Space Sci Rev (2012) 172:397–409
DOI 10.1007/s11214-011-9764-1

Nonmodal Linear Theory for Space Plasmas

Enrico Camporeale

Received: 2 August 2010 / Accepted: 9 February 2011 / Published online: 18 March 2011

Abstract The nonmodal approach is a linear theory formalism that emphasizes the transient evolution of a perturbed equilibrium. It differs from the normal-mode analysis by not assuming an exponential behavior of physical perturbations.

We discuss works that have applied the nonmodal formalism to the problem of solar wind heating and acceleration. We briefly review the methodology of the Kelvin formalism and of the Generalized Stability Theory, and discuss the cases of both sheared and non-sheared plasmas.

The results and methodology reviewed in this paper could form the basis for a trend of research in solar wind dynamics that has not been yet systematically explored.

Keywords Solar wind · Linear theory

1 Introduction: When Normal Mode Analysis Fails

Linear theory is certainly the easiest and most widely applied tool for studying the stability of a physical equilibrium. The *linearization* of a set of nonlinear partial differential equations is a straightforward procedure where all the physical quantities of interest are written as a sum of a term which describes the equilibrium and a term that quantifies the deviation from that equilibrium (the perturbation). If one assumes that the latter term is small compared to the first one, for all the physical variables, then all the non-linear combinations of perturbed quantities can be neglected. The nonlinear equations can be generally cast, in their linearized form, as a set of equations:

$$\frac{d\phi(t)}{dt} = \mathbf{A}\phi(t), \tag{1}$$

where $\mathbf{A}$ is a linear operator (that we will assume is not a function of time), and $\phi(t)$ is the state vector that incorporates all the perturbed variables that describe the system. Clearly

E. Camporeale (✉)
T-5 Applied Mathematics and Plasma Physics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, 87545, USA
e-mail: enrico@lanl.gov

the complexity of the operator $\mathbf{A}$, and the number of variables contained in ϕ depend on the particular model employed. If the system is spatially homogeneous, or can be discretized in space, $\mathbf{A}$ can be thought as a $n \times n$ complex matrix.

One of the major goals of the linear theory is to study whether the assumed equilibrium is unstable, that is if the application of a small perturbation would drive the system away from its original equilibrium, or not. In order to do so, the traditional way of dealing with (1) is to perform a so-called normal mode analysis. This means that the perturbation is assumed to vary in time as $\phi(t) \sim \exp(\omega t)$. Such assumption transforms the set of linear differential equations (1), into the eigenvalue problem:

$$\omega\phi(t) = \mathbf{A}\phi(t), \tag{2}$$

where in general $\omega = \omega_r + i\omega_i$ is a complex quantity. Assuming that one is able to solve (2), the stability problem reduces to computing the sign of the real part of the eigenvalues ω. If there exists at least one solution for which $\omega_r > 0$, then the system is unstable, as the perturbed quantities will grow exponentially in time, eventually breaking the assumption of linearity. On the other hand, if all the solutions have $\omega_r < 0$, then the amplitude of the perturbation will vanish in time, and the initial equilibrium will be restored: the equilibrium is stable.

A few considerations about the normal mode approach are in order, before moving on the description of the nonmodal approach. The choice of an exponential behavior in time for the quantity $\phi(t)$ stems from the fact that the formal solution of (1) is:

$$\phi(t) = e^{\mathbf{A}t}\phi(0), \tag{3}$$

where the exponential of a matrix is defined via its Taylor expansion: $e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{1}{2}(\mathbf{A}t)^2 + \cdots$. Assuming that $\mathbf{A}$ has a complete set of linearly independent eigenvectors (i.e. it is not defective), one can define a matrix $\mathbf{V}$ as the one composed columnwise by the eigenvectors v_j of $\mathbf{A}$ (with $j = 1, \ldots, n$), and the initial condition $\phi(0)$ can be expressed as a weighted combination of eigenvectors:

$$\phi(0) = \mathbf{V}\alpha = \sum_{j=1}^{n} \alpha_j v_j, \tag{4}$$

where α is a column vector which represents the weights. From the identity $e^{\mathbf{A}t} = \mathbf{V}e^{\mathbf{D}t}\mathbf{V}^{-1}$, where $\mathbf{D}$ is the matrix containing the eigenvalues ω_j on its diagonal, it follows that the solution in (3) can be rewritten as

$$\phi(t) = \sum_{j=1}^{n} \alpha_j e^{\omega_j t} v_j. \tag{5}$$

Therefore it is evident that, not only the *ansatz* $\phi(t) \sim e^{\omega t}$ is well-grounded, but that the time-asymptotic solution of (1) will precisely evolve as $e^{\omega_j t}$, with ω_j being the eigenvalue with largest real part. In other words that ansatz is exactly valid only in a time-asymptotic limit, while at a given finite time the solution is a sum of exponentials.[1] A peculiar characteristic of this approach, which is strictly related to its time-asymptotic nature, is that the

[1] This is generally not true if the matrix $\mathbf{A}$ is defective.

choice of the initial condition (4) does not affect the time evolution of the solution, if not by a rescaling coefficient. Moreover, by focusing on the eigenvalue with largest real part (the most rapidly growing mode in an unstable system, or the most slowly decaying in a stable system), the normal mode approach makes also the somewhat hidden assumption that the initial condition is the most general possible. In particular, it assumes that the state vector at time $t = 0$ has a finite projection on the eigenvector that corresponds to the least stable eigenvalue: it is a *worst-case scenario*.

