Electro-Magnetic Earthquake Bursts
and Critical Rupture of Peroxy Bond Networks in Rocks

F. Freund$^1$ and D. Sornette$^{2,3}$

$^1$Department of Physics, San Jose State University San Jose, CA 95192-0106 and NASA Ames Research Center, Mail Stop 242-4, 515 N. Whisman Road, CA 94035-1000 and SETI Institute, 2035 Landing Drive, Mountain View, CA 94043, USA (e-mail: ffreund@mail.arc.nasa.gov)

$^2$Chair of Entrepreneurial Risks, D-MTEC. ETH Zurich (Swiss Federal Institute of Technology Zurich), CH-8032 Zurich Switzerland (email: dsornette@ethz.ch)

$^3$Laboratoire de Physique de la Matière Condensée, CNRS UMR6622 and Université des Sciences, Parc Valrose 06108 Nice Cedex 2, France

(Dated: May 9, 2006)

We propose a mechanism for the low frequency electromagnetic emissions and other electromagnetic and electric phenomena which have been associated with earthquakes. The mechanism combines the critical earthquake concept and the concept of crust acting as a charging electric battery under increasing stress. The electric charges are released by activation of dormant charge carriers in the oxygen anion sublattice, called peroxy bonds or positive hole pairs (PHP), where a PHP represents an $O_3^X/Y_2O_3^Y$ with $X,Y = Si^{4+}, Al^{3+},...$, i.e. $O^-$ in a matrix of $O^{2-}$ of silicates.

We propose that PHP are activated by plastic deformations during the slow cooperative build-up of stress and the increasingly correlated damage culminating in a large “critical” earthquake. Recent laboratory experiments indeed show that stressed rocks form electric batteries which can release their charge when a conducting path closes the equivalent electric circuit. We conjecture that the intermittent and erratic occurrences of EM signals are a consequence of the progressive build-up of the battery charges in the Earth crust and their erratic release when crack networks are percolating throughout the stressed rock volumes, providing a conductive pathway for the battery currents to discharge. EM signals are thus expected close to the rupture, either slightly before or after, that is, when percolation is most favored. The proposed mechanism should be relevant for the broader understanding of fractoemissions.

PACS numbers:

INTRODUCTION

There are many reports of electric and electro-magnetic radiations associated with earthquakes. The observations include (NASDA-CON-960003, 1997; Johnston, 1997) low-frequency electromagnetic emissions (Fujinawa et al., 2001; Fujinawa and Takahashi, 1990; Gershenzon and Bambakidis, 2001; Gokhberg et al., 1982; Molchanov and Hayakawa, 1998a, b; Nitsan, 1977; Vershinin et al., 1999; Yoshida et al., 1994; Yoshino and Tomizawa, 1989; Taylor and Purucker, 2002), local magnetic field anomalies (Fujinawa and Takahashi, 1990; Gershenzon and Bambakidis, 2001; Ivanov et al., 1976; Kopytenko et al., 1993; Ma et al., 2003; Zlotnicki and Cornet, 1986), increases in radio-frequency noise (Bianchi and al., 1984; Hayakawa, 1989; Martelli and Smith, 1989; Pulinets et al., 1994), earthquake lights (Derr, 1973; King, 1983; St-Laurent, 2000; Tsukuda, 1997; Yasui, 1973), seismic electric signals (Varotsos, 2005; Debate on VAN, 1996) and so on. The ability of the Earth’s crust in areas of active deformation to transform mechanical energy into electromagnetic radiations may be due to one or several mechanisms. The mechanisms which have been proposed in the literature include piezoelectric effects, streaming potentials, Lenard splashing, plasma generation, electron emission, separating electrification, pseudo-piezoelectric effects, crack surface friction and so on. Here, we examine in some details the mechanism based on the release of electric charges by activation of dormant charge carriers in the oxygen anion sublattice, called peroxy bonds or positive hole pairs (PHP).

We first present a short review of the hypothesis proposed by the first author and his co-authors that electric charges are activated in rocks as stress increases, making the crust an electric battery which can discharge. The discharge currents may lead to electromagnetic (EM) emissions associated with the closing of the electric circuit by the percolation of conducting paths most probably occurring around the time of (before, during and after) large earthquakes. We then briefly review the critical earthquake concept proposed some time ago (Sornette and Sornette, 1990; Sornette and Sammis, 1995; Saleur et al., 1996, Bowman et al., 1998; Huang et al., 1998; Sammis and Sornette, 2002), which views large earthquakes as the culmination of a growing collective organization of damage similar to a critical phase transition (see the recent review (Sornette, 2005) for its description in the context of material rupture). We then combine the two concepts to propose an explanation for why powerful low frequency EM emissions would
become observable under certain conditions but remain unobservable in other cases. Eftaxias et al. have previously found that EM signals reveal characteristic signs of an approaching critical point (Eftaxias et al., 2002; 2004; Kapiris et al., 2004a,b; Contoyiannis et al., 2005). Here, we continue along this line of thought by proposing a specific mechanism for electric charge generation at the basis of EM emissions, which is combined with the theory of a critical organization of the crust before large earthquakes. This allows us to suggest some observational consequences for EM precursors of earthquakes.

INCIPIENT DAMAGE, PEROXY BONDS AND POSITIVES HOLES

Recent work has brought forth evidence of electric signatures associated with mechanical disturbances (Freund, 2000; Freund et al., 2006a,b). These experiments have led to the following understanding that we briefly summarize.

