Cnoidal electron hole propagation: Trapping, the forgotten nonlinearity in plasma and fluid dynamics

Hans Schamel[a]

Physikalisches Institut, Universität Bayreuth, D-95440 Bayreuth, Germany

In this review a plaidoyer is held for a specific form of nonlinearity, the trapping nonlinearity (TN), which arises due to a capture of particles and/or fluid elements in an excited coherent structure. This is of some importance since it appears that TN has not yet taken roots hitherto, neither in turbulence nor in anomalous transport models. The present state of knowledge about wave excitation, seen numerically and experimentally, especially at space craft, however, speaks a different language suggesting that current wave models are constructed too narrowly to reflect reality. The focus is on traveling cnoidal electron holes (CEHs) in electrostatically driven plasmas and the physical world associated with these. As a result a new wave concept emerges, in which the low amplitude dynamics is nonlinearly controlled by TN. © *2012 American Institute of Physics.* [doi:10.1063/1.3682047]

The kinetic theory of periodic electron hole equilibria or phase space vortices is thereby reexamined in the small amplitude limit. Use is made of the pseudo-potential method extended into the kinetic regime. A set of two macroscopic equations, the nonlinear dispersion relation (NDR), in charge of the wave velocity, and the pseudo-potential, responsible for the spatial structure, is derived and analyzed. The obtained modes are of cnoidal type, i.e., are described by Jacobian elliptic functions, and encompass solitary as well as periodic wave equilibria of essentially rarefaction type. These structures, which are moving with electron thermal velocity and slower, are reminiscent of the van Kampen wave continuum, but are intrinsically nonlinear even in the infinitesimal amplitude limit. Only very exceptionally, a wavepacket composed of the latter modes can come close to the present structures, but still suffers damping and is ill-behaved in phase space. It is the microscopic particle distribution function (DF) at resonant velocity which marks the difference. On the other hand, linear packets exhibit filamentary and more or less singular distributions in this region, and the present structures are distinguished by sufficiently smooth and regular trapped particle distributions, being typically depressed and hence hole-like.

In plasma physics, the following innovations are presented:

(1) The intrinsic nature of TN, as a plasma property independent of the strength of wave amplitude, is proved via several approaches.
(2) Although identical in shape and speed, the van Kampen modes (and all other linear superpositions supposed to describe equilibria) are excluded from the class of solutions of the Vlasov-Poisson (VP) system, the true modes being nonlinear and controlled by TN. This holds up to the infinitesimal amplitude limit. In other words, it

appears rather doubtful to confront reality with solutions of a kinetic equation, which turns out to be invalid because of truncation.
(3) In the spectrum of nonlinear waves there are three localized solutions, two referring to rarefactive waves, the solitary electron hole (SEH), and the cnoidal electron hole wavelet (CEHWL); the other one is solitary too, but refers to a compressional solution (solitary potential dip, SPD). The former two have been identified in space experiments as bipolar, tripolar, or multipolar spikes of $E_\parallel$, respectively, the latter one possibly in a lab experiment.
(4) A linear instability for the generation of these modes is not needed. They can be found in linearly stable regimes as well, as a result of a nonlinear instability, e.g., in a current-carrying, noisy plasma, being guided by the zero-energy concept. Space observations bear witness of this.
(5) Rayleigh's group velocity concept ceases to be valid and applicable in the presence of trapping due to the lack of a linear carrier mode, the correct speed, e.g., of a CEHWL, being given by the NDR. Moreover, several generalizations and theoretical extensions are presented in this review.

This altogether challenges electrostatic turbulence and anomalous transport models by questioning their linear basis due to trapping. The omnipresent and intrinsic feature of trapping at all levels of wave intensity holds true for any collisionless or weakly collisional, fluid-like plasma driven electrostatically by currents, beams, or inhomogeneities.

In fluid dynamics, the mathematical analogy between the VP system and the 2D, incompressible shear flow equations suggests a transfer of these trapping ideas to excited flows as well. Secondary equilibria involving localized patches of vorticity, in which fluid elements are trapped, pave the way for a better theoretical understanding the onset

a)Electronic mail: hans.schamel@uni-bayreuth.de. URL: www.hans-scha-mel.de.

of turbulence, as exemplarily indicated for the plane Couette flow problem.

In summary, the underlying TN is expected to play a central, universal role in the general setting of plasma and fluid turbulence beyond that of one-dimensional, electrostatically driven plasmas.

I. INTRODUCTION

Phase space holes, well-known for a couple of decades, are long living electrostatic structures in driven, collisionless plasmas far away from thermodynamic equilibrium, exhibiting typically a dip in the phase space distribution function at resonant velocity. Owing their existence to the trapping of particles in the wave's potential, these structures require a kinetic Vlasov-Poisson description. Experimentally, they have been first discovered in Q-machines by Refs. 1 and 2 and are now ubiquitously found in laboratories and space. Holes are, therefore, of paramount importance in modern plasma physics, affecting transport anomalously via intermittent turbulence.

Theoretically, the first consistent description of (solitary) electron holes (EHs) has been presented by Ref. 3, based on a method to solve the Vlasov-Poisson system,[4] which differs from the BGK method.[5] Solitary EH solutions, obtained by the original BGK method, have been presented, e.g., by Refs. 6–8. Reviews on EHs and other potential structures, such as ion holes (IHs) or double layers (DLs), have been published in Refs. 9–12. In these monographs elaborated lists of hole observations in numerical, laboratory, and satellite experiments can be found, besides generalized solutions of holes in three-dimensional (3D), non-uniform, and magnetized plasmas,[13,14] in relativistic plasmas[15] or in non-thermal plasmas.[16]

A typical EH structure consists of a solitary potential hump that moves near electron thermal velocity v_{te} and appears as a saturated state of a linear two-stream instability. This characterization, however, turns out to be too narrow, as trapped electron structures have meanwhile been observed, which appear in the form of wavelets or even potential dip structures that can move with velocities well below v_{te} and can surprisingly be found in linearly stable plasmas, see e.g., Refs. 17 and 18. Moreover, the nonlinear extension of the group velocity of an electron plasma wave in Refs. 19 and 20 indicates that a new element is at work the moment particle trapping comes into play.

This richness in its manifestations challenges plasma wave theory. It is hence worthwhile to take these novelties as motivation to study the periodic extension of EHs, introduced in Refs. 21 and 10, more thoroughly.

The goal in the present paper, which for long stretches reviews the subject, therefore, is to present and evaluate the analysis in a self-contained manner, to uncover new properties of experimental and theoretical interest, e.g., in the magnetic reconnection process, to discuss its implications on plasma wave theory, e.g., on the group velocity of a wavelet, and to give an outlook on other dynamical, collective systems, such as shear flows, as a consequence of the underlying trapping nonlinearity.

II. BASIC THEORY

We are looking for stationary, 1D, electrostatic waves, which are traveling with wavespeed v_0 in a collisionless, unperturbed, thermal plasma. The electron motion in phase space is governed by the Vlasov equation, which reads in the frame moving with v_0, i.e., in the wave frame,

$$[v\partial_x + \Phi'(x)\partial_v]f(x,v) = 0, \tag{1}$$

where normalized quantities have been used, based on the density n_0, and the temperature T_e of the unperturbed plasma.

An appropriate solution is given by the following Ansatz:[3,4,22]

$$f(x,v) = \frac{1 + k_0^2\Psi/2}{\sqrt{2\pi}}\left[\theta(\epsilon)exp\left[-\frac{1}{2}(\sigma\sqrt{2\epsilon} + v_0)^2\right] + \theta(-\epsilon)exp\left(-\frac{v_0^2}{2}\right)exp(-\beta\epsilon)\right], \tag{2}$$

where $\theta(\epsilon)$ represents the Heavyside step function, $\sigma = sg(v)$ is the sign of the velocity, $\epsilon := \frac{v^2}{2} - \Phi(x)$ is the single particle energy, Φ the electrostatic potential, and v_0 the not yet known phase velocity of the expected structure. The first part in Eq. (2) represents the free (or untrapped), the second part the trapped electrons, distinguished in phase space by the contour $\epsilon = 0$, the separatrix. Depending on the two constants of motion, σ and ϵ, $f(x, v)$ is a solution of Eq. (1).

Note that this distribution function is continuous in phase space, especially across the separatrix. It is suggested by the replacement of $v \equiv \sigma\sqrt{v^2}$ through $\sigma\sqrt{2\epsilon}$ and by the demand that it represents a shifted Maxwellian at the point where the trapped particles are absent, i.e., at $\Phi = 0$. We have assumed w.l.o.g. $0 \leq \Phi(x) \leq \Psi$, where Ψ represents the amplitude of the perturbation. With this we have correctly incorporated the unperturbed plasma state given by $\Psi \equiv 0$, being represented by the shifted Maxwellian: $f_M(v) = \frac{1}{\sqrt{2\pi}} exp[-\frac{1}{2}(v + v_0)^2]$. The parameter β in Eq. (2) controls the amount of trapped particles and turns out to be a necessary requisite for obtaining a closed self-consistent description. A dip in the distribution function in the trapped particle region in phase space, $\epsilon < 0$, is thereby provided by β negative. The special form of the normalization in Eq. (2) will be commented upon later.

To close the system, i.e., to find a self-consistent solution, we have to solve the second part, the Poisson equation, which in the immobile ion limit becomes

$$\Phi''(x) = \int f(x,v)dv - 1 =: -V'(\Phi). \tag{3}$$

In the second step of Eq. (3) we have introduced the pseudo-potential (often called Sagdeev potential), because its knowledge allows by a quadrature to obtain the final shape of the potential structure $\Phi(x)$ via the pseudo-energy,

$$\frac{1}{2}\Phi'(x)^2 + V(\Phi) = 0, \tag{4}$$

where we w.l.o.g. assumed $V(0) = 0$.

The electron density in Eq. (3), valid for small amplitudes, $\Psi \ll 1$, can be obtained in two ways: either by a direct integration of Eq. (2) and a subsequent Taylor expansion of the full nonlinear density expression, as was done in Ref. 4 or by a Taylor expansion of Eq. (2) first, followed by the velocity integration, as was done in Refs. 10, 21, 23, and 21. If performed correctly, both procedures, of course, must end up in an identical expression, which in the present case becomes

$$n(\Phi) = \left[1 + \frac{k_0^2 \Psi}{2} - \frac{1}{2} Z_r'\left(v_0/\sqrt{2}\right)\Phi - \frac{4}{3}b(\beta, v_0)\Phi^{3/2} + \ldots \right].$$

(5)

It should be mentioned explicitly that in Eq. (5), where higher order terms are neglected inclusively the ordinary square nonlinearity term, the only remaining nonlinear term, which stems from trapping, is $O(\Phi^{3/2})$ and is hence the ruling nonlinearity throughout the whole paper. In this density expression the term $-\frac{1}{2}Z_r'(v_0/\sqrt{2})$ can be interpreted as an electronic shielding term[9] and is defined by $-\frac{1}{2}Z_r'(v_0/\sqrt{2}) := P \int \frac{1}{v} \partial_v f_M(v) dv$, where P stands for principal value, and $Z_r(x)$ represents the real part of the complex plasma dispersion function for real arguments. The trapping effects are incorporated in $b(\beta, v_0)$, which is defined by

$$b(\beta, v_0) = \frac{1}{\sqrt{\pi}}(1 - \beta - v_0^2)exp(-v_0^2/2).$$

(6)

A plot of the function $-\frac{1}{2}Z_r'(x)$, displayed in Fig. 1, shows that it has a zero transition at $x_0 = 0.924$ ($\sqrt{2}x_0 = 1.307$), a minimum of -0.285 at $x_{min} = 1.506$ ($\sqrt{2}x_{min} = 2.13$), and is positive for $x < x_0$ and negative for $x > x_0$ and vanishes at infinity. Since, according to Eq. (5), it holds $x = v_0/\sqrt{2}$, there hence exist two separated regions for the phase velocity: a slow one with $0 \le v_{0s} \le 2.13$ and a fast one with $2.13 \le v_{0f}$.

A solution of our problem is then obtained by demanding
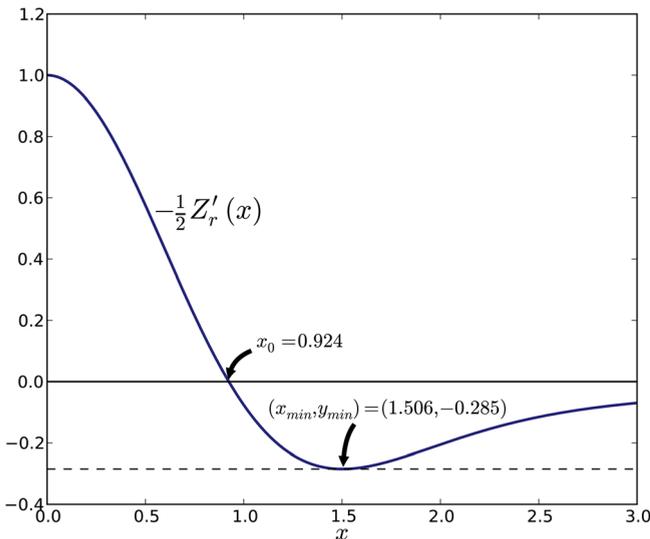
(i) $V(\Phi) < 0$ in $0 < \Phi < \Psi$ and

FIG. 1. (Color online) The function $-\frac{1}{2}Z_r'(x)$.

