TUTORIAL | MARCH 01 2017

Electron holes in phase space: What they are and why they matter

Special Collection: Reviews and Tutorials in Basic Plasma Phenomena, Waves, and Instabilities

I. H. Hutchinson

Phys. Plasmas 24, 055601 (2017)
https://doi.org/10.1063/1.4976854

Electron holes in phase space: What they are and why they matter

I. H. Hutchinson[a),b)]
Plasma Science and Fusion Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(Received 4 November 2016; accepted 5 December 2016; published online 1 March 2017)

This is a tutorial and selective review explaining the fundamental concepts and some currently open questions concerning the plasma phenomenon of the electron hole. The widespread occurrence of electron holes in numerical simulations, space-craft observations, and laboratory experiments is illustrated. The elementary underlying theory is developed of a one-dimensional electron hole as a localized potential maximum, self-consistently sustained by a deficit of trapped electron phase-space density. The spatial extent of a hole is typically a few Debye lengths; what determines the minimum and maximum possible lengths is explained, addressing the key aspects of the as yet unsettled dispute between the integral and differential approaches to hole structure. In multiple dimensions, holes tend to form less readily; they generally require a magnetic field and distribution-function anisotropy. The mechanisms by which they break up are explained, noting that this transverse instability is not fully understood. Examples are given of plasma circumstances where holes play an important role, and of recent progress on understanding their holistic kinematics and self-acceleration. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4976854]

I. INTRODUCTION: HOLE CHARACTER AND OCCURRENCE

The term electron hole refers to a localized plasma region where the electron density is lower than the surrounding plasma, because of reduced phase-space density on trapped electron orbits. The decreased electron density causes a local maximum in charge density and hence in electric potential, which is self-consistently responsible for electron trapping. It is a type of Bernstein, Greene, and Kruskal (BGK) mode,[1] and has also been referred to by a variety of related terms: "phase-space(-density) holes (or structures),"[2–8] "electrostatic solitary (or coherent) waves (or structures)"[9–18] (especially ESW in space-plasma literature), "phase-space vortices,"[19] "Debye-scale structures,"[20,21] "coherent electric field structures,"[22,23] and "time domain structures."[24] Many of these more diverse names appropriately represent ambiguity in identifying the precise nature of observed phenomena, but for simplicity, I will refer to all by the name electron hole. The defining characteristics of the BGK subset implied by this name are: first, that the electron hole is a *localized positive*-potential structure so that electrons are the trapped species (in contrast with ion holes which are negative potential), and second that the potential is predominantly *symmetric* about the potential peak, thus ruling out from our present topic "double layers" and strongly asymmetric holes that can be considered a combination of a hole and a double-layer. See Fig. 1.

Electron holes were observed to form in the earliest one-dimensional kinetic computational simulations of the two-stream instability,[25,26] and in countless simulations since. To illustrate, Fig. 2 shows frames from a simulation using the simple one-dimensional particle in cell (PIC) code XES1, distributed with the textbook of Birdsall and Langdon.[27] Two separated electron streams of opposite velocity and equal density give rise to instability. The wave grows until a non-linear regime is entered and reflection and trapping of some of the electrons occur, by the strong potential perturbations. Four electron holes are formed in this case, but they quite quickly merge in pairs (by encircling one another) until only one big hole is left. The particle density (density of points) is low in the center of the phase-space vortex, and enhanced at its rim, giving rise to a positive charge peak with adjacent slightly negative wings: just what is needed for a positive potential peak. An animation of this progress is available at the URL in Fig. 2. This illustrates the fact that essentially any collisionless electron two-stream instability in one-dimension gives rise to electron holes in the long-time limit.

In Fig. 3 are shown the orbits in phase-space of electrons and ions in a schematic positive-potential hole. The orbits are contours of constant distribution function for collisionless particles governed by the Vlasov equation, and are contours of constant total energy (sum of kinetic and potential energy), when the potential is time-invariant, as it is in the rest-frame of the hole. The electron orbits are of two types, passing, open orbits, which connect to the distant plasma,

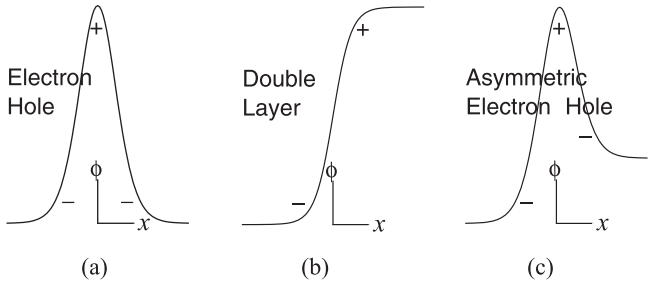

FIG. 1. Schematic illustration of the potential shape of simple generic types of BGK-modes.

Note: Paper CT2 1, Bull. Am. Phys. Soc. **61**, 62 (2016).
[a)]Invited speaker.
[b)]Electronic mail: ihutch@mit.edu

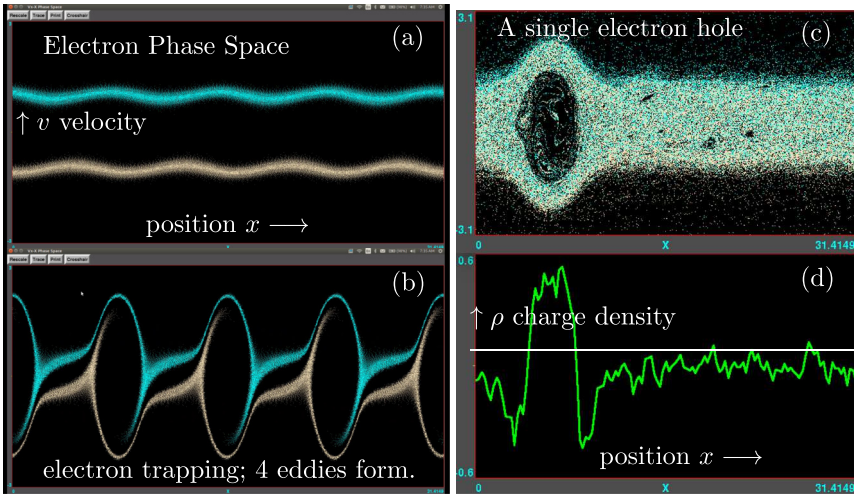

FIG. 2. Two-stream instability PIC simulations develop into electron holes. Simulation particles are plotted as points in phase space (v versus x). Initial small amplitude instability (a), grows with time until particles are trapped and phase-space eddies form (b). The resulting (four) holes eventually merge into one (c), whose charge density is shown in panel (d). (Multimedia view) [URL: http://dx.doi.org/10.1063/1.4976854.1]

and trapped, closed orbits, localized to the hole region and to a small speed. The separatrix is the boundary between these two phase-space regions. Ions, having the opposite charge, are repelled by the positive potential. All their orbits are then open. No ions are locally trapped, but some may be reflected; others are passing. The much heavier ions generally have a velocity spread much less than the electrons. However, holes usually move substantially faster than the ion thermal speed; so in the hole rest-frame, the ions can be considered to be a stream with a relatively narrow spread of velocities with values near to minus the hole speed (in the ion frame). In that case, negligibly few ions exist on reflected orbits.

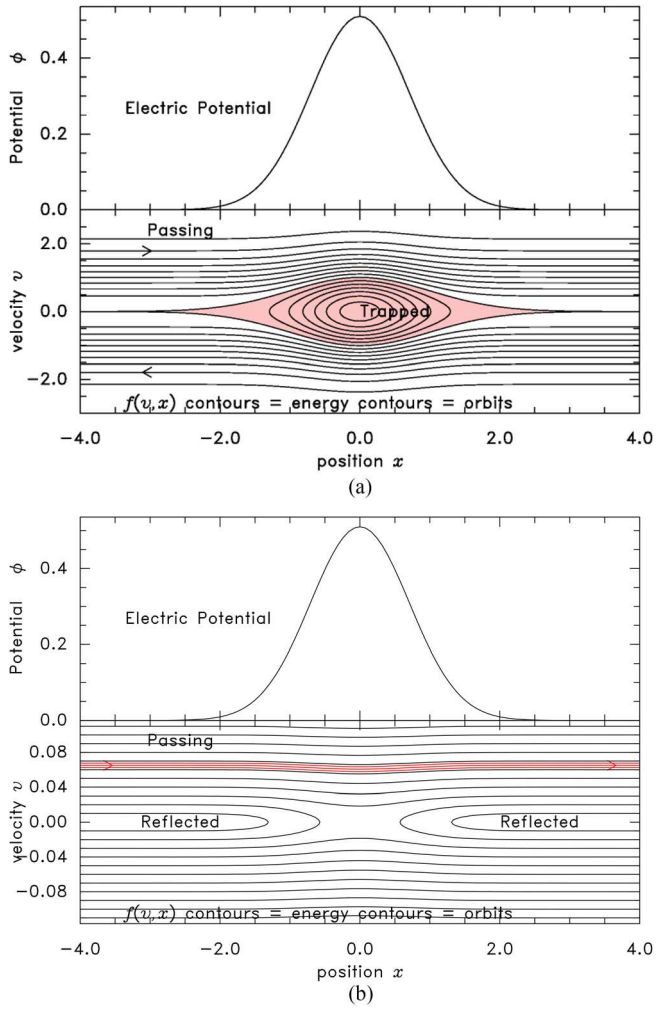

FIG. 3. In a steady potential (in the rest-frame of the hole), the orbits of electrons (a) and ions (b) are contours of constant total energy. The collisionless distribution-functions f_e and f_i are constant along orbits, the functions of energy. Positive potential traps electrons, but not ions.

