Backus and Gilbert Method for Vector Fields

Rolando Grave de Peralta Menendez* and Sara L. Gonzalez Andino

Functional Brain Mapping Lab, Department of Neurology, Geneva University Hospital, Geneva, Switzerland

◆ ━━ ◆

Abstract: This report describes the theory of Backus and Gilbert with special emphasis for the case of vector fields as required for the solution of the electromagnetic inverse problem. A description of the method is presented with the detailed mathematical derivation of the coefficients that determine the solution for the retrieval of vector fields. Such derivation, to our knowledge, has never been reported in the literature. We also identify some crucial points that can (and had) lead to misuse of this solution and describe some disadvantages of this theory for the case of vector fields suggesting some alternatives to deal with them. *Hum. Brain Mapping 7:161–165, 1999.*   © 1999 Wiley-Liss, Inc.

Key words: Backus Gilbert method; inverse problem; bioelectromagnetism; bioelectromagnetic inverse problem

◆ ━━ ◆

INTRODUCTION

The necessity of solving the linear inverse problem with discrete data appears on several fields of application. One example is the determination of the current distribution within the human brain from scalp-recorded electric and/or magnetic data. One of the multiple alternatives that exist to solve the bioelectromagnetic inverse problem is the Backus and Gilbert (BG) method. The basic idea of the BG method, already applied in bioelectromagnetism [Robinson and Rose, 1992; Grave de Peralta et al., 1996; Lütkenhönner and Grave de Peralta Menendez, 1997] is to try to achieve the maximum spatial resolution at each site from the available data.

The basic elements of the Backus and Gilbert theory appeared in three articles: Backus and Gilbert [1967, 1968, 1970]. This report emphasizes the exposition of the method for the retrieval of scalar fields. Although the estimation of vector fields was briefly mentioned in Backus and Gilbert [1968], this problem has since received little attention. Since the bioelectromagnetic inverse problem is usually stated in terms of the current density vector, the application of this method requires a careful analysis because the extension of this theory to vector fields is not completely straightforward as was suggested in Grave de Peralta et al. [1996]. This is evident from some erroneous derivations of the method that have already appeared in the literature. In addition, some disadvantages of this theory have been identified from application to the retrieval of vector fields [Grave de Peralta Menendez et al., 1996, 1997a].

Here, we briefly present the BG theory with special emphasis on the case of vector fields. We offer a detailed description of the procedure to compute the coefficients that determine the BG solution, which to our knowledge has not been described before in the literature. We note the inconveniences of this method

together with alternatives to circumvent them. Possible mistakes that might or have been already committed in the application of the method to the case of vector fields are also described.

BASIC EQUATIONS

Boldface uppercase letters indicate matrices, and boldface lowercase letters denote vectors.

The Backus and Gilbert method is an alternative inversion procedure for the linear inverse problem with discrete data determined by the system of linear equations:

$$\mathbf{d}_i = \int_R \mathbf{l}_i(\mathbf{r}) \circ \mathbf{j}(\mathbf{r})d\mathbf{r} \qquad i = 1 .. N_s \qquad (1)$$

where $\mathbf{d}_i$ represents the i-th measurement obtained after the integration over region R of the product between the lead field function $\mathbf{l}_i(\mathbf{r})$ and the unknown function $\mathbf{j}(\mathbf{r})$. The symbol ($\circ$) represents the internal (scalar) product between vectors, and N_s stands for the number of measurements.

A well-known theorem of functional analysis from Banach states that if the inverse of a linear operator exists, it also will be a linear operator. On this basis, if the inverse operator determined by the system of equations (1) has an inverse, it can be written as:

$$\hat{\mathbf{j}}(\mathbf{r}_0) = \sum_{i=1}^{N_s} \mathbf{c}_i(\mathbf{r}_0)\mathbf{d}_i \qquad (2)$$

Since $\hat{\mathbf{j}}(\mathbf{r}_0)$ is a vector field of dimension, n (e.g., n = 1, 2, 3), equation (2) makes sense only if the coefficients $\mathbf{c}_i(\mathbf{r}_0)$ are vectors of the same dimension as $\mathbf{j}$. Note the dependence of the coefficients on the solution point $\mathbf{r}_0$. In this way the problem is reduced to the selection of the functions $\mathbf{c}_i(\mathbf{r}_0)$, that express the linear relationship between the measurement data vector and the solution at point $\mathbf{r}_0$.

