Discussion of the Origin of the Terrestrial and Solar Magnetic Fields

By H. ALFVÉN, Royal Institute of Technology, Stockholm.

(Manuscript received 21 January 1950).

Abstract

The origin of the terrestrial and solar magnetic fields is discussed. An original meridional field can by hydrodynamic motion get an east-west component. It is pointed out that this component cannot exceed the order of magnitude of the original field. If it does, the field becomes unstable and an increase of the meridional component is caused. The possibility of finding an amplifying mechanism of this type is considered.

1. The origin of the terrestrial and solar magnetic fields is still one of the most puzzling problems of cosmical physics. Of the many different hypotheses that have been proposed, LARMOR's idea that the fields are due to internal motions seems to be the most attractive one [1]. Recent work has shown that a magnetic field could be amplified very much by hydrodynamic motions [2] [3] [4]. For example, it seems possible to account for the strong magnetic fields of sunspot by assuming that motions in the interior of the sun [2] have amplified the rather weak general solar field. In a similar way it is likely that in interstellar space an originally weak field has been amplified by hydrodynamic motions [4]. In analogy with these processes it seems reasonable that it should be possible to account even for the terrestrial and solar magnetic fields in terms of some sort of magneto-hydrodynamic process.

2. In the discussion of the general magnetic fields a paper by COWLING has attracted much interest [5]. COWLING draws the attention to the fact that the main problem is how new magnetic lines of force are created. A certain line of force may be deformed to any configuration by hydrodynamic motion and in this way a given magnetic field may be amplified. At the same time, however, the structure of the field becomes more and more complicated. The only way to amplify a field without changing its general character is to breed new lines of force. Under the conditions in the interior of the earth and sun COWLING is inclined to consider such a breeding as impossible.

Extensive investigations have been carried out by ELSASSER and BULLARD [6]. They doubt that COWLING's conclusion is valid without restrictions and try to find a rather complicated system of whirls by which an amplification should be possible. Their attempts have had no definite success so far. Upon the whole one has the impression that a phenomenon of such a general character as the terrestrial and solar fields should be due to a mechanism more simple than those proposed by these authors.

The kinematics of a field amplification

3. In order to find how the general magnetic fields of the earth and the sun are generated and maintained, it is advisable to start by trying to find a cyclic process by which a magnetic field could be amplified.

