Computational electromagnetism and geometry:
Building a finite-dimensional "Maxwell's house"

Alain Bossavit, Électricité de France
1, Av. Gal de Gaulle, 92141 Clamart
Alain.Bossavit@der.edfgdf.fr

(1): Network equations

INTRODUCTION

In this new series,* we shall highlight some benefits of a geometrical approach to the Maxwell equations as regards their *numerical treatment*.

The first series, devoted to a discussion of possible formalisms for the mathematical description of electromagnetism, focused on one of them, in which the Maxwell equations appear as

$$(1)\quad \partial_t b + \mathrm{d}e = 0, \qquad (2)\quad -\partial_t \tilde{d} + \mathrm{d}\tilde{h} = \tilde{\jmath},$$

$$(3)\qquad b = \mu \tilde{\star} \tilde{h}, \qquad (4)\qquad \tilde{d} = \epsilon \tilde{\star} e.$$

The meaning of this notation will soon be recalled in detail. For the time being, it's enough to remember that the above equations are just, thinly disguised, the familiar ones,

$$(1')\quad \partial_t \mathbf{B} + \mathrm{rot}\,\mathbf{E} = 0, \quad (2')\quad -\partial_t \mathbf{D} + \mathrm{rot}\,\mathbf{H} = \mathbf{J},$$

$$(3')\qquad \mathbf{B} = \mu \mathbf{H}, \qquad (4')\qquad \mathbf{D} = \epsilon \mathbf{E},$$

which hold for non-moving isotropic media and a given current density **J**. We want to "discretize" these equations, which implies that each field, at any instant, will be represented by a *finite* number of real parameters, time-dependent, and that ordinary differential equations, in equal number, will somehow be derived for these "degrees of freedom" (DoF).

A key remark in this respect: the main equations (1)(2) or $(1')(2')$, are "conservation laws" of sorts. More precisely, (1) and (2) are local, differential versions of the following integral forms of the Faraday and Ampère laws,

$$(1'')\qquad \partial_t \int_\Sigma b + \int_{\partial\Sigma} e = 0,$$

$$(2'')\qquad -\partial_t \int_{\tilde{\Sigma}} \tilde{d} + \int_{\partial\tilde{\Sigma}} \tilde{h} = \int_{\tilde{\Sigma}} \tilde{\jmath},$$

to be satisfied for all surfaces Σ or $\tilde{\Sigma}$, inner- and outer-oriented respectively. Equation $(2'')$, for instance, says that the flux of the total current $\partial_t \tilde{d} + \tilde{\jmath}$ across $\tilde{\Sigma}$ matches the magnetomotive force, relative to the rim $\partial\tilde{\Sigma}$, of the magnetic field $\tilde{h}$, for all smooth surfaces $\tilde{\Sigma}$. Such a conservation statement can be seen as *one* equation relative to $\tilde{d}$ and $\tilde{h}$—one equation for each surface—so we have an infinity of equations, as befits unknown entities which have infinite dimension.

This is enough to suggest a method: Instead of requesting $(1'')$ and $(2'')$ for *all* surfaces, we shall be content with enforcing these balance relations over a *finite* set of surfaces, those generated by the facets of some finite-element mesh, and we shall attribute one degree of freedom to each "cell" (facet or edge, as the case may be) of this mesh. "Network equations", discrete analogues to (1)(2), will thus be found—forced on us, in fact. (Their analogy with the Kirchhoff equations for ordinary networks will be obvious.)

Appropriate relations between DoF's will also be needed in order to transcribe the constitutive laws, (3)(4) or $(3')(4')$. In other words, some "discrete Hodge operator" will have to be defined, leading to what one may call "network constitutive laws". There, in contrast, we'll have a large freedom of choice: There are good and less good discrete Hodges, and hence, choosing one will be the only really difficult part of the whole process.

This being done, the number of independent equations will, as we shall see, automatically

* Published in **J. Japan Soc. Appl. Electromagn. & Mech., 7** (1999), pp. 150-9 (no 1), pp. 294-301 (no 2), pp. 401-8 (no 3), and **8** (2000), pp. 102-9 (no 4), **8** (2000), pp. 203-9 (no 5), **8** (2000), pp. 372-7 (no 6). Some minor corrections have been done for the present printout.

match the number of DoF's, hence an "equivalent network", described by a system of ordinary differential equations. The latter can be solved by standard methods. (We shall concentrate on one of them, a simple "leapfrog scheme", on the model of the well-known Yee scheme [Ye].) This will leave us with the following questions: (1) How do we recover the fields, or rather, approximations of the exact fields, from the computed degrees of freedom? (2) How far are these approximate fields from the true ones?

Such questions are the bread-and-butter of numerical analysis. They arise for all methods. What is special here is the way the discrete equations are derived: To a large extent, finite elements are *not* needed for this. Equations are set up in a way which is quite reminiscent of the old "finite differences" approach (or of its modern "finite volumes" avatar). Yet, instead of being confined to hexahedral bricks, we can use cells of complex shape, thus accommodating bodies with curved or contorted boundaries at will—best of both worlds. Moreover, the method can be explained to anyone: No previous familiarity with finite elements is required, to the point that a complete, explicit recipe—one that any competent programmer can implement—can be formulated without ever mentioning finite elements!

This does not mean they are useless. But their rôle is confined to, mainly, assessing the value of the discrete Hodge operator by way of error estimates, and, secondarily, helping to reconstruct fields from their degrees of freedom, in the post-processing phase.

This rapid description of the approach is enough to perceive the benefits of our previous geometrization. In standard theory of electrical networks, there is a neat distinction between what may be referred to as *topology* (the way the network is connected, encoded in the coefficients of the so-called "node equations" and "loop equations") and what belongs to *metric* in our sense, that is, the values of the impedances of the network branches. One has all reasons in the world to maintain such orderly distinctions, for instance when branch impedances may vary with time, whereas the structure of the network doesn't change. In the theory to be developed here (which can be understood as the construction of two interlocked networks), a similar separation exists: Network equations, stemming from the metric-free equations (1) and (2), only depend on the combinatorial properties of the underlying mesh(es), and network constitutive laws, like the Hodge operator from which they derive, encompass metric information, as well as material properties.