Substituting (1) into (2) gives:

$$\hat{\mathbf{j}}(\mathbf{r}_0) = \int_R \sum_{i=1}^{N_s} \mathbf{c}_i(\mathbf{r}_0) \otimes \mathbf{l}_i(\mathbf{r})\mathbf{j}(\mathbf{r})d\mathbf{r} = \int_R \mathbf{A}(\mathbf{r}_0, \mathbf{r})\mathbf{j}(\mathbf{r})d\mathbf{r} \quad (3)$$

where $\otimes$ stands for the external product between vectors and the kernel of the resolution operator determined by the *n* by *n* matrix:

$$\mathbf{A}(\mathbf{r}_0, \mathbf{r}) = \sum_{i=1}^{N_s} \mathbf{c}_i(\mathbf{r}_0) \otimes \mathbf{l}_i(\mathbf{r}) \qquad (4)$$

is called the averaging or resolution kernel. The key point of the Backus and Gilbert theory [1968] is the selection of the coefficients c_i by optimizing a function of the resolution kernel $\mathbf{A}(\mathbf{r}_0, \mathbf{r})$ that measures its "deltaness" subject to an auxiliary unimodularity constraint.

It is important to note that finding the coefficients for each target point $\mathbf{r}_0$, according to equation (2), will determine a function all over the solution space. In general, such a function will not fulfill the system of equations (1), although as noted from equations (2) and (3), each particular value $\hat{\mathbf{j}}(\mathbf{r}_0)$ is a weighted average shared by all the solutions. For this reason some authors refer to the BG method as one of solution appraisal, or to estimate the information content of the given data set instead of a method for solution construction.

Scalar fields

For simplicity, we first consider the one-dimensional case, i.e., when the dimension of vector $\mathbf{j}$ is 1 (n = 1). In that case, a general analysis can be carried out considering the following measure of closeness between the delta function and the resolution kernel $\mathbf{A}(\mathbf{r}_0, \mathbf{r})$.

$$\min \int_R [\mathbf{A}(\mathbf{r}_0, \mathbf{r}) - \delta(\mathbf{r}_0 - \mathbf{r})]^2 W(\mathbf{r}_0, \mathbf{r})d\mathbf{r}$$

$$= \min \int_R [\mathbf{A}^2(\mathbf{r}_0, \mathbf{r}) + \delta^2(\mathbf{r}_0 - \mathbf{r})$$

$$- 2\mathbf{A}(\mathbf{r}_0, \mathbf{r})\delta(\mathbf{r}_0 - \mathbf{r})]W(\mathbf{r}_0, \mathbf{r})d\mathbf{r} \quad (5)$$

Using the standard formalism for the delta function [Roach, 1970] and considering only terms that depend upon the unknown coefficients, problem (5) is equivalent to:

$$\min \int_R \mathbf{A}^2(\mathbf{r}_0, \mathbf{r})W(\mathbf{r}_0, \mathbf{r})d\mathbf{r} - 2\mathbf{A}(\mathbf{r}_0, \mathbf{r}_0)W(\mathbf{r}_0, \mathbf{r}_0)$$
$$(6)$$

From this expression, the typical criteria for the selection of the coefficients can be derived. If $W(\mathbf{r}_0, \mathbf{r}_0) = 0$, then an additional condition is needed to obtain the "classical" BG method as often described [e.g., Backus and Gilbert, 1968, 1970; Grave de Peralta Menendez et al., 1996]. It corresponds to the so-called Q-criterion:

$$\min \int_R \mathbf{A}^2(\mathbf{r}_0, \mathbf{r})W(\mathbf{r}_0, \mathbf{r})d\mathbf{r} \qquad \text{st:} \int_R \mathbf{A}(\mathbf{r}_0, \mathbf{r}) = 1 \quad (7)$$

Here, it is important to discuss the selection of function *W*. After having pointed out that several alternatives

are available, Backus and Gilbert [1968] defined the nonpathological weight function W as a monotone function of the distance between the two arguments and such that does not become constant with the increase of this distance and such that $W(r, r) = 0$. They concluded that if the solution obtained for A with a nonpathological weight W does not resemble a blurred delta function, then none of the linear combinations of the $l_i(r)$ will be near to the delta function and thus it is not possible to obtain an approximation to an average near the target point r_0 from the given data.

