(3): Convergence

Where we stand

Trying to discretize the Maxwell equations by using a pair of interlocked meshes, we have obtained a system of ordinary differential equations

$$\partial_t \mathbf{b} + \mathbf{R}\mathbf{e} = 0, \quad -\partial_t \tilde{\star}_\epsilon \mathbf{e} + \mathbf{R}^t \tilde{\star}_\mu^{-1} \mathbf{b} = \tilde{\mathbf{j}}. \tag{1}$$

There, **e** and **b** are arrays of degrees of freedom, meant to approximate the emf's and magnetic fluxes associated with edges and facets of the primal mesh, and **R** is the edge-to-facets incidence matrix. Intensities through dual facets, which compose the array $\tilde{\mathbf{j}}$, are supposed to be known, as functions of time. Symbols $\tilde{\star}_\epsilon$ and $\tilde{\star}_\mu$ denote square symmetric positive-definite matrices which encode the constitutive laws in "discrete" form.

We chose such strange-looking symbols in order to emphasize the connection with the Hodge operator $\tilde{\star}$ of differential geometry, which appears in the "continuous" form of the constitutive laws,

$$b = \mu \tilde{\star} \tilde{h}, \qquad \tilde{d} = \epsilon \tilde{\star} e.$$

We shall rewrite this as $b = \tilde{\star}_\mu \tilde{h}$ and $\tilde{d} = \tilde{\star}_\epsilon e$ from now on, to stress the parallel between the compound operators $\tilde{\star}_\mu = \mu\tilde{\star}$ and $\tilde{\star}_\epsilon = \epsilon\tilde{\star}$ and the "discrete Hodges" $\tilde{\star}_\mu$ and $\tilde{\star}_\epsilon$.

The form (1) of the discrete equations was practically forced on us—as soon as we decided to use meshes in duality, that is. But what $\tilde{\star}_\epsilon$ and $\tilde{\star}_\mu$ should be, in detail, was left to our choice. We found several criteria about what constitutes a "good discrete Hodge". In the case of mutually orthogonal meshes (the "orthogonal construction" of Part 1, cf. *JSAEM, 7,* 2, 1999, pp. 150-9), a likely candidate was identified: a diagonal discrete Hodge, which in the case of $\tilde{\star}_\mu$ had the following diagonal entries (indexed over facets f of the primal mesh):

$$\tilde{\star}_\mu^{ff} = \mu \frac{\text{area}(f)}{\text{length}(\tilde{f})}, \tag{2}$$

where $\tilde{f}$ is the edge dual to f in the dual mesh (Fig. 1). Similarly, $\tilde{\star}_\epsilon^{ee} = \epsilon\,\text{area}(\tilde{e})/\text{length}(e)$. Let's immediately emphasize that there are other ways to build a discrete Hodge, and we'll consider a few of them later. But this one, being particularly simple, should be a good test of the validity of the whole approach: Can we prove that, when the meshes are refined in some well-specified way, the solution of (1) converges, in some reasonable sense, towards the solution of the Maxwell equations? This is the subject of the present installment.

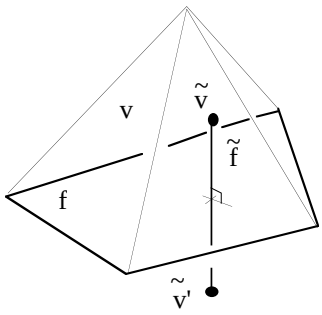

Figure 1. A facet f and its dual edge $\tilde{f}$ in the orthogonal construction ($\tilde{v}$ and $\tilde{v}'$ are the dual nodes which lie inside the volumes v and v' just above and just below f). From $\tilde{v}$, all boundary facets of v can directly be seen at right angle, but we don't require more: $\tilde{v}$ is neither v's barycenter nor the center of its circumscribed sphere, if there is such a sphere.

3.1 The static case

Obviously, the discrete Hodge of (2) would have no virtue if it didn't work satisfactorily in *static* situations. So let's begin with that, which will take the major part of this paper. Then we shall briefly return to the full Maxwell system.

Magnetostatics, in the context of our original model problem, is this: Given a time-independent current density **J** in the bounded

domain of Fig. 2, find **B** and **H** such that

$$(3') \qquad \begin{array}{c} \mathrm{div}\,\mathbf{B} = 0,\ \mathbf{B} = \mu\,\mathbf{H},\ \mathrm{rot}\,\mathbf{H} = \mathbf{J}, \\ \nu \cdot \mathbf{B} = 0 \text{ on } S^e,\ \nu \times \mathbf{H} = 0 \text{ on } S^h, \end{array}$$

which our geometric language of differential forms expresses as

$$(3) \qquad \begin{array}{c} \mathrm{d}b = 0,\ b = \tilde{\star}_\mu \tilde{h},\ \mathrm{d}\tilde{h} = \tilde{\jmath}, \\ tb = 0 \text{ on } S^e,\ t\tilde{h} = 0 \text{ on } S^h. \end{array}$$

The corresponding network equations are

$$(4) \qquad \mathbf{Db} = 0,\ \mathbf{b} = \tilde{\star}_\mu \tilde{\mathbf{h}},\ \mathbf{R}^t \tilde{\mathbf{h}} = \tilde{\mathbf{j}},$$

with $\mathbf{b} = \{\mathbf{b}_f : f \in \mathcal{F}\}$ and $\tilde{\mathbf{h}} = \{\tilde{\mathbf{h}}_f : f \in \mathcal{F}\}$ indexed over the set $\mathcal{F}$ of "active" facets, i.e., all of them except those lying in S^e. We might establish (4) from first priciples, as we did for the network equations (1). But since we have these already, let's rather derive (4) from (1), as the steady-state equations for infinite t, under the hypothesis that $\tilde{\mathbf{j}}$ does not depend on time. For this, we let the facets-to-volumes incidence matrix **D** act on the first eq. (1): since **DR** = 0, $\partial_t(\mathbf{Db}) = 0$, hence **Db** = 0 at all times, since **b** was null at time zero.

