On the design and scaling of ARPEGE diffusions

Pierre Bénard
CNRM/GMAP/ALGO 11 July 2002
file : memohdf.tex

1 Introduction

This document summarizes the principles for implementing an horizontal diffusion in a stretched model and lists the problems encountered in the development and validation of HDF scheme, and its current status. One of the goals is to have a written material in order to avoid further errors. Lack of formalization can sometimes be a source of problems with very serious consequences.

The usual form of a HD operator is:

$$\frac{\partial X}{\partial t} = -K\nabla^r X \tag{1}$$

where K is called "diffusion coefficient" r is the diffusion order, and the meaning of ∇ is explained in next paragraph. Let X_0 be an eigenmode of the ∇ operator, with the associated eigenvalue λ. We can define the damping-time of the HD operator for this eigenmode by $\tau(X_0) = 1/(K\lambda^r)$. For a given discretized-model configuration, we define τ_s as the damping time of the smallest resolved scale due to the diffusion. This notion is useful when examining the resolution-dependency properties of the diffusion operators in subsequent sections.

For a limited area model such as Aladin, the mapping factor remains close from one, provided the domain is not too large, hence the geographical ∇ operator and the reduced ∇' operator (on the projected geometry) are very close. However, for a global stretched model as ARPEGE, if the stretching factor is chosen significantly larger than 1, the definition of the ∇ operator must be precised. Traditionally the geographical derivative operator is still noted ∇, while the reduced derivative operator (in the stretched geometry) is noted ∇'. For ARPEGE, we have

$$\nabla = \mathrm{m}\nabla'$$

where m is the local mapping factor.

In this paper we will recall the following concepts:

- "Numerical" diffusion: A diffusion will be said to be "numerical" if it is expressed as a pure power of the reduced operator ∇'. The spectral numerical diffusion of ARPEGE is HDI. This operator equally damps all spectral components which have the same total wavenumber in the spectral space. Hence this diffusion will damp more strongly a wave with a given physical wavelength near the Antipode than near France.

  The numerical diffusion operator writes:

  $$\frac{\partial X}{\partial t} = -K\nabla'^r X = -K(\nabla^r/\mathrm{m}^r)X \tag{2}$$

- "Geographical" diffusion: A diffusion will be said to be "geographical" if it is expressed as a pure power of the geographical derivative operator ∇. This operator equally damps all waves which have the same total physical wavenumber in the real world.

  The geographical diffusion operator writes:

  $$\frac{\partial X}{\partial t} = -K\nabla^r X \tag{3}$$

Important: In the following, for stretched geometries, we will call "geographical" diffusion-coefficient the coefficient K of the geographical derivative operator ∇ in the above equation (3). This is to avoid the ambuiguous definition of the general word "diffusion-coefficient", which is sometimes applied to the coefficient of the numerical operator ∇' (e.g. in ARPEGE documentations). It should be noted that for a pure geographical diffusion operator, the geographical diffusion-coefficient is a space-independent number, while for non-geographical diffusion operators, the geographical diffusion-coefficient may be a space-dependent function

2 HD coefficients in non-stretched models

In this preliminary section, we first examine how must be specified the dependency of HD coefficents to the resolution, for non-stretched discretized models (this will be useful as guideline when going to stretched models).
Let assume that a diffusion is well tuned for a given discretized-model configuration. When we decide to increase the resolution, we must specify a new diffusion tuning. Three solutions are possible as depicted on Fig. 1:

(i) we can choose to have the same damping-time of the HD operator for the same physical wave, i.e. the geographical diffusion coefficient K is kept constant. The formulation of the operator is then :

$$\frac{\partial X}{\partial t} = -k\nabla^r X \tag{4}$$

where k is a pure constant. In this case τ_s is proportional to Δx^r. However, this results in too much diffusion for small waves, and the model is found to rapidly loose energy when increasing the resolution.

(ii) Hence we could choose to have the same damping-time for the smallest wave (whatever size be it according to the resolution change), i.e. the geographical diffusion coefficient is proportional to Δx^r, i.e. τ_s is constant. The formulation of the operator is then :

$$\frac{\partial X}{\partial t} = -k\Delta x^r\nabla^r X \tag{5}$$

where k is a pure constant. But in this case, the model is found to have too much energy in the new resolved small scales when increasing the resolution.

(iii) The behaviour is found experimentally to be optimal when the geographical diffusion coefficient is proportional to Δx^{r-1}. Then the operator writes:

$$\frac{\partial X}{\partial t} = -k\Delta x^{r-1}\nabla^r X \tag{6}$$

where k is a pure constant. In this case, τ_s is proportionnal to Δx. The "pertinent" coefficient k which describes the diffusion in a resolution-independent way then has the dimension of a speed. In ARPEGE and Aladin, this particular solution is chosen for the specification of the dependency of the geographical diffusion coefficient to the resolution, as well as in many other centres.

N.B: It should be noted that all the results of this section were purely experimental, and there is no obvious "scientific" justification accessible by just examining the equations.
Summary: For a non-stretched model, the diffusion coefficient should be proportional to Δx^{r-1} where r is the order of the diffusion. As a consequence, the damping-time of the shortest wave should be proportional to Δx.

3 "Target" diffusion operator for stretched models

In the stretched context, we will now show that the diffusion operator cannot be purely numerical or geographical. On contrary, it must must be "hybrid", i.e. intermediate between numerical and geographical. We will see that the nature of this hybrid character can be determined using the experimental results presented in the previous section.

The key idea is to specify that for any location, the diffusion operator of the stretched model should behave like the diffusion operator of a non-stretched model with the same horizontal resolution.
Let m be the local map-factor of the stretched geometry at a given location. The local resolution of the stretched model is then:

$$\Delta x(\mathrm{m}) = \Delta x_0/\mathrm{m}$$

where Δx_0 is the mesh of the stretched model at the location where $m = 1$. The diffusion operator of a non-stretched model with the same resolution would be given by (6):

$$\frac{\partial X}{\partial t} = -k\,\Delta x(\mathrm{m})^{r-1}\,\nabla^r X$$

where k is a constant. The diffusion operator of the stretched model must follow the same form, hence we must have:

$$\frac{\partial X}{\partial t} = -k\Delta x_0^{r-1}\,\mathrm{m}^{1-r}\,\nabla^r X \tag{7}$$

We see that the resolution-dependency of the Aladin model diffusion **coefficient** implies a resolution-dependency of the stretched model diffusion **operator**. As a result, the optimal diffusion operator for a stretched geometry is neither the numerical operator ∇'^r nor the geographical operator ∇^r, but an hybrid operator: $\mathrm{m}^{1-r}\nabla^r$.
Generally, the diffusion operator of the stretched model is specified in terms of the numerical derivative operator ∇'. The target diffusion operator for the stretched model then writes:

$$\frac{\partial X}{\partial t} = -k\Delta x_0^{r-1}\,\mathrm{m}\nabla'^r X$$

Summary: For a stretched model, the space-dependency of the diffusion operator is fully constrained by the choice made for the resolution-dependency of geographical diffusion coefficient for non-stretched models. Thus for stretched ARPEGE the diffusion operator must have (or should approach) the following form:

$$\frac{\partial X}{\partial t} = -K\mathrm{m}^{1-r}\nabla^r X = K\mathrm{m}\nabla'^r X$$

where K is a space-independent diffusion coefficient. Hence the geographical diffusion coefficient of the stretched ARPEGE must be space-dependent, and proportional to m^{1-r}.

This "ideal" operator will be called "target" operator hereafter.

4 Notations for the ARPEGE stretched geometry

The stretched ARPEGE model uses a particular analytical stretching transformation which allows an algebraic treatment of the equations in the stretched geometry.
Let θ be the latitude on the non-stretched sphere, and Θ the latitude on the stretched sphere. We note $\mu = \sin\Theta$. The ARPEGE stretching transformation is defined by:

$$\theta(\mu) = \arcsin\left(\frac{a\mu + b}{a + b\mu}\right) \tag{8}$$

where $a = (c^2+1)/2c$ and $b = (c^2-1)/2c$ and c is the stretching factor. The local map factor is given by:

$$\mathrm{m}(\mu) = \frac{\partial\Theta}{\partial\theta} = a + b\mu \tag{9}$$

The ARPEGE grid is almost regularly spaced in Θ ($\Delta\Theta \approx \mathrm{Cst}$), hence the θ discretized latitude are non uniformly distributed. Locally, the grid-spacing are related by:

$$\Delta\Theta(\mu) = m(\mu)\Delta\theta(\mu) \tag{10}$$

we have:

$$\nabla = \frac{\partial}{\partial\theta} = \frac{\partial\Theta}{\partial\theta}\frac{\partial}{\partial\Theta} = m\frac{\partial}{\partial\Theta} = m\nabla' \tag{11}$$

5 Graphical examination of the spatial behaviour of diffusion operators in ARPEGE

The hybrid character of any diffusion operator can be graphically presented by comparison with the target formulation (7) as shown e.g. in Fig. 2 for the purely numerical and geographical diffusion operators.
The abcissa axis represents the sine of latitude in the stretched sphere, μ. At the pole of interest $\mu = 1$, at the antipode, $\mu = -1$, and $\mu = 0$ is the equator of the transformed sphere. The ordinate axis represents the strength of the diffusion for an arbitrary physical wavelength, scaled by the target one, assuming that all diffusion are tuned in such a way that their strength exactely coincide at the pole of interest. Note that the ordinate axis is in logarithmic scale. The strength of the diffusion is given by the inverse of the damping time of the wave.
The curves for the various operators are obtained by writing them under the form:

$$\frac{\partial X}{\partial t} = -K f(\mu) \nabla^r X$$

and by making $K\nabla^r = 1$ in the above equation. The result is then scaled by the target operator m^{1-r}, and all functions are assumed equal at $\mu = 1$, hence the plotted function is :

$$f_{\text{plotted}}(\mu) = \frac{f(\mu)\ \mathrm{m}(1)^{1-r}}{f(1)\ \mathrm{m}(\mu)^{1-r}}$$

A curve superimposed to the abcissa axis means that the corresponding diffusion has everywhere the correct hybridicity i.e. the ideal spatial diffusion repartition in the sense of (7).
The geographical diffusion is seen to have not enough damping away from the pole of interest, compared to the target operator. Conversely, the numerical diffusion is too much damping away from the pole of interest, although the discrepancy with the reference is much smaller (roughly a factor 10 instead of 1/2000).
The figure is shown for $c = 3.5$ and $r = 4$. In comparison, all these discrepancies would increase for increasing values of c and r.

6 The HDU diffusion in ARPEGE

In this section the behaviour of the HDU spectral hybrid diffusion in ARPEGE is examined. The "pure" (i.e. space-continuous) HDU is first examined. We show that this "pure" form is already a first approximation of the above target operator. Then the "practical" (i.e. space-discretized) HDU is presented. This "practical" implementation makes use of a second approximation in the definition of the map-factor used to compute the geographical component. We will show that this second approximation partly compensate the effect of the first one. At last, we present a "new HDU" possible formulation, which exactely matches the target operator for both the "pure" and "practical" forms.

