Reconnection onset in the tail of Earth’s magnetosphere

M. I. Sitnov,1,4 A. S. Sharma,1 P. N. Guzdar,2 and P. H. Yoon3

Received 29 October 2001; revised 7 February 2002; accepted 19 April 2002; published 27 September 2002.

[1] We present the nonlocal kinetic linear stability analysis of the self-consistent isotropic collisionless plasma equilibrium with strongly stretched magnetic field lines (the so-called modified Harris sheet) with respect to the tearing mode, which provides the onset of laminar reconnection in the system. The stability problem is solved using the finite element technique and the drift-kinetic description for the electron species with additional averaging over the bounce motion of the trapped electrons. The mode is found unstable for ion-to-electron temperature ratios typical for the tail current sheet of Earth’s magnetosphere when this sheet is sufficiently long, so that the electrons leaving it may be treated as transient particles. Comparison of the theory with earlier fluid modeling shows that the onset of reconnection is controlled by the Hall effect and a purely kinetic effect arising from different responses of the trapped and transient electrons. Geophysical implications such as the formation of the near-Earth neutral line and thin current sheets during substorms are discussed. INDEX TERMS: 2772 Magnetospheric Physics: Plasma waves and instabilities; 2744 Magnetospheric Physics: Magnetotail; 2788 Magnetospheric Physics: Storms and substorms; 3230 Mathematical Geophysics: Numerical solutions; KEYWORDS: reconnection onset, tearing instability, thin current sheet, substorm


1. Introduction

[2] Magnetic reconnection in plasmas is of primary importance for many laboratory and astrophysical phenomena [e.g., Biskamp, 2000]. While originally finite resistivity was considered as the key to reconnection, in high-temperature plasmas such as in the solar corona or Earth’s magnetotail the resistivity is negligibly small, and the conventional resistive MHD approach becomes inadequate. Recent studies have revealed a number of distinctive features of the collisionless reconnection process, specifically, the importance of the Hall effect and whistler turbulence in the high-beta case [Biskamp, 2000; Birn et al., 2001, and references therein]. These studies are limited however to the magnetic field configurations with the X-line pattern being already formed. Thus, they leave open the problem of the reconnection onset, which requires the stability study of configurations where no magnetic field line is reconnected yet. These configurations, potential subject of spontaneous reconnection, can be divided into two basic classes [e.g., Galeev, 1984], namely, the sheared magnetic field, which is typical in particular for fusion devices [e.g., Biskamp, 2000] and Earth’s magnetopause [e.g., Russell, 1995], and the stretched magnetic field, which is a distinctive feature of Earth’s magnetotail [e.g., Fairfield and Ness, 1970]. While the onset of reconnection in the sheared magnetic field is understood relatively well [e.g., Galeev, 1984, and references therein], its mechanism in tail-like magnetic field configurations has been a big puzzle for more than a quarter of century preventing proper modeling of many interesting space phenomena. This strongly hinders in particular reliable modeling and forecasting of magnetospheric substorms.

[3] The instability responsible for the onset of reconnection in tail-like magnetic field configurations, namely the tearing instability, was first considered for the collisionless plasma case relevant to Earth’s magnetotail by Coppi et al. [1966] in the simplest geometry with antiparallel magnetic field lines [Harris, 1962]. The instability was described as a growth of the negative energy wave (due to the mutual attraction of parallel current filaments) with the dissipation provided by the Landau resonance of the nonmagnetized electrons near the neutral plane, where the magnetic field has a null. Later it was revealed that the tail magnetic field lines prior to the substorm onset are not antiparallel but rather sharply curved due to the finite component $B_n$ of the magnetic field normal to the sheet plane [Fairfield and Ness, 1970; Nishida and Nagayama, 1973; Russell and McPherron, 1973]. This component was found to be large enough to magnetize electrons and prevent their Landau dissipation [Schindler, 1974]. To resolve the problem Schindler [1974] proposed a similar dissipation provided by the ion species, with electrons treated merely as a charge neutralizing background taking into account the large ion-
to-electron temperature ratio $T_i/T_e$ typical for the magnetotail. It was found however that the magnetized electrons can also change the sign of the mode energy from negative to positive, thus making any dissipation ineffective for the instability excitation [Galeev and Zelenyi, 1976]. Then, assuming all the electrons were magnetized and trapped inside the current sheet, Lembge and Pellat [1982] obtained the sufficient stability criterion for the tearing mode in an isotropic current sheet to be $kLB_0/B_0 > 4/\pi$, where $k$ is the wave number, $L$ is the thickness of the sheet, and $B_0$ is the magnetic field outside it. In contrast to the conjecture of Schindler [1974], this criterion does not depend on the temperature ratio. It is also very restrictive as it coincides with the WKB approximation used in the stability analysis. The matter is that the magnetic field line tension in the isotropic current sheet equilibria with stretched magnetic field lines must be balanced by the pressure gradient along the direction of stretching, and the appropriate inhomogeneity scale $L_x \sim LB_0/B_0$ must exceed the WKB tearing wavelength.

[4] The stabilization effect arises due to the drift motion of electrons in the crossed fields $B_n$ and $E_{1y}$, the dawndusk component of the tearing electric field (we use here standard system of reference with the $X$-axis directed antiparallel to the direction of the magnetic field line stretching, $Z$-axis parallel to the magnetic field component $B_n$, and the unit vector in $Y$-direction $e_y = e_x \times e_z$). Drifting electrons drag ions along the $X$-axis, and the energy expended may exceed the free energy available for the tearing mode due to mutual attraction of parallel current filaments. The physical effect looks quite universal and difficult to circumvent. In fact, a number of attempts to destabilize the tearing mode either by the external diffusion [Coroniti, 1980; Esarey and Molvig, 1987] or by dynamical chaos in electron orbits [Büchner and Zelenyi, 1987] have failed [Pellat et al., 1991; Brittnacher et al., 1994, 1998; Quest et al., 1996]. The specific arguments are based on the energy principle where the stabilizing term is reduced to the expression $\int dxdy|\langle n_i \rangle|^2/\langle n_0 \rangle$, containing the equilibrium and perturbed electron densities $\langle n_0 \rangle$ and $\langle n_i \rangle$, integrated over the flux tube $\langle n_i \rangle = \int ds nj(s)/B(s)$. Then the denominator is estimated as $\langle n_0 \rangle \approx 2Ln_0/B_n$ by limiting the integration to the region of trapped motion, while the numerator is obtained using the drift-kinetic theory and additional bounce averaging over the trapped electron orbits.

[5] All authors, advocating the tearing stability in the case $B_n \neq 0$ [Lembege and Pellat, 1982; Galeev, 1984; Pellat et al., 1991; Wang and Bhattacharjee, 1993; Brittnacher et al., 1994, 1998; Quest et al., 1996], considered electrons as a single fluid. Similar fluid approach was successfully used later to describe the fast reconnection close to an existing X-line [Biskamp, 2000; Birn et al., 2001 and references therein]. It turns out however that the fluid approach completely forbids the formation of the X-line itself. Meanwhile the fluid description of the electrons is not quite correct, when the self-consistent models of the current sheet used to describe the initial state by [Schindler, 1974] and [Lembege and Pellat, 1982] are considered. These isotropic models, with the plasma density constant along the field line, require the existence of a population of transient electrons. Transient electrons spend a small time in the current sheet, and their response to the tearing perturbation is adiabatic [Sitnov et al., 1998]. As a result they can shield the electrostatic potential created by differences in the motions of ions and trapped electrons.

[6] Possible destabilizing effects involving transient electrons were considered for the first time by Goldstein and Schindler [1978] in terms of short-circuiting the electrostatic field of the tearing mode by the conducting ionosphere. Many details of this effect were elaborated later by Swift [1986]. Goldstein and Schindler [1978] and Swift [1986] ignored any direct shielding effect of transient electrons and considered only the effect of the electrodynamic coupling between the tail current sheet and the ionosphere associated with the generation of field-aligned currents. However, this effect cannot strongly influence the tearing instability as the appropriate time constants of magnetosphere-ionosphere interaction exceed the required tearing growth time. Irby et al. [1979] suggested the reduction of the tearing stability threshold by taking into account the reduction of the untrapped electron compression by sound wave propagation along the field lines. The latter effect is valid however only for collisional plasmas. The destabilizing role of transient electrons was studied also by Wang and Bhattacharjee [1993] assuming all the electrons to be effectively transient because of the large bounce period exceeding the tearing growth time.

[7] The first systematic investigation of the role of trapped and transient electron populations in the tearing stability analysis was performed by Lakhina [1992]. He reported the destabilizing effect of transient electrons, which increases with the increase of their relative amount. However, this destabilization was not connected with any shielding effect since the Poisson’s equation was not taken into account, and arose instead from the increased contribution of the transient electrons to the nonadiabatic component of the perturbed current. A similar effect was reported by Lakhina [1993] for the more complex current sheet structure with the thin current sheet embedded within the thicker plasma sheet, and with Poisson’s equation being taken into account. However, both analyses [Lakhina, 1992, 1993] had a significant deficiency, namely, the use of the global parameter $\eta(1 - \eta)$ to assess the relative contribution of trapped (transient) particles to the non-total density of electrons. This made the description inconsistent because such a parameter should be essentially local depending on the distance from the neutral plane with $\eta \approx 0$ outside the sheet and $\eta \approx 1$ near the neutral plane $z = 0$.

