nuary 1981

0036-1445/81/2301-0005 $01.00/0

THEMATICAL REMARKS ON THE VAN ALLEN RADIATION BELT: A SURVEY OF OLD AND NEW RESULTS*

MARTIN BRAUN†

Abstract. In this paper we present the new and important results obtained for the Störmer problem in ast decade, together with their implications for the existence of the Van Allen radiation belt. Specifically, rove the existence of quasi-periodic solutions for both small and large values of energy, and we give a plete topological description of all orbits near the dipole singularity. This leads, in turn, to a model for the plete flow. This model, aside from its purely theoretic interest, indicates that the Störmer Problem is ntegrable, and that the flow is ergodic in a suitable region of phase-space.

Introduction. In 1956, satellites launched by the National Aeronautics and Space ministration brought back conclusive evidence of charged particles oscillating back forth in the earth's atmosphere. These particles, which became known as the Van en radiation belt, were seemingly "trapped" by the magnetic field of the earth. ed, some of these particles were observed to have a lifetime of several years.

The motion of a charged particle under the influence of the earth's magnetic field ng been of interest to both mathematicians and physicists in connection with the of the polar aurora and cosmic rays. The mathematical formulation of this em was given by Störmer [27] as early as 1907, and it is often referred to as the mer problem. Subsequently, many mathematicians, such as De Vogelaere [9], [10],] and Goddart [16], devoted much of their work to this problem.

Particle motion in a magnetic field, and the earth's magnetic field in particular, ceived renewed significance with the discovery of the Van Allen radiation belt. The ain reason for this is that the Van Allen experiments exhibited the capability of a agnetic field to "trap" particles moving through it. This result intrigued the U. S. Air orce. Suppose, their experts argued, that the Russians detonated a nuclear device over reenland, thereby trapping many charged particles in the atmosphere. Could these arged particles interfere with our Ballistic Early Warning Radar System, thus leaving s vulnerable to a surprise attack?

Plasma physicists working in controlled thermonuclear fission were also keenly terested in the capabilities of magnetic fields to entrap particles for long periods of me. Essentially, the aim of plasma physicists is to create a very hot plasma by confining harged particles to a bounded region for a long period of time.

Section 1 of this paper describes an asymptotic theory, known as "adiabatic ivariance," for the entrapment of charged particles in a magnetic field. Adiabatic heory leads us, in a very natural way, to the Arnold, Kolmogorov, Moser (KAM) heory which rigorously establishes the confinement of particles for infinite time. The KAM theory is described in § 2, and then applied in § 3 to establish the existence of Van Allen particles for all time. Section 4 describes the behavior of all solutions in a neighborhood of the "singularity" of the earth's magnetic field, and this result is then used in § 5 to derive a model for the "global flow" of the Störmer problem. This model, which was verified numerically by Dragt [13], puts the finishing touches on a very interesting story, and sheds a great deal of light on the question of the global behavior of a dynamical system in the neighborhood of a homoclinic point.

* Received by the editors November 13, 1978, and in revised form April 11, 1980. This work was partially supported by the National Science Foundation under Grant MCS77-03999.

† Department of Mathematics, Queens College, Flushing, New York 11367.

This paper is completely self-contained. In addition to solving several important problems in geophysics, it is an excellent introduction to KAM theory, and how it is used in applied problems.

1. Adiabatic theory. A particle moving under the influence of a magnetic field **B** experiences a force

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}, \tag{1.1}$$

where q is its charge, and **v** its velocity. Since the force is perpendicular to the direction of motion, the energy

$$E = \tfrac{1}{2}m|\mathbf{v}|^2 \tag{1.2}$$

of the particle is constant. In general, the energy (1.2) is the only nontrivial integral of motion, and the trajectory of the particle is quite complicated. The only simple case is when the magnetic field is constant in magnitude and direction. Then, it is easily seen that the particle moves in a circular helix (see Fig. 1) about a field line. The radius

$$a = \frac{m|\mathbf{v}|}{q|\mathbf{B}|} \tag{1.3}$$

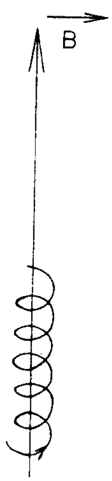

FIG. 1. *Particle motions in a constant field.*

of the helix is called the "radius of gyration" of the particle, and the particle oscillates about the field line with the cyclotron frequency

$$\omega = \frac{q|\mathbf{B}|}{m}.$$

Adiabatic theory begins with the assumption that the magnetic field varies slowly in space and time. Essentially, this means that in the course of one gyration about a magnetic field line, the particle sees an approximately constant magnetic field. In a slowly varying field, the particle moves approximately in a circle whose center drifts slowly across the lines of force, and moves rapidly along the lines. (This is the so-called "guiding center" or "adiabatic" approximation.) Under these circumstances, it was shown by Alfvén [1] that the magnetic moment

$$\mu = \frac{mv_\perp^2}{2|\mathbf{B}|}, \tag{1.4}$$

where $v_\perp$ is the component of the velocity perpendicular to the magnetic field, is an approximate constant of the motion. More precisely, the magnetic moment μ is constant to first order in the radius of gyration (1.3).

This result is of fundamental importance because it provides a mechanism for the confinement of charged particles in a bounded region of space. Suppose that the field strength $B = |\mathbf{B}|$ is a convex function along each line of force. Then, a particle moving along a line of force will be "reflected" backwards at the point P_0, where

$$\mu B(P_0) = E.$$

Thus, to first order in the radius of gyration, the guiding center of a particle oscillates periodically along a line of force, between two "mirror points." In this case, it has been shown by Northrop [24] that the quantity

$$J = \oint P_\parallel \, ds$$

is also an adiabatic invariant, where $P_\parallel$ is the guiding center momentum parallel to the lines of force, and the integral is taken over a complete oscillation from one mirror point to the other and back again.

Let us see what this all means in the case of the earth's magnetic field. We approximate the earth's magnetic field by that of a magnetic dipole situated at the center of the earth. Such a field can be described in cylindrical coordinates ρ, z, ϕ by the equations

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A}, \\ \mathbf{A} &= \frac{M\rho}{r^3}\hat{\phi}, \\ B = |\mathbf{B}| &= \frac{M}{r^3}(1 + 3\sin^2\lambda)^{1/2}, \end{aligned} \tag{1.5}$$

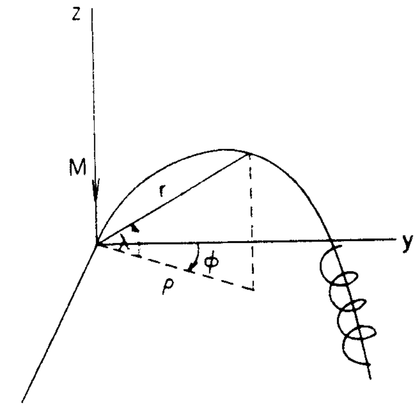

FIG. 2

(see Fig. 2), where M is the moment of the magnetic dipole, which points in the negative z direction, and $\hat{\phi}$ is a unit vector in the ϕ direction. The plane $\lambda = 0$ is the equatorial plane, and the magnetic lines of force are given by

$$r = r_0 \cos^2 \lambda, \qquad \phi = \text{constant}. \tag{1.6}$$

Consequently, along a line of force, the field strength

$$B = \frac{M(1+3\sin^2\lambda)^{1/2}}{r_0^3 \cos^6 \lambda}$$

is a convex function, and we have an asymptotic mechanism for the confinement of Van Allen particles.

However, adiabatic theory cannot account for the long time periods in which Van Allen particles are trapped. It must also be put on a more rigorous foundation if it is to have any application to plasma physics, where the goal is to trap particles in a bounded region for a period of time that encompasses many millions of gyrations.

To this end, Hellwig [18] proved in 1955 that the magnetic moment was constant to second order in the radius of gyration, and in 1957 Kruskal [19] showed that it was constant to all orders. Finally, in 1959, Gardner [15] showed that the longitudinal adiabatic invariant was constant to all orders, and he presented a general method of obtaining formal asymptotic expansions for all the adiabatic invariants.

One of the aims of this paper is to show that the phase space of a particle moving under the influence of the earth's magnetic field contains a region where series, analogous to the formal expansions of Gardner, are actually *convergent* expansions on a Cantor subset. This will be accomplished by using KAM theory. In this manner we will show that particles which are adiabatically trapped are, in fact, rigorously trapped for all time. This possibility was first pointed out by Arnold [2] in 1962. Unfortunately, the regions of phase space and/or the energies for which one can prove convergence are as yet far too small to have any real physical significance.

2. KAM theory. The application of KAM theory to the confinement of charged particles in a magnetic field rests on the following geometric theorem of Moser [21].

THEOREM. *Let M be an area preserving mapping of the plane into itself which, in polar coordinates r, θ, has the form*

$$M: \begin{cases} r_1 = r + \varepsilon^\sigma f(r, \theta, \varepsilon), \\ \theta_1 = \theta + \alpha + \varepsilon^\rho \gamma(r) + \varepsilon^\sigma g(r, \theta, \varepsilon). \end{cases} \tag{2.1}$$

We assume, specifically, that $\rho < \sigma$, f and g are sufficiently smooth, and $\gamma'(r) \neq 0$. Then, for sufficiently small ε, the mapping M possesses infinitely many closed invariant curves of nonzero measure, and these curves form a Cantor set.

Moser's theorem has the following geometric interpretation. Suppose that f and g are both zero. Then, each circle $r = \text{constant}$ is invariant, and the mapping M "twists" each point on $r = r_0$ by an amount $\alpha + \varepsilon^\rho \gamma(r_0)$. If $\gamma'(r) \neq 0$, then the twist varies from circle to circle. Finally, if the perturbation is small enough, i.e., the twist varies over a region that is large compared to the perturbation, then a Cantor set of closed invariant curves still persists under perturbation.

Moser's theorem has tremendous applications to the construction of invariant tori for Hamiltonian systems of differential equations. A Hamiltonian system of differential equations is a system of the form

$$\dot{x} = H_y, \qquad \dot{y} = -H_x, \tag{2.2}$$

where $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n)$ and H is a scalar function of x and y and possibly t. The function H is called the Hamiltonian and is usually synonomous with the energy of the system. The vectors x and y are called canonical coordinates, and we say the y is conjugate to x. If H does not depend explicitly on the time t, then it is conserved along any orbit. In this case we will call H the energy of our system.

Suppose now that $n=2$, and we can find canonical coordinates R_1, R_2, θ_1, and θ_2 such that our Hamiltonian H can be written in the form

$$H = H_0(R_1, R_2, \varepsilon) + \varepsilon^l H_1(R_1, R_2, \theta_1, \theta_2, \varepsilon), \tag{2.3}$$

where

$$\dot{\theta}_i = H_{R_i}, \qquad \dot{R}_i = -H_{\theta_i}, \tag{2.4}$$

and H_1 is periodic in θ_1, θ_2, with period 2π. If H_1 is identically zero, then both R_1 and R_2 are constant. In this case, every orbit of (2.4) lies on an invariant torus $R_1 = R_1^0$, $R_2 = R_2^0$, and on this torus

$$\theta_i(t) = \omega_i t + \theta_i(0),$$

where $\omega_i = H_{R_i}(R_1^0, R_2^0, \varepsilon)$. The numbers ω_1, ω_2 are called the frequencies of the motion. Such a system is said to be *integrable*, since R_1 and R_2 are two integrals.

Suppose now that $H_1 \neq 0$. In this case, the system (2.4) is almost definitely (see Poincaré [25]) not integrable. However, a Cantor set of invariant tori will still persist under perturbation if ε is small enough. To see this, we first solve for

$$R_1 = \Phi(R_2, \theta_1, \theta_2, \varepsilon, E) \tag{2.5}$$

on the energy surface $H = E$. Using θ_1 instead of t as the independent variable, and setting $R_2 = R$, $\theta_2 = \theta$, we see that

$$\frac{dR}{d\theta_1} = \frac{-H_\theta}{H_{R_1}}, \qquad \frac{d\theta}{d\theta_1} = \frac{H_R}{H_{R_1}}.$$

It is easily verified that on the surface $H = E$ these equations take the form

$$\frac{dR}{d\theta_1} = \Phi_\theta, \qquad \frac{d\theta}{d\theta_1} = -\Phi_R, \tag{2.6}$$

where Φ is defined by (2.5).