6.1 The "pure" current HDU operator

For the spectral HDU diffusion of ARPEGE, the hybrid character is currently introduced in an approximate way by a weighted average between the pure numerical and geographical operators.

$$\frac{\partial X}{\partial t} = -K(\alpha_n \nabla'^r + \alpha_g \nabla^r)X = -K\left(\frac{\alpha_n}{\mathrm{m}^r} + \alpha_g\right)\nabla^r X \tag{12}$$

The HDU operator thus has two "components", a numerical one and a geographical one.
Let $\beta = \alpha_n/\alpha_g$ be the relative proportion of numerical and geographical diffusion. A correct balance of the diffusion strength between the two poles is obtained if the ratio of the geographical diffusion-coefficients between the two poles is the same as for the target operator. This writes:

$$\frac{\frac{\alpha_n}{c^r} + \alpha_g}{\alpha_n c^r + \alpha_g} = \frac{c^{1-r}}{(1/c)^{1-r}}$$

$\Rightarrow$

$$\frac{\frac{\beta}{c^r}+1}{\beta c^r+1}=\frac{c^{1-r}}{(1/c)^{1-r}}$$

$\Rightarrow$

$$\beta=\frac{c^r-(1/c)^r}{c-(1/c)}\approx c^{r-2} \tag{13}$$

The "pure" current HDU operator thus write:

$$\frac{\partial X}{\partial t}=-K\left(\frac{\beta}{\mathrm{m}^r}+1\right)\nabla^r X \tag{14}$$

i.e.:

$$\frac{\partial X}{\partial t}=-K\left(\frac{\beta}{\mathrm{m}^r}+1\right)\nabla^r X \tag{15}$$

The ratio of the diffusion strength between the two poles is correct compared to the target operator, but not for other locations. The amount of incorrectness for this "pure" HDU is depicted in Fig. 3 for $c=3.5$ and $r=4$. It indicates that in the hemisphere of interest, the "pure" HDU behaves too much as a geographical diffusion due to the choice of the formulation. A correct tuning of diffusion for the pole of interest results in a diffusion too weak by a factor 3–5 near the equator of the transformed sphere.

6.2 The "practical" current HDU operator

The spectral diffusion is necessarily formulated with the ∇' operator, by construction. The introduction of the exact m^r in (15) would require the spectral multiplication by a r-order polynomial in μ. Since the method is similar to the "interaction-coefficient" method, this would lead to an expensive algorithm which would involve $(2r+1)$-diagonal banded matrices. Moreover, spectral components with large total wave-numbers $n>\mathtt{NSMAX}-r$ cannot be treated properly with this method due to the finite truncation of the model, and this may lead to a slight "aliasing" for these high spectral components.

Hence the current "practical" implementation of HDU makes use of an important approximation which substantially modifies its hybrid behaviour: the exact exponentiated map factor m^r used to compute the geographical component in spectral space is replaced by an approximation $\widehat{\mathrm{m}^r}$ consisting in a second-order polynomial in μ:

$$\widehat{\mathrm{m}^r}=a_0+a_1\mu+a_2\mu^2$$

Moreover, the proportion of numerical and geographical components is also specified in an approximated way. We saw above that the correct balance the diffusion strength between the two poles requires $\alpha_n=\beta\alpha_g\approx c^{r-2}\alpha_g$. For the operational $r=4$, this means $\beta\approx c^2$. In HDU, the relative proportion is set by default to c for the numerical diffusion and $(1/c)$ for the geographical diffusion. The operational NAMELIST values also specify this proportion. This choice $(\beta=c^2)$ enforces an almost correct balance of diffusion strength between the two poles, when $r=4$, but not for other diffusion orders (however, the operational diffusion order is set to 4, hence all this paragraph is of no significant practical consequence in the present discussion).

The resulting scheme then writes:

$$\frac{\partial X}{\partial t}=-K\left(c^r+\frac{\widehat{\mathrm{m}^r}}{c}\right)\nabla'^r X=-K\left(\frac{c^r}{\mathrm{m}^r}+\frac{\widehat{\mathrm{m}^r}}{c\,\mathrm{m}^r}\right)\nabla^r X$$

These approximations brings a further modification of the hybrid character for the "practical" HDU, as seen in Fig. 3. The result is to partly compensate the initial lack of diffusion between the two poles. The lack of diffusivity is now reduced to a maximum value of 40 %.

Summary: The current practical HDU diffusion has an almost correct balance for the two poles compared to the target operator. However, it has slight lack of diffusivity between the two poles, compared to the target operator. This lack is probably with no significant consequence.

6.3 Proposal for a "new" HDU operator

The hybrid character could be simply introduced in the spectral diffusion by:

$$\frac{\partial X}{\partial t} = -Km\nabla'^{r} X$$

The HDU operator would then have only one "component", the hybrid one. This formulation has sereval advantages compared to the current one:

- it is exactely equal to the target operator.
- no approximation of m is required.
- it is simpler since the involved matrices are tri-diagonal whatever is the value of r
- it makes simpler the definition of space-, resolution-, and stretching-independent default values.

The spatial behaviour for the "new" HDU is not plotted since it coincides with the target operator by construction.

Simulating the new-HDU diffusion in the current one is very simple:
- Set the numerical part of current HDU to zero (HDIRDIV=0)
- Enforce $\widehat{\mathrm{m}^r}$ to m by setting: $a_0 = a$, $a_1 = b$, $a_2 = 0$.
- Modify HDURDIV to the correct value

7 The HDF diffusion in ARPEGE

In this section the behaviour of the Fourier-space hybrid diffusion HDF in ARPEGE is examined. The "pure" (i.e. space-continuous) current HDF operator is first presented. As for the current HDU, this operator is shown to be an approximation of the target operator, with the correct balance at the two poles. Then the "practical" (i.e. space-discretized) is examined. The effect of the space-discretization is to modify the spectral response, resulting in less selectivity.
It should be noted that the order of HDF diffusion cannot be chosen and is set to 4. This will be assumed in all this section.

7.1 The "pure" current HDF operator

For HDF, the hybrid character is controlled by a NAMELIST parameter REXPKDX, noted ϵ here:

$$\epsilon = \texttt{REXPKDX}$$

This controls the definition of a modified stretching coefficient (called "pseudo-stretching" coefficient hereafter):

$$c^*(\epsilon) = c^{\epsilon/4}$$

This new stretching coefficient c^* is then used to define a pseudo-stretched latitude θ^* by:

$$\theta^*(\mu) = \arcsin\left(\frac{a^*\mu + b^*}{a^* + b^*\mu}\right)$$

where $a^* = (c^{*2}+1)/2c^*$ and $b^* = (c^{*2}-1)/2c^*$. The map-factor is then:

$$\mathrm{m}^*(\mu) = \frac{\partial\Theta}{\partial\theta^*} = a^* + b^*\mu \tag{16}$$

The pseudo-stretched latitude θ^* is then used to define a pseudo-stretched derivative operator by:

$$\nabla^* = \frac{\partial}{\partial\theta^*} = \frac{\partial\Theta}{\partial\theta^*}\frac{\partial}{\partial\Theta} = \mathrm{m}^*\frac{\partial}{\partial\Theta} = \mathrm{m}^*\nabla' \tag{17}$$

The diffusion operator then writes:

$$\frac{\partial X}{\partial t} = -K\nabla^{*4}X = -K\mathrm{m}^{*4}\nabla'^{4}X \tag{18}$$

If REXPKDX=0, then $\mathrm{m}^* = 1$ and the diffusion is purely numerical.
If REXPKDX=4, then $\mathrm{m}^* = \mathrm{m}$ and the diffusion is purely geographical.
The parameter REXPKDX is currently set to:

$$\mathtt{REXPKDX} = 1 \tag{19}$$

The corresponding operator is not identical to the target one because $\mathrm{m}^*(\mu) \neq \mathrm{m}(\mu)^{1/4}$. However, the ∇^* operator coincides with the target operator at the two poles, hence HDF has a correct global balance for the two poles compared to the target operator (but not at other locations).
The discrepancy between HDF and target operators can be seen in Fig. 4 for $c = 3.5$ and $r = 4$. The discrepancy is in the opposite direction than with HDU: a correct tuning of the diffusion for the pole of interest results in a diffusion too strong by 30–40 % near the equator of the transformed sphere.

7.2 The "practical" current HDF operator

The Laplacian operator used to compute the HDF diffusion is defined through a finite-difference discretization in the grid-point space. As a consequence, the spectral response of the obtained grid-point HDF operator $\widetilde{\nabla^{*4}}$ is weaker for the shortest waves than the one of the exact or spectral one $m^{*4}\nabla'^{4}$. This effect, which has no significant consequence on the spatial response of the operator, hence the above plot also applies to the "practical" HDF operator.
The spectral response factor of the HDF operator scaled by the exact response is given by:

$$\left(\frac{\sin(\pi x/2)}{\pi x/2}\right)^4$$

where x is the reduced horizontal wave-number: for a linear grid, $x \in [0,1]$, and for a quadratic grid, $x \in [0,(2/3)]$. The spectral response of the HDF operator is shown in Fig. 7. For the shortest wave, and with a linear grid, the HDF operator has 6 times too weak response compared to the exact operator.

In order to partially compensate this effect, the whole HDF scheme is enhanced by a factor $(\pi/2)^\rho$ where ρ is represented by the NAMELIST parameter RSPVSGP, and is currently set to 1. The intensity of the diffusion for the exact, initial HDF, and enhanced HDF operator is shown in Fig. 8 as a function of x. The HDF operator, be it enhanced, appears to suffer from a significant lack of scale-selectivity compared to the exact or spectral operator. The enhancement allows to have a satisfactory response up to the $4\Delta x$ wave, but all the short waves are insufficiently damped.
This lack of scale selectivity a a well-known disadvantage of grid-point discretizations, and was the initial justification for using a spectral method in ARPEGE and IFS.
Finally, the "practical" HDF diffusion operator writes:

$$\frac{\partial X}{\partial t} = -K\widetilde{\nabla^{*4}}X \tag{20}$$

where $\widetilde{\nabla^{*4}}$ is the grid-point approximation of $m^{*4}\nabla'^{4}$.

Summary: The HDF has a correct balance for the two poles compared to the target operator. However, it has a slight execedent of diffusivity between the two poles compared to the target operator. As for HDU, this slight discrepancy is probably with no practical consequences. Moreover, HDF has a lack of scale selectivity inherent in any grid-point operator.

7.3 Alternative HDF operator

We saw that for high diffusion orders r and large stretching coefficients c, the spatial repartition of the ∇^* operator defined by (17) could deviate significantly from the target operator. However, this is due to the particular definition of ∇^* through the specification of the pseudo-stretched geometry. Here, as for the HDU diffusion above, we propose a definition of the HDF operator which matches the target one in the space-continuous context.
For this, a new latitude pseudo-stretched latitude is defined through:

$$d\Theta(\mu) = \mathrm{m}(\mu)^{(1/4)} d\theta^*(\mu) \tag{21}$$

As a consequence, the derivative operator ∇^* built with this new θ^* latitude writes:

$$\nabla^* = \frac{\partial}{\partial\theta^*} = \frac{\partial\Theta}{\partial\theta^*}\frac{\partial}{\partial\Theta} = \mathrm{m}^{(1/4)}\frac{\partial}{\partial\Theta} = \mathrm{m}^{(1/4)}\nabla' \tag{22}$$

The associated diffusion scheme then writes:

$$\frac{\partial X}{\partial t} = -K\nabla^{*4}X = -K\mathrm{m}\nabla'^4 X \tag{23}$$

which leads exactely the target operator for the diffusion scheme built on this new ∇^{*4}.