[8] The modification of the original approach by Lembege and Pellat [1982] using the self-consistent description of trapped and transient electrons has been performed by Sitnov et al. [1998, 1999]. It has been shown in particular that the sufficient stability criterion taking into account the transient electron population has the form: $(3T_i/T_e)^{1/2}kLB_0/B_n > 4/\pi$. It has two distinctive features. First, it shows that the stabilizing role of electrons is progressively reduced with the decrease of their temperature as was first conjectured by Schindler [1974]. Second, it allows for instability gaps within the WKB approximation. In other words, it permits the onset of reconnection. The destabilization effect was found also to be independent of any free parameter, which would control the relative number of transient electrons. The important point is that in isotropic self-consistent current sheet models considered by Schindler [1974] and Lembege and Pellat [1982] the number density of transient
electrons is not a free parameter. Its local value depends on the basic sheet parameters such as $B_n/B_\perp$ and on the distance $z$ from the neutral plane.

Attempts to resolve the tearing stability problem in configurations with $B_n \neq 0$ using particle simulations [Swift and Allen, 1987; Zwingmann et al., 1990; Pritchett, 1994; Pritchett and Coroniti, 1994, 1995; Dreher et al., 1996; Pritchett et al., 1997] are beset with problems such as the spurious temperature anisotropy at $T/T_e > 1$ [Pritchett, 1994], unrealistic mass ratios ($m_i/m_e \leq 64$), and inconclusive simulation results. While unstable tearing modes were found by Zwingmann et al. [1990], complete stability of these modes was reported by Pritchett [1994] and Dreher et al. [1996]. Besides, the simulation boxes reported by Zwingmann et al. [1990], Pritchett [1994], and Dreher et al. [1996] were too short to model transient particles, and the use of reflecting or reintroducing particle boundary conditions artifically made all electrons to be trapped. In this context, of particular interest are simulations [Pritchett and Coroniti, 1982, 1994, 1995; Pritchett et al., 1997], where the onset of reconnection was detected when the loss cone was explicitly introduced.

Thus, the reconnection onset problem remains open. The result of Sitnov et al. [1998], while relaxing the criterion derived by Lembege and Pellat [1982], is still based on a sufficient stability condition and a simplified energy principle, and does not reveal the actual stability threshold. In this paper we resolve this issue and obtain the explicit solution of the tearing eigenvalue problem in this context. In contrast, the detailed drift-kinetic analysis by Sitnov et al. [1998] [see also Coroniti, 1980, equations (23) and (25)], the perturbed distribution has the form

$$\tilde{f}_e = T_e^{-1} f_0 (\phi - (\phi_1)) + T_e^{-1} f_0 \left( \frac{e}{c} v \cdot B_1 + \mu B_0 \right), \quad (5)$$

where $\langle \ldots \rangle$ denotes the averaging over the bounce period. This result is a fully sufficient description of the trapped electrons for the specific form of the energy principle used by Lembege and Pellat [1982], which is based on the assessment of the perturbed electron density $\tilde{n}_1$ integrated over the flux tube. It is not sufficient however for solving the explicit eigenmode problem for the tearing instability. In fact, (5) describes only the slowly varying part of the perturbed distribution in the drift wave approximation. According to the detailed drift-kinetic analysis by Lembege and Pellat [1982] [see also Coroniti, 1980, equations (23) and (25)], the perturbed distribution has the form

$$\tilde{f}_1 = T_e^{-1} f_0 \phi_1 + g_0 + g_0^{(1)}, \quad (6)$$

where

$$g_0 \approx -T_e^{-1} f_0 \phi_1 + T_e^{-1} f_0 \left( \frac{e}{c} v \cdot B_1 + \mu B_0 \right) \quad (7)$$

is actually the slowly varying part, which depends on the perturbed field through the integrals of the adiabatic motion of trapped electrons, while

$$g_0^{(1)} \approx -\omega_e^{-1} (v \cdot n) (\nabla g_0 - \gamma f_0 A_1 / e T_e), \quad (8)$$

where $\omega_e = eB_0 m_e c$ is the electron cyclotron frequency. The latter correction to the perturbed electron distribution contains the rapidly varying velocity factor. As a result it gives small contribution to the perturbed electron density, which was used in the analysis of Lembege and Pellat [1982]. It may give however a considerable contribution to the Ampere’s equation (1) because there the perturbed distribution $f_1$ is convolved with the additional electron velocity factor $v$. This correction term is important also in other forms of the energy principle analysis such as the ones used by Coroniti [1980] and Kaznetsova and Zelenyi...
In particular, its omission in the term $\Delta W_e$ in the energy balance equation used by Kuznetsova and Zelenyi [1991] led the authors to an underestimate of the stabilizing effect of trapped electrons and the incorrect conclusion about the possible destabilization of the tearing mode by the pitch angle diffusion. (Although the latter conclusion may be true, it requires more detailed kinetic description of electrons including their transient population [Sitnov and Lui, 1999].)

The main advance in our present kinetic description, compared to essentially fluid approaches by Lembege and Pellat [1982], Coroniti [1980], Galeev [1984], Pellat et al. [1991], Brittnacher et al. [1994, 1998], and Quest et al. [1996], is the explicit inclusion of the response of the transient electrons with the magnetic moment $\mu < \mu_{\text{min}} = W/B_0$, where $W$ is the electron kinetic energy, namely, the condition [Sitnov et al., 1998]

$$g^{(\text{trans})}(\mu < \mu_{\text{min}}) = 0 \quad (9)$$

It is valid as long as the time necessary for electrons leaving the current sheet to bounce from the ionosphere is large compared to the time they spend in the current sheet interacting with the tearing perturbation. Condition (9) may be substantiated using the approach similar to that used for trapped electrons. Let us assume that the transient electrons are reflected far outside the current sheet and their bounce period $\tau_{b_{\text{(trans)}}}$ considerably exceeds that of the electrons trapped inside the current sheet $\tau_{b_{\text{(trap)}}}$ but is still less than the tearing growth time, i.e., $\tau_{b_{\text{(trans)}}} \ll \tau_{b_{\text{(trap)}}} \ll \gamma^{-1}$. The latter assumption is appropriate in particular for the thermal electrons in the midtail with $v_T \sim 100 R_E/\text{min}$ reflected from the ionosphere. Then these quasi-trapped electrons can be described by the same equations (6)–(8) as the trapped electrons. However, since they spend most of their bounce period outside the current sheet all bounce averaged terms in (6) should contain additional small factor $\delta_s = \tau_{b_{\text{(trans)}}}/\tau_{b_{\text{(trap)}}}$ as compared to similar terms for the trapped electrons

$$\left(\Psi_{\text{trans}}^{(\text{trans})}(\Psi_{\text{trap}}) = \delta_s \int_{S_{\text{trans}}} S_{\text{trap}}(s) v_{\|}^{-1} ds = \int_{S_{\text{trap}}} \Psi(s) v_{\|}^{-1} ds, \quad (10)$$

where $S_{\text{trans}}$ and $S_{\text{trap}}$ are the lengths of the bounce orbits of transient and trapped electrons, and $v_{\|}$ is the component of the particle velocity along the field line. Since the tearing eigenmodes are localized around the current sheet (for instance, for thin current sheet $\Psi(z) \propto \text{cosh}(z/L)^{-1/2}$) and the width of actual interaction region is even smaller than that of the current sheet [e.g., Coroniti, 1980, pp. 6726–6727], the first integral in (10) is smaller than or of the order of the second one for the same kinetic energy. Hence $g^{(\text{trans})} = O(\delta_s)$, and can be neglected in the subsequent analysis. If the reflection of transient electrons from the ionosphere is not ideal due to collisions, approximation (9) should be still valid as a consequence of the Boltzmann distribution of the thermalized transient particles in the electromagnetic field disturbed by the tearing mode. Approximation (9) is also consistent with the earlier estimate of the perturbed distribution of transient electrons made by Lakhina [1992]: $g^{(\text{trans})} \propto (1 - \eta) A_1(s)$, where $\eta$ is the fraction of trapped particles (so that their distribution $f^{(\text{trans})}_{\text{trans}} = \eta n_0(z) v_{\|}^{-3} \exp\left(-W/T_e\right)$. When used as a local estimate it shows that $g^{(\text{trans})}$ is small inside the interaction region because the local parameter $(1 - \eta)$ is small there. Note however that the electrostatic contribution to $g^{(\text{trans})}$ was completely ignored by Lakhina [1992].