The system (2.6) is again Hamiltonian, with one degree of freedom ($n=1$), but nonautonomous. To eliminate the independent variable we follow the solutions from $\theta_1 = 0$ to the next intersection with $\theta_1 = 2\pi$. This defines a mapping M, which by Liouville's Theorem is area preserving. This mapping M will have the form

$$R(2\pi) = R(0) + O(\varepsilon^l),$$
$$\theta(2\pi) = \theta(0) + \alpha + \varepsilon^k \gamma(R) + O(\varepsilon^l).$$

Thus, if $k < l$, and $\gamma'(R) \neq 0$, then Moser's theorem guarantees the existence of infinitely many invariant curves of M, for ε sufficiently small. Each of these invariant curves generates an invariant torus for the system (2.4) if we take all solutions which issue forth from the invariant curve. More important though, consider two such tori generated from two concentric invariant curves. Any point between these two curves must remain between these two curves under all iterates of M. This implies that any orbit starting between the two tori generated by these curves must remain trapped between these two tori for all time. This is our desired stability result.

The above analysis hinges on our ability to approximate a given Hamiltonian by an integrable Hamiltonian, to a desired order. This is not as difficult as it may seem, since to transform the equations of motion we need only transform the Hamiltonian H. This is a consequence of the fact (Siegel [26]) that if x and y are canonical coordinates and if ξ and η are defined via the equations

(2.7) $$\xi = S_\eta(x, \eta), \qquad y = S_x(x, \eta),$$

for some *scalar* function $S(x, \eta)$, then ξ and η are again canonical coordinates. The system of equations (2.2) is then transformed into the new Hamiltonian system

$$\dot{\xi} = K_\eta, \qquad \dot{\eta} = -K_\xi,$$

where

(2.8) $$K(\xi, \eta) = H(x(\xi, \eta), y(\xi, \eta)).$$

Let us assume, now, that we can write our Hamiltonian $H(R_1, R_2, \theta_1, \theta_2, \varepsilon)$ in the form

(2.9) $$H = \omega_1 R_1 + \omega_2 R_2 + \varepsilon H_1(R_1, R_2, \theta_1, \theta_2) + \varepsilon^2 H_2(R_1, R_2, \theta_1, \theta_2) + \cdots,$$

where H has period 2π in θ_1 and θ_2. There is a very nice method known as Lindstedt's method [25] for approximating H to higher and higher order in ε by an integrable Hamiltonian. Lindstedt's method is quite easy to implement. Let us first try to find new canonical variables R_i', θ_i' so that H is independent of θ_1' and θ_2' through order one. To this end, we take a generating function $S = S(R', \theta, \varepsilon)$ of the form

$$S = R_1'\theta_1 + R_2'\theta_2 + \varepsilon S_1(R_1', R_2', \theta_1, \theta_2),$$

and we write the new Hamiltonian

$$K(R', \theta', \varepsilon) = H(R(R', \theta', \varepsilon), \theta(R', \theta', \varepsilon), \varepsilon)$$

in the form

$$K = \omega_1 R_1' + \omega_2 R_2' + \varepsilon K_1(R_1', R_2') + \cdots.$$

From (2.7) we see that

$$R_i = S_{\theta_i} = R_i' + \varepsilon \frac{\partial S_1}{\partial \theta_i}, \qquad \theta_i' = S_{R_i'} = \theta_i + \varepsilon \frac{\partial S_1}{\partial R_i'}.$$

Equating terms of order ε in (2.8) (and using θ and R' as variables) we see that

(2.10) $$\omega_1 \frac{\partial S_1}{\partial \theta_1} + \omega_2 \frac{\partial S_1}{\partial \theta_2} = K_1(R_1', R_2') - H_1(R_1', R_2', \theta_1, \theta_2).$$

In order for S_1 to be periodic in θ_1 and θ_2 we must require that the right-hand side of (2.10) have mean value zero. This then determines

$$K_1(R_1', R_2') = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} H_1(R_1', R_2', \theta_1, \theta_2)\, d\theta_1\, d\theta_2.$$

To solve for S_1 now, we expand the right-hand side of (2.10) in the Fourier series

$$\sum_{k \neq 0} a_k(R')\, e^{i(k, \theta)},$$

where $(k, \theta) = k_1\theta_1 + k_2\theta_2$. Then,

(2.11) $$S_1 = \sum_{k \neq 0} \frac{a_k(R')}{i(k, \omega)} e^{i(k, \theta)}.$$

However, this series will not converge, in general, because of the presence of the small divisors

$$(k, \omega) = k_1\omega_1 + k_2\omega_2.$$

To overcome this difficulty, we assume that ω_1 and ω_2 satisfy the infinitely many inequalities

$$|(j, \omega)| \geqq \gamma |j|^{-\tau}, \tag{2.12}$$

for all integers j_1, j_2 and with some constants γ and $\tau > 1$. Then (see Moser [21]) the series (2.11) will converge.

Repeating this process inductively, we see that we can approximate H to order n by n successive applications of Lindstedt's method.

If $H_0 = H_0(R_1, R_2)$, then we can achieve $H_0 = \omega_1 R_1 + \omega_2 R_2$ by choosing R_1^0 and R_2^0 so that

$$\omega_i = \frac{\partial H_0(R_1^0, R_2^0)}{\partial R_i}$$

are rationally independent numbers satisfying (2.12), and then restricting R_i to the region

$$|R_i - R_i^0| = O(\varepsilon^{n+1}).$$

Finally, observe that if $\omega_2 = 0$, then Lindstedt's method is very easy to apply to eliminate the dependence of H on θ_1.

3. The existence of invariant tori. To analyze the equations of motion for the Störmer problem, it is convenient to write the equations in Hamiltonian form with Hamiltonian[1]

$$H = \frac{1}{2m}\left[p_\rho^2 + p_z^2 + \left(\frac{p_\phi}{\rho} - q|\mathbf{A}|\right)^2\right],$$

(see Goldstein [17]), where

$$\begin{aligned} p_\rho &= m\dot\rho, & \dot p_\rho &= -H_\rho, \\ p_z &= m\dot z, \qquad \text{and} & \dot p_z &= -H_z, \\ p_\phi &= m\rho^2\dot\phi + q\rho|\mathbf{A}|, & \dot p_\phi &= -H_\phi. \end{aligned} \tag{3.1}$$

Since H is independent of time, the energy

$$H = \tfrac{1}{2}m|\mathbf{v}|^2 = E$$

is a constant of the motion. A second integral of the motion is obtained by noting that H is independent of ϕ. Hence, the electromagnetic angular momentum

$$p_\phi = qM\Gamma,$$

where Γ is defined by this equation, is a constant of the motion. (The integration constant Γ has the dimension of reciprocal length.) Our three-dimensional problem is now reduced to the simpler problem of finding the two-dimensional motion of a particle in the ρ, z-plane under the influence of the potential

$$V(\rho, z) = \frac{q^2M^2}{2m}\left(\frac{\Gamma}{\rho} - \frac{\rho}{r^3}\right)^2. \tag{3.2}$$

[1] The relativistic case follows by setting $m = \gamma m_0$.

Once $\rho(t)$ and $z(t)$ have been found, $\phi(t)$ is then determined from the equation

$$\dot{\phi} = -H_{p_\phi},$$

which yields

$$\phi(t) = \phi(0) + \int_0^t \left(\frac{qM\Gamma}{\rho^2} - \frac{qM}{r^3}\right) d\tau.$$

The sign of Γ plays a crucial role in determining the general properties of the trajectories of (3.1). Observe that the radial derivative of V is given by

$$\mathbf{r} \cdot \nabla V = \frac{-q^2M^2}{m}\left(\frac{\Gamma}{\rho} - \frac{\rho}{r^3}\right)\left(\frac{\Gamma}{\rho} - \frac{2\rho}{r^3}\right).$$

This quantity is strictly less than zero for Γ negative. A negative radial derivative for the potential V corresponds to a repulsive radial force, since $-\mathbf{r} \cdot \nabla V$ is the component of the force in the radial direction. Hence, all trajectories characterized by a negative value of Γ must extend to infinity, and cannot be trapped. Note also, that all trajectories must be confined to the region $V(\rho, z) \leqq E$. This region is indicated in Fig. 3. Observe too, from this figure, that all trajectories remain bounded away from the singularity $r = 0$.

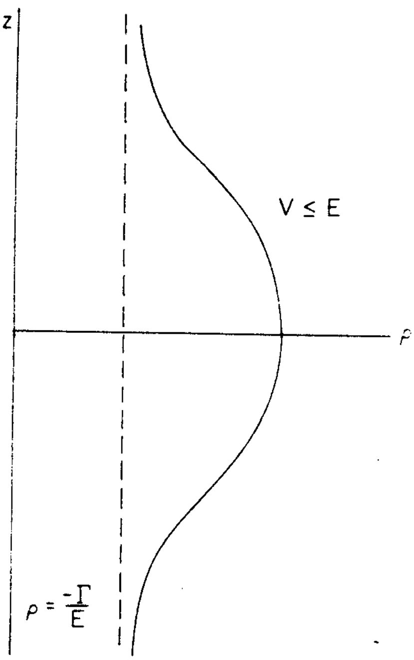

FIG. 3. *The region* $V \leqq E$ *for* $\Gamma < 0$.

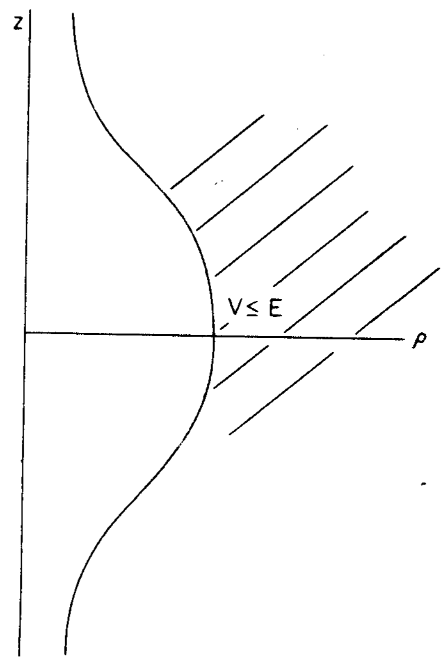

FIG. 4. *The region* $V \leqq E$ *for* $\Gamma = 0$.

The situation for $\Gamma = 0$ is very similar. The radial derivative of V is again negative, but this time only for $\rho \neq 0$. In this case, $\rho = 0$ is an orbit, and along this orbit,

(3.3a) $$z(t) = \sqrt{2mE}t + z_0, \qquad z_0 < 0;$$

(3.3b) $$z(t) = -\sqrt{2mE}t + z_0, \qquad z_0 > 0.$$

Note that (3.3a) corresponds to a trajectory which runs into the singularity in finite time from below the equator, while (3.3b) corresponds to a trajectory running into the singularity, again in finite time, from above the equator. Finally, observe that all orbits starting in the shaded region of Fig. 4 must extend to infinity.

For the study of bounded trajectories, therefore, we restrict ourselves to the case $\Gamma > 0$. It is convenient at this point to introduce the dimensionless variables

$$z' = \Gamma z, \qquad \phi' = \phi,$$
$$\rho' = \Gamma \rho, \qquad t' = \frac{\Gamma^3 qM}{m} t.$$

Then, the equations of motion for these dimensionless variables are derived from the new Hamiltonian

$$H = \frac{1}{2}(p_\rho^2 + p_z^2) + \frac{1}{2}\left(\frac{1}{\rho} - \frac{\rho}{r^3}\right)^2,$$

where we have omitted the primes for convenience. In this system of units, the particle has the dimensionless velocity $w_0 = \frac{1}{4}\gamma_1^2$, where

$$\gamma_1^4 = \frac{1}{16}\left(\frac{qM}{m|\mathbf{v}|}\right)^2 \Gamma^4.$$

The dimensionless constant γ_1 is that used by Störmer [27]. Again, the particle is restricted to lie in the region $V \leqq E$. This region assumes three different forms, depending on whether E is less than, equal to, or greater than $\frac{1}{32}$.

