Discretization of Electromagnetic Problems: The "Generalized Finite Differences" Approach

Alain Bossavit

Laboratoire de Génie Électrique de Paris,
11 Rue Joliot-Curie,
91192 Gif-sur-Yvette Cedex,
France
E-mail address: Bossavit@lgep.supelec.fr

Numerical Methods in Electromagnetics
Special Volume (W.H.A. Schilders and E.J.W. ter Maten, Guest Editors) of
HANDBOOK OF NUMERICAL ANALYSIS, VOL. XIII
P.G. Ciarlet (Editor)

ISSN 1570-8659
DOI 10.1016/S1570-8659(04)13002-0

Contents

CHAPTER I

Preliminaries: Euclidean Space

What we shall do in this preliminary chapter (Sections 1–5, out of a total of 25) can be described as "deconstructing Euclidean space". Three-dimensional Euclidean space, denoted by E_3 here, is a relatively involved mathematical structure, made of an affine 3D space (more on this below), equipped with a metric and an orientation. By taking the Cartesian product of that with another Euclidean space, one-dimensional and meant to represent Time, one gets the mathematical framework in which most of classical physics is described. This framework is often taken for granted, and should not.

By this we do not mean to challenge the separation between space and (absolute) time, which would be getting off to a late start, by a good century. Relativity is not our concern here, because we won't deal with moving conductors, which makes it all right to adopt a privileged reference frame (the so-called laboratory frame) and a unique chronometry. The problem we perceive is with E_3 itself, too rich a structure in several respects. For one thing, orientation of space is *not* necessary. (How could it be? How could physical phenomena depend on this social convention by which we class right-handed and left-handed helices, such as shells or staircases?) And yet, properties of the cross product, or of the curl operator, so essential tools in electromagnetism, crucially depend on orientation. As for metric (i.e., the existence of a dot product, from which norms of vectors and distances between points are derived), it also seems to be involved in the two main equations, $\partial_t \mathrm{B} + \mathrm{rot}\,\mathrm{E} = 0$ (Faraday's law) and $-\partial_t \mathrm{D} + \mathrm{rot}\,\mathrm{H} = \mathrm{J}$ (Ampère's theorem), since the definition of rot depends on the metric. We shall discover that it plays no role there, actually, because a change of metric, in the description of some electromagnetic phenomenon, would change *both* rot *and* the vector fields E, B, etc., in such a way that the equations would stay unchanged. Metric is no less essential for that, but its intervention is limited to the expression of constitutive laws, that is, to what will replace in our notation the standard $\mathrm{B} = \mu\mathrm{H}$ and $\mathrm{D} = \varepsilon\mathrm{E}$.[1]

Our purpose, therefore, is to separate the various layers present in the structure of E_3, in view of using exactly what is needed, and nothing more, for each subpart of the Maxwell system of equations. That this can be done is no news: As reported by POST [1972], the metric-free character of the two main Maxwell equations was pointed out by Cartan, as early as 1924, and also by KOTTLER [1922] and VAN DANTZIG [1934]. But the exploitation of this remark in the design of numerical schemes is

[1] We shall most often ignore Ohm's law here, for shortness, and therefore, treat the current density J as a data. It would be straightforward to supplement the equations by the relation $\mathrm{J} = \sigma\mathrm{E} + \mathrm{J}^s$, where only the "source current" J^s is known in advance.

a contemporary thing, which owes much to (again, working independently) TONTI [2001], Tonti (see TONTI [1996], MATTIUSSI [2000]) and Weiland (see EBELING, KLATT, KRAWCZYK, LAWINSKY, WEILAND, WIPF, STEFFEN, BARTS, BROWMAN, COOPER, DEAVEN and RODENZ [1989], WEILAND [1996]). See also SORKIN [1975], HYMAN and SHASHKOV [1997], TEIXEIRA and CHEW [1999]. Even more recent (BOSSAVIT and KETTUNEN [1999], MATTIUSSI [2000]) is the realization that such attention to the underlying geometry would permit to soften the traditional distinctions between finite-difference, finite-element, and finite-volume approaches. In particular, it will be seen here that a common approach to error analysis applies to the three of them, which does rely on the existence of finite elements, but not on the variational methods that are often considered as foundational in finite element theory. These finite elements, moreover, are not of the Lagrange (node based) flavor. They are differential geometric objects, created long ago for other purposes, the Whitney forms (WHITNEY [1957]), whose main characteristic is the interpretation they suggest of degrees of freedom (DoF) as integrals over geometric elements (edges, facets, ...) of the discretization mesh.

As a preparation to this deconstruction process, we need to recall a few notions of geometry and algebra which do not seem to get, in most curricula, the treatment they deserve. First on this agenda is the distinction between vector space and affine space.

1. Affine space

A *vector space*[2] on the reals is a set of objects called *vectors*, which one can (1) add together (in such a way that they form an Abelian group, the neutral element being the null vector) and (2) multiply by real numbers. No need to recall the axioms which harmonize these two groups of features. Our point is this: The three-dimensional vector space (for which our notation will be V_3) makes an awkward model of physical space,[3] unless one deals with situations with a privileged point, such as for instance a center of mass, which allows one to identify a spatial point x with the translation vector that sends this privileged point to x. Otherwise, the idea to add points, or to multiply them by a scalar, is ludicrous. On the other hand, taking the midpoint of two points, or more generally, barycenters, makes sense, and is an allowed operation in affine space, as will follow from the definition.

An *affine space* is a set on which a vector space, considered as an additive group, acts effectively, transitively and regularly. Let's elaborate.

A group G *acts* on a set X if for each $g \in G$ there is a map from X to X, that we shall denote by a_g, such that a_1 is the identity map, and $a_{gh} = a_g a_h$. (Symbol 1 denotes

[2]Most definitions will be implicit, with the defined term set, on first appearance, in *italics* style. The same style is also used, occasionally, for emphasis.

[3]Taking $\mathbb{R}^3$, the set of triples of real numbers, with all the topological and metric properties inherited from $\mathbb{R}$, is even worse, for this implies that some basis $\{\partial_1, \partial_2, \partial_3\}$ has been selected in V_3, thanks to which a vector v writes as $v = \sum_i v^i \partial_i$, hence the identification between v and the triple $\{v^i\}$ of components (or coordinates of the point v stands for). In most situations which require mathematical modelling, no such basis imposes itself. There may exist privileged directions, as when the device to be modelled has some kind of translational invariance, but even this does not always mandate a choice of basis.

the neutral element, and will later double for the group made of this unique element.) The action is *effective* if $a_g = 1$ implies $g = 1$, that is to say, if all nontrivial group elements "do something" to X. The *orbit* of x under the action is the set $\{a_g(x)\colon\ g \in G\}$ of transforms of x. Belonging to the same orbit is an equivalence relation between points. One says the action is *transitive* if all points are thus equivalent, i.e., if there is a single orbit. The *isotropy group* (or stabilizer, or little group) of x is the subgroup $G_x = \{g \in G\colon\ a_g(x) = x\}$ of elements of G which fix x. In the case of a transitive action, little groups of all points are conjugate (because $g_{xy} G_y = G_x g_{xy}$, where g_{xy} is any group element whose action takes x to y), and thus "the same" in some sense. A transitive action is *regular* (or *free*) if it has no fixed point, that is, if $G_x = 1$ for all x. If so is the case, X and G are in one-to-one correspondence, so they look very much alike. Yet they should not be identified, for they have quite distinctive structures. Hence the concept of *homogeneous space*: A set, X here, on which some group acts transitively and effectively. (A standard example is given by the two-dimensional sphere S_2 under the action of the group SO_3 of rotations around its center.) If, moreover, the little group is trivial (regular action), the only difference between the homogeneous space X and the group G lies in the existence of a distinguished element in G, the neutral one. Selecting a point 0 in X (the origin) and then identifying $a_g(0)$ with g (and hence 0 in X with the neutral element of G) provides X with a group structure, but the isomorphism with G thus established is not canonical, and this group structure is most often irrelevant, just like the vector-space structure of 3D space.

Affine space is a case in point. Intuitively, take the n-dimensional vector space V_n, and forget about the origin: What remains is A_n, the affine space of dimension n. More rigorously, a vector space V, considered as an additive group, acts on itself (now considered as just a set, which we acknowledge by calling its elements *points*, instead of vectors) by the mappings[4] $a_v = x \to x + v$, called *translations*. This action is transitive, because for any pair of points $\{x, y\}$, there is a vector v such that $y = x + v$, and regular, because $x + v \neq x$ if $v \neq 0$, whatever x. The structure formed by V as a set equipped with this group action is called the *affine space A associated with V*. Each vector of V has thus become a point of A, but there is nothing special any longer with the vector 0, as a point in A. Reversing the viewpoint, one can say that an affine space A is a homogeneous space with respect to the action of some vector space V, considered as an additive group. (Points of A will be denoted x, y, etc., and $y - x$ will stand, by a natural notational abuse, for the vector that carries x to y.) The most common example is obtained by considering as equivalent, in some vector space V, two vectors u and v such that $u - v$ belong to some fixed vector subspace W. Each equivalence class has an obvious affine structure (W acts on it regularly by $v \to v + w$). Such a class is called an *affine subspace* of V, *parallel to* W[5] (see Fig. 1.1) Of course, no vector in such an

[4] We'll find it convenient to denote a map f by $x \to \mathrm{Expr}(x)$, where Expr is the defining expression, and to link name and definition by writing $f = x \to \mathrm{Expr}(x)$. (The arrow is a "stronger link" than the equal sign in this expression.) In the same spirit, $X \to Y$ denotes the set of all maps "of type $X \to Y$", that is, maps from X to Y, not necessarily defined over all X. Points x for which f is defined form its *domain* $\mathrm{dom}(f) \subset X$, and their images form the *codomain* $\mathrm{cod}(f) \subset Y$, also called the *range* of f.

[5] Notice how the set of all affine subspaces parallel to W also constitutes an affine space under the action of V, or more pointedly – because then the action is regular – of the quotient space V/W. A "point", there, is a whole affine subspace.

