity) depends on the complicated propagation of waves generated by the particle in the whole space, this wave being itself constrained by not-less-complicated remote boundary conditions, and also influenced by the whole

history of the system (particle and boundaries). The trajectory of a particle past the slit is therefore not necessarily a straight line with uniform velocity in dual walkers experiments, due to faint perturbations linked to the history of the particle itself, or to slight perturbations of the wave structure resulting from remote boundaries.

Newton (and Feynman) says that a particle moves with a uniform motion in the absence of force, but Couder and Fort (2006) dual walkers experiments suggest that the concept of 'absence of force' does not exist in practice, in the real world, since, due to frightfully complex and sensitive evolution laws, the slightest perturbation from a far remote source or a past history event is sufficient to create a significant modification in the long term evolution (i.e. a non-zero force). In dual walkers experiments, uniform motion does not exist. Only exist appallingly complex deterministic laws applying to wave-particle systems spreading in the whole space/time domain. This complexity may lead to the interpretation that determinism is lost, but in reality the determinism is hidden by a tremendously high level of chaos. It is really enlightening to understand in deep detail the mechanics of dual-walker double-slit experiment, and to re-read the section 1.5 of Feynman's book, identifying with an extreme ease where are the weaknesses in his interpretation of experimental results. In particular the sentence 'It is all quite mysterious' sounds outrageously naive and obsolete when considering Couder and Fort experiment, in which everything is immediate and simple to interpret.

Incidentally, these macroscopic experiments are also able to reproduce quantized orbit radii for pairs of orbiting particles, and various other striking macroscopic analogues of Quantum Mechanics behaviours¹. The general feeling at reading these articles is that Quantum behaviour which was claimed by Copenhagen School to be inherent to the microscopic world is indeed perfectly reproducible in a macroscopic world, leading to the idea that finally, strange and fundamentally undeterministic (as alleged) Quantum behaviours, might very well be the result of the frightfully sensitive character of deterministic evolution laws, leading to permanent chaos, even when applied to (apparently) very simple systems. In this view, the evolution is therefore deterministic, but the chaos is permanent and everywhere, even for simple systems. So that Newtonian Mechanics is one particular version of a statistical physics applied to the real world, whereas Quantum Mechanics is another –more refined– particular version of a statistical physics, both representations falling short of truly describing the universe.

Adopting another phraseology, one could say that there is still some place for the evolution laws of Universe to be local as BVE (save for some gauges conditions for ψ) and deterministic, but that the concept of isolated system does simply not exist, so that nobody is allowed to speak of an 'isolated system'. Thus, all the apparent complexity of microscopic behaviour comes from the unreductible complexity of the unique field describing the Universe and permanently covering the whole time-space, and this field can –in any respect– never be considered as 'local' by neglecting the influence of its remote-wings oscillations. When considering an 'apparently simple system' (i.e. a situation where the magnitude of the field is comparatively high only in some very limited region) ignoring non-zero magnitude of remote wings of the field leads to approximations that lead to false conclusions (e.g. photon trajectories are straight lines past the slits or in vacuum), and thereby, to a false interpretation of the reality. The striking thing in this view is that the above fragment '...isolated system simply does not exist...' is very reminiscent of some Quantum Mechanics phraseology (any measure influences the evolution,...), but is used here to go into a completely different direction. The views expressed here are so vague that experiments around quantic entanglement, and debates around realism and locality do not constitute objections at this stage.

One may imagine that finally, the Universe is too complex for us to understand everything in complete detail and predict actual evolutions even in simplest situations, just because simple situations do not exist indeed. It is therefore not impossible that the cursor of 'who is wrong / who is right' could be some day shifted back toward Einstein in this Bohr-Einstein debate, although the complexity of the subject will probably not allow clear answers to emerge before decades or even centuries. But at least one may consider a small crack has reopened in Copenhagen dogma, a crack that still would make possible that, at the end, "God does not play dice with the Universe".

1. A justification for the presence of this personal digression in this memoir, is that at some time, around 2010, I officially asked my hierarchy to be allowed to work in the exploration of dual walkers behaviour by numerical means. This was refused as totally fanciful. Given the complexity of the subject and of underlying issues, it was maybe a wise decision...

8.2 Flow on the rotating sphere with $\eta = \eta_0$

In the case considered here, the surface S is a sphere of radius a , with a uniform angular rotation Ω with respect to fixed stars, around a ‘polar’ axis stationary with respect to fixed stars. Only the case with a constant depth $\eta = \eta_0$, corresponding to the standard BVE equation on the rotating sphere is considered here. The longitude and latitude are denoted λ and φ respectively. The governing equations are therefore

$$\frac{du}{dt} = -\frac{1}{a \cos \varphi} \frac{\partial(p/\rho_0)}{\partial \lambda} + fv \quad (8.31)$$

$$\frac{dv}{dt} = -\frac{1}{a} \frac{\partial(p/\rho_0)}{\partial \varphi} - fu \quad (8.32)$$

$$\frac{\partial u}{\partial \lambda} + \frac{\partial(v \cos \varphi)}{\partial \varphi} = 0 \quad (8.33)$$

where $f = (2\Omega \sin \varphi)$ is the Coriolis parameter. A streamfunction ψ is defined, so that

$$u = -\frac{1}{a} \frac{\partial \psi}{\partial \varphi} \quad (8.34)$$

$$v = \frac{1}{a \cos \varphi} \frac{\partial \psi}{\partial \lambda} \quad (8.35)$$

The relative vorticity is

$$\zeta = \frac{1}{a \cos \varphi} \left[\frac{\partial v}{\partial \lambda} - \frac{\partial(u \cos \varphi)}{\partial \varphi} \right] = \Delta \psi \quad (8.36)$$

and the potential vorticity is

$$q = \frac{1}{\eta_0} (\zeta + f). \quad (8.37)$$

The potential vorticity is conserved for a parcel, which leads to the following standalone equation for ψ

$$\frac{\partial \Delta \psi}{\partial t} + \frac{1}{a^2 \cos \varphi} \left(\frac{\partial \psi}{\partial \lambda} \frac{\partial \Delta \psi}{\partial \varphi} - \frac{\partial \psi}{\partial \varphi} \frac{\partial \Delta \psi}{\partial \lambda} \right) + \frac{2\Omega}{a^2} \frac{\partial \psi}{\partial \lambda} = 0 \quad (8.38)$$

From the formal point of view, this equation is exactly the same as for the classical BVE system, valid for a free atmosphere. This is not a surprise, since phenomenologically, the BVE system is derived from the shallow-water system under the additional constraint that the flow is maintained non-divergent at any time, which indeed is equivalent to impose a constant-height for the top of the free layer, which consequently is not really free. However, in the paradigm of the BVE system, the notion of pressure inside the fluid is absent as well as the notion of suction locally exerted by a vortex. Although the evolution is the same, both systems do not represent exactly the same conceptual situation.

8.2.1 Comparison with Charney’s equation

It is clear that in the case of two parallel spherical surfaces, the pressure perturbation p' field may not be used as a relevant parameter for a NWP forecast, since both a cyclonic vortex and an anticyclonic vortex produce a negative pressure perturbation. If the system is meant to predict the evolution of surface pressure of a shallow atmosphere with a free surface, another approach should be sought.

Charney used the following prognostic standalone equation

$$\frac{\partial \Delta z}{\partial t} + \frac{g}{f} \frac{1}{a^2 \cos \varphi} \left(\frac{\partial z}{\partial \lambda} \frac{\partial \Delta z}{\partial \varphi} - \frac{\partial z}{\partial \varphi} \frac{\partial \Delta z}{\partial \lambda} \right) + \frac{2\Omega}{a^2} \frac{\partial z}{\partial \lambda} = 0 \quad (8.39)$$

where z is an estimate of the height of the 500 hPa surface of the real world. Indeed, in Charney 1950, through a proper argumentation, this quantity z is just related to ζ by

$$\zeta = \frac{g}{f} \Delta z, \quad (8.40)$$

so that ultimately, z is simply related to ψ by

$$z = \frac{g}{f} \psi. \quad (8.41)$$

By inserting this latter expression in (8.39), an equation for ψ is obtained, and this equation is exactly the same as (8.38). So the evolution that would be given by Charney's system (filtered geopotential equation, FGE) would be the same as in the BVE system, if the aim was to predict the flow itself (ψ or ζ). But instead, the system predicts another quantity z .

A possible explanation for this choice is as follows. At these times, fields related to the column-averaged flow velocity were not possible to access experimentally, and only the surface pressure was easily available, from which, the height of the 500 hPa surface was diagnosed using standard profiles. This is maybe why Charney chose an equation that predicts the evolution of z . Note that Charney could very well (and equivalently) have performed the whole forecast with ψ as a prognostic variable, and at the last moment, translate the final ψ field into a height field by mean of (8.41). In this method, the initial ψ field would also have been diagnosed from the surface pressure field.