So here is, in a few words, what we may expect: a unification of field theory with network theory, via discretization methods. An enticing goal, worth a long journey. Let's take the first steps of it.

1. NETWORK EQUATIONS

It's a good thing to keep in mind a representative of the family of problems one wishes to model. Here, we shall have wave-propagation problems in view, and the following example is typical.

1.1 A model problem

In a closed cavity with metallic walls (Fig. 1), which has been free from any electromagnetic field till time t = 0, suppose a flow of electric charge is created in an enclosed antenna after this instant, by some unspecified agency. An electromagnetic field then develops, propagating at the speed of light towards the walls which, as soon as they are reached by the wavefront, begin to act as secondary antennas. Dielectric parts inside the cavity, too, may scatter waves. Hence a complex evolution, which one may imagine simulating by numerical means.

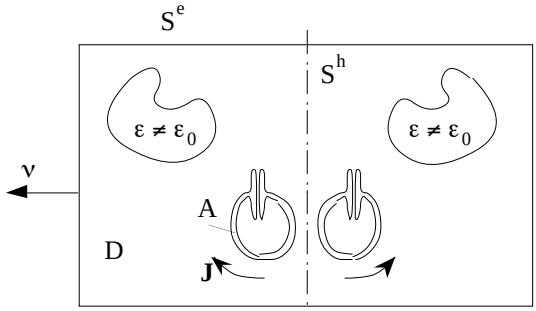

Figure 1. Situation and notation (dimension 3). Region D is the left half of the cavity. Its boundary S has a part S^e in the conductive wall and a part S^h in the symmetry plane. Region A, the left "antenna", is the support of the given current density **J** (mirrored on the right), for which some generator, not represented and not included in the modelling, is responsible.

For the sake of generality, let's assume a symmetry plane, and a symmetrically distributed current. The computation will thus be restricted

to a spatial domain D coinciding with one half of the cavity, on the left of the symmetry plane, say. Calling S its surface (in two parts, S^h and S^e, as Fig. 1 shows) and ν the outward directed field of normal unit vectors on S, the relevant equations, first expressed in standard notation, are $(1')$–$(4')$ above. Coefficients ϵ and μ are real, constant in time, but not necessarily equal to their vacuum values ϵ_0 and μ_0, and may therefore depend on variable x, which denotes the spatial position. (They could be tensors, too, without any serious extra difficulty.) The current density **J** is given, and assumed to satisfy $\mathbf{J}(t,x) = 0$, at all points x in D for $t \leq 0$. Other fields, unknown, are also supposed to be null before $t = 0$, hence initial conditions, $\mathbf{E}(0,x) = 0$ and $\mathbf{H}(0,x) = 0$ for all x. At the boundary,

$$(5')\ \ \nu \times \mathbf{E} = 0 \text{ on } S^e, \quad (6')\ \ \nu \times \mathbf{H} = 0 \text{ on } S^h.$$

Condition $(5')$ amounts to considering the cavity walls as perfect conductors, and $(6')$ comes from the mirror symmetry of **J**, which entails the *skew* symmetry of **H** ([B3], p. 28).

The mathematical theory of $(1')$–$(6')$ is not our concern here. Suffice it to say that, under reasonable assumptions about **J**, there is a unique solution $\{\mathbf{E}, \mathbf{H}\}$, that satisfies some standard requirements, apparently of mathematical nature, but actually dictated by physics. For instance, one wants the fields to have finite energy. This is translated, in mathematical terms,[1] by "**E**, at any time, should belong to the space, called $L^2(D)$, of square-summable vector fields over D", i.e., those for which the integral $\int_D |\mathbf{E}|^2$ is finite, and the same about **H**. For similar reasons, one requires $\mathrm{rot}\,\mathbf{E}$ and $\mathrm{rot}\,\mathbf{H}$ to be in $L^2(D)$, hence the convenience of the notation $L^2_{\mathrm{rot}}(D)$ for square-summable vector fields whose curl, too, is in $L^2(D)$. So both **E** and **H** are sought for in $L^2_{\mathrm{rot}}(D)$, at all times. Fields of this space happen to have a well-defined tangential part (one says a tangential "trace") on smooth surfaces. This gives sense to $(5')$–$(6')$, and entails the tangential continuity of **E** and **H** at material interfaces: Such "transmission conditions", often explicitly added to the set of equations, are here automatically enforced by the sole virtue of restricting the search to $L^2_{\mathrm{rot}}(D)$.

Equations $(1')$ and $(2')$ then imply that both **D** and **B** are in $L^2(D)$ with a divergence also in $L^2(D)$, and again the notation $L^2_{\mathrm{div}}(D)$ for such fields comes handy.[2] As a consequence, normal continuity of **D** and **B** at interfaces is enforced. Notice—it will be important later—that *exact* enforcement of the equations is necessary for this.

Now, assuming both **J** and $\mathrm{div}\,\mathbf{J}$ in $L^2(D)$, at all times, plus some smoothness of **J** with respect to time (details on these side issues can be found in [B1]), one can prove existence and uniqueness of the $\{\mathbf{E}, \mathbf{H}\}$ pair. Notice that $\mathrm{div}\,\mathbf{J} = 0$ is *not* assumed: some electric charge may accumulate at places in the antenna, in accordance with the charge-conservation equation $\partial_t q + \mathrm{div}\,\mathbf{J} = 0$, which results from (1) and from the relation $q = \mathrm{div}\,\mathbf{D}$, where q denotes charge density. This charge is not an independent data, but derives from **J** by integration in time, $q(t,x) = -\int_0^t (\mathrm{div}\,\mathbf{J})(s,x)ds$.

