Primordial Magnetic Fields, Right Electrons, and the Abelian Anomaly

M. Joyce$^{1,2}$ and M. Shaposhnikov$^1$

$^1$Theory Division, CERN, 1211 Geneve 23, Switzerland
$^2$School of Mathematics, Trinity College, Dublin 2, Ireland

(February 1997)

Abstract

In the standard model there are charges with abelian anomaly only (e.g. right-handed electron number) which are effectively conserved in the early universe until some time shortly before the electroweak scale. A state at finite chemical potential of such a charge, possibly arising due to asymmetries produced at the GUT scale, is unstable to the generation of hypercharge magnetic field. We argue that quite large magnetic fields ($\sim 10^{23}$ gauss at $T \sim 3$ TeV with typical inhomogeneity scale up to $\sim \frac{10^6}{T}$) can be generated. These fields may be of cosmological interest, potentially acting as seeds for amplification to larger scale magnetic fields through non-linear mechanisms.
It is usually assumed that the early Universe at temperatures above the electroweak scale and below, say, $10^{12} - 10^{16}$ GeV (depending on the model of inflation) consists of (almost) equilibrium primordial plasma of elementary particles, where any long ranged fields, such as the magnetic one, are absent. There are a few exceptions, though. The old problem of generating galactic magnetic fields may require the presence of primordial seed magnetic fields, amplified then by a galactic dynamo mechanism (see, e.g., [1]). The creation of long ranged magnetic fields requires the breaking of the conformal invariance in the coupling of the electromagnetic field to gravity [2], and a number of mechanisms based on different ideas about this breaking have been proposed so far [2,3].

In this paper we are going to argue that there may be a relation between several, apparently completely different, phenomena: (i) Smallness of the electron Yukawa coupling constant, (ii) possible lepton asymmetry of the Universe, and (iii) magnetic fields in the early Universe.

In short, the logic goes as follows. There are three exact conservation laws in the standard electroweak theory. They can be written as $N_i = L_i - \frac{1}{3} B$, where $L_i$ is the lepton number of $i$th generation and $B$ is the baryon number. The fourth possible combination, $B + \sum_i L_i$ is not conserved because of electroweak anomalous processes, which are in thermal equilibrium in the range $100 \text{ GeV} < T < 10^{12} \text{ GeV}$ [4]. Now, if $f_R = 0$, where $f_R$ is the right electron Yukawa coupling constant, then the electroweak theory on the classical level shows up a higher symmetry, associated with the chiral rotation of the right electron field. For the small value of the Yukawa coupling ($f_R \sim 10^{-6}$ in the MSM) this symmetry has an approximate character. At temperatures higher than $T_R \simeq 3 \text{ TeV}$ the perturbative processes with the right electron chirality flip are slower than the rate of the Universe expansion [5]. (In alternative cosmologies such as that considered in [6] in which the expansion rate before nucleosynthesis is different to the standard one, $T_R$ can be well below the electroweak scale.) Thus, this symmetry may be considered as an exact one on the classical level at $T > T_R$. (The importance of this symmetry for the consideration of the wash-out of the GUT baryon asymmetry by anomalous electroweak $B$ and $L$ non-conserving reactions was realized in refs.
Suppose now that an excess of right electrons over positrons was created by some means at $T > T_R$ (say, by the GUT mechanism for baryogenesis, the exact source of this asymmetry is not essential in what follows). Recall that right electron number is anomalous in the MSM:

$$\partial_\mu j^\mu_R = -\frac{g'^2 y_R^2}{64 \pi^2} f_{\mu\nu} \tilde{f}^{\mu\nu},$$

where $f(\tilde{f})$ are the $U_Y(1)$ hypercharge field strengths (and their duals) respectively, $g'$ is the associated gauge coupling, $y_R = -2$ is the hypercharge of right electron. Thus, the number of the right electrons $N_R$ can be changed by the change of the Chern-Simons number of the hypercharge field configuration, $\Delta N_R = \frac{1}{2} y_R^2 \Delta N_{cs}$ with

$$N_{cs} = -\frac{g'^2}{32 \pi^2} \int d^3 \vec{x} \epsilon_{ijk} f_{ij} b_k,$$

where $b_k$ is the hypercharge field potential. The energy density of the fermion number “sitting” in fermions is of the order $\mu_R^2 T^2$, and the fermion number density is of the order $\mu_R T^2$ where $\mu_R$ is the chemical potential of right electrons. At the same time, the lepton number which can be absorbed by the hypercharge field is of the order $g'^2 k b^2$, and its energy is of the order $k^2 b^2$, where $k$ is the momentum of the classical hypercharge field and $b$ is its amplitude.