A second alternative called the D criterion corresponds to the case of constant weighting function, i.e., $W(r_0, r) = $ constant. Because of the metric associated to it, it is also called the least-square criterion and is determined by the following optimization problem:

$$\min \int_R [A(r_0, r) - \delta(r_0, r)]^2 dr$$

$$= \int_R A^2(r_0, r) dr - 2A(r_0, r_0) \quad (8)$$

It is not difficult to check that the solution obtained corresponds to the minimum norm solution as described in the appendix of Backus and Gilbert [1968], Tarantola [1982], and others. In practice, some authors modify this criterion, including the unimodularity constraint as in (7) [Oldenburg, 1981].

A third criterion is the so-called S method and corresponds to minimize the difference between A and the delta function in a determined Sobolev space. It is not considered here.

Vector fields

The following optimization problem is a natural (but not unique) extension of problem (5) to the case of vector fields:

$$\min \int_R \|A(r_0, r) - I\delta(r_0 - r)\|^2 W(r_0, r) dr \quad (9a)$$

With the auxiliary constraints (when needed):

$$\int_R A(r_0, r) dr = I \quad (9b)$$

where I represents the identity matrix and $\|X\|$ represent the Frobenius norm of matrix X, i.e., the sum of all the elements of X squared.

Expresion (9a) was considered for the discrete case in Grave de Peralta Menendez et al. [1997a]. There, we

proved that for the case of a constant weighting function, i.e., $W(r, r) = $ constant, there is no need for an additional constraint and the solution obtained corresponds to the minimum norm solution and not to the Backus and Gilbert solution. This mistake appeared in Riera et al. [1997] where the authors attempted to compare the BG method with the minimum norm, but derived the BG solution with a constant weighting function and without auxiliary constraint. It is hard to explain why they found differences in their results when the derivation used for BG exactly produces the minimum norm solution. The other common error when deriving the BG solution is to assign the same estimator for all the n components of the vector field at the target point.

In the original work of Backus and Gilbert [1968], they consider instead of (9) a simpler expression that can be obtained like a particular case of:

$$\min \int_R \|\text{diag}\{A(r_0, r) - I\delta(r_0 - r)\}\|^2 W(r_0, r) dr$$

$$+ \int_R \|A(r_0, r) - \text{diag}\{A(r_0, r)\}\|^2 dr \quad (10)$$

where diag (X) represents a diagonal matrix with main diagonal as X. Denoting

$$c_i(r_0) = (c_i^1(r_0), c_i^2(r_0), \ldots, c_i^n(r_0))^t$$

$$\text{and } l_i(r) = (l_i^1(r), l_i^2(r), \ldots, l_i^n(r))^t \quad (11)$$

and using (4), problem (10) can be rewritten as:

$$\min \int_R \sum_{k=1}^{n} \left[\left\{ \sum_{i=1}^{N_s} c_i^k(r_0) l_i^k(r) - \delta(r_0 - r) \right\}^2 \right.$$

$$\left. \times W(r_0, r) + \sum_{\substack{p=1 \\ p \neq k}}^{n} \left\{ \sum_{i=1}^{N_s} c_i^k(r_0) l_i^p(r) \right\}^2 \right] dr \quad (12)$$