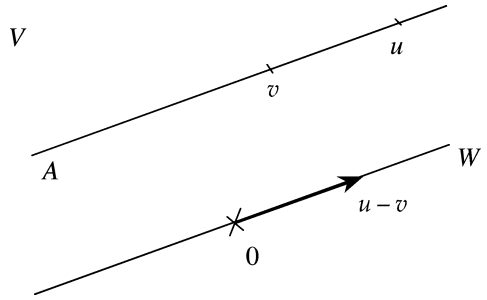

FIG. 1.1. No point in the affine subspace A, parallel to W, can claim the role of "origin" there.

affine subspace qualifies more than any other as origin, and calling its elements "points" rather than "vectors" is therefore appropriate.

At this stage, we may introduce the *barycenter* of points x and y, with weights λ and $1-\lambda$, as the translate $x+\lambda(y-x)$ of x by the vector $\lambda(y-x)$, and generalize to any number of points. The concepts of affine independence, dimension of the affine space, and affine subspaces follow from the similar ones about the vector space. *Barycentric coordinates*, with respect to $n+1$ affinely independent points $\{a_0,\ldots,a_n\}$ in A_n are the weights $\lambda^i(x)$ such that $\sum_i \lambda^i(x)=1$ and $\sum_i \lambda^i(x)(x-a_i)=0$, which we shall feel free to write $x=\sum_i \lambda^i(x)a_i$. *Affine maps* on A_n are those that are linear with respect to the barycentric coordinates. If x is a point in affine space A, vectors of the form $y-x$ are called *vectors at* x. They form of course a vector space isomorphic to the associate V, called the *tangent space at* x, denoted T_x. (I will call *free* vectors the elements of V, as opposed to vectors "at" some point, dubbed *bound* (or *anchored)* vectors. Be aware that this usage is not universal.) The tangent space to a curve or a surface which contains x is the subspace of T_x formed by vectors at x tangent to this curve or surface.[6] Note that vector fields are maps of type *POINT* $\rightarrow$ *BOUND_VECTOR*, actually, subject to the restriction that the value of v at x, notated $v(x)$, is a vector at x. The distinction between this and a *POINT* $\rightarrow$ *FREE_VECTOR* map, which may seem pedantic when the point spans ordinary space, must obviously be maintained in the case of tangent vector fields defined over a surface or a curve.

Homogeneous space is a key concept: Here is the mathematical construct by which we can best model humankind's *physical* experience of spatial homogeneity. Translating from a spatial location to another, we notice that similar experiments give similar results, hence the concept of invariance of the structure of space with respect to the group of such motions. By taking as mathematical model of space a homogeneous space relative to the action of this group (in which we recognize V_3, by observing how translations compose), we therefore acknowledge an essential *physical* property of the space we live in.

REMARK 1.1. In fact, translational invariance is only approximately verified, so one should perhaps approach this basic modelling issue more cautiously: Imagine space as

[6] For a piecewise smooth manifold (see below), such a subspace may fail to exist at some points, which will not be a problem.

a seamless assembly (via smooth transition functions) of patches of affine space, each point covered by at least one of them, which is enough to capture the idea of *local* translational invariance of physical space. This idea gets realized with the concept of smooth manifold (see below) of dimension 3. What we shall eventually recognize as the metric-free part of the Maxwell's system (Ampère's and Faraday's laws) depends on the manifold structure only. Therefore, postulating an affine structure is a *modelling decision*, one that goes a trifle beyond what would strictly be necessary to account for the homogeneity of space, but will make some technical discussions easier when (about Whitney forms) barycentric coordinates will come to the fore.

There is no notion of distance in affine space, but this doesn't mean no topology: Taking the preimages of neighborhoods of $\mathbb{R}^n$ under any one-to-one affine map gives a system of neighborhoods, hence a topology – the same for all such maps. (So we shall talk loosely of a "ball" or a "half ball" in reference to an affine one-to-one image of $B = \{\xi \in \mathbb{R}^n\colon \sum_i (\xi^i)^2 < 1\}$ or of $B \cap \{\xi\colon \xi^1 \geqslant 0\}$.) Continuity and differentiability thus make sense for a function f of type $A_p \to A_n$. In particular, the derivative of f at x is the linear map $\mathrm{D}f(x)$, from V_p to V_n, such that $|f(x+v) - f(x) - \mathrm{D}f(x)(v)|/|v| = \mathrm{o}(|v|)$, if such a map exists, which does not depend on which norms $||$ on V_p and V_n are used to check the property. The same symbol, $\mathrm{D}f(x)$, will be used for the *tangent map* that sends a vector v anchored at x to the vector $\mathrm{D}f(x)(v)$ anchored at $f(x)$.

2. Piecewise smooth manifolds

We will do without a formal treatment of manifolds. Most often, we shall just use the word as a generic term for lines, surfaces, or regions of space ($p = 1, 2, 3$, respectively), piecewise smooth (as defined in a moment), connected or not, with or without a boundary. A 0-manifold is a collection of isolated points.

For the rare cases when the general concept is evoked, suffice it to say that a p-dimensional manifold is a set M equipped with a set of maps of type $M \to \mathbb{R}^p$, called *charts*, which make M look, for all purposes, but only locally, like $\mathbb{R}^p$ (and hence, like p-dimensional affine space). *Smooth* manifolds are those for which the so-called *transition functions* $\varphi \circ \psi^{-1}$, for any pair $\{\varphi, \psi\}$ of charts, are smooth, i.e., possess derivatives of all orders. (So-called C^k manifolds obtain when continuous derivatives exist up to order k.) Then, if some property P makes sense for functions of type $\mathbb{R}^p \to X$, where X is some target space, f from M to X is reputed to have property P if all composite functions $f \circ \varphi^{-1}$, now of type $\mathbb{R}^p \to X$, have it. A manifold M *with boundary* has points where it "looks, locally, like" a closed half-space of $\mathbb{R}^p$; these points form, taken together, a (boundaryless) $(p-1)$-manifold ∂M, called the *boundary* of M. Connectedness is not required: A manifold can be in several pieces, all of the same dimension p.

In practice, our manifolds will be glued assemblies of *cells*, as follows.

First, let us define "reference cells" in $\mathbb{R}^p$, as illustrated on Fig. 2.1. These are bounded convex polytopes of the form

$$K_p^\alpha = \left\{ \xi \in \mathbb{R}^p\colon \xi^l \geqslant 0 \ \forall l = 1, \ldots, p,\ \sum_{j=1}^{p} \alpha_j^i \xi^j \leqslant 1 \ \forall i = 1, \ldots, k \right\}, \tag{2.1}$$

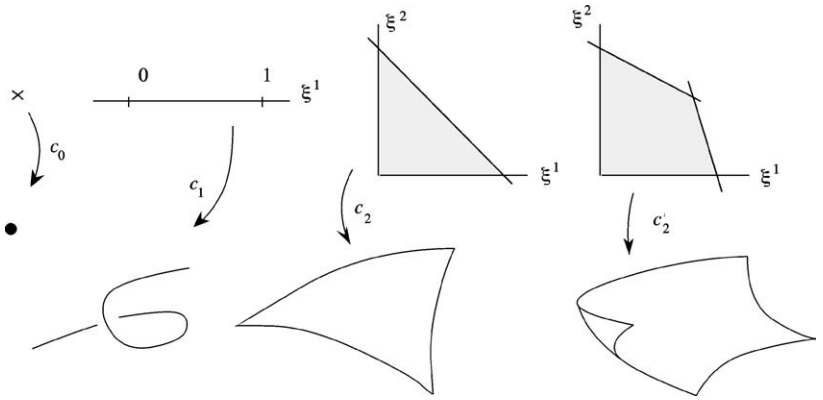

FIG. 2.1. Some cells in A_3, of dimensions 0, 1, 2.

where the α^i_j's form a rectangular $(k \times p)$-matrix with nonnegative entries, and no redundant rows.

Now, a *p-cell* in A_n, with $0 \leqslant p \leqslant n$, is a smooth map c from some K^α_p into A_n, one-to-one, and such that the derivative $\mathrm{D}c(\xi)$ has rank p for all ξ in K^α_p. (These restrictions, which qualify c as an *embedding*, are meant to exclude double points, and cusps, pleats, etc., which smoothness alone is not enough to warrant.) The same symbol c will serve for the map and for the image $c(K^\alpha_p)$. The *boundary* ∂c of the cell is the image under c of the topological boundary of K^α_p, i.e., of points ξ for which at least one equality holds in (2.1). Remark that ∂c is an assembly of $(p-1)$-cells, which themselves intersect, if they do, along parts of their boundaries.

Thus, a 0-cell is just a point. A 1-cell, or "path", is a simple parameterized curve. The simplest 2-cell is the triangular "patch", a smooth embedding of the triangle $\{\xi\colon\ \xi^1 \geqslant 0,\ \xi^2 \geqslant 0,\ \xi^1 + \xi^2 \leqslant 1\}$. The definition is intended to leave room for polygonal patches as well, and for three-dimensional "blobs", i.e., smooth embeddings of convex polyhedra.

We shall have use for the *open* cell corresponding to a cell c (then called a *closed* cell for contrast), defined as the restriction of c to the interior of its reference cell.

A subset M of A_n will be called a *piecewise smooth* p-manifold if (1) there exists a finite family $\mathcal{C} = \{c_i\colon\ i = 1, \ldots, m\}$ of p-cells whose union is M, (2) the open cell corresponding to c_i intersects no other cell, (3) intersections $c_i \cap c_j$ are piecewise smooth $(p-1)$-manifolds (the recursive twist in this clause disentangles at $p = 0$), (4) the cells are properly joined at their boundaries,[7] i.e., in such a way that each point of M has a neighborhood in M homeomorphic to either a p-ball or half a p-ball.

Informally, therefore, piecewise smooth manifolds are glued assemblies of cells, obtained by topological identification of parts of their respective boundaries. (Surface S in Fig. 4.1, below, is typical.)

[7] This is regrettably technical, but it can't be helped, if M is to be a manifold. The assembly of *three* curves with a common endpoint, for instance, is not a manifold. See also HENLE [1994] for examples of 3D-spaces obtained by identification of facets of some polyhedra, which fail to be manifolds. Condition (2) forbids self-intersections, which is overly drastic and could be avoided, but will not be too restrictive in practice.

Having introduced this category of objects – which we shall just call manifolds, from now on – we should, as it is the rule and almost a reflex in mathematical work, deal with maps between such objects, called *morphisms*, that preserve their relevant structures. About cells, first: A map between two images of the same reference cell which is bijective and smooth (in both directions) is called a *diffeomorphism*. Now, about our manifolds: There is a *piecewise smooth diffeomorphism* between two of them (and there too, we shall usually dispense with the "piecewise smooth" qualifier) if they are homeomorphic and can both be chopped into sets of cells which are, two by two, diffeomorphic.

