ADAPTATION OF SPECTRAL METHODS TO NON-UNIFORM MAPPING (GLOBAL AND LOCAL)

J-F Geleyn
Météo-France - CNRM/GMAP
Toulouse, France

Summary: This lecture note is serving two purposes. First it attempts to be a brief introduction to spectral methods in general, insisting more on their links with other aspects of NWP than on their intrinsic properties. Second it shows, theoretically as well as practically, how they can be applied to two problems of non-uniform mapping-type, namely a global zoom and an imbedded LAM.

1. INTRODUCTION

1.1 Foreword

The structure of this lecture note is not very usual. It is neither a full course on spectral methods with historical background, full mathematical justifications of their properties and details about their algorithmics nor a justification of their operational use through comparative results with respect to other horizontal discretisation techniques. Indeed it was thought that spectral methods, at least in their global homogeneous «classical» variant, are now rather well known and that their use in a clear majority of global NWP or GCM applications speaks for itself. We rather elected to concentrate on two aspects: (i) the link with other numerical constraints in NWP (time schemes and horizontal diffusion; parameterisation schemes; data assimilation) and, as the core part of this lecture, (ii) the more recent use of spectral techniques for inhomogeneous representations of the earth, be it globally (zoom) or locally (LAM). Of course, some quick review of the basic properties of spectral modelling is still necessary, even in this particular framework; we simply tried to keep it to a minimum and to make it as intuitive as possible. For instance neither precision, convergence nor consistency will be studied here; the same applies to the cumbersome question of the wind vector representation and of the associated choice of spectral transforms. To justify these by-passes let us recall here the words of *Temperton* (1991a) in an otherwise very specialised article: «*This illustrates one of the nice features of the spectral method: although there is certainly more than one way to organise the computations, there is fundamentally no argument about what is to be done*»! On may indeed wonder whether the observed very strong pro and con opinions with respect to spectral methods do not simply depend whether one feels relieved or jobless at the idea not to handle the link between a Jacobian's stencil and energy or enstrophy conservation!

The reader interested in more details about the intrinsic properties of the spectral technique is anyhow referred to two previous ECMWF Seminar Lectures (*Jarraud and Simmons*, 1983; *Machenhauer*, 1991) and to technical documents like *Machenhauer* (1979) or *Rochas and Courtier* (1992).

Given the above-mentioned choices about the subject of this lecture, it was rather natural to try and mix all types of information, from basic equations up to very practical implementation details. Since the author has been heavily involved with both (zoom and LAM) parts of operational non-uniform spectral applications, the practical aspects will be drawn from experiences obtained with ARPEGE (the «French» side of the joint ECMWF/Météo-France IFS/ARPEGE project, operational in its variable mesh version since October 1993, *Bénichou and Legrand* (1992)) and ALADIN (the LAM counterpart of ARPEGE, operational in several partner National Meteorological Services from 1996 onwards, *Members of the ALADIN international team* (1997)). This deliberate choice to link the obviously pro-spectral argumentation of the discussions with operational aspects is in itself a guarantee of sincerity: nothing will be said that has not been put to the acid test of a decision process with operational consequences.

The associated drawback for the lecture will be a rather fragmented structure (19 Sections!) and the implied need to frequently anticipate some results to be justified later. In this sense, the main tutorial challenge is to keep the reader interested enough to reach the point where an overview of the patchwork of rather simple items will hopefully start to make sense. If this fails, let us hope that at least a few of the most practical aspects of the text will be of some use.

Finally, the question about the adequacy of spectral methods to distributed memory computers will not be treated here. Indeed the proof made by ECMWF when converting IFS to its new computing platform was hopefully convincing enough to stop a debate that should never have been launched, but for the above-mentioned too emotional reactions prompted by the word «spectral».

1.2 Basic theorem

The use of a spectral method simply amounts to do grid-point calculations with an infinite precision in the computation of derivatives and to automatically eliminate aliasing-generated noise.

1.3 Classification of the space discretisation methods for partial derivative equations

1.3.1 *Question number one:*

Do we express the searched solution as the expansion of a sum of functions defined everywhere (Galerkin method) or do we express the solution only at the computational nodes (discrete method)?

Remark 1: a discrete method may be considered as a Galerkin method using Dirac functions, but this semantic exercise does not have any practical application;

Remark 2: by construction, a Galerkin method will always involve some inversion of a linear operator, in opposition to the basic explicit character of a discrete method;

Remark 3: in principle, in the case of a Galerkin method, most operators necessary for the solution of meteorological equations (so-called «primitive equations» or Euler equations) could be computed in the mathematical space of the basic functions the combination of which creates the resulting fields; however this «interaction coefficients» method is highly inefficient; Galerkin methods started to become useful when *Eliasen et al.* (1970) and *Orszag* (1970) independently proposed the «transform method» recalled in the above-mentioned theorem, the first practical full application being that of *Bourke* (1974).

1.3.2 *Question number two:*

Do we seek to have a global (i.e. the whole computational domain) or a local representation of the discretisation operators?

Remark 4: when one uses local methods, the question of accuracy is mainly determined through the order of the approximation, i.e. the choice of a computational stencil; this problem automatically disappears in spectral methods where the equivalent degree of freedom (the choice of the basic set of functions) is removed because of the central character of the horizontal Laplacian operator in meteorological equations;

Remark 5: one usually makes a confusion between so-called «grid-point» methods and the use of local discretisation operators; as already stated in «1.2», this is completely misleading;

Remark 6: the fact that the discretisation operators may have a global character does not mean that the representation of the results on the model grid is less local than if the opposite choice had been made.

	GALERKIN	**DISCRETE**
GLOBAL	Spectral	XXX
LOCAL	Finite elements	Finite differences

Table 1: Classification of the discretisation numerical methods

Since finite elements methods are a kind of compromise between the two other possibilities (as seen in Table 1, the other intermediate alternative does not make sense) and since our aim is here mainly to describe spectral methods, we shall compare the latter exclusively with finite differences methods, in order to obtain a clearer distinction and a better evaluation of their advantages and disadvantages. For more details about the third and very flexible approach offered by finite elements when applied in meteorological science, the reader is referred to *Temperton* (1991b).

1.4 Separation of the horizontal and vertical discretisations

Attempts to do anything else than finite differences in the vertical have had little success up to now (finite elements seem anyhow to be the only alternative) ; this is probably due (i) to the need to have a very irregular level spacing and (ii) to the fact that the upper boundary condition (infinite geopotential gradients for a finite pressure jump) cannot be fitted without distortion in any Galerkin method.

Hence we shall deal from now on only with the horizontal discretisation part of the problem. For the complementing part see for example a previous ECMWF Seminar Lecture by *Arakawa* (1983) or *Bubnova* (1998) in the same Proceedings.

The next introductory Section will be treated in one dimension, for the sake of simplicity, since the bi-dimensional applications anyhow have characteristics that strongly depend on the geometry of the problem (spherical or plane-transformed) and will be treated later on. Furthermore, only the case of Fourier series will be treated, the irregular character of other representations (e.g. Chebyshev polynoms) bringing in general more new problems than solving previous ones.

2. GENERALITIES

If one expresses any quantity involved in the solution of a 1-D version of the meteorological equations following:

$$Y(x) = \sum_{j=-M}^{j=M} y_j . e^{i.j.x} \tag{1}$$

with $y_{\pm m}$ being complex conjugate numbers, the space derivation operator obviously becomes:

$$\frac{\partial . Y(x)}{\partial . x} = \sum_{j=-M}^{j=M} i . j . y_j . e^{i.j.x} \tag{2}$$

and the Laplacian operator:

$$\frac{\partial^2 . Y(x)}{\partial . x^2} = \sum_{j=-M}^{j=M} -j^2 . y_j . e^{i.j.x} \tag{3}$$

The value M is called the truncation of the Fourier series that is used to obtain a set of basic functions for the representation of Y, provided the latter is periodic with respect to x on the interval $[0 , 2\pi]$.

Linear combinations and products of quantities Y_1 and Y_2 (or of n^{th} order derivatives of them) can in principle be handled by arithmetic operations on the expansion coefficients y_{1j} and y_{2j}. In the former case the algorithm

is trivial. In the latter case it involves the classical trigonometric equations that link a product of sine/cosine with the sine/cosine associated to the sum and to the difference of the wave numbers.

Since divisions and (when considering physical forcing functions) even more complicated mathematical operators are involved in the solution of simulation problems with the meteorological equations, this very cumbersome «interaction coefficients» method is anyhow never used. One rather uses the transform method that can be described as follows:

+ Choose a regular discretisation of the interval $[0\ ,\ 2\pi]$ with L points per period ($L \geq 2\,M + 1$ so that Fourier transforms can be performed in practice);

+ Make the Fourier transforms that starts with the spectral representation of Y_1 and Y_2 and finish with the set of values on the model grid $Y_{1/2}\ (x_{i,\ i=1,L})$, preferably with a FFT (Fast Fourier Transforms) code;

+ Perform the mathematics of the operator that involves Y_1 and Y_2 independently in each grid-point and put the result in $Y_3\ (x_{i,\ i=1,L})$;

+ Make the reverse Fourier transform that brings back Y_3 to spectral space, again preferably with a FFT code.

Obviously the transform method is not used for each individual operation, but only once per time-step of the model.

Inverse Laplacian operators (see Sections 4 and 8) and advection operators $-u.\partial\psi/\partial x$ are the two main ingredients of the horizontal part of the meteorological equations. Since the advection operators correspond to pure products in spectral space and have to be treated as exactly as possible, the number L must be chosen such as the transform method gives then the same result as an analytical computation. One can show that this suppression of any aliasing on quadratic terms is obtained if one has $L \geq 3\,M + 1$.

One can empirically show that it is unnecessary to further increase L for dealing with the infinite order aliasing associated with divisions and other complex operators, probably owing to the dominant role of advection and geostrophic adjustment in dynamical meteorology calculations.

With the more stringent condition on L expressed just above, the use of a forward-backward set of Fourier transforms on a given quantity Y (without any other operation) will not bring back the original values on the grid when applied for the first time (as it would have been the case with the less stringent first condition linking M and L). This «spectral fitting» is sometimes considered as a handicap in terms of accuracy for the spectral method. Since it is applied to the final result of complex computations that would otherwise suffer from aliasing noise, it may however rather be considered as a distinct advantage of the method.

The latter point is well illustrated on Figure 1 (obtained from a rerun under conditions as close as possible to those of the first study). An early version of ALADIN was run (with a mesh size of ***8.8 km***) on the flash-flood case of Vaison-la-Romaine (22/9/92), with the aim to study the sensitivity of the forecast rainfall pattern to the intensity of the horizontal diffusion (*Gabersek* (personal communication)). While there are deficiencies in the simulation of the event (the rainfall maximum is situated on the western side of the Rhone valley rather than on the first slopes of the Alps, the timing is not fully satisfying, ...), the «warning» aspect about a localised risk of very heavy precipitations must be put to the credit of the model. This quite local simulated structure is a first answer to the recurrent claim that spectral methods cannot represent sharp atmospheric patterns because of their global algorithm. But the most interesting part of the story corresponds to the sensitivity study: opposite to intuitive expectations, an increase of intensity in the horizontal diffusion (by factors of *2*, *4* and *8*) leads on average to an even higher maximum of rainfall and hence to a sharper precipitation pattern (for the intense spots only, otherwise the expected smoothing effect acts). This paradox, as confirmed in the study of *Marku* (1998), is intrinsically linked to the positive action of the «spectral fitting»: the latter, partly enhanced by the application of the horizontal diffusion operator, destroys the most unstructured («noisy») part of the simulated flow, thus allowing a better consistency (and hence efficiency) of the slightly larger scale motions. When horizontal diffusion is enhanced (in reasonable proportions, of course) this effect becomes even more apparent.

