Three-dimensional modelling of ocean electrodynamics using gauged potentials

Robert H. Tyler,1 Frédéric Vivier2 and Shenghui Li1

1Ocean Physics Department, Applied Physics Laboratory, University of Washington, 1013 NE 40th St, Seattle, WA 98105, USA
E-mail: tyler@apl.washington.edu
2Laboratoire d’Océanographie Dynamique et de Climatologie, Université Pierre et Marie Curie, Paris Cedex 05, France

Accepted 2004 March 19. Received 2004 March 5; in original form 2003 October 7

SUMMARY
It is appreciated that one of the ‘bottlenecks’ in understanding low-frequency electrodynamical processes in the solid earth, oceans, and lower atmosphere is the limitations of current numerical models capable of simulating realistic 3-D aspects of these phenomena. The limitations have primarily been set by the computational demands of such calculations, as well as by the fact that the numerical implementation of the Maxwell equations can be at times cumbersome when compared with formulations such as the heat equation, which afford simple intuitive numerical implementation based on simple flux conservations. Here we discuss the advantages of formulations based on the electromagnetic gauge potentials. This formulation appears to offer several advantages. Some, for instance the continuity of these potentials (as opposed to the discontinuous electric field, for example), are well known. Others, however, do not appear—at least explicitly—to have been previously examined. For frequency-domain studies, we note that the gauge formulation can be written as a generic set of coupled elliptic equations, allowing a simple modular numerical approach using simple conservation principles. More importantly, by writing the governing equations using gauge potentials, an extra degree of freedom is introduced into the system. We show that this extra degree of freedom can be exploited to increase performance, accuracy, and simplicity. We have implemented the gauge approach in a 3-D numerical model which we call MOED (Model for Ocean ElectroDynamics). We validate the formulation as well as the software by comparing MOED results with several analytical and numerical results taken from oceanographic and geophysical applications.

Key words: electrodynamics, electromagnetics, induction, magnetohydrodynamics, numerical methods, ocean.

1 INTRODUCTION
For calculations involving the classical electrodynamics of conductive media, various formulations describing these dynamics are available. These range from the standard formulation of Maxwell (Maxwell 1891) to more exotic forms such as Weber electrodynamics (Assis 1994). The formulation we will discuss involves the ‘gauge potentials’, which are also known as the ‘scalar electric and vector magnetic’ potentials. We feel that for numerical problems these potentials offer subtle but important advantages that have not been fully recognized.

One advantage of the gauge potential formulation that is immediately obvious is that the gauge potentials deliver more complete information on the electromagnetic (EM) fields. (As an example, the gauge potentials include a description of the electric potential, which is not immediately available from the electric and magnetic field vectors.) The gauge potentials are also continuous and typically smoother than the electric or magnetic fields, possibly affording higher stability in numerical solution schemes.

More subtle, however, is the following. The gauge formulation typically has an extra degree of freedom in the system—the freedom in the specification of the gauge. For many practical problems the gauge can be specified somewhat arbitrarily, and this introduces into the numerical representation of a problem a degree of freedom that could be exploited. In the finite-difference method that we discuss, for example, there is also a degree of freedom in choosing how to represent the parameters and variables on one or more grids. The subtle advantage of the gauge formulation is that these two sets of specifications can be chosen together to increase accuracy and/or performance. While in a Maxwell formulation, for example, the easiest or most efficient gridding scheme may make it difficult during discretization to obtain a conservative representation of the partial differential equations (PDEs), in the gauge formulation the unspecified gauge is available to help out. One might start, for example, by choosing the simplest grid, and then, during discretization of the PDEs, a thoughtful specification of the arbitrary gauge might assure conservation of important physical quantities in the numerical representation. A ‘conservative’ representation of the PDEs can
be understood to be a numerical scheme that is consistent with the PDEs not only on a point-by-point basis but also when integrated over arbitrary volumes. Underlying the desired preservation of these integrals is typically a conservation principle such as the conservation of charge.

The gauge formulation, with its exploitation of the gauge freedom, forms the basis of the Model for Ocean ElectroDynamics (MOED), which we have developed for use in various applications in the oceanic, atmospheric, and geophysical sciences. Our primary applications have involved the generation and propagation of EM fields in and around the ocean. Cases with a variety of artificial and natural sources have been considered, including the generation of EM fields by the flow through magnetohydrodynamic interaction between the ocean flow and the Earth’s main magnetic field; EM fields due to controlled and stray (e.g. due to the ubiquitous 50/60 Hz power grids, electric trains, vessel corrosion currents, etc.) artificial sources; and the propagation into the ocean of EM fields due to sources in the atmosphere, ionosphere, and magnetosphere (e.g. Schumann resonant fields, telluric fields). In a recent application, we used MOED routines to calculate the magnetic fields at satellite altitude due to the global ocean M2 tides, and the MOED predictions were verified with observations (Tyler et al. 2003). This paper does not describe further details of these various applications, but primarily describes the formulation and shows why it is generally well suited to such applications in the earth sciences. Specific examples are discussed only for validation purposes.

Realistic 3-D simulations of the processes described above are usually not tractable analytically, and numerical schemes have been sought using various techniques (Wannamaker et al. 1984; Mackie et al. 1994; Everett & Schultz 1996; Wannamaker & Zhdanov 1999), including finite-difference, finite-element, finite-volume, and integral-equation approaches. It is evident from the outset that no single method will be optimal for all applications, and that there is some trade-off between efficiency, versatility, and ease in configuring. The goal of this paper will be to support the assertion that the numerical approach we present based on the free gauge approach is generally well suited for such applications in the earth sciences. Specific examples are presented only for validation purposes.

The use of the gauge formulation in numerical EM applications is not new. It has been used in engineering applications for quite some time (Biró & Preis 1989, 1990; Groiss et al. 1996; Biró & Preis 2000). More recently, it has also been used in geophysical applications (Everett & Schultz 1996; Tyler & Sanford 1998; Weiss & Everett 1998; Badea et al. 2001; Everett et al. 2001). A recurring advantage noted in these studies is that the gauge potentials can be made to be continuous across material interfaces. In contrast, formulations involving the electric field vector must contend with possible discontinuities in this variable.

MOED is specifically a 3-D, conservative, finite-difference, frequency-domain numerical model for calculating the EM fields generated by arbitrary sources which include zero-frequency ‘DC’ as a special case. MOED models the entire domain (which may include insulators as well as conductors) in the same way, and solves the same set of equations throughout the domain. This can of course be adjusted for efficiency but the default is to follow this simple and versatile approach. The gridding system allows arbitrary orthogonal grids to be used as the discretized formulation includes consideration of the grid metrics in a conservative finite-volume sense. The electrical conductivity in the domain may vary widely to represent insulators (the conductivity of air, for example, is about 14 orders of magnitude lower than that of sea water) as well as conductors. In the present version, however, the conductivity must be a scalar field rather than a tensor. We have retained the displacement and polarization currents in the formulation and numerical implementation despite the fact that in the applications that we consider conduction currents dominate. Our reason for doing this is that the numerical formulation then conserves electric charge exactly. Moreover, it is included at negligible numerical or programming cost. Hence, MOED may be applied to high-frequency cases as well as to the low-frequency cases we consider. The only current limitation is that the electrical parameters are assumed to be 3-D scalar fields rather than tensors (which might be needed in space physics applications). MOED comprises a collection of MATLAB modules. By default, the whole process of configuration, solution, and plotting of results can be done in the MATLAB environment and the model is therefore very easy to use. Alternatively, the user can regard MOED simply as an interface for configuring and plotting the problem. That is, the primary effort in MOED is in building the linear system of equations. Finding solutions to the linear system (inverting a large sparse matrix) can of course be done outside MATLAB using other software if this is considered to be more practical. By default, MOED will solve the linear system within the MATLAB environment and it allows a choice of various solution methods including direct inversion and iterative Krylov-space methods. Hence, MOED does not sacrifice numerical performance for the versatility and simplicity of coding in the MATLAB environment. The MOED software together with a user’s manual and validation examples can be obtained at http://sirena.apl.washington.edu.