Electron holes have been observed both in laboratory experiments and in space measurements. A much cited Q-machine experiment[28] succeeded in 1979 in generating electron holes that were observed with an array of electric probes. It was only in the 1990s that electron holes were unambiguously identified in spacecraft measurements. Earlier spectral analysis had identified "broadband electrostatic noise" extending up to the plasma frequency, for example, in the geomagnetic tail.[29] But only when sufficiently rapid sampling was able to resolve the time behavior and identify individual bipolar pulsed parallel electric fields,[9] was it established that electron holes were a major constituent, see Fig. 4(a). Satellite measurements are generally performed at a single location, and determine the electric field by the potential difference between two Langmuir probes, or between the spacecraft and a probe. A single time trace measurement cannot then determine the spatial scale, since the hole speed is unknown. But some satellites obtain two synchronized signals from potential difference between the craft and two probes, sufficiently separated to observe a time delay. From it, the hole speed and spatial extent can be deduced, see Fig. 4(b). Fast "cadence" (i.e., fast sampling, optimally every ∼0.1 ms or faster) electric field measurements from various satellites continue to show electron holes in many different space environments: the Earth's auroral region (FAST),[20] bow shock (Wind)[30] magnetopause (GEOTAIL),[31] magnetosheath (Cluster)[16,32] (MMS),[33] plasma sheet (THEMIS)[34] (Cluster),[35] outer radiation belt (Van Allen Probes)[36,37] and also in the free solar wind, at interplanetary shocks, and current sheets (Wind).[18,38–40]

Electron holes have been deliberately generated in the laboratory in a pure-electron plasma by a chirped autoresonant field oscillation that drags a bucket of low-phase-space-density into the electron distribution.[41,42] More recent laboratory experiments have also observed electron holes[8,17] in quasineutral plasmas. Fig. 5 shows an example obtained with extremely small (10 μm) Langmuir probes on the linear

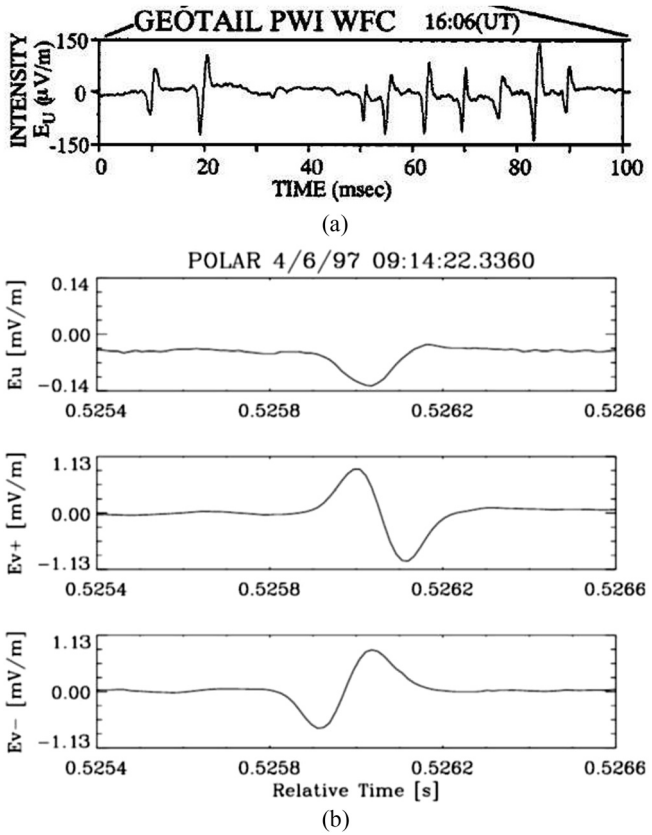

FIG. 4. Examples of the electric field pulses of electron holes in space. (a) GEOTAIL data (adapted from Matsumoto *et al.*, Geophys. Res. Lett. **21**, 2915–2918 (1994), Copyright 1994 John Wiley and Sons[9]) showing the first satellite-measured resolved electron holes. (b) POLAR data (adapted from J. R. Franz, P. M. Kintner, and J. S. Pickett, Geophys. Res. Lett. **25**, 1277–1280 (1998). Copyright John Wiley and Sons[22]) showing (inversion and) a time delay between two probes extended in opposite parallel directions (bottom two panels), from which the hole speed can be deduced. The perpendicular electric field is in the top panel of (b).

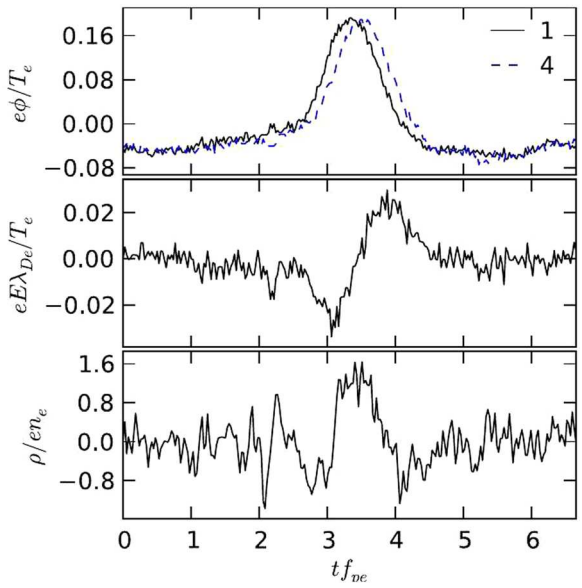

FIG. 5. Electron hole measurements with small Langmuir probes on the LAPD facility (adapted from B. Lefebvre *et al.*, Phys. Rev. Lett. **105**, 115001 (2010). Copyright 2010 American Physical Society[17]) show a localized potential peak and the accompanying electric field and charge density. There is a small phase shift between the measurements at adjacent probes (1 and 4), indicating the hole speed.

LAPD basic plasma facility.[17] Two probes separated by 1.7 times the Debye length ($\lambda_{De} \approx 60\mu$m) show a phase shift that indicates hole speed. The electrons are strongly magnetized, the ratio of electron-cyclotron frequency to plasma frequency is $\Omega/\omega_p = 4$. Phase shifts between adjacent probes have similarly been observed in experiments on the toroidal experiment VTF,[8,43] in which the holes are generated through the process of magnetic reconnection. Reconnection is implicated in many space-plasma situations as the context of electron hole occurrence;[31,44] and holes occur in some reconnection simulations.[45,46]

II. HOLE STRUCTURE: 1-D EQUATIONS AND SOLUTIONS

The equations governing the one-dimensional electron hole are the Vlasov equation and Poisson's equation. A hole is a collisionless phenomenon but absolutely requires a kinetic theory description, hence Vlasov's equation. In some cases, the ion response must also be considered, but to begin with we will assume the ion density to be simply a uniform neutralizing background. Electron holes are predominantly electrostatic, described by an electrostatic potential ϕ governed by Poisson's equation. In some (e.g., relativistic) situations, full Maxwell equations ought to be used, but in this tutorial, we will consider only the electrostatic case.

It simplifies the algebra and expresses the physics more directly to use scaled dimensionless parameters as follows. Length x is measured in units of the Debye length $\lambda_D = \sqrt{\epsilon_0 T_e/n_\infty e^2}$, and time t in units of the inverse plasma frequency $\omega_p^{-1} = \sqrt{\epsilon_0 m_e/n_\infty e^2}$, (consequently velocity v is units $\sqrt{T_e/m_e}$). Potential ϕ is measured in units of T_e/e, and density n in units of n_∞. Here the electron temperature T_e and density n_∞ are some fixed values representative of the background plasma. In terms of these dimensionless parameters, the Vlasov equation governing the distribution function $f(v, x)$ is

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} - \frac{\partial \phi}{\partial x}\frac{\partial f}{\partial v} = 0. \tag{1}$$

It expresses the fact that f is constant along the particle orbits (which are the characteristics of this hyperbolic equation). We generally use that fact rather than explicitly solving the differential equation.

Poisson's equation relates the potential to the charge density ρ

$$\frac{\partial^2 \phi}{\partial x^2} = -\rho = -n_i + n_e, \tag{2}$$

where for simplicity we assume the ions are singly charged, and in this section $n_i = n_\infty$ (= 1 in dimensionless units).

In the rest frame of an electron hole of constant size and shape, the potential is steady $\frac{\partial}{\partial t} = 0$. Then the total electron energy is $\mathcal{E} = \frac{1}{2}v^2 - \phi$, constant on particle orbits. A further algebraic convenience in performing velocity space integrals

is gained by using an alternative velocity $u = v/\sqrt{2}$ (which can be considered to be the velocity scaled to the alternative form of the thermal velocity $\sqrt{2T_e/m_e}$) because then $\mathcal{E} = u^2 - \phi$ and various tiresome factors of 2 are eliminated from the equations.

When a steady potential is given, we therefore know the orbits in phase space to be the contours of constant energy, and that f is constant on them. Referring to Fig. 3(a), we see that if the distribution function $f(u)$ is known at the peak ($\phi = \psi$) of the potential, at $x = 0$, then it is known on *all* orbits. Denoting that distribution function as f_0, so $f(u, x = 0) = f_0(u) = f_0\left(\sqrt{\mathcal{E} + \psi}\right)$, the distribution function anywhere else is simply a function of the potential (and velocity)

$$f(u,x) = f_0\left(\sqrt{u^2 - \phi(x) + \psi}\right). \tag{3}$$

Consequently the density is

$$n_e(\phi) = \int f(u,\phi)du = \int f_0\left(\sqrt{u^2 - \phi + \psi}\right)du, \tag{4}$$

which provides us with the charge density $\rho(\phi)$, also a function only of potential. It should be noted that being constant on orbits, the trapped part of the distribution (for $\mathcal{E} < 0$) must be *symmetric*, i.e., independent of the sign of velocity; in contrast, the passing part ($\mathcal{E} > 0$) need not be symmetric, depending on the symmetry of the external distribution.

The solution task is to find self-consistent $\phi(x)$ and $f_0(u)$ that satisfy Eqs. (2) and (4). It is helpful to portray it using Fig. 6, which shows (in the rest frame of the hole) four quantities related to one another, with appropriately matched axes. The right hand column has a spatial position x as its abscissa (x-axis), and shows half of the hole potential, panel (c), and the corresponding contours of constant energy (and hence constant f), panel (d). The lower row has a velocity u as its ordinate (y-axis). Panel (a) plots the distribution function $f_0(u)$ (at $x = 0$), but it does so sideways so as to match panel (d). The upper row has an electric potential as its ordinate, and panel (b) will be explained later.

A. Integral solution approach

There are two different ways to obtain hole solutions. These are called by Bernstein, Greene, and Kruskal[1] the "integral equation" and the "differential equation" methods. In the literature, the integral approach is sometimes called the BGK approach, and the differential is often called the Sagdeev, or Schamel (after influential authors), or Classical Potential approach; but actually BGK's paper introduced both approaches.