The relevance of $(1')$–$(6')$ as a realistic model-problem may be discussed, on several counts. For one, a term $\sigma\mathbf{E}$ might be introduced at the right-hand side of $(2')$ in order to account for the presence of conductive bodies inside the cavity. We refrain from this easy generalization for the sake of simplicity. Perhaps the assumption of a given current density in the antenna (which is routinely done) is a more serious issue, because the antenna is not insensitive to the reaction of its own radiated field, so we can't, in full rigor, know the antenna current in advance.

This point is made in [SS] in a comment on [UM]. Should one then feel compelled to model the dynamics of whatever drives charges in the antenna, including possibly the electrical network in background, and why not the whole universe? Authors of [UM] sensibly argue against that in their rejoinder [Um], and justify their informed guess of **J**. But even when **J** cannot be guessed about in advance with enough accuracy, $(1')$–$(6')$ can be considered

[1] There is no denying that mathematicians express things this way, in part, for their own comfort, because "Hilbert spaces", of which $L^2(D)$ is a well-known example, have nice properties. Notwithstanding, the close adequation of such abstractions to physics, often marvelled about [Wi], is especially obvious in the case of Maxwell's equations.

[2] Note that the same symbol, $L^2(D)$, serves here for a space of *scalar* fields and a space of *vector* fields, which is tolerable abuse.

as part of a coupled problem, for which it's legitimate to adopt **J** as an "interface parameter", the solution of $(1')$–$(6')$ thus becoming a subroutine in some higher-level iterative loop. All things considered, our model problem appears realistic enough, while being as simple as possible, as wave propagation problems go.

1.2 The model problem, in terms of differential forms

Next, let us translate this problem in the geometric language we acquired in the previous columns. Instead of the "proxy" vector fields **E** and **H**, we consider the differential forms e and $\tilde{h}$ they stand for.

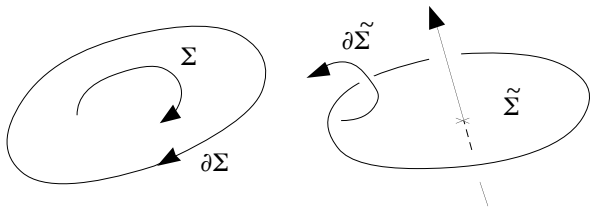

Figure 2. Inner- and outer-oriented surfaces Σ and $\tilde{\Sigma}$, with boundaries oriented in accordance.

Differential forms (DF's), one will remember, come in two varieties: the straight ones, like e and b, meant to be integrated over lines and surfaces with inner orientation, and the twisted ones, whose integrals make sense over lines or surfaces equipped with an outer orientation (a direction across the surface, in the case of $\tilde{d}$ and $\tilde{\jmath}$, a "way to turn around" lines, in the case of $\tilde{h}$). Therefore, eqs. $(1'')$ and $(2'')$ should be satisfied for all *inner* oriented surfaces Σ and *outer* oriented surfaces $\tilde{\Sigma}$ respectively (Fig. 2). As we know, these equations are equivalent, thanks to the Stokes theorem, to (1) and (2), themselves a translation of $(1')$–$(2')$. The nice thing about them is the absence of any reference to metric concepts. The latter are isolated in (3) and (4), where the Hodge operator $\tilde{\star}$, whose knowledge is equivalent to knowing the scalar product, maps ordinary p-forms to twisted $(n-p)$-forms and the other way round, where n is the dimension of ambient space.

What precedes is enough to let us carry on, and the rest of the present Section can be skipped. A few points about the relationships between DF's and their proxies, however, may be at their right place here.

In differential-geometric language, conditions $(5')$ and $(6')$ can be expressed very compactly, and without the recourse to metric (and orientation) that the " $\times$ " symbol may seem to imply. Starting from a 1-form e, consider the covectors defined, at points x of S, by $v \to \langle e(x), v\rangle$ for only those vectors v at x which are tangent to S. This defines a DF of degree 1, living on S, called the *trace* of e, and denoted $t_S e$, or of course just te in non-ambiguous cases. Its vector proxy is the tangential part of **E**, that we shall denote by $\mathbf{E}_S$, equal to $-\nu \times (\nu \times \mathbf{E})$. Same considerations about **H**. Conditions $(5')$–$(6')$ thus amount to

$$(5)\quad te = 0 \quad \text{on } S^e, \quad (6)\quad t\tilde{h} = 0 \quad \text{on } S^h.$$

This is more natural than $(5')$–$(6')$, in fact. Physically, it's indeed $\mathbf{E}_S$ that must vanish at a perfect conductor's boundary (and $\mathbf{H}_S$ at "magnetic walls", another case where $(6')$ would hold). People use "$\nu \times \mathbf{E} = 0$" as a way to say "$\mathbf{E}_S = 0$" without having to break pace to introduce this notation, or to use the more accurate but clumsy "$-\nu \times (\nu \times \mathbf{E}) = 0$".

The trace tb of b is, similarly, the field of 2-covectors $\{v, w\} \to \langle b(x); v, w\rangle$, defined for tangent vectors v and w. As a 2-form living on a 2-dimensional manifold, tb must have a *scalar* proxy, which one easily identifies as $\nu \cdot \mathbf{B}$, since $\langle b(x); v, w\rangle = \langle {}^2\mathbf{B}; v, w\rangle = \mathbf{B} \cdot v \times w = (\nu \cdot \mathbf{B})\ \nu \cdot (v \times w)$. Note the *unification* we have achieved: tangential part of this, normal part of that, are proxies for one and the same thing actually, the trace of a form. Imposing the trace of a form on a surface, as in (5) or (6), is a generalization to DF's of the "Dirichlet boundary condition" for functions, and thus deserves to be called that.

As an exercise, let us investigate the relations between t and d, in the case of 1-forms. (It goes the same way for higher degrees, and for twisted forms.) Let e be the 1-form, and Σ some surface. By the very definition of the integral (cf. [B3], p. 235), $\int_\Sigma t\,\mathrm{d}e = \int_\Sigma \mathrm{d}e$. By Stokes, this is $\int_{\partial\Sigma} e$, again equal, the way the integral was defined, to $\int_{\partial\Sigma} te$, which is $\int_\Sigma \mathrm{d}\,te$, by Stokes again. Since this holds for all smooth 1-forms and surfaces, we conclude that

$$(7)\qquad t\mathrm{d} = \mathrm{d}t.$$

Applying this to (5), and t to (1), we find (integrate in time and assume $b = 0$ at time 0) that $tb = 0$ on S^e: A perfectly conductive surface ($te = 0$) is also a barrier to the magnetic

flux ($tb = 0$). A nice example of notational tidying up!