The normal mode approach may be inadequate if a non-exponential transient behavior becomes physically more relevant than the time-asymptotic solution $\phi(t) \sim e^{\omega t}$. The normal mode approach indeed misses any transient. This inadequacy has started to become evident, in the late 1980s, in the hydrodynamics community. Indeed, many neutral fluid experiments found that sheared flows are very often observed to be in a turbulent status, even for Reynolds numbers much lower than the one predicted to trigger an instability. In other words, equilibria that are linearly stable, and therefore are supposed to maintain a laminar regime are instead able to develop a turbulent status. It is evident that this means that in such systems the time-asymptotic prediction of normal mode analysis (i.e. linear stability), is of little interest and does not reflect the transient evolution of some perturbations (Trefethen et al. 1993).

The disagreement with the linear theory has fueled the study of the nonmodal approach, which is now a mature and well-understood tool (Schmid 2007). It is now known that the transient evolution of a linear system, which is the focus of the nonmodal approach, is characterized by different spectral properties of the operator $\mathbf{A}$. This is in contrast with the fact that the normal mode approach focuses only on the eigenvalues of the operator. In particular, there exists a class of linear operators, that are called *non-normal*, that can exhibit a transient growth of some quantities even if the set of eigenvalues is completely contained in the stable half-plane. Transient growth has its origin in the fact that the eigenvectors of a non-normal operator are not mutually orthogonal. This gives rise to strong interactions and interference between different modes. Indeed, an easy inspection of the solution formulated as in (5) reveals that the sum of non-orthogonal vectors, which are all exponentially decreasing in time, can result in a state vector that can grow for a finite amount of time, before it aligns with the most slowly decaying eigenvector and then starts to decay (see Sect. 3). The magnitude of a transient growth can be very large in amplitude, with obvious effects on the dynamics of the perturbation. For instance, non-linear effects could be excited. It has been proposed that the nonmodal amplification of small disturbances in a laminar, stable flow (such as plane Couette or Poiseuille flows) could produce a *by-pass transition* to a turbulent status (Schmid 2000; Reshotko 2001).

Despite the interesting implications of a nonmodal linear theory, this approach has so far been only marginally appreciated in the plasma community.

In this paper we will focus on transient growth that might have an impact on the physics governing the expansion, acceleration, and heating of the solar wind. We present a review of ideas proposed in the last fifteen years that make use of the nonmodal approach, and propose future trends of research.

The paper is organized as follows. Section 2 presents the nonmodal approach for shear flows, briefly discusses the *Kelvin formalism* and some articles that have applied such formalism to the solar wind acceleration problem. Sections 3 and 4 are devoted to non-sheared plasmas. In particular we introduce and explain the effect of transient growth in Sect. 3 and present pseudospectra in Sect. 4. Conclusions are drawn in Sect. 5.

2 Nonmodal Approach: Shear Flows in the Solar Wind

Shear flows are flows that present an inhomogeneity in the velocity field. The linearized operator that describes the evolution of perturbations in shear flows is non-normal. Most of the earliest studies that employed the nonmodal approach have focused their attention on shear flows. Since they are also often observed in the solar wind, this is a relevant topic for the understanding of its acceleration and heating.

In this Section we will review the present state of research, and we will sketch the general idea behind the mathematical formalism.

Most of the works that study shear flows in the solar wind deal with an ideal MHD description of the plasma. A typical setting of the problem is the following. The background magnetic field $\mathbf{B}_0$ is assumed constant and homogeneous, the flow is parallel to the magnetic field, and the velocity profile is linearly sheared in the direction perpendicular to $\mathbf{B}_0$. In Cartesian coordinates (x, y, z):

$$\mathbf{B} = (\mathbf{B}_0, 0, 0); \qquad V = (Sy, 0, 0). \tag{6}$$

The shear constant S parametrizes the inhomogeneity. It is argued that a linear shear profile is generally good enough, for the purpose of linear stability, as long as the typical scale of the shear is much smaller than the length scale of the flow. All the other physical quantities that characterize the equilibrium are assumed homogeneous.

There are two ways to tackle this problem with a nonmodal approach. One is purely computational and consists in discretizing in space the set of linear partial differential equations. This leads to the formulation of a large but probably sparse (depending on the method of discretization) matrix $\mathbf{A}$. This method has not been followed for shear flows, but we will discuss it in the next section, for non-sheared plasmas.

The alternative, and more elegant, method is referred as *Kelvin formalism* and consists in transforming the spatial inhomogeneity into a temporal one, by means of the transformation of variables in the Lagrangian frame of reference:

$$x' = x - Syt; \qquad y' = y; \qquad z' = z; \qquad t' = t. \tag{7}$$

In this way one can perform a Fourier analysis with respect to the new variables, and the state vector is assumed to vary as:

$$\phi = \tilde{\phi}(t) e^{i(k'_x x' + k'_y y' + k'_z z')}. \tag{8}$$

The nonmodal character of this new ansatz consists in the fact that the solution is not assumed to vary as $e^{\omega t}$. The Fourier modes defined in this way are called Spatial Fourier Harmonics (SFH), and the wavenumbers are function of time:

$$k'_y(t) = k_y(0) - k_x St. \tag{9}$$

Chagelishvili et al. (1996) have applied this approach to a 2D MHD unbounded parallel shear flow, and they showed that the Kelvin formalism allows to recast the problem in terms of two ordinary differential equations which describe coupled oscillators:

$$\frac{d^2 X}{dt} + \omega_X^2 X + c(t) Y = 0, \tag{10}$$

$$\frac{d^2 Y}{dt} + \omega_Y^2 Y + c(t) X = 0, \tag{11}$$

where X and Y are linear combinations of the perturbed density and magnetic field, ω_X and ω_Y are the eigenfrequencies of the oscillators X and Y, which are defined as:

$$\omega_X = 1, \tag{12}$$

$$\omega_Y = \sqrt{\left(\frac{c_A}{c_s}\right)^2 + \left[1 + \left(\frac{c_A}{c_s}\right)^2\right]\left[\frac{k'_y}{k'_x} - St'\right]^2}, \tag{13}$$

where c_S and c_A are sound and Alfven speed, respectively. $c(t) = -k'_y/k'_x + St$ is the coupling coefficient, which is a function of k'_y and of the shearing parameter S. In the 2D case, this system admits a solution in which the normal mode frequencies correspond to the slow and fast magnetosonic waves. The peculiarity of this approach, however, is that those frequencies are themselves function of time. The physical picture that corresponds to this solution is that of a perturbation that oscillates with a given frequency on a given SFH, but that *evolves* in time changing its frequency, as the SFH is stretched out in space. Chagelishvili et al. (1996) have further shown that a complete analogy with coupled oscillators exist, and in particular that the effect of the shear flow is to couple and mutually transform slow to fast magnetosonic waves, along a single SFH mode. They point out that the linear transformation mechanism induced by the shear flow is of a different nature from transformations induced by an inhomogeneity in the density profile (Swanson 1998).

Consistently with the mechanical analogy of coupled oscillators, they also found that two conditions are necessary for an effective transformation between linear modes to take place. First, the uncoupled frequencies must be close: there must be a *degeneration region*, where $|\omega_X^2 - \omega_Y^2| \leq |c|$. Note that since the frequencies and the coupling coefficient c are all functions of time, this condition can hold only for a finite amount of time. Indeed, the second condition is that the degeneration region must be passed through slowly, i.e. the oscillators must remain long enough in the region to effectively interact.

When this approach is extended to a 3D MHD flow, the set of (10–11) is coupled to a third oscillator, and the third eigenfrequency corresponds to an Alfven wave. Chagelishvili et al. (1997) have shown that in this case many different mutual transformations between the three modes are possible, for different initial conditions. They have also reported the amplification of the initial perturbation, at the expenses of the energy of the background flow. It is worth noting that Chagelishvili et al. (1997) concludes by suggesting that MHD turbulence, in presence of shear flows, should be formulated as a mixture of the three modes, rather than on the basis of purely Alfvenic fluctuations.

The nonmodal amplification and mutual transformation of linear waves in shear flows have subsequently been proposed as mechanisms responsible (at least to some extent) to the heating and acceleration of the solar wind. Poedts et al. (1998) have verified that such mechanism can be effective under solar wind conditions. Kaghashvili (1999) has studied the transition from lightly damped Alfven waves to high-frequency fast magnetosonic waves, which can be more efficiently dissipated through collisionless damping.

The emerging scenario is that the redistribution of energy, and therefore the heating, is possible, in the presence of a shear, through different mechanisms associated with nonmodal transformation of linear waves. In particular, it has been argued that some part of the energy flow can be effectively extracted and converted into amplified linear waves, that can subsequently change their properties from lightly damped Alfvenic fluctuations to cyclotron resonant high-frequency waves or to Landau damped slow waves, which will ultimately heat the flow. This scenario has been proposed in a number of papers, and its plausibility in different regions of the solar wind and the corona has been discussed (Kaghashvili and Esser 2000; Rogava 2004; Chen 2005; Li et al. 2006; Shergelashvili et al. 2006).

The main disadvantage of the Kelvin formalism is that the solution is given in terms of oscillations of Spatial Fourier Harmonics. Since these modes change their spatial structure in time, it is difficult to visualize how the nonmodal effects so far described would appear in physical space. Such an effort has been undertaken by Bodo et al. (2001), by means of numerical simulations. These authors have shown that a wavepacket, initially composed prevalently with slow magnetosonic waves evolves in time rotating in the physical space (due to the evolution of the wavevectors of the SFH), and partially transforming to fast waves. The identification of slow and fast waves has been done in Bodo et al. (2001) through the phase relations between velocity and magnetic field.

A somewhat more rigorous (but less straightforward) formalism that describe the coupling of linear waves in shear flows has been devised by Gogoberidze et al. (2004), through the use of a *transition matrix* and *transformation coefficients*, in analogy to a quantum mechanical formalism. They have carried out a systematic study of how these coefficients (that quantify the effectiveness of a linear transformation) vary as a function of the plasma β, wavenumbers and shear coefficient S. They have also shown the possibility of *overreflection*, which is analogous to a transient growth. The application of this formalism to the solar wind suggests the result that the transformation of Alfven to fast and slow magnetosonic waves is likely to take place in the energy-containing low frequency range (Gogoberidze et al. 2007). This suggestion is certainly relevant to the understanding and modeling of the solar wind turbulent cascade.

