

The Inverse Problem of Magnetoencephalography: Source Localization and the Shape of a Ball

By *Erkki Somersalo*

The human brain is a collection of specialized cells of different types, such as astrocytes and neurons. The approximately 10^{11} neurons are excitable cells whose resting state is characterized by a cross-membrane voltage difference. Neurons have long filaments, called axons, that allow the propagation of electromagnetic signals as voltage variations, known as “action potentials.”

The transfer of a signal from one neuron to another occurs in a chemical process performed by neurotransmitters over the synaptic gap, creating a postsynaptic potential in the receiving neuron. The postsynaptic potential, which can be fairly stable over a period of milliseconds, can be modeled as a localized current dipole. Neurons are organized in bundles; when thousands of neighboring neurons are simultaneously in the postsynaptic excitation state, the postsynaptic potentials together give rise to a localized current approximately parallel to the neuron bundle. This elementary weak current, called the “impressed current,” can be strong enough to create an observable electromagnetic field outside the skull. The inverse problem of magnetoencephalography (MEG) is to estimate impressed currents from observations of magnetic fields outside the skull.

The extracranial magnetic fields induced by neural activity are extremely weak. Normal brain activity induces fields of order 10^{-15} Tesla; fields during epileptic seizures can reach the 10^{-13} Tesla level. Measuring fields of this size is a great challenge, requiring the use of SQUID (for superconducting quantum interference device) magnetometers. Another technological challenge is isolation of the measuring environment from external magnetic noise. To understand the importance of this isolation, consider that the magnetic field of the Earth is of order $5 \cdot 10^{-5}$ Tesla, and that an ordinary compass is readily disturbed by normal home electronics or large nearby metallic objects.

A common way to model impressed currents is to generate a grid that covers the region of interest in the brain and attach three mutually orthogonal electric dipoles with unknown amplitudes at each grid point. The electromagnetic fields obey Maxwell’s equations; because of this linearity, the calculated magnetic field is a linear combination of the magnetic fields induced by each individual dipole. If $\vec{J}(x)$ denotes the impressed current field and \vec{q}_j is the j th unit dipole located at x_j , we have a discretized model,

$$\vec{J}(x) = \sum \delta(x - x_j) \alpha_j \vec{q}_j, \quad b = L\alpha + e,$$

where α is a vector containing the dipole amplitudes α_j , b is a vector containing the measured magnetic field components at the magnetometers, e represents the observation noise, and, finally, L is the “lead field matrix” whose j th column contains the calculated observations to which the j th dipole \vec{q}_j gives rise at the magnetometers.

Apart from the technological challenges, the MEG inverse problem poses great computational problems. One of the central difficulties is the existence of silent currents—that is, impressed currents that cause no signal in the detectors. In terms of our linear discretized model, this means that the lead field matrix L has nontrivial null space. To choose a meaningful solution vector from all possible solutions, we need to use a selection rule that is based on the available prior information concerning the impressed currents.

The earliest MEG reconstruction algorithms were based on a straightforward idea: Because the data gives no information about the null space of the lead field matrix, the null space component should be set to zero. This reasoning leads to the concept of the “minimum-norm solution,” α_{MN} , which is commonly formulated in the MEG literature as

$$\alpha_{\text{MN}} = \arg \min \|\alpha\|, \text{ subject to } \|L\alpha - b\| = d,$$

where $d > 0$ is the approximate norm of the error vector e . In other words, we are seeking the coefficient vector with minimum euclidian norm, satisfying the data up to the error discrepancy.

The minimum-norm estimate results from a selection principle that is justified more by algebraic considerations than by a priori beliefs about the impressed current. It turns out that the minimum-norm estimates are typically spread out spatially—that is, a large number of elementary dipoles are simultaneously active. Such predictions often contradict what we believe to be happening in reality. When a subject moves a finger, for instance, well-localized activity should appear on the motor cortex and the MEG system should see a few active dipoles rather than wide-spread activity.

An alternative selection principle that leads to strongly localized solutions is based on the idea that the data should be explained with the minimum number of active dipoles. Such a solution, called the “minimum-current estimator,” α_{MC} , can be defined as

$$\alpha_{\text{MC}} = \arg \min \sum |\alpha_j|, \text{ subject to } \|L\alpha - b\| = d.$$

Here we are minimizing the ℓ^1 norm, rather than the euclidian ℓ^2 norm, of the coefficient vector. This seemingly innocent change has a dramatic effect on the estimated solution. Figure 1 shows the minimum-norm and minimum-current estimates. A single dipole was used to simulate the data, and noise was added.

Is there a simple way to understand why the estimates behave so differently? To gain insight, we assume that we are solving an underdeter-

mined system, $Ax = y$, using either the minimum- ℓ^1 criterion or the minimum- ℓ^2 criterion.

Schematically, we can imagine that x is a point in the plane, with $Ax = y$ defining a line in the space. For the minimum-norm solutions, we look for the smallest ball that touches this line; the touching point is the solution. The ℓ^2 ball in this case is simply a circle, while the ℓ^1 ball is a diamond with corners on the coordinate axes (see Figure 2). The ℓ^1 ball, because it extends along the coordinate axes, is likely to touch the line $Ax = y$ on one of the coordinate axes. But this means exactly that the minimum- ℓ^1 solution explains the data with a single non-zero component. This simple geometric picture can be extended to higher dimensions, and it is possible—if not simple—to imagine what the solutions look like.

Interestingly, minimum- ℓ^1 solutions have received lots of publicity recently, in the context of restoring piecewise constant signals and images from incomplete frequency samples. The basic explanation of the functioning of these algorithms is the same as the one given here for the MEG inverse problem.

Erkki Somersalo is a professor of mathematics at Helsinki University of Technology.

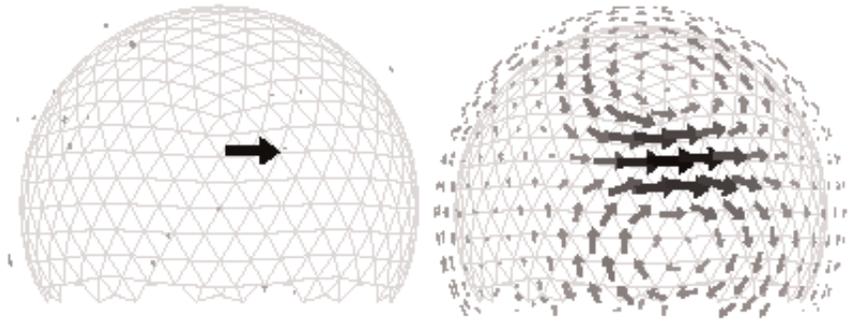


Figure 1. Minimum-norm (left) and minimum-current (right) estimates. A seemingly innocent change is responsible for the very different behaviors.

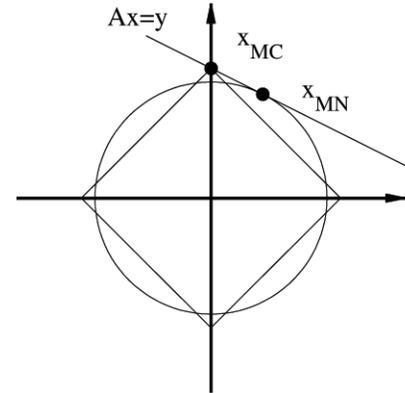


Figure 2. A simple geometric picture helps explain why the estimates of Figure 1 behave so differently.