DOE/ET-53088-280 IFSR #280

The Free Energy Principle, Negative Energy Modes, and Stability

P.J. Morrison and *M. Kotschenreuther*
Department of Physics and Institute for Fusion Studies
The University of Texas at Austin
Austin, Texas 78712

October 1989

Except for minor modifications, the first three sections of this paper were completed in 1986. This material was presented at several meetings, including the following invited presentations: American Physical Society Division of Plasma Physics Meeting, November 1986, Baltimore, Maryland, Bull. Am. Phys. **31**, 1609 (1986); Currents in Geophysics and Hydrodynamics, University of California, June 1986, Lake Arrowhead, CA; International Atomic Energy Agency Meeting on Controlled Thermonuclear Research, November 1986, Kyoto, Japan, [c.f. Ref. 7]. Section 4 contains more recent material, as documented in Refs. 8 and 9.

THE FREE ENERGY PRINCIPLE, NEGATIVE ENERGY MODES, AND STABILITY

P. J. Morrison and M. Kotschenreuther

Department of Physics and Institute for Fusion Studies
University of Texas at Austin
Austin, TX 78712

Presented at the IV International Workshop
"Nonlinear and Turbulent Processes in Physics"
October 9-22, 1989 Kiev, U.S.S.R.

I. INTRODUCTION

1.1. Overview

This paper is concerned with instability of equilibria of Hamiltonian, fluid and plasma dynamical systems. Usually the dynamical equilibrium of interest is not the state of thermodynamic equilibrium, and does not correspond to a free energy minimum. The relaxation of this type of equilibrium is conventionally considered to be initiated by linear instability. However, there are many cases where linear instability is not present, but the equilibrium is nonlinearly unstable to arbitrarily small perturbations. This paper is about general free energy expressions for determining the presence of linear or nonlinear instabilities. These expressions are simple and practical, and can be obtained for all equilibria of all ideal fluid and plasma models. By free energy, we mean the energy change upon perturbations of the equilibrium that respect dynamical phase space constraints. This quantity is measured by a self-adjoint quadratic form, called $\delta^2 F$. The free energy can result in instability when $\delta^2 F$ is indefinite; i.e. there exist accessible perturbations that lower the free energy of the system.

A primary purpose of this paper is to tie together three manifestations of what we will refer to as *negative energy modes*. The first is the conventional plasma physics notion of negative energy mode that is based on the definition of the energy in a homogeneous dielectric medium[1-3]. A negative energy mode is a normal mode of the medium (plasma) that possesses negative dielectric energy. The second manifestation occurs in finite degree-of-freedom Hamiltonian normal form theory[4,5]. The quadratic part of a Hamiltonian in the vicinity of an equilibrium point, which possesses only distinct oscillatory eigenvalues, has an invariant signature. Thus in cases where the quadratic form is indefinite, it is natural to refer to the modes corresponding to the negative signature as negative energy modes. The third, and most general, definition of a negative energy mode[6-9] relies on the free energy functional that we have termed $\delta^2 F$. An equilibrium has a negative energy mode if it is linearly stable and has indefinite $\delta^2 F$, and further we require this to be true in all frames of reference. This later proviso is required since energy is not a covariant quantity.

The definition of negative energy modes based on $\delta^2 F$ is a generalization of the dielectric definition in that it is applicable to arbitrary equilibria. It does not require Fourier transformation in space or time, nor does it require the existence of a dielectric function. The δ^2F expressions do not contain unknown frequency dependence and in general are easier to use. The δ^2F definition obviously generalizes the normal form definition since the systems of interest are generally partial differential equations. We point out, however, that it includes the previous two. All three definitions are equivalent for systems with a complete discrete spectrum and a notion of dielectric function, such as the fluid description of the beam-plasma instability.

A second primary purpose of this paper is to describe a conjecture that we refer to as the *free energy principle*, a principle that yields criteria which are sufficient and, in a sense, necessary for stability. When δ^2F is definite an equilibrium is stable. When δ^2F is indefinite and we are in a reference frame where the equilibrium has minimum energy, there are two possibilities: (1) the equilibrium possesses linear instability or (2) there is linear stability with negative energy modes. Possibility (2) is a restatement of a definition, but in this case there are two avenues for instability: (a) the equilibrium can be unstable to an arbitrarily small perturbation because of nonlinearity or (b) the equilibrium can be structurally unstable to the inclusion of dissipation in the dynamical model. There are many examples that illustrate (a), in particular, a simple example due to Cherry[10] (which will surface in Sec. 3) and important calculations of nonlinear electrostatic instabilities in homogeneous plasma[11-15]. Illustrations of (b) include the well-known Thompson-Tait theorem[16], and we point out an early plasma physics example[17]. It is conjectured that (a) and (b) are generic; i.e. although there exist cases where δ^2F is indefinite and the system is stable, and there exist special types of dissipation that do not result in instability, these are conjectured to be exceptional. If we accept the conjecture and ignore these possibilities because of their rarity, we obtain a sense in which δ^2F provides a "necessary" and sufficient condition for stability. This is what is meant by the free energy principle.

The fluid and plasma literature contains a large number of sufficient conditions for the stability of ideal equilibria that depend upon the positive definiteness of some quadratic form. These criteria are often obtained by ad hoc means and sometimes are devoid of physical interpretation. Generally these conditions, unlike the magnetohydrodynamic (MHD) energy principle, are believed to yield no information when indefinite. In fact a great deal of effort has been spent in the unsuccessful attempt to obtain necessary and sufficient conditions for the stability of general equilibria. In all cases known to the authors, definiteness of these forms amounts to definiteness of the free energy. Thus the notion of δ^2F greatly clarifies the situation in that there is a framework for interpreting and constructing these criteria. Also it puts to rest the question of necessary and sufficient conditions. The best to be hoped for in general is the "necessary" condition described above, because in general, Hamiltonian systems do not possess δW type criteria.

In the remainder of the Introduction (Sec. 1.2) we discuss two descriptions of equations that describe ideal plasma, the Lagrangian variable description, which possesses an action principle[18-20], and the Eulerian variable description, which posseses the noncanonical Hamiltonian structure[21-29]. Both the Lagrangian[9] and Hamiltonian[5-7,29] formalisms have been used to obtain free energy expressions. We will dwell somewhat on the later, but quote results obtained from the former. Section 3 contains an example that typifies finite Larmour radius stabilization. Here we see in a simple system of ordinary differential equations, how a system can be linearly stable, yet unstable, and how the inclusion of dissipation can result in linear instability by shifting a real frequency into the complex plane. In Sec. 4 the Vlasov-Poisson system is treated. A general expression for the free energy of arbitrary equilibria is stated.

1.2 Review

In the Lagrangian variable description of continuous media, for example fluids, the complete state of a system is determined by a continuum of "particle" positions, $\mathbf{x}(\mathbf{x}_0,t)$, where $\mathbf{x}_0$ is the particle position at $t=0$. The Lagrangian variable description is model independent in the sense that it arises in kinetic[20] as well as fluid systems[18,19] with a generality of force laws. In the case of fluid systems, $\mathbf{x}$ corresponds to the position of a "fluid element". This case differs from the usual description of a fluid in terms of Eulerian variables, where the density and velocity fields are expressed as functions of space and time. In contrast to the Eulerián variable description, when one describes a fluid as a collection of particles, as in the Lagrangian variable description, it is not surprising that the equations of motion possess the form of Newton's second law, and therefore are derivable from Hamilton's principle of mechanics.

As an example consider the nonlinear equations of ideal magnetohydrodynamics (MHD), which have the following form:

$$\frac{\partial^2 \mathbf{x}}{\partial t^2} = -\frac{\delta W[\mathbf{x}]}{\delta \mathbf{x}} \quad , \qquad (1)$$

where the right hand side is the functional or variational derivative of a potential energy functional, $W[\mathbf{x}]$. The MHD energy principle exists because of the Hamiltonian form of Eq. (1), for if we linearize by letting $\mathbf{x} = \mathbf{x}_e(\mathbf{x}_0) + \boldsymbol{\xi}(\mathbf{x}_0, t)$, where $\mathbf{x}_e$ satisfies $\delta W[\mathbf{x}_e]/\delta \mathbf{x} = 0$, then $\boldsymbol{\xi}$ satisfies

$$\frac{\partial^2 \boldsymbol{\xi}}{\partial t^2} = O\boldsymbol{\xi} \quad , \tag{2}$$

where the linear operator O arises from the second variation of $W[\mathbf{x}]$,

$$\delta^2 W = \frac{1}{2} \int \boldsymbol{\xi} \cdot \frac{\delta^2 W[\mathbf{x}_e]}{\delta \boldsymbol{\xi} \delta \boldsymbol{\xi}} \cdot \boldsymbol{\xi} \; d^3x = \frac{1}{2} \int \boldsymbol{\xi} \cdot O \cdot \boldsymbol{\xi} \; d^3x \; . \tag{3}$$

From Eq. (3) it is evident that O is self-adjoint since it is the second variation of W. (The quantity $\delta^2 W$ is usually misleadingly referred to as δW.) This self-adjointness, with its root traced back to the Hamiltonian form of the original nonlinear system, is the crucial element that gives rise to the necessary and sufficient δW energy criteria for static MHD equilibria.

Consider now the Eulerian variable description. One thing that is immediately evident is that there is a difference between the classes of Eulerian and Lagrangian equilibria. For example, in ideal MHD the dynamical variables are the velocity field $\mathbf{v}(\mathbf{x},t)$, the density $\rho(\mathbf{x},t)$, the magnetic field $\mathbf{B}(\mathbf{x},t)$, and the entropy per unit mass $s(\mathbf{x},t)$. If we define equilibria by the vanishing of the first time derivative of the dynamical variables, then Eulerian equilibria correspond to some (not all) functions of space. In particular note that the equilibrium velocity field, $\mathbf{v}_e(\mathbf{x})$, need not be constant nor vanish. This should be contrasted to Lagrangian equilibria where $\partial \boldsymbol{\xi}/\partial t = 0$, which implies that there is no equilibrium flow. Thus we see that Lagrangian equilibria are static equilibria, while the class of Eulerian equilibria includes stationary equilibria as well.

Another difference between the Eulerian and Lagrangian variable descriptions is the apparent difference in the number of dynamical variables needed to specify the state of the system. The map from Lagrangian to Eulerian variables is not a one-to-one transformation, and evidently it is not a canonical transformation. In spite of the "pathology" of this map, the Hamiltonian structure survives, albeit in what we have called the noncanonical Hamiltonian form.

For simplicity we discuss a finite degree-of-freedom Hamiltonian system, which for ease in generalizing to the noncanonical description we write in terms of the 2N dynamical variables z^i, the first N of which are the canonical coordinates while the second N are the conjugate momenta. Hamilton's canonical equations take the form

$$\frac{dz^i}{dt} = [z^i, H] = J_c^{ij} \frac{\partial H}{\partial z^j} \quad , \tag{4}$$

where the Poisson bracket is defined by

$$[f,g] = \frac{\partial f}{\partial z^i} J_c^{ij} \frac{\partial g}{\partial z^i} \tag{5}$$

with

$$(J_c^{ij}) = \begin{bmatrix} 0 & I_N \\ -I_N & 0 \end{bmatrix} \quad , \tag{6}$$

which is a $2N \times 2N$ matrix and I_N is the $N \times N$ unit matrix. Here repeated indices are to be summed to 2N. The quantity (J_c^{ij}) is a second order contravariant tensor that is called the cosymplectic form.