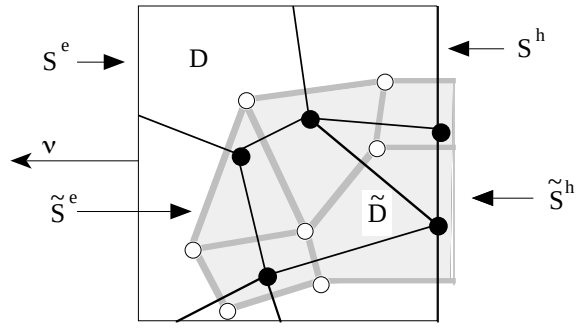

Figure 2 (to be imagined in dimension 3). A reminder of our notations: Domain D, bounded by S, stands for the left half of the cavity. D is paved by the primal cells. The dual cells pave a slightly different domain $\tilde{D}$. Dual $(n-p)$-cells and primal p-cells are orthogonal, two by two. Primal cells in the electric boundary S^e, as well as dual cells in the magnetic boundary $\tilde{S}^h$, are discarded, because degrees of freedom they would bear are a priori zero. (The primal and dual pavings are "closed modulo" S^e and $\tilde{S}^h$, cf. *JSAEM, 7*, 2, 1999, p. 155.) The ν of (3′) is the field of outgoing normals.

Our first concern is whether (4) determines a unique pair $\{\mathbf{b}, \tilde{\mathbf{h}}\}$. Let's denote by N, E, F, V the numbers of active primal nodes, edges, facets, and volumes. (Again, active nodes and edges are those not in S^e.) In (4), we count $2F$ unknowns, the components of **b** and $\tilde{\mathbf{h}}$, and $V + E + F$ equations, for **D** has V rows (one per volume), **R** has E columns, indexed over the set $\mathcal{E}$ of active edges, and $\mathbf{b} = \tilde{\star}_\mu \tilde{\mathbf{h}}$ provides F equations. By a basic result of topology, the *Euler–Poincaré formula*, we know that, whichever way the primal mesh was constructed,

$$(5) \qquad N - E + F - V = \chi,$$

where χ is a constant (equal to 0 in the case we consider) which only depends on the topology of D, S^e, and S^h. This leaves us with $V + E + F \equiv N + 2F$ equations for $2F$ unknowns. So it's not so obvious that (4) has a unique solution.

However, there are precisely N constraints on the data $\tilde{\mathbf{j}}$, owing to current conservation: For each dual volume, that is, for each active primal node n, currents entering this volume should cancel out, hence N relations on the $\tilde{\mathbf{j}}_e$s, of the form $\sum_{e \in \mathcal{E}} \mathbf{G}_{en}\, \tilde{\mathbf{j}}_e = 0$, where **G** denotes the (primal, and active) nodes-to-edges incidence matrix. They can simultaneously be expressed as $\mathbf{G}^t \tilde{\mathbf{j}} = 0$. (We could as well have derived this necessary condition from the last eq. (4), since $\mathbf{G}^t \mathbf{R}^t = 0$, by transposing the combinatorial relation **RG** = 0.)

So by the removal of N redundant data, and of the corresponding equations, we could fall back on a square system, which we should still prove regular. This can be done, but the following indirect approach will be more instructive: We shall construct a linear system *equivalent* to (4), the regularity of which will be obvious.

Let's recall that, with the simple topology we assume here, not only **RG** = 0 and **DR** = 0, but the kernels ker(**R**) and ker(**D**) coincide with the ranges of **G** and **R**. By transposition, ker($\mathbf{G}^t$) is the range of $\mathbf{R}^t$, and ker($\mathbf{R}^t$) is the range of $\mathbf{D}^t$. So if $\mathbf{G}^t \tilde{\mathbf{j}} = 0$, there exists an $\mathcal{F}$-indexed array $\tilde{\mathbf{h}}^{\mathbf{j}}$ such that $\mathbf{R}^t \tilde{\mathbf{h}}^{\mathbf{j}} = \tilde{\mathbf{j}}$. (It's not unique, and need not be explicitly constructed, though that would be a trivial task. That there be one is enough for our purpose.) Now, $\mathbf{R}^t(\tilde{\mathbf{h}} - \tilde{\mathbf{h}}^{\mathbf{j}}) = 0$, so there is a DoF-array $\tilde{\varphi}$, indexed over volumes, such that $\tilde{\mathbf{h}} = \tilde{\mathbf{h}}^{\mathbf{j}} + \mathbf{D}^t \tilde{\varphi}$, and (4) reduces to

$$(6) \qquad \mathbf{D} \tilde{\star}_\mu \mathbf{D}^t \tilde{\varphi} = -\mathbf{D} \tilde{\star}_\mu \tilde{\mathbf{h}}^{\mathbf{j}}.$$

Now this is a square symmetric linear system, with respect to $\tilde{\varphi}$, with a regular matrix,

because $\tilde{\star}_\mu$ is regular on the one hand, and (this is the non-obvious part) $\ker(\mathbf{D}^t) = \{0\}$ on the other hand. Indeed, $\mathbf{D}^t\tilde{\psi} = 0$ means that $\sum_v \mathbf{D}_{vf}\tilde{\psi}_v = 0$ for all primal facets f. But for each such f, there are at most two incident volumes v and v', one on each side of f, and their incidence numbers $\mathbf{D}_{vf}$ and $\mathbf{D}_{v'f}$ have opposite signs. Therefore, $\tilde{\psi}_v = \tilde{\psi}_{v'}$, and $\mathbf{D}^t\tilde{\psi} = 0$ implies that all components of $\tilde{\psi}$ are equal, as soon as the paved domain is connected. Moreover (and now, this is a characteristic of the present situation, where $\tilde{S}^h$ is not empty, not an always valid property), there are facets f with only one adjacent volume (Fig. 2), hence this common value must be zero for all $\tilde{\psi}_v$s. So $\tilde{\varphi} = 0$ in (6) if $\tilde{\mathbf{h}}^{\mathbf{j}} = 0$.

