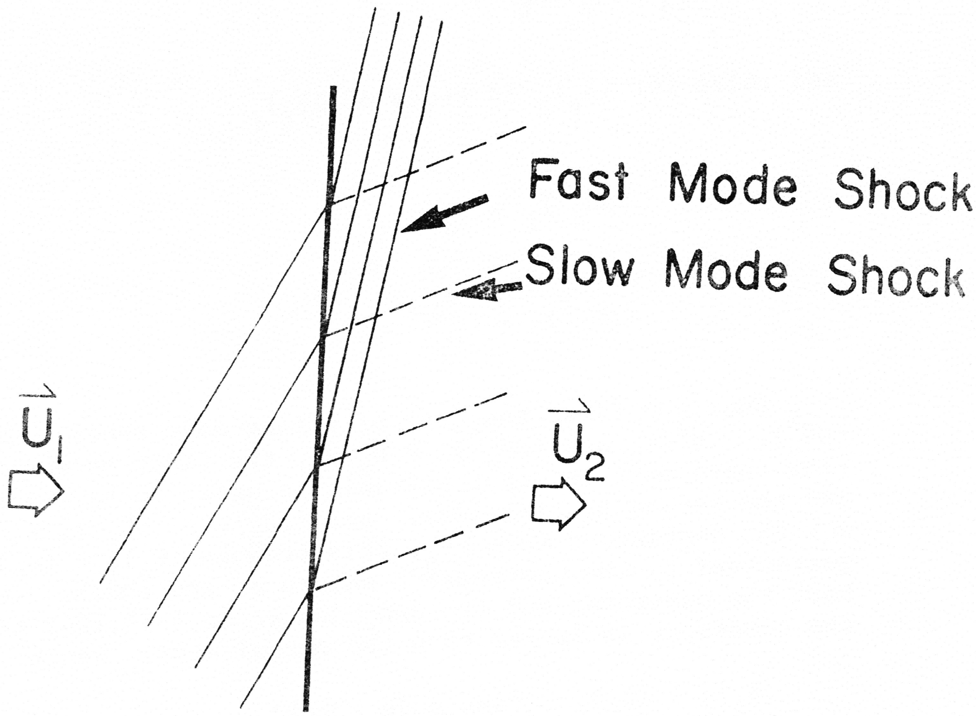

Figure III.4 Magnetic signatures of fast and slow mode shock waves

All three MHD shock modes have been observed in solar system plasmas, especially in the solar wind. Slow mode shocks play an expecially important role in some models of magnetic merging (Vasyliunas, 1975).

IV. MHD INSTABILITIES

This section reviews the four basic instabilities to which MHD fluids are subject. Two of the four are common to all fluids, namely the Kelvin-Helmholtz instability, colloqually known as the "wind over water" instability, and the Rayleigh-Taylor instability, which is called the flute instability in plasma physics and the interchange instability in magnetospheric physics. The other two, the firehose instability and the mirror instability, are caused by differences between $p_{\parallel}$ and $p_{\perp}$ in an anisotropic magnetized plasma, and are therefore peculiar to collisionless MHD fluids. We begin with instabilities driven by pressure anisotropy.

IV.1 The Firehose and Mirror Instabilities

In the theory of small amplitude MHD waves in an anisotropic plasma, these instabilities present themselves in the form of non-propagating, purely exponentially growing waves. The firehose instability is an exponential growth of the intermediate mode, and the mirror instability as an exponential growth of the slow mode. The mirror instability materializes first and grows fastest when the normal to the wave front $\hat{k}$ is oriented nearly perpendicular to the magnetic field.

To confirm the existence of these instabilities, to determine the conditions under which they occur, and to establish their properties, it is necessary to derive the dispersion relation for small amplitude MHD waves in an anisotropic plasma. The analysis proceeds analogously to that carried out in Section III.1 for small amplitude MHD wave in an isotropic plasma, except that in the Euler equation (III.2) the gradient of the scalar pressure is replaced by the expressions for the divergence of the anisotropic pressure tensor given in Section I.16, and the scalar adiabatic relation (III.3) is replaced by the double adiabatic relations (II.80 and 81). Explicitly, this procedure leads to the following form for the plane wave representation of the pressure perturbation

$$\begin{aligned}\nabla\cdot\overleftrightarrow{P} \longrightarrow &- \left[\frac{\delta\rho}{\rho_o}(p_\perp)_o + \frac{\delta B_z}{B_o}(p_\parallel)_o\right] k_x\hat{x} \\ &- \left[\frac{\delta B_z}{B_o}(p_\perp)_o + 3\left(\frac{\delta\rho}{\rho_o} - \frac{\delta B_z}{B_o}\right)(p_\parallel)_o\right] k_z\hat{z} \\ &+ \frac{(p_\parallel - p_\perp)_o}{B_o^2/\mu_o}\,\vec{J}\times\vec{B}\end{aligned} \qquad \text{(IV.1)}$$

in which the unit imaginary number i has been suppressed since it ultimately multiplies all terms, and $\vec{J}\times\vec{B}$ has not been replaced by its wave perturbation form, because this reduction has already been given in (III.14). The equation is referenced to the coordinate system which was used previously and which is shown in Figure III.1.

The wave perturbation forms of the continuity equation and Faraday's induction law (eq's III.13 and 15) are now used to eliminate the density and field perturbations, leading to a single equation for the velocity perturbation, corresponding to eq. (III.20) for the isotropic pressure case.

$$\begin{aligned}\omega^2\delta\vec{V} = &\left[C_\perp^2 k_x(\vec{k}\cdot\delta\vec{V}) - C_\parallel^2 k_x^2\delta V_x\right]\hat{x} \\ &+ \left[C_\perp^2 k_x k_z\delta V_x + 3C_\parallel^2 k_z^2\delta V_z\right]\hat{z} \\ &+ \xi_o C_A^2\left[k^2\delta V_x\hat{x} + k_z^2\delta V_y\hat{y}\right]\end{aligned} \qquad \text{(IV.2)}$$

in which we have defined for notational convenience the (pseudo) anisotropic sound speeds

$$C_{\perp}^2 = \frac{(p_{\perp})_o}{\rho_o} , \quad C_{\parallel} = \frac{(p_{\parallel})_o}{\rho_o} \qquad \text{(IV.3)}$$

The symbols C_A^2 and ξ have been defined previously (eq's III.17 and 64).

As before the procedure from this point is to write out explicitly the three vector components of (IV.2), identify the coefficient matrix of the dependent variables δV_x, δV_y and δV_z, and set the determinant of that matrix to zero to obtain the dispersion equation for the waves. Again however the y-component contains only the independent variable δV_y, and therefore is itself a pure mode. One obtains the anisotropic form of the dispersion equation for the intermediate mode wave directly from the y-component,

$$\omega_i^2 = \xi_o \; C_A^2 \; k_z^2 \qquad \text{(IV.4)}$$

Comparison with the isotropic form of this equation (III.21) shows that the effect of anisotropy is the introduction of the multiplicative factor ξ_o. Whereas the isotropic form of this equation is positive definite, ξ_o may be negative in which circumstance the intermediate mode will exhibit non-propagating, pure exponential growth. This is the firehose instability.