8 Spatial comparison of current HDU and HDF operators

We saw in the previous sections that HDU and HDF have similar profiles in their spatial repartition.
In this section we take the operational HDU with $c = 3.5$, and $r = 4$ as a reference and we examine how the spatial repartition of HDF (with $\epsilon = 1$) deviates from this reference. This deviation is depicted in Fig. 5, still assuming an absolute setting such as the two diffusions are equal at the pole of interest.
The HDF operator is seen to be close from HDU almost everywhere. The discrepancy reaches a maximum of 20% in the antipode hemisphere.

9 Initial error in the specification of the hybrid character of the HDF

Implementing the Fourier diffusion for the stretched model was done in two steps:
- write a pure (non-hybrid) diffusion operator
- modify the algorithm to allow an hybrid behaviour

However, a first error arised in the second part of the work: due to a misunderstanding between the conceptors of the HDF scheme, the modification toward an hybrid operator was done as if the starting point was a pure numerical one, whilst it was in fact a pure geographical one! As a consequence, the resulting hybrid diffusion was "over-geographic" instead of being hybrid: the diffusion was very strong near the pole of interest, and very weak at the antipode. Still assuming a correct diffusion at the pole of interest, the lack of diffusion at the antipode was by a factor: c^8 ≈22000.
Moreover, the scaling of the damping-time at the pole of interest was also erroneous, since it involved a equator-to-pole transformation assuming a stretching coefficient $c^{5/4}$ instead of $c^{1/4}$.
These two errors resulted in a HDF operator 110 times too strong near the pole of interest, and 205 times too weak near the antipode, which is fully consistent with the parallel suit results (better scores but more noise over orography in interest hemisphere).

10 Definition of the pertinent parameter for the diffusion in stretched models

As for non-stretched models, it is useful to know how must be specified the diffusion coefficient when the absolute resolution (i.e. the spectral truncation) and the stretching coefficient are changed. The goal of this exercise is to find what is the space-, resolution-, and stretching-independent "pertinent" parameter for defining the diffusion of the stretched model in such a way that this diffusion should always be locally equivalent to the diffusion of the set of locally coinciding non-stretched models.

10.1 space-independent pertinent parameter

We saw above that the space-independency of the the geographical diffusion coefficient for a given resolution and stretching was satisfied by simply taking a diffusion operator with the correct hybridicity.
Hence with the following definition of the stretched diffusion operator:

$$\frac{\partial X}{\partial t} = -k_1 \mathrm{m}^{1-r} \nabla^r X = -k_1 \mathrm{m} \nabla'^r X$$

the k_1 parameter is automatically space-independent (but not resolution- and stretching-independent).

10.1.1 What to call "resolution" for diffusion operator

Now we would like to define a pertinent "mean-resolution-independent" parameter for the diffusion. However, for spectral models as ARPEGE, the actual meaning of the word "mean-resolution" must be carefully defined. As long as the spectral-ratio NDLON/NSMAX is kept constant, the definition is straightforward: the "mean-resolution" which is pertinent for the horizontal diffusion could be indifferently defined by:

$$\frac{\texttt{NDLON}}{2\pi a} \quad \text{or by :} \quad \frac{\texttt{NSMAX}}{\pi a}$$

which is the inverse of the length of the smallest resolved feature at the location where $\mathrm{m} = 1$ (a is the earth radius). But when the spectral ratio can be modified (e.g. by changing the quadratic-grid to a linear-grid), then the definition cannot be set indifferently, because these two definitions cease to be always proportional. We will not address this problem here, because it is much wider than the specific problem of diffusion itself. After some experimentation, the following definition of the pertinent "mean-resolution" for the horizontal diffusion purpose has been adopted in ARPEGE:

$$\frac{\texttt{NDLON}}{2\pi a}(1 + 0.5\,\texttt{RNLGINC})^{-2.5} = \frac{\texttt{NSMAX}}{\pi a}(1 + 0.5\,\texttt{RNLGINC})^{-1.5}$$

where RNLGINC is 1 for a quadratic-grid and 0 for a linear-grid.
In all this section this corrective factor $(1 + 0.5\,\texttt{RNLGINC})$ for non-linear grids is omitted, for conciseness. However, it is correctly introduced in all the following sections, for correctness.

10.2 space- and resolution-independent pertinent parameter

We are now looking for a pertinent "mean-resolution-independent" parameter for defining the diffusion. This means that the diffusion operator specified with this parameter should be unchanged when the mean-resolution is changed while the stretching factor remains constant. Here we thus assume that the stretching is kept constant while the mean-resolution is changed, and we ignore the problem of changing the spectral-ratio of the grid, as explained above.
Let assume the following form for the stretched diffusion operator:

$$\frac{\partial X}{\partial t} = -k_2 f(\Delta x_0) \mathrm{m}^{1-r} \nabla^r X$$

where Δx_0^{-1} represents the mean-resolution (i.e. the resolution at the location where $\mathrm{m} = 1$), and f is an arbitrary function.
The resolution of the locally coinciding non-stretched model is $(\Delta x_0/\mathrm{m})^{-1}$ hence its diffusion operator writes:

$$\frac{\partial X}{\partial t} = -k_2 \left(\frac{\Delta x_0}{\mathrm{m}}\right)^{1-r} \nabla^r X$$

For the two operators to be identical, we must therefore have $f(\Delta x_0) = \Delta x_0^{r-1}$
Finally, the stretched-model diffusion must write:

$$\frac{\partial X}{\partial t} = -k_2 \Delta x_0^{r-1} \mathrm{m}^{1-r} \nabla^r X$$

the k_2 parameter is then space- and resolution-independent (but not stretching-independent).

10.3 space-, resolution-, and stretching-independent pertinent parameter

Here and we assume that the stretching is changed.
Let assume the following form for the stretched diffusion operator:

$$\frac{\partial X}{\partial t} = -k_3 \Delta x_0^{r-1} f(c) \mathrm{m}^{1-r} \nabla^r X$$

where f is an arbitrary function.
The resolution of the locally coinciding non-stretched model is $(\Delta x_0/\mathrm{m})^{-1}$ hence its diffusion operator writes:

$$\frac{\partial X}{\partial t} = -k_3 \left(\frac{\Delta x_0}{\mathrm{m}}\right)^{1-r} \nabla^r X$$

For the two operators to be identical, we must therefore have $f(c) = 1$
Finally, the stretched-model diffusion must write:

$$\frac{\partial X}{\partial t} = -k_3 \Delta x_0^{r-1} \mathrm{m}^{1-r} \nabla^r X \tag{24}$$

the k_3 parameter is then space-, resolution-, and stretching-independent. k_3 is thus the "pertinent" parameter which characterize the strength of the stretched model-diffusion in an absolute way. As for non-stretched models, the pertinent k_3 parameter has the dimension of a speed.
It should be noted that this definition is indifferently valid for grid-point and spectral diffusions (HDU and HDF). If we adopt a definition of the mean-resolution based on the spectral truncation N, as done in ARPEGE, then the final form of the diffusion operator must be:

$$\frac{\partial X}{\partial t} = -k \left(\frac{N}{a}\right)^{1-r} \mathrm{m}^{1-r} \nabla^r X = -k \left(\frac{N}{a}\right)^{1-r} \mathrm{m} \nabla'^r X \tag{25}$$

The NAMELIST parameter governing the diffusion of the stretched ARPEGE should thus be chosen in such a way that the dependency of the diffusion operator to the mean-resolution and stretching should be as in (25).

10.4 Comparison of the constraints linked to the global balance vs. the non-dependence of the pertinent parameter

Here we will establish a hierarchy between the two types of constraints that we have presented for defining a diffusion operator:

(i) The first type of constraint answers to the wish to have a correct global balance for the two poles, compared to the target operator [see e.g. (13) or (19)]

(ii) The second type of constraint answers to the wish to have pertinent tuning parameter for the diffusion fully-independent of the geometry of the grid [see e.g. (25)].

The constraints of type (i) are used only for operators which are approximated with respect to the target operator, and they are used as a way to somehow minimize the discrepancy between the approximated operator and the target operator. This involves the values of parameter β for the current HDU in (13), and the exponent ϵ for HDF in (19).
The constraint adopted in this paper for this type (i), is to enforce a correct balance for the two poles. This minimizes by construction the discrepancies to the target at the two poles. However, many other constraint could have been legitimately chosen to minimize the discrepancy to the target operator. For instance we could have chosen the values of β and ϵ which minimizes the mean discrepancy over the whole globe, etc...
All this is to say that the values of β and ϵ could legitimately be modified, hence the constraints of type (i) appear to be quite weak.

It is interesting to note that the new proposed HDU does not have such a constraint of type (i) since it uses the exact target operator without any approximation.

The constraints of type (ii) are linked to the possibility of having a "geometry"-independent tuning parameter for defining the strength of the diffusion in an absolute way. Of course, this constraint exists for both approximated and exact operators.

Obviously also, a diffusion will work properly even if this type (ii) of constraint is not satisfied. For instance neither HDU nor HDF fully satisfy these constraints, yet they work properly. If this type of constraint is not satisfied, the consequence is that each change in the "geometry" (truncation, stretching, and even order of diffusion) requires a new tuning of the diffusion parameter. Hence the diffusion parameter is not "pertinent" if constraints of type (ii) are not satisfied.
However, in opposition to the constraint of type (i) no counterpart advantage can be expected from a setting which does not satisfy the constraints of type (ii).
All this is to say that the constraints of type (ii) should definitely be considered as stronger than constraints of type (i), since they do not result from an aribtrary choice.

11 Comments on the time-discretization of diffusion schemes

In this section we examine how the time-continuous response of a diffusion scheme is modified by the time-discretization. The current fully-implicit time-discretization of ARPEGE and IFS is shown to result in a lack of selectivity compared to the time-continuous scheme for long time-steps as used in operational. This lack of selectivity makes obscure the conceptual effect of the diffusion, because its magnitude depends on many parameters.
A new time-discretization for the diffusion schemes is proposed in order to remove this weakness. This requires a spectral treatment in the basis of the eigenmodes of the desired diffusion operator ∇^{*r}. For non-stretched spectral models this is readily done in the normal spectral space because the eigenmodes of the ∇ operator are identical to the decomposition basis of the spectral transform (i.e. the spherical harmonics). For stretched models, the method requires to compute the eigenmodes of ∇^{*r} during the set-up process, and to project forth and back the state-variables at each time-step on this basis. The computing cost is slightly increased, but in counterpart the storage is decreased, when compared to the current method.