Now, at $b > T/g'^2$ and $k \sim \mu_R T^2 / (g'^2 b^2)$ the gauge field configuration has the same fermion number as the initial one, but smaller energy. Thus, the hot matter with excess of right electrons is unstable against the generation of hypercharge magnetic field, which tends to “eat up” real fermions. (The line of reasoning presented here is similar to consideration of the cold fermionic matter with anomalous charges in [9].) It is important that at temperature $T > T_R$ the electroweak symmetry is “restored”, and that $U(1)$ hypercharge magnetic field is massless at that time. (No term like $m_Y^2 b^2$ is generated in any order of perturbation theory in abelian gauge theory at high temperature [10]; the lattice study in [11] confirmed this expectation for $SU(2) \times U(1)$ EW theory beyound perturbation theory). If the hypercharge magnetic fields survive until the time of the EW phase transition ($T \sim 100$ GeV), then they will give rise to ordinary magnetic fields because of electroweak mixing. In the rest of this

3
paper we present quantitative estimates of the parameters of the (hypercharge) magnetic fields produced.

Let us discuss first the origin and the magnitude of the right electron number asymmetry $\delta_R = e_R/s$, where $s = \frac{2}{45}\pi^2 T^3 N_{\text{eff}}$ is the entropy density, $N_{\text{eff}} = N_b + \frac{7}{8} N_f = 106.75$ is the total effective number of degrees of freedom of the MSM. In principle $\delta_R$ produced by out of equilibrium decay at the GUT scale can be as large as $\sim 10^{-2} - 10^{-4}$ (for a review see, e.g. [12]). This is quite consistent with the magnitude of the final baryon asymmetry $\delta_B$ being that observed since there is no simple general relation between the two numbers. In theories like those discussed in [5,8] with $L$ violating processes at intermediate scales one has $\delta_B \sim \delta_R$, at least in the case that the $L$ violating processes go out of equilibrium before the $e_R$ violating ones come into equilibrium. In [8] the case is considered where the L violation continues for just long enough to reduce the final $\delta_B$ to the observed one from an initially larger value fixed by $\delta_R$. And, in a simple GUT like $SU(5)$ which has no $B - L$ violation, we can have $\delta_B = 0$ at the electroweak scale irrespective of the value of $\delta_R$ during the time it is effectively conserved.

Below we will simply assume the existence of a primordial density of $e_R$, with its chemical potential as a free input parameter, assumed only small enough to be treated perturbatively.

We need to determine first whether the instability against magnetic field generation has time to develop on a time scale shorter than that associated with the conservation of $e_R$ i.e. before $T_R$. To do this we take the following effective Lagrangian for the $U_Y(1)$ gauge field:

$$\mathcal{L} = -\frac{1}{4}\sqrt{-g}f_{\mu\nu}f^{\mu\nu} + \mu_0 \epsilon_{ijk} f_{ij} b_k. \quad (3)$$

This is simply the covariant generalization to curved space time of the Maxwell action plus a term obtained by integrating out the fermions at finite chemical potential [13]. This latter term simply describes how winding the gauge fields to give Chern-Simons number changes the energy of the system because it changes the number of fermions as described by the anomaly equation. We fix the constant $\mu_0$ in the FRW universe (with scale-factor $a$) by relating it to the appropriate linear combination of chemical potentials at $a_o = 1$ and
temperature $T_o$:

\[ \mu_o \equiv \mu a \equiv \delta T_o = \frac{g^2}{4\pi^2} \mu R a; \quad (4) \]

\[ \mu_R = \frac{2}{45} \pi^2 N_{eff} \left[ \frac{783}{88} \delta R - \frac{201}{88} \delta_1 + \frac{15}{22} (\delta_2 + \delta_3) \right] T. \]

Here $\delta_i$ are the asymmetries in the conserved numbers $N_i$. $\mu$, which scales as $1/a$, is defined since it is proportional to the usual chemical potential. The numerical relation given between this and $\delta_i$, $\delta_R$ is that obtained by doing a local thermal equilibrium calculation in the EW theory with three fermionic generations and one scalar doublet, with the conserved charges assumed to be $N_i$ and $e_R$.

