THE EARTH'S DYNAMO

H.K. MOFFATT
Department of Applied Mathematics and Theoretical Physics
University of Cambridge
Silver Street
CAMBRIDGE CB3 9EW,
United Kingdom

ABSTRACT. Some key landmarks in the development of dynamo theory are indicated. Kinematic dynamo theory applied to the Earth favours magnetic modes which do not satisfy the condition of magnetostrophic force balance (the Taylor constraint). Current efforts to address and overcome this difficulty are described.

1. Introduction

Since this Workshop was conceived in honour of Leon Mestel, I'd like to start by taking you back to the situation in dynamo theory around 1960 when I was privileged to be a student in Leon's graduate class in Cosmical Electrodynamics, and when my interest in the Dynamo Problem was first aroused. Just two years earlier, the first examples of velocity fields in a fluid sphere capable of dynamo action had been found [Herzenberg 1958, Backus 1958] and I well recall the interest that Leon stimulated through his recognition of the importance of these contributions and through his lectures that unravelled their intricacies.

Until the appearance of these papers, despite the considerable achievements of Elsasser (1946 and later papers), Bullard (1949) and particularly Bullard & Gellman (1954), there remained a lurking suspicion that the anti–dynamo theorem of Cowling (1934), who had proved the impossibility of dynamo maintenance of an axisymmetric magnetic field, might in fact be generalisable to *all* magnetic fields. This suspicion had perhaps been fuelled by the (erroneous) comment of Chapman & Bartels(1940) in relation to Larmor's (1921) famous conjecture that magnetic fields might be self–excited by dynamo action: "Cowling, however, has shown that such self–excitation is not possible". Fortunately, Elsasser, and the generations following him, ignored this unwarranted generalisation!

Parker's 'cyclonic events' paper of (1955) (see also Parker 1979) may now be recognized as the paper which, following principles laid out by Elsasser, sowed the seeds of the dynamo theory that was to be developed in subsequent decades. Cowling comments on Parker's theory in his Interscience tract 'Magnetohydrodynamics' which gave such a boost to the subject when published in 1957; but he did not

D. Lynden-Bell (ed.), Cosmical Magnetism, 1–10.

find the arguments convincing, and could only conclude that Parker's suggestion "deserves a good deal of attention". Cowling's discussion of the dynamo problem was inconclusive, reflecting the widespread uncertainty that still persisted in 1957 concerning the viability of the dynamo mechanism.

The papers of Herzenberg and Backus in 1958 were therefore of the greatest importance in demonstrating the possibility of dynamo action with an acceptable degree of mathematical rigour. It may be useful to draw certain parallels between the two papers, which, on the face of it, appear totally dissimilar. Herzenberg considered two spherical rotors of radius a whose centres are separated by vector distance $\mathbf{R}$; the rotors are imbedded in a medium of the same conductivity, and they rotate with steady angular velocities $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2$, the condition

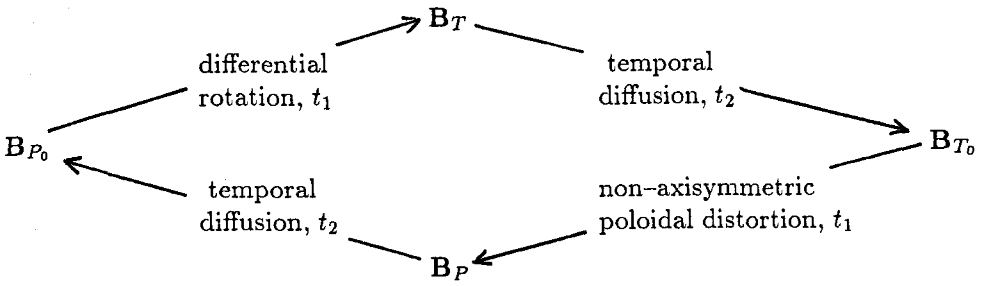
$$\mathbf{R} \cdot (\boldsymbol{\omega}_1 \wedge \boldsymbol{\omega}_2) \neq 0 \tag{1}$$

being satisfied. The dynamo mechanism is represented by the diagram:

$$\begin{tikzcd}
 & \mathbf{B}_T^{(1)} \arrow[dr, "\text{spatial diffusion}"] & \\
\mathbf{B}_P^{(1)} \arrow[ur, "\text{differential rotation}"] & & \mathbf{B}_P^{(2)} \arrow[dl, "\text{differential rotation}"] \\
 & \mathbf{B}_T^{(2)} \arrow[ul, "\text{spatial diffusion}"] &
\end{tikzcd}$$