In the integral approach, one starts with a specified potential shape $\phi(x)$ (panel c) which defines orbits (panel d)

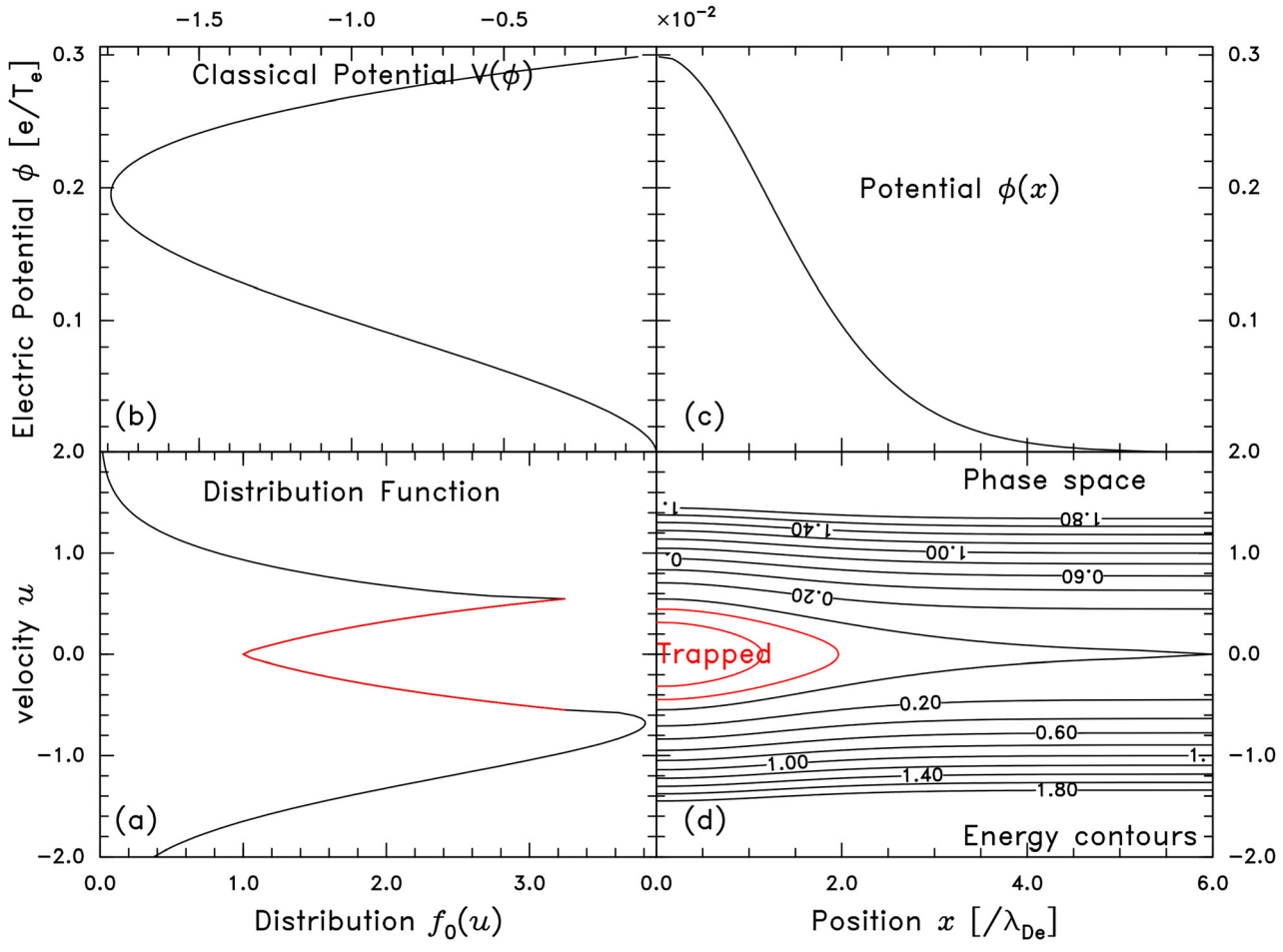

FIG. 6. Illustration of the dependencies in the solution of an electron hole.

and also gives $\rho(\phi) = -\frac{d^2\phi}{dx^2}$ and consequently the total electron density $n_e(\phi)$. The background distribution function $f(u, \infty)$ is also presumed given. It then determines everywhere the *passing* particle part of the distribution f_p, and the density attributable to it n_p. The remainder of the electron density $n_e - n_p = n_t$ is the trapped density; and

$$n_t(\phi) = \int f_t(u)du = \int_{-\phi}^{0} f_0\left(\sqrt{\mathcal{E}+\psi}\right)\frac{d\mathcal{E}}{\sqrt{\mathcal{E}+\phi}}. \quad (5)$$

This is an integral equation that must be solved to find the trapped part of the distribution function ($f_t(u) = f_0\left(\sqrt{\mathcal{E}+\psi}\right)$ for $\mathcal{E} < 0$) knowing $n_t(\phi)$. Fortunately it is Abel's integral equation, whose solution is well known

$$f_0(u_0) = \int_0^{\psi-u_0^2} \frac{dn_t}{d\phi}\frac{d\phi}{\pi\sqrt{\psi-u_0^2-\phi}}. \quad (6)$$

That result fills in the trapped part[47] of panel (a), the passing part was known from the presumed background distribution and the passing orbits (d). Panel (b) is not needed for the integral approach.

The remarkable thing about this result is that it appears to show that any monotonic potential shape ($\phi(x)$) hole can be constructed. There are, however, constraints that must be satisfied. One is that f_t cannot be negative. Others relate to the smoothness of f, which is more subtle. We'll return to these points later.

B. Differential solution approach

The alternative solution approach is to start by specifying the entire distribution $f_0(u_0)$, panel (a). That provides us with the charge density $\rho(\phi)$. Poisson's (differential) equation can then be solved by multiplying it by $\frac{d\phi}{dx}$

$$-\frac{d\phi}{dx}\frac{d^2\phi}{dx^2} = -\frac{1}{2}\frac{d}{dx}\left(\frac{d\phi}{dx}\right)^2 = \frac{d\phi}{dx}\rho, \quad (7)$$

to obtain (taking $\phi = 0$ at $d\phi/dx = 0$)

$$-\frac{1}{2}\left(\frac{d\phi}{dx}\right)^2 = \int_0^{\phi}\rho(\phi)d\phi \equiv V(\phi), \quad (8)$$

which defines what is called a Classical or Sagdeev potential V, simply the integral of charge density with respect to potential. It is plotted (sideways so as to match ordinates) in panel (b). The spatial form of ϕ is then implicitly the second integral, obtained from the square root of Eq. (8), that is $\pm\int\frac{d\phi}{\sqrt{-2V(\phi)}} = [x]$.

The boundary conditions need a careful consideration. A hole (half-)solution generally has two positions where $d\phi/dx = 0$, so $V = 0$. One such boundary is the distant edge of the hole, where it merges into the background plasma. We can choose ϕ to be zero there. The final merging takes place at $|x| \to \infty$, provided $V \sim \phi^p$ with $p \geq 2$. This is the usual presumption although it is mathematically possible to have a hole-edge at finite distance when $p < 2$.

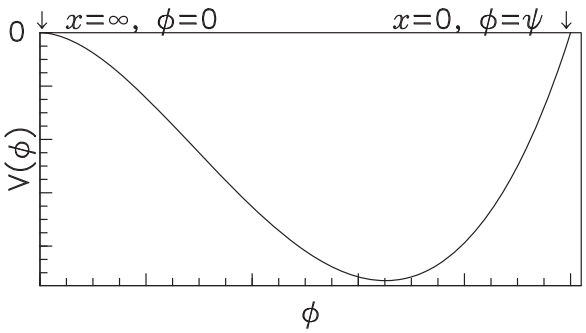

FIG. 7. Form of $V(\phi)$ defining boundaries.

The other $V = 0$ position ($x = 0$) is the potential peak where $\phi = \psi$. This limit must be localized, requiring $V \sim (\psi - \phi)^p$ with $p < 2$. Consequently $\sqrt{-V(\phi)}$ emerges from $V(\psi)$ with an infinite slope, and the form of V is as shown in Fig. 7. Then

$$\int_{\phi}^{\psi}\frac{d\phi'}{\sqrt{-2V(\phi')}} = x(\phi). \quad (9)$$

The differential approach thus starts at panel (a) of Fig. 6, and moves through (b), and (c), to (d) defining the orbits. There is, however, no guarantee that the features of the originally specified $f_0(u_0)$ actually align with the orbits that are obtained. In particular, the distribution function of an electron hole is generally depleted in the trapped region and specified by the external plasma in the passing region. There is therefore a slope discontinuity or distinctive feature at the separatrix. If we start with an arbitrary hole "width" (separatrix u-value at the peak $u_s(x = 0)$) and arbitrary hole "depth" (degree of depletion of the trapped distribution, $f_0(u_s) - f_0(0)$) the orbits will not in fact align at the boundary between panels (a) and (d). The solution will be inconsistent, not with the equations, but with the implicit presumptions about how the specified f_0 shape is related to the trapped and untrapped regions. A fully consistent hole calculation using the differential approach must therefore solve an eigenproblem: only certain related combinations of hole width and depth are permitted that satisfy a consistency condition. The consistency condition can conveniently be expressed by noticing that the total charge of the (half-)hole must be zero, which is the same thing as requiring V to be zero at both boundaries:

$$0 = \int_0^{\psi}\rho d\phi = \int_0^{\psi} n_i(\phi) - n_p(\phi)d\phi - \int_0^{\psi} 2\int_0^{\sqrt{\phi}} f_0\left(\sqrt{u^2-\phi+\psi}\right)dud\phi. \quad (10)$$

The passing (ion+electron) contribution $n_i - n_p$ is determined by external conditions. The final trapped integral must zero the total. In general satisfying this consistency condition requires that the width and depth are approximately proportional to one another $u_s \propto f_0(u_s) - f_0(0)$. The constant of proportionality is determined by the presumed shape of the hole distribution function in the trapped region and by the external distribution function f_p.

A very fruitful hole model introduced and explored by Schamel[3,48–53] and others is to take the shape of the trapped distribution to be the exponential of a parabola $\propto u^2$. It can be argued[2] that this shape, which is sometimes called a Maxwell-Boltzmann hole, is the state of minimum entropy subject to certain constraints. However, since many holes do not have this shape, and in a collisionless situation any approach to minimum energy is extremely slow, it is probably better to regard the form as representing the first two terms of a Taylor expansion of the hole shape: a mathematical ansatz. In a shifted-Maxwellian external distribution, the distribution function is then

$$f(u) = \begin{cases} \frac{1}{\sqrt{\pi}} \exp\left(-\left[\pm\sqrt{u^2-\phi}+U\right]^2\right), & \mathcal{E} > 0 \\ \frac{1}{\sqrt{\pi}} \exp\left(-\beta\left[u^2-\phi\right]-U^2\right), & \mathcal{E} \le 0. \end{cases} \tag{11}$$

Here, β is the inverse "temperature" (in units of T_e) of the trapped region, a negative quantity; U is minus the shift of the external distribution in the hole frame (the hole velocity in the stationary Maxwellian frame).