The word 'equivalently' used in the previous paragraph is not totally exact however. The traditional BVE equation (involving ψ) may be used on the whole sphere with no problem, and with a forecast range as long as wished. The modified equation used by Charney (involving z) may be used only in a limited-area domain not containing the equatorial area (since the spatial operators involved in the z equation are not defined at Equator, where f is zero). As a secondary consequence the equation may also not be used for an arbitrary forecast range, since in a long term, travelling vortices may cross the Equator.

In some sense, the choice of Charney for the governing equation may therefore seem a bit stupid. However, to mitigate this remark, one should underline that the aim of Charney was only to perform a 24-hour forecast, moreover with lateral boundary conditions (BCs) consisting in fixed a value for z , which amounts to impose rigid lateral BCs (walls), for which vortices coming from beyond the edges are not possible to represent. Finally, the restricted-validity equation used by Charney was relevant to the problem to solve. But, considering the full-validity equation is not more difficult to solve, the choice appears as strange, overall.

8.3 Flow on the rotating sphere with $\eta = \eta(\lambda, \varphi)$

The general case on a rotating sphere is derived in this section. The continuity equation is

$$\frac{\partial \eta u}{\partial \lambda} + \frac{\partial (\eta v \cos \varphi)}{\partial \varphi} = 0. \quad (8.42)$$

Hence, the vector field $\eta \mathbf{v}$ is non-divergent on the sphere, and a streamfunction ψ may be defined as follows

$$\eta u = -\frac{1}{a} \frac{\partial \eta \psi}{\partial \varphi} \quad (8.43)$$

$$\eta v = \frac{1}{a \cos \varphi} \frac{\partial \eta \psi}{\partial \lambda}. \quad (8.44)$$

The expression of the relative vorticity is therefore

$$\zeta = \frac{1}{a^2 \cos^2 \varphi} \frac{\partial}{\partial \lambda} \left(\frac{1}{\eta} \frac{\partial \eta \psi}{\partial \lambda} \right) + \frac{1}{a^2 \cos \varphi} \frac{\partial}{\partial \varphi} \left(\frac{\cos \varphi}{\eta} \frac{\partial \eta \psi}{\partial \varphi} \right) = \Delta_2 \psi \quad (8.45)$$

where, here again, Δ_2 is a modified Laplacian operator in which η plays the role of a metric factor :

$$\Delta_2 \psi = \frac{1}{a^2} \left[\frac{1}{\cos^2 \varphi} \frac{\partial}{\partial \lambda} \left(\frac{1}{\eta} \frac{\partial \eta \psi}{\partial \lambda} \right) + \frac{1}{\cos \varphi} \frac{\partial}{\partial \varphi} \left(\frac{\cos \varphi}{\eta} \frac{\partial \eta \psi}{\partial \varphi} \right) \right] \quad (8.46)$$

The potential vorticity is

$$q = \frac{1}{\eta}(\zeta + f). \quad (8.47)$$

The potential vorticity is conserved for a parcel, which leads to the following standalone equation for ψ

$$\frac{\partial \Delta_2 \psi}{\partial t} + \frac{1}{a^2 \cos \varphi} \left[\frac{\partial \eta \psi}{\partial \lambda} \frac{\partial}{\partial \varphi} \left(\frac{\Delta_2 \psi}{\eta} \right) - \frac{\partial \eta \psi}{\partial \varphi} \frac{\partial}{\partial \lambda} \left(\frac{\Delta_2 \psi}{\eta} \right) + \frac{\partial \eta \psi}{\partial \lambda} \frac{\partial}{\partial \varphi} \left(\frac{f}{\eta} \right) \right] = 0 \quad (8.48)$$

This equation describes a slightly more general situation than the classical BVE on the rotating sphere (8.38) since a representation of some ‘orography’ is allowed. A single vortex similar to (8.30) travelling along the Equator would for instance be distorted when passing a local perturbation of η located near its travelling path.

Chapitre 9

Appendice 2 : Gravitational potential in simple cases

9.1 Motivation

Unapproximated laws of Newtonian Mechanics are physically consistent. Consequently general equations of Fluid Mechanics as e.g. Euler Equations for a perfect fluid are also physically consistent. These equations may be written in general form, independent of a particular coordinate system and expressing conservation laws. When dealing with Meteorology, or Planetary Fluid Mechanics one must then specify some data, considered as given since they are not part of the atmospheric evolution (mainly the shape of the planet and of its gravity field, but possibly also some other details). Since the exact shape of actual Earth data is very complicated as well as the morphology of its gravitational field, simplifying approximations have to be applied, and the problem of ‘dynamical consistency’ appears right here. Note that the dynamical consistency discussed here is not the dynamical consistency of a discretization scheme, which is another —completely different— problem.

Dynamical consistency may then be defined as the proven existence of an analogue conservation law for any conservation law of the unapproximated system (being understood here that the conserved quantity and the shape of the conservation law in the approximated system must reasonably reflect unapproximated ones). Dynamical consistency with this definition may be proven with relatively little difficulty for simple cases (simple geometry of approximated planetary data ; simple coordinate systems...). However, showing the dynamical consistency of a mathematically-complicated approximation may rapidly become difficult, if not out of reach.

In some particular cases however, a simpler alternative avenue may be proposed, as explored below : if the approximated mathematical system may be shown to be formally equivalent to the mathematical system of an unapproximated problem (derived from a thought experiment), then the approximated system is necessarily dynamically consistent. This approach is illustrated here for simple but not uninteresting cases.

**

Consider a planet constituted by concentric spherical strata of heavy mater, the density of which is uniform inside each infinitesimal stratum, so that the mass-density variations are purely radial $\rho = \rho(r)$. As a consequence, the surface of the planet also has a spherical shape . It is well-known that the field of the Newtonian gravitational geopotential ϕ_N in the entire outer space of this planet is fully determined, upon specification of the surface value. This latter specification of the surface value just acts as an arbitrary and uninteresting gauge condition since ϕ_N is involved in physical processes only through its derivative. This result is also well-known in electrostatic, where it is usual to establish from Stokes’s Theorem that the outside electrostatic potential of a uniformly-charged sphere is exactly the same as if all the electric charge was concentrated at the centre.

The practical consequence of this well-known result is that if a spherical surface encloses all the mass or all the electric charge of a system, and is also isopotential, then one may completely ignore the details of the distribution hidden inside, since any distribution leading to this situation will provide exactly the same potential field in the outer space.

Returning to the gravity problem, a similar but less well-known result has been established by Stokes (1849) for spheroids : if the spheroidal shape of **one** isopotential surface is known and encloses all the heavy matter, then the geopotential field is fully-determined in the whole outer space of this spheroidal surface (still upon an arbitrary and uninteresting gauge condition). This implies that here again we don't need to know the hidden details of the inner distribution of mass leading to this situation, for being able to determine the structure of the geopotential field in the outer space. In other words, this means that any mass distribution making this given spheroidal surface isopotential, will generate the same Newtonian potential field in the outer space.

Armed with this solid result, one may seek simple mass distributions leading to the various approximations of geopotential fields and planet shapes used or useful in Meteorology, for which the 'dynamical consistency' has sometimes been questioned in the literature. For instance, it has often been questioned in the past whether assuming a spherical rotating planet with a non-uniform surface gravity (with a pole/equator gradient, as observed on Earth) is dynamically-consistent or not. Clearly, as stated in Bénard, 2014, the answer is 'yes', but publishing such a statement in a peer-reviewed publication required a struggle for weeks with the reviewers, who were in the 'no' camp, before becoming convinced. See White et al., 2005 (p. 2087) for an initial statment of inconsistency ; and Wood and White 2012 (section 6, paragraph 4) for a reiteration of the claim of inconsistency. See Bénard 2014, Bénard 2015, p. 196 (especially the footnote of the second column) for the claim of consistency. See Staniforth et White, 2015a (section 4.6, p. 661) for the U-turn in opinion, and a definitive and now shared consensus for the claim of consistency. Producing a concrete mass-distribution leading to this situation provides an irrefutable proof of the consistency of this approximation. The aim of this Appendix is to concretely examine this kind of situations on very simple and concrete examples.

9.2 Generalities

In this appendix, the first-order shape of geopotential surfaces is determined for some simple cases, in order to give a concrete illustration and justification of the various geometric approximations used in meteorology. The origin of space (representing the location of a planet centre) is denoted O , the polar axis of the planet is denoted Oz and the equatorial plane is described by (O, x, y) , as usual. The simplified cases considered here will consist in distributions of mass with zonal symmetry and concentrated in the vicinity of O , say, inside a small sphere of radius δ . Moreover we will only consider the geopotential field at 'large' distances of O , either in the equatorial plane or along the polar axis. Therefore, we will consider the geopotential field at points $(X, 0, 0)$ and $(0, 0, Z)$, with $(X, Z) \gg \delta$. The total mass of the distribution is denoted M .