The threshold for the onset of the instability is $\xi_o = 0$. In general we may classify the behavior of the mode according to whether ξ_o is positive, zero or negative.

$\xi_o > 0$, Propagating intermediate mode

$\xi_o = 0$, Non-propagating, non-growing perfectly inelastic perturbations

$\xi_o < 0$, Firehose instability (non-propagating, pure exponential growth)

In the propagating wave regime ($\xi_o > 0$), the phase speed (and the group velocity) can be greater or less than its value for the isotropic case (eq. III.23) depending on whether ξ_o is greater than or less than unity.

$$(V_{ph})_i \text{ (anisotropic)} = \sqrt{\xi_o} \, (V_{ph})_i \text{ (isotropic)} \qquad \text{(IV.5)}$$

To understand the physical reason for the dependence of ω_i on ξ_o in (IV.4), it is useful to rewrite the expression for ξ_o in the form

$$\xi_o - 1 = \frac{(p_{\perp} - p_{\parallel})_o}{B_o^2/\mu_o} \qquad \text{(IV.6)}$$

The right hand side of (IV.6), which contains the full effect of pressure anisotropy on this mode, shows that one must consider the contributions that all three pressures, $p_\perp$, $p_\parallel$ and $B_o^2/2\mu_o$, make to the frequency of the wave. Recall that in the isotropic case, the frequency is fixed by balancing the inertial force exerted by a volume of plasma that is oscillating transversely to the magnetic field against the magnetic tension the motion engenders in stretching the field. Since an increase in the restoring force increases the frequency, and the right hand side of (IV.6) measures the change in frequency resulting from pressure anisotropy, it is evident that $p_\perp$ acts to increase the restoring force and $p_\parallel$ acts to decrease it. When the two pressures are equal (i.e. isotropy) their effects cancel. It can be seen qualitatively from Figure IV.1 that both effects can be described in terms of a centrifugal force. The centrifugal force exerted by the thermal motions of particles moving parallel to the bent flux tube acts against the magnetic tension, which is attempting to straighten the tube. The centrifugal force exerted by the thermal motions of particles gyrating perpendicular to the flux tube acts against the tension on the oustide but with the tension on the inside. However, the bend increases the density of particles gyrating on the inside and decreases it on the outside, thus there is a net force tending to straighten the tube.

When $p_\perp > p_\parallel$ the combined pressure effect and magnetic tension increase the restoring force over that of pure isotropy, the frequency and wave speed therefore increase. When $p_\parallel > p_\perp$, the reverse occurs, and if the imbalance $p_\parallel - p_\perp$ should exceed the magnetic tension B^2/μ_o, the net restoring force becomes negative, and the bend grows under the force the bend itself produces.

Consider next the dispersion equation for the two compressive modes in an anisotropic plasma. Setting to zero the determinant of the coefficient matrix obtained from the x and z-components of (IV.2) gives

$$(\omega_{f,s}^2 - 3C_\parallel^2 k_z^2)\ [\omega_{f,s}^2 - (2C_\perp^2 + C_\parallel^2)\ k_x^2 - \xi_o\ C_A^2\ k_z^2] - C_\perp^4\ k_x^2\ k_z^2 = 0 \tag{IV.7}$$

Before isolating the mirror instability, it is instructive to see how parallel propagation and perpendicular propagation are modified by the anisotropy. As in the isotropic case, when $k_x = 0$, equation (IV.7) has two solutions which can be written in terms of the phase speed.

$$\vec{k}\,||\,\vec{B}_o\ :\ (V_{ph}^2)_{f,s} = \begin{cases} 3C_\parallel^2 \\ \xi_o C_A^2 \end{cases} \tag{IV.8}$$

The first of these corresponds to a sound wave propagating parallel to the field in a gas for which $\gamma = 3$, which is the value appropriate to one degree of thermal freedom. The second is the solution for the intermediate mode wave (eq. IV.5) with $\theta = 0$. When $k_z = 0$, equation (IV.1) again has two solutions

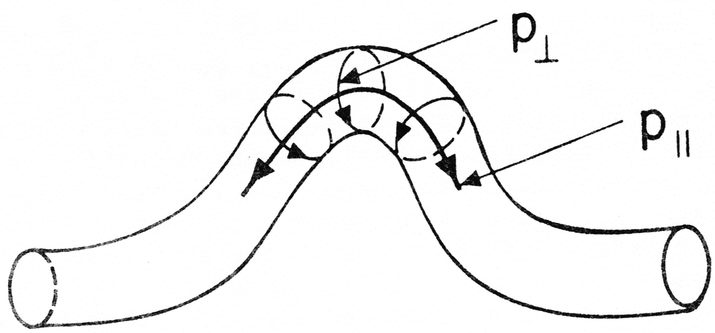

Figure IV.1 Magnetic flux tube in "firehose" configuration.

$$\vec{k} \perp \vec{B}_o : (V_{ph})^2_{f,s} = \begin{cases} 2C_\perp^2 + C_A^2 \\ 0 \end{cases} \tag{IV.9}$$

The first corresponds to the perpendicular fast mode wave, which has a phase speed equal to the Pythagorean sum of the Alfvén speed and the sound speed of a gas with two degrees of freedom. The second solution corresponds to the slow mode wave, which does not propagate perpendicular to the magnetic field.

That a slow mode wave with its wave normal oriented nearly perpendicular to the magnetic field can be unstable is seen by utilizing the approximations that apply to this mode and this orientation. These are $\omega_s \sim 0$ (and therefore the ω_s^4 term can be dropped compared to the ω_s^2 term) and $k_z^2/k_x^2 \ll 1$. Then the dispersion equation, can be solved for ω_s^2/k^2

$$\frac{\omega_s^2}{k^2} = \frac{3C_\parallel^2(2C_\perp^2 + C_A^2) - C_\perp^4}{2C_\perp^2 + C_A^2} \tag{IV.10}$$

The right hand side is negative, and therefore the wave is unstable, if

$$\frac{C_\perp^2}{C_\parallel^2} > 6\left(1 + \frac{C_A^2}{2C_\perp^2}\right) \tag{IV.11}$$