(ii) $V(\Psi) = 0$,

the latter expression representing zero electric field at potential maximum. After substitution of Eq. (5) into Eq. (3) and a subsequent Φ-integration we get $V(\Phi)$ and from it the two equations,

$$k_0^2 - \frac{1}{2}Z_r'(v_0/\sqrt{2}) = B,$$

(7)

$$-V(\Phi) = \frac{k_0^2}{2}\Phi(\Psi - \Phi) + \frac{B}{2}\Phi^2\left(1 - \sqrt{\frac{\Phi}{\Psi}}\right).$$

(8)

In Eq. (7) B is defined by

$$B = \frac{16}{15}b(\beta, v_0)\sqrt{\Psi},$$

(9)

and represents the condition (ii) and is called the nonlinear dispersion relation since it determines v_0 in terms of B and k_0, i.e., the phase speed v_0 is a derived quantity here. (We note in parenthesis that Eq. (7) is called later more precisely quasi-nonlinear dispersion relation. The reason is that k_0 is generally not identical with k, the wave number of the periodic structure. Only in the harmonic or near-harmonic limit both coincide.[4,10]) Equation (8) determines the spatial wave form (or spectral content), as mentioned. It is constrained by (i). Note that in the present theory of a zero background current there are three independent parameters, k_0, B, and Ψ, which control everything. The other two parameters are derived ones: v_0 is determined by Eq. (7) and β follows from the definition of B, namely from Eq. (9) with Eq. (6).

As was shown in Ref. 21 one can incorporate a further free parameter, the drift velocity v_D between electrons and ions in an unperturbed, current-carrying plasma and hence extend the present theory by a fourth independent parameter, allowing a broader physical application. To catch this case, we have simply to replace v_0 in Eqs. (7) and (9) by $v_D - v_0$. Strictly speaking, we in addition have to assume that ion trapping effects are negligible and that the phase velocity is well above ion sound velocity. If the latter assumption is not made, allowing for finite ion mass (and temperature) effects, we get an additional term on the l.h.s. of Eq. (7), given by $-\frac{\theta}{2}Z_r'\left(\sqrt{\frac{\theta}{2\delta}}v_0\right)$, where $\theta = T_e/T_i$ and $\delta = m_e/m_i$.[10] It can be interpreted as an ionic shielding term. We shall come back later in Sec. VI to this situation.

We mention another generalization. Replacing $f_M(\xi)$ by an arbitrary free particle distribution $f_0(\xi)$ and $\frac{1}{\sqrt{2\pi}}\exp(-v_0^2/2)\exp(\beta\xi^2/2)$ by an arbitrary trapped particle distribution $f_t(|\xi|)$, Eq. (2) can be generalized to

$$f(x, v) = (1 + k_0^2\psi/2)[\theta(\epsilon)f_0(\xi) + \theta(-\epsilon)f_t(|\xi|)],$$

(10)

where ξ is the generalized velocity, defined by $\xi := \sigma\sqrt{|2\epsilon|}$.[21,23] By Taylor expansion around $\xi = 0$ it is then found, see e.g., Appendix of Ref. 21, that $b(\beta, v_0)$ becomes

$$b(\beta, v_0) = -\sqrt{2}[f_0''(0) + f_t''(0)].$$

(11)

The density expression generalizing Eq. (5) is then obtained by the new expressions, provided that continuity at the

separatrix, $f_0(0) = f_t(0)$, is assumed. If we lifted the latter assumption, allowing a jump of f at the separatrix, an additional term of $O(\sqrt{\phi})$ in the density expression would arise, changing the entire analysis and physical background. Equations (7) and (8), which are our main result, will be analyzed further in the following sections. Before this, however, we call attention to the relationship between linear and nonlinear wave analyses, especially in the infinitesimal wave amplitude limit $\Psi \rightarrow 0^+$.

III. INTRINSIC NONLINEARITY

First, as can be seen from the $n(\Phi)$ and $V(\Phi)$ expressions, the independent parameter k_0 introduces a spatial periodicity, noting that it holds $\Phi''(x) = k_0^2 \Psi$ at potential minimum, where $\Phi = 0$. The curvature is, therefore, nonzero there and positive as long as $k_0 \neq 0$, representing a finite wavelength structure. (There is one exception to be mentioned later.)

Second, Eqs. (7) and (8) have a very appealing simple form, reminiscent of the van Kampen description[24] of linear waves, more concretely of the wave continuum found there. In the latter, the perturbed distribution function is given by (see also Ref. 10),

$$f_1(x,v) = \left[-P \frac{f_0'(v)}{v - v_0} - \lambda \delta(v - v_0) \right] \Phi(x). \quad (12)$$

We would, however, be too rash in our judgement if we stopped here and declared the present structures as the nonlinear extension of van Kampen modes, as found in the literature. The truth is somewhat more subtle, demanding a closer look especially at the microscopic details.

First of all, van Kampen modes or wave packets composed of a superposition of van Kampen modes (or similar linear modes, such as the damped modes of Ref. 25) do not solve the correct equation but an equation, which has been truncated by linearization. The common belief is that in the infinitesimal amplitude limit this does not matter. That this is not so is seen from the corresponding DF at resonant velocity (see Fig. 2). The DF for a linear wave packet is composed of more or less singular functions, representing partially ballistic particles, and is hence rough and filamentary. Such wave packets generally decay in time due to phase mixing exhibiting Landau damping, as Ref. 26 has pointed out. Even if one relaxes this severely restricted construction by choosing less singular, or nonsingular, but still linear functions, as was done by Ref. 25, a damping remains albeit less strong.

To obtain zero-damped structures the DFs have to satisfy the full Vlasov equation without limitation. Moreover, they should be smooth and continuous to be physically meaningful, such as the asymptotic coarse-grained distributions of Ref. 27 in the nonlinear Landau damping scenario (Ref. 28). Our distribution $f(x,v)$ in (2) is of this quality.

In other words, although from a macroscopic point of view our small amplitude solutions are reminiscent of van Kampen modes, the underlying microscopic physics indicates a qualitative difference with respect to the latter, which cannot be removed by taking the infinitesimal amplitude limit, a conclusion first drawn by Ref. 29.

We now deepen this conclusion by collecting further evidence of the nonlinear nature of the present structures. We shall call them PROOFs to be interpreted in the theoretical physical sense. They are not independent of each other but illuminate the subject from different viewpoints.

Proof 1. Consider the most simple case, $B = 0$, which corresponds to monochromatic, or as we say harmonic waves. In this case our solution exactly reproduces the van Kampen solution, as long as the focus is restricted on Eqs. (7) and (8) only. Microscopically, however, although the most singular part, the δ-function contribution to f_1 is absent in van Kampen's solution, there still remains the
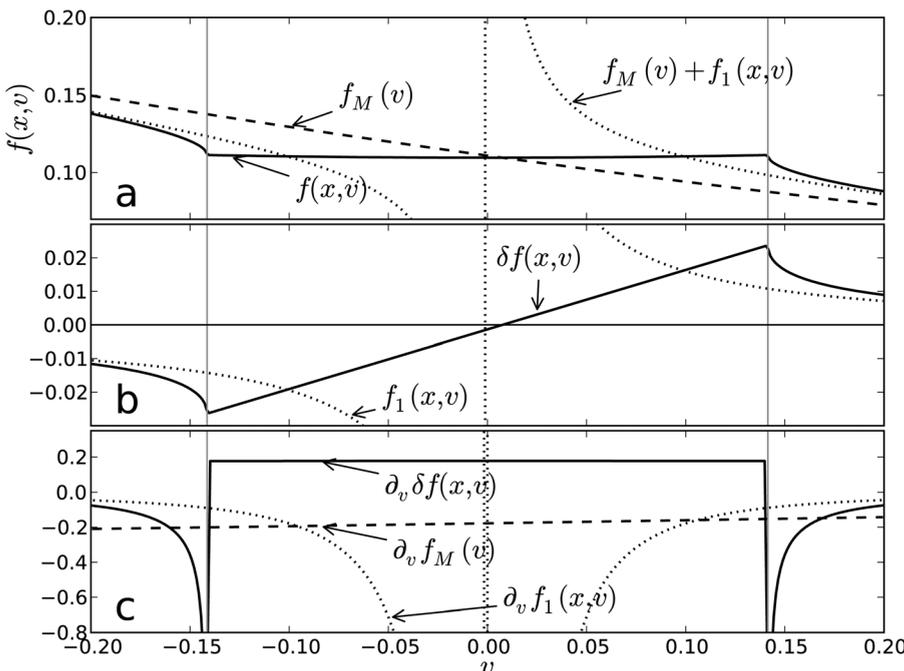

FIG. 2. Zoom on the resonant distribution functions of linear and nonlinear theory in velocity space. (a) The unperturbed (Maxwellian) velocity distribution function $f_M(v)$, the van Kampen's linear distribution function $f_M(v) + f_1(x,v)$ as well as the nonlinear distribution function $f(x,v)$ as given by Eq. (2) with $\Psi = 1/100$. Note the divergence of the linear perturbation and the trapped range of $f(x,v)$ in $(-\sqrt{2\Psi}, \sqrt{2\Psi})$. (b) Perturbations to the Maxwellian in the linear and nonlinear theories. (c) Velocity derivatives of the Maxwellian and the linear and nonlinear perturbations.

singularity stemming from the principal value integral, whereas our distribution is smooth and well behaved. The latter is seen very easily if we anticipate the results from the next section. For $B = 0$ the phase velocity of the (slow) harmonic waves lies in the range $1.307 < v_0 < 2.13$, from which we can deduce that the trapped particle parameter β, given from Eqs. (6) and (9) by $\beta = 1 - v_0^2$, is negative, yielding a depressed trapped particle region and an altogether well behaved distribution. Note that B can become zero without taking the infinitesimal amplitude limit $\Psi \to 0^+$.

A typical distribution function $f(x, v)$ for such a harmonic wave at potential maximum is plotted in Fig. 2(a) (solid line) together with the Maxwellian (dashed line) and the corresponding van Kampen distribution (dotted line). The vertical dotted line indicates the location of the phase velocity. (For simplicity the normalization factor $1/\sqrt{2\pi}$ has been replaced by unity.) The selected numbers are $v_0 = 1.6$ and $\Psi = 1/100$, corresponding to $k_0 = 0.354$ and $\beta = -1.56$. For this β the dip in the trapped particle distribution, albeit present, is almost invisible. Fig. 2(a) shows that $f(x, v)$ is smooth and well-behaved contrary to van Kampen's distribution with its giant principal value singularity at phase velocity (not to speak about its negativity in the lower velocity range close by).

We conclude that even in the monochromatic wave limit, where on the macroscopic basis both solutions, the van Kampen mode solving the truncated linearized Vlasov-Poisson (lVP) system and our sinusoidal EH solution solving the full Vlasov-Poisson system, cannot be distinguished, there are obvious differences on the microscopic level and on the time-behavior of the wave packet when subject to the correct evolution equation (see later). Since the physics takes place in phase space, it is our solution which deserves the term "physical" rather than the van Kampen solution. The merit of the latter, in spite of its deficiency not to be a solution of the correct equation, is that it comes closest to the real one, as long as one ignores microscopic and temporal details.

Since this is the simplest wave structure to be analyzed most easily, let us, therefore, repeat our main assertion: whereas our solution is stationary and well-behaved in phase space, the van Kampen solution, albeit stationary within the lVP-system, is ill-behaved in phase space and will decay in time, because it has to be subject to the full VP-system. There is obviously no link connecting the two worlds, the function spaces of lVP and of VP.

Proof 2. A similar conclusion can be drawn for $B \neq 0$, which reads in van Kampen's theory $\lambda \neq 0$. Through it one can shift the phase velocity to an arbitrary value, producing thereby the van Kampen wave continuum. A single mode is now faced with the stronger δ-function singularity. For a wave packet, obtained by a superposition of such modes of the same phase velocity, the distribution function will be less singular (or perhaps be even nonsingular) but its roughness will remain, being associated with damping, e.g., by phase mixing, when subject to real physics. This is in contrast to our solution, which is smooth and well-behaved and does not decay in time. From the next section we will see that a concrete wave form and a concrete phase velocity v_0 are associated with a given k_0 and B. For a given B and v_0, the

parameter β, which is determined through Eqs. (6) and (9), will generally be negative, giving rise to a smooth, depressed distribution. It will still depend on the amplitude Ψ and even in the infinitesimal amplitude limit, when β becomes proportional to $-1/\sqrt{\Psi}$, the distribution function will remain smooth, exhibiting a tiny dip at resonant velocity merely. There is no unusual or unphysical behavior associated with.