11.1 Current fully-implicit discretization

The current time-discretization of the diffusion in ARPEGE (and IFS) is an uncentered fully-implicit (UFI) scheme:

$$\frac{X^{+}-X^{0}}{\Delta t}=-K\nabla^{*r}X^{+}$$

where the subscript notation for time-levels is the standard ARPEGE one, and ∇^{*r} is the considered diffusion operator, not precised here. Let X be an eigenmode of ∇^{*2}, with the associated eigenvalue $-n^2$. The evolution for the mode due to the time-discretized diffusion is:

$$X(p\Delta t)=X(0)\left(\frac{1}{1+Kn^{r}\Delta t}\right)^{p}=X(0)\exp\left[-\frac{\ln(1+Kn^{r}\Delta t)}{\Delta t}.(p\Delta t)\right]$$

The time-discretized damping-time τ_{TD} for this mode is then given by:

$$\tau_{\mathrm{TD}}=\frac{\Delta t}{\ln(1+Kn^{r}\Delta t)}$$

The time-continuous damping-time τ_{TC} for this mode is:

$$\tau_{\mathrm{TC}}=\frac{1}{Kn^{r}}$$

Hence, this time-discretization scheme has a significant lack of spectral selectivity as soon as Δt is large enough to have:

$$\frac{\Delta t}{\tau_{s_T C}}=Kn_{\max}^{r}\Delta t>1 \tag{26}$$

where $n_{\max}$ is the maximum eigenvalue modulus of ∇^{*}.

It should be noticed that for a given model, this implicitness-parameter does not change dramatically when changing the resolution, provided the time-step is chosen proportional to the mesh (as most often) and provided the diffusion coefficient is chosen proportional to Δx^{r-1} as explain in section 2. In effect, under these conditions:

$$K n_{\max}^{r} \Delta t = (K_0 \Delta x^{r-1})(n_{max0}^{r} \Delta x^{-r})(\Delta t_0 \Delta x)$$

where quantities with subscript "0" refers to the initial configuration. As a consequence, the implitness-parameter is a characteristic of the accuracy of the time-treatment of the diffusion scheme which evoluates quite few for a given application when changing the resolution.
In ARPEGE this "implicitness parameter" $K n_{\max}^{r} \Delta t$ is much larger than one, near the pole of interest: in the current operational HDU scheme, for the divergence variable at the pole of interest, at the lowest level, the value of this "implicitness parameter" is $\approx c^4 \Delta t / \mathtt{HDURDIV} \approx 17.7$ (see details for this computation in section 16.2). For higher levels this parameter becomes even larger due to the enhancement of the diffusion with height. At 1 hPa, the value of the parameter reaches 17700.
As an illustration, the Fig. 9 shows a comparison of the spectral response of time-continuous and the UFI time-discretized diffusion operators based on ∇^4, for an implicitness parameter $K n_{\max}^{4} \Delta t = 17.7$. The time-discretization is responsible for a dramatic decrease in the strength of the diffusion for large wavenumbers. This can be interpreted physically as a lack of scale-selectivity.
For IFS current operational configuration, the implicitness parameter does not depend on the location or on the variable. At the lowest level its value is $\Delta t / \mathtt{HDIRDIV} = 1200/1500 = 0.8$, as can be seen from section 16.4. The Fig. 9 shows that the lack of scale-selectivity due to the UFI time-discretization is much smaller than for ARPEGE, and can be considered as acceptable.
For Aladin, the situation is intermediate. The value of the parameter is $\Delta t / \mathtt{HDIRDIV} = 415/155 = 2.7$, as can be seen from section 16.1.

In summary, we can say that due to the UFI time-discretization, large diffusion strength, and long operational time-steps, the ARPEGE diffusion does not possess the scale-selectivity of its time-continuous spatial operator, but a dramatically smaller one.
This contributes to making confused the effects of the diffusion in the stretched ARPEGE model. In other words, the UFI time-discretized diffusion of ARPEGE does not do what we believe it does, and moreover, it does not correspond to what a diffusion should do.
This is not necessarily very dangerous, because "compensative-tuning" is most of times able to compensate the weaknesses of an ill-formulated algorithm, but if a solution without this weakness is possible at slight or no cost, it is worth to be considered. Moreover, for diffusion purposes, one of our goals is to seek a formulation which does not require a lot of tuning at each change of configuration of the model, which is undoubtedly not the case with this UFI time-discretization in ARPEGE.

11.2 Proposal for a new time-discretization of spectral diffusions

The weakness of the UFI time-discretization for very long time-steps can be avoided if the eigenmodes of the time-continuous chosen diffusion operator can be explicitely accessed, and if the projection on this basis can be obtained at low cost during the integration of the model.
Let ∇^* be the desired diffusion operator, and let X be an eigenvector of ∇^{*r}, with the associated eigenvalues λ^r. The proposed time-discretization consists in writing, for each eigenmode:

$$X^{+} = X^{0} \exp(-K \lambda^{r} \Delta t) \tag{27}$$

or equivalently:

$$\frac{X^{+} - X^{0}}{\Delta t} = -X^{0} \frac{[1 - \exp(-K \lambda^{r} \Delta t)]}{\Delta t} \tag{28}$$

The sign of the RHS of (27) is always negative, as required for unconditional stablity.

For non-stretched spectral models, the diffusion operator is diagonal in the spectral space, hence the eigen-modes of the diffusion operator are simply the basis functions of the spectral decomposition, and this time-discretisation can be applied directly by replacing the implicit formulation by the exact formulation for each spectral coefficient.
For the stretched spectral ARPEGE, the diffusion operator ∇^{*r} is not diagonal in spectral space, as seen above. However, for a given zonal wavenumber m, the ∇^{*r} operator has no dependencies from the others zonal wavenumbers m'. The ∇^{*r} operator can thus be arranged as a block-diagonal operator, each block

involving a given m. For a given m the corresponding block is a square matrix of order $(N-m+1)$ describing the dependency between the spherical harmonics Y_n^m with the same m and total wavenumbers $n \in [m, N]$.
For the new-HDU, this dependency results in a tri-diagonal matrix because the multiplication by m of a spectral component (n, m) involves only the two neighbouring components $(n-1, m)$ and $(n+1, m)$, and for the current HDU scheme, the block matrices are penta-diagonal (see Yessad and Benard,1995 for details). For each m, the block matrix ∇_m^{*r} of the ∇^{*r} operator can be diagonalized in the spectral space:

$$\nabla_m^{*r} = \mathrm{A}_m.\Delta_m.\mathrm{A}_m^{-1}$$

We will now work in the space of eigenmodes of ∇_m^{*r} for a given m. For one mode X_n^m, with the associated eigenvalue $(\lambda_n^m)^r$, we have:

$$\nabla_m^{*r}.X_n^m = (\lambda_n^m)^r \, X_n^m$$

Hence we write for each mode, according to (27):

$$X_n^{m+} = X_n^{m0} \exp\left[-K \, (\lambda_n^m)^r \, \Delta t\right] \tag{29}$$

This writes, matrically:

$$\mathrm{HDU}_m = \mathrm{A}_m.\exp\left[-K\Delta t \Delta_m\right].\mathrm{A}_m^{-1} \tag{30}$$

where HDU_m is the matrical operator of the new-HDU diffusion for the zonal wave-number m in spectral space, and $\exp\left[-K\Delta t\Delta_m\right]$ follows the mathematical definition of the matrix exponentition.
The value of this diagonal exponential matrix can be pre-computed and stored during the set-up of the model (it must be stored for each value of the K coefficient, i.e. for each variable and each level if the strength of the diffusion is made level-dependent). Similarly, the eigenmode matrices A_m and their inverse can be pre-computed and stored during the set-up. Here, these matrices do not depend on K, hence they need to be stored only once whatever is the value of the K coefficient. This method has the advantage of allowing an exact scale-selectivity whatever is the length of the time-step and the local strength of the diffusion.
In terms of computational cost, this method should be comparable to the current one. In the current method, two inversions of tri-diagonal matrices and then the product of these two full matrices are performed. In the new method, the product of two full matrices and a diagonal matrix is performed.
In terms of storage, the new method is cheaper, since only one diagonal matrix is stored for each m, each level, and each variable. In the current method, two tri-diagonal matrices are stored.

11.3 Extension to HDF diffusion

The previous method can be extended to the HDF diffusion.
In Fourier space, for a given m, we diagonalize the spatial operator ∇_m^{*4} of the HDF diffusion, similarly to above:

$$\nabla_m^{*4} = \mathrm{A}_m.\Delta_m.\mathrm{A}_m^{-1}$$

Then for a given value of $K\Delta t$, we write the time-discretization of the diffusion in the following form:

$$\mathrm{HDF}_m = \mathrm{A}_m.\exp\left[-K\Delta t \Delta_m\right].\mathrm{A}_m^{-1} \tag{31}$$

where HDF_m is the matrical operator of the HDF diffusion for the zonal wave-number m in Fourier space, and $\exp\left[-K\Delta t\Delta_m\right]$ is equal to the the original diagonal matrix Δ_m, but now with exponentiated elements, following the mathematical definition.

Now, the matrix of the HDF diffusion, taken as a whole in spectral space, is built for the zonal wave-number m. This writes:

$$\mathrm{HDFS}_m = \mathrm{LD}_m.\mathrm{HDF}_m.\mathrm{LI}_m \tag{32}$$

where LD_m is the direct Legendre transform, LI_m the inverse Legendre transform, and HDF_m is the matrix of the HDF diffusion in the Fourier space for this zonal wave-number m. It should be noted that neither LD_m nor LI_m are square matrices. However we have:

$$
\begin{aligned}
\mathrm{LD}_m.\mathrm{LI}_m &= I_{(N-m+1)} \\
\mathrm{LI}_m.\mathrm{LD}_m &\approx I_{\mathtt{NDGL}}
\end{aligned}
$$

where I_p is the p-order identity matrix. For a given m, $\mathrm{LI}_m.\mathrm{LD}_m$ is the projection on the space $Y_m^m, ..., Y_N^m$ in Fourier space, and represents the zonal impact of the spectral-fit for this particular zonal-wavenumber m.

Equation (32) was valid for the chosen value of $K\Delta t$. We will now examine how it transforms when doubling this value. In this case, the new matrical operator $\mathrm{HDFS'}_m$ writes:

$$
\begin{aligned}
\mathrm{HDFS'}_m &= \mathrm{LD}_m.A_m.\exp\left[-2K\Delta t\Delta_m\right].A_m^{-1}.\mathrm{LI}_m \\
&\approx \mathrm{LD}_m.A_m.\exp\left[-K\Delta t\Delta_m\right].A_m^{-1}.\mathrm{LI}_m.\mathrm{LD}_m.A_m.\exp\left[-K\Delta t\Delta_m\right].A_m^{-1}.\mathrm{LI}_m \\
&= \mathrm{HDFS}_m^2
\end{aligned}
$$

The impact of replacing $\mathrm{HDFS'}_m$ by HDFS_m^2 can be seen by writing:

$$
\begin{aligned}
\mathrm{HDFS}_m^2 &= \mathrm{LD}_m.\mathrm{HDF}_m.\mathrm{LI}_m.\mathrm{LD}_m.\mathrm{HDF}_m.\mathrm{LI}_m \\
&= \mathrm{LD}_m.\mathrm{HDF}_m.\mathrm{LI}_m.\mathrm{LD}_m.\mathrm{HDF}_m.\mathrm{LI}_m.\mathrm{LD}_m.\mathrm{LI}_m \\
&= \mathrm{LD}_m.\left(\mathrm{HDF}_m.\mathrm{LI}_m.\mathrm{LD}_m\right)^2.\mathrm{LI}_m
\end{aligned}
$$

Hence we see that doing the above approximation is equivalent to replacing the exact form:

$$
\mathrm{HDFS'}_m = \mathrm{LD}_m.\left(\mathrm{HDF}_m\right)^2.\mathrm{LI}_m
$$

by the approximated form:

$$
\mathrm{HDFS'}_m \approx \mathrm{LD}_m.\left(\mathrm{HDF}_m.\mathrm{LI}_m.\mathrm{LD}_m\right)^2.\mathrm{LI}_m
$$

Physically, the approximation consists in iterating not only the diffusion, but also the spectral fit operator. This approximation can be considered as acceptable since the only source of unfitted components in this process is the HDF diffusion itself, and it can reasonably be considered as a very small source of unfitted components.