3. Orientation

To get oneself oriented, in the vernacular, consists in knowing where is South, which way is uptown, etc. To orient a map, one makes its upper side face North. Pigeons, and some persons, have a sense of orientation. And so forth. *Nothing* of this kind is implied by the mathematical concept of orientation – which may explain why so simple a notion may be so puzzling to many. Not that mathematical orientation has no counterpart in everyday's life, it has, but in something else: When entering a roundabout or a circle with a car, you know whether you should turn clockwise or counterclockwise. *That* is orientation, as regards the ground's surface. Notice how it depends on customs and law. For the spatial version of it, observe what "right-handed" means, as applied to a staircase or a corkscrew.

3.1. Oriented spaces

Now let us give the formal definition. A *frame* in V_n is an ordered n-tuple of linearly independent vectors. Select a basis (which is thus a frame among others), and for each frame, look at the determinant of its n vectors, as expressed in this basis, hence a *FRAME* $\rightarrow$ *REAL* function. This function is basis-dependent, but the equivalence relation defined by "$f \equiv f'$ if and only if frames f and f' have determinants of the same sign" does not depend on the chosen basis, and is thus intrinsic to the structure of V_n. There are two equivalence classes with respect to this relation. Orienting V_n consists in designating one of them as the class of "positively oriented" frames. This amounts to defining a function, which assigns to each frame a label, either *direct* or *skew*, two equivalent frames getting the same label. There are two such functions, therefore two possible orientations. An *oriented vector space* is thus a pair $\{V, Or\}$, where Or is one of the two orientation classes of V. (Equivalently, one may define an oriented vector space as a pair {*vector space*, *privileged basis*}, provided it's well understood that this basis plays no other role than specifying the orientation.) We shall find convenient to extend the notion to a vector space of dimension 0 (i.e., one reduced to the single element 0), to which also correspond, by convention, two oriented vector spaces, labelled $+$ and $-$.

REMARK 3.1. Once a vector space has been oriented, there are direct and skew *frames*, but there is no such thing as direct or skew *vectors*, except, one may concede, in dimension 1. A vector does not acquire new features just because the space where it belongs has been oriented! Part of the confusion around the notion of "axial" (vs. "polar") vectors stems from this semantic difficulty (BOSSAVIT [1998a, p. 296]). As axial vectors

will not be used here, the following description should be enough to deal with the issue. Let's agree that, if *Or* is one of the orientation classes of V, the expression $-Or$ denotes the other class. Now, form pairs $\{v, Or\}$, where v is a vector and *Or* any orientation class of V, and consider two pairs $\{v, Or\}$ and $\{v', Or'\}$ as equivalent when $v' = -v$ and $Or' = -Or$. *Axial vectors* are, by definition, the equivalence classes of such pairs. (*Polar* vectors is just a redundant name, inspired by a well-minded sense of equity, for vectors of V.) Notice that axial *scalars* can be defined the same way: substitute a real number for v. Hence axial vector fields and axial functions (more often called "pseudo-functions" in physics texts). The point of defining such objects is to become able to express Maxwell's equations in *non*-oriented Euclidean space, i.e., V_3 with a dot product but no specific orientation. See BOSSAVIT [1998b] or [1999] for references and a discussion.

An affine space, now, is oriented by orienting its vector associate: a *bound frame* at x in A_n, i.e., a set of n independent vectors at x, is direct (respectively skew) if these n vectors form a direct (respectively skew) frame in V_n.

Vector subspaces of a given vector space (or affine subspaces of an affine space[8]) can have their own orientation. Orienting a line, in particular, means selecting a vector parallel to it, called a *director* vector for the line, which specifies the "forward" direction along it.

Such orientations of different subspaces are a priori unrelated. Orienting 3D space by the corkscrew rule, for instance, does not imply any orientation in a given plane. This remark may hurt common sense, for we are used to think of the standard orientation of space and of, say, a horizontal plane, as somehow related. And they are, indeed, but only because we think of vertical lines as oriented, bottom up. This is the convention known as *Ampère's rule*. To explain what happens there, suppose space is oriented, and some privileged straightline is oriented too, on its own. Then, any plane *transverse* to this line (i.e., thus placed that the intersection reduces to a single point) inherits an orientation, as follows: To know whether a frame in the plane is direct or skew, make a list of vectors composed of, in this order, (1) the line's director, (2) the vectors of the planar frame; hence an enlarged spatial frame, which is either direct or skew, which tells us about the status of the plane frame.

More generally, there is an interplay between the orientations of complementary subspaces and those of the encompassing space. Recall that two subspaces U and W of V are *complementary* if their *span* is all V (i.e., each v in V can be decomposed as $v = u + w$, with u in U and w in W) and if they are *transverse* ($U \cap W = \{0\}$, which makes the decomposition unique). We shall refer to V as the "ambient" space, and write $V = U + W$. If both U and W have orientation, this orients V, by the following convention: the frame obtained by listing the vectors of a direct frame in U first, then those of a direct frame in W, is direct. Conversely, if both U and V are oriented, one may orient W as follows: to know whether a given frame in W is direct or skew, list its vectors behind those of a direct frame of U, and check whether the enlarged frame thus obtained is direct or skew in V. This is a natural generalization of Ampère's rule.

[8] An affine subspace is oriented by orienting the parallel vector subspace. A point, which is an affine subspace parallel to $\{0\}$, can therefore be oriented, which we shall mark by apposing a sign to it, $+$ or $-$.

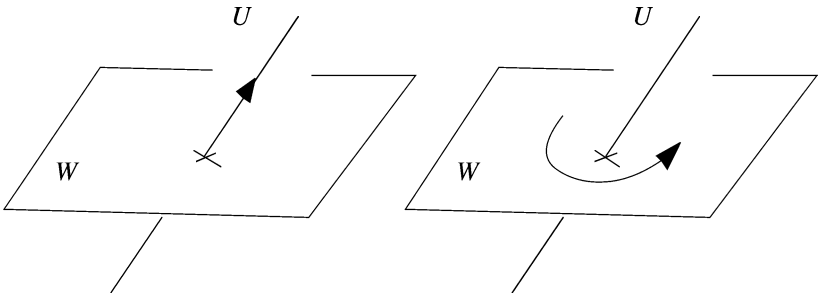

FIG. 3.1. Left: Specifying a "crossing direction" through a plane W by inner-orienting a line U transverse to it. Right: Outer-orienting U, i.e., giving a sense of going around it, by inner-orienting W.

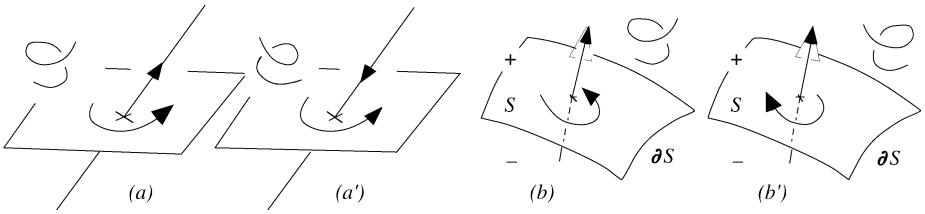

FIG. 3.2. Left: How an externally oriented line acquires inner orientation, depending on the orientation of ambient space. (Alternative interpretation: if one knows both orientations, inner and outer, for a line, one knows the ambient orientation.) Right: Assigning to a surface a crossing direction (here from region "$-$" below to region "$+$" above) will not by itself imply an inner orientation. But it does if ambient space is oriented, as seen in (b) and (b′). Figs. 3.2(a) and 3.2(b) can be understood as an explanation of Ampère's rule, in which the ambient orientation is, by convention, the one shown here by the "right corkscrew" icon.

Now what if U is oriented, but ambient space is not? Is U's orientation of any relevance to the complement W? Yes, as Fig. 3.1 suggests (left): For instance, if W has dimension $n-1$, an orientation of the one-dimensional complement U can be interpreted as a crossing direction relative to W, an obviously useful notion. (Flow of something through a surface, for instance, presupposes a crossing direction.) Hence the concept of *external*, or *outer orientation* of subspaces of V: Outer orientation of a subspace is, by definition, an orientation of one[9] of its complements. Outer orientation of V itself is thus a sign, $+$ or $-$. (For contrast and clarity, we shall call *inner* orientation what was simply "orientation" up to this point.) The notion (which one can trace back to Veblen (VEBLEN and WHITEHEAD [1932]), cf. VAN DANTZIG [1954] and SCHOUTEN [1989]) passes to affine subspaces of an affine space the obvious way.

Note that *if* ambient space is oriented, outer orientation determines inner orientation (Fig. 3.2). But otherwise, the two kinds of orientation are independent. As we shall see, they cater for different needs in modelling.

[9]Nothing ambiguous in that. There is a canonical linear map between two complements W_1 and W_2 of the same subspace U, namely, the "affine projection" π_U along U, thus defined: for v in W_1, set $\pi_U(v) = v + u$, where u is the unique vector in U such that $v + u \in W_2$. Use π_U to transfer orientation from W_1 to W_2.

3.2. Oriented manifolds

Orientation can be defined for other figures than linear subspaces. Connected parts of affine subspaces, such as polygonal facets, or line segments, can be oriented by orienting the supporting subspace (i.e., the smallest one containing them). Smooth lines and surfaces as a whole are oriented by attributing orientations to all their tangents or tangent planes in a consistent way.

"Consistent"? Let's explain what that means, in the case of a surface. First, subspaces parallel to the tangent planes at all points in the neighborhood $N(x)$ of a given surface point x have, if $N(x)$ is taken small enough, a common complement, characterized by a director $n(x)$ (not the "normal" vector, since we have no notion of orthogonality at this stage, but the idea is the same). Then $N(x)$ is consistently oriented if all these orientations correspond via the affine projection along $n(x)$ (cf. Note 9). But this is only *local* consistency, which can always be achieved, and one wants more: *global* consistency, which holds if the surface can be covered by such neighborhoods, with consistent orientation in each non-empty intersection $N(x) \cap N(y)$. This may not be feasible, as in the case of a Möbius band, hence the distinction between (internally) orientable and non-orientable manifolds.