### 2.3. Ion Dynamics

We assume that ions are not magnetized by the field $B_\mu$. Following Coroniti [1980], we neglect any nonadiabatic ion contribution to the perturbed number density. The corresponding contribution to the current density cannot be neglected as it provides the main dissipation mechanism for the tearing mode. However, we shall additionally simplify the ion current contribution using the thin current sheet model [Pritchett et al., 1991], which is valid for current sheets with the thickness $L$ comparable to the thermal ion gyroradius $\rho_0$ in the field $B_0$. Then the resonant current density has the same profile as the equilibrium density and simply modifies the coefficient in the second term of Ampere’s equation (1): $2/L^2 \rightarrow 2b/L^2$ with

$$b = 1 - \gamma/\gamma_0 k L \quad (11)$$

where

$$\gamma_0 = \left(2v_T^2 e^2/\sqrt{\text{m}_e} \omega_0 L^3\right)/(1 + 2\rho_0^2/L^2) \approx (2/3\sqrt{\pi}) \omega_0 (\rho_0/L)^3 \quad (12)$$

This approximation gives a simple analytical formula for the tearing growth rate when the stabilization effect by Lembege and Pellat [1982] is neglected [Pritchett et al., 1991]

$$\gamma = \gamma_0 k L (1 + kL/2)(1 - kL) \quad (13)$$

According to this expression, thin current sheets with $L \sim \rho_0$ should be particularly unstable. They are considered to be a distinctive feature of the magnetotail current sheet prior to and during substorms [Sergeev et al., 1998, and references therein]. Note however that the threshold stability condition, which is the main goal of our research, does not depend upon the specific form of the ion dissipation.

### 2.4. Final Set of Equations

The basic equations (1) and (2) may be now rewritten as follows

$$\nabla^2 A_1 + \frac{2e A_{1y}}{L^2 \text{cosh}^2(z/L)} = \frac{4\pi}{c} \left(j_{\text{trans}}^{(\text{trans})} + j_{\text{trans}}^{(\text{ad})} + j_{\text{ad}}^{(\text{ad})}\right) \quad (14)$$

$$\rho_{1e} + \rho_{1e}^{(\text{trans})} + \rho_{1e}^{(\text{ad})} = 0 \quad (15)$$

where

$$j_{\text{trans}}^{(\text{trans})} = -e v_T n f_{\text{trans}} \phi d^3 \nabla \cdot \left(\omega_{\text{Ly}}^{-1} \phi (n \times \nabla) g_0 d^3 v \right) \quad (16)$$

$$-e \int_{\Omega_{\text{trans}}} \omega_{\text{Ly}}^{-1} \phi (n \times \nabla) g_0 d^3 v \quad (16)$$
Similarly (15) can be rewritten in the form

\[ j_{1i}^{(\text{ad})} = -\frac{e^2}{T_i} \int_{\Omega_{\text{trap}}} f_{0e} \phi_1 d^3v \]  

\[ j_{1i}^{(\text{ad})} = -\frac{e^2}{T_i} \int_{\Omega_{\text{trap}}} f_{0e} \phi_1 n_0 d^3v \]  

\[ \rho_{1e}^{(\text{trans})} = -\frac{e}{T_e} \int_{\Omega_{\text{trap}}} f_{0e} \phi_1 d^3v + e \int_{\Omega_{\text{trap}}} g_{1e} d^3v \]  

\[ \rho_{1e}^{(\text{trans})} = -\frac{e^2}{T_e} \int_{\Omega_{\text{trap}}} f_{0e} \phi_1 d^3v \]  

\[ \rho_{1i}^{(\text{ad})} = -\frac{e^2}{T_i} \int_{\Omega_{\text{trap}}} f_{0e} \phi_1 d^3v \]  

Note that the Maxwell equations imply additional effective averaging over the Larmor period due to the integration over the velocity space. In Poisson’s equation this has resulted in the disappearance of the term from \( g_0 \). In Ampere’s equation the expression for \( g_0 \) can be simplified by omitting the term \(-ye f_0 A_{1/e} / T_e\) because its scalar product with the factor \( (\mathbf{v} \times \mathbf{n}) \) only involves the \( v_x \) and \( v_z \) components, which gives no contribution after integration over the velocity space. Using similar arguments we replaced \( (\mathbf{v} \times \mathbf{n}) \nabla \) by \( v_y (\mathbf{n} \cdot \nabla) \), in (16). The Poisson’s equation is further simplified considering that the relevant scales exceed the Debye length and \( \nabla^2 \varphi_1 \approx 0 \). The equations (14)–(21) represent the complete set of the tearing mode equations to address the stability problem.

[17] Using the relation between the ion and electron drift speeds \( u_i/u_e = -T_i/T_e \) and the quasineutrality condition one can further simplify the right hand side of (14) to get

\[ \frac{\langle \psi \rangle_j (x_1, z_1)}{v_{Te}} - \frac{\langle \psi \rangle_{\phi_1} (x_1, z_1)}{v_{Te}} + \frac{\langle \psi \rangle_{\phi_1} (x_1, z_1)}{v_{Te}} = \frac{e^2 v_{Te}}{T_e} n_{0e} \]

\[ \frac{\langle \psi \rangle_j (x_1, z_1)}{v_{Te}} = \frac{e^2 v_{Te}}{T_e} n_{0e} \]

\[ \frac{\langle \psi \rangle_{\phi_1} (x_1, z_1)}{v_{Te}} = \frac{e^2 v_{Te}}{T_e} n_{0e} \]

\[ \frac{\langle \psi \rangle_{\phi_1} (x_1, z_1)}{v_{Te}} = \frac{e^2 v_{Te}}{T_e} n_{0e} \]

\[ \frac{\langle \psi \rangle_{\phi_1} (x_1, z_1)}{v_{Te}} = \frac{e^2 v_{Te}}{T_e} n_{0e} \]

\[ \frac{\langle \psi \rangle_{\phi_1} (x_1, z_1)}{v_{Te}} = \frac{e^2 v_{Te}}{T_e} n_{0e} \]

where the average have been taken over the volume \( \Omega_{\text{trap}} \) occupied by trapped electrons as

\[ \langle \psi \rangle_j (x_1, z_1) = \int_{\Omega_{\text{trap}}} (v_j / v_{Te}) \psi f_{0e} / n_{0e} d^3v \]

\[ \langle \psi \rangle_{\phi_1} (x_1, z_1) = \int_{\Omega_{\text{trap}}} (v_j / v_{Te}) \phi_1 f_{0e} (\mathbf{n} \cdot \nabla) \psi f_{0e} / n_{0e} d^3v \]

Similarly (15) can be rewritten in the form

\[ \sigma \langle (\phi_1) \rangle_{\mu} - \phi_1 = \sigma \langle (\psi_j) \rangle_{\mu} \]

with \( \sigma = T_i / T_e \) and

\[ \langle \psi \rangle_j (x_1, z_1) = \int_{\Omega_{\text{trap}}} \psi (f_{0e} / n_{0e}) d^3v \]

\[ \langle \psi \rangle_{\phi_1} (x_1, z_1) = \int_{\Omega_{\text{trap}}} \psi (f_{0e} / n_{0e}) d^3v \]

The explicit expressions for terms \( \langle (\phi_1) \rangle_{\mu} \), \( \langle (\psi_j) A_1 + (\psi_{\phi_1} n) B_1 \rangle_{\mu} \), \( \langle (\psi_j) A_1 + (\psi_{\phi_1} n) B_1 \rangle_{\mu} \), and \( \langle (\psi_j c) A_1 + (\psi_{\phi_1} c) B_1 \rangle_{\mu} \) which appear in (22) and (25), are given in appendices B and C. Using these expressions and introducing the z-profiles of the perturbed fields, which are parametrized by the local equilibrium magnetic field \( b = B_{0z} \), as \( A_1 = A(z) \exp(i k x_0 + i \theta(b)) \) and \( \phi_1 = \phi (z) \exp(i k x_0 + i \theta(b)) \) with \( \theta(b) = b_k L \cos(z(b)/L) \), \( b_k = B_{0k} / B_{0z} \), \( \phi = \phi \phi_1 \), and \( \phi_1 = (B_0 / B_{0z}) u_i / c \), we obtain the following set of equations

\[ \hat{L}_0 (A) + L_{11} (A) = 0 \]

\[ L_{21} (A) = 0 \]

\[ L_{22} (A) = 0 \]

\[ L_{12} (A) = 0 \]

\[ L_{12} (\phi) = \sigma \langle (\phi_1) \rangle_{\mu} - \phi \]

\[ \Xi = \frac{i k L}{b} A_1 (1 + p) - \frac{L}{b} \frac{\partial A_1}{\partial z} \left( \sqrt{b^2 - 1} - 1 \right) \]

\[ T(b, b_{m}) = \int_1^{b_{m}} \frac{b^{1/2} \sqrt{b^2 - 1} - b b^2 db}{b^2 - 1} \]

\[ P = 2 (b_{m} - b) (b^2 - 1) / b, \chi = (2 - i k L b_{m}) L^2 \cos^2(z/L), \]

\[ b = \sqrt{1 + b^2} \]