For the simple cases considered here, due to the assumed zonal symmetry of mass distribution, and the large distance (X, Z) , geopotential surfaces are concentric spheres at zero-th order and concentric spheroids at first-order (although zonally-symmetric, these surfaces are generally **not** exact spheroids). Since we are only interested here by the first-order approximation of geopotential surfaces, the determination of the equatorial and polar radii of an isopotential surface is sufficient to claim that the shape of the whole geopotential surface is known at first-order : it is the unique zonally-symmetric spheroid possessing the same two radii as the actual isopotential surface. In the following, a and c will respectively denote the equatorial and polar radius of a geopotential surface, regardless of which of the two values is the largest.

In this Appendix, as in Chapter 4 we use the 'meteorological' convention for geopotential and gravity signs and radial variations : the geopotential ϕ is positive and increases above a planet surface, and the gravity g is positive and decreases to zero at infinite. The Newtonian geopotential around a single punctual mass M at O is

$$\phi_N(r) = \frac{GM}{\sqrt{x^2 + y^2 + z^2}} = \frac{GM}{r}, \quad (9.1)$$

and the (positive) gravity g is given by

$$g(r) = -\frac{\partial\phi}{\partial r} = \frac{GM}{r^2}. \quad (9.2)$$

Obviously, geopotential iso-surfaces (also called 'geopotential surfaces' or simply 'geopotentials' hereafter) are concentric spheres in this case, thus satisfying $a = c$ for any particular geopotential surface.

Now, for the assumed distribution of mass in the vicinity of the centre O , the Newtonian geopotential only depart from the above expression by a small amount, and the shape of a geopotential surface of equatorial radius a will slightly depart for the sphere of radius a , so that at pole the radius of this surface will be $c = a(1 - \varepsilon)$. The value of ε is a measure of the ellipticity (oblateness or prolateness) of each geopotential surface, approximated by a spheroid. If $\varepsilon > 0$, c is the minor axis of the (oblate) spheroid following the most usual notation, but if $\varepsilon < 0$, then the polar radius c becomes the major axis of the (prolate) spheroid, and the equatorial radius a becomes the minor axis.

For a frame with angular velocity Ω_{CEN} (around Oz axis, and with respect to fixed stars), the centrifugal geopotential is

$$\phi_C(x, y, z) = \frac{\Omega_{\text{CEN}}^2}{2} (x^2 + y^2).$$

In this rotating frame, the total (or apparent) geopotential is given by

$$\phi = \phi_N + \phi_C. \quad (9.3)$$

By similarity with a notation used in previous studies dealing with this topic, the following symbol is introduced :

$$m(a) = \Omega_{\text{CEN}}^2 a^3 / GM \quad (9.4)$$

This non-dimensional parameter $m(a)$ represents a measure of the ratio of centrifugal and Newtonian forces in the equatorial plane of a given geopotential surface of equatorial radius a . The only (slight) difference with the symbol m introduced in White et al. (2008), Bénard (2014) and subsequent papers on the topic, is that in these papers, m was a single number, valid at the surface of a planet, whereas here, the symbol is defined for any geopotential surface enclosing the distribution of mass (but indeed only valid at large distances of the centre O , i.e. valid for $a/\delta \gg 1$).

9.3 Spherical geopotential approximation

The spherical geopotential approximation (SGA) is the simplest approximation for global meteorology, and is used in almost all current operational NWP applications to date. This approximation is indeed twofold : the first point stipulates that the Earth and all its outer-space geopotential surfaces are concentric spheres ; the second point stipulates that the value of Coriolis parameter is the observed one. The consistency of these two aspects may be questioned since the first point is reminiscent of a non-rotating planet, whereas the second refers to a rotating planet. This particular question is examined here.

For a rotating frame, the centrifugal contribution ϕ_C implies increasing oblateness of apparent geopotential surfaces in the outer space when the distance to the centre O increases. Therefore a situation where all outer-space isopotential surfaces of the total geopotential ϕ are concentric spheres may occur in one only one case : the frame is non-rotating ($\phi_C = 0$) and the surface of the planet is spherical. If the planet is rotating and spherical, then any outer-space iso- ϕ surface will be oblate. Hence the hypothesized geopotential field ϕ of the SGA approximation corresponds to the conceptual model of a non-rotating planet with a spherical symmetry of all masses layers inside the planet.

Yet, SGA approximation retains Coriolis forces with the same magnitude as observed on the actual (rotating) Earth. Hence the governing equation retain terms such as fu or fv in spherical coordinates, and consequently we are forced to introduce a ‘Coriolis’ angular velocity Ω_{COR} , and set it to the actually-observed value. Finally when using SGA we end-up with a model of planet which involves two definitions of the angular velocity Ω_{CEN} and Ω_{COR} , with two non-identical values. Is this a problem for physical consistency ?

Not at all, and in order to demonstrate this, we may produce a simple (unapproximated) physical model in which such a situation may arise, thereby proving consistency. Let consider a non-rotating spherical planet plunged inside a uniform magnetic field \mathbf{B} parallel to its polar axis, and assume the atmosphere contains electrically-charged parcels all of them with a positive charge density ρ_e , proportional to the mass density ρ , as e.g. positive ions uniformly diluted in a neutral gaseous or liquid atmosphere, so that ρ_e/ρ is a constant (one may assume that the solid interior of the planet contains the counterbalancing

negative charge). Of course, the gravity field must be strong enough to keep the atmosphere around the planet, but such details are irrelevant here, and are simply ignored : the charge density may be assumed arbitrarily weak, and the magnetic field arbitrarily strong. The situation of this thought experiment is depicted in Fig. 9.1.

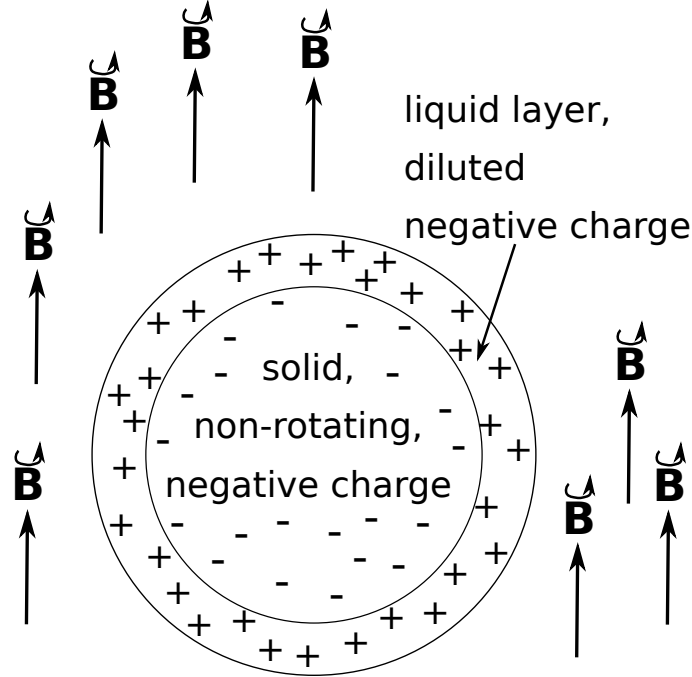


FIGURE 9.1 – Thought experiment to justify the dynamical consistency of setting $\Omega_{CENT} = 0$ together with $\Omega_{COR} \neq 0$ in SGA. A planet is plunged inside a uniform magnetic field \mathbf{B} (e.g. produced by a neighbouring star). The planet body is a non-rotating solid sphere with a uniform mass density and a uniform negative electric charge density. The ‘atmosphere’ is a thin layer of liquid or gaseous inviscid fluid, with also uniform densities of mass and electric charge, such that the total charge of the entire system is zero. The ratio of repulsive/attractive Coulomb forces to magnetic Lorentz forces may be made arbitrarily small by assuming arbitrary small electric charge densities and an arbitrarily large magnetic field magnitude B , so that *in fine* Coulomb forces may be simply neglected. Being purely radial, they may alternatively be incorporated into the apparent geopotential, to be exact. In this unapproximated system, governing equations are exactly identical to that of the SGA geometry in meteorology.

On this planet, the geopotential is exactly that of the SGA approximation, yet, the governing momentum equation for the atmosphere (e.g. $d\mathbf{v}/dt = \dots$) must include a (mass-specific) Lorentz force $(\rho_e/\rho) \mathbf{v} \times \mathbf{B}$, which is formally identical to the Coriolis force that would occur on a rotating planet $2\mathbf{v} \times \boldsymbol{\Omega}_{COR}$. Thus there is no fundamental inconsistency to consider governing equations with an exactly spherical geopotential field but with a non-zero Coriolis-type force imposed to atmospheric parcels.