In the next section we present the gauge formulation and show that after some manipulation the governing equations can be written as a simple set of coupled elliptic equations all having the same generic form. This simplifies the numerical implementation greatly as it allows the model to be written in a modular form and also allows an efficient block structure of the coefficient matrix to be exploited. In later sections we discuss issues related to the numerical discretization of these equations, and finally we present several examples serving towards an exhaustive validation of the method and numerical software.

2 FORMULATION OF GOVERNING EQUATIONS

2.1 Maxwell equations

The governing equations for macroscopic averages of the electric and magnetic fields in the presence of non-magnetic, non-relativistic, possibly fluid material can be written as

\[ \nabla \cdot \mathbf{B} = 0, \] (1)

\[ \nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0, \] (2)

\[ \nabla \times \mathbf{B} - \mu_0 \mathbf{J} - \mu_0 \mathbf{E} = 0, \] (3)

\[ \nabla \cdot (\mathbf{J} + \epsilon \partial_t \mathbf{E}) = 0, \] (4)

\[ \mathbf{J} = -\sigma (\mathbf{E} + \mathbf{u} \times \mathbf{B}) = 0, \] (5)

where \( \mathbf{B} \) (T) is the magnetic field, \( \mathbf{E} \) (V m\(^{-1}\)) is the electric field, \( \mathbf{J} \) (A m\(^{-2}\)) is the electric current density, \( \mathbf{u} \) (m s\(^{-1}\)) is the velocity of the medium, \( \mu_0 = 4\pi \times 10^{-7} \) (H m\(^{-1}\)) is the vacuum permeability, \( \epsilon \) (F m\(^{-1}\)) is the dielectric constant, and \( \sigma \) (S m\(^{-1}\)) is the electrical conductivity. We allow \( \sigma \) to be a general function of time and space, but as the displacement current effects typically occur on timescales much more rapid than that of the environmental fluctuations, we assume that \( \epsilon \) is a function of space but not time.
Assuming that the flow does not depend on the EM fields, Ohm’s law (5) is simply a diagnostic equation for J, and the other equations are essentially the Maxwell equations. The two scalar equations are not independent of the vector equations, as (4) can be obtained by taking the divergence of (3), and (1) is similarly evident from (2) and the assumption that ∇ · B vanished once. Hence, the system to be solved above is essentially the two vector equations (2, 3) for two vector variables. The set then constitutes six equations for six unknowns that must be solved with appropriate boundary conditions.

2.2 Motional and controlled sources

Let us now describe the EM sources. Throughout this work we will make the assumption that the motional induction source term ξ u × B ≈ u × F, where both the velocity and the main field F are prescribed. This requires that the EM forces on the flow are negligible and that the amplitude and spatial gradients of the magnetic fields generated are small compared with those of the main field. For controlled sources using an impressed electric field E, or electric current Jc, one can make the substitution u × B → E, = Jc/σ. To allow for either or both possibilities, let us define the motional or controlled electric current:

$$J_p = \sigma u \times F + J_c.$$  

(6)

Hereafter, Jp is used to represent the prescribed current due to either motional or impressed sources.

2.3 Helmholtz potentials

Under natural boundary conditions, a vector is fully determined once its divergence and curl are specified. The associated Helmholtz decomposition of a vector a has the form $a = \nabla \times c - \nabla d$, where c and d are vector and scalar potentials to be determined.

A general decomposition for the EM variables in terms of the Helmholtz potentials, denoted by A, \(\psi\), G, \(\eta\), is

$$B = \nabla \times A - \nabla \psi,$$

(7)

$$E = \nabla \times G - \nabla \eta.$$  

(8)

These expressions can be used to replace the variables E, B in the above equations. Note, however, that in so doing we will replace six unknowns with eight unknowns. We will therefore need eight equations. The two ‘extra’ equations we must specify reflect our freedom to specify arbitrarily ∇ · A and ∇ · G without affecting the validity of any of the above equations. One approach, which we will follow, is to retain the two non-independent equations (1, 4) and impose the arbitrary choices for ∇ · A and ∇ · G as conditions such that each of the governing equations in the set (1) and (4) becomes independent. Substituting using (7) and (8), the set (1) to (4) becomes

$$\nabla^2 \psi = 0,$$

(9)

$$\nabla^2 G - \nabla(\nabla \cdot G) - \nabla \times \nabla \psi = 0,$$

(10)

$$\nabla^2 A - \nabla(\nabla \cdot A) + \sigma_0 \sigma(\nabla \times G - \nabla \eta)$$

$$+ \sigma_0 \sigma \nabla \psi = 0,$$

(11)

$$\nabla \times (\sigma \nabla \eta + \sigma_0 \sigma_0 \nabla \eta - \sigma \nabla \times G - \mu_0 \sigma \nabla \psi + \psi \nabla \eta - \nabla \nabla \psi - \nabla \psi = 0.$$  

(12)

If the formulation in terms of potentials is useful, it is typically because we can make certain choices for ∇ · A and ∇ · G or the free boundary conditions that uncouple the equations in the set above. Under assumptions of a uniform medium and source-free static conditions, for the simplest example, one can extract from the set above simple Laplacian equations for the magnetic and electric scalar potentials. More relevant to what follows, let us specify $\psi = 0$ on the boundaries (this ‘natural’ boundary condition follows the assumption of vanishing fields in the far field, and the specification of this boundary value is regarded as an aspect of specifying the divergence of A).