This model makes two obvious assumptions which we will relax below:

(i) The induced electric and magnetic fields cause perturbations in the fermionic particles in ways which are not modelled by chemical potential. In particular currents are induced which are damped by scattering in the plasma. That this is very significant for the dynamics of magnetic fields in the early universe is well known: the universe is a very good conductor and tends to freeze magnetic fields.

(ii) Once the instability develops, the Chern-Simons number grows and begins to “eat-up” the fermion number as described by the anomaly. The calculation therefore breaks down when the change to $\mu$ becomes significant. We will estimate below when this happens and discuss how the problem should be treated in more general.

Taking the gauge $b_0 = 0$ the equations of motion are, in conformal time coordinates

\[ \tau = \int a^{-1}(t) dt, \]

\[ \partial_\tau \partial_i b_i = 0, \quad (5) \]

\[ \partial^2 b_i - \partial_j (\partial_j b_i - \partial_i b_j) - 2\mu a e^{ijk} (\partial_j b_k - \partial_k b_j) = 0. \quad (6) \]

The equations are the same as in Minkowski space except for the scale factor in the anomaly term, which breaks the conformal invariance.

Taking a Fourier mode decomposition of these linear equations $\vec{b}(\vec{x}) = \int d^3 k b(k) e^{-i\vec{k}\cdot\vec{x}}$ and taking $\vec{k} = k\vec{e}_3$ it is easy to show that
\[-\partial^2_\tau b_\pm(k) = k(k \mp 4\mu a)b_\pm(k), \quad (7)\]
\[\vec{b}(k, \tau)e^{-ikz} + \text{c.c.} = b_+(k)(\vec{e}_1 \sin kz + \vec{e}_2 \cos kz) + b_-(k)(\vec{e}_1 \cos kz + \vec{e}_2 \sin kz). \quad (8)\]

where, for simplicity, we have omitted a constant relative phase between the two modes. For \( k < 4\mu a \) the mode \( b_+ \) is unstable, growing with exponent \( \sqrt{k(4\mu a - k)\tau} \). It has the property \( \vec{B} = k\vec{b} \) (where \( \vec{B} = \frac{1}{2} \epsilon^{ijk} f_{jk} \)) so that, as discussed above, \( f^2 \sim k^2 b^2 \) and \( n_{cs} \sim \vec{b} \cdot \vec{B} \sim k^2 b^2 \), leading clearly to a net negative mass squared for sufficiently small \( k \). The maximally growing mode is \( k = 2\mu a \) and the instability which this represents becomes significant when \( 2\mu a \tau \sim 1 \) i.e. when the unstable mode \( k \sim 2\mu a \) comes inside the horizon. Using \( a\tau = H^{-1} = \frac{M_0}{T^2} \) (where \( M_0 = M_{pl}/1.66\sqrt{N_{eff}} \simeq 2 \times 10^{18} \text{ GeV} \)) we find that the temperature at which the unstable mode begins to grow is \( T \sim \delta M_0 \). Thus, for right-handed electron number which is approximately conserved by perturbative processes until \( T_R \), the instability appears to be relevant for asymmetries as small as \( \delta \sim 10^{-15} \).

We will limit ourselves here to improving this first approximation by including the dynamical back reaction of the currents with a term \( \sqrt{-g} j_\mu b^\mu \) in the action, taking \( j^\mu = (0, \sigma a \partial_\tau b^i) \), where \( \sigma \) is the sum of the conductivities of the particles in the plasma. This results in the addition of a damping term \( \sigma a \partial_\tau b_i \) to (6). This leaves the modes unchanged but alters the exponents, so that the growing mode still exists for \( k < 4\mu a \) but with the smaller exponent \( \frac{k(4\mu a - k)}{\sigma a} \). We assume here that \( \sigma a \gg \sqrt{|k - 4\mu a|} \) since \( \sigma \approx 3 \times 10^2 T \) [14]. The growing instability now starts to develop at \( T \sim T_g \) where we define \( T_g \) to be
\[8(\frac{\mu(\tau)}{T})^2 \frac{1}{\sigma/T} \frac{M_0}{T_g} = 1. \quad (9)\]

Requiring that \( T_g \) be greater than \( T_R \) gives a minimum value \( \delta_{\text{crit}} = 2 \times 10^{-7} \) for \( \delta \) in order for the instability to be relevant.