The reasoning is now extended to any real positive exponent, which is necessary in order to be able to apply the diffusion for all possible values of the parameter $K\Delta t$ occuring in the model. This extension is made by extending the previous reasoning to numbers of the form $(1/n)$, then to all rational positive numbers, then, by density arguments, to all real positive numbers. This process is not detailed here, and we will assume it to be entirely valid for our particular case (it may be difficult to prove some identities for exponentiation of non-square matrices).
Finally, for a new parameter $K'\Delta t$ satisfying $K'\Delta t = \alpha K\Delta t$, the HDFS' operator will be written:

$$
\mathrm{HDFS'}_m = \left(\mathrm{HDFS}_m\right)^\alpha
$$

This equation will be used in practice under the diagonalized from:

$$
\mathrm{HDFS'}_m = \mathrm{BD}_m.\left(\exp\left[-K\Delta t\Delta_m\right]\right)^\alpha.\mathrm{BI}_m
$$

where $\mathrm{BD}_m = \mathrm{LD}_m.\mathrm{A}_m$ and $\mathrm{BI}_m = \mathrm{A}_m^{-1}.\mathrm{LI}_m$ are the two non-square matrices for passage from spectral space to Fourier-eigenmode space, and back. The size of these matrices is $\mathtt{NDGL} * (N-m+1)$. It should be noted that BI_m is not the "inverse" of BD_m (anyway, the notion of inverse is not straightforward for non-square matrices).
In terms of computational cost, this method is roughly equivalent to its counterpart for spectral new-HDU diffusion. The storage is increased because the passage matrices BD and BI are bigger than A and A^{-1} of the previous subsection.

12 Choice of the diffusion parameter(s) in the current HDU

The namelist parameter for scaling the reference level of diffusion in the current HDU is the pair (HDIRDIV, HDURDIV). The scheme is then written as:

$$\frac{\partial X}{\partial t} = -\left[\frac{a^2}{N(N+1)}\right]^{r/2}\left(\frac{1}{\mathtt{HDIRDIV}}\nabla'^r X + \frac{1}{\mathtt{HDURDIV}}\widehat{\mathrm{m}^r}\nabla'^r X\right) \tag{33}$$

where N is the truncation order and a is the earth radius. This specification calls for three remarks:
First, the possibility of having two NAMELIST parameters (HDIRDIV and HDUDIV) to specify the proportion of numerical vs. geographical diffusion violates the constraint imposed by (15) stating that the proportion is imposed by the stretching factor c through β, in order to respect the correct balance at the two poles.
Second, the definition of the NAMELIST parameters (HDIRDIV and HDUDIV) are not in agreement with the constraint imposed by (25). The consequence is that when changing the resolution (truncation) of the model, the NAMELIST parameter must be changed accordingly. In other words, it is not possible to run ARPEGE with various resolution while keeping an empty diffusion NAMELIST.
Third, the choices made for the specification of the HDU fails to make it stretching-independent, for all orders of diffusion r. Moreover the global balance is correct only for $r = 4$. Having an HDU operator independent of c and with a correct global balance for all r would require:

$$\alpha_n = \frac{c^{r-1} - (1/c)^{r-1}}{c^r - (1/c)^r} \approx \frac{1}{c}$$

and:

$$\alpha_g = \frac{c - (1/c)}{c^r - (1/c)^r} \approx \frac{1}{c^{r-1}}$$

in (12), since in this case, we would have, at the two poles:

$$\left(\frac{1}{c} + \frac{\widehat{\mathrm{m}^r}(\mu = \pm 1)}{c^{r-1}}\right)\nabla'^r = \mathrm{m}(\mu = \pm 1)\nabla'^r$$

as for the target operator. Instead of this, $\alpha_n = c$ and $\alpha_g = (1/c)$ are imposed for all values of r. This makes the diffusion stretching-dependent and moreover, the proportion of numerical vs. geographical diffusions is not correct as soon as $r \neq 4$, resulting in an incorrect global balance of HDU diffusion in this case.
Fourth, when changing the truncation/grid-ratio (e.g. linear or quadratic) the value of the (HDIRDIV and HDUDIV) NAMELIST parameters has to be changed.

Summary: The choice of the diffusion parameter(s) in HDU is not "pertinent". Moreover the global balance and the "geometry" independency with this setting is not always insured.

13 Choice of the diffusion parameter in new HDU

We proposed above to formulate HDU in a simpler but more correct way. In order to insure the non-dependency of the diffusion operator, the formulation should be done in agreement with (25) under the form:

$$\frac{\partial X}{\partial t} = -\frac{k}{\pi}\left(\frac{N}{a}\right)(1 + 0.5\,\mathtt{RNLGINC})^{-1.5}\left[\frac{a^2}{N(N+1)}\right]^{r/2}\mathrm{m}\nabla'^r X \tag{34}$$

The π factor is inroduced in order to give the same meaning of k than in HDF (see below). We propose, in consistency with HDF setting, to call this pertinent parameter k under the name RRDXTAU.
We have $\Delta x_0 = (1 + 0.5\,\mathtt{RNLGINC})^{-1} a\pi/N$. Hence, the diffusion operator writes:

$$\frac{\partial X}{\partial t} = -\frac{\mathtt{RRDXTAU}}{\Delta x_0}(1 + 0.5\,\mathtt{RNLGINC})^{-2.5}\left[\frac{a^2}{N(N+1)}\right]^{r/2}\mathrm{m}\nabla'^r X \tag{35}$$

Under the above form, the physical meaning of RRDXTAU appears clearly:

In new-HDU, $\texttt{RRDXTAU}^{-1}$ **is the ratio of the time-continuous local damping-time of the shortest spectral wave divided by the local linear-grid mesh-size.**

This physical definition calls for some comments:

- Strictly looking to (35) one could think that the ratio must be taken at the location where Δx is equal to Δx_0, i.e. at the location where $\mathrm{m} = 1$. However, due to the fact that the operator new-HDU is equal to the target one, this ratio is in fact independent of the location due to the factor m in the operator. In other words, the damping-time of the shortest spectral wave is *exactely* proportional to the resolution. This is why the above definition can be understood anywhere on the sphere.
- For the current UFI time-discretziation, the practical damping-time can be dramatically larger than the time-continuous one for long time-steps, as discussed above. For an exact time-discretization this dependency to the time-step lentgh disappears, hence the above definition can be understood as for the practical time-discretized damping-time as well.
- The ratio in the above definition must finally be understood as the one for a linear grid. For a non-linear (e.g. quadratic) grid, the factor linked to $(1 + 0.5\,\texttt{RNLGINC})$ modifies the damping-time.

The new HDU diffusion should then finally be written as:

$$\frac{\partial X}{\partial t} = -\frac{\texttt{RRDXTAU}}{\pi}\frac{N}{a}(1 + 0.5\,\texttt{RNLGINC})^{-1.5}\left[\frac{a^2}{N(N+1)}\right]^{r/2} \mathrm{m}\nabla'^{r} X \tag{36}$$

14 Choice of the diffusion parameter in HDF

The diffusion parameter for the current implementation of HDF is called RRDXTAU.
The current implementation of HDF in ARPEGE writes:

$$\frac{\partial X}{\partial t} = -\frac{1}{c^{*4}}\left(\frac{\pi}{2}\right)^{\rho}\frac{\texttt{RRDXTAU}}{a\Delta\theta_{\min}}(1 + 0.5\,\texttt{RNLGINC})^{-2.5}\left[\frac{a^2}{N(N+1)}\right]^2 \widetilde{\nabla^{*4}} X$$

where $(\pi/2)^{\rho}$ is the corrective factor to compensate the fact that the response of the grid-point operator $\widetilde{\nabla^{*2}}$ is weaker than the one of the space-continuous $m^{*2}\nabla'^2$ for short waves, as discussed above.
Since $\Delta\theta_{\min} = (1 + 0.5\,\texttt{RNLGINC})^{-1}\pi/(Nc)$, this rewrites:

$$\frac{\partial X}{\partial t} = -\frac{c}{c^{*4}}\frac{\texttt{RRDXTAU}}{\pi}\frac{N}{a}\left(\frac{\pi}{2}\right)^{\rho}(1 + 0.5\,\texttt{RNLGINC})^{-1.5}\left[\frac{a^2}{N(N+1)}\right]^2 \widetilde{\nabla^{*4}} X$$

We see that this is almost in agreement with (25), except by the factor c^{*4}/c in the beginning. The effect is to have an incorrect stretching dependency as soon as $c^{*4} \neq c$ (i.e. $\texttt{REXPKDX} \neq 1$). It would be better to replace the c^{*4} by c, since the current setting violates a constraint of type (ii). Furthermore, correcting this error would allow to replace the complicated factor $1/(c^{*4}a\Delta\theta_{\min})$ by the simpler factor $(1 + 0.5\,\texttt{RNLGINC})N/(\pi a)$ exactely as done in (34) for the new proposed HDU. Hence this would allow a much more deeper harmonization of the settings for these two schemes.
Since $\Delta x_0 = (1 + 0.5\,\texttt{RNLGINC})^{-1}\pi a/N$, the HDF operator rewrites (with corrected (c^{*4}/c) factor):

$$\frac{\partial X}{\partial t} = -\frac{\texttt{RRDXTAU}}{\Delta x_0}(1 + 0.5\,\texttt{RNLGINC})^{-2.5}\left[\frac{a^2}{N(N+1)}\right]^2 \widetilde{\nabla^{*4}} X \tag{37}$$

The exact physical meaning of RRDXTAU appears clearly from (37):

In HDF, $\texttt{RRDXTAU}^{-1}$ **is the ratio of the time-continuous local damping-time of the shortest spectral wave divided by the local linear-grid mesh-size, this ratio being computed at one of the two poles, and assuming a perfect spectral response for the space-discretized $(\pi/2)^{\rho}\widetilde{\nabla^{*4}}$ operator compared to the space-continuous ∇^{*4} operator.**

This physical definition calls for some comments:

- Due to the various imperfections of HDF, the physical meaning of RRDXTAU is much less clear than for the new-HDU.

- In opposition to what occured in new-HDU, the HDF operator $\widetilde{\nabla^{*4}}$ spatially deviates from the target operator $\mathrm{m}\nabla'^4$, especially for high diffusion orders and high stretching coefficients. Hence, the above definition of $\mathtt{RRDXTAU}^{-1}$ can be defined only at the places where the space-continuous operator ∇^{*4} coincides with the target operator $\mathrm{m}\nabla'^4$. These two operators coincide at the two poles, because at the poles $\mathrm{m}^*(\pm 1)$ is exactely equal to $\mathrm{m}^{1/4}(\pm)$ [see e.g. (16)].