3. Finite Element Analysis

[18] To solve the complex eigenvalue problem for the set of equations (27)–(37) one can use the finite element
approach [Sharma, 1983; Chen and Lee, 1985; Brittnacher et al., 1995, 1998; Daughton, 1998, 1999]. According to this approach, the potentials are expanded into a complete series of basis functions $\Psi_n$

$$A(z) = \sum_{n=1}^{N} C_n \Psi_n(z), \quad \phi(z) = \sum_{n=1}^{N} C_{N+n} \Psi_n(z),$$  

(38)

which is convenient for the given problem, and in particular obeys the appropriate boundary conditions. For the slab geometry of the modified Harris sheet the Hermite functions are the appropriate basis functions for the representation of the potentials

$$\Psi_n(z) = \left(n!2^{n+1/2}\right)H_n(z)\exp\left(-z^2/2\right)$$  

(39)

Then the original set of integro-differential equations can be reduced to a set of algebraic equations

$$M_{ij}C_j = 0$$  

(40)

with the matrix $M_{ij}$ being a generalized inner product between $N$ basis functions and the two equations (27) and (28)

$$\int_{-\infty}^{\infty} \Psi_i(z) \tilde{L}(A,\phi)(z)dz$$  

(41)

[19] Earlier the finite element approach was used in tearing stability studies for the case $B_n = 0$ and Harris [1962] equilibrium [Daughton, 1999], single species models [Brittnacher et al., 1995], or the energy principle analysis limited by trapped electrons [Brittnacher et al., 1998]. One should mention also a number of closely related nonlocal tearing stability studies based on the fluid plasma models [Terasawa, 1983; Harrold et al., 1995].

[20] The results of the numerical solution of the system (27)–(37) are shown in Figures 1–4. The upper curve in Figure 1 represents the growth rate of the tearing mode in a simplified case with $B_n = 0$, corresponding to the solution (13) of the equation $L_0(A) = 0$ [Pritchett et al., 1991]. This is in good agreement with the result of Pritchett et al. [1991] and provides a useful benchmark for our code. The thick solid curve represents the growth rate of the tearing mode with $B_n/B_0 = 0.1$, and the temperature ratio $\sigma = 7$, typical for the tail of the Earth’s magnetosphere [Baumjohann et al., 1989]. This curve shows in particular that there is a considerable region of unstable WKB tearing modes. Lower curves represent a way to model the results of Lembege and Pellat [1982]. Earlier, using the energy principle analysis, Sitnov et al. [1999] showed that one can recover these results in the limit of the infinitely long flux tube $b_w \to \infty$ (see in particular their Figure 4). The latter limiting condition cannot be imposed in the present case as it would reduce the whole eigenvalue problem to the case $B_n = 0$. We used instead the limit $\sigma \to 0$ in the Poisson’s equation (28) as it allows to mimic the fluid relationship between the flux tube averaged perturbed electron density and the electromagnetic potential (see equation (30), Coroniti [1980], and discussions below). Figure 1 shows that in this limit the maximum possible parameter $k_{ms}L$ of the unstable solutions, corresponding to the marginally stable wave number $k_{ms}$ actually approaches the WKB boundary. Based on the previous analytical assessment one could expect that $k_{ms}L$ should be inside the non-WKB region as the stability condition of Lembege and Pellat [1982] is sufficient. The analysis of this discrepancy shows however that the above explicit kinetic theory contains an additional effect, compared to the simplified single-species electron description by Lembege and Pellat [1982] and Coroniti [1980], which arises because of the finite length of the flux tube in a self-consistent equilibrium with $B_n \neq 0$ and $B|z| \to \infty = Const = B_0$ for the current sheet embedded into a lobe magnetic field. In particular, the expression for the drift velocity contains the additional factor $(1 - b_w^{-2}(b^2 - 1))$. Such factors occur in several integrals in the stabilizing term in the energy principle and increase the maximum wave number of unstable waves $k_{ms}$.

[21] Finding an instability gap within the WKB region seems to be at variance with the results of a similar nonlocal kinetic stability analysis performed recently by Brittnacher et al. [1998]. In that work a finite element approach was combined with the energy principle analysis, and no instability was found for realistic tail plasma parameters. A closer examination shows however that our results are consistent with those of Brittnacher et al. [1998]. In their analysis only trapped particles were considered due to the reflecting boundary conditions, and they argued that the size of their simulation region was large enough to neglect “lost” (transient in our terminology) particles. Below we show, however, that the size of the simulation region in the model of Brittnacher et al. [1998] was too short to properly describe the transient population of electrons, which crucially affects the tearing stability as indicated by our present analysis. To prove the latter statement it is sufficient to assess the maximum excursion $dx$ of trapped particles with $|z| < L$ along the X-direction.
Using the parabolic approximation of the Harris magnetic field we obtain $dx \sim (1/2)(B_0/B_n)dz^2/L \sim (1/2k)(B_0/B_n)kL$. This value must exceed the tearing wavelength within the whole WKB region $kLB_0/B_n \gg 1$. Moreover, in contrast to trapped electrons, the use of reflecting boundary conditions for transient electrons requires the simulation region to be even longer than $dx$, because transient electrons are supposed to spend most of their bounce period outside the current sheet region $|z| < L$. However, the length of the simulation box in the analysis of Brittnacher et al. [1998] was limited to much shorter values and in particular to just two wavelengths for the case $B_n/B_0 = 0.1$. In fact, the only region where the comparison between our results and those of Brittnacher et al. [1998] might be relevant is marked by the question sign in Figure 1 of their work. This is however a non-WKB region. Besides, the above estimate for $dx$ together with the well-known scaling relation of modified Harris equilibria $L/L_x \sim B_0/B_n$ shows that the detailed description of transient particles requires taking into account the current sheet inhomogeneity along the tail axis. In our approach we circumvented the latter problem by using a reduced description of transient electrons, namely, condition (9). Note that the above arguments are particularly relevant to the case of thin current sheets with $L/\delta \lesssim 0.1R_E$.

A priori it is not clear why the population of transient electrons is so important to the stability problem, and yet their relative amount does not appear explicitly in the...
tearing growth rate. In particular, the relative number of transient and trapped particles \( n_{\text{trans}}/n_{\text{trap}} \) at the neutral plane is proportional to \( B_n/B_0 \ll 1 \), suggesting a minor role of transient electrons. However, \( n_{\text{trans}}/n_{\text{trap}} \) is neither free nor a global parameter of the system. It changes from a small value \( \sim B_n/B_0 \) at \( z = 0 \) to infinity at \( |z| > L \), where all particles are transient. These local values are determined by the structure of the given self-consistent equilibrium, which has no independent parameter to control \( n_{\text{trans}}/n_{\text{trap}} \). Moreover, as shown in appendix D, in the case \( B_n \ll B_0 \) the corresponding global, parameter representing the relative amount of transient and trapped particles in the modified Harris sheet within the region \( |z| < L \) (it is often termed as a trapped region), is approximately constant independent of \( B_n/B_0 \). The latter result clearly shows that the transient population is actually an important ingredient of isotropic current sheet equilibria, whose role was strongly underestimated earlier.

[23] The tearing eigenfunctions, with electromagnetic and electrostatic components, for the sheet with \( B_n/B_0 = 0.1 \), and the temperature ratio \( \sigma = 7 \) are shown in Figure 2. This figure shows in particular that the electrostatic component of the tearing mode increases close to the marginal stability boundary (bottom panel). When compared to similar results of the fluid nonlocal stability analysis [Harrold et al., 1995], the kinetic tearing eigenmodes resemble the resistive modes shown in Figure 4 of that paper. The effective resistivity is provided in our case by the ion Landau dissipation, and the instability is of the ion-tearing type [Schindler, 1974]. Comparison of the electromagnetic potential eigenfunctions with the corresponding analytical solution for the case \( B_n/B_0 = 0 \) [Pritchett et al., 1991] reveals also the increase of deviations from that solution near the marginal stability boundary. Note that within the framework of the simplified model of the ion dynamics used in our work and in the analysis by Pritchett et al. [1991] the electrostatic potential for the case \( B_n/B_0 = 0 \) is identically zero.

[24] Figure 3 shows the actual dependence of the tearing growth rate on the temperature ratio. It reveals the mode destabilization when electrons temperature is smaller compared to that of ions. The location of stable and unstable regions in the space of parameters \( T_e/T_i \) and \( B_n/B_0 \) and its comparison with earlier theoretical estimates is shown in Figure 4. It shows that for high enough, yet realistic values of the temperature ratio, the wave numbers \( k_{\text{ms}} \) corresponding to the marginal stability boundary (they are also maximum possible wave numbers within the unstable region) are far outside the non-WKB region determined by the estimate of Lembege and Pellat [1982], \( k_{\text{WKB}} = (4/\pi) B_n/B_0 \). Moreover, the relative size of the WKB region of unstable tearing modes compared to the non-WKB region increases with the decrease of the ratio \( B_n/B_0 \). The inset shows the comparison of the actual marginal stability boundary and its previous estimates [Lembege and Pellat, 1982] and [Sitnov et al., 1998] based on the energy principle analysis. One can see that this boundary is at variance with the fluid estimates by Lembege and Pellat [1982]. It is however quite consistent with the more recent kinetic result [Sitnov et al., 1998] obtained for \( T_e/T_i \gg 1 \), although the specific dependence of \( k_{\text{ms}} \) on \( T_e/T_i \) turns out to be slower. Thus, the onset of reconnection in collisionless plasmas is possible without any additional effects like the dynamical chaos or diffusion induced by external turbulent sources. However, its proper description requires the consistent kinetic description of electrons taking into account both trapped and transient populations.