The first step is to recognize that electronic charge carriers exist in igneous and high-grade metamorphic rocks in a dormant, electrically inactive form. These charge carriers are defect electrons in the oxygen anion sublattice, also known as positive holes, and their dormant precursors are positive hole pairs (PHP) (see Figure 1). Among the mechanisms that activate the PHPs and generate the highly mobile positive holes are processes that are expected to occur on a large scale in rock volumes experiencing an increasing stress due to tectonic forces. The two main processes are plastic flow and microfracturing, both that are thought to occur ubiquitously in the earth crust. During plastic deformation that underlies plastic flow, dislocations, which sweep through the crystal structures of the constituent minerals, pass through PHPs and cause them instantly to release positive hole charge carriers. The stress field associated with microfractures can cause additional activation of PHPs within the surrounding microvolume. In a rock experiencing massive compressive stress such as in thrust setting characteristic of subduction zones or in shear zones under global tectonic forces, these processes are predicted to lead to a large-scale activation of positive hole charge carriers that would flow out of the “source volume”. Both, plastic deformation and microfracturing of rocks are expected to mark the on-going preparation of rock deformation towards global failure, i.e. prior to initiation of an earthquake.

Let us now step back for a while to justify the key role that PHP, we believe, play in the signature of large damage processes. Martens et al. (1976) reported early that hydroxyl pairs in a model mineral, MgO, converted to molecular H₂ while oxidizing two regular lattice oxygen anions from the 2- to the 1- state, hence to peroxo:

\[
\text{OH} + \text{OH} \rightleftharpoons \text{H}_2 + \text{O}_2^2.
\]  

This redox conversion was confirmed spectroscopically (Freund and Wengeler, 1982) and through electrical conductivity and dielectric polarization studies (Freund et al., 1993). The same redox conversion was found in olivine which, though nominally anhydrous, contains traces of “water” in form of hydroxyl, O₃XOH, where \(X = \text{Si}^{4+}, \text{Al}^{3+}\), and so on (Freund and Oberheuser, 1986). It has long been recognized that all nominally anhydrous minerals (meaning nearly all minerals in common igneous rocks) contain traces of dissolved “water” (Bell and Rossman, 1992; Martin and Wyss, 1975; Rossman, 1996; Smyth et al., 1991 and revue of Sornette, 1999). What had not been recognized is that XOH pairs in these minerals appear to undergo the same redox reaction as OH⁻ pairs in MgO, giving rise to peroxo links:

\[
\text{O}_3\text{XOH}...\text{HOXO}_3 \rightleftharpoons \text{H}_2 + \text{O}_3\text{X}^+/\text{YO}_3.
\]

This reaction occurs below 400 – 500°C. The geoscience community at large has not yet taken notice of this redox conversion. One of the prime reason seems to be the assumption that oxygen is never in any other oxidation state but 2-. Accepting the possibility of oxygen in the 1- oxidation state in peroxo links opens a road towards understanding earthquake-related electrical phenomena. Figure 1 presents a schematic representation of these reactions corresponding to the conversion of hydroxyl pairs in MgO to peroxo (PHP) plus H₂ and the “tearing apart” of the PHP to generate p-holes. As argued below, we propose that analogous reactions occur for Si-OO-Si in silicates.

Peroxy and positive holes (positive electronic charge carriers) are defect electrons or “holes.” When occurring in the \(O^2\)-sublattice of oxides or silicate minerals, they are called “positive holes” or p-holes for short and designated \(k\). Chemically, a positive hole is an \(O\). While an \(O\) is a radical, two \(O\) can stabilize by pairing and tying an OO bond. An OO bond is a peroxy link, \(O_3X^{\text{OO}}\text{YO}_3\) with \(X, Y = \text{Si}^{4+}, \text{Al}^{3+}\) and so on. As we have said, in physical terms, a peroxy link is a positive hole pair (PHP).

Magmatic and high-grade metamorphic rocks consist primarily or entirely of nominally anhydrous minerals that crystallized or recrystallized in \(H_2O\)-laden environments. Invariably, such minerals incorporate small amounts of
The dissolved water occurs in form of hydroxyls, e.g. as OH or XOH (Bell and Rossman, 1992; Martin and Wyss, 1975; Rossman, 1996; Smyth et al., 1991):

\[(H_2O)_{\text{dissolved}} + (O^2^-)_{\text{structure}} \Leftrightarrow (OH^-)_{\text{structure}} + (OH^-)_{\text{structure}} \]
\[(H_2O)_{\text{dissolved}} + (X/O/Y)_{\text{structure}} \Leftrightarrow (X/OH \text{ HO}/Y)_{\text{structure}} .\]

A major fraction of these hydroxyls is believed to undergo this particular redox conversion by which \(OH^-\) or \(X-\text{OH}\) pairs rearrange so as to reduce two \(H^+\) to \(H_2\) and oxidize two \(O^2^-\) to \(O^-\) which tie a peroxy link:

\[(OH^-)_{\text{structure}} + (OH^-)_{\text{structure}} \Leftrightarrow (H_2)_{\text{structure}} + (O^2^-)_{\text{structure}} \]
\[(X/OH \text{ HO}/Y)_{\text{structure}} \Leftrightarrow (H_2)_{\text{structure}} + (X/OO\text{ }Y)_{\text{structure}} .\]

Even minerals that come from highly reducing crustal and upper mantle environments and contain reduced transition metal cations can acquire peroxy. This may appear paradoxical but peroxy and reduced transition metals are not mutually exclusive. This is because the redox conversion takes place during cooling, at relatively low temperatures, below 400–500°C. At these low temperatures, all major solid state processes are frozen and non-equilibrium conditions prevail (Kathrein et al., 1984; Sornette, 1999).