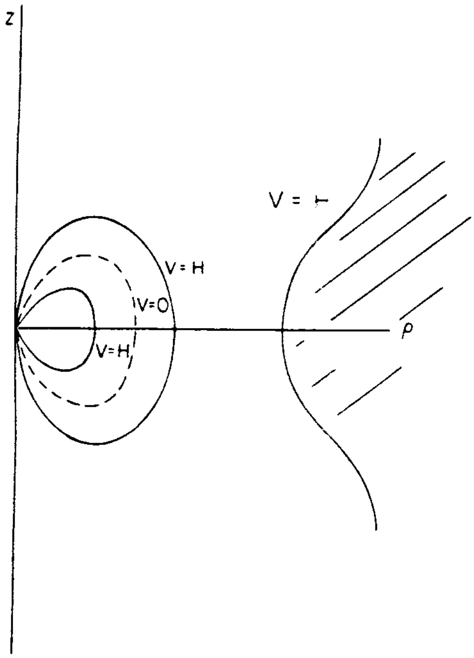

FIG. 5. *The region* $V(\rho, z) \leqq E$, *for* $E < \frac{1}{32}$.

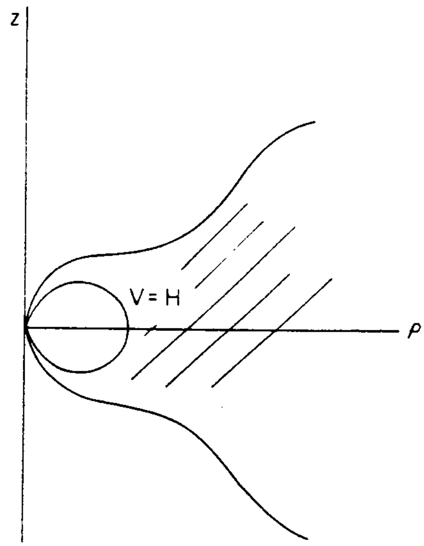

FIG. 6. *The region* $V(\rho, z) \leqq E$, *for* $E > \frac{1}{32}$.

From Fig. 5, we see that any trajectory starting in the oval-like region surrounding the curve $V = 0$ ($r = \cos^2 \lambda$), with initial energy less than $\frac{1}{32}$, can *never* leave this region. (The curve $r = \cos^2 \lambda$ corresponds in our old coordinates to the line of force $r =$

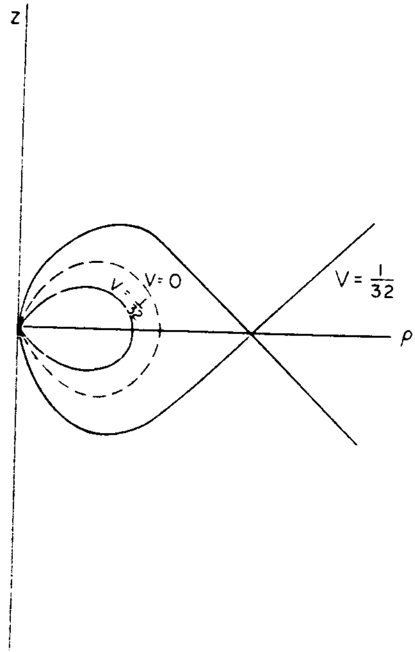

FIG. 7. *The region* $V(\rho, z) \leqq \frac{1}{32}$.

$r^{-1}\cos^2\lambda$.) Thus, it would appear that our problem is solved. However, particles which penetrate to high latitudes are lost because their energy is decreased due to collisions with various atoms in the upper atmosphere. Thus, our problem is not only containment, but containment between fixed latitudes. And to prove this, we must roll out some heavy artillery.

Remark. It would appear that values of $E > \frac{1}{32}$ are uninteresting since the region $V \leqq E$ extends to infinity. However, these trajectories are important for just this reason. To wit, those particles which influence the polar aurora[2] (Störmer [27]) are the particles which can penetrate from infinity to the singularity. This is possible only when $E > \frac{1}{32}$.

(A) Invariant tori near the equatorial plane. Adiabatic theory leads us to the intuitive notion that a particle motion (for small velocities, at least) for the Störmer problem is essentially a gyration about a line of force and an oscillation along the line of force. Therefore, it is natural to introduce new coordinates $a(\rho, z)$ and $b(\rho, z)$: The curve $a(\rho, z) = \text{constant}$ should define a magnetic line of force, and the $b(\rho, z)$ should describe where we are on this line of force. The coordinate $a(\rho, z)$ is simply $r/\cos^2\lambda$, since the lines of force are given by the equation

$$a = \frac{r}{\cos^2\lambda}, \qquad a = \text{constant}. \tag{3.4}$$

An appropriate b is obtained by requiring that $a(\rho, z)$ and $b(\rho, z)$ be an orthogonal system of coordinates. This then leads to

$$b(\rho, z) = \frac{\sin\lambda}{r^2}. \tag{3.5}$$

[2] Störmer's theory of the polar aurora is no longer believed to be correct.

New canonical variables a, b, p_a, p_b can now be obtained from the generating function

$$F(\rho, z, p_a, p_b) = a(\rho, z)p_a + b(\rho, z)p_b.$$

Then,

$$a = \frac{\partial F}{\partial p_a} = a(\rho, z), \qquad p_\rho = \frac{\partial F}{\partial \rho} = \frac{\partial a}{\partial \rho} p_a + \frac{\partial b}{\partial \rho} p_b,$$

$$b = \frac{\partial F}{\partial p_b} = b(\rho, z), \qquad p_z = \frac{\partial F}{\partial z} = \frac{\partial a}{\partial z} p_a + \frac{\partial b}{\partial z} p_b.$$

In terms of these new variables the Hamiltonian takes the form

(3.6) $$H(a, b, p_a, p_b) = \frac{1}{2}\left(\frac{p_a^2}{h_a^2} + \frac{p_b^2}{h_b^2}\right) + \frac{1}{2}\frac{(a-1)^2}{a^4 \cos^6 \lambda},$$

where

(3.7) $$h_a^2 = \frac{\cos^6 \lambda}{1 + 3\sin^2 \lambda} \quad \text{and} \quad h_b^2 = \frac{a^6 \cos^{12} \lambda}{1 + 3 \sin^2 \lambda}.$$

It is understood of course, that $\sin \lambda$ and $\cos \lambda$ are to be expressed in terms of a and b.

We now wish to restrict ourselves to a region in phase space where the energy H will be small, so that adiabatic theory will be valid. To this end, we make the substitution

(3.8) $$a = 1 + \varepsilon^2 \alpha, \qquad b = \varepsilon \beta,$$
$$p_a = \varepsilon^2 p_\alpha, \qquad p_b = \varepsilon^3 p_\beta,$$

where ε is a small parameter. The first equation of (3.8) restricts the particle to a small neighborhood of the guiding field line $r = \cos^2 \lambda$, while the last equation guarantees that the motion along this field is much slower than the motion across it. Strictly speaking, the transformation (3.8) is not canonical. However, it is easily verified that our system is still Hamiltonian if we choose (see Siegel [26])

(3.9) $$H(\alpha, \beta, p_\alpha, p_\beta) = \frac{1}{\varepsilon^4} H.$$

Our next step is to approximate H by an integrable Hamiltonian. To this end we first show that $\lambda = \lambda(a, b)$ is an analytic function of a and b for b small. Squaring (3.4) and multiplying by (3.5) yields

(3.10) $$a^2 b = \frac{\sin \lambda}{(1 - \sin^2 \lambda)^2}.$$

The derivative of the right-hand side of (3.10) with respect to λ is one at $\lambda = 0$. Hence, we are guaranteed that $\lambda(a, b)$ is an analytic function of $z = a^2 b$, at least for $|z|$ small. Squaring (3.10) we easily find (Braun [4]) that

$$\sin^2 \lambda = \beta^2 \varepsilon^2 + 4(\alpha\beta^2 - \beta^4)\varepsilon^4 + G_1(\alpha, \beta, \varepsilon),$$
$$\cos^{-6} \lambda = 1 + 3\beta^2 \varepsilon^2 + 6(2\alpha\beta^2 - \beta^4)\varepsilon^4 + G_2(\alpha, \beta, \alpha),$$

where

$$G_i = \varepsilon^6 \bar{G}_i(\alpha, \beta, \varepsilon),$$

with $\bar{G}_i$ analytic in the variables α, β, ε.

The Hamiltonian (3.9) can now be written in the form

$$H=H_0+H_2+H_4+H_6,$$

where

$$H_0=\frac{\alpha^2+p_\alpha^2}{2},$$

$$H_2=\frac{\varepsilon^2}{2}(6\beta^2p_\alpha^2-4\alpha^3+p_\beta^2+3\beta^2\alpha^2),$$

$$H_4=\frac{3\varepsilon^4}{2}(24\alpha\beta^2p_\alpha^2-3\beta^4p_\alpha^2-2\alpha p_\beta^2+3\beta^2p_\beta^2+2\alpha^4-2\alpha^2\beta^4),$$

$$H_6=\varepsilon^6\bar{H}_6(\alpha,\beta,p_\alpha,p_\beta,\varepsilon),$$

with $\bar{H}_6$ analytic in all its variables.

Notice that H_0 depends on $\alpha^2+p_\alpha^2$ alone. If we make the canonical substitution

$$\alpha=\sqrt{2R}\sin\theta,\qquad p_\alpha=\sqrt{2R}\cos\theta,$$

then $H_0=R$ is independent of θ. Hence, we can use Lindstedt's method to average out the dependence of H on θ to any order in ε. Following Braun [4], and omitting the primes on α, β, p_α and p_b, we find that

$$H=H_0+H_2+H_4+H_6,$$

where

$$H_0=R,$$

$$H_2=\frac{\varepsilon^2}{2}(9R\beta^2+p_\beta^2),$$

$$H_4=\frac{\varepsilon^4}{2}\left(9\beta^2p_\beta^2-21R^2-\frac{69R\beta^4}{4}\right),$$

$$H_6=\varepsilon^6\bar{H}_6.$$

Next observe that for fixed R, the curves

$$9R\beta^2+p_\beta^2=\text{constant}$$

are ellipses in the β, p_β-plane. These curves may be transformed into circles (with the same area) by the generating function

$$F(\theta,\beta,R_1,p'_\beta)=(9R_1)^{1/4}\beta p'_\beta+\theta R_1.$$

Setting

$$\beta'=\sqrt{2R_2}\sin\theta_2,\qquad p'_\beta=\sqrt{2R_2}\cos\theta_2,$$

we see that H_2 is independent of θ_2. Hence, we may employ Lindstedt's method to average out the θ_2 dependence of H to order ε^6. In terms of new canonical variables, which we again call R_1, θ_1, R_2, θ_2 the Hamiltonian H assumes the form

$$H=R_1+3\varepsilon^2R_2\sqrt{R_1}+\frac{\varepsilon^4}{2}\left(-21R_1^2+\frac{13R_2^2}{8}\right)+H_6,$$

where

$$H_6 = \varepsilon^6 \bar{H}_6(R_1, R_2, \theta_1, \theta_2, \varepsilon),$$

with $\bar{H}_6$ analytic in all its variables.

Finally, we apply Moser's theorem to establish the existence of infinitely many invariant tori. As described in the previous section we solve for $R_1 = \Phi(R_2, \theta_1, \theta_2, \varepsilon)$ on the energy surface $H = E$, and take θ_1 instead of t as independent variable. We then follow all solutions from $\theta_1 = 0$ to their next intersection with $\theta_1 = 2\pi$. This defines an area-preserving mapping M. Since

$$\frac{d\theta_1}{d\theta_2} = \frac{H_{R_2}}{H_{R_1}} = 3\varepsilon^2\sqrt{R_1} - \tfrac{23}{8}\varepsilon^4 R_2 + O(\varepsilon^6)$$

and

$$R_1^{1/2} = E^{1/2} - \tfrac{3}{2}\varepsilon^2 R_2 + O(\varepsilon^4),$$

the mapping M, expressed in coordinates $R = R_2$, $\theta = \theta_2$, has the form

$$M: \quad \begin{cases} R(2\pi) = R + \varepsilon^6 f(R, \theta, \varepsilon), \\ \theta(2\pi) = \theta + 3\varepsilon^2 E^{1/2} - \tfrac{59}{8}\varepsilon^4 R + \varepsilon^6 g(R, \theta, \varepsilon), \end{cases} \tag{3.12}$$

with $R = R(0)$, $\theta = \theta(0)$ and f, g analytic in all their variables. This mapping is of the *form* (2.1) with $\rho = 4$, $\sigma = 6$, and $\gamma'(R) = -\frac{59}{8} \neq 0$. Hence, Moser's theorem guarantees the existence of infinitely many invariant curves of (3.12), and this in turn guarantees the existence of infinitely many invariant tori, of nonzero measure, on each energy surface $H = E$.