[25] This conclusion, which is the main result of our stability analysis, gives a new impetus to studying other aspects of the reconnection onset problem and in particular the more detailed investigation of the ion dissipation. In fact, the assumption that ions are unmagnetized by the field \( B_n \), requires the instability growth rate \( \gamma \) to exceed the
corresponding gyrofrequency \( \omega_{ci} = eB/m_e c \) [Schindler, 1974; Galeev and Zelenyi, 1976]. This condition is satisfied in our model assuming that the sheet is sufficiently thin \( L \leq \rho_0 \). Then, according to (12) and the numerical result shown in Figure 1, the maximum possible growth rate \( \gamma_{\text{max}} \sim 0.1\omega_{ci}(\rho_0/L)^3 \sim \omega_{ci}(\rho_0/L)^3 \). In our present study we neglected however the influence of \( B_\| \) component on the ion dissipation and even further simplified its description using the thin current sheet approximation [Pritchett et al., 1991]. These features of the ion response were studied earlier using both the finite element approach [Brittnacher et al., 1995] and particle simulations [Ambrosiano et al., 1983; Pritchett et al., 1991; Hesse and Winske, 1993]. In particular, as shown by Pritchett et al. [1991], the normal component \( B_\perp \geq 0.06 B_0 \) stabilizes the tearing mode for the flaring tail model. However, taking into account the effects of the pitch angle diffusion in the ion dynamics increases the region of unstable tearing modes up to \( B_\perp = 0.1 B_0 \) [Brittnacher et al., 1995]. It should be noted also that, in contrast to the soft tearing mode excitation due to the transition through the marginal stability boundary shown in Figure 4 and associated with the reduction of the electron compressibility effect, the transition to instability connected with the ion demagnetization represents a hard excitation process as it starts from the finite growth rate right above the threshold. As shown by Brittnacher et al. [1995], this may provide quite realistic growth time of the tearing mode (~5 min) consistent with typical substorm onset timescales.

4. Comparison to MHD Modeling

[25] To better reveal the physics of the reconnection onset we compare our basic equations (27) and (28) with their fluid counterpart, which has been most explicitly described by Galeev [1984]. Consider the third operator in the Ampere’s equation (27). In the limiting case of infinitely long flux tube it may be written in the form

\[
\mathcal{L}_{12}(\phi) = \frac{2 - i kLb_s \omega_p^2}{c\omega_p L \cosh^2(z/L)} \langle \phi_1 \rangle_{\mu}
\]  

(42)

where

\[
\langle f(b_m) \rangle_{\mu} = \frac{3}{4} b_s^2 \int_0^{\infty} f(b_m) b_m^2 \sqrt{b_m - b_s} \, db_m
\]  

(43)

represents averaging over the dimensionless magnetic moment. Numerical analysis of (27)–(35) shows that the second term in brackets (42) dominates the tearing eigenvalue problem. Keeping in (42) only this dominating term and using the relation \( \phi_1 = -(T/e\eta_0) n_1 \) for the perturbed ion distribution [see also Coroniti, 1980, equation (31)] we obtain the leading nonadiabatic current contribution to Ampere’s equation in the form

\[
\delta J = \frac{c}{4\pi} \mathcal{L}_{12}(\phi) = \frac{ikcT}{B_s} \langle \langle n_1 \rangle \rangle_{\mu}
\]  

(44)

This corresponds to the equation (90) of Galeev [1984] and represents the bounce-averaged Hall effect in the tearing mode.

[27] Now consider the role of Poisson’s equation. Let us neglect for a moment the kinetic term arising from a consistent description of trapped and transient electrons in the isotropic equilibrium (4), which corresponds to the first term in \( \mathcal{L}_{22} \) given by (33). Using only the leading term in (35) \( ikA_2 (n_0 - b) (b^2 - 1)/b_s^2 \) we reconstruct the second basic relation (equation (88)) of the fluid model [Galeev, 1984]

\[
n_1 = ikn_0 \langle \langle A_1 \rangle \rangle_{\mu} / B_s
\]  

(45)

because the additional form-factor

\[
F(b, b_m, b_s) = \frac{2(b_m - b)(b^2 - 1)}{b_s b_s^2}
\]  

(46)

do not effect the value of this density averaged over the flux tube

\[
\langle n_1 \rangle_{\mu} = ikn_0 \langle \langle A_1 \rangle \rangle_{\mu} / B_s
\]  

(47)

Thus practically all elements of Galeev’s [1984] fluid model are captured in our kinetic theory. The only essential difference between this fluid model and our kinetic theory, which however plays a crucial role in the stability of the tearing mode, is the first term in \( \mathcal{L}_{22} \) given by (33) in the Poisson’s equation (28)

\[
\sigma \left( \langle \phi \rangle - \phi \right)
\]  

(48)

This term arises due to different responses of trapped and transient electrons to the tearing perturbation, but it does not exhibit dependence on the relative fraction of transient particles. However, according to (B11), it depends on the parameter \( B_{\|}/B_\perp \), which controls the local value of that fraction close to the neutral plane, as shown in appendix D. According to Lembege and Pellat [1982] and Coroniti [1980], the term (48) disappears after the integration over an infinitely long flux tube, making their models independent of the temperature ratio \( \sigma \) and eliminating the destabilizing effect of transient electrons. However, the latter operation is not consistent with the rest of the energy principle analysis. As was shown by Sitnov et al. [1998], if the integral in the denominator of the stabilizing term (equation (C6) of Lembege and Pellat [1982] or equation (3) of Coroniti [1980]) were taken over an infinite flux tube, it would go to infinity, thereby making the whole stabilizing contribution zero. In contrast, in the self-consistent isotropic equilibria (4) the flux tube is limited to the current sheet boundaries \( z < L \), and this limitation, according to Sitnov et al. [1998], radically changes the results of the energy principle analysis. It should be noted that the stability properties of the tearing mode as a nonlocal current sheet perturbation are determined not by the local values of the terms like (48) but rather by their moments involving integration over the current sheet \( z < L \). This explains why the local value of the fraction of transient particles formally disappears from some of our stability results. As shown in appendix D, this fraction itself, when averaged over the region \( z < L \), is approximately constant, independent of any parameter including \( B_\|/B_\perp \).
The model of Galeev [1984], which has been inferred above from our kinetic analysis, completely reproduces the earlier stability result by Lembege and Pellat [1982], which is consistent with our result in the limit \( \sigma \to 0 \) as shown in Figure 1. Note that this model is also fully consistent with the Hall-MHD models, which have been used for the description of the plasma behavior in the vicinity of the X-line [Biskamp, 2000; Birn et al., 2001]. However, as we have shown in our above kinetic analysis, this fluid description is incomplete and misleading when applied to the reconnection onset problem. For realistic values of the temperature ratio, the kinetic analysis reveals the instability gaps in the WKB region. This result is consistent with earlier conjecture made by Schindler [1974] about the negligible role of cold (\( \sigma \to \infty \)) electrons in the stability problem.

Similar kinetic effects have been found for kinetic ballooning and internal kink modes in tokamak plasmas [Cheng, 1982; Antonsen and Bondeson, 1993] in the form of a stabilizing influence of the thermal trapped electrons and progressive mode destabilization with the increase of the temperature ratio \( T_i/T_e \) [Antonsen and Bondeson, 1993]. Although these modes do not reconnect the field lines, the corresponding kinetic effects are quite interesting and relevant as they provide the destabilization due to the short-circuiting of the electrostatic potential in the absence of conducting boundaries. These effects however do not qualitatively change the stability picture in fusion devices. In contrast, the role of the kinetic term (48) in (28) is dramatic indeed. While the fluid relations (44) and (45) reproduce the fluid results by Lembege and Pellat [1982], Coroniti [1980], and Galeev [1984] about the complete stability of the tearing mode, the kinetic term creates considerable instability gaps in the WKB region of the parameter \( kL \). Thus it becomes a key to the onset of reconnection in realistic tail-like configurations and shows that this is actually a kinetic process.

This is only a key to the reconnection onset. It is quite possible that after the linear stage of the tearing instability the corresponding perturbation may saturate because of nonlinear effects, thereby not reaching large enough amplitude to actually tear the magnetic field lines. On the other hand, it may develop to a stage to provide the X-line geometry after which other important effects around the X-line such as the Hall current [Terasawa, 1983; Daughton, 1999; Birn et al., 2001] may come into the game and strongly change the reconnection onset process. We believe nevertheless that the mechanism explored above is an important new ingredient of this process.