Basic information about PHP’s and positive hole charge carriers, \(h\), was obtained by studying \(MgO\), theoretically (King and Freund, 1984) and experimentally by dc conductivity, dielectric polarization (Freund et al., 1993), thermal expansion, refractive index measurements etc. (Freund et al., 1994) and fracture experiments (Dickinson et al., 1986). Information about PHP’s in upper mantle olivine was derived from dielectric polarization and fracture mass spectroscopy experiments (Freund and Ho, 1996). Peroxy links in fused \(SiO_2\) were studied theoretically by Edwards and Fowler (1982) and Ricci et al. (2001), and experimentally by electron spin resonance spectroscopy (Friebele et al., 1979) and dielectric polarization (Freund et al., 1993). Fracture experiments with \(MgO\) crystals suggested that the stress fields and the acoustic waves emitted during fracture activate \(h\) charge carriers (Dickinson et al., 1986; Freund et al., 2002). Recent experiments using slow velocity impacts have shown the validity of the concept that PHPs are activated under the effect of time-dependent stresses (Freund, 2002). The experiments used steel ball bearings impacting at 100 m/sec on quartz-free gabbro (\(\approx 80\%\) plagioclase, \(\approx 20\%\) pyroxene) and low-quartz diorite cut to cylindrical cores. Sensors using magnetic pick-up coils, photodiodes, capacitive sensors and contact electrodes clearly demonstrated the appearance of positive carriers that were created by the impact. The conditions under impact resemble those experienced microscopically by the mineral grains in a large rock volume that is slowly compressed with intermittent damage. Though the overall stress change may be small, the strain will tend to discharge “explosively” on the microscopic scale by bursts of dislocation movements and microfractures that open and close on short time scales. These bursts, though numerous, occur distributed over time and space. Another recent experiment by Freund et al. (2006b) shows that, when stress is applied to one end of a block of igneous rocks, two currents flow out of the stressed rock volume. One current is carried by electrons and it flows out through a Cu electrode directly attached to the stressed rock volume. The other current is carried by p-holes, i.e. defect electrons on the oxygen anion sublattice, and it flows out through at least 1 m of unstressed rock. The two currents are coupled via their respective electric fields and fluctuate. Evidence of these two types of currents include the positive sign of the charges accumulating on the rock surface which can be directly measured. In sum, the stressed rock volume acts as a genuine electric battery, which can discharge only when the two poles are connected via a conductor, i.e., when the electric circuit is closed on itself.

Dry rocks are usually good insulators but become p-type semiconductors as a consequence of the break-up of the peroxy bonds. The p-hole charge carriers have not been recognized earlier probably due to the fact that a \(h\) is (chemically speaking) an \(O\) which is highly oxidizing and reacts with any reduced gas and is thus annihilated. All the studies on the electrical conductivity of minerals and rocks noticed the unusually high conductivity in the 400 – 600°C window but assumed it away as a result of surface contamination; indeed, in all these studies, the anomaly was eliminated by annihilating the \(h\) charge carriers by “equilibrating” their minerals and rocks for long time in reduced gas until the conductivity returned to “normal”!

Applying the insight gained from these laboratory experiments (Freund, 2002; Freund et al., 2006b) to the field, where large volume of rocks are subjected to ever increasing stress, has led us to suggest the existence of transient, fluctuating currents of considerable magnitude that could build up in the Earth’s crust prior to major earthquakes. An important problem however is to understand how large scale currents can develop and under what circumstances. For this up-scaling purpose, we propose the critical earthquake model, that we now briefly review before combining it with the “battery” hypothesis.
THE CRITICAL-EARTHQUAKE MODEL

As often in earthquake physics, we face the challenge of up-scaling, i.e., the transition from the laboratory scale to the scale of the Earth crust. In order to explain EM signals which appear at large scales, it seems necessary to identify global signatures of the cooperative damage occurring at small scales. This leads us to proposing the relevance of the critical-earthquake model. We emphasize that a local analysis is inherently incompatible with the critical earthquake hypothesis. The problem is similar to the prediction of incipient failure in materials in the laboratory or in engineering structures (Anifrani et al., 1995; Johansen and Sornette, 2000): only by integrating the information on many damaged elements can one develop efficient diagnostic for failure (Andersen and Sornette, 2005; Sornette and Andersen, 2006), in contrast with the detailed description of individual cracking events which are like the trees hiding the forest. Here, we describe a physical mechanism that converts damage at the micro-scale into electric signals, such that the cooperative and progressive damage occurring over a large spatial region preceding the main shock should translate into large-scale observable electro-magnetic signatures.

We start with a cartoon of the mechanical processes occurring under stress, first on a small scale of individual grains, then on a larger scale within the crust taking into account the variation of essential parameters such as temperature, ductility and electrical conductivity. Figure 2 shows schematically the response of a cube of rock, confined within a larger rock volume and subjected to increasing stress, (a) if the sample is brittle, and (b) if the sample is ductile.

If a brittle sample is subjected to stress beyond its elastic limit, the deformations become nonlinear. On an atomic scale, dislocations move and new dislocations are generated. As the dislocations per unit volume become ever more numerous, they begin to coalesce, leading to microfractures. As the microfractures become ever more numerous, they too coalesce leading to larger cracks. These larger cracks will eventually lead to catastrophic failure.