Canonical transformations, by definition, preserve the form of the Poisson bracket, or equivalently the form of the tensor (J_c^{ij}). This is not the case for the transformation between Lagrangian and Eulerian variables, since this transformation is noncanonical and moreover degenerate. For finite degree-of-freedom systems, the noncanonical Hamiltonian form is the same as that of Eqs. (4)-(6), except (J_c^{ij}) is replaced by a tensor (J^{ij}), which may be odd dimensional, have vanishing determinant and depend upon the z^i. However, in spite of these changes in the cosymplectic form, the algebraic properties of bilinearity, antisymmetry and the Jacobi identity of the Poisson bracket, are preserved. These properties are the essence of the Hamiltonian description.

For MHD the Eulerian remnant of the form given by Eq. (1) is

$$\frac{\partial \psi^i}{\partial t} = J^{ij} \frac{\delta H}{\delta \psi^j} , \tag{7}$$

where the ψ^i ($i = 1,\ldots 8$) correspond to the field variables **v**, **B**, ρ and s; the Hamiltonian $H = \int[\rho v^2/2 + B^2/2 + \rho U(\rho,s)]\, d^3x$, where U is the internal energy per unit mass; and the cosymplectic operator, (J^{ij}), was given in Ref. [25].

A feature of noncanical Poisson brackets, not present in ordinarily Poisson brackets, is the existence of special constants of motion called Casimir invariants. A phase space function C is a Casimir invariant if it commutes with the Hamiltonian as well as with any function F of the dynamical variables z^i ($i = 1,\ldots M$) describing the system, i.e.

$$[C, F(z)] = 0 \ . \tag{8}$$

A consequence of this definition of the Casimirs, using Eq. (5), is

$$\frac{\partial C}{\partial z^i} J^{ij} \frac{\partial F}{\partial z^j} = 0 \tag{9}$$

but F is arbitrary and therefore

$$J^{ij} \frac{\partial C}{\partial z^j} = 0 , \qquad i = 1,\ldots,M. \tag{10}$$

Thus, the phase space gradient of a Casimir ($\partial C/\partial z^j$) is a null eigenvector of (J^{ij}). In fact, it can be shown[21] that the null space of (J^{ij}) is spanned by null eigenfunctions that are gradients. Clearly, nontrivial Casimirs (i.e. not constants) exist only if

$$\det(J^{ij}) = 0 \tag{11}$$

and the number of independent Casimirs is equal to the corank of (J^{ij}). In the case that (J^{ij}) is canonical, it has the structure given in Eq. (5) and the determinant is unity.

Therefore in the canonical Hamiltonian formalism there are no nontrivial Casimirs. When (J^{ij}) has null eigenvectors, the phase space can be described by hyperplanes, or symplectic leaves, which are labeled by the Casimirs. A trajectory must remain in the leaf of phase space as determined by the specification of the initial conditions. This follows from the fact that the generalized Poisson bracket cannot generate flow, i.e. trajectories in phase space in the direction of these null eigenvectors.

Hamiltonian systems possess a built-in sufficient condition for Liapunov stability. This kind of stability is a stronger than conventional linear or spectral stability. The most common example of nonlinear stability occurs in the case where the Hamiltonian has the standard form: $H = p^2/2 + V(q)$. Here it is well-known that positivity of the curvature of V is sufficient (and under some mild conditions also necessary[5]) for stability of equilibria given by $p^i = 0$, $\partial V/\partial q^i = 0$. This is the finite degree-of--freedom analogue of the MHD energy principle.

The above curvature condition is a special case of a more general condition for Hamiltonian systems. The topology of energy surfaces in the vicinity of an equilibrium point is determined by the curvature of the entire Hamiltonian: $\partial^2 H/\partial z^i \partial z^j$. If this quantity is either positive or negative definite then the energy surfaces near the equilibrium point are nested, closed and bounded surfaces. Since trajectories are confined to the energy surface, those with initial conditions sufficiently close to the equilibrium point will remain in an arbitrarily small neighborhood of the equilibrium point. Evidently, for standard Hamiltonians, positive definiteness of H is insured by that of V.

For noncanonical Hamiltonian systems there is an analogous sufficient condition. To begin with, unlike conventional Hamiltonian systems, extremals of the Hamiltonian are not the only possible equilibria. In fact these equilibria, which correspond to the lowest possible energy states, are generally quite trivial. For example, varying the MHD energy with respect to $\mathbf{v}$, $\mathbf{B}$, etc. yields an equilibrium with zero flow, magnetic field, pressure and temperature. If one extremizes the Hamiltonian subject to the constancy of certain of the Casimirs, then interesting equilibria that are constrained away from the lowest energy state arise. That these are equilibria is evident from the following:

$$\frac{dz^i}{dt} = [z^i, H] = [z^i, H + C] = J^{ij}\frac{\partial F}{\partial z^j} \quad . \tag{12}$$

Thus vanishing of $\partial F/\partial z^i$ implies that $dz^i/dt = 0$. If we define

$$\delta^2 F \equiv \frac{1}{2}\frac{\partial^2 F(ze)}{\partial z^i \partial z^j}\,\delta z^i \delta z^j \quad , \tag{13}$$

then definiteness of the quadratic form $\delta^2 F$ implies that surfaces of constant F near the equilibrium point are topologically spheres and thus we have stability. In the next section we physically interpret this quantity in the noncanonical context.

The free energy functional $\delta^2 F$ can also be obtained in the Lagrangian context[9]. In fact, since all equilibria are not extremal points of $\delta F = 0$ in the noncanonical formalism, the most general expressions have been obtained by beginning with Lagrangian action principles. Free energy expressions obtained in either way are equivalent for equivalent equilibria.

2. THE FREE ENERGY - $\delta^2 F$

Now consider the physical interpretation of $\delta^2 F$. Here we show that $\delta^2 F$ is the energy change resulting from perturbations of the equilibrium that obey the dynamical constraints. Further, we show that the use of $\delta^2 F$ considerably simplifies the computation of the energy change. Also we discuss the connection between $\delta^2 F$ and the dielectric definition of negative energy modes.

It is useful to consider the energy needed to create a small perturbation with the constraint that the motion remain in a symplectic leaf. For small δz we have

$$\Delta H \equiv H(z_e + \delta z) - H(z_e) = \frac{\partial H}{\partial z^i}\delta z^i + \frac{1}{2}\frac{\partial^2 H}{\partial z^i \partial z^j}\delta z^i \delta z^j + \ldots \tag{14}$$

$$\Delta C \equiv C(z_e + \delta z) - C(z_e) = \frac{\partial C}{\partial z^i}\delta z^i + \frac{1}{2}\frac{\partial^2 C}{\partial z^i \partial z^j}\delta z^i \delta z^j + \ldots , \tag{15}$$

where z_e is an equilibrium given by $\partial F/\partial z^i = 0$. For δz such that $\Delta C = 0$, we can add $\Sigma \lambda^i \Delta C^i$ to Eq. (14) in order to obtain the energy change at constant Casimir invariant,

$\Delta H|_C$. Thus

$$\Delta H|_C = \frac{1}{2}\delta^2(H + \Sigma\,\lambda^i C^i) \equiv \delta^2 F \,. \tag{16}$$

This $\Delta H|_C$ is second order in δz, as expected for a Hamiltonian near an equilibrium point. Since $\Delta H|_C$ is also a constant of motion it can serve as a Liapunov functional. Moreover, $\Delta H|_C$ has physical significance: it is the energy increment contained in a perturbation which is accessible to the dynamics given the Casimir constraints. Note that $\delta^2 F$ depends on the equilibrium through the choice of λ^i, so the same δz contains different accessible energies for each equilibrium. This is intuitively satisfying since one would expect the energetic favorability of δz to depend upon the equilibrium. For example, in MHD the energetic favorability and stability of a kink mode perturbation depends on the equilibrium current.

The expression $\delta^2 F$ is a most convenient way to to compute $\Delta H|_C$. The reason for this is that accessible perturbations, δz, must satisfy $\Delta C = 0$. To compute ΔH to second order directly requires $\Delta C = 0$ to second order. It is relatively easy (often trivial) to find trial functions δz_T that satisfy $\Delta C = 0$ to first order in δz. In general a variation δz satisifying $\Delta C = 0$ to higher order would require a calculation to obtain $\delta z = \delta z_T + O(\delta z_T^2)$. This is usually difficult to satisfy for all Casimir invariants simultaneously. However, $\Delta H|_C$ can be computed to second order accuracy by substituting δz_T directly into $\delta^2 F$. We do this for the Vlasov equation in Sec. 4.

It is easy to show that $\delta^2 F$ is the Hamiltonian for the linearized equations of motion[29,6], but more importantly one can show that the work performed by an external agent, which observes the constraints while creating a linear perturbation, is $\delta^2 F$. This is the usual definition of the energy content of a wave in dielectric theory. Suppose an external source is applied to the system which adds a transient term H_{ext} to the Hamiltonian for times between 0 and t_0. The J^{ij} is unchanged by this; thus the Casimir invariants (which depend only on J^{ij}) remain constant. For linear perturbations the new

Hamiltonian is

$$H' = \delta^2 F + H_{ext} \cdot$$

A convenient form for H_{ext} is $z^i s_j(t)$, the explicit time dependence appearing only in the source functions s_j. Standard Hamiltonian mechanics yields

$$\frac{dH'}{dt} = \frac{\partial H}{\partial t} = z^j \frac{\partial s_j}{\partial t} \; . \tag{17}$$

Integrating Eq. (17) in time from $t = 0$ to $t = t_0$ yields, upon integration by parts on the right, and the assumption that $\mathbf{s} = 0$ for $t \leq 0$ and $t \geq t_0$,

$$\delta^2 F = -\int_0^{t_0} (dz^j/dt)\, s^j dt \; . \tag{18}$$

The right side is the usual expression for the work done on a system, i.e. the integrated input power, by an external agent.

As an example, consider the one-dimensional Vlasov-Poisson equation. The external agent in this case is an externally imposed potential φ_{ext} and the external portion of the Hamiliton is

$$H_{ext} = \int f \, \varphi_{ext} \, dx \, dv \, .$$

Here $f(x,v,t)$ is the phase space density. In this case we have for the power

$$\begin{aligned} (dz^j/dt)\, s^j &= \int (\partial f/\partial t)\, \varphi_{ext} \, dx \, dv = \int (\partial \rho/\partial t)\, \varphi_{ext} \, dx \\ &= \int \mathbf{J} \cdot \nabla \varphi_{ext} \, dx \; . \end{aligned} \tag{19}$$

The last equality of Eq. (19) follows by making use of the continuity equation, $\partial \rho/\partial t + \nabla \cdot \mathbf{J} = 0$, and integrating by parts. This expression is the standard one for the power input from an external field. Thus $\delta^2 F$ is indeed the energy needed to create a perturbation δf.

In dielectric theory the energy content[1-3] in a linear wave is defined by the work performed by an external agent in producing the wave. The energy content is found to be $[\partial(\omega\varepsilon)/\partial\omega]\,|E_k|^2$, where ω is the wave frequency, k is the mode number, $\varepsilon(k,\omega)$ is the dielectric function, and E_k is the electric field amplitude for mode k. This brings us to the important topic of positive and negative energy waves; in dielectric theory these have $\partial(\omega\varepsilon)/\partial\omega > 0$ and $\partial(\omega\varepsilon)/\partial\omega < 0$ respectively. As noted in the Introduction, negative energy waves have the property that they are spectrally stable, but their presence decreases the energy of the medium. If both types of waves exist in the medium then δ^2F is positive for some perturbations, δz, and negative for others, and is thus indefinite.

Now suppose that δ^2F is indefinite. What conclusions can be drawn about the spectral properties of the system? For finite degree-of-freedom systems there are two possibilities: either there is spectral instability or the system has a mixture of positive and negative energy waves.

