10

Energy conversion in planetary magnetospheres

VYTENIS VASYLIŪNAS

10.1 Introduction

Planetary magnetospheres, by their very nature, provide plenty of possibilities for the development of energy conversion processes. Fundamentally a planetary magnetosphere (see e.g. Vol. I, Chapter 10) is simply the interface between two distinct regions: on the outside, the solar wind; on the inside, the ionosphere, atmosphere, and surface of the planet. The quite different motions of matter within the two regions, together with the role of the magnetic field in mediating the interaction between them, lead (almost unavoidably, it seems) to configurations of changing energy; the changes occur on a variety of time scales, ranging from quasistatic to explosive.

In keeping with the general approach adopted in this series of textbooks, this chapter aims to present energy conversion in planetary magnetospheres in general terms as part of a sub-branch of physics, namely the discipline of magnetospheric physics (which in turn is a sub-branch of heliophysics). Many of the concepts and basic results, however, originate from specific observations at and near Earth; accordingly, the chapter begins (Section 10.2) with a phenomenological overview of geophysical processes related to space storms and radiation. The physical description of energy conversion processes is then developed (Sections 10.3, 10.4, 10.5) and applied to interpret the phenomenology of energy-conversion events, both at Earth (Section 10.6) and at other planets (Section 10.7). The chapter concludes (Section 10.8) with a sketch of a possibly universal process.

10.2 Overview of disturbances in Earth's space environment

Of the observed phenomena related to energy conversion processes in outer space, the polar aurora is the earliest known (with records and traces in history, mythology, literature, and the arts reaching back millennia; see Chapter 2,

Heliophysics: Space Storms and Radiation: Causes and Effects, eds. Carolus J. Schrijver and George L. Siscoe. Published by Cambridge University Press.

and e.g. Eather, 1980) and the easiest to observe, even without instruments. Next come disturbances of the Earth's magnetic field, detectable with relatively simple instruments available by the mid-nineteenth century. By the early twentieth century, the two phenomena were known to be connected, and the concept of *magnetic storm* was already current: geomagnetic disturbance of wide (global) extent on time scales of hours to days, unusually intense storms associated with occurence of aurora at unusually low latitudes, evidence of connection with solar activity. More localized auroral manifestations and intense geomagnetic disturbances at high latitudes, on time scales of minutes to hours, were studied under a variety of designations and synthesized much later (1960s) into the concept of *magnetospheric substorm*, with the help of *in situ* outer space observations which were becoming available and proved essential to establish the physical nature of the phenomenon. For brief historical accounts, see e.g. Chapman (1969; one of the key participants), Siscoe (1980), Egeland (1984), and Stern (1991).

The magnetic storm is defined nowadays (Gonzalez *et al.*, 1994) by the time variation of the geomagnetic Dst (disturbance storm time) index, illustrated schematically in Fig. 10.1. The Dst index (see e.g. Mayaud, 1980) is a measure of

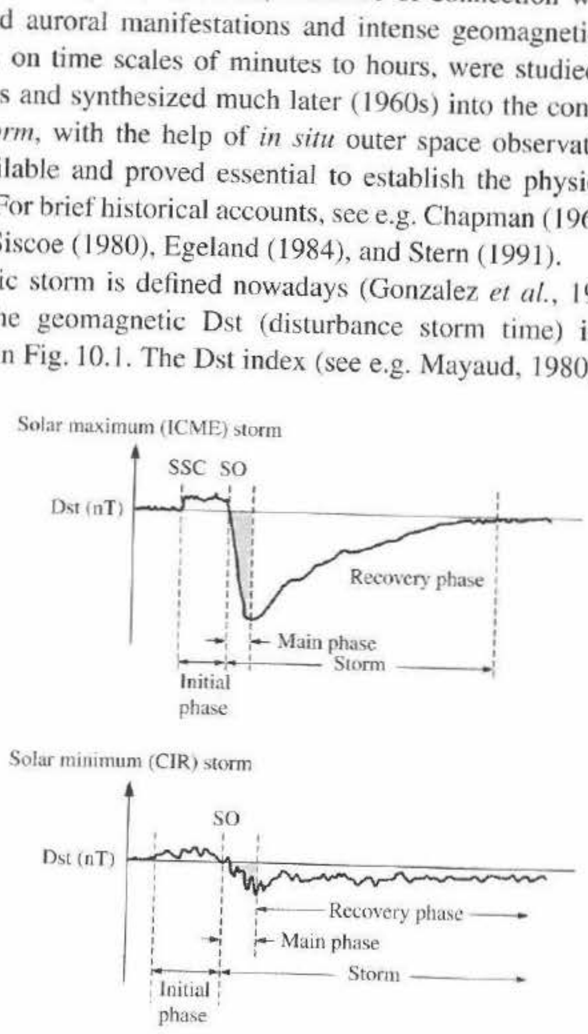

Fig. 10.1. Schematic time history of geomagnetic field variation for two characteristic magnetic storms. Time range: several days. Vertical variation range: ~100–200 nT. SSC: storm sudden commencement. SO: storm onset. The top panel shows the storm development in response to a characteristic interplanetary coronal mass ejection (ICME), and the bottom panel that for the passage of a corotating interaction region (CIR). (Figure adapted from Tsurutani *et al.*, 2006.)

a quasi-uniform magnetic disturbance field near the Earth, aligned with the dipole axis (northward for Dst > 0), such as would be produced by a ring of electric current (westward if Dst < 0) near the equatorial plane. A prolonged (hours to days) interval of negative Dst values constitutes a magnetic storm. The peak negative excursion is often taken as a measure of storm intensity: Dst −30 nT to −50 nT are weak storms, −50 nT to −100 nT moderate, and over −100 nT intense; storms over −300 nT occur at most a few times during a solar cycle (Earth's dipole field at the equator is about 31 000 nT, for comparison). The storm sudden commencement and the initial phase of positive Dst, which accompany many but not all storms, are no longer considered necessary ingredients of the storm concept.

As discussed in Section 10.6.2, the field depression quantified by Dst is the result of plasma pressure that inflates the dipole field. The essential phenomenon of the magnetic storm is thus the addition of a large amount of plasma energy to the dipolar field region of the magnetosphere. Furthermore, it is now well established that this energy addition results from a particular condition in the solar wind: "a sufficiently intense and long-lasting interplanetary convection electric field" (Gonzalez *et al.*, 1994), meaning $-\mathbf{v} \times \mathbf{B}/c$, for the IMF's southward component.

In contrast to the magnetic storm, there is much less unanimity on what defines a magnetospheric substorm (Rostoker *et al.*, 1980, 1987). Probably the most spectacular phenomenon, and the one most widely used as a unifying concept, is the auroral substorm, summarized in the classic figure of Akasofu (1964) reproduced here in Fig. 10.2, which illustrates schematically, by a time sequence of polar views of Earth, the development of the auroral forms (light-emitting regions) during what is called the *expansion phase* of the substorm: beginning with an initial brightening at the lowest latitudes near midnight (*onset*), the aurora intensifies greatly, becomes very complex in spatial structure (*auroral breakup*) and expands, predominantly westward and poleward but also eastward, eventually subsiding in a *recovery phase*. This auroral development is accompanied by strong geomagnetic disturbances (commonly reaching ~1500 nT and more), with a spatial distribution almost as complex as that of the aurora but describable roughly as equivalent to a current above the Earth (*auroral electrojet*) that is westward near and before midnight and eastward after midnight. Note: although the development shown in Fig. 10.2 is in the Northern Hemisphere only, essentially the same sequence also occurs simultaneously in the Southern Hemisphere, at the (more or less) magnetically conjugate locations (the resemblance to a two-ribbon solar flare, with ribbons of opposite magnetic polarity, has been repeatedly remarked upon).

Within the magnetosphere, the substorm expansion phase is marked by (1) greatly enhanced intensities and energies of charged particles, (2) changes of the magnetic field in the nightside magnetosphere and magnetotail, the initially tail-like field becoming more dipolar (*dipolarization*), and (3) fast ($\sim v_A$) bulk flows

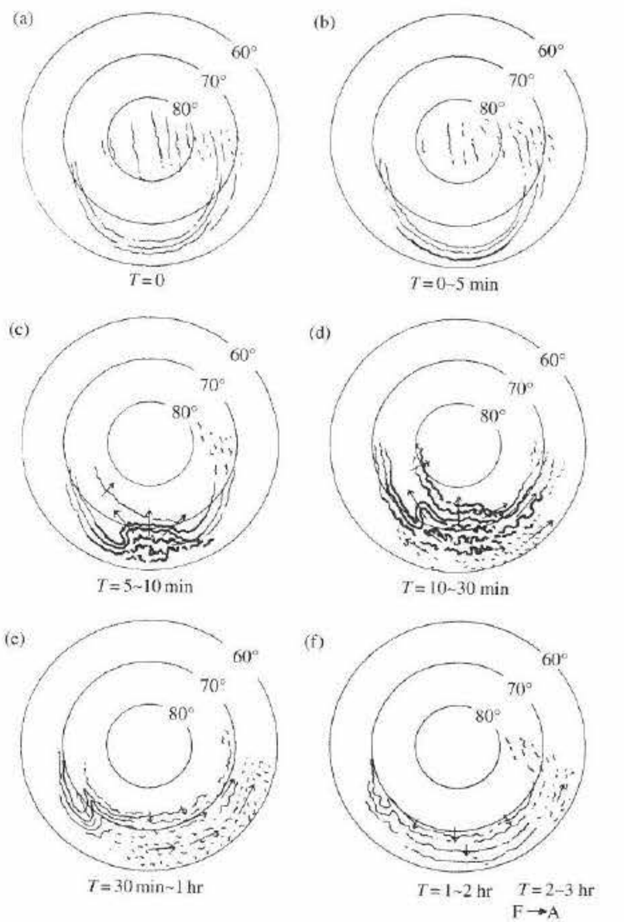

Fig. 10.2. Schematic diagram of an auroral substorm. View from above the North Pole, circles of constant geomagnetic latitude, Sun toward the top (Akasofu, 1964).

of plasma in the magnetotail, predominantly away from Earth at larger distances. This is the merest sketch of substorm phenomenology; for more detailed accounts, see e.g. Akasofu (1977), Kennel (1995) and Syrjäsuo and Donovan (2007).