The only solution that satisfies (9) then is $\psi = 0$. Then, using $\nabla \times \nabla \psi = -\nabla^2 G + \nabla(\nabla \cdot G)$, (10) can be written as

$$\nabla \times G + \partial_t A + \nabla \psi = 0,$$  

(13)

where $\psi$ is a new scalar. Substituting to remove G in (9)–(12), a new set of equations can be written:

$$\nabla^2 \phi = -\partial_t \nabla \cdot A,$$  

(14)

$$\nabla^2 A = \mu_0 \varepsilon_0 \nabla \phi - \nabla(\nabla \cdot A) - \nabla \phi \nabla \phi$$

$$- \mu_0 \varepsilon_0 \nabla \psi = 0,$$  

(15)

$$\nabla \cdot (\sigma \nabla \phi + \mu_0 \varepsilon_0 \nabla \phi + \partial_t \nabla A + \mu_0 \varepsilon_0 \nabla \psi - J_p) = 0,$$  

(16)

with the magnetic B and electric E fields now given as

$$E = -\partial_t A - \nabla \psi$$  

(17)

and

$$B = \nabla \times A.$$  

(18)

where $\phi = \xi + \eta$. Eq. (14) can be regarded as simply a diagnostic equation for $\psi$, as the other equations are uncoupled from $\xi$. The same is now true for the variable G, which is governed by (13) but is no longer needed in calculating any other variables. Therefore, we do not need to calculate $\xi$ or G. In summary, we have distilled the two governing equations (15) and (16), which constitute four equations in four unknowns. The scalar equation (16) is, however, not independent, as it can be obtained from the divergence of (15). By defining $\nabla \cdot A$ and making substitutions in one of the equations, the equations become independent. This definition specifies the gauge and closes the freedom introduced by using the Helmholtz potentials. In the following subsection, we discuss the ‘damped Lorentz’ gauge, while in Appendix A we discuss the advantages of alternative gauge choices.

2.4 ‘Damped Lorentz’ gauge potentials

As described in (15) and (16), the equations of electrodynamics for a moving conductive medium can be written in terms of the potentials (A, $\phi$). So far, however, these potentials are not unique. Once we have prescribed the divergence of A, we will refer to these potentials as ‘gauge’ (or better would be ‘gauged’) potentials. As we asserted in the Introduction, the optimal choice of the gauge depends on both mathematical and numerical considerations. For the co-located grid that will be discussed in this paper, the following ‘damped Lorentz’ gauge will be used:

$$\partial_t \phi + c^2 \nabla \cdot A + c^2 K^{-1} \phi = 0.$$  

(19)

Above, $c = (\mu_0 \varepsilon_0)^{-1/2}$ is the speed of light in the material medium, $K = (\mu_0 \varepsilon_0)^{-1}$ is the magnetic diffusion coefficient, and we will refer to $\phi$ and A as either the scalar and vector gauge potentials, or as the electric potential and magnetic vector potential, respectively. With the above gauge, the governing equation (15) for A can be written as

$$\partial_t A = K \nabla^2 A + K(\phi \nabla K^{-1} + \partial_t \nabla c^{-2}) + J_p/\sigma.$$  

(20)

© 2004 RAS, GJI, 158, 874–887
where $\Box^2$ is the D’Alembertian wave operator $\Box^2 = \nabla^2 - c^{-2} \partial_t \partial_t$.
Eqs (19) and (20) form the governing equations for the electrodynamics. By exercising our freedom in choosing a form for the gauge, we have given the governing equations the particular form above, which, as we will see, has certain advantages. To give some initial insight into the governing equations above we briefly consider their form in two limiting cases.

### 2.4.1 Gauge potentials for good conductors

We can make an important simplification in good conductors based on assuming that the squared ratio of the diffusion speed $K/L$ (where $L$ is a typical length scale for the EM fields) to the speed of light $c$ in the medium vanishes, while also assuming that any gradients $\nabla c^{-2}$ remain finite. In this case, the gauge (19) simplifies to

$$\nabla \cdot \mathbf{A} + K^{-1} \phi = 0,$$

while the vector potential equation (20) becomes

$$\partial_t \mathbf{A} = K \nabla^2 \mathbf{A} + \mathbf{J}_p / \sigma + \phi \nabla \ln \sigma.$$  \hspace{1cm} (22)

Or, using (21) to substitute for $\phi$ and (18),

$$\partial_t \mathbf{A} = K \nabla^2 \mathbf{A} + \mathbf{J}_p / \sigma - K \nabla \cdot \mathbf{A} \nabla \ln \sigma,$$

which, provided $u$ is prescribed, is an equation entirely in terms of one vector variable $\mathbf{A}$.

### 2.4.2 Gauge potentials for poor conductors

In the limit $\sigma \to 0$ ($K \to \infty$), the gauge (19) returns to the usual Lorentz gauge [the last term in (19) vanishes], and (20) becomes simply a wave equation depending on the dielectric properties of the medium through $c$ and its gradients.

### 2.5 Governing equations for the frequency domain

We will assume here that sources and fields have a time dependence of the form $e^{-i\omega t}$, where $\omega$ is the frequency and it is understood that we take only the real part of any solutions as physical. By defining a complex conductivity

$$\sigma^* = \sigma - i \omega \epsilon,$$  \hspace{1cm} (24)

we can write the damped Lorentz gauge as

$$\nabla \cdot \mathbf{A} + \mu_0 \sigma^* \phi = 0,$$  \hspace{1cm} (25)

which applied to eqs (15) and (16) gives

$$\nabla^2 \mathbf{A} + i \omega \mu_0 \sigma^* \mathbf{A} + \phi \mu_0 \nabla \sigma^* + \mu_0 \mathbf{J}_p = 0,$$  \hspace{1cm} (26)

$$\nabla \cdot (\sigma^* \nabla \phi) + i \mu_0 \sigma^* \omega \sigma^* \phi - i \omega \mathbf{A} \cdot \nabla \sigma^* - \nabla \cdot \mathbf{J}_p = 0.$$  \hspace{1cm} (27)

In this case, all four of the vector-component or scalar EM equations can be written in one form:

$$\nabla \cdot (D \nabla \mathbf{A}_\mu) + a A_\mu + p_\mu + f_\mu = 0,$$  \hspace{1cm} (28)

where the 4-form $A_\mu = (A, \phi)$ includes all four of the gauge potential components, and the coupling and forcing vectors are, respectively,

$$p_\mu = \langle \phi \mu_0 \nabla \sigma^*, -i \omega A \cdot \nabla \sigma^* \rangle,$$  \hspace{1cm} (29)

$$f_\mu = \langle \mu_0 J_p, -\nabla \cdot \mathbf{J}_p \rangle.$$  \hspace{1cm} (30)

The coefficients for $\mu = 1, 2, 3$ are $D = 1$ and $a = i \omega \mu_0 \sigma^*$; and for $\mu = 4$ they are $D = \sigma^*$ and $a = i \mu_0 \sigma^* \omega \sigma^*$.

The general form of the governing equation (28) is also retained under other gauge choices. The coefficients and terms for three different gauges are given in Table 1: the damped Lorentz gauge ($\nabla \cdot \mathbf{A} + \mu_0 \sigma^* \phi = 0$); the usual, sometimes termed ‘high-frequency’, Lorentz gauge ($\nabla \cdot \mathbf{A} - i \omega \mu_0 \phi = 0$); and the Coulomb gauge ($\nabla \cdot \mathbf{A} = 0$).

### 3 Discretization of Partial Differential Equations and Solution of the Linear System

#### 3.1 Basic flow of operations

In the last section we showed that a sufficient set of governing equations describing electrodynamics in the frequency domain can be written as the coupled set of elliptic equations (28) which all have the same form (when the coupling term $p_\mu$ is ignored initially or taken as prescribed). Beyond simple mathematical elegance, this symmetry simplifies the coding in MOED and also allows maximum versatility, as MOED can obviously be used to solve problems outside electrodynamics as well. For this reason, care is taken in separating kernel operations in MOED that treat the PDEs in their generic form (28) from the parts that are more application-dependent.