5. Discussion and Conclusion

The tearing instability was first proposed as a mechanism of the energy release in Earth’s magnetotail long before the basic substorm models were proposed [e.g., McPherron et al., 1973; Lui et al., 1992]. It has been persistently considered as a basis of the particular near-Earth neutral line (NENL) model of substorms [McPherron et al., 1973; Hones, 1979; Baker et al., 1996], and the advances and failures of the tearing stability analysis were all the time tightly intertwined with the debates on substorm models. It is not surprising therefore that the conclusion about the universal stability of the tearing mode [Lembege and Pellat, 1982; Pellat et al., 1991; Brittnacher et al., 1994; Quest et al., 1996] stimulated the development of the substorm models other than the NENL model. In particular, based on the dominance of earthward hot plasma flows during substorm expansion within \( \sim 20 R_E \) in the tail [Lui et al., 1977; Baumjohann et al., 1989] it was concluded that near-Earth neutral line is rarely if ever formed during the initial phase of the substorm [Lui, 1992]. Then the substorm activity in the near-Earth tail can be alternatively explained as a result of the instabilities driven by the kinetic energy of the dawn-dusk flows of unmagnetized ions near the neutral plane [Lui et al., 1991]. The instabilities are supposed to develop earthward of the near-Earth neutral line, and its formation is considered as a consequence of the disruption of the original tail current because of these instabilities. Such scenario is known now as the current disruption (CD) model of substorms [e.g., Lui, 1996]. The corresponding tail current sheet stability studies later transformed into a more realistic nonlocal analysis of drift-kink, drift-sausage and lower hybrid drift instabilities [Lui et al., 1995; Yoon and Lui, 1996; Zhu and Winglee, 1996; Pritchett and Coroniti, 1996; Pritchett et al., 1996; Lapenta and Brackbill, 1997; Hesse et al., 1998; Yoon et al., 1998], which differ from the tearing mode in that the perturbations develop largely in the dawn-dusk direction (\( k_x \neq 0 \)), while the subsequent reconnection resembles Sweet-Parker current sheet (due to the effective turbulent resistivity) rather than the X-line pattern. Another family of substorm models originated from the so-called Kiruna conjecture [Kennel, 1992] that the auroral substorm onset is very closely coupled to events in the geosynchronous region. The main free energy source of the corresponding ballooning-interchange instabilities [Miura et al., 1989; Roux et al., 1991; Ohtani and Tamao, 1993; Bhattacharjee et al., 1998; Cheng and Lui, 1998; Hurricane et al., 1999; Lee, 1999; Wong et al., 2001] is the pressure gradient along the tail axis.

Meanwhile recent Geotail observations [Mukai et al., 1998; Nagai et al., 1998] have confirmed that the magnetic reconnection may take place in the near-Earth tail plasma sheet prior to substorm onset detected on the ground. On the other hand, simulations have shown that the substorm events in the geosynchronous region may arise because the earthward flows caused by reconnection at \( X \lesssim -20 R_E \), are abruptly braked at the near-Earth edge of the tail current sheet [Birn and Hesse, 1996; Shiokawa et al., 1997]. Also, recent 3D particle simulations [Pritchett, 2001; Zeiler et al., 2002] demonstrated the persistent stability of the X-line patterns with respect to \( k_x \neq 0 \) instabilities. These findings strongly support the NENL model of substorms with its laminar X-line structure of the reconnection region, which can be formed due to the tearing instability. The main result of our nonlocal kinetic analysis, that this instability can develop for realistic tail current sheet conditions (plasma isotropy, \( B_p/B_0 \) and ion-to-electron temperature ratio), provides the firm theoretical basis for this model.

It is interesting however that our results are not at variance with the above alternative models of substorms, with the emphasis on the processes in the transition region between the dipole and tail-like magnetic fields. The key issue is that the length of this transition region is comparable to the current sheet thickness. As a result, all particles,
which are transient for such a current sheet will be reflected right away from the dipole field region. This is exactly the situation when the assumptions of the model [Lembege and Pellat, 1982] are valid. As first noticed by Syrovatskii [1971] and later elaborated by Kulcsrud and Hahm [1982] and Schindler and Birn [1993], boundary perturbations of the current sheet either violate conservation of magnetic topology or continuity. This results in the formation of either X-lines or Y-lines, respectively [e.g., Biskamp, 2000, and references therein]. Far from the transition region, where the length of the tail current sheet is much larger than its thickness, our analysis predicts the spontaneous change of its topology due to tearing mode growth as shown schematically in Figure 5. Closer to the Earth the sheet is tearing-stable and perturbations due either to the immediate solar wind trigger or to quasi-static changes in the process of the magnetospheric convection should result in the formation of discontinuities. Their kinetic analogues are thin current sheets [Sitnov and Sharma, 2000, and references therein], complex non-MHD structures, often embedded within thicker plasma sheets and presumably more susceptible to current-driven instabilities because of the bulk velocity shear [Isoon and Lui, 1996]. This picture is fully consistent with spacecraft observations. In particular, Geotail measurements showed that the formation of the near-Earth neutral line during substorms, detected by the tailward ion flows and southward magnetic field, usually starts in the premidnight sector of the magnetotail between $X_{GSM} = -20R_E$ and $X_{GSM} = -30R_E$ prior to an onset signature identified with Pi 2 pulsations on the ground [Nagai et al., 1998]. On the other hand, thin current sheets are detected as a distinctive feature of the near-Earth magnetotail between $X_{GSM} \sim -7R_E$ and $X_{GSM} \sim -18R_E$ prior to and during substorms [McComas et al., 1986; Kaufmann, 1987; McPherron et al., 1987; Mitchell et al., 1990; Lui et al., 1992; Sergeev et al., 1993, 1998; Sunny et al., 1994; Pulkkinen et al., 1994]. It thus appears that both types of the reconnection, involving the formation of X- and Y-lines happen in the Earth’s magnetosphere.

[34] The impact of the results of Lembege and Pellat [1982] extends far beyond the modeling of substorms. In fact, at present collisionless reconnection studies have been polarized into two mutually inconsistent directions. One line of investigation studies the linear and nonlinear alternatives to the tearing instability to describe the reconnection onset [e.g., Daughton, 1999; Shinohara et al., 2001], which results in turbulent reconnection in the Sweet-Parker-Syrovatskii regime with Y-lines connected by thin current sheets. The other large group of researchers study reconnection in the X-line vicinity, merely overlooking the unresolved problem of the X-line formation [e.g., Biskamp, 2000; Birn et al., 2001, and references therein]. It is assumed that the X-line can be formed as a result of an external perturbation of the initially stretched or anti-parallel magnetic field configuration. However, as argued above, this really happens if the change of topology provided by the tearing mode is energetically favorable. Otherwise, the system evolves toward Y- rather than X-line pattern. This is why the very interesting Hall-MHD reconnection physics revealed in the X-line vicinity cannot be directly applied to model the reconnection in real tail-like systems. In particular, coming back to the magnetospheric physics, the Hall-MHD models cannot be utilized in global MHD simulations unless the kinetic physics of the reconnection onset is taken into account. The nonlocal kinetic stability analysis presented in this paper is thus believed to serve as a link connecting these two large areas of magnetospheric modeling.

[35] Of course, more work is necessary to meet the requirements of current MHD and Hall-MHD simulations of the magnetosphere. In the present analysis we had to greatly simplify the tail current sheet model. One can expect in particular considerable changes of our stability picture in case of the nonzero $B_z$ component of the initial magnetic field [Wang et al., 1990; Hernandes et al., 1993; Wang and
Bhattacharjee, 1993; Brittnacher et al., 1995; Kuznetsova et al., 1996) or when this component appears as a part of the more general tearing perturbation [Terasawa, 1983; Daughton, 1999]. As shown by Zhu and Parks [1993], Chapman and Rowlands [1998], and Delcourt et al. [2000], the $B_z$ component of the initial current sheet magnetic field may strongly change the single particle dynamics and in particular increase the number of transient particles [Chapman and Rowlands, 1998]. The latter result suggests that the finite $B_z$ component may amplify the effect, which we describe in the present paper, leading to further destabilization of the tearing mode. External diffusion and chaos effects in either electron or ion dynamics may also influence the current sheet stability [Hernandes et al., 1993; Brittnacher et al., 1995, 1998] and plasma transport properties [Horton and Tajima, 1991; Hernandes et al., 1993], and this influence is shown to increase through the kinetic mechanism considered in the present paper [Sitnov and Lui, 1999]. Finally, a more detailed description of transient electrons and the determination of the specific location of the future X-line position in the tail will require going beyond the WKB approximation [e.g., Goldstein and Schindler, 1982]. We plan to address these problems in our future publications.