Exercise. The reader is invited to work out the concept of "Neumann boundary condition", i.e., to try and define the *normal component* of a form on one side of a surface. (Hint: Play with the expression $\tilde{\star}t\tilde{\star}$.)

Although this is not crucial either to what follows, let's take the time to say how a priori restrictions such as "$\mathbf{E} \in L^2_{\mathrm{rot}}(D)$", and so forth, are expressed in terms of DF's. Taking the wedge product of the 1-form e with the twisted 2-form $\tilde{\star}e$, we obtain a twisted 3-form $e \wedge \tilde{\star}e$, which can be integrated over the 3-dimensional region D, hence a number $\int_D e \wedge \tilde{\star}e$, the square root of which is defined as the *norm* of the 1-form e. This works for all degrees (and of course relies on the metric). Notation $L^2(D)$ then takes on a new meaning: it's the Hilbert space of square-integrable p-forms. One may then define $L^2_{\mathrm{d}}(D)$ as the space of p-forms ω in $L^2(D)$ such that $\mathrm{d}\omega \in L^2(D)$. This is isomorphic, by passing to the vector proxies, to $L^2_{\mathrm{rot}}(D)$ if $p = 1$, and to $L^2_{\mathrm{div}}(D)$ if $p = 2$. Having no existence proofs in view, we shall not develop this, but it's comforting to know that all theorems, based on the Lax–Milgram lemma and similar things, that one can prove about vector-field solutions of Maxwell's equations, have precise counterparts, often much more compactly stated and proved, in terms of differential forms.

1.3 Primal mesh

Let's define what we shall call a "cellular paving". This is hardly different from a finite-element mesh, just a bit more general, but we need to be more fussy than usual about some details. Let's recall that V_n and A_n denote the real vector space of dimension n and its affine associate, and that when V_n is endowed with a dot product, whence a norm $|v| = \sqrt{v \cdot v}$, the distance this induces in A_n turns it into E_n, Euclidean space. Of course $n = 3$ in the sequel. By $\overline{A}$, we mean the *closure* of a set A in E_n, i.e., the set formed of all points whose distance to A is 0. *Closed* sets are those such that $\overline{A} = A$. Open sets are their complements, and the largest open set contained in A is its *interior*.

Start from the *open unit ball* B_p in V_p, that is, all vectors v such that $|v| < 1$. A *cell* of dimension p, or p-cell, for $1 \leq p \leq n$, is the image of B_p in A_n under some mapping c, piecewise smooth in both directions. (Notation c may refer to the image of the ball, or to the mapping itself, as convenient.) "Piecewise" leaves room for some irregularity: a 1-cell can well be a broken line, a 2-cell may have the shape of a triangle, a 3-cell the shape of a brick, etc. (Fig. 3).[3] Note that cells are *not* closed, for they don't contain their own boundaries. The case $p = 0$ is special: a 0-cell, by definition, is just a point,[4] that we shall call a *node*. A 1-cell will be an *edge*, a 2-cell a *facet* (we reserve "face" for another usage), and a 3-cell a *volume*.

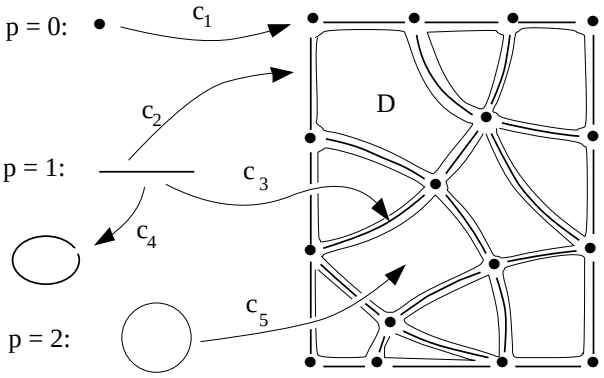

Figure 3. A few p-cells (caution, c_4 is *not* one of them), contributing to a closed cellular paving of D. (This should be imagined in dimension 3.)

Now, a *cellular paving* of some region R of space is a finite set of p-cells such that (1) Two distinct cells never intersect, (2) The union of all cells is R, (3) If the closures of two cells c and c' meet, their intersection is the closure of some (unique) cell c''. It may well happen that c'' is c, or c'. In such a case, e.g., if $\overline{c} \cap \overline{c'} = \overline{c}$, we say that c is *a face of* c'. For instance, on Fig. 3, c_3 is a face of c_5. If c is a face of c' which itself is a face of c'', then c is a face of c''.

We'll say we have a *closed* paving if R is closed. (Fig. 3 gives a two-dimensional example, where $R = \overline{D}$.) But it need not be so. Closed pavings are not necessarily what is needed in practice, as one may rather wish to discard some cells in order to deal with boundary conditions. Hence the usefulness of

[3] This notion of cell is slightly more restrictive than topologists would have it [HW]. For instance, to map B_1 (the segment $]-1,+1[$) to a closed loop minus one point (cf. c_4 in Fig. 3) would *not* make a cell in our sense.

[4] No inconsistency there: V_0 reduces to a single element, the null vector, and $B_0 \equiv V_0$.

the following notion of "relative closedness": C being a closed part of R, we shall say that a paving of R is *closed modulo* C if it can be obtained by removing, from some closed paving, all the cells which map into C. Fig. 4 displays the case we shall actually need, of a paving of $R = \overline{D} - S^e$ which is closed modulo S^e. Informally said, "pave $\overline{D}$ first, then remove all cells from the electric boundary".