Differential rotation acting on the poloidal field $\mathbf{B}_P^{(1)}$ in the neighbourhood of sphere 1 generates a toroidal field $\mathbf{B}_T^{(1)}$ which diffuses to the neighbourhood of sphere 2, there providing a poloidal ingredient $\mathbf{B}_P^{(2)}$ (relative to the rotation vector $\boldsymbol{\omega}_2$); and $\mathbf{B}_P^{(1)}$ is 'regenerated' by a reciprocal process. Thus the cycle is closed and self-excitation is possible. Herzenberg assumed $a/R \ll 1$, and developed a perturbation scheme based on this small parameter. The condition (1) may be recognized as a manifestation of 'lack of reflexional symmetry' in the velocity field considered.

Backus, by contrast, relied on temporal rather than spatial diffusion. His dynamo mechanism may be represented in similar diagrammatic form:



We start with the fundamental dipole mode $\mathbf{B}_{P_0}$ and act on this with a short burst (duration t_1) of differential rotation $\mathbf{u}_T$ generating a toroidal field $\mathbf{B}_T$; this is then allowed to decay during a period t_2 of 'stasis' at the end of which the slowest decaying toroidal mode $\mathbf{B}_{T_0}$ dominates. We then act with a burst (again of duration t_1) of poloidal distortion $\mathbf{u}_P$ (necessarily non–axisymmetric to thwart Cowling's theorem) to generate a poloidal field $\mathbf{B}_P$; finally we have a second period of stasis t_2 to cleanse this poloidal field of its higher harmonics, leaving essentially only the fundamental harmonic $\mathbf{B}_{P_0}$, which is reinforced by the whole cycle provided the ingredients $\mathbf{u}_P$ and $\mathbf{u}_T$ are well chosen. Again there is a small parameter $(t_1/t_2 \ll 1)$ on which a perturbation scheme can be developed, and again there is a lack of reflexional symmetry manifest through the 'time–shifted' helicity

$$< \mathbf{u}_T \cdot \mathrm{curl}\ \mathbf{u}_P > \neq 0 \tag{2}$$

(the average being over any sphere $r = $ const.).

Key landmarks in the subsequent development of dynamo theory, all of which have a crucial bearing on all current attempts to develop a dynamically self–consistent theory of the geodynamo, are the following:

(i) Taylor's (1963) treatment of the dynamics of a sphere of fluid in which the dominant force balance (now described as 'magnetostrophic') is that between Coriolis, Lorentz and buoyancy forces;

(ii) Braginsky's (1964 et seq.) development of a scenario for the geodynamo, based on *nearly* axisymmetric buoyancy–driven motions associated with slow solidification of the inner core (see §4 below);

(iii) Steenbeck, Krause & Rädler's (1966 et seq.) development of mean–field electrodynamics (based on scale separation) and their recognition of the fact that random small–scale motions may generate an electromotive force having a component parallel to the mean magnetic field (the α–effect);

(iv) Malkus & Proctor's (1975) attempt to reconcile the field structures that emerge from dynamo action based on the α–effect with the need to satisfy the 'Taylor constraint' associated with the condition of magnetostrophic force balance. This paper heralded the dawn of the 'dynamic' (as opposed to kinematic) era of geodynamo theory, and identified the crucial dynamical problems that must be addressed if a fully satisfactory theory is to emerge.

