Inductive-Dynamic Coupling of the Ionosphere With the Thermosphere and the Magnetosphere

P. Song

Center for Atmospheric Research and Department of Physics, University of Massachusetts Lowell, Massachusetts, USA

V. M. Vasyliūnas

Center for Atmospheric Research and Department of Physics, University of Massachusetts Lowell, Massachusetts, USA Max-Planck-Institut für Sonnensystemforschung, Katlenburg-Lindau, Germany

The central issue of magnetosphere-ionosphere-thermosphere coupling is the long-range coupling of the system, from the solar wind and magnetosphere to the ionosphere and thermosphere; coupling between the ionosphere and thermosphere is known and appropriately modeled by local coupling via collisions. The longrange coupling is mostly through the electromagnetic force in addition to direct flow. In steady state, the electromagnetic force may be represented by "mapping" of the electric potential along the magnetic field and field-aligned currents with Ohm's law accommodating the local coupling in the ionosphere. To correctly describe the long-range coupling on time scales from longer than few seconds to less than 30 min, the inductive effect (Faraday's law) as well as the dynamic effect (acceleration terms) need to be considered. When the inductive and dynamic effects are included, the physical description of the long-range coupling is drastically changed: the electric potential no longer suffices to describe the electric field, and the electric currents, including Birkeland (magnetic field-aligned) currents, become a secondary derived quantity. All these are replaced by electromagnetic perturbations with associated fluid perturbations, propagating and reflecting between the magnetosphere and ionosphere and from one hemisphere to the other. The perturbations in one part of the ionosphere can propagate into other ionospheric regions, relatively rapidly, which, in turn, couple back to other regions of the magnetosphere. In this article, we discuss the key physical and mathematical differences and compare the physical quantities according to each of the coupling schemes.

1. INTRODUCTION

In a partially ionized gas, the electrons, the ions, and the neutrals each move under the action of different forces. Electromagnetic forces act only on the charged particles but not on the neutrals. The charged and the neutral particles interact through collisions, which transfer energy and momentum from one species to another while converting energy from one type to another. Electrons and ions can be

Modeling the Ionosphere-Thermosphere System Geophysical Monograph Series 201 © 2013. American Geophysical Union. All Rights Reserved. 10.1029/2012GM001308 treated together as plasma. Depending on time scales, the plasma and the neutrals may be treated as moving together when the collision times are much shorter than the time scale of interest or as moving separately when the collision times are much longer than the time scale of interest. On intermediate time scales, one species may accelerate or drag the other, while the whole fluid is heated by frictional heating due to collisions among them. Note that collisions are a local process, whereas long-range coupling, which connects the processes and variations in one region to those in another (say, from the magnetosphere to each altitude in the ionosphere) is by electromagnetic forces and by direct flow (if it exists) between the two regions. Magnetosphere-ionosphere/ thermosphere (M-IT) coupling is an example of such longrange coupling processes, connecting the highly ionized nearly collisionless magnetospheric plasma to the weakly ionized highly collisional gases of the ionosphere and thermosphere with local coupling (at a given height) between the ionosphere and thermosphere. In this respect, a particular challenge for M-IT coupling theories arises from the rapid increase in collision frequency as one goes from the magnetosphere to the ionosphere/thermosphere, with the result that, for any given perturbation, the dominant process changes from collisionless to strongly collisional over a very short distance, the location and spatial extent of this transition depending on the time scale of the perturbation.

Conventional M-IT coupling models are all based on the classical magnetosphere-ionosphere (M-I) coupling theory, developed by *Vasyliūnas* [1970] and *Wolf* [1970] from earlier models of *Fejer* [1964] and others, which is discussed in section 2.1. Most of the models may be assigned to one of three general categories: an electrostatic ionosphere, a height-integrated ionosphere, and a structured ionosphere with wave propagation. Since the focus of this *Monograph* is the ionosphere/thermosphere (IT) subsystem, we discuss the first category in more detail below.

Most of the models developed in the ionospheric community focus on the region up to no more than a few thousand kilometers above the Earth. Wave propagation or communication time over this relatively small distance is sufficiently short so that one can ignore wave propagation effects, while the full dynamics of the neutral wind can be retained. The most important assumption invoked in these models is that the time-variability of the magnetic field is negligible, hence the inductive term in Faraday's law may be dropped and the electric field assumed curl-free. As a result, the models become electrostatic, which leads to a significant simplification mathematically because the electric potential and the Birkeland current will now suffice to calculate the long-range coupling. The mathematical convenience of using electric fields and currents is elevated to a physical description ("E-J approach") in many discussions. In addition, the time derivative (or dynamic) term in the plasma momentum equation is neglected in most models. This formalism may be adequate to describe the IT system alone (as long as inputs from or coupling to the magnetosphere play no role), although even there, as argued by *Vasyliūnas* [2012], its equations are only conditions for quasisteady state equilibrium and may not provide insight into temporal and causal sequences. For global M-IT coupling, its validity is even more limited: an often overlooked key point is that this scheme, restricted to describing M-IT coupling in quasisteady state equilibrium, cannot be applied to explain transient processes, such as substorms and auroral brightenings, during which the magnetic perturbations and plasma acceleration are not negligibly small.

A key parameter is the time scale that separates the dynamic and the quasisteady state, or equivalently the transient time, required for the coupled M-IT system to change from one quasisteady state to another. Theoretical estimates of this time scale (discussed in section 2.4) depend on the spatial scale of the phenomenon of interest, ranging from a few seconds for purely ionospheric events to 15-30 min for changes on scales of magnetospheric dimensions. The latter is the approximate time scale for many observed important M-IT processes, such as substorms and magnetospheric transients [e.g., Earle and Kelley, 1987; Russell and Ginskey, 1993, 1995; Zesta et al., 2000; Bristow et al., 2003; Huang et al., 2008, 2010]; the focus of our paper is on time scales from a few minutes to 20-30 min. It is, thus, not surprising that, as pointed out by Schunk [2013], conventional ionospheric quasisteady state equilibrium models, e.g., thermosphere ionosphere electrodynamics global circulation model [Richmond et al., 1992], coupled thermosphere ionosphere plasmasphere electrodynamics model [Fuller-Rowell et al., 1996], and the global ionosphere-thermosphere model (GITM) [Ridley et al., 2006], describe well the large-scale slow variations or "climatology" of the ionosphere-thermosphere system but do not work well for the rapidly changing phenomena or "weather." (The SAMI 2 ionospheric model [Huba et al., 2000] does include the time derivatives in the field-aligned component of the momentum equation but does not include the inductive field, so the long-range coupling in the model is still electrostatic.)

On the other hand, many of the M-I coupling models developed in the magnetospheric community, including almost all the global MHD simulation models, retain the inductive term and plasma dynamics in the magnetosphere (see more discussion in the work of *Song et al.* [2009]). They are coupled to the ionosphere through the height-integrated quasisteady state equilibrium Ohm's law, with the ionosphere treated as a height-integrated layer at the lower boundary.

There is an intrinsic inconsistency here: a full inductive and dynamic magnetosphere is being combined with an electrostatic ionosphere. Many of the (magnetospheric) wave models take the same approach, including wave propagation within the magnetosphere and reflection between the magnetosphere and the ionosphere but not within the ionosphere because the "height-integrated" ionosphere has been idealized as having zero extent in altitude. Also, most of those models do not include neutral wind dynamics, except for corotation of the atmosphere as a whole.

As exceptions, there are a few models that include inductive effects as well as plasma dynamics with a structured ionosphere. Taking advantage of powerful Fourier analysis techniques, these models treat time-dependent effects as waves [e.g., *Hughes*, 1974; *Hughes and Southwood*, 1976] and include propagation and reflection effects within the ionosphere. This approach, in general, is able to describe the dynamics of M-I coupling. (Recall that a step function, a mathematical representation of a changing condition, can be decomposed into a wave spectrum, and each frequency can propagate and reflect, but converting the perturbation as a sum of individually propagated and reflected waves from the frequency domain with phase information back to the time domain is by no means a trivial task.)