Rewriting the instability in terms of pressure ratios, (IV.11) becomes

$$\left(\frac{p_\perp}{p_\parallel}\right)_o > 6\left(1 + \frac{1}{\beta_\perp}\right) \quad \text{(MHD)} \tag{IV.12}$$

in which the symbol β is defined in general in plasma physics as the ratio of the particle pressure to the magnetic field pressure. In this case it is

$$\beta_{\perp} \equiv \frac{(p_{\perp})_o}{B_o^2/2\mu_o} \qquad (IV.13)$$

The criterion for the existence of the mirror instability given by eq. (IV.12) was derived with the use of the double adiabatic invariants of MHD. The qualification (MHD) has accordingly been affixed to the result. A treatment of this instability by the use of plasma kinetic theory, which takes into account diabatic heating by parallel heat flux, arrives at a similar result, but the factor of six in this case is replaced by unity (e.g. Krall and Trivelpiece, 1973)

$$\left(\frac{p_{\perp}}{p_{\parallel}}\right)_o > 1 + \frac{1}{\beta_{\perp}} \quad \text{(Kinetic Theory)} \qquad (IV.14)$$

A sketch of mirror geometry in the magnetic field produced by an obliquely propagating compressive mode wave is shown in Figure IV.2. In the case of a slow mode wave, the particle pressure is strongest where the magnetic pressure is weakest, namely, in the middle of the magnetic bottles formed by the periodic constriction of each flux tube. If the mirror instability criterion is met, the increase in the distabilizing component of the pressure, $p_{\perp}$, that attends an oblique slow mode perturbation exceeds the increase in the restraining tensions in the field and in $p_{\parallel}$, and the perturbation grows.

IV.2. The Kelvin-Helmholtz Instability

The pressure perturbations that arise in a fluid when it is forced to flow over a wavy wall are such that if the wall were non-rigid the amplitude of the wave in the wall would tend to grow. The example commonly given to illustrate the preceding statement is the waves that form on the surface of open water on a windy day. There are two equivalent ways of understanding the origin of the destabilizing perturbations in this situation. In one we are to notice that the centrifgual force generated in the fluid as it moves along the wall in a serpentine motion to conform to waves in the wall are such as to apply extra force on the wall in the troughs and to diminish the force on the wall in the crests. The troughs are thereby impelled to deepen and the crests to heighten.

An alternative (but, it should be emphasized, equivalent) way to understand the phenomenon focuses on the change in the pressure in an element of fluid next to the wall as it descends into a trough or surmounts a crest of the wave. In the frame of reference of the wall, these pressure changes are of course time stationary and are characteristic of troughs and crests generally. The essential physical principle

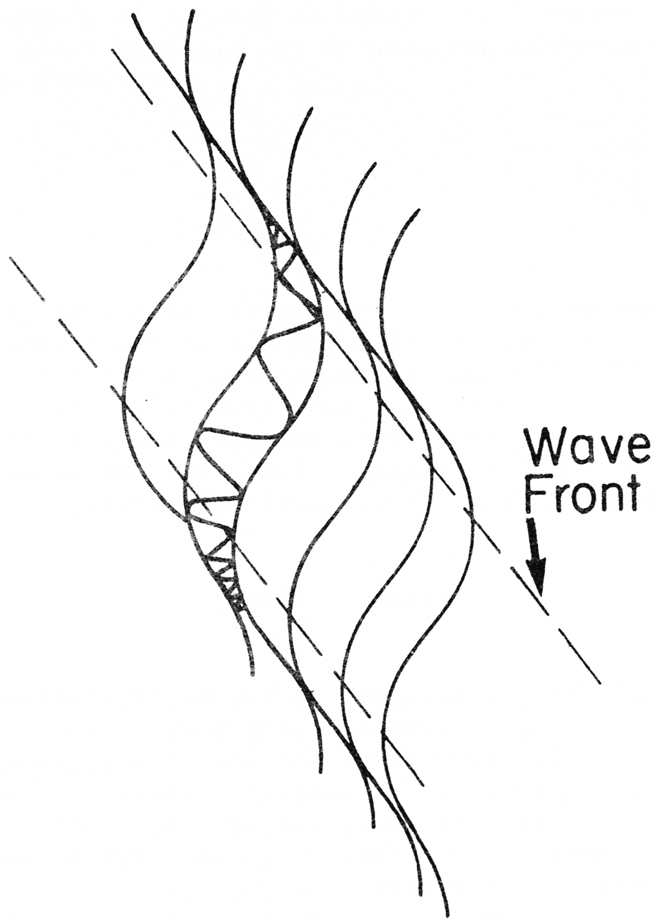

Figure IV.2 Oblique slow mode wave showing the "mirror" geometry prerequisite for the mirror instability. The path of a particle bouncing between two mirror points is indicated.

is most clearly seen if we consider the case of an incompressible fluid, that is one for which the density is constant in space and time. The distinguishing equation for an incompressible fluid is

$$\nabla . \vec{V} = 0 \tag{IV.15}$$

which follows from the constant density condition and the continuity equation (I.22). In ordinary (non-magnetic) fluid dynamics in which gravity and viscosity are ignored, an incompressible fluid satisfies the following Bernoulli equation (from I.89 and 92 with the isometric polytropic index $n = \infty$)

$$\frac{1}{2} \rho V^2 + p = \rho W \tag{IV.16}$$

The right hand side of (IV.16) is a constant. This equation expresses the well-known Venturi effect: when along a given streamline the velocity of the fluid increases the pressure drops, and vice versa.

The Venturi effect explains why the pressure distribution in a fluid moving along a wavy wall acts in the sense to amplify the wave. As we shall see shortly, the perturbation caused by the wavy wall decreases exponentially away from the wall with a scale length equal to $\lambda/2\pi$, where λ is the wave length of the wave in the wall. For many purposes a problem involving a perturbation that decreases exponentially with a scale length ℓ can be replaced by a problem with a constant perturbation extending over the same distance ℓ, but with zero perturbation beyond this distance. Adopting for a moment the second description, one can say that the zero-order area through which the wave-affected part of the fluid flows is given by some constant width in the direction parallel to the wave troughs and crests multiplied by a height equal to the exponential scale length $\lambda/2\pi$. Consider then how the perturbation in the area caused by the wavy wall effects the flow velocity of the fluid. In a steady incompressible flow the total quantity of fluid passing over a trough must be equal to the total quantity of fluid passing over a crest in the same interval of time. But a trough increases the area available to the flow and a crest decreases it. The velocity must then change to compensate for the change in area to keep the total flow continuous. Thus, the velocity over a trough must be less than the velocity over a crest. By the Venturi effect, the pressure is therefore increased over a trough and decreased over a crest. As in the case of the explanation in terms of the centrifugal force, the differential force on the wall exerted by the pressure acts to amplify the wave. The actual force that the fluid exerts on the wall is the pressure force. A demonstration of the equivalence between the explanation in terms of centrifugal force and pressure force follows upon deriving an explicit expression for the pressure perturbation and comparing it with the centrifugal force associated with the fluid motion over the wavy wall.