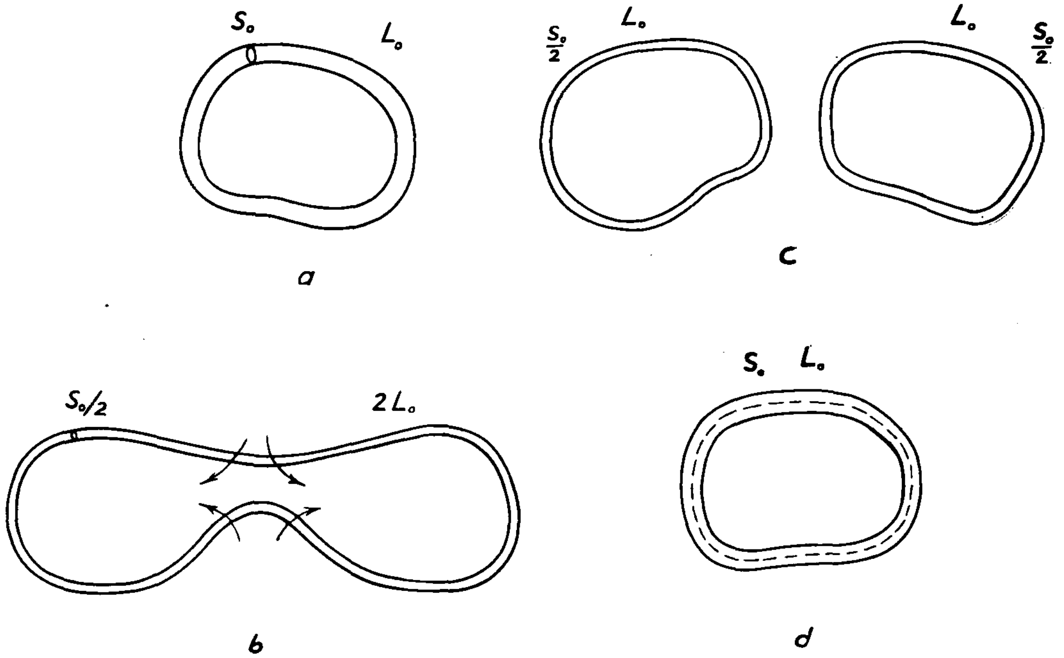

Fig. 1. The magnetic field in a closed tube of force can be amplified by hydrodynamic motion in the following way. The initial tube (*a*) is drawn out to the double length (*b*) and by a motion of the type marked by arrows this tube is split up into two tubes (*c*). These are placed side by side thus forming a tube (*d*) of the same shape as the initial tube but with the magnetic field doubled.

It is rather easy to outline a kinematically possible process but to find the dynamics leading to this process is much more difficult. Consider a closed infinitely thin tube of force in an incompressible and infinitely conducting liquid (fig. 1 a). The circumference of the lines of force be L_0 and the cross-section of the tube S_0. To simplify the discussion we suppose that S_0 is constant. Then the volume of the tube of force is $V_0 = L_0 \cdot S_0$. As the liquid has an infinite conductivity no motion perpendicular to the lines of force is possible, and as the liquid is incompressible the volume U must remain constant. If the magnetic field is H_0, the flux is $\Phi_0 = S_0 H_0$.

We want to find a process which ends in a tube of the same dimensions but with an increased magnetic field. This could be accomplished in the following way. Let hydrodynamic motions lengthen the tube to $2L_0$. As V_0 and Φ_0 remain constant, this means that the cross-section diminishes to $S_0/2$ and the magnetic field increases to $2\,H_0$. By a motion of the type shown in fig. 1 b it is possible to split up this ring into two rings, each with the length L_0 (fig. 1 c). These two rings are placed side by side, thus forming one ring with the length L_0 and cross-section S_0, but with the magnetic field $2\,H_0$ (fig. 1 d). This process could be repeated indefinitely thus leading to an unlimited amplification of H.

4. The process considered refers to lines of force which are closed in the liquid, but it is also possible to increase the magnetic field by a process starting by open lines of force. In order to show this let us suppose that in an infinitely conducting sphere of incompressible liquid there is a uniform tube of flux from one pole to the other (fig. 2 a). If the upper half of this is rotated in relation to the lower half the magnetic lines of force will be twisted, so that the axial field H_0 is superimposed by a tangential field H_t. When the twisting exceeds a certain limit, the tube of force tends to make a loop (Fig. 2 b). The mechanical analogy of this phenomenon is well-known: if a wire is twisted, a series of loops are formed. (This is easily demonstrated by rotating the micro-

phone of an ordinary table telephone and observing the telephone cord.)

The order of magnitude conditions for loop formation can easily be calculated. Suppose that a cylindrical tube of force with the length L and radius $= R$ ($\ll L$) is twisted in one of its ends by an angle φ. In a volume element at the axial distance r the axial field H_a is superimposed by a tangential field

$$H_t = H_a \frac{r\varphi}{L} \quad (1)$$

Thus the total energy of the tube of force is

$$W_a = \frac{L}{8\pi}\int_0^R (H_a^2 + H_t^2)\, 2\pi r \,\mathrm{d}r = \frac{H_a^2 R^2}{8} \cdot Y \quad (2)$$

with

$$Y = L + \frac{R^2 \varphi^2}{2L} \quad (3)$$

If the tube of force changes its shape by making a loop, its length increases by at least $\Delta L = 2\pi R$. At the same time the twist changes by $\Delta\varphi = -2\pi$. As the volume $\pi R^2 L$ remains unchanged the radius changes by $\Delta R = -\frac{R}{2L}\Delta L = \frac{\pi R^2}{L}$. Further the flux $\pi R^2 H_a$ also remains constant, which means that the magnetic field changes by $\Delta H_a = \frac{2\pi H_a R}{L}$. The condition that the total energy should not change gives

$$\Delta W_a = 0 \quad (4)$$

or

$$\varphi = (\sqrt{5} - 1)\; L/R \quad (5)$$

When φ exceeds this value the conditions for loop formation are favourable. By introducing (5) into (1) this approximate calculation shows that as soon as the tangential field exceeds the order of magnitude of the axial field a loop will be formed. Unlike the conditions in a telephone cord this loop will separate because the lines of force of the initial string will close again and so will the loop itself (fig. 2 c).