For the case of canonical Hamiltonian systems it is straight forward to show this assertion. The theory of normal forms by now is well understood[5]. In this case δ^2F is the linearized Hamiltonian, i.e. the second variation of the full Hamiltonian. It is a quadratic form in δq_i and δp_i. We will consider the case where H has the following form:

$$H = \sum_i \alpha_i(\delta p_i)^2 + \beta_i(\delta q_i)^2 \quad . \tag{20}$$

If the H of Eq. (20) is indefinite, then for a least one degree-of-freedom one of the following holds: (i) α_i has a different sign than β_i, which corresponds to instability, or (ii) the pair of α_i and β_i has a different sign from other pairs. The latter case corresponds to a mixture of positive and negative energy waves. It is evident from the discussion of Sec. 1.2 that the same results apply for finite noncanonical Hamiltonian systems.

It is clear from the above that there is no such thing as a spectrally unstable negative energy mode. One might think this would correspond to a negative α_i and a positive β_i, but since $q_i \to -p_i$ and $p_i \to q_i$ is a canonical transformation we see that there is no distinction between this case and that of instability.

However, we have conjectured that a system with both positive and negative energy waves is nonlinearly unstable in the general case. This can arise because there is a low order resonance in the system, which results in explosive instability, i.e. divergence in finite time (c.f. Sec. 3), as occurs in the well-known three-wave interaction. For two degree-of-freedom systems where no low order resonance exists, the KAM theorem indicates stability. For systems of more than two degrees-of-freedom, instability can still be present by the slow mechanism known as Arnold diffusion. For infinite degree-of-freedom systems with continuous spectra, there may be additional avenues for nonlinear instability.

There are many continuum systems for which the concept of a dielectric function is not well defined, or if defined in principle, is very difficult to calculate in practice. However, $\delta^2 F$ can be easily computed for such systems. Thus $\delta^2 F$ provides a practical generalized definition of the concept of a negative energy wave. It can easily be generalized to the concept of non-wave-like phenomena.

Before closing this section, let us consider a general bifurcation property concerning negative energy modes. Suppose we have a sequence of noncanonical Hamiltonian equilibria parameterized by a continuous variable η. Assume that $\delta^2 F$ is positive definite for $\eta < \eta_0$, but indefinite for $\eta > \eta_0$, because a positive energy mode becomes a negative energy mode (the system being spectrally stable on both sides of the threshold). We now show that the frequency of such a mode must go through zero at $\eta = \eta_0$, at least for finite dimensional systems.

The tensor $\partial^2 F/\partial z^i \partial z^j$ is symmetric, and for $\eta < \eta_0$ positive definiteness implies it has all positive eigenvalues. At least one eigenvalue becomes negative for $\eta > \eta_0$, and thus traverses zero at $\eta = \eta_0$. The associated eigenvector, $\delta z_0{}^i$, is also a zero frequency eigenvector for the dynamical system; i.e. the existence of a zero eigenvalue of $\partial^2 F/\partial z^i \partial z^j$ implies the existence of a zero frequency mode. This follows from linearization of Eq. (12) about an equilibrium z_e,

$$i\omega \delta z_0{}^i = J^{ij} \frac{\partial^2 F}{\partial z^j \partial z^k} \delta z_0{}^k = 0 , \tag{21}$$

where J^{ij} and $\partial^2 F/\partial z^j \partial z^k$ are evaluated at z_e.

Note that this result can also be derived from the definition of wave energy density, which as noted above has its sign determined by the factor $\partial(\omega\varepsilon)/\partial\omega$. Since the wave frequency satisifies $\varepsilon = 0$ this factor becomes $\omega\partial\varepsilon/\partial\omega$. Thus as the wave energy goes from positive to negative at $\eta = \eta_0$ we have two possibilities: either ω goes through zero or $\partial\varepsilon/\partial\omega$ does. The latter possibility can be excluded by showing it leads to a contradiction. Assume $\varepsilon = 0$ and $\partial\varepsilon/\partial\omega = 0$ at $\omega = \omega_0 \neq 0$. Taylor expanding ε about this point yields

$$\varepsilon = \sum_{n=2}^{\infty} \frac{(\omega - \omega_0)^n}{n!} \frac{\partial^n \varepsilon}{\partial\omega^n} + (\eta - \eta_0)\frac{\partial\varepsilon}{\partial\eta} + \ldots \tag{22}$$

Since $\partial\varepsilon/\partial\omega$ is assumed to change sign at $\eta = \eta_0$, one can show that the first nonvanishing term on the Taylor series must have even n. One can also show that if $\partial\varepsilon/\partial\eta \neq 0$, then the cross terms neglected in Eq. (22) are asymptotically negligible compared with those kept when $\omega \approx \omega_0$ and $\eta \approx \eta_0$.

The condition $\varepsilon = 0$ allows us to solve for the mode frequency near ω_0 by keeping the first nonvanishing term in the series; hence

$$(\omega - \omega_0)^n = n!(\eta - \eta_0)\frac{(\partial\varepsilon/\partial\eta)}{(\partial^n\varepsilon/\partial\omega^n)} . \tag{23}$$

Since n is even, this implies that instabilities must exist for either $\eta < \eta_0$ or $\eta > \eta_0$, depending on the signs of the ε partial derivatives.

Thus, as equilibrium parameters are varied, equilibria with only positive energy modes can only acquire negative energy modes (which are by definition stable) when the mode frequency passes through zero. Furthermore, the δ^2F and dielectric definitions of energy agree on the value of η for which this occurs.

3. A MODEL OF FLR STABILIZATION

We consider a model that has a general form that occurs in models of finite Larmor radius stabilization. This model is canonical in nature and displays the form of negative energy modes discussed in Sec. 2 for finite degree-of-freedom systems. The

Lagrangian for a particle in a uniform magnetic field B in the z-direction, subject to a potential which to leading order is harmonic and inverted, is

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{eB}{2c}(\dot{y}x - \dot{x}y) + \frac{k}{2}(x^2 + y^2) - V_3, \qquad (24)$$

where $k > 0$ and V_3 is an anharmonic contribution to the potential that is a cubic in x and y. The inverted harmonic potential could arise because of MHD equilibrium[16] or one could be interested in the motion of particles in a uniformly charged cylindrical column.

Upon Legendre transforming (21) and writing the resulting Hamiltonian in terms of the following canonical variables:

$$q_1 = x/\sqrt{m} \qquad\qquad q_2 = y/\sqrt{m}$$
$$p_1 = m^{3/2}(\dot{x} - \omega_L y) \qquad\qquad p_2 = m^{3/2}(\dot{y} + \omega_L x) \qquad (25)$$

yields

$$H = \frac{1}{2}(p_1{}^2 + p_2{}^2) + \omega_L(q_2 p_1 - q_1 p_2)$$
$$+ \frac{1}{2}(\omega_L{}^2 - \omega_0{}^2)(q_1{}^2 + q_2{}^2) + V_3, \qquad (26)$$

where ω_L is the Larmor frequency, $eB/2mc$, and ω_0 is the natural frequency of the harmonic potential, $\sqrt{k/m}$.

Consider now the spectrum of the equilibrium point $q_1 = q_2 = p_1 = p_2 = 0$. Neglecting V_3 and supposing that the dynamical variables vary as $\exp(i\omega t)$, we obtain a fourth-order characteristic polynomial that has the roots

$$\omega = \pm\,\omega_0[\sqrt{\eta - 1} \pm \sqrt{\eta}\,], \qquad (27)$$

where $\eta = \omega_L{}^2/\omega_0{}^2$. Equation (24) displays the symmetries of the spectrum of a Hamiltonian system; i.e. pure real and pure imaginary frequencies appear in pairs, while in general frequencies occur in quartets of the form $\pm a \pm i b$.

Let us trace the frequencies in the complex ω-plane as we increase the magnetic field away from zero. At $\eta = 0$ we have two doubly degenerate pure imaginary roots, $\omega = \pm i\omega_0$. It is not surprising that there is instability in this case since the particle is free to fall off the hill. Upon increasing η the degenerate eigenvalues split and move respectively clockwise and counterclockwise on a circle of radius ω_0 toward the ω_r axis. When $\eta = 1$, $\omega = \pm\omega_0$ and the eigenvalues again become doubly degenerate and experience a Krein crash (See Refs. [4] and [5]), with one pair of real eigenvalues approaching the origin and one pair receding. After the Krein crash the eigenvalues have magnitude

$$\omega_1 = -\omega_L + \sqrt{\omega_L^2 - \omega_0^2}$$

$$\omega_2 = \omega_L + \sqrt{\omega_L^2 - \omega_0^2}\,. \qquad (28)$$

Observe that $0 < \omega_1 \leq \omega_L$ and $\omega_2 \geq \omega_L$. Subsequently we will see that ω_1, the smaller frequency, is a negative energy mode. It has negative signature while ω_2 has positive signature, as required by Krein's theorem.

Consider now the canonical transformation generated by the following mixed variable generating function:

$$F_2(q_1, q_2, P_1, P_2) = \mu(q_1 P_1 + q_2 P_2) + P_1 P_2 + \mu^2 q_1 q_2/2, \qquad (29)$$

where $\mu = [4(\omega_L^2 - \omega_0^2)]^{1/4}$. In terms of the new canonical variables P_1, P_2, Q, and Q_2, the Hamiltonian becomes

$$H = -\frac{1}{2}\omega_1 (P_1^2 + Q_1^2) + \frac{1}{2}\omega_2 (P_2^2 + Q_2^2) + V_3\,, \qquad (30)$$

where V_3 is now a cubic function of the new coordinates and momenta. We see from Eq. (27) that this Hamiltonian has the standard form for a negative energy mode.

Let us now see what happens to the spectrum when we add negative definite dissipation to this system; i.e. dissipation that removes energy. This will happen if we dissipate the negative energy mode. Since V_3 does not effect the spectrum we can consider the negative energy oscillator independently. With the inclusion of dissipation this oscillator is governed by

$$\dot{Q}_1 = -\omega_1 P_1$$
$$\dot{P}_1 = \omega_1 Q_1 + 2\upsilon P_1 \,. \tag{31}$$

To see that the term involving υ corresponds to dissipation note that, for $\upsilon > 0$, $dH/dt = -2\omega_1 \upsilon P_1^2 \leq 0$. The eigenfrequency corresponding to this oscillator becomes

$$\omega = -i\upsilon \pm \sqrt{\omega_1^2 - \upsilon^2}\,, \tag{32}$$

which corresponds to exponential growth for $\upsilon > 0$. Thus dissipation destabilizes the negative energy modes.

Now consider the other way that negative energy modes lead to instability; i.e. nonlinear destabilization. In the case where there is a resonance this phenomenon is tractable. Let us suppose an order three resonance condition: $2\omega_1 = \omega_2$. This occurs when $3\omega_0 = 2\sqrt{2}\omega_L$. The important terms of V_3 are those that drive the resonance. These can be obtained by the standard Hamiltonian perturbation method of averaging. This procedure yields

$$H = \frac{1}{2}\omega_1 (P_1^2 + Q_1^2) - \omega_1 (P_2^2 + Q_2^2) + \frac{\alpha}{2}[Q_2(Q_1^2 - P_1^2) - 2Q_1 P_1 P_2]\,, \tag{33}$$

where α is a constant. We have reversed the sign of time in order to show that the Hamiltonian of Eq. (33) is that due to Cherry[14], who observed that a two-parameter solution set is given by

$$Q_1 = \frac{\sqrt{2}}{\alpha(\omega_1 t - \varepsilon)}\sin(\omega_1 t + \gamma)$$

$$P_1 = \frac{-\sqrt{2}}{\alpha(\omega_1 t - \varepsilon)}\cos(\omega_1 t + \gamma)$$

$$Q_2 = \frac{1}{\alpha(\omega_1 t - \varepsilon)}\sin(2\omega_1 t + \gamma)$$

$$P_2 = \frac{-1}{\alpha(\omega_1 t - \varepsilon)}\cos(2\omega_1 t + \gamma)\,, \tag{34}$$

where ε and γ are determined by the initial conditions. The interesting thing about this solution is that any neighborhood of the origin contains initial conditions of solutions that diverge in finite time. Thus this system is spectrally stable, but in reality unstable. The nonlinearity diverts energy from the negative energy mode to the positive energy mode.