- In opposition to what occured in new-HDU, the HDF operator $\widetilde{\nabla^{*4}}$ spectrally deviates from the target operator $\mathrm{m}\nabla'^4$, especially for shortest waves, even with the introduced correction factor $(\pi/2)^\rho$. Since the above definition of $\mathtt{RRDXTAU}^{-1}$ specifically involves the shortest spectral wave, it can be applied only by pretending that the spectral response of $(\pi/2)^\rho\widetilde{\nabla^{*4}}$ is correct compared to its space-continuous counterpart ∇^{*4}.

- As for new-HDU the damping-time in the above definition must be understood as the time-continuous one if the current UFI time-discretisation is used, and can be understood as the practical time-discretized one as well if the exact time-discretization is used.

- As for new-HDU, the ratio in the above definition must finally be understood as the one for a linear grid. For a non-linear (e.g. quadratic) grid, the factor $(1 + 0.5\,\mathtt{RNLGINC})$ modifies the damping-time.

The physical meaning of RRDXTAU appears to be more unclear than the one of new-HDU, due to the various unperfections of HDF concerning the used operator and the space-discretization. However, if these imperfections are neglected, the meaning becomes the same as for the new-HDU.

The HDF diffusion should then finally be written as:

$$\frac{\partial X}{\partial t} = -\frac{\mathtt{RRDXTAU}}{\pi}\,\frac{N}{a}(1 + 0.5\,\mathtt{RNLGINC})^{-1.5}\left[\frac{a^2}{N(N+1)}\right]^{r/2}\widetilde{\nabla^{*4}}X \tag{38}$$

This expression is directly comparable to the proposed new-HDU one (36).

Summary: With a very slight modification in the formulation of HDF settings, the RRDXTAU parameter would become a pertinent parameter for this diffusion, with a physical meaning approaching the one of its counterpart in new-HDU.

15 Choice of the diffusion parameter in proposed new-HDF

The formulation of the HDF would remains as in (38) except that $\widetilde{\nabla^{*4}}$ now represents the grid-point approximation of the target operator $\mathrm{m}\nabla'^4$. The RRDXTAU parameter is still pertinent, but have a simpler meaning than in the previous section because the operator ∇^{*4} is now correct:
In the new proposed HDF, $\mathtt{RRDXTAU}^{-1}$ **is the ratio of the time-continuous local damping-time of the shortest spectral wave divided by the local linear-grid mesh-size, assuming a perfect spectral response for the space-discretized** $(\pi/2)^\rho\widetilde{\mathrm{m}\nabla'^4}$ **operator compared to the space-continuous** $\mathrm{m}\nabla'^4$ **operator.**

16 Comparison of operational diffusion absolute intensities

Here we compare the absolute intensity of ALADIN, HDU, HDF, and IFS diffusion operators. For this we examine the damping time of a given wave of wavelentgh $\lambda = 38000$ m (this is the shortest wave of current operational stretched ARPEGE at the pole of interest). These comparisons are made with the linear-grid settings for all models (hence the factor $(1 + 0.5\,\mathtt{RNLGINC})$ is always set to 1).

16.1 Damping-time in operational ALADIN

The diffusion equation writes:

$$\frac{\partial X}{\partial t} = -\left(\frac{n^2}{N^2}\right)^{r/2} \frac{1}{\mathtt{HDIRDIV}} X \tag{39}$$

we have in operational:

- $r = 4$
- $\Delta t = 415.$
- $N = 149$ is the truncation
- $\mathtt{HDIRDIV} = 155$

The spectral coefficient n is given by:

- $n = L_x/\lambda = 75 \quad \Rightarrow \quad (N/n) = 2$

reflecting the fact that the considered wave is not the shortest one for Aladin, but the $4\Delta x$ one.
Hence the damping-time is:

- $\tau = \mathtt{HDIRDIV} * (N/n)^r = 155 * 16 = 2480.$
- $\Delta t/\tau_s = 415/155 = 2.7$

The effective damping-time for this wave is given by:

- $\tau_{\mathrm{TD}} = \Delta t/\log(1 + \Delta t/\tau_s) * (N/n)^r = 5072.$

However, since the resolutions of operational ALADIN and operational ARPEGE are not equal, the damping times cannot be directly compared. Hence we will determine the damping time of this wave for another Aladin model with the local resolution of ARPEGE at the pole of interest ($\Delta x = 19000m$) and called "reference" ALADIN hereafter

16.2 Damping-time in operational HDU

The diffusion equation writes:

$$\frac{\partial X}{\partial t} = -\left[\frac{n(n+1)}{N(N+1)}\right]^{r/2} \left(\frac{1}{\mathtt{HDIRDIV}} + \frac{\widehat{m^r}}{\mathtt{HDURDIV}}\right) X \tag{40}$$

we have in operational:

- $r = 4$
- $\Delta t = 831.$
- $N = 298$
- $\mathtt{HDIRDIV} = 620$
- $\mathtt{HDURDIV} = 7595$

The spectral coefficient n is given by:

- $n = 2\pi a/\lambda c = 298 \quad \Rightarrow \quad (N/n) = 1$

reflecting the fact that the wave is the shortest one.
Hence the damping-time is:

- $\tau = \tau_s = \left[\frac{N(N+1)}{n(n+1)}\right]^{r/2} \left(\frac{1}{\mathtt{HDIRDIV}} + \frac{\widehat{m^r}}{\mathtt{HDURDIV}}\right)^{-1} \approx \mathtt{HDURDIV}/c^4 = 7595/150 = 46.8$
- $\Delta t/\tau_s = 831/46.8 = 17.7$

The effective damping-time for this wave is given by:

- $\tau_{\mathrm{TD}} = \Delta t/\log(1 + \Delta t/\tau_s) = 284.$

16.3 Damping-time in "reference" ALADIN

The diffusion equation writes:

$$\frac{\partial X}{\partial t} = -\left(\frac{n^2}{N^2}\right)^{r/2} \frac{1}{\mathtt{HDIRDIV}} X \tag{41}$$

According to section 2, since the resolution of the "reference" Aladin is twice poorer than the operational ALADIN, the $K = N^r * \mathtt{HDIRDIV}$ coefficient must be multiplied by 2^{r-1}, hence the HDIRDIV parameter would have to be set to:

$\bullet$HDIRDIV $= 155 * 2 = 310$ for having a consistent diffusion between operational and reference Aladin models.
we have for the reference Aladin:

$\bullet r = 4$

$\bullet \Delta t$ would be 831.

$\bullet (n/N) = 1$

reflecting the fact that now the wave is the shortest one.
Hence the damping-time is:

$\bullet \tau = \tau_s = \mathtt{HDIRDIV} * (N/n)^r = 310.$

Thus reflecting the fact that the τ_s must be proportional to the resolution (τ_s for operational Aladin was 155.).

$\bullet \Delta t/\tau_s = 831/310 = 2.7$, as above.

The effective damping-time for this wave is given by:

$\bullet \tau_{\mathrm{TD}} = \Delta t/\log(1 + \Delta t/\tau_s) = 635.$

16.4 Damping-time in the "reference" IFS

As for Aladin, the resolution of the operational IFS model is not the same as the operational ARPEGE. Hence, an imaginary "reference" IFS model with the same resolution as ARPEGE at the pole of interest has to be defined in order to allow an absolute comparison of the diffusions strengths between these models.
The diffusion of the IFS model writes:

$$\frac{\partial X}{\partial t} = -\left[\frac{n(n+1)}{N(N+1)}\right]^{r/2}\left(\frac{1}{\mathtt{HDIRDIV}}\right)X \tag{42}$$

According to section 2, since the resolution of the operational IFS model is poorer than the operational ARPEGE, the HDIRDIV parameter for the target resolution has to be found. The truncation for a non-stretched model with same local resolution as ARPEGE one over the pole of interest is NSMAX = 298*3.5=1043. The current operational IFS truncation is 511, hence the current HDIRDIV=1500 parameter has to be extrapolated for this truncation. We assume here that IFS model respects the constraint of proportionality to the mesh applicable to a non-stretched model: HDIRDIV= (511/1043)*1500=734. Hence:

$\bullet r = 4$

$\bullet \Delta t$ would be 1200.*(511/1043) = 588.

$\bullet (n/N) = 1$

Hence the damping-time is:

$\bullet \tau = \tau_s = \mathtt{HDIRDIV} = 734.$

The diffusion in IFS model is weaker than ARPEGE one in the absolute.

$\bullet \Delta t/\tau_s = 588/734 = 0.8$

The effective damping-time for this wave is given by:

$\bullet \tau_{\mathrm{TD}} = \Delta t/\log(1 + \Delta t/\tau_s) = 1000.$

16.5 Damping-time in HDF

We have seen that the spatial profile of HDU and HDF profiles were significantly different. This prevent a clean comparison of the absolute strengths of these two operators by just examining the damping-time at the pole of interest.
The HDF diffusion writes:

$$\frac{\partial X}{\partial t} = -\frac{1}{c^{*4}}\left(\frac{\pi}{2}\right)^{\rho}\frac{\mathtt{RRDXTAU}}{a\Delta\theta_{\min}}\left[\frac{a^2\widetilde{\nabla^{*2}}}{N(N+1)}\right]^2 X$$

The attenuated spectral response is ignored here since we seek to compare the intensity of diffusions in a global way. We will thus assume that the $\pi/2$ factor exactely compensate the response factor of $\widetilde{\nabla^{*4}}$. Hence, we have:

$$\frac{\pi}{2}\left[\frac{a^2\widetilde{\nabla^{*2}}}{N(N+1)}\right]^2 = \left[\frac{a^2\mathrm{m}^{*2}\nabla'^2}{N(N+1)}\right]^2 = \left[\frac{\mathrm{m}^{*2}n(n+1)}{N(N+1)}\right]^2 = \mathrm{m}^{*4} = m = c = 3.5$$

we have :

•$\Delta t = 831.$

•$\mathtt{RRDXTAU}$ = 61.45

•$a\Delta\theta_{\rm min} = \Delta x_{\rm min}$ = 19000

At the pole of interest we have:

•$c^{*4} = c$

Hence:

•$\tau =_t au_s = \Delta x_{\rm min}/\mathtt{RRDXTAU}$ = 310.

•$\Delta t/\tau_s = 831/310 = 2.7$

The effective damping-time for this wave is given by:

•$\tau_{\rm TD} = \Delta t/\log(1 + \Delta t/\tau_s)$ = 635.

16.6 Spectral comparison of operational HDU vs. reference IFS

Since the scale-selectivity of the current HDU and IFS diffusions are not the same, the unique digit given above is not enough to have an idea of the comparative strength of the two diffusions. Here we depict in Fig. the spectral response at the pole of interest.

The intensity of the two diffusions as a function of the adimensionalized wave-number is depicted in Fig. 10 and 11 , for the divergence and the other variables. The intensity is scaled by 1 for the shortest wave in HDU. The HDU diffusion is always stronger than IFS one. For the divergence, the ratio is around 17 for long waves to 4 for smallest waves. For other variables, the ratio is around 2–3.