The only place where there is a real disadvantage of spectral fitting is when representing orography, a field that must be equivalent in grid-point and in spectral space. Hence the number of associated degrees of freedom is reduced (by a factor 1.5 in the 1-D case and $9/\pi$ in the 2-D case) and Gibbs waves are thus affecting the final field, this being particularly detrimental over the oceans where the orography is obviously known to be exactly flat. A similar problem may arise when treating bounded fields like specific humidity and (even more) liquid or ice water concentration. In the case of semi-Lagrangian modelling some solutions, avoiding the spectral fitting for purely conservative parameters, are however already available, even if their implementation may create new problems (see Section 17).

If one applies the spectral fitting operator a second time to any already fitted field, the result now remains untouched. The spectral fitting operator is therefore called a «projector». In other words, doing a spectral transform is equivalent to search for the set of complex spectral coefficients that will minimise in the least square sense the distance between the original representation and the «fitted» one. This property is central to several intrinsic advantages of the spectral method. It is also linked to the orthogonality of the set of Fourier functions and to the fact that the spectral coefficients obtained by Fourier transform are the scaled integrals on $[0\ ,\ 2\pi]$ of the product of Y by the relevant basic functions.

The «projector» property has also a nice application to minimise the «Gibbs» detrimental effect explained just before. Since the spectral fit of the orography is performed once and for ever for any model configuration, one

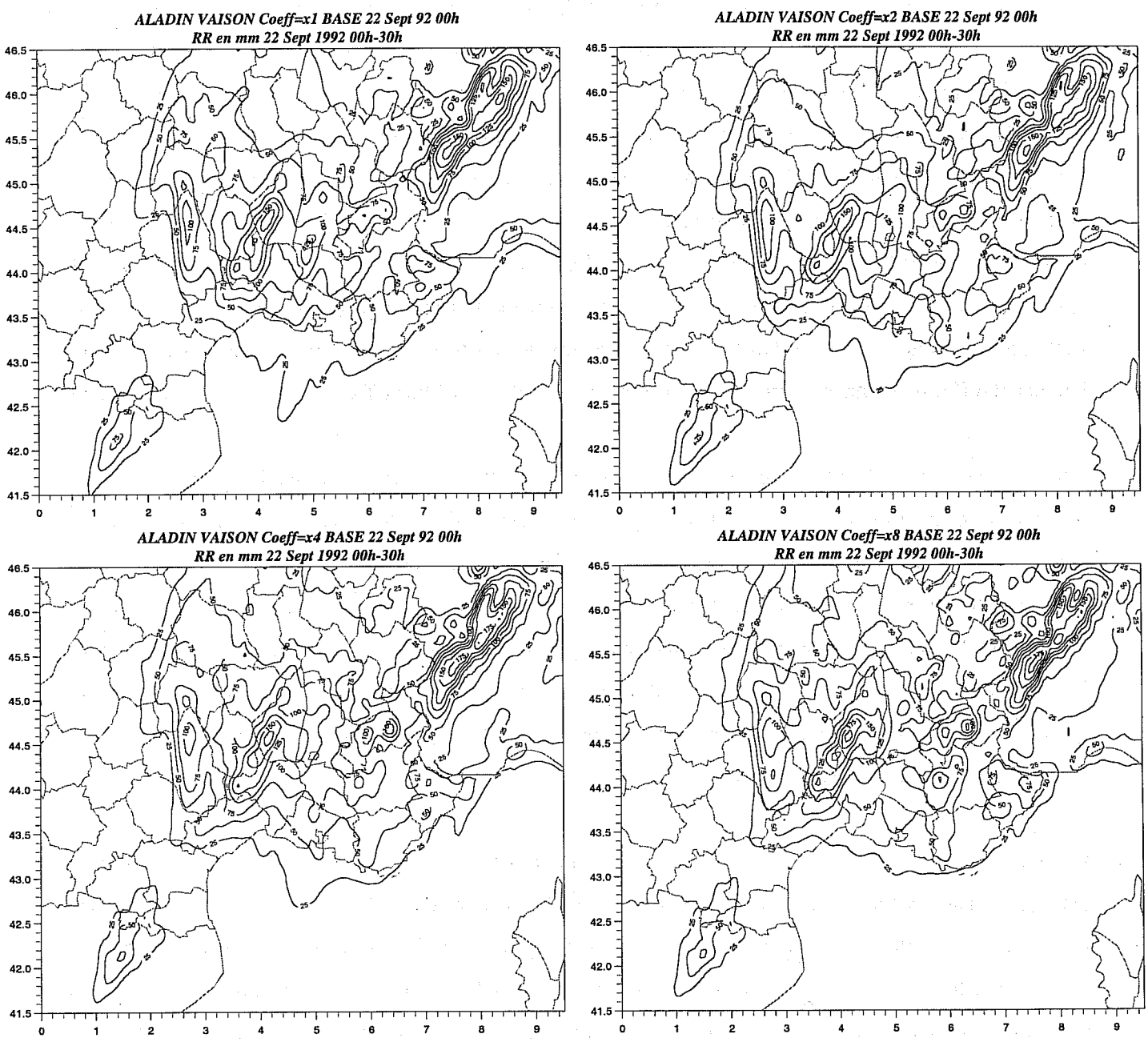

Figure 1: 30 hour accumulated precipitations for the «Vaison-la-Romaine» case of 22/9/92 00 UTC. Forecasts with standard horizontal diffusion coefficients (top left) and increased values by factors of 2 (top right), 4 (bottom left) and 8 (bottom right)

may replace the «equal weight» minimisation problem that corresponds to the orthogonal projection by a complex «variable weight» minimisation problem that can be solved iteratively via an expensive adjoint method. The new minimisation weights can then be chosen to alleviate the Gibbs problem (*Bouteloup*, 1995). Provided one has indeed at disposal a minimisation procedure for variational applications, this method, very similar in its aims to the one proposed by *Navarra et al.* (1994), offers a total flexibility of use and a stronger theoretical background.

Equation (3) shows that the set of basic functions of the Fourier spectral representation are the eigen-vectors of the 1-D Laplacian operator. Hence any Helmholtz operator will become diagonal in spectral space. Given the crucial role of this operator in eliminating or slowing down fast propagating waves (gravity and/or acoustic ones), the time-stepping algorithm of a spectral model must be organised so that the Helmholtz equation is solved in spectral space.

The eigen-values of the same Laplacian operator behave (see Equation (3)) like the squares of the wave numbers. Hence any additional operation that must be scale-selective (like horizontal diffusion) is easily implemented also in spectral space and with the maximum possible flexibility. It is sufficient to multiply the final result of the other computations by a damping coefficient that depends on the wave number, either the analytical counterpart of the exact use of the Laplacian, or in fact any arbitrary monotonous function of the wave number. On the other hand this has the intrinsic disadvantage that the damping rate must then be the same everywhere in real space.

3. COMPARED PROS & CONS WITH FINITE DIFFERENCES (*Gustafsson and Mc Donald*, 1996)

Following the two above-mentioned authors (from which this Section is almost entirely inspired), the question of the ratio between accuracy and computing price will not be considered here. Indeed they claim that it is very similar between both techniques. Other authors, however, give an advantage to spectral methods, for that very crucial item of computing efficiency (*Jarraud and Girard*, 1983). The matter is delicate since it refers to the equivalent finite difference resolution of a spectral truncation, a problem that can be solved in at least half a dozen equally valid manners. Since this is not in the scope of this summary about spectral methods, the matter will not be further dealt with here.

3.1 **Pros**

* in Eulerian mode, the spectral method has no linear phase error;

* there is no source of non-linear instability in a spectral model;

* the problem of evaluating the pros and cons of grids A, B, C, D or E (in the Arakawa classification) disappears in a spectral model; the most simple grid (the A one) is used without any negative impact; this

creates additional benefits, when using the semi-Lagrangian technique, by avoiding different trajectories (and hence interpolations) for different prognostic quantities;

* on the sphere, and even more with a triangular truncation (see Section 4), there is no pole problem with the spectral method;

* the spectral method has no aliasing error for the computation of quadratic terms;

* Helmholtz equations are trivially solved in a spectral model and this «diagonalisation» property is also useful for variational data assimilation;

* implicit (hence absolutely stable) horizontal diffusion is also trivial and infinitely tunable in a spectral model;

* spectral fitting is anyhow a powerful way to remove any numerical noise;

* the reduction of the number of degrees of freedom in spectral space is a quite favourable property for such various applications as post-processing or variational computations.

3.2 Cons

* sharp features may be less well represented (Gibbs effect) in a spectral model;

* spectral fitting of bounded quantities can lead to unrealistic results through over- or undershoot through the Gibbs phenomenon (negative specific humidities for instance);

* the orography is less well represented in spectral models for the same accuracy of the «free-atmospheric» part;

* it is impossible in practice to apply a flow dependent operator for horizontal diffusion in a spectral model (too expensive transforms if one wants to get away from the linear selective mechanism of a diagonal operator in spectral space);

* the method cannot be efficiently applied in the global case for very high truncations (see Section 4).

4. GLOBAL UNIFORM APPLICATION

The earth (assimilated to a sphere) being periodic in all directions, it is easy to see that the spectral method is ideally suited for global applications. However the need to efficiently perform 2-D spectral transforms imposes a factorisation between two orthogonal directions: the longitude λ and the sine of the latitude μ. The Fourier functions are, like in the 1-D case, the natural expansion basis in λ. The bounded character of the geometry in $\mu \in [-1, 1]$ imposes another choice for the second component of the basic set of functions.

The property that the combined basic functions must be eigen-vectors of the spherical horizontal Laplacian operator (spherical harmonics) imposes one single choice. The «μ» functions are the associated Legendre polynoms. The transform grid is regular in λ and follows the set of «North-South Gaussian latitudes» in μ. The transforms in the North-South direction are performed by Gaussian quadrature, i.e. by matricial

multiplication of the «fitted field» by the values of the associated Legendre polynoms evaluated on the Gaussian grid.

The 2-D spherical equivalent of Equation (1) thus becomes:

$$Y(\lambda,\mu)=\sum_{n=0}^{n=N}\sum_{m=-n}^{m=n}Y_n^m.P_n^m(\mu).e^{i.m.\lambda} \qquad (4)$$

if one decides to use the so-called «triangular truncation» for limiting the expansion series in both ***n*** (total wave number) and ***m*** (zonal wave number).

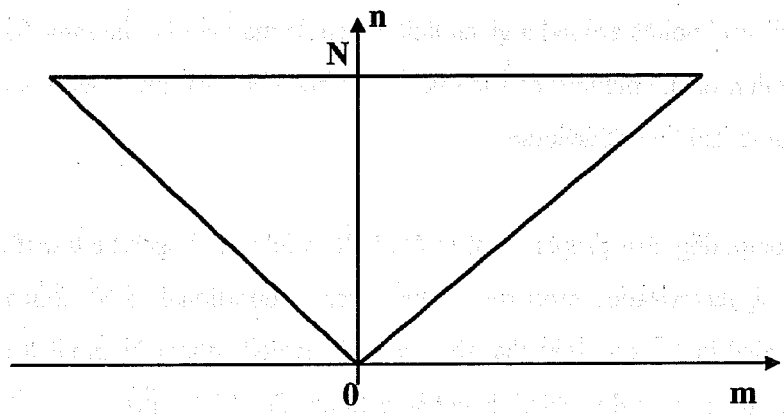

Figure 2: The triangular truncation's arrangement of allowed bi-dimensional wave numbers

The eigenvalue of the spherical Laplacian for one «spherical harmonic» associated to the complex conjugate coefficient Y_n^m is $-n.(n+1)/a^2$, a being the radius of the earth. One therefore understands the advantage of the triangular truncation: the highest total wave number basic functions all have the same bi-dimensional scale, whatever value the zonal wave number ***m*** takes. In other words, the representation with a triangular truncation T_N is homogeneous and isotropic.