In a ductile sample, the dislocations anneal as fast as they are generated under stress. The result is plastic flow without microfractures and, hence, without cracks and catastrophic failure. It is important to note, however, that even a ductile material can become brittle and develop fractures, if the stresses reach very high values. The reason is that the brittle-ductile transition is controlled by two time-dependent processes: by the rate at which dislocations are generated and by the rate at which dislocations anneal. If, at high stress rates, more dislocations are generated per unit time and unit volume than can be annealed, dislocations can pile up and coalesce even in a ductile material, leading to microfractures and cracks. This is graphically represented on the right hand side of Figure 2.

In Figure 3a, we apply this small-scale concept to infer the qualitative mechanical behavior of favorable domains for the nucleation and propagation of future earthquakes. We keep in mind the fact that the stress field is probably highly heterogeneous (Shamir et al., 1990; Hickman et al., 2000). Figure 3 represents the effective rheology of a large sample converging to rupture. In the upper left of Figure 3 we show a cross section through two crustal blocks which we assume to have collided at time $t_0$ along the dashed line. The two blocks comprise the cool, brittle upper crust and the increasingly hot, ductile mid- to lower crust. We further assume that, due to tectonic forces, the two blocks are being pushed against each other at a constant speed (constant strain). Also shown within the blocks are two cross-hatched zones: one where the rocks change from brittle to ductile and another one where their conductivity changes from p-type to n-type. Indeed, we include one important electrical parameter: the fact that rocks in the cool upper crust tend to be ever so slightly p-type conducting but turn n-type conducting deeper into the crust when the temperatures reach or exceed 500 – 600°C (Freund, 2003).

In Figure 3b, we show that the strain increases linearly with time. In Figure 3c we plot in a simplified way strain versus stress. In the elastic range, the strain is linear with stress. In the inelastic range, which is of interest here, the strain, i.e. deformation, increases non-linearly, due to the generation and movement of dislocations. Eventually, under the onslaught of continuing deformation at a constant strain rate, stress increases even more rapidly and the system becomes mechanically unstable. In Figure 3d, we plot, again schematically, how the viscosity changes with increasing depth and, hence, with increasing temperature. The deeper we go into the crust, the more the temperature increases. Hence, the rocks become ever more ductile. The brittle to ductile transition is marked by a cross-hatched region. However, the position of the brittle-to-ductile transition zone is not fixed. The reason is, as explained in the context of Figure 2a, a ductile response at low stress levels (slow deformation) may turn into a brittle response when high levels of stress are applied (rapid deformation).

Finally, in Figure 3e, we combine the information derived from Figures 3a-c with the representation of two crustal blocks colliding at a constant speed. In Figure 3f, we plot the deformation volume versus time, i.e. the rock volume that reaches a certain level of deformation as time progresses. We can arbitrarily choose as the level of deformation the on-set of microfracturing as depicted in Figure 3g. Because of the non-linear increase of stress (Figure 3h), combined with the decrease in viscosity (Figure 3i), we know that the deformation volume, which satisfies this condition, will expand outward and downward with increasing stress as indicated by the white lines in the two-block model (Ashby
and Jones, 1980; Frost and Ashby, 1982; Nechad et al., 2005(a,b)). The points \( t_0, t_1 \) and \( t_2 \) on the time axis of Figure 3 mark the initial stress build-up. We assume that the deformation volume reaches the transition zone from brittle to ductile at \( t_3 \) and the transition zone from p-type to n-type behavior at \( t_4 \).

While we expect the brittle-to-ductile transition zone to move downward in the crust when stresses become high and deformations become rapid, the transition from p-type to n-type is controlled by temperature only and independent of the stress level (Freund 2003). Thus, we can envision situations where, with ever increasing stress levels, the transition to a ductile behavior is pushed deeper and deeper into the crust. This would allow the zone of brittle fracturing to extend downward, eventually overlapping with the zone where the rocks become pervasively n-type. At this point, we believe, a crucial electrical connection is made which enables the electrons in the battery, i.e. in the stressed rock volume, to massively flow downward and create a condition where the p-hole currents follow suit by also flowing out massively.

Uncorrelated percolation (Stauffer and Aharony, 1992) provides a starting modeling point valid in the limit of very large disorder (Roux et al., 1988). For realistic systems, long-range correlations transported by the stress field around defects and cracks make the problem much more subtle. Time dependence is expected to be a crucial aspect in the process of correlation building in these processes. As the damage increases, a new “phase” appears, where micro-cracks begin to merge leading to screening and other cooperative effects. This should lead to an overall acceleration of the seismicity at large scales. However, by the very nature of the damage processes, this acceleration is expected to be very intermittent and strongly fluctuating from case to case. With respect to both mechanical and electric (and therefore electro-magnetic) precursors, the big unknown is still the role of water in gouge-filled faults or aquifers. In addition, given the complexity of the crust, large-scale processes will most likely be influenced as well by factors such as the types of rocks in the stressed volume. Finally, the main fracture is formed causing global failure. This scenario is very different from the nucleation model (Dieterich, 1992), which envisions instead a local preparation or nucleation zone with a typical size of meters to maybe a few kilometers at most. Within the nucleation model, precursors are not expected to occur over very large spatial regions.

The idea, that great earthquakes could be “critical” events in a technical sense of the term explained below, is gaining momentum in the geophysical community, even if it is far from proven as its consequences are still debated. The critical earthquake model provides an original view of the organization of the crust prior to great earthquakes that, as we shall see, allows one to explain several paradoxical observations, which pose problem within the nucleation model (Dieterich, 1992).