Equation (6) thus appears as a way to solve (4), with guaranteed existence and uniqueness: having $\tilde{\varphi}$, we set $\tilde{\mathbf{h}} = \tilde{\mathbf{h}}^{\mathbf{j}} + \mathbf{D}^t\tilde{\varphi}$, and $\mathbf{b} = \tilde{\star}_\mu\tilde{\mathbf{h}}$. This is known as the *finite volume* approach to magnetostatics, with one degree of freedom per volume of the (primal) mesh, which one may of course interpret as the value of a magnetic potential at the dual node. Many researchers have analyzed the convergence of (6), for various mesh designs and various choices of $\tilde{\star}_\mu$. (See, e.g., [2, 8, 10, 11, 14, 19].) Why then not rely on their results? Because system (4), with its symmetrical and balanced treatment of $\mathbf{b}$ and $\tilde{\mathbf{h}}$, will lend itself to a much simpler error analysis than (6), and one which does the job for several apparently distinct formulations, in one stroke.

For there are other systems equivalent to (4), that we shall indicate before carrying on, by following up on this symmetry idea. Since $\ker(\mathbf{D})$ is the range of $\mathbf{R}$, one may look for $\mathbf{b}$, which has to be in it, in the form $\mathbf{b} = \mathbf{R}\mathbf{a}$, where the DoF-array $\mathbf{a}$ is indexed over $\mathcal{E}$. Then (4) is equivalent to the following linear system, in terms of $\mathbf{a}$,

$$\mathbf{R}^t\tilde{\star}_\mu^{-1}\mathbf{R}\mathbf{a} = \tilde{\mathbf{j}}. \tag{7}$$

No uniqueness, this time,[1] because $\ker(\mathbf{R})$ does not reduce to 0, but there are solutions, thanks to the condition $\mathbf{G}^t\tilde{\mathbf{j}} = 0$, which guarantees that $\tilde{\mathbf{j}} = 0$ lies in the range of $\mathbf{R}^t$, and $\mathbf{b} = \mathbf{R}\mathbf{a}$ is the same for all these solutions. So solving (7), thus getting a unique $\mathbf{b}$, and setting $\tilde{\mathbf{h}} = \tilde{\star}_\mu^{-1}\mathbf{b}$, is equivalent to solving (4).

This is not all. If we refrain to eliminate $\tilde{\mathbf{h}}$ in the reduction of (4) to (7), but still use $\mathbf{b} = \mathbf{R}\mathbf{a}$, we get an intermediate two-equation system,

$$\begin{pmatrix} -\tilde{\star}_\mu & \mathbf{R} \\ \mathbf{R}^t & 0 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{h}} \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} 0 \\ \tilde{\mathbf{j}} \end{pmatrix}, \tag{8}$$

often called a *mixed* algebraic system. The same manipulation in the other direction (eliminating $\tilde{\mathbf{h}}$ by $\tilde{\mathbf{h}} = \tilde{\mathbf{h}}^{\mathbf{j}} + \mathbf{D}^t\tilde{\varphi}$, but keeping $\mathbf{b}$) gives

$$\begin{pmatrix} -\tilde{\star}_\mu^{-1} & \mathbf{D}^t \\ \mathbf{D} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{b} \\ \tilde{\varphi} \end{pmatrix} = \begin{pmatrix} -\tilde{\mathbf{h}}^{\mathbf{j}} \\ 0 \end{pmatrix}. \tag{9}$$

Systems (6), (7), (8), and (9) differ in size and in sparsity, but give the same solution pair $\{\mathbf{b}, \tilde{\mathbf{h}}\}$, so which one effectively to solve is a matter of algorithmics that need not concern us here.[2] The important point is, the error analysis we shall perform applies to all of them.

3.2 Consistency

A notational point, before going further. Last time, we used $\mathcal{M}$ to denote the primal mesh. We shall subscript by $\mathcal{M}$, when necessary, all mesh-related entities. For instance, the largest diameter of all cells, primal and dual, will be denoted $\gamma_\mathcal{M}$ (with a mild abuse, since it also depends on the metric of the dual mesh, $\widetilde{\mathcal{M}}$), and called the "grain" of the pair of meshes. The computed solution $\{\mathbf{b}, \tilde{\mathbf{h}}\}$ will be $\{\mathbf{b}_\mathcal{M}, \tilde{\mathbf{h}}_\mathcal{M}\}$ when we wish to mark its dependence on the mesh-pair. And so on.

Our purpose can informally be stated as "study $\{\mathbf{b}_\mathcal{M}, \tilde{\mathbf{h}}_\mathcal{M}\}$ when $\gamma_\mathcal{M}$ tends to 0". Alas, this lacks definiteness, because how the *shape*

[1] Whether to "gauge" $\mathbf{a}$ in this method, that is, to impose a condition such as $\mathbf{G}^t\mathbf{a} = 0$ that would select a unique solution, remains to these days a contentious issue. It depends on which method is used to solve (7), and on how well the necessary condition $\mathbf{G}^t\tilde{\mathbf{j}} = 0$ is implemented. With iterative methods such as the conjugate gradient and its variants, and if one takes care to set up an array $\tilde{\mathbf{h}}^{\mathbf{j}}$ such that $\mathbf{R}^t\tilde{\mathbf{h}}^{\mathbf{j}} = \tilde{\mathbf{j}}$, and to use $\mathbf{R}^t\tilde{\mathbf{h}}^{\mathbf{j}}$ instead of $\tilde{\mathbf{j}}$ in (7), then it's better *not* to gauge. See [18].

[2] Assigning b to *dual* facets and $\tilde{h}$ to *primal* edges would generate a similar family of equivalent systems, but not equivalent to (6)–(9), thus yielding *complementary* information. See [5], Chap. 6.

of the cells changes in the process matters a lot. In the case of triangular 2D meshes, for instance, there are well-known counter-examples [1] showing that, if one tolerates too much "flattening" of the triangles as the grain tends to 0, convergence may not occur. Hence the following definition: A family $\mathcal{M}$ of (pairs of interlocked) meshes is *uniform* if there is a *finite* catalogue of "model cells" such that any cell in any $\mathcal{M}$ or $\widetilde{\mathcal{M}}$ of the family is similar to one of them. The notation "$\mathcal{M} \to 0$" will then refer to a sequence of meshes, all belonging to some definite uniform family, and such that their $\gamma_{\mathcal{M}}$s tend to zero. Now we redefine our objective: Show that the error, whatever one means by that, incurred by taking $\{\mathbf{b}_{\mathcal{M}}, \tilde{\mathbf{h}}_{\mathcal{M}}\}$ as a substitute for the real field $\{b, \tilde{h}\}$, tends to zero when $\mathcal{M} \to 0$.