The special character of the triangular representation is made even more spectacular through its invariance by rotation: a field exactly represented by Equation (4) and having been subject to a pole rotation on the sphere can be exactly represented again by a similar expression in the new (λ',μ') coordinates, the only change being that of the expansion coefficients. All this explains why nearly all spherical spectral applications are nowadays using a triangular truncation and not a rhomboidal or trapezoidal one (earlier proposals).

The FFT algorithms require the number of grid points around any latitude circle to be of the arithmetic form $2^i.3^j.5^k$, to ensure good factorisation properties. Here is for example a list of those numbers up to 800, nicely showing how irregular the distribution of such integers might be (there are sometimes frustrating jumps in the search of a fitting truncation for a given problem!):

2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 18, 20, 24, 25, 27, 30, 32, 36, 40, 45, 48, 50, 54, 60, 64, 72, 75, 80, 81, 90, 96, 100, 108, 120, 125, 128, 135, 144, 150, 160, 162, 180, 192, 200, 216, 225, 240, 243, 250, 256, 270, 288, 300, 320, 324, 360, 375, 384, 400, 405, 432, 450, 480, 486, 500, 512, 540, 576, 600, 625, 640, 648, 675, 720, 729, 750, 768, 800.

Often, for symmetry reasons, it is decided to only use even FFT numbers ($i > 0$). In such a case the above list reduces to:

2, 4, 6, 8, 10, 12, 16, 18, 20, 24, 30, 32, 36, 40, 48, 50, 54, 60, 64, 72, 80, 90, 96, 100, 108, 120, 128, 144, 150, 160, 162, 180, 192, 200, 216, 240, 250, 256, 270, 288, 300, 320, 324, 360, 384, 400, 432, 450, 480, 486, 500, 512, 540, 576, 600, 640, 648, 720, 750, 768, 800.

In fact the number of grid points around a given latitude circle may not be the same for all latitudes. Indeed in IFS there is a reduction of the number of such points as one goes towards each pole (*Hortal and Simmons*, 1991) without any practical loss of accuracy.

The FFTs have a computing cost proportional to $N^2.ln(N)$ while the Legendre transforms (made without any «fast» algorithm using factorisation properties) have a cost proportional to N^3. Since the rest of the model's computations are scaled in N^2, one sees that the spectral method cannot be applied efficiently for very high resolutions. However the practical problem is not so serious. The FFTs' price growth is of little concern, that of Legendre transforms looks like becoming serious for N of the order of 1000 and mathematicians are anyhow working hard to produce FLT (Fast Legendre Transforms) algorithms (see for instance *Mohlenkamp* (1997)).

5. GLOBAL VARIABLE RESOLUTION APPLICATION

There exists one and only one non-trivial conformal sphere to sphere projection (*Courtier and Geleyn*, 1988), the one proposed by *Schmidt* (1977), with only three degrees of freedom: the latitude and longitude on the real sphere of the so-called pole of dilatation (i.e. the centre of interest in the case of a meteorological application like the ARPEGE one at Météo-France (*Courtier et al.*, 1991)) and the stretching coefficient «c» (map factor value at the pole of dilatation and inverse of the map factor value at the antipode). The associated transform possesses a number of very interesting mathematical properties, detailed in the two above-mentioned papers, but for the purpose of this lecture note it is surely better to start with a quasi-graphical explanation (that surprisingly enough was absent from the original paper of Schmidt). In the following, primed values usually correspond to the so-called «transformed sphere» and non-primed ones to the real sphere.

The geometrical transform can best be expressed, after rotation of the coordinate's «pseudo north pole» to the real pole of dilatation by:

$$tan(cos^{-1}(\mu')/2) = c \, . \, tan(cos^{-1}(\mu)/2) \qquad (5)$$

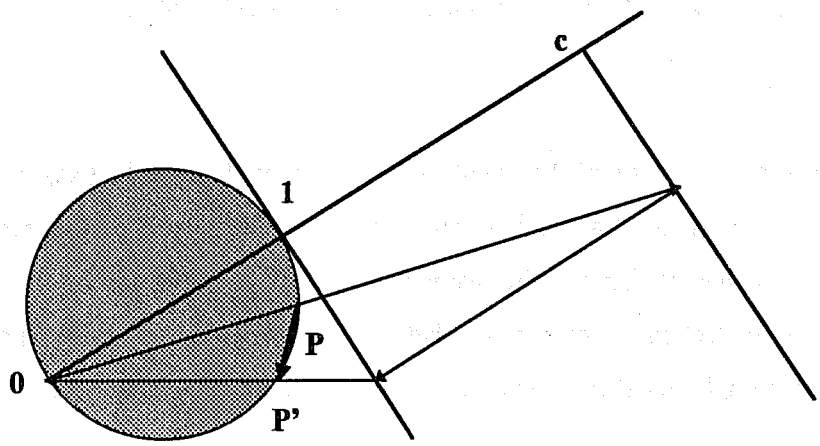

Figure 3: Geometrical interpretation of the Schmidt transform or how to go from *P* on the real sphere to *P'* on the transformed sphere (*Courtier and Geleyn*, 1988)

While it would be easy to also apply this transform for a «zoom» effect in finite differences (or finite elements) modelling, it has two special properties that make it particularly suitable for spectral applications:

+ the conformal property of the geometrical transform implicit in Equation (5) (polar stereographic projection, homothecy of factor «*c*» in the tangent plane and inverse stereographic projection all being conformal operations) ensures that the homogeneity and isotropy properties of a triangular spectral representation will be locally preserved;

+ the map factor is a spherical harmonic of degree «one» (no λ' variation, linear variation in μ' from $1/c$ at -1 to c at $+1$) on the so-called «transformed sphere».

One can therefore either write the horizontal part of the meteorological equations (with transformed operators) on the transformed sphere defined by (λ',μ') or still on the (λ,μ) sphere with the classical formalism of a map factor introduced in the derivatives and in the wind components. Since the advective terms then involve the square of the map factor (of order two), the only additional need for the anti-aliasing choice of the truncation is to replace the value one by the value three in the formula linking L to M (if one returns for a second to the 1-D framework and to its notations), i.e. $s.M + 1 + \underline{2} = s.M + 3$ replaces $s.M + 1$, whatever the choice for s might be (3, 2, 2.5, ...), the underlined term corresponding to the square of the map factor.

Another problem linked to the variation of the map factor is the choice of the linearisation procedure for the semi-implicit scheme. One may (like for instance in ALADIN) choose to have a constant map factor (equal to its maximum value and hence here to «*c*») in the linear model, but this creates accuracy problems at the antipode even for reasonable values of the stretching parameter, especially when using the relatively long time-steps allowed by the semi-Lagrangian algorithms. Alternatively, the map factor may enter the Laplacian-type operator to be inverted but the latter then ceases to be purely diagonal (*Yessad and Bénard*, 1996). Fortunately, thanks again to the order one representation of m, the North/South part of the matricial operator

becomes only pentadiagonal, while the East/West part obviously remains diagonal. Hence, the computational overhead is of the order of a few percents only.

Provided one is carefully choosing which parts of the calculations have to be done in the (λ',μ') framework and which ones in the (λ,μ) one, a global spectral variable resolution model can be created in a rather simple way, while preserving all important characteristics of the unstretched version, including for instance the use of a reduced grid near the poles (see Section 5), even if the latter are not anymore the geographical ones. *Hardiker* (1997) describes another implementation of basically the same idea at FSU Talahassee.

6. DEGREES OF FREEDOM OF THE SCHMIDT TRANSFORM

As already mentioned there are only three of them:

- Stretching factor «c» (map factor at the pole of dilatation);
- Sine of the latitude of the pole of dilatation;
- Longitude of the pole of dilatation.

With the additional choice that the reference meridian in the projection is the one pointing to the (geographical) south, the geometry of the transformed earth is perfectly defined with the three above-mentioned parameters.

Technical details about the grid-point (*Geleyn*, 1988; *Clochard*, 1990) and spectral (*Rochas et al.*, 1991) transformations from the real sphere to the stretched one (or the reverse) will not be detailed here, since they would not help getting any additional understanding about the Schmidt method, beyond the results mentioned in Section 5.

Concerning the current operational choices at Météo-France, Figure 4 (a, b and c) shows (with only every fourth point plotted in both directions) respectively the area of maximum interest of ARPEGE, the opposite hemispheric view and the former with the ALADIN-France grid-points added. On Figure 4a the drawn «equator of the transformed sphere» indicates the area of local zoom of the Schmidt transform (*Courtier and Geleyn*, 1988). On Figure 4c the ALADIN-France grid (***9.9 km***) is roughly equal to half that of ARPEGE in the same area following the strategy advocated in *Caian and Geleyn* (1997).

7. PROBLEMS LINKED WITH THE USE OF THE STRETCHED GEOMETRY

Even if they were numerous at the beginning of the ARPEGE project, problems associated with the irregular character of the grid when returning to the real sphere (for instance when computing trajectories in semi-Lagrangian applications, or when changing geometry to obtain new initial conditions or to perform post-processing) have all been solved. Remaining problems are all associated with the spectral representation on the transformed earth.