Not shown in Fig. 10.2 is the substorm *growth phase* which, was not identified until some years after 1964: a time interval (~0.5–1 h) preceding the substorm onset, during which magnetospheric convection (see Vol. I, Section 10.4.3) observed in the ionosphere is enhanced, the amount of open magnetic flux in the magnetosphere increases, and quiet-time auroral forms move equatorward (to reach their locations shown in panel A of Fig. 10.2). Generally, the beginning of the growth phase is associated with a southward turning (or an enhancement of a pre-existing southward component) of the interplanetary magnetic field B_{sw}.

What changes, if any, of B_{sw} or other solar wind parameters are associated with the substorm onset and expansion phase is a still unsettled controversy; the two extreme positions are that the onset (1) is triggered by a northward turning of B_{sw} or (2) is a purely internal development of magnetospheric dynamics.

Because the presence of a southward component of B_{sw} (opposite to the dipole field in the Earth's equatorial plane) appears to be a prerequisite for the occurrence of both storms and substorms, the question may be raised: do magnetic storms and magnetospheric substorms constitute two physically distinct phenomena, or are they merely different-time-scale manifestations of a single phenomenon? Aside from matters of time scale and sequence, one essential conceptual difference is that the defining signature of a magnetic storm represents an enhanced *storage* of plasma energy, while that of a magnetospheric substorm represents in essence (independent of arguments about what it is in detail) an enhanced *dissipation* of energy.

In summary, geomagnetic and auroral phenomena involve particle energy, stored in the magnetosphere (e.g. to inflate the magnetic field) or transferred to the atmosphere (e.g. to excite the aurora); there are related changes of magnetic field configuration, and an evidently significant role is played by the component of the interplanetary magnetic field that can reconnect with the Earth's dipole field. A physical description of energy conversion in a general heliophysical context must also include other magnetospheres (see e.g. Vol. I, Chapters 10 and 13) in which the rotation of the planet may be more important than the solar wind.

10.3 Fundamentals of energy storage, transfer, and loss

10.3.1 Forms of energy

Throughout this chapter, I take a fundamental physical approach, treating energy as a *field* quantity, localizable to any point $(\mathbf{r}, t)$ of space and time (in contrast to an engineering approach, with energy assigned to a particular device, e.g. flywheel, capacitor, or inductor). For each form of energy, one has an *energy density* $U(\mathbf{r}, t)$ and an *energy flux density* $\mathbf{S}(\mathbf{r}, t)$, which satisfy the conservation equation

$$\frac{\partial U}{\partial t} + \nabla \cdot \mathbf{S} = \text{conversion rate}, \tag{10.1}$$

where the right-hand side represents the rate per unit volume of conversion of energy into or out of the particular form.

Three forms of energy are of direct importance for heliophysics: kinetic energy of motion, electromagnetic energy, and gravitational energy. The latter two are the energies of the two long-range fields which (as discussed in Vol. I, Chapter 1) act to organize matter in the cosmos. Nuclear energy (associated with the short-range fundamental forces) is of course the ultimate source of energy that powers the luminosity of the Sun and other stars, but its direct presence is confined to deep stellar interiors (fusion reactions) and, to a minor extent, planetary interiors (radioactivity); elsewhere it only appears in any significant amounts after conversion to other forms.

Kinetic energy of motion includes both energy of bulk flow and energy of thermal motions; the total, including both, is conveniently referred to as *mechanical* energy, for which the conservation equation (10.1) takes the form

$$\frac{\partial}{\partial t} U_{\text{mech}} + \nabla \cdot \left[\mathbf{v} U_{\text{mech}} + \mathsf{P} \cdot \mathbf{v} + \mathbf{q}\right] = \mathbf{E} \cdot \mathbf{J} + \rho \mathbf{v} \cdot \mathbf{g}, \tag{10.2}$$

$$U_{\text{mech}} \equiv \tfrac{1}{2} \rho v^2 + \epsilon, \qquad \epsilon = \text{Trace}\,(\mathsf{P})\,.$$

For electromagnetic energy, the conservation equation is given by Poynting's theorem

$$\frac{\partial}{\partial t} \frac{1}{8\pi}\left[B^2 + E^2\right] + \nabla \cdot \left[\frac{c}{4\pi} \mathbf{E} \times \mathbf{B}\right] = -\mathbf{E} \cdot \mathbf{J}. \tag{10.3}$$

For gravitational energy, an approximate expression adequate for most purposes of heliophysics and magnetospheric physics (Siscoe, 1983) is

$$\frac{\partial}{\partial t} \left[\rho \Phi_G\right] + \nabla \cdot \left[\rho \mathbf{v} \Phi_G\right] = -\rho \mathbf{v} \cdot \mathbf{g}. \tag{10.4}$$

(In the above equations, ρ is the mass density, $\mathbf{v}$ the bulk flow velocity, P the pressure tensor, $\mathbf{q}$ the heat flux vector, $\mathbf{B}$ the magnetic and $\mathbf{E}$ the electric field, $\mathbf{J}$ the electric current density, Φ_G the gravitational potential, and $\mathbf{g} = -\nabla \Phi_G$ the gravitational acceleration.)

The conversion rates between different forms of energy are given by

$$\begin{array}{ll} \mathbf{E} \cdot \mathbf{J} > 0 & \text{electromagnetic} \longrightarrow \text{mechanical} \\ \mathbf{E} \cdot \mathbf{J} < 0 & \text{mechanical} \longrightarrow \text{electromagnetic} \end{array} \tag{10.5}$$

and

$$\begin{array}{ll} \rho \mathbf{v} \cdot \mathbf{g} > 0 & \text{gravitational} \longrightarrow \text{mechanical} \\ \rho \mathbf{v} \cdot \mathbf{g} < 0 & \text{mechanical} \longrightarrow \text{gravitational;} \end{array} \tag{10.6}$$

there is no direct conversion between electromagnetic and gravitational energy (at least as long as general relativistic effects are neglected). If all the energy equations (10.2), (10.3), and (10.4) are added together, the conversion terms on the right-hand sides add to zero, implying conservation of total energy:

$$\frac{\partial}{\partial t} U_{\text{total}} + \nabla \cdot \mathbf{S}_{\text{total}} = 0. \tag{10.7}$$

Note that the conversion rates are *not* independent of frame of reference. All three quantities – energy density, energy flux density, and energy conversion rate – vary with choice of frame of reference, in such a way that the *form* of the energy equation remains invariant. Sometimes a profound significance is claimed for the sign of $\mathbf{E} \cdot \mathbf{J}$, regions with $\mathbf{E} \cdot \mathbf{J} < 0$ or > 0 being identified as "dynamo" or "load", respectively; since physics is frame-independent, this distinction cannot be fundamental.

10.3.2 Sources of energy for magnetospheres

Strictly speaking, there can be no source of energy as such: according to Eq. (10.7), energy can neither be created nor be destroyed but can only be converted from one form to another or transported from one region to another. For a region bounded in space such as a planetary magnetosphere, however, the term *energy source* is often applied to energy transported into the region across the boundary. The external source (in this sense) of energy for a planet and its associated system is the Sun, which supplies energy in two forms: electromagnetic radiation and the solar wind. The power carried by the electromagnetic radiation (solar luminosity) is observed to exceed that carried by the solar wind by a factor $\sim 10^6$; with $v_{\text{sw}}/c \sim 10^{-3}$, this implies that the rate at which the Sun is losing mass through relativistic energy-equivalent mass removal by the solar radiation is comparable to the mass outflow by the solar wind (Axford, 1985).

The solar radiation is the dominant energy source for the planet, the atmosphere, and part of the ionosphere. For the magnetosphere and the upper regions of the ionosphere, on the other hand, the solar wind is the only significant external source of energy available; solar radiation does not interact at all with these regions, where the density of matter is sufficiently low to make the mean free path for interaction with photons vastly larger than the size (column depth) of the system. (For the same reason, dynamics of the solar wind and the interplanetary magnetic field can be treated without reference to the omnipresent solar visible radiation: at 1 AU, for instance, one discusses magnetic fields typically of order $\sim$10 nT and electric fields $\sim$4 mV m^{-1}, while ignoring magnetic fields $\sim 10^3$ nT and electric fields $\sim$30 V m^{-1} that are simultaneously present – albeit oscillating at $\sim 10^{15}$ Hz.)