Practical implications of achieving this are as follows. If, for a given $\mathcal{M}$, the computed solution $\{\mathbf{b}_{\mathcal{M}}, \tilde{\mathbf{h}}_{\mathcal{M}}\}$ is not deemed satisfactory, one must *refine* the mesh and redo the computation, again and again. If the refinement rule guarantees that all meshes such a process can generate will belong to some definite uniform family, then the convergence result means "you may get as good an approximation as you wish by refining this way", a state of affairs we are more or less happy to live with.[3]

Fortunately, such refinement rules do exist (this is an active area of research [3, 4, 7, 15]). Given a pair of coarse meshes to start with, there are ways to subdivide the cells so as to keep bounded the number of different cell-shapes that appear in the process, hence a potential infinity of refined meshes, which do constitute a uniform family. (A refinement process for tetrahedra is illustrated by Fig. 3. As one can see, at most five different shapes can occur, for each tetrahedral shape present in the original coarse mesh.)

[3] Effective error *bounds*, for a given $\mathcal{M}$, would of course be more satisfying. Such bounds can be obtained with the complementarity methods alluded to in Note 2.

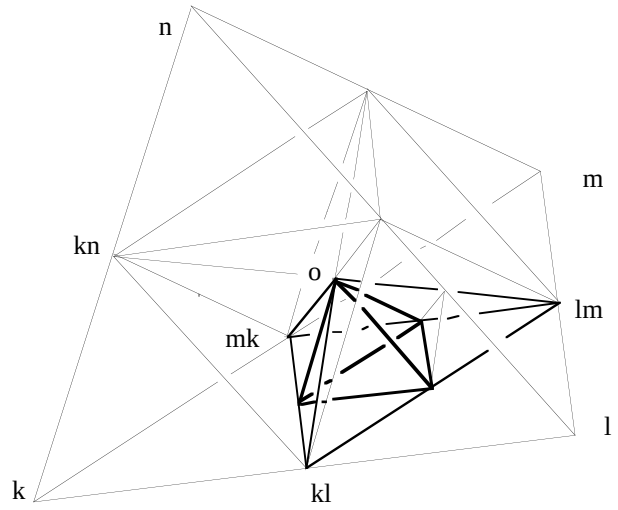

Figure 3. Subdivision rule for a tetrahedron T = $\{k, l, m, n\}$. (Mid-edges are denoted kl, lm, etc., and o is the barycenter.) A first halving of edges generates four small tetrahedra and a core octahedron, which itself can be divided into eight "octants" such as O = $\{o, kl, lm, mk\}$, of at most four different shapes. Now, octants like O should be subdivided as follows: divide the facet in front of o into four triangles, and join to o, hence a tetrahedron similar to T, and three peripheral tetrahedra. These, in turn, are halved, as shown for the one hanging from edge $\{o, lm\}$. Its two parts are similar to O and to the neighbor octant $\{o, kn, kl, mk\}$ respectively.

Back to the comparison between $\{\mathbf{b}_{\mathcal{M}}, \tilde{\mathbf{h}}_{\mathcal{M}}\}$ and $\{b, \tilde{h}\}$, a natural idea is to compare the computed DoF arrays, $\mathbf{b}_{\mathcal{M}}$ and $\tilde{\mathbf{h}}_{\mathcal{M}}$, with arrays of the same kind, $r_{\mathcal{M}} b = \{\int_f b : f \in \mathcal{F}\}$ and $r_{\mathcal{M}} \tilde{h} = \{\int_{\tilde{f}} \tilde{h} : f \in \mathcal{F}\}$, composed of the fluxes and mmf's of the (unknown) solution of (3). This implicitly defines two operators with the same name, $r_{\mathcal{M}}$: one that acts on 2-forms, giving an array of facet-fluxes, one that acts on twisted 1-forms, giving an array of dual-edge mmf's. (No risk of confusion, since the name of the operand, b or $\tilde{h}$, reveals its nature.)

Since $\mathrm{d}b = 0$, the flux of b across the boundary of any primal 3-cell v must vanish, hence the sum of facet fluxes $\sum_f \mathbf{D}_{vf} \int_f b$ must vanish for all v. Similarly, $\mathrm{d}\tilde{h} = \tilde{\jmath}$ yields the relation $\sum_f \mathbf{R}_{fe} \int_{\tilde{f}} \tilde{h} = \int_{\tilde{e}} \tilde{\jmath}$, by integration over a dual 2-cell. In matrix form, all this becomes

$$\mathbf{D} r_{\mathcal{M}} b = 0, \quad \mathbf{R}^t r_{\mathcal{M}} \tilde{h} = \tilde{\mathbf{j}}, \tag{10}$$

since the entries of $\tilde{\mathbf{j}}$ are precisely the intensities across the dual facets. Comparing with (4), we obtain

$$\mathbf{D}(\mathbf{b}_{\mathcal{M}} - r_{\mathcal{M}} b) = 0, \quad \mathbf{R}^t(\tilde{\mathbf{h}}_{\mathcal{M}} - r_{\mathcal{M}} \tilde{h}) = 0, \tag{11}$$

and

$$(\mathbf{b}_{\mathcal{M}} - r_{\mathcal{M}} b) - \tilde{\star}_\mu(\tilde{\mathbf{h}}_{\mathcal{M}} - r_{\mathcal{M}} \tilde{h}) = (\tilde{\star}_\mu r_{\mathcal{M}} - r_{\mathcal{M}} \tilde{\star}_\mu)\tilde{h} \equiv \tilde{\star}_\mu(r_{\mathcal{M}} \tilde{\star}_\nu - \tilde{\star}_\nu r_{\mathcal{M}})b. \tag{12}$$

(Here, of course, ν and $\tilde{\star}_\nu$ stand for the inverses of μ and $\tilde{\star}_\mu$, and $\tilde{\star}_\nu$ for $\tilde{\star}_\mu^{-1}$. We have no more use for ν as a normal vector, so no confusion should ensue.)