It is convenient for the purpose of fixing ideas with a concrete example to begin a treatment of the Kelvin-Helmholtz instability with a discussion of the flow of fluid past a wavy wall. In retrospect one can see that the essential effect does not depend on the presence of a physical wall, but rather it depends only on the existence of a shear in the flow velocity of the fluid. If there exists a frame of reference in which one layer of the fluid is at rest (representing the wall) while an adjacent layer is moving, any naturally occurring deviation from smoothness in the interface between the two layers will grow under the resulting pressure perturbation. Manifestations of the instability therefore should be nearly as plentiful as appearances of sheared flows. However in many situations fluids exhibit some degree of elasticity and resist the growth of the initial perturbation. Elastic resistance to the Kelvin-Helmholtz instability results from surface tension in the case of wind-over-water waves, stable stratification of the atmosphere in the case of atmospheric gravity waves, and magnetic tension in the case of surface MHD waves. In each of these instances the shear in velocity must exceed some threshold value before the onset of unstable growth occurs.

To illustrate the general nature of the treatment of the Kelvin-Helmholtz instability in a MHD context, we consider here the problem of a sheared flow between two incompressible, isotropic, inviscid MHD fluids separated by a tangential discontinuity. Gravity will be ignored. The example has application to the magnetopauses of planetary magnetospheres, the quasi-equatorial current sheet in the solar wind and the myriad accidental tangential discontinuities lacing the solar wind.

On both sides of the discontinuity, the continuity equation for an incompressible fluid given by (IV.15) applies as well as the Euler and hydromagnetic equations, which we repeat here for convenience

$$\rho \frac{d\vec{V}}{dt} + \nabla p = - \frac{\vec{B} x \nabla x \vec{B}}{\mu_o} \qquad \text{(IV.17)}$$

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{V} x \vec{B}) \qquad \text{(IV.18)}$$

As in the treatment of the MHD plane waves, the variable parameters, $\vec{V}$, p, and $\vec{B}$, are decomposed into zero-order and perturbation parts.

$$\vec{V} = \vec{V}_o + \delta\vec{V} \qquad \text{(IV.19)}$$

$$p = p_o + \delta p \qquad \text{(IV.20)}$$

$$\vec{B} = \vec{B}_o + \delta\vec{B} \qquad \text{(IV.21)}$$

Let (x,y,z) be a Cartesian coordinate system in which the z-axis is normal to the plane of the discontinuity. Then $\vec{V}_o$ and $\vec{B}_o$ lie in the (x,y) plane. Assume the perturbations have the form of propagating plane waves along the surface of the discontinuity, but decay in strength away from the discontinuity. This is the cannonical form of linearized plane surface waves. That is the perturbation of any quantity Q has the explicit space and time dependence given by

$$\text{Perturbation (Q)} = \delta Q\, e^{i(\omega' t - \vec{k}_t \cdot \vec{r}) - k_z z} \qquad \text{(IV.22)}$$

where δQ is the (possibly complex) amplitude of the perturbation, ω' is the Doppler shifted frequency of the wave (that is the frequency in our frame of reference), $\vec{k}_t \equiv k_x\hat{x} + k_y\hat{y}$ is the component of the propagation vector that lies in the plane of the discontinuity, and k_z is the reciprocal of the scale length for the exponential decay of the strength of the perturbation in the direction normal to the surface of the discontinuity. Thus, we must have $k_z>0$ for $z>0$ and $k_z<0$ for $z<0$.

Substitution of the perturbation forms of the dependent variables with the plane surface wave assumption into the continuity, Euler and hydromagnetic equations gives respectively

$$\vec{\kappa}_{\pm} \cdot \delta\vec{V} = 0 \qquad \text{(IV.23)}$$

$$\omega\rho\delta\vec{V} - \vec{\kappa}_{\pm}\,\delta p = -\,(\vec{B}_o \cdot \vec{k})\,\frac{\delta\vec{B}}{\mu_o} + \left(\frac{\vec{B}_o}{\mu_o} \cdot \delta\vec{B}\right)\vec{\kappa}_{\pm} \qquad \text{(IV.24)}$$

$$\omega\delta\vec{B} = -\,(\vec{B}_o \cdot \vec{k})\,\delta\vec{V} \qquad \text{(IV.25)}$$

where $\omega = \omega' - \vec{V}_o \cdot \vec{k}$ is the frequency in the plasma rest frame (cf. eq. III.19), and the complex wave number $\kappa_{\pm}$ is defined by

$$\vec{\kappa}_{\pm} \equiv k_x\hat{x} + k_y\hat{y} \pm ik_z\hat{z} \qquad \text{(IV.26)}$$

in which the + sign applies for $z>0$ and the - sign for $z<0$. By this definition of $\vec{\kappa}_{\pm}$, we have arranged for k_z to be everywhere positive. In the combinations $\vec{V}_o \cdot \vec{k}$ and $\vec{B}_o \cdot \vec{k}$, the subscript t on $\vec{k}$ is superfluous.

The special properties of the boundary waves are determined by the applicable continuity relations. The total pressure is continuous across a tangential discontinuity (eq. III.66) and the displacement of the two fluids in the z-direction must be continuous in order to avoid separation or interpenetration of the fluids. Let P_T designate the total pressure (i.e. $P_T = p + (B^2/2\mu_o)$). Then the two continuity relations can be expressed as

$$[[P_T]] = 0 \qquad \text{(IV.27)}$$

$$[[\delta z]] = 0 \qquad \text{(IV.28)}$$

Now

$$\delta P_T = \delta p + \frac{1}{\mu_o}\vec{B}_o \cdot \delta\vec{B} \qquad \text{(IV.29)}$$

The quantities δp and $\delta\vec{B}$ can be eliminated in favor of $\delta\vec{V}$ through the use of (IV.24 and 25) resulting in

$$\delta P_T \vec{\kappa}_{\pm} = [\omega^2 - (\vec{V}_A \cdot \vec{k})^2]\,\frac{\rho\delta\vec{V}}{\omega} \qquad \text{(IV.30)}$$

where $\vec{V}_A = \vec{B}_o/\sqrt{\mu_o\rho}$ is the Alfvén velocity. The scalar product of (IV.30) with $\vec{\kappa}_{\pm}$ together with (IV.23) show that

$$\delta P_T\,\kappa_{\pm}^2 = 0 \qquad \text{(IV.31)}$$

Thus either $\delta P_T = 0$ or $\kappa_{\pm}^2 = 0$. The first option leads directly to the usual relations for an intermediate mode MHD wave, as can be seen from eq. (IV.30), which in this case gives $\omega^2 = (\vec{V}_A \cdot \vec{k})^2$. The second option is specific to surface waves, and yields the important relation

$$k_t^2 = k_z^2 \tag{IV.32}$$

This expresses the condition stated earlier that the decay length for the strength of the perturbation away from the discontinuity (or wall) is $k_z^{-1} = k_t^{-1} = \lambda/2\pi$, where λ is the wave length of the surface wave.