When considering the solar wind as the energy source, only the kinetic energy of plasma bulk flow is of importance; the thermal and magnetic energies of the solar wind can be neglected, for a reason somewhat more subtle than might appear at first. They are small compared to the kinetic energy of the bulk flow, but not necessarily small compared to energies dissipated in the magnetosphere; the reason they are not important is that at the bow shock they are overwhelmed by additional thermal and magnetic energies extracted from the flow. Furthermore, to transfer magnetic energy across the magnetopause requires a normal component of the Poynting vector, hence a tangential component of the electric field, which interacts with the magnetopause current to extract more mechanical energy from the plasma; thus the Poynting vector just inside the magnetopause is in general completely different from the Poynting vector just outside (and also from the Poynting vector in the solar wind). The interplanetary magnetic field does exert a dominant influence on energy conversion processes in a planetary magnetosphere, but primarily by control of magnetic reconnection processes and open field lines – not by entry of the solar-wind Poynting flux into the magnetosphere.

An interior source of energy available for a planetary magnetosphere is planetary rotation (other sources of energy internal to the planet, e.g. heating by radioactivity or by slow contraction, have in general no direct interaction with the magnetosphere).

10.3.3 Energy loss and dissipation processes

Analogously to "energy source", the term *energy loss* (or *sink*) is often used to denote a process in which energy is transported out of the region under consideration, or else transformed into a form that allows it to escape from the system with no further interaction. A related concept is that of *energy dissipation*, a process in which energy is transformed into heat in the thermodynamic sense, with increase of entropy (for a detailed discussion in relation to the energy and momentum equations, see e.g. Vasyliūnas and Song, 2005; the dissipation rate so defined, unlike the energy conversion rate, is independent of frame of reference).

The following are among the principal loss and dissipation processes in planetary magnetospheres, energy being lost primarily to the atmosphere in (1) and (2) and being removed outside the system (to "infinity") in (3) and (4):

(1) **Collisional and Joule heating in the ionosphere** If the bulk flow of plasma differs from the bulk flow of the neutral atmosphere (usually as a consequence of magnetospheric dynamics), there is energy dissipation given by $\mathbf{E}^* \cdot \mathbf{J}$, where $\mathbf{E}^*$ is the electric field in the frame of reference of the neutral atmosphere. This is commonly referred to as "ionospheric Joule heating", but Vasyliūnas and Song (2005) have shown that in fact it is primarily frictional heating by collisions between plasma and neutral particles; Joule heating in the true physical sense ($\mathbf{E}' \cdot \mathbf{J}$, where $\mathbf{E}'$ is the electric field in the frame of reference of the plasma) contributing only a small fraction of the total. The energy is removed from the magnetic field and converted (via kinetic energy of relative bulk flow as an intermediary) to heat (thermal energy), with the heating rate per unit volume partitioned approximately equally between plasma and neutrals.

(2) **Charged-particle precipitation** Energetic charged particles that enter the atmosphere from above are usually said to be *precipitating*. They penetrate into the atmosphere to a depth that increases with increasing energy, until their energy is lost, going partly into heating the atmosphere and partly into ionization or other interactions.

One source of precipitating particles is simple loss from the radiation belts (see Chapter 11) or from the ring current and plasma sheet regions; the energy deposited in the atmosphere is taken from the mechanical (thermal) energy of the respective magnetospheric particle populations. In addition to these particles that precipitate merely because their velocity vectors are oriented in the appropriate direction, there are other sources of precipitating charged particles, in which the energy and the intensity of the particles have been enhanced by an acceleration process. In particular, the auroral phenomena that occur in nearly all of the planetary magnetospheres observed to date are generally interpreted as resulting from some special acceleration process that supplies the required intensities of precipitating charged particles. A widely accepted model, developed from extensive studies at Earth and applied to aurora at Jupiter and at Saturn, ascribes auroral acceleration to Birkeland (magnetic-field-aligned) electric currents accompanied by electric fields parallel to the magnetic field; the rate of energy supply to the precipitating particles is $E_\parallel J_\parallel$, hence the added energy is taken out of the magnetic field (in this model, aurora occurs only when the Birkeland current is directed upward, corresponding to electron motion downward). Auroral acceleration has also been associated with intense Alfvénic turbulence (which contains fluctuating Birkeland currents). For detailed reviews, see e.g. Paschmann *et al.* (2003) and references therein.

(3) **Emission of electromagnetic radiation** A variety of processes in planetary magnetospheres produce electromagnetic radiation of various types: atomic and molecular line emissions (from the aurora and from magnetospheric interactions with plasma and neutral tori), radio waves (wideband and narrowband), a veritable zoo of plasma waves, and even X-rays (bremsstrahlung from precipitating electrons and, possibly, nuclear line emissions excited by very energetic precipitating particles). Some aspects are discussed in Chapter 4; the emissions are of course of great interest for remote sensing of the associated processes. As far as the energetics of planetary magnetospheres are concerned, however, the amount of energy involved is negligibly small for most emissions, with only a few exceptions (UV radiation from the Io torus at Jupiter).

(4) **Energetic neutral particle escape** Neutral particles that remain within a magnetosphere must be gravitationally bound to the planet; plasma particles within the magnetosphere, on the other hand, typically have speeds that exceed (often by a large factor) the gravitational escape speed – plasma is held within the

magnetosphere by the magnetic field, not by gravity (the magnetic field itself, however, must be anchored to the planet by its gravity, as discussed in Vol. I, Chapter 1). Charge-exchange collisions between ions and neutrals, in which the outgoing neutral has the velocity of the incoming ion and vice versa, thus produce fast neutrals that escape from the system immediately, with their (newly acquired) kinetic energy. This process represents a loss (generally by quite significant amounts) both of neutral particles and of energy from the magnetosphere.

(5) **Dissipation processes in the magnetosphere** In regions of the magnetosphere with major departures from the MHD approximation (particularly where magnetic reconnection is occurring) dissipative processes such as Joule heating associated with effective resistivity may be significant. The primary effect is not energy loss but enhancement of conversion from magnetic to thermal energy.

10.3.4 Reservoirs of energy

The field approach to energy implies that energy may be regarded as *stored* in space, the energy density of the various forms being given by the terms that are time-differentiated in the energy equations (10.2), (10.3), and (10.4). The primary reservoir of stored mechanical energy in a planetary magnetosphere is the thermal energy of its various plasma structures, especially the *plasma sheet* of the magnetotail or magnetodisk, the *ring current*, and the plasma and neutral *tori* associated with the planet's moons (see e.g. the description of the structures in Vol I, Section 10.5.3); the kinetic energy of bulk flow of magnetospheric plasma also plays a role, particularly for plasma tori and in the case of rapid changes discussed in Section 10.5.

The primary reservoir of stored electromagnetic energy of importance for a planetary magnetosphere is the energy of the magnetic field; except for high-frequency radiation, which does not interact with the magnetosphere, the energy in the electric field is negligible in comparison to that in the magnetic field. Because the energy of the planetary dipole field itself does not change (except on time scales of the secular variation, $\sim 10^2 - 10^3$ years for Earth) and thus has no effect on the energetics of the magnetosphere, a convenient measure of stored electromagnetic energy is the energy of the total magnetic field minus the (unchanging) energy of the dipole field:

$$\frac{1}{8\pi}\int\left[B^2 - \left(B_{\text{dipole}}\right)^2\right]\mathrm{d}V.$$

The stored gravitational energy can be changed only by a net radial displacement of matter; any such effects in the magnetosphere are for the most part negligible in comparison to changes of mechanical or magnetic energy.

10.4 Energy budget of magnetospheres

The topic of this chapter may now be formulated as follows: the primary sources of energy for a planetary magnetosphere being the kinetic energy of bulk flow, both exterior (solar wind flow) and interior (planetary rotation) to the magnetospheric volume, by what process and in what form does the energy enter the magnetosphere, what are its flow paths and conversions within the magnetosphere, what are its ultimate sinks, and what determines the time history of these developments? In this section, I first consider these questions without reference to explicit time variations, with particular attention to the role of stress balance and magnetic flux transfer (some of the issues are briefly discussed also in Vol. I, Chapter 11), leading to the construction (Section 10.4.3) of a schematic diagram for the magnetospheric global energy budget in a quasi-steady or time-averaged context. Then I consider in Section 10.5 the time-varying energy conversion processes, many of which can be described as consequences of time offsets or delays in the interactions corresponding to particular branches of the average energy budget diagram.

10.4.1 Extracting energy from bulk flow

Bulk flow of a medium carries not only kinetic energy but also linear momentum; extracting kinetic energy from the flow necessarily means also extracting linear momentum, which requires a force to be applied to the medium. Similarly, rotation of a body carries angular momentum; extracting kinetic energy from the rotation necessarily means also extracting angular momentum, which requires a torque to be applied to the body.