In contrast to the "E-J" approach mentioned above, the "B-V" approach in which the magnetic field and the plasma bulk flow are treated as the primary quantities has always been a basic principle of MHD [e.g., Cowling, 1957; Dungey, 1958], and its importance for the magnetosphere and the ionosphere has been urged particularly by Parker [1996, 2000, 2007]. A new impetus was provided when Buneman [1992] (in a laboratory-plasma paper unknown to the M-IT community) and, independently, Vasyliūnas [2001] showed mathematically that in a dynamic process, for a plasma sufficiently dense so that $V_A^2 \ll c^2$, the electric field by itself cannot produce plasma bulk flow, but plasma flow does produce an electric field self-consistently. Tu et al. [2008] later confirmed this argument with particle simulations. Vasyliūnas [2005a, 2005b, 2011, 2012] investigated the basics of the two approaches extensively and showed that, contrary to widespread assumption, the "B-V" approach is not derived exclusively (or even primarily) from MHD but is, in essence, a consequence of charge quasineutrality. Song et al. [2005] and Vasyliūnas and Song [2005] derived a general set of equations governing long-range coupling for partially ionized collisional flows with electromagnetic fields, solved simple situations for 1-D steady state, and derived the dispersion relations for wave-type solutions. Song et al. [2009] further solved the 1-D time-dependent problem for M-IT coupling and found that, for M-IT coupling, plasma flow and magnetic tension force (rather than electric field and current)

are indeed the primary coupling mechanisms. More recently, the new approach and formalism has been applied to study heating processes in the solar chromosphere [*Song and Vasyliūnas*, 2011] and in the terrestrial IT system [*Tu et al.*, 2011].

2. INDUCTIVE-DYNAMIC COUPLING VERSUS ELECTROSTATIC COUPLING

2.1. Differences in Governing Equations

To illustrate the differences between the conventional steady state coupling and the inductive-dynamic coupling, we first examine how each is formulated mathematically. The system consists of plasma, neutral medium, and electromagnetic fields. For simplicity of the comparison, we assume the plasma and neutral medium each consists of multiple species and is represented by its averaged mass and bulk velocity, as well as temperature. The momentum equations for electrons, ions, and neutrals can be combined and rewritten [e.g., *Song et al.*, 2005, 2009; *Vasyliūnas and Song*, 2005; *Vasyliūnas*, 2012] as the generalized Ohm's law, the plasma momentum equation and the neutral momentum equation. Under the approximations listed below, their leading terms are:

$$0 = N_e e(\mathbf{E} + \mathbf{V} \times \mathbf{B}) - \mathbf{J} \times \mathbf{B} - (m_e/e) \mathbf{v}_e \mathbf{J}, \qquad (1)$$

$$\rho \frac{d\mathbf{V}}{dt} = \mathbf{J} \times \mathbf{B} - \rho_i \mathbf{v}_{in} (\mathbf{V} - \mathbf{u}_n) - \nabla p + \rho \mathbf{g}, \qquad (2)$$

and

$$\rho_n \frac{\mathrm{d}\mathbf{u}_n}{\mathrm{d}t} = \rho_n \mathbf{v}_{ni} (\mathbf{V} - \mathbf{u}_n) - \nabla p_n + \rho_n \mathbf{g}, \qquad (3)$$

where e, ρ_{η} , p_{η} , **J**, **V**, **u**_n, **E**, **B**, **g**, N_e , and $v_{\eta\xi}$, are the elementary electric charge, the mass density and pressure of species η , the electric current, the bulk velocities of the plasma and neutrals, the electric and the magnetic fields, gravitational acceleration, electron concentration (number density), and the collision frequency between particles of species η and ξ , respectively; $v_e \equiv v_{ei} + v_{en}$. Subscripts *i*, *e*, and *n* denote ions, electrons, and neutrals; densities and pressures without subscripts refer to the plasma (electrons plus ions). Time derivatives in equations (2) and (3) are convective derivatives in the flow of the respective medium. Note that these equations hold locally. The set of equations is completed by the addition of continuity and energy equations and by Faraday's and Ampère's laws

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \tag{4}$$

204 INDUCTIVE-DYNAMIC COUPLING

$$\mathbf{J} = \nabla \times \mathbf{B} / \boldsymbol{\mu}_0, \tag{5}$$

which provide long range forces for global coupling. The principal approximations invoked in deriving equations (1)-(5) are the following:

(a) Neglect of $\partial \mathbf{J}/\partial t$ term and of $\partial \mathbf{E}/\partial t$ (displacement current) term on the left-hand sides of equations (1) and (5), respectively; this applies on time scales that are long compared to inverse plasma frequency and length scales that are large compared to electron inertial length and ensures the validity of charge quasineutrality (for a detailed discussion, see, e.g., the work of Vasyliūnas [2005a], and references therein).

(b) Neglect of electron inertia and pressure terms in equation (1) (but **not** necessarily in equation (2)), valid on length scales that are large compared to electron gyroradius.

(c) Neglect of all nonisotropic components of the pressure tensor (including viscous stresses) in equations (2) and (3).

Approximation (a) is the essence of the "B-V" approach; its consequence is that E and J must be calculated at a fixed time from equations (1) and (5), respectively, whereas **B** and V evolve in time according to the prognostic equations (4) and (2). As noted above, it applies on sufficiently large time and length scales, which excludes some small-scale ionospheric phenomena (e.g., the Farley-Buneman instability), but includes most aspects of M-IT coupling. Approximations (b) and (c) are made principally for simplicity. Furthermore, the electron-collision (resistive) term in equation (1) is significant only at the lowest altitudes (below approximately 100 km) and can be neglected in the magnetosphere and for many purposes in the ionosphere as well.

Equations (1)–(5) are the general governing equations for M-IT coupling. To be added are continuity and energy equations for each species, to determine densities and pressures, but their discussion is beyond the scope of this paper, in which the key terms of interest are the left-hand-side term in equations (2) and (4), representing the dynamic and the inductive aspects, respectively.

Intuitively, the inductive term in equation (4) should be negligible if the time-varying magnetic field either is sufficiently small (the argument is sometimes made that, in the ionosphere, the perturbed magnetic field is much smaller than the background field) or is varying on a sufficiently long time scale. However, equation (4) (the only equation in the entire set that is exact under all conditions and contains no approximation whatsoever) has only two terms, one on each side; they must, therefore, always be equal, and neither can be small compared to the other. To derive a criterion for neglecting the inductive effect, it is necessary to estimate the electric field from some other equation and compare it to the non-curl-free part of the electric field (which can be

calculated by standard techniques, most simply as the induced field in the Coulomb gauge) from equation (4). Vasyliūnas [2012] used the ionospheric Ohm's law together with Ampère's law to show that $\nabla \times \mathbf{E}$ can be considered zero if the magnetic field is changing on a time scale that is long compared to

$$\tau = \mu_0 \Sigma_p L, \tag{6}$$

where Σ_p is the Pedersen conductance (height-integrated conductivity), and L is the length scale of the curl. With $\Sigma_n = 10$ mho, $\tau \approx s$ for L = 200 km (local ionospheric scale), and $\tau \approx 15$ min for $L = 10 R_E$ (global magnetospheric scale). The time scale defined by equation (6) has been discussed in several M-I coupling contexts [Coroniti and Kennel, 1973; Holzer and Reid, 1975; Vasyliūnas and Pontius, 2007].

When the inductive effect is neglected, the electric field can be derived from an electrostatic potential. With further neglect of acceleration terms, the governing equations become

$$0 = \mathbf{J} \times \mathbf{B} - \rho \mathbf{v}_{in} (\mathbf{V} - \mathbf{u}_n) - \nabla p + \rho \mathbf{g},$$
(7)

$$\mathbf{E} = -\nabla\Phi \tag{8}$$

and

$$\nabla \bullet \mathbf{J} = 0 \tag{9}$$

where Φ is the electrostatic potential; if the resistive term in equation (1) is neglected, magnetic field lines are equipotentials. (Some coupling models introduce additional fieldaligned potential drops, mainly in relation to physics of the aurora; these do not change fundamentally the nature of the coupling as long as the potential remains electrostatic.) Equations (8) and (9) follow from equations (4) and (5), in this approximation.