Cells, as defined above, are inner orientable, thanks to the fact that Dc does not vanish. For instance (cf. Fig. 3.3), for a path c, i.e., a smooth embedding $t \to c(t)$ from $[0, 1]$ to A_n, the tangent vectors $\partial_t c(t)$ determine consistent orientations of their supporting lines, hence an orientation of the path. (The other orientation would be obtained by starting from the "reverse" path, $t \to c(1-t)$.) Same with a patch $\{s, t\} \to S(s, t)$ on the triangle $T = \{\{s, t\}\colon\ 0 \leqslant s,\ 0 \leqslant t,\ s + t \leqslant 1\}$: The vectors $\partial_s S(s, t)$ and $\partial_t S(s, t)$, in this order, form a basis at $S(s, t)$ which orients the tangent plane, and these orientations are consistent.

As for piecewise smooth manifolds, finally, the problem is at points x where cells join, for a tangent subspace may not exist there. But according to our conventions, there must be a neighborhood homeomorphic to a ball or half-ball, which *is* orientable, hence a way to check whether tangent subspaces at regular points in the vicinity of x have consistent orientations, and therefore, to check whether the manifold as a whole is or is not orientable.

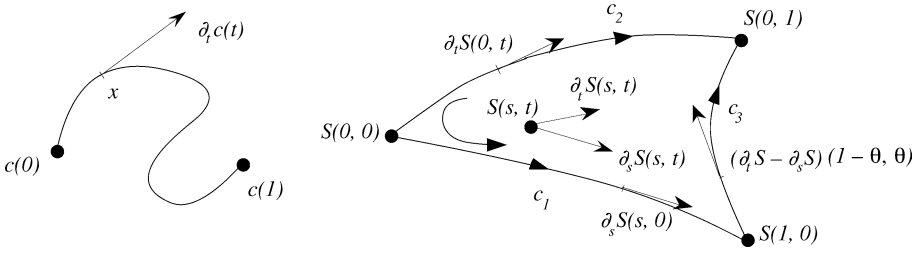

FIG. 3.3. A path and a patch, with natural inner orientations. Observe how their boundaries are themselves assemblies of cells: $\partial c = c(0) - c(1)$ and $\partial S = c_1 - c_2 + c_3$, with a notation soon to be introduced more formally. Paths c_i are $c_1 = s \to S(s, 0)$, $c_2 = t \to S(0, t)$, and $c_3 = \theta \to S(1 - \theta, \theta)$, each with its natural inner orientation.

Similar considerations hold for external orientation. Outer-orienting a surface consists in giving a (globally consistent) crossing direction through it. For a line, it's a way of "turning around" it, or "gyratory sense" (Fig. 3.1, right). For a point, it's an orientation of the space in its neighborhood. For a connected region of space, it's just a sign, $+$ or $-$.

3.3. Induced orientation

Surfaces which enclose a volume V (which one may suppose connected, though the boundary ∂V itself need not be) can always be outer oriented, because the "inside out" crossing direction is always globally consistent. Let us, by convention, take this direction as defining the canonical outer orientation of ∂V. No similarly canonical *inner* orientation of the surface results, as could already be seen on Fig. 3.2, since there are, in the neighborhood of each boundary point, two eligible orientations of ambient space. But if V is inner oriented, this orientation can act in conjunction with the outer one of ∂V to yield a natural inner orientation of V's boundary about this point. For example, on the left of Fig. 3.4, the 2-frame $\{v_1, v_2\}$ in the tangent plane of a boundary point is taken as direct because, by listing its vectors behind an outward directed vector ν, one gets the direct 3-frame $\{\nu, v_1, v_2\}$. Consistency of these orientations stems from the consistency of the crossing direction. Hence V's inner orientation *induces* one on each part of its boundary.

The same method applies to manifolds of lower dimension p, by working inside the affine p-subspace tangent to each boundary point. See Fig. 3.4(b) for the case $p = 2$. The p-manifold, thus, serves as ambient space with respect to its own boundary, for the purpose of inducing orientation.

In quite a similar way (Fig. 3.5), *outer* orientation of a manifold induces an *outer* orientation of each part of its boundary. (For a volume V, the induced outer orientation of ∂V is the inside-out or outside-in direction, depending on the outer orientation, $+$ or $-$, of V.)

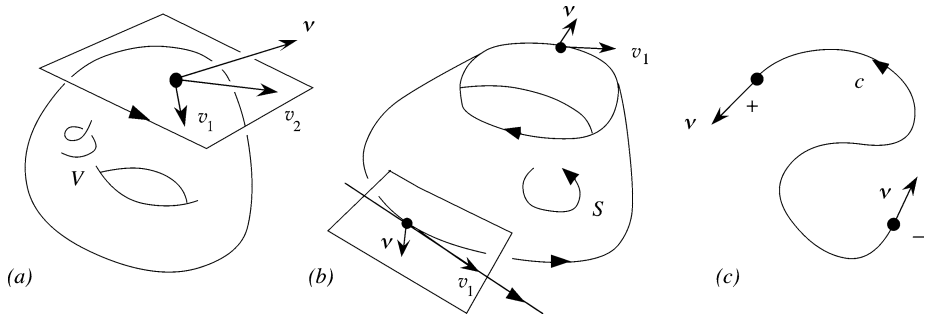

FIG. 3.4. Left: Induced orientation of the boundary of a volume of toroidal shape (v_1 and v_2 are tangent to ∂V, ν points outwards). Middle: The same idea, one dimension below. The tangent to the boundary, being a complement of (the affine subspace that supports) ν, with respect to the plane tangent to the surface S (in broken lines), inherits from the latter an inner orientation. Right: Induced orientation of the endpoints of an oriented curve.

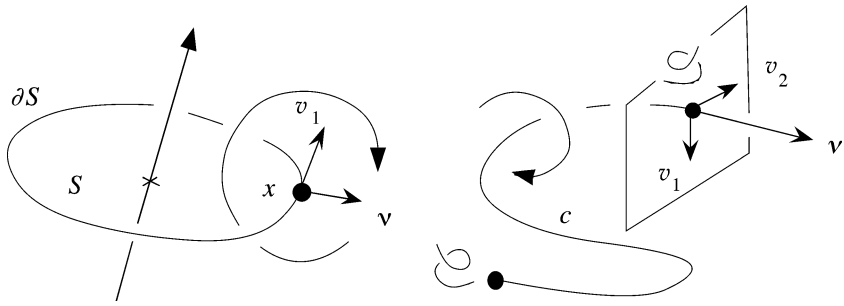

FIG. 3.5. Left: To outer-orient ∂S is to (consistently) inner-orient complements of the tangent, one at each boundary point x. For this, take as direct the frame $\{v_1, \nu\}$, where $\{v_1\}$ is a direct frame in the complement of the plane tangent to S at x, and ν an outward directed vector tangent to S. That $\{v_1\}$ is direct is known from the outer orientation of S. Right: Same idea about the boundary points of line c. Notice that ν is now appended *behind* the list of frame vectors. Consistency stems from the consistency of ν, the inside-out direction with respect to S. The icons near the endpoints are appropriate, since outer orientation of a point is inner orientation of the space in its vicinity.

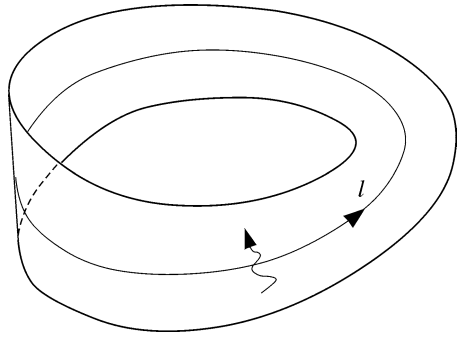

FIG. 3.6. Möbius band, not orientable. As the middle line l does not separate two regions, it cannot be assigned any consistent crossing direction, so it has no outer orientation with respect to the "ambient" band.

3.4. Inner vs outer orientation of submanifolds

We might (but won't, as the present baggage is enough) extend these concepts to submanifolds of ambient manifolds other that A_3, including non-orientable ones. A two-dimensional example will give the idea (Fig. 3.6): Take as ambient manifold a Möbius band M, and forget about the 3-dimensional space it is embedded in for the sake of the drawing. Then it's easy to find in M a line which (being a line) is inner orientable, but cannot consistently be outer oriented. Note that the band by itself, i.e., considered as its own ambient space, can be outer oriented, by giving it a sign: Indeed, outer orientation of the tangent plane at each point of M, being inner orientation of this point, is such a sign, so consistent orientation means attributing the same sign to all points. (By the same token, any manifold is outer orientable, with respect to itself as ambient space.)

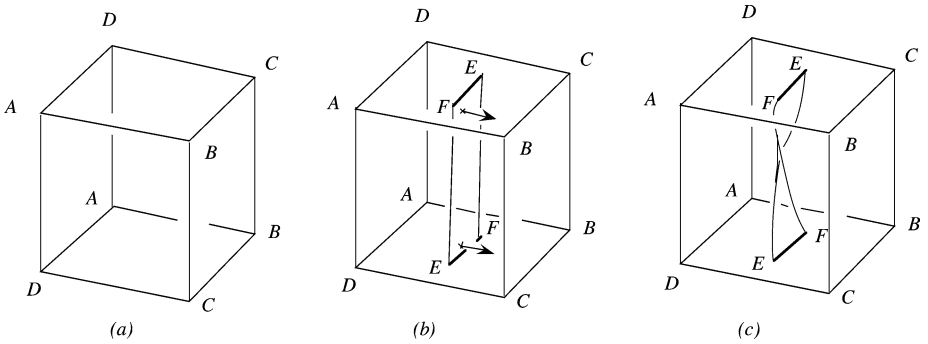

FIG. 3.7. Left: Non-orientable 3-manifold with boundary: Identify top and bottom by matching upper A with lower A, etc. Middle: Embedded Möbius band, with a globally consistent crossing direction. Right: Embedded ribbon.

For completeness, let us give another example (Fig. 3.7), this time of an outer-orientable surface without inner orientation, owing to non-orientability of the ambient manifold. The latter (whose boundary is a Klein bottle) is made by sticking together the top and bottom sides of a vertical cube, according to the rule of Fig. 3.7(a). The ribbon shown in (b) is topologically a Möbius band, a non-(inner)orientable surface. Yet, it plainly has a consistent set of transverse vectors. (Follow the upper arrow as its anchor point goes up and reenters at the bottom, and notice that the arrow keeps pointing in the direction of AB in the process. So it coincides with the lower arrow when this passage has been done.) Contrast with the ordinary ribbon in (c), orientable, but not outer orientable with respect to this ambient space.