3 Non-Sheared Plasma: Transient Growth

If the plasma is spatially homogeneous the linearized set of equations can be Fourier decomposed in the traditional way, and the matrix $\mathbf{A}$ in (1) is understood to be also a function of the wavevector k. Alternatively, as we anticipated in the previous section, in the presence of a dependence of some quantities on one or more spatial variables, the operator can be discretized in space, and the problem can still be cast as an ordinary differential equation. In both cases the problem reduces to the formulation of a complex (possibly large and sparse) matrix $\mathbf{A}$, and the evolution of the perturbation is completely determined by the propagator $e^{\mathbf{A}t}$.

It becomes a matter of numerical/computational ability to study the structure of the matrix $\mathbf{A}$, and the physical information associated with its non-normality. In this section we will introduce the mathematical tools that characterise the nonmodal evolution of a linear perturbation. This approach has been called Generalized Stability Theory (Farrell and Ioannou 1996).

The non-normality of a matrix can be readily checked by the condition $\mathbf{A}\mathbf{A}^* \neq \mathbf{A}^*\mathbf{A}$, where $\mathbf{A}^*$ indicates the adjoint of $\mathbf{A}$. A non-normal matrix does not commute with its adjoint. The transient growth of a particular initial condition $\phi(0)$ is characterized by the growth function $G(t)$:

$$G(t) = \frac{\|\phi(t)\|}{\|\phi(0)\|} = \frac{\|e^{\mathbf{A}t}\phi(0)\|}{\|\phi(0)\|}, \tag{14}$$

where $\|\cdot\|$ denotes the 2-norm: $\|\phi\| = \sqrt{\sum_{i=1}^{n} |\phi_i|^2}$.

Note that the maximum value of $G(t)$, at a given time, for any possible choice of the initial perturbation $\phi(0)$ coincides with the definition of $\|e^{\mathbf{A}t}\|$, i.e.:

$$G(t) \leq \|e^{\mathbf{A}t}\|. \tag{15}$$

The function $\|e^{\mathbf{A}t}\|$ therefore is the upper bound for the evolution of any possible initial condition. It is interesting to compare the behavior of $\|e^{\mathbf{A}t}\|$ for normal and non-normal matrices. If $\mathbf{A}$ is normal, then

$$\|e^{\mathbf{A}t}\| = e^{\gamma t}, \tag{16}$$

for any time t, where γ is the largest real part of the set of eigenvalues (it is called *spectral abscissa*). For a non-normal operator, (16) holds only in the limit $t \to \infty$. This explains the fact that modal and non-modal approaches converge to the same result, but generally only in the time-asymptotic limit.

The growth function $G(t)$ is meaningful only for a given initial condition, but, embracing the worst-case scenario philosophy which is common to the normal mode analysis, one could ask which particular initial condition maximizes the function $G(t)$ at a given time θ. In other words, what is the initial value $\phi(0)$ such that $G(\theta) = \|e^{\mathbf{A}\theta}\|$.

The answer comes from the singular-value decomposition of the matrix $e^{\mathbf{A}\theta} = \mathbf{U\Sigma V}^*$, where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\mathbf{\Sigma}$ is the diagonal matrix containing the singular values of $e^{\mathbf{A}\theta}$. The column vector of $\mathbf{V}$ associated with the highest singular value is the initial condition that at time θ will reach an amplification exactly equal to $\|e^{\mathbf{A}\theta}\|$, that is the maximum possible amplification at that time for any initial condition.

In order to illustrate the effect of transient growth for a non-normal operator, let us make an example with trivial 2×2 real matrices:

$$\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} -1 & 10 \\ 0 & -2 \end{pmatrix}. \tag{17}$$

It is straightforward to check that $\mathbf{A}$ is normal and $\mathbf{B}$ is non-normal. If one studies the stability of two systems whose evolution are given by the matrices $\mathbf{A}$ and $\mathbf{B}$ with a normal-mode approach, i.e. looking at the sign of the eigenvalues, the result will be the same for both matrices. They indeed have the same eigenvalues $(-1, -2)$ that are both negative, and therefore both systems are stable. Clearly the large off-diagonal term in the matrix $\mathbf{B}$ is responsible for its non-normality and the evolution of a perturbation will be largely affected.

The solution in (5) will be in this case:

$$\phi(t) = \alpha_1 e^{-t} v_1 + \alpha_2 e^{-2t} v_2, \tag{18}$$

but while the eigenvectors of $\mathbf{A}$ are the orthogonal vectors $v_1 = (1,0)$, $v_2 = (0,1)$, the eigenvectors of $\mathbf{B}$ are $v_1 = (1,0)$, $v_2 = (-0.995, 0.0995)$, which are non-orthogonal.

As a consequence some initial conditions $\phi(0)$ can undergo a transient growth in the latter case. We show in Fig. 1 the value of $\|e^{\mathbf{A}t}\|$ (dashed line) and $\|e^{\mathbf{B}t}\|$ (solid line) as a function of time. Since $\mathbf{A}$ is normal, it follows that $\|e^{\mathbf{A}t}\| = e^{-t}$. As we said, the function $G(t)$, that measures the relative growth or decay of an initial perturbation will always be below those curves. It is evident that for the case of matrix $\mathbf{B}$, $G(t)$ can grow for a finite time before decaying.