ARPEGE T199 C3.5 operational grid

1 grid-point over 16 plotted

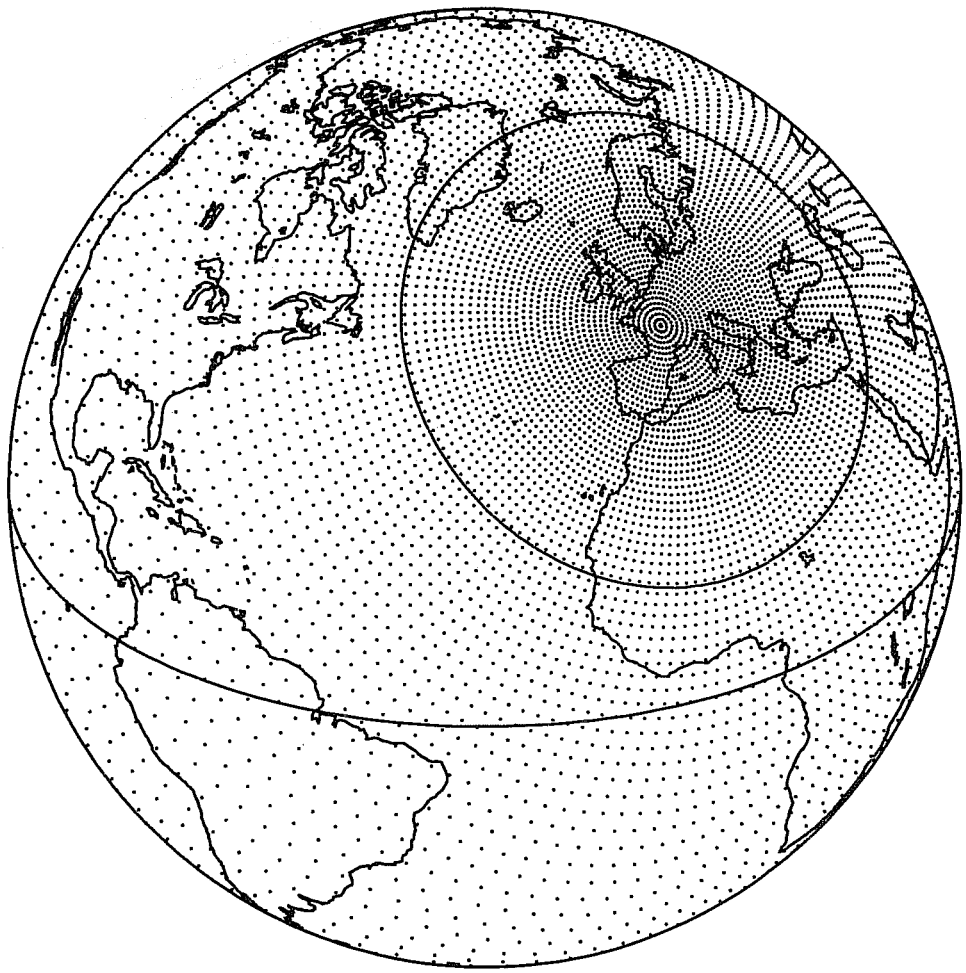

Figure 4 a

ARPEGE T199 C3.5 operational grid

1 grid-point over 16 plotted

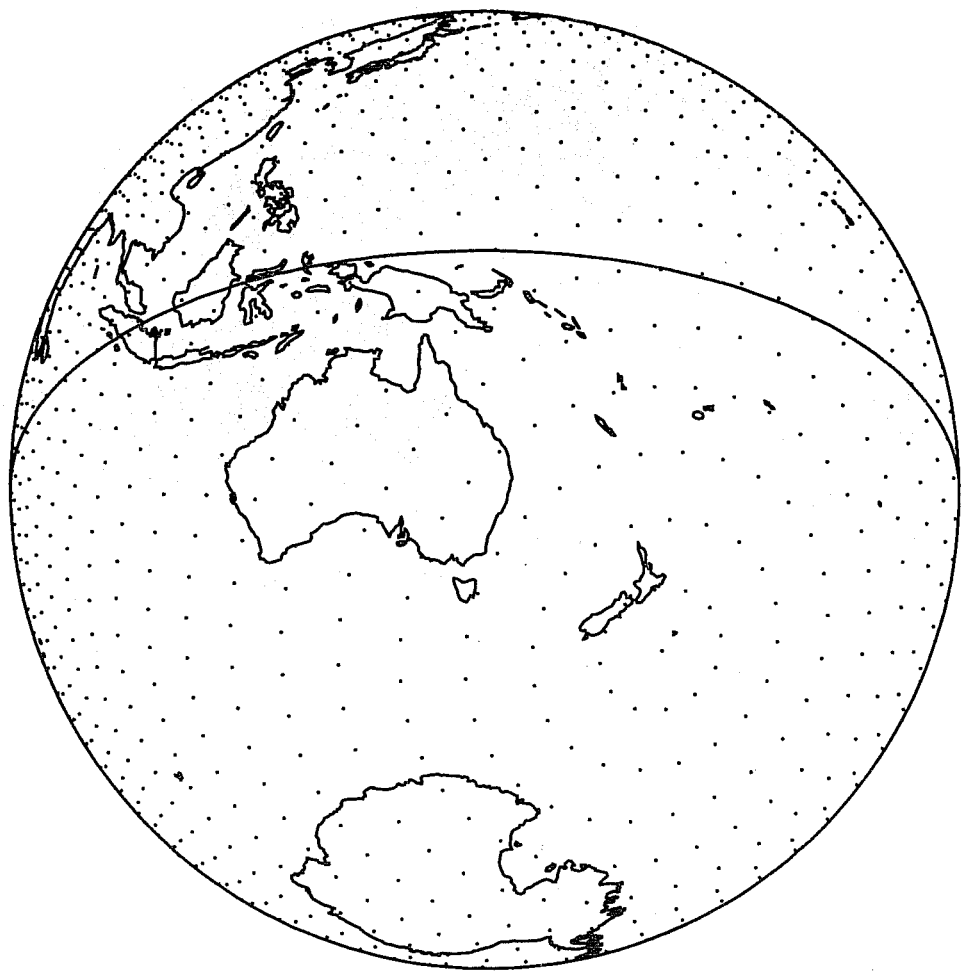

Figure 4 b

ARPEGE (T199 C3.5) and ALADIN-FRANCE (9.9 km) operational grids

1 grid-point over 16 plotted

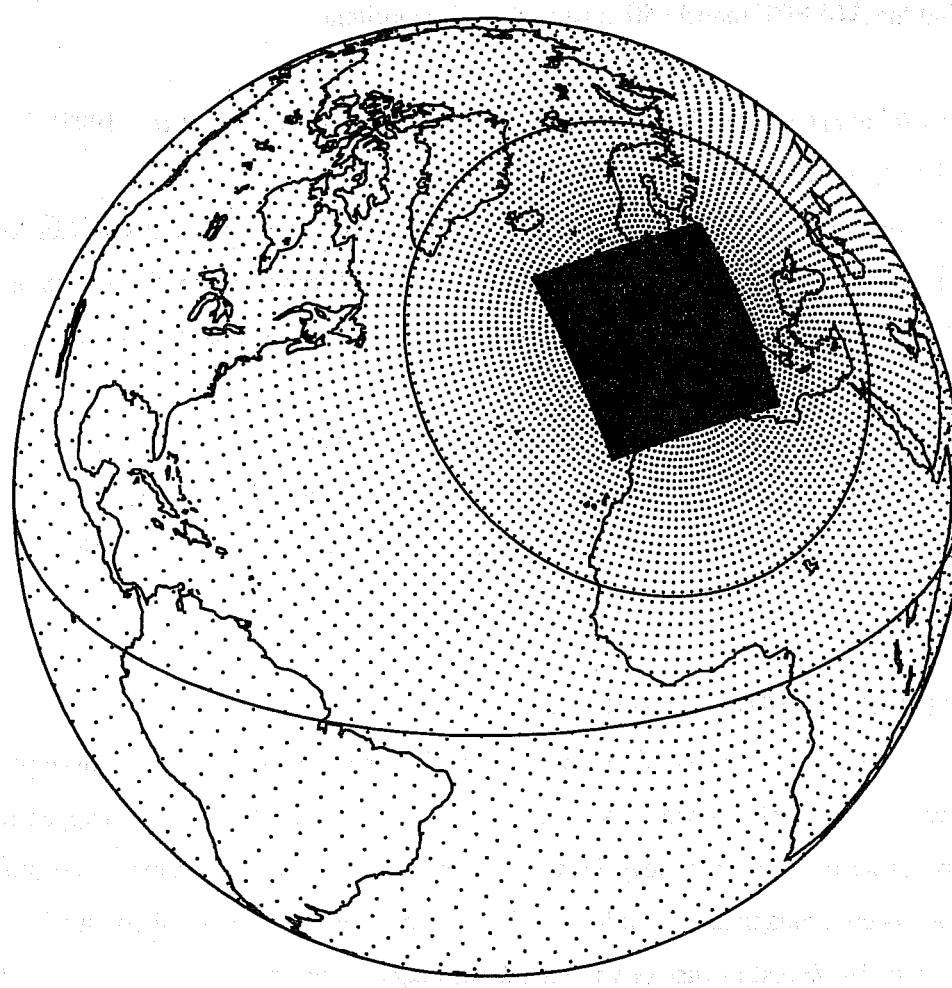

Figure 4 c

A first category is associated with the horizontal diffusion, that has to be performed in spectral space and that ideally should have an intensity depending on the local resolution. This subject will be treated in detail in Sections 14 and 15.

The second and last category concerns all instances when one has to go back to the real earth while remaining in spectral space. One may want to do so either to get rid of the stretching (for problems where homogeneity of the representation is crucial) or of the tilt (when one wants separability of the representation of the Coriolis parameter) or of both together. Currently the operational ARPEGE suite does not need to use any more such a procedure. However it will probably be used again when the stretching will be introduced in the incremental variational data assimilation (see Section 18). We may then expect problems similar to those we encountered when initialising the ARPEGE model with a normal mode algorithm.

Those problems (eliminated through the choice to use digital filters (*Lynch et al.*, 1997) for initialising the ARPEGE fields) are of two types:

* from a practical point of view, the memory requirements are very penalising (for keeping an exact equivalent to the analytical results one has to multiply the original truncation by the stretching parameter and by a «security factor» of about *1.2*) and huge matrices have to be precomputed and stored to keep the computational burden of the transform acceptable;

* from a scientific point of view, the aliasing introduced through the multiplication by the map factor is not completely controlled by the «*1 => 3*» extra security in the choice of the number of grid points and some accumulation of errors at the very end of the spectrum is noticeable when iterating the normal mode calculations.

8. LIMITED AREA MODELLING (LAM) APPLICATIONS

LAM applications are not as obviously fitted to the spectral framework as global ones, since periodicity is now missing. However, if this problem can be solved, and we will now see that this can happen in at least three ways, the plane projection geometry may be more advantageous than the spherical one (no pole mapping problem for the spectral functions, i.e. full periodicity in both directions). Furthermore it has been shown (*Caian and Geleyn*, 1997) that the use of very high stretching coefficients in the Schmidt transform cannot be a suitable substitute to the application of LAMs for rather local problems.

8.1 How to eliminate the non-periodicity problem, variant A

Here as well as later, we shall assume that the problem of «coupling», i.e. the introduction at the lateral boundaries of information coming from a larger scale model and the elimination of the outgoing information produced by the LAM, is solved like in a finite difference model, e.g. with the method of *Davies* (1976).

In this variant, tried first at ECMWF (*Hoyer*, 1987) and later developed at NCEP Washington (*Juang and Kanamitsu*, 1994), the LAM forecast fields are the departures (over the whole LAM domain) from the results of the integration of the coupling model, assumed to obey to exactly the same equations as the LAM:

+ *advantages*: the departure being exactly zero on the edges and its derivatives being close to zero also (thanks to the coupling procedure) the application of the spectral technique to a truly periodic pattern may be similar to that on the sphere;

+ *disadvantages*: the part on which the evolution equations at the LAM scale are not acting differs from the part of the flow that is determined by the boundary values and by the inversion of the basic Helmholtz operator (the so-called «harmonic solution», *Chen and Kuo* (1992)); it is sometimes argued that the problem is negligible since the latter can only be of large scale (like the «forced» background); examination of the spectra of a relevant field, especially in presence of finer scale orography, immediately proves the opposite. This ceases to be true if one artificially forces the fields on the edge to be very smooth but then the problem of regenerating useful small scale information in the vicinity of those edges replaces the previous one and becomes the weak point of the method in «replacement» of the above-mentioned consequence of the split between the full- and the truly spectral computations (*Machenhauer and Haugen*, 1987).

8.2 How to eliminate the non-periodicity problem, variant B

In this variant, developed at JMA Tokyo (*Tatsumi*, 1986), there is one more component in the basic set of «spectral» functions, namely the harmonic solution:

+ *advantages*: this formally leads to a truly independent LAM, even if in practice the distinction from variant A is small (the latter can probably be run with any coupling model without much harm being done); the application of the spectral technique is straightforward;

+ *disadvantages*: the set of basic functions is not any more orthogonal (even if nearly) and hence some of the nice properties of the spectral technique are lost; the bi-periodicised fields are continuous on the edge (zero values), but, unlike in the previous case, their derivatives are discontinuous and this reduces the similarity with the global case, with some additional negative effects (*Kuo and Williams*, 1992).

8.3 How to eliminate the non-periodicity problem, variant C

In this variant, proposed by *Machenhauer and Haugen* (1987) and developed in the HIRLAM (*Haugen and Machenhauer*, 1993) and ALADIN (*Bubnova et al.*, 1995) groups in Europe, the grid-point fields are made bi-periodic by the creation of an artificial extension zone in which they are interpolated between the values at the opposite edges; a pure bi-Fourier spectral representation can then be applied to these «extended» fields:

+ *advantages*: this leads again to a truly independent LAM; the application of the spectral technique is, even more than in variant A, the mirror image of that on the sphere, i.e. the most natural one, with purely orthogonal fully periodic functions;

+ *disadvantages*: the operator «extension, spectral fitting and reduction to the useful domain», the functional equivalent of the spectral fitting in the global case, is not a projector (one can however find a technical solution to only compute the extension values for the interpolated coupling fields, thus reducing the associated problem to a minimum); there is a reasonably small additional computing cost since the number of spectral degrees of freedom increases with respect to the two other variants (with the careful treatment of the semi-implicit and semi-Lagrangian algorithms proposed by *Radnoti* (1995), there is indeed no need to do the grid point computations in the extension zone).