Fig. 8 illustrates this distribution, and some other features to be explained in Section II C. The specified shape allows one to write integral relations between the hole depth (represented by β) and width (represented by ψ). If the hole is narrow, in the sense $\psi \ll 1$, (and hence shallow) an approximate analytic solution by expansion is[54]

$$n_e = 1 - \frac{1}{2}Z_r'(U)\phi - \frac{4}{3}b\phi^{3/2} + \frac{1}{32}Z_r'''(U)\phi^2 + O\left(\phi^{5/2}\right), \tag{12}$$

where

$$b = \exp(-U^2)\left(1-\beta-2U^2\right)/\sqrt{\pi}. \tag{13}$$

The resulting analytic spatial form using terms to order $\phi^{3/2}$ is $\phi(x) = \psi\,\mathrm{sech}^4\left(\left[b\sqrt{\psi}/15\right]^{1/2}x\right)$, and the consistency condition, which is often referred to as the "Nonlinear Dispersion Relation" is

$$-\frac{1}{2}Z_r'(U) = \frac{16}{15}b\sqrt{\psi}. \tag{14}$$

Here, $Z_r'(u)$ is the derivative of the real part of the Fried-Conte plasma dispersion function.[55] The combination $-\frac{1}{2}Z_r'(U)$ is equal to 1 at $U=0$ and falls monotonically to zero at $U=0.92$. A slow hole ($U \ll 1$) thus has $\frac{16}{15}b\sqrt{\psi} \simeq 1$, and gives a hole spatial form

$$\phi(x) = \psi\,\mathrm{sech}^4(x/4). \tag{15}$$

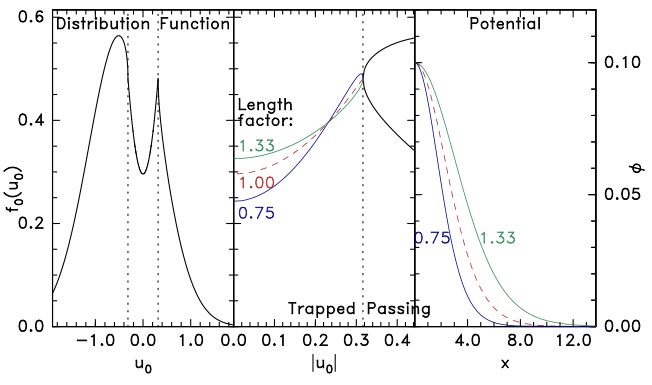

FIG. 8. Example Schamel distribution shape Eq. (11), for $\psi = 0.1$, $U = 0.4$. Center panel shows in detail the low-speed region, with the nominal trapped distribution drawn dashed. Two adjacent additional distributions are shown; they are (integral approach) solutions for longer or shorter than nominal hole length, with spatial form plotted in the third frame.

C. Comparing approaches, the hole length

Over the years, there has been some controversy about the relative merits of the integral and differential approaches. As far as mathematical convenience is concerned, the integral approach is more direct, and starts with what generally emerges from the observations: the potential shape (and the background or passing $f(v)$). It also appears to permit almost any hole length (spatial extent). By contrast, the differential approach requires an eigenproblem to be solved; and it appears to *prevent* long hole solutions (except near $U=0.92$). Setting aside mathematical or practical convenience, the interesting question is why there appear to be different conclusions about the allowable hole length.

There is no controversy about *minimum* hole length. It is easy to understand. Minimum length is constrained by non-negativity of f_e. Requiring $f_t \ge 0$ gives an upper limit on the hole depth and hence the charge density ρ. Consequently the ϕ-curvature $|d^2\phi/dx^2|$ in the region near $x=0$ has an upper limit. It cannot be greater than $\rho = n_i - n_p \simeq \left(2/\sqrt{\pi}\right)\sqrt{\psi}$ (ignoring n_i variation) obtained by putting the trapped f_t equal to zero. If we set the curvature constant, equal to that limit, an inverted parabolic shape of $\phi(x)$ is obtained whose half-length L (where $\phi = 0$ at most) must be at least such that $|d^2\phi/dx^2|L^2 \simeq \left(2/\sqrt{\pi}\right)\sqrt{\psi}L^2 \sim \psi$. This implies

$$L_{min} \sim \psi^{1/4}. \tag{16}$$

The precise coefficient of proportionality depends[14,56] upon the external distribution which determines $n_p(\phi)$. There is observational statistical evidence that approximately confirms this minimum size.[57,58]

The *maximum* allowable hole length is where opinions are divided. There is no mathematical limit on hole length unless additional constraints on f are enforced. But there are physical arguments that might lead one to enforce them; so we briefly explore the mathematical consequences of hole length increases.

Consider the slow Schamel hole Eq. (15). Why does it have the asymptotic ($x \to \infty$ and $\phi \to 0$) form $\phi \sim \mathrm{sech}^4(x/4) \to \exp(-x)$? The answer is that, asymptotically, the potential is governed by approximately Debye shielding: $n_e \propto e^{\phi} \to 1 + \phi$. As $\phi \to 0$, the trapped region becomes infinitesimally narrow, $u_s = \sqrt{\phi} \to 0$. Moreover, if the trapped distribution $f_0(u)$ has a finite slope at the separatrix (u_s), then the local depth of

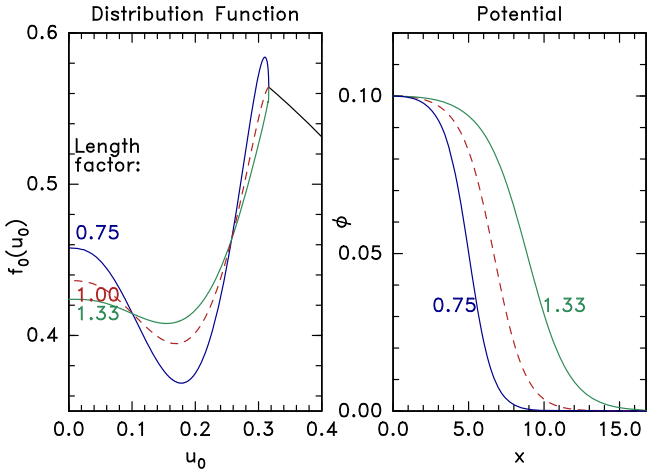

FIG. 9. Numerical solution of the BGK integral equation to obtain the trapped electron distribution function. When the length factor L is increased or decreased from its nominal value 1, slope singularities appear at the separatrix, $u_s = \sqrt{\psi} = \sqrt{0.1}$. Also this flattened-top potential requires a non-monotonic inner f_0.

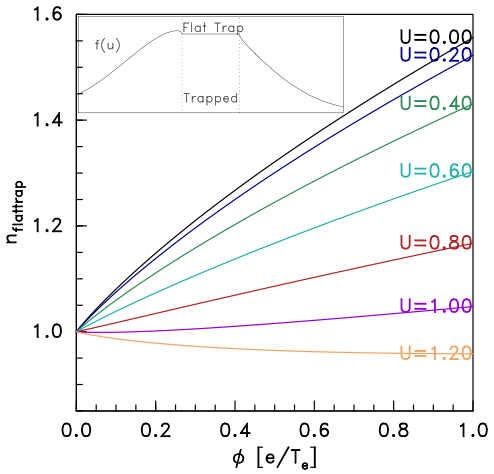

FIG. 10. Flat-trapped distribution total density in a Maxwellian external distribution, for different hole speeds U. Inset illustrates the velocity distribution $f(u)$ for $U = 0.2$, $\phi = 0.1$.

the trapped region $f_0(u_s,\phi) - f_0(0,\phi)$ becomes small, and the trapped distribution can be considered to be asymptotically flat, i.e., $f_t(u)$ independent of u. The total density of a flat-trapped $f_t(u) = const. = f_p(u_s)$ distribution, when the background passing distribution is Maxwellian, is readily evaluated as $n_e = n_{flattrap} = n_\infty[2\phi^{1/2}/\sqrt{\pi} + e^\phi \mathrm{erfc}(\phi^{1/2})]$ (for zero U) and has the asymptotic form $n(\phi) \to 1 + \phi$. This linear asymptotic dependence is because there is an exact cancellation between the $\sqrt{\phi}$ terms in n_e that arise from the trapped (first term) and passing (second term) densities.

In order to change the asymptotic exponential fall-off length, one must change the trapped density dependence on ϕ in such a way that a different coefficient of ϕ arises in the total density. That requires there to be a slope singularity of trapped $f(u)$ at the separatrix. This effect is clearly visible in the central panel of Fig. 8, where the nominal hole length is changed by the factors indicated. The required distribution function is obtained by a numerical implementation of the integral approach in Figs. 8 and 9. Fig. 9, shows the distribution solution for a potential shape (right panel) $\phi = \psi[e^F + 1]/[e^F + \cosh^4(x/4L)]$ (and zero hole velocity $U = 0$, in external Maxwellian distribution). The parameter F prescribes a flattening of the top of the potential. Positive values of F such as $F = 4$, used in Fig. 9, are approximately the length of the flattened region. Negative values of F quickly remove any flattening and return to a $\psi\ \mathrm{sech}^4(x/4L)$ form. The consequences of F will be discussed in a moment. The length factor, L, changes the asymptotic potential fall-off at large x. The first key aspect of Figs. 8 and 9 is to observe (top right of left panel of Fig. 9) the slope singularities at $u \to u_s$, when $L \neq 1$.

The slope singularity requirement can be demonstrated analytically as follows. If we suppose the trapped distribution to have a difference from flat of $f_t - f_p(u_s) = \delta f = \alpha(-\mathcal{E})^{1/2} + \beta\mathcal{E}$ ($\mathcal{E}$ being the energy, which is negative), then it can readily be shown that the trapped density difference relative to a flat distribution is $\delta n = \frac{\pi}{2}\alpha\phi + \frac{4}{3}\beta\phi^{3/2}$. Thus, a trapped contribution to density difference proportional to ϕ requires an f contribution proportional to $(-\mathcal{E})^{1/2}$, which has a slope singularity at $\mathcal{E} = 0$. And this is the key point: Schamel's form disallows such an infinite-slope distribution.[59] It enforces a finite f-slope constraint amounting to $\alpha = 0$. Therefore, its asymptotic decay-length is fixed.

This type of analysis also explains the variation of Schamel hole length with speed, which shifts the external Maxwellian in the hole frame. A more general Debye shielding calculation obtains $\phi \propto \exp(\phi/\lambda)$, where $\lambda = (dn/d\phi)^{-1/2}$, which follows immediately from $\delta n \approx \phi\, dn/d\phi$ and $d^2\phi/dx^2 = \delta n$. As hole speed U increases, λ increases because $dn_{flattrap}/d\phi$ decreases, as illustrated by Fig. 10 (in which the slope of the curves at $\phi \to 0$ is equal to $-\frac{1}{2}Z_r'(U)$). The hole length in the differential approach also increases $\sim \lambda$. In fact, in Fig. 8, for which $U = 0.4$, the nominal hole length was $\lambda \simeq 1.18$. It becomes infinite (for Maxwellian) at $U = v_{hole}/\sqrt{2T_e/m_e} = 0.92 = U_{max}$. Large λ-enhancement occurs only for U close to (but below) U_{max}.

Beyond that speed, no holes exist because the sign of $dn/d\phi$ for small ϕ becomes negative, being dominated by passing particles far from the separatrix. This hole speed limit can theoretically be overcome only by non-thermal distortions of the background distribution function that enhances the electron population at a slow speed in the hole frame. Non-maxwellian f_p also changes the hole length λ, in a way that has been calculated for various external distributions and Schamel type holes[60] with a view of trying to deduce something about non-thermal f_p from the hole length.