It is known that plasma physics is characterized by multiple scales in time and spaces, and many different models can be employed depending on the scales and phenomena of interest. Hence, in the non-modal framework it becomes interesting to understand which

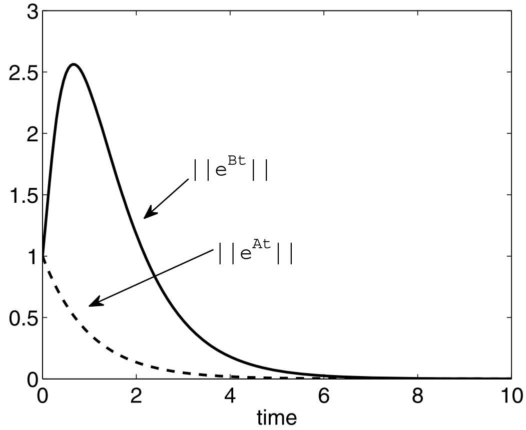

Fig. 1 Time evolution of the norm of the exponential matrix for the matrices **A** and **B**. The *curves* shown bound by above the growth function $G(t)$ for any possible initial perturbation. The matrix **B**, being non-normal can support a transient growth

model is characterized by a non-normal linear operator, and whether the transient effects can reach large amplitudes and/or last for sufficiently long time.

It has been suggested by Camporeale et al. (2009) that the effects induced by non-normality, such as transient growth and mode coupling, become more important when the plasma approaches kinetic scales and higher plasma β. This has been demonstrated using a Landau fluid model (Passot and Sulem 2006, 2007) that is a hierarchical fluid model, closed at the level of fourth order moments, that incorporates Landau damping and finite Larmor radius effects. The fact that a kinetic description of the plasma can more easily accommodate non-normal effects is easily understandable in terms of free-energy. In order for some quantity to transiently grow, there must be a reservoir of free energy available. Camporeale et al. (2009) have shown that a transient growth is possible even for a stable Maxwellian plasma: in this case the small initial perturbations of the distribution function plays the role of free energy, that is redistributed into the plasma via a transient growth of perturbed density, magnetic field, velocity, etc.

The Landau fluid model used in Camporeale et al. (2009) has been shown to support transient growth that, depending on the wavevector k and plasma β, can reach amplitudes of the order of $G(t) \sim 10^3$–10^4, and that can last for times of 10–100 ion gyroperiods. It is very well arguable that such transients can affect the dynamics of the plasma and therefore that the normal mode analysis will be misleading, in this context.

In a more recent paper, Camporeale et al. (2010) have applied the nonmodal approach to the interpretation of space plasma observations. Protons and electrons are very often observed in the solar wind with a certain degree of temperature anisotropy (with respect to the background magnetic field). Depending on the plasma β, large values of anisotropy would trigger kinetic instabilities such as firehose (if $T_\parallel > T_\perp$) or mirror modes (if $T_\parallel < T_\perp$) (see, e.g., Gary and Madland 1985; Li and Habbal 2000; Gary and Karimabadi 2006). The primary effect of such linear instabilities is to reduce the anisotropy and hence decrease the amount of free energy, making the plasma closer to the marginal stability condition (Camporeale and Burgess 2008). However, it has been known for a long time that the values of temperature anisotropy observed is well below the predicted threshold for instabilities (i.e. the marginal stability state) (Kasper et al. 2002; Stverak et al. 2008). This is in contradiction with the fact that the temperature anisotropy is supposed to steadily increase due to the expansion of the wind. Therefore, the combined effect of the expansion and kinetic instabilities should result in the plasma ending up in a state close to the marginal stability condition (Camporeale and Burgess 2010). Moreover, it is known that small amplitude, short wavelength, magnetic fluctuations are persistently observed in the solar wind, even in linearly stable conditions (Bale et al. 2009).

Camporeale et al. (2010) have suggested that those apparent contradictions between linear theory and observations could be re-conciliated by using a non-modal approach. The scenario proposed is one where magnetic fluctuations, amplified non-modally, would scatter particles and consequently reduce the temperature anisotropy even in conditions far from the linear threshold for kinetic instabilities.

Therefore the particle temperature anisotropy would be regulated via locally generated transient fluctuations, without the need of any kinetic instability. The plausibility of such scenario has been demonstrated studying a very large ensemble of initial perturbations. The results of Camporeale et al. (2010) showed that transient growth are not a sporadic events associated with a peculiar choice of initial conditions, but rather an intrinsic feature of space plasmas.

4 Non-Sheared Plasma: Pseudospectra

An important tool that has been developed for the understanding of the dynamics associated with non-normal operators is represented by the pseudospectrum and pseudoeigenvalues. Those are a generalization of the standard concepts of spectrum and eigenvalues of a matrix. The pseudospectrum is particularly useful to visualize and understand how sensitive to small perturbations a linear operator is.