Now some algebra, which requires further notation. Last time, we chose to denote by $\langle \mathbf{b}, \tilde{\mathbf{h}} \rangle$ a sum such as $\sum_{f \in \mathcal{F}} \mathbf{b}_f \tilde{\mathbf{h}}_f$. We shall use the shorthands $|\tilde{\mathbf{h}}|_\mu$ and $|\mathbf{b}|_\nu$ for the square roots of the quantities $\langle \tilde{\star}_\mu \tilde{\mathbf{h}}, \tilde{\mathbf{h}} \rangle$ and $\langle \mathbf{b}, \tilde{\star}_\nu \mathbf{b} \rangle$, and call them the "$\mu$-norm" and "$\nu$-norm" of these DoF arrays. (Notice their connection with what we called last time "discrete energy".) We want to compute the μ-norm of both sides of (12).

Doing this, "square" and "rectangle" terms appear, as usual. The rectangle term for the left-hand side is $2\langle \mathbf{b}_{\mathcal{M}} - r_{\mathcal{M}} b, \tilde{\mathbf{h}}_{\mathcal{M}} - r_{\mathcal{M}} \tilde{h} \rangle$, but since $\mathbf{D}(\mathbf{b}_{\mathcal{M}} - r_{\mathcal{M}} b) = 0$ implies the existence of some $\mathbf{a}$ such that $\mathbf{b}_{\mathcal{M}} - r_{\mathcal{M}} b = \mathbf{R}\mathbf{a}$, we have

$$\langle \mathbf{b}_{\mathcal{M}} - r_{\mathcal{M}} b, \tilde{\mathbf{h}}_{\mathcal{M}} - r_{\mathcal{M}} \tilde{h} \rangle = \langle \mathbf{R}\mathbf{a}, \tilde{\mathbf{h}}_{\mathcal{M}} - r_{\mathcal{M}} \tilde{h} \rangle = \langle \mathbf{a}, \mathbf{R}^t(\tilde{\mathbf{h}}_{\mathcal{M}} - r_{\mathcal{M}} \tilde{h}) \rangle = 0,$$

after (11), by the same transposition trick as last time. Only square terms remain, and we get

$$|\mathbf{b}_{\mathcal{M}} - r_{\mathcal{M}} b|_\nu^2 + \|\tilde{\mathbf{h}}_{\mathcal{M}} - r_{\mathcal{M}} h|_\mu^2 = \|(\tilde{\star}_\nu r_{\mathcal{M}} - r_{\mathcal{M}} \tilde{\star}_\nu)b\|_\mu^2 \equiv |(\tilde{\star}_\mu r_{\mathcal{M}} - r_{\mathcal{M}} \tilde{\star}_\mu)\tilde{h}\|_\nu^2, \tag{13}$$

which will be the cornerstone of the convergence proof.

So at last we have found a plausible measure for what we called earlier "the error incurred by taking $\mathbf{b}_{\mathcal{M}}$ as a substitute for the real field b": the μ-norm of $\mathbf{b}_{\mathcal{M}} - r_{\mathcal{M}} b$. Components of this array are what can be called the "residual fluxes" $\mathbf{b}_f - \int_f b$, i.e., the difference between the computed flux across face f and the genuine (but unknown) flux $\int_f b$. It makes sense to try and *bound* this norm. (Parallel considerations apply to $\tilde{h}$, with mmf's along $\tilde{f}$ instead of fluxes.) So let's focus on the right-hand side of (13), for instance on its second expression, in terms of $\tilde{h}$.

By definition of $r_{\mathcal{M}}$, the f-component of $r_{\mathcal{M}} \tilde{\star}_\mu \tilde{h}$ is the flux of $b = \tilde{\star}_\mu \tilde{h}$ across f. On the other hand, the flux-array $\tilde{\star}_\mu r_{\mathcal{M}} \tilde{h}$ is the result of applying the discrete Hodge operator to the mmf-array $r_{\mathcal{M}} \tilde{h}$, so the compound operators $r_{\mathcal{M}} \tilde{\star}_\mu$ and $\tilde{\star}_\mu r_{\mathcal{M}}$ will not be equal: they give different fluxes when applied to a generic $\tilde{h}$. This contrasts with the equalities $(\mathbf{D} r_{\mathcal{M}} - r_{\mathcal{M}} \mathrm{d})b = 0$ and $(\mathbf{R}^t r_{\mathcal{M}} - r_{\mathcal{M}} \mathrm{d})\tilde{h} = 0$, which stem from the Stokes theorem. The mathematical word to express such equalities is "conjugacy": $\mathbf{D}$ and d are conjugate via $r_{\mathcal{M}}$, and so are $\mathbf{R}^t$ and d, too.

Thus, $\tilde{\star}_\mu$ and $\tilde{\star}_\mu$ are *not* conjugate via $r_{\mathcal{M}}$—and this is, of course, the reason why discretizing entails some error. Yet, in the case we are examining (the diagonal Hodge defined by (2)), $r_{\mathcal{M}} \tilde{\star}_\mu$ and $\tilde{\star}_\mu r_{\mathcal{M}}$ *do* coincide for *some* $\tilde{h}$s, those that have piecewise constant vector proxies, since this is how formula (2) was motivated. Since all smooth fields look constant at a small enough scale, we may expect "asymptotic conjugacy", in the sense that the right-hand side of (13) will tend to 0 with $\mathcal{M}$, for a smooth b or $\tilde{h}$. This property, which we rewrite informally but suggestively as

$$\begin{aligned} &\tilde{\star}_\nu r_{\mathcal{M}} - r_{\mathcal{M}} \tilde{\star}_\nu \to 0 \text{ when } \mathcal{M} \to 0, \\ &\tilde{\star}_\mu r_{\mathcal{M}} - r_{\mathcal{M}} \tilde{\star}_\mu \to 0 \text{ when } \mathcal{M} \to 0 \end{aligned} \tag{14}$$

(two equivalent statements), is called *consistency* of an approximation scheme in Numerical Analysis (approximation of $\tilde{\star}_\mu$ and $\tilde{\star}_\nu$ by $\tilde{\star}_\mu$ and $\tilde{\star}_\nu$, here). To prove it, we need to estimate the right-hand side of (13).