The other generic possibility for increasing the hole length is to *flatten* the top of the potential in the center. That reduction in $|d^2\phi/dx^2|$ requires a reduction in the positive charge at the hole center, by an increase of electron density there, back towards the external density. So, as illustrated in Fig. 9, in order for the electron density in the flat region to return toward the external value, the phase-space density $f(u)$ on the innermost (lowest energy) trapped orbits must be greater than on the intermediate-energy orbits responsible for the negative-ρ region of the $\phi(x)$ profile. Thus, $f_0(u)$ must have a local minimum at an intermediate speed, u_m, such that $0 < u_m < u_s$. The Schamel form of hole excludes

that possibility by its presumed monotonic $f_t(u)$ shape; so it cannot accommodate hole lengthening by central flattening.

The extreme limit of flattening, when there is literally a finite region $|x| < x_f$, where $d\phi/dx = 0$, requires a zero charge density there, with the electron density equal to the (presumably unperturbed) ion density. Such a situation can be considered to be two double-layers, back to back. There is evidence for the existence of such extended hole structures in space.[61] The inner shape of the distribution $f_0(u)$ required for such extreme flatness looks hardly different from that illustrated in Fig. 9 for $L = 0.75$.

Whether or not finite separatrix f-slope or other constraints such as monotonicity on the shape of f_t, implied by a differential approach, ought in fact to be enforced are the real physics questions behind the integral versus differential debate. They are not fully resolved. To do so requires theoretical considerations beyond those implied by the presumption of constant-f on steady orbits. One must take into account the initial (unsteady) conditions of hole formation, and the dynamics of particles close to the separatrix, whose transit time becomes theoretically infinite, and for which f being a function only of energy therefore becomes poorly justified.

III. MULTIPLE DIMENSIONS AND TRANSVERSE STABILITY

Everything discussed so far is one-dimensional. What happens in the real three-dimensional world? Clearly holes do exist under some circumstances, because they have been observed. But what are the conditions for their existence, what happens if the conditions are violated, and what is their multidimensional character when they do exist? The summary answer is that for holes to exist, there must be a strong enough magnetic field and distribution-function anisotropy. There is considerable numerical simulation evidence that holes in a magnetic field weaker than to satisfy $\Omega \gtrsim \omega_b$ (where $\omega_b \sim \sqrt{\psi}\,\omega_p$ is the trapped-electron bounce-frequency) are broken up by a transverse "kinking" instability, but no convincing analysis has determined the detailed mechanism or threshold. Moreover, even if the magnetic field is strong, there is computational evidence of transverse modulation or break-up of holes by whistlers.

Analytic existence criteria in 3-d have a much shorter and sparser literature than simulations. But it has been argued,[62] by reference to early papers[63–65] that the non-existence of unmagnetized isotropic-distribution holes is simple, arising from the fact that the trapped-electron phase-space volume is proportional to $\psi^{N_d/2}$ where N_d is the number of dimensions. The depletion of trapped density is therefore too small for $N_d \geq 2$ to furnish the positive charge required to sustain a hole. Although spherically symmetric holes (without B-field) have been shown to exist whose passing (as well as trapped) velocity distributions are anisotropic, depending upon angular momentum as well as energy,[66,67] they imply implausibly pathological background distributions at infinity.[68] Having an imposed magnetic field bypasses these non-existence proofs, which seem essential for multidimensional holes. With a strong magnetic field, reduced-dimensionality (e.g., gyrokinetic) plasma treatments can yield tractable *equilibria* (e.g., Refs. 12 and 69) but the question still remains whether such equilibria are or are not *stable*. If they are unstable on electron transit timescales, then they can be observed, if at all, only transiently.

It was found in the earliest PIC research[25] that in unmagnetized two-stream instabilities, holes did not visibly form in two- and three-dimensional simulations, though they did very clearly in 1-d. If a magnetic field is included (still using electrostatic response) then multidimensional simulations (using far greater computational power) in the 1990s and later, showed hole formation, but a subsequent hole modulation and break up.[11,70–76]

Studies of an initially two-stream electron distribution are illustrated by Fig. 11. The simulation has a strong magnetic field $\Omega/\omega_p = 5$. Its snap-shots show the initial instability (at a nonlinear stage), the individual electron holes into which it evolves, extended in the transverse (y) direction, and a final state with strong "whistlers." These are near-horizontal ($k \sim k_\perp$) striations that have the effect of partially breaking up the holes into shorter individual "blobs" of lower electron density, that are sometimes almost spherical.

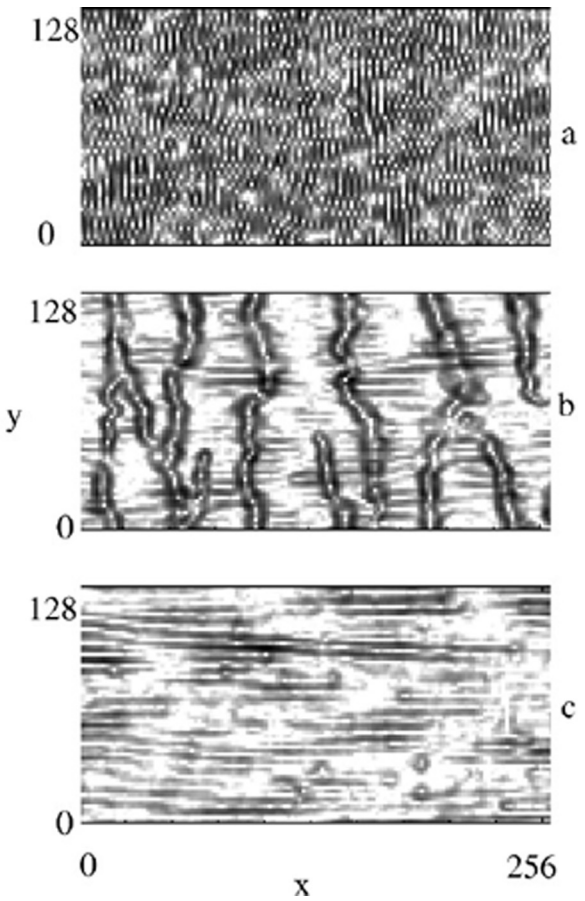

FIG. 11. Simulations of two-stream instability showing a hole formation and break up adapted from M. Oppenheim, D. L. Newman, and M. V. Goldman, Phys. Rev. Lett. **83**, 2344–2347 (1999). Copyright 1999 American Physical Society.[72] The magnetic field is parallel to the x-axis, and at early time $t = 24\omega_p^{-1}$ (a) nonlinear unstable waves arise, which quite soon, $t = 448\omega_p^{-1}$ (b), merge into discrete holes with long transverse extent. At a much later time, $t = 1920\omega_p^{-1}$ (c), "whistler" fluctuations with transverse wavevector (striations almost aligned with x) predominate, and the holes are less evident in the quantity contoured: $\log|E|^2$.

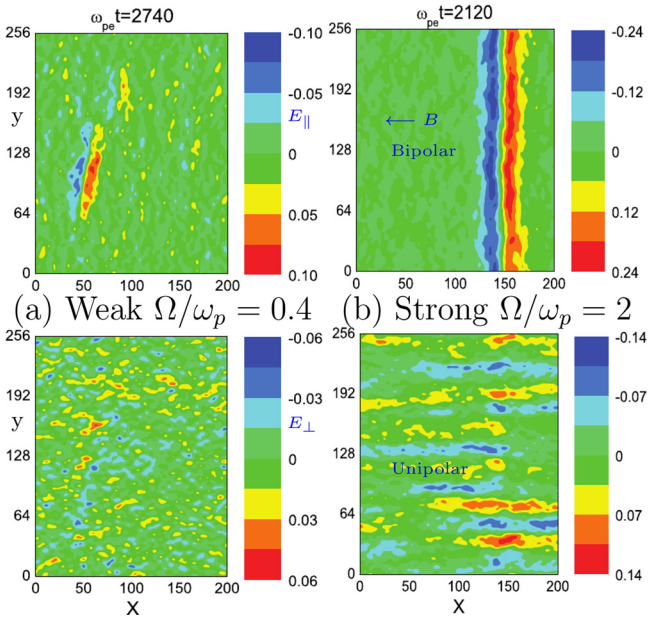

FIG. 12. Long-time behavior of two-stream instability-generated holes in two dimensions for stronger (a) and weaker (b) magnetic field strength, adapted from Q. M. Lu *et al.*, J. Geophys. Res. **113**, A11219 (2008). Copyright John Wiley and Sons.[76] Upper panels $E_\parallel$; lower panels $E_\perp$.

The effect of the magnetic field strength is illustrated in Fig. 12, also a two-stream instability simulation, but showing separately $E_\parallel$ and $E_\perp$, for two cases: weaker and stronger B-field: $\Omega/\omega_p = 0.4, 2.0$. The strongly magnetized case (right) shows even after $t = 2100$ that an elongated hole persists in $E_\parallel$. The whistlers are also very evident in $E_\perp$ at a comparable field intensity. The weaker magnetic field case (left) has holes with limited transverse y-extent, although still present at $t > 2700$. The whistlers are hardly evident.

The two-stream studies are of rather violent instability, and show a complicated nonlinear evolution caused by multiple processes. In order to try to understand the processes better, it is valuable to run simulations that start with an already-formed isolated hole. In effect, this is an exploration in multiple dimensions of the stability of simplest equilibrium, namely, an equilibrium independent of the two transverse cartesian coordinates (which is precisely the 1-D hole). This approach has yielded insight into the transverse instability to which holes are susceptible. Fig. 13 shows frames from a simulation that is 2-d in space but 3-d in velocity ("2d3v") as is necessary for a magnetized plasma. It starts from a $\mathrm{sech}^4(x/4)$ Schamel-form hole (without any applied perturbation other than noise) in a stationary Maxwellian plasma (with immobile ions) and has a fairly weak magnetic field such that $\Omega/\omega_p = 0.2$. It is initialized uniform in the y ("axis-2") direction perpendicular to B at $t = 0$. A long wavelength (constrained by the box size) kink, consisting predominantly of hole displacement $\delta x(y) \propto \exp(iky)$ of the hole along the B-field (x-direction), with little change in potential shape or height, grows to a significant amplitude at $t = 204$. As time progresses, the perturbation causes a substantial reduction in the hole potential height together with a separation of the kinked region into two different y-humps. The full time evolution as