16.7 summary

The damping times for the considered wave are summarized in the following table for ALADIN, IFS, and ARPEGE HDU and HDF, all models having been put to the same resolution at the pole of interest.

	$\tau_{\rm TC}$ (DIV)	$\tau_{\rm TC}$ (VOR)	$\tau_{\rm TD}$ (DIV)	$\tau_{\rm TD}$ (VOR)
reference ALADIN	310	1550	635	1924
reference IFS	734	734	1000	1000
operational HDU	46.8	234	284	549
current HDF	310	1550	635	1924

•The HDU diffusion is intrinsically 2–4 times stronger than the Aladin one over France

•The IFS diffusion is 4 to 17 times weaker than HDU for DIV and 2 to 3 times weaker for the VOR, depending on the wave-number.

•These diffusions cannot easily be compared with only one number since they have quite different spatial profiles and spectral response.

17 Conclusion

17.1 Recommendations

In order to rationalize the stretched ARPEGE diffusions, we recommend:

- Formulate HDU in terms of the "new-HDU" operator proposed above.

- Remove the current HDU from the ARPEGE code.

- Make the specification of settings for new-HDU in agreement with this document: only one pertinent parameter, with a correct "geometry" dependency.

- Make the specification of settings for HDF in agreement with this document: this require to correct the stretching dependency.

- Formulate multiplicative coefficients of RRDXTAU in term of the simple $N/(a\pi)$ factor instead of the complicated $\Delta\theta_{\min}$ or Δx_0, for new-HDU and new HDF.

- Formulate the time scheme of new-HDU as an exact one, and then retune the diffusion parameter to a less diffusive value, dur to better spectral response.

- In future, after careful evaluation, remove HDF from ARPEGE code and use the proposed new-HDU (see next sub-section).

17.2 Comparison of new-HDU vs. HDF advantages

It should be noted that HDF is more general than the HDU diffusion: HDU diffusion assumes by nature that the stretched geometry is obtained through the Schmidt transformation, while HDF could be used with any geometric transformation. This can be obtained by simply modifying locally the metrics in the denominator of the grid-point operator in order to introduce the correct hybrid space-dependency. Hence the non-stretched metrics Δx in the denominator has to be simply replaced by $\mathrm{m}^(1/r)\Delta x$.
HDF has no advantage compared to new-HDU for ARPEGE:

- HDF has a lack of selectivity

- HDF is an approximation of the target operator

- HDF is complicated in terms of code, namely it interfers badly with the transform package.

- Passing to a diagonalized formulation is much more heavy than for new-HDU in terms of code.

new-HDU has no disadvantage:

- new-HDU code is very simple

- new-HDU operator is the exact target operator

- new-HDU has negligible "aliasing" potential problem. The possible source of aliasing is twice smaller than the LSIDG algorithm, which is used operationally.

- Passing to a diagonalized formulation is much more easy than for HDF in terms of code.

- If the target operator was to be changed after new experimental results with non-stretched models, the modification would be easy in the current framework.

This is why we recommend to remove HDF and current-HDU from the ARPEGE code, and to keep only the proposed new-HDU code.

17.3 Absolute intensity of ARPEGE diffusion

The general level of HDU diffusion used in the curent operational ARPEGE is too high, compared to IFS and Aladin diffusions. Further work should be undertaken in order to be able to reduce this general level without degradation of the results, which is not possible currently. We think that this should be one of the next priorities of the group to answer the question: "why do we need so much diffusion in ARPEGE?".

18 Figures

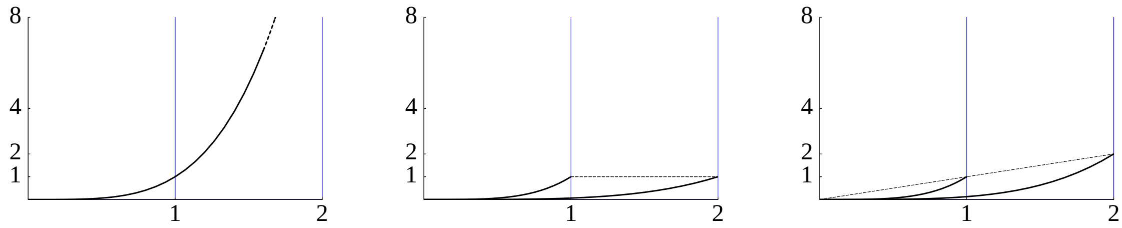

Figure 1: Three possible extensions of diffusion when doubling the resolution: strength of diffusion as a function of the wave-number. From left to right: solution (i), (ii), (iii).

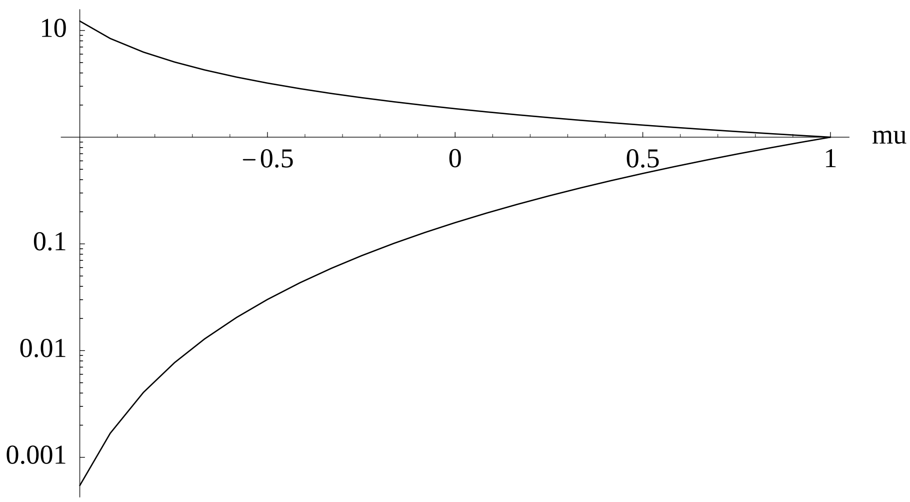

Figure 2: Latitudinal profile of diffusion strength scaled at the pole of interest by the reference for HDU. Top: numerical operator, bottom: geographical operator.

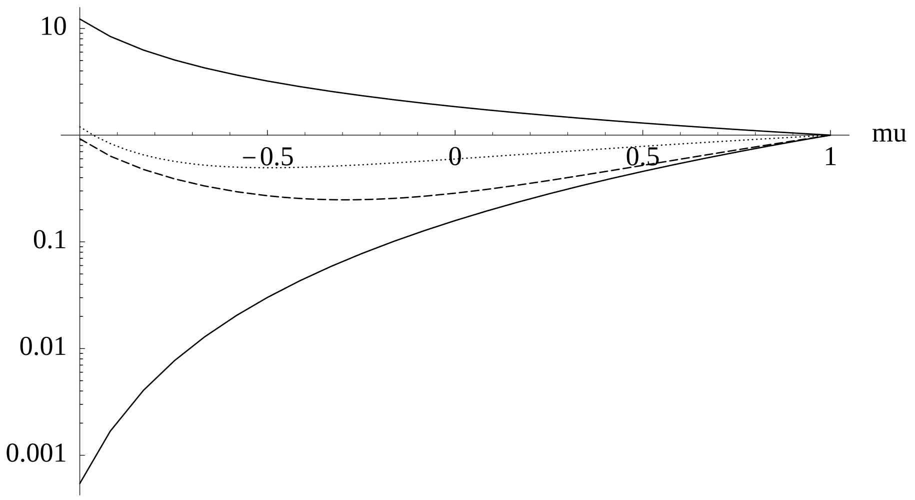

Figure 3: Latitudinal profile of diffusion strength scaled at the pole of interest by the reference for HDU. Dashed: "analytical" HDU, dotted "practical" HDU, solid: numerical and geographical operators.

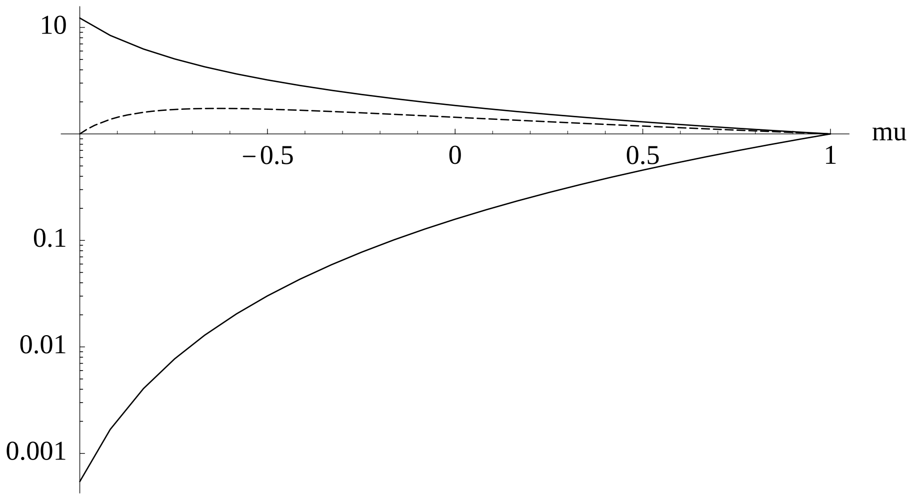

Figure 4: Latitudinal profile of diffusion strength scaled at the pole of interest by the reference for HDF. Dashed: HDF operator, solid: numerical and geographical operators

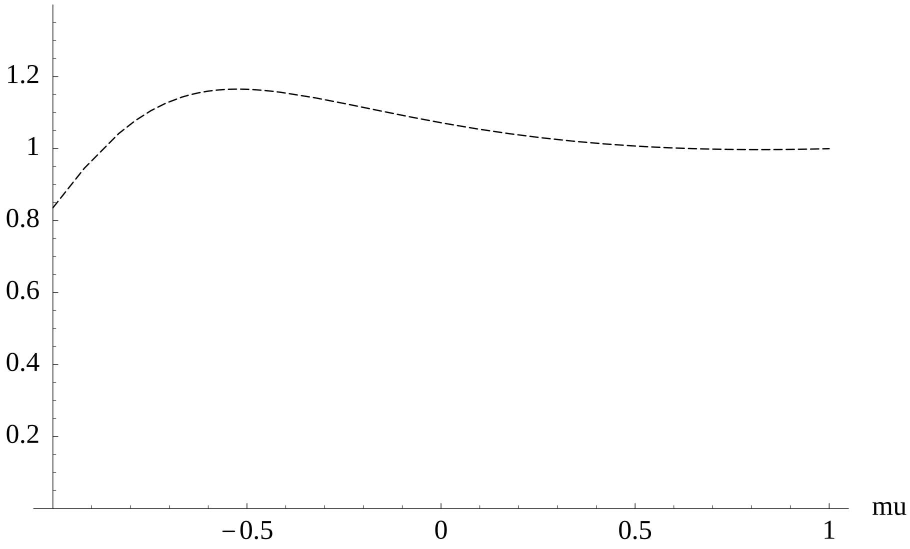

Figure 5: Latitudinal profile of diffusion strength of HDF scaled by the strength of HDU for $c = 3.5$ and $r = 4$, assuming an equal strength at the pole of interest.