In the superposition of van Kampen modes of the same phase velocity v_0 there is an exceptional case, namely when the expansion coefficients are chosen such that they exactly reproduce the actual CEH structure $\Phi(x)$, see Ref. 10. With this one again has a macroscopic identity with Eqs. (7) and (8) and hence again closest approach between both descriptions; the microscopic distinction and the different t-behavior, however, remain. And, of course, the very specific values of the infinite Fourier expansion coefficients, which decrease slowly and alternate in sign with increasing harmonic number, cannot be suggested without the knowledge of the CEH solution. There is hence no chance to meet our physical structures by a method different from the actual one, which not only applies for linear wave packets but also for the nonlinear solutions obtained by the original BGK method.[5] (Some more details are presented in Sec. IV.)

Proof 3. Let us define the nonlinear perturbation of the distribution function by $\delta f := f(x, v) - f_M(v)$, where f_M is again the unperturbed shifted Maxwellian. Taking its velocity derivative, it is easily seen that there are regions in phase space where the inequality $|\partial_v \delta f| \ll |\partial_v f_M|$ is violated, as would be required for linear wave theory to be applicable. Near the center of the trapped region at constant x, for example near $\epsilon = -\Phi$, the derivatives are indeed of equal size since $\partial_v f(x, v)$ vanishes or is very small there. For an illustration, coming back to the harmonic wave in Fig. 2, we plot the nonlinear δf (solid line) and the linear f_1 (dotted line) in Fig. 2(b) and the velocity-derivatives of δf, f_M, and f_1 in Fig. 2(c). We see that $|\partial_v \delta f|$ is of equal size as $|\partial_v f_M|$ and, more seriously, that $|\partial_v f_1|$ becomes giantly large and singular near the phase velocity, such that its neglect in a linear wave theory is more than suspect. The anomaly of the velocity derivative of δf just outside the trapped range in Fig. 2(c) is a relic of the $\sqrt{\epsilon}$ dependence of the free particle distribution function and appears for any chosen function as long as $v_0 \neq 0$. Since it is restricted to a small area of $\epsilon = 0^+$ it will be wiped out by processes like coarse graining or faint collisions and is hence not considered as an obstacle for the physical relevance of our present structures. On the other hand, these processes, when applied to the linear distribution, will not prevent the wave from damping. The reason is the trapping oscillations performed by the free particles after capture, which extract wave energy in favor of particle energy,[27] whereas in the nonlinear theory the trapped particle region represents a "buffer" for balancing the energy exchange during the wave-particle interaction.

We infer that the term $\Phi'(x)\partial_v \delta f$ in the Vlasov equation is crucial and nonnegligible for this kind of structures.

Proof 4. Another proof is of numerical nature and has been given by Refs. 21 and 11. If we take our solution as initial data and solve the Vlasov-Poisson system in time, the full Vlasov-Poisson solver exactly reproduces the

time-stationary solution, whereas the linearized Vlasov-Poisson solver provides a damped solution only. Not only the correct description in phase and real space is necessary but also the correct evolution equation in time must prevail. A linear code does not know how to handle a correct wave input precisely.

Proof 5. Moreover, it could be shown again numerically in Refs. 11 and 30–35 that in a current-carrying plasma, i.e., when a drift between electrons and ions exists in the unperturbed state, the system can be nonlinearly unstable by the excitation of electron and ion phase space holes, even in a linearly stable situation, i.e., below threshold of linear two-stream instability. This nonlinear instability preferentially excites holes of zero-or negative energy being triggered by seed fluctuations in a noisy plasma background and requires mobile ions to occur,[11,32,33,35] as given in a real, driven plasma.

Proof 6. And last but not least, coming back to our starting equations, we can provide a more direct proof of the intrinsic nonlinearity. In a typical cnoidal hole structure, all three terms in Eq. (7) are finite and of the same order. Hence, k_0^2 and B are comparable quantities and the last term in Eq. (8), which reflects nonlinearity, is of the same order as the preceding linear term, namely $O(\Psi^2)$, a fact that remains true up to the infinitesimal amplitude limit. There is no threshold value for Ψ below which the last term would be negligible and hence linearity would prevail. This can also be seen from the density expression (5), in which all three perturbed terms contribute at the same level, namely at $O(\Psi)$, noting that with (9) it holds $b(\beta, v_0) \sim B/\sqrt{\Psi}$.

In the following sections, we will give further arguments in favor of nonlinearity valid even for a weakly dissipative kinetic system.

In conclusion, linear wave theory is generally nonsuited as a starting point for the investigation of plasmas that are excited by long-living electron hole structures. Plasmas excited by such persistent trapped particle structures are not only kinetic but also fundamentally nonlinear.

Third, there is another point worth to be mentioned. The simple form of the determining equations (7) and (8) is achieved by elimination of v_0 from $V(\Phi)$ and using Ψ as the independent external parameter, as mentioned. Often other ways of dealing with trapped particle solutions of the one or the other sort are found, which use v_0 instead of Ψ as the external parameter and focus on the wave profile, ignoring the second part, the solution of the NDR. They unfortunately appear, as a consequence, incomplete and less transparent, due to this different handling of the VP system. Moreover, they are often faced with the problem that ψ comes out O(1) even though a small amplitude theory is being intended.

IV. WAVE PROPERTIES IN DETAIL

The region in the (B, k_0^2)-parameter space for which coherent wave solutions exist is limited through the solubility of the NDR (7) and the positivity of $-V$ in $0 < \Phi < \Psi$. It is found to be

$$k_0^2 = 0 \quad 0 < B \leq 1, \tag{13}$$

$$0 < k_0^2 < 0.095 \quad -2k_0^2 \leq B \leq 1 + k_0^2, \tag{14}$$

$$0.095 \leq k_0^2 \quad k_0^2 - 0.285 \leq B \leq 1 + k_0^2. \tag{15}$$

The corresponding wave spectrum includes solitary electron holes ($k_0^2 = 0$) given by Refs. 3, 4, 9

$$\Phi(x) = \Psi sech^4\left(\frac{x}{L}\right), \tag{16}$$

with a half-width of $L = 4/\sqrt{B}$, $B > 0$ purely harmonic waves ($B = 0$, sometimes mistaken for linear waves),

$$\Phi(x) = \frac{\Psi}{2}[cos(k_0 x) + 1], \tag{17}$$

as well as a special solitary potential dip (SPD; $B = -2k_0^2$), represented by the special pseudo-potential,

$$-V(\Phi) = \frac{k_0^2}{2}\left(\Phi(\Psi - \Phi) - 2\Phi^2\left(1 - \sqrt{\frac{\Phi}{\Psi}}\right)\right), \tag{18}$$

and given by Eq. (3.33) in Ref. 21. (We quote in parenthesis that the $sech^4$ expression for a solitary trapped particle structure has first been obtained by Ref. 36 for small amplitude ion bipolar structures.) All the other waves are periodic in space and are represented by Jacobian elliptic functions,[4,21] such as $cn(x)$, justifying the notation cnoidal electron holes (CEHs).

They can be classified by means of the parameter S defined by $S := B/k_0^2$, also called "steepness parameter," because it stands for the steepness or distortion of the wave form from sinusoidal shape, the latter being given by $S = 0$. A constant S would be a straight line going through the origin in the (B, k_0^2) space, i.e., in the existence diagram presented later. The range of S is thereby bounded by $-2 \leq S$, where $S = -2$ and $S \to \infty$ represent the solitary SPD and SEH structures, respectively. It is seen later that $S > 0$ is a rarefactive wave, having an extended potential valley, whereas $-2 \leq S < 0$ is a compressional wave with an extended potential hill region.

Of special interest are periodic, rarefactive solutions close to and including SEH. It holds for $1/4 \leq S$,[4,21]

$$\Phi(x) = \Psi\left[1 - K\frac{1 - cn(u|m)}{1 + cn(u|m)}\right]^2, \tag{19}$$

where we have defined

$$K = \sqrt{1 + 2/S} \qquad m = \frac{1}{2}\left[1 + \frac{1 + 2S}{2\sqrt{S(S + 2)}}\right]$$

$$u = [S(S + 2)]^{\frac{1}{4}}(k_0 x)/2.$$

Fig. 3(a), in which $\Phi(x)/\Psi$ is plotted, gives an impression of that structure for $S = 8$, $B = 1$, $k_0^2 = 1/8$ and $v_0 = 0.37$, corresponding to $K = 1.118$, $m = 0.975$, and $u = 0.53\ x$. We recognize a periodic potential structure with a somewhat broadened valley and a weakly contracted hill. Here we have assumed that the periodicity interval (box) is repeated infinitely, which implies identical trapping conditions for each

repeated interval. Since, however, each potential hump is isolated from its neighbor concerning electron trapping and because in different regions usually different trapping scenarios are in action during the creation of the structure, a wavelet structure, as indicated in Fig. 3(b), may be more typical. In such a wavelet structure, all humps do have the same values of k_0, B, and v_0, i.e., move with the same phase velocity and hence the "wave packet" will not disperse. Nevertheless, each hump can have its individual pair of (ψ, β)-values belonging to the given B and v_0 according to Eq. (9). The selected values for a five hump, symmetric wavelet structure, as shown in Fig. 3(b), are $S = 16$, $B = 1$, $k_0 = 0.25$, and $v_0 = 0.3$ corresponding to $K = 1.061$, $m = 0.986$, and $u = 0.515x$. For the central peak we have chosen: $(\psi_0 = 1/100, \beta_0 = -16.5)$, whereas the two neighboring peaks are determined by $(\psi_1 = 2\psi_0/3, \beta_1 = -20.4)$ and $(\psi_2 = \psi_0/3, \beta_2 = -29.2)$, respectively, but any pair of (ψ, β) could have been chosen as well, as long as they follow from Eq. (9) for given B and v_0. We like to stress again that all humps in Fig. 3(b), which in the general case have not to be distributed symmetrically, have the same phase velocity v_0 such that the velocity of the whole packet is still given by v_0 as in the unmodulated cnoidal wave case, Fig. 3(a). The "group velocity" of this wavelet is still given by v_0 and has hence nothing to do with the standard group velocity concept, as described more thoroughly in Sec. VI C. We also like to mention that the amplitude of a hump is smaller, the deeper the trapped electron region is excavated, a fact to which we will discuss later in Sec. VI B.

We notice that a CEH wavelet, defined this way with only few humps (2, 3, 4... humps) and abbreviated for the sake of brevity by "CEHWL," represents a new localized nonlinear trapped particle state in addition to the SEH in the rarefaction regime. (We remember that in the compressional regime another localized member exists, the SPD solution.) That a different trapping state and a different amplitude can still yield the same phase velocity is not new and has already been found for SEHs in Ref. 37 (see, e.g., Fig. 3 of that paper). It is expected that both structures, which are distinguished by the parameter k_0, i.e., essentially by the shape, can be excited under the same conditions, and may, therefore, appear in company. In Sec. VI B we will present first examples of their simultaneous existence in space and simulations.

Note that the SEH, obeying $S \to \infty, k_0^2 \to 0$ and $\sqrt{S}k_0 = \sqrt{B}$, is simply obtained from Eq. (16) by setting $K = 1, m = 1, u = \sqrt{B}x/2$, and $cn(u|1) = sechu$.

From these expressions for $\Phi(x)$, which are too sophisticated to be guessed *a priori*, it becomes clear that the BGK method,[5] which presumes the knowledge of $\Phi(x)$, is not suited for finding our solution. Only the kinetically generalized pseudo-potential method of Refs. 4, 10, and 23 is able to provide it. A BGK solution is typically discontinuous at the separatrix or the searched trapped particle distribution f_{et} possesses a singularity or a singular velocity derivative there,[10,16,23] which would be the case if we, for example, had chosen the function $e^{\beta\sqrt{-\epsilon}}$ in the second part of Eq. (2) instead of $e^{\beta(-\epsilon)}$. More seriously, f_{et} may even be negative.[38] In other words, such a f(x,v) would be less smooth and less physically meaningful. The usage of the terminology "BGK holes" for the present structures would, therefore, be too crude and undifferentiated. With respect to regularity deficits of BGK holes we refer to a forthcoming publication.[39]

Fig. 4 shows the existence diagram of nonlinear wave solutions in the (B, k_0^2)-parameter space, which is bounded by thick solid lines, representing the inequalities (13–15), as mentioned. Waves with constant phase velocity v_0 lie on a straight line defined by

$$D := B - k_0^2 = const. \qquad (20)$$

Two wave branches can be distinguished, separated by $v_0 = 2.13$ (resp. $D = -0.285$). The fast branch with
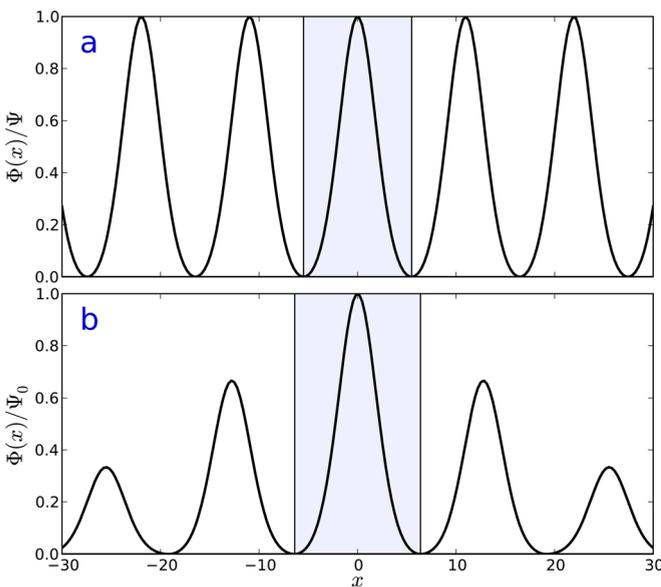