The conventional ionospheric Ohm's law can be derived by combining equations (1) and (7) to eliminate V and also neglecting the pressure gradient and gravitational forces, to obtain

$$\mathbf{J} = \sigma_p(\mathbf{E}_{\perp} + \mathbf{u}_n \times \mathbf{B}) + \sigma_H(\mathbf{E} + \mathbf{u}_n \times \mathbf{B}) \times \mathbf{B}/B + \sigma_{\parallel} \mathbf{E}_{\parallel}, \qquad (10)$$

where $\sigma_p = \frac{\rho_i v_{in}}{B^2} \left(1 + \frac{v_{in}^2}{\Omega_i^2} \right)^{-1}$ is the Pedersen conductivity, $\sigma_H = -\frac{\rho_i \Omega_i}{B^2} \left(1 + \frac{v_{in}^2}{\Omega_i^2} \right)^{-1}$ is the Hall conductivity, and σ_{\parallel}

 $m_e v_e / N_e e^2$ is the parallel conductivity; see more detailed

Modeling the lonosphere-Thermosphere, edited by J. D. Huba, et al., American Geophysical Union, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/uml/detail.action?docID=1647732 Created from uml on 2018-04-05 12:43:20

discussion by Song et al. [2001]. Equation (10) superficially looks like (and is often described as) an Ohm's law in the conventional sense of linear relation between current and electric field. Physically, however, J given by equation (10) is not an ohmic current but a stress-balance current, a distinction emphasized by Vasyliūnas and Song [2005] and Vasyliūnas [2011, 2012]. The key assumption on which equation (10) depends is neglect in the plasma momentum equation of all forces except $\mathbf{J} \times \mathbf{B}$ and collisions with neutrals as well as of all acceleration terms; it does not depend on the electrostatic approximation. When pressure gradient and gravitational forces or dynamic terms are included, additional currents appear, in general, although some of the effects can be accommodated by redefining the conductivities as functions of frequency [Hughes, 1974] or even [Strangeway, 2012] by introducing a new quantity to replace the electric field. Such modifications preserve the mathematical form of equation (9) but not the physical meaning.

Equations (7)–(10) are the governing equations for electrostatic coupling. If the resistive term in equation (1) is assumed negligible, the Birkeland current cannot be calculated from equation (10) but must be determined from current continuity, equation (9), which relates perpendicular to field-aligned currents:

$$\frac{\partial J_{\parallel}}{\partial_{\parallel}} = -\nabla_{\perp} \bullet \mathbf{J}_{\perp} = -\nabla_{\perp} \bullet [\sigma_{p}(\mathbf{E} + \mathbf{u}_{n} \times \mathbf{B}) + \sigma_{H}(\mathbf{E} + \mathbf{u}_{n} \times \mathbf{B}) \times \mathbf{B}/B].$$
(11)

The electric field between the highly collisional and the collisionless regions is connected by equation (8) along the magnetic field lines, an assumption that is not valid when the inductive effect is significant. Equations (8) and (11) now replace equations (4) and (5) to provide long-range coupling for quasisteady state electrostatic M-I coupling models. Some models introduce additional empirical or semiempirical terms to describe the effects of precipitation particles and Birkeland currents on the conductivities, but these modifications do not change the electrostatic nature of the coupling.

Similarly to the derivation of equation (10), equations (1) and (7) can also be combined to eliminate \mathbf{J} instead of \mathbf{V} to obtain

$$0 = e(\mathbf{E} + \mathbf{V} \times \mathbf{B}) - m_i \mathbf{v}_{in} (\mathbf{V} - \mathbf{u}_n), \qquad (12)$$

an equation that is very useful in relating plasma flow and electric field. It has, however, caused considerable confusion: contrary to a common interpretation, the electric field does not produce motion of the plasma but is the result of the motion, linking the electron and ion flows to preserve charge quasineutrality. The description that the electric field produces plasma motion is fundamentally incorrect [*Buneman*, 1992; *Vasyliūnas*, 2001, 2011, 2012; *Tu et al.*, 2008; *Song et al.*, 2009], although it is convenient sometimes and has been used for a long time, as argued by *Kelley* [2009]. When an (external) electric field is imposed onto a plasma, an internal electric field is produced by charge separation, as described in the first chapter of any plasma physics textbook. If the plasma is sufficiently dense, the internal electric field will largely cancel out the imposed one, leaving nearly no electric field inside the plasma.

2.2. Differences in Physical Description of the Coupling Processes

It may be instructive to first consider the coupling process qualitatively in more detail. The time-varying interaction between the solar wind, the magnetosphere, and the ionosphere proceeds primarily by continual creation and relaxation of imbalances between the various force terms in the momentum equation (2). Force imbalance produces plasma motions, which affect all the other quantities, creating a perturbation that can be described to first order as a set of waves, which propagate (and possibly reflect at boundaries, depending on boundary conditions), thereby relaxing the imbalance. The essential difference between electrostatic and dynamic-inductive coupling is that the former neglects all wave effects, whereas the latter explicitly includes them.

When a model, such as SAMI 2 [*Huba et al.*, 2000], includes the dynamic term only but neglects the inductive term, it will contain sonic and gravity waves, produced by force imbalances associated with pressure gradient and gravity terms in equation (2). The sound waves are isotropic, and the gravity waves act primarily in the vertical direction. The two wave modes can be coupled, particularly when the magnetic field is inclined. The wave amplitude increases with altitude unless strong damping is present.

When inductive and dynamic terms are both included in a model, imbalances associated with all the forces, which now include the electromagnetic force, produce MHD waves. With gravity neglected, in a uniform MHD fluid, this gives rise to three wave modes: slow, intermediate or Alfvén, and fast modes. (Note: the term "Alfvén wave," used here for the intermediate mode, is sometimes applied to MHD waves in general.) Different modes not only have different propagation speeds but also different perturbation relations, which can be used to identify the wave mode [*Song et al.*, 1994], as well as to associate particular modes with specific types of changes. The fast mode is relatively more isotropic, i.e., it can easily mitigate the force imbalances both parallel and perpendicular to the magnetic field. The slow and Alfvén modes are more anisotropic and more effective in the

direction along the magnetic field. Obliquely propagating Alfvén waves are the only ones (at least in a homogeneous medium) that can carry field-aligned (Birkeland) currents. Between the magnetosphere and the ionosphere, where the plasma beta (ratio of thermal pressure to magnetic pressure) is small, the fast and Alfvén modes are the most effective, and the Alfvén mode is the dominant mechanism in M-I coupling where Birkeland currents play an essential role. For further discussion of the functions of various modes, see the work of Song and Vasyliūnas [2010]. For example, the fast mode, since it is the only mode (at least in a homogeneous medium) that propagates in the direction perpendicular to the magnetic field, can directly communicate changes from the subsolar magnetopause to the low-latitude ionosphere. In the high-altitude ionosphere, the fast mode is most efficient in mitigating horizontal force imbalances. Since the fast mode speed is here nearly the same as the Alfvén speed, $\sim 10^3$ to 10^4 km s⁻¹, magnetospheric changes are felt nearly instantaneously throughout the whole polar cap as well as into the closed field line regions, for example. The ionosphere can then, in turn, influence other magnetospheric regions where direct intermagnetospheric propagation, especially perpendicular to the magnetic field, might take longer (an example is discussed at the end of section 2.2).

As a simple specific illustration, consider the canonical problem of M-I coupling: plasma flow imposed in the outer magnetosphere (or at the magnetopause). In electrostatic coupling, this is described as an imposed electric field, with the plasma flow calculated from it (E-J description); the electric potential in the magnetosphere (constant along a magnetic field line) is directly mapped into the ionosphere (at every height). As noted in sections 1 and 2.1, this description is not valid in dynamic processes because the inductive effect can no longer be neglected.

In dynamic-inductive coupling, if (according to the B-V description) the electric field is not the primary agent to couple the magnetosphere with the ionosphere/thermosphere, how is coupling achieved? i.e., when the magnetospheric motion changes, how does the ionosphere know about it, and why does the ionosphere/thermosphere move accordingly? (particularly nonintuitive if the magnetospheric motion is entirely parallel to the ionosphere and perpendicular to the magnetic field). Figure 1 (adapted from the discussion by Song and Vasyliūnas [2011] of an analogous problem in the solar atmosphere) schematically illustrates the process in a simple 1-D stratified ionosphere/thermosphere with a vertical magnetic field. The plasma and neutrals are initially (t = 0) at rest (or at a common background velocity). A (horizontal) plasma motion (not an electric field!) is imposed on the top boundary at t = 0+. With the plasma inside the domain linked by the frozen-in magnetic field, the plasma motion creates a kink on the magnetic field line at the top boundary, which exerts a tension force on the plasma below the boundary and makes it move, creating another kink at the lower altitude, and so on. The net result is a magnetic perturbation front propagating downward along the magnetic field line at the Alfvén speed V_A . Since the neutrals are not subject to the electromagnetic force, a velocity difference between plasma and neutrals is produced as the perturbation front passes. The relative motion causes momentum transfer through ion-neutral collisions, accelerating the neutrals. This is the primary mechanism of momentum coupling from the magnetosphere to the ionosphere. The collisions generate frictional heating of plasma and neutrals and also tend to slow down the plasma motion,



Figure 1. Schematic illustration of the plasma and neutral motion in a partially ionized and magnetized fluid. Closed red dots indicate ions and, open circles, neutrals. The imposed plasma motion at the top boundary produces a magnetic field perturbation that propagates downward, which in turn drives the motion of the plasma further below. Adapted and modified from *Song and Vasyliūnas* [2011].

which, however, continues as long as the motion at the top boundary is sustained. On time scales shorter than $1/v_{in}$, where v_{in} is the ion-neutral collision frequency, the plasma is little affected by the neutrals, and on time scales longer than $t \sim 1/v_{ni}$, where v_{ni} is the neutral-ion collision frequency, the neutrals are catching up with the plasma. In the absence of other forces acting on the neutrals, the system will eventually $(t \rightarrow \infty)$ reach a state in which the plasma and neutrals have a common velocity equal to the driving velocity. The motion of the plasma creates an electric field, and the distortion of the magnetic field corresponds to a current.