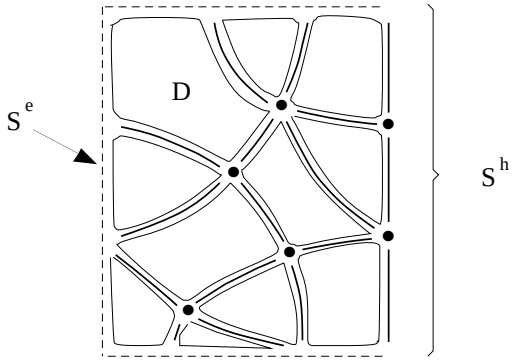

Figure 4. A culled paving, now "closed relative to" S^e. This is done in anticipation of the modelling we have in mind, in which cells of S^e would carry null DoF's, so they won't be missed.

Each cell is provided with an inner orientation of its own: Each edge has a "forward direction", each face has a notion of "turning clockwise" in it, each volume its own "corkscrew rule". These orientations are arbitrary and independent. For reasons soon to be disclosed, nodes must be oriented, too (a possibility we mentioned earlier, without using it). This consists in attributing a sign, $+$ or $-$, to each of them. (For simplicity, we may assume they all bear a $+$ sign.) We shall denote by $\mathcal{N}, \mathcal{E}, \mathcal{F}, \mathcal{V}$, the sets of oriented p-cells of the paving, and by N, E, F, V the number of cells in each of them.

Two cells σ and c, of respective dimensions p and $p+1$, are assigned an *incidence number*, equal to ± 1 if σ is a face of c, and to 0 otherwise. The sign, $+$ or $-$, depends on whether orientations "match" or not, a concept we have met before (cf. [B3], Section 3.3).[5] The boundary of c has a natural *outer* orientation, corresponding to the "inside to outside" crossing direction. Since σ belongs to this boundary, the inner orientation of c and this induced outer orientation of σ cooperate[5] in defining an *inner* orientation of σ. If this coincides with σ's *own* inner orientation, we say that the orientations of σ and c *match* (see Fig. 5). This understood, the sign rule is: $+$ if orientations match, $-$ if they don't.

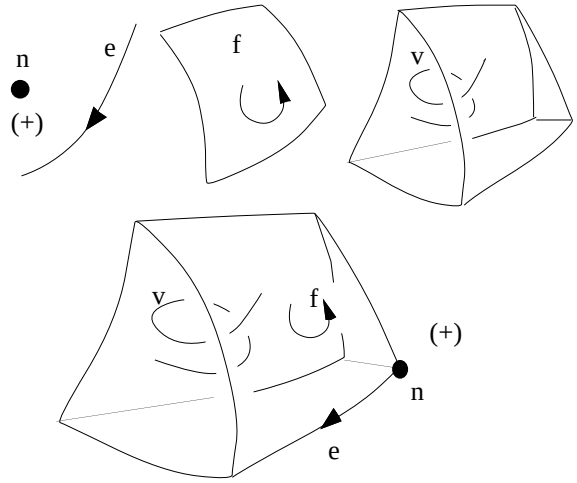

Figure 5. Top: individual oriented cells. Bottom: the same, as part of a paving, showing respective orientations. Here, orientations of v and f match, those of f and e, or of e and n, don't. So $\mathbf{G}_{ne} = -1, \mathbf{R}_{fe} = -1, \mathbf{D}_{vf} = 1$.

Collecting these numbers in arrays, we obtain rectangular matrices **G, R, D**, called *incidence matrices* of the tesselation. For instance (Fig. 5), the incidence number for edge e and facet f is denoted $\mathbf{R}_{fe}$, and makes one entry in matrix **R**, whose rows and columns are indexed over facets and edges, respectively. The entry $\mathbf{G}_{en}$ of **G** is -1, as explained in Note 5. Symbols **G, R, D** are of course intentionally reminiscent of grad, rot, div, but we still have a long way to go to fully understand the connection. Yet, one thing should be conspicuous already: contrary to grad, rot, div, the incidence matrices are *metric-independent* entities, so the analogy cannot be complete. Matrices **G, R, D** are more akin to the (metric-independent) operator d from this viewpoint, and the generic symbol **d**, indexed by the dimension p if needed, would make cleaner notation: $\mathbf{d}_0$ = **G**, $\mathbf{d}_1$ = **R**, $\mathbf{d}_2$ =

[5] Let's recall the essentials. An inner orientation of σ is a way to decide whether p independent vectors, tangent to σ at one of its points, form a direct frame of a skew frame. Take, at this point, a vector which goes *outward* with respect to c's boundary (an unambiguous notion), and list the p given vectors behind it. If the $(p+1)$-frame thus obtained is direct, with respect to the inner orientation of c, we decide the original p-frame was direct, and the other way round. If $p = 0$, σ is a node, c is an edge, and the rule specializes as follows: attribute the sign $+$ to σ if the outgoing vector tangent to c, at the end-point σ, goes in the same direction as c itself.

D. The mnemonic value of **G, R, D**, however, justifies the abuse.

Let's only point out that, just as rot ∘ grad = 0 and div ∘ rot = 0, one has **GR** = 0 and **DR** = 0. Indeed, for an edge e and a volume v, the $\{v, e\}$-entry of **DR** is $\sum_{f \in \mathcal{F}} \mathbf{D}_{vf}\mathbf{R}_{fe}$. Nonzero terms occur, in this sum over facets, only for facets which at once contain e and are a face of v, which happens only if e belongs to $\overline{v}$. In that case, there are exactly two facets f and g of v hinging on e (Fig. 6), and hence two nonzero terms. As Fig. 6 shows, they have opposite signs, whatever the orientations of the individual cells, hence the result, **DR** = 0. By a similar proof, **RG** = 0, and more generally, $\mathbf{d}_{p+1}\mathbf{d}_p = 0$.[6]

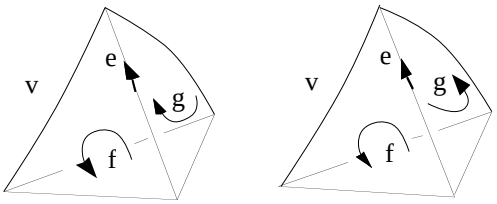

Figure 6. Opposition of incidence numbers, leading to **DR** = 0, whatever the orientations.