From this discussion, we conclude that the values of angular velocity for centrifugal forces Ω_{CEN} and for Coriolis forces Ω_{COR} may be set separately in a completely arbitrary manner, and that they need not be identical. Setting two different values do not jeopardize physical consistency in any respect. This result has no reason to be particular to SGA geometry, and is admitted to hold for non-spherical geopotential fields as well.

9.4 Geopotential by two punctual masses on the polar axis

Two punctual masses $M/2$ are located at $z = \pm \delta$ along the polar axis. We determine the Newtonian geopotential ϕ_N at $(X, 0, 0)$ in the equatorial plane, and at $(0, 0, Z)$ on the polar axis, assuming these two remote points are located at far distance of the two masses, i.e. $(X, Z) \gg \delta$.

At point $(X, 0, 0)$ we have

$$\phi_N^E(X) = \frac{GM/2}{\sqrt{X^2 + \delta^2}} + \frac{GM/2}{\sqrt{X^2 + \delta^2}} = \frac{GM}{X} \frac{1}{\sqrt{1 + \delta^2/X^2}}$$

\Rightarrow

$$\phi_N^E(X) \approx \frac{GM}{X} \left(1 - \frac{1}{2} \frac{\delta^2}{X^2}\right)$$

At point $(0, 0, Z)$ we have

$$\phi_N^P(Z) = \frac{GM/2}{\sqrt{(Z - \delta)^2}} + \frac{GM/2}{\sqrt{(Z + \delta)^2}} = \frac{GM}{Z} \frac{1}{1 - \delta^2/Z^2}$$

\Rightarrow

$$\phi_N^P(Z) \approx \frac{GM}{Z} \left(1 + \frac{\delta^2}{Z^2}\right).$$

If the equatorial radius of a given geopotential surface is a , the ellipticity of this surface will be ε given by

$$\frac{GM}{a(1 - \varepsilon)} \left(1 + \frac{\delta^2}{a^2(1 - \varepsilon)^2}\right) = \frac{GM}{a} \left(1 - \frac{1}{2} \frac{\delta^2}{a^2}\right)$$

\Rightarrow

$$a^2(1 - \varepsilon)^2 + \delta^2 = (1 - \varepsilon)^3 \left(a^2 - \frac{1}{2} \delta^2\right)$$

and consequently, at first order

$$\varepsilon = -(1/2) \frac{\delta^2}{a^2}.$$

The ellipticity of the Newtonian geopotential surface of equatorial radius a is thus $\varepsilon = (-3/2)(\delta^2/a^2)$. The ellipticity is negative, so that, as expected, at large distance, Newtonian geopotential surfaces are close to prolate spheroids in this case. However, the prolateness decreases with distance a , so that at very large distances Newtonian geopotential surfaces tend toward concentric spheres. The Newtonian geopotential field is depicted in Fig. 9.2

**

In a rotating frame with angular velocity Ω , the total geopotential at points X and Z is therefore

$$\begin{aligned} \phi^E(X) &\approx \frac{GM}{X} \left(1 - \frac{1}{2} \frac{\delta^2}{X^2}\right) + \frac{\Omega_{\text{CEN}}^2 X^2}{2} \\ \phi^P(Z) &\approx \frac{GM}{Z} \left(1 + \frac{\delta^2}{Z^2}\right). \end{aligned} \quad (9.5)$$

In this rotating frame, the ellipticity ε of the geopotential surface possessing the equatorial radius a obeys

$$\phi^P[a(1 - \varepsilon)] = \phi^E(a)$$

\Rightarrow

$$\frac{GM}{a(1 - \varepsilon)} + \frac{GM\delta^2}{a^3(1 - \varepsilon)^3} = \frac{GM}{a} - \frac{1}{2} \frac{GM\delta^2}{a^3} + \frac{\Omega_{\text{CEN}}^2 a^2}{2}$$

\Rightarrow

$$a^2(1 - \varepsilon)^2 + \delta^2 = a^2(1 - \varepsilon)^3 - \frac{1}{2} \delta^2(1 - \varepsilon)^3 + \frac{m}{2} a^2(1 - \varepsilon)^3,$$

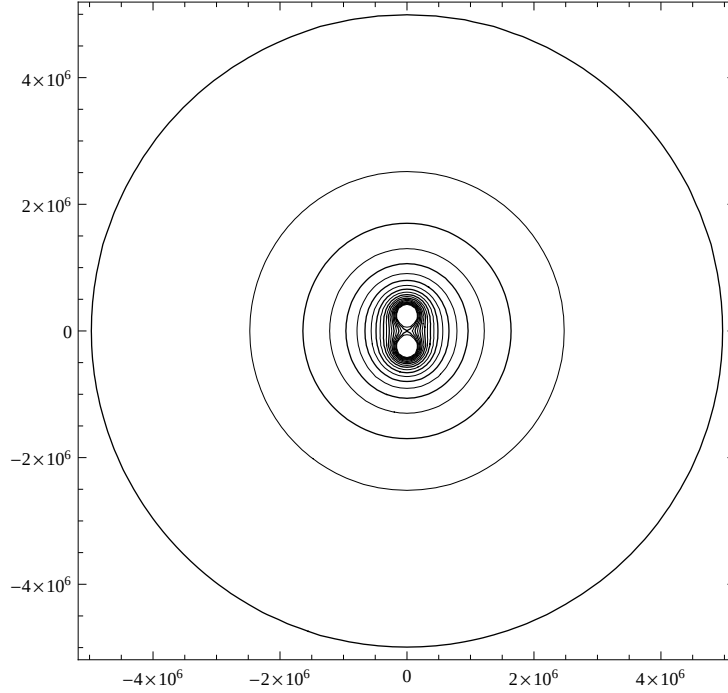


FIGURE 9.2 – Arbitrary contours –in (O, x, z) plane– of the Newtonian geopotential field ϕ_N created by the first mass distribution (two masses on the polar axis). Each mass is half the Earth’s mass, and is separated from O by $\delta = 263$ km. Note that a depiction of the total geopotential ϕ would lead to a graphically indistinguishable picture, for angular velocity magnitudes comparable to Earth’s one. The outermost contour, roughly corresponding to the Earth surface is a very slightly prolate spheroid for ϕ_N , so that when rotation is added, the corresponding surface for ϕ becomes a sphere (at first order).

where m is the non-dimensional ratio defined in (9.4). After neglecting high-order terms in m and δ^2/a^2 , the solution of this equation is, at first order

$$\varepsilon = -\frac{3}{2} \frac{\delta^2}{a^2} + \frac{m}{2}. \quad (9.6)$$

We see that (logically) the rotation contributes to make the originally-prolate geopotential surface of equatorial radius a less prolate, or even oblate if m is sufficiently large, which certainly occurs when a increases, so that at very large distance, all geopotential surfaces have an oblate shape.

The gravity along the x and z axes is given by

$$\begin{aligned} g^E(X) &= -d\phi(X)/dX \\ g^P(Z) &= -d\phi(Z)/dZ. \end{aligned}$$

Hence,

$$\begin{aligned} g^E(X) &= \frac{GM}{X^2} + \frac{3}{2} \frac{GM\delta^2}{X^4} - \Omega_{\text{CEN}}^2 X \\ g^P(Z) &= \frac{GM}{Z^2} + 3 \frac{GM\delta^2}{Z^4}. \end{aligned} \quad (9.7)$$

We may now determine the ratio of equatorial and polar gravity on any geopotential surface. The equatorial radius of the geopotential surface is a , hence along the x axis the gravity is

$$g^E(a) = \frac{GM}{a^2} - \frac{3}{2} \frac{GM\delta^2}{a^4} - \Omega_{\text{CEN}}^2 a = \frac{GM}{a^2} \left(1 - \frac{3}{2} \frac{\delta^2}{a^2} - m \right)$$

The polar radius of the surface is $c = a(1 - \varepsilon)$, that is $c = a \left[1 + (3/2)(\delta^2/a^2) - m/2 \right]$, hence along the z axis the gravity is

$$g^P(c) = \frac{GM}{c^2} - \frac{3}{2} \frac{GM\delta^2}{c^4}$$

Inserting the expression of $c = a(1 - \varepsilon)$, the gravity becomes, at first order in m and δ^2/a^2 ,

$$g^P(c) = \frac{GM}{a^2}(1 + m)$$

The ratio g^P/g^E on this geopotential surface (of equatorial radius a) is therefore

$$[g^P/g^E](a) = g(c)/g(a) = (1 + m) / \left(1 - \frac{3}{2} \frac{\delta^2}{a^2} - m \right) = 1 + 2m + \frac{3}{2} \frac{\delta^2}{a^2}$$

Finally, by inserting the expression (9.6) for ε

$$[g^P/g^E](a) = 1 + \frac{5}{2}m(a) - \varepsilon(a) \quad (9.8)$$

This equation is in agreement with Clairaut's (1743) Theorem, although it is valid for any value of a at large distance of the centre O , whereas, strictly speaking, Clairaut's originally established this relationship only at the surface of a rotating planet. However, this 'generalization' of Clairaut's Theorem to any geopotential surface has nothing surprizing, keeping in mind the result of Stokes (1849) mentioned above. As a consequence of Stokes' result, there is no distinction in nature between the planet surface in itself (which is the smallest geopotential surface enclosing all attracting material), and any other geopotential surface laying above the planet surface (which is just another –larger– surface enclosing all attracting material). After Stokes' statement, any mathematical property valid at the (isopotential) surface of a planet must also be valid at any other geopotential surface located above the ground. So must therefore be Clairaut's Theorem, as seen in (9.8).