The word “critical” describes a system at the boundary between order and disorder, that is characterized by both extreme susceptibility to external factors and long-range correlations between different parts of the system (Sornette, 2004). Examples of such systems are magnets close to their Curie point, where the system progressively orders under small external changes. Critical behavior is fundamentally a cooperative phenomenon, resulting from the repeated interactions between “microscopic” elements which progressively “phase up” and construct a “macroscopic” self-similar state.

This idea can be traced back to the critical branching model of Vere-Jones (1977) and to the percolation model of rupture using the real-space renormalization group approach (Allegre et al., 1982). The Russian school has also extensively developed this concept (see (Keilis-Borok, 1990) and references therein). Sornette and Sornette (1990) and Sornette and Sammis (1995) identified specific measurable signatures of the predicted critical behavior in terms of a time-to-failure power-law acceleration of physical properties such as electric precursors (Sornette and Sornette, 1990) or the Benioff strain, previously interpreted differently with empirical mechanical-damage laws (Sykes and Jaume, 1990; Bufe and Varnes, 1993).

The critical-earthquake concept has been further strengthened by showing that a strong heterogeneity of the elastic and/or failure properties of rocks is necessary for a critical behavior of rupture to occur (Andersen et al., 1997; Heimpel, 1997; Sornette and Andersen, 1998; Lei et al., 2000; 2004). This was anticipated early by Mogi (1969), who noticed that, for experiments on a variety of materials, the larger the disorder, the stronger and more useful are the precursors to rupture. In a heterogeneous material, the fracture behavior will be determined by the block that is most brittle and fails in a catastrophic way. The disorder may not need to be only frozen (pre-existing) as it may be generated during the deformation processes (Bouchaud and Mezard, 1994), in particular in the transition from the generation/movement of dislocations to coalescence/microfracturing. Numerical simulations have confirmed that, near the global failure point, the cumulative elastic energy released during fracturing of heterogeneous solids with long-range elastic interactions exhibit a critical behavior with observable log-periodic corrections (Sahimi and Arbatii, 1996; Johansen and Sornette, 1998; Zhou and Sornette, 2002). Recent experiments measuring acoustic emissions associated with the rupture of fiber-glass composites (Garcimartin et al., 1997; Moura and Yukalov, 2002; Yukalov et al., 2004, Nechad et al., 2005(a,b), on kevlar and carbon fiber composites (Johansen and Sornette, 2000) and on rocks (Lei et al., 2003; Moura et al., 2005) have confirmed the critical scenario. Sammis and Sornette (2002) have
generalized the “critical point” concept for large earthquakes in the framework of so-called “finite-time singularities,” and have proposed that the singular behavior associated with accelerated seismic release results from a positive feedback of the seismic activity on its release rate. The most important mechanisms for such positive feedback include (i) stress load in subcritical crack growth, (ii) geometrical effects in creep rupture, (iii) weakening by damage, (iv) stress redistribution in a percolation model of asperity failures and (v) fragmentation of the stress-shadow (cast by the last large earthquake) by the increasing tectonic stress. The critical-earthquake model has also been tested with positive results on rockbursts occurring in deep South African mines (Ouillon and Sornette, 2000), thus offering an example with an intermediate range of scales between the laboratory and the earth crust.

As predicted by the critical earthquake model, there are many reported observations of increased intermediate magnitude seismicity before large events (Ellsworth et al., 1981; Jones, 1994; Keilis-Borok et al., 1988; Knopoff et al., 1996; Lindh, 1990; Main, 1996; Mogi, 1969; Raleigh et al., 1982; Sykes and Jaumé, 1990; Toccher, 1959). Because these precursory events occur over an area much greater than is predicted for elasto-dynamic interactions, they are not considered to be classical foreshocks (Jones and Molnar, 1979). While the observed long-range correlations in seismicity can not be explained by simple elasto-dynamic interactions, they can be understood by analogy to the statistical mechanics of a system approaching a critical point for which the correlation length is only limited by the size of the system (Wilson, 1979). There are large fluctuations of the precursory seismicity from event to event. Therefore, the statistical significance of these precursors has not yet been established firmly (Huang et al., 2000, Helmstetter and Sornette, 2003). Another ingredient for large scale correlations of the precursory signals is that the tectonic plates are being moved or dragged along by large-scale mantle convections and, hence, stresses act coherently on large sections of the crust.

If the crust does behave such that larger earthquakes are “critical” events, stress correlation lengths should grow in the lead-up to large events and drop sharply once these occur. In the critical model, a large or great earthquake dissipates a sufficient proportion of the accumulated regional strain to destroy these long wavelength stress correlations, which need to be re-established to prepare for the next big one. According to this model, large earthquakes are not just scaled-up version of small earthquakes but play a special role as “critical points” (Bowman et al., 1998; Brehm and Braile, 1999a,b; Jaume and Sykes, 1999). However, this evolution in stress correlation lengths is very difficult to observe directly. Recently, Mora and Place (2000) have shown, using the lattice solid model to describe discontinuous elasto-dynamic systems subjected to shear and compression presented in (Place et al., 1999), that cumulative strain and correlation lengths do exhibit a critical evolution in these model systems. Huang et al. (1998) have shown that the concept of critical earthquake is compatible with the concept of self-organized criticality (Bak and Tang, 1989; Sornette and Sornette, 1989) in a hierarchical structured crust (Ouillon et al., 1996). This has been further elaborated in terms of the concept of “intermittent criticality” (Bowman and Sammis, 2004; Sammis et al., 2004).