The basic flow of the operations in the model essentially involves three steps: (1) pre-processing routines convert the prescribed physical parameters to the needed coefficients in the generic system (28); (2) kernel routines solve the generic system; (3) the generic solutions are passed through the post-processing routines to generate physical solutions. The important point is that the only routines that ‘know’ anything about the specific application (e.g. EM) are contained in the pre- and post-processors. This provides a conceptually clean and numerically efficient and versatile separation of tasks in the model. Importantly, this also allows a modularization of the numerical routines consistent with the modern approach in numerical modelling.

---

© 2004 RAS, GJI, 158, 874–887

---

**Table 1.** Coefficients and terms for the generic governing equation (28) under various gauges. The first column gives the name of the gauge choice, the second gives the equation number(s) $\mu = 1, 2, 3, 4$ on which the coefficients are defined in the remaining columns.

<table>
<thead>
<tr>
<th>Gauge</th>
<th>$\mu$</th>
<th>$D$</th>
<th>$a$</th>
<th>$p_\mu$</th>
<th>$f_\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damped Lorentz</td>
<td>$\mu = 1, 2, 3$</td>
<td>$1$</td>
<td>$i \omega \mu_0 \sigma^*$</td>
<td>$\mu_0 \mathbf{J}_p$</td>
<td>$\mu_0 \mathbf{J}_p$</td>
</tr>
<tr>
<td>$\mu = 4$</td>
<td>$\sigma^*$</td>
<td></td>
<td>$-i \omega A \cdot \nabla \sigma^*$</td>
<td>$-\nabla \cdot \mathbf{J}_p$</td>
<td></td>
</tr>
<tr>
<td>Lorentz</td>
<td>$\mu = 1, 2, 3$</td>
<td>$1$</td>
<td>$i \omega \mu_0 \sigma^*$</td>
<td>$-i \omega \mu_0 \nabla (\epsilon \phi) - \mu_0 \sigma^* \nabla \phi$</td>
<td>$\mu_0 \mathbf{J}_p$</td>
</tr>
<tr>
<td>$\mu = 4$</td>
<td>$\sigma^*$</td>
<td></td>
<td>$0$</td>
<td>$-i \omega \nabla \cdot (\sigma^* \mathbf{A})$</td>
<td>$-\nabla \cdot \mathbf{J}_p$</td>
</tr>
<tr>
<td>Coulomb</td>
<td>$\mu = 1, 2, 3$</td>
<td>$1$</td>
<td>$i \omega \mu_0 \sigma^*$</td>
<td>$-\mu_0 \sigma^* \nabla \phi$</td>
<td>$\mu_0 \mathbf{J}_p$</td>
</tr>
<tr>
<td>$\mu = 4$</td>
<td>$\sigma^*$</td>
<td></td>
<td>$0$</td>
<td>$-i \omega \nabla \cdot (\sigma^* \mathbf{A})$</td>
<td>$-\nabla \cdot \mathbf{J}_p$</td>
</tr>
</tbody>
</table>
3.2 Discretization of the governing partial differential equations

The set of governing equations (28) is discretized using the finite-difference method. First the domain is discretized by creating a collection of grid nodes on which parameter and variables are defined. The values on the nodes then represent the average value over a local volume around each grid node. The specified grid defines the location of these nodes, and we will refer to the collection of nodes thus created as simply the ‘grid’.

It is not necessary that all parameters/variables be defined on the same grid. Indeed, one can often obtain increased efficiency or accuracy by using more than one grid. Optimal use of these so-called ‘staggered grids’ typically depends on the application. One must consider not only the representation of the PDEs (e.g. preservation of physical conservation principles), but also how the input parameters are defined. Obviously, if the forcing function, for example, is an analytically prescribed function then this can be described on any grid. More generally, however, one may need to consider interpolation errors introduced while bringing the input data onto the chosen grid. More generally, however, one may need to consider interpolation errors introduced while bringing the input data onto the chosen grid(s), as well as whether the resulting discretized equations are consistent with the original governing equations. If calculation of post-processed variables is required, the choice of grid may also be important in that it controls the locations at which the meta-solutions are available.

As described earlier, the formulation in MOED anticipates the potential for exploiting the gauge in one or more of its possible forms. Here we discuss only the simplest ‘co-located’ grid in which all quantities are defined on the same grid and the damped Lorentz gauge is chosen for suitability. In this implementation, the method we use may also be considered to be a finite-volume method because control surfaces around the grid nodes are considered when discretizing the PDEs to obey physical conservation principles. These control surfaces are neither labelled nor stored, however, but are simply calculated in reference to the grid nodes (see Fig. 1). The first term in (28) is discretized, for example, by integrating over the control volume and using Gauss’s Theorem. The fluxes involve first differences, which are suitably calculated at the half gridpoint locations by differencing values at adjacent grid nodes. In this sense, we find it appropriate to refer to the method as a ‘conservative finite-difference method’. As we have stated previously, the discretization in MOED allows for general curvilinear coordinates and the control volumes can therefore be more distorted than in Fig. 1.

We see that the result of the discretization is that functions of space $f(x)$ are represented by discrete functions of the grid nodes $f(i,j,k)$. For describing directional orientation in the grid we will associate the directions ‘east’, ‘north’, ‘up’ with positive increases in $i$, $j$, $k$ respectively.

3.3 Discretization of boundary conditions

As opposed to the generic form (28) we found for PDEs, the form for the boundary conditions can be highly application-dependent. The inclusion of arbitrary boundary conditions involves, of course, representing boundary equations in the coefficient matrix. As a convenient versatile default, MOED uses the general mixed boundary condition of the form

$$C_0^{(μ)} \frac{∂}{∂x_μ} A_μ = C_1^{(μ)} A_μ = C_2^{(μ)}, \quad (31)$$

where $∂/∂x_μ$ refers to the outward normal derivative, and one chooses the various coefficients $C$ appropriate for each of the $μ = 1, 2, 3, 4$ components. The primary restriction of this approach is that it assumes that the boundary conditions are uniform (although possibly different) on each of the boundary faces. The most physically appropriate boundary conditions are typically natural boundary conditions ($A_μ → 0$ in the far field), and this can be achieved by selecting $C_1 = 1$ and $C_2 = 0$. Often, however, it is either impractical to include an appropriately large domain to apply this condition, or the boundary condition itself may be driving the solution. Moreover, one may often know an appropriate boundary condition in terms of $E$, $B$ or $J$ and want to choose appropriate coefficients in (31) reflecting this.

![Figure 1. Control volume around node $(i,j,k)$. The control surfaces are defined at points midway between gridpoints.](https://academic.oup.com/gji/article-abstract/158/3/874/553518/fig1)
The boundary equation (31) is discretized by approximating $\partial_n$ at the boundary points using a backward difference. Hence, the discretized (31) produces equations involving linear combinations of the unknowns located at boundary nodes and nodes located one gridpoint step towards the interior.

### 3.4 Construction of the coefficient matrix for the linear system

The discretization of (28) together with the boundary equation (31) produce a sequence of algebraic equations that can be sorted arbitrarily. We can order the four unknowns based on the vertex (i.e. we can order all the four unknowns for one node and then move on to the next node), or we can use a lexicographic scheme in which we order one unknown for all the nodes and then move to the next unknown. We adopt the latter scheme as this produces a block structure. The block structure of a matrix can be valuable for very large problems where memory limitations become evident. Some operations on the whole coefficient matrix (described below) can be replaced by operations on individual blocks, making larger problems accessible.