Let hydrodynamic motion draw out part of this loop to the surface (fig. 2 d), where

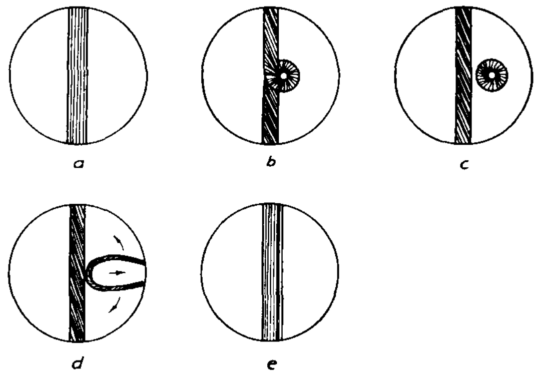

Fig. 2. Amplification of a magnetic field by loop formation. A magnetic tube of force within a conducting sphere (*a*) is twisted. This leads to the formation of a loop (*b*) which separates from the initial tube (*c*). By a hydrodynamic motion the loop may be split up and brought side by side with the initial tube (*d*) thus restoring the shape of the tube (*e*) but with an amplified magnetic field.

it will break, and move the ends to the poles, so that the original loop is deformed into a uniform tube which is parallel to the initial tube (fig. 2 e). At the loop formation the volume of the initial tube has decreased by $2\pi^2 R^3$ so that as the length is unchanged its cross-section has diminished by factor $\frac{2\pi R}{L}$. As the axial flux remains the same, the field has increased by a factor $1 + \frac{2\pi R}{L}$. If the tube deriving from the loop is added to the initial tube, the result is a tube with the initial volume and shape but with a field which has increased by $1 + \frac{2\pi R}{L}$. An iteration of the process will bring us up to any value of H_a.

The dynamics of a field amplification

5. In the preceding paragraphs it has been shown that an amplification of a magnetic field within a conducting liquid is kinematically possible. (This means also that the breeding of new lines of force is no really serious problem.) What remains in order to explain the generation of the terrestrial and solar magnetic fields is to find the dynamics which produces the postulated motions.

The dynamical problem is much more difficult because the behaviour of a set of electric currents in a fluid leads to mathematically very complicated problems, which have not been solved so far. By analogy with current-carrying wires certain conclusions could be drawn but they are only tentative.

Two parallel wires attract each other if their currents flow into the same direction. If the wires make an angle they tend to move so as to become parallel. A single loop of arbitrary form tends to expand and if free it will form itself into a circle. If a circular current is given, magnetic dipoles in its neighbourhood tend to move to the surface of the circle and place themselves in such a position as to amplify the magnetic field. If a multitude of current-carrying wires are placed near together they tend to move into such positions as to make the resulting dipole moment a maximum.

If the same laws hold for currents in a conducting liquid—but this could be established only after rigid calculations—currents in different parts of the liquid should adjust themselves so as to amplify each other. Eventually the result may be that the dipole moment becomes a maximum. Hence if only new lines of force are bred, they would move into such positions as to amplify the already existing field.

6. In §§ 3 and 4 we have only treated the magnetic lines of force without mentioning the currents producing them. In the ideal case of a single tube of flux the field is produced by currents flowing around the tube in an infinitely thin surface. In the case of a twisted flux tube there is also a current parallel to the tube. In the case of a uniformly twisted tube of force discussed in § 4 a current with homogeneous density $i = H_T/r$ flows inside the tube, the circuit being closed by a return current in an infinitely thin layer at the surface of the tube.