The two continuity relations (IV.27 and 28) will now be used together with the expression for the perturbation in the total pressure (IV.30) to obtain the dispersion equation for the surface waves. The mathematical procedure has the following structure. The quantity $\delta\vec{V}$ in (IV.30) will be eliminated in favor of δz to arrive at a relation that has the algebraic form

$$\delta P_T = A\ \delta z \tag{IV.33}$$

where the quantity A will be determined below. Then the condition $[[P_T]] = 0$ gives $[[A\delta z]] = 0$. But since δz is also continuous it can be factored out, resulting in

$$[[A]] = 0 \tag{IV.34}$$

Equation (IV.34) with A made explicit is the dispersion equation.

To find the expression represented by A, scalar multiply equation (IV.30) by $\hat{z}$ and solve for δP_T

$$\delta P_T = \frac{1}{\pm i k_z} [\omega^2 - (\vec{V}_A \cdot \vec{k})^2] \frac{\rho\ \delta\vec{V}\cdot\hat{z}}{\omega} \tag{IV.35}$$

But

$$\delta\vec{V}\cdot\hat{z} = \frac{d\delta z}{dt} = \frac{\partial\delta z}{\partial t} + \vec{V}_o \cdot \nabla\delta z = i(\omega' - \vec{V}_o \cdot \vec{k})\ \delta z = i\omega\delta z \tag{IV.36}$$

Thus

$$A = \mp \rho[\omega^2 - (\vec{V}_A \cdot \vec{k})^2] \tag{IV.37}$$

The common factor k_z has been dropped from the final expression since it would cancel out in the continuity relation (IV.34).

If we denote quantities that refer to the plasma above the discontinuity ($z>0$ corresponding to the + sign) by the subscript 1 and quantities that refer to the plasma below the discontinuity by the subscript 2, the discontinuity relation (IV.34) becomes

$$\rho_1[\omega_1^2 - (\vec{V}_{A1}\cdot\vec{k})^2] + \rho_2[\omega_2^2 - (\vec{V}_{A2}\cdot\vec{k})^2] = 0 \qquad \text{(IV.38)}$$

This can be rewritten as a quadratic equation for the frequency ω' in the initially chosen frame of reference. This is the observed frequency and it must be the same for both sides of the discontinuity. The solution of the quadratic equation for ω' is

$$\omega' = \frac{\rho_1\vec{V}_1\cdot\vec{k}+\rho_2\vec{V}_2\cdot\vec{k}}{\rho_1+\rho_2} \pm \frac{1}{(\rho_1+\rho_2)}\{(\rho_1+\rho_2)[\rho_1(\vec{V}_{A1}\cdot\vec{k})^2 + \rho_2(\vec{V}_{A2}\cdot\vec{k})^2] - \rho_1\rho_2[(\vec{V}_1-\vec{V}_2)\cdot\vec{k}]^2\}^{\frac{1}{2}} \qquad \text{(IV.39)}$$

in which the subscripted zeros on $\vec{V}_1$ and $\vec{V}_2$ have been dropped.

It is evident that ω' will have an imaginary part corresponding to the functioning of the Kelvin-Helmholtz instability if the argument of the radical is negative. It can be quickly verified from this equation that the condition for the operation of the instability can be expressed in terms of a threshold condition on the velocity shear $\Delta\vec{V} \equiv \vec{V}_1-\vec{V}_2$, namely

$$(\Delta\vec{V}\cdot\vec{k})^2 > \frac{1}{\mu_o}\left(\frac{1}{\rho_1}+\frac{1}{\rho_2}\right)[(\vec{B}_1\cdot\vec{k})^2 + (\vec{B}_2\cdot\vec{k})^2] \quad \text{(Kelvin-Helmholtz instability)} \qquad \text{(IV.40)}$$

in which $\vec{B}_1$ and $\vec{B}_2$ are the zero-order fields on the two sides of the discontinuity.

In order to make the left hand side of (IV.40) as large as possible for a given $\Delta\vec{V}$, choose $\vec{k}$ to be parallel (or antiparallel) to $\Delta\vec{V}$. Then it is apparent the value which $\Delta\vec{V}$ must exceed in order for the instability to operate depends on the size of components of $\vec{B}_1$ and $\vec{B}_2$ parallel (or antiparallel) to $\Delta\vec{V}$. This is readily understood to be a consequence of the stretching by the wave of that component of $\vec{B}$ (on either side) that lies parallel to $\vec{k}$. The stretching occurs because if there is a component of $\vec{B}$ parallel to $\vec{k}$, the field lines cut across the troughs and crests of the wave and are stretched in length, according to their obliquity relative to k, in proportion to the ratio that the area of a wavy surface makes to a smooth one. Since the tension inherent in a magnetic field resists any force acting to stretch the field, the velocity shear is required to exceed a certain value given by (IV.40) in order to overcome this resistance. Note that if $\Delta\vec{V}$ is perpendicular to $\vec{B}$ on both sides of the discontinuity, the right hand side of (IV.40) is zero for a wave propagating parallel to $\Delta\vec{V}$. In this case the wave does not stretch the field lines and the surface is unstable for arbitrarily small values of ΔV.

As a final observation concerning the properties of MHD surface waves, consider the wavelength dependence of the instability. In the expression for the threshold criterion (IV.40), the wave vector $\vec{k}$ multiplies all terms. The expression would be unaffected numerically therefore if $\vec{k}$ were replaced by the unit vector pointing in the propagation direction, $\hat{k}$. The threshold criterion is thereby seen to be independent of the wavelengths of the surface wave. That is, if the surface is unstable for one wavelength, it is unstable for all wavelengths. On the other hand, the dispersion equation (IV.39) shows that when the instability criterion is met, the imaginary part of the frequency is directly proportional to k. Thus the growth rate of the instability is greatest for small wavelength waves. However, we have treated only the linear problem. The result concerning the growth rate therefore only implies that the short wavelength waves reach their nonlinear form more quickly than the long wavelength waves. The linear treatment can not predict which wavelength waves will have the largest amplitude after they have evolved into the nonlinear domain.

IV.3 The Magnetospheric Interchange Instability

The magnetospheric interchange instability is a particular type of Rayleigh-Taylor instability in a MHD fluid. The Rayleigh-Taylor instability occurs whenever an adiabatic interchange of fluid parcels results in a reduction of stored energy, whether the energy be stored as potential energy or as kinetic energy of compression. The kinetic energy associated with the interchange motion can then be supplied by the release of stored energy, the motion becomes self-propelled, and the initial arrangement of fluid parcels proves to be unstable.