Appendix A: Bounce-Averaged Part of Perturbed Electron Distribution

[36] Let us take explicitly into account the conservation of the first and second adiabatic invariants of the particle motion. These invariants together with the constant dawn-dusk component of the canonical momentum $P_z$ form a complete set of integrals of motion to represent the general solution of Vlasov equation as an arbitrary function of these invariants. Then the perturbed distribution may be written in the form

$$f = f_0(P_z, H_0(I_1, I_2, P_y))$$

(A1)

Here $f_0$ is the initial equilibrium distribution. $H_0$ is the initial Hamiltonian as a function of the initial invariants, while their new functional dependence on the point in the phase space $(r, v) \rightarrow (r, W, vB_{(0)}(B_{(0)})(W$ is the particle energy, $B_{(0)} = curL_0$ is the equilibrium magnetic field) are now determined by the perturbed Hamiltonian

$$H = \frac{mv_y^2}{2} + \frac{mv_z^2}{2} + \frac{1}{2m}P_z^2 + q/\epsilon A^2 + q\phi_1 \equiv W + q\phi_1$$

(A2)

with $A = A_0 + A_1$. The first and second adiabatic invariants may be written, respectively, as follows

$$I_1 = \frac{mc}{q} (\mu + \Delta \mu) = \frac{1}{2\pi} \int v' dW'$$

$$= \frac{1}{2\pi} \int \sqrt{2mW - (P_y')^2 - \left(P_z - q\epsilon A^2\right)^2} - q(\phi_1' - \phi_1) dW'$$

$$I_2 = \frac{1}{2\pi} \int P_y' ds = \frac{1}{2\pi} \int \sqrt{2m(W - \mu' B' - q(\phi_1' - \phi_1))} dW'$$

(A3)

Here $\mu = mv_y^2/2B$ is a conventional definition of the magnetic moment with $n = B/B$, $B = B_{(0)} + B_1$, $P_\parallel = P \cdot b$, $P_\perp = P \cdot n$. In (A3) we use the local coordinate system $(e_x, e_y, e_z)$ with $e_z = n$ and $dl = e_x dx' + e_y dy' + e_z dz'$, while $dW'$ is the differential distance along the magnetic field line. The index $(\ldots)'$ marks the point on the Larmor orbit, $(\ldots)$ marks the point of the guiding center, and the potentials without indices characterize the fixed point in the phase space. Note that the invariant $I_1$ given by (A3) is conserved at each point of the bounce orbit, while the definition of the invariant $\mu$ is relevant only in the considered point $(r, v)$ of the phase space.

[37] We can obtain now the linear part $f_1$ of the perturbation $f(A_1, \phi_1) - f(0, 0)$ by variation of (A1) with respect to $A_1$ and $\phi_1$

$$f_1 = \frac{\partial f_0}{\partial A_1} A_1 + \frac{\partial f_0}{\partial \phi_1} \phi_1$$

(A5)

where the bounce averaging in the right hand side is performed over the unperturbed particle orbits. The variation of the second invariant (A4) as follows

$$\delta I_2 = -\frac{1}{2\pi} \int m \left(\frac{\partial B_0}{\partial P_y} + \frac{\partial c}{\partial I_1} \frac{\partial P_y}{\partial A_1} \right) ds$$

$$\approx -\frac{1}{2\pi} \left(\langle \partial B_0 \rangle + \langle \partial c \rangle \langle \partial A_1 \rangle \right)$$

(A6)

where $\langle \ldots \rangle$ denotes the averaging over the period of bounce-oscillations $\tau_b$.

[38] To calculate $\delta I_1$ in (A6) we express the variation of the first invariant (A3) for an arbitrary point of the bounce orbit

$$\delta I_1 = -\frac{1}{2\pi} \int m v_y B_0 + P_\parallel B_1 + q/c(A_1 c - qA_0^2)A_1'' ds$$

(A7)

For simplification of this expression we expand the potential relative to the guiding center $R: r \equiv R + \rho N A_1$ to obtain

$$\delta I_1 \approx \frac{mc}{q} \frac{v_y}{v_0} + \frac{\tau_L}{2\pi} \left(\frac{q}{c} v_0 (A_1 c - (\phi_1' - \phi_1)) - v_0 \partial_0 B_y \right)$$

(A8)

where

$$v_0 = \frac{1}{2\pi} \int v_0'' ds$$

(A9)

is $v_y$ averaged over the Larmor period $\tau_L$ (denoted as $\langle \ldots \rangle_\tau_L$). We also take into account that $\langle I_1 \rangle = \delta I_1$ due to adiabatic invariance,

$$\frac{\partial H_0}{\partial I_1} = \left(\frac{2\pi}{\tau_L}\right) \epsilon \frac{\partial H_0}{\partial I_2} = \frac{2\pi}{\tau_b}$$

(A10)

due to specific properties of action variables $I_1$ and $I_2$, and

$$\frac{\partial H_0}{\partial P_y} = \langle v_y \rangle = \langle v_0 \rangle = 0$$

(A11)

due to the specific geometry of the equilibrium magnetic field. Taking into account these relations and substituting...
(A6) and (A8) into (A5), we obtain the bounce averaged perturbed electron distribution in the form

\[
\bar{f}_i = f_i - \frac{\partial f_i}{\partial p_i} \frac{\partial p_i}{\partial A_0} A_0 = \frac{\partial f_i}{\partial H_0} \left[ q(\phi_i - \langle \phi_i \rangle) + \left\{ \frac{q}{c} v_{d} A_1 - \mu B_1 \right\} \right]
\]

(A12)

which was obtained first by Lembege and Pellat [1982] using the drift wave approximation.

**Appendix B: Integral Terms in Poisson’s Equation**

[40] Let us start from the first term in the Poisson’s equation (25)

\[
\langle \langle \phi \rangle \rangle_0(x, z) = \frac{4 \pi B_0}{m_e^2} \int_0^\infty (f_{hi}/n_0)(W) dW
\]

\[
\times \int_{W/B_m}^{w/L} \frac{dW}{W_m} \tau_0(W, \mu) \int \frac{dW}{W_m} \int \frac{dW}{W_m} \phi(s') \\
\]  

(B1)

Following Coroniti [1980], we introduce new variables for description of the integration over the bounce orbit

\[
b(s) = \frac{B(x, z)}{B_m} = \sqrt{1 + b^2 s \tanh (\frac{z}{L})}
\]

(B2)

where \(b_s = B_0/B_m\), to get

\[
x(s) = x_0 + b_s L \ln \left( \frac{\cosh (\frac{z(b)}{L})}{1} \right),
\]

where \(x_0 = x(b = 1)\) is the x-coordinate of the flux tube,

\[
\frac{db}{ds} = \frac{b_s}{L} \sqrt{b^2 - 1} \left( 1 - b^2 s \right)
\]

and

\[
\tau(W, \mu) = \tau_0(W, \mu) \frac{B_0}{4 L} \frac{2 \mu}{m_e B_m} \\
= \int b_s \frac{b^2 db}{(1 - b^2 s)(b^2 - 1)} \sqrt{b_m - b \sqrt{b^2 - 1}}
\]

\[
\equiv \tau(b_m)
\]

(B4)

(B5)

where \(b_m = W(\mu B_m)\). Now with \(W = W/T_e\) the key expressions in (B1) can be rewritten as

\[
\mu = \frac{w_T e}{b_s B_m}, \int_{W/B_m}^{w/L} dW = \frac{w_T e}{B_m} \int b_s db
\]

(B6)

where \(b = \sqrt{1 + b^2 s}\),

\[
\int dW = \frac{L}{b_s} \int b_s \frac{b^2 db}{(1 - b^2 s)(b^2 - 1)} \sqrt{b^2 - 1}
\]

\[
v_{\mu}(x, W, \mu) = \sqrt{\frac{2 \mu B_e}{m_e} (b_m - b_s)} = v_{Te} \sqrt{\frac{w}{b_m - b_s}}
\]

(B7)

(B8)

Using the above expressions and simplifying the equilibrium distribution function for electrons noting that the asymmetry is very small, \(\omega_e/\omega_T = \rho_w L \ll 1\) the integral term (B1) can be presented as

\[
\langle \langle \phi \rangle \rangle_0(x, z) = \frac{b_s}{2} \int_{b_s}^{b} \frac{db}{\tau(b_m) b_s^{1/2}} \sqrt{b_m - b_s}
\]

\[
\times \int \frac{b_s^2 db}{\sqrt{b_m - b \sqrt{b^2 - 1}} (1 - b^2 s)(b^2 - 1)} \\
\cdot \phi(x(b), z(b)) \]

\[
(B11)
\]

Consider now the expression \(\langle \langle (v_{Te} e) A_1 + (\mu/e) nB_1 \rangle \rangle_0\). Using the definition of the drift velocity, it may be written in the form [Lembege and Pellat, 1982]

\[
v_{x} = \frac{w_T e}{2 \omega_c B_m L} \frac{b^3 - b - 2b_s}{b^4} \left( 1 - b^2 s \right)
\]

(B13)

where \(\omega_c = eB_0/m_e c\). Note that for the parabolic field (with \(1 - b_s^2(b^2 - 1) \rightarrow 1\)) this expression for the drift velocity coincides with the equation (7) used by Coroniti [1980].