Remark. The orbits we have just generated are quasiperiodic (see Siegel [26]) with two frequencies, and all lie near the equatorial plane. They gyrate very tightly about the guiding field line $r = \cos^2 \lambda$. Moreover, these orbits all cross the equatorial plane at an angle near 90°, since the velocity of the particle parallel to the magnetic field is much smaller than the total velocity. This follows from the fact that $p_b = \varepsilon^3 p_\beta$ is third order small, while $p_a = \varepsilon^2 p_\alpha$ is second order small.

(B) Invariant tori which are not close to the equatorial plane. Our goal in this section is to prove the existence of invariant tori which need not lie near the equatorial plane. To this end, we consider, instead of (3.8), the change of coordinates

$$a = 1 - \varepsilon\alpha, \qquad b = \varepsilon\beta,$$

$$p_a = \varepsilon p_\alpha, \qquad p_b = \varepsilon p_\beta,$$

where ε is a small parameter, and β is of order $1/\varepsilon$. Specifically, let M and N be two fixed constants, with N very large. We consider all variables to be complex, and restrict ourselves to the region T defined by

$$|\alpha| + |p_\alpha| + |p_\beta| \leq M,$$

$$|\mathrm{Re}\, b| \leq N, \qquad |\mathrm{Im}\, b| \leq \delta(M, N, \varepsilon),$$

where δ depends on M, N, and ε, and will be chosen so that the Hamiltonian H is analytic in the region T, for ε sufficiently small. Analyticity of H is required for two reasons. First, we need to expand H in a power series, and thus derivatives are certainly required. More important, though, is the fact that the functions f and g in (2.1) must be *uniformly* small, together with their derivatives. This is most easily proven if f and g are analytic.

To find the singularities of H in the complex four-space, we consider again the equation

$$a^2 b = \frac{\sin\lambda}{(1-\sin^2\lambda)^2}.$$

With $z = a^2 b$ and $y = \sin\lambda(a, b)$, we see that

$$\frac{dz}{dy} = \frac{1+3y^2}{(1-y^2)^3}.$$

Hence, the only singular points are $y = \pm 1$, and $y = \pm i/\sqrt{3}$. It is clear that for fixed M and N, δ can be chosen so that $\sin\lambda \neq \pm 1$ for $|\mathrm{Im}\, b| \leqq \delta$ and ε sufficiently small. The points $y = \pm i/\sqrt{3}$ correspond to the points

$$z = a^2 b = \pm i\frac{3\sqrt{3}}{16}.$$

Setting $a^2 = a_1 + ia_2$, $b = b_1 + ib_2$, we see that

$$a^2 b = a_1 b_1 - a_2 b_2 + i(a_1 b_2 + a_2 b_1).$$

For fixed M and N, we restrict δ and ε still further so that

$$|a_1 b_2 + a_2 b_1| \leqq \tfrac{3}{16}.$$

In this manner we exclude the points $z = \pm i3\sqrt{3}/16$, and the new Hamiltonian

$$\begin{aligned} H(\alpha, \beta, p_\alpha, p_\beta) &= \frac{1}{\varepsilon^2} H(a, b, p_a, p_b) \\ &= \frac{1}{2}\left(\frac{p_\alpha^2}{h_a^2} + \frac{p_\beta^2}{h_b^2}\right) + \frac{\alpha^2}{2(1+\varepsilon\alpha)^4 \cos^6\lambda} \end{aligned} \tag{3.13}$$

is now analytic in all its variables in the region T. Note that N is fixed; however, it may be chosen initially as large as desired. Large values of N, i.e., large values of b, correspond to points far along a field line. In other words, orbits which achieve large values of b are orbits which contain points far from the equatorial plane.

Our next task is to approximate the Hamiltonian H by an integrable one. We cannot expand H in powers of $\varepsilon\alpha$ and $\varepsilon\beta$ as we did in the previous section, because now, $\beta \sim 1/\varepsilon$. Instead, we expand H in powers of $a - 1 = \varepsilon\alpha$ alone. Since $\sin\lambda$ and $\cos^{-1}\lambda$ are analytic in T, we may write

$$\begin{aligned} 1 + 3\sin^2\lambda &= K_1(b) + F_1(a, b), \\ \cos^{-6}\lambda &= C_1(b) + F_2(a, b), \end{aligned}$$

where $F_i(a, b) = O((a-1))$. Then, the Hamiltonian (3.13) may be written in the form $H = H_0 + H_1$, where

$$H_0 = \frac{C_1}{2}(K_1 p_\alpha^2 + \alpha^2) + \frac{K_1 C_1^2}{2} p_\beta^2$$

and

$$H_1 = \varepsilon \bar{H}_1(\alpha, \varepsilon\beta, p_\alpha, p_\beta, \varepsilon),$$

with $\bar{H}_1$ analytic in all its variables.

We now work on the Hamiltonian H_0. Our first step is to find new canonical variables α', p'_α so that H_0 is a function of $(\alpha')^2+(p'_\alpha)^2$ alone. This is accomplished by taking

$$F(\alpha, \beta, p'_\alpha, p'_\beta) = [K_1(\varepsilon\beta)]^{-1/4}\alpha p'_\alpha + \frac{f(\varepsilon\beta)p'_\beta}{\varepsilon}.$$

It would suffice to take $f=\varepsilon\beta$. However, as we shall see shortly, a judicious choice of f will enable us to express H_0 explicitly, and thereby greatly simplify the ensuing calculations. The new canonical variables α', p'_α, β', p'_β are determined from the relations

$$p_\alpha = \frac{\partial F}{\partial \alpha} = K_1^{-1/4} p'_\alpha,$$

$$p_\beta = \frac{\partial F}{\partial \beta} = f^{(1)} p'_\beta - \frac{\varepsilon}{4} K_1^{-5/4} K_1^{(1)} \alpha p'_\alpha,$$

$$\alpha' = \frac{\partial F}{\partial p'_\alpha} = K_1^{-1/4}\alpha, \qquad \beta' = \frac{\partial F}{\partial p'_\beta} = \frac{f}{\varepsilon},$$

where $K^{(1)}$ and $f^{(1)}$ denote the derivatives of K and f with respect to $b=\varepsilon\beta$.

A point in the ρ, z-plane is determined from the coordinates α', β' in the following manner. Given α' and β', the point lies on the magnetic field line $r = a\cos^2\lambda$, where

$$a = 1 + \varepsilon\alpha'[1+3(\varepsilon\beta')^2]^{1/4}.$$

The coordinate b on this field line is then found from the relation

$$f(b) = f(\varepsilon\beta) = \varepsilon\beta'. \tag{3.14}$$

We cannot solve (3.14) explicitly for b. However, it shall be seen for the f we shall determine, that as $\varepsilon\beta'$ increases (decreases) to 1 (-1), b increases (decreases) to $+\infty$ $(-\infty)$, and that the latitude λ, for $a=1$ is precisely

$$\sin^{-1}(\varepsilon\beta').$$

Thus, $\beta'(t)$ describes, essentially, the motion along a line of force, for small ε.

In terms of the primed variables, the Hamiltonian (3.13) assumes the form

$$H = \frac{C_1\sqrt{K_1}}{2}[(\alpha')^2 + (p'_\alpha)^2 + C_1\sqrt{K_1}(f^{(1)})^2(p'_\alpha)^2] + H_1,$$

with $H_1 = O(\varepsilon)$. We would like to show that H_0 is actually *integrable*. To this end, we introduce canonical coordinates R, θ in place of α', p'_α by the relations

$$\alpha' = \sqrt{2R}\sin\theta, \qquad p'_\alpha = \sqrt{2R}\cos\theta.$$

Now, H_0 is independent of θ, so that to lowest order, R (which is proportional to the magnetic moment (Braun [4])) is a constant of the motion. Also, to lowest order,

$$\dot\theta = \omega_1 = C_1\sqrt{K_1}.$$

It is shown in [4] that $C_1\sqrt{K_1}$ is the magnitude of the magnetic field along the line of force $r=\cos^2\lambda$. Thus, we have already confirmed our intuitive notion that to lowest order the particle gyrates about its field line with the so-called cyclotron frequency ω_1.

It is convenient at this point to introduce a new time scale τ so that the frequency ω_1 becomes unity. This is accomplished by setting

$$\tau = \int \omega_1 \, dt.$$

As observed first by Poincaré [25], our system is still Hamiltonian if we consider only the zero energy solutions of the new Hamiltonian

$$F = \frac{1}{\omega_1}(H - h) = F_0 + F_1,$$

where h is the constant value of the energy H along our trajectory, and

$$F_0 = R + \frac{1}{2\omega_1}[\omega_1^2 (f^{(1)})^2 (p'_\beta)^2 - 2h],$$

$$F_1 = \frac{H_1}{\omega_1}.$$

Notice that to lowest order we have decoupled the motion in α', p'_α, and β', p'_β. This puts into evidence the integrability of the truncated system.

To lowest order, the motion in the β', p'_β plane is described by the Hamiltonian

$$\Phi = \tfrac{1}{2}\omega_1 (f^{(1)})^2 (p'_\beta)^2 - \frac{h}{\omega_1}.$$

Let us now define the function $f(b)$ by the relation

$$b = \frac{f}{(1 - f^2)^2}.$$

We choose this function f because now the Hamiltonian Φ can be calculated explicitly. To wit, write

$$\sin \lambda = f_1(b) + O(a - 1).$$

It follows from

$$a^2 b = \frac{\sin \lambda}{(1 - \sin^2 \lambda)^2}$$

that $f_1 = f$. Next, compute

$$\omega_1 = \frac{\sqrt{1 + 3\gamma^2}}{(1 - \gamma^2)^3}, \qquad \gamma = \varepsilon \beta'$$

and

$$f^{(1)} = \frac{(1 - f^2)^3}{1 + 3f^2} = \frac{(1 - \gamma^2)^3}{1 + 3\gamma^2} = \frac{1}{\omega_1 \sqrt{1 + 3\gamma^2}}.$$

Hence,

$$\Phi = \frac{1}{2\omega_1}\left[\frac{(p'_\beta)^2}{1 + 3\gamma^2} - 2h\right] = \frac{(1 - \gamma^2)^3}{2\sqrt{1 + 3\gamma^2}}\left[\frac{(p'_\beta)^2}{1 + 3\gamma^2} - 2h\right]. \tag{3.15}$$

To analyze the orbits determined by Φ we must consider the level curves

$$c=\frac{-h(1-\gamma^2)^3}{(1+3\gamma^2)^{1/2}}+\frac{1}{2}\frac{(1-\gamma^2)^3}{(1+3\gamma^2)^{3/2}}(p'_\beta)^2, \tag{3.16}$$

in the β', p'_β-plane. For $c>0$, these curves are *not* closed, and asymptotically approach the line $\gamma=\pm 1$. The curves corresponding to $c=0$ are the hyperbolas

$$\tfrac{1}{2}(p'_\beta)^2-3h\gamma^2=h. \tag{3.17}$$

The interesting case is when $c<0$, for then the components of these curves in $|\gamma|<1$ are all *closed*, and lie inside the domain bounded by $\gamma=\pm 1$ and the hyperbolas (3.17) (see Fig. 8). The point s in Fig. 8 is determined from the equation

$$\frac{-h}{c}=\frac{(1+3s^2)^{1/2}}{(1-s^2)^3}.$$

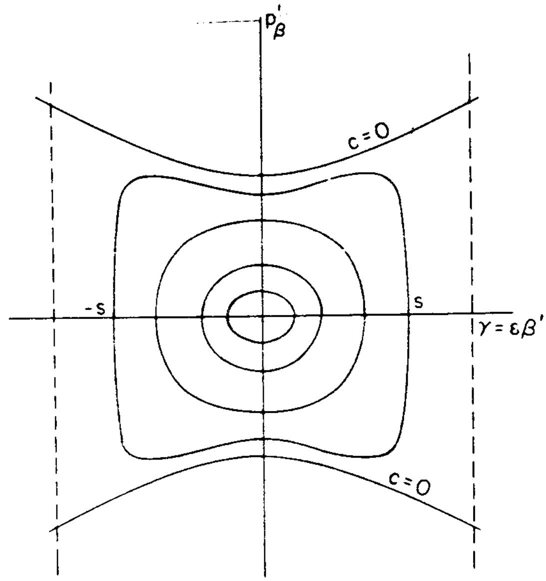

FIG. 8. *The level curves of* (3.16) *for* $c<0$.