(1) **Relation between global energy input rate and force/torque** The net rate of energy extraction (power) $\mathcal{P}_{\text{sw}}$ from solar wind flow is equal to the difference of the solar wind kinetic energy flux across two surfaces perpendicular to the Sun–planet line, surface 1 ahead of the bow shock and surface 2 far downstream of the entire interaction,

$$\begin{aligned}\mathcal{P}_{\text{sw}} &= \frac{1}{2}\int_1 \rho v^3\,\mathrm{d}A - \frac{1}{2}\int_2 \rho v^3\,\mathrm{d}A \\ &= \frac{1}{2}\int \rho v(\overline{v}_1^2 - \overline{v}_2^2)\,\mathrm{d}A \\ &= S_{\text{ft}}\,\overline{v}\Delta v \end{aligned} \tag{10.8}$$

(subscripts sw on ρ and v have been omitted, for simplicity), and the total force F is similarly equal to the difference of the linear momentum flux,

$$F = \int_1 \rho v^2\,\mathrm{d}A - \int_2 \rho v^2\,\mathrm{d}A = S_{\text{ft}}\Delta v, \tag{10.9}$$

where $\Delta v \equiv \overline{v}_1 - \overline{v}_2$, $\overline{v} \equiv (\overline{v}_1 + \overline{v}_2)/2$ (bars indicate suitable averages) and

$$S_{\rm fl} = \int_1 \rho v \, \mathrm{d}A \simeq \int_2 \rho v \, \mathrm{d}A \tag{10.10}$$

is the amount of mass per unit time flowing through the region of interaction between the solar wind and the magnetosphere, to be distinguished from $S_{\rm sw}$, the mass input rate from the solar wind into the magnetosphere discussed in Vol. I, Section 10.6.2. (Note: magnetic and thermal contributions to solar wind energy and momentum flux have been neglected as small in comparison to those of the bulk flow.) Combining Eqs. (10.8) and (10.9) yields a relation between the power and the force (in the direction of solar wind flow),

$$\mathcal{P}_{\rm sw} = F \overline{v}, \tag{10.11}$$

which was used first by Siscoe (1966) and Siscoe and Cummings (1969) to estimate the energy input into the terrestrial magnetosphere, under the assumption that the relevant force F is the tangential (magnetotail) force acting primarily on the nightside, $F_{\rm MT}$ (see detailed discussion of forces in Vol. I, Section 10.3.2). (Note: if F is equated to the pressure force $F_{\rm MP}$ on the entire magnetopause, it can be shown that the associated $\mathcal{P}$ does not go into the magnetosphere but represents the power expended in irreversible heating at the bow shock (see also Section 10.4.3).)

Calculating the power extracted from planetary rotation is somewhat simpler. The angular momentum of the rotating planet is $\mathcal{I}\Omega_0$ and the kinetic energy of rotation is $\frac{1}{2}\mathcal{I}\Omega_0{}^2$, where $\mathcal{I}$ is the moment of inertia and Ω_0 the angular frequency of rotation (the subscript 0 designates the rotation frequency of the planet, as distinct from, e.g., the atmosphere or the magnetosphere). With $\mathcal{T}$ the torque on the planet (component along the rotation axis),

$$\mathcal{P}_{\rm rot} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}\mathcal{I}\Omega_0{}^2\right) = \Omega_0 \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{I}\Omega_0\right) = \mathcal{T}\Omega_0, \tag{10.12}$$

a relation between the power and the torque, completely analogous to Eq. (10.11). (In principle, Ω_0 decreases with time as the result of the torque, but in practice the rate of decrease is completely negligible. The time for the present magnetospheric torque to reduce appreciably the planet's rate of rotation is several orders of magnitude longer than the Hubble time, both at Jupiter and at Earth; for the latter, this implies that the magnetospheric torque is much smaller than the lunar tidal torque.)

(2) Implications for linear/angular momentum The linear or angular momentum that is extracted together with the kinetic energy is a conserved quantity; it cannot simply disappear, and its further transport must be accounted for.

What happens to the linear momentum extracted from the solar wind flow is well understood: it is transferred to and exerts an added force on the massive planet (Siscoe, 1966; Siscoe and Siebert, 2006; Vasyliūnas, 2007; see discussion in Vol. I, Section 10.3.2). The angular momentum extracted from the rotation of the planet, on the other hand, can only be removed to "infinity", and identifying the mechanism by which it is transported away is indispensable for understanding the interaction. There are several possibilities:

(a) In magnetospheres with a significant interior source S of plasma (from moons or planetary rings), angular momentum can be advected by the outward transport of mass. For the simple example of plasma corotating rigidly out to a distance $R_{\rm H}$ and coasting freely beyond $R_{\rm H}$ (an approximation to the partial-corotation model discussed in Vol. I, Section 10.4.4), angular momentum is transported outward at the rate $S R_{\rm H}{}^2\Omega_0$, hence from Eq. (10.12) the extracted power is

$$\mathcal{P}_{\rm rot} \simeq S\,\Omega_0{}^2\,R_{\rm H}{}^2, \tag{10.13}$$

one half of which goes into the kinetic energy of bulk flow of the outflowing plasma (in this model), and the remainder is available for powering other magnetospheric processes (proposed for the magnetosphere of Jupiter by Dessler, 1980, and by Eviatar and Siscoe, 1980).

(b) If the solar wind exerts a tangential force on the magnetosphere, it will also exert a torque whenever the distribution of the force is not symmetric about the plane containing the solar wind velocity and the planetary rotation axis. The torque may be estimated as $\mathcal{T} \sim R_{\rm MP}\Delta F$, where $R_{\rm MP}$ is the distance to the dayside magnetopause and ΔF is the difference between the force on the dawn and on the dusk side; this gives the ratio of power from rotation to power from solar wind flow as

$$\mathcal{P}_{\rm rot}/\mathcal{P}_{\rm sw} \sim (\Delta F/F)\,(\Omega_0 R_{\rm MP}/v_{\rm sw})\,. \tag{10.14}$$

In a slowly rotating magnetosphere such as Earth, $\Omega_0 R_{\rm MP}/v_{\rm sw} \equiv \epsilon \ll 1$ and one also expects $\Delta F/F$ to scale as $\sim\epsilon$; hence the power extracted from rotation by the solar wind torque is negligible.

(c) In a rapidly rotating open magnetosphere, on the other hand, magnetic field lines that extend from the planet into the solar wind may become twisted (by a process analogous to the formation of the Parker spiral in the solar wind), creating a Maxwell stress that transports angular momentum outward into the solar wind. This mechanism of extracting energy from planetary rotation was proposed by Isbell *et al.* (1984) for Jupiter (where it is now considered not important in comparison to mass outflow) and by Hill *et al.* (1983) for Uranus.

(d) If the magnetic moment of the planet is tilted relative to the rotation axis, electromagnetic waves that carry away angular momentum may be generated by the rotation. This is generally believed to be the primary mechanism for energy loss from pulsars but is negligible for systems that are very small in comparison to

c/Ω_0, the radius of the speed-of-light cylinder (which is the case for all planets in our solar system and their magnetospheres).

10.4.2 Role of magnetic flux transport

To extract kinetic energy from bulk flow, whether exterior (solar wind) or interior (planetary rotation), and inject it into the magnetosphere, the first step is to slow down the flow by the action of magnetic force at the interface. For the solar wind, this is sketched in Fig. 10.3a, which should be looked at in the context of a more complete representation of the open magnetosphere (e.g. Fig. 10.3 or Fig. 13.4 in Vol. I). As the plasma flows through the current layer implied by the sharp turn of the magnetic field, it is slowed down by the $\mathbf{J} \times \mathbf{B}$ force, by an amount Δv readily estimated as the Alfvén speed based on the internal field B_T and the external density ρ,

$$\Delta v \simeq B_T / (4\pi\rho)^{1/2}, \tag{10.15}$$

and the (initially mechanical) energy flux density incident on the outside continues into the inside of the magnetotail as an electromagnetic energy flux density (Poynting vector). The interface is here idealized as a thin magnetopause, but in reality it must have appreciable thickness so that the amount of plasma S_{ft} flowing through the interaction region carries sufficient energy to account for the energy input into the magnetotail. The energy input rate from Eq. (10.11) with the force equal to F_{MT} given by Eq. (10.7) in Vol. I is

$$\mathcal{P}_{sw} \simeq \left(B_T^2/8\pi\right) A_T\, v, \tag{10.16}$$

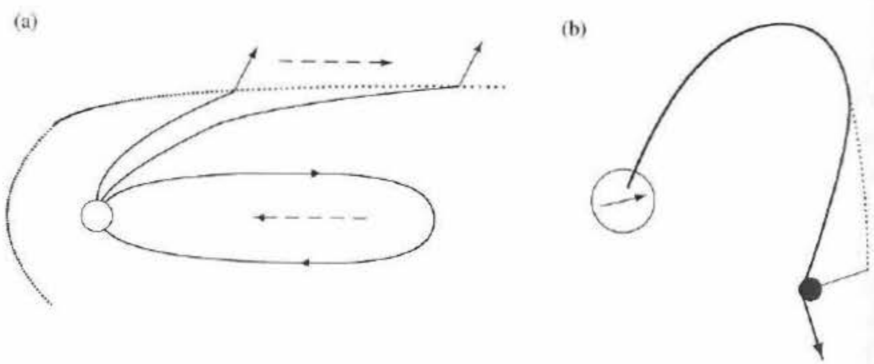

Fig. 10.3. (a) Deformation of magnetotail field by external plasma flow. Solid lines: magnetic field lines. Dashed arrows: plasma flow direction. Dotted line: magnetopause. (b) Deformation of planetary magnetic field by torque from magnetospheric plasma element (black sphere). Solid line: actual magnetic field line. Dashed line: undistorted magnetic field line. Arrow on planet's surface: direction of rotational motion.

which combined with Eqs. (10.8) and (10.15) gives

$$S_{\mathrm{ft}} \simeq \rho\, A_T\, v \left[B_T / \left(16\pi\rho v^2\right)^{1/2}\right]; \tag{10.17}$$

at Earth. This implies that mass flow through the interaction region must be a significant fraction ($\sim 1/4$) of solar wind flux through an area equal to the cross section of the magnetotail (for a more detailed discussion, see e.g. Vasyliūnas, 1987).