In the most commonly cited example of a Rayleigh-Taylor instability, one incompressible fluid overlies another which has a smaller mass density, or more informally expressed, a heavy fluid overlies a light one. Then an interchange of a parcel of the heavy fluid and an equal volume of light fluid results in lowering their common center of gravity, thereby releasing stored gravitational energy. Such interchanges therefore will occur spontaneously, overturning the unstable configuration until the light fluid completely overlies the heavy fluid. In contrast to the initial state, the final state is stably stratified.

For this simple example it is a trivial matter to write down a mathematical criterion which must be satisfied in order for the instability to occur

$$\vec{g} \cdot \nabla \rho < 0 \quad \text{(unstable)} \qquad \text{(IV.41)}$$

In (IV.41), $\vec{g}$ is the force of gravity, but it could as well represent any inertial force such as the centrifugal force. The combination of gravitational and centrifugal forces, which is called the geopotential force, is used to describe the "effective" gravitational force in the frame of reference corotating with a planet. Later in this subsection we will need to use the geopotential force and we will denote it by $\vec{g}^*$. Then $\vec{g}$ in (IV.41) is replaced by $\vec{g}^*$.

The Rayleigh-Taylor instability manisfests itself in stellar and planetary atmospheres as well, but here the compressibility of the fluid (in this case a gas) must be taken into account. To guarantee stability, it is not sufficient that a less dense gas overlie more dense gas since in a vertical interchange of gas parcels, a descending parcel is compressed adiabatically and becomes denser as it moves through increasing atmospheric pressure. Conversely, an ascending parcel expands adiabatically and becomes less dense as it moves through decreasing atmospheric pressure. Thus, instead of comparing mass densities at two different levels, it is necessary to compare the masses in volumes that increase with height according to the adiabatic relation

$$V = c\ (p_{st})^{-\frac{1}{\gamma}} \tag{IV.42}$$

which follows from eq. (I.82) and the definition of a fluid parcel, which entails a volume of fixed total mass and thus which obeys ρV=const. The pressure p in (IV.42) is meant to be the actual pressure of the atmosphere. To make this designation explicit, we have used the subscript st, which denotes "structual". For a given, fixed value of the parameter c, atmospheric parcels with volumes given by (IV.42) can be interchanged vertically with no change in the volume of the surrounding gas. Thus no work in the form of compression attends such interchanges, and the only change in energy can result from a change in the gravitational potential. It is now evident by direct analogy with (IV.41) the instability criterion in this case is

$$\vec{g} \cdot \nabla\ (\rho_{st} V) < 0 \quad \text{(unstable)} \tag{IV.43}$$

in which $\rho_{st}V$ is the mass of volume-equivalent fluid parcels. Again $\vec{g}$ can be replaced by $\vec{g}^*$, which represents the generalized inertial force.

It is convenient to cast (IV.43) in terms of more readily available observables. To do this first note that

$$\rho_{st} V = c\ \rho_{st} (p_{st})^{-1/\gamma} \tag{IV.44}$$

and recall that the specific entropy s is proportional to $\ell n(p/\rho^{\gamma})$. In terms of s_{st}, the instability criterion becomes (upon multiplying (IV.43) by $-\gamma/\rho_{st}V$)

$$\vec{g} \cdot \nabla\ s_{st} > 0 \quad \text{(unstable)} \tag{IV.44}$$

Equation (IV.44) states that an atmosphere is unstable if the specific entropy decreases with height. Of course, the specific entropy is also not a readily available observable, but equation (IV.44) reveals that the condition of isentropy (s_{st} = const.) divides stable from unstable atmospheres. This suggests that the instability criterion should be expressed by reference to a convenient variable in an adiabatically stratified atmosphere. Since the variation of pressure with height in an atmosphere changes in time much less than the vertical variation of temperature does, the vertical pressure profile is usually regarded as fixed and known. Then the family of temperature profiles can be used for reference purposes that satisfys the adiabatic relation with respect to the given structural pressure profile (see eq's I.62 and 82)

$$T_{ad}(p_{st})^{-\frac{\gamma-1}{\gamma}} = \text{const.} \tag{IV.45}$$

where the constant on the right hand side is the family parameter. To arrive at an instability criterion in terms of a comparison between the vertical profiles of the structural temperature and the adiabatic temperature given by (IV.45), eliminate ρ_{st} on the right hand side of (IV.44) by use of the ideal gas law (I.62) to find

$$\frac{1}{\rho_{st}V} = c_1\, T_{st}(p_{st})^{-\frac{\gamma-1}{\gamma}} = c_2\, T_{st}(T_{ad})^{-1} \tag{IV.46}$$

in which the second equality follows from (IV.45), and c_1 and c_2 are constants. Equation (IV.46) is substituted into (IV.43) and evaluated with T_{ad} initiallized to T_{st}. Then the instability criterion becomes

$$\vec{g} \cdot (\nabla T_{st} - \nabla T_{ad}) > 0 \quad \text{(unstable)} \tag{IV.47}$$

In atmospheric application, instead of referencing directions of change to $\vec{g}$, it is usual to use height, z. Then noting that $\vec{g} = -g\hat{z}$, we may write (IV.47) in the most common form for the instability criterion

$$\frac{dT_{st}}{dz} < \frac{dT_{ad}}{dz} \quad \text{(unstable)} \tag{IV.48}$$

Since temperature normally decreases with height, equation (IV.48) states that an atmosphere is unstable if the temperature decreases in it more rapidly with height than in an adiabatic atmosphere.

With the incompressible and compressible forms of the gravitational Rayleigh-Taylor instability treated as preliminary examples, we turn next to consider the algebraically somewhat more complicated case of the magnetospheric interchange instability. We are here concerned with a planetary magnetosphere or portion of magnetosphere for which for mathematical convenience the following idealizations and approximations may be made. The magnetic field is a pure dipole field. The dipole axis is parallel to the rotation axis of the planet. The kinetic energy density of the plasma in the magnetosphere is small compared to the magnetic energy density. This last condition is to ensure that the magnetic field retains its dipole geometry, that is that the distortions caused by magnetospheric currents may be ignored. As noted by Gold (1959), the magnetic field can nevertheless undergo interchange motions as a result of forces generated by the plasma. This type of motion merely involves the interchange of entire magnetic flux tubes, all of which enclose the same quantity of magnetic flux. Such interchange motions can occur without causing any change in the magnetic field configuration, and thus entail no change in magnetic energy. It is clear that interchange motions take the form of circulations, since each flux tube that moves to fill the place of another must have its place filled in turn. An example of an interchange circulation pattern is shown in Figure IV.3.

It should be noted explicitly that the notion of interchanging flux tubes applies to a plasma for which the hydromagnetic approximation is valid. Then by the freezing law, the flux tube plays the role of a fluid parcel. It retains a constant quantity of plasma within it as it moves in interchanging circulations. As we shall see, an important difference between this and the previous compressible case we studied is that the volume of a flux tube is fixed by the quantity of magnetic flux it contains and its position in the magnetosphere. By our "low β" assumption, the volume of the flux tube is independent of the pressure of the gas it contains.