This sequence, depicted in Figure 1, goes from an initial state of rest, through a transient phase of plasma-neutral flow difference, to a final state of equal plasma-neutral flow, all a consequence of an Alfvén wave launched by imposing a small velocity step at the boundary. A small step function can be decomposed into a spectrum of waves (as noted in section 1), and any time profile of perturbations at the upper boundary, including oscillations, can be built up as a series of such small steps. If the perturbations continue without end but always vary more slowly than the neutral-plasma collision time, the system remains in a slowly varying quasisteady state, with a small but nonzero velocity difference between plasma and neutrals and consequent heating by collisions.

The above description is, of course, highly simplified. First, since the perturbation at the top boundary is velocity and not displacement, the field line in the second panel of Figure 1 bends more gradually than is shown in the figure. Second, if the magnetic field is inclined at an angle, there will be a component propagating horizontally, which produces compressible modes. Third, with strictly vertical field and horizontal velocity perturbations, for field-aligned propagation vector, the Alfvén mode and the (compressional) fast mode are indistinguishable. Fourth, horizontal nonuniformity, either in the medium or in the driver, may also generate compressible modes. Velocity shear and field-aligned flow will create additional effects.

In inductive-dynamic coupling, the ionosphere is not merely a passive recipient of driving forces from the magnetosphere but can also impose motions on the magnetosphere, by the process described above but proceeding in the reversed direction in a similar manner to that in the solar chromosphere [Song and Vasyliūnas, 2011]. Because the horizontal distances in the ionosphere are much smaller than spatial scales of the magnetosphere, and communication across them is at the fast-mode speed, which in the upper parts of and above the ionosphere is faster than the Alfvén speed in the magnetosphere, it is possible for an imposed magnetospheric change, after propagating to the ionosphere along the local field lines, to propagate quickly across field lines to different regions of the ionosphere, setting the ionosphere there into motion accordingly. This ionospheric motion then propagates back to the magnetosphere at the Alfvén speed and modifies the plasma flow in other magnetospheric regions that have not yet been affected by direct signals from the initially imposed change, as shown in Figure 2.



Figure 2. Illustration of the active role of the ionosphere in affecting magnetospheric convection. (left) magnetospheric plasma on open field lines convects along the magnetopause at the Alfvén velocity over a long distance. (right) the corresponding ionospheric feet of the field lines move at fast mode speed over a short distance, imposing motion on adjacent closed field lines. The motion of the ionosphere propagates to the magnetosphere at the Alfvén speed and modifies magnetospheric convection on closed field lines.

2.3. Effect of Reflection

The discussion in section 2.2 provides a basic understanding of how the plasma and neutrals respond in a time-dependent fashion to the solar wind and/or magnetospheric driver, but it describes only the initial phase, with one-way propagation of the perturbation. Afterward, the density gradients of the ionosphere/thermosphere (and to a lesser degree the gradient of the magnetic field) cause partial reflection of the downward propagating perturbation, modifying the basic picture shown in Figure 1. The net result is that the wave is reflected back to the magnetosphere, where it can be re-reflected at the magnetopause or at the conjugate ionosphere. The system, thus, experiences multiple reflections before reaching its steady state. Such a multiple-reflection model has been previously studied by Holzer and Reid [1975], who invoked surface reflection from a thin (height-integrated) ionosphere. The actual reflection process is undoubtedly more complicated; since ionospheric parameters vary gradually, reflection should not occur at a simple boundary surface but in a continuous manner over height. As the incident perturbation penetrates deeper down into the ionosphere, the effective wave propagation speed becomes lower, due to the neutralinertia-loading effect [Song et al., 2005], and the velocity perturbation decreases drastically at and below the altitude where the ion-neutral collision frequency becomes greater than the ion gyrofrequency, and thus, only a small velocity difference is needed if the collision term in equation (2) is to balance the $\mathbf{J} \times \mathbf{B}$ force. That wave reflection is affected by these processes and adds a key ingredient to time-dependent M-IT coupling is an important lesson from our analyses [Song et al., 2009; Tu et al., 2011].

With reflection present, a perturbation of any quantity (velocity, magnetic field, electric field, Poynting vector, etc.) will include contribution from both incident and reflected waves, having different phases. For velocity and magnetic field (the key parameters of inductive-dynamic coupling)

$$\mathbf{V} = \mathbf{V}_{inc} + \mathbf{V}_{ref}, \quad \delta \mathbf{B} = \delta \mathbf{B}_{inc} + \delta \mathbf{B}_{ref}.$$
(13)

The velocity and the magnetic perturbations locally are connected by an equation specific to the wave mode, which in regions where collision effects and the Hall term can be neglected takes the simple form of the Walén relation $\delta \mathbf{B} = \pm \mathbf{V}B/V_A = \pm \sqrt{\rho\mu_0}\mathbf{V}$, derived from equations (1), (2), (4), and (5). The negative (positive) sign is for propagation parallel (antiparallel) to the background magnetic field. In general, it is extremely difficult to separate reflected from incident perturbations, either in simulations or in observations, particularly if the reflection occurs in a gradual manner. The reflected perturbation may either enhance or reduce the strength of the incident perturbation, depending on the phase delay. The net perturbation (incident plus reflected) is, thus, expected to be more variable than the incident perturbation alone. Since the ionosphere, as discussed above, acts primarily to reduce the net velocity perturbation, the reflected velocity will tend to subtract from the incident one, and therefore (because of the sign change in the Walén relation), the reflected magnetic field will tend to add to the incident one. This is why reflection may produce overshoots in the magnetic field perturbations, which in turn result in enhanced currents, profound consequences for evaluating the power and heating rate (proportional to the square of the perturbation amplitude). (All these effects, it goes without saying, are absent in the electrostatic models, which contain neither propagation nor reflection.)

2.4. Differences in Physical Quantities

Below, we compare electrostatic coupling described by equations (7)–(11) and inductive-dynamic coupling described by equations (1)–(5), for an initial velocity perturbation V_0 perpendicular to the magnetic field B_0 , at a location L_0 away from the ionosphere. Note that collision effects are dominant below 130 km [*Song et al.*, 2001]. A summary of the comparison is given in Table 1.

2.4.1. Communication speed along a magnetic field line. In the absence of field-aligned bulk flow, communication in long-range coupling is accomplished primarily via electromagnetic coupling described by Maxwell's equations. In quasisteady state models with electrostatic coupling (equation (8)), the coupling is treated as instantaneous as far as the model is concerned (formally, the communication speed is infinite, although it is, of course, admitted that, in reality, the magnetosphere communicates with the ionosphere at the Alfvén speed). For inductive-dynamic coupling, it follows explicitly from equations (2) and (4) that the coupling (communication) speed is the Alfvén speed. For slow variations when $\omega < v_{in}$, the neutral wind affects the waves, by a process referred to as neutral-inertia-loading process [Song et al., 2005]; when collisions are important, the fluid appears heavier than the plasma alone. For perturbations at frequencies less than the ion-neutral collision frequency, the propagation speed decreases continuously with perturbation frequency, and when the frequency becomes less than the neutral-ion collision frequency v_{ni} , approaches $\alpha^{1/2} V_A$ (where $\alpha = \rho_i / (\rho_i + \rho_n)$ is the ionization fraction), i.e., Alfvén speed calculated from total (plasma plus neutrals) mass density. For periods of interest, $>1 \text{ min}^{-1}$. the inertia loading effect becomes appreciable below 400 km [Song et al., 2005]. With decreasing altitude, as