A qualitative but more physical way of looking at the interaction is to note that the flow of the solar wind plasma (massive in comparison to plasma in the magnetotail) is carrying the open magnetic field lines with it, while at the same time the feet of these field lines remain anchored to the planet (although free to move laterally, cf. Vol. I, Section 10.4.1); the length of a magnetic flux tube is thus increasing, but its cross-sectional area remains nearly constant (the field magnitude is fixed by the external pressure), hence the volume and with it the magnetic energy content is increasing. (For the plasma, the process is approximately a free expansion; hence the plasma energy content does not change much and is small in any case.) From this point of view, the energy input is closely related to the transport rate of open magnetic flux, from reconnection in the dayside to the magnetotail in the nightside; the question of the amount of energy involved is connected to the fundamental question of the length the magnetotail – how far can an open field line be stretched before it must reconnect and flow back as a closed field line?

Also sketched in Fig. 10.3a is a closed field line, flowing toward the planet and carrying the return magnetic flux. The volume of the flux tube is decreasing, and the plasma energy (greatly enhanced already by the reconnection process from the open to the closed field) is being increased by adiabatic compression. This can be shown to be a conversion of energy from magnetic to mechanical and is further discussed in Section 10.4.3 (energization by adiabatic compression is equivalent to energization by particle drift along the electric field; Hines, 1963).

For planetary rotation, the conversion of kinetic into magnetic energy is sketched in Fig. 10.3b. The prerequisite is a mechanical torque (directed against the rotation) in the magnetosphere; most easily visualized is simple inertia of a plasma element, which holds back the equatorial segment of a magnetic field line, while the feet of the field line at the planet continue to corotate, thus creating an azimuthal magnetic field and increasing the magnetic energy. If the plasma element were to remain at a fixed radial distance indefinitely, it would ultimately be brought up to full corotation and the azimuthal field would disappear; the outward transport process (see Vol. I, Sections 10.5.2, 13.2.1, and 13.3.4), however, removes the plasma in a finite time. The energy input rate thus depends on the rate of mass outflow, which in turn is coupled to circulation of magnetic flux (Vol. I, Chapter 10 and references therein).

Note that in both cases the kinetic energy is first converted into magnetic energy. The energy extracted from planetary rotation can be transported outward, at altitudes just above the top of the ionosphere, *only* by the Poynting vector – with the low density of matter in this region, any mechanical energy flux density is simply too small. That the energy input from the solar wind enters the magnetosphere predominantly in magnetic form is confirmed, at Earth, by the observation that the energy input rate is an order of magnitude larger than the mass input rate multiplied by $\frac{1}{2}v_{sw}^2$ (Hill, 1979).

10.4.3 Energy budget diagram

A schematic diagram for the principal energy flow and transformation processes in a planetary magnetosphere–ionosphere system interacting with the solar wind is shown in Fig. 10.4. This is a simplified synthesis of more detailed energy flow charts derived for two extreme cases, solar wind energy source (Earth) and planetary-rotation energy source with planetary-moon mass source (Jupiter); for the more complicated case of Saturn, where these two sources are of comparable importance, detailed studies of the magnetosphere have only recently become possible (see Section 10.7.1). Here, I concentrate on energy aspects only; for a

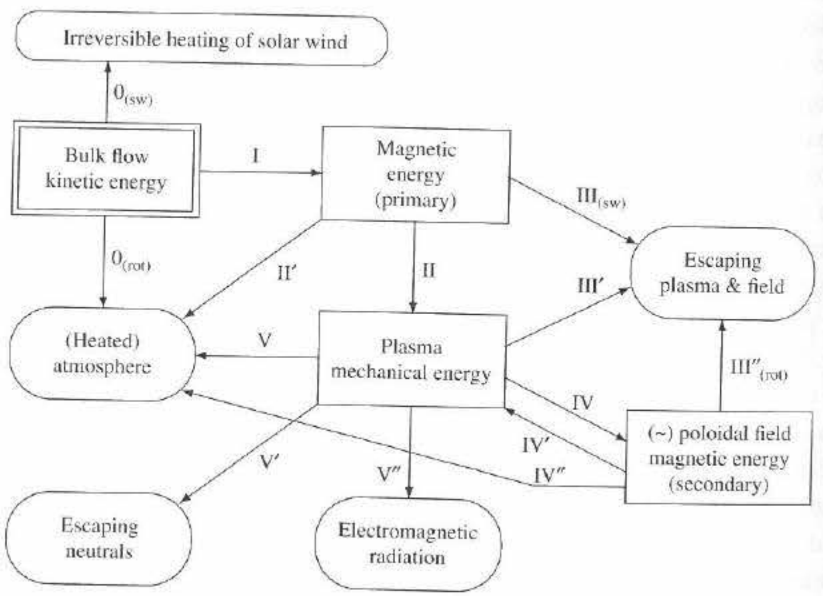

Fig. 10.4. (Simplified) general energy flow chart for planetary magnetospheres and ionospheres. Rectangular boxes: energy reservoirs. Rounded boxes: energy sinks. Lines: energy flow/conversion processes. (Note: only the energy-flow paths are shown, not the mass-flow paths.)

more general discussion, including plasma sources and transport, see Chapters 10, 11, and 13 in Vol. I.

In Fig. 10.4, the primary source of energy – bulk-flow kinetic energy, either of the solar wind or of planetary rotation – is shown by the double-lined box. Energy reservoirs are shown by rectangular boxes (with form of energy identified). Energy sinks, where energy is dissipated or leaves the system, are shown by rounded boxes (for all forms). Plasma mechanical energy is shown as a single reservoir, lumping together the various plasma regions and their thermal and bulk-flow energies. Magnetic energy is shown as two reservoirs: one coupled directly to the source (magnetotail, for the solar wind; wrapped-up field, for rotation), the other from inflation of the dipole field by plasma within the magnetosphere. Energy conversion processes are shown by connecting lines, with arrows indicating the direction of energy flow. The labels on the lines are keyed to the discussion of the corresponding process in the text; subscript (sw) or (rot) – process important only for dominant solar-wind or dominant rotation source, respectively; no subscript – important for both. Processes of minor importance on the scale of the entire magnetosphere (e.g. direct particle precipitation into the atmosphere from the magnetosheath or from the solar wind) have been left out.

(1) **Process I** is the initial conversion of bulk-flow kinetic energy into magnetic energy, described in Section 10.4.2. With the solar wind source, the magnetic energy is stored predominantly in the magnetotail. With the planetary-rotation source, the magnetic energy is stored predominantly in the azimuthal magnetic field, which is in the direction of lagging behind rotation of the planet (like the Parker spiral relative to rotation of the Sun).

(2) **Process II** is conversion of magnetic energy into mechanical energy that is then stored in magnetospheric plasma. It includes formation and energization of the plasma sheet (by magnetic reconnection and adiabatic compression, see Section 10.4.2) and energization of the plasma in the ring current region (predominantly by adiabatic compression during inward transport; although energy is also removed of course by adiabatic expansion during outward transport, the net effect is energy addition as long as there is a net inflow of plasma, to increase the ring current or maintain it against losses). Additionally, in magnetospheres with a significant interior source of plasma from a moon (Io at Jupiter, Enceladus at Saturn) there is the *pick-up* process: ionization of slow-moving (Keplerian) neutrals in the presence of flowing (nearly corotating) plasma, which imparts both flow and thermal energy to the ions.

(3) **Process II′** is also conversion of magnetic energy into mechanical energy which, however, goes directly into the ionosphere and the aurora (these are not shown in Fig. 10.4 – as far as the energy budget is concerned, they are simply

intermediaries in the process by which magnetic energy is converted to heat of the atmosphere). The process, closely associated with Birkeland currents, consists primarily of collisional and Joule heating of the ionosphere as well as auroral acceleration and precipitation (Section 10.3.3).

(4) **Processes III, III′, III″** represent the loss of energy (mechanical and magnetic) by outflow down the distant magnetotail; they all involve magnetic reconnection, since the magnetic field lines must sooner or later become disconnected from the planetary dipole. In a rotation-dominated magnetosphere, processes III′ and III″ are related to the formation of the planetary/magnetospheric wind (see Vol. I, Sections 10.4.4 and 13.2.4).

(5) **Processes IV, IV′** describe energy conversions between plasma in the inner regions of the magnetosphere (ring current, plasma torus) and the nearly dipolar but perturbed magnetic field. Process IV represents the deformation of the magnetic field by plasma pressure in the ring current region as well as by corotational stresses. Process IV′, in turn, represents energization of the plasma by adiabatic compression during inward transport related to the distorted poloidal magnetic field (rather than to the magnetotail field, as in process II; the distinction between the two is not always clear-cut).

(6) **Process IV″** is conversion of magnetic energy into heating of the atmosphere, via the ionosphere and the aurora, analogous to process II′ but taking energy from the distorted magnetic field in the outer magnetosphere rather than from the magnetotail or the wrapped-up azimuthal field. The related auroral processes may be important for substorm onset (Section 10.6).