FIG. 3. (Color online) Cnoidal wave structures. In (a) a periodic train of cnoidal EH structures is shown, all humps being characterized by the same set of parameters. In (b) all humps have the same speed but different amplitudes and trapping parameters. This often met structure in experiments and simulations is called cnoidal electron hole wavelet (CEHWL).
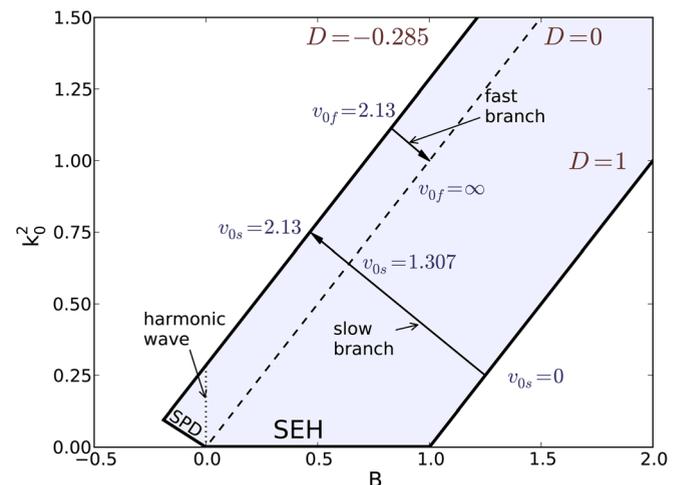

FIG. 4. (Color online) Existence diagram: solutions exist in the shaded region, i.e., between the lines $D = 1$ and $D = -0.285$ and the lines of SEHs and SPDs. Lines of $D = constant$ mark waves of the same phase velocity v_0.

$2.13 < v_{0f}$ includes the classical Langmuir branch modified by trapping, and the slow branch with $0 \leq v_{0s} \leq 2.13$ represents phase velocities in the thermal range. Whereas Langmuir waves survive in the linear limit (as indicated later in Fig. 5 for $B = 0$), the slow branch as a linear solution does not; for its existence trapping, and hence nonlinearity, turns out to be essential.

Fig. 5 shows the quasi-NDR, $\omega_0 := k_0 v_0$ as a function of $\sqrt{2}k_0$ for the spatially periodic CEHs. Lines of constant B are drawn. Of special interest is the line $B = 0$ not only because it includes the Langmuir branch, as predicted by Ref. 40, but also because it separates the harmonic waves ($B = 0$) from the multi-harmonic nonlinear wave branches ($B \neq 0$). B is, therefore, also called spectral parameter. Note that for $B > 0$ and the fast branch there is a cut-off at lower k_0 given by $\sqrt{B} < k_0$ which is determined by $D = 0$. The dashed line, defined by $D = -0.285$, is the separating line between the fast and slow branches. For $B < 0$ no solution exists for $k_0 < \sqrt{\frac{-B}{2}}$, i.e., below the SPD solution, which is indicated by crosses. The minimum possible B is given by $B_{min} = -0.19$, which is located at $\sqrt{2}k_0 = 0.436$.

An important contribution of the trapping nonlinearity to the spectrum of electrostatic waves are the acoustic modes of the slow branch and $B \geq 0$, as these modes entirely lie in the thermal wing of the distribution function and exist due to a proper modification of the latter at phase velocity, ruling out Landau damping. Because of their relevance in experiments (see also later) they have gotten an extra notation,[3,9,41,43] and are termed slow electron acoustic waves (SEAWs). They are thus the periodic extension of SEHs and are sometimes referred to as the acoustic modes underlying solitary electron holes.

Note that the general expression of the electron density in terms of (B, k_0^2) is given by

$$n_e(\Phi) = 1 + \frac{k_0^2}{2}\Psi + (B - k_0^2)\Phi - \frac{5B}{4\sqrt{\Psi}}\Phi^{\frac{3}{2}} + \cdots \quad (21)$$
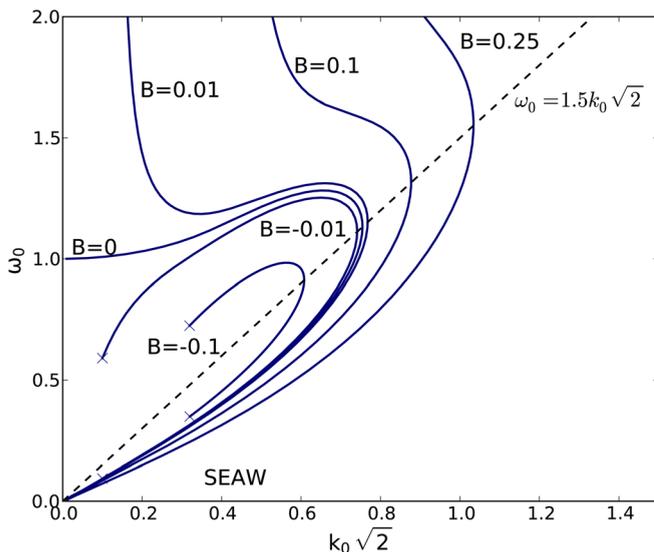

FIG. 5. (Color online) Quasi-nonlinear dispersion relation (NDR) with $\omega_0 := k_0 v_0$.

From this follows that all structures are rarefactive in the sense that $n_e(\Psi) < n_e(0)$. Since, however, the SPD has zero potential at its center its central density is increased in comparison with its neighboring values, especially with its asymptotic value, where $\Phi = \Psi$ and $n_e(\Psi) = 1$. The SPD structure with its opposite polarity is hence effectively compressional. More generally, the dividing line between rarefactive and compressional waves in the latter sense is essentially given by $B = 0$, the harmonic wave line or by $S = 0$, as mentioned.

To see the consequences of the slow branch for the particle distribution function, e.g., in case of a SEH with $k_0 = 0$, take a B with $0 < B \leq 1$, the corresponding v_0 with $0 \leq v_0 < 1.307$ and a Ψ satisfying $\Psi \ll 1$ and insert them into β, which follows from Eq. (9) with Eq. (6):

$$\beta = 1 - v_0^2 - \frac{15}{16}\sqrt{\frac{\pi}{\Psi}}B \exp\left(\frac{v_0^2}{2}\right). \quad (22)$$

It is found that β is negative in this domain and hence stands for a depression of the distribution function at phase velocity.

(To be somewhat more concrete, take $\psi = 1/100$ and get: $\beta = 15.6, 14.3, 12.3, 9.7, 6.2,$ and 0.7 for $B = 1, 0.8, 0.6, 0.4, 0.2,$ and 0, respectively. The corresponding v_0-values are $v_0 = 0, 0.5, 0.71, 0.89, 1.1,$ and 1.307. With increasing phase velocity β decreases magnitude-wise, i.e., the dip becomes less pronounced.)

On the other hand, if we are in the compressional region $B < 0$ within the allowed range (the triangle in Fig. 4) and take the fast branch with a velocity well above 2.13, we would find from Eq. (22) a very large positive β, corresponding to a beam-type distribution function. A stationary, compressional, nonlinear Langmuir wave, therefore, would need a trapped beam to become nonlinearly existent. This result should, however, be treated with caution since typically $|-\beta\epsilon|$ would no longer be small, such that the Taylor expansion used to get Eq. (5) breaks down. For the fast branch a full amplitude analysis, such as in Ref. 4, is usually called for.

We finally mention that within our restriction of immobile ions and of small amplitudes, a beam-type electron distribution admits (besides periodic waves) the SPD solution only, a DL solution being not included. To have access to the latter, we must allow beam-type mobile ions and finite amplitudes, as shown by Refs. 43 and 44 for the so-called strong DL.

V. MISCELLANEOUS

We are aware that the offered solutions represent only a small section in the general class of nonlinear wave solutions of the VP system, being characterized by a simple $(x - v_0 t)$-behavior and restrained further by treating electron dynamical effects only. Since our main concern is to show that solutions are thrown away by linearization, which are then missing in a general plasma dynamical context, we have concentrated on this most simple configuration because already there the primacy of the trapping nonlinearity in comparison

with the ordinary nonlinearity, being quadratic in the wave amplitude, comes out relatively easily, as shown.

On the other hand, since there exists already a large body of physical and mathematical literature dealing with similar aspects, we are aware that at least some of them should be mentioned in a review. Therefore, to narrow this gap, and thereby broaden the basis, we refer to some adjacent topics, such as accessibility and stability of these structures, higher dimensional and magnetic effects, etc., and give some further experimental demonstrations of these structures in different physical environments.

Of course, a more general review should also cover simulations and a still deeper look at the space and laboratory observations. Unfortunately, such an undertaking would need much more effort and time. We, therefore, apologize to those whose work has not or not adequately been mentioned and refer to some already existing reviews such as Ref. 11 or 12.

A. Access, existence, and more dynamics

A plasma driven by an external agent shows, depending on the driving strength, typically a rather complex space-time behavior that is characterized by many spatial and temporal scales. In such a general picture a solution seems to be out of reach such that simplifications are unavoidable.

One approach is that after the elapse of some time a dynamical evolution emerges, which is characterized by long scales only. As shown in Ref. 45 (see also Ref. 9) a coherent wave in this long time regime is governed dynamically by the Schamel equation, being a modified Korteweg-de Vries-type evolution equation, in which the square nonlinearity is replaced by the stronger 3/2-power trapping nonlinearity. Solitary EHs propagating near the critical velocity $v_{0c} = 1.307$ are found to be stationary solutions of this wave equation. Moreover, since this equation is not completely integrable, as suggested by the works of Refs. 46–48, the coalescence of SEHs, copropagating with nearly the same speed as seen in simulations and experiments (see, e.g., the Risø experiment[2]), does come as no great surprise.

A different Ansatz for solving the VP system as an initial value problem has been employed by Ref. 49. These authors developed a new procedure for the analysis of the long-time behavior for a certain class of initial conditions by decomposing the VP problem into a transient part and a time-asymptotic part. Focusing on the time-asymptotic part they could solve a corresponding bifurcation problem with the initial condition and the transient field playing the role of parameters. Their main result is that *if* the VP system possesses a nonzero small-amplitude time-asymptotic solution the corresponding field is given at leading order by a superposition of traveling-wave modes associated with the roots of a "time-asymptotic" Vlasov dispersion relation. In other words, they underpinned mathematically the existence and accessibility of our present traveling modes in the $B = 0$ limit (if one accepts a kind of coarse graining for the latter). For some reason, however, they miss the $B \neq 0$ branch, which contains as important members the physically relevant SEHs and CEHWLs, respectively. A presumption is that the

chosen C^1-class of functions may be too narrow to encompass also the $B \neq 0$ class, noting that our functions do not fall within this C^1-class. In this respect we refer to another mathematical paper published recently by Lin and Zeng[50] in which a Sobolev space with a low regularity requirement has been chosen instead to approve the absence of nonlinear Landau damping and the existence of long term BGK-like states for a certain class of initial conditions, the latter being in accord with our numerical findings, presented, e.g., in Ref. 35, in which non-Landau damped initial fluctuations in a noisy plasma are found to be the cause of our states.

Note also that the time-asymptotic superposition of phase-space holes (or streets of them) in Vlasov plasmas is nowadays a well established numerical fact.[51–53]

B. Stability and magnetic effects

This is a vast field, which can be considered in a few points only. An important issue is the stability of CEHs, especially of SEHs.

Theoretically, the stability analysis of inhomogeneous phase space equilibria is a formidable task because integrations along non-straight line particle orbits (characteristics of the Vlasov equation) are invoked, which are hard (if not impossible) to be performed analytically. Again, simplifications seem to be unavoidable.

An outcome in case of linear transverse instability of a 1D, solitary EH in an unmagnetized plasma ($B_0 = 0$) has been found by the author[54] (see also Ref. 9) by making use of a general framework developed by Ref. 55 (see also Ref. 56). Solving a nonlocal linear eigenvalue problem in the so-called fluid limit and subsequent truncation (for which a justification is missing), he found longitudinal stability but unconditional, transverse instability of a SEH that explains the numerical observations of Refs. 57 and 58, namely the disappearance of vortex structures in spatially higher dimensions. (We add in parenthesis that a CEHWL instead of a SEH may turn out longitudinally unstable in 1D because of the coalescence instability, being a slow process.)