In the case where there is no resonance and the system possesses a general cubic term with explicit time dependence the system is most likely unstable by the mechanism of Arnold diffusion. Preliminary numerical calculations show that the system Arnold diffuses until a separatrix is reached, at which point the growth is rapid.

4. VLASOV EQUILIBRIA

In this section we discuss the free energy for the Vlasov equation. We begin by reviewing in some detail the noncanonical Hamiltonian formalism for the one-dimensional Vlasov Poisson system[30] and its associated variational principle for equilibria. From this $\delta^2 F$ is obtained, but within this formalism only monotonic and isotropic equilibria are obtained as extremals of the variational principle, and only sufficient conditions for the stability of this limited class of equilibria are obtained. In recent work expressions for $\delta^2 F$ that generalize to arbitrary equilibria, have been obtained. The result for the Vlasov-Poisson equation is stated here; the reader is referred to the references for its derivation, along with the general Maxwell-Vlasov result.

The Vlasov-Poisson equation is

$$\frac{\partial f(z,t)}{\partial t} = -v\frac{\partial f}{\partial x} + \frac{e}{m}\frac{\partial \varphi(x;f)}{\partial x}\frac{\partial f}{\partial v} \tag{35}$$

where, as usual, $f(z,t)$ is the phase space density at the phase space point $z=(x,v)$ for a charged species of particles with charge e and mass m. Now we consider only a single species, but the results are readily generalizable. The electrostatic potential φ is to be viewed as a functional of f determined via Poisson's equation $\varphi_{xx} = -e\int f\,dz$; thus $\varphi(x;f) = e\int V(z,\tilde{z})\, f(\tilde{z})\, d\tilde{z}$, where $V(x,\tilde{x})$ is the single particle potential (assumed spatially invariant). The Hamiltonian for this system is the energy functional

$$H[f] = \int T(z)\, f(z)\, dz + \frac{e^2}{2} \int\int V(z,\tilde{z})\, f(z)\, f(\tilde{z})\, d\tilde{z}\, dz, \quad (36)$$

where $T(z) = mv^2/2$ is the particle kinetic energy. This system possesses the following noncanonical Poisson bracket:

$$\{F,G\} = \frac{1}{m} \int f(z) \left[\frac{\delta F}{\delta f}, \frac{\delta G}{\delta f} \right] dz \quad (37)$$

where the inner bracket [,] is defined by $[k,h] = k_x h_v - k_v h_x$. Note that $\delta H/\delta f = T + e\varphi \equiv E$, where E is the total particle energy. Evidently,

$$\frac{\partial f}{\partial t} = \{f,H\} = -[f,E] \quad (38)$$

where $-[f,E]$ is equivalent to the right hand side of Eq. (35).

As discussed in Sec. 1.2 stationary points of the free energy $F = H + C$ are equilibria. In the case of the Vlasov-Poisson equation, H is given by Eq. (36) and C is the well-known Liouville constraint, which is given by the following:

$$C[f] = \int c(f)\, dz\,, \quad (39)$$

where $c(f)$ is an arbitrary function. Upon varying F we obtain

$$\begin{aligned} \delta F[f;\delta f] &= \int (E + \partial c/\partial f)\, \delta f\, dz \\ &= \int (mv^2/2 + e\varphi + \partial c/\partial f)\, \delta f\, dz\,. \end{aligned} \quad (40)$$

Thus equilibria (f_e) are given by

$$E + \partial c(f)/\partial f = 0\,. \quad (41)$$

There are two things to notice about these equilibria that are obtained as extremals of F: firstly, in order to solve for $f_e(E)$, the quantity $\partial c/\partial f$ must be monotonic and therefore its inverse must also be monotonic. This gives

$$f_e = f_e(E) \qquad , \tag{42}$$

where $f_e(E)$ is a monotonic function of E, and we see that there is a one-to-one correspondence between an equilibrium and a choice for the function c. Secondly, since f_e is a function of E, only velocity symmetric distributions $[f_e(v) = f_e(-v)]$ are obtained. Evidently, extremals of F only make up a subclass of equilibria of the Vlasov-Poisson equation, since this system is known to possess nonmonotonic and velocity asymmetric (for untrapped particles) equilibria.

For the above restricted class of equilibria we can obtain a criterion for stability[31-34] by taking the second variation of F; viz,

$$\delta^2F[f] = \frac{1}{2}\left(e^2\int\int V(z,\tilde{z})\,\delta f(z)\,\delta f(\tilde{z})\,d\tilde{z}\,dz + \int \partial^2 c/\partial f^2\,(\delta f)^2\,dz\right) , \tag{43}$$

Observe that the first term of Eq. (43) is positive definite (it corresponds to the second variation of the electrostatic energy which goes as the square of the electric field), while the second term will be positive definite provided $\partial^2c/\partial f^2 > 0$. For stability this must be true over the entire domain of integration when f is set equal to f_e, since we can make δ^2F negative by choosing δf such that the first term of Eq. (43) vanishes and such that δf is localized where $\partial^2c/\partial f^2 < 0$. This statement translates into a statement about f_e: upon differentiating Eq. (41) with respect to E we obtain

$$\partial f_e/\partial E = -[\partial^2 c(f_e)/\partial f^2]^{-1}. \tag{44}$$

Therefore we have stability if f_e is <u>any</u> monotonic decreasing function of the energy.

Note that if we attempted to apply the formula of Eq. (43) to nonmonotonic equilibria then δ^2F diverges unless δf vanishes at places where $\partial f_e/\partial E = 0$. If we restrict δf to the Casimir surfaces then this problem is avoided. In the notation of Sec. 2

the δz_T is given by

$$\delta f = [f_e, g] = (\partial f_e/\partial E)[E,g] \ , \tag{45}$$

where g is an arbitrary function. One can show that δf as given by Eq. (45) preserves the Casimir constraint to first order. Inserting Eq. (45) into Eq. (43), upon making use of Eq. (44), yields

$$\delta^2 F[f] = \frac{1}{2}\left(e^2 \int\int V(z,\tilde{z})\,\delta f(z)\,\delta f(\tilde{z})\,d^3\tilde{z}\,d^3z + \int [g,f_e][E,g]\,d^3z\right) . \tag{46}$$

Observe that the monotonicity condition for stability is replaced by positivity of the second integrand, for all g. When this integrand is not positive definite, and cannot be made so by a frame change, then there exist negative energy modes when there is linear stability. The result of Eq. (46) was derived, within the Lagrangian variable context, in Refs. [9]. It is applicable to arbitrary Vlasov-Poisson equilibria. The use of Lagrangian variables formally circumvents the limitation associated with the Eulerian description. Also in Refs. [9] the free energy expression for arbitrary Maxwell-Vlasov equilibria is given.

ACKNOWLEDGMENT

This work was supported by the U.S. Department of Energy contract #DE-FG05-80ET-53088.

REFERENCES

1. von Laue, M., Ann. Physik 18, 523 (1905)
2. Brillouin, L., Wave Propagation and Group Velocity, (Academic Press, New York ,1960).
3. Landau, L. D. and Lifshitz, E. M., Electrodynamics of Continuous Media, (Pergmon Press, Reading, 1960) p. 253.
4. Moser, J. K., Comm. Pure Appl. Math. 11, 81 (1958); Mem. Am. Math. Soc. 81, 1 (1968).

5. Arnold, V. I., Kozlov, V. V. and Neishtadt, A. I., in Encyclopedia of Mathematical Sciences 3, Dynamical Systems III, Ed. Arnold, V. I. (Springer-Verlag, Berlin,1988).
6. Morrison, P. J. and Eliezer, S., Phys. Rev. A 33, 4205 (1986).
7. Kotschenreuther, M., *et al.*, Plasma Physics and Controlled Nucl. Fusion Res. 2, 149 (1987) (International Atomic Energy Agency Publication).
8. Morrison, P. J., Z. Naturforsch. 42a, 1115 (1987).
9. Morrison, P. J. and Pfirsch, D., Phys. Rev. A 40, 3898 (1989); Institute for Fusion Studies, Austin, Texas, Report, IFSR #401, Oct. 1989. Submitted to Physics of Fluids.
10. Sturrock, P. A., J. Appl. Phys. 31, 2052 (1960).
11. Dikasov, V. M., Rudakov, L. I., and Ryutov, D. D., Zh. Eksperim. i Teor. Fiz. 48, 913 (1965). [English trans. Soviet Phys. -JETP 21, 608 (1966).]
12. Aamodt, R. E. and Sloan, M. L., Phys. Rev. Lett. 19, 1227 (1967).
13. Sagdeev, R. Z. and Galeev, A. A., in T. M. O'Niel and D. L. Book (eds.), Nonlinear Plasma Theory, (W. A. Benjamin, New York, 1969).
14. Coppi, B., Rosenbluth, M. N., and Sudan, R., Ann. Phys. 55, 207 (1968).
15. Cherry, T. M., Trans. Cambridge Philos. Soc. 23, 199 (1925); See Whittaker, E. T., Analytical Dynamics (Cambridge, London, 1937), Sec. 182, p. 142; and Wintner, A.,The Analytical Foundations of Celestial Mechanics (Princeton University, Princeton, 1947), Sec. 136, p. 101.
16. Thompson, W. and Tait, P. G., Treatise on Natural Philosophy, part 1, (Cambridge University Press, Campridge,!921), p. 388.
17. Greene, J. M. and Coppi, B., Phys. Fluids 8, 1745 (1965).
18. Newcomb, W. A., Nuclear Fusion: Supp. part 2, 451 (1962).
19. VanKampen, N. G. and Felderhoff B. U., Theoretical Methods in Plasma Physics (North-Holland, Amsterdam, 1967).
20. Low, F. E., Proc. Roy. Soc. 248A, 282 (1958).
21. This formalism has roots that trace back to Sophus Lie, see for instance, Caratheodory, C., Calculus of Variations and Paritial Differential Equations of the First Order (Holden-Day, San Francisco, 1965).
22. Pauli, W., Il Nouvo Cimento 10, 648 (1953).
23. Martin, J. L., Proc. Roy. Soc. 251A, 536 (1959).
24. Arnold, V.I., Inst. Fourier Grenoble **16**, 319 (1966).
25. Morrison P.J. and Greene, J.M., Phys. Rev. Lett. 45, 790 (1980).

26. Morrison, P.J., in Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems, AIP Conf.. Proc. 88, 13 (1982).

27. Sudarshan, E.C.G. and Mukunda, N. Classical Mechanics - a Modern Perspective, (Krieger, Melbourne, Florida, 1983) Second Edition.

28. Arnold, V. I., Ann. Math. Soc. Transl. **79**, 267 (1969).

29. In the following reference many sufficient conditions for stability were reviewed and presented. These expressions, when restricted to the constraint surfaces, correspond to the free energy. Holm, D. D., Marsden, J. E., Ratiu, T. and Weinstein, A., Phys.Reps.123, 1 (1985).

30. Morrison, P.J., Phys. Lett. 80A, 383 (1980).

31. Newcomb, W. appendix to I. B. Bernstein, Physical Review **109**, 10 (1958).

32. Kruskal, M. D. and Oberman, C., Physics of Fluids **1**, 275 (1958).

33. Gardner, C. S., Phys. Fluids **6**, 839 (1963).

34. Fowler, T. K., J. Math. Phys. **4**, 559 (1963); also paper in A. Simon and W. B. Thompson (eds.), Advances in Plasma Physics, (Wiley, New York, 1968) vol.1, p. 201.