An alternative proposal, called the Stress Accumulation model sees the accelerating seismicity sequences as resulting from the tectonic loading of large fault structures through aseismic slip in the elasto-plastic lower crust (Bowman and King, 2001a,b; King and Bowman, 2003; Levin et al., 2006). According to this view, the upper crust could be driven in significant part from below (lower crust and mantle) with its associated plastic and plastic-ductile localization controlling in part the organization of faulting (Regenauer-Lieb and Yuen, 2003) and of seismicity in the upper crust. According to this view, major earthquakes would reflect this tectonic driving, while triggered earthquakes would be mostly “witnesses” (and relatively minor “actors”) of the physics of stress redistribution and delayed rupture nucleation (according to stress corrosion, damage theory and/or state and velocity dependent friction).

We think that reality is probably a mixture of critical damage build-up and stress accumulation due to the lower crust loading. While it seems unrealistic to neglect (as often done) the strength of the plastic and plastic-ductile lower crust and its mechanical coupling with the upper brittle crust, it also seems that rupture in highly heterogeneous rocks does occur generically according to the critical rupture scenario. The important point for our purpose is that both mechanisms combine to produce large scale coherent stress loading and damage processes.

PUTTING TOGETHER THE ELECTRIC BATTERY AND CRITICAL EARTHQUAKE CONCEPTS FOR LARGE SCALE ELECTRO-MAGNETIC SIGNATURES

When stress is applied, dislocations that move through the mineral grains cause the PHPs to break up and to release p-hole charge carriers. The p-holes “live” in the valence band. They travel by passing from oxygen anion to oxygen anion using energy levels provided at the upper edge of the valence band. Following their concentration gradient between stressed and unstressed rock, the p-holes spread out of the stressed rock volume. They propagate as charge waves with an estimated group velocity on the order of $100 - 300$ m/s (Freund, 2002). At the laboratory scale, this phenomenon is observed as the arrival of a charge cloud building up a concentration of positive charges at the
surface of the sample, with electric fields of the order of $10^5 - 10^6$ V/cm. Such high fields may lead to field-ionization of air molecules, the injection of ions into the air and eventually to corona discharges accompanied by light emissions that have indeed been measured (Freund, 2000).

Therefore, if there is a large scale region subjected to increasing damage at the microscopic scale, it will act as a net source of p-holes charge carriers flowing out to the surface and creating a net positive group potential. This charge accumulation can be counteracted in two ways: (1) by an outflow of electrons from the stressed rock volume and (2) by an inflow of protons possibly available from free water present in the crust (Freund et al., 2004). In a steady state, the two charge fluxes will balance and no net effect should be observed. Also, when no percolating conduction path exists, the charge accumulation increases as in the charge of a battery with no possibility for conduction. When a large earthquake becomes incipient, the critical model predicts that the damage should accelerate at the microscopic scale (which may not be usually detectable by increasing seismicity at observable levels or with other mechanical signatures), corresponding to an increasingly stronger source of p-holes charge carriers released from the rupture of peroxy bonds. In these circumstances, one might be able to observe the effects associated with a sudden increase of outflow currents, when the conditions are such that the circuit loop closes on itself allowing p-holes and electrons in the battery volume to flow out. These p-hole and electron currents are coupled through their respective electric fields and, hence, fluctuating. Another effect of the stress activation of these electronic charge carriers is the build-up of a net positive ground potential over the broad regional areas surrounding an impending large earthquake.

Essentially all the electro-magnetic manifestations reported anecdotically in the literature (Park et al., 1993) can be rationalized by our theory, once we understand that igneous and high-grade metamorphic rocks contain dormant electronic charge carriers. Stress “awakens” these charge carriers, p-holes and electrons. As a result the stressed rock volume turns into a battery from where electric currents can flow out. For these currents to flow in any sustained manner, however, certain conditions have to be fulfilled. If the build-up of stress occurs somewhere in the cool, brittle upper to middle crust, the p-holes can always spread out laterally into the surrounding unstressed rocks. By contrast, the electrons are blocked from entering the surrounding unstressed rocks. They require a conductivity path downward into the hot, n-type conductive lower crust. Once this downward connection is established, the circuit loop closes allowing both currents, electrons and p-holes, to flow out. Depending on conditions that are still poorly characterized, the outflows may occur in bursts giving rise to pulses of low-frequency electromagnetic radiation. In addition, the two outflow currents can be expected to couple via the electric fields that the propagating charges build up. Such a coupling should give rise to fluctuations, which would be another source of low-frequency electromagnetic emissions.

Small earthquakes at depth within the crust are expected to generate individually electric signals which are too weak to be detected. But their collective electric effect during the build-up phase towards the critical point, in which many small events below the detection threshold are triggered, may be significant and detectable at the surface. Indeed, there are reasons to believe (Sornette and Werner, 2005a,b) that there is an immense swarm of small earthquakes below detection thresholds which, notwithstanding their small size, may dominate the triggering of detectable earthquakes. Collectively, these events can produce the conditions for the build-up of the electric charges and their circulation. In this respect, mine rockburst offer interesting analogs and an intermediate scaling range to test this idea. First, the power law acceleration of Benioff strain often reported for earthquakes (Keilis-Borok and Malinovskaya, 1964; Sykes and Jaumé, 1990; Bufe and Varnes, 1993; Saleur et al., 1996; Bowman et al., 1998; Jaumé and Sykes, 1999) is also a characteristic of the space-time organization of rockbursts prior to their largest counterparts (Ouillon and Sornette, 2000). Second, it was suggested that electro-magnetic radiation also exhibits a similar acceleration and can thus be useful for the prediction of rockbursts (Frid and Vozoff, 2005). We also note that the organization of seismicity prior to mainshocks seems similar, with the same statistical signatures of accelerated Benioff strain and Gutenberg-Richter distribution (Rabinovitch et al., 2001), both for mine rockbursts as mentioned above, to earthquakes and all the way to the scale of neutron stars, whose quakes are observed in the form of soft gamma-ray repeaters (Kossobokov et al., 2001; Sornette and Helmstetter, 2002).