In a magnetized plasma ($B_0 \neq 0$) the situation is still more complex because, besides the adaptation of the equilibrium to this case (see later), particle orbits exhibiting now both bouncing and gyration are involved. A numerical approach to this case has been presented in Ref. 59, which simulated with a particle-in-cell (PIC) code the transverse (in)stability of a non-propagating BGK electron hole. Their general findings for an anisotropic background plasma ($T_{e\perp} < T_{e\parallel}$) are that SEHs are stable against transverse perturbations provided that $\Omega_e > \omega_{be}$, where the cyclotron and bounce frequency, respectively, are given by $\Omega_e = eB_0/m_e/\omega_p$ and $\omega_{be} = \sqrt{\frac{\psi}{L^2}} = \frac{\sqrt{B\psi}}{4}$, respectively, ω_p being the plasma frequency and ψ, L the amplitude and width, respectively, of a SEH (see Eq. (16)). This means that lowering the product $B\psi \propto (1 - \beta - v_0^2)\exp(-v_0^2/2)\psi^{3/2} > 0$ increases the stability of a SEH in a magnetized plasma, i.e., a smaller amplitude, a larger propagating speed (lowering $B < 1$, see Fig. 4) and a less deep hole in the distribution

favor stability. The stability increases further when the plasma becomes more isotropic.

A partial confirmation of this numerical result could be given in Ref. 60, which studied the stability of a 1D SEH in the presence of a background magnetic field, using a fully linearized Vlasov description for the small perturbations. They found within a broad range of plasma parameters a destabilization through the resonant interaction of the wave with the trapped and free particles. In case of the high-frequency cyclotron resonance, however, i.e. when the wave frequency ω, being within the bounce frequency ($\omega \approx \omega_{be}$), is close to the gyrofrequency Ω_e, a transition from instability to stability could be seen when ω_{be} becomes smaller than Ω_e, in accordance with the numerical results of Ref. 59. For low-frequency resonances, an instability was always seen, leading to the conclusion that long-living structures observed by the satellite missions are likely to be fully 3D (or quasi 3D) and that for their stability behavior probably the full $\perp$ dynamics of electrons and ions needs to be accounted for.

The often met belief in a complete disintegration of 1D SEHs in higher dimensional, magnetized plasmas must hence at least be modified and be restricted to low frequency resonances. By the way, this finding with respect to high frequency perturbations is confirmed by the observations made in space, namely that the most intense fields are observed close to the Earth where the magnetic field is strongest, as pointed out in Ref. 59. The former is also in accord with earlier simulations of Ref. 61, where two classes of coherent structures were seen. While the large structures were quickly unstable, the weaker ones were stable. A similar finding was more recently reported in Ref. 62.

So, even in weakly magnetized plasmas ($\Omega_e < 1$), EHs might exist as weak structures, although not necessarily as 1D structures.

This brings us to the next issue, the construction of 2D and 3D EH equilibria in magnetized plasmas.

When the magnetic field is very strong (effectively infinite) the 1D EH solution is mostly applicable, not only because of its proved stability but also because of its very definition (particle orbits in the 2D phase space (x,v)). For weaker magnetic fields corrections are indispensible.

This is of course a difficult task again and can only be performed approximately.

One reason is that in the presence of B_0 and higher dimensions further conserved quantities, other than the energy and the velocity sign of untrapped particles, are needed, which are, however, virtually impossible to calculate, since the spatial distribution of the electrostatic potential is not known in ad\roundvance to get a solution of the Vlasov equation in the 6D phase space. A simplification is to reduce the dimensions by invoking the drift-kinetic approximation for the electrons, as done in Ref. 14. Looking for traveling wave solutions and utilizing an electron distribution function similar to Eq. (2), these authors were able to derive and solve a generalized Poisson's equation valid for SEHs in a magnetized plasma. Two different solutions could be found and explored numerically. One depends only on $(z, v_{\parallel})$ and $R(x, y)$, where z is now the coordinate along the straight-line magnetic field and x, y are the perpendicular

coordinates. This so-called "parallel solution" is found to be ellipsoidal in shape with comparable size of the $\perp$ and $\parallel$ scales, when $\Omega_e > 1$ (e.g., in Earth magnetosphere). When $\Omega_e < 1$ (e.g., deep in the magnetotail) the perpendicular scale is found larger by the factor $1/\Omega_e$. The second solution, called "oblique solution," is 2D in a reference frame that is tilted relative to the z axis. In this case the parallel electron momentum provides a further invariant. This oblique 2D solution is cylindrical in shape with a larger $\parallel$ than $\perp$ scale. For more details, especially for the interpretation of satellite data, we refer to the original publication and to the review of Ref. 12. Note that in these solutions the parameters are chosen such that ion dynamics is negligible, such as in the whistler or upper-hybrid frequency regime.

This is different for holes being related to the fast magnetic reconnection process, as discussed by Jovanović and Shukla.[63–65] Presenting a drift-kinetic model for the electron dynamics, which is, specifically for the lower hybrid regime, coupled with the linear dynamics of unmagnetized ions, and solving the former again by taking into account distributions of type (2), they were able to derive and solve numerically two nonlinearly coupled equations, for the scalar and the vector potential. With this they could show that nonlinear currents and charges can efficiently mediate the topological transformations of magnetic field lines, yielding a chain of magnetic islands coupled with a double chain of hydrodynamic vortices.[12] We shall come back in Sec. VI B to the magnetic reconnection problem again.

To summarize, holes of higher dimension in magnetized plasmas definitely would provide a better approach to realistic plasmas than the 1D holes, which are strictly appropriate for unmagnetized or strongly magnetized plasmas only. Since, however, the available analytical description of them is less rigorous and since their description is based on the same fundament (quasi-potential method with distributions of the kind (2)), it seems at least not implausible to transfer the results obtained with the 1D modes to the more general 2D, 3D magnetized case, notably their intrinsically nonlinear nature, their stability and easy excitation potential, the latter not requiring anymore a linear instability mechanism. This admittedly somewhat lax handling is the more justified the stronger the magnetic field and the noisier the background plasma are, which supports nonlinear hole excitation. Another support stems from the fact that hole stability is increased whenever one of the three properties prevails, namely small amplitude, finite propagation velocity, and enhanced particle trapping, all within the limitations set by the equilibrium theory.

C. Evidence in other physical systems

We find it worth to at least mention the occurrence and identification of essentially 1D hole solutions in other physical systems, as well.

Firstly, CEHs have been identified in laser-plasma interaction experiments, predominantly periodic holes of essentially SEAW type ($k_0 \neq 0; B = 0$) in Ref. 66 and SEHs ($k_0 = 0; B > 0$) in Ref. 67.

Secondly, signatures of cnoidal hole solutions of the type ($k_0 \neq 0; B > 0$), especially of the upper branch, have

been seen in fullerene pair plasmas[68,69] and of the type $(k_0 \neq 0; B = 0)$ in dusty plasmas,[70] noting that in both cases a NDR similar to Eq. (7) for the corresponding mode[71] applies.

In the pair plasma experiment of Ref. 69, to be more explicit, longitudinal electrostatic modes along B-field lines in a dust-and electron-free fullerene (hydrogen) pair ion plasma could be generated by a cylindrical exciter with the following properties: the structure is periodic with a well-defined frequency-wave number relationship, i.e., it is characterized by a nonlinear dispersion relation $\omega = \omega(k; \psi)$, as shown in Fig. 2 of Ref. 69, three regions of ω can be distinguished, namely (i) a low frequency, acoustic region (mistakenly identified by the authors as linear ion acoustic waves, despite the general knowledge that the latter cannot exist in such a plasma), (ii) an intermediate frequency wave region (IFW) in which, after a turn over, the frequency increases with decreasing k, and (iii) a high frequency region, which extends to large ω exhibiting a wave number cutoff at small k.

A qualitative explanation of all of these properties can be given in a natural way-without entreating the presence of additional linear effects-by cnoidal nonlinear modes that exist due to ion trapping, as explored in Ref. 72 and shown in Fig. 3 of that paper.

And thirdly, holes (and humps) on coasting (and bunched) beams are a common, well-known structural phenomenon in circular accelerators and storage rings, where again trapped particle equilibria of the VP system can be made responsible for Refs. 73–76. The "mystery" of a hump seen recently on one of the beams in LHC in CERN (Ref. 77) may find an explanation within our theory, the reason for this self-generated structure being probably a misregulated power supply in the RF system (by courtesy of Frank Zimmermann).

These are in short some references dealing with CEHs in general.

In the last section we dwell on recently published works in some more detail, because of their benchmarking character, as we believe. We will provide in (A) a deeper view at a recent experiment, give in (B) evidence of SEH and CEHWL observations in space and simulations especially in the Earth's magnetic tail region, and comment in (C) on the concept of the group velocity in the presence of particle trapping. Finally, in (D) we will make an excursion into hydrodynamics and show that the trapping of fluid elements gives rise to a similar trapping scenario.

VI. A FIRST APPLICATION

This last section is hence devoted, as a first application, to an updated interpretation of some recent observations of phase space holes, based on the present achievements.

A. Holes in a laboratory experiment

In Ref. 18 the exitation of SEHs has been reported, initiated by a fast rising positive voltage pulse applied to a metallic disk electrode that has been immersed in a low pressure argon plasma. When the pulse width τ_p was below $3f_i^{-1}$, where f_i is the ion plasma frequency, a virtual source in front

of the electrode could be seen from which solitary potential humps started to propagate in two opposite directions: a left one, propagating towards the electrode with a speed of approximately $v_{0l} = 0.4$ and a right one propagating with approximately $v_{0r} = 1.3$, respectively. The virtual source region was initially identical with a localized ion rich region corresponding to a positive potential hump region.

The emanation of the two SEHs can then be qualitatively understood as follows.

Initially the spontaneously generated ion rich region is not adapted to a standing SEH and hence cannot persist in time. Instead, an evolution starts in which the potential hump is split into two counterpropagating SEHs subject to the laws of energy and momentum conservation for the global system. This splitting scenario is well known from evolution equations of Boussinesq-type (i.e., from equations extended to second order and formulated in lab frame) and applies to the one with a trapping nonlinearity as well,[45] the latter being characteristic for SEHs,[9] as mentioned. These self-consistently generated SEHs satisfy the NDR (7) with $k_0 = 0$ and are prescribed by $B_l = 0.83$ and $B_r = 0.01$, respectively. The corresponding trapping parameters, assuming $\psi = 1/100$, are found to be $\beta_l = -14$ and $\beta_r = -1$, respectively. We obviously have met here two typical members of SEHs. Note that an increase of the amplitude reduces the size of the trapping parameter but not its sign.[37]

We mention in parenthesis that in this experiment also a negative potential region (a well) was seen near the biased electrode. Assuming that all structures are mainly caused by the dynamics of electrons, one is tempted to associate this negative potential region with our solitary potential dip solution. This, however, requires that the life time of the structure is sufficiently long and that a beam-type trapped particle distribution prevails self-consistently, properties that have still to be confirmed experimentally.

When the pulse width was prolonged to $\tau_p > 3f_i^{-1}$ no virtual source was seen anymore. Instead the propagation of a single potential hump away from the exciter with approximately $v_0 = 1.3$ could be seen for a while. But suddenly a slowing down to $v_0 = 0.4$ occurred. This slower structure, however, must be of different origin, as no transition can be imagined that is valid within the collisionless theory. Instead, weak collisions supplemented by ion mobility and a pump DC electric field could possibly provide the conditions for a new dissipative structural EH state, as proposed in Refs. 11, 21, and 78. Our solutions, therefore, seem to provide the appropriate theoretical background for this experiment.

B. Holes at a reconnection side of Earth's magnetotail

It has become evident in the last decades that EHs belong to the most ubiquitous structures found in space, generated by large-scale parallel currents or electron beams in auroral plasmas,[79,80] in magnetopause[81] and in the Earth's magnetotail.[82,83] With respect to the latter, an important issue is as to whether EHs in the vicinity of a reconnection side in the Earth's magnetotail can shed more light on the small-scale electron dynamics and possibly on the dissipation mechanism accounting for collisionless magnetic

reconnection. Accumulated by several teams,[17,82–87] the observational facts can be summarized, as follows:

(i)   small amplitude bipolar, tripolar, and multipolar spikes in the parallel electric field $E_\parallel$ can be identified simultaneously, which appear

(i)   near the outer edge of the plasma sheet within a cavity and occur

(iii) during intervals of narrow electron beams or counter-streaming beams, for which at best and speculatively the Buneman instability has been made responsible. Often, if not mostly, however, they are found in the Buneman stable, i.e., linearly stable, regime.

Before becoming more concrete let us first report in more detail the experimental facts.

In Ref. 84 solitary holes and hole wavelets associated with magnetic reconnection have been observed by Cluster and Geotail in the diffusion region at dayside magnetopause and in the magnetotail region providing evidence that EHs may play a role in magnetic reconnection. In Fig. 1 of their paper, which refers to Geotail measurements at the dayside magnetopause, three types of electrostatic wave patterns in the Broadband Electrostatic Noise (BEN) have been seen, denoted by A, B, and C. Whereas A refers to a SEH (bipolar spike in $E_\parallel$), B and C have the characteristics of two and more humps in ϕ (tripolar and multipolar spikes), showing close similarity with structures found in particle simulations of the electron beam instability in which phase space vortices confirm their microscopic nature. A further structure D with even more equidistant humps in ϕ (AMEW standing for Amplitude Modulated Electrostatic Waves) could be seen as well, but a proof from simulations that these are due to particle trapping is missing.