To lowest order, the constant c is $-R$, since we are only considering the zero energy solutions of F. Thus, to lowest order, the particle "mirrors" at a latitude $\lambda=\sin^{-1} s$ where

$$\frac{h}{R}=\frac{(1+3s^2)^{1/2}}{(1-s^2)^3}.$$

This confirms our intuitive notion that the guiding center of the particle oscillates along a line of force between two mirror points.

Finally, since the curves (3.16) are closed, we may introduce the familiar action-angle variables J, ϕ in place of β', p'_β, where

$$J=\frac{4\sqrt{2}}{\varepsilon}\int_0^s\left[h(1+3x^2)+c\frac{(1+3x^2)^{3/2}}{(1-x^2)^3}\right]^{1/2}dx$$

is the area of the closed curves (3.16) in the β', p'_β-plane. In terms of the variables R, θ, J,

ϕ, the Hamiltonian F_0 takes the form

(3.18)
$$F_0 = R + c(\varepsilon J).$$

Thus, we have succeeded in showing that the Hamiltonian F_0 is *integrable*.

We are now at the stage where we can apply Moser's theorem to the full Hamiltonian

$$F = R + c(\varepsilon J) + \varepsilon F_1(R, \varepsilon J, \theta, \phi, \varepsilon).$$

On the energy surface $F = 0$ we solve for

$$R = \Psi(\varepsilon J, \theta, \phi, \varepsilon),$$

and take θ instead of t as our independent variable. We then follow all solutions from $\theta = 0$ to their next intersection with $\theta = 2\pi$. This defines a mapping M, which in the coordinates $r = \varepsilon J$, ϕ is given by

$$M: \quad \begin{cases} r_1 = r + \varepsilon^2 f(r, \phi, \varepsilon), \\ \phi_1 = \phi + \varepsilon c'(r) + \varepsilon^2 g(r, \phi, \varepsilon), \end{cases}$$

with f, g analytic in all its arguments. Now, M is not area preserving; however, each closed curve must intersect its image curve (since M *is* area preserving in the coordinates J, ϕ) and this is the main requirement on M in Moser's theorem.

The mapping M is of the form (2.1) with $\rho = 1$, $\sigma = 2$, and $\gamma(r) = c'(r)$. To apply Moser's theorem, we need only verify the condition

$$\gamma'(r) = c''(r) \not\equiv 0.$$

Unfortunately, we cannot calculate $\gamma(r)$ explicitly and must be content with numerical calculations. Fig. 9 below shows the graph of $\gamma(r)$ versus c. It is clear that $\gamma(r)$ has a single stationary point at $c = -.7$. Thus if we stay away from this point, we can apply Moser's theorem to prove the existence of infinitely many invariant tori on each energy surface. These tori contain points far from the equatorial plane. And indeed, some tori penetrate very close to the dipole singularity. Of course, the larger we choose N, the smaller we must take ε; i.e., the tighter the particle gyrates about its guiding field line.

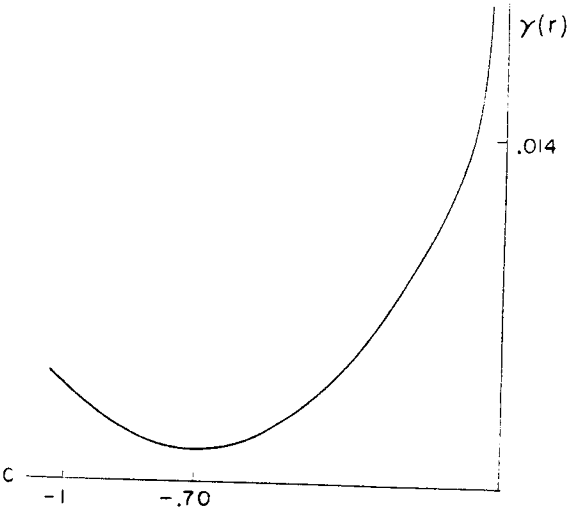

FIG. 9. *The graph of $\gamma(r)$ versus c.*

(C) Near equatorial orbits with large values of energy. The invariant tori we found in the last two sections are interesting because they indicate the possibility of confinement of charged particles. However these tori lie on energy surfaces with values of $H = E$ which are far too small to be of any practical importance. This is because the parameter ε in Moser's theorem must be extremely small. In the following two sections we would like to describe two different classes of invariant tori which lie on energy surfaces which are physically relevant. These two families were recently discovered by Braun [6], and the reader is referred to this paper for more detailed proofs.

There is an interesting distinction between the existence proofs in (A) and (B) and the existence proofs which follow. In the previous proofs we first manipulated the Hamiltonian H (in a suitable region of phase space) so that the associated Poincaré section map M had the "correct form". In the following, we will first construct the associated Poincaré section map M (again in suitable regions of phase space) and then manipulate the mapping M so that it assumes the form of a twist mapping.

If a particle starts in the equatorial plane ($\rho, \dot{\rho}$-plane), then it must remain there forever. In other words, if $z = \dot{z} = 0$ initially, then $z(t) = \dot{z}(t) = 0$, for all t. The motion of the particle in the $\rho, \dot{\rho}$-plane is then determined from the one-dimensional Hamiltonian

$$H = \frac{1}{2}\left[p_\rho^2 + \left(\frac{1}{\rho} - \frac{1}{\rho^2}\right)^2\right], \tag{3.19}$$

where $p_\rho = \dot{\rho}$. Since the energy H is conserved, we see that the qualitative properties of all equatorial orbits can be determined from the integral curves

$$E = \frac{1}{2}\left[p_\rho^2 + \left(\frac{1}{\rho} - \frac{1}{\rho^2}\right)^2\right]. \tag{3.20}$$

These integral curves are described in Fig. 10.

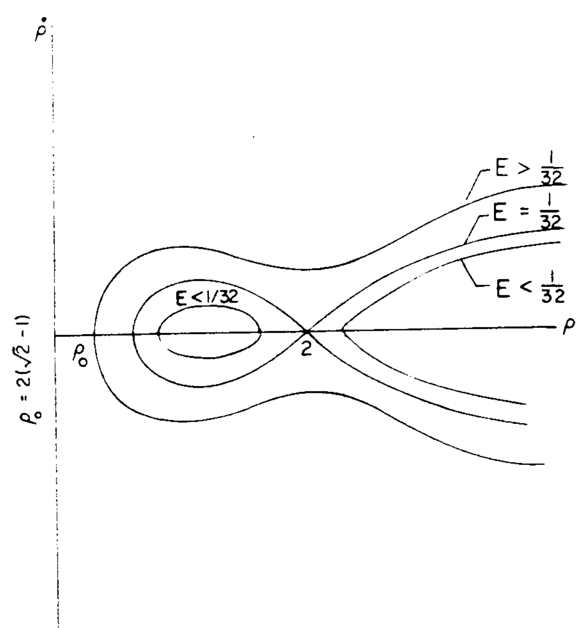

FIG. 10. *Integral curves in the* ...

The important thing to realize is that these curves are *closed* for $E<\frac{1}{32}$ and $\rho_0<\rho<2$. If we now introduce action-angle variables R, θ in place of ρ, p_ρ (where R is just the area enclosed by these closed curves), then the Hamiltonian H can be written in the form

$$H=C(R).$$

Let us now return to the full Hamiltonian

$$H=\frac{1}{2}\left[p_\rho^2+p_z^2+\left(\frac{1}{\rho}-\frac{\rho}{r^3}\right)^2\right]. \tag{3.21}$$

Expanding (3.21) in powers of z, we see that H can be written in the form

$$H=C(R)+[H_1(R,\theta)z^2+\tfrac{1}{2}p_z^2]+H_4,$$

where $H_4=O(z^4)$. For fixed R_0 (to be chosen later) we make the substitution

$$\begin{aligned} R&=R_0+\varepsilon^2R' \qquad & z&=\varepsilon z',\\ \theta&=\theta', & p_z&=\varepsilon p_z'. \end{aligned} \tag{3.22}$$

This substitution guarantees that we stay near the equatorial plane. It is not canonical; however, the differential equations of motion are still Hamiltonian with Hamiltonian $H=H/\varepsilon^2$. Neglecting the primes, we write

$$H=\frac{C(R_0+\varepsilon^2R)}{\varepsilon^2}+H_1(R_0+\varepsilon^2R,\theta)z^2+\tfrac{1}{2}p_z^2+O(\varepsilon^2).$$

Equivalently, we may take

$$H=\omega_1R+H_1(R_0,\theta)z^2+\tfrac{1}{2}p_z^2+O(\varepsilon^2), \tag{3.23}$$

where

$$\omega_1=\left.\frac{\partial C}{\partial R}\right|_{R=R_0}.$$

The constant term $C(R_0)/\varepsilon^2$ has been omitted because it does not affect the equations of motion.

On the energy surface $H=E$ we now solve for

$$R=\Phi(\theta,z,p_z,\varepsilon). \tag{3.24}$$

The motion in the z, p_z-plane is again Hamiltonian, with $H=\Phi$. However, our system is not autonomous. To eliminate the dependence of (3.24) on θ, we follow all solutions from $\theta=0$ to $\theta=2\pi$. This gives rise to a mapping M which can be written in the form

$$\begin{bmatrix} z\\ p_z\end{bmatrix}\to A\begin{bmatrix} z\\ p_z\end{bmatrix}+O(\varepsilon^2), \tag{3.25}$$

where A is a 2×2 matrix. The crucial point now is that the eigenvalues of A are on the unit circle for $\gamma_1>1.3137$, where

$$E=\frac{1}{32\gamma_1^4}.$$

This follows from the extremely important result of DeVogelaere [9]. Thus we can find new coordinates ξ, $\bar{\xi}$ such that the mapping M assumes the form

$$\xi_1=e^{i\alpha}\xi+\varepsilon^2f(\xi,\bar{\xi},\varepsilon),\qquad \bar{\xi}_1=e^{-i\alpha}\bar{\xi}+\varepsilon^2\bar{f}(\bar{\xi},\xi,\varepsilon). \tag{3.26}$$

The coordinates ξ, $\bar{\xi}$ can be chosen so that (3.26) remains area preserving and $\alpha \neq 2\pi p/q$ for any integers p and q. Following Birkhoff [26] we now introduce new coordinates which we again call ξ, $\bar{\xi}$, such that f only depends on $|\xi|^2$ to lowest order in ε. Specifically, we introduce new coordinates so that (3.26) assumes the form

$$\xi_1 = \exp(i[\alpha + \varepsilon^2 \alpha_1 |\xi|^2]\xi + O(\varepsilon^4). \tag{3.27}$$

The coefficient α_1 in (3.27) comes from the z^4 term in the expansion of (3.21). Since this term is nonzero, it follows in the implementation of Birkhoff's method that $\alpha_1 \neq 0$. Setting

$$r = |\xi|, \qquad \theta = \arg \xi,$$

we see that (3.27) has the form

$$\begin{aligned} r_1 &= r + O(\varepsilon^4), \\ \theta_1 &= \theta + \alpha + \varepsilon^2 \alpha_1 r^2 + O(\varepsilon^4). \end{aligned} \tag{3.28}$$

This is a twist mapping with $\rho = 2$ and $\sigma = 4$. Thus, for ε sufficiently small, (which in no way affects the energy E) the mapping M possesses infinitely many closed invariant curves, and these invariant curves give rise to invariant tori if we follow all solutions which emanate from them. It should be noted, though, that the smallness of ε implies *that the* region of phase space in which there are known to be invariant curves is quite small, and thus has minimal physical significance.