Using the lexicographic order scheme, and a logical sequencing of the nodal values following the permutation $i, j, k$, the discretized solution vector has the ordering: $A(i = 1, j = 1, k = 1), A1(i = 1, j = 1, k = 1) . . . A_4(i = M, j = N, k = P - 1), A_4(i = M, j = N, k = P)$. A similar sorting is used to create the discretized force vector $f$. The total number of grid nodes is $MN$, where $M$, $N$, and $P$ are sequentially the number of gridpoints along the $i, j, k$ directions.

Under such an arrangement, equations on all the nodes can be assembled as the sparse linear system

$$
\begin{bmatrix}
\nabla(D_1\nabla) + a_1 & 0 & 0 & p_1 \\
0 & \nabla(D_2\nabla) + a_2 & 0 & p_2 \\
0 & 0 & \nabla(D_3\nabla) + a_3 & p_3 \\
p_{41} & p_{42} & p_{43} & \nabla(D_4\nabla) + a_4 \\
\end{bmatrix}

\times

\begin{bmatrix}
f_1 \\
f_2 \\
f_3 \\
f_4 \\
\end{bmatrix}

= \begin{bmatrix}
f \\
f \\
f \\
f \\
\end{bmatrix},

(32)

where $p_{4i}$ is the $i$th component of $p_4$.

Once discretized, the system (32) can be written in an abbreviated form as $GA = f$ where the coefficient matrix $G$ is typically a very large sparse matrix with complex entries. The size of $G$ depends on the number of nodes and unknowns. Assuming, as in our examples, that we are solving four PDEs, the size of $G$ is $(4MNP)^2$. The sparsity pattern of the matrix depends on how the unknowns are ordered. Using the lexicographic order, $G$ can be dissected into small blocks. The matrix $G$ is sparse but not necessarily diagonally dominant or symmetric. It can be immediately appreciated that increases in the amplitudes of the coupling terms $p_n$ will typically decrease diagonal dominance.

### 3.5 Pre-conditioning

An important part of the Krylov iterative methods for solving the linear system (to be described in the next section) is to first apply a suitable pre-conditioner to $G$. These iterative methods can show very slow convergence when $G$ is ill-conditioned, and a good pre-conditioner will improve the situation. Among the many preconditioners we examined, the block incomplete LU pre-conditioner of Barrett et al. (1994) is considered to be a good candidate for our system.

Conceptually, the blocks we choose stem from considering the individual linear systems that would arise if the coupling terms $p_n$ were neglected. For each $\mu$, a block represents the discretization of the operators $\nabla(D_n\nabla) + a_n$ in (32) and is limited to a $2 \times M \times N$-wide diagonal band. The advantage is that the incomplete LU (ILU) decomposition is then performed more quickly on the individual blocks, with the decomposed parts eventually reassembled to the full size of $G$. Aside from the advantage of speed, this method allows a decomposition for large problems that would normally be beyond the memory capacity of the computer system. The incomplete LU decomposition requires choices for the degree of ‘incompleteness’. Typically, we use the ILU method with zero fill in, which is faster and uses less memory, but higher fill in becomes more appropriate for cases where $G$ is not diagonally dominant.

### 3.6 Solving the linear system

The tasks so far have involved casting the appropriate equations in the mutually common form (28) and discretizing these PDEs to replace them with a linear system of algebraic equations. The next step is the somewhat generic task of finding the solution to the linear system $G \cdot A = f$. This step can of course be approached with a wide variety of methods and need not necessarily be done in the MATLAB environment. In this sense, the fact that MOED is written in MATLAB is of very little relevance when considering the efficiency of solving the linear system as this can be done by exporting $G$ and $f$ and using any desired routine to solve for $A$. The primary responsibility of MOED is in interpreting a physical problem to deliver the linear system described by $G$, and this is done quite conveniently in the MATLAB environment.

Which method is the ideal one for solving a linear system can usually only be decided formally for simple matrices having special properties, and much of the progress on linear system solution methods has been made through some degree of trial and error, and with results that are specific to the behaviour (i.e. locations and relative values) of the entries in the coefficient matrix. The only general features we can predict for $G$ that are not case-dependent are that it is very sparse, has a blocked structure, is possibly complex, and by construction is diagonally dominant when the coupling terms are ignored.

In working with linear systems (derived from similar problems) having a structure similar to our $G$, Aruliah et al. (2001) compared the performance of several pre-conditioned Krylov-space methods. Their study showed that BICGSTAB (BiConjugate Gradients Stabilized Method) performs better than GMRES (Generalized Minimum Residual Method) or QMR (Quasi-Minimal Residual Method). Remaining within the MATLAB environment (as described above, this is convenient but not necessary), we have tested BICGSTAB with the block ILU pre-conditioner on our system and found that a fast convergence rate can always be achieved when the matrix is much smaller than (less than 1/5 of) the system memory. However, BICGSTAB with the block ILU may require more memory than the system can provide when the size of the matrix is close to 1/4 of the system memory. While this is clearly a limitation of MATLAB, and not of the algorithm itself, it is also understood that, for very large problems, the memory constraints become an
issue for any environment. For the largest problems, we found success using the remarkably simple ILU solution method, which has previously been used Everett & Schultz (1996) in large geophysical problems discretized using finite elements. Compared with the pre-conditioned BICGSTAB, the ILU solver can solve much bigger systems (3 times as big), and has been shown to be very stable for many ill-conditioned systems. However, the ILU solver is much slower than the BICGSTAB solver. To reach a specified error tolerance, the ILU solver may need 100 times more computing time than the BICGSTAB solver. Our current practice for large problems is, then, to use the pre-conditioned BICGSTAB solver unless \( G \) is larger than 1/5 of the computer memory, in which case we use the ILU solver.

Although the ILU solver can handle large, ill-conditioned systems, it may converge to an incorrect solution. A good initial solution ‘guess’ for the system can not only reduce the time to reach a specified error tolerance but it can also provide a more reliable convergence. Moreover, a good initial guess can provide the system with accurate solutions on the boundaries and further constrain the solution space. To obtain a good initial solution for the ILU solver, we can use a multi-grid technique Trottenberg et al. (2001). In typical 3-D applications, we start by considering an equivalent configuration with only one axis of variations (a layered 1-D model) and solve the fields for this model on a very coarse grid. This can be done quickly since we do not need high resolution horizontally, and fast solvers can be deployed. After we have accurate solutions for the 1-D model, we interpolate these solutions to the real 3-D model and use the interpolated solution as the initial solution for the ILU solver. By doing this, we provide the ILU solver with an initial guess that is close to the true solution—particularly near the boundaries.