 Table 1. Comparison of Steady State Coupling and Dynamic Coupling

Quantity	Electrostatic Coupling	Dynamic Coupling
Communication V_{phase} Transient time δt	$\stackrel{\infty}{0}$	$\delta t = \begin{cases} V_A \sim \alpha^{1/2} V_A \\ t_A & 2 \min \text{ (Alfvén time)} \\ 10 \sim 15 t_A & 30 \min \text{ (magnetosphere-} \\ & \text{ionosphere equilibrium)} \\ 1/\gamma_{ni} & 1 \sim 3h \text{ (neutral accel)} \end{cases}$
Reflection	No reflection, $\mathbf{V} = \mathbf{V}_{inc}$, $\delta \mathbf{B}$ not consistent with \mathbf{V}	total = $I + R$ for both $\delta \mathbf{B}$ and \mathbf{V} , relative sign $\delta \mathbf{B}/V$ changes on reflection $\mathbf{V} = \mathbf{V}_{inc} - \mathbf{V}_{ref} \ \delta \mathbf{B} = \pm \delta \mathbf{B}_{inc} \pm \delta \mathbf{B}_{ref}$
Velocity perturbation V	$V = V_0 B_0 D_0 / BD$	$V = (\rho_0/\rho)^{1/4} V_0$ for single Alfvén wave parallel propagation
Magnetic perturbation $\delta \mathbf{B}$	not considered as part of model evolution, calculated from J, not consistent with V	$\delta \mathbf{B} = \pm \mathbf{V} B / V_A = \pm \sqrt{\rho \mu_0} \mathbf{V}$ $ \delta \mathbf{B} = (\rho \rho_0 \mu_0^2)^{1/4} V_0 = (\rho / \rho_0)^{1/4} \delta \mathbf{B}_0 $ for single Alfvén wave parallel propagation
Electric field E	$E = E_0 D_0 / D$	$E = (\rho_0/\rho)^{1/4} (B/B_0)^{1/2} E_0$ for single Alfvén wave parallel propagation
Current J Heating rate q	$\mathbf{J}_{\perp} = \overleftarrow{\boldsymbol{\sigma}} \cdot (\mathbf{E} + \mathbf{u}_n \times \mathbf{B}), \nabla \cdot \mathbf{J} = 0$ $q = \sigma_p (\mathbf{E} + \mathbf{u}_n \times \mathbf{B})^2$	$\mathbf{J} = \nabla \times \mathbf{B}/\mu_0, \text{ resulting in current conservation}$ $q = \mathbf{J} \cdot (\mathbf{E} + \mathbf{V} \times \mathbf{B}) + v_{in}\rho(\mathbf{u}_n - \mathbf{V})^2$
	$= \left[\left(1 + \frac{\mathbf{v}_{in}^2}{\Omega_i^2} \right)^{-1} + \frac{\Omega_i \Omega_e}{\mathbf{v}_e \mathbf{v}_{in}} \left(\frac{\mathbf{v}_{en}}{\mathbf{v}_e} + \frac{\Omega_e^2}{\mathbf{v}_{en} \mathbf{v}_e} \right)^{-1} \right] \\ \times \mathbf{v}_{in} \rho(\mathbf{E}/B + \mathbf{u}_n \times \mathbf{b})^2 \approx \mathbf{v}_{in} \rho(\mathbf{V} - \mathbf{u}_n)^2$	$\approx \left(1 + \frac{\mathbf{v}_e \mathbf{v}_{in}}{\Omega_e \Omega_i}\right) \mathbf{v}_{in} \rho \langle (\mathbf{V} - \mathbf{u}_n)^2 \rangle$ $-\frac{1}{2} \left[\Delta(\rho V^2) + \Delta(\rho_n u_n^2)\right] / \Delta t$
Poynting vector S	Not considered explicitly, DC part included implicitly in dissipation	$\mathbf{S} = \frac{1}{2} \frac{\delta B^2}{\mu_0} \mathbf{V}_A = \frac{1}{2} \rho V^2 \mathbf{V}_A \text{ (single wave)}$ Fluctuating component only, DC part may be a problem

the ion-neutral collision frequency increases rapidly and the ionization fraction decreases, the inertia-loading effect can reduce the propagation speed by an order of magnitude.

2.4.2. Transient time. The original formulations of M-I coupling theory [Fejer, 1964; Vasyliūnas, 1970; Wolf, 1970] in terms of quasisteady state electrostatic models were based on the assumption that time scales of interest (in particular, flow times of magnetospheric convection, as stated explicitly by Vasyliūnas [1970]) are long compared to wave propagation times within the system. The governing equations (1), (7)–(11) of electrostatic coupling do not contain any time derivatives and, thus, in principle, describe self-consistently only the time-independent solutions for fixed given boundary conditions and system parameters. In practice, however, models are produced (particularly when implemented as numerical simulations) that do evolve with time. This is done by allowing the boundary conditions and/or system parameters to vary with time, e.g., introducing temporal

variations at the boundaries of global MHD simulation models, from measurements, or from semiempirical models such as AMIE [Richmond and Kamide, 1988]. This approach has been extremely successful in describing many processes in the ionosphere/thermosphere. Nevertheless, the model output still is only a time series of successive but disjointed steady state solutions (the model temporal variations reflecting merely the variability of the input, not the dynamics of coupling) and cannot be considered an adequate approximation to the evolving solutions of the time-dependent equations unless the effective time step of the model is much longer than the time scales of all the transient processes. (It is an illusion to suppose that such models are self-consistently time-dependent and that more dynamic effects can be derived by using higher time resolution data as input or smaller and smaller time steps of the simulations.)

In the electrostatic coupling models, with all transients neglected and perturbations in the magnetosphere mapped instantaneously along field lines to (every altitude of) the ionosphere, the only intrinsic time scales are two time scales associated with collision processes. The ionosphere affects the thermosphere on a time scale $1/v_{ni}$, where v_{ni} is the neutral-ion collision frequency, and the thermosphere affects the ionosphere on a time scale $1/v_{in}$, where v_{in} is the ion-neutral collision frequency. Note that v_{in} and v_{ni} are different quantities; the two collision frequencies are related by momentum conservation during collisions, $\rho_n v_{ni}$ = $\rho_i v_{in}$. As the neutral density is much higher than the plasma density below 800 km [e.g., Tu et al., 2011], it takes a very long time before the neutral wind is affected by the plasma motion, although the plasma motion can be affected by the neutral wind relatively quickly. Neutral dynamics can, therefore, be included in a quasisteady state treatment of the ionosphere. (For plasma, collisions are important when v_{in} is comparable or much greater than the ion gyrofrequency Ω_{i} .)

In inductive-dynamic coupling, there are three intrinsic time scales: the Alfvén (communication) time, the time for the M-I system (thermosphere not included!) to reach equilibrium or a quasisteady state, and the time for the complete M-IT system to reach equilibrium.

Neglecting collisions, the coupling time from equations (2), (4), and (5) is the Alfvén time $t_A = L/V_A$ during which the magnetic perturbation propagates from the magnetosphere to the ionosphere over a distance *L* along the magnetic field. The Alfvén time, or magnetosphere-ionosphere communication time, is well known and has been considered by some to be the transient time for M-I coupling. The time for M-I equilibrium, however, is strongly affected by collisions and consequent wave reflection, as discussed in section 2.3.

The condition for M-I equilibrium state is negligible time derivative in equation (2), which implies a current (required to balance the collision term) $J = v_{in}\rho_i V/B$ and associated magnetic perturbation $|\delta \mathbf{B}| = \mu_0 \Delta L J = \mu_0 \Delta L v_{in}\rho_i V/B$ (ΔL is the altitude range over which the current is distributed, and all quantities are evaluated in the ionosphere). Taking the curl of the electric field from the magnetopause to the ionosphere gives, from equations (4) and (1), the time required to produce this magnetic perturbation, in order of magnitude,

$$\delta t \sim \frac{|\delta \mathbf{B}|L}{\Delta E} \sim \left(\frac{\Delta L \Omega_i}{V_A}\right) \left[\left(\Omega_i / \nu_{in}\right)^2 + 1\right]^{-1/2} t_A \tag{14}$$

(essentially the same time scale as in equation (6)). The dimensionless coefficient of t_A in equation (14) is determined by parameters in the high-collision region and is small when the collision frequency is small. Using typical values of the ionosphere [e.g., *Song et al.*, 2001] gives the

transient time for dynamic processes of order of 10~15 t_A , consistent with values quoted by Song et al. [2009] and Tu et al., [2011] as well as those obtained by wave analysis [Lysak and Dum, 1983]. This is substantially longer than the Alfvén time, indicating that the waves have to reflect several times before the system reaches equilibrium (in agreement with the discussion in section 2.3). For $L \sim 20$ R_E and $V_A \sim 1000$ km s⁻¹, the Alfvén time is 120 s, and the transient time is about 20~30 min; the latter is the observed time scale for many important M-IT processes, such as the substorms and magnetospheric transients [e.g., Earle and Kelley, 1987; Russell and Ginskey, 1993, 1995; Zesta et al., 2000; Bristow et al., 2003; Huang et al., 2008, 2010]. The time scale (14) is obtained by self-consistently joining longrange coupling with ionosphere-thermosphere local coupling and cannot be derived from local coupling alone such as that modeled by Strangeway [2012]. The fundamental reason for this additional time scale is the inertia of the ionosphere, irrelevant when the ionosphere is treated as being in a steady state.