9.5 Geopotential by a lineic density along an equatorial circle

The previous case in section (9.4) represented a zonally-symmetric distribution of mass with an excess along the polar axis. Here, by contrast, we consider a zonally-symmetric distribution of mass with an excess in the equatorial plane. The zonal symmetry therefore imposes that the simplest distribution is a lineic density ρ along a circle of radius δ , so that $M = 2\pi\rho\delta$, thereby ensuring a total mass M .

At point $(X, 0, 0)$ we have (see Fig. 9.3)

$$\phi_N^E(X) = \int_0^{2\pi} \frac{G\rho\delta d\lambda}{\sqrt{(X - \delta \cos \lambda)^2 + (\delta \sin \lambda)^2}} = \frac{G\rho\delta}{X} \int_0^{2\pi} \left(1 - 2\frac{\delta}{X} \cos \lambda + \frac{\delta^2}{X^2} \right)^{-1/2} d\lambda$$

\Rightarrow

$$\phi_N^E(X) \approx \frac{GM}{X} \left(1 + \frac{1}{4} \frac{\delta^2}{X^2} \right)$$

At point $(0, 0, Z)$ we have

$$\phi_N^P(Z) = \int_0^{2\pi} \frac{G\rho\delta d\lambda}{\sqrt{Z^2 + \delta^2}}$$

\Rightarrow

$$\phi_N^P(Z) \approx \frac{GM}{Z} \left(1 - \frac{1}{2} \frac{\delta^2}{Z^2} \right).$$

If the equatorial radius of a given geopotential surface is a , the ellipticity of this surface will be ε given by

$$\frac{GM}{a(1 - \varepsilon)} \left(1 - \frac{1}{2} \frac{\delta^2}{a^2(1 - \varepsilon)^2} \right) = \frac{GM}{a} \left(1 + \frac{1}{4} \frac{\delta^2}{a^2} \right)$$

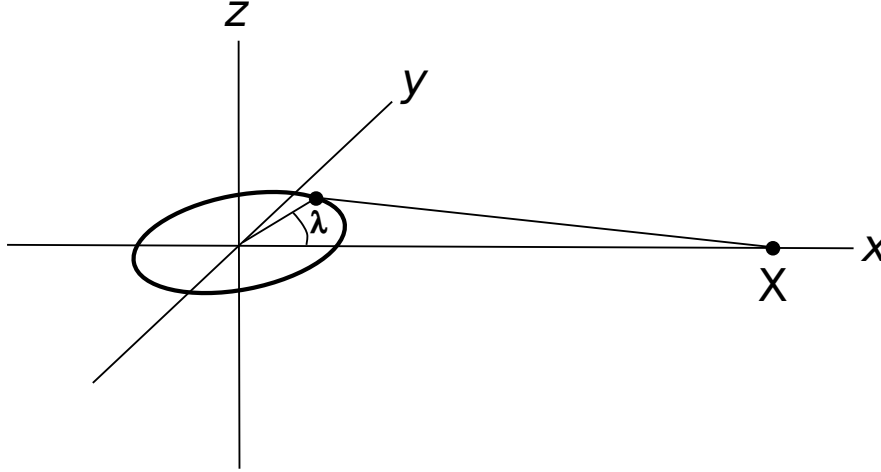


FIGURE 9.3 – Integral path along angle λ for the second mass distribution (depicted with a heavy solid line. The point where is determined the Newtonian geopotential is at $(X, 0, 0)$, in the right part of the figure.

\Rightarrow

$$a^2(1 - \varepsilon)^2 - \frac{1}{2} \delta^2 = (1 - \varepsilon)^3 \left(a^2 + \frac{1}{4} \delta^2 \right)$$

and consequently, at first order

$$\varepsilon = (3/4) \frac{\delta^2}{a^2}.$$

The ellipticity is positive, so that, as expected, at large distance geopotential surfaces are close to oblate spheroids in this case. However, the oblateness decreases with distance a , so that at very large distances geopotential surfaces tend toward concentric spheres. The Newtonian geopotential field is depicted in Fig. 9.4

**

In a rotating frame with angular velocity Ω_{CEN} , the total geopotential at points X and Z is therefore

$$\begin{aligned} \phi^E(X) &\approx \frac{GM}{X} \left(1 + \frac{1}{4} \frac{\delta^2}{X^2} \right) + \frac{\Omega_{\text{CEN}}^2 X^2}{2} \\ \phi^P(Z) &\approx \frac{GM}{Z} \left(1 - \frac{1}{2} \frac{\delta^2}{Z^2} \right). \end{aligned} \quad (9.9)$$

In this rotating frame, the ellipticity ε of the geopotential surface possessing the equatorial radius a obeys

$$\phi^P[a(1 - \varepsilon)] = \phi^E(a)$$

\Rightarrow

$$\frac{GM}{a(1 - \varepsilon)} - \frac{GM\delta^2}{2a^3(1 - \varepsilon)^3} = \frac{GM}{a} + \frac{1}{4} \frac{GM\delta^2}{a^3} + \frac{\Omega_{\text{CEN}}^2 a^2}{2}$$

\Rightarrow

$$a^2(1 - \varepsilon)^2 - \frac{1}{2} \delta^2 = a^2(1 - \varepsilon)^3 + \frac{1}{4} \delta^2(1 - \varepsilon)^3 + \frac{m}{2} a^2(1 - \varepsilon)^3,$$

After neglecting high-order terms in m and δ^2/a^2 , the solution of this equation is, at first order

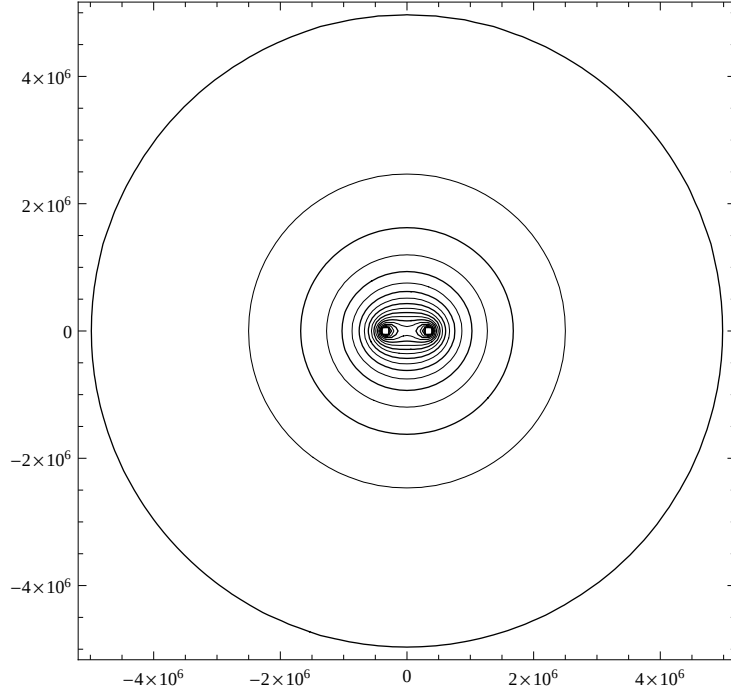


FIGURE 9.4 – Arbitrary contours –in (O, x, z) plane– of the Newtonian geopotential field ϕ_N created by the second mass distribution, along a circle in the equatorial plane. The total mass is the Earth’s mass, and the radius of the circle is $\delta = 348$ km. Note that a depiction of the total geopotential ϕ would lead to an indistinguishable picture, for angular velocity magnitudes comparable to Earth’s one. The outermost contour is almost an oblate spheroid of same shape as the Earth.

$$\varepsilon = \frac{3}{4} \frac{\delta^2}{a^2} + \frac{m}{2}. \quad (9.10)$$

We see that (logically) the rotation contributes to make the originally-oblate geopotential surface of equatorial radius a even more oblate at all distances. It may be noted also that the oblateness of ϕ_N surfaces always decrease with distance from O , and since on contrary, adding a rotation increases the oblateness with distance, a proper rotation of the frame makes possible that at some distance, the oblateness is locally-independent of the distance, i.e. the gravity is uniform on the whole geopotential surface.