3.7 Post-processing

The generic solutions \( A_i \) are delivered to the post-processor, which ascribes the physical interpretation to the solutions and does any needed post-calculations. In our application, the generic solutions are first interpreted physically as the magnetic vector and electric potentials \( (A = A_{1,2,3} \text{ and } \phi = A_4) \). Then \( E \) and \( B \) are calculated through differentiation of the potentials following (17) and (18). This is most easily performed using direct first-order differencing. In making estimates for \( E \) and \( B \) using only first-order differences, however, we may be ‘short changing’ the intrinsic second-order accuracy of the potentials. Indeed, the accuracy of the estimates for \( E \) and \( B \) depends not only on the accuracy of the gauge potential solutions obtained but also on the post-processing algorithm used in calculating \( E \) and \( B \). Luckily, there are some options available in choosing how \( E \) and \( B \) are calculated from the potentials, and we have strived to make the choices that retain the most accuracy.

In our experience, the greatest challenge has been in selecting a method for calculating \( E \), as this field can be discontinuous when the media involves (as is typical) discontinuities in the conductivity. As will be seen with the ‘three-segment case’ described in the next section, we have found a method that gives accurate estimates for \( E \) even at the material discontinuities where \( E \) is discontinuous. Using (25) we compute the term \( \nabla \phi \) in (18) using second-order differencing as

\[
\nabla \phi = \frac{\nabla(\nabla \cdot A_i) + \mu_i \phi \nabla \sigma^* \times \mu_i \sigma^*}{-\mu_i \sigma^*}. \tag{33}
\]

4 E X A M P L E S

In this section, we first present three examples that have been chosen for validation purposes rather than for realism or physical content. The first and second examples test the accuracy of MOED in simple 1-D and 2-D configurations for which we have analytical solutions for comparison. The third example (a conductive block illuminated by EM plane waves) represents a simple 3-D case that has been thoroughly studied numerically in previous work using a standard Maxwell formulation. A fourth example is included to illustrate the use of MOED in a global domain.

4.1 Plane wave incident on uniform ocean

We first present the case of plane waves propagating down from the atmosphere into a uniform, infinitely deep ocean. Assuming that a horizontally directed magnetic field \( B_o \) is prescribed on the upper boundary, the analytical solutions for the magnetic and electric fields are

\[
B_z = B_o e^{(1-i\gamma)/\delta}, \tag{34}
\]

\[
E_z = -\left( \frac{\omega}{2 \sigma \mu} \right)^{1/2} (1-i) \hat{\gamma} \times B_o e^{(1-i\gamma)/\delta}, \tag{35}
\]

where \( \delta = \left( \frac{2}{\omega \sigma \mu} \right)^{1/2} \) is the skin depth.

To model this case, we consider a \((10 \text{ km})^2\) cube of ocean with a uniform ocean conductivity of \( 4 \text{ S m}^{-1} \). The upper boundary conditions are \( \partial_z A_{(i=0)} = B_o \times \hat{\gamma} \), where \( B_o \) is the horizontal magnetic field prescribed at the top boundary [which we will take to be \( B_o = 1 \times 1\text{S(T)} \)], and the physical condition \( J_z = 0 \) can be satisfied by \( A_z = 0 \). On the horizontal boundaries \( \partial_x A_x = 0 \) because of the lateral invariance of the plane wave. On the bottom boundary we take \( \partial_z A_z = A_z = 0 \), while for the components \( A_1, A_2 \) we use the condition for downward propagation, for example \( \partial_z A_1 - 1/\sigma A_1 = 0 \). Because of the horizontal invariance, we can reduce the number of horizontal gridpoints to the minimum \((3 \times 3)\) permissible with Neumann boundary conditions. An adequate resolution for the purposes here is given with 50 vertical gridpoints with a concentration of points near the sea surface.

Comparison of our numerical results and analytical solution for two frequencies are shown in Figs 2 and 3, and an excellent agreement is found.

As a next step of sophistication (while still affirming an analytical solution for comparison—see Appendix B), we include an abrupt transition in electrical conductivity. We extend the last case to include air and sediment layers in addition to the ocean layer. \( B_o \) is now imposed at 10 km above sea level, and the air conductivity is set to \( 10^{-11} \text{ S m}^{-1} \). We use \( 0.1 \text{ S m}^{-1} \) to represent the sediment 5 km below sea level, and use the natural boundary condition (i.e. \( A_z = 0 \)) at the bottom of the model under the reasonable assumption that little energy reaches the lower boundary. As this 1-D case imposes no computational demands, a large number of vertical nodes (1000) can easily be used such that we can expect very good agreement even near the interfaces. In Fig. 4 the numerical solutions are seen to be essentially identical to the analytical solutions.

As shown in Fig. 4, the amplitude of the signal quickly attenuates in the ocean, with the phase progressing downwards. Numerical results show excellent agreement with the analytical solutions. Using formulae as, for example, in Tyler et al. (1998), we should expect 10 wavelengths in the ocean and 1.56 wavelengths in the sediment, which is also evident from the figure.
4.2 Three-segment model

The three-segment model is a classic 2-D example for evaluating numerical algorithms and has been adopted in many studies (Wannamaker et al. 1984; Weaver et al. 1985; Aprea et al. 1997). The model includes three adjacent conducting regions on top of an infinitely conductive basement, and analytical solutions are available for both the transverse electric (TE) and the transverse magnetic (TM) modes. In the TE mode, electric currents flow along these interfaces and a weak response is expected. In the TM mode, however, currents flow perpendicular to these interfaces. When electric currents pass through a conductivity interface, spatial charge densities are created near the interface, which can have dramatic effects on the solution. Moreover, because the TM mode provides a more demanding validation than the TE mode, we focus in this example on only the TM mode.

To make a comparison with the case described in Weaver (1994), we let \( \omega = 1 \), and then the second segment conductivity \( \sigma_2 \) is chosen...
4.3 3-D magnetotelluric example

In this section, we present a 3-D example. The configuration is shown in Fig. 6. A conductive block (0.1 S m⁻¹) is illuminated by downward propagating plane waves (e.g. generated by ionospheric/magnetospheric sources). We compare our results with numerical results from the algorithm described in Mackie et al. (1994) for the same test configuration. The downward plane waves are created by specifying the upper boundary conditions for $B_x$ and $B_z$ 50 km above the Earth's surface. Based on the dimensions of the conductive prism (6 km $\times$ 6 km $\times$ 3 km), we set the horizontal boundaries at 40 km from the centre (same for west, east, north and south). To discretize the domain, we use 41 gridpoints for the horizontal and 43 nodes for the vertical, with the grid spacing increasing exponentially away from the centre of the domain. This choice for the number of grid nodes reflects an attempt to deliver a domain resolution comparable to that of the Mackie study ($39 \times 39 \times 33$) when considering the implicitly higher resolution of the staggered grid approach.

To avoid the interference between the source field and the fields due to the inhomogeneity inside the Earth, the source needs to be placed far above the Earth’s surface. In Mackie et al. (1993) it is recommended that this distance be at least 75 km. We have experimented with the source distance and found no appreciable difference between a source at 50 km and one at 75 km, and present results for the 50 km case. To define a 3-D impedance, $B$ and $E$ fields from the different sources are required. We apply $B$, $B_z$ sources to the model and then combine these results.