Thus the effect of the conductivity of the plasma is to freeze modes which would otherwise grow. Instead of evolving when it enters the horizon, an unstable mode must wait until its wavelength is \( \sim \delta \) times the horizon size. The unstable modes associated with right
asymmetries less than $\delta_{\text{crit}}$ are frozen on the time-scale characterizing conservation of the charge, and no non-trivial dynamics results from the presence of such a chemical potential. The bounds derived on lepton number violating processes in [5,8] treating the right-handed electron number as a simple conserved particle number are therefore unaffected, since they typically involve values of $\delta_R$ of order of the observed baryon asymmetry $\sim 10^{-10}T$. An exception to this is the case considered in [8], where the bound on such operators is further weakened by allowing for the possibility that the $L$ violating processes remain in equilibrium for a time after $T_R$, just long enough to reduce an initially large $\delta_B \sim \delta_R$ to the observed value. For example, the upper bound on a Majorana neutrino mass can be increased from $\sim 5keV$ to $\sim 15keV$ by increasing the value of the primordial asymmetry for $e_R$ from $\sim 10^{-7}$ to $\sim 10^{-2}$.

To determine the evolution of the instability in the case that the chemical potential is large enough for it to develop before $T_R$ we first simply evolve the system in the linear (i.e. assumed constant chemical potential) regime, and self-consistently determine when this approximation breaks down. To do this it is sufficient to calculate the Chern-Simons number, both to determine the magnitude of the magnetic field and (by comparing the Chern-Simons number to the fermionic number) to determine when the linear approximation breaks down. The Chern-Simons number density (per co-moving volume) can be written

$$n_{cs}(\tau) = -\frac{g^2}{32\pi^2} \langle \epsilon_{ijk} \tilde{f}_{jk}(\tau) b_i(\tau) \rangle$$

$$\approx -\frac{g^2}{64\pi^4} \int_0^{4\mu a} dke^2 \frac{k^{(4\mu a-k)\tau}}{\sigma a} k^2 f(k),$$

where we have neglected all but the growing mode. We have also taken $\langle b_i(\vec{k}, \tau) b_j^*(\vec{l}, \tau) \rangle |_{\tau=0} = \delta^3(\vec{k} - \vec{l}) \delta_{ij} < b^2(k) >_0$, assuming translational and rotational invariance of the initial perturbations. We also assumed that the perturbations are thermal in origin, with the appropriate normalizations, $< b^2(k) >_0 = \frac{1}{2k(2\pi)^3} f(k)$ where $f(k) = (e^{\frac{k}{T_0}} - 1)^{-1}$ is the bosonic distribution function and $T_0$ the temperature at which we define $a_0 = 1$. Defining $\epsilon$ by $\frac{1}{2} g_R^2 n_{cs} \equiv \epsilon \Delta e_R a^3$, where $\Delta e_R = \frac{88}{783} \mu_R T^3$ i.e. the difference between the right electron density in the initial state and the $\mu_R = 0$ state, the linear approximation breaks down
when $\epsilon \sim 1$. Evaluating the integral in (10) we find $\epsilon \approx 2 \times 10^{-6} \delta \frac{1}{\sqrt{\alpha}} e^\alpha$ where $\alpha = T_g/T$ with $T_g$ as in (9). Thus for quite a few expansion times after the mode starts growing at temperature $T_g$ we have $\epsilon < 1$ and the linear approximation is valid. The corresponding physical magnetic field $B_{phy}$ can be estimated by putting $|n_{cs}| \approx g' \frac{1}{16\pi} k^2 \approx \frac{1}{2} \epsilon \Delta e a^3$ and using $kb = a^2 B_{phy}$, where $k \sim 2\mu a$ (i.e. assuming the maximal growing mode to dominate).

Putting in the numbers this gives a physical magnetic field of strength $B_{phy} \approx 3 \times 10^2 \sqrt{\epsilon \delta T^2}$ at a physical length scale $(\frac{k}{a})^{-1} \sim \frac{1}{23T}$. For $\delta \sim 10^{-6}$ we are in the linear regime until $T_R \sim 3$ TeV and at that point therefore have a magnetic field of strength $B_{phy} \sim 10^{23}$ gauss (1 GeV$^2 \equiv 10^{20}$ gauss) at a length scale of $\sim 10^6/T$ (compared to a horizon scale of $\sim 10^{16}/T$).

These are very large fields: if frozen in until the electroweak scale they are $\sim 10^{20}$ gauss which compares, for example, with a field generated at bubbles walls of $10^{-2}$ gauss which were suggested as seeds for amplification through turbulence in [15].