A set of force tubes of this ideal type will not interact. In reality, however, the currents spread in the environment, producing a magnetic coupling between the different parts of the liquid. This gives rise to mechanical forces which change the state of motion of the liquid. Accepting the analogy with the current-carrying wires we should expect the net result to be that the currents amplify each other.

Generation of magnetic fields in celestial bodies

7. We shall now try to apply what has been stated above to the generation of the general magnetic fields. We consider the liquid core of the earth as the place where the earth's magnetic field is generated. This core has a radius of about half the earth's radius. In the case of the sun the whole sun is gaseous so that the same considerations could be applied to the whole sun. In a liquid with the conductivity σ we have

$$\vec{i} = \sigma\left(\vec{E} + \frac{1}{c}\vec{v} \times \vec{H}\right) \qquad (1)$$

$$\text{rot } \vec{E} = -\frac{1}{c}\frac{\partial \vec{H}}{\partial t} \qquad (2)$$

Combining these equations and integrating along a closed line we find

$$\frac{\partial \Phi}{\partial t} = \int (\vec{v} \times \vec{H})\, \vec{ds} - c \int \frac{\vec{i}}{\sigma} \vec{ds} \qquad (3)$$

where Φ means the magnetic flux through the surface limited by the line, and ds the line element.

The first term at the right hand side of (3) means the net flux transported by hydro-

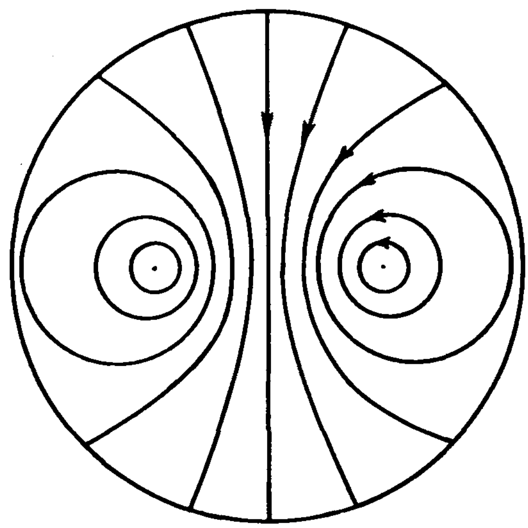

Fig. 3. Simple model of the magnetic field in the interior of the earth (or sun) with meridional field only.