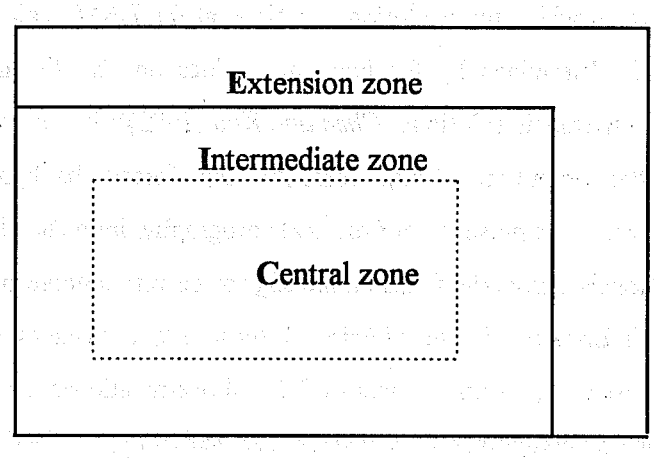

Figure 5: Schematic representation of the Machenhauer-Haugen geometry

The third variant, from the point of view of advantages and disadvantages, seems to be the most «consensual» choice from the three solutions, especially if compatibility with spherical applications is searched. Hence we shall use it alone from now on to study the LAM specificities of spectral computations, notwithstanding the fact that both other variants can lead to the construction of robust spectral LAMs.

8.4 Spectral representation

The bi-Fourier expansion obviously creates eigen-vectors of the LAM horizontal Laplacian and FFTs can be economically applied for both «W-E» and «S-N» transforms. The equivalent of Equation (1), in analogy with Equation (4), becomes:

$$Y(x_1,x_2)=\sum_{n=-N}^{n=N}\sum_{m=-M'(n)}^{m=M'(n)}Y_m^n.e^{i.(m.x_1+n.x_2)} \tag{6a}$$

or:

$$Y(x_1,x_2)=\sum_{m=-M}^{m=M}\sum_{n=-N'(m)}^{n=N'(m)}Y_m^n.e^{i.(m.x_1+n.x_2)} \tag{6b}$$

where the truncation wave numbers in the two directions M and N are related to the number of grid points of the total domain (extension zone included) by the same relationship as in the 1-D case, these numbers being of course FFT favourable ones.

8.5 Truncation

Apart from the trivial use of the physical domain's size in the Laplacian operator's expression, the only remaining problem is that of the choice of $M'(n)$ (or $N'(m)$) of Equations (6). The solution must have the advantage of an homogeneous isotropic truncation. Even if it is not exactly possible (there is no total wave number but a Pythagorean combination of two relative ones), a good approximation is the elliptic truncation:

$$\left(\frac{n}{N}\right)^2+\left(\frac{m}{M}\right)^2 \leq 1 \qquad (7)$$

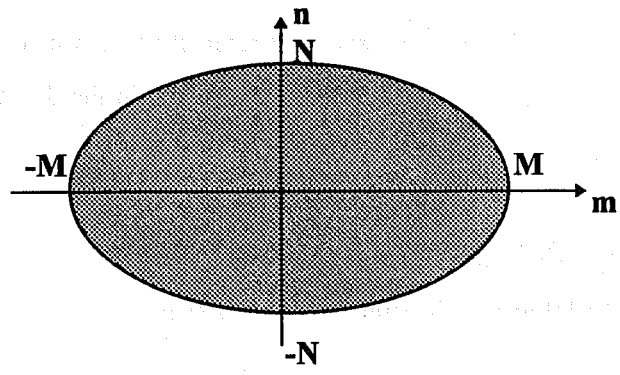

Figure 6: The elliptic truncation's arrangement of allowed bi-dimensional wave numbers

9. CHOICE OF THE PROJECTED GEOMETRY FOR THE LAM DOMAIN

For ALADIN, we may have elected, like in the HIRLAM case, to choose the same equivalence than in finite difference models, i.e. the lat/lon system with a rotated pole (in order to bring the pseudo-equator in the middle of the domain of interest) as counterpart of a portion of the sphere. Following the work of *Joly* (1992), we decided not to do so. One of the main advantages of the spectral models (*Temperton* (1991a), *Rochas and Courtier* (1992)) is to be able to play on the duality between the two wind representations (either (u,v) components or vorticity/divergence) in order to optimise the arrangement of transforms and spectral computations. But this duality, if one wants to have an exact expression for the metric terms on the sphere (sometimes called «curvature terms» in the description of the primitive equations), requires a representation locally invariant for any rotation of the chosen coordinate system. Hence the pseudo-latitude of the rotated lat/lon system should in principle undergo a change of variable that, not surprisingly, leads to the equations of the Mercator system (corresponding to the only projection conformal and invariant by «∝» translation). In other words, while in finite difference models the functional LAM equivalent to a portion of the sphere is obviously of the lat/lon type, in spectral models a conformal projection is needed to obtain the same duality!

This point of course reinforces the similitude between the (conformal) Schmidt «sphere to sphere» transform and the ALADIN application of the Machenhauer-Haugen concept; it also has nice consequences for the

maintenance of maximum code compatibility between ARPEGE and ALADIN: in both cases a single map factor defines the metric and hence the way to go from pseudo-wind to pseudo-vorticity/divergence. The only difference is that in ARPEGE the longitudinal derivative of the map factor vanishes when computing the metric terms, which is only the case in ALADIN if one chooses the Mercator projection. A few unnecessary additions of zero or multiplications by one in IFS/ARPEGE are obviously enough to iron out this slight difference.

It was therefore decided to use only conformal projections for the plane representation of the LAM domain of ALADIN:

* there are three possibilities: polar stereographic, conformal Lambert and Mercator;

* the poles and reference meridian of the earth representation on which these three projections are applied may be chosen arbitrarily (rotated lat/lon concept as first step in the definition of the projection, this requiring the additional choice of two parameters).

If (to simplify and it is usually sufficient) one however elects to remain in classical (λ,φ) coordinates, the geometry of one particular ALADIN domain is defined by 16 parameters:

- Sine k of the «tangential» latitude of the projection:
  1 for polar stereographic, 0 for Mercator, in between for Lambert;
- Latitude φ_0 of the reference point of the projection (where the map factor is 1);
- Reference longitude λ_0 of the projection («N/S» axis for the domain), irrelevant in the Mercator case;
- Latitude φ_1 of the «SW» corner of the C+I domain;
- Longitude λ_1 of the «SW» corner of the C+I domain;
- Latitude φ_2 of the «NE» corner of the C+I domain;
- Longitude λ_2 of the «NE» corner of the C+I domain;

- Number of δx intervals in the «W/E» direction for the C+I+E domain;
- Number of δy intervals in the «S/N» direction for the C+I+E domain;
- Position in x direction of the «SW» corner of the C+I domain;
- Position in y direction of the «SW» corner of the C+I domain;
- Position in x direction of the «NE» corner of the C+I domain;
- Position in y direction of the «NE» corner of the C+I domain;

- x direction related truncation M;
- y direction related truncation N;
- Truncation n_t for the map factor's representation (see Section 11).

Remark 1: It is a good habit to choose φ_0 so that the projection is really tangential ($k = sin\ (\varphi_0)$). The value *1* for the map factor then becomes the minimum value of the latter in the C+I domain. The arbitrary multiplying parameter for the map factor then disappears as «unnecessary» degree of freedom;

Remark 2: Both total C+I+E numbers of intervals must be «good» FFT numbers of the form $2^i.3^j.5^k$;

Remark 3: In practice it is convenient to use the set-up of Figure 5 by putting the C+I-zone in the bottom left corner of the C+I+E-zone. This removes in practice two «apparent» degrees of freedom;

Remark 4: Empirically the «ideal» width of the extension zone has been estimated to *12* grid intervals. This value must anyhow be considered as a minimum;

Remark 5: It is recommended to choose the parameters so that $\delta x = \delta y$, in accordance with the conformal character of the projection. If it is done, one «hidden» degree of freedom also disappears;

Remark 6: The «*s*» ratio between grid-point and spectral dimensions (a number between *2* and *3*) must logically be the same in the two directions. Hence the choices for M and N are not independent and another «unused» degree of freedom disappears here;

Remark 7: It is recommended that n_t (used as the same value in both geographical directions, see Section 11) be the integer equal or immediately above the maximum domain size divided by *12* in the Eulerian case and the minimum truncation wave number minus *10* in the semi-Lagrangian case (*Le Moigne* and *Jerczynski* (personal communications));

Remark 8: Depending on the chosen constraints there are thus between 8 and 11 «real» degrees of freedom in the choice of the 16 above-mentioned parameters (7 and 10 in the special Mercator case where the reference longitude becomes arbitrary or 6 and 9 if one chooses, whatever projection is used, to minimise the variations of the map factor by having the reference point at the exact centre of the domain). Another way of figuring out the basic six degrees of freedom is to think in terms of coordinates of the reference projection point (two), mesh size (one), distance between the projection point and the domain's edges in number of grid sizes (two) and ratio between grid-point and spectral dimensions (one). The three optional degrees of freedom are the two widths of the extension zone and the cut-off wave number n_t.

10. BIPERIODICISATION

Running the Machenhauer-Haugen solution requires the existence of one linear «biperiodicisation» operator that allows to fill (before the spectral fit, if the latter is necessary) values in the E-zone, knowing only values in the C+I-zone.

The wished properties of this operator must be simplicity, regularity and isotropy. There is no obvious a-priori mathematical solution to obtain at the same time all three properties. Heuristic choices are here unavoidable.

One way to evaluate the quality of such an operator is to measure the degree of «non-projectibility» of the operator «biperiodicisation, spectral fitting and reduction to the useful domain». With respect to this score the relevant ALADIN operator (invented by *Batka* (personal communication)) is excellent since even the solution of a variational problem using all C+I values as entry cannot beat it within a reasonable number of iterations!

This ALADIN operator is the combination of two steps:

* using the values at the edges and in the next two rows on each side a spline function is constructed to produce preliminary values in the extension zone; this operation is done first in the N/S direction and then in the E/W one (the inverse procedure would give the same result owing to the linearity of the spline operator);

* a 9-point «Laplacian» smoothing operator is applied iteratively on the resulting field, the number of iterations increasing like the distance to the edge in number of grid-sizes.

As an example of the properties of the method, Figure 7 shows the biperiodicised orography of the current ALADIN-LACE domain. One notices the spectacular «bridge» between the Scottish Highlands and the Atlas Mountains as well as the nice isotropy of the prolongation of the Anatolian Plateau.

Since the values produced by biperiodicisation are essentially useful for coupling purposes, it is wishable to find a way to avoid doing such rather expensive computations too often (i.e. at each time step). If one notices that the above-described operator is linear, as well as the way to interpolate in time boundary conditions provided by the coupling model, the solution becomes obvious. One commutes the two operations by applying biperiodicisation only on the coupling files and interpolating in time all values, including those in the E-zone. Hence, the model integration does not know this operator (mainly used when preparing lateral boundary conditions), except in the set-up phase for preparing the truncation of the map factor (see Section 11).

The biperiodicisation operator has also to be used when preparing constant «geographical» fields, at least for orography. If one thinks of further research on «spectral coupling» (see Section 12), it is also useful to do it for all surface fields, which is indeed the case (this time without spectral fit, of course). For similar reasons, the Coriolis parameter and the geometrical parameters governing the calculation of the solar zenith angle are also treated, together with the map factor, in the set-up phase of the model integration. Hence all dynamical and physical computations may be performed in the extension zone, even if their result is currently irrelevant.