For a matrix $\mathbf{A}$ we define the spectrum $\Lambda(\mathbf{A})$ as the set of complex numbers z such that $\det(z\mathbf{I}-\mathbf{A})=0$. By convention we define $\|(z\mathbf{I}-\mathbf{A})^{-1}\|=\infty$. The ε-pseudospectrum $\Lambda_\varepsilon(\mathbf{A})$ is defined as (Trefethen and Embree 2005):

$$\Lambda_\varepsilon(\mathbf{A})=\{z\in\mathbb{C}:\|(z\mathbf{I}-\mathbf{A})^{-1}\|>1/\varepsilon\}.$$

This definition is equivalent (see, e.g., Trefethen and Embree 2005 for the proof) to say that

$$\Lambda_\varepsilon(\mathbf{A})=\{z\in\mathbb{C}:z\in\Lambda(\mathbf{A}+\mathbf{E})\text{ for some }\mathbf{E}\text{ with }\|\mathbf{E}\|<\varepsilon\}.$$

In words, the ε-pseudospectrum is the open set that contains the eigenvalues of the perturbed operator $\mathbf{A}+\mathbf{E}$, with the smallness of $\mathbf{E}$ measured by ε. Clearly, the pseudospectrum $\Lambda_\varepsilon(\mathbf{A})$ is defined such that it reduces to the spectrum $\Lambda(\mathbf{A})$ when $\varepsilon\to 0$.

Why are we interested in pseudospectra? Because a peculiar characteristic of non-normal matrices is that they are highly sensitive to perturbations, in the sense that their set of eigenvalues can be easily distorted in the complex plane, when the original operator is slightly perturbed. Such distortion is exactly measured by the concept of pseudospectra.

Reasoning in more physical terms one could argue that when the stability of an equilibrium is studied using linear theory, the linear operator always represents an ideal configuration. For instance, inhomogeneities of some physical quantities and noise could be thought to be represented as small deviations of the linear operator. Such small deviations are not to be confused with the linear perturbations applied to the equilibrium. In this way the possible distortion of the spectrum might become more informative than the original set of eigenvalues. In particular, when the ε-pseudospectrum for a small value of ε contains more than one eigenvalue, the dynamics (and also the time-asymptotic evolution) of the system will be given by pseudoeigenvalues that can lie anywhere within the

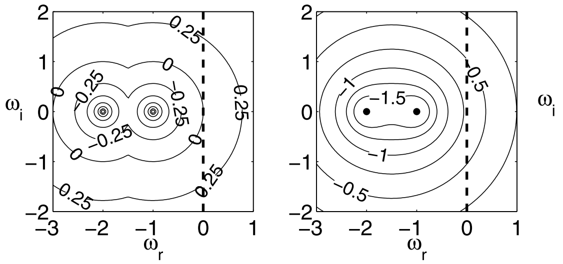

Fig. 2 Pseudospectra for the matrices **A** (*left*) and **B** (*right*). The values of $\log \varepsilon$ for the isolevels are reported on the contours

pseudospectrum contour. This means that the original eigenvalues (the unperturbed spectrum) become meaningless, because they can be easily dislocated to different locations in the complex plane, with consequent changes in the damping rate and frequency associated with them.

In order to better understand the fact that non-normal operators have a spectrum that is more easily distortable than normal operators, we show in Fig. 2 the contours of the pseudospectra for the matrices **A** and **B**, in the left and right panel respectively. The contours are given as $\log_{10} \varepsilon$. A dashed line divides the plane in the stable (left) and unstable (right) halves. The contours in the two plots are not qualitatively different. The main difference is in the values of ε attached to similar contours. For example the smallest contours that contains both eigenvalues $(-2, -1)$, is $\varepsilon = 10^{-0.25}$ for the normal matrix **A**, and $\varepsilon = 10^{-1.5}$ for the non-normal matrix **B**, which indicates that the spectrum of **B** is more easily deformable.

Clearly our artificial example is not related to any physical model. However, the concept of pseudospectrum has been used in plasma physics by Borba et al. (1994) to solve the *Alfven paradox* for the resistive MHD spectrum. It was known that the discrete spectrum of resistive MHD does not converge to the continuous spectrum of ideal MHD, when the resistivity tends to vanish. Borba et al. (1994) have suggested that the two spectra indeed converge to the same solution when the resistive spectrum is extended to its ε-pseudospectrum. Indeed, for any $\varepsilon > 0$, the ε-pseudospectrum contains the ideal spectrum, for sufficiently small resistivity.

Plots of the pseudospectra for plasma physics models have also been presented in Camporeale et al. (2009) for the Landau fluid model, and in van Dorsselaer (2002) for the resistive MHD. In the context of space plasmas, it would certainly be interesting to study the pseudospectrum of the Vlasov-Maxwell linear equations. However, a complication related to such equations is that they cannot be cast in terms of linear ordinary differential equation of the form in (1). Consequently, the normal mode analysis of the Vlasov-Maxwell linear equations does not reduce to a linear eigenvalue problem of the form as in (2). The set of equations are instead cast as a nonlinear eigenvalue problem: $\mathbf{D}(\omega, k) \cdot \phi = 0$ (see, e.g., Stix 1992).

Fortunately, the definition of ε-pseudospectrum has been extended to nonlinear eigenvalue problems (see, e.g., Tisseur and Higham 2001; Michiels et al. 2006). A more systematic study of the pseudospectrum of the Vlasov-Maxwell linear equations will be reported elsewhere, but for the purpose of this review we show here few examples and comment on their physical meaning.