This can be done by estimating the contribution of a single facet f, that is

$$\tilde{\star}_\nu^{ff}(\tilde{\star}_\mu^{ff} \int_{\tilde{f}} \tilde{h} - \int_f \tilde{\star}_\mu \tilde{h})^2, \tag{15}$$

and we may even pretend that $\tilde{f}$ is entirely in the volume v to do so (cf. Fig. 1), since there are two parts in the contribution of f, one for each adjacent volume. Then we may assume a constant μ inside v, and work in terms of the vector proxies $\mathbf{H}$ and $\mathbf{B} = \mu\mathbf{H}$. Now $P \equiv \int_{\tilde{f}} \tilde{h}$ is the circulation of $\mathbf{H}$ along $\tilde{f}$ and $Q \equiv \int_f \tilde{\star}_\mu \tilde{h}$ is μ times the flux of $\mathbf{H}$ across f. Since (15) vanishes for a constant $\mathbf{H}$ we may, suppose that $\mathbf{H} = 0$ vanishes at the intersection $f \cap \tilde{f}$ (Fig. 1), and select this point as origin. Then $\mathbf{H}(x)$ is bounded by $C\gamma_{\mathcal{M}}$ over v, where C is a constant which depends on $\mathbf{H}$, but not

on the mesh. (All such constants, whatever their value, will uniformly be denoted by C from now on.) A very crude[4] bound for Q is then $C\mu\gamma_{\mathcal{M}}$area(f). Similarly, P is bounded by $C\gamma_{\mathcal{M}}$length($\tilde{f}$), and hence $|\tilde{\star}_\mu^{ff}P - Q| \leq C\mu\gamma_{\mathcal{M}}$ area(f), after (2). Now, to obtain the desired estimate, square this, divide by $\tilde{\star}_\mu^{ff}$, which gives C length($\tilde{f}$) area(f)$\gamma_{\mathcal{M}}^2$, and finally, sum over f, hence $C\gamma_{\mathcal{M}}^2$ volume(D) as bound for the right-hand side of (13). This proves (14).

Going back to (13), we conclude that both the ν-norm of the residual flux-array and the μ-norm of the residual mmf-array tend to 0 as fast as $\gamma_{\mathcal{M}}$.

3.3 Stability

Although this is considered by many as sufficient in practice, we can't be satisfied with such "discrete energy" estimates. To really prove convergence, one should build from the DoF-arrays $\mathbf{b}_{\mathcal{M}}$ and $\tilde{\mathbf{h}}_{\mathcal{M}}$ an approximation $\{b_{\mathcal{M}}, \tilde{h}_{\mathcal{M}}\}$ of the pair of differential forms $\{b, \tilde{h}\}$, and prove that both the magnetic energy of the discrepancy $b_{\mathcal{M}} - b$ and the magnetic coenergy of $\tilde{h}_{\mathcal{M}} - \tilde{h}$ tend to 0 with $\mathcal{M}$.

To deal with such things, let us denote by $|\tilde{h}|_\mu$ and $|b|_\nu$, on the model of the previous $|\tilde{\mathbf{h}}|_\mu$ and $|\mathbf{b}|_\nu$, the square roots of the quantities $\int_D \tilde{\star}_\mu \tilde{h} \wedge \tilde{h}$ and $\int_D b \wedge \tilde{\star}_\nu b$.

So we are after some map, that we shall denote by $p_{\mathcal{M}}$, that would transform a flux-array $\mathbf{b}$ into a 2-form $p_{\mathcal{M}}\mathbf{b}$ and an mmf-array $\tilde{\mathbf{h}}$ into a twisted 1-form $p_{\mathcal{M}}\tilde{\mathbf{h}}$, and a satisfactory result would be that both $|b - p_{\mathcal{M}}\mathbf{b}_{\mathcal{M}}|_\nu$ and $|\tilde{h} - p_{\mathcal{M}}\tilde{\mathbf{h}}_{\mathcal{M}}|_\mu$ tend to 0 with $\mathcal{M}$ (convergence "in energy"). As next paragraph will show, sufficient conditions on $p_{\mathcal{M}}$ to this effect are the obvious consistency conditions:

$$(16)\qquad \begin{array}{l} p_{\mathcal{M}} r_{\mathcal{M}} b \to b, \text{ in energy, when } \mathcal{M} \to 0, \\ p_{\mathcal{M}} r_{\mathcal{M}} \tilde{h} \to \tilde{h}, \text{ in energy, when } \mathcal{M} \to 0, \end{array}$$

and the following inequalities:

$$(17)\qquad \alpha |p_{\mathcal{M}}\mathbf{b}|_\nu \leq |\mathbf{b}|_\nu, \quad \alpha |p_{\mathcal{M}}\tilde{\mathbf{h}}|_\mu \leq |\tilde{\mathbf{h}}|_\mu$$

for all $\mathbf{b}$ and $\tilde{\mathbf{h}}$, where the constant $\alpha > 0$ does not depend on $\mathcal{M}$. Since $|\mathbf{b}|_\nu$ and $|\tilde{\mathbf{h}}|_\mu$ depend on the discrete Hodge, this is a property of the approximation scheme, called *stability*.

Indeed, with both (14)(16) and (17), convergence is straightforward, thanks to (13): First, $p_{\mathcal{M}}(\mathbf{b} - r_{\mathcal{M}} b) \to 0$, by (17), then $p_{\mathcal{M}}\mathbf{b} \to b$, thanks to (16), all that "in energy". Same argument about $\tilde{h}$. This is Lax's celebrated folk theorem: *consistency + stability = convergence*.