(D) Invariant tori which need not lie close to the equatorial plane. Another way of obtaining invariant tori for Hamiltonian systems of two degrees of freedom is to first find a periodic solution which is stable. This means that all solutions which start nearby remain close for all future time. Störmer [27] did extensive work in this area, and DeVogelaere [11] generalized his results. Recently Dragt [14] succeeded in numerically locating these solutions, and by a clever topological argument, found additional periodic solutions.

Once a stable periodic solution has been found, we can usually find a family of invariant tori nearby. This is accomplished in the following manner. First, we restrict ourselves to constant energy surfaces $H = E$, so that our phase space is three-dimensional. Second, we construct a Poincaré section map M for a suitable surface of section. (A suitable surface of section is one for which M is area preserving.) If x and y are our coordinates in the section, and if we translate the axes so that the given periodic solution crosses the section at the origin, then the mapping M can be written in the form

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} + f_2(x, y) + f_3(x, y) + \cdots. \tag{3.29}$$

The area-preserving property of M forces the eigenvalues of A to be reciprocals of each other, and the stability of the periodic solution forces the eigenvalues to lie on the unit circle. In this case, we can find a complex coordinate ξ such that

$$\xi_1 = e^{i\alpha_0}\xi + g_2(\xi, \bar{\xi}) + g_3(\xi, \bar{\xi}) + \cdots. \tag{3.30}$$

If $\alpha \neq 2\pi p/q$, $q \leqq n$, then we can put the mapping (3.30) into the normal form

$$z_1 = \exp(i[\alpha_0 + \alpha_1 |z\bar{z}| + \alpha_2 |z\bar{z}|^2 + \cdots + \alpha_n |z\bar{z}|^n])z + O(z^{n+3}).$$

As shown in Braun [6], this mapping reduces to a twist mapping if any of the α_i, for $i \geqq 1$, is unequal to zero.

It is usually an elementary though rather tedious exercise to compute $\alpha_1, \cdots, \alpha_n$. This is because the matrix A and the functions $f_2, f_3, \cdots$, can be calculated explicitly (Braun [6]). However, in the case of the Störmer problem, we are spared these calculations since the numerical work of Dragt [14] shows that the mapping (3.29) is indeed a twist.

Remark. The invariant tori obtained by this method are not physically relevant by themselves, since they occupy a very narrow region in phase space. However, they are very important because they indicate the strong possibility of trapping particles with "large" energy in regions of phase space which need not lie close to the equatorial plane.

4. The structure of solutions near the singularity. As was seen in the previous sections, the construction of an appropriate Poincaré section map is an important tool in analyzing the structure of a dynamical system. It was shown by DeVogelaere [10] that the equatorial plane is a surface of section for the Störmer problem. More precisely, let us fix an energy surface $H = E$, and follow all solutions from the equatorial plane $z = 0$ as they travel toward either pole and then return again for the first time to the equatorial plane. This defines a mapping M of the equatorial plane onto itself, and this area-preserving mapping has been the subject of much of the research (Dragt [13], [14]) on the Störmer problem. Specifically, what *is* M? An answer to this question will provide us with a global picture of the orbits for the Störmer problem.

The mapping M has been studied numerically with the following result: There exist closed invariant curves of M in that region of the equatorial plane which gives rise to near-equatorial orbits. More precisely, successive iterates of M, acting on some initial point, behave in a regular way and appear to fall on a smooth closed curve, at least if they are not examined with too great precision. This result, of course, is to be expected, since any invariant torus will intersect a surface of section in a closed invariant curve. This "integrable" or "stable" behavior changes as we move to points $(\rho, \dot{\rho})$ which give rise to orbits that penetrate to higher and higher latitudes. Specifically, the iterates of such a point, under M, do not lie on a closed curve. Rather, they seem to ramble over a "fuzzy" region. The natural conclusion, in this case, is that the invariant curves of M have disintegrated, and we have approached a region of "instability" or "nonintegrability" for the Störmer problem. However, one cannot be so sure. To wit, it is extremely difficult to accurately compute solutions which penetrate to high latitudes. This is because we must use a time step in our numerical schemes which is small compared to the period of one gyration about a magnetic field line, and a particle oscillates faster and faster as it rises to higher and higher latitudes.

To resolve this question, it is necessary to analyze the structure of all solutions in the neighborhood of the singularity $r = 0$. This is an extremely difficult problem, because the singularity is so awful. In 1907, Störmer constructed an asymptotic expansion on each energy surface, which, if it converged would represent an orbit which ran into the singularity. It was not until 1943 that Malmquist [20], in a difficult paper, proved the existence, on each energy surface, of at least one orbit which entered the singularity. The problem remained at this stage until 1970, when Braun [5] finally succeeded in proving the uniqueness of this orbit and in obtaining the structure of all solutions in the neighborhood of the singularity.

We would now like to describe these results and to indicate how they were obtained. Let us go back to adiabatic theory for a moment. In this approximation, a particle gyrates about a guiding field line, and oscillates along the line. Now, it is quite difficult to picture how all these different orbits fit together. Indeed, the only conceivable picture on each energy surface is if we *unfold* the motion along a line of force.

Then, the orbit structure should have the following form (see Fig. 11). All orbits lie on the cylinders $y_1^2+y_2^2=\text{constant}$, and move up these cylinders in the direction of increasing q.

Now, here is the key idea. If we believe that Fig. 11 is the only possible picture, and if we can find a "model case" which gives this picture, then the model case should provide the answer for the Störmer problem also. Specifically, once we see what the important variables are in the model case, and how to "unfold" the oscillation, then we can do the same for the Störmer problem. This is a really neat idea, and what is even nicer about it is that it works.

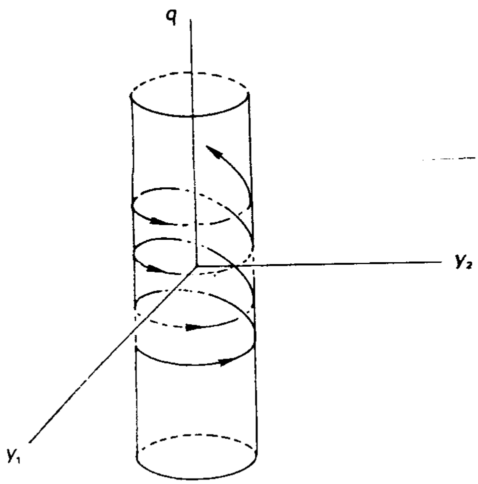

FIG. 11

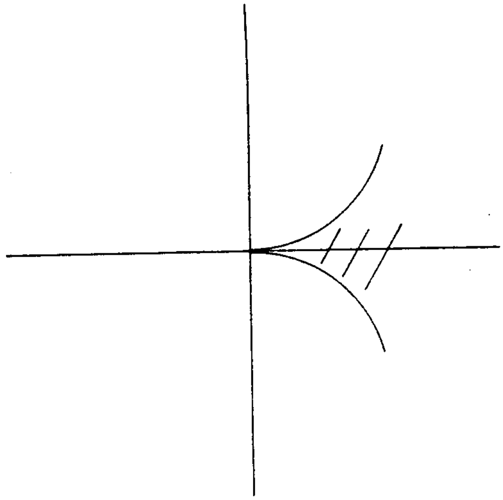

FIG. 12. *The region* $h(\theta)/r^2 \leqq E$.

A model case is provided by the Hamiltonian

$$H=\frac{1}{2}\left[p_r^2+\frac{p_\theta^2}{r^2}\right]+\frac{1}{2}\frac{h(\theta)}{r^2}, \tag{4.1}$$

where r, θ are polar coordinates and $h(\theta)$ is a convex function of θ, with $h(0)=0$. The region $h(\theta)/r^2 \leqq E$ is described in Fig. 12. Notice the similarity between this region and the region $V(\rho, z) \leqq E$ in Fig. 5.

Since we are only interested in the region

$$\frac{h(\theta)}{r^2} \leqq E, \qquad \left(H=\frac{E}{2}\right),$$

we can disregard the periodicity of $h(\theta)$ and set $h(\theta)=\theta^2$. Thus, we take

$$H=\frac{1}{2}\left[p_r^2+\frac{p_\theta^2}{r^2}\right]+\frac{1}{2}\frac{\theta^2}{r^2}. \tag{4.2}$$

There are two key steps now, which will be repeated for the Störmer problem. The first is to change the time scale along orbits to

$$\tau=\int\frac{dt}{r^2}.$$

Then, our solutions are the zero energy solutions of the Hamiltonian

$$K = \tfrac{1}{2}[r^2(p_r^2 - E) + p_\theta^2 + \theta^2].$$

The second, and major step, is the unfolding

$$q = rp_r.$$

If $y_1 = \theta$, and $y_2 = p_\theta$, then the equations of motion become

$$(4.3)\qquad \frac{dq}{d\tau} = q^2 + |y|^2, \qquad \frac{dy}{d\tau} = -iy, \qquad y = y_1 + iy_2.$$

Note that all orbits of (4.3) lie on the cylinders $|y| = c$.

Our next step is to show that the orbit structure of (4.3) is unchanged under a small perturbation. More generally, consider the system of equations

$$(4.4)\qquad \frac{dq}{dt} = r^a + f(q, y, \bar{y}), \qquad \frac{dy}{dt} = ih(q, y, \bar{y})y + g(q, y, \bar{y}),$$

where $r^2 = q^2 + |y|^2$ and h is *real*. We have the following important theorem (Braun [5]).

THEOREM. *Consider the system* (4.4), *and define*

$$\|\phi\|_R = \sup_{r<R} \sum_{k=0}^{2} r^k |\partial^k \phi|.$$

Assume that

$$\|f\|_R = O(R^{a+\varepsilon}), \quad \|g\|_R = O(R^N), \quad \|h - h(0,0,0)\|_R = O(R^c),$$

and that

$$(4.5)\qquad N > \max(a+1, 2a-c, 3a-2c-1).$$

Then the orbit structure of (4.4) *is topologically equivalent to the orbit structure of the system of equations*

$$(4.6)\qquad \frac{dq}{dt} = r^a, \qquad \frac{dy}{dt} = ih(q, y, \bar{y})y.$$

The condition (4.5) is strange, and we have no explanation for it. The only thing that is intuitively clear is that N must be greater than a. If this is not so, then the orbit structure of (4.4) may differ drastically from the orbit structure of (4.6). For example, consider the system of equations

$$(4.7)\qquad \frac{dq}{dt} = q^2 + s^2, \quad s^2 = y_1^2 + y_2^2, \qquad \frac{dy_1}{dt} = y_2 + 2qy_1, \qquad \frac{dy_2}{dt} = -y_1 + 2qy_2.$$

For this system, (with $a=2, N=2$)

$$\frac{ds}{dt}=\frac{y_1(dy_1/dt)+y_2(dy_2/dt)}{[y_1^2+y_2^2]^{1/2}}=2qs.$$

Hence for $\xi=q+s$, $\eta=q-s$ we have

$$\frac{d\xi}{dt}=\xi^2, \qquad \frac{d\eta}{dt}=\eta^2,$$

with the solutions

$$\xi=\frac{\xi_0}{1-\xi_0 t}, \qquad \eta=\frac{\eta_0}{1-\eta_0 t}.$$

This implies that any orbit of (4.7) with $\xi_0 \leqq 0$ and $\eta_0 \leqq 0$ must run into the origin as $t \to \infty$. The locus of these points is a right circular cone. Hence, an entire cone runs into the origin under (4.7) (this phenomenon is known as funneling), while a single curve $(s=0, q<0)$ only runs into the origin under (4.3).