INSTITUTE FOR FUSION STUDIES
THE UNIVERSITY OF TEXAS AT AUSTIN
Robert Lee Moore Hall · Austin, Texas 78712-1060 · (512) 471-1322

airmail - 1st class

October 26, 1989

Prof. E.A. Kuznetsov
Inst. of Automatics and Electrometry
630090, Novosibirsk
University Avenue, 1
U.S.S.R.

Dear Zhenya:

Enclosed is a new original and two (2) copies of my manuscript for the NTPP proceedings that will be published by World Scientific. After returning home, I discovered a few typographical errors, which have been corrected.

Best wishes,

Phil

Philip J. Morrison
Prof. of Physics

encl: 1 original and 2 copies manuscript

suz
p.s. I will send, under separate cover, the papers pertinent to my talk that you requested.

I. INTRODUCTION

1.1. Overview

This paper is concerned with instability of equilibria of Hamiltonian, fluid and plasma dynamical systems. Usually the dynamical equilibrium of interest is not the state of thermodynamic equilibrium, and does not correspond to a free energy minimum. The relaxation of this type of equilibrium is conventionally considered to be initiated by linear instability. However, there are many cases where linear instability is not present, but the equilibrium is nonlinearly unstable to arbitrarily small perturbations. This paper is about general free energy expressions for determining the presence of linear or nonlinear instabilities. These expressions are simple and practical, and can be obtained for all equilibria of all ideal fluid and plasma models. By free energy, we mean the energy change upon perturbations of the equilibrium that respect dynamical phase space constraints. This quantity is measured by a self-adjoint quadratic form, called $\delta^2 F$. The free energy can result in instability when $\delta^2 F$ is indefinite; i.e. there exist accessible perturbations that lower the free energy of the system.

A primary purpose of this paper is to tie together three manifestations of what we will refer to as *negative energy modes*. The first is the conventional plasma physics notion of negative energy mode that is based on the definition of the energy in a homogeneous dielectric medium[1-3]. A negative energy mode is a normal mode of the medium (plasma) that possesses negative dielectric energy. The second manifestation occurs in finite degree-of-freedom Hamiltonian normal form theory[4,5]. The quadratic part of a Hamiltonian in the vicinity of an equilibrium point, which possesses only distinct oscillatory eigenvalues, has an invariant signature. Thus in cases where the quadratic form is indefinite, it is natural to refer to the modes corresponding to the negative signature as negative energy modes. The third, and most general, definition of a negative energy mode[6-9] relies on the free energy functional that we have termed $\delta^2 F$. An equilibrium has a negative energy mode if it is linearly stable and has indefinite $\delta^2 F$, and further we require this to be true in all frames of reference. This later proviso is required since energy is not a covariant quantity.

The definition of negative energy modes based on $\delta^2 F$ is a generalization of the dielectric definition in that it is applicable to arbitrary equilibria. It does not require Fourier transformation in space or time, nor does it require the existence of a dielectric

Except for minor modifications, the first three sections of this paper were completed in 1986. This material was presented at several meetings, including the following invited presentations: American Physical Society Division of Plasma Physics Meeting, November 1986, Baltimore, Maryland, Bull. Am. Phys. **31**, 1609 (1986); Currents in Geophysics and Hydrodynamics, University of California, June 1986, Lake Arrowhead, CA; International Atomic Energy Agency Meeting on Controlled Thermonuclear Research, November 1986, Kyoto, Japan, [c.f. Ref. 7]. Section 4 contains more recent material, as documented in Refs. 8 and 9.

function. The δ^2F expressions do not contain unknown frequency dependence and in general are easier to use. The δ^2F definition obviously generalizes the normal form definition since the systems of interest are generally partial differential equations. We point out, however, that it includes the previous two. All three definitions are equivalent for systems with a complete discrete spectrum and a notion of dielectric function, such as the fluid description of the beam-plasma instability.

A second primary purpose of this paper is to describe a conjecture that we refer to as the *free energy principle*, a principle that yields criteria which are sufficient and, in a sense, necessary for stability. When δ^2F is definite an equilibrium is stable. When δ^2F is indefinite and we are in a reference frame where the equilibrium has minimum energy, there are two possibilities: (1) the equilibrium possesses linear instability or (2) there is linear stability with negative energy modes. Possibility (2) is a restatement of a definition, but in this case there are two avenues for instability: (a) the equilibrium can be unstable to an arbitrarily small perturbation because of nonlinearity or (b) the equilibrium can be structurally unstable to the inclusion of dissipation in the dynamical model. There are many examples that illustrate (a), in particular, a simple example due to Cherry[10] (which will surface in Sec. 3) and important calculations of nonlinear electrostatic instabilities in homogeneous plasma[11-15]. Illustrations of (b) include the well-known Thompson-Tait theorem[16], and we point out an early plasma physics example[17]. It is conjectured that (a) and (b) are generic; i.e. although there exist cases where δ^2F is indefinite and the system is stable, and there exist special types of dissipation that do not result in instability, these are conjectured to be exceptional. If we accept the conjecture and ignore these possibilities because of their rarity, we obtain a sense in which δ^2F provides a "necessary" and sufficient condition for stability. This is what is meant by the free energy principle.

The fluid and plasma literature contains a large number of sufficient conditions for the stability of ideal equilibria that depend upon the positive definiteness of some quadratic form. These criteria are often obtained by ad hoc means and sometimes are devoid of physical interpretation. Generally these conditions, unlike the magnetohydrodynamic (MHD) energy principle, are believed to yield no information when indefinite. In fact a great deal of effort has been spent in the unsuccessful attempt to obtain necessary and sufficient conditions for the stability of general equilibria. In all cases known to the authors, definiteness of these forms amounts to definiteness of the free energy. Thus the notion of δ^2F greatly clarifies the situation in that there is a framework for interpreting and constructing these criteria. Also it puts to rest the question of necessary and sufficient conditions. The best to be hoped for in general is the "necessary" condition described above, because in general, Hamiltonian systems do not possess δW type criteria.

In the remainder of the Introduction (Sec. 1.2) we discuss two descriptions of equations that describe ideal plasma, the Lagrangian variable description, which possesses an action principle[18-20], and the Eulerian variable description, which posseses the noncanonical Hamiltonian structure[21-29]. Both the Lagrangian[9] and Hamiltonian[5-7,29] formalisms have been used to obtain free energy expressions. We will dwell somewhat on the later, but quote results obtained from the former. Section 3 contains an example that typifies finite Larmour radius stabilization. Here we see in a simple system of ordinary differential equations, how a system can be linearly stable, yet unstable, and how the inclusion of dissipation can result in linear instability by shifting a real frequency into the complex plane. In Sec. 4 the Vlasov-Poisson system is treated. A general expression for the free energy of arbitrary equilibria is stated.

1.2 Review

In the Lagrangian variable description of continuous media, for example fluids, the complete state of a system is determined by a continuum of "particle" positions, $\mathbf{x}(\mathbf{x}_0,t)$, where $\mathbf{x}_0$ is the particle position at $t=0$. The Lagrangian variable description is model independent in the sense that it arises in kinetic[20] as well as fluid systems[18,19] with a generality of force laws. In the case of fluid systems, $\mathbf{x}$ corresponds to the position of a "fluid element". This case differs from the usual description of a fluid in terms of Eulerian variables, where the density and velocity fields are expressed as functions of space and time. In contrast to the Eulerian variable description, when one describes a fluid as a collection of particles, as in the Lagrangian variable description, it is not surprising that the equations of motion possess the form of Newton's second law, and therefore are derivable from Hamilton's principle of mechanics.

As an example consider the nonlinear equations of ideal magnetohydrodynamics (MHD), which have the following form:

$$\frac{\partial^2 \mathbf{x}}{\partial t^2} = -\frac{\delta W[\mathbf{x}]}{\delta \mathbf{x}} \quad , \tag{1}$$

where the right hand side is the functional or variational derivative of a potential energy functional, $W[\mathbf{x}]$. The MHD energy principle exists because of the Hamiltonian form of Eq. (1), for if we linearize by letting $\mathbf{x} = \mathbf{x}_e(\mathbf{x}_0) + \boldsymbol{\xi}(\mathbf{x}_0,t)$, where $\mathbf{x}_e$ satisfies $\delta W[\mathbf{x}_e]/\delta \mathbf{x} = 0$, then $\boldsymbol{\xi}$ satisfies

$$\frac{\partial^2 \boldsymbol{\xi}}{\partial t^2} = O\boldsymbol{\xi} \quad , \tag{2}$$

where the linear operator O arises from the second variation of $W[\mathbf{x}]$,

$$\delta^2 W = \frac{1}{2} \int \boldsymbol{\xi} \cdot \frac{\delta^2 W[\mathbf{x}_e]}{\delta\boldsymbol{\xi}\delta\boldsymbol{\xi}} \cdot \boldsymbol{\xi} \; d^3x = \frac{1}{2} \int \boldsymbol{\xi} \cdot O \cdot \boldsymbol{\xi} \; d^3x \; . \tag{3}$$

From Eq. (3) it is evident that O is self-adjoint since it is the second variation of W. (The quantity $\delta^2 W$ is usually misleadingly referred to as δW.) This self-adjointness, with its root traced back to the Hamiltonian form of the original nonlinear system, is the crucial element that gives rise to the necessary and sufficient δW energy criteria for static MHD equilibria.

Consider now the Eulerian variable description. One thing that is immediately evident is that there is a difference between the classes of Eulerian and Lagrangian equilibria. For example, in ideal MHD the dynamical variables are the velocity field $\mathbf{v}(\mathbf{x},t)$, the density $\rho(\mathbf{x},t)$, the magnetic field $\mathbf{B}(\mathbf{x},t)$, and the entropy per unit mass $s(\mathbf{x},t)$. If we define equilibria by the vanishing of the first time derivative of the dynamical variables, then Eulerian equilibria correspond to some (not all) functions of space. In particular note that the equilibrium velocity field, $\mathbf{v}_e(\mathbf{x})$, need not be constant nor vanish. This should be contrasted to Lagrangian equilibria where $\partial\boldsymbol{\xi}/\partial t = 0$, which implies that there is no equilibrium flow. Thus we see that Lagrangian equilibria are static equilibria, while the class of Eulerian equilibria includes stationary equilibria as well.

Another difference between the Eulerian and Lagrangian variable descriptions is the apparent difference in the number of dynamical variables needed to specify the state of the system. The map from Lagrangian to Eulerian variables is not a one-to-one transformation, and evidently it is not a canonical transformation. In spite of the "pathology" of this map, the Hamiltonian structure survives, albeit in what we have called the noncanonical Hamiltonian form.

For simplicity we discuss a finite degree-of-freedom Hamiltonian system, which for ease in generalizing to the noncanonical description we write in terms of the 2N dynamical variables z^i, the first N of which are the canonical coordinates while the second N are the conjugate momenta. Hamilton's canonical equations take the form

$$\frac{dz^i}{dt} = [z^i, H] = J_c^{ij} \frac{\partial H}{\partial z^j} \quad , \tag{4}$$

where the Poisson bracket is defined by

$$[f,g] = \frac{\partial f}{\partial z^i} J_c^{ij} \frac{\partial g}{\partial z^i} \tag{5}$$

with

$$(J_c^{ij}) = \begin{bmatrix} 0 & I_N \\ -I_N & 0 \end{bmatrix} \quad , \tag{6}$$

which is a $2N \times 2N$ matrix and I_N is the $N \times N$ unit matrix. Here repeated indices are to be summed to 2N. The quantity (J_c^{ij}) is a second order contravariant tensor that is called the cosymplectic form.