The third time scale, the time for the M-IT system to reach equilibrium, in inductive-dynamic coupling as well as in electrostatic coupling, is the neutral wind acceleration time which is $1/v_{ni}$ (minimum value in the *F* layer, of an order of a few hours) [e.g., *Song et al.*, 2009]. On time scales longer than a significant fraction of this time, the neutral acceleration cannot be neglected. On still longer time scales, the neutrals and the plasma are tightly coupled, and the three fluids then can be considered a single fluid with collisions as internal dissipation [*Song and Vasyliūnas*, 2011]. The neutral dynamic effects are included in most ionospheric-thermospheric models but neglected in most M-I coupling models.

In the following discussion of magnitudes of the physical quantities, estimates from electrostatic coupling (in equations (15) and (18)) are applicable only to times much longer than the second (M-I equilibrium) transient time scale. For inductive-dynamic processes, estimates in equations (16), (17), and (19) are for a single (incident) wave and thus apply to times below and up to the first (Alfvén-wave communication) transient time scale. For times between the first and the second transient time scales, the quantities become sums of incident and reflected waves and are no longer equal to the values given by the equations. For times much longer than the second transient time scale, the results from inductivedynamic coupling (after multiple reflections) are essentially the same as those from electrostatic coupling. (All the estimates here are for weakly collisional regions and do not include damping and neutral inertia loading effects, thus do not apply in the lower part of the ionosphere, say, below 130 km.)

Modeling the lonosphere-Thermosphere, edited by J. D. Huba, et al., American Geophysical Union, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/uml/detail.action?docID=1647732. Created from uml on 2018-04-05 12:43:20.

2.4.3. Velocity mapping in weakly collisional regions. In electrostatic coupling, the velocity variation along the magnetic field is mapped according to the electric potential mapping (equation (8)), or

$$V = E_0 D_0 / BD = (B_0 D_0 / BD) V_0 \sim (B_0 / B)^{1/2} V_0, \quad (15)$$

where D and D_0 are the distance between two field lines in the direction perpendicular to the flow and the (background) magnetic field at two heights. In the ionosphere, since the magnetic field strength does not change significantly over height, the flow velocity does not change as long as it is not affected by collisions. The last expression in equation (15) is an approximation, derived from magnetic flux conservation by ignoring the dependence of the ratio D/D_0 on orientation relative to the magnetic meridian.

In inductive-dynamic coupling, on the other hand, for a single (incident) wave, the velocity may be determined from energy conservation when collisions are weak. The Poynting flux of the perturbations contained in a flux tube

is $F = SA \propto S/B = \frac{1}{2} \frac{|\delta \mathbf{B}|^2}{\mu_0} \frac{V_A}{B} = \frac{1}{2} \rho V^2 \frac{V_A}{B}$, where A is the cross section of the flux tube, and the last expression follows from the Walén relation. Energy conservation here implies that F is constant, which gives

$$V = (\rho_0 / \rho)^{1/4} V_0. \tag{16}$$

Note the completely different velocity variation along a field line given by the two descriptions, equations (15) and (16). The density dependence of the velocity (slower flow in regions of higher density) during the dynamic stage in inductive-dynamic coupling can be understood as due to plasma inertia.

2.4.4. Magnetic perturbation. In electrostatic coupling, the magnetic perturbation does not play any role in the calculation but is derived from Ampère's law (equation (5)) in order to compare with observations. In inductive-dynamic coupling, the magnetic perturbation for a single wave (in a collisionless or at most weakly collisional region) can be related locally to the velocity perturbation according to the Walén relation discussed in section 2.3. Combined with equation (16), this gives the mapping relation for the magnitude of the magnetic perturbation

$$|\delta \mathbf{B}| = (\rho \rho_0 \mu_0^2)^{1/4} V_0 = (\rho / \rho_0)^{1/4} |\delta \mathbf{B}_0|$$
(17)

showing that the magnetic perturbation tends to increase in regions of higher density, also a plasma inertia effect. (Note: the increase at very low frequencies of the effective ion density in the expression by adding part or all of the neutral density due to neutral inertia loading [*Song et al.*, 2009; *Song and Vasyliūnas*, 2011] becomes appreciable only at low altitudes where the Walén relation is no longer a good approximation.)

2.4.5. *Electric field mapping along a field line*. In electrostatic coupling, the electric field variation along the magnetic field is calculated from the potential,

$$E = E_0 D_0 / D \sim (B/B_0)^{1/2} E_0, \qquad (18)$$

where D is the perpendicular distance between two field lines (defined the same way and with the same approximations as in equation (15) for the velocity).

The electric field in inductive-dynamic coupling is calculated from the generalized Ohm's law (1), which gives

$$E = (\rho_0/\rho)^{1/4} (B/B_0)^{1/2} E_0, \tag{19}$$

if electron collisions and Hall effect are neglected, and the velocity is taken from equation (16). Again, the plasma inertia effect results in additional density dependence.

2.4.6. Electric current. In electrostatic coupling, the current is calculated from Ohm's law (equation (10)), with the electric field either given by some other model or obtained by solving equation (11) with the parallel current derived from a global MHD simulation model. The derived temporal variation of the current is intrinsically in conflict with the assumption that the magnetic field is constant. In inductive-dynamic coupling, the current is derived from equation (5), using the calculated magnetic perturbation. The Birkeland (magnetic field-aligned) current is determined automatically by this procedure; no special process or time delay to connect the horizontal currents in the ionosphere with Birkeland currents is needed.

2.4.7. Poynting vector: The importance and measurements of the Poynting vector have been promoted in recent years in the community, which may be considered progress in the understanding of the coupling processes. There are, however, two distinct forms of the Poynting vector. One, the "DC" form is calculated from the (quasi)steady electric and magnetic fields and is obviously the only form applicable in the electrostatic coupling approach. In this form, the Poynting vector is always perpendicular to the (background) magnetic field and, in general, flows into or out of the ionosphere in regions far away from Birkeland currents, even when the energy input to the ionosphere is concentrated in regions where the Birkeland currents are located. The other, the "AC" form, is calculated from perturbation electric and magnetic fields and obviously can be described only in the inductive-dynamic approach. Much of the recent observational evidence for enhanced Poynting-flux energy input into the ionosphere refers to this form.

In inductive-dynamic coupling, the (AC) Poynting vector is $S = \frac{1}{2\mu_0} |\delta \mathbf{B}|^2 V_A = \frac{1}{2} \rho V^2 V_A$ for fluctuating perturbations that obey the Walén relation and propagate all in the same direction along the background magnetic field (which is now, in contrast to the DC form, the direction of the Poynting vector); the factor O comes from averaging over fluctuations. If there are perturbations propagating in both directions (as with incident and reflected waves), the total Poynting vector, being quadratic in the perturbation amplitudes, contains also cross terms of incident and reflected waves. The sign difference between the two in the Walén relation, however, ensures that the cross terms always cancel, independent of any amplitude or phase difference. The calculated (AC) Poynting vector, thus, represents the net flow (up minus down) of electromagnetic energy along the direction of the background magnetic field, as long as the assumptions concerning the Walén relation hold. Energy carried by incident or reflected waves flows down or up, respectively; hence, the measured Poynting vector, which represents the superposition of incident and reflected, is subject to the same uncertainties and ambiguities discussed above for other quantities.

The Poynting vector can be derived from local measurements of the magnetic field and electric field or plasma velocity, and its direction determined from the phase difference between the magnetic and electric or velocity perturbations. The energy flow is downward if the magnetosphere is driving the ionosphere and upward if otherwise. As discussed in section 2.2, ionospheric regions (e.g., low latitudes) that are not directly driven by the external source of magnetospheric convection can still be driven, via fast mode propagation, by ionospheric convection in other regions (e.g., high latitude and polar cap). This secondary ionospheric convection can then locally drive the magnetosphere, a process that sometimes is misleadingly called "penetration electric field." In these regions, the Poynting vector can be upward, similarly to a proposed process for chromospheric heating [Song and Vasyliūnas, 2011]. The above description assumes that the driving accelerates the plasma flow. If the driving decelerates the flow instead, e.g., due to a reversal of the IMF direction, the Poynting vector is reversed (often described as a "flywheel effect").