The two concepts of orientation are therefore essentially different.

In what follows, we shall use the word "twisted" (as opposed to "straight") to connote anything that is to do with outer (as opposed to inner) orientation.

4. Chains, boundary operator

It may be convenient at times to describe a manifold M as an assembly of several manifolds, even if M is connected. Think for example of the boundary of a triangle, as an assembly of three edges, and more generally of a piecewise smooth assembly of cells. But it may happen – so will be the case here, later – that these various manifolds have been *independently* oriented, with orientations which may or may not coincide with the desired one for M. This routinely occurs with boundaries, in particular. The concept of chain will be useful to deal with such situations.

A *p-chain* is a finite family $\mathcal{M} = \{M_i\colon\ i = 1, \ldots, k\}$ of oriented connected p-manifolds,[10] to which we shall loosely refer below as the "components" of the chain, each loaded with a weight μ^i belonging to some ring of coefficients, such as $\mathbb{R}$ or $\mathbb{Z}$ (say $\mathbb{R}$ for definiteness, although weights will be signed integers in most of our examples). Such a chain is conveniently denoted by the "formal" sum $\sum_i \mu^i M_i \equiv \mu^1 M_1 + \cdots + \mu^k M_k$,

[10] For instance, cells. But we don't request that. Each M_i may be a piecewise smooth manifold already.

thus called because the + signs do not mean "add" in any standard way. On the other hand, chains themselves, as whole objects, can be added, and there the notation helps: To get the sum $\sum_i \mu^i M_i + \sum_j \nu^j N_j$, first merge the two families $\mathcal{M}$ and $\mathcal{N}$, then attribute weights by adding the weights each component has in each chain, making use of the convention that $\mu M'$ is the same chain as $-\mu M$ when M' is the same manifold as M with opposite orientation. If all weights are zero, we have the *null chain*, denoted 0. All this amounts, as one sees, to handling chains according to the rules of algebra, when they are represented via formal sums, which is the point of such a notation. *Twisted* chains are defined the same way, except that all orientations are external. (Twisted and straight chains are not to be added, or otherwise mixed.)

If M is an oriented piecewise smooth manifold, all its cells c_i inherit this orientation, but one may have had reasons to orient them on their own, independently of M. (The same cell may well be part of several piecewise smooth manifolds, for instance.) Then, it is natural to associate with M the chain $\sum_i \pm c_i$, also denoted by M, with ith weight -1 when the orientations of M and c_i differ. (Refer back to Fig. 3.3 for simple examples.)

Now, the boundary of an oriented piecewise smooth $(p+1)$-manifold M is an assembly of p-manifolds, each of which we assume has an orientation of its own. Let us assign each of them the weight ± 1, according to whether its orientation coincides with the one inherited from M. (We say the two orientations *match* when this coincidence occurs.) Hence a chain, also denoted ∂M. By linearity, the operator ∂ extends to chains: $\partial(\sum_i \mu^i M_i) = \sum_i \mu^i \partial M_i$. A chain with null boundary is called a *cycle*. A chain which is the boundary of another chain is called, appropriately, a *boundary*. Boundaries are cycles, because of the fundamental property

$$\partial \circ \partial = 0, \tag{4.1}$$

i.e., the boundary of a boundary is the null chain. A concrete example, as in Fig. 4.1, will be more instructive here than a formal proof.

REMARK 4.1. Beyond its connection with assemblies of oriented cells, no too definite intuitive interpretation of the concept of chain should be looked for. Perhaps, when $p = 1$, one can think of the chain $\sum_i \gamma_i c_i$, with integer weights, as "running along each c_i, in turn, $|\gamma_i|$ times, in the direction indicated by c_i's orientation, or in the reverse direction, depending on the sign of γ_i". But this is a bit contrived. Chains are better conceived as algebraic objects, based on geometric ones in a useful way – as the example in Fig. 4.1 should suggest, and as we shall see later. However, we shall indulge in language abuse, and say that a closed curve "is" a 1-cycle, or that a closed surface "is" a 2-cycle, with implicit reference to the associated chain.

So boundaries are cycles, after (4.1). Whether the converse is true is an essential question. In affine space, the answer is positive: A closed surface encloses a volume, a closed curve (even if knotted) is the boundary of some surface (free of self-intersections, amazing as this may appear), called a Seifert surface (SEIFERT and THRELFALL [1980], ARMSTRONG [1979, p. 224]). But in some less simple ambient manifolds, a cycle need not bound. In the case of a solid torus, for instance, a meridian circle is a boundary, but a parallel circle is not, because none of the disks it bounds in A_3 is entirely contained in

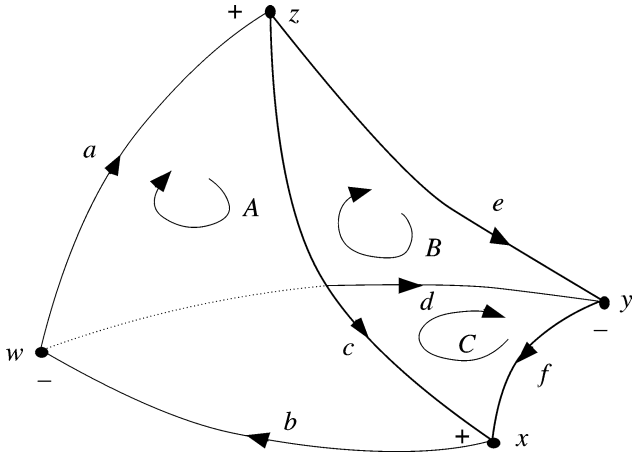

FIG. 4.1. Piecewise smooth surface S, inner oriented (its orientation is taken to be that of the curved triangle in the fore, marked A), represented as the chain $A - B - C$ based on the oriented curved triangles A, B, C. (Note the minus signs: B's and C's orientations don't match that of S.) One has $\partial A = a + b + c$, $\partial B = e + a - d$, $\partial C = b + d + f$, where a, b, c, d, e, f are the boundary curves, arbitrarily oriented as indicated. Now, $\partial S = \partial(A - B - C) = c - e - f$: Observe how the "seams" a, b, c automatically receive null weights in this 1-chain, whatever their orientation, because they appear twice with opposite signs. Next, since $\partial c = x - z$, $\partial e = -y - z$, and $\partial f = x + y$, owing to the (arbitrary) orientations assigned to points w, x, y, z, one has $\partial\partial S = \partial(c - e - f) = 0$, by the same process of cancellation by pairs. The reader is invited to work out a similar example involving twisted chains instead of straight ones.

the torus. Whether cycles are or aren't boundaries is therefore an issue when investigating the global topological properties of a manifold. Chains being algebraic objects then becomes an asset, for it makes possible to harness the power of algebra to the study of topology. This is the gist of *homology* (HENLE [1994], HILTON and WYLIE [1965]), and of algebraic topology in general.

5. Metric notions

Now, let us equip V_n with a dot product: $u \cdot v$ is a real number, linearly depending on vectors u and v, with symmetry ($u \cdot v = v \cdot u$) and strict positive-definiteness ($u \cdot u > 0$ if $u \neq 0$). Come from this, first the notions of orthogonality and angle, next a norm $|u| = (u \cdot u)^{1/2}$ on V_n, then a distance $d(x, y) = |y - x|$, translation-invariant by construction, between points of the affine associate A_n.

DEFINITION 5.1. Euclidean space, E_n, is the structure composed of A_n, plus a dot product on its associate V_n, plus an orientation.

Saying "the" structure implies that two realizations of it (with two different dot products and/or orientations) are isomorphic in some substantial way. This is so: For any other dot product, "$\cdot$" say, there is an invertible linear transform L such that

$u \cdot v = Lu \cdot Lv$. Moreover,[11] one may have L "direct", in the sense that it maps a frame to another frame of the same orientation class, or "skew". Therefore, two distinct Euclidean structures on A_n are linked by some L. In the language of group actions, the linear group GL_n, composed of the above L's, acts transitively on Euclidean structures, i.e., with a unique orbit, which is our justification for using the singular. (These structures are said to be *affine equivalent*,[12] a concept that will recur.) The point can vividly be made by using the language of group actions: the isotropy group of $\{\cdot, Or\}$ "cannot be any larger". (More precisely, it is maximal, as a subgroup, in the group of direct linear transforms.)

In dimension 3,[13] dot product and orientation conspire in spawning the *cross product*: $u \times v$ is characterized by the equality

$$|u \times v|^2 + (u \cdot v)^2 = |u|^2 |v|^2 \tag{5.1}$$

and the fact that vectors u, v and $u \times v$ form, in this order, a direct frame. The 3-*volume* of the parallelotope built on vectors u, v, w, defined by $\mathrm{vol}(u, v, w) = (u \times v) \cdot w$, is equal, up to sign, to the above volumic measure, with equality if the frame is direct.[14] Be well aware that $\times$ doesn't make any sense in *non*-oriented three-space.

We shall have use for the related notion of *vectorial area* of an outer oriented triangle T, defined as the vector $\vec{\mathrm{T}} = \mathrm{area}(\mathrm{T})n$, where n is the normal unit vector that provides the crossing direction. (If an ambient orientation exists, two vectors u and v can be laid along two of the three sides, in such a way that $\{u, v, n\}$ is a direct frame. Then, $\vec{\mathrm{T}} = \frac{1}{2} u \times v$. Fig. 6.1 gives an example.) More generally, an outer oriented surface of E_3 has a vectorial area: Chop the surface into small adjacent triangular patches, add the vectorial areas of these, and pass to the limit. (This yields 0 for a closed surface.)