Canonical transformations, by definition, preserve the form of the Poisson bracket, or equivalently the form of the tensor (J_c^{ij}). This is not the case for the transformation between Lagrangian and Eulerian variables, since this transformation is noncanonical and moreover degenerate. For finite degree-of-freedom systems, the noncanonical Hamiltonian form is the same as that of Eqs. (4)-(6), except (J_c^{ij}) is replaced by a tensor (J^{ij}), which may be odd dimensional, have vanishing determinant and depend upon the z^i. However, in spite of these changes in the cosymplectic form, the algebraic properties of bilinearity, antisymmetry and the Jacobi identity of the Poisson bracket, are preserved. These properties are the essence of the Hamiltonian description.

For MHD the Eulerian remnant of the form given by Eq. (1) is

$$\frac{\partial \psi^i}{\partial t} = J^{ij} \frac{\delta H}{\delta \psi^j} \quad , \tag{7}$$

where the ψ^i $(i = 1,...8)$ correspond to the field variables $\mathbf{v}$, $\mathbf{B}$, ρ and s; the Hamiltonian $H = \int [\rho v^2/2 + B^2/2 + \rho U(\rho,s)] d^3x$, where U is the internal energy per unit mass; and the cosymplectic operator, (J^{ij}), was given in Ref. [25].

A feature of noncanical Poisson brackets, not present in ordinarily Poisson brackets, is the existence of special constants of motion called Casimir invariants. A phase space function C is a Casimir invariant if it commutes with the Hamiltonian as well as with <u>any</u> function F of the dynamical variables z^i $(i = 1,...M)$ describing the system, i.e.

$$[C, F(z)] = 0 \quad . \tag{8}$$

A consequence of this definition of the Casimirs, using Eq. (5), is

$$\frac{\partial C}{\partial z^i} J^{ij} \frac{\partial F}{\partial z^j} = 0 \tag{9}$$

but F is arbitrary and therefore

$$J^{ij} \frac{\partial C}{\partial z^j} = 0 \ , \qquad i = 1,...,M. \tag{10}$$

Thus, the phase space gradient of a Casimir $(\partial C/\partial z^j)$ is a null eigenvector of (J^{ij}). In fact, it can be shown[21] that the null space of (J^{ij}) is spanned by null eigenfunctions that are gradients. Clearly, nontrivial Casimirs (i.e. not constants) exist only if

$$\det (J^{ij}) = 0 \tag{11}$$

and the number of independent Casimirs is equal to the corank of (J^{ij}). In the case that (J^{ij}) is canonical, it has the structure given in Eq. (5) and the determinant is unity.

Therefore in the canonical Hamiltonian formalism there are no nontrivial Casimirs. When (J^{ij}) has null eigenvectors, the phase space can be described by hyperplanes, or symplectic leaves, which are labeled by the Casimirs. A trajectory must remain in the leaf of phase space as determined by the specification of the initial conditions. This follows from the fact that the generalized Poisson bracket cannot generate flow, i.e. trajectories in phase space in the direction of these null eigenvectors.

Hamiltonian systems possess a built-in sufficient condition for Liapunov stability. This kind of stability is a stronger than conventional linear or spectral stability. The most common example of nonlinear stability occurs in the case where the Hamiltonian has the standard form: $H = p^2/2 + V(q)$. Here it is well-known that positivity of the curvature of V is sufficient (and under some mild conditions also necessary[5]) for stability of equilibria given by $p^i = 0$, $\partial V/\partial q^i = 0$. This is the finite degree-of--freedom analogue of the MHD energy principle.

The above curvature condition is a special case of a more general condition for Hamiltonian systems. The topology of energy surfaces in the vicinity of an equilibrium point is determined by the curvature of the entire Hamiltonian: $\partial^2 H/\partial z^i \partial z^j$. If this quantity is either positive or negative definite then the energy surfaces near the equilibrium point are nested, closed and bounded surfaces. Since trajectories are confined to the energy surface, those with initial conditions sufficiently close to the equilibrium point will remain in an arbitrarily small neighborhood of the equilibrium point. Evidently, for standard Hamiltonians, positive definiteness of H is insured by that of V.

For noncanonical Hamiltonian systems there is an analogous sufficient condition. To begin with, unlike conventional Hamiltonian systems, extremals of the Hamiltonian are not the only possible equilibria. In fact these equilibria, which correspond to the lowest possible energy states, are generally quite trivial. For example, varying the MHD energy with respect to **v**, **B**, etc. yields an equilibrium with zero flow, magnetic field, pressure and temperature. If one extremizes the Hamiltonian subject to the constancy of certain of the Casimirs, then interesting equilibria that are constrained away from the lowest energy state arise. That these are equilibria is evident from the following:

$$\frac{dz^i}{dt} = [z^i, H] = [z^i, H + C] = J^{ij} \frac{\partial F}{\partial z^j} \quad . \tag{12}$$

Thus vanishing of $\partial F/\partial z^i$ implies that $dz^i/dt = 0$. If we define

$$\delta^2 F \equiv \frac{1}{2} \frac{\partial^2 F(z_e)}{\partial z^i \partial z^j} \delta z^i \delta z^j \quad , \tag{13}$$

then definiteness of the quadratic form $\delta^2 F$ implies that surfaces of constant F near the equilibrium point are topologically spheres and thus we have stability. In the next section we physically interpret this quantity in the noncanonical context.

The free energy functional $\delta^2 F$ can also be obtained in the Lagrangian context[9]. In fact, since all equilibria are not extremal points of $\delta F = 0$ in the noncanonical formalism, the most general expressions have been obtained by beginning with Lagrangian action principles. Free energy expressions obtained in either way are equivalent for equivalent equilibria.

2. THE FREE ENERGY - $\delta^2 F$

Now consider the physical interpretation of $\delta^2 F$. Here we show that $\delta^2 F$ is the energy change resulting from perturbations of the equilibrium that obey the dynamical constraints. Further, we show that the use of $\delta^2 F$ considerably simplifies the computation of the energy change. Also we discuss the connection between $\delta^2 F$ and the dielectric definition of negative energy modes.

It is useful to consider the energy needed to create a small perturbation with the constraint that the motion remain in a symplectic leaf. For small δz we have

$$\Delta H \equiv H(z_e + \delta z) - H(z_e) = \frac{\partial H}{\partial z^i} \delta z^i + \frac{1}{2} \frac{\partial^2 H}{\partial z^i \partial z^j} \delta z^i \delta z^j + \ldots \tag{14}$$

$$\Delta C \equiv C(z_e + \delta z) - C(z_e) = \frac{\partial C}{\partial z^i} \delta z^i + \frac{1}{2} \frac{\partial^2 C}{\partial z^i \partial z^j} \delta z^i \delta z^j + \ldots , \tag{15}$$

where z_e is an equilibrium given by $\partial F/\partial z^i = 0$. For δz such that $\Delta C = 0$, we can add $\Sigma \lambda^i \Delta C^i$ to Eq. (14) in order to obtain the energy change at constant Casimir invariant,

$\Delta H|_C$. Thus

$$\Delta H|_C = \frac{1}{2}\delta^2(H + \Sigma\, \lambda^i C^i) \equiv \delta^2 F \,. \tag{16}$$

This $\Delta H|_C$ is second order in δz, as expected for a Hamiltonian near an equilibrium point. Since $\Delta H|_C$ is also a constant of motion it can serve as a Liapunov functional. Moreover, $\Delta H|_C$ has physical significance: it is the energy increment contained in a perturbation which is accessible to the dynamics given the Casimir constraints. Note that $\delta^2 F$ depends on the equilibrium through the choice of λ^i, so the same δz contains different accessible energies for each equilibrium. This is intuitively satisfying since one would expect the energetic favorability of δz to depend upon the equilibrium. For example, in MHD the energetic favorability and stability of a kink mode perturbation depends on the equilibrium current.

The expression $\delta^2 F$ is a most convenient way to to compute $\Delta H|_C$. The reason for this is that accessible perturbations, δz, must satisfy $\Delta C = 0$. To compute ΔH to second order directly requires $\Delta C = 0$ to second order. It is relatively easy (often trivial) to find trial functions δz_T that satisfy $\Delta C = 0$ to first order in δz. In general a variation δz satisifying $\Delta C = 0$ to higher order would require a calculation to obtain $\delta z = \delta z_T + O(\delta {z_T}^2)$. This is usually difficult to satisfy for all Casimir invariants simultaneously. However, $\Delta H|_C$ can be computed to second order accuracy by substituting δz_T directly into $\delta^2 F$. We do this for the Vlasov equation in Sec. 4.

It is easy to show that $\delta^2 F$ is the Hamiltonian for the linearized equations of motion[29,6], but more importantly one can show that the work performed by an external agent, which observes the constraints while creating a linear perturbation, is $\delta^2 F$. This is the usual definition of the energy content of a wave in dielectric theory. Suppose an external source is applied to the system which adds a transient term H_{ext} to the Hamiltonian for times between 0 and t_0. The J^{ij} is unchanged by this; thus the Casimir invariants (which depend only on J^{ij}) remain constant. For linear perturbations the new

Hamiltonian is

$$H' = \delta^2 F + H_{ext} \,.$$

A convenient form for H_{ext} is $z^i S_j(t)$, the explicit time dependence appearing only in the source functions S_j. Standard Hamiltonian mechanics yields

$$\frac{dH'}{dt} = \frac{\partial H}{\partial t} = z^j \frac{\partial S_j}{\partial t} \,. \tag{17}$$

Integrating Eq. (17) in time from $t = 0$ to $t = t_0$ yields, upon integration by parts on the right, and the assumption that $\mathbf{S} = 0$ for $t \leq 0$ and $t \geq t_0$,

$$\delta^2 F = -\int_0^{t_0} (dz^j/dt) S^j dt \,. \tag{18}$$

The right side is the usual expression for the work done on a system, i.e. the integrated input power, by an external agent.

As an example, consider the one-dimensional Vlasov-Poisson equation. The external agent in this case is an externally imposed potential φ_{ext} and the external portion of the Hamiliton is

$$H_{ext} = \int f \, \varphi_{ext} \, dx \, dv \,.$$

Here $f(x,v,t)$ is the phase space density. In this case we have for the power

$$\begin{aligned}(dz^j/dt) S^j &= \int (\partial f/\partial t) \varphi_{ext} \, dx \, dv = \int (\partial \rho/\partial t) \varphi_{ext} \, dx \\ &= \int \mathbf{J} \cdot \nabla \varphi_{ext} \, dx \,.\end{aligned} \tag{19}$$

The last equality of Eq. (19) follows by making use of the continuity equation, $\partial \rho/\partial t + \nabla \cdot J = 0$, and integrating by parts. This expression is the standard one for the power input from an external field. Thus $\delta^2 F$ is indeed the energy needed to create a perturbation δf.

In dielectric theory the energy content[1-3] in a linear wave is defined by the work performed by an external agent in producing the wave. The energy content is found to be $[\partial(\omega\varepsilon)/\partial\omega]|E_k|^2$, where ω is the wave frequency, k is the mode number, $\varepsilon(k,\omega)$ is the dielectric function, and E_k is the electric field amplitude for mode k. This brings us to the important topic of positive and negative energy waves; in dielectric theory these have $\partial(\omega\varepsilon)/\partial\omega > 0$ and $\partial(\omega\varepsilon)/\partial\omega < 0$ respectively. As noted in the Introduction, negative energy waves have the property that they are spectrally stable, but their presence decreases the energy of the medium. If both types of waves exist in the medium then δ^2F is positive for some perturbations, δz, and negative for others, and is thus indefinite.

Now suppose that δ^2F is indefinite. What conclusions can be drawn about the spectral properties of the system? For finite degree-of-freedom systems there are two possibilities: either there is spectral instability or the system has a mixture of positive and negative energy waves.