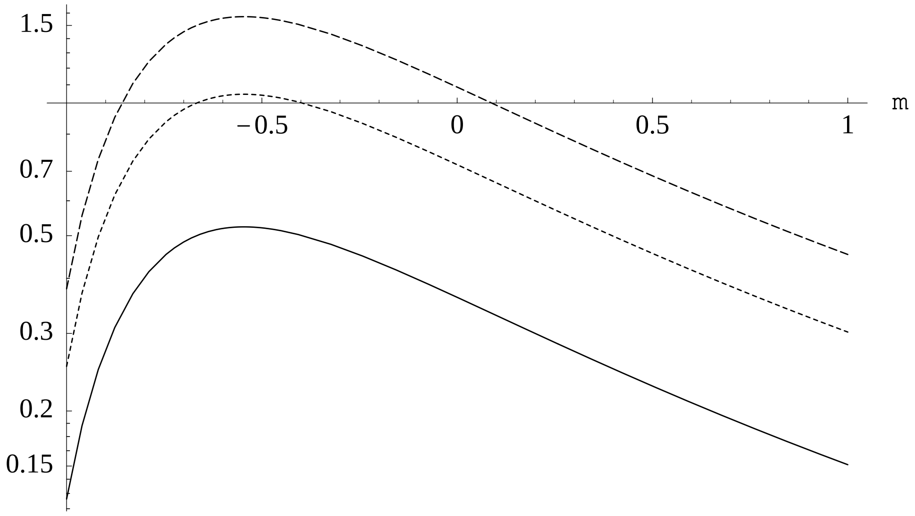

Figure 6: Latitudinal profile of diffusion strength of HDF scaled by the strength of operational HDU for $c = 3.5$ and $r = 4$. Solid: current setting `RRDXTAU` $= 61.45$; short-dash: with HDF strength multiplied by 2; long-dash: with HDF strength multiplied by 3.

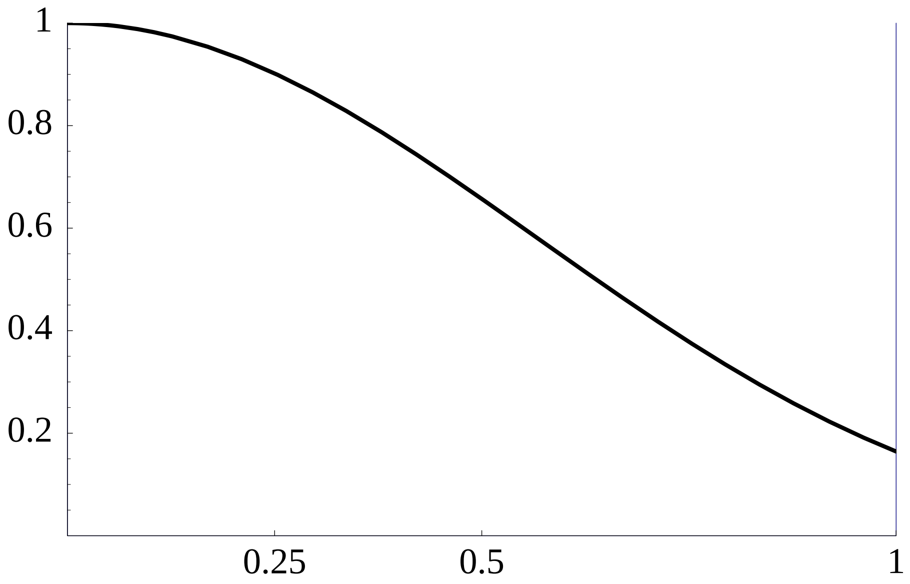

Figure 7: Spectral response factor of the grid-point discretized operator ∇^4 (the spectral response factor of the spectral operator ∇^4 is 1 for all waves).

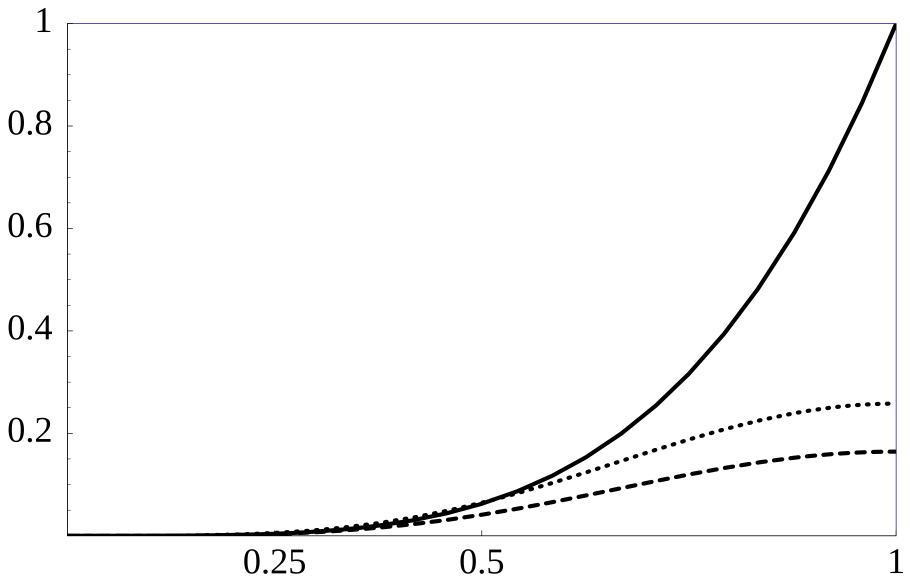

Figure 8: Intensity of the diffusion as a function of the adimensionalised wave-number. Solid: exact (or spectral) operator; dashed: original grid-point HDF operator; dotted: enhanced grid-point HDF operator.

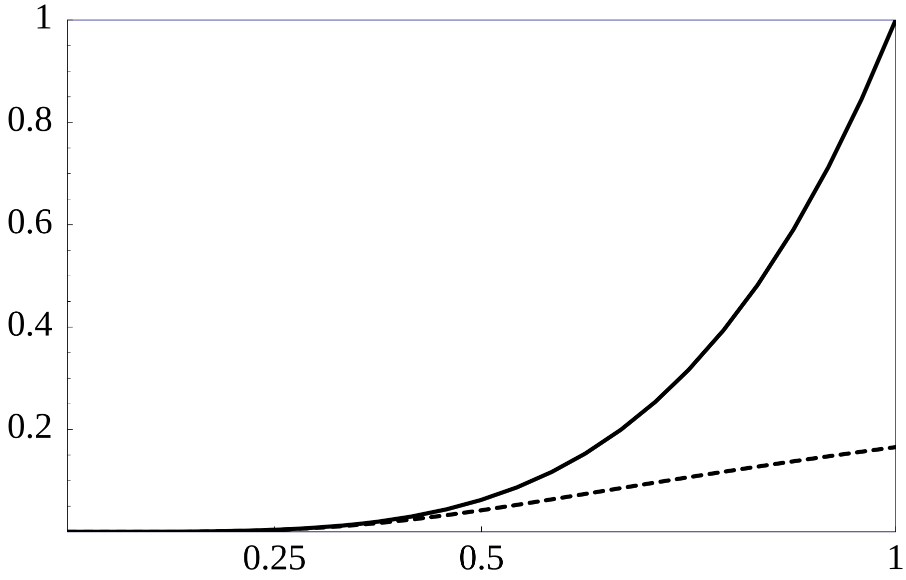

Figure 9: Intensity of the diffusion as a function of the adimensionalised wave-number. Solid: exact (time-continuous) ∇^4 operator; dashed: UFI time-discretized ∇^4 operator.

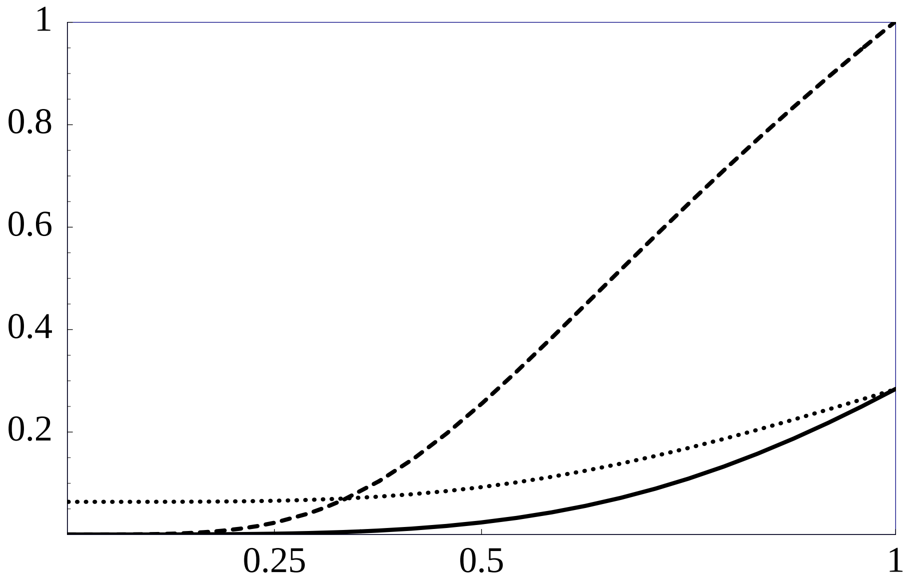

Figure 10: Intensity of the diffusion as a function of the adimensionalised wave-number for the divergence variable. Solid: reference IFS; dashed: operational HDU in ARPEGE; dotted: ratio IFS/HDU.

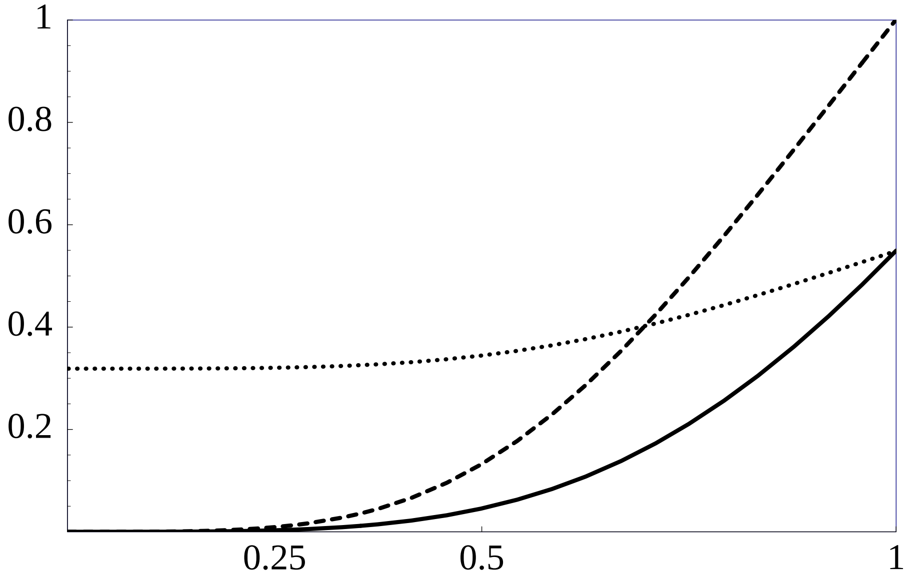

Figure 11: Intensity of the diffusion as a function of the adimensionalised wave-number for the other variables. Solid: reference IFS; dashed: operational HDU in ARPEGE; dotted: ratio IFS/HDU.