From the theoretical point of view, as a first conclusion, it is not difficult to see that our CEHWLs, possibly extended by trapped ion effects, provide an explanation of the A, B, C, and AMEW structures, and that it is the task of numerical/experimental investigations to confirm (or disprove) the supremacy of the trapping nonlinearity for the latter.

Similarly, observations of Geotail (Fig. 3 of Ref. 84) in the magnetotail region revealed bi- and tripolar spikes in $E_\parallel$ both near the neutral sheet and around the plasma sheet boundary. The Cluster measurements during several plasma sheet encounters near the X-line region captured monopolar (DL), bipolar and tripolar pulses, seen when an intense narrow electron beam or narrow counter-streaming beams occurred. They are hence accompanied by enhanced fluxes of high-energy electrons flowing along the ambient magnetic field. Comparisons with PIC simulations[88] indicate that a Buneman instability ($v_0 \propto (m_e/m_i)^{1/3} v_{De}$) may be at work although the predicted hole speed of about 900–3500 km/s was typically higher than the observed one of 700–2500 km/s. In other words, the measured current density was roughly speaking a quarter lower than that required for a Buneman instability scenario. There were obviously events in which the Buneman instability was not active. In contrast to Ref. 88, who needed a guide magnetic field, EHs could be detected with and without a guide magnetic field.

A similar observation has been made earlier in Ref. 82 with Geotail, where the observed localized structures were found to be of three types: bipolar (type A), tripolar (type B) and multipolar (type C), as seen in their Fig. 4, the whole scenario including trapping being supported again by PIC simulations (Fig. 5).

Cluster observations of SEHs in association with magnetotail reconnection and comparison with simulations have been provided in Ref. 83, too. The hole structures were seen near the outer edge of the plasma sheet within a density cavity and occurred during intervals of narrow electron beams. In the numerical simulations these localized structures developed only relatively late after the beams driven by reconnection became strong. "They evolve from periodic fluctuations and grow to large amplitude with the largest-amplitude fluctuations coalescing into localized structures." The holes in this 2D simulation have a velocity of 0.2 of the electron streaming velocity, being roughly consistent with the Buneman instability. However, the predicted velocities were again higher than the observed ones, such that an under critical excitation cannot be excluded.

We also mention that bi- and tripolar pulses of $E_\parallel$ have been seen by the four Cluster space craft in the auroral zone[85] and by Polar and Cluster in and near the Earth's magnetosphere, magnetopause, and bowshock.[86] Another Polar observation of both structures in the high altitude polar magnetosphere has been reported in Ref. 87.

And last but not least the recent *in situ* observations of slow EHs at a reconnection site in the Earth's magnetotail in Ref. 17 substantiate this picture further. The four Cluster spacecraft measurements of EHs revealed the following properties:

(1) they are weak ($\psi \approx 0.02 - 0.07$) and their speed is close to the ion thermal or sound velocity, i.e., $v_0 \sim O(\sqrt{\delta})$,

(2) they typically appear simultaneously in the form of solitary holes and hole wavelets (Fig. 2 of Ref. 17), and

(3) the measured current density is an order of magnitude lower than the one required for a linear two-stream instability, such as for the ion acoustic or the Buneman instability (third page of Ref. 17).

With reference to Ref. 17, intense localized short-lived currents, which are produced by the coalescence of two magnetic islands and have the expected direction, namely opposite to the down-to-dusk propagation direction of EHs, are responsible for EH generation. Whether, however, the Buneman instability or some other current driven excitation mechanism can be made responsible, remained undecided. "This (the Buneman instability) cannot be reliably verified from the data." Also, the presence of such high drifts in the electron measurements could not been verified, because data were not available.

We conclude that the excitation mechanism is at least an open question and offer in the following an alternative, which requires lower drifts and has thus less problems with item 3). We will show now that the properties (1–3) can be straightforwardly understood by the present theory when it is extended by mobile ion effects and when use is made of the zero energy concept.

As mentioned at the end of Sec. II, the present theory in case of a current-carrying plasma can be extended to low hole velocities as low as ion acoustic velocities with $v_0 \sim O(\sqrt{\delta})$. One simply has to add in Eq. (7) an ionic shielding term and replace v_0 by $v_D - v_0$, while ion trapping effects will still be ignored. With this extension the energy difference Δw between the energy of the actual plasma, excited for simplicity by a single SEH only, and the energy of the unperturbed plasma was found in Ref. 32 to be

$$\Delta w = \frac{\psi}{2}\left[1 + \frac{1}{2}Z_r'\left(u_0/\sqrt{2}\right)\left(1 - u_0^2\right)\right], \qquad (23)$$

where $u_0 = v_0\sqrt{\theta/\delta}$. Expression (23) is zero at $u_0 = 2.12$ and is negative (positive) for $u_0 > 2.12$ ($u_0 < 2.12$). Applying energy conservation during the growth of the structure, $\Delta w = 0$, we then expect that zero-energy holes with $u_0 = 2.12$ are preferentially generated, in agreement with the observation. The reason is that in the absence of collisions the plasma resides in its total energy status transferring merely kinetic electron energy to potential energy during the establishment of the structure. The corresponding drift velocity v_D, noting that with u_0 also v_0 is already determined, is given by the NDR, which reads for SEHs and negligible ion trapping effects

$$-\frac{1}{2}Z_r'((v_D - v_0)/\sqrt{2}) - \frac{\theta}{2}Z_r'(u_0/\sqrt{2}) = \frac{16}{15}b(\beta, v_D - v_0)\sqrt{\psi}. \qquad (24)$$

In the Earth's magnetotail, where $\theta = 1$, and for zero-energy SEHs, where $u_0 = 2.12$ and hence $v_0 = u_0\sqrt{\delta} \approx 0$, it simplifies to

$$-\frac{1}{2}Z_r'(v_D/\sqrt{2}) - 0.285 = \frac{16}{15}b(\beta, v_D)\sqrt{\psi} =: \tilde{B}, \qquad (25)$$

which has to be satisfied by v_D for a given $\tilde{B}$. We add that the corresponding threshold for the Buneman instability would be $v_c \approx 1.3$.

For the remainder we show that there is no problem to find trapping conditions under which v_D can be appreciably below that value, e.g., one quarter below 1.3 or even a magnitude lower than 1.3.

This can be directly seen from Fig. 2 of Ref. 32, which displays Eq. (25) in the (θ, v_D) plane for several values of $\tilde{B}$ together with the critical drift velocity v_c for linear instability. We observe that $\tilde{B} = 0.01$ ($\tilde{B} = 0.7$, respectively) yields $v_D = 1$ ($v_D = 0.13$, respectively), which proves our claim.

For $\tilde{B} = 0.01$ it holds $(-\beta)\sqrt{\psi} = 0.0274$, which can be solved, e.g., by the pair ($\beta = -0.685, \psi = 1/625$), yielding a specific trapping condition for a given amplitude. For larger values of ψ the trapping parameter becomes less negative, e.g., ($\beta = -0.274, \psi = 0.01$) or ($\beta = -0.193, \psi = 0.02$). This implies, keeping the robust parameters $\tilde{B}$, v_D, or v_0 constant, that a gradual filling up of the trapped particle region, e.g., by processes such as scattering or phase space diffusion, is accompanied by an increase of the wave amplitude. This explains, besides the nonlinear trigger of a seed SEH, the underlying process of nonlinear growth, at least partially.

To complete the picture, we note that a $\tilde{B} = 0.7$ implies $(0.983 - \beta)\sqrt{\psi} = 1.173$, which yields $\beta = -28.3$ for $\psi = 1/625$. A deep excavation of the distribution in the resonant trapped particle region is thus in accord with drift velocities well below the critical linear drift velocity. (We point out in parentheses that long living excavations of this depth have been seen on debunched particle beams at Fermi lab in the mid 1990s[89] (see also Refs. 73 and 90 and references therein for more details).) Moreover, it is worth mentioning that the inclusion of ion trapping strengthens this picture, as it reduces the drift velocity further.[33]

The trapping nonlinearity, therefore, provides a natural explanation of the observed low speed structures in the Earth's magnetotail region with no need for invoking a linear instability for their excitation.

We may, therefore, state that these observations are consistent with the theoretically predicted nonlinear destabilization mechanism in linearly undercritical plasmas, as seen by the formation of CEHs.

We add that this, of course, does not answer the question whether SEHs do play an active role in the fast collisionless magnetic reconnection process. Indeed, recent 3D simulations[91] indicate that reconnection can take place without them, being mediated by anomalous viscosity, which is due to a filamentary instability, rather than by anomalous resistivity. To be excited, this instability requires a guide magnetic field of about 1/2 of the reconnecting magnetic field. However, as long as these simulations cannot reproduce the experimental facts, namely the SEH generation and especially their excitation also at low guide fields, it appears not implausible to assume that the last word about the role of electron holes in the magnetic reconnection process has not been spoken yet.

C. Group velocity of hole wavelets

In the preceding experiment, assuming that our interpretation applies, the velocity of the wavelet as a whole essentially coincides with the common phase velocity of each potential hump, such that no distinction can be made between the velocity of an individual hump and that of the group. This is in strong contrast with the usual group velocity concept, introduced by Rayleigh (see Refs. 92 and 93, and references therein), in which the group velocity, obtained by modulation of a train of a linear carrier wave, is given by $v_g = \partial\omega/\partial k$. The approach of calculating the wavelet velocity by $\partial\omega_0/\partial k_0$ and using the NDR $\omega_0(k_0; B)$, as depicted in Fig. 5, would, however, not make any sense and would fail.

The lack of an underlying linear carrier wave, therefore, questions all attempts to relate the wavelet velocity with Rayleigh's expression or more sophisticated elaborations, e.g., via variational approaches (see, e.g., Refs. 94–96) or via incorporation of nonlinear counterparts of linear elements, such as Landau damping, as has been proposed, e.g., by Refs. 19 and 20.

It is, therefore, by no means surprising when authors in Refs. 19 and 20 find strong deviations of the group velocity as a result of particle trapping and conclude "that the group

velocity of an essentially undamped wave, calculated by using the very definition of Rayleigh, is found to significantly differ from $\partial\omega/\partial k$ "or" that, surprisingly enough, the main nonlinear change in v_g occurs once the wave is effectively undamped!"

Phase space vortices or holes are intrinsically nonlinear and since no linear wave can be associated with the usual group velocity concept, the latter and elaborated extensions have no foundation anymore.

In fact, this failure of the group velocity concept could be added in Sec. III as a further "proof" of the intrinsic nonlinearity of cnoidal hole structures.

D. Vortex defect equilibria in hydrodynamic shear flows

As a last application we make use of the well-known mathematical analogy between VP plasmas and incompressible, 2D shear flows in hydrodynamics, Refs. 11, 97, 98, and references therein. In conformity with the electron hole solutions, formulated in phase space, we are going to construct explicitly self-consistent equilibria of vortex defects in ideal plane Couette, formulated in the 2D real space. The expectation is that, when subject to real physics, i.e., by inclusion of viscosity effects, etc., this will provide a better starting point for finding analytically critical Reynolds numbers Re_c for the onset of turbulence than the use of the linear shear flow profile. This is of interest since Re_c is infinity for the latter, such that the flow should be stable against infinitesimal perturbations for all Reynolds numbers in contradiction to the observational facts.[99]

To describe the ideal incompressible Couette flow in 2D we represent the flow velocity $\mathbf{u}(x, y, t)$ by the stream function $H(x, y, t)$ through

$$\mathbf{u} = \nabla H \times \hat{\mathbf{z}}, \tag{26}$$

which satisfies $\nabla \cdot \mathbf{u} = 0$. Taking the z-component of the curl of the Euler equation, $(\partial_t + \mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\varrho}\nabla \mathbf{p}$, we get for the vorticity field,

$$f(x,y,t) := (\nabla \times \mathbf{u}) \cdot \hat{\mathbf{z}} = -\nabla^2 H, \tag{27}$$

the evolution equation,

$$\partial_t f + [f, H] = 0, \tag{28}$$

where $[\cdot, \cdot]$ stands for the Poisson bracket in the two canonical variables (x, y). (Note that in case of a finite viscosity ν the term $\nu\nabla^2 f$ must be added on the rhs of (28).)

Equation (28), being also called convective cell equation in connection with Eq. (27), is a Liouville-type equation and expresses the Hamiltonian nature of the flow dynamics, as the Vlasov equation does for the plasma dynamics.

Steady state solutions, satisfying $[f, H] = 0$, are found by expressing f in terms of H, $f(H)$. The linear shear velocity profile is then represented by $f_0 = -1$ and $H_0 = y^2/2$, yielding $\mathbf{u_0} = y\hat{\mathbf{x}}$. To get a vortex defect equilibrium embedded in a constant shear flow profile we set $f = -1 + \delta f$ and $H = y^2/2 + \delta H =: \varepsilon$, in which case a solution is given by $\delta f = \delta f(\varepsilon)$. As in the Vlasov case ε can admit both signs

with $\varepsilon = 0$ representing the contour in the (x, y)-plane, which separates the laminar free flow from the one with self-trapped fluid elements, called the separatrix. In constructing δf a lot of freedom is available.