What happens in the case that the linear approximation breaks down before $T_R$? In [9] this question is analyzed in a simple model in Minkowski space. By minimizing the total energy of the fermions (calculated now in the background of the unstable mode) and the gauge fields, it was found that the system has infra-red unstable behaviour. The reason for this can again be traced back to the fact that the energy in the gauge fields scales as $k^2$ and the Chern-Simons number as $k$: The fermionic energy is minimized by putting $n_{cs} \sim k b^2 \sim e_R$; but the gauge field energy then is $\sim k e_R$, which can always be reduced by going to smaller $k$. How do we fix this cut-off which determines the final state? In an expanding universe it is natural to ask whether the cut-off might be the horizon scale. To see whether there is any indication that a transfer to larger length scales may occur once the linear approximation breaks down, we naively model the back reaction of the growth of the condensate by making $\mu$ a function of time, calculated from the depleted number of ‘real’ fermions left at each moment (a marginal approximation since $\mu$ and the fields change on comparable time-scales). The growth factor $A(k, \tau)$ for a mode $k$ is

$$A(k, \tau) = \exp \left( \int_0^\tau \frac{k(4\mu(\tau)a(\tau) - k)}{\sigma a} \right),$$

(11)
where
\[
\mu(\tau) \approx \mu(0) - \frac{783}{176(2\pi)^4} \left(\frac{g^2}{4\pi}\right)^2 T \int_0^{4\mu(0)a} \frac{k}{a} A^2(k, \tau) d\left(\frac{k}{a}\right).
\]

When the linear approximation breaks down $\mu$ starts significantly decreasing and the growth of modes is shifted to longer wavelengths. It appears that this is a procedure which should continue, growth of any mode eventually turning itself off and increasing the growth coefficient of modes at larger scales. However we can see that there is a minimum value $\mu_m$ of $\mu$ which can be reached at any given time $\tau$ (and, correspondingly, a maximum physical scale for the sourced fields): It is simply that given by (9), solved for $\mu$ with $T_g$ replaced by the temperature $T(\tau)$ i.e it is just the minimal chemical potential required to drive a growing mode at that time in the linear approximation. If the chemical potential reaches this value all modes which can evolve on the relevant time-scale are damped rather than amplified, thus increasing $\mu$ again. We would thus expect that at time $\tau$ the Chern-Simons number will be such that the initial chemical potential is cancelled (i.e. $\epsilon \sim 1$) but stored predominantly in a mode of $k \sim \text{few} \times (\mu_m a)$. This corresponds to a field on a length scale larger by $\sim \frac{\mu(0)}{\mu_m}$, but smaller in magnitude by $\sim \sqrt{\frac{\mu_m}{\mu(0)}}$, than the field present at the point where the linear approximation breaks down.

How does the evolution of the fields continue after $T_R$, and what are the fate and role of these magnetic fields? Once we reach $T_R$ the remaining chemical potential rapidly disappears as the right-handed electrons equilibrate. In the simple linear analysis (taking $\mu = 0$) all modes then simply decay because of the plasma conductivity, and the magnetic fields have plenty of time to damp away on the relevant length scales before $T_{ew}$. If this is the case, the only possible cosmological consequence from the existence of those fields we could see is the modification of the amplitudes of the different particle physics processes, such as enhancement of the heavy neutrino decay [16].

However, if one considers the problem more fully it is not at all clear that this is what occurs. With the full set of MHD equations (which include the velocity of the fluid which we have neglected) there is a transition to a turbulent regime when the magnetic Reynold’s
number $R = \sigma L v$ is large [17]. Since we have here $\sigma \sim 10^2 T$ and magnetic fields which begin to grow on length scales $L$ up to $\sim 10^6 / T$ we expect to enter the turbulent regime if there are bulk velocities of greater than $\sim 10^{-8}$. A recent study of this phenomenon [17] suggests that the effect of this turbulence is to transfer the magnetic energy to larger length scales, thus evading the Silk argument [18]. If true, the fields generated by the mechanism under discussion may play a role of the seed galactic magnetic fields. It is worth mentioning also the peculiar structure of the magnetic fields appearing because of the abelian anomaly. The Chern-Simons wave (8) has non-zero value of $\vec{B} \cdot \vec{\partial} \times \vec{B}$ and thus breaks parity. Could it be that the rotation of galaxies are related to this?

It remains to be seen if the magnetic fields generated from right electron asymmetry and abelian anomaly have observable consequences. The clarification of this issue requires the solution of the entire set of MHD equations with addition of anomaly terms.

We are grateful to M. Giovannini, A. Kusenko and L. McLerran for interesting discussions.
REFERENCES