We show in Fig. 3 the contour of the Vlasov-Maxwell pseudospectra for the following parameters: $k\rho_i = 1$, $\theta = 80°$, $\beta = 0.1$, $v_A/c = 0.01$, where ρ_i is the ion gyroradius, θ is the angle formed between the wavevector k and the background magnetic field, v_A is the Alfven velocity, and c is the speed of light. The portion of the complex plane shown in Fig. 3 is

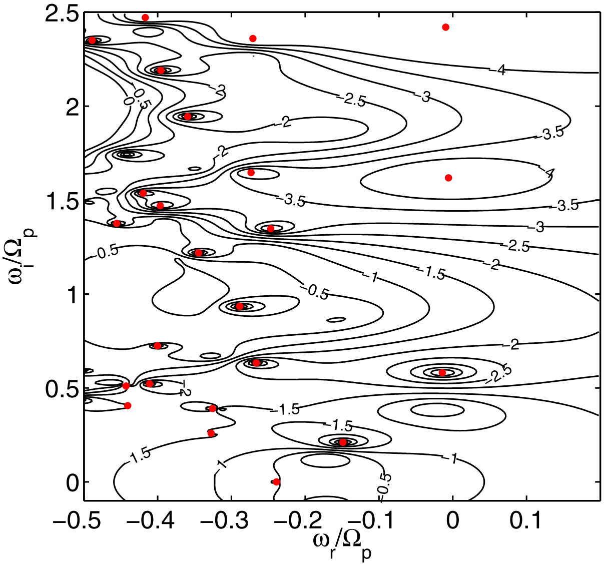

Fig. 3 Pseudospectra of the linear Vlasov-Maxwell operator for the parameters: $k\rho_i = 1$, $\theta = 80°$, $\beta = 0.1$, $v_A/c = 0.01$ (definitions in the text). The *red dots* indicate the eigenvalues, and the contours are expressed as $\log \varepsilon$. Ω_p is the proton gyrofrequency

$\omega_r \in [-0.5, 0.2]$, $\omega_i \in [-0.1, 2.5]$. Frequencies are normalized to the proton gyrofrequency. The eigenvalues are indicated with red points. The contours shown in Fig. 3 denote isolevels of $\log \varepsilon$.

The pseudoeigenvalues for a given value of ε can lie everywhere inside an ε-contour. All the contours are closed lines, although the may close outside the portion of complex plane shown in the figure. The point that we want to make is that there are many contours that connect points of the complex plane which are quite distant, and however pass close to an eigenvalue. This suggest that a large distortion of the spectrum is possible. Also, it is worth noting that different eigenvalues have different behavior and generally the most heavily damped modes are the ones that need a smaller disturbance to be relocated far away from their original position. Heavily damped modes are generally regarded as useless for the description of the evolution of a linear system. The usual argument is that the fact that they are quickly damped makes those modes virtually impossible to detect, and therefore they are assumed not to give any contribution to the dynamics. However, in the light of the results shown in Fig. 3 one could argue that heavily damped modes could be relocated under the effect of small perturbations of the linear operator. They could possibly have to be taken in account for instance for the understanding of damping turbulent fluctuations in the solar wind. Figure 4 shows another contour plot of pseudospectra for the same parameters as in Fig. 3, but for parallel propagation $\theta = 0°$. It is interesting that in this case all of the modes appear to be more resilient to perturbations, except for the two purely oscillating modes. The different behavior of pseudospectra for different angle of propagation will need further investigation.

At the present the results of Figs. 3 and 4 are just qualitative and not predictive. However they clearly show that the dynamics of the linear Vlasov-Maxwell equations should be understood in terms of a nonmodal stability theory. The large distortion of the spectrum shown is indeed typical of non-normal operators. Moreover the role played by highly damped modes should be investigated.

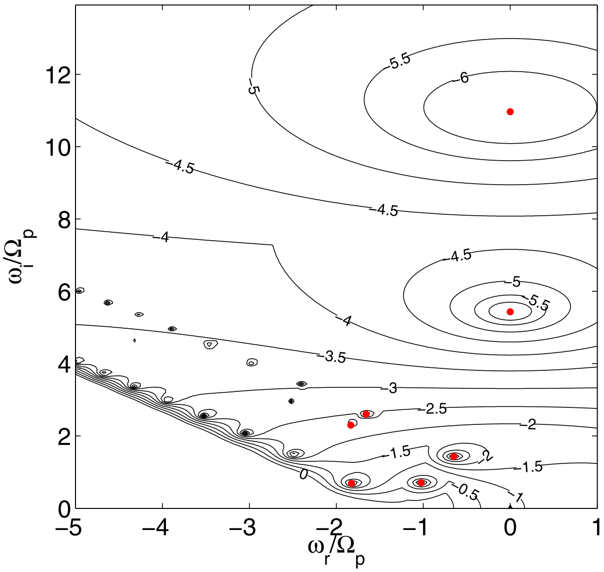

Fig. 4 Pseudospectra of the linear Vlasov-Maxwell operator for the parameters: $k\rho_i = 1$, $\theta = 0°$, $\beta = 0.1$, $v_A/c = 0.01$ (definitions in the text). The *red dots* indicate 7 of the least damped eigenvalues, and the contours are expressed as $\log \varepsilon$

5 Conclusions

It is now clear that the normal-mode analysis should be complemented by a non-modal approach when the linear operator of interest can support transient growth due to its non-normality.

In this paper we have reviewed the main results obtained through the application of the nonmodal approach to a simplified solar wind model.

There are issues related to the choice of an appropriate norm in (14), that we have not discussed, and we refer to Schmid (2007) for a thorough discussion.