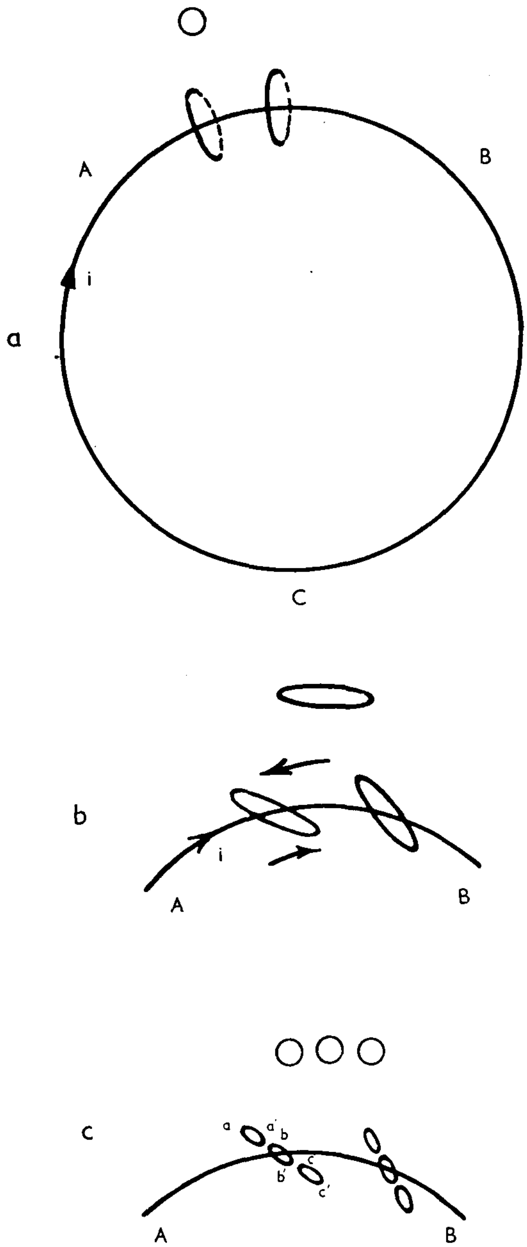

Fig. 4. Amplification of terrestrial field by means of the process of fig. 1. The neutral line ABC (fig. 4 a) is surrounded by circular magnetic lines of force (cross-section shown above). A differential hydrodynamic flow indicated by arrows in (*b*) deforms the circles into ellipses. These split up for example into three almost circular lines (*c*).

dynamic motion over the line of integration. If the flux Φ shall not diminish this transport must at least compensate the second term to the right, which denotes the consumption of flux due to the finite conductivity.

If we integrate over a circle in the equatorial plane of the earth, the net flux within this circle is directed southward. In order to amplify the field we must have a transport of southward flux into the circle, or a transport of northward flux out of the circle, or both.

Fig. 3 shows the simplest possible model of the field in the terrestrial core or in the sun. The field has meridional symmetry and is produced by a current which is more or less concentrated in a circle in the equatorial plane. Near this circle there is a circle called the "neutral line" where the magnetic field is zero.

If our integration line coincides with the neutral line there could obviously be no transport of flux over the circle because $H = 0$. Hence in this model the magnetic field cannot be maintained or amplified in a stationary state.

This difficulty, which has been stressed by Cowling [5], could be overcome by making the model asymmetrical and non-stationary. In fact, a neutral line $H = 0$ exists only because of the assumed symmetry. In a magnetic field of arbitrary form there exist singular points $H = 0$ and the neutral curve is not a continuous line but an assembly of such singular points.

The possibility of a continuous transport of flux over a given circle in the equatorial plane is connected with the breeding of new lines of force. Without such a breeding the transport could only go on until all flux going southward through the equatorial plane has been brought within the circle and all northward flux has been brought outside it. Thus in order to maintain or amplify the field it is necessary to breed northward flux inside, or southward flux outside the circle, or both.

8. In order to study the process more in details we start from the model of fig. 3. Admittedly it has been shown that this cannot be exactly correct but it is useful as a first approximation and the necessary deviations may be considered as disturbances.

The breeding of new lines of force may go on in the whole volume of the terrestrial core and of the sun but the processes are somewhat different in different parts. In this paragraph we shall consider a process which may be active near the neutral line whereas in the next paragraph we shall consider the breeding near the axis.

In the absence of a tangential field the lines of force near the neutral line are circles around this line (fig. 4 a). If the angular velocity of a tangential motion in the equatorial plane differs outside and inside the neutral line, the lines of force will be dragged out (Fig. 4 b). In this way mechanical motion is converted into magnetic field energy. If the line of force is prolonged very much it may divide into two or more loops. Speaking of the currents the dragging out corresponds to the production of a radial sheet of current due to the polarisation field $\vec{E} = \frac{1}{c}\,\vec{v} \times \vec{H}$ produced by the differential motion. The splitting of the prolonged loop means that the current density in the sheet is no longer uniform. It can be shown that an infinite sheet of current in a conducting liquid is unstable and will have a tendency to break up into separate lines of current. In this way southward lines of force are created at a' outside the line of integration ABC and northward lines of force are created at c inside the same line. A transport of the region a' inwards and of the region c outwards will give an increase of southward flux inside the line of integration. How this transport is effected dynamically is very difficult to analyse. The breaking up of the extended lines of force into smaller loops corresponds to the main current along the line ABC being twisted. If this twisted current is reshaped into a circle the expected transport is effected. It seems reasonable to assume that the current has a tendency to return to the circular shape. It should be observed that the breaking up of an extended line of force occurs when the tangential field surpasses the order of magnitude of the meridional field.