Again, if $W(r, r) = 0$, an additional constraint is needed and similarly to (6) and (7), we have:

$$\min \int_R \sum_{k=1}^{n} \left[\left\{ \sum_{i=1}^{N_s} c_i^k(r_0) l_i^k(r) \right\}^2 \right.$$

$$\left. \times W(r_0, r) + \sum_{\substack{p=1 \\ p \neq k}}^{n} \left\{ \sum_{i=1}^{N_s} c_i^k(r_0) l_i^p(r) \right\}^2 \right] dr$$

$$\text{st: } \int_R \left\{ \sum_{i=1}^{N_s} c_i^k(r_0) l_i^k(r) \right\} dr = 1, \text{ for } k = 1 \ldots n \quad (13)$$

Note in (12) and (13) that the problem can be independently solved for each value of k = 1 .. n. Problem (13) corresponds to the formula presented in Backus and Gilbert [1968] and is used in the following section to determine the coefficients.

Computing coefficients

As noted earlier, problem (13) can be solved for each value of k. In order to derive the coefficients that solve (13), we consider the estimation of the component k = m of a vector field of dimension n at point r_0. If we define:

$$\mathbf{f} = (c_1^m(\mathbf{r}_0), c_2^m(\mathbf{r}_0), \ldots, c_{N_s}^m(\mathbf{r}_0))^t \quad (14)$$

$$\mathbf{M}_{ij} = \int_R l_i^m(\mathbf{r}) l_j^m(\mathbf{r}) W(\mathbf{r}_0, \mathbf{r}) d\mathbf{r}$$

$$+ \sum_{\substack{p=1 \\ p \neq m}}^{n} \int_R l_i^p(\mathbf{r}) l_j^p(\mathbf{r}) d\mathbf{r} \quad (15)$$

$$\mathbf{u}_i = \int_R l_i^m(\mathbf{r}) d\mathbf{r} \quad (16)$$

then, the unknown coefficients $\mathbf{f}$ are the solution of a quadratic optimization problem with a linear constraint that is equivalent to problem (13) for k = m:

$$\min \mathbf{f}^t \mathbf{M} \mathbf{f} \qquad st: \mathbf{f}^t \mathbf{u} = 1 \quad (17)$$

that, if $\mathbf{M}$ is invertible, has the solution:

$$\mathbf{f} = \frac{\mathbf{M}^{-1} \mathbf{u}}{\mathbf{u}^t \mathbf{M}^{-1} \mathbf{u}} \quad (18)$$

These coefficients can be substituted in (2) to obtain the BG estimator for the m-th coordinate of $\mathbf{j}(\mathbf{r}_0)$. Note that $\mathbf{f}$ comprises only the m-th coordinate of each $c_i(\mathbf{r}_0)$. Repeating the solution process (14)–(18) for different values of k = 1 .. n, we can completely determine the coefficients $c_i(\mathbf{r}_0)$ for the point r_0, and by changing the target point r_0, we can compute the BG estimate for the whole solution space.

Note that in the discrete case, i.e., when instead of system (1) we have a system of algebraic equations, equations (15) and (16) change accordingly to sums over the rows of the lead field matrix. Similar to the scalar case described in Backus and Gilbert [1967, 1968, 1970], to consider noisy data, it is enough to add an additional term to the elements of matrix $\mathbf{M}$ in (15).

DISCUSSION

It is interesting to discuss some important aspects of formulas (9) and (13). Consider the estimation of the component k of a vector field of dimension n at point r_0. From equation (3), is clear that the averaging kernel of the k-th component has factors multiplying the k-th component of the vector field for each point $\mathbf{r} \neq r_0$ and factors multiplying the other components p (p ≠ k). Then, the resolution kernel can be decomposed in the "main resolution kernel" associated to the k-th component and the others "cross resolution kernels" associated to the p (p ≠ k) components. If we consider one row of $\mathbf{A}(\mathbf{r}_0, \mathbf{r})$ in (4), then the main resolution kernel will be the function in the diagonal and the cross resolution kernels are determined by the off-diagonal elements of that row.