So what about $p_{\mathcal{M}}$? Later, we shall find a systematic way to construct it, at least in the case of a tetrahedral primal mesh, the so-called *Whitney map*. If we don't insist right now on generality, there is an easy way to find this map in the case of DoF arrays $\mathbf{b}$ that satisfy $\mathbf{Db} = 0$, and luckily, only these do matter. The idea is to find a vector proxy $\mathbf{B}$ which be uniform inside each tetrahedron and such that its flux across each facet f be equal to $\mathbf{b}_f$. (Then, div$\mathbf{B} = 0$ all over D.) This, which would not be possible with cells of arbitrary shapes, can be done with tetrahedra, for there are, for each tetrahedral volume v, three unknowns (the components of $\mathbf{B}$) to four fluxes linked by one linear relation, $\sum_f \mathbf{D}_{vf}\mathbf{b}_f = 0$, so the problem has a solution. Hence $p_{\mathcal{M}}\mathbf{b}$.

We do have $p_{\mathcal{M}} r_{\mathcal{M}} b \to b$, then. This was proved long ago [9], by an argument which relies on mesh uniformity [17], and is very close to the one we now invoke to establish the stability condition (17). One has $|p_{\mathcal{M}}\mathbf{b}|_\nu^2 = \int_D \mu^{-1}|\mathbf{B}|^2$, which is obviously some quadratic form with respect to the facet fluxes, which we may therefore denote by $\langle \mathbf{b}, \mathbf{Nb} \rangle$, with $\mathbf{N}$ some square regular matrix. Now, suppose first a *single* tetrahedron in the mesh $\mathcal{M}$, and consider the Rayleigh-like quotient $\langle \mathbf{b}, \tilde{\star}_\nu \mathbf{b} \rangle / \langle \mathbf{b}, \mathbf{Nb} \rangle$. Its lower bound, strictly positive, depends only on the *shape* of the tetrahedron, not on its size. Uniformity of the family of meshes, then, allows us to take for α in (17) the smallest of these lower bounds, which is strictly positive and independent of $\mathcal{M}$. We may thereby conclude that $p_{\mathcal{M}}\mathbf{b}_{\mathcal{M}}$ converges towards b in energy.

No similar construction on the side of $\tilde{h}$ is

[4] Any known regularity of the mesh can be exploited, at this level, to obtain sharper bounds. In particular, for the kind of paving that Yee used, or generalizations of it [13], not only the constant part but the *linear* part of $\mathbf{H}$ gives a vanishing contribution. This accounts for a higher order of convergence in FDTD ($\gamma_{\mathcal{M}}^2$ rather than $\gamma_{\mathcal{M}}$, typically [16]) than what we find here.

available, but this is not such a handicap: if $p_{\mathcal{M}}\mathbf{b}_{\mathcal{M}} \to b$, then $\tilde{\star}_\nu p_{\mathcal{M}}\mathbf{b}_{\mathcal{M}} \to \tilde{h}$. This amounts to setting $p_{\mathcal{M}}$ on the dual side equal to $\tilde{\star}_\nu p_{\mathcal{M}}\tilde{\star}_\mu$. The problem with that is, $p_{\mathcal{M}}\tilde{\mathbf{h}}_{\mathcal{M}}$ fails to have the continuity properties we expect from a magnetic field: its vector proxy **H** is not tangentially continuous across facets, so one cannot take its curl. (One says of such a $p_{\mathcal{M}}$ that it constitutes a "non-conformal" approximation.) But never mind: In the case of a tetrahedral primal mesh, we have succeeded in proving the convergence in energy of $b_{\mathcal{M}}$ and $\tilde{h}_{\mathcal{M}}$ to b and $\tilde{h}$, which was our objective. And no Sobolev space has been invoked!

3.4 The dynamic case

Let us finish with a sketch of the convergence proof for the generalized Yee scheme of last issue.

First, linear interpolation in time between the values of the DoF arrays, as output by the Yee scheme, provides DoF-array-valued functions of time which converge, when δt tends to zero, towards the solution of the "spatially discretized" equations (1). This is not difficult.

Next, linearity of the equations permits to pass from the time domain to the frequency domain, via a Laplace transformation. Instead of studying (1), therefore, we may examine the behavior of the solution of

$$-p\,\tilde{\mathbf{D}} + \mathbf{R}^t\tilde{\mathbf{H}} = \tilde{\mathbf{J}}, \quad p\,\mathbf{B} + \mathbf{R}\mathbf{E} = 0, \tag{18}$$

$$\tilde{\mathbf{D}} = \tilde{\star}_\epsilon \mathbf{E}, \quad \mathbf{B} = \tilde{\star}_\mu \tilde{\mathbf{H}}, \tag{19}$$

when $\mathcal{M} \to 0$. Here, $p = \xi + \mathrm{i}\omega$, with $\xi > 0$, and small capitals denote Laplace transforms, which are arrays of *complex*-valued DoFs. If one can prove uniform convergence with respect to ω (which the requirement $\xi > 0$ makes possible), convergence of the solution of (1) will ensue, by inverse Laplace transformation. The main problem, therefore, is to compare E, B, $\tilde{\mathrm{H}}$, $\tilde{\mathrm{D}}$, as given by (18)(19), with $r_{\mathcal{M}}\mathrm{E}, r_{\mathcal{M}}\mathrm{B}, r_{\mathcal{M}}\tilde{\mathrm{H}}, r_{\mathcal{M}}\tilde{\mathrm{D}}$, where small capitals, again, denote Laplace transforms, but of differential forms this time.