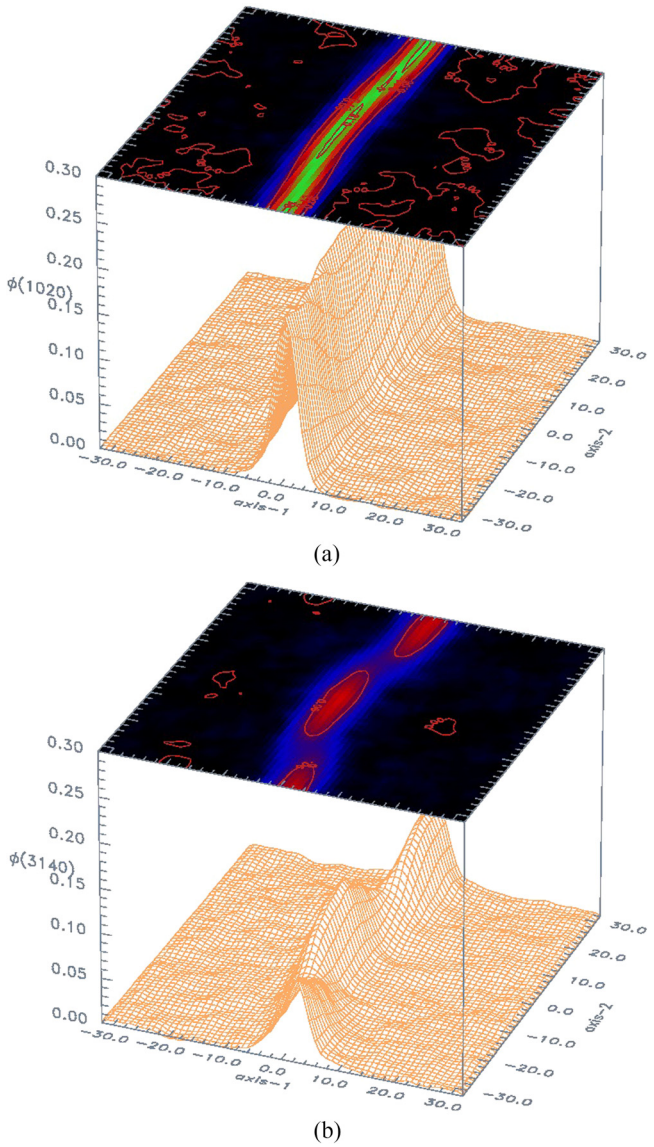

FIG. 13. Transverse instability of a 2-D electron hole at $\omega_p t = 204$ (a) and $\omega_p t = 628$ (b). PIC simulation, 25 600 particles per Debye cell for low noise. (Multimedia view) [URL: http://dx.doi.org/10.1063/1.4976854.2]

an animation is available at the URL in Fig. 13. There are no signs of the "whistler" perturbations.

Simulations like those of Fig. 13 with different parameters often result in almost a complete destruction of the hole, as was observed in early studies of the transverse instability of an isolated prepared hole using a PIC code by Muschietti *et al.*[73] Those authors observed that the stability boundary was approximately at $\Omega/\omega_b = 1$. Larger values suppressed the kinking. The trapped particle bounce frequency is $\omega_b \simeq \omega_p\sqrt{\psi}/L$ for hole length L. Later simulations of the same cases found very similar results.[77] The amplitude and length of holes measured at the magnetopause,[32] for example, are observed to obey the stability criterion $\Omega/\omega_b > 1$. Muschietti *et al.* also proposed single-particle focusing by the transverse electric field of the kink as the instability mechanism, a supposition that has been taken up by subsequent authors. However, their numerical results showed unexplained dependence upon the details of the external distribution function

such as the effective transverse temperature. My own investigations (unpublished, using the code COPTIC[78,79] which also produced the results of Fig. 13), while they confirm the stability criterion within approximately a factor of 2, and show a self-stabilization of the holes by amplitude (ψ) reduction, do not reproduce the charge density perturbations that were much of the evidence that Muschietti *et al*. invoked for their mechanism identification. Moreover, analytic investigations of 2-d hole stability[80] have not yielded eigenmodes consistent with the simulation's observations. Therefore, overall, the detailed mechanism of transverse instability and analytic understanding of the stability criterion appear still to be open questions.

IV. HOLE CAUSES AND CONSEQUENCES

Broadly speaking, if there is sufficiently a strong magnetic field, electron holes can be and often are formed whenever an electron velocity distribution arises that is electrostatically (Penrose[81]) unstable. Holes are the natural long-term non-linear states of many "bump on tail" (beam-plasma), and "warm bistream" (dimpled) distributions[10,11] as well as the well-separated two-stream instability distributions illustrated by Fig. 2. They can be expected whenever the instability is dominated by a single wavelength or narrow range of wavelengths. The alternative (non-hole) saturated final state of collisionless instability is the generation of a wide range of wavelengths incoherent with each other, giving rise to quasi-linear diffusion and flattening of the distribution in the unstable region. Simulations suggest that hole formation is at least as important as quasi-linear flattening in many cases. It may be that ion parameters, (e.g., ion temperature) influence the predominance of electron holes. However, one must be cautious in concluding so from simulations with artificially low mass ratio m_i/m_e, because such treatments (often adopted to decrease computational cost) reduce and largely remove the velocity separation of ion and electron waves. Electron (and ion) holes are observed in simulations to result from electron-ion instabilities[5,82] of the kinetic Buneman type, i.e., caused by a relative flow of electrons and ions. This instability can be enhanced by the presence of a distortion of f_e in the form of an electron beam[83] and lead to initially very slowly moving holes. Holes have been observed in space measurements of reconnection,[44,84,85] where they appear to result from electron beams of small angular spread; their occurrence is considered an indicator that an ambient guide magnetic field is present. The kinetic-electron simulation of reconnection is a very fast-developing field[46,86,87] where electron holes are now a frequent phenomenon.

I now use as an illustration, a different context where electron holes prove to be key. It is in the cross-field flow of plasma past an obstacle. One can speak of this as the flow of the solar wind past the moon, although the physics applies to other situations too. When this happens, the plasma wake is highly depleted immediately behind the obstacle, but fills in by mostly parallel (cross-wake) flow of electrons (quickly) and ions (more slowly), forming a negative potential hill that repels electrons and attracts ions. In this wake, the collisionless electron distribution function acquires a "dimple" (or

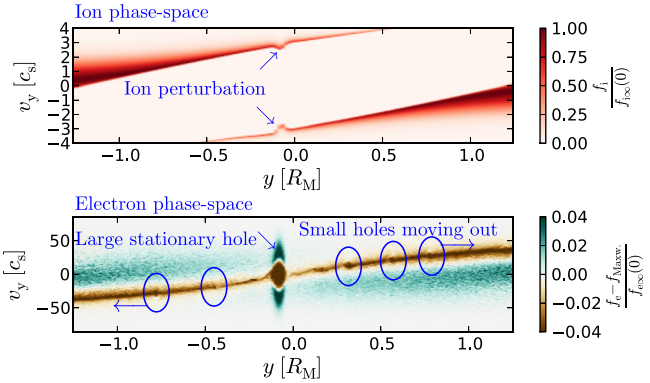

FIG. 14. Holes growing in a cross-field plasma wake. Most holes move out at the electron orbit speed. One here remains stationary and grows large enough to perturb the ions. (Multimedia view) [URL: http://dx.doi.org/10.1063/1.4976854.3]

"notch") in the vicinity of orbits that are *only just* reflected or passed by the repelling potential hill.[88] This dimple is at a different local velocity depending upon the transverse position in the wake, which determines how far down the transverse wake potential profile the location is. A dimple near the center of the electron distribution is sometimes referred to as giving rise to a "warm bistream instability."[10,32,70] Such a distribution arises from counter-propagating streams of electrons of comparable density whose thermal speed is comparable to the stream speed. The wake does represent counter-propagating streams of electrons entering from either side along the B-field. But at positions away from the wake symmetry axis, electrons reflected from the repelling wake potential are also important "counter-propagating" contributors to the shape, and the dimple is not centered. In any case, the dimple is (Penrose) unstable, and in 1-d, PIC simulations are observed[89] to spawn electron holes. Fig. 14 illustrates the parallel phase space of ions (top) and electrons (bottom). The space coordinate, y, here is along the B-field which is across the wake, whose width is approximately twice the obstacle radius R_M. The simulation has a very small Debye length $\lambda_{De} = 0.00125R_M$. The mass ratio for this plot, $m_i/m_e = 459$, assists the visibility of the features; but similar results are obtained at a physical mass ratio (1836). The electron color contours are of the difference between the actual PIC f_e and a Maxwellian of the same density (even though this is a full-f simulation). This difference more clearly shows the dimple, along the diagonal separatrix in phase-space that arises from the wake's electron-repelling potential structure. Along the dimple, barely visible small electron holes are continually spawned. Most of them move out of the wake almost exactly following the individual electron phase-space orbits (that is, following the separatrix) before experiencing much growth. But some holes (one in this case) remain almost stationary and during their long dwell time in the wake, grow in amplitude, often sufficiently to perturb (and later to disrupt) the counter-propagating ion streams shown in the top panel. This full process can be viewed as an animation at the URL in Fig. 14. The growth of pre-formed coherent holes is therefore crucial to understand how the wake eventually re-thermalizes the ions. The growth mechanism has been identified[90] as, in

essence, the result of movement of a hole from a lower to a higher density background plasma. Since the trapped phase-space density is fixed while the external phase-space density rises, the hole deepens, and the amplitude (velocity width) therefore increases. This is an example of a situation where the entire outcome depends upon how an electron hole *behaves as a composite object*: how it moves or does not move, and how it grows or does not.

V. HOLISTIC KINEMATICS AND SELF-ACCELERATION

In view of the importance of the behavior of holes as composite objects, the question arises as to what determines the hole's lumped parameters and their evolution. Most notably what determines the hole speed. Contrary to some implications in the literature, the "non-linear dispersion relation" arising from the consistency condition does not fix the speed. For a given speed, it relates depth and width, and determines length (under the differential approach). But there is great freedom for a hole to adjust these other parameters to accommodate different speeds up to the limit of U_{max}. It is known that the influence of the ions is important, and slow holes interact with the ion response in a way that has been referred to as a "coupled electron hole and ion soliton,"[91,92] which can sometimes lead to splitting of slow holes. Perhaps the most remarkable consequence is that holes deliberately formed at effectively zero initial speed with respect to a Maxwellian ion distribution are observed in simulations to self-accelerate quickly to high speed. Fig. 15 shows an example of a (continuum) Vlasov 1-d simulation,[19,93] where a hole is initialized with an electron distribution of the Schamel type, but uniform ions. As time evolves, the ions are repelled by the hole and suddenly at time $t \simeq 130$, the hole heads away at a speed approximately half the electron thermal speed, leaving the ion perturbation (and a smaller corresponding electron perturbation) behind. A heuristic understanding of this process is simple. Initially near-stationary uniform ions repelled outward by the sudden application of the hole potential give rise to a negative central charge density that repels the hole itself (which has an effective charge-to-mass ratio that of electrons). But the question arises whether one can predict such phenomena *quantitatively*. Dupree[2,94] developed expressions for the charge, momentum, and energy of holes conceived as holistic objects. But his interest was mostly ion holes in the context of simulations of extremely low mass ratio (4). Consequently, he emphasized passing particle reflection as causing hole deceleration and consequent growth. Realistic

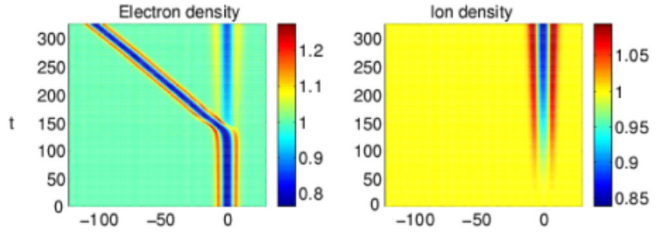

FIG. 15. Self-acceleration of an electron hole, adapted from B. Eliasson and P. K. Shukla, Phys. Rev. Lett. **93**, 045001 (2004). Copyright 2004 American Physical Society.[93]

mass-ratio electron holes usually (with some notable exceptions) move fast enough that the ion thermal velocity spread can be ignored. We have recently developed a comprehensive "kinematics" of electron hole's momentum conservation,[95] which shows that the important ion momentum transfer terms are usually those of second order in the hole potential ($\propto\psi^2$), neglected by Dupree. The new expressions of momentum, which are explicit integrals over the hole, but are more complicated than is useful to present here, explain *semi-quantitatively* hole self-acceleration such as that of Fig. 15 and *quantitatively* the self-acceleration of holes initialized less violently. The agreement with individual hole simulations[96] is gratifying, as illustrated by Fig. 16.