With this convention note that in (9a) all the differences determined by the main and the cross resolution kernels minus the delta and the null function, respectively, are weighted. In addition, the constraints in (9b) force the main resolution kernels to have an integral equal to 1 and the cross resolution kernels to have an integral equal to zero. In contrast, in (13) only the difference of the main resolution kernels and the delta is weighted, i.e., the difference between the cross resolution kernels and the null function has a constant weight equal to 1. In addition, the constraints of problem (13) only force the main resolution kernels to have an integral equal to 1, without any constraint on the cross resolution kernels. In our opinion, these are two weak points of the BG method for vector fields as presented in Backus and Gilbert [1968]. Other problems, such as the inability to synthesize a delta or the presence of negatives sidelobes, cannot be considered as shortcomings of this method since they appear in all the linear inverse solutions and mainly depend on the set of functions $l_i(\mathbf{r})$.

The first aspect was considered in Grave de Peralta Menendez et al. [1997a] where some changes are introduced in the weighting strategy to obtain a quadratic optimization problem without additional constrains, i.e., $W(\mathbf{r}, \mathbf{r}) \neq 0$. In addition, the differences between the cross resolution kernels and the null function are weighted to reduce the contribution of faraway points. This approach, termed the Weighted Resolution Optimization (WROP) method, has proved to be a quite flexible approach containing as particular cases some well-known solutions and providing encouraging results in the analysis of experimental data [Grave de Peralta Menendez et al., 1997b].

The second point, i.e., imposing the condition to the sum of the cross resolution kernels, was considered

in Ory and Pratt [1995]. To our knowledge, there is no work considering both aspects together. For an alternative to suppress the negative sidelobes of the resolution kernels, see Huestis [1987].

Conclusions

We present the Backus and Gilbert theory for the case of vector fields including the detailed derivation of the coefficients that determine the solution. Two main drawbacks of this theory are identified together with some typical mistakes that appear in the literature when trying to derive the Backus and Gilbert solution. Alternatives to deal with the drawbacks are suggested, driving the attention to the references that deal with them.

ACKNOWLEDGMENTS

REFERENCES

Backus GE, Gilbert JF. 1967. Numerical applications of a formalism for geophysical inverse problems. Geophys J Roy Astron Soc 13:247–276.

Backus GE, Gilbert JF. 1968. The resolving power of gross earth data. Geophys J Roy Astron Soc 16:169–205.

Backus GE, Gilbert JF. 1970. Uniqueness in the inversion of gross earth data. Phil Trans R Soc 266:123–192.

Grave de Peralta Menendez R, Gonzalez Andino S, Lütkenhöner B. 1996. Figures of merit to compare linear distributed inverse solutions. Brain Topography 9:117–124.

Grave de Peralta Menendez R, Hauk O, Gonzalez Andino S, Vogt H, Michel CM. 1997a. Linear inverse solutions with optimal resolution kernels applied to the electromagnetic tomography. Hum Brain Mapp 5:454–467.

Grave de Peralta Menendez R, Gonzalez Andino S, Hauk O, Spinelli L, Michel CM. 1997b. A linear inverse solution with optimal resolution properties: WROP. Biomedizinische Technik 42:53–56.

Huestis SP. 1987. Construction of non-negative resolving kernels in Backus-Gilbert theory. Geophys J R Astr Soc 90:495–500.

Lütkenhönner B, Grave de Peralta Menendez R. 1997. The resolution field concept. Electroencephal Clin Neurophysiol 102:326–334.

Oldenburg DW. 1976. Calculation of Fourier transforms by the Backus Gilbert method. Geophys J R Astr Soc 44:413–432.

Ory J, Pratt RG. 1995. Are our parameter estimators biased? The significance of finite-difference regularization operators. Inverse Problems 11:397–424.

Riera JJ, Valdes PA, Fuentes ME, Oharriz Y. 1997. Explicit Backus and Gilbert EEG inverse solution for spherical symmetry. Biomedizinische Technik 42:53–56.

Roach GF. 1970. Green functions. Introductory theory with applications. New York: Van Nostrand Reinhold.

Robinson S, Rose D. 1992. Current source image estimation by spatially filtered MEG. In: Hoke M, Erne S, Okada Y, Romani G, editors. Biomagnetism: clinical aspects. Amsterdam: Excerpta Medica, p 761–765.