11. REPRESENTATION OF THE BIPERIODIC MAP FACTOR

When going from the sphere to the plane projection (or from sphere to pseudo-sphere in the case of the stretched ARPEGE geometry), we introduce as map factor m the locally defined ratio «mesh size in the projection» divided by «mesh size on the real sphere». The use of m and of the associated reduced winds ($u'=u/m$, $v'=v/m$) helps writing the adiabatic equations in a regular metric on the projection. Please beware

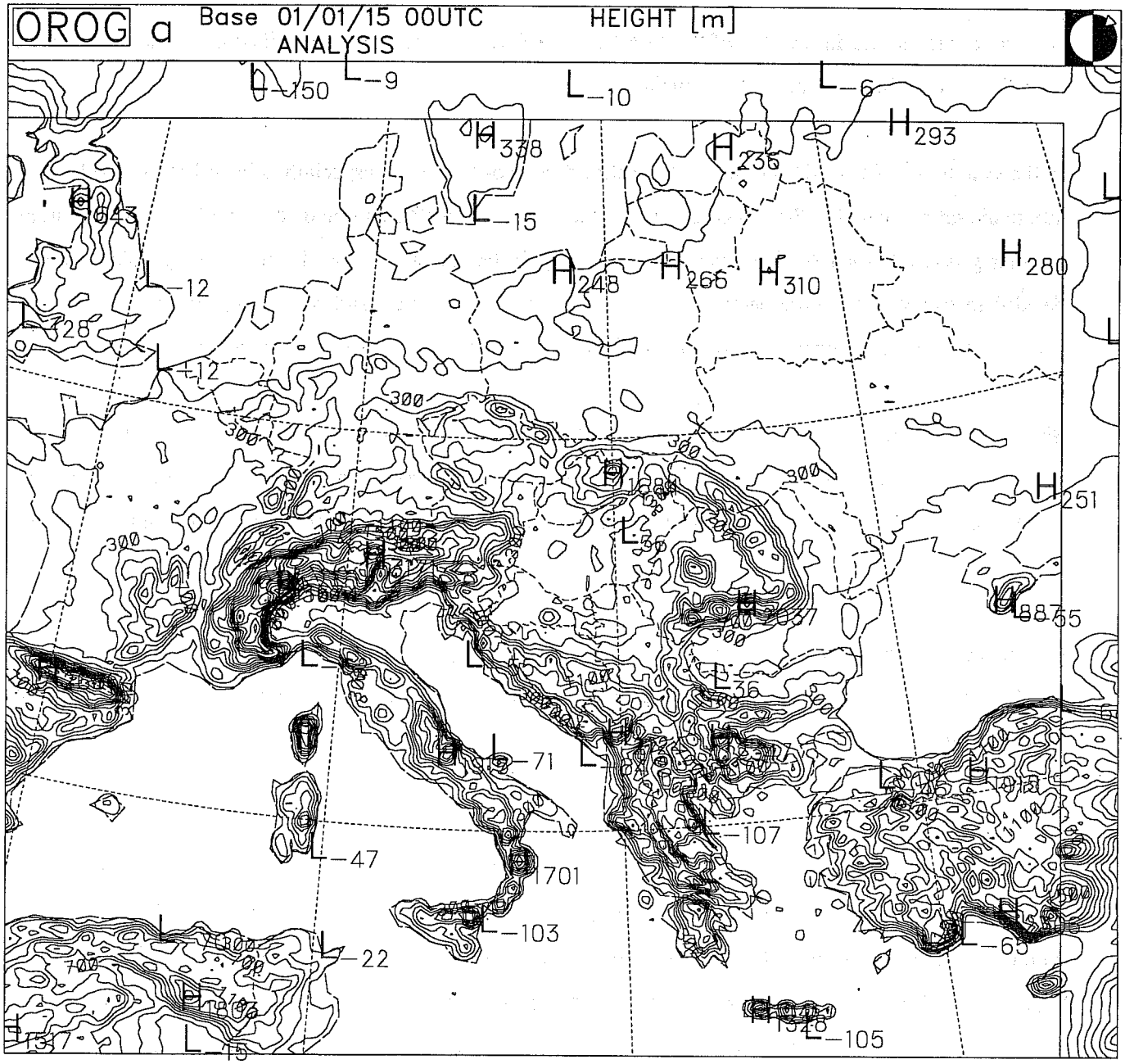

Figure 7: The biperiodicised orography of ALADIN-LACE
The extension-zone is on the top and the right of the picture

that the notation m for the map factor has nothing to do with the previous use of the same symbol (zonal/x wave number).

However, in the case of a spectral model, this creates an additional source of aliasing, since the advection equation for the wind itself takes the fourth-order form:

$$\frac{\partial.\vec{u}'}{\partial.t} = -m^2.\vec{u}'.\frac{\partial.\vec{u}'}{\partial.x'/y'} \tag{8}$$

As already mentioned, in the case of the stretched ARPEGE geometry, m is described by a basic function of order one so that the consequences are minimal.

In the case of any ALADIN projection, a spectral decomposition of m, especially after it has been subject to biperiodicisation, will fill the whole spectrum with non-zero coefficients and there will be a lot of weak aliasing generated when solving Equation (8), mainly in the Eulerian case. To solve this problem, it was decided to truncate the representation of m between biperiodicisation and return to grid point space. For historical reasons, the truncation is circular and not elliptic (one single cut-off wave number n_t).

Since one may not want to be dependent about the choice of n_t in the handling of the historical file representation of ALADIN, it was decided that this external representation would be performed in terms of real winds. Hence the choice of the truncation cut-off is acting only (and strictly only) when integrating the model.

In practice n_t can be chosen bigger when running a semi-Lagrangian version of ALADIN than for an Eulerian one (for reasons similar to those that will be developed in Section 13) and this leads to both recommended choices of Section 9.

12. GEOMETRICAL ASPECTS OF THE COUPLING PROBLEM

Coupling is performed in ALADIN according to the method of *Davies* (1976). This means that, at the end of each time step, the values obtained without any instantaneous influence of the coupling model are combined with the values interpolated on the ALADIN grid starting from the coupling model. The weight of the coupling value in this linear combination, α (the one for the solution of the LAM coupled model being $(1 - \alpha)$) is obviously zero in the C-zone and one in the E-zone. So the problem treated here reduces to the determination of α in the I-zone.

For any coupled field the bi-dimensional α function depends only on 4 tuning parameters p, n, m and γ, used in the following expressions:

$$\alpha = (p+1).Z^p - p.Z^{p+1} \qquad \text{if } p \geq 2 \tag{9}$$

$$\alpha = 1 - (p+1).(1-Z)^{-p} + p.(1-Z)^{1-p} \qquad \text{if } p \leq -2 \tag{10}$$

where Z is the scaled distance between the edges inside the I-zone.

The value of Z is obvious on the sides of the domain (i.e. equal to either a X or a Y equivalent).

In the corner parts of the I-zone, Z is obtained by solving the following set of implicit equations:

$$Z^e = X^e + Y^e \tag{11}$$

$$e.(n+m)^{n+m}.Z^n.(1-Z)^m = \gamma.n^n.m^m \tag{12}$$

This formulation is chosen in order to have a smooth continuity with the 0 and 1 values in the C- and E-zones. The coefficient p determines the asymmetry of the representation inside the I-zone, while n, m and γ are controlling the shape of the function in the corners. In ALADIN, one always has $n=3$, $m=1$, $\gamma=2$ while p depends of the parameter to be coupled. Please beware that the n and m of this paragraph have nothing to do with previous uses of the same notation (wave numbers and, in the case of m, map factor).

The value of p was adjusted by *Janiskova* (1994), in Eulerian mode, in order to minimise unwanted wave reflections at the edge of the domain. For wind and mass variables, the chosen value is 2.16 (i.e. near symmetry) while for specific humidity it is 5.52, this denoting a sharp variation of α near the edge of the C+I domain (as little coupling as possible for this very fluctuating prognostic quantity). The latter choice confirms the results obtained by *Yessad* (1993) in the framework of the finite-difference model PERIDOT.

The recommended width of the I-zone associated to this tuning is 8 grid-lengths.

All previous considerations in this Section are based on the assumption that the coupling is performed in grid-point space. However, in a spectral model, there also exists the possibility to perform a «spectral coupling», i.e. to impose the largest horizontal scales as those of the coupling model, to let the smallest horizontal scales be entirely determined (even in the E- and I-zone in the ALADIN case) by internal computations and to have a spectral transition zone with a mixed solution in between. Experience shows that the performances of this kind of coupling are becoming worse than those of the grid-point coupling very early in the forecast (*Pescaru* (personal communication)).

There is at least however one area of research where this technique might find a niche of application: in the case of 4D-Var data assimilation, the fact to assume that the «trajectory» found (for the longest waves of the LAM) to be the ideal one (in the coupling model in fact) must be preserved during the optimisation of the

LAM's own trajectory might become an attractive solution to the vexing problem of the inclusion of boundary conditions in the so-called «control variable» of the LAM. From the LAM point of view, the method could be called «decremental» (the minimisation does not concern the largest scales); but it could of course be combined with some «incremental» characteristics (the LAM smallest scales also remain unanalysed), this leading to a control variable limited to intermediate scales (*Gustafsson et al.*, 1997).

13. LINK WITH TIME-STEPPING SCHEMES

As already mentioned, the spectral method is very well suited for use in Eulerian semi-implicit methods, since the Helmholtz operator is inverted via the (anyhow necessary) FFTs and since local derivatives are computed with an infinite precision.

When going to semi-Lagrangian schemes, the situation is slightly modified; the advection part of the equations (in 1-D):

$$\frac{\partial.\psi}{\partial.t} = -u.\frac{\partial.\psi}{\partial.x} \tag{13}$$

is replaced by the differential expression:

$$\frac{\psi_G^{t+\Delta t} - \psi_O^{t-\Delta t}}{2.\Delta t} = 0 \tag{14}$$

where the value of ψ on the target grid point «*G*» at the end of the time step is taken equal to the value of the same quantity at the beginning of the time step at the origin point «*O*» of a trajectory finishing in «*G*».

The value at the origin point (which is surely not a point of the grid) has to be interpolated from values at the neighbouring grid points. This interpolation procedure is already too costly with Lagrange-polynom methods to make any additional meaningful use of the derivatives of the interpolated field on the points of the grid. Hence one of the advantages of the spectral method (the infinite accuracy of the derivatives' computation) is lost.

There is however a hidden potential counterpart to this disadvantage. Replacing Equation (13) by Equation (14) is equivalent, from the point of view of grid point calculations, to replace a product by a linear combination (the interpolation operator is always a linear one). Hence the degree of aliasing of the advection part of the computation is reduced from two to one. If one still admits that the aliasing is dominated by these advection terms (this may be empirically studied), one may now accept to relax the condition on the number of grid points ($L \geq 3\,M + 1$) towards the original «linear» one ($L \geq 2\,M + 1$) or, probably better (since there is still some small aliasing), towards a «semi-linear» compromise like ($L \geq 2.5\,M + 1$).

In fact what is done is not to decrease L at fixed M, but the reverse, i.e. to increase M at fixed L. Hence one can obtain an increase of the spectral resolution (in particular for the orographic forcing, see Section 2) with the same computational grid and hence nearly the same cost.

Finally, if one replaces Equation (13) by Equation (8) and does again the exercise of replacing products by interpolations, it becomes obvious why the recommended values of n_t (truncation cut-off for the map factor m) are so different between the Eulerian and semi-Lagrangian cases (see Section 9, Remark 7).