If an interchange motion such as the one depicted in the figure results in reducing the amount of energy stored in the plasma which is enclosed by the participating flux tubes, it is reasonable to assume that the motion will occur spontaneously, driven by the released energy. By analogy with the previous examples it is easy to anticipate the general structure that the resulting instability criterion will take. If the interchange affected only the geopotential energy, as previously, the criterion would have the general form given by eq. (IV.43) with $\vec{g}$ replaced by $\vec{g}^*$ and V taken to be the volume of equi-flux flux tubes. (The apparent inconsistency between the local nature of the equation and the global nature of V is resolved by specifying the point of application to be the equatorial plane.) However, in this case the interchange can also result in a net change in the kinetic energy of compression. This was not true in the previous example because there each interchange involved equal and opposite changes in the volumes of the interchanging parcels as they passed in opposite directions through the identical pressure variation in the surrounding atmosphere. Thus, the work done on a descending parcel was identically cancelled by the work done by an ascending parcel, and there was no net change in energy of compression in the system of interchanging gas parcels. In the present situation, the volume is governed by the geometry of the dipole field. The pressure within each flux tube is arbitrary, in principle, as long as it satisfies the low β requirement. Thus, while the changes in the volumes of interchanging flux tubes are equal and opposite, the pressure attending the volume changes can be different for rising and sinking flux tubes, or, to switch to magnetospheric parlance, for outward

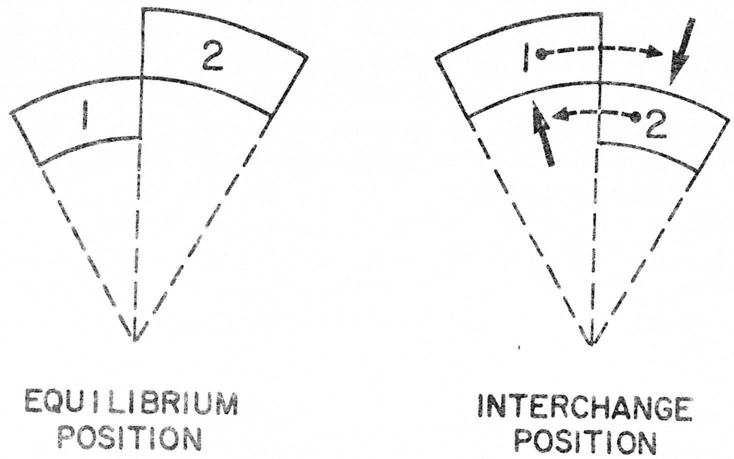

Figure IV.3 Sketch of interchange motion in the equatorial plane of a dipolar magnetosphere. The interchange involves two radially moving elements labeled 1 and 2. The radial displacements are indicated by solid arrows. The positions vacated by elements 1 and 2 are filled by azimuthal motions of the elements which formerly occupied the positions into which 1 and 2 moved, as indicated by the dashed arrows.

and inward moving flux tubes. Since inward moving flux tubes absorb energy of compression and outward moving flux tubes release energy of compression, it is evident that a magnetosphere that is radially stratified such that the compressed energy per unit magnetic flux decreases outward is unstable to interchange motions, as far as the criterion based on stored energy of compression is concerned (Gold, 1959). That is, in this situation more energy will be released by outward moving flux tubes than is absorbed by the ones moving inward to replace them.

A criterion for the interchange instability based on the energy principle has been given by Sonnerup and Laird (1963) in which both geopotential and compressional energies are included (see also Melrose 1967). The approach to be adopted here is based on the method used in discussions of the flute instability in plasma physics. It serves thereby to demonstrate the equivalence between the flute instability of plasma physics and the MHD interchange instability.

Figure IV.3 represents two situations in the equatorial plane of a magnetosphere. The equilibrium state is assumed to be static and to consist of radially stratified flux shells, that is, there are no motions initially and no variations in plasma parameters in the azimuthal direction. The radial profile of plasma parameters is assumed to be known and, as before, will be designated by a subscript st. The equilibrium is maintained by azimuthally flowing electrical currents that provide the ponderomotive force to balance the pressure gradient and geopotential forces. The requisite current is found by solving the momentum balance equation for $\vec{J}$,

$$\vec{J}_\perp = \frac{\vec{B}}{B^2} \times (\nabla p - \rho \vec{g}^*) \tag{IV.49}$$

in which the subscript $\perp$ on $\vec{J}$ merely makes explicit what is implicit in the right hand side of the equation, that the ponderomotive current flows perpendicular to the magnetic field, and has no parallel component. The pressure p in (IV.49) is taken to be a scalar to simplify the discussion.

After the interchange has occurred the radial profile of plasma parameters becomes locally adiabatic, and will be designated by a subscript ad. The current required to provide force balance as given by (IV.49) will now be discontinuous across the azimuthal interfaces between the interchanged and the ambient plasmas. In the discussion of the flute instability, the discontinuity in current builds up electric space charges on the two azimuthal walls of a radially displaced interchange element. The space charges will have opposite algebraic signs on the two sides of the element by the symmetry of the problem. An azimuthal electric fields is thereby generated. In the presence of the magnetic field the electric field will produce an $\vec{E} \times \vec{B}$ drift which is either radially out or in. If $\vec{E} \times \vec{B}$ is in the direction of the initial displacement that gave rise to the $\vec{E}$ field, the initial stratification is unstable, and all such perturbations will continue to grow. If $\vec{E} \times \vec{B}$ is in the direction opposite to the initial displacement, the plasma is stably stratified.

The only change to the above discussion that is called for because of the magnetospheric setting of the problem is the replacement of space charge build up by parallel currents. The space charge is discharged through the conducting ionosphere by flowing down the magnetic field lines. However, in this process also an electric field is generated across the flux tube as a result of the discharge current crossing the finite electrical resistance of the ionosphere. This electric field is in the same direction as the field that the space charge would have created. Thus, the instability criterion is the same in both cases.