(7) **Energy sinks**: loss of magnetic and mechanical energy by outflow down the magnetotail (processes III, III′, III″) and loss of magnetic energy by ionospheric and auroral processes into the atmosphere (processes II′, IV″) have already been discussed. In addition, mechanical energy of plasma in the magnetosphere is lost to the atmosphere by particle precipitation (**process V**); it is lost to "infinity" by escape of fast neutrals from charge exchange (**process V′**) and by electromagnetic radiation (**process V″**). In the case of the Io torus at Jupiter, radiation (mainly UV) produced by atomic/molecular collision and excitation processes carries an amount of energy that is significant (possibly even dominant) for the magnetospheric energy budget (Thomas *et al.*, 2004, and references therein). Electromagnetic radiation from the atmosphere (including auroral emissions) is not shown in Fig. 10.4, the power involved being in general negligible on the scale of the magnetospheric energy budget; discussions of energy supply in the aurora usually refer to energy in the precipitating particles that excite the auroral emissions, not energy in the emissions themselves.

(8) **Processes 0** extract energy from bulk flow but do not put it into the magnetosphere. For the solar wind source, the net energy extracted from bulk flow as the solar wind is slowed down and forced to go around the magnetospheric obstacle represents the power expended in irreversible heating of solar wind plasma at the bow shock: when the plasma that has been compressed and heated at the bow shock expands far downstream to its initial (ambient solar wind) pressure, it has a higher temperature (because of the increase of the entropy) and hence (by Bernoulli's law) a slower velocity. For the planetary-rotation source, the torque that extracts energy from rotation is transmitted by magnetic stresses, which can act only as far as the bottom of the ionosphere; farther down, between the ionosphere and the planet, the torque must be transmitted purely by stresses in the neutral medium – effective viscous stresses from velocity shear, accompanied in general by energy dissipation and thus heating of the neutral atmosphere. In both cases, the power involved may exceed by an order of magnitude the entire energy input into the magnetosphere. This is most obvious for Earth, where (see Section 10.4.1) F_{MP} exceeds F_{MT} typically by a factor ~10; for Jupiter, this is suggested by the inference that, when ionospheric plasma slips relative to corotation, the neutral atmosphere very nearly moves together with the plasma (Huang and Hill, 1989; Pontius, 1995), implying a large velocity shear below the ionosphere. Processes 0 may thus constitute the largest energy dissipation processes in the entire interaction of the solar wind with a magnetized planet.

10.4.4 *Overview of rates and constraints*

It is of interest to consider what can be said about the energy conversion rates for the various paths of Fig. 10.4 in the two extreme cases, solar wind energy source (Earth) and planetary-rotation energy source with planetary-moon mass source (Jupiter), with particular attention to the constraining processes; these play an essential role in the origin of time-varying energy releases discussed in Sections 10.5, 10.6, and 10.7.

(1) **Solar-wind-dominated magnetosphere** The total power $\mathcal{P}_{total}$ supplied by the solar wind energy source can be considered to be a known quantity, fixed by the solar wind parameters and the size of the magnetosphere; in order of magnitude it is equal to the flux of solar wind kinetic energy through an area equal to the cross section of the magnetotail, $\frac{1}{2}\rho_{sw}(v_{sw})^3 A_T$. The power in paths 0_{sw} and I is fixed by force balance considerations: $\mathcal{P}_{0_{(sw)}}$ and $\mathcal{P}_I$ are obtained from Eq. (10.11) with F set equal to F_{MP} and F_{MT}, respectively. This gives $\mathcal{P}_{0_{(sw)}}$ nearly equal to (but of necessity slightly less than) $\mathcal{P}_{total}$, as discussed in Section 10.4.3, and $\mathcal{P}_I$ equal to $\mathcal{P}_{sw}$ of Eq. (10.16), smaller than $\mathcal{P}_{total}$ by an order of magnitude and determined to

a large extent by the amount of open magnetic flux $\sim B_T A_T$, in agreement with the discussion of Section 10.4.2.

For the remaining paths, there are no obvious general estimates of the expected power. There have been numerous empirical estimates, however, of the power in paths II and II′, along with a search for its dependence on solar wind parameters (e.g. Weiss *et al.*, 1992; Koskinen and Tanskanen, 2002, and references therein). The ratio $\mathcal{P}_{II}/\mathcal{P}_{II'}$ of dissipated to stored energy is uncertain (estimates range from ~0.1 to >0.5), and most studies concentrate on the sum $\mathcal{P}_{II} + \mathcal{P}_{II'}$, which is found to vary with the rate of open magnetic flux transport, similarly to $\mathcal{P}_I$ on the average. Several empirical formulas for the dependence on solar wind parameters have been proposed (Burton *et al.*, 1975; Perrault and Akasofu, 1978, and others; review by Gonzalez, 1990); the differences are not very significant in view of the uncertainties. The magnitude of $\mathcal{P}_{II} + \mathcal{P}_{II'}$, however, is in general nearly an order of magnitude smaller than $\mathcal{P}_I$ estimated from Eq. (10.16), for comparable solar wind conditions. This implies that, at least on the average, a large part of the power $\mathcal{P}_I$ supplied to the magnetosphere escapes down the magnetotail, via paths $III_{(sw)}$ and III′, and only a fraction enters near-Earth space – the space weather effects discussed in Chapter 2 are produced by something like a small percentage of the total power in the solar wind interaction with the Earth system.

(2) **Rotation-dominated magnetosphere with internal mass source** In this case, the total power supplied by the rotational energy source *cannot* be considered a quantity known a priori: it is determined by the applied torque, which depends in detail on the dynamics of the magnetosphere (in contrast to the solar wind case, where the mere deflection of the solar wind specifies the dominant force). What can be considered as known is the internal mass source of the magnetosphere: the total rate S of *mass* (not energy) flow associated with path III′. The requirement of outward transport of mass S determines, among other magnetospheric parameters, the torque and thence, by Eq. (10.12), the total power $\mathcal{P}_{rot}$ extracted from planetary rotation. An example is provided by the simple model of Eq. (10.13) in which R_H, given by Eq. (10.23) in Vol. I, itself depends on S and other magnetospheric and ionospheric parameters. Note that the power along path $\mathcal{P}_{0_{(rot)}}$, direct heating of the atmosphere, is contained in the total, leaving only the difference $\mathcal{P}_{rot} - \mathcal{P}_{0_{(rot)}}$ as the power supplied to the magnetosphere. Because plasma flow is coupled to magnetic flux transport, maintaining the given outflow S imposes self-consistency constraints on other energy flow paths besides III′.

At Jupiter, the average loss rates of energy by radiation (path V″) and by escape of neutral particles (path V′) have been empirically determined and the associated collisional/radiative processes extensively modeled (Thomas *et al.*, 2004; Vol. I, Chapter 10 and references therein). The energy loss in precipitating particles that produce the observed aurora has also been empirically estimated (Clarke *et al.*, 2004, and references therein); the main auroral oval is generally attributed to Birkeland currents of partially corotating plasma (Cowley and Bunce, 2001; Hill, 2001) and is thus part of path II′. Little can as yet be said about power in paths I and II, other than inferences from summing the empirical loss rates.

10.5 What leads to explosive energy releases?

The discussion so far has ignored time variations and has proceeded on the tacit assumptions that all the energy supply, conversion, and dissipation processes are more or less in balance. There is no general requirement for this to be the case, and in fact often it is not the case, as evidenced by the occurrence of rapid or even explosive processes (e.g. substorm onset at Earth). Energy balance presupposes a more general equilibrium of the entire system; as the system evolves in response to, for instance, the changing external boundary, the various terms initially in balance may change differently, so that the system no longer is in equilibrium but varies in time (possibly much faster than the variation of the boundary conditions).

The prototypical example is kinetic energy from the solar wind being converted into magnetic energy of the magnetotail at an increased rate due to enhanced dayside reconnection (in response to changed solar wind conditions), but the rate of removal by conversion of magnetic energy into mechanical energy of magnetospheric plasma plus escape down the magnetotail not being equally enhanced (for reasons that need to be identified). In this case, the magnetic energy reservoir increases with time and reaches a point at which (again, for reasons that need to be identified) the magnetic energy content can no longer be maintained but must be converted to other forms.

10.5.1 Magnetic topological changes

As noted in Section 10.4.2, magnetic flux transport and the increase of magnetic energy by stretching the field play an important role in supplying energy to the magnetosphere. Non-equilibrium configurations of the magnetotail that change the magnetic topology and allow different paths of flux transport are therefore of particular interest. (For a discussion of magnetic topology, see e.g. Vol. I, Chapter 4.)