Our theory also opens the door toward understanding, at least in principle, other pre-earthquake phenomena that have been widely reported in the literature. Prior to major earthquakes, the ionosphere displays remarkable perturbations (Molchanov et al. 1993; Molchanov and Hayakawa 1998b, c; Alperovich and Fedorov 1999; Chen et al. 1999; Cliverd et al. 1999; Liu et al. 2000, 2001; Pulinets et al. 2000; Sorokin et al. 2001; Pulinets and Boyarchuk 2004). These perturbations extend over large areas, on the order of 500-1000 km. They require changes of the electric field at ground level, strong enough to induce a recognizable reaction in the ionosphere. If the p-hole charge carriers are able to spread out of a stressed rocks volume in the upper to middle crust, they can be expected to reach the Earth’s surface. If they get trapped at the surface as laboratory experiments suggest (Freund, et al., 1993; Freund, et al., 1994), the surface will acquire a positive charge. In the field, this would be equivalent to a change in the Earth’s ground potential over the region where the stress accumulation takes place. The Earth-ionosphere system can be considered a capacitor of which one plate, fixed at the Earth’s surface, becomes positively charged. The ionosphere,
representing the other, flexible plate, will then react and produce a mirror image of the charge accumulation on the ground.

Other effects derive from the electric fields that are built up on a microscopic scale at the rock-air interface when p-holes arrive at the surface. Theory predicts that the p-holes form charge layers at the interface, $10 - 100$ nm thick, with charge densities sufficient to produce surface potentials on the order of $0.1 - 1$ V (King and Freund, 1984). Experimentally, surface potentials in the range of $50$ mV to $10 - 20$ V have been observed (Freund, et al., 2004). The associated electric field, calculated for a flat surface (King and Freund, 1984), then fall into the range of $5 \cdot 10^5 - 10^7$ V/cm. At sharp points, the electric fields can be expected to reach values high enough to cause field ionization of air molecules at the interface and injection of positive ions into the atmosphere. These processes would in turn be consistent with dielectric breakdown of the air and with glow or corona discharges, i.e. with luminous phenomena and high-frequency electromagnetic emissions in the range of tens of kHz.

Yet another widely reported pre-earthquake phenomenon may find its physical explanation through a better understanding of the p-hole charge carriers: the observation that, prior to major earthquakes in semi-arid parts of the world, large areas of the land surface around the future epicenter tend to emit infrared radiation equivalent to a temperature increase on the order of $2 - 4^\circ$C (Dey and Singh, 2003; Srivastav, et al., 1997; Tronin, 1996; Tronin, et al., 2004; Qiang, et al., 1999; Freund, et al., 2003; Ouzounov and Freund, 2004; Liu, et al., 2000; Naaman, et al., 2001; Ohta, et al., 2001; Ondoh, 1998; Vershinin, et al., 1999; Pulinets, et al., 2005; Trigunait et al., 2004; Yen, et al., 2004). These “thermal anomalies” come and go rapidly. After the earthquake and any aftershocks they disappear. If p-holes spread out from a stressed rock volume, travel though thick layers of rock and reach the Earth’s surface, they can be expected to recombine to form peroxy links similar to those that were broken deep below during application of stress. During hole-hole recombination, energy is released. A consequence of this process would be the formation of vibrationally excited O-O bonds. These excited bonds can dissipate their excess energy by emitting IR photons at the characteristic frequencies of the O-O stretching vibration or by transferring energy onto their Si-O and Al-O neighbors causing them to become excited and to emit at their characteristic frequencies. Laboratory experiments have produced results that are consistent with this line of reasoning. They show that, when one end of a large block of rock is subjected to stress, a free surface $50$ cm away instantly begins to emit in the infrared. The emission spectrum is composed of sharp bands at the characteristic frequencies of O-O, Si-O and Al-O vibrations (Freund et al., 2003).

A further piece of evidence for an outflow of p-holes from a stressed rock volume is provided by the serendipitous observation of strong magnetic field variations along the $110$ km long Chelungpu Fault in Taiwan that ruptured during the M=7.7 Chi-Chi earthquake on Sept. 21, 1999. For seven week before the main event and for several weeks during the intense aftershock period, the local magnetometer network recorded rapid, pulse-like variations of the magnetic field, each lasting for several hours (Yen et al., 2004). Telluric currents on the order to $10^8$ Amp would have been necessary to produce such magnetic field excursions (Freund et al., 2004).

Another consequence of our theory is the following. As the crust “battery” produces a current, the outflow of p-holes through the rock is equivalent to electrons hopping back into the stressed rock volume. These electrons will reconstitute the positive hole pairs inside the stressed rock volume. Therefore, the battery is (theoretically) inexhaustible: the current can run until the battery is “empty.” As stress is varying in time due to intermittent earthquakes and the slow tectonic loading, the rock again charges up. The battery should be reusable many, many times, paralleling the many, many seismic cycles over millions of years of a given tectonic plate. In this vein, let us mention a recent study supporting the idea that electric signals in the crust (ultra-low frequency ground electric signals from stations operated by the China Seismological Bureau over the last 20 years) have a statistically significant pre-seismic component and are thus linked with the preparatory stage before earthquakes (Zhuang et al., 2005).