In the absence of space charge build up, the divergence of the total current is zero. The equation for the parallel current is then

$$\frac{\partial J_\parallel}{\partial z} = - \nabla_\perp \cdot \vec{J}_\perp = - \nabla_\perp \cdot [\frac{\vec{B}}{B^2} \times (\nabla p - \rho \vec{g}^*)] \tag{IV.50}$$

in which the z direction is defined by $B = B\hat{z}$. The subscript $\perp$ on the del operator denotes a two dimensional divergence in the equatorial plane. If we adopt a cylindrical polar coordinate system (r,ϕ,z), the two dimensional divergence of the cross product of any two vectors $\vec{G}$ and $\vec{H}$ can be written explicitly as

$$\nabla_{\perp}\cdot(\vec{G}\times\vec{H}) = \vec{H}\cdot\nabla\times\vec{G} - \vec{G}\cdot\nabla\times\vec{H} + G_{\phi}\frac{\partial H_r}{\partial z} - G_r\frac{\partial H_{\phi}}{\partial z} - H_{\phi}\frac{\partial G_r}{\partial z} + H_r\frac{\partial G_{\phi}}{\partial z} \tag{IV.51}$$

By use of (IV.51), equation (IV.50) can be evaluated with the aid of the following conditions which are appropriate to the equatorial plane and our initial assumptions: $\nabla \times \vec{B} = 0$, and therefore $\partial B_r/\partial z = \partial B_z/\partial r$, $B = B_z$, $\partial B/\partial r = -3B/r$ (dipole field), $B_r = B_\phi = g^*_\phi = 0$, $g^* = g^*(r)\hat{r}$ and $\nabla \times \vec{g}^* = 0$ ($\vec{g}^*$ is given by the gradient of a potential). Eq. (IV.50) then reduces after several intermediate steps to

$$\frac{\partial J_{\parallel}}{\partial z} = \frac{1}{rB}\frac{\partial}{r\partial\phi}(3p + r\rho g^*) \tag{IV.52}$$

Thus as stated above, an azimuthal contrast in p or ρ brought about by a radial interchange motion gives rise to field-aligned currents. It should perhaps be noted that equation (IV.52) for $J_{\parallel}$ is general (for an isotropic plasma in the equatorial plane of a rotation-aligned dipole field) and does not depend on how the azimuthal variations in p and ρ are created.

The instability criterion can now be deduced from (IV.52). By our definition of the z-direction ($\vec{B} = B\hat{z}$) and the fact that (r,ϕ,z) forms a right handed coordinate system, we see that a positive E_ϕ produces an outward $\vec{E} \times \vec{B}$ drift. Thus, the magnetospheric stratification is unstable if an outward adiabatic interchange displacement produces a positive E_ϕ. The sign of E_ϕ across the interchange element will be the same as the sign of $\partial J_{\parallel}/\partial z$ on the clockwise (smaller ϕ) side of the element, since a positive $\partial J_{\parallel}/\partial z$ would correspond to a build up of positive charge in the absence of a field aligned current. From this we concluded that the stratification is unstable if $(3p + r\rho g^*)$ is greater on the adiabatic side of the azimuthal interface created by an outward displacement than it is on the ambient side. Mathematically the criterion can be expressed by

$$\frac{d}{dr}(3p + r\rho g^*)_{st} < \frac{d}{dr}(3p + r\rho g^*)_{ad} \quad \text{(unstable)} \tag{IV.53}$$

In the case of an isotropic plasma that uniformly fills the flux tubes, the adiabatic gradients can be given explicitly, since the volume of a dipolar flux tube is to a good approximation proportional to r^4. In an adiabatic displacement pV^γ = constant and ρV = constant (where V is the volume of the flux tube). The adiabatic gradient therefore can be written as

$$\frac{d}{dr}(3p + r\rho g^*)_{ad} = -4\gamma\frac{p}{r} - 4\frac{r\rho g^*}{r} + \rho\frac{drg^*}{dr} \quad \text{(IV.54)}$$

If the radial profiles of the structural pressure and density are expressed also as power-law variations

$$p_{st} \propto r^{\lambda}, \quad \rho_{st} \propto r^{\mu} \quad \text{(IV.55)}$$

then the instability criterion becomes

$$3(\lambda+4\gamma)p + (\mu+4)r\rho g^* < 0 \quad \text{(unstable)} \quad \text{(IV.56)}$$

(Since p_{ad} and ρ_{ad} are initiallized to the values of p_{st} and ρ_{st} at the same value of r, the subscripts are unnecessary in (IV.54 and 56). The term $\rho\, drg^*/dr$ is common to both sides of (IV.53), and therefore cancels out in IV.56). It remains only to give an explicit expression for g^*. In the equatorial plane g^*, the combination of the gravitational and centrifugal accelerations is given by

$$g^* = -g\frac{R_p^2}{r^2} + \Omega_p^2 r \quad \text{(IV.57)}$$

where g is the gravitational acceleration at the surface of the planet, R_p is the radius of the planet and Ω_p is the angular velocity associated with the rotation of the planet.

The criterion (IV.56) can be modified readily to make it applicable to situations in which the plasma does not fill the flux tube completely, such as in the case of anisotropic pressure or when the plasma is confined to the equatorial plane by the centrifugal force. The plasma formation in Jupiter's magnetosphere which is composed of matter originating on Io exhibits equatorial confinement. In this case the volume occupied by the plasma varies with distance more nearly as r^3. The factor four that appears in (IV.56) should then be replaced by the factor three. In the Jovian case there is also a background of energetic particles for which the factor four is appropriate. The terms in (IV.56) must be evaluated by combining both populations to determine whether or not the stratification of the Jovian magnetosphere is stable.

References:

Alfvén, H.: 1942, "On the existence of Electromagnetic-hydrodynamic waves", Nature, 150, 405.

Gold, T.: 1959, "Motions in the magnetosphere of the earth", J. Geophys. Res., 64, pp. 1219-1224.

Hudson, P. D.: 1970, "Discontinuities in an anisotropic plasma and their identification in the solar wind", Planet. Space Sci., 18, pp. 1611-1622.

Kantrowitz, A. F., and Petschek, H. E.: 1964, "MHD characteristics and shock waves", AVCO-Everett Research Report 185.

Krall, N. A., and Trivelpiece, A. W.: 1973, Principles of Plasma Physics, McGraw-Hill Book Company.

Melrose, D. B.: 1967, "Rotational effects on the distribution of thermal plasma in the magnetosphere of Jupiter", Planet. Space Sci., 15, pp. 381-393.

Rossi, B., and Olbert, S.: 1970, Introduction to Space Physics, McGraw-Hill Book Company, New York.

Sonnerup, B. U. Ö., and Laird, M. J.: 1963, "On the magnetospheric interchange instability", J. Geophys. Res., 68, pp. 131-139.

Sonnerup, B. U. Ö.: 1979, "Transport mechanisms at the magnetopause", in Dynamics of the Magnetosphere, edited by S.-I. Akasofu, D. Reidel Publishing Co., Dordrecht-Holland, pp. 77-100.

Spitzer, L., Jr.: 1956, Physics of Fully Ionized Cases, Interscience Publishers, Inc., New York.

Stern, D. P.: 1966, "The motion of magnetic field lines", Space Sci. Rev., 6, pp. 147-173.

Vasyliunas, V. M.: 1975, "Theoretical models of magnetic field line merging, 1", Rev. Geophys. Space Phys., 13, pp. 303-336.