A simple case is given by assuming $\delta f = \theta(-\varepsilon)(-\hat{\beta}\varepsilon)$ with a constant $\hat{\beta}$ valid for $\varepsilon < 0$ and by setting $\delta f = 0$ for $\varepsilon > 0$, which means that the free flow remains unperturbed. We note that due to the perturbation there is now a region in which fluid elements will be trapped, the so-called trapped region $\varepsilon < 0$.

To satisfy self-consistency, $f = -\nabla^2 H$, in this region we set $\delta H = -(y^2/2 + \hat{\Phi})$ and get

$$\nabla^2\hat{\Phi} - \hat{\beta}\hat{\Phi} + 1 = 0, \tag{29}$$

valid for $\hat{\Phi} = -\varepsilon > 0$.

Equation (29) can be simplified by the Ansatz $\hat{\Phi} = 1/\hat{\beta} + \varphi(\zeta)$ with $\zeta := x^2 + y^2$, which yields

$$\zeta\varphi''(\zeta) + \varphi'(\zeta) - \frac{\hat{\beta}}{4}\varphi = 0. \tag{30}$$

The solution of Eq. (30) can be found by a series expansion and is given by

$$\varphi(\zeta) = a_0 \sum_{n=0}^{\infty} \frac{1}{(n!)^2}\left(\frac{\hat{\beta}\zeta}{4}\right)^n, \tag{31}$$

which is a quickly decaying series. An inspection shows that the streamlines $H = -\hat{\Phi} = -1/\hat{\beta} - \varphi(\zeta) = const$ are concentric circles around the origin, $\zeta = 0$, and that at the origin $\hat{\Phi}(0) = 1/\hat{\beta} + a_0$ holds. A positive $\hat{\Phi}$ thus requires $a_0 > -1/\hat{\beta}$ and the possibility of a zero of $\hat{\Phi}(\zeta)$, $\hat{\Phi}(\zeta_0) = 0$, asks for a negative $\hat{\beta}$, in which case Eq. (31) is an alternating, quickly converging series. The strength of the perturbation is controlled by the parameter $\hat{\Psi} := a_0 + 1/\hat{\beta} > 0$, which can be considered as the amplitude of the perturbation. In terms of $\hat{\Psi}$ the zero of $\hat{\Phi}(\zeta)$ is in first approximation obtained for $\zeta_{01} = \frac{-4\hat{\Psi}}{a_0\hat{\beta}} > 0$. The corresponding radius of the perturbed region $r_{01} = \sqrt{\zeta_{01}}$ thus shrinks with decreasing amplitude.

For the explicit case of $\hat{\beta} = -4$, and $a_0 = 5/4$ we obtain $\hat{\Phi}(\zeta) = 1 - \frac{5}{4}\zeta(1 - \zeta/4 + \zeta^2/36 - \zeta^3/576 + ...)$, which has a zero around $\zeta_0 \approx 1$ ($\zeta_{01} = 0.8$, $\zeta_{02} = 1.11, ...$). The admissible region, where the laminar flow is perturbed by a vortex defect, is hence given by $x^2 + y^2 \leq \zeta_0 \approx 1$. The perturbed vorticity in this region is given by $\delta f = -\hat{\beta}\varepsilon < 0$, i.e., the total vorticity becomes $f = -1 - \hat{\beta}\varepsilon$ and is enhanced magnitudewise. We mention that at the separatrix, $\varepsilon = 0$, the flow is discontinuous, which is in accordance with the zero viscosity assumption.

The local enhancement of the vorticity, therefore, provides a new, secondary equilibrium state, in which fluid elements are trapped, similar to the electron trapping scenario in plasma physics. This new equilibrium with an embedded vorticity perturbation is supposed to be unstable against perturbations (linearly and/or nonlinearly). For small amplitudes it can also serve as a tiny seed, which grows in

space-time and spreads over the whole flow. Forthcoming investigations are expected to shed more light on these dynamical processes that result in turbulence. This would explain why an inviscid plane Couette flow is turbulent rather than laminar.

At finite Reynolds numbers the use of the Navier-Stokes equation with a finite viscosity term for finding the corresponding secondary equilibrium with localized trapped fluid elements (provided that it exists) would be a straightforward, but mathematically more delicate challenge. For high Re this could perhaps be achieved by a boundary layer theory around the separatrix. Also, 3D secondary equilibria will probably be more successful to better approach the spatial-temporal "puff" like states, which are expected to be the origin of destabilization in analogy to the pipe flow problem (see next section). In any case, it could be an interesting task to approach the onset of turbulence analytically, i.e., to find the critical Reynolds number Re_c for more realistic situations.

For the cylindrical pipe flow problem, the dynamics and turbulence are of similar nature. Again, the pressure driven flow should be stable for all Reynolds numbers according to linear theory for the simple first order parabolic flow.[99] A secondary equilibrium with a localized vortex defect could again be an escape from this dilemma and provide an understanding of the onset of turbulence theoretically.

From an experimental point of view, the conditions for the onset of turbulence in a long pipe flow have recently been clarified by series of experiments, performed by Hof and collaborators[100–104] (see also Eckhardt[105]). These authors observed that for Reynolds numbers below about 2300 the turbulence remains localized in short "puffs" that move downstream without any change in form, having a finite lifetime that increases with Re. On the other hand, Nishi et al.[106] could show by numerical simulations that puffs can split, the splitting time being decreasing with Re, as confirmed also by the experiments. "In particular, if a puff manages to split before it decays, the sibling may carry on the turbulence, spatial and temporal couplings become important and there may always be some turbulence somewhere along the pipe."[105] Equating the lifetime of a single puff, created by injection of a localized water jet into the flow, and the time of splitting, Hof and collaborators determined $Re_c = 2040$ as the critical Reynolds number for the pipe flow problem. The coherent structures thus provide a scaffold that supports turbulent dynamics by creating a multitude of connections between these states.[105] Turbulence is thus found to be triggered by localized patches, which are transient and spatially coupled, placing the pipe flow dynamics in the larger theoretical framework of spatial-temporal intermittency.[107]

VII. CONCLUSIONS AND OUTLOOK

The subject of the present review has been the trapping nonlinearity as evident in persistent, 1D, weak, electrostatic structures of cnoidal electron hole type. The latter propagate at bulk velocity in a thermal, collisionless plasma, i.e., in a region where standard linear wave theory predicts nonexistence due to strong Landau damping. The ubiquitous

evidence of these structures in laboratory, space and numerical experiments, however, provides a different picture and indicates that something must be wrong with wave theories that rely on a linearization of the governing equations in the small amplitude limit.

In the present review, we have collected arguments which prove that such standard wave theories cannot claim general validity and must be given up in favor of an updated wave theory in which nonlinearity prevails from the outset, no matter how small the amplitude is. It is the velocity region at phase velocity—the resonant or trapped particle region—for which special care must be taken not to miss the present electron hole structures. The latter automatically come to light if self-consistent Vlasov distribution functions, rather than linearized Vlasov distribution functions, are involved that are sufficiently smooth near phase velocity as, for instance, coarse-grained distributions.

We argued that the function space of the VP solutions and that of the lVP solutions are two different worlds that cannot be bridged, a fact being independent of the wave amplitude.

Here, we concentrated on the simplest possible description of trapped particle modes, the cnoidal electron holes propagating in a neutralizing immobile ion background, and analyzed their properties in some detail. As mentioned, the neglect of ion dynamics and especially of ion trapping effects is admissible only for not too slow phase velocities, i.e., the EHs have to exceed the ion sound speed $v_0 > \sqrt{m_e/m_i}$. Otherwise ionic effects come into play, which increase the variety of trapped particle modes, such as treated in Refs. 9–12, 37, 108, etc. A direct application of this generalization has been made for the magnetic reconnection process in the Earth's magnetotail.

The birth of these structures demands a driving mechanism, which can be a current, as mentioned, but also wave launching, beam injection, or inhomogeneities can be the cause. The easiest and often met wave excitation mechanism is a linear two-stream instability in which holes are generated out of thermal noise to become nonlinear states, saturated by particle trapping. This process, however, is not restricted to an ideal collisionless plasma but can take place in a weakly collisional plasma, as well. As seen numerically in Refs. 78 and 11 for a linearly two-stream unstable plasma, electron hole structures (generally of finite amplitude and with a β nearly zero) can emerge out of seed fluctuations in a noisy plasma and resist collisions, provided that ions are treated mobile and a DC electric field is present. The result is a driven, structurally excited, dissipative equilibrium state far away from the thermodynamic one. In such an excited, structurally embossed plasma, the transport properties are no longer determined by classical transport coefficients as have been derived by Ref. 109 for a Maxwellian plasma, but have now to be rederived from distributions of the kind (2).

Moreover, for a collisionless, current-carrying plasma, numerical simulations revealed that holes can be excited nonlinearly even in situations of a linear two-stream stability,[11,30–35] as mentioned and as applied for the magnetic reconnection process in the present paper. There is, hence, no chance to understand the numerical findings on the basis of linearized wave theories.

In other words, the presence of electron holes and their interpretation open a door into a new world of anomalous plasma processes.

As discussed also, this knowledge about trapping can directly be transferred to the solenoidal fluid dynamics of a 2D, incompressible shear flow, as met in the plane Couette flow, in the cylindrical pipe flow, in geostrophic fluids,[110,111] etc., where the convective cell equation as the underlying evolution equation for the vorticity bears the same Hamiltonian structure. A secondary flow equilibrium with a localized region of trapped fluid elements offers a new way for a theoretical understanding of the onset of turbulence.

Furthermore, strongly magnetized plasmas, being ruled by evolution equations of similar type, such as the Hasegawa-Mima equation[112] or the Hasegawa-Wakatani equation,[113] expose the same type of nonlinearity, the latter resulting from the perpendicular, ExB driven ion (plasma) dynamics coupled with the parallel electron dynamics. In this case the trapping effect can enter even twice, hence making any linear wave approach as a meaningful description obsolete (see also Refs. 11, 12, 14, 114 for more details).

We finish by mentioning that this generalized view of the intrinsic nonlinearity in plasma and fluid dynamics, the moment trapping is involved, receives further support by a numerical simulation of Drake *et al.*[115] Solving the 3D collisional drift-wave turbulence in a sheared magnetic field these authors found that the dynamics is self-sustained by the excitation of nonlinear structures even in situations where all linear modes are damped. The trapping scenario, explored in some detail for electrostatic waves in the present paper, therefore, gives a hint for a deeper understanding of the complex plasma and fluid dynamics and may thus contribute to the resolution of a longstanding mystery in the dynamical evolution processes.

And finally, the overwhelming body of evidence of hole observations, especially in linearly stable regions, suggests that the theoretical treatment of plasma (and fluid) dynamics should be released from the "Prokrustes bed" of linear wave theory. Except for the initial state of a linear instability and the relatively rare events of quiet plasmas, subject initially to sufficiently small, analytic perturbations, for which linear Landau theory and a linear dispersion relation do apply, an electrostatically driven plasma evolves as a rule nonlinearly and is governed even in the small wave amplitude limit by the trapping nonlinearity, the propagation of cnoidal electron holes being merely an example.