A simple sketch of a model widely invoked to interpret magnetospheric substorms at Earth is shown in Fig. 10.5 (Vasyliūnas, 1976), which displays a time sequence of magnetospheric configurations. Each panel shows the magnetic field line configuration in the noon–midnight meridian plane (left) as well as the configuration of magnetic singular X- and O-lines in the equatorial plane (right)

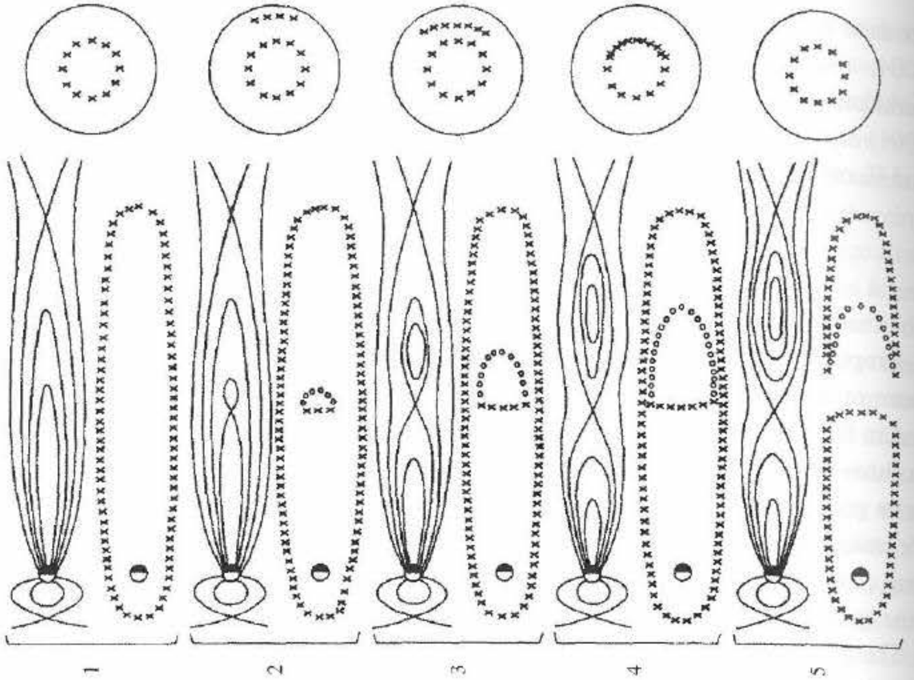

Fig. 10.5. Possible changes of the magnetic field topology in the magnetotail of a solar-wind-dominated magnetosphere. The diagram (from Vasyliūnas, 1976) is shown rotated to facilitate comparisons with diagrams of filament eruptions in e.g. Chapter 6: the solar wind here blows from bottom to top, rather than from left to right as in the original and in the analogous figures of Vol. I, Chapter 10. Each panel in the sequence shows a side view of the magnetic field (left), the outline of the X-lines seen from above the north pole (right), and a top-down view of the mapping of the reconnection region onto the Earth (top).

and projected to the ionosphere (top); the equatorial projection, absent in many later versions (e.g. Hones, 1977), is essential for describing the three-dimensional structure of the magnetic field. Panel 1 is the simplest topology of the open magnetosphere (cf., e.g., Fig. 10.3 in Vol. I). In panel 2, a small volume usually called a *plasmoid* appears deep within the closed-field-line region, bounded on the earthward side by a newly formed *near-Earth X-line* (NEXL) and threaded by magnetic field lines that encircle the attached O-line; ideally, the field lines are confined within the plasmoid and connect neither to the Earth nor to the solar wind (what the real topology is, however, is still uncertain). For the ideal topology, the plasmoid can be visualized in three dimensions as shaped roughly like a banana, oriented approximately dawn to dusk and tapering to zero thickness at both ends, with the X-line on its surface and the O-line running through the middle of its volume. The plasmoid grows (panel 3) by magnetic reconnection until it touches the separatrix

of the open field lines (panel 4, *onset of lobe reconnection*); afterwards (panel 5), the plasmoid is on interplanetary field lines and is carried away (presumably) by the solar wind.

A model of topological changes for a rotation-dominated magnetosphere has been developed (Vasyliūnas, 1983) and likewise widely invoked to interpret events at Jupiter and Saturn believed analogous to magnetospheric substorms at Earth. Shown in Fig. 10.7 in Vol. I, it is in essence a direct adaptation of Fig. 10.5, different only in three respects: (1) the time sequence has been translated into an azimuthal-angle sequence, (2) field lines are stretched by the outflow of plasma from an internal magnetospheric source (planetary/magnetospheric wind) rather than by the flow of the solar wind past the magnetosphere, (3) there are no counterparts to panels 4 and 5, since field lines connected to the solar wind are not considered.

Numerous examples of magnetic topological changes, some similar to those of Fig. 10.5, others more complicated (possible in the absence of a strongly constraining planetary dipole field) have been discussed in relation to solar flares and coronal mass ejections (see Chapter 6).

10.5.2 Role of instabilities

Instabilities have attracted much attention as a possible way of inducing rapid change from equilibrium to non-equilibrium configurations – an alternative to straightforward evolution to non-equilibrium as the result of changing boundary conditions. (Actually, the two possibilities are related: if a system evolves from equilibrium to non-equilibrium, the configuration at the transition point is one of unstable equilibrium.) Specific types of instabilities have been invoked to interpret particular aspects of rapid energy conversion processes in planetary magnetospheres, especially at Earth.

(1) **Tearing-mode instabilities** "Tearing mode" is a generic term for instabilities that result in the reconnection of initially oppositely directed magnetic fields. They are obvious candidates for initiating topological changes of the magnetotail (in particular, those envisaged in Fig. 10.5), as proposed by Schindler (1974) and others; see e.g. Wang and Bhattacharjee (1993); Chapter 10 of Schindler (2007), and references therein.

(2) **Current-driven instabilities** The concept that a sufficiently intense electric current may bring about its own breakdown, by creating conditions that impede current flow, was first suggested by Alfvén and Carlquist (1967) as a model for solar flares. Under the name "current disruption" it has been widely discussed as a model for substorm onset and expansion. Various instabilities that develop when

the current density exceeds some threshold value have been proposed; see e.g. Lui (1996, 2004), and references therein.

(3) **Interchange and ballooning instabilities** Interchange instabilities that do not appreciably change the magnetic field are thought to be essential for plasma transport in rotation-dominated magnetospheres (see Vol. I, Sections 10.5.2 and 13.2.1). Ballooning instabilities can be viewed roughly as interchange that does change the magnetic field. As a model for substorms, they have been invoked particularly at the transition between the dipole field and the magnetotail, in several variants; see e.g. Hurricane *et al.* (1998), Samson (1998), Cheng (2004), and references therein.

10.6 Applications: Earth

Fundamentally, time-varying energy conversion events in the magnetosphere are produced when the various energy-flow paths in Fig. 10.4 are not in balance. The task is to understand which paths are out of balance, on what time scales, and for what physical reasons. The fact that the incident solar wind is itself always varying on many different time scales ensures the occurrence of a whole spectrum of time-varying magnetospheric phenomena, but it also makes it difficult to determine the extent to which they are governed either by internal dynamics of the magnetosphere or by changing solar wind conditions.

10.6.1 Magnetospheric substorms

The phenomenological description of the magnetospheric substorm (sketched briefly in Section 10.2) leads to a physical description that can be summarized (equally briefly) as a two-stage process. Stage 1 (growth phase): as a consequence of a southward interplanetary magnetic field, the configuration of the magnetosphere changes, its magnetic field becoming highly stretched (increased magnetic flux in the magnetotail, reduced flux in the nightside equatorial region). Stage 2 (expansion phase, initiated by the onset): the magnetic field changes to more nearly dipolar (increased flux on the nightside), and there is enhanced energy input and dissipation to the inner magnetosphere and the ionosphere/atmosphere; the process occurs on dynamical time scales (comparable to or shorter than wave travel times) and is accompanied (most probably) by changes of magnetic topology.

In terms of energy flow paths of Fig. 10.4: during stage 1, $\mathcal{P}_{\mathrm{I}}$ (power in path I) is enhanced and is appreciably larger than the sum $\mathcal{P}_{\mathrm{II}} + \mathcal{P}_{\mathrm{II'}} + \mathcal{P}_{\mathrm{III}_{(\mathrm{sw})}}$. During stage 2, $\mathcal{P}_{\mathrm{II}}$ and particularly $\mathcal{P}_{\mathrm{II'}}$ are enhanced; $\mathcal{P}_{\mathrm{III}_{(\mathrm{sw})}}$ and $\mathcal{P}_{\mathrm{III'}}$ presumably are enhanced in connection with topological changes exemplified by Fig. 10.5.

The substorm growth phase is in essence the increase of open magnetic flux in the magnetosphere, which occurs for a two-fold reason. First, the flux addition rate at the dayside reconnection region increases as the solar wind transports more magnetic flux, of the sense opposite to the terrestrial dipole flux, toward the magnetosphere; the reasons for this are assumed to lie in the physics of magnetic reconnection (e.g. Vol. I, Chapter 5). Second, the flux return rate at the nightside reconnection region does *not* increase to match the addition rate; the reasons for this are not at all well understood. One obvious possibility is to assume that the nightside reconnection rate is controlled by local solar wind conditions, just like the dayside rate, so that any increase is delayed by the solar wind flow time to reach the distant X-line of the open magnetosphere (e.g. Fig. 10.3 in Vol. I), but this is unlikely for at least two reasons: (a) in most models, the distant X-line is located well within the magnetotail, not in direct contact with solar wind plasma; (b) any effect of enhanced dayside reconnection can be communicated to the magnetotail by wave propagation within the magnetosphere much faster than by advection in the solar wind. Another possibility is related to stress balance in the magnetotail: the earthward-directed magnetic tension force is opposed by a tailward-directed total (plasma plus field) pressure gradient force, which may impede the earthward flow of plasma and hence the return of magnetic flux. Within the magnetosphere, the net effect of the substorm growth phase is to remove magnetic flux from the nightside magnetosphere by flow toward the dayside reconnection region and to add magnetic flux to the magnetotail (enhanced stretching of magnetotail field lines).