Remark. The answer to the natural question, "then, is the kernel **R** equal to the range of **G** ?", is "yes" here, because $\overline{D} - S^e$ has simple topology. Otherwise, this would lead us far into *homology*, a branch of topology which studies the global topological properties of manifolds by first chopping them into cells, then looking at the algebraic properties of the incidence matrices. (See, e.g., [Ar].) ♢

The whole algebraic structure composed of the sets $\mathcal{N}, \mathcal{E}, \mathcal{F}, \mathcal{V}$ and of the incidence matrices is called a "cell complex". This, plus the (later quite necessary) details about each cell's map, forms what we shall denote $\mathcal{M}$, and call, informally, a *mesh of* D, the *primal* mesh in the theory.

1.4 Dual mesh

The *dual* mesh of D is also a cellular paving, though not of the same region exactly, and with *outer* orientation of cells. Let's explain.

To each p-cell c of the primal mesh, we assign a unique $(n-p)$-cell, meeting c at a single point, called the *dual* of c, and denoted $\tilde{c}$. Hence a 1–1 correspondence between cells of complementary dimensions. Thus, for instance, facet f is pierced by the dual edge $\tilde{f}$ (a line), node n is inside the dual volume $\tilde{n}$, and so forth. Since, at the common point, the tangent spaces to the primal cell c and the dual cell $\tilde{c}$ are complementary ([B3], p. 26), the inner orientation of c provides an outer orientation for $\tilde{c}$ (Fig. 7). Incidence matrices $\tilde{\mathbf{G}}, \tilde{\mathbf{R}}, \tilde{\mathbf{D}}$ can then be defined, as above, the sign of each nonzero entry depending on whether outer orientations match or not.

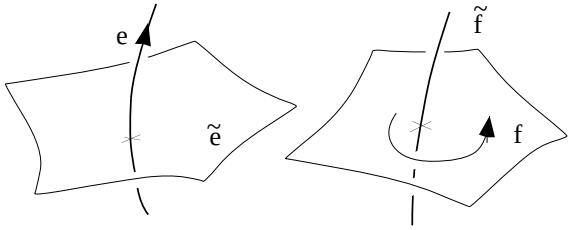

Figure 7. Inner orientations of edge e and facet f, respectively, give crossing direction through $\tilde{e}$ and gyratory sense around $\tilde{f}$.

Moreover, it is required that, when c is a face of c', the dual $\tilde{c}'$ be a face of $\tilde{c}$, and the other way round. This has two consequences. First, we don't really need new names for the dual incidence matrices. Indeed, consider for instance edge e and facet f, and suppose $\mathbf{R}_{fe} = 1$, i.e., e is a face of f and their orientations match: Then the dual edge $\tilde{f}$ is a face of the dual facet $\tilde{e}$, whose outer orientations match, too. So what we would otherwise denote $\tilde{\mathbf{R}}_{\tilde{e}\tilde{f}}$ is equal to $\mathbf{R}_{fe}$. Same reasoning with the opposite signs, and for other kinds of cells, from which we conclude that the would-be dual incidence matrices $\tilde{\mathbf{G}}, \tilde{\mathbf{R}}, \tilde{\mathbf{D}}$ are just the transposes $\mathbf{D}^t, \mathbf{R}^t, \mathbf{G}^t$ of the primal ones.

Second consequence, there is no gap between dual cells, which thus form a cellular paving of a connected region $\tilde{R}$, the interior $\tilde{D}$ of which is nearly D, but not quite (Fig. 8). A part of its boundary is paved by dual cells: We name it $\tilde{S}^e$, owing to its kinship with S^e (not so obvious on our coarse drawing! but the finer the mesh,[7]

[6] It's no accident if this evokes the proof of Stokes theorem we saw in [B3], Section 3.4. The same basic observation, "the boundary of a boundary is zero" [TW, KW], underlies both proofs.

[7] A *refinement* of a paving is another paving of the same

the closer $\tilde{S}^e$ and S^e will become). The other part is denoted $\tilde{S}^h$. So the cellular paving we now have is closed modulo $\tilde{S}^h$, whereas the primal one was closed modulo S^e. The whole structure, again, is called the *dual mesh*, denoted by $\widetilde{\mathcal{M}}$.

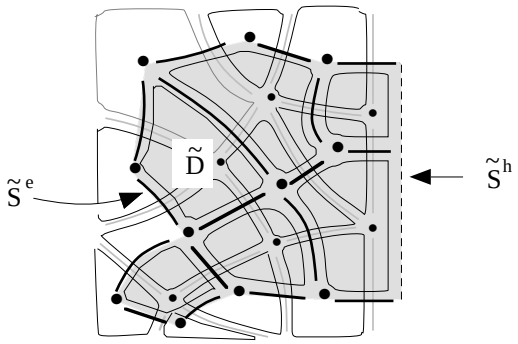

Figure 8. A dual paving, overlaid on the primal one.

Given $\mathcal{M}$, all its duals have the same *combinatorial* structure (the same incidence matrices), but can differ as regards *metric*, which leaves much leeway to construct dual meshes. Two approaches are noteworthy, which lead to the "barycentric dual" and the "Voronoi–Delaunay dual". We shall present them as special cases of slightly more general procedures, the "star construction" and the "orthogonal construction" of meshes in duality. We shall consider only *polyhedral* meshes (those with polyhedral 3-cells), which is not overly restrictive in practice.

The orthogonal construction is an old idea, developed by the young Maxwell[8] [Ma]. It works for "straight" primal cells, i.e., with straight edges and plane polygonal facets. Let dual cells be straight, too, each orthogonal to its primal partner. Figure 9 gives a 2D example. A particular case is the *Voronoi–Delaunay* tesselation: Start from a set $\mathcal{N}$ of would-be primal nodes; for each node n, located at point x_n, build the *Voronoi cell*

$$\tilde{n} = \{x : |x - x_n| < |x - x_m| \ \ \forall\, m \neq n\},$$

comprising all points closer to n than to other nodes;[9] for two nodes m and n, take

$$\{x : |x-x_n| = |x-x_m| < |x-x_l| \ \forall\, l \neq n, l \neq m\}$$

which, if non-empty, will be the dual 2-cell for edge $\{m, n\}$, and so on. If things go well, two meshes in duality are thus obtained. (See [B2], p. 107, for more details and references.) Generically,[10] the primal mesh is then a *simplicial* one, as its volumes are tetrahedra.