9. The breeding may also occur far away from the neutral line by means of the process outlined in § 4. This seems to be especially important near the axis. If the lines of force in this region are twisted by hydrodynamic motions, loops of magnetic flux will be formed. If a series of loops are bred symmetrically around the axis they will join to a toroidal field associated with a circular current in the equatorial plane. This current will expand due to its own magnetic pressure and also because of the attraction from the initial current near the neutral line. It is possible that the mutual attraction may prevent the initial current from expanding which it naturally should do because of its own magnetic pressure. If so the two current rings may coalesce so that there is a certain current (near the neutral line) which from time to time is fed by new rings. It is also possible that the general field derives from a series of expanding current rings, moving outwards in a long succession. In both cases there is a northward flux just outside the expanding currents which when passing a given line of integration effects the necessary transport of flux.

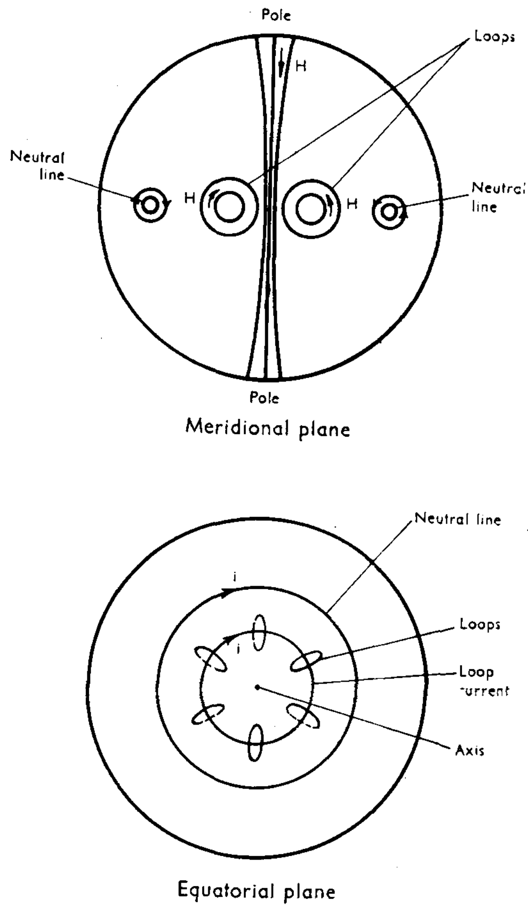

Fig. 5. Amplification of terrestrial field by the process of fig. 2. By twisting a tube of force loops are formed. The "loop current" of a symmetrical set of loops amplifies the initial current near the neutral line.

The Earths magnetic field

10. The process which we have considered is of a very general type. It should occur as soon as the average energy due to hydrodynamic motion within a part of the body exceeds the magnetic energy. The equality of these two sorts of energy is characteristic for the magneto-hydrodynamic waves. If such a wave is started by a hydrodynamic motion, half of the kinetic energy is transformed into magnetic energy.

Starting from a meridional field of a simple structure, the hydrodynamic motions may drag out the lines of force and effect a breeding of new lines of force directly, but a more obvious process is that the motions at first give the field a tangential component, which corresponds to a twisting of the lines of force. When this exceeds the meridional component new lines of force are bred and in this way the meridional field is amplified.

If the average field within a body is H the average motion which is necessary has the kinetic energy

$$\frac{\varrho v^2}{2} = \frac{H^2}{8\pi}$$

or

$$v = \frac{H}{\sqrt{4\pi\varrho}}$$

where ϱ is the density and v the velocity. For the earth's core we may put $H = 20$ gauss, $\varrho = 10$ which gives $v = 2$ cm/sec. For the sun $H = 100$, $\varrho = 1$ gives $v = 30$ cm/sec.

In order to maintain the general field at a certain strength the hydrodynamic motion must be supplied continuously with energy. ELSASSER and BULLARD suppose that the sourse of energy is nuclear reactions. Radioactivity in the earth's core produces a heating which gives rise to convection. In the solar core there is a very large release of nuclear energy.