The boundary conditions for this model are relatively simple. At the top of the model, we apply $A_z = 0$, $\partial_z \phi = 0$. Combining this condition with either $\partial_y A_y = 1$, $\partial_y A_z = 0$ or $\partial_y A_z = 0$, $\partial_y A_z = -1$, we produce a $B_y$ or $B_z$ source. On the horizontal boundaries, $\partial_y A_y = 0$ because of the lateral invariance of the plane wave. On the bottom boundary we apply natural boundary conditions, i.e. $A_\mu = 0$. As mentioned above, the natural boundary condition can only be used when the bottom of the domain is deep enough so that no significant energy will be reflected. This condition is met provided...
3-D modelling using gauged potentials

5

that the bottom boundary is located many skin depths below any sources. We model the response of four frequencies ranging from 3 Hz to 0.1 Hz. For the highest frequency (3 Hz) the skin depth in the 0.01 S m\(^{-1}\) host medium is about 2.9 km. In this case, we extend the bottom of the domain to a depth of 50 km (about 17 times the skin depth). With similar considerations, the bottom boundary for the low-frequency case (0.1 Hz; skin depth \(\approx 16\) km) is chosen to be 150 km.

We use a Dell Precision 340 workstation to simulate the MT response of the 3-D model. This machine has 1 GB of in-core memory and a 2-GHz CPU. Each run requires about 25 per cent of memory and lasts less than 1 hour. Use of a finer mesh and more iterations will certainly improve the results, but we prefer to illustrate the fact that the program can produce reasonable results with modest computing resources and a short time.

In this example, we will not discuss the detailed behaviour of the 3-D fields but merely compare the responses from simulations with different models. The apparent resistivity and phase response of this model are shown in Fig. 7. Although these two models are based on different formulations, their results agree very well. Except for the nodes near the edge of the prism, the differences in resistivity are less than 2 per cent, and the \(XY\) mode agrees better than \(XY\) mode.

Similarly, the phase responses from these two models also agree well. The phase differences for all four frequencies are less than 2\(^\circ\), and these differences are comparable to those reported by Mackie et al. (1993). Notice that the offset in phase occurs for 3 Hz and that our results are closer to the true phase (45\(^\circ\)).

4.4 Electric potential of ocean circulation

The examples we have presented have been chosen primarily for validation purposes and are therefore idealized configurations. To simply illustrate a more realistic case in spherical coordinates, we show in Fig. 8 the electric potential generated by the 1992–2002 mean ocean circulation of the ECCO ocean model (1-degree resolution adjoint version; for more information see: www.ecco-group.org). Similar cases have been previously calculated and discussed, and it is not necessary for the purposes of this paper to present the details of this calculation as the procedure we used here is much the same as in Tyler et al. (1997a,b), for example. We wish merely to indicate that the solution shown in Fig. 8 is similar to that shown in earlier work, and this then provides a modest validation for a global-scale case with motional sources.

5 DISCUSSION

In this paper we have asserted that there are advantages in using the gauge potential formulation in numerical calculations of geophysical (including solid earth, oceanic, and atmospheric) electrodynamics. Historically, the gauge potentials were often viewed simply as mathematically convenient intermediary variables devoid of physical content. Despite later experiments that indicated that this view is erroneous, there remained a reluctance to use the gauge potentials as—at least in macro-scale studies involving material media—the specification of the gauge was viewed as being arbitrary and therefore the gauge variables were not unique. We take an alternative view and assert that this non-uniqueness confers an advantage in numerical methods. The reasoning is that the freedom in choosing the gauge can be combined with freedom in other choices (such as
Figure 7. Comparison of MT responses of the conductive prism defined in Fig 6. In all the above plots, the solid line is the solution of Mackie3D and the circles are the solution of our program. Both the apparent resistivity and phase for four frequencies are shown.

Figure 8. Electric potential generated by mean ocean circulation. Contour interval is 0.1 V.

which type of grid to use) to optimize performance, accuracy, and simplicity in the numerical approach. Performance is increased primarily by choices that decrease the coupling terms in the governing equations—which usually has the effect of improving the condition number of the coefficient matrix. Accuracy is increased by choices that retain conservation of certain integrals. Simplicity is increased by choices that allow simplified grids (such as the maximally simple co-located grid we focus on) and simple forms for the governing equations that allow a modularized numerical approach.

We mainly discuss the combination of a co-located grid with the damped Lorentz gauge, but also outline (in an appendix) how a conservative approach can be retained when using a staggered grid if one adopts instead a Lorentz gauge. Use of a staggered grid will increase the implicit resolution per number of grid nodes, allowing...
larger problems to be tackled and/or higher resolution to be obtained. The trade-off is that it is more clumsy to use as the variables are assigned on different grids. Special care must be taken when interpolating variables between the grids if one wishes to retain the property conservations implicit in the native-grid solutions. A further consideration in choosing a gauge requires consideration of the input fields. For example, as we often use input forcing fields that come from ocean-flow model output data, an attraction of the gauge approach and gridding scheme we have used is that it is compatible (often using the same kinds of finite volume flux conservation principles) with the models of ocean flow.

We demonstrate these ideas in the numerical model MOED. MOED is designed to allow a variety of gauge choices but the default we present uses the co-located grid with theamped Lorentz gauge, which to our mind has the highest total score in performance, accuracy, and simplicity. We have validated MOED with several examples taken from oceanic and geophysical applications and see that MOED performs quite well and produces accurate solutions in close agreement with analytical and previous numerical results.

Possible disadvantages of the gauge approach may include the following. First, compared with the electric and magnetic field vectors, the gauge potentials are less familiar and are not usually measured directly. Second (if this is really a disadvantage), the gauge approach forces the user to understand more completely what physical process he/she is attempting to model. A Maxwell approach, for example, may allow the user to rather blindly set the electric field to be some value on the boundary. The electric field, however, is really a combination of two distinct physical phenomena. An electric field can be produced either by distributions of spatial charge, or by the acceleration of electric charge. The first is a Galvanesque effect, while the second is an induction effect. In the gauge formulation, the electric field is a combination of two terms. One involves the gradient of the electric potential (established by the spatial distribution of electric charge), while the other involves the time rate of change of the magnetic vector potential (accelerating electric charges tend to ‘entrain’ other charges—even if on the other side of an insulator). Hence, to specify the electric field on the boundary using the gauge approach, one is forced to decide what ‘kind’ of electric field one really means.

ACKNOWLEDGMENTS

This work was conducted with support from the US Office of Naval Research (N00014-99-1-0407) and the US National Aeronautics and Space Administration (NAG5-10570).

REFERENCES


APPENDIX A: CHOOSING A GAUGE FOR MACROSCOPIC ELECTRODYNAMICS

As we have asserted, there is an arbitrariness in the choice of the gauge which should be exploited in our solution strategy. The choice of gauge controls to some degree the form of the governing equation...
for $A_\mu$, but more importantly how the four equations are coupled. These choices will affect the accuracy and efficiency of the numerical solutions.

Theoretically, the choice of gauge can be linked to the physical interpretation of the gauge potentials and the requirement of certain invariances following relativity. On these terms, perhaps the most fundamental gauge is the Lorentz gauge, which, at least in vacuum electrodynamics, is required to retain consistency with special relativity theory. When the gauge potentials are assumed to reflect macroscopic averages in the presence of material media, however, the Lorentz gauge is no longer a required or obvious choice even formally because the introduction of material creates a ‘preferred’ reference frame in which the simplest constitutive relationships hold. In material macroscopic electrodynamics, gauge preferences on physical grounds can follow as a result of the material macroscopic electrodynamics.