Let us now apply all this information to the Störmer problem. Our first thought is to try to map our Hamiltonian into the model case (4.1). However, this fails because of a $4q$ term which appears in the unfolding equations (4.3). This term appears because our new coordinates are not orthogonal. This leads us to introduce new orthogonal coordinates

$$a=a(r,\theta)=\frac{r-\cos^2\theta}{r}, \qquad b=b(r,\theta)=\frac{r}{\sin^{1/2}\theta}, \tag{4.8}$$

via the generating function

$$F(r,\theta,p_a,p_b)=\left(1-\frac{\cos^2\theta}{r}\right)p_a+\frac{r}{\sin^{1/2}\theta}p_b.$$

Then the Hamiltonian

$$H=\frac{1}{2}\left(p_r^2+\frac{p_\theta^2}{r^2}\right)+\frac{1}{2r^2\cos^2\theta}\left(\frac{r-\cos^2\theta}{r}\right)^2$$

for the Störmer problem in polar coordinates is transformed into the Hamiltonian

$$\begin{aligned} H=&\left[\frac{4(1-a)}{r^3}-\frac{3(1-a)^2}{r^2}\right]\frac{p_a^2}{2}+\frac{a^2}{2(1-a)r^3} \\ &+\left[\frac{1}{[1-r(1-a)]^{1/2}}+\frac{1}{4}\frac{r(1-a)}{[1-r(1-a)]^{3/2}}\right]\frac{p_b^2}{2}, \end{aligned} \tag{4.9}$$

where r is assumed to be expressed in terms of a and b.

Taking a cue now from our model case, we now change the time scale along orbits to

$$\tau=\int\frac{dt}{r^3}.$$

Our trajectories are now the zero energy solutions of the Hamiltonian

$$K=r^3\left(H-\frac{E}{2}\right), \tag{4.10}$$

where $E/2$ is the constant value of the energy H. Solving for $r=r(a,b)$ from the equation

$$b(1-a)=\frac{r(1-a)}{[1-r(1-a)]^{1/4}}.$$

we see that (4.10) can be written in the form

$$K=\frac{a^2+4p_a^2}{2}+b^3\frac{(p_b^2-E)}{2}+k_1(a, b, p_a, p_b, E), \tag{4.11}$$

where

$$K=\sum_{j,k,l} c_{jkl}(p_b, E)a^j b^k p_a^l,$$

with

$$j+\frac{2k}{3}+l\geqq \tfrac{8}{3}.$$

For our convenience later on, we make the transformation

$$a\to\sqrt{2}\,a, \qquad p_a\to\frac{p_a}{\sqrt{2}},$$

so that

$$a^2+4p_a^2\to 2(a^2+p_a^2).$$

Taking a second cue from our model case, we make the unfolding

$$q=b^{3/2}p_b.$$

Then, the Hamiltonian (4.11) gives rise to the equations

$$\begin{aligned}\frac{dq}{d\tau}&=\frac{3}{2E^{1/6}}(q^2+2|y|^2)^{7/6}+f,\\ \frac{dy}{d\tau}&=-2iy+g,\end{aligned} \tag{4.12}$$

with $y=a+ip_a$.

Now, the system (4.12) is essentially (4.4) with $a=\frac{7}{3}$. Unfortunately, though, we cannot apply our theorem, since the function g satisfies

$$\|g\|_R=O(R^{5/3}).$$

However, the theorem can still be salvaged by the clever trick (see Braun [5]) of making K a function of $y_1^2+y_2^2$ alone through an order greater than $\frac{17}{3}$. This is accomplished via a takeoff on the Birkhoff averaging method (see Siegel [26]). The normal form terms are then lumped into an $h(q,|y|)$ term, and it is easily verified that our theorem holds. Thus, all solutions (on each constant energy surface) have the form described in Fig. 11, close to the singularity $r=0$.

5. A model of the global flow for the Störmer problem. In the previous section we showed that on each energy surface, near the dipole singularity, all orbits moved on two dimensional invariant cylinders. These cylinders when followed backwards in time, will intersect the equatorial plane $z=0$ in closed curves. The center of these curves is the intersection with the equatorial plane with the orbit that enters the singularity from the north. Exactly the same situation prevails for the orbit that enters the singularity from the south. Now, if the orbit that leaves the singularity from the north were to intersect the equatorial plane at right angles, then it would be the orbit that entered the singularity from the south. However, the orbit that leaves the singularity from the north does not cross the equatorial plane at right angles: it misses by a little bit (Dragt [14]). Consequently, the cylinders from the north and the cylinders from the south do not match up, and we would expect some chaotic behavior.

Let M be the equatorial section map described in the previous section. If we transform the closed curves which result from the intersection of the "northern" cylinders with the equatorial plane into circles, then a good model for M can be obtained as follows. First, let $\alpha(r)$ be a C^1 function satisfying

$$-r\alpha'(r) \to \infty \quad \text{as } r \to 0,$$

and set ϕ to be the mapping

$$\phi : (r, \theta) \to (r, \theta + \alpha(r)). \tag{5.1}$$

The mapping ϕ rotates circles centered about the origin by an angle $\alpha(r)$ which tends to ∞ as $r \to 0$. Next let τ be the translation

$$\tau : (x, y) \to (x + \varepsilon, y), \tag{5.2}$$

where ε is a fixed positive number. The mapping M is defined to be the composition of τ with ϕ, i.e.,

$$M = \tau \cdot \phi. \tag{5.3}$$

It has been verified numerically by Dragt [13] that (5.3) is indeed the case for M. And the analyses of the previous section would say that $\alpha(r) = r^{-4/3}$ is the best bet for $\alpha(r)$.

The mapping M in (5.3) is very similar to the mapping M studied by Braun [7], and the linked twist mappings studied by Bowen [3], Devaney [8] and Thurston [28]. And as we shall soon see, its properties are essentially the same.

At first glance it would appear that the iterates of a point under M are unbounded, on account of the translation τ. However, this is not the case. The further we move from the origin, the more the mapping M resembles a twist mapping. Hence, by Moser's theorem, the mapping M possesses closed invariant curves (for α sufficiently smooth) far enough from the origin.

Near the origin, though, the mapping M is quite erratic, as is to be expected. It possesses infinitely many unstable periodic points, as well as infinitely many homoclinic and heteroclinic points. This is a consequence of the following theorem, whose proof is due to Moser [23].

THEOREM. *Given any positive integer N, there exists an $\varepsilon(N)$ such that for $0<\varepsilon<\varepsilon(N)$ the mapping M possesses the shift on N symbols as a subsystem.*

Proof. Our first step is to construct an annulus with the property that points on the inner boundary are rotated about the origin N times more than points on the outer boundary. To this end, we choose two constants c_1, c_2, such that

$$\tfrac{1}{2} < c_1 < c_2 < \tfrac{3}{2}, \tag{5.4}$$

and consider the annulus

$$B = \{(r, \theta) | c_1\varepsilon \leqq r \leqq c_2\varepsilon\}.$$

The restriction (5.4) is imposed so that B and τB intersect in two components, and the boundaries of these components intersect under a nonzero angle, as shown in Fig. 13.

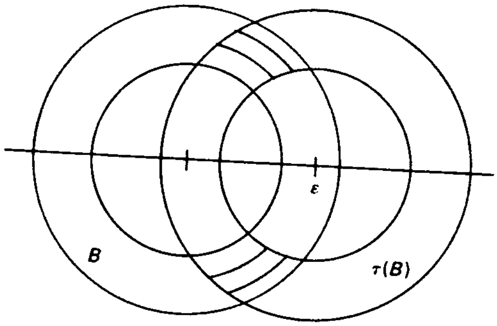

FIG. 13

Next, we form the annulus

$$A_m = \{(r, \theta) | (2m-1)\pi \leqq \alpha(r) \leqq (2m+1)\pi\},$$

and set, with some n still to be determined,

$$A = \bigcup_{s=0}^{N-1} A_{n+s}.$$

We will require that A be contained in B. To this end, note that the outer boundary of A is rotated by an amount $(2n-1)\pi$, while the inner boundary of A is rotated by an amount

$$(2(n+N-1)+1)\pi = (2n+2N-1)\pi.$$

Therefore,

$$(2n+2N-1)\pi < \alpha(c_1\varepsilon) \quad \text{and} \quad (2n-1)\pi > \alpha(c_2\varepsilon).$$

Combining these inequalities gives

(5.5) $$\alpha(c_2\varepsilon) < (2n-1)\pi < \alpha(c_1\varepsilon) - 2N\pi.$$

To satisfy (5.5) we first assume that $\alpha(r)$ is a monotonic decreasing function of r. Second, we choose ε small enough so that

$$\alpha(c_1\varepsilon) - \alpha(c_2\varepsilon) > (2N+2)\pi.$$

Finally, we choose an integer n large enough so that (5.5) holds. To show that this is possible observe that $-r\alpha' > L$ for r small enough. Then,

$$\alpha(c_1\varepsilon) - \alpha(c_2\varepsilon) = -\int_{c_1\varepsilon}^{c_2\varepsilon} \alpha'(r)\, dr > L \ln\left(\frac{c_2}{c_1}\right).$$

Now, we define the region

$$Q = A \cap \tau A \cap \{y > 0\},$$

which will play the role of the square in [22]. It is clear from Fig. 14 that the image $\phi(Q)$ spirals N times around in A and therefore intersects the region $\tau^{-1}Q$ in N components. Consequently, $M(Q) = \tau \cdot \phi(Q)$ also intersects Q in N components, $U_0, U_1, \cdots, U_{N-1}$, which are so labeled that U_0 is the outer one and U_{N-1} the inner one. (See Fig. 15.) Similarly, we see that

$$M^{-1}(Q) \cap Q = \bigcup_{s=0}^{N-1} V_s$$

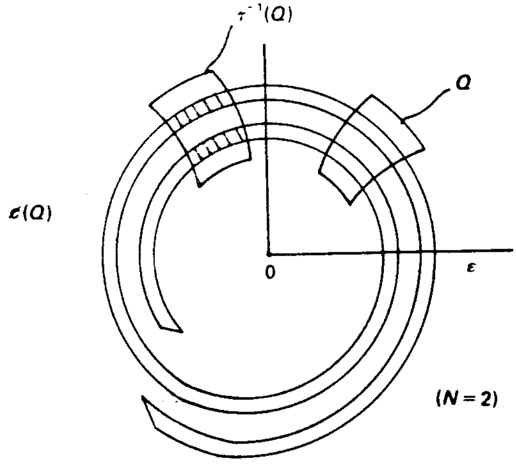

FIG. 14

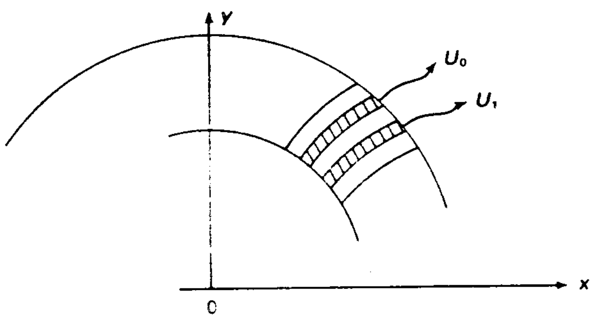

FIG. 15

has N components, labeled so that

$$M(V_s) = U_s, \qquad s = 0, 1, \cdots, N-1.$$

Finally, one also verifies the boundary conditions on V_s and U_s, as required in [22].

It remains to show that M contracts a bundle of sectors. For this purpose, let r^+ denote the distance from the point (x, y) to $(\varepsilon, 0)$, and introduce r, r^+ as coordinates in Q. This is possible since $Q \subset B \cap \tau(B)$. Then Q appears as a square contained in the larger square $c_1\varepsilon < r, r^+ < c_2\varepsilon$. Let λ be a number greater than one, and introduce the bundle of sectors

$$\text{(5.6)} \qquad \begin{aligned} S^+&: |\delta r^+| < \lambda^{-1}|\delta r|, \\ S^-&: |\delta r| < \lambda^{-1}|\delta r^+|; \end{aligned}$$

as well as the larger sectors

$$\text{(5.6)}' \qquad \begin{aligned} \Sigma^+&: |\delta r^+| < \lambda|\delta r|, \\ \Sigma^-&: |\delta r| < \lambda|\delta r^+|. \end{aligned}$$

We have the following crucial lemma:

LEMMA. *If ε is sufficiently small, and*

$$dM: (\delta r, \delta r^+) \to (\delta r_1, \delta r_1^+),$$

then,

$$dM(\Sigma^+) \subset S^+ \quad \textit{and} \quad |\delta r_1| > \lambda |\delta r| \quad \textit{for } (\delta r, \delta r^+) \in S^+,$$

and

(5.7) $$(dM)^{-1}(\Sigma^{-1}) \subset S^- \quad \textit{and} \quad |\delta r_1| < \lambda^{-1} |\delta r| \quad \textit{for } (\delta r_1, \delta r_1^+) \in S^+.$$

Proof of Theorem: From the lemma,

$$dM(S^+) \subset dM(\Sigma^+) \subset S^+,$$

and S^+ is invariant under dM as is S^- under $(dM)^{-1}$. Thus, we have verified hypothesis H3 of [22], which proves our theorem.