An order-of-magnitude estimation, extrapolating from the laboratory at the scale of rock blocks of $1$ meter long, suggests very large currents in the range of thousands to millions of amperes flowing out of the hundreds to thousands of cubic kilometers of rock, which experience increasing levels of stress prior to the catastrophic rupture. Conceptually, these strong currents can be concentrated in narrow conducting paths. This begs the question of why we do not observe these currents and EM radiations more easily. To address this question, recall that our proposed concept is that of the crust acting like a battery being electrically “charged” by the application of stress. Inside the stressed rock volume, both electrons and defect electrons (p-holes) are activated. The p-holes can spread out into the surrounding unstressed rocks because p-holes move along the upper edge of the valence band (which is dominated by energy states of O2s and O2p character). In other words, unstressed rock are (every so slightly) p-type conducting. We believe that, as a result, p-holes can travel very far. The electrons, however, are stuck inside the stressed rock volume. They cannot leave it unless an n-type contact is provided, which is easily done in the laboratory by sticking a Cu tape on the stressed rock volume or using the steel pistons as contacts. In the laboratory, a wire from the stressed rock volume is run to any place on the unstressed rock, which closes the loop and allows the battery to discharge. Out in nature, we propose that the n-type contact can be provided by rocks deeper in the crust that exceed $\sim 500^\circ$C and thereby
become n-type conductive. Another possibility is through the percolation of fractures filled with small amounts of conductive brine, which is expected to occur close to the critical point. Therefore, the answer to the question, “why we don’t “see” powerful currents more often and more easily”, is that the battery is there, ready to release hundreds of thousands of amperes per km$^3$, but the electric loop is normally not closed. The loop closes only when massive stresses build up that reach from the upper crust down to the hot lower crust or when damage percolates through a conductive path, closing the electric circuit.

Our theory is falsifiable (Kagan, 1999) as it makes specific tests that can be refuted. Our preferred crucial test would consist in deploying a large array of stations capable of measuring the potential difference of the crust at the scale of tens of kilometers or more. To the best of our knowledge, such measurements have never been done before but constitute one of the best probe for detecting anomalous accumulation of positive charges at locations related to large impending earthquakes.

Acknowledgements: We thank Xinglin Lei and another anonymous referee for useful comments which helped improve the manuscript.

Sornette, D. and Werner, M., 2005b. Apparent Clustering and Apparent Background Earthquakes Biased by Undetected
Sykes, L. R. and Jaum´e, S., 1990. Seismic activity on neighboring faults as a long-term precursor to large earthquakes in
Sornette, D. and C.G. Sammis, 1995. Complex critical exponents from renormalization group theory of earthquakes :
National Academy of Sciences USA, V99 SUPP1, 2501-2508.
Sornette, A. and D. Sornette, 1990. Earthquake rupture as a critical point: Consequences for telluric precursors. Tectono-
physics 179, 327-334.
Sensing, 17, 1439-1455.
Science, 72, 880-884 (1997).
Sykes, L. R. and Jaumé, S., 1990. Seismic activity on neighboring faults as a long-term precursor to large earthquakes in
the San Francisco Bay area, Nature (London) 348, 595-599.
Taylor, P. and M. Purucker, 2002. Searching for a magnetic signature from earthquakes in the ionosphere, 76 pp., Final
Report, GEOS Requirements Definition Study.
732-735.
Trahinat, A.P., S. Pulinets and F. Li, 2004. Variations of the ionospheric electron density during the Bhuj seismic event,
Annales Geophysicae, 22, 4123-4131.
Sensing, 17, 1439-1455.
Sensing, 25, 2649-2655.
Phys. Earth, 45: 73-82.
FIG. 1: Simple picture sequence for the conversion of hydroxyl pairs in MgO to peroxy (PHP) plus \( H_2 \) and the “tearing apart” of the PHP to generate a p-hole. The circles indicate the delocalization of the charge associated with the p-hole and with the defect that is left behind at the site of the new broken peroxy bond. The case of Si-OO-Si in silicates, while more complex to represent, proceeds analogously.
FIG. 2: Brittle and ductile response at the level of small rock volumes. (a) Increasing stress in the brittle regime causes dislocation movement, coalescence of dislocations to microfractures and merging of microfractures to cracks. (b) Increasing stress in the ductile regime causes plastic deformation but, if the stresses become very high and increase rapidly, even ductile materials can develop cracks.
FIG. 3: Cross section through two crustal blocks colliding at a constant speed. (a) Strain versus time. (b) Stress versus time for a brittle material being deformed at a constant speed. (c) Viscosity decreasing with increasing depth and temperature. (d) Volume of rock undergoing brittle fracture as a function of time assuming deformation at a constant speed and, hence, increasing levels of stress.
This figure "Figure_1_Critical_Rupture_lowQ.jpg" is available in "jpg" format from:

http://arXiv.org/ps/physics/0603205
This figure "Figure_2_Critical_Rupture_lowQ.jpg" is available in "jpg" format from:

http://arXiv.org/ps/physics/0603205
This figure "Figure_3_Critical_Rupture_lowQ.jpg" is available in "jpg" format from:

http://arXiv.org/ps/physics/0603205