For later use, we state the relations between the structures induced by $\{\cdot, Or\}$ and $\{\boldsymbol{\cdot}, \mathbf{Or}\}$, where $\mathbf{Or} = \pm Or$, the sign being that of $\det(L)$. (There is no ambiguity about "$\det(L)$", understood as the determinant of the matrix representation of L: its value is the same in any basis.) The norm $(u \cdot u)^{1/2}$ will be denoted by $|u|$. The corresponding cross product $\boldsymbol{\times}$ (boldface) is defined by $|u \boldsymbol{\times} v|^2 + (u \cdot v)^2 = |u|^2 |v|^2$ as in (5.1) (plus the request that $\{u, v, u \boldsymbol{\times} v\}$ be **Or**-direct), and the new volume is $\mathbf{vol}(u, v, w) = (u \boldsymbol{\times} v) \boldsymbol{\cdot} w$. It's a simple exercise to show that

$$|u| = |Lu|, \quad L(u \boldsymbol{\times} v) = Lu \times Lv, \quad \mathbf{vol}(u, v, w) = \det(L)\, \mathrm{vol}(u, v, w). \tag{5.2}$$

(It all comes from the equality $\det(Lu, Lv, Lw) = \det(L) \det(u, v, w)$, when u, v, w, and L are represented in some basis, a purely affine formula.) Notice that, for any w,

[11] L is not unique, since UL, for any *unitary* U (i.e., such that $|Uv| = |v|\ \forall v$), will work as well. In particular, one might force L to be self-adjoint, but we won't take advantage of that.

[12] Such equivalence is what sets Euclidean norms apart among all conceivable norms on V_n, like for instance $|v| = \sum_i |v^i|$. As argued at more length in BOSSAVIT [1998a], choosing to work in a Euclidean framework is an acknowledgment of another observed symmetry of the world we live in: its *isotropy*, in addition to its homogeneity.

[13] A binary operation with the properties of the cross product can exist only in dimensions 3 and 7 (SHAW and YEADON [1989], ECKMANN [1999]).

[14] An n-volume could directly be defined on V_n, as a map $\{v_1, \ldots, v_n\} \to \mathrm{vol}(v_1, \ldots, v_n)$, multilinear and null when two vectors of the list are equal. Giving an n-volume implies an orientation (direct frames are those with positive n-volumes), but no metric (unless $n = 1$).

one has $L^a L(u \times v) \cdot w = L(u \times v) \cdot Lw = \det(L)(u \times v) \cdot w$, where L^a denotes the *adjoint* of L (defined by $Lu \cdot v = u \cdot L^a v$ for all u, v), hence an alternative formula:

$$u \times v = \det(L)(L^a L)^{-1}(u \times v). \tag{5.3}$$

As for the vectorial area, denoted $\vec{\mathbf{T}}$ in the "bold" metric, one will see that

$$\vec{\mathbf{T}} = \big|\det(L)\big|(L^a L)^{-1}\vec{\mathrm{T}}, \tag{5.4}$$

with a factor $|\det(L)|$, not $\det(L)$, because $\vec{\mathrm{T}}$ and $\vec{\mathbf{T}}$, both going along the crossing direction, point towards the same side of T.

We shall also need a topology on the space of p-chains, in order to define differential forms as *continuous* linear functionals on this space. As we shall argue later, physical observables such as electromotive force, flux, and so forth, can be conceived as the values of functionals of this kind, the chain operand being the idealization of some measuring device. Such values don't change suddenly when the measurement apparatus is slightly displaced, which is the rationale for continuity. But to make precise what "slightly displaced" means, we need a notion of "nearness" between chains – a topology.[15]

First thing, nearness between manifolds. Let us define the distance $d(M, N)$ between two of them as the greatest lower bound (the infimum) of $d_\phi(M, N) = \sup\{x \in M\colon |x - \phi(x)|\}$ with respect to all orientation-preserving piecewise smooth diffeomorphisms (OPD) ϕ that exist between M and N. There may be no such OPD, in which case we take the distance as infinite, but otherwise there is symmetry between M and N (consider ϕ^{-1} from N to M), positivity, d can't be zero if $M \neq N$, and the triangle inequality holds. (*Proof*: Take M, N, P, select OPDs ϕ and ψ from P to M and N, and consider x in P. Then $|\phi(x) - \psi(x)| \leqslant |\phi(x) - x| + |x - \psi(x)|$, hence $d_{\psi\circ\phi^{-1}}(M, N) \leqslant d_\phi(M, P) + d_\psi(N, P)$, then minimize with respect to ϕ and ψ.) Nearness of two manifolds, in this sense, does account for the intuitive notion of "slight displacement" of a line, a surface, etc. The topology thus obtained does not depend on the original dot product, although d does.

Next, on to chains. The notion of convergence we want to capture is clear enough: a sequence of chains $\{c_n = \sum_{i=1,\ldots,k} \mu_n^i M_{i,n}\colon n \in \mathbb{N}\}$ should certainly converge towards the chain $c = \sum_{i=1,\ldots,k} \mu^i M_i$ when the sequences of components $\{M_{i,n}\colon n \in \mathbb{N}\}$ all converge, in the sense of the previous distance, to M_i, while the weights $\{\mu_n^i\colon n \in \mathbb{N}\}$ converge too, towards μ^i. But knowing some convergent sequences is not enough to know the topology. (For that matter, even the knowledge of *all* convergent sequences would not suffice, see GELBAUM and OLMSTED [1964, p. 161].) On the other hand, the finer the topology, i.e., the more open sets it has, the more difficult it is for a sequence to converge, which tells us what to do: Define the desired topology as the finest one which (1) is compatible with the vector space structure of p-chains (in particular, each neighborhood of 0 should contain a convex neighborhood) (2) makes all sequences of the above kind converge.

[15] What follows is an attempt to bypass, rather than to face, this difficult problem, to which Harrison's work on "chainlet" spaces (nested Banach spaces which include chains and their limits with respect to various norms, HARRISON [1998]), provides a much more satisfactory solution.

The space of straight [respectively twisted] p-chains, as equipped with this topology, will be denoted by $\mathcal{C}_p$ [respectively $\tilde{\mathcal{C}}_p$]. Both spaces are purely affine constructs, independent of the Euclidean structure, which only played a transient role in their definition.

It now makes sense to ask whether the linear map ∂ is continuous from $\mathcal{C}_p$ to $\mathcal{C}_{p-1}$. The answer is by the affirmative, thanks to the linearity of ∂ and the inequality $d(\partial M, \partial N) \leqslant d(M, N)$. [*Proof*: The restriction to ∂M of an OPD ϕ is an OPD which sends it to ∂N, so $d(\partial M, \partial N) \leqslant \inf_\phi \sup\{x \in \partial M\colon\ |\phi(x) - x|\} \leqslant \inf_\phi \sup\{x \in M\colon |\phi(x) - x|\} = d(M, N)$.]

CHAPTER II

Rewriting the Maxwell Equations

Deconstruction calls for reconstruction: We now resettle the Maxwell system in the environment just described, paying attention to what makes use of the metric structure and what does not. In the process, differential forms will displace vector fields as basic entities.

6. Integration: Circulation, flux, etc.

Simply said, differential forms are, among mathematical objects, those meant to be integrated. So let us revisit Integration.

In standard integration theory (HALMOS [1950], RUDIN [1973], YOSIDA [1980]), one has a set X equipped with a measure dx. Then, to a pair $\{A, f\}$, where A is a part of X and f a function, integration associates a number, denoted $\int_A f(x)\,dx$ (or simply $\int_A f$, if there is no doubt on the underlying measure), with additivity and continuity with respect to both arguments, A and f. In what follows, we operate a slight change of viewpoint: Instead of leaving the measure dx in background of a stage on which the two objects of interest would be A and f, we consider the whole integrand $f(x)\,dx$ as a single object (later to be given its proper name, "differential form"), and A as some piecewise smooth manifold of A_3. This liberates integration from its dependence on the metric structure: The integral becomes a map of type *MANIFOLD* $\times$ *DIFFERENTIAL_FORM* $\rightarrow$ *REAL* (by linearity, *CHAIN* will eventually replace *MANIFOLD* there), which we shall see is the right approach as far as Electromagnetics is concerned. The transition will be in two steps, one in which the Euclidean structure is used, one in which we get rid of it.

The dot product of E_n induces measures on its submanifolds: By definition, the Euclidean measure of the parallelotope built on p vectors $\{v_1, \ldots, v_p\}$ anchored at x, i.e., of the set $\{x + \sum_i \lambda^i v_i\colon\ 0 \leqslant \lambda^i \leqslant 1,\ i = 1, \ldots, p\}$, is the square-root of the so-called Gram determinant of the v_i's, whose entries are the dot products $v_i \cdot v_j$, for all i, j from 1 to p. One can build from this, by the methods of classical measure theory (HALMOS [1950]), the p-dimensional measures, i.e., the lineal, areal, volumic, etc., measures of a (smooth, bounded) curve, surface, volume, etc. (what Whitney and his followers call its "mass", WHITNEY [1957]). For $p = 0$ not to stand out as an exception there, we attribute to an isolated point the measure 1. (This is the so-called *counting measure*, for which the measure of a set of points is the number of its elements.)

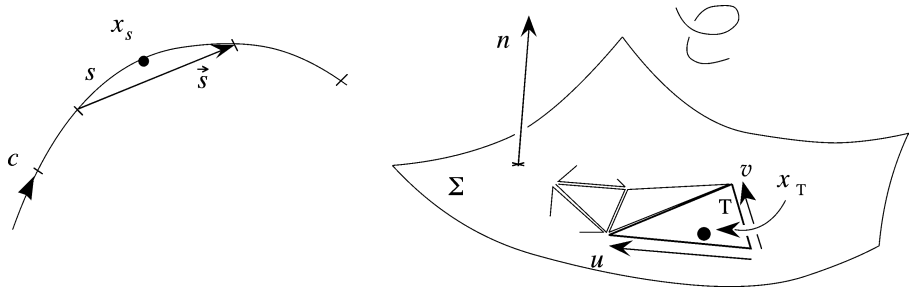

FIG. 6.1. Forming the terms of Riemann sums. Left: generic "curve segment" s, with associated sampling point x_s and vector $\vec{s}$. Right: generic triangular small patch T, with sampling point x_T. Observe how, with the ambient orientation indicated by the icon, the vectorial area of T happens to be $\frac{1}{2}u \times v$.

We shall consider, corresponding to the four dimensions $p = 0, \ldots, 3$ of manifolds in E_3, four kinds of integrals which are constantly encountered in Physics. Such integrals will be defined on cells first, then extended by linearity to chains, which covers the case of piecewise smooth manifolds.

First, $p = 0$, a point, x say. The integral of a smooth function φ is then[16] $\varphi(x)$. If the point is inner oriented, i.e., if it bears a sign $\varepsilon(x) = \pm 1$, the integral is by convention $\varepsilon(x)\varphi(x)$.