Proof of Lemma. We will only verify the first part of (5.7) since the other part follows in exactly the same manner. Stretching the variables, we set $\rho = \varepsilon^{-1} r$, $\rho^+ = \varepsilon^{-1} r^+$. The coordinate ρ^+ can be expressed in terms of ρ, θ by the relation

(5.8) $$\rho^+ = f(\rho, \theta) = 1 + \rho^2 - 2\rho \cos \theta.$$

Note that ρ^+ is independent of ε. The domain $B \cap \tau(B)$ is now described by $c_1 \leqq \rho, \rho^+ \leqq c_2$. Moreover, we can find a constant $c > 1$, which depends only on c_1, c_2 such that

$$|f_1|, |f_2| \quad \text{and} \quad |f_2^{-1}| \leqq c \quad \text{in } Q,$$

where $f_1 = \partial f / \partial \rho$ and $f_2 = \partial f / \partial \theta$.

The smallness condition we impose on ε is that

(5.9) $$-\varepsilon \alpha'(\varepsilon \rho) > 2c(\lambda + c) \quad \text{in } B \cap \tau(B).$$

Equation (5.9) can easily be satisfied since $-r\alpha'(r) \to \infty$ as $r \to 0$.

Now, if $(\delta\rho, \delta\rho^+) \subset \Sigma^+$, i.e., $|\delta\rho^+| < \lambda |\delta\rho|$, then from the relation

$$\delta\rho^+ = f_1 \delta\rho + f_2 \delta\theta.$$

we deduce that

$$|\delta\theta| \leqq |f_2^{-1}| |\delta\rho^+ - f_1 \delta\rho|$$
$$\leqq c(\lambda + c)|\delta\rho| = k|\delta\rho|,$$

with $k = c(\lambda + c)$. Thus Σ^+ is contained in the sector $|\delta\theta| < k|\delta\rho|$.

Next, we apply ϕ, which maps (ρ, θ) into

$$\rho_* = \rho, \qquad \theta_* = \theta + \alpha(\varepsilon\rho),$$

so that in A

$$\delta\theta_* = \delta\theta + \varepsilon\alpha'(\varepsilon\rho)\delta\rho_*,$$

and in the previous sector we have over A that

(5.10) $$|\delta\theta_*| > (\varepsilon\alpha'(\varepsilon\rho) - k)|\delta\rho_*| > k|\delta\rho_*|,$$

since $\varepsilon\alpha'(\varepsilon\rho) > 2k$ by (5.9).

We express this condition in terms of ρ, ρ^- in $\tau^{-1}(Q)$, where $\rho^- = \varepsilon^{-1} r^-$, with r^- being the distance from (x, y) to $(-\varepsilon, 0)$. To this end, observe that

$$\rho^- = f(\rho, \pi - \theta),$$

so that

(5.11) $$\delta\rho_*^- = f_1 \delta\rho_* - f_2 \delta\theta_*.$$

Therefore,

$$|\delta\theta_*| \leq c[|\delta\rho_*^-| + c|\delta\rho_*|],$$

and from (5.10) we see that

$$k|\delta\rho_*| < c|\delta\rho_*^-| + c^2|\delta\rho_*|,$$

or

$$(k - c^2)|\delta\rho_*| < c|\delta\rho_*^-|.$$

Since $k = c(\lambda + c)$ we conclude that

$$|\delta\rho_*| < \lambda^{-1}|\delta\rho_*^-|, \quad \text{or } |\delta r_*| < \lambda^{-1}|\delta r_*^-|.$$

Finally, we apply the translation τ. Since

$$d\tau(\delta r_*^-, \delta r_*) = (\delta r_1, \delta r_1^+),$$

where $\delta r_1 = \delta r_*^-$, $\delta r_1^+ = \delta r_*$, we find that

$$|\delta r_1^+| < \lambda^{-1}|\delta r_1|,$$

as was claimed in (5.7). To get the inequality $|\delta r_1| > \lambda|\delta r|$, observe from (5.10) that for $(\delta r, \delta r^+) \in \Sigma^+$,

$$|\delta\theta_*| > k|\delta\rho|.$$

Therefore, from (5.11)

$$\begin{aligned}
|\delta\rho_*^-| &\geq c^{-1}[|\delta\theta_*| - c|\delta\rho_*|] \\
&= c^{-1}[|\delta\theta_*| - c|\delta\rho|] \\
&> c^{-1}(k - c)|\delta\rho| \\
&= (c^{-1}k - 1)|\delta\rho| \\
&= (\lambda + c - 1)|\delta\rho| \\
&> \lambda|\delta\rho|,
\end{aligned}$$

since $c > 1$. Hence

$$|\delta r_*^-| > \lambda|\delta r|,$$

and applying τ leads to

$$|\delta r_1| > \lambda|\delta r|,$$

as claimed. Q.E.D.

As was mentioned previously, the above theorem implies that M possesses infinitely many homoclinic points. It has been conjectured that area-preserving mappings are transitive in a global neighborhood of a homoclinic point. Numerical experiments for the special case of the two-circle rotation studied by Braun [7], are suggestive of and consistent with transitivity. Interestingly enough, the mapping M also appears to be transitive, as is seen in Fig. 16, and this figure is exactly the same as the instability observed in [7].

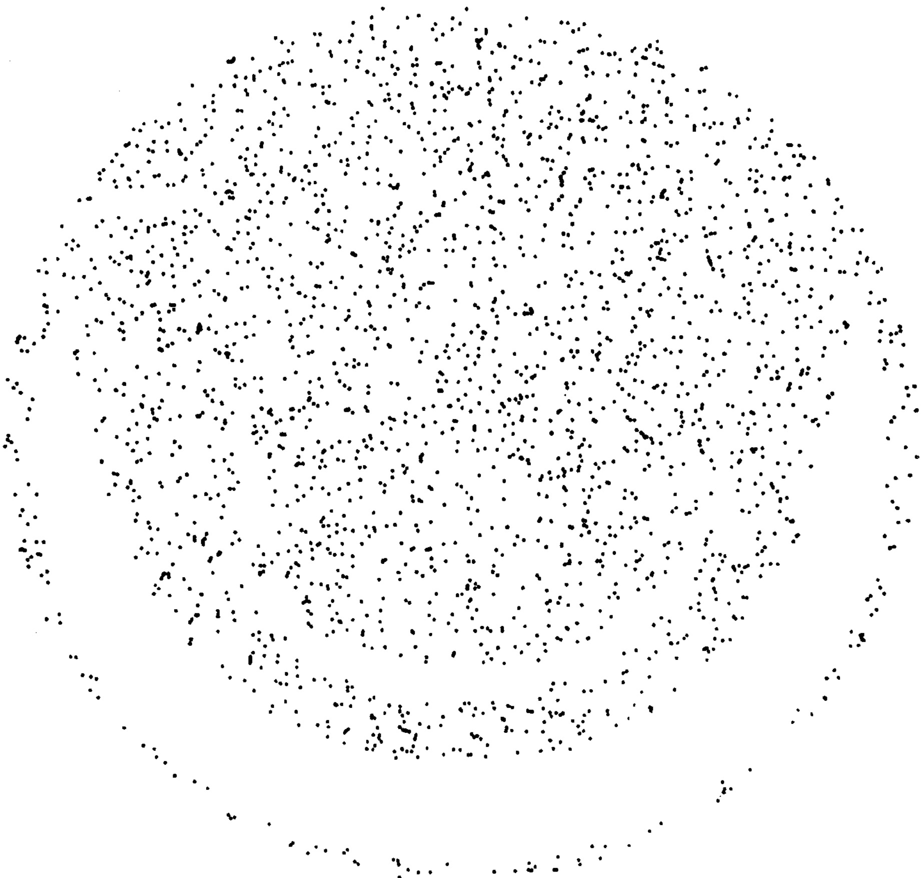

FIG. 16. 2500 *iterates of the point* $(.1, 0)$ *for* $\alpha(r) = r^{-4/3}$ *and* $\varepsilon = .01$.

REFERENCES

[1] H. ALFVÉN, *Cosmical Electrodynamics*, Clarendon Press, Oxford, 1950.

[2] V. I. ARNOLD, *Small divisor and stability problems in classical and celestial mechanics*, Uspekhi Mat. Nauk., 18 (114) (1963), pp. 85–191.

[3] R. BOWEN, *On Axiom A Diffeomorphisms*, CBMS Regional Conference Series on Mathematics 35, American Mathematical Society, Providence, RI, 1978.

[4] M. BRAUN, *Particle motions in a magnetic field*, J. Differential Equations, 8 (1970), pp. 294–332.

[5] ———, *Structural stability and the Störmer problem*, Indiana Univ. Math. J., 20 (1970), pp. 469–497.

[6] ———, *On the stability of the Van Allen radiation belt*, SIAM J. Appl. Math., 37 (1979), pp. 464–468.

[7] ———, *Invariant curves, homoclinic points, and ergodicity in area preserving mappings*, to appear.

[8] R. L. DEVANEY, *Subshifts of finite type in linked twist mappings*, preprint.

[9] R. DEVOGELAERE, *L'Equation de Hill et le problème de Störmer*, Canad. J. Math., 2 (1950), pp. 440–456.

[10] ———, *Surface de section dans le problème de Störmer, Acad. Roy. Belg. Bulletin Classe des Sciences*, 40 (1954), pp. 705–714.

[11] ———, *On the structure of symmetric periodic solutions of conservative systems, with applications*, in Contributions to the Theory of Nonlinear Oscillations, Princeton University Press, 1958, pp. 53–84.

[12] A. J. DRAGT, *Trapped orbits in a magnetic dipole field*, Rev. Geophys., 3 (1965), pp. 255–298.

[13] A. J. DRAGT AND JOHN M. FINN, *Insolubility of trapped particle motion in a magnetic dipole field*, J. Geophys. Res., 81 (1976), pp. 2327–2340.

[14] ———, *Stability of charged particle motion in a magnetic dipole field*, preliminary report.

[15] C. S. GARDNER, *The adiabatic invariants of periodic classical systems*, Phys. Rev., 115 (1959), pp. 791–794.

[16] O. GODART, *Periodic orbits, stability and resonances*, in Proceedings of Symposium at the University of Sao Paulo, G. E. Giacaglia, ed., Reidel, Dordrecht, 1970.
[17] H. GOLDSTEIN, *Classical Mechanics*, Addison-Wesley, Reading MA, 1959.
[18] G. HELLWIG, *Uber die Bewegung geladener Teilchen in schwach veränderlicher Magnetfeldern*, Zeits Naturforsch, 10a (1955), pp. 508–516.
[19] M. KRUSKAL, *Rendiconti del Terzo Congresso Internationale sui Fenomeni D'Ionezzaziane nei Gas Tennuto a Venezia*, Societe Italiana di Fiscia, Milan, 1957.
[20] J. MALMQUIST, *Sur les systèmes d'equations differentielles*, Ark. Math. Astr. och Fysik. 30A, 5 (1944), pp. 1–8.
[21] J. K. MOSER, *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Göttingen: Math. Phys. Kl., IIa (1962), pp. 1–20.
[22] ———, *Stable and Random Motions in Dynamical Systems*, Annals of Mathematical Studies, No. 77, Princeton University Press, Princeton, NJ, 1973.
[23] ———, Private communication.
[24] T. G. NORTHROP, *The Adiabatic Motion of Charged Particles*, Interscience, New York, 1963.
[25] H. POINCARÉ, *Methodes nouvelle de la mécanique célèste*, Gauthier-Villars, Paris, 1893, Vol. 2, Chapter 2.
[26] C. L. SIEGEL, *Lectures on Celestial Mechanics*, Springer-Verlag, New York, 1970.
[27] C. STÖRMER, *The Polar Aurora*, Clarendon Press, Oxford, 1955.
[28] W. THURSTON, *On the geometry and dynamics of diffeomorphisms of surfaces*, to appear.