Moreover, the hole kinematics explains an important feature of the phenomena of Fig. 14. The question is, what keeps the stationary holes that grow large from moving out of the wake? The answer can be summarized by saying that holes cannot, without violence, overtake ion streams. The kinematic calculations show that if a hole velocity smoothly approaches the velocity of an ion stream, it experiences a momentum transfer that becomes larger and larger, and opposes the approach. The strength of that ion interaction is more than sufficient to overcome the hole acceleration from background electric field, which is what causes the small holes to move as they do, along the separatrix in the wake. The stationary hole(s) are those whose velocities lie between the velocities of the two ion streams, shown in the upper panel. Those streams act as velocity barriers, imprisoning the hole velocity between them, preventing the hole from escaping from the wake. Thus it is a complex interaction between the streaming character of the ion distribution and the holistic kinematics of the electron hole that explains the slow holes.

The reason why this holistic hole analysis works so well is that the hole is a long-lived object whose self-consistent structure is remarkably rigid for times longer than the electron transit timescale. There is a very wide range of timescales from the period of hole-length oscillation, of order the transit or bounce time $t_t \sim \lambda_D/v \sim \omega_b^{-1}$, out to the long collisional relaxation time t_c of the distribution function. What is more, provided it does not experience transverse instability, a hole rather well preserves its potential height, length, and shape within quite wide limits of hole velocity changes. For phenomena such as acceleration or amplitude growth having timescales τ between, but well separated from, the two extremes, $t_t \ll \tau \ll t_c$, the dynamics, in the hole, of electrons and ions can be well described in an approximation of short

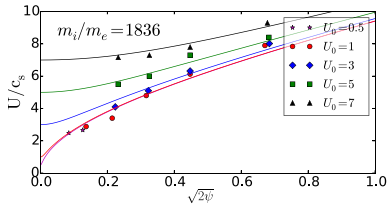

FIG. 16. Self-acceleration of an electron hole, adapted from C. Zhou and I. H. Hutchinson, Phys. Plasmas **23**, 082102 (2016). Copyright 2016 AIP Publishing LLC.[96] PIC simulation observations (points) are compared with with the kinematic theory prediction (lines). The final hole speed U, in the ion frame, is shown for different initial speeds U_0, in units of ion sound speed $c_s = \sqrt{T_e/m_i}$.

transit time compared with the typical hole perturbation time-scale. Yet collisional effects still remain negligible.

VI. CONCLUSION

Electron holes are now known to occur widely in magnetized plasmas subject to electrostatic instability arising from violation of the Penrose criterion. They may perhaps occur even more widely than already observed, because their spatial extent, of order a few Debye lengths, is very short compared with what can typically be resolved in confined plasmas; and the frequency of their potential variation in the lab frame, up to the plasma frequency, is higher than can typically be time-resolved by laboratory plasma measurements.

We know that they are implicated at the subscale of many phenomena such as reconnection, shocks, and plasma wakes. In some cases, e.g., wakes, they appear to play an important role in the macroscopic evolution. In others, e.g., space plasma reconnection, in which it is hard to know where a spacecraft is relative to macroscopic features such as the current sheet, they might offer valuable diagnostic signals of the satellite's whereabouts. Electron holes' importance in controlling or diagnosing such phenomena has yet to be fully explored.

The theory of electron holes has a 60 year history, but it is a history that has developed relatively slowly because of the difficulty of the non-linear analysis and the paucity, until fairly recently, of experimental observations. Multidimensional questions are by no means settled, despite the insights that have come from numerical simulations; the mechanisms of transverse instability are not firmly established. The subtleties of what counts as reasonable trapped distribution functions, and therefore what can be the hole lengths under what conditions, are still open questions. The quantitative analysis of holes as composite objects, for example, their kinematic behavior, is only recently beginning to be understood. The merging or splitting of holes (not covered in this tutorial) is also a holistic topic that might benefit from creative analysis to improve and establish quantitative understanding of how and when holes merge.

Simulation has played an important role in electron hole physics from the beginning. Current computational power is now sufficient that many of the tougher questions can be addressed numerically as well as analytically. Indeed, it seems likely that comparison with simulation might be the most available resource for testing new analytic developments. These, together with increasingly comprehensive satellite measurements, and perhaps innovative laboratory experiments, offer opportunities for major progress in the near future.

ACKNOWLEDGMENTS

My understanding of electron holes has benefitted from students and colleagues too numerous to list, through many enjoyable papers and conversations. This work was partially supported by the NASA, Grant No. NNX16AG82G.

[1]I. B. Bernstein, J. M. Greene, and M. D. Kruskal, Phys. Rev. **108**, 546 (1957).
[2]T. H. Dupree, Phys. Fluids **25**, 277 (1982).
[3]H. Schamel, Phys. Rep. **140**, 161 (1986).
[4]J. R. Franz, P. M. Kintner, C. E. Seyler, J. S. Pickett, and J. D. Scudder, Geophys. Res. Lett. **27**, 169, doi:10.1029/1999GL010733 (2000).
[5]M. V. Goldman, D. L. Newman, and R. E. Ergun, Nonlinear Processes Geophys. **10**, 37 (2003).
[6]D. Y. Eremin and H. L. Berk, Phys. Plasmas **11**, 3621 (2004).
[7]P. Guio and H. L. Pecseli, Ann. Geophys. **23**, 853 (2005).
[8]W. Fox, M. Porkolab, J. Egedal, N. Katz, and A. Le, Phys. Plasmas **19**, 032118 (2012).
[9]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).
[10]Y. Omura, H. Matsumoto, T. Miyake, and H. Kojima, J. Geophys. Res. **101**, 2685, doi:10.1029/95JA03145 (1996).
[11]T. Miyake, Y. Omura, H. Matsumoto, and H. Kojima, J. Geophys. Res. **103**, 11841, doi:10.1029/98JA00760 (1998).
[12]L.-J. Chen and G. K. Parks, Geophys. Res. Lett. **29**, 41, doi:10.1029/2001GL013570 (2002).
[13]L.-J. Chen, D. Thouless, and J.-M. Tang, Phys. Rev. E **69**, 055401 (2004).
[14]V. L. Krasovsky, H. Matsumoto, and Y. Omura, J. Geophys. Res.: Space Phys. **108**, 1117, doi:10.1029/2001JA000277 (2003).
[15]F. Califano, L. Galeotti, and C. Briand, Phys. Plasmas **14**, 052306 (2007).
[16]J. Pickett, L.-J. Chen, R. Mutel, I. Christopher, O. Santolík, G. Lakhina, S. Singh, R. Reddy, D. Gurnett, B. Tsurutani, E. Lucek, and B. Lavraud, Adv. Space Res. **41**, 1666 (2008).
[17]B. Lefebvre, L. J. Chen, W. Gekelman, P. Kintner, J. Pickett, P. Pribyl, S. Vincena, F. Chiang, and J. Judy, Phys. Rev. Lett. **105**, 115001 (2010).
[18]D. M. Malaspina, D. L. Newman, L. B. Willson, K. Goetz, P. J. Kellogg, and K. Kerstin, J. Geophys. Res.: Space Phys. **118**, 591, doi:10.1002/jgra.50102 (2013).
[19]B. Eliasson and P. Shukla, Phys. Rep. **422**, 225 (2006).
[20]R. E. Ergun, C. W. Carlson, J. P. McFadden, F. S. Mozer, L. Muschietti, I. Roth, and R. J. Strangeway, Phys. Rev. Lett. **81**, 826 (1998).
[21]S. D. Bale, A. Hull, D. E. Larson, R. P. Lin, L. Muschietti, P. J. Kellogg, K. Goetz, and S. J. Monson, Astrophys. J. **575**, L25 (2002).
[22]J. R. Franz, P. M. Kintner, and J. S. Pickett, Geophys. Res. Lett. **25**, 1277, doi:10.1029/98GL50870 (1998).
[23]D. Jovanovic and P. K. Shukla, Phys. Rev. Lett. **84**, 4373 (2000).
[24]F. S. Mozer, O. V. Agapitov, A. Artemyev, J. F. Drake, V. Krasnoselskikh, S. Lejosne, and I. Vasko, Geophys. Res. Lett. **42**, 3627, doi:10.1002/2015GL063946 (2015).
[25]R. L. Morse and C. W. Nielson, Phys. Rev. Lett. **23**, 1087 (1969).
[26]H. L. Berk, C. E. Nielsen, and K. V. Roberts, Phys. Fluids **13**, 980 (1970).
[27]C. K. Birdsall and A. B. Langdon, *Plasma Physics via Computer Simulation* (IOP Publishing, Bristol, 1991).
[28]K. Saeki, P. Michelsen, H. L. Pecseli, and J. J. Rasmussen, Phys. Rev. Lett. **42**, 501 (1979).
[29]F. L. Scarf, L. A. Frank, K. L. Ackerson, and R. P. Lepping, Geophys. Res. Lett. **1**, 189, doi:10.1029/GL001i005p00189 (1974).
[30]S. D. Bale, P. J. Kellogg, D. E. Larsen, R. P. Lin, K. Goetz, and R. P. Lepping, Geophys. Res. Lett. **25**, 2929, doi:10.1029/98GL02111 (1998).
[31]H. Matsumoto, X. H. Deng, H. Kojima, and R. R. Anderson, Geophys. Res. Lett. **30**, 1326, doi:10.1029/2002GL016319 (2003).
[32]D. B. Graham, Y. V. Khotyaintsev, A. Vaivads, and M. André, J. Geophys. Res.: Space Phys. **121**, 3069, doi:10.1002/2015JA021527 (2016).
[33]F. S. Mozer, O. A. Agapitov, A. Artemyev, J. L. Burch, R. E. Ergun, B. L. Giles, D. Mourenas, R. B. Torbert, T. D. Phan, and I. Vasko, Phys. Rev. Lett. **116**, 145101 (2016).
[34]L. Andersson, R. E. Ergun, J. Tao, A. Roux, O. Lecontel, V. Angelopoulos, J. Bonnell, J. P. McFadden, D. E. Larson, S. Eriksson, T. Johansson, C. M. Cully, D. N. Newman, M. V. Goldman, K. H. Glassmeier, and W. Baumjohann, Phys. Rev. Lett. **102**, 225004 (2009).
[35]C. Norgren, M. André, A. Vaivads, and Y. V. Khotyaintsev, Geophys. Res. Lett. **42**, 1654, doi:10.1002/2015GL063218 (2015).
[36]D. M. Malaspina, L. Andersson, R. E. Ergun, J. R. Wygant, J. W. Bonnell, C. Kletzing, G. D. Reeves, R. M. Skoug, and B. A. Larsen, Geophys. Res. Lett. **41**, 5693, doi:10.1002/2014GL061109 (2014).
[37]I. Y. Vasko, O. V. Agapitov, F. Mozer, A. V. Artemyev, and D. Jovanovic, Geophys. Res. Lett. **42**, 2123, doi:10.1002/2015GL063370 (2015).
[38]A. Mangeney, C. Salem, C. Lacombe, J.-L. Bougeret, C. Perche, R. Manning, P. J. Kellogg, K. Goetz, S. J. Monson, and J.-M. Bosqued, Ann. Geophys. **17**, 307 (1999).