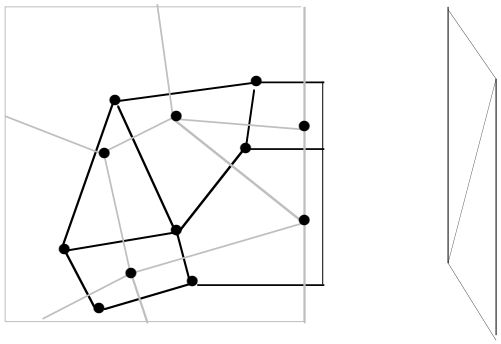

Figure 9. Left: Orthogonal dual mesh. (Same graphic conventions as in Fig. 8, slightly simplified.) Right: Likely the simplest example of a 2D mesh with no orthogonal dual.

Alas, such pavings, whose virtues will be obvious when we study the discrete Hodge operator, are notoriously difficult to build. Even the less stringent condition of orthogonality can be impossible to satisfy, if the primal mesh is imposed (Fig. 9). If one starts from a simplicial primal with only *acute* dihedral angles,[11] all goes well. But this property, which we shall see is desirable, is not so easily obtained, and certainly not warranted by common mesh generators.

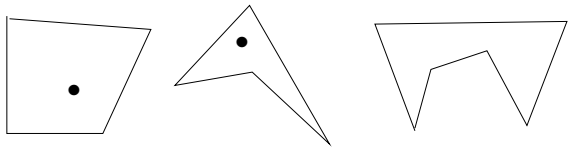

Figure 10. Two star-shaped polygons (left), and one which is not (right). Dots mark eligible centers.

Hence the usefulness of the star construction, more general, because it applies to any primal mesh with star-shaped cells. Recall that a part A of E_3 is *star-shaped* (Fig. 10) if it contains

region, which restricts to a proper cellular paving of each original cell.

[8] Working on elasticity, Maxwell had done the same kind of separation between non-metric and metric notions that we rediscover nowadays in analysing Maxwell equations. Read the masters ...

[9] The idea dates back to Dirichlet [Di].

[10] Meaning that, if not so, a slight displacement of nodes will make it so.

[11] The angles between facets.

a point a, that we shall call a *center*, such that the whole segment $[a, x]$ belongs to A when x belongs to A. Now, take a center for each primal cell (the center of a primal node is itself), and join it to centers of all faces of its cell. This way, simplicial *subcells* are obtained (tetrahedra and their faces, in 3D). One gets the dual mesh by rearranging them, as follows: for each primal cell c, build its dual by putting together all k-subcells, $k \leq n - p$, which have one of their vertices at c's center, and other vertices at centers of cells incident on c. Figures 11 and 12 give the idea.

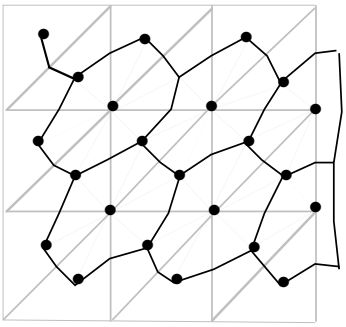

Figure 11. Star construction of a dual mesh (quite close, here, to a barycentric mesh). Notice the isolated dual edge, and the arbitrariness in shaping dual cells beyond S^h.

Remark. The recipe is imprecise about cells dual to those of S^h, whose shape outside D can be as one fancies (provided the requirements about duality are satisfied). Nothing there to worry about: Such choices are just as arbitrary as the selection of the cell centers. It's all part of the unavoidable approximation error, which can be reduced at will by refinement (insofar as computing resources are there). $\diamondsuit$

In the case of a simplicial mesh, cell barycenters make convenient centers, hence the *barycentric dual*. It's a well-known structure, but visualizing it in three dimensions may be not so easy. Cf. Fig. 12.

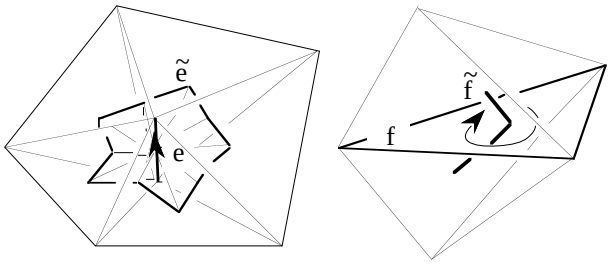

Figure 12. A dual facet and a dual edge, in the case of a simplicial mesh and of its barycentric dual. Observe the orientations.

Remark. If the primal mesh has been obtained by restriction of a closed one, as suggested above ("pave $\overline{D}$ first ..."), subcells built from the latter form a refinement of *both* the primal mesh and the dual mesh. The existence of this "underlying simplicial complex" will be important later. $\diamondsuit$

1.5 The network equations

We now want to apply the principle described in the Introduction: Satisfy the balance equations $(1'')$–$(2'')$ for a selected *finite* family of surfaces.