Next ($p = 1$), let c be a 1-cell. At point $x = c(t)$, define the *unit tangent vector* $\tau(x)$ as the vector at x equal to $\partial_t c(t)/|\partial_t c(t)|$, which inner-orients c. Given a smooth vector field u, the dot product $\tau \cdot u$ defines a real-valued function on the image of c. We call *circulation* of u, along c thus oriented, the integral $\int_c \tau \cdot u$ of this function with respect to the Euclidean measure of lengths.

REMARK 6.1. Integrals (of smooth enough functions) are limits of Riemann sums. In the present case, such a sum can be obtained as suggested by Fig. 6.1, left: Chop the curve into a finite family $\mathcal{S}$ of adjacent curve segments s, pick a point x_s in each of them, and let $\vec{s}$ be the vector, oriented along c, that joins the extremities of s. The Riemann sum associated with $\mathcal{S}$ is then $\sum_{s\in\mathcal{S}} \vec{s} \cdot u(x_s)$, and converges towards $\int_c \tau \cdot u$ when $\mathcal{S}$ is properly refined.

Further up ($p = 2$), let Σ be a 2-cell, to which a crossing direction has been assigned, and choose the parameterization $\{s, t\} \to \Sigma(s, t)$ in such a way that vectors $\eta(s, t) = \partial_s \Sigma(s, t) \times \partial_t \Sigma(s, t)$ point in this direction. Then set $n(x) = \eta(s, t)/|\eta(s, t)|$, at point $x = \Sigma(s, t)$, to obtain the outer-orienting *unit normal field*. Given a smooth vector field u, we define the *flux* through Σ, thus outer oriented, as the integral $\int_\Sigma n \cdot u$ of the real-valued function $n \cdot u$ with respect to, this time, the Euclidean measure of

[16]This is also its integral over the set $\{x\}$, with respect to the counting measure, in the sense of Integration Theory. The integral over a *finite* set $\{x_1, \ldots, x_k\}$, in this sense, would be $\sum_i \varphi(x_i)$. Notice the difference between this and what we are busy defining right now, the integral on a 0-chain, which will turn out to be a weighted sum of the reals $\varphi(x_i)$.

areas. (No ambiguity on this point, since the status of Σ as a surface has been made clear.)

REMARK 6.2. For Riemann sums, dissect Σ into a family $\mathcal{T}$ of small triangular patches T, whose vectorial areas are $\vec{\mathrm{T}}$, pick a point x_T in each of them, and consider $\sum_{\mathrm{T}\in\mathcal{T}} \vec{\mathrm{T}} \cdot u(x_\mathrm{T})$.

Last, for $p = 3$, and a 3-cell V with outer orientation $+$, the integral of a function f is the standard $\int_V f$, integral of f over the image of V with respect to the Lebesgue measure. This is consistent with the frequent physical interpretation of $\int_V f$ as the quantity, in V, of something (mass, charge, ...) present with density f in V. With outer orientation $-$, the integral is $-\int_V f$. Thus, outer orientation helps fix bookkeeping conventions when f is a rate of variation, like for instance, heat production or absorption. The inner orientation of V is irrelevant here.

Now, let us extend the notion to chains based on oriented cells. In dimension 0, where an oriented point is a point-cum-sign pair $\{x, \varepsilon\}$, a 0-chain m is a finite collection $\{\{x_i, \varepsilon_i\}\colon\ i = 1, \ldots, k\}$ of such pairs, each with a weight μ^i. The integral $\int_m \varphi$ is then defined as $\sum_i \mu^i \varepsilon_i \varphi(x_i)$.[17] In dimension 1, the circulation along the 1-chain $c = \sum_i \mu^i c_i$ is $\int_c \tau \cdot u = \sum_i \mu^i \int_{c_i} \tau \cdot u$. The flux $\int_\Sigma n \cdot u$ through the *twisted* (beware!) chain $\Sigma = \sum_i \mu^i \Sigma_i$ is defined as $\sum_i \mu^i \int_{\Sigma_i} n \cdot u$. As for dimension 3, a twisted chain manifold V is a finite collection[18] $\{\{V_i, \varepsilon_i\}\colon\ i = 1, \ldots, k\}$ of 3D blobs-with-sign, with weights μ^i, and $\int_V f$ is, by definition, $\sum_i \mu^i \varepsilon_i \int_{V_i} f$.

Note that we have implicitly defined integrals on piecewise smooth manifolds there, since these can be considered as cell-based chains with "orientation matching weights" (1 if the cell's orientation and the manifold's match, -1 if they don't).

Thus the most common ways[19] to integrate things in three-space lead to the definition of integrals over *inner* oriented manifolds or chains in cases $p = 0$ and 1 and *outer* oriented ones[20] in cases $p = 2$ and 3. An unpleasant asymmetry. But since we work in *oriented* Euclidean space, where one may, as we have seen, derive outer from inner orientation, and the other way round, this restores the balance, hence finally *eight* kinds of integrals, depending on the dimension and on the nature (internal or external) of the orientation of the underlying chain.

Thus we have obtained a series of maps of type *CHAIN* $\rightarrow$ *REAL*, but in a pretty awkward way, one must admit. Could there be an underlying unifying concept that would make it all simpler?

[17] One might think, there, that orientation-signs and weights do double duty. Indeed, a convention could be made that all points are positively oriented, and this would dispose of the ε_is. We won't do this, for the sake of uniformity of treatment with respect to dimension.

[18] Again, one might outer-orient such elementary volumes by giving them all a $+$ sign, reducing the redundancy, and we refrain to do so for the same reason.

[19] Others reduce to one of these. For instance, when using Cartesian coordinates x–y–z, $\int_c f(x, y, z)\,\mathrm{d}x$ is simply the circulation along c, in the sense we have defined above, of the field of x-directed basis vectors magnified by the scalar factor f.

[20] A tradition initiated in FIRESTONE [1933] distinguishes between so-called "across" and "through" physical quantities (KOENIG and BLACKWELL [1960], BRANIN [1961]), expressible by circulations and fluxes, respectively. As we shall see, this classification is not totally satisfying.

7. Differential forms, and their physical relevance

Indeed, these maps belong to a category of objects that can be defined without recourse to the Euclidean structure, and have thus a purely affine nature:

DEFINITION 7.1. A straight [respectively twisted] differential form of degree p, or p-form, is a real-valued map ω over the space of straight [respectively twisted] p-chains, linear with respect to chain addition, and continuous in the sense of the above-defined topology of chains (end of Section 5).

Differential forms, thus envisioned, are dual objects with respect to chains, which prompts us to mobilize the corresponding machinery of functional analysis (YOSIDA [1980]): Call $\mathcal{F}^p$ [respectively $\widetilde{\mathcal{F}}^p$] the space of straight [respectively twisted] p-forms, as equipped with its so-called "strong" topology.[21] Then $\mathcal{C}_p$ and $\mathcal{F}^p$ [respectively $\widetilde{\mathcal{C}}_p$ and $\widetilde{\mathcal{F}}^p$] are *in duality* via the bilinear bicontinuous map $\{c, \omega\} \to \int_c \omega$, of type *$p$-CHAIN $\times$ p-FORM $\to$ REAL*. A common notation for such duality products being $\langle c; \omega \rangle$, we shall use that as a convenient alternative[22] to $\int_c \omega$. A duality product should be *non-degenerate*, i.e., $\langle c'; \omega \rangle = 0\ \forall c'$ implies $\omega = 0$, and $\langle c; \omega' \rangle = 0\ \forall \omega'$ forces $c = 0$. The former property holds true by definition, and the latter is satisfied because, if $c \neq 0$, one can construct an ad hoc smooth vector field or function with nonzero integral, hence a nonzero form ω such that $\langle c; \omega \rangle \neq 0$.

The above eight kinds of integrals, therefore, are instances of differential forms, which we shall denote (in their order of appearance) by ${}^0\varphi$, ${}^1u$ (circulation of u), ${}^2\tilde{u}$ (flux of u), ${}^3\tilde{\varphi}$, and ${}^0\tilde{\varphi}$, ${}^1\tilde{u}$, ${}^2u$, ${}^3\varphi$. This is of course ad hoc notation, to be abandoned as soon as the transition from fields to forms is achieved. Note the use of the pre-superscript p, accompanied or not by the tilde as the case may be, as an *operator*, that transforms functions or vector fields into differential forms (twisted ones, if the tilde is there). This operator, being relative to a specific Euclidean structure is as a rule metric- and orientation-dependent. (We'll use $\mathbf{P}$, and $\tilde{}$, versus ^{p}, and $\tilde{}$, to distinguish[23] the $\{\cdot, \mathbf{Or}\}$ and the $\{\cdot, Or\}$ structure.) For instance, the 2 in ${}^2u$ means that, given the straight 2-chain S, one uses both the inner orientation of each of its components

[21] Differential forms converge, in this topology, if their integrals converge uniformly on bounded sets of chains. (A *bounded* set B is one that is *absorbed* by any neighborhood V of 0, i.e., such that $\lambda B \subset V$ for some $\lambda > 0$.) We won't have to invoke such technical notions in the sequel. (Again, see HARRISON [1998] for *norms* on (Banach) spaces of differential forms.) Note the generic use of "differential form" here: Whether an object qualifies as differential form depends on the chosen topology on chain spaces.

[22] In line with the convention of Note 4, we shall denote by ω the map $c \to \langle c; \omega \rangle$, and feel free to write $\omega = c \to \langle c; \omega \rangle$. Of course, the symmetric construct $c = \omega \to \langle c; \omega \rangle$ is just as valid. Maps of the latter kind, from forms to reals, were called *currents* in DE RHAM [1960]. (See DE RHAM [1936, p. 220], for the physical justification of the term.) There are, a priori, much more currents than chains (or even chainlets, HARRISON [1998]), and one should not be fooled by the expression "in duality" into thinking that the dual of $\mathcal{F}^p$, i.e., the bidual of $\mathcal{C}_p$, is $\mathcal{C}_p$ itself.

[23] This play on styles is only a temporary device, not to be used beyond the present Chapter. Later we shall revert to the received "musical" notation, which assumes a single, definite metric structure in background, and cares little about ambiguity: $\sharp u$ denotes the vector proxy of form u, and $\flat$U is the form represented by the vector field U.

and the ambient orientation to define a crossing direction, then the metric in order to build a normal vector field n in this direction, over each component of the chain. Then, $\langle S;\, {}^2u\rangle = \int_S n \cdot u$ defines ${}^2u$, a straight 2-form indeed. (Notice that $\langle S;\, {}^2u\rangle$ does *not* depend on the ambient orientation.)