The substorm expansion phase does return the magnetic flux, rapidly and spectacularly, from the magnetotail to the nightside magnetosphere (dipolarization of a previously stretched tail-like field); given that plasma in the magnetotail beyond a distance typically ~15–20 Earth radii is observed to flow away from Earth, the process must almost unavoidably proceed by topological changes of the type sketched in Fig. 10.5. The energy input into plasma, energetic charged particles, and the aurora can be largely accounted for by adiabatic compression and Birkeland current effects. What remains highly controversial is how does the process start and why is it so sudden and catastrophic. Two distinct views have been in contention for decades. One (commonly, albeit inaccurately, called "current disruption model" or sometimes "inside-out scenario") postulates that the substorm onset begins deep within the magnetosphere, at or near the interface between the tail-like and the dipolar magnetic fields, most likely as a result of one or more of the current-driven or ballooning instabilities mentioned in Section 10.5.2; topological changes of the magnetotail are regarded as consequences of the onset. The other ("NEXL model" or "outside-in scenario") postulates the topological sequence of Fig. 10.5 (or some equivalent) as the essential phenomenon and regards the inner-magnetosphere and auroral effects as consequences; the onset itself is identified either with the appearance of the plasmoid (panel 2) or, less commonly, with the onset of lobe reconnection (panel 4). For references and discussion of physical issues distinguishing the models, see e.g. Vasyliūnas (1998).

A further complication is the question of external versus internal influences. That the growth phase is initiated by changing solar wind conditions is the consensus view. The onset and expansion phase, on the other hand, are regarded by the majority as basically the result of internal dynamical processes, although subject to solar wind influences (e.g. if the system is evolving toward instability, it may be pushed over the threshold by a change in the solar wind). A substantial minority, however, considers the substorm onset intrisically as triggered by a solar wind change (typically toward a more northward interplanetary magnetic field).

10.6.2 Magnetic storms

Our understanding of magnetic storms has been decisively influenced by a remarkable theoretical result, the Dessler–Parker–Sckopke theorem, which relates the external magnetic field at the location of a dipole to properties of the plasma trapped in the field of the dipole. First derived by Dessler and Parker (1959) for special pitch-angle distributions and extended to any distribution by Sckopke (1966), the theorem states that $\mathbf{b}(0)$, the magnetic disturbance field of external origin at the location of a dipole of moment $\boldsymbol{\mu}$, satisfies

$$\boldsymbol{\mu} \cdot \mathbf{b}(0) = 2U_K, \qquad (10.18)$$

where U_K is the total kinetic energy content of plasma in the magnetosphere. What is remarkable is that the right-hand side does not depend on the spatial distribution, the partition between bulk-flow and thermal energy, or any properties of the energy spectrum.

Originally derived by Biot–Savart integration of axially symmetric drift currents, the theorem was subsequently derived from a virial-theorem argument and thereby considerably generalized, with the addition of a few terms on the right-hand side (which, however, are mostly ignored in practice except for a negative contribution from solar wind dynamic pressure, $\rho_{sw} v_{sw}^2$); see Carovillano and Siscoe (1973), Vasyliūnas (2006), and references therein.

Although $\mathbf{b}(0)$ nominally is evaluated at the center of the Earth, it is also equal to the (vector) average of $\mathbf{b}(\mathbf{r})$ over the surface of the globe (by a theorem for solutions of Laplace's equation, satisfied within the globe by each Cartesian component). The Dst index is the average, over a low-latitude strip of the globe, of the disturbance field component aligned with the dipole; after some corrections (chiefly removing the contribution from induced earth currents), −Dst may be considered a reasonable proxy for the left-hand side of Eq. (10.18), as long as Dst < 0. The Dessler–Parker–Sckopke theorem then provides a method of inferring the plasma energy content – the energy contained in the box "plasma mechanical energy" in Fig. 10.4 – simply from the value of the Dst index. (The empirical estimates of $\mathcal{P}_{II}$ mentioned in Section 10.2 were obtained largely by this method from observed time variations of Dst.) Direct *in situ* observations have established that the greater part of the energy resides in what is called the ring current region (see Vol. I, Section 10.5.3).

Geomagnetic storms, particularly the intense ones, are characterized by unusually large amounts of energy stored as mechanical energy of plasma in the ring current region, in comparison to other storage regions. This implies that during the development of an intense storm the power in path II is unusually large, on the average. Whether this enhanced conversion rate from magnetic energy into mechanical energy of ring current plasma results from a different interaction process or simply from a different time sequence of solar wind parameters is an unresolved question. More specifically, can the energy for storms be suplied by a sequence of substorms (perhaps unusually frequent and/or unusually intense), or is some other process required? A related question is that of *geoeffectiveness*: when interplanetary structures such as CMEs (see e.g. Chapter 6) impinge on the Earth, under what conditions do they produce intense magnetic storms? (prolonged southward B_{sw} is one that is well established). For discussion and references, see e.g. Tsurutani *et al.* (1997), Kamide *et al.* (1998), and Song *et al.* (2001).

10.7 Applications: other planets

10.7.1 Survey of processes

Our knowledge of energy conversion processes in the magnetosphere of planets other than Earth is strongly conditioned by available observations. For the most part, these are measurements of magnetic fields and charged particles by instruments on spacecraft on flyby trajectories and, more recently, in orbit around the planet. (Remote sensing, e.g. of the aurora, is mostly limited to special campaigns; sufficient observations have been accumulated to establish a reasonable picture of the general morphology of aurora at Jupiter and Saturn, but detailed studies of the time-varying aspects, with the use of concurrent *in situ* observations, are just beginning.) Energy conversion events observed at other planets so far have been generally classified as analogous or at least similar to terrestrial magnetospheric substorms, on the basis of features in the data that resemble what is observed at Earth.

Substorm-like events were first described at Mercury (Siscoe *et al.*, 1975); they occurred during the first Mariner 10 flyby in 1974 and were identified on the basis of observed energetic electron and magnetic field changes that were similar in almost every respect to substorm-related changes in the Earth's magnetotail, except for a much shorter time scale (~20 times faster at Mercury than at Earth). This is in agreement with the supposition that the magnetosphere of Mercury is essentially just a scaled-down version of that of Earth (Ogilvie *et al.*, 1977). (Note: results from the Messenger spacecraft currently in orbit around Mercury are just beginning to be available and have not been taken into account in writing this chapter.)

The magnetosphere of Uranus has been investigated only once, during the flyby of Voyager 2 in 1986. Observed temporal variations of plasma (McNutt *et al.*, 1987), energetic particles (Mauk *et al.*, 1987), and magnetic field (Behannon *et al.*, 1987) were interpreted (on the basis of similarity to observations at Earth) as suggestive of substorm-like events.

By contrast, at Jupiter the extensive data set, from six flybys and above all from the Galileo orbiter, has made it possible to establish unambiguously the existence of characteristic energy conversion events and to determine their main features. These include: magnetic field change, first stretched or more tail-like, followed by relaxed or more nearly dipolar; enhanced plasma flow along approximately the radial direction, alternating between toward and away from the planet; increase of energetic particle intensities, interpreted as heating of the plasma. The duration of an event is typically one to a few hours; there is some indication of a possible recurrence tendency at an interval of a few days. The most common interpretation is that these are rotational counterparts of the terrestrial substorm, involving topological changes similar to those of Fig. 10.5 but driven by the rotational stresses of outflowing plasma rather than by the solar wind drag on open field lines, hence described by some variant (possibly time-dependent or small-scale) of Fig. 10.7 in Vol. I. For references and more detailed description see e.g. Krupp *et al.* (2004); I discuss the physics of the energy conversion briefly in Section 10.7.2.

At Saturn, following two flybys, the accumulation of data by the Cassini orbiter is still in progress. Substorm-like events quite similar to those at Jupiter and interpreted by the same basic concepts have been reported (Jackman *et al.*, 2007; Hill *et al.*, 2008).

10.7.2 Analogs of magnetospheric substorms in strongly rotating magnetospheres

Because the observations at Jupiter and Saturn suggest that the substorm-like events may represent a two-stage process, we may ask how this can be accounted for by imbalances of paths in the energy flow diagram, Fig. 10.4. In a rotation-dominated magnetosphere with internal mass source, the rate of mass flow S along path III′ may be considered as given (Section 10.4.4). Plasma outflow carries magnetic flux with it and would (in the absence of flux return) increase the energy in the magnetic field by stretching the field lines; hence the outward transport magnetic flux may be associated with path IV and the return flux with path IV′. An explosive energy release can now occur in a way that closely parallels the two stages of the terrestrial magnetospheric substorm as described in Section 10.6.1: first, magnetic flux is transported outward, but the return flux is impeded, for a reason to be identified (possibly by the adverse pressure gradient of a stretched-out field, as discussed for Earth); second, a fast return of the accumulated flux is initiated by some process, to be identified (possibly an instability of some type).

10.8 Concluding remarks

Magnetospheric substorms at Earth, analogous events at Jupiter and Saturn, and solar flares and other events discussed in Chapter 6, most of which are interpreted as explosive releases of energy stored in the magnetic field, may perhaps be viewed as manifestations of an underlying universal process, which I summarize tentatively as follows:

(1) The process occurs in two steps: first, mechanical stresses deform the magnetic field (on the Sun, the emergence of new flux – as flux ropes – from below the surface, associated, of course, with a plasma flow, often plays a part in this active-region environment) into a configuration of increased energy; second, the magnetic configuration becomes unsustainable and changes quickly, releasing the energy. Both steps are in general associated with magnetic topological changes.

(2) In most cases, the mechanical stress is related to plasma flow, which transports magnetic flux and, with field lines attached to a massive body, increases the magnetic energy.

(3) Why the magnetic configuration becomes unsustainable and what causes the quick change remain highly disputed questions; many possibilities can be imagined, and there may not be a universal answer.

(4) A potentially universal aspect is magnetic flux return: inability to return the flux smoothly seems to play a role (for Earth at least).