Let's first adopt a finite, approximate representation of the fields. Consider b, for instance. As a 2-form, it is meant to be integrated over oriented surfaces. So one may consider the integrals $\int_f b$, denoted $\mathbf{b}_f$, for all facets f, as a kind of "sampling" of b, and take the "DoF-vector" $\{\mathbf{b} = \mathbf{b}_f : f \in \mathcal{F}\}$, indexed over primal facets, as a finite representation of b. This does not tell us about the *value* of the field at any given point, of course. But is that the objective? Indeed, all we know about a field is what we can measure, and we don't measure point values. These are abstractions. What we do measure is the *flux* of b—or rather, its variations, but never mind—by reading off the induced emf along the loop of a small enough magnetic probe. The above sampling thus consists in having each facet of the mesh play the role of such a probe, and the smaller the facets, the better we know the field. Conceivably, the mesh may be made so fine that the $\mathbf{b}_f$'s are *sufficient information* about the field, in practice. So one may be content with a method that would yield the four meaningful arrays of degrees of freedom, listing

- the edge emf's, $\mathbf{e} = \{\mathbf{e}_e : e \in \mathcal{E}\}$,
- the facet fluxes, $\mathbf{b} = \{\mathbf{b}_f : f \in \mathcal{F}\}$,
- the dual-edge mmf's, $\tilde{\mathbf{h}} = \{\tilde{\mathbf{h}}_f : f \in \mathcal{F}\}$,
- and the dual-face displacement currents, $\tilde{\mathbf{d}} = \{\tilde{\mathbf{d}}_e : e \in \mathcal{E}\}$,

all that from a similar sampling, across dual facets, of the given current $\tilde{j}$, encoded in the DoF array $\tilde{\mathbf{j}} = \{\tilde{\mathbf{j}}_e : e \in \mathcal{E}\}$.

Next, suppose the surface Σ in $(1'')$ is the simplest possible one in the present context, that is, a *single* primal facet, f. The (inner) orientation of f confers an orientation to its boundary ∂f. The integral of e along ∂f,

by linearity, would be the sum of its integrals along edges that make f, if the orientations of these were compatible with the orientation of ∂f. But when orientations don't match, multiplying by $\mathbf{R}_{fe}$ restores the right value, by the very definition of these incidence numbers. Therefore, eq. $(1'')$ applied to f yields

$$\partial_t \mathbf{b}_f + \sum_{e \in \mathcal{E}} \mathbf{R}_{fe} \mathbf{e}_e = 0. \tag{8}$$

There is one equation like this for each facet of the primal mesh, that is—thanks for having discarded facets in S^e, for which the flux is known to be 0—one for each genuinely unknown facet-flux of b. We may now express (8) in matrix form, like this:

$$\partial_t \mathbf{b} + \mathbf{Re} = 0, \tag{9}$$

the first group of our *network* differential *equations*.

Finally, the same reasoning about each dual facet $\tilde{e}$ (the simplest possible outer-oriented surface that $\tilde{\Sigma}$ in $(2'')$ can be) yields

$$-\partial_t \tilde{\mathbf{d}}_e + \sum_{f \in \mathcal{F}} \mathbf{R}_{fe} \tilde{\mathbf{h}}_f = \tilde{\mathbf{j}}_e, \tag{10}$$

for all e in $\mathcal{E}$, i.e., in matrix form,

$$-\partial_t \tilde{\mathbf{d}} + \mathbf{R}^t \tilde{\mathbf{h}} = \tilde{\mathbf{j}}, \tag{11}$$

the second group of network equations.

If a field $e, b, \tilde{h}, \tilde{d}$, described by its DoF-arrays $\mathbf{e}, \mathbf{b}, \tilde{\mathbf{h}}, \tilde{\mathbf{d}}$, satisfies (10) and (11), it automatically satisfies the balance equations $(1'')$–$(2'')$ for, respectively, all inner-oriented surfaces made of 2-cells and outer-oriented surfaces made of dual 2-cells. Our first objective is thus achieved.

WHAT NEXT?

There are F scalar equations in (10), and E in (11), for $2(E+F)$ unknowns. So we miss $E+F$ equations: one relation between $\mathbf{e}$ and $\tilde{\mathbf{d}}$ for each edge, one relation between $\mathbf{b}$ and $\tilde{\mathbf{h}}$ for each facet. How to get them most simply? This is the problem of the discrete Hodge operator, next column's subject.

References

[Ar] M.A. Armstrong: **Basic Topology,** McGraw–Hill (London), 1979.

[B1] A. Bossavit: "Existence and uniqueness in the full Maxwell system", in **Int. Workshop on Electric and Magnetic Fields** (Proc., Liège, 28–30 Sept. 1992), A.I.M. (31 Rue St-Gilles, Liège), 1992, pp. 35.1-4.

[B2] A. Bossavit: **Computational Electromagnetism,** Academic Press (Boston), 1998.

[B3] A. Bossavit: "On the geometry of electromagnetism", **J. JSAEM, 6** (1998), pp. 17-28, 114-23, 233-40, 318-26.

[Di] G.L. Dirichlet: "Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen", **J. Reine Angew. Math., 40** (1850), p. 209

[HW] P.J. Hilton, S. Wylie: **Homology Theory,** An Introduction to Algebraic Topology, Cambridge U.P. (Cambridge), 1965.

[KW] A. Kheyfets, J.A. Wheeler: "Boundary of a Boundary and Geometric Structure of Field Theories", **Int. J. Theor. Phys., 25,** 6 (1986), pp. 573-80.

[Ma] J.C. Maxwell: "On Reciprocal Figures, Frames, and Diagrams of Forces", **Trans. Roy. Soc. Edinburgh, 26** (1869), pp. 1-40.

[SS] J.B. Smyth, D.C. Smyth: "Critique of the paper 'The Electromagnetic Radiation from a Finite Antenna'", **Am. J. Phys., 45,** 6 (1977), pp. 581-2.

[TW] E.F. Taylor, J.A. Wheeler: **Spacetime physics,** Freeman (New York), 1992.

[UM] M.A. Uman, D.K. McLain, E.Ph. Krider: "The Electromagnetic Radiation from a Finite Antenna", **Am. J. Phys., 43**, 1 (1975), pp. 33-8.

[Um] M.A. Uman: "Reply to Smyth and Smyth", **Am. J. Phys., 45,** 6 (1977), p. 582.

[Wi] E.P. Wigner: "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", **Comm. Pure Appl. Math., 13** (1960), pp. 1-14.

[Ye] K.S. Yee: "Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media", **IEEE Trans., AP-14,** 3 (1966), pp. 302-7.