2.4.8. Heating rate. The local heating rate in electrostatic coupling is

$$q = \mathbf{j} \cdot (\mathbf{E} + \mathbf{u}_n \times \mathbf{B}) = \sigma_p (\mathbf{E} + \mathbf{u}_n \times \mathbf{B})^2.$$
(20)

This is often referred to as Joule heating, on the basis of the definition of Joule heating as $\mathbf{j} \cdot \mathbf{E}^*$ where \mathbf{E}^* is an electric field. Note, however, that the electric field and, hence, also the Joule heating so defined varies with the chosen frame of reference, whereas the true heating rate does not. There are several frames of reference in our problem: Earth, ionospheric plasma, and thermospheric wind. The conventional treatment chooses the neutral wind frame. Vasyliūnas and Song [2005] argued that the Joule heating, as the mechanism that converts electromagnetic energy to thermal energy, should properly be defined in the plasma frame of reference, and they showed that heating given by equation (20) is predominantly collisional and not electromagnetic. That most of the heating occurring in M-IT coupling is frictional and not Joule heating was confirmed by Tu et al. [2011]. To obtain the heating rate (equation (20)) observationally, one first uses the measured plasma velocity to calculate the electric field, then the electric potential, the conductivities, the current... etc., each step involving uncertainties and approximations. For weak collisions, the heating rate is

$$q = \left[\left(1 + \frac{\mathbf{v}_{in}^2}{\Omega_i^2} \right)^{-1} + \frac{\Omega_i \Omega_e}{\mathbf{v}_e \mathbf{v}_{in}} \left(\frac{\mathbf{v}_{en}}{\mathbf{v}_e} + \frac{\Omega_e^2}{\mathbf{v}_{en} \mathbf{v}_e} \right)^{-1} \right] \times \mathbf{v}_{in} \rho(\mathbf{E}/B + \mathbf{u}_n \times \mathbf{b})^2 \approx \mathbf{v}_{in} \rho(\mathbf{V} - \mathbf{u}_n)^2.$$
(21)

The local heating rate for inductive-dynamic heating is from *Vasyliūnas and Song* [2005],

$$q = \mathbf{J} \cdot (\mathbf{E} + \mathbf{V} \times \mathbf{B}) + v_{in} \rho_i (\mathbf{u}_n - \mathbf{V})^2.$$
(22)

For weak collisions and averaged over a long period of time, the heating rate is

$$q \approx \left(1 + \frac{\mathbf{v}_{e} \mathbf{v}_{in}}{\Omega_{e} \Omega_{i}}\right) \mathbf{v}_{in} \rho_{i} \langle (\mathbf{V} - \mathbf{u}_{n})^{2} \rangle - \frac{1}{2} [\Delta(\rho_{i} V^{2}) + \Delta(\rho_{n} u_{n}^{2})] / \Delta t \sim \mathbf{v}_{in} \rho \langle (\mathbf{V} - \mathbf{u}_{n})^{2} \rangle.$$
(23)

The heating rates from the two coupling approaches are superficially similar but differ in several respects. First, equation (21) is valid only for quasisteady state and, therefore, does not include heating associated with waves, whereas equation (23) does include wave heating, as indicated by the average signs. In inductive-dynamic coupling, the velocity includes both incident and reflected perturbations; due to reflection, the magnitude of the velocity can be significantly larger than its steady state value during the dynamic period because the average of the velocity perturbations can be close to zero. Second, the physical picture is much clearer for the latter approach: heating is primarily generated by the frictional heating between the plasma and the neutrals. If the velocity difference between the two is measured or otherwise known, the heating rate can be derived relatively easily.

3. CONCLUSIONS AND DISCUSSION

To model global M-IT coupling processes on time scales shorter than about 30 min, the dynamic term in the momentum equation and the inductive term in Faraday's law need to be included. Although the E-J and B-V schemes give the same results in electrostatic and quasisteady state coupling, time-dependent processes can be dealt with only with the B-V scheme because the E-J scheme cannot describe the time evolution self-consistently.

There are major differences between the conventional M-I and M-IT coupling approaches and the inductive-dynamic approach. First, the transition times are estimated differently. Conventionally, the Alfvén time between the magnetosphere and ionosphere or from one hemisphere to the other is considered the transition time scale of dynamic coupling. However, a fully coupled M-IT system cannot complete its transition in one Alfvén time because the ionosphere has its own inertia, and longer time is needed to accelerate the magnetosphere- ionosphere system from one state to another. This transient time is about 10~15 times Alfvén time or typically 20~30 min.

Second, the propagation and reflection of waves, a most important phenomenon during the transition, is not included in the conventional theory. (The term "waves" here refers generally to any type of perturbations.) This process cannot be resolved within the ionosphere/thermosphere if it is treated as a height-integrated layer. This issue is very confusing and has often been overlooked or misunderstood. A question one may ask is: Why is the wave reflection so important? Or, how much does a model miss without including wave reflection? In the presence of incidence and reflection, a physical quantity at a given location and time is the superposition of the two. Depending on the reflection coefficient and the phase difference between the incident and reflected perturbations, the resultant (measured) perturbation can range from as small as zero to as large as twice the incident amplitude both positively and negatively. For a quantity such as energy or heating rate, proportional to the square of the amplitude, the difference can range from zero to a factor 4. Furthermore, because the phase difference between incidence and reflection may vary with time and location, the resulting (measured) signals are actually fluctuating within this range. It is, thus, easy to understand why the observations of a dynamic process (e.g., a substorm) often contain many large-amplitude fluctuations or overshoots, which then may diminish over a transition period after conditions stabilize. The fluctuations can, of course, be averaged out over the transition period, but these intense fluctuations (such as those during substorms or auroral brightenings) are often of great physical significance and constitute the subject of the study. Averaging them out is tantamount to eliminating the phenomenon of primary interest.

Third, the conventional theory does not appropriately describe the ionospheric dynamic responses to the magnetosphere and the coupling within the ionosphere. In fact, because the magnetic field variations in the ionosphere is assumed time constant and the plasma dynamic terms in the perpendicular direction are not considered, MHD modes in the ionosphere itself do not play a role in the theory. In particular, the fast mode, the most effective mechanism to mitigate force imbalance in the horizontal direction, to communicate efficiently between different parts of the ionosphere and to produce ionospheric convection, does not appear in the theoretical description; only some of its effects in the ionosphere are mimicked by the so-called "penetration electric field" postulated for the purpose. The motion of ionospheric plasma in the closed field line region is a result not of imposed electric field in the open field region but because of magnetospheric convection plasma flow there that, due to the requirement of continuity, induces motion in the closed field line regions by action of fast mode waves.

Finally, quantitatively, there are many differences between the electrostatic and dynamic couplings, summarized in Table 1. The coupling speeds, although very different between the two schemes, do not to produce significant differences if the region of interest is only up to a few thousand kilometers above the Earth. The inertia effect of the plasma plays a role in determining the perturbations of the flow velocity, magnetic field, and electric field along the magnetic field. The physical picture for the current is completely different in the two coupling schemes. In quasisteady state coupling, the system is coupled via Birkeland (field-aligned) currents; as the Birkeland currents vary, the Pedersen and Hall currents in the ionosphere are assumed to always vary so as to connect properly to the Birkeland currents. In inductive-dynamic coupling, the currents are carried by the field and flow variations; field-aligned currents are formed in regions where plasma condition and flow vary, as required for self-consistency, and propagate down with magnetic field variations from the magnetosphere to the ionosphere. The Joule heating rate calculated in the conventional theory is actually mostly frictional heating (not Joule heating in the proper physical sense) and neglects the effects of wave heating.

214 INDUCTIVE-DYNAMIC COUPLING

Acknowledgments. The authors thank Dr. J.-N. Tu for comments. This work was supported by the National Science Foundation under grant AGS-0903777 and NASA under grant NNX12AD22G/ NSR:3030708.