The approach is similar to what we did in statics. First establish that

$$\begin{aligned} p\tilde{\star}_\mu(\tilde{\mathbf{H}} - r_{\mathcal{M}}\tilde{\mathrm{H}}) + \mathbf{R}(\mathbf{E} - r_{\mathcal{M}}\mathrm{E}) \\ = p(r_{\mathcal{M}}\tilde{\star}_\mu - \tilde{\star}_\mu r_{\mathcal{M}})\tilde{\mathrm{H}}, \end{aligned} \tag{20}$$

$$\begin{aligned} -p\tilde{\star}_\epsilon(\mathbf{E} - r_{\mathcal{M}}\mathrm{E}) + \mathbf{R}^t(\tilde{\mathbf{H}} - r_{\mathcal{M}}\tilde{\mathrm{H}}) \\ = -p(r_{\mathcal{M}}\tilde{\star}_\epsilon - \tilde{\star}_\epsilon r_{\mathcal{M}})\mathrm{E}. \end{aligned} \tag{21}$$

Then, right-multiply (20) by $(\tilde{\mathbf{H}} - r_{\mathcal{M}}\tilde{\mathrm{H}})^*$ and the conjugate of (21) by $-(\mathbf{E} - r_{\mathcal{M}}\mathrm{E})$, add. The middle terms (in $\mathbf{R}$ and $\mathbf{R}^t$) cancel out, and energy estimates follow. The similarity between the right-hand sides of (12), on the one hand, and (20)(21), on the other hand, shows that no further consistency requirements emerge. Stability, thanks to $\xi > 0$, holds there if it held in statics. What is a good Hodge discrete operator in statics, therefore, is a good one in transient situations. We may tentatively promote this remark as a heuristic principle:

As regards discrete constitutive laws, *what makes a convergent scheme for* static *problems will, as a rule, make one for the Maxwell evolution equations as well.*

From the perspective of the present series ("to build a finite-dimensional Maxwell house"), this is noteworthy. The idea was to replace all infinite-dimensional entities by finite-dimensional ones in consistent fashion: differential forms by DoF arrays, operator d by **G**, **R**, **D** and their transposes, depending on the degree, and the μ- and ϵ-related Hodge stars by appropriate square symmetric, positive definite matrices. All that in the hope that simple substitution of such "discrete" objects to "continuous" ones in the equations would generate valid approximation schemes. This working programme has succeeded, to some extent: We have a consistent diagonal discrete Hodge, at least for orthogonal meshes, and a convergence proof for the Yee-like scheme based on it, at least for tetrahedral primal meshes.

But the weak spot in all that is now apparent: We need a systematic way to pass from DoF arrays to differential forms—the $p_{\mathcal{M}}$ operator. Not only to interpolate inside volumes (this we could do without), but as a way to assess stability, in the above sense. Whitney forms, which will now enter the scene, provide this mechanism.

Acknowledgments

It's a pleasure to thank R. Hiptmair [12], L. Kettunen [6], and E. Tonti [20] for discussions on the topics treated here.

References

[1] I. Babuška, A.K. Aziz: "On the angle condition in the finite element method", **SIAM. J. Numer. Anal., 13,** 2 (1976), pp. 214-26.

[2] R.E. Bank, D.J. Rose: "Some Error Estimates for the Box Method", **SIAM J. Numer. Anal., 24,** 4 (1987), pp. 777-87.

[3] E. Bänsch: "Local Mesh Refinement in 2 and 3 Dimensions", **Impact of Computing Sci. Engrg., 3** (1991), pp. 181-91.

[4] J. Bey: "Tetrahedral Grid Refinement", **Computing, 55,** 4 (1995), pp. 355-78.

[5] A. Bossavit: **Computational Electromagnetism,** Academic Press (Boston), 1998.

[6] A. Bossavit, L. Kettunen: "Yee-like schemes on a tetrahedral mesh, with diagonal lumping", **Int. J. Numer. Modelling, 12,** (1999), pp. 129-42.

[7] H.L. de Cougny, M.S. Shephard: "Parallel refinement and coarsening of tetrahedral meshes", **Int. J. Numer. Meth. Engng., 46,** 7 (1999), pp. 1101-25.

[8] B. Courbet, J.P. Croisille: "Finite volume box schemes on triangular meshes", **M^2AN, 32,** 5 (1998), pp. 631-49.

[9] J. Dodziuk: "Finite-difference approach to the Hodge theory of harmonic forms", **Amer. J. Math., 98,** 1 (1976), pp. 79-104.

[10] T. Gallouet, J.P. Vila: "Finite volume element scheme for conservation laws of mixed type", **SIAM J. Numer. Anal., 28,** 6 (1991), pp. 1548-73.

[11] B. Heinrich: **Finite Difference Methods on Irregular Networks,** Akademie-Verlag (Berlin), 1987.

[12] R. Hiptmair: "Discrete Hodge operators", to appear.

[13] R. Holland: "Finite-Difference Solution of Maxwell's equations in generalized nonorthogonal coordinates", **IEEE Trans., NS-30,** 6 (1983), pp. 4589-95.

[14] Huang Jianguo, Xi Shitong: "On the finite volume element method for general self-adjoint elliptic problems", **SIAM J. Numer. Anal., 35,** 5 (1998), pp. 1762-74.

[15] J.M. Maubach: "Local bisection refinement for N-simplicial grids generated by reflection", **SIAM J. Sci. Stat. 16,** 1 (1995), pp. 210-27.

[16] P. Monk, E. Süli: "A convergence analysis of Yee's scheme on nonuniform grids", **SIAM J. Numer. Anal., 31,** 2 (1994), pp. 393-412.

[17] W. Müller: "Analytic torsion and R-torsion of Riemannian manifolds", **Advances in Mathematics, 28** (1978), pp. 233-305.

[18] Z. Ren: "Autogauging of vector potential by iterative solver—Numerical evidence", in **3d Int. Workshop on Electric and Magnetic Fields**, A.I.M. (31 Rue St-Gilles, Liège), 1996, pp. 119-24.

[19] E. Süli: "Convergence of finite volume schemes for Poisson's equation on nonuniform meshes", **SIAM J. Numer. Anal., 28,** 5 (1991), pp. 1419-30.

[20] E. Tonti: "Algebraic Topology and Computational Electromagnetism", in **4th Int. Workshop on Electric and Magnetic Fields**, A.I.M., 1998, pp. 284-294.