For the case of canonical Hamiltonian systems it is straight forward to show this assertion. The theory of normal forms by now is well understood[5]. In this case δ^2F is the linearized Hamiltonian, i.e. the second variation of the full Hamiltonian. It is a quadratic form in δq_i and δp_i. We will consider the case where H has the following form:

$$H = \sum_i \alpha_i(\delta p_i)^2 + \beta_i(\delta q_i)^2 \quad . \tag{20}$$

If the H of Eq. (20) is indefinite, then for a least one degree-of-freedom one of the following holds: (i) α_i has a different sign than β_i, which corresponds to instability, or (ii) the pair of α_i and β_i has a different sign from other pairs. The latter case corresponds to a mixture of positive and negative energy waves. It is evident from the discussion of Sec. 1.2 that the same results apply for finite noncanonical Hamiltonian systems.

It is clear from the above that there is no such thing as a spectrally unstable negative energy mode. One might think this would correspond to a negative α_i and a positive β_i, but since $q_i \rightarrow -p_i$ and $p_i \rightarrow q_i$ is a canonical transformation we see that there is no distinction between this case and that of instability.

However, we have conjectured that a system with both positive and negative energy waves is nonlinearly unstable in the general case. This can arise because there is a low order resonance in the system, which results in explosive instability, i.e. divergence in finite time (c.f. Sec. 3), as occurs in the well-known three-wave interaction. For two degree-of-freedom systems where no low order resonance exists, the KAM theorem indicates stability. For systems of more than two degrees-of-freedom, instability can still be present by the slow mechanism known as Arnold diffusion. For infinite degree-of-freedom systems with continuous spectra, there may be additional avenues for nonlinear instability.

There are many continuum systems for which the concept of a dielectric function is not well defined, or if defined in principle, is very difficult to calculate in practice. However, $\delta^2 F$ can be easily computed for such systems. Thus $\delta^2 F$ provides a practical generalized definition of the concept of a negative energy wave. It can easily be generalized to the concept of non-wave-like phenomena.

Before closing this section, let us consider a general bifurcation property concerning negative energy modes. Suppose we have a sequence of noncanonical Hamiltonian equilibria parameterized by a continuous variable η. Assume that $\delta^2 F$ is positive definite for $\eta < \eta_0$, but indefinite for $\eta > \eta_0$, because a positive energy mode becomes a negative energy mode (the system being spectrally stable on both sides of the threshold). We now show that the frequency of such a mode must go through zero at $\eta = \eta_0$, at least for finite dimensional systems.

The tensor $\partial^2 F/\partial z^i \partial z^j$ is symmetric, and for $\eta < \eta_0$ positive definiteness implies it has all positive eigenvalues. At least one eigenvalue becomes negative for $\eta > \eta_0$, and thus traverses zero at $\eta = \eta_0$. The associated eigenvector, $\delta z_0{}^i$, is also a zero frequency eigenvector for the dynamical system; i.e. the existence of a zero eigenvalue of $\partial^2 F/\partial z^i \partial z^j$ implies the existence of a zero frequency mode. This follows from linearization of Eq. (12) about an equilibrium z_e,

$$i\omega \delta z_0{}^i = J^{ij} \frac{\partial^2 F}{\partial z^j \partial z^k} \delta z_0{}^k = 0 , \tag{21}$$

where J^{ij} and $\partial^2 F/\partial z^j \partial z^k$ are evaluated at z_e.

Note that this result can also be derived from the definition of wave energy density, which as noted above has its sign determined by the factor $\partial(\omega\varepsilon)/\partial\omega$. Since the wave frequency satisifies $\varepsilon = 0$ this factor becomes $\omega\partial\varepsilon/\partial\omega$. Thus as the wave energy goes from positive to negative at $\eta = \eta_0$ we have two possibilities: either ω goes through zero or $\partial\varepsilon/\partial\omega$ does. The latter possibility can be excluded by showing it leads to a contradiction. Assume $\varepsilon = 0$ and $\partial\varepsilon/\partial\omega = 0$ at $\omega = \omega_0 \neq 0$. Taylor expanding ε about this point yields

$$\varepsilon = \sum_{n=2}^{\infty} \frac{(\omega - \omega_0)^n}{n!} \frac{\partial^n \varepsilon}{\partial \omega^n} + (\eta - \eta_0) \frac{\partial \varepsilon}{\partial \eta} + \ldots \tag{22}$$

Since $\partial\varepsilon/\partial\omega$ is assumed to change sign at $\eta = \eta_0$, one can show that the first nonvanishing term on the Taylor series must have even n. One can also show that if $\partial\varepsilon/\partial\eta \neq 0$, then the cross terms neglected in Eq. (22) are asymptotically negligible compared with those kept when $\omega \approx \omega_0$ and $\eta \approx \eta_0$.

The condition $\varepsilon = 0$ allows us to solve for the mode frequency near ω_0 by keeping the first nonvanishing term in the series; hence

$$(\omega - \omega_0)^n = n!(\eta - \eta_0) \frac{(\partial\varepsilon/\partial\eta)}{(\partial^n\varepsilon/\partial\omega^n)} . \tag{23}$$

Since n is even, this implies that instabilities must exist for either $\eta < \eta_0$ or $\eta > \eta_0$, depending on the signs of the ε partial derivatives.

Thus, as equilibrium parameters are varied, equilibria with only positive energy modes can only acquire negative energy modes (which are by definition stable) when the mode frequency passes through zero. Furthermore, the $\delta^2 F$ and dielectric definitions of energy agree on the value of η for which this occurs.

3. A MODEL OF FLR STABILIZATION

We consider a model that has a general form that occurs in models of finite Larmor radius stabilization. This model is canonical in nature and displays the form of negative energy modes discussed in Sec. 2 for finite degree-of-freedom systems. The

Lagrangian for a particle in a uniform magnetic field B in the z-direction, subject to a potential which to leading order is harmonic and inverted, is

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{eB}{2c}(\dot{y}x - \dot{x}y) + \frac{k}{2}(x^2 + y^2) - V_3, \qquad (24)$$

where $k > 0$ and V_3 is an anharmonic contribution to the potential that is a cubic in x and y. The inverted harmonic potential could arise because of MHD equilibrium[16] or one could be interested in the motion of particles in a uniformly charged cylindrical column.

Upon Legendre transforming (21) and writing the resulting Hamiltonian in terms of the following canonical variables:

$$q_1 = x/\sqrt{m} \qquad\qquad q_2 = y/\sqrt{m}$$
$$p_1 = m^{3/2}(\dot{x} - \omega_L y) \qquad\qquad p_2 = m^{3/2}(\dot{y} + \omega_L x) \qquad (25)$$

yields

$$H = \frac{1}{2}(p_1{}^2 + p_2{}^2) + \omega_L(q_2 p_1 - q_1 p_2)$$
$$+ \frac{1}{2}(\omega_L{}^2 - \omega_0{}^2)(q_1{}^2 + q_2{}^2) + V_3\,, \qquad (26)$$

where ω_L is the Larmor frequency, $eB/2mc$, and ω_0 is the natural frequency of the harmonic potential, $\sqrt{k/m}$.

Consider now the spectrum of the equilibrium point $q_1 = q_2 = p_1 = p_2 = 0$. Neglecting V_3 and supposing that the dynamical variables vary as $\exp(i\omega t)$, we obtain a fourth-order characteristic polynomial that has the roots

$$\omega = \pm\,\omega_0[\sqrt{\eta - 1} \pm \sqrt{\eta}\,]\,, \qquad (27)$$

where $\eta = \omega_L{}^2/\omega_0{}^2$. Equation (24) displays the symmetries of the spectrum of a Hamiltonian system; i.e. pure real and pure imaginary frequencies appear in pairs, while in general frequencies occur in quartets of the form $\pm a \pm i\,b$.

Let us trace the frequencies in the complex ω-plane as we increase the magnetic field away from zero. At $\eta = 0$ we have two doubly degenerate pure imaginary roots, $\omega = \pm i\omega_0$. It is not surprising that there is instability in this case since the particle is free to fall off the hill. Upon increasing η the degenerate eigenvalues split and move respectively clockwise and counterclockwise on a circle of radius ω_0 toward the ω_r axis. When $\eta = 1$, $\omega = \pm\omega_0$ and the eigenvalues again become doubly degenerate and experience a Krein crash (See Refs. [4] and [5]), with one pair of real eigenvalues approaching the origin and one pair receding. After the Krein crash the eigenvalues have magnitude

$$\omega_1 = -\omega_L + \sqrt{\omega_L^2 - \omega_0^2}$$

$$\omega_2 = \omega_L + \sqrt{\omega_L^2 - \omega_0^2}\,. \qquad (28)$$

Observe that $0 < \omega_1 \leq \omega_L$ and $\omega_2 \geq \omega_L$. Subsequently we will see that ω_1, the smaller frequency, is a negative energy mode. It has negative signature while ω_2 has positive signature, as required by Krein's theorem.

Consider now the canonical transformation generated by the following mixed variable generating function:

$$F_2(q_1, q_2, P_1, P_2) = \mu(q_1 P_1 + q_2 P_2) + P_1 P_2 + \mu^2 q_1 q_2/2, \qquad (29)$$

where $\mu = [4(\omega_L^2 - \omega_0^2)]^{1/4}$. In terms of the new canonical variables P_1, P_2, Q, and Q_2, the Hamiltonian becomes

$$H = -\frac{1}{2}\,\omega_1 (P_1^2 + Q_1^2) + \frac{1}{2}\,\omega_2 (P_2^2 + Q_2^2) + V_3\,, \qquad (30)$$

where V_3 is now a cubic function of the new coordinates and momenta. We see from Eq. (27) that this Hamiltonian has the standard form for a negative energy mode.

Let us now see what happens to the spectrum when we add negative definite dissipation to this system; i.e. dissipation that removes energy. This will happen if we dissipate the negative energy mode. Since V_3 does not effect the spectrum we can consider the negative energy oscillator independently. With the inclusion of dissipation this oscillator is governed by

$$\dot{Q}_1 = -\omega_1 P_1$$
$$\dot{P}_1 = \omega_1 Q_1 + 2\upsilon P_1 \; . \qquad (31)$$

To see that the term involving υ corresponds to dissipation note that, for $\upsilon > 0$, $dH/dt = -2\omega_1 \upsilon P_1{}^2 \leq 0$. The eigenfrequency corresponding to this oscillator becomes

$$\omega = -i\upsilon \pm \sqrt{\omega_1{}^2 - \upsilon^2} \, , \qquad (32)$$

which corresponds to exponential growth for $\upsilon > 0$. Thus dissipation destabilizes the negative energy modes.

Now consider the other way that negative energy modes lead to instability; i.e. nonlinear destabilization. In the case where there is a resonance this phenomenon is tractable. Let us suppose an order three resonance condition: $2\omega_1 = \omega_2$. This occurs when $3\omega_0 = 2\sqrt{2}\omega_L$. The important terms of V_3 are those that drive the resonance. These can be obtained by the standard Hamiltonian perturbation method of averaging. This procedure yields

$$H = \frac{1}{2}\,\omega_1 (P_1{}^2 + Q_1{}^2) - \omega_1 (P_2{}^2 + Q_2{}^2)$$
$$+ \frac{\alpha}{2}\,[Q_2(Q_1{}^2 - P_1{}^2) - 2Q_1 P_1 P_2] \; , \qquad (33)$$

where α is a constant. We have reversed the sign of time in order to show that the Hamiltonian of Eq. (33) is that due to Cherry[14], who observed that a two-parameter solution set is given by

$$Q_1 = \frac{\sqrt{2}}{\alpha(\omega_1 t - \varepsilon)} \sin(\omega_1 t + \gamma)$$

$$P_1 = \frac{-\sqrt{2}}{\alpha(\omega_1 t - \varepsilon)} \cos(\omega_1 t + \gamma)$$

$$Q_2 = \frac{1}{\alpha(\omega_1 t - \varepsilon)} \sin(2\omega_1 t + \gamma)$$

$$P_2 = \frac{-1}{\alpha(\omega_1 t - \varepsilon)} \cos(2\omega_1 t + \gamma) \, , \qquad (34)$$

where ε and γ are determined by the initial conditions. The interesting thing about this solution is that any neighborhood of the origin contains initial conditions of solutions that diverge in finite time. Thus this system is spectrally stable, but in reality unstable. The nonlinearity diverts energy from the negative energy mode to the positive energy mode.