REFERENCES

- Bristow, W. A., G. J. Sofko, H. C. Stenbaek-Nielsen, S. Wei, D. Lummerzheim, and A. Otto (2003), Detailed analysis of substorm observations using SuperDARN, UVI, ground-based magnetometers, and all-sky imagers, *J. Geophys. Res.*, 108(A3), 1124, doi:10.1029/2002JA009242.
- Buneman, O. (1992), Internal dynamics of a plasma propelled across a magnetic field, *IEEE Trans. Plasma Sci.*, 20, 672–677.
- Coroniti, F. V., and C. F. Kennel (1973), Can the ionosphere regulate magnetospheric convection?, J. Geophys. Res., 78(16), 2837–2851.
- Cowling, T. G. (1957), *Magnetohydrodynamics*, Interscience, New York.
- Dungey, J. W. (1958), Cosmic Electrodynamics, p. 10, Cambridge Univ. Press, London, U. K.
- Earle, G. D., and M. C. Kelley (1987), Spectral studies of the sources of ionospheric electric fields, J. Geophys. Res., 92(A1), 213–224.
- Fejer, J. A. (1964), Theory of the geomagnetic daily disturbance variations, J. Geophys. Res., 69(1), 123–137.
- Fuller-Rowell, T. J., D. Rees, S. Quegan, R. J. Moffett, M. V. Codrescu, and G. H. Millward (1996), A coupled thermosphereionosphere-plasmasphere model, in *STEP Handbook on Ionospheric Models*, edited by R. W. Schunk, pp. 217–238, Utah State Univ., Logan.
- Holzer, T. E., and G. C. Reid (1975), The response of the day side magnetosphere-ionosphere system to time-varying field line reconnection at the magnetopause, 1. Theoretical model, *J. Geophys. Res.*, 80(16), 2041–2049.
- Huang, C.-S., K. Yumoto, S. Abe, and G. Sofko (2008), Lowlatitude ionospheric electric and magnetic field disturbances in response to solar wind pressure enhancements, *J. Geophys. Res.*, *113*, A08314, doi:10.1029/2007JA012940.
- Huang, C.-S., F. J. Rich, and W. J. Burke (2010), Storm time electric fields in the equatorial ionosphere observed near the dusk meridian, *J. Geophys. Res.*, 115, A08313, doi:10.1029/ 2009JA015150.
- Huba, J. D., G. Joyce, and J. A. Fedder (2000), Sami2 is Another Model of the Ionosphere (SAMI2): A new low-latitude ionosphere model, *J. Geophys. Res.*, 105(A10), 23,035–23,053.
- Hughes, W. J. (1974), The effect of the atmosphere and ionosphere on long period magnetospheric micropulsations, *Planet. Space Sci.*, *22*, 1157–1172.
- Hughes, W. J., and D. J. Southwood (1976), The screening of micropulsation signals by the atmosphere and ionosphere, J. Geophys. Res., 81(19), 3234–3240.
- Kelley, M. C. (2009), *The Earth's Ionosphere: Plasma Physics and Electrodynamics*, 2nd ed., Academic Press, San Diego, Calif.

- Lysak, R. L., and C. T. Dum (1983), Dynamics of magnetosphereionosphere coupling including turbulent transport, J. Geophys. Res., 88(A1), 365–380.
- Parker, E. N. (1996), The alternative paradigm for magnetospheric physics, J. Geophys. Res., 101(A5), 10,587–10,625.
- Parker, E. N. (2000), Newton, Maxwell, and magnetospheric physics, in *Magnetospheric Current Systems*, *Geophys. Monogr. Ser.*, vol. 118, edited by S. Ohtani et al., pp. 1–10, AGU, Washington, D. C., doi:10.1029/GM118p0001.
- Parker, E. N. (2007), Conversations on Electric and Magnetic Fields in the Cosmos, Princeton Univ. Press, Princeton, N. J.
- Richmond, A. D., and Y. Kamide (1988), Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique, J. Geophys. Res., 93(A6), 5741–5759.
- Richmond, A. D., E. C. Ridley, and R. G. Roble (1992), A thermosphere/ionosphere general circulation model with coupled electrodynamics, *Geophys. Res. Lett.*, 19(6), 601–604.
- Ridley, A. J., Y. Deng, and G. Toth (2006), The global ionospherethermosphere model, J. Atmos. Sol. Terr. Phys., 68, 839–864.
- Russell, C. T., and M. Ginskey (1993), Sudden impulses at low latitudes: Transient response, *Geophys. Res. Lett.*, 20(11), 1015–1018.
- Russell, C. T., and M. Ginskey (1995), Sudden impulses at subauroral latitudes: Response for northward interplanetary magnetic field, J. Geophys. Res., 100(A12), 23,695–23,702.
- Schunk, R. W. (2013), Ionosphere-thermosphere physics: Current status and problems, in *Modeling the Ionosphere-Thermosphere System, Geophys. Monogr. Ser.*, doi:10.1029/2012GM001351, this volume.
- Song, P., and V. M. Vasyliūnas (2010), Aspects of global magnetospheric processes, *Chin. J. Space Sci.*, 30(4), 289–311.
- Song, P., and V. M. Vasyliūnas (2011), Heating of the solar atmosphere by strong damping of Alfvén waves, J. Geophys. Res., 116, A09104, doi:10.1029/2011JA016679.
- Song, P., C. T. Russell, and S. P. Gary (1994), Identification of lowfrequency fluctuations in the terrestrial magnetosheath, J. Geophys. Res., 99(A4), 6011–6025, doi:10.1029/93JA03300.
- Song, P., T. I. Gombosi, and A. J. Ridley (2001), Three-fluid Ohm's law, J. Geophys. Res., 106(A5), 8149–8156, doi:10.1029/2000JA 000423.
- Song, P., V. M. Vasyliūnas, and L. Ma (2005), Solar windmagnetosphere-ionosphere coupling: Neutral atmosphere effects on signal propagation, *J. Geophys. Res.*, 110, A09309, doi:10. 1029/2005JA011139.
- Song, P., V. M. Vasyliūnas, and X.-Z. Zhou (2009), Magnetosphereionosphere/thermosphere coupling: Self-consistent solutions for a one-dimensional stratified ionosphere in three-fluid theory, J. Geophys. Res., 114, A08213, doi:10.1029/2008JA013629.
- Strangeway, R. J. (2012), The equivalence of Joule dissipation and frictional heating in the collisional ionosphere, J. Geophys. Res., 117, A02310, doi:10.1029/2011JA017302.
- Tu, J., P. Song, and B. W. Reinisch (2008), On the concept of penetration electric field, *AIP Conf. Proc.*, 974, 81–85.
- Tu, J., P. Song, and V. M. Vasyliūnas (2011), Ionosphere/thermosphere heating determined from dynamic magnetosphere-ionosphere/

thermosphere coupling, J. Geophys. Res., 116, A09311, doi:10. 1029/2011JA016620.

- Vasyliūnas, V. M. (1970), Mathematical models of magnetospheric convection and its coupling to the ionosphere, in *Particles and Fields in the Magnetosphere*, edited by B. M. McCormack, pp. 60–71, D. Reidel, Dordrecht, The Netherlands.
- Vasyliūnas, V. M. (2001), Electric field and plasma flow: What drives what?, *Geophys. Res. Lett.*, 28(11), 2177–2180.
- Vasyliūnas, V. M. (2005a), Time evolution of electric fields and currents and the generalized Ohm's law, Ann. Geophys., 23, 1347–1354.
- Vasyliūnas, V. M. (2005b), Relation between magnetic fields and electric currents in plasmas, Ann. Geophys., 23, 2589–2597.
- Vasyliūnas, V. M. (2011), Physics of magnetospheric variability, Space Sci. Rev., 158, 91–118, doi:10.1007/s11214-010-9696-1.
- Vasyliūnas, V. M. (2012), The physical basis of ionospheric electrodynamics, Ann. Geophys., 30, 357–369.
- Vasyliūnas, V. M., and D. H. Pontius Jr. (2007), Rotationally driven interchange instability: Reply to André and Ferrière, J. Geophys. Res., 112, A10204, doi:10.1029/2007JA012457.

- Vasyliūnas, V. M., and P. Song (2005), Meaning of ionospheric Joule heating, J. Geophys. Res., 110, A02301, doi:10.1029/ 2004JA010615.
- Wolf, R. A. (1970), Effects of ionospheric conductivity on convective flow of plasma in the magnetosphere, J. Geophys. Res., 75 (25), 4677–4698.
- Zesta, E., H. J. Singer, D. Lummerzheim, C. T. Russell, L. R. Lyons, and M. J. Brittnacher (2000), The Effect of the January 10, 1997, pressure pulse on the magnetosphere-ionosphere current system, in *Magnetospheric Current Systems*, *Geophys. Monogr. Ser.*, vol. 118, edited by S. Ohtani et al., pp. 217–226, AGU, Washington, D. C., doi:10.1029/GM118p0217.

P. Song, Center for Atmospheric Research and Department of Physics, University of Massachusetts Lowell, MA, USA. (Paul_Song@uml.edu)

V. M. Vasyliūnas, Max-Planck-Institut für Sonnensystemforschung, 37191 Katlenburg-Lindau, Germany.