REMARK 7.1. In the foregoing example, it would be improper to describe $\langle S;\, {}^2u\rangle$ as the flux of u "through" S, since the components of S, a straight chain, didn't come equipped with crossing directions. These were derived from the ambient orientation, part of the Euclidean structure, instead of being given as an attribute of S's components. To acknowledge this difference, we shall refer to $\int_S n \cdot u$ as the flux "embraced by" S. This is not mere fussiness, as will be apparent when we discuss magnetic flux.

One may wonder, at this point, whether substituting the single concept of differential form for those of point-value, circulation, flux, etc., has gained us any real generality, besides the obvious advantage of conceptual uniformity. Let us examine this point carefully, because it's an essential part of the deconstruction of Euclidean space we have undertaken.

On the one hand, the condition that differential forms should be continuous with respect to deformations of the underlying manifolds doesn't leave room, in dimension 3, for other kinds of differential forms than the above eight. First, it eliminates many obvious linear functionals from consideration. (For instance, γ being an outer-oriented curve, the *intersection number*, defined as the number of times γ crosses S, counted algebraically (i.e., with sign – if orientations do not match), provides a linear map $S \to S \wedge \gamma$, which is not considered as a bona fide differential form. Indeed, it lacks continuity.) Second, it allows one, by using the Riesz representation theorem, to build vector fields or functions that reduce the given form to one of the eight types: For instance, given a 1-form ω, there is[24] a vector field Ω such that $\langle c;\, \omega\rangle = \int_c \tau \cdot \Omega$, which is our first example of what will later be referred to as a "proxy" field: A scalar or vector field that stands for a differential form. For other degrees, forms in 3D are representable by vector fields ($p = 1$ and 2) or by functions ($p = 0$ and 3).

However, the continuity condition requires less regularity from the proxy fields than the smoothness we have assumed up to now. Not to the point of allowing them to be only piecewise smooth: What is required lies in between, and should be clear from Fig. 7.1, which revisits a well known topic from the present viewpoint. As one sees, the contrived "transmission conditions", about tangential continuity of this or normal continuity of that, are implied by the very definition of forms as continuous maps.

Last, the generalization is genuine in spatial dimensions higher than 3: A two-form in 4-space, for instance, has no vector proxy, as a rule.

So, although differential forms do extend a little the scope of integration, this is but a marginal improvement, at least in the 3D context. The real point lies elsewhere, and will

[24]The proof is involved. From a vector field v, build a 1-chain $\sum_i \mu_i s_i$, akin to the graphic representation of v by arrows, i.e., s_i is an oriented segment that approximates v in a region of volume μ_i. Apply ω to this chain, go to the limit. The real-valued linear map thus generated is then shown, thanks to the continuity of ω, to be continuous with respect to the L^2 norm on vector fields. Hence a Riesz vector field Ω, which turns out to be a proxy for ω.

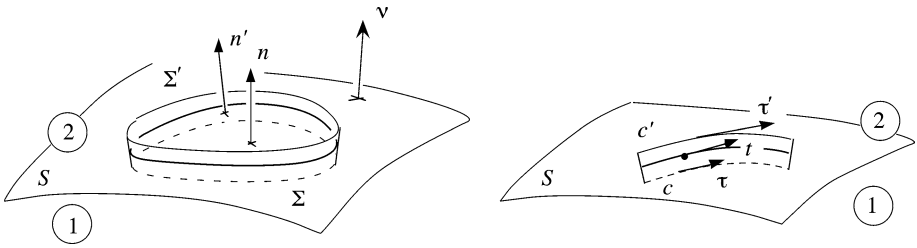

FIG. 7.1. The interface S, equipped with the unit normal field ν, separates two regions where the vector field u is supposed to be smooth, except for a possible discontinuity across S. Suppose Σ or c, initially below S, is moved up a little, thus passing into region 2. Under such conditions, the flux of u through Σ (left) and circulation of u along c (right) can yet be *stable*, i.e., vary continuously with deformations of c and Σ, provided u has some partial regularity: As is well known, and easily proven thanks to the Stokes theorem, *normal* continuity (zero jump $[\nu \cdot u]$ of the normal component across the interface) ensures continuity of the flux $\int_\Sigma n \cdot u$ with respect to Σ (left), while *tangential* continuity of u (zero jump $[u_S]$ of the tangential component across the interface) is required for continuity of the circulation $\int_c \tau \cdot u$ (right) with respect to c. Forms ${}^0\varphi$ and ${}^0\tilde{\varphi}$ require a continuous φ. Piecewise continuity of the proxy function φ is enough for ${}^3\varphi$ and ${}^3\tilde{\varphi}$.

now be argued: Which differential form is built from a given (scalar or vector) field depends on the Euclidean structure, *but the physical entity one purports to model via this field does not*, as a rule. Therefore, the entity of physical significance is the form, conceived as an affine object, and not the field. Two examples will suffice to settle this point.

Consider an electric charge, Q coulombs strong, which is made to move along an oriented smooth curve c, in the direction indicated by the tangent vector field τ. We mean a *test* charge, with Q small enough to leave the ambient electromagnetic field $\{\mathrm{E}, \mathrm{B}\}$ undisturbed, and a *virtual* motion, which allows us to consider the field as frozen at its present value. The work involved in this motion is Q times the quantity $\int_c \tau \cdot \mathrm{E}$, called the *electromotive force* (e.m.f.) *along* c, and expressed in volts (i.e., joules per coulomb). No unit of length is invoked in this description.

Then why is E expressed in volts *per meter* (or whatever unit one adopts)? Only because a vector v such that $|v| = 1$ is one meter long, which makes $\mathrm{E} \cdot v$, and the integral $\int_c \tau \cdot \mathrm{E}$ as well, a definite amount of *volts*, indeed. This physical data, of course, only depends on the field and the curve, not on the metric structure. Yet, change the dot product, from $\cdot$ to $\bullet$ (recall that $u \bullet v = Lu \cdot Lv$), which entails a change in the measure of lengths (hence a rescaling of the unitary vector, now $\boldsymbol{\tau}$ instead of τ), and the circulation of E is now[25] $\int_c \boldsymbol{\tau} \bullet \mathrm{E} = \int_c \tau \cdot L^a L\mathrm{E}$, a different (and physically meaningless) number. On the other hand, there *is* a field $\mathbf{E}$ such that $\int_c \boldsymbol{\tau} \bullet \mathbf{E} = \int_c \tau \cdot \mathrm{E}$, namely $\mathbf{E} = (L^a L)^{-1}\mathrm{E}$. Conclusion: *Which vector field encodes the physical data* (here, e.m.f.'s along all curves) *depends on the chosen metric, although the data themselves do not.* This metric-dependence of E is the reason to call it a vector *proxy*: It merely *stands* for

[25]On the left of the equal sign, the integral and the symbols $\bullet$ and $\boldsymbol{\tau}$ are boldface. (One should see the difference, unless something is amiss in the visualization chain.) So the circulation of E is with respect to the "bold" measure of lengths on the left. The easiest way to verify this equality (and others like it to come) is to work on the above Riemann sums $\sum_s v_s \bullet \mathrm{E}(x_s)$ of the "bold" circulation of E: One has, for each term (omitting the subscript), $v \bullet \mathrm{E} = Lv \cdot L\mathrm{E} = v \cdot L^a L\mathrm{E}$, hence the result.

the real thing, which is the mapping $c \to \langle$e.m.f. along $c\rangle$, i.e., a differential form of degree 1, which we shall from now on denote by e.

Thus, summoning all the equivalent notations introduced so far,

$$e = {}^1\mathrm{E} = {}^1\mathbf{E} = c \to \langle c; e\rangle, \quad \text{where } \langle c; e\rangle \equiv \int_c e = \int_c \tau \cdot \mathrm{E} = \int_c \boldsymbol{\tau} \cdot \mathbf{E}. \tag{7.1}$$

This (straight) 1-form is the right mathematical object by which to represent the electric field, for it tells all about it: Electromotive forces along curves are, one may argue (TONTI [1996]), all that can be observed as regards the electric field.[26] To the point that one can get rid of all the vector-field-and-metric scaffolding, and introduce e directly, by reasoning as follows: The *1-CHAIN → REAL* map we call e.m.f. depends linearly and continuously, *as can experimentally be established*, on the chain over which it is measured. But this is the very definition of a 1-form. Hence e is the minimal, necessary and sufficient, mathematical description of the (empirical) electric field.

REMARK 7.2. The chain/form duality, thus, takes on a neat physical meaning: While the form e models the field, chains are abstractions of the *probes*, of more or less complex structure, that one may place here and there in order to measure it.

The electric field is not the whole electromagnetic field: it only accounts for forces (and their virtual work) exerted on non-moving electric charges. We shall deal later with the magnetic field, which gives the motion-dependent part of the Lorentz force, and recognize it as a 2-form. But right now, an example involving a *twisted* 2-form will be more instructive.

So consider current density, classically a vector field J, whose purpose is to account for the quantity of electric charge, $\int_\Sigma n \cdot \mathrm{J}$, that traverses, per unit of time, a surface Σ in the direction of the unit normal field n that outer-orients it. (Note again this quantity is in ampères, whereas the dimension of the proxy field J is A/m^2.) This map, $\Sigma \to \langle$intensity through $\Sigma\rangle$, a twisted 2-form (namely, ${}^2\tilde{\mathrm{J}}$), is what we can measure and know about the electric current, and the metric plays no role there. Yet, change $\cdot$ to $\bullet$, which affects the measure of areas, and the flux of J becomes[27] $\int_\Sigma \mathbf{n} \bullet \mathrm{J} = |\det(L)| \int_\Sigma n \cdot \mathrm{J}$. The "bold" vector proxy, therefore, should be $\mathbf{J} = |\det(L)|^{-1}\mathrm{J}$, and then ${}^2\tilde{\mathbf{J}} = {}^2\tilde{\mathrm{J}}$. Again, different vector proxies, but the same twisted 2-form, which thus appears as the invariant and physically meaningful object. It will be denoted by j.

This notational scheme will be systematized: Below, we shall call e, h, d, b, j, a, etc., the differential forms that the traditional vector fields E, H, D, B, J, A, etc., represent.

[26] Pointwise values cannot directly be measured, which is why they are somewhat downplayed here, but of course they do make sense, at points of regularity