ACKNOWLEDGMENTS

[1]K. Saeki, P. Michelsen, H. L. Pécseli, and J. J. Rasmussen, Phys. Rev. Lett. **42**, 501 (1979).
[2]J. P. Lynov, P. Michelsen, H. L. Pécseli, J. J. Rasmussen, K. Saeki, and V. A. Turikov, Phys. Scr. **20**, 328 (1979).
[3]H. Schamel, Phys. Scr. **20**, 336 (1979).
[4]H. Schamel, Plasma Phys. **14**, 905 (1972).
[5]I. B. Bernstein, J. M. Greene, and M. D. Kruskal, Phys. Rev. **108**, 546 (1957).
[6]V. A. Turikov, Phys. Scr. **30**, 73 (1984).
[7]L. Muschietti, I. Roth, R. E. Ergun, and C. W. Carlson, Nonlinear Processes Geophys. **6**, 211 (1999).
[8]L.-J. Chen, J. Pickett, P. Kintner, J. Franz, and D. Gurnett, J. Geophys. Res. **110**, A09211, doi:10.1029/2005JA011087 (2005).
[9]H. Schamel, Phys. Rep. **40**, 161 (1986).
[10]H. Schamel, Phys. Plasmas **7**, 4831 (2000).
[11]A. Luque and H. Schamel, Phys. Rep. **415**, 261 (2005).
[12]B. Eliasson and P. K. Shukla, Phys. Rep. **422**, 225 (2006).
[13]D. Jovanović and F. Pegoraro, Phys. Rev. Lett. **84**, 95 (2000).
[14]D. Jovanović, P. K. Shukla, L. Stenflo, and F. Pegoraro, J. Geophys. Res. [Space Phys.] **107**, 1110, doi:10.1029/2001JA900180 (2002).
[15]B. Eliasson and P. K. Shukla, Phys. Lett. A **338**, 237 (2005).
[16]M. V. Goldman, D. L. Newman, and A. Mangeney, Phys. Rev. Lett. **99**, 145002 (2007).
[17]Yu. V. Khotyaintsev, A. Vaivads, M. Andre, M. Fujimoto, A. Retino, and C. J. Owen, Phys. Rev. Lett. **105**, 165002 (2010).
[18]S. Kar, S. Mukherjee, G. Ravi, and Y. C. Saxena, Phys. Plasmas **17**, 102113 (2010).
[19]D. Bénisti, D. J. Strozzi, L. Gremillet, and O. Morice, Phys. Rev. Lett. **103**, 155002 (2009).
[20]D. Bénisti, O. Morice, L. Gremillet, E. Siminos, and D. J. Strozzi, Phys. Plasmas **17**, 082301 (2010).
[21]J. Korn and H. Schamel, J. Plasma Phys. **56**, 307 (1996).
[22]D. Montgomery and G. Joyce, J. Plasma Phys. **3**, 1 (1969).
[23]H. Schamel, J. Plasma Phys. **13**, 139 (1975).
[24]N. G. van Kampen and B. U. Felderhof, *Theoretical Methods in Plasma Physics* (North-Holland, Amsterdam, 1967).
[25]G. Belmont, F. Mottez, T. Chust, and S. Hess, Phys. Plasmas **15**, 052310 (2008).
[26]K. M. Case, Ann. Phys. **7**, 349 (1959).
[27]T. O'Neil, Phys. Fluids **8**, 2255 (1965).
[28]G. Manfredi, Phys. Rev. Lett. **79**, 2815 (1997).
[29]J. P. Holloway and J. J. Dorning, Phys. Rev. A **44**, 3856 (1991).
[30]R. H. Berman, D. J. Tetreault, T. H. Dupree, and T. B. Ghali, Phys. Rev. Lett. **48**, 1249 (1982).
[31]R. H. Berman, D. J. Tetreault, and T. H. Dupree, Phys. Fluids **28**, 155 (1985).
[32]J.-M. Grießmeier and H. Schamel, Phys. Plasmas **9**, 2462 (2002).
[33]J.-M. Grießmeier, A. Luque, and H. Schamel, Phys. Plasmas **9**, 3816 (2002).
[34]A. Luque, H. Schamel, B. Eliasson, and P. K. Shukla, Phys. Plasmas **12**, 122307 (2005).
[35]H. Schamel and A. Luque, Space Sci. Rev. **121**, 313 (2005).
[36]A. V. Gurevich, Sov. Phys. JETP **26**, 575 (1968).
[37]S. Bujarbarua and H. Schamel, J. Plasma Phys. **25**, 515 (1981).
[38]H. Schamel, Plasma Phys. **13**, 491 (1971).
[39]H. Schamel, A. Luque, I. Kourakis, and N. S. Saini, "On the non-existence of travelling, weak solitary electron holes with shapes different from sech^4-profile," Phys. Plasmas (to be published).
[40]A. A. Vlasov, J. Phys. (USSR) **9**, 25 (1945).
[41]T. H. Stix, *The Theory of Plasma Waves* (McGraw-Hill, New York, 1962).
[42]H. Schamel, Z. Naturforsch. **38a**, 1170 (1983).
[43]H. Schamel, Phys. Scr. **T2/1**, 228 (1982).
[44]H. Schamel and S. Bujarbarua, Phys. Fluids **26**, 190 (1983).
[45]H. Schamel, J. Plasma Phys. **9**, 377 (1973).
[46]M. W. Coffey, J. Phys. A **24**, L1345 (1991).
[47]A. Ramani and B. Grammaticos, J. Phys. A **24**, 1969 (1991).
[48]F. Verheest and W. Hereman, Phys. Scr. **50**, 611 (1994).
[49]C. Lancelotti and J. J. Dorning, J. Math. Phys. **40**, 3895 (1999).
[50]Z. Lin and C. Zeng, Commun. Math. Phys. **306**, 291 (2011).
[51]M. Buchanan and J. J. Dorning, Phys. Rev. Lett. **70**, 3732 (1993).
[52]M. Buchanan and J. J. Dorning, Phys. Rev. E **50**, 1465 (1994).
[53]E. Fijalkow and L. Nocera, J. Plasma Phys. **71**, 401 (2005).
[54]H. Schamel, Phys. Rev. Lett. **48**, 481 (1982).
[55]H. R. Lewis and K. R. Symon, J. Math. Phys. **20**, 413 (1979).
[56]J. Schwarzmeier, H. R. Lewis, B. Abraham-Shrauner, and K. R. Symon, Phys. Fluids **22**, 1747 (1979).
[57]K. V. Roberts and H. L. Berk, Phys. Rev. Lett. **19**, 297 (1967).
[58]R. L. Morse and C. W. Nielson, Phys. Rev. Lett. **23**, 1087 (1969).

[59]L. Muschietti, I. Roth, C. W. Carlson, and R. E. Ergun, Phys. Rev. Lett. **85**, 94 (2000).

[60]D. Jovanović and H. Schamel, Phys. Plasmas **9**, 5079 (2002).

[61]F. Mottez, S. Perraut, and A. Roux, J. Geophys. Res. **102** 11399, doi:10.1029/97JA00385 (1997).

[62]M. Wu, Q. Lu, C. Huang, and S. Wang, J. Geophys. Res. **115**, A10245, doi:10.1029/2009JA015235 (2010).

[63]D. Jovanović and P. K. Shukla, Phys. Plasmas **11**, 4946 (2004).

[64]D. Jovanović and P. K. Shukla, Phys. Rev. Lett. **93**, 015002 (2004).

[65]D. Jovanović and P. K. Shukla, Phys. Plasmas **12**, 052114 (2005).

[66]D. S. Montgomery, R. J. Focia, H. A. Rose, D. A. Russell, J. A. Cobble, J. C. Fernández, and R. P. Johnson, Phys. Rev. Lett. **87**, 155001 (2001).

[67]G. Sarri, M. E. Dieckmann, C. R. D. Brown, C. A. Cecchetti, D. J. Hoarty, S. F. James, R. Jung, I. Kourakis, H. Schamel, O. Willi, and M. Borghesi, Phys. Plasmas **17**, 010701 (2010).

[68]W. Oohara, Y. Kuwabara, and R. Hatakeyama, Phys. Rev. E **75**, 056403 (2007).

[69]W. Oohara and R. Hatakeyama, Phys. Plasmas **14**, 055704 (2007).

[70]P. Bandyopadhyay, G. Prasad, A. Sen, and P. K. Kaw, Phys. Lett. A **368**, 491 (2007).

[71]H. Schamel, Phys. Plasmas **16**, 113709 (2009).

[72]H. Schamel and A. Luque, New J. Phys. **7**, 69 (2005).

[73]H. Schamel, Phys. Rev. Lett. **79**, 2811 (1997).

[74]J.-M. Grießmeier, H. Schamel, and R. Fedele, Phys. Rev. ST Accel. Beams **5**, 024201 (2002).

[75]M. Blaskiewicz, J. Wei, A. Luque, and H. Schamel, Phys. Rev. ST Accel. Beams **7**, 044402 (2004).

[76]H. Schamel and A. Luque, New J. Phys. **6**, 113 (2004).

[77]www.cerncourier.com/cws/article/cern/42331, The LHC's new frontier, CERN Courier May 5, 2010; also CERN Courier, **50**(4), 27 (2010).

[78]J. Korn and H. Schamel, J. Plasma Phys. **56**, 339 (1996).

[79]R. Boström, G. Gustafsson, B. Holback, G. Holmgren, H. Koskinen, and P. Kintner, Phys. Rev. Lett. **61**, 82 (1988).

[80]R. E. Ergun, C. W. Carlson, J. P. McFadden, F. S. Mozer, G. T. Delory, W. Peria, C. C. Chaston, M. Temerin, I. Roth, L. Muschietti, R. Elphic, R. Strangeway, R. Pfaff, C. A. Cattell, D. Klumpar, E. Shelley, W. Peterson, E. Moebius, and L. Kistler, Geophys. Res. Lett. **25**, 2041, doi:10.1029/98GL00636 (1998).

[81]H. Matsumoto, X. H. Deng, H. Kojima, and R. R. Anderson, Geophys. Res. Lett. **30**, 1326, doi:10.1029/2002GL016319 (2003).

[82]H. Matsumoto, H. Kojima, T. Miyatake, Y. Omura, M. Okada, I. Nagano, and M. Tsutsui, Geophys. Res. Lett. **21**, 2915, doi:10.1029/94GL01284 (1994).

[83]C. Cattell, J. Domebeck, J. Wygant, J. F. Drake, M. Swisdak, M. L. Goldstein, W. Keith, A. Fazakerley, M. André, E. Lucek, and A. Balogh, J. Geophys. Res. **110**, A01211, doi:10.1029/2004JA010519 (2005).

[84]X. H. Deng, R. X. Tang, H. Matsumoto, H. Kojima, W. Baumjohann, A. Coates, R. Nakamura, D. A. Gurnett, and Z. X. Liu, Ad\roundv. Space Res. **37**, 1373 (2006).

[85]J. S. Pickett, S. W. Kahler, L.-J. Chen, R. L. Huff, O. Santolik, Y. Khotyaintsev, P. M. E. Décréau, D. Winningham, R. Frahm, M. L. Goldstein, M. L. Lakhina, B. T. Tsurutani, B. Lavraud, D. A. Gurnett, M. André, A. Fazakerley, A. Balogh, and H. Réme, Nonlinear Processes Geophys. **11**, 183 (2004).

[86]C. Cattell, C. Neiman, J. Dombeck, J. Crumley, J. Wygant, C. A. Kletzing, W. K. Peterson, F. S. Mozer, and M. André, Nonlinear Processes Geophys. **10**, 13 (2003).

[87]J. R. Franz, P. M. Kintner, and J. S. Pickett, Geophys. Res. Lett. **25**, 1277, doi:10.1029/98GL50870 (1998).

[88]J. Drake, M. Swisdak, C. Cattell, M. A. Shay, B. N. Rogers, and A. Zeiler, Science **299**, 873 (2003).

[89]P. L. Colestock and L. K. Spentzouris, in *The Tamura Symposium Proceedings, Austin, Texas, 1994 AIP Conference Proceedings* (AIP, Woodbury, New York, 1996), Vol. 356.

[90]H. Schamel, Phys. Scr. **T75**, 23 (1998).

[91]H. Che, J. F. Drake, and M. Swisdak, Nature (London) **474**, 184 (2011).

[92]W. R. Hamilton, Proc. R. Ir. Acad. **1**, 341 (1839).

[93]L. Brillouin, *Wave Propagation and Group Velocity* (Academic, New York, 1960).

[94]M. J. Lighthill, J. Inst. Math. Appl. **1**, 1 (1965).

[95]G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974).

[96]C. Decker and W. M. Mori, Phys. Rev. E **51**, 1364 (1995).

[97]H. L. Berk and K. V. Roberts, Phys. Fluids **10**, 1595 (1967).

[98]N. J. Balmforth, D. Del-Castillo-Negrete, and W. R. Young, J. Fluid Mech. **333**, 197 (1997).

[99]P. Drazin and W. Reid, *Hydrodynamic Stability* (Cambridge University Press, Cambridge, 1981).

[100]B. Hof, C. van Doorne, J. Westerweel. F. T. M. Nieuwstadt, H. Faisst, B. Eckhardt, H. Wedin, R. R. Kerswell, and F. Waleffe, Science **305**, 1594 (2004).

[101]B. Hof, J. Westerweel, T. M. Schneider, and B. Eckhardt, Nature (London) **443**, 59 (2006).

[102]B. Hof, A. de Lozar, D. J. Kuik, and J. Westerweel, Phys. Rev. Lett. **101**, 214501 (2008).

[103]A. de Lozar and B. Hof, Philos. Trans. R. Soc. London, Ser. A **367**, 589 (2009).

[104]K. Avila, D. Moxey, A. de Lozar, M. Avila, D. Barkley, and B. Hof, Science **333**, 192 (2011).

[105]B. Eckhardt, Science **333**, 165 (2011).

[106]M. Nishi, B. Ünsal, F. Durst, and G. Biswas, J. Fluid Mech. **614**, 425 (2008).

[107]P. Manneville, Phys. Rev. E **79**, 025301 (2009).

[108]H. Schamel and S. Bujarbarua, Phys. Fluids **23**, 2498 (1980).

[109]S. I. Braginskii, *Transport Processes in a Plasma, Reviews of Plasma Physics* (Consultants Bureau, New York, 1965), Vol. 1.

[110]J. G. Charney, J. Atmos. Sci. **28**, 1087 (1971).

[111]J. Pedlovsky, *Geophysical Fluid Dynamics* (Springer, New York, 1987).

[112]A. Hasegawa and K. Mima, Phys. Fluids **21**, 87 (1978).

[113]A. Hasegawa and M. Wakatani, Phys. Rev. Lett. **50**, 682 (1983).

[114]D. Jovanović and P. K. Shukla, Phys. Rev. Lett. **84**, 4373 (2000).

[115]J. F. Drake, A. Zeiler, and D. Biskamp, Phys. Rev. Lett. **75**, 4222 (1995).