14. GENERAL FORMULATION OF THE HORIZONTAL DIFFUSION PROBLEM

There are a lot of options and complex formulations in the horizontal diffusion part of ARPEGE and ALADIN thanks to the fact that the spectral formulation allows in principle any choice for a diagonal implicit selective damping of the form:

$$X^{+}_{(m,n)} = \frac{1}{1 + K_X(r,l).\Delta t} X^{-}_{(m,n)} \tag{15}$$

where X is the prognostic variable to be «diffused», K_X a coefficient (dimensioned to the inverse of a time quantity) attributed to this variable and depending on the vertical level l and on the relative wave number r:

* for unstretched ARPEGE $$r = \sqrt{\frac{n(n+1)}{N(N+1)}} \tag{16}$$

* for ALADIN $$r = \sqrt{\left(\frac{n}{N}\right)^2 + \left(\frac{m}{M}\right)^2} \tag{17}$$

while no similar simple expression can be shown for stretched ARPEGE owing to the too large variation of the map factor (for ALADIN the expression is slightly wrong since the map factor is also varying, but this rather small variation is systematically neglected in all considerations concerning horizontal diffusion).

If the dependency of K with respect to r is of the form:

$$K \propto r^q \tag{18}$$

then Equation (15) is the equivalent of:

$$\left(\frac{\partial X}{\partial t}\right)_{hdiff} = -K'.\left(-\Delta X\right)^{q/2} \tag{19}$$

where q, the so-called order of the diffusion, is in principle an even integer (at least in Equation (19)).

Basically it is only in ALADIN that a formulation of the type of Equation (19) is used. Thereafter we shall thus only briefly review the ARPEGE situation and give more technical details for the ALADIN implementation.

There are however currently three common points between all NWP applications inside ARPEGE/ALADIN:

* the coefficients are equal for vorticity, temperature and specific humidity and 9 times bigger for divergence (a translation to spectral horizontal diffusion of the finite difference «divergence damping»); there is obviously no diffusion on surface pressure; in the case of the non-hydrostatic version of ALADIN, the diffusion coefficient should be the same for the vertical divergence as for the horizontal divergence and there should not be any diffusion on the pressure perturbation;

* to avoid the too strong artificial mixing along mountain slopes of air parcels having different temperatures for similar potential temperatures, it is not T that is diffused but $T - \beta \ln(p_s)$ where β is a level-dependent constant, optimised in the conditions of the standard atmosphere;

* the dependency of K on l takes the form:

$$K \propto \frac{y_0}{\min(y_0, p_{ST}(l) / p_{REF})} \tag{20}$$

where y_0 is a tuning parameter to be chosen between 0 and 1, $p_{ST}(l)$ the pressure of the standard atmosphere at level l and p_{REF} the reference pressure of the same standard atmosphere.

Hence K is constant for $\eta \geq y_0$ and increases aloft as the inverse of pressure. The latter choice is made to partly damp fine scale gravity waves propagating towards the top of the model where they will be unduly reflected back (a variant of the so-called «sponge» technique).

15. HORIZONTAL DIFFUSION IN STRETCHED ARPEGE

15.1 Numerical diffusion on the transformed sphere

It is the equivalent of Equations (15), (16) and (18) on the transformed sphere and it is activated for purely numerical reasons, namely the cleaning of the consequences of residual aliasing and of accumulation of variance at the truncation's edge. The order of the operator is usually $q = 4$ for high resolution applications.

Obviously it becomes the only applied diffusion in the unstretched case, even if in practice the additional schemes that will now be described degenerate to it in the case $c=1$.

15.2 «Geographical» part of the so-called «unified diffusion»

In the stretched case one has to combine the effects of the scheme mentioned in «15.1» with those of a scheme that mimics what would be a diffusion operator of homogeneous order and constant intensity on the geographical sphere. Except in the special case $q = 2$ where the problem can be treated exactly, the chosen solution is to approximate the implicit variation of the K' value on the transformed sphere by a polynom of second order in function of the sine μ' of the pseudo-latitude ($a0 + a1.\mu' + a2.\mu'^2$). The operator one tries to approximate is here usually also of fourth order for high resolution applications.

The horizontal diffusion operator then ceases to be diagonal for the pseudo-N/S dependency and becomes by construction pentadiagonal (in the exact case $q = 2$ like in the approximate case $q \neq 2$). The solution of the associated linear system remains at a still acceptable price (like for the semi-implicit case, see Section 5) and the algorithm keeps its crucial implicit character.

15.3 Geographical horizontal diffusion in Fourier space

The difficulties associated with the use of the solution described in «15.2» and its approximate character can be eliminated if, following *Li et al.* (1994), one elects to do the horizontal diffusion in the intermediate step between the Fourier and Legendre transforms of the model time step. The operator remains spectral and hence diagonal in the space of the Fourier coefficients for one given pseudo-latitude row but the problem ceases to be spectral in the pseudo-N/S direction. When treated by a finite element method the problem reduces to the combined inversion of tridiagonal matrices, with the advantage that the local intensity of the operator can be chosen such as to obtain an exact equivalent of Equation (19), again with $q = 4$, the only currently coded version (*Tolstykh* (personal communication)).

Like the previous one, this scheme (not yet in operational status) is currently combined with a component of the type described in «15.1» (for numerical reasons), even if we are here considering a possible merge of the two components.

16. HORIZONTAL DIFFUSION IN ALADIN

As already mentioned it exactly follows Equations (15), (17) and (18) in the elliptic truncation (we assume here $\delta x = \delta y$). The order is $q = 4$ here also. For the vertical variation one has $y_0 = 1$, i.e. a variation of the intensity of K starting right from the bottom.

An important issue (that also exists in ARPEGE but is obscured there by the coexistence of two methods) is that of the tuning of K when changing model resolution, i.e. truncation and/or stretching in ARPEGE or δx (= δy) in ALADIN.

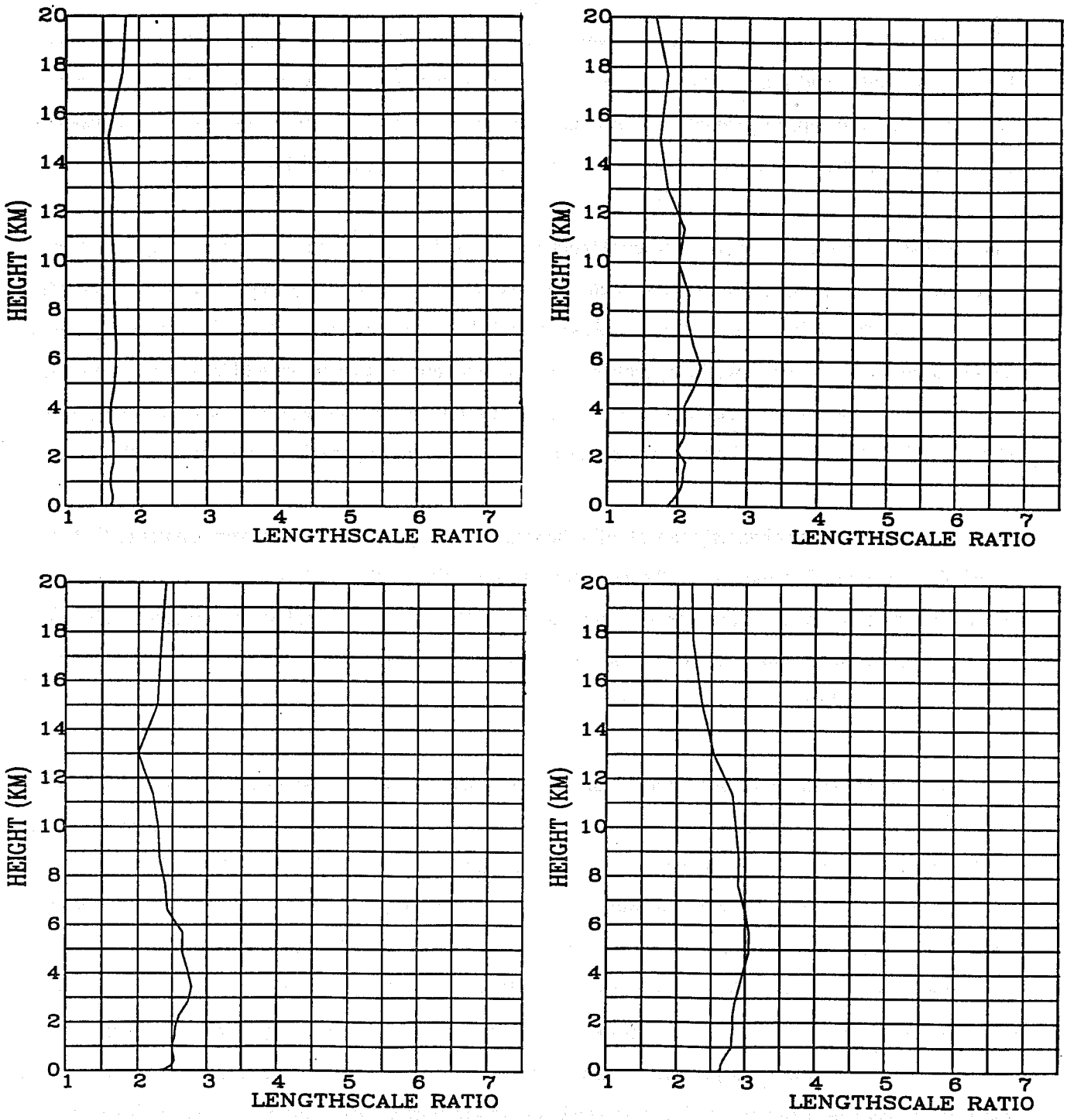

Figure 9: Ratios between the maximum and minimum horizontal length scales of the structure functions of a model with c=3.5; top left:stream function; top right: velocity potential; bottom left: temperature; bottom right: specific humidity

space in which the minimisation is performed (i.e. «control variable»). For the sake of simplicity, this will also be the case for the first 4D-Var application of ARPEGE. But the contradiction between the then very entangled model- and assimilation aspects will become paramount and some degree of stretching will have to be introduced soon after in the control variable and/or the structure functions. Given the above mentioned uncertainty about the ideal stretching of the latter, one hope is that it will be sufficient to act only on the control variable, letting the dynamics of an intermediately stretched tangent linear model recover the variations of the structure functions during the iterative forward-backward integration process.

However, even in this case (in fact in any case where the degree of stretching will differ between the structure functions and the control variable), we shall be back with the very expensive use of the dilatation matrices once necessary for normal mode initialisation. Additional problems may thus be expected again in this framework (see Section 7). No similar problem is expected with ALADIN, provided of course that the integration domains remain reasonably small. Indeed, like for horizontal diffusion (see Section 14), the hypothesis that one can partly neglect the horizontal variations of the map-factor will then remain a valid one. On the other hand, new problems will arise concerning the coupling and the way to handle it in the variational framework (see for instance Section 12).

The second basic argument in favour of spectral methods is linked to the enormous computational burden that 3D-Var (in terms of memory) and 4D-Var (in terms of memory and computing time) are requiring. The only currently acceptable way to escape this problem is to work in «incremental mode» (*Courtier et al.*, 1994). One elects to have less degrees of freedom in the control variable than in the forecasting model. In the case of the spectral method this can be achieved with one single tuning parameter, by reducing the truncation of the representation: only the largest scales are updated in the analysis step of the data assimilation process, the smaller ones being supposed to adapt during the forecasting part of the cycle. Of course multi-grid methods may play the same role in the case of a local representation of the fields, but the simplicity of the spectral method surely remains an advantage here.

19. SUMMARY

Spectral methods have nothing «magic», they simply are another way to discretise in the horizontal, with no unnecessary ancillary question to solve and with nice computational properties. From the point of view of algorithms, the «respiration» of a spectral time-step is quasi-perfect in semi-implicit Eulerian mode. The interaction with the semi-Lagrangian methods destroys part (but not all) of this crucial advantage.

The mathematics behind the spectral method is merely an application of linear algebra theorems. The practical application however strongly depends on the existence of powerful Fast Fourier Transform (FFT) algorithms.

When considered from a practical point of view, at equal computing costs, the advantages of the spectral method somehow outweigh those of the finite differences method. This is probably also true with respect to the finite elements method, albeit in a less «clear cut» manner.