The gravity along the x and z axes is found by differentiating $-\phi$ with respect to X and Z respectively

$$\begin{aligned} g^E(X) &= \frac{GM}{X^2} + \frac{3}{4} \frac{GM\delta^2}{X^4} - \Omega_{\text{CEN}}^2 X \\ g^P(Z) &= \frac{GM}{Z^2} - \frac{3}{2} \frac{GM\delta^2}{Z^4}. \end{aligned} \quad (9.11)$$

We may now determine the ratio of equatorial and polar gravity on any geopotential surface. The equatorial radius of the geopotential surface is a , hence along the x axis the gravity is

$$g^E(a) = \frac{GM}{a^2} + \frac{3}{4} \frac{GM\delta^2}{a^4} - \Omega_{\text{CEN}}^2 a = \frac{GM}{a^2} \left(1 + \frac{3}{4} \frac{\delta^2}{a^2} - m \right)$$

The polar radius of the surface is $c = a(1 - \varepsilon)$, that is $c = a [1 - (3/4)(\delta^2/a^2) - m/2]$, hence along the z axis the gravity is

$$g^P(c) = \frac{GM}{c^2} - \frac{3}{2} \frac{GM\delta^2}{c^4}$$

Inserting the expression of c the gravity becomes, at first order in m and δ^2/a^2 ,

$$g^P(c) = -\frac{GM}{a^2}(1+m)$$

The ratio g^P/g^E on this geopotential surface (of equatorial radius a) is therefore

$$[g^P/g^E](a) = g(c)/g(a) = (1+m)/\left(1 + \frac{3\delta^2}{4a^2} - m\right) = 1 + 2m - \frac{3\delta^2}{4a^2}$$

Finally, by inserting the expression (9.10) for ε , we obtain, here again, the generalized result of Clairaut :

$$[g^P/g^E](a) = 1 + \frac{5}{2}m(a) - \varepsilon(a) \quad (9.12)$$

9.6 Discussion

Let denote by a_{\oplus} , c_{\oplus} , Ω_{\oplus} and GM_{\oplus} the actual observed values of Earth's equatorial and polar radii, angular velocity and GM respectively. Observed values are

$$\begin{aligned} a_{\oplus} &\approx 6378137. \text{ m} \\ c_{\oplus} &\approx 6356752. \text{ m} \\ \varepsilon_{\oplus} &\approx 0.00335 \\ \Omega_{\oplus} &\approx 7.292 \times 10^{-5} \text{ rad/s} \\ GM_{\oplus} &\approx 3.985 \times 10^{14} \text{ m}^3 \cdot \text{s}^{-2} \\ g_{\oplus}^E &\approx 9.78 \text{ m} \cdot \text{s}^{-2} \\ g_{\oplus}^P &\approx 9.83 \text{ m} \cdot \text{s}^{-2} \end{aligned}$$

where g_{\oplus}^E and g_{\oplus}^P denote the observed terrestrial surface values of gravity at equator and poles respectively. We now consider the situation of the first mass distribution (with two masses on the polar axis). As seen above, the ellipticity of geopotential surfaces for this mass distribution is

$$\varepsilon = -\frac{3\delta^2}{2a^2} + \frac{\Omega_{\text{CEN}}^2 a^3}{2GM}$$

For the observed value of GM_{\oplus} , the geopotential surface with equatorial radius a_{\oplus} is –at first order– a spheroid with vanishing ellipticity (i.e. a sphere, with $a = c$) if

$$\delta = \left[\frac{\Omega_{\text{CEN}}^2 a_{\oplus}^5}{3GM_{\oplus}} \right]^{1/2},$$

One may now determine the values of Ω_{CEN} and δ needed to obtain the observed Earth values of g_{\oplus}^E and g_{\oplus}^P . For this, one must have (exploiting $a = c$)

$$\begin{aligned} g_{\oplus}^E &= \frac{GM_{\oplus}}{a_{\oplus}^2} \left(1 + \frac{\Omega_{\text{CEN}}^2 a_{\oplus}^3}{2GM_{\oplus}} \right) - \Omega_{\text{CEN}}^2 a_{\oplus} \\ g_{\oplus}^P &= \frac{GM_{\oplus}}{a_{\oplus}^2} \left(1 + \frac{\Omega_{\text{CEN}}^2 a_{\oplus}^3}{2GM_{\oplus}} \right) \end{aligned}$$

⇒

$$\Omega_{\text{CEN}} = \left(\frac{g_{\oplus}^P - g_{\oplus}^E}{a_{\oplus}} \right)^{1/2},$$

The resulting numerical values are

$$\delta \approx 263 \text{ km}$$

$$\Omega_{\text{CEN}} \approx 8.8 \times 10^{-5} \text{ rad. s}^{-1}$$

Hence it has been shown here that two identical mass-distributions with spherical symmetry slightly separated along the polar axis, and a centrifugal angular velocity close to the observed one, lead to a spherical planet with variations of the surface gravity field in agreement with observations. Since this situation might be encountered in the Nature, this proves the consistency of the controversial approximation where the planet surface is spherical, but the surface gravity is not uniform and varies according to the observations. If realistic meteorological forecasts are to be performed, the value of Ω_{COR} must be set, independently, to the observed value Ω_{\oplus} .

**

Now, the second mass distribution is considered, and the conditions for an oblate-spheroidal planet (a, c) with a uniform gravity are sought. For a spheroidal planet in the rotating frame, the relationship between ε , δ and Ω_{CEN} is

$$\varepsilon = -\frac{3}{2} \frac{\delta^2}{a^2} + \frac{\Omega_{\text{CEN}}^2 a^3}{2GM}. \quad (9.13)$$

The requirement that pole and equator surface gravity are equal leads to

$$\left(1 + \frac{3}{4} \frac{\delta^2}{a^2} - m\right) = (1 + m)$$

\Rightarrow

$$m = \frac{3}{8} \frac{\delta^2}{a^2}$$

Inserting this value in (9.13) provides the value of the radius of the mass-distribution

$$\delta = \left(\frac{8a_{\oplus}^2 \varepsilon}{9}\right)^{1/2}$$

The resulting numerical values are

$$\delta \approx 348 \text{ km}$$

$$\Omega_{\text{CEN}} \approx 4.1 \times 10^{-5} \text{ rad. s}^{-1}$$

When $\Omega_{\text{CEN}} = 0$, the oblateness of geopotential surfaces decreases with distance to the centre O , so that very far from O , geopotentials tend to spheres. In this case, the surface gravity at pole is smaller than its equator counterpart. In order to restore identical values at pole and equator, a non-zero value of Ω_{CEN} is required, but smaller than the observed value, since this observed value eventually leads to a larger polar gravity than the equatorial gravity, as observed.

Hence it has been shown here that a zonally-uniform density of mass-distributions with spherical symmetry but slightly separated from the centre O in the equatorial plane, and a centrifugal angular velocity somewhat smaller than the observed one lead to a spheroidal planet of similar oblate shape as the Earth, but with a quasi-uniform value of the surface gravity. Since this situation might be encountered in the Nature, this proves the consistency of this strange approximation where the planet surface is oblate spheroidal, but the surface gravity is uniform. Here also if realistic meteorological forecasts are to be performed, the value of Ω_{COR} must be set, independently, to the observed value Ω_{\oplus} .

**

Finally a quartet of geopotential approximations useful for meteorology, depicted in Fig. 9.5 may be set as follows :

- (i) All geopotential surfaces are concentric spheres (at planet's surface and in the whole outer space. This is the SGA approximation, which requires to set $\Omega_{\text{CEN}} = 0$ and $\Omega_{\text{COR}} \neq 0$. It has been shown that this 'discrepancy' is not a problem for dynamical consistency. The surface gravity is a constant. In this case spherical coordinates makes the expression of governing equations straightforward.
- (ii) The planet is exactly spherical, but outer-space geopotential surfaces are concentric oblate spheroids. The variation of oblateness is set so as to ensure that surface gravity is exactly the observed one. In this case, Ω_{CEN} has to be set to a value somewhat larger than the observed one (and here again Ω_{COR} , set to the observed value, has to depart from Ω_{CEN}). Since the geometry of the planet surface is exactly spherical, two sub-cases are to be considered. The first sub-case is the shallow-atmosphere (SA) approximation, where the metric in the outer-space is taken identical to the surface metric. Then spherical coordinates may still be used, and the expression of governing equations remain straightforward. In the sub-case of the deep-atmosphere (DA) framework, governing equations must be expressed in a non-spherical system of coordinates (e.g. GREAT coordinate, see Staniforth 2015b) if one want to stick to the usual guideline of orthogonal curvilinear coordinates which follow horizontal and vertical directions. Note, that it is always possible to express governing equations in a spherical coordinate system, but then gravity force will have non-zero components along \mathbf{e}_λ and \mathbf{e}_φ directions, which could arguably make an accurate expression of discrete evolution equations possibly problematic.
- (iii) The planet has the observed oblate spheroid figure, but the surface gravity is uniform. This requires a smaller value of Ω_{CEN} than the actual one. Here, an expression of governing equations in spheroidal coordinates seems the most logical approach, whatever SA or DA framework is chosen.
- (iv) The planet has the same figure as the actual one and the pole/equator ratio of surface gravity is the same as the actual one as well. In this case Ω_{CEN} and Ω_{COR} both have the same value as the actual one. Here again, an expression of governing equations in spheroidal coordinates seems the most logical approach, whatever SA or DA framework is chosen. This approach is termed EGA approximation (for ellipsoidal geopotential approximation). Although more accurate than the SGA framework, EGA is still an approximation because in reality none of the geopotential surface is an exact spheroid.