[39]J. D. Williams, L. J. Chen, W. S. Kurth, D. A. Gurnett, M. K. Dougherty, and A. M. Rymer, Geophys. Res. Lett. **32**, L17103, doi:10.1029/2005GL023079 (2005).
[40]L. B. Wilson, C. A. Cattell, P. J. Kellogg, K. Goetz, K. Kersten, J. C. Kasper, A. Szabo, and M. Wilber, J. Geophys. Res.: Space Phys. **115**, A12104, doi:10.1029/2010JA015332 (2010).
[41]W. Bertsche, J. Fajans, and L. Friedland, Phys. Rev. Lett. **91**, 265003 (2003).
[42]L. Friedland, F. Peinetti, W. Bertsche, J. Fajans, and J. Wurtele, Phys. Plasmas **11**, 4305 (2004).
[43]W. Fox, M. Porkolab, J. Egedal, N. Katz, and A. Le, Phys. Rev. Lett. **101**, 255003 (2008).
[44]C. Cattell, J. Dombeck, 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).
[45]J. F. Drake, M. Swisdak, C. Cattell, M. A. Shay, B. N. Rogers, and A. Zeiler, Science (New York, NY) **299**, 873 (2003).
[46]M. V. Goldman, D. L. Newman, and G. Lapenta, Space Sci. Rev. **199**, 651 (2016).
[47]It is highly advantageous when evaluating numerically the integrals, to perform the solution not for the entire f_t but for the difference $f_t(u) - f_0(u_s)$ between the trapped distribution and a "flat-trapped" distribution that is independent of u and continuous with the passing distribution at the separatrix velocity u_s.
[48]H. Schamel, Plasma Phys. **13**, 491 (1971).
[49]H. Schamel, Plasma Phys. **14**, 905 (1972).
[50]H. Schamel, J. Plasma Phys. **9**, 377 (1973).
[51]H. Schamel, Phys. Scr. **20**, 336 (1979).
[52]H. Schamel, Phys. Plasmas **7**, 4831 (2000).
[53]J. Korn and H. Schamel, J. Plasma Phys. **56**, 307 (1996).
[54]The ϕ^2 term in this expansion was calculated by Korn and Schamel,[53] but contained a factor of 2 error, corrected here. Their equation A4 omits a term $-\frac{1}{2v^2}f'\phi^2$ because it uses an incomplete expansion in the immediately preceding equation, which ought to be continued to second order and read $\xi = v - \phi/v - \phi^2/2v^3 + O(\phi^3)$. The expression for zero hole velocity given in Ref. 50 is similarly in error; substituting $Z_r'''(0) = 8$ into Eq. (12) gives the correct equation. These errors are not of great significance.
[55]B. D. Fried and S. D. Conte, *The Plasma Dispersion Function* (Academic Press, San Diego, 1961).
[56]V. A. Turikov, Phys. Scr. **30**, 73 (1984).
[57]R. E. Ergun, C. W. Carlson, L. Muschietti, I. Roth, and J. P. McFadden, Nonlinear Processes Geophys. **6**, 187 (1999).
[58]L. J. Chen, J. Pickett, P. Kintner, J. Franz, and D. Gurnett, J. Geophys. Res.: Space Phys. **110**, A09211, doi:10.1029/2005JA011087 (2005).
[59]The justification for forbidding infinite slope *trapped* distributions is substantially weakened by the observation (see Fig. 8) that the *passing* distribution of Schamel anyway *has* a slope singularity. The passing singularity was raised as an issue in Ref. 53.
[60]M. V. Goldman, D. L. Newman, and A. Mangeney, Phys. Rev. Lett. **99**, 145002 (2007).
[61]I. Roth, L. Muschietti, C. W. Carlson, F. S. Mozer, and R. E. Ergun, J. Geophys. Res.: Space Phys. **107**, 1239, doi:10.1029/2001JA900175 (2002).
[62]V. L. Krasovsky, H. Matsumoto, and Y. Omura, J. Geophys. Res.: Space Phys. **109**, A04217, doi:10.1029/2003JA010198 (2004).
[63]I. B. Bernstein and I. Rabinowitz, Phys. Fluids **2**, 112 (1959).
[64]A. Gurevich, Geomagn. Aeron. **3**, 185 (1963).
[65]Y. L. Al'pert, A. V. Gurevich, and L. P. Pitaevskii, *Space Physics with Artificial Satellites* (Consultants Bureau, New York, 1965).
[66]C. S. Ng and A. Bhattacharjee, Phys. Rev. Lett. **95**, 245004 (2005).
[67]C. S. Ng, A. Bhattacharjee, and F. Skiff, Phys. Plasmas **13**, 055903 (2006).
[68]For example, the distribution used by Ng and Bhattacharjee[66] to construct the spherically symmetric holes can be written $f(\mathbf{u}) \propto \exp(-u^2)\left[1 - d\,\exp\left(-u_\perp^2 r^2/l_0^2\right)\right]$. At distant radii, r, it has a Maxwellian form except for the depletion (by a fraction d) of particles whose velocity is directed toward the hole center within a narrow angle $\sim l_0/r_u$.
[69]D. Jovanović, P. K. Shukla, L. Stenflo, and F. Pegoraro, J. Geophys. Res.: Space Phys. **107**, SMP 15-1, doi:10.1029/2001JA900180 (2002).
[70]F. Mottez, S. Perraut, A. Roux, and P. Louarn, J. Geophys. Res. **102**, 11399, doi:10.1029/97JA00385 (1997).
[71]M. V. Goldman, M. M. Oppenheim, and D. L. Newman, Geophys. Res. Lett. **26**, 1821, doi:10.1029/1999GL900435 (1999).
[72]M. Oppenheim, D. L. Newman, and M. V. Goldman, Phys. Rev. Lett. **83**, 2344 (1999).
[73]L. Muschietti, I. Roth, C. W. Carlson, and R. E. Ergun, Phys. Rev. Lett. **85**, 94 (2000).
[74]M. M. Oppenheim, G. Vetoulis, D. L. Newman, and M. V. Goldman, Geophys. Res. Lett. **28**, 1891, doi:10.1029/2000GL012383 (2001).
[75]N. Singh, S. M. Loo, and B. E. Wells, J. Geophys. Res. **106**, 21183, doi:10.1029/2001JA900056 (2001).
[76]Q. M. Lu, B. Lembege, J. B. Tao, and S. Wang, J. Geophys. Res. **113**, A11219, doi:10.1029/2008JA013693 (2008).
[77]M. Wu, Q. Lu, C. Huang, and S. Wang, J. Geophys. Res.: Space Phys. **115**, A10245, doi:10.1029/2009JA015235 (2010).
[78]I. H. Hutchinson, Phys. Rev. Lett. **107**, 095001 (2011).
[79]I. H. Hutchinson, e-print, arXiv:1105.1356.
[80]D. Jovanović and H. Schamel, Phys. Plasmas **9**, 5079 (2002).
[81]O. Penrose, Phys. Fluids **3**, 258 (1960).
[82]D. L. Newman, M. V. Goldman, R. E. Ergun, and A. Mangeney, Phys. Rev. Lett. **87**, 255001 (2001).
[83]C. Norgren, M. André, D. B. Graham, Y. V. Khotyaintsev, and A. Vaivads, Geophys. Res. Lett. **42**, 7264, doi:10.1002/2015GL065390 (2015).
[84]Y. V. Khotyaintsev, A. Vaivads, M. André, M. Fujimoto, A. Retinò, and C. J. Owen, Phys. Rev. Lett. **105**, 165002 (2010).
[85]J. B. Tao, R. E. Ergun, D. L. Newman, J. S. Halekas, L. Andersson, V. Angelopoulos, J. W. Bonnell, J. P. McFadden, C. M. Cully, H.-U. Auster, K.-H. Glassmeier, D. E. Larson, W. Baumjohann, and M. V. Goldman, J. Geophys. Res. **117**, A03106, doi:10.1029/2011JA017364 (2012).
[86]G. Lapenta, S. Markidis, A. Divin, M. Goldman, and D. Newman, Phys. Plasmas **17**, 082106 (2010).
[87]M. V. Goldman, D. L. Newman, G. Lapenta, L. Andersson, J. T. Gosling, S. Eriksson, S. Markidis, J. P. Eastwood, and R. Ergun, Phys. Rev. Lett. **112**, 145002 (2014).
[88]I. H. Hutchinson, J. Geophys. Res.: Space Phys. **117**, A03101, doi:10.1029/2011JA017119 (2012).
[89]C. B. Haakonsen, I. H. Hutchinson, and C. Zhou, Phys. Plasmas **22**, 032311 (2015).
[90]I. H. Hutchinson, C. B. Haakonsen, and C. Zhou, Phys. Plasmas **22**, 032312 (2015).
[91]K. Saeki and J. J. Rasmussen, J. Phys. Soc. Jpn. **60**, 735 (1991).
[92]K. Saeki and H. Genma, Phys. Rev. Lett. **80**, 1224 (1998).
[93]B. Eliasson and P. K. Shukla, Phys. Rev. Lett. **93**, 045001 (2004).
[94]T. H. Dupree, Phys. Fluids **26**, 2460 (1983).
[95]I. H. Hutchinson and C. Zhou, Phys. Plasmas **23**, 082101 (2016).
[96]C. Zhou and I. H. Hutchinson, Phys. Plasmas **23**, 082102 (2016).