Against this it may be objected that if these processes were the cause of the general fields we should expect a much larger difference in strength. The field at the surface of the earth's core can be estimated to about 5 gauss and this is of the same order as the field at the solar surface (10—20 gauss). Hence even the dipole moment per unit volume is of the same order. It is very unlikely that two fields of about the same strength should be caused by the release of nuclear energy, which is enormously much larger in the sun than in the earth. Admittedly it may be unfair to compare the earth's core with the whole sun because the magnetic field may be generated mainly in the solar core, but even so the difference in magnetic field would be very much smaller than the difference in energy production.

11. As an alternative there is the possibility that the energy is supplied by the non-uniform rotation which means that the ultimate cause is the braking of the rotating bodies by interaction with the environment. In the case of the earth the main braking is due to the tides, the action of which lengthen the day by 10^{-5} sec. per year. The tides brake the crust and if the braking is transmitted from the crust to the core by electromagnetic forces the energy released by the braking of the core may be converted into magneto-hydrodynamic energy. BULLARD [6] means that it is impossible that a differential rotation in the earth's core could be large enough to account for the magnetic field, because the magnetic field itself would force the core to rotate with about the same angular velocity. This conclusion depends upon what assumptions we make about the shape of the magnetic field in the core. If the field of the earth's core has the shape of fig. 3, that part of it from which magnetic lines of force go out to the crust, must rotate with almost the same velocity as the crust, according to FERRARO's theorem [7]. That toroidal part of the core, however, from which no magnetic lines of force go out to the crust has no direct electro-magnetic contact with this and may very well have a different angular velocity.

Suppose that the terrestrial dipole moment is a and the core has the radius R. If within the core there is a fairly uniform current density i in the tangential direction, this is of the order

$$i = a\,R^{-4}$$

The Joule energy developed per second by this current is of the order

$$p = i^2\,R^3\,\sigma^{-1} = a^2\,R^{-5}\,\sigma^{-1}$$

where σ is the electric conductivity. This should be compared with the energy developed in transferring rotational energy from the core to the crust. In order to find the order of magnitude, let us suppose that the core rotates as a rigid sphere thus neglecting that a part of it is magnetically bound to rotate with the same velocity as the crust. The moment of inertia be I_1 and the angular velocity ω_1, whereas the corresponding quantities of the crust are I_2 and ω_2. When the angular momentum $I_1\, d\omega_1 = -I_2 d\omega_2$ is transferred, the energy

$$dW = -\left[d\left(\frac{I_1\omega_1^2}{2}\right) + d\left(\frac{I_2\omega_2^2}{2}\right)\right] = $$
$$= (\omega_2 - \omega_1)\, I_1\, d\omega_1$$

goes into heat. If the "friction" is due to electro-magnetic forces, this energy is at first converted into magnetic energy, which finally is dissipated into Joule heat. Supposing that most of the magnetic energy is used to produce the general magnetic field we may put

$$p = \frac{dW}{dt}$$

or

$$a^2 R^{-5} \sigma^{-1} = I_1\, \omega' \frac{d\omega_1}{dt}$$

where

$$\omega' = \omega_2 - \omega_1$$

is the differential angular velocity between the crust and the core. Putting $I_1 = \frac{2}{5} \varrho R_1^5$ (ϱ = density) we find

$$\omega' = \frac{5}{2} \frac{a^2}{\sigma \varrho R_1^{10} \frac{d\omega_1}{dt}}$$

We put $a = 8 \cdot 10^{25}$ gauss cm^3, $\varrho = 10$ g cm^{-3}, $R_1 = 3.5 \cdot 10^8$ cm. For $\frac{d\omega_1}{dt}$ we choose $3 \cdot 10^{-22}$ sec^{-2} which is the value of the retardation of the crust due to the tides, thus assuming that a stationary state is reached so that the core is retarded at the same rate as the crust. Accordind to BULLARD we have $\sigma = 3 \cdot 10^{-6}$. This gives

$$\omega' = 7 \cdot 10^{-8} \text{ sec}^{-1}$$

which corresponds to an excess rotational velocity of the core of about 0.1 %. There seems to be no certain information about the non-uniform rotation of the earth's core so this value cannot be checked at the present stage but does not seem unreasonable.