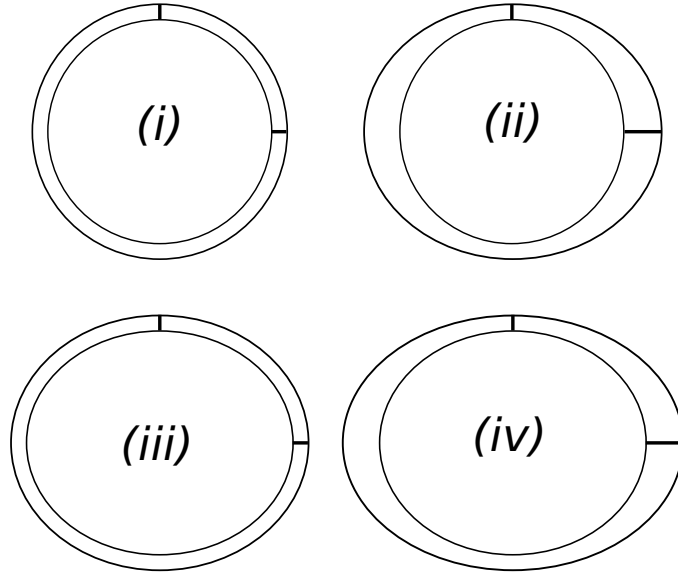


FIGURE 9.5 – Quartet of dynamically-consistent geopotential approximations of interest for NWP, as discussed in the text. The inner closed curve represents the planet surface, and the outer one represents an outer-space geopotential surface (e.g. the mean free-surface for a shallow-water system, or any mean geopotential surface as $\phi_{500}, \phi_{250}, \dots$ for a 3D system). Oblateness of spheroids has been exaggerated for graphical purpose. The intensity of surface gravity at pole and equator is illustrated by the length of short thick line segments (shorter lengths imply larger values of surface gravity)

Case (i) is the most widely used in NWP, because it is the simplest one, and the improvement resulting from refining the framework would still be speculative (although Bénard 2015 showed that in long-range forecasts, the impact in quality might begin to become detectable). Case (ii) and (iii) are mainly of scientific interest, since they may be used to evaluate which of two factors (the oblateness of Earth's surface, and the pole/equator variation of gravity) has the biggest expected impact on the quality of NWP forecasts (especially in long-range applications). In this point of view, the overall improvement when passing from (i) to (iv) may be splitted into two partial improvements, i.e. from (i) to (ii), and from (i) to (iii). Therefore, this couple of approximations (ii)-(iii) may be not only of scientific interest, but also of practical interest because if gravity variations reveals to be preponderant (or at least not negligible) in the total improvement, then the possibility of improving long-range NWP forecasts by using a spherical planet with varying surface gravity could be considered. Once again, this modification is straightforward because spherical coordinates may still be used in the same way as in the SGA approximation, with all its advantages.

References

- Baines, P, 1976 : The stability of planetary waves on a sphere. *J. Fluid Mech.*, **73**, 193-213.
- Bénard, P, Marki, A, Neytchev, PN, and Prtenjak, MT, 2000 : Stabilization of nonlinear vertical diffusion schemes in the context of NWP models. *Mon. Wea. Rev.*, **128**, 1937-1948.
- Bénard, P., 2003 : Stability of Semi-Implicit and Iterative Centered-Implicit Time Discretizations for Various Equation Systems Used in NWP. *Mon. Wea. Rev.*, **131**, 2479-2491.
- Bénard, P, R Laprise, J Vivoda, P Smolíková, 2004 : Stability of Leap-Frog Constant-Coefficients Semi-Implicit Schemes for the Fully Elastic System of Euler Equations. Flat-Terrain Case. *Mon. Wea. Rev.*, **132**, 1306-1318.
- Bénard, P., 2004 : On the Use of a Wider Class of Linear Systems for the Design of Constant-Coefficients Semi-Implicit Time-Schemes in NWP. *Mon. Wea. Rev.*, **132**, 1319-1324.
- Bénard, P, J Mašek, and J Vivoda, 2005 : Stability of leap-frog constant-coefficients semi-implicit schemes for the fully elastic system of Euler equations. Case with orography. *Mon. Wea. Rev.*, **133**, 1065-1075.
- Bénard, P, J Vivoda, and J. Mašek, P. Smolíková, K. Yessad, Ch. Smith, R. Brožková and J.-F. Geleyn, 2010 : Dynamical kernel of the Aladin-NH spectral limited-area model : revised formulation and sensitivity experiments *Q. J. R. Meteorol. Soc.* , **136** , 155-169.
- Bénard, P, 2014 : An oblate-spheroid geopotential approximation for global meteorology. *Q. J. R. Meteorol. Soc.*, **140**, 170-184.
- Bénard, P, 2015 : An assessment of global forecast errors due to the spherical geopotential approximation in the shallow-water case. *Q. J. R. Meteorol. Soc.*, **141**, 195-206.
- Bénard, P, and Ginton, MR, 2019 : Circumventing the pole problem of reduced lat/lon grids with local schemes. Part I : Analysis and model formulation. *Q. J. R. Meteorol. Soc.*, **145**, 1377-1391.
- Bénard, P, 2019a : Numerical investigation of Rossby waves for nonlinear Shallow-Water equations on the sphere. *Q. J. R. Meteorol. Soc.*, **145**, 1461–1473.
- Bénard, P, 2020 : Stability of Rossby-Haurwitz waves. *Q. J. R. Meteorol. Soc.*, **146**, 613– 628
- Bouttier, F and Buizza, R, 2018 : La prévision d'ensemble. *La Météorologie* **100**, 29-36.
- Bubnová, R, G Hello, P Bénard, and J-F Geleyn, 1995 : Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. *Mon. Wea. Rev.*, **123**, 515-535.
- Bush, JWM, Y Couder, T Gilet, PA Milewski, and A Nachbin, 2018 : Introduction to focus issue on hydrodynamic quantum analogs. *Chaos*, **096001**, doi : 10.1063/1.5055383 .
- Charney, J. G. ; Fjørtoft, R. ; von Neumann, J., 1950 : Numerical Integration of the Barotropic Vorticity Equation. *Tellus*, **2**, 237-254.
- Charron, M, Zadra, A, and Girard, C, 2014 : Four-dimensional tensor equations for a classical fluid in an external gravitational field. *Q. J. R. Meteorol. Soc.*, **140**, 908-916.
- Charron, M and Zadra, A, 2014 : On the dynamical consistency of geometrically approximated equations for geophysical fluids in arbitrary coordinates. *Q. J. R. Meteorol. Soc.*, **140**, 2078-2083.
- Coiffier, J, Ph Chapelet, and N Marie, 1987 : Study of various quasi-Lagrangian techniques for numerical models. *Proc. ECMWF Workshop*, Shinfield Park, Reading, United Kingdom, ECMWF, 1987, 19-46.
- Coiffier, 2011 : Fundamentals of Numerical Weather Prediction *Cambridge University press*, 340 pp.