The spectral method can be equally well employed for global uniform, global «zoomed» or LAM applications, in the latter case with a marginal additional problem that can be solved in at least three different ways (for example through the use of a «toroidal» geometry in the described ALADIN case).

The Schmidt transform represents the only practical way to solve the variable mesh problem in a global spectral framework.

In LAM spectral applications there is no reason to leave the isotropic framework of conformal projections; the analogy with the global case is then paradoxically stronger than for the (rotated) lat/lon geometry.

In the link between model's dynamics and physics the spectral fitting acts as a welcome filter (less grid-point storms for the convective parameterisation, for example). Consequently, at high resolution, the concept of linear (or semi-linear) grids for semi-Lagrangian integrations appears less attractive.

The spectral method is the most natural way to define the «control variable» of variational methods, especially when this technique has to be applied in the «incremental» way. For global applications, the use of the ARPEGE stretched geometry in this framework will however represent a tough challenge.

Last but not least, there are now at least seven operational applications of the ALADIN concept and one of the stretched ARPEGE one. Owing to the difficulties to push an initial idea up to this ultimate «proof of the pudding», Machenhauer and Haugen on one hand and Schmidt on the other hand deserve a lot of credits.

ACKNOWLEDGEMENTS

The author wishes to thank all people that have been associated with either of the three parts of the IFS/ARPEGE/ALADIN effort and that helped therefore to demonstrate the feasibility of modern operational and research NWP applications of spectral methods, in many geometrical conditions. Special thanks are due to Gérald Desroziers, Alain Joly and Metodi Marku for allowing the use of unpublished results. Eric Bazile reran the sensitivity study about the «Vaison-la-Romaine» case. Jean-Marcel Piriou prepared the operational grid pictures. Jean Pailleux and Jean-Noël Thépaut helped to improve the manuscript. With his spectral HIRLAM experience, the help of Nils Gustafsson in the early stages of the ALADIN project was invaluable. Finally the patience of Michel Rochas, Bennert Machenhauer, Michel Jarraud and Philippe Courtier in guiding the author over the years in the maze of the mathematical basis of spectral computations is gracefully acknowledged here.

REFERENCES

Arakawa, A., 1983: Vertical differencing of the primitive equations. ECMWF Seminar Proceedings on «Numerical Methods for Weather Prediction», 5-9 September 1983, vol. I, 207-223.

Bénichou, P. and E. Legrand, 1992: Validation of ARPEGE/IFS for operational short-range predictions. ECMWF Seminar Proceedings on «Validation of Models over Europe», 7-11 September 1992, vol. II, 41-74.

Bourke, W., 1974: A multi-level spectral model. Part I: Formulation and hemispheric integrations. Mon. Wea. Rev., 102, 687-701.

Bouteloup, Y., 1995: Improvement of the spectral representation of the earth topography with a variational method. Mon. Wea. Rev., 123, 1560-1573.

Bubnova, R., 1998: Vertical discretisation problems for very high resolution LAM; ARPEGE/ALADIN as example. This issue of the ECMWF 1998 Seminar Proceedings.

Bubnova, 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.

Caian, M. and J.-F. Geleyn, 1997: Some limits to the variable-mesh solution and comparison with the nested-LAM solution. Q. J. R. Meteorol. Soc., 123, 743-766.

Chen, Q.-S. and Y.-H. Kuo, 1992: A harmonic-sine series expansion and its application to partitioning and reconstruction problems in a limited area. Mon. Wea. Rev., 120, 91-112.

Clochard, J., 1990: Changement de repère local et transformation du vent dans la transformation «ARPEGE». Note ARPEGE N° 14, available from Météo-France, 5 pp..

Côté, J. and A. Staniforth, 1988: A two-time level semi-Lagrangian semi-implicit scheme for spectral models. Mon. Wea. Rev., 116, 2003-2012.

Courtier, P., C. Freydier, J.-F. Geleyn, F. Rabier and M. Rochas, 1991: The ARPEGE Project at Météo-France. ECMWF Seminar Proceedings on «Numerical Methods in Atmospheric Models», 9-13 September 1991, vol. II, 193-231.

Courtier, P. and J.-F. Geleyn, 1988: A global numerical weather prediction model with variable resolution: Application to the shallow-water equations. Quart. J. Roy. Meteor. Soc., 114, 1321-1346.

Courtier P., J.-N. Thépaut and A. Hollingsworth, 1994: A strategy for the operational implementation of 4D-Var, using an incremental approach. Quart. J. Roy. Meteor. Soc., 120, 1367-1387.

Da Silva, A., J. Pfaendtner, J. Guo, M. Sienkiewicz and S. Cohn, 1995: Assessing the effects of data selection with DAO's Physical-space Statistical Analysis System. Proceedings of the Second International WMO Symposium on Assimilation of Observations in Meteorology and Oceanography, Tokyo, 13-17 March 1995. WMO/TD N° 651, 273-278.

Davies, H.C., 1976: A lateral boundary formulation for multi-level prediction models. Quart. J. Roy. Meteor. Soc., 102, 405-418.

Desroziers, G., V. Mathiot, F. Orain et P. Bernardet, 1995: Estimation locale des covariances des erreurs de prévision d'un modèle spectral sur la sphère: Application au modèle à résolution variable ARPEGE. Note GMME N° 34, available from Météo-France, 50 pp..

Eliasen, E., B. Machenhauer and E. Rasmussen, 1970: On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. Report N° 2, Institut for Teoretisk Meteorologi, University of Copenhagen.

Geleyn, J.-F., 1988: Transformation de coordonnées pour aller de la sphère terrestre réelle à la sphère fictive de la grille de collocation et vice-versa. Note ARPEGE N° 3, available from Météo-France, 8 pp..

Gustafsson, N., P. Lönnberg and J. Pailleux, 1997: Data assimilation for high-resolution limited-area models. «Data assimilation in meteorology and oceanography: Theory and practice». Special issue (WMO Symposium 13-17/3/95) of the J. Meteor. Soc. Japan, 257-272.

Gustafsson, N. and A. Mc Donald, 1996: A comparison of the HIRLAM gridpoint and spectral semi-Lagrangian models. Mon. Wea. Rev., 124, 2008-2022.

Hardiker, V., 1997: A global numerical weather prediction model with variable resolution. Mon. Wea. Rev., 125, 59-73.

Haugen, J.E. and B. Machenhauer, 1993: A spectral limited-area formulation with time-dependent boundary conditions applied to the shallow-water equations. Mon. Wea. Rev., 121, 2618-2630.

Hortal, M., 1998: Aspects of the numerics of the ECMWF model. This issue of the ECMWF 1998 Seminar Proceedings.

Hortal, M. and A.J. Simmons, 1991: Use of reduced Gaussian grids in spectral models. Mon. Wea. Rev., 119, 1057-1074.

Hoyer, J.-M., 1987: The ECMWF spectral limited area model. ECMWF Workshop Proceedings on «Techniques for Horizontal Discretisation in Numerical Prediction Models», 2-4 November 1987, 343-359.

Janiskova, M., 1994: Study of coupling problem. Note ALADIN N° 1, available from Météo-France, 17 pp..

Jarraud, M. and C. Girard, 1983: An extensive quasi-operational comparison between a spectral and a grid point model. ECMWF Seminar Proceedings on «Numerical Methods for Weather Prediction», 5-9 September 1983, vol. II, 61-111.

Jarraud, M. and A.J. Simmons, 1983: The spectral technique. ECMWF Seminar Proceedings on «Numerical Methods for Weather Prediction», 5-9 September 1983, vol. II, 1-59.

Joly, A., 1992: ARPEGE/ALADIN: Adiabatic model equations and algorithm. Internal documentation note, available from Météo-France, CNRM/GMAP.

Juang, H.-M. H. and M. Kanamitsu, 1994: The NMC nested regional spectral model. Mon. Wea. Rev., 122, 3-26.

Kuo, H.-C. and R.T. Williams, 1992: Boundary effects in regional spectral models. Mon. Wea. Rev., 120, 2986-2992.

Li, Y., S. Moorthi and J.R. Bates, 1994: Direct solution of the implicit formulation of fourth order horizontal diffusion for grid point models on the sphere. NASA Technical Memorandum N° 104606.

Lynch, P., D. Giard and V. Ivanovici, 1997: Improving the efficiency of a digital filtering scheme for diabatic initialisation. Mon. Wea. Rev., 125, 1976-1982.

Machenhauer, B., 1979: The spectral method. Numerical methods used in atmospheric models. GARP publication series N° 17, 121-275.

Machenhauer, B., 1991: Spectral methods. ECMWF Seminar Proceedings on «Numerical Methods in Atmospheric Models», 9-13 September 1991, vol. I, 3-85.

Machenhauer, B. and J.E. Haugen, 1987: Test of a spectral limited area shallow water model with time dependent lateral boundary conditions and combined normal mode/semi-Lagrangian time integration schemes. ECMWF Workshop Proceedings on «Techniques for Horizontal Discretisation in Numerical Prediction Models», 2-4 November 1987, 361-377.

Marku M., 1998: Etude dans ALADIN de l'impact d'un passage en grille de calcul linéaire ou semi-linéaire pour une advection de type semi-Lagrangien. Note de travail de l'ENM N° 625, available from Météo-France, 71 pp..

Members of the ALADIN international team, 1997: The ALADIN project: Mesoscale modelling seen as a basic tool for weather forecasting and atmospheric research. WMO Bulletin, 46-4, 317-324.

Mohlenkamp, M.J., 1997: A fast transform for spherical harmonics. PhD Thesis. Yale University.

Navarra, A., W.F. Stern and K. Miyakoda, 1994: Reduction of the Gibbs oscillation in spectral model simulations. J. Climate, 7, 1169-1183.

Orszag, S.A., 1970: Fourier series on spheres. Mon. Wea. Rev., 102, 56-75.

Parrish, D. and J. Derber, 1992: The National Meteorological Center's Spectral Statistical-interpolation Analysis System. Mon. Wea. Rev., 120, 1747-1763.

Radnoti, G., 1995: Comments on «A spectral limited-area formulation with time-dependent boundary conditions applied to the shallow-water equations». Mon. Wea. Rev., 123, 3122-3123.

Ritchie, H., C. Temperton, A. Simmons, M. Hortal, T. Davies, D. Dent and M. Hamrud, 1995: Implementation of the semi-Lagrangian method in a high-resolution version of the ECMWF forecast model. Mon. Wea. Rev., 123, 489-514.

Rochas, M. et P. Courtier, 1992: La méthode spectrale en météorologie. Note ARPEGE N° 30, available from Météo-France, 58 pp..

Rochas, M., K. Yessad et J. Clochard, 1991: Transformations de sphère à sphère; Rotations et dilatations dans l'espace spectral. Note ARPEGE N° 19, available from Météo-France, 51 pp..

Schmidt, F., 1977: Variable fine mesh in spectral global model. Beitr. Phys. Atmosph., 50, 211-217.

Tatsumi, Y., 1986: A spectral limited-area model with time dependent lateral boundary condition and its application to a multi-level primitive equation model. J. Meteor. Soc. Japan, 64, 637-664.

Temperton, C., 1991a: On scalar and vector transform methods for global spectral models. Mon. Wea. Rev., 119, 1303-1307.

Temperton, C., 1991b: Finite element methods. ECMWF Seminar Proceedings on «Numerical Methods in Atmospheric Models», 9-13 September 1991, vol. I, 103-117.

Yessad, K., 1993: Spin-up like effect diagnosed with the theoretical invariant concept. Beitr. Phys. Atmosph., 66, 201-221.

Yessad, K. and P. Bénard, 1996: Introduction of a local mapping factor in the spectral part of the Météo-France global variable mesh numerical weather forecast model. Quart. J. Roy. Meteor. Soc., 122, 1701-1719.