In assessing the potential advantages of using a staggered grid, there are other options. In particular, a staggered grid approach might be useful for computational efficiency, but this raises the question, ‘is there a gauge that will provide a conservative scheme on this grid?’ The answer is ‘yes’. This specific topic is treated in detail in the next subsection.

### Benefits of a staggered grid?

The current version of MOED co-locates all parameters and variables on the same grid. This can be contrasted with approaches using staggered grids, in which variables and parameters can be defined on different grids. A common example is the staggered grid associated with the Maxwell equations discretized on the Yee (1966) computation lattice.

While staggered grids demonstrate a clear advantage in the solution of the Maxwell equations, a similar demonstration has not been given for the gauge equations. Nor does reference to the staggered grid Maxwell equations approach immediately indicate that there would be an advantage, because the gauge formulation involves higher-order equations in fewer variables. Analogies with staggered grids (e.g. the Arakawa C-grid) used in geophysical fluid dynamics, however, do suggest a possible benefit to adopting a staggered grid in MOED. Essentially, staggered grids offer the potential advantage of increasing computational efficiency. The disadvantage is that the staggered grid may be more complicated to configure and the solution products may be defined on different grids, requiring careful post-processing.

In assessing the potential advantages of using a staggered grid, several factors must be considered. The most basic consideration is of the gridding system that will most efficiently allow for the solution of the components of $A_\mu$. This, however, depends on the form of the governing equations which, as we have seen, vary with the choice of the gauge. Hence, the choices of staggered grid and

<table>
<thead>
<tr>
<th>Gauge</th>
<th>$\mu$</th>
<th>$D$</th>
<th>$a$</th>
<th>$p_\mu$</th>
<th>$f_\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damped Lorentz</td>
<td>$\mu = 1, 2, 3$</td>
<td>1</td>
<td>$i\omega \mu \sigma^*$</td>
<td>$\mu_\phi \nabla \sigma^*$</td>
<td>$\mu_\phi J_\mu$</td>
</tr>
<tr>
<td></td>
<td>$\mu = 4$</td>
<td>$\sigma^*$</td>
<td>$i\omega \mu \sigma^*$</td>
<td>$-i\omega \nabla \cdot (\sigma^* A)$</td>
<td>$\nabla \cdot J_\mu$</td>
</tr>
<tr>
<td>Lorentz</td>
<td>$\mu = 1, 2, 3$</td>
<td>1</td>
<td>$i\omega \mu \sigma^*$</td>
<td>$-i\omega \nabla \cdot (\sigma^* A)$</td>
<td>$\mu_\phi J_\mu$</td>
</tr>
<tr>
<td></td>
<td>$\mu = 4$</td>
<td>$\sigma^*$</td>
<td>0</td>
<td>$-i\omega \nabla \cdot (\sigma^* A)$</td>
<td>$\nabla \cdot J_\mu$</td>
</tr>
<tr>
<td>Coulomb</td>
<td>$\mu = 1, 2, 3$</td>
<td>1</td>
<td>$i\omega \mu \sigma^*$</td>
<td>$-i\omega \nabla \cdot (\sigma^* A)$</td>
<td>$\mu_\phi J_\mu$</td>
</tr>
<tr>
<td></td>
<td>$\mu = 4$</td>
<td>$\sigma^*$</td>
<td>0</td>
<td>$-i\omega \nabla \cdot (\sigma^* A)$</td>
<td>$\nabla \cdot J_\mu$</td>
</tr>
</tbody>
</table>
gauge should be done simultaneously. A second consideration is that what is often wanted in the end are accurate descriptions of E and B (which are calculated from \( A_n \)). Hence, a second consideration is that ideally the solutions for the \( A_n \) components should occur on grids that allow for appropriate differencing to be conducted to produce accurate estimates of E and B without the need for first interpolating.

A staggered grid system that provides this is the following (see Fig. A1). We first assume that the parameters—which are prescribed—can be provided on any grid as necessary without reducing accuracy. We assume the Lorentz gauge. The electric potential \( (A_x) \) is assigned on a central grid. Each of the components \( A_i \) is assigned on grids shifted by a half gridpoint spacing along the axis \( i \) such that the Lorentz gauge (which has the form of a conservation equation) can be efficiently discretized by considering fluxes across the faces of a grid point control volume. It can be shown that on this grid the governing equations for \( A_n \) with the Lorentz gauge can be discretized in a straightforward manner. Furthermore, the four component solutions of \( A_n \) (all on different grids) are in appropriate locations for directly calculating both E and B without first interpolating the solutions to another grid. The components of E are obtained on the same grids as the components of \( A_i \). The components of B are obtained on three different grids located at the corners of the computational lattice.

**APPENDIX B: ANALYTICAL SOLUTION FOR TRANSVERSE EM MODES IN A THREE-LAYER EARTH**

When we consider the equations of transverse EM modes, such as discussed in Weaver (1994), analytical solutions are readily obtained for the case of a three-layer earth (see Fig. A2). The solutions for the electric and magnetic fields with \( B_1/\mu_0 = 1 \text{A/m} \) at the surface are given as

\[
E_i = C_i e^{ik_i z} + D_i e^{-ik_i z},
\]

\[
B_j = \frac{k_j}{\omega_0} (C_j e^{ik_j z} - D_j e^{-ik_j z}),
\]

where \( i \) is the layer number, \( k_i = \sqrt{\mu_0 \sigma / \rho} \), \( z \) is the depth, and \( C_i, D_i \) are the coefficients for each layer. Note that \( D_3 = 0 \) for the solutions to remain bounded.

Using the matching conditions at \( h_1 \) and \( h_2 \) \( (E_1 = E_2^+, B_1^+ = B_2^+ \) and \( E_2 = E_3^+, B_2^+ = B_3^+ ) \), we can solve the coefficients of \( C_i \) and \( D_j \):

\[
C_1 = \frac{1}{k_1} \left( 1 - \frac{1 - \frac{1}{\mu_0} R}{1 + \frac{1}{\mu_0} R} e^{2ik_1 h_1} \right),
\]

\[
D_1 = \frac{1}{1 + \frac{1}{\mu_0} R} e^{2ik_1 h_1} C_1,
\]

\[
C_2 = C_1 e^{ik_1 h_1} + F_1 e^{-ik_1 h_1} \left( e^{ik_2 h_2} + K_{23} e^{ik_3 h_3} \right),
\]

\[
D_2 = k_2 e^{ik_2 h_2} C_2,
\]

\[
C_3 = \frac{k_2}{k_3} (1 - k_2) e^{ik_3 (h_1 + h_2)} C_2,
\]

where \( R = \frac{1 - k_1 e^{2ik_1 h_1}}{1 + k_1 e^{2ik_1 h_1}} \), \( K_{23} = \frac{\sqrt{2} - 1}{\sqrt{2} + 1} \) and \( F_1 = \frac{1 - k_1 R}{1 + k_1 R} e^{ik_1 h_1} \).

![Figure A1. Location of variables on a staggered grid.](image)

![Figure A2. Three-layer earth with \( B_1 \) prescribed at the surface.](image)