2. The favoured magnetic modes of kinematic dynamo theory

The simplest mean–field models give an induction equation for the mean field $\mathbf{B}(\mathbf{x}, t)$ of the form

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \wedge (\mathbf{U} \wedge \mathbf{B}) + \nabla \wedge (\alpha \mathbf{B}) + \eta \nabla^2 \mathbf{B} \quad , \tag{3}$$

where $\mathbf{U}$ is the mean velocity (i.e. differential rotation plus meridional circulation) and α is the 'α–effect parameter' closely related to the helicity $< \mathbf{u} \cdot \mathrm{curl}\, \mathbf{u} >$ of small–scale motions; η is the magnetic diffusivity of the medium (augmented possibly by the turbulent diffusivity associated with the same small–scale motions). If $\mathbf{U}, \alpha$ and η are steady in time, then (3) admits solutions of the form

$$\mathbf{B}(\mathbf{x}, t) = \mathrm{Re}\,(\hat{\mathbf{B}}(\mathbf{x}) e^{pt}) \quad , \tag{4}$$

where

$$p\hat{\mathbf{B}} = \nabla \wedge (\mathbf{U} \wedge \hat{\mathbf{B}}) + \nabla \wedge (\alpha \hat{\mathbf{B}}) + \eta \nabla^2 \hat{\mathbf{B}} \quad . \tag{5}$$

In conjunction with the conditions that (i) $\hat{\mathbf{B}}$ is regular at $r = 0$, and (ii) $\hat{\mathbf{B}}$ matches smoothly to a potential field at the core–mantle boundary $r = R_C$, (5) constitutes an eigenvalue problem. Let p_0 be the eigenvalue with largest real part and let $\hat{\mathbf{B}}_0(\mathbf{x})$ be the corresponding eigenfunction. Then, if Re $p_0 > 0$, we have dynamo action and (at least according to this kinematic theory), the field $\hat{\mathbf{B}}_0(\mathbf{x})$ will emerge for $|p_0 t| \gg 1$, starting from arbitrary initial conditions. This field may be described as the 'favoured' magnetic mode. If Im $p_0 = 0$, then $\hat{\mathbf{B}}_0(\mathbf{x})$ is real; if Im $p_0 \neq 0$, then both real and imaginary parts of $\hat{\mathbf{B}}_0(\mathbf{x})$ are relevant via eqn. (4).

Many examples of such dynamos (for different choices of $\mathbf{U}(\mathbf{x})$ are now well documented (see for example Krause & Rädler 1980). However it should be recognized that we are not really free to choose both $\mathbf{U}(\mathbf{x})$ and $\alpha(\mathbf{x})$ independently, since both differential rotation and the α–effect find their origins in the same convective (radial) motions which on the one hand transport angular momentum, and on the other provide the helicity that is responsible for the α–effect. We need therefore to address the dynamical processes that control the structure and influence of such convective motions.

3. Magnetostrophic force balance

The governing dynamical equation is the Navier–Stokes equation including Coriolis, Lorentz and buoyancy forces:

$$\rho \frac{D\mathbf{U}}{Dt} + 2\rho \boldsymbol{\Omega} \wedge \mathbf{U} = -\nabla p + \delta\rho \mathbf{g} + \mathbf{j} \wedge \mathbf{B} + \rho\nu \nabla^2 \mathbf{U} \quad . \tag{6}$$

Here ρ is the mean fluid density and $\delta\rho$ any density perturbation associated with variations of temperature and/or chemical composition; $\boldsymbol{\Omega}$ is the angular velocity of the solid Earth ($\sim 7 \times 10^{-5} \mathrm{s}^{-1}$), p is pressure, $\mathbf{g}$ the local acceleration of gravity, $\mathbf{j} = \mu_0^{-1} \nabla \wedge \mathbf{B}$ the current distribution, and ν the kinematic viscosity of the fluid. Equation (6) is augmented by the incompressibility (or Boussinesq) condition $\nabla \cdot \mathbf{U} = 0$.