In the case where there is no resonance and the system possesses a general cubic term with explicit time dependence the system is most likely unstable by the mechanism of Arnold diffusion. Preliminary numerical calculations show that the system Arnold diffuses until a separatrix is reached, at which point the growth is rapid.

4. VLASOV EQUILIBRIA

In this section we discuss the free energy for the Vlasov equation. We begin by reviewing in some detail the noncanonical Hamiltonian formalism for the one-dimensional Vlasov Poisson system[30] and its associated variational principle for equilibria. From this $\delta^2 F$ is obtained, but within this formalism only monotonic and isotropic equilibria are obtained as extremals of the variational principle, and only sufficient conditions for the stability of this limited class of equilibria are obtained. In recent work expressions for $\delta^2 F$ that generalize to arbitrary equilibria, have been obtained. The result for the Vlasov-Poisson equation is stated here; the reader is referred to the references for its derivation, along with the general Maxwell-Vlasov result.

The Vlasov-Poisson equation is

$$\frac{\partial f(z,t)}{\partial t} = -v\frac{\partial f}{\partial x} + \frac{e}{m}\frac{\partial \varphi(x;f)}{\partial x}\frac{\partial f}{\partial v} \tag{35}$$

where, as usual, $f(z,t)$ is the phase space density at the phase space point $z = (x,v)$ for a charged species of particles with charge e and mass m. Now we consider only a single species, but the results are readily generalizable. The electrostatic potential φ is to be viewed as a functional of f determined via Poisson's equation $\varphi_{xx} = -e\int f\,dz$; thus $\varphi(x;f) = e\int V(z,\tilde{z})\,f(\tilde{z})\,d\tilde{z}$, where $V(x,\tilde{x})$ is the single particle potential (assumed spatially invariant). The Hamiltonian for this system is the energy functional

$$H[f] = \int T(z)\, f(z)\, dz + \frac{e^2}{2}\int\int V(z,\tilde{z})\, f(z)\, f(\tilde{z})\, d\tilde{z}\, dz, \quad (36)$$

where $T(z) = mv^2/2$ is the particle kinetic energy. This system possesses the following noncanonical Poisson bracket:

$$\{F,G\} = \frac{1}{m}\int f(z)\left[\frac{\delta F}{\delta f}, \frac{\delta G}{\delta f}\right] dz \quad (37)$$

where the inner bracket [,] is defined by $[k,h] = k_x h_v - k_v h_x$. Note that $\delta H/\delta f = T + e\varphi \equiv E$, where E is the total particle energy. Evidently,

$$\frac{\partial f}{\partial t} = \{f,H\} = -[f,E] \quad (38)$$

where $-[f,E]$ is equivalent to the right hand side of Eq. (35).

As discussed in Sec. 1.2 stationary points of the free energy $F = H + C$ are equilibria. In the case of the Vlasov-Poisson equation, H is given by Eq. (36) and C is the well-known Liouville constraint, which is given by the following:

$$C[f] = \int c(f)\, dz\,, \quad (39)$$

where $c(f)$ is an arbitrary function. Upon varying F we obtain

$$\begin{aligned} \delta F[f;\delta f] &= \int (E + \partial c/\partial f)\, \delta f\, dz \\ &= \int (mv^2/2 + e\varphi + \partial c/\partial f)\, \delta f\, dz\,. \end{aligned} \quad (40)$$

Thus equilibria (f_e) are given by

$$E + \partial c(f)/\partial f = 0\,. \quad (41)$$

There are two things to notice about these equilibria that are obtained as extremals of F: firstly, in order to solve for $f_e(E)$, the quantity $\partial c/\partial f$ must be monotonic and therefore its inverse must also be monotonic. This gives

$$f_e = f_e(E) \qquad , \qquad (42)$$

where $f_e(E)$ is a monotonic function of E, and we see that there is a one-to-one correspondence between an equilibrium and a choice for the function c. Secondly, since f_e is a function of E, only velocity symmetric distributions $[f_e(v) = f_e(-v)]$ are obtained. Evidently, extremals of F only make up a subclass of equilibria of the Vlasov-Poisson equation, since this system is known to possess nonmonotonic and velocity asymmetric (for untrapped particles) equilibria.

For the above restricted class of equilibria we can obtain a criterion for stability[31-34] by taking the second variation of F; viz,

$$\delta^2 F[f] = \frac{1}{2}\left(e^2 \int\int V(z,\tilde{z})\,\delta f(z)\,\delta f(\tilde{z})\; d\tilde{z}\, dz + \int \partial^2 c/\partial f^2\, (\delta f)^2\, dz\right) , \qquad (43)$$

Observe that the first term of Eq. (43) is positive definite (it corresponds to the second variation of the electrostatic energy which goes as the square of the electric field), while the second term will be positive definite provided $\partial^2 c/\partial f^2 > 0$. For stability this must be true over the entire domain of integration when f is set equal to f_e, since we can make $\delta^2 F$ negative by choosing δf such that the first term of Eq. (43) vanishes and such that δf is localized where $\partial^2 c/\partial f^2 < 0$. This statement translates into a statement about f_e: upon differentiating Eq. (41) with respect to E we obtain

$$\partial f_e/\partial E = -[\partial^2 c(f_e)/\partial f^2]^{-1} . \qquad (44)$$

Therefore we have stability if f_e is <u>any</u> monotonic decreasing function of the energy.

Note that if we attempted to apply the formula of Eq. (43) to nonmonotonic equilibria then $\delta^2 F$ diverges unless δf vanishes at places where $\partial f_e/\partial E = 0$. If we restrict δf to the Casimir surfaces then this problem is avoided. In the notation of Sec. 2

the δz_T is given by

$$\delta f = [f_e, g] = (\partial f_e/\partial E)[E,g] \; , \qquad (45)$$

where g is an arbitrary function. One can show that δf as given by Eq. (45) preserves the Casimir constraint to first order. Inserting Eq. (45) into Eq. (43), upon making use of Eq. (44), yields

$$\delta^2 F[f] = \frac{1}{2}\left(e^2 \int\int V(z,\tilde{z})\,\delta f(z)\,\delta f(\tilde{z})\,d^3\tilde{z}\,d^3z + \int [g,f_e][E,g]\,d^3z\right) . \qquad (46)$$

Observe that the monotonicity condition for stability is replaced by positivity of the second integrand, for all g. When this integrand is not positive definite, and cannot be made so by a frame change, then there exist negative energy modes when there is linear stability. The result of Eq. (46) was derived, within the Lagrangian variable context, in Refs. [9]. It is applicable to arbitrary Vlasov-Poisson equilibria. The use of Lagrangian variables formally circumvents the limitation associated with the Eulerian description. Also in Refs. [9] the free energy expression for arbitrary Maxwell-Vlasov equilibria is given.

ACKNOWLEDGMENT

This work was supported by the U.S. Department of Energy contract #DE-FG05-80ET-53088.

REFERENCES

1. von Laue, M., Ann. Physik 18, 523 (1905)
2. Brillouin, L., Wave Propagation and Group Velocity, (Academic Press, New York ,1960).
3. Landau, L. D. and Lifshitz, E. M., Electrodynamics of Continuous Media, (Pergmon Press, Reading, 1960) p. 253.
4. Moser, J. K., Comm. Pure Appl. Math. 11, 81 (1958); Mem. Am. Math. Soc. 81, 1 (1968).

5. Arnold, V. I., Kozlov, V. V. and Neishtadt, A. I., in Encyclopedia of Mathematical Sciences 3, Dynamical Systems III, Ed. Arnold, V. I. (Springer-Verlag, Berlin,1988).
6. Morrison, P. J. and Eliezer, S., Phys. Rev. A 33, 4205 (1986).
7. Kotschenreuther, M., *et al.*, Plasma Physics and Controlled Nucl. Fusion Res. 2, 149 (1987) (International Atomic Energy Agency Publication).
8. Morrison, P. J., Z. Naturforsch. 42a, 1115 (1987).
9. Morrison, P. J. and Pfirsch, D., Phys. Rev. A 40, 3898 (1989); Institute for Fusion Studies, Austin, Texas, Report, IFSR #401, Oct. 1989. Submitted to Physics of Fluids.
10. Sturrock, P. A., J. Appl. Phys. 31, 2052 (1960).
11. Dikasov, V. M., Rudakov, L. I., and Ryutov, D. D., Zh. Eksperim. i Teor. Fiz. 48, 913 (1965). [English trans. Soviet Phys. -JETP 21, 608 (1966).]
12. Aamodt, R. E. and Sloan, M. L., Phys. Rev. Lett. 19, 1227 (1967).
13. Sagdeev, R. Z. and Galeev, A. A., in T. M. O'Niel and D. L. Book (eds.), Nonlinear Plasma Theory, (W. A. Benjamin, New York, 1969).
14. Coppi, B., Rosenbluth, M. N., and Sudan, R., Ann. Phys. 55, 207 (1968).
15. Cherry, T. M., Trans. Cambridge Philos. Soc. 23, 199 (1925); See Whittaker, E. T., Analytical Dynamics (Cambridge, London, 1937), Sec. 182, p. 142; and Wintner, A.,The Analytical Foundations of Celestial Mechanics (Princeton University, Princeton, 1947), Sec. 136, p. 101.
16. Thompson, W. and Tait, P. G., Treatise on Natural Philosophy, part 1, (Cambridge University Press, Campridge,!921), p. 388.
17. Greene, J. M. and Coppi, B., Phys. Fluids 8, 1745 (1965).
18. Newcomb, W. A., Nuclear Fusion: Supp. part 2, 451 (1962).
19. VanKampen, N. G. and Felderhoff B. U., Theoretical Methods in Plasma Physics (North-Holland, Amsterdam, 1967).
20. Low, F. E., Proc. Roy. Soc. 248A, 282 (1958).
21. This formalism has roots that trace back to Sophus Lie, see for instance, Caratheodory, C., Calculus of Variations and Paritial Differential Equations of the First Order (Holden-Day, San Francisco, 1965).
22. Pauli, W., Il Nouvo Cimento 10, 648 (1953).
23. Martin, J. L., Proc. Roy. Soc. 251A, 536 (1959).
24. Arnold, V.I., Inst. Fourier Grenoble **16**, 319 (1966).
25. Morrison P.J. and Greene, J.M., Phys. Rev. Lett. 45, 790 (1980).

26. Morrison, P.J., in Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems, AIP Conf.. Proc. 88, 13 (1982).

27. Sudarshan, E.C.G. and Mukunda, N. Classical Mechanics - a Modern Perspective, (Krieger, Melbourne, Florida, 1983) Second Edition.

28. Arnold, V. I., Ann. Math. Soc. Transl. **79**, 267 (1969).

29. In the following reference many sufficient conditions for stability were reviewed and presented. These expressions, when restricted to the constraint surfaces, correspond to the free energy. Holm, D. D., Marsden, J. E., Ratiu, T. and Weinstein, A., Phys.Reps.123, 1 (1985).

30. Morrison, P.J., Phys. Lett. 80A, 383 (1980).

31. Newcomb, W. appendix to I. B. Bernstein, Physical Review **109**, 10 (1958).

32. Kruskal, M. D. and Oberman, C., Physics of Fluids **1**, 275 (1958).

33. Gardner, C. S., Phys. Fluids **6**, 839 (1963).

34. Fowler, T. K., J. Math. Phys. **4**, 559 (1963); also paper in A. Simon and W. B. Thompson (eds.), Advances in Plasma Physics, (Wiley, New York, 1968) vol.1, p. 201.