What has been clearly established in the past few years, from a theoretical point of view, is that nonmodal mechanisms are able to couple, transform and transiently amplify linear modes in the solar wind, both in shear flows and in bi-Maxwellian plasmas, and that predictions based on the normal-mode analysis can be very often misleading.

Some realistic scenarios relevant to the heating, the acceleration, and the control of macroscopic parameters in the expanding solar wind have been proposed in the papers reviewed in this article. In particular, it has been suggested the possibility that a fraction of the heating can be provided by the damping of nonmodally amplified modes.

What is still missing to make those scenarios useful is probably a more quantitative predictions based on nonmodal analysis, that would be able to interpret some satellite observations. In particular, the concept of pseudospectra, presented in Sect. 4, although very useful to understand the spectral properties of a linear operator, still lacks a connection with physical observables.

We suggest that a better understanding of those issues will be a trend of research which is worth exploring in a near future, and that it could help resolving the long time open problem about the solar wind heating and acceleration.

Acknowledgements I acknowledge the kind hospitality of the International Space Science Institute (Bern), and I would like to thank Gian Luca Delzanno and David Burgess for useful discussions on the topic of nonmodal linear theory. This work was partially supported by STFC grant ST/H002731/1.

References

S.D. Bale, J.C. Kasper, G.G. Howes, E. Quataert, C. Salem, D. Sundkvis, Phys. Rev. Lett. **103**, 211101 (2009)
G. Bodo, S. Poedts, A. Rogava, P. Rossi, Astron. Astrophys. **374**, 337–347 (2001)
D. Borba, K.S. Riedel, W. Kerner, G.T.A. Huysmans, M. Ottaviani, P.J. Schmid, Phys. Plasmas **1**, 3151–3160 (1994)
E. Camporeale, D. Burgess, J. Geophys. Res. **113**, A07107 (2008)
E. Camporeale, D. Burgess, Astrophys. J. **710**, 1848 (2010)
E. Camporeale, D. Burgess, T. Passot, Phys. Plasmas **16**, 030703 (2009)
E. Camporeale, T. Passot, D. Burgess, Astrophys. J. **715**, 260–270 (2010)
G.D. Chagelishvili, A.D. Rogava, D.G. Tsiklauri, Phys. Rev. E **53**, 6028–6031 (1996)
G.D. Chagelishvili, R.G. Chanishvili, J.G. Lominadze, A.G. Tevzadze, Phys. Plasmas **4**, 259–269 (1997)
Y. Chen, Phys. Plasmas **12**, 5 (2005)
B.F. Farrell, P.J. Ioannou, J. Atmos. Sci. **53**, 2025–2040 (1996)
S.P. Gary, C.D. Madland, J. Geophys. Res. **90**(A8), 7607–7610 (1985)
S.P. Gary, H. Karimabadi, J. Geophys. Res. **111**, A11224 (2006)
G. Gogoberidze, G.D. Chagelishvili, R.Z. Sagdeev, D.G. Lominadze, Phys. Plasmas **11**, 4672–4685 (2004)
G. Gogoberidze, A. Rogava, S. Poedts, Astrophys. J. **664**, 549–555 (2007)
E.K. Kaghashvili, Astrophys. J. **512**, 969–974 (1999)
E.K. Kaghashvili, R. Esser, Astrophys. J. **539**, 463–468 (2000)
J.C. Kasper, A.J. Lazarus, S.P. Gary, Geophys. Res. Lett. **29**, 1839 (2002)
X. Li, S.R. Habbal, J. Geophys. Res. **105**(A12), 27377–27385 (2000)
J.W. Li, Y. Chen, Z.Y. Li, Phys. Plasmas **13**, 042101 (2006)
W. Michiels, K. Green, T. Wagenknecht, S.I. Niculescu, Linear Algebra Appl. **418** (2006)
T. Passot, P.L. Sulem, J. Geophys. Res. **111**, 4203 (2006)
T. Passot, P.L. Sulem, Phys. Plasmas **14**(8), 082502 (2007)
S. Poedts, A.D. Rogava, S.M. Mahajan, Astrophys. J. **505**, 369–375 (1998)
E. Reshotko, Phys. Fluids **13**, 1067–1075 (2001)
A. Rogava, Astrophys. Space Sci. **293**, 189–196 (2004)
P.J. Schmid, Phys. Plasmas **7**, 1788–1794 (2000)
P.J. Schmid, Annu. Rev. Fluid Mech. **39**, 129–162 (2007)
B.M. Shergelashvili, S. Poedts, A.D. Pataraya, Astrophys. J. **642**, L73–L76 (2006)
T.H. Stix, *Waves in Plasmas* (American Institute of Physics, New York, 1992)
S. Stverak, P. Travnicek, E. Maksimovic, E. Marsch, A.N. Fazakerley, E.E. Scime, J. Geophys. Res. **113**, A03103 (2008)
D.G. Swanson, *Theory of Mode Conversion and Tunneling in Inhomogeneous Plasmas* (Wiley, New York, 1998)
F. Tisseur, N.J. Higham, SIAM J. Matrix Anal. Appl. **23**, 1 (2001)
L.N. Trefethen, A.E. Trefethen, S.C. Reddy, T.A. Driscoll, Science **261**, 578–584 (1993)
L.N. Trefethen, M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators* (Princeton Univ. Press, Princeton, 2005)
J.L.M. van Dorsselaer, SIAM J. Sci. Comput. **24**, 3 (2002)