[41] The second term in \(\langle \langle (v_{Te} e) A_1 + (\mu/e) nB_1 \rangle \rangle_0\) can be written, in our notations, as

\[
\frac{\mu}{e} n B_1 = \frac{w_T e}{2 \omega_c B_m L} \int \frac{1}{b_s} \frac{\partial A_1}{\partial \mu} - \frac{(b^2 - 1)^{1/2}}{b} \frac{\partial A_1}{\partial b}
\]

(B14)

Using now the definition (26), notations (B5)–(B10), the expressions (B13) and (B14), and the approximation of the electron distribution, which we have used in derivation of (B11) we obtain

\[
\langle \langle \langle \frac{\mu}{e} A_1 + \frac{\mu}{e} n B_1 \rangle \rangle_0 \rangle \rangle_0(x, z) = \frac{3}{8} \frac{v_{Te} \sqrt{b_s B_e}}{L \omega_c} \int_{b_s}^{b} \frac{db}{\tau(b_m) b_m \sqrt{b_m - b_s}}
\]

\[
\times \int \frac{b_s^2 db}{\sqrt{b_m - b \sqrt{b^2 - 1}} (1 - b^2 s)(b^2 - 1)} \\
\cdot \frac{b^3 - b - 2b_s}{b^4} \left( 1 - b^2 s \right) A_1 + \frac{L}{b} \frac{\partial A_1}{\partial b}
\]

(B15)

It is convenient for our analysis to further transform the latter expression using the identity
\[
\frac{\partial}{\partial b} \left( \frac{\sqrt{b^2 - 1} - \sqrt{b_m - b}}{b} \right) = -\frac{1}{2} \frac{b^3 + b - 2b_m}{b^2 \sqrt{b^2 - 1} \sqrt{b_m - b}} \quad (B16)
\]

and integrating by parts we get
\[
\phi \left( A_1 + \frac{\mu}{e} n B_1 \right)_p (x, y, z) = \frac{3}{8} \frac{v_b^3 b_t}{c_0 \omega_m} \int_0^b \frac{db_m}{\tau_b(b_m) b_m \sqrt{b_m - b}} \int_1^{b_m} \frac{b_m^2 db_m}{\sqrt{b_m - b} \sqrt{b - b_m} - 1} \left( 1 - b_m^2 (b - b_m^2) \right) \frac{\partial A_1}{\partial z} (x, y, z) \quad (B17)
\]

**Appendix C: Integral Terms in Ampere’s Equation**

[42] According to the definition (23), the term \( \langle \phi \rangle_j \) can be presented in the form
\[
\langle \phi \rangle_j (x, y, z) = \frac{2 \sqrt{2} |B|^{3/2}}{m^2 v_r} \int_0^\infty dW \int_0^{2\pi} d\theta \sin \theta \int_{W/B_i}^{1/2} \mu^{1/2} d\mu \times \frac{F_0 (W, \mu, 0)}{\tau_b (W, \mu, 0) |\nu| (x, W, \mu)} \int_{|\nu|}^{\eta} \frac{d\nu'}{|\nu'| (x', W, \mu')} \phi(x', y', z'), \quad (C1)
\]
or using notations (B5)–(B10)
\[
\langle \phi \rangle_j (x, y, z) = \frac{b_t^2}{2 \pi} \int_0^\infty \int_0^{2\pi} d\theta \sin \theta \int_0^b \frac{db_m}{\tau_b(b_m) b_m \sqrt{b_m - b}} \times \exp \left[ -w + wb_b b_m^{1/2} \sin^2 \theta \right] \left[ \left( w^{1/2} b_t^2 b_m^{-1/2} \sin^2 \theta - \epsilon \right)^2 \right] \int_1^{b_m} \frac{b_m^2 db_m}{\sqrt{b_m - b} \sqrt{b - b_m} - 1} \left( 1 - b_m^2 (b - b_m^2) \right) \Phi(x, y, z), \quad (C2)
\]
where \( \epsilon = u_s \sqrt{\tau_c} = \rho_{0L}. \) In contrast to the Poisson’s equation, we cannot use in the above formula the approximation of symmetric electron distribution as it gives a zero contribution. Let us take into account however that the \( \epsilon \) is very small. Even in thin CS with \( L \sim \rho_{0L} \) we have \( \epsilon \sim \langle mTC/mT \rangle \sim 10^{-2}. \) Also, the parameter \( w^{1/2} b_t^2 b_m^{-1/2} \sin \theta \) is small because \( b_j < b_m. \) Therefore we can neglect in (C2) the factor \( \exp (-\epsilon) \), while another exponent \( \exp (-2\epsilon w^{1/2} b_t^2 b_m^{-1/2} \sin \theta) \) may be approximated as \( 1 - 2\epsilon w^{1/2} b_t^2 b_m^{-1/2} \sin \theta. \) As a result (C2) may be rewritten as
\[
\langle \phi \rangle_j (x, y, z) = \frac{3 \rho_{0L} b_t^2}{4L} \int_0^b \frac{db_m}{\tau_b(b_m) b_m \sqrt{b_m - b}} \int_1^{b_m} \frac{b_m^2 \Phi(x, y, z) db_m}{\sqrt{b_m - b} \sqrt{b - b_m} - 1} \quad (C3)
\]

Similarly, using the above approximation, the expressions (B13) and (B14), and the identity (B16) we obtain
\[
\langle \frac{\nu_t}{c} A_1 + \frac{\mu}{e} n B_1 \rangle (x, y, z) = \frac{15}{16} \frac{b_t^2 v_e}{L c_0 \omega_m \omega_0} \int_0^b \frac{db_m}{\tau_b(b_m) b_m \sqrt{b_m - b}} \int_1^{b_m} \frac{b_m^2 db_m}{\sqrt{b_m - b} \sqrt{b - b_m} - 1} \left( 1 - b_m^2 (b - b_m^2) \right) \frac{\partial A_1}{\partial z} (x, y, z) \quad (C4)
\]

Asymmetry of the electron distribution can be again neglected in calculating the expressions \( \langle \phi \rangle_j \) and \( \langle \nu_t/c \rangle A_1 + \langle \mu/e \rangle n B_1 \rangle, \) as this approximation gives nonzero contribution. Then using the identity \( (1/\omega_m) \times \nabla \chi \chi (x_0) = (1/\omega_m) \times (\partial/\partial \delta (x_0)) \times (x_0), \) where \( x_0 = x \), \( b = 1 \), is the x-coordinate of the specific flux tube, and that according to (B3), \( (\partial/\partial \delta) g (x_0) = ik g_0 \), the first of these expressions can be presented in the form
\[
\langle \phi \rangle_j = \frac{3 \rho_{0L} b_t^2 v_e}{8} \int_0^b \frac{db_m}{\tau_b(b_m) b_m \sqrt{b_m - b}} \int_1^{b_m} \frac{b_m^2 db_m}{\sqrt{b_m - b} \sqrt{b - b_m} - 1} \left( 1 - b_m^2 (b - b_m^2) \right) \phi(x, y, z), \quad (C5)
\]
while
\[
\langle \nu_t A_1 + \frac{\mu}{e} n B_1 \rangle = i \frac{15}{32} \frac{b_t^2 v_e}{c_0 \omega_m \omega_0} \int_0^b \frac{db_m}{\tau_b(b_m) b_m \sqrt{b_m - b}} \int_1^{b_m} \frac{b_m^2 db_m}{\sqrt{b_m - b} \sqrt{b - b_m} - 1} \left( 1 - b_m^2 (b - b_m^2) \right) \frac{\partial A_1}{\partial z} (x, y, z) \quad (C6)
\]

**Appendix D: Evaluation of Relative Amount of Transient Electrons**

[43] Using the notations of appendix B and the condition \( b_0 \gg 1 \), the relative amount of transient and trapped electrons at the neutral plane \( z = 0 \) can be presented in the form
\[
\frac{n_{trans}}{n_{trans}} \approx \left( \int_{b_0}^{\infty} \frac{db_m}{b_m^3 \sqrt{b_m - 1}} \right)^{-1}\left( \int_{1}^{b_0} \frac{db_m}{b_m^3 \sqrt{b_m - b}} \right)^{-1} \approx \frac{1}{2b_0}, \quad (D1)
\]
On the other hand, the similar ratio of densities averaged over the region $|z| < L$ can be estimated as follows:

$$
\frac{b_c}{b_{\text{H}}^\text{avg}} \approx 1 + \frac{1}{2} \int_{b_c}^{b_{\text{H}}} \frac{b_c^2}{b_{\text{H}}^2 - b_c^2} \frac{db_c}{db_{\text{H}}} \left( \int_{b_c}^{b_{\text{H}}} \frac{db_{\text{H}}}{db_c} \right)^{-1}
$$

,$$
L = \int_{b_c}^{b_{\text{H}}^\text{avg}} \frac{db_c}{db_{\text{H}}}
$$

In the limit $b_{\text{H}} \gg 1$, after the change of sequence of integration, the numerator in the right hand side of (D2) can be estimated as $(2/3) b_{\text{H}}$, while the denominator as $(4/3) b_{\text{H}}$, giving the estimate value 1/2 for the relative amount of transient and trapped electrons within the region $|z| < L$, which thus does not depend on the parameter $b_{\text{H}}$.

[D44] Acknowledgments. This research is supported by NSF grant ATM-9901733 and NASA grant NAGS10298.

References


Lui, A. T. Y., C.-L. Chang, and P. H. Yoon, Preliminary nonlocal analysis of