Thus even the braking of the earth suffices to maintain the terrestrial field. It is difficult to judge whether the energy is supplied in this way or by nuclear reactions or in both ways.

The difference between slow and rapid motion

12. If a force acts perpendicular to a string during a time which is short compared to the period of oscillation of the string a wave is produced which travels along the string. On the other hand, if the force is applied smoothly and acts during a time much longer than the period, no wave is formed but the string is statically deformed. The same holds for a magnetic line of force. A rapidly changing disturbance produces a magneto-hydrodynamic wave but a static or semistatic force produces a constant or slowly varying change in the shape of the line of force. It is likely that the factors producing the general magnetic field vary very slowly compared to the time of travel of magneto-hydrodynamic waves through the earth's core or through the sun. Considering the magnetic lines of force as strings, in this case we have not actually to do with the oscillations of these strings but with the forces stretching the strings.

In this way the production of the general magnetic field and the production of sunspot waves in the sun may go on rather independently, the former being due to slowly varying forces and the latter to forces of pulse character. This problem can be settled only when there is a general theory of the internal circulation taking account also of electro-magnetic effects.

Finally the relation with the magneto-hydrodynamic theory of sunspots should be discussed. In the development of this theory it was supposed that the magnetic field in the interior of the sun was homogeneous inside a sphere R_1 ($\sim 1.8 \cdot 10^{10}$ cm) and a dipole field outside R_1. This model is not very remote from the model we have discussed here. It should be observed that a superimposed tangential field does not affect the theory of sunspots because the velocity $\vec{V}$ of the magneto-hydrodynamic waves is proportional to $\vec{H}$ so that the motion in the meridional plane depends only on the meridional field component.

A loop formation in the solar core must affect the motion of sunspot waves. Whether this has any connection with the sunspot cycle cannot be decided at present.

I wish to thank Mr. Herlofson and Mr. Lundquist for many discussions.

REFERENCES

1. LARMOR, J., 1919: How could a Rotating Body such as the Sun become a Magnet? *Rep. of the 87th meeting of the British Ass. for the advancement of science.*
2. ALFVÉN, H., 1950: Cosmical Electrodynamics. Chapters 4—5. Oxford University Press.
3. LUNDQUIST, S., 1949: Experimental Demonstration of Magneto-hydrodynamic Waves. *Nature* **164**, p. 145.
— 1950: Experimental Investigations of Magneto-hydrodynamic Waves. *Physical Review* (in the press).
4. ALFVÉN, H., RICHTMYER, R. D., TELLER, E., 1949: On the Origin of Cosmic Rays. *Physical Review* **75**, 5, p. 892.
FERMI, E., 1949: On the Origin of the Cosmic Radiation. *Physical Review* **75**, 8, p. 1169.
RICHTMYER, R. D., TELLER, E., 1949: On the Origin of Cosmic Rays. *Physical Review* **75**, 11, p. 1729.
ALFVÉN, H., 1949: On the Solar Origin of Cosmic Radiation. *Physical Review* **75**, 11, p. 1732.
— 1950: On the Solar Origin of Cosmic Radiation II. *Physical Review* (in the press).
5. COWLING, T. G., 1934: The Magnetic Field of Sunspots. *Mon. Not. R. Astr. Soc.* **94**, p. 39.
6. ELSASSER, W. M., 1946: Induction Effects in Terrestrial Magnetism. *Physical Review* **69**, p. 106.
— 1946: Induction Effects in Terrestrial Magnetism II. *Physical Review* **70**, p. 202.
— 1947: Induction Effects in Terrestrial Magnetism III. *Physical Review* **72**, p. 821.
7. BULLARD, E. C., 1949: The Magnetic Field within the Earth. *Proc. of the Royal Society*, **A 197**, p. 433.
8. FERRARO, V. C. A., 1937: The Non-uniform Rotation of the Sun and its Magnetic Field. *Mon. Not. R. Astr. Soc.* **97**, p. 458.