- Côté, J, and Staniforth, A, 1988 : A two-time-level semi-Lagrangian semi-implicit scheme for spectral models. *Mon. wea. rev.*, **116**, 2003-2012.
- Côté, J, M Béland, and A Staniforth, 1983 : Stability of vertical discretization schemes for semi-implicit primitive equation models : theory and application. *Mon. Wea. Rev.*, **111**, 1189-1207.
- Courtier, Ph, and Geleyn, J-F, 1988 : A global numerical weather prediction model with variable resolution : Application to the shallow water equations. *Q. J. R. Meteorol. Soc.*, **114**, 1321-1346.
- Courtier, P, and Naughton, M, 1994. A pole problem in the reduced Gaussian grid. *Q. J. R. Meteorol. Soc.*, **120**, 1389-1407.
- Cullen, MJP, 2001 : Alternative implementations of the semi-Lagrangian semi-implicit schemes in the ECMWF model. *Q. J. R. Meteorol. Soc.*, **127**, 2787-2802.
- Cullen, MJP, T Davies, MH Mawson, J. A. James and S. C. Coulter, 1997 : An overview of numerical methods for the next generation UK NWP and climate model. *Numerical Methods in Atmospheric Modelling*, Canadian Meteorological and Oceanographical Society, 581 pp.
- Feynman, RP, RB Leighton and M Sands, 1965 : The Feynman Lectures on Physics; Vol. III, Quantum Mechanics. Addison-Wesley, Reading, Massachusetts.
- Galewsky J, Scott RK, Polvani LM, 2004 : An initial-value problem for testing numerical models of the global shallow-water equations. *Tellus A*, **56**, 429-440.
- Gates WL, 2004 : Derivation of the equations of atmospheric motion in oblate spheroidal coordinates. *J. Atmos. Sci.*, **61**, 2478-2487.
- Gill, AE, 1982 : Atmosphere-Ocean dynamics (*International Geophysics Series*). Academic Press.
- Girard, C., and Delage, Y., 1990 : Stable schemes for nonlinear vertical diffusion in atmospheric circulation models. *Mon. wea. rev.*, **118**, 737-745.
- Glinton, MR and Bénard, P, 2019 : Circumventing the pole problem of reduced lat-lon grids with local schemes. Part II : Validation experiments. *Q. J. R. Meteorol. Soc.*, **145**, 1392-1405.
- Haurwitz, B, 1940 : The motion of atmospheric disturbances on the spherical Earth. *J. Mar. Research*, **3**, 254-267.
- Hoskins, BJ, 1973 : Stability of the Rossby-Haurwitz wave. *Q. J. R. Meteorol. Soc.*, **99**, 723-745.
- Hoskins, BJ, and Simmons, AJ, 1975 : A multi-layer spectral model and the semi-implicit method. *Q. J. R. Meteorol. Soc.*, **101**, 637-655.
- Kalnay, E, and Kanamitsu, M, 1988 : Time schemes for strongly nonlinear damping equations. *Mon. wea. rev.*, **116**, 1945-1958.
- Kwizak, M, and Robert, AJ, 1971 : A semi-implicit scheme for grid point atmospheric models of the primitive equations. *Mon. Wea. Rev.*, **99**, 32-36.
- Lorenz, EN, 1972 : Barotropic instability of Rossby wave motion. *J. Atm. Sci.*, **29**, 258-265
- Lynch, P, 2004 : Richardson's Forecast : What Went Wrong? *Symposium on the 50th Anniversary of Operational Numerical Weather Prediction*, University of Maryland, College Park, MD, 14-17 June 2004.
- Marchuk, GI, 1965 : A new approach to the numerical solution of differential equations of atmospheric processes. *W.M.O. Technical Note*, **66**, 286-294.
- Morrison, PJ, 1988 : Hamiltonian description of the ideal fluid *Rev. Mod. Phys.*, **70**, 467-521
- Phillips, NA, 1959 : Numerical integration of the primitive equations on the hemisphere. *Mon. Wea. Rev.*, **87**, 333-345.
- Qian, J-H, F. HM Semazzi, and JS Scroggs, 1998 : A global nonhydrostatic semi-Lagrangian atmospheric model with orography. *Mon. Wea. Rev.*, **126**, 747-771.
- Richtmyer. RD, 1957 : Difference methods for initial-value problems. *Interscience*.
- Robert, AJ, J Henderson, and C Turnbull, 1972 : An implicit time integration scheme for baroclinic models of the atmosphere. *Mon. Wea. Rev.*, **100**, 329-335.
- Rossby, CG and collaborators, 1939 : Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *J. Marine Sci.*, **2**, 38-55.

- Scott, RK, Harris, LM and Polvani, LM, 2016 : A test case for the inviscid shallow-water equations on the sphere. *Q.J.R. Meteorol. Soc.*, **142**, 488-495.
- Salmon, R, 1988 : Hamiltonian fluid dynamics. *Ann. Rev. Fluid. Mech.*, **20**, 225-256.
- Schmidt, F, 1977 : Variable fine mesh in spectral global models. *Beitr. Phys. Atmos.*, **50**, 211-217.
- Shuman, FG, 1970 : On certain truncation errors associated with spherical coordinates. *J. Appl. Meteor.*, **9**, 564-570.
- Seity, Y, Brousseau, P, Malardel, S, Hello, G, Bénard, P, Bouttier, F, Lac, Ch and V Masson , 2011 : The AROME-France convective scale operational model *Mon. Wea. Rev.* , **139** , 976-991.
- Simmons, AJ, B Hoskins, and D Burridge, 1978 : Stability of the semi-implicit method of time integration. *Mon. Wea. Rev.*, **106**, 405-412.
- Simmons, AJ, and D Burridge, 1981 : An Energy and Angular-Momentum Conserving Vertical Finite-Difference Scheme and Hybrid Vertical Coordinates. *Mon. Wea. Rev.*, **109**, 758-766.
- Simmons, AJ, C Temperton, 1997 : Stability of a two-time-level semi-implicit integration scheme for gravity wave motion. *Mon. Wea. Rev.*, **125**, 600-615.
- Skamarock, WC, PK Smolarkiewicz, and JB Klemp, 1997 : Preconditioned conjugate-residual solvers for Helmholtz equations in nonhydrostatic models. *Mon. Wea. Rev.*, **125**, 587-599.
- Staniforth, A and Wood, N. (2005). An Unsuspected Boundary-Induced Temporal Computational Mode in a Two-Time-Level Discretization. Monthly Weather Review - *Mon. Wea. Rev.*, **133**, 712-720.
- Staniforth, A and White, A, 2015a : The shallow-water equations in non-spherical geometry with latitudinal variation of gravity. *Q.J.R. Meteorol. Soc.*, **141**, 655-662.
- Staniforth, A, and White, A, 2015b : Geophysically Realistic, Ellipsoidal, Analytically Tractable (GREAT) coordinates for atmospheric and oceanic modelling. *Q.J.R. Meteorol. Soc.*, **141**, 1646-1657.
- Stokes GG. 1849. On the variation of gravity at the surface of the Earth. *Transactions of the Cambridge Philosophical Society*, **8** : 672-695. Available at : <http://books.google.co.uk/books?id=2vsIAAAAIAAJ>
- Tanguay, M, A Robert, and R Laprise, 1990 : A semi-implicit semi-Lagrangian fully compressible regional forecast model. *Mon. Wea. Rev.*, **118**, 1970-1980.
- Temperton, C, and Staniforth, A, 1987 : An efficient two-time-level semi-Lagrangian semi-implicit integration scheme. *Q. J. R. Meteorol. Soc.*, **113**, 1025-1039.
- Temperton, C, Hortal, M, and Simmons, A, 2001 : A two-time-level semi-Lagrangian global spectral model. *Q. J. R. Meteorol. Soc.*, **127**, 111-127.
- Thomas, SJ, C Girard, R Benoit, M Desgagné, and P Pellerin, 1998 : A new adiabatic kernel for the MC2 model. *Atmos. Ocean*, **36 (3)**, 241-270.
- Thuburn, J and Li, Y, 2000 : Numerical simulations of Rossby-Haurwitz waves. *Tellus A*, **52**, 181-189.
- Tort, M, and Dubos, T, 2014 : Dynamically consistent shallow atmosphere equations with a complete Coriolis force. *Q. J. R. Meteorol. Soc.*, **140**, 2388-2392.
- Tort, M, and Dubos, T, 2014 : Usual approximations to the equations of atmospheric motion : A variational perspective. *J. Atmos. Sci.*, **71**, 2452-2466.
- Veronis, G, 1973 : Large scale ocean circulation. *Adv. Applied Mech.*, **13**, 92 pages.
- Voitus, 2004
- White, AA, Hoskins, BJ, Roulstone, I, and Staniforth, A, 2005 : Consistent approximate models of the global atmosphere : shallow, deep, hydrostatic, quasi-hydrostatic and non-hydrostatic. *Q. J. R. Meteorol. Soc.*, **131**, 2081-2107.
- White AA, Staniforth A, Wood N 2008 (WSW08) : Spheroidal coordinate systems for modelling global atmospheres. *Q. J. R. Meteorol. Soc.*, **134**, 261-270.
- White, AA and Wood, N, 2012 : Consistent approximate models of the global atmosphere in non-spherical geopotential coordinates. *Q. J. R. Meteorol. Soc.*, **138**, 980-988.

- Williamson DL, Drake JB, Hack JJ, Rüdiger J, Swarztrauber PN, 1992 : A standard test set for numerical approximations to the shallow water equations in spherical geometry. *J. Comput. Phys.*, **102**, 211-224.
- Yessad, K, and Bénard, P,1996 : Introduction of a local mapping factor in the spectral part of the Météo-France global variable mesh numerical forecast model. *Q. J. R. Meteorol. Soc.*, **122**, 1701-1719.

Publications essentielles

- Yessad and Bénard, 1996
- Bénard, 2003
- Bénard, 2004
- Bénard et al., 2004
- Bénard et al., 2010