Now the velocities in the liquid core are at most of order 1 mm/s on relevant scales of the order $L \sim 100$ km and greater. The ratio of Coriolis to convective acceleration is then

$$\frac{|2\boldsymbol{\Omega} \wedge \mathbf{U}|}{|D\mathbf{U}/Dt|} \sim \frac{2\Omega L}{U} \gtrsim 10^4 \quad , \tag{7}$$

so that the convective acceleration is completely negligible (this is the case even down to scales of the order of 100m). Moreover the viscous term is even smaller, and viscous effects are believed to be important only in very thin boundary layers and/or internal shear layers. In the bulk of the flow, equation (6) is then approximated by the equation of magnetostrophic balance

$$2\rho \boldsymbol{\Omega} \wedge \mathbf{U} = -\nabla p + \delta\rho \mathbf{g} + \mathbf{j} \wedge \mathbf{B} \quad , \tag{8}$$

i.e. a quasi–static balance between Coriolis, buoyancy and Lorentz forces.

Consider now the ϕ–component of (8) in cylindrical polar coordinates (s, ϕ, z). Since $(\boldsymbol{\Omega} \wedge \mathbf{U})_\phi = \Omega U_s$, this is

$$2\rho\Omega U_s = -s^{-1} \partial p / \partial \phi + (\mathbf{j} \wedge \mathbf{B})_\phi \quad . \tag{9}$$

Integrating this over (the fluid part of) the surface of a cylinder $C(s)$ of radius s and generators parallel to $\boldsymbol{\Omega}$, and noting that $\int_C U_s dA = 0$, it follows that

$$\mathcal{T}(s) = \int_{C(s)} (\mathbf{j} \wedge \mathbf{B})_\phi dA = 0 \quad . \tag{10}$$

This is the Taylor (1963) constraint, which any field participating in the dynamic condition of magnetostrophic balance must satisfy.

The potential conflict between kinematic and dynamic theory is now apparent: the favoured mode $\hat{\mathbf{B}}_0(\mathbf{x})$ of kinematic theory, together with its associated current distribution $\hat{\mathbf{j}}_0(\mathbf{x}) = \mu_0^{-1}\nabla \wedge \hat{\mathbf{B}}_0(\mathbf{x})$, will not in general satisfy the Taylor constraint (10). Only if $\mathbf{U}(\mathbf{x})$ and $\alpha(\mathbf{x})$ are 'tuned' in an appropriate way will the favoured kinematic mode have this property. The most commonly adopted view is that a 'geostrophic' component $V(s)\mathbf{e}_\phi$ of $\mathbf{U}$ develops and adjusts itself so that the favoured mode of (3) is just such as to satisfy $\mathcal{T}(s) \equiv 0$. This mechanism must however have implications for the helicity distribution and so also for the function $\alpha(\mathbf{x})$ in (3). The need to develop a dynamic theory for the α–effect (rather than a prescriptive theory based partly on guess–work) forces us to focus on small–scale dynamical processes (down to scales of the order of a few kilometres) which find their origin in buoyancy–induced instabilities in the 'mushy zone' near the inner–core boundary (ICB). This point of view finds its genesis in the scenario for the geodynamo first developed by Braginsky (1964) which we now describe.

4. Braginsky's (1964) scenario for the geodynamo

This scenario is based on the supposition that the solid inner core of the Earth is slowly growing as a result of heat loss from the core and consequential solidification of the liquid from the centre outwards. The present radius R_I of the inner core is $R_I \sim 1200$ km, and if we suppose that the growth has been more or less uniform over the whole lifetime of the Earth, then the rate of growth is

$$\dot{R}_I \sim 10^{-11}\mathrm{m/s} \quad . \tag{11}$$

The outer (liquid) core consists of an alloy, predominantly iron plus an admixture of lighter elements. When this alloy is cooled below its freezing point, the iron preferentially solidifies leaving an excess of the lighter elements in the mushy zone – a layer of thickness estimated to be of order 1 km in which the solidification process is located. The density ρ_s of the solid is thus greater than the density ρ of the overlying liquid, the usual estimate (based on seismological inference) being

$$\frac{\rho_s - \rho}{\rho} \sim 0{\cdot}05 \quad . \tag{12}$$

The downward mass flux is provided by the buoyancy flux in the liquid; thus

$$f\, w\, \delta\rho \sim (\rho_s - \rho)\dot{R}_I \quad , \tag{13}$$

where w is the typical velocity of rise of an element with density defect $\delta\rho$, and $f(0 < f < 1)$ is the fraction of the surface of a sphere just outside the ICB occupied by such rising elements.

A second relation is given by the condition of magnetostrophic balance applied to rising blobs (which may be assumed to be of scale ~ 1 km or greater). Buoyancy and Coriolis forces are comparable provided

$$2\rho\Omega w \sim \delta\rho\, g \quad . \tag{14}$$

Thus, combining (13) and (14), we obtain the estimates

$$\delta\rho/\rho \sim \left(\frac{0{\cdot}05\Omega\dot{R}_I}{fg}\right)^{1/2} \sim 10^{-8} \quad , \tag{15}$$

$$\text{and } w \sim \left(\frac{0{\cdot}05g\dot{R}_I}{f\Omega}\right)^{1/2} \sim 6\times 10^{-4}\text{m/s} \quad . \tag{16}$$

Here, we have used $g \sim 3$ m/s^2 (the value near the ICB) and a plausible estimate $f \sim 0{\cdot}1$. The results (15) and (16) were obtained (although without the f–factor) by Moffatt (1989) and have since been exploited in a detailed treatment of the dynamics of individual rising blobs by Moffatt & Loper (1993).

These rising blobs acquire helicity through the influence of the Coriolis force; at the same time they transport angular momentum so that, over a suitably long time–scale, a field of differential rotation $w(r,\theta)$ is established in the liquid core. The present challenge is to combine the microdynamics of rising blobs, the macrodynamics of large–scale differential rotation and meridional circulation, and the dynamo equation (3) in a self–consistent manner. The separate ingredients are now in place, but the blending of these ingredients has not yet been satisfactorily achieved.

5. Estimate of the core field

The poloidal field at the core-mantle boundary (CMB) can be estimated by simply extrapolating downwards the geopotential field at the surface (assuming that the current source is confined to the core). The dipole ingredient at the CMB is thus a factor of about 8 times stronger than at the surface, i.e. about 3G (and field harmonics of order n are a factor $\sim 2^{n+2}$ stronger than at the surface). The radial component of the field at the CMB has been charted out by Bloxham et al (1989) at various epochs since magnetic records first accumulated; these charts show a number of bipolar loops (analogous to the field of sunspot pairs) of which one below the vicinity of Madagascar is the most prominent. This loop has drifted in a westward direction over the last hundred years, suggesting that it is indeed associated with an upwelling with a similar westward drift, in the manner of Bullard (1949).

The toroidal field $\mathbf{B}_T$ in the core may be estimated again through use of the equation of magnetostrophic balance, on the assumption that the dynamo mechanism intensifies the field until the Lorentz force is of the same order of magnitude as the buoyancy and Coriolis forces. At the macroscopic level, the radial component of $\mathbf{j} \times \mathbf{B}$ of order $\mu_0^{-1} L^{-1} B_T^2$ so that this condition becomes

$$\mu_0^{-1} L^{-1} B_T^2 \sim 2\rho\Omega U \quad . \tag{17}$$

Here L is the macro–scale, say 2000 km (the depth of the liquid core) and $U \sim w \sim 6 \times 10^{-4}$ m/s. Hence we obtain

$$B_T \sim (2\mu_0 \rho L \Omega U)^{\frac{1}{2}} \sim 300\ G \quad . \tag{18}$$

Alternatively, at the blob micro–scale, the Lorentz force intervenes when

$$|\mathbf{j} \times \mathbf{B}| \sim \sigma B_T^2 w \sim 2\rho |\mathbf{\Omega} \times \mathbf{u}| \tag{19}$$

and this gives the smaller estimate

$$B_T \sim (2\rho\Omega\mu_0\eta)^{\frac{1}{2}} \sim 20\ G \quad . \tag{20}$$

There is thus a major uncertainty in estimates of B_T reflecting uncertainty in the relative importance of macro– and micro–dynamics.

6. Magnetostrophic microdynamics

The problem of a single buoyant blob or plume rising under magnetostrophic conditions (in which buoyancy, Coriolis and Lorentz fores are all of the same order of magnitude) appears to play a crucial role in these considerations. The magnetic Reynolds number associated with a blob of mean radius a is

$$R_m = \frac{aw}{\eta} \lesssim 20 \quad \text{if} \quad a \lesssim 100 \text{ km} \quad , \tag{21}$$

so that we may start from the helpful standpoint of a low-R_m theory in which

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{b} \quad , \quad |\mathbf{b}| = O(R_m)|\mathbf{B}_0| \quad . \tag{22}$$

At lowest order in R_m, the equations relating the buoyancy field $\theta(= \Delta\rho/\rho)$, the magnetic field perturbation $\mathbf{b}$ and velocity $\mathbf{u}$ are then quasi–static and linear:

$$\left.\begin{aligned} 2\mathbf{\Omega} \times \mathbf{u} &= -\nabla P - \mathbf{g}\theta + (\mu_0\rho)^{-1} \mathbf{B}_T \cdot \nabla \mathbf{b} \quad , \\ 0 &= \mathbf{B}_T \cdot \nabla \mathbf{u} + \eta \nabla^2 \mathbf{b} \quad , \\ \nabla \cdot \mathbf{u} &= \nabla \cdot \mathbf{b} = 0, \end{aligned}\right\} \tag{23}$$

so that, for given $\theta(\mathbf{x})$, $\mathbf{u}(\mathbf{x})$ and $\mathbf{b}(\mathbf{x})$ may be obtained by Fourier transform techniques.

Nonlinearity and time–dependence enter only through the advection–diffusion equation for θ:

$$\frac{D\theta}{Dt} \equiv \frac{\partial\theta}{\partial t} + \mathbf{u}\cdot\nabla\theta = \kappa\nabla^2\theta \quad . \tag{24}$$

Here κ is the diffusivity of the field θ, i.e. of the lighter elements relative to the background iron alloy of the liquid core. The molecular contribution to κ is very small; the Reynolds number of rise is however

$$Re = (\eta/\nu)R_m \quad , \tag{26}$$

and this may be of order 10^6 or greater if $\eta/\nu \sim 10^6$; we should then expect some turbulent entrainment on scales $\ll a$, although such turbulence may be strongly inhibited by the ambient magnetic field. Here again is an area of continuing major uncertainty: the role of turbulence on scales down to the order of metres in the liquid core. If this turbulence is significant, then blobs may mix quite thoroughly with their environment and lose their identity as they rise; if, on the other hand, turbulent mixing is weak, they will retain their identity until they reach the CMB at which they will cause local eruptions of field (like the sub–Madagascar eruption) and spread out forming a stable layer.

A primary objective of this theory is to calculate the mean helicity $\mathcal{H}(r,\theta)$ associated with a random field of rising blobs, and the associated α–effect $\alpha(r,\theta)$. Current investigations indicate that it is necessary to go to second order in R_m ~~to~~ in order to obtain the first non–zero contributions to both $\mathcal{H}$ and α, and that at this level, the value of α obtained is amply sufficient to drive the geodynamo (in conjunction with differential rotation $\omega(r,\theta)$). The problems that present themselves in this context are extremely challenging, and significant progress may be expected over the next few years.

References

Backus, G.E. (1958) *Ann. Phys.* **4**, 372–447

Bloxham, J., Gubbins, D. & Jackson, A. (1989) *Phil. Trans. Roy. Soc. Lond.* A **329**, 415–502

Braginsky, S.I. (1964) *Geomag. Aeron.* **4**, 698–711

Bullard, E.C. (1949) *Proc. Roy. Soc.* A **197**, 433–453

Bullard, E.C. & Gellman, H. (1954) *Phil. Trans. Roy. Soc.* A **247**, 213–278

Chapman, S. & Bartels, J. (1940) Geomagnetism. Oxford, Clarendon Press

Cowling, T.G. (1934) *Mon. Not. Roy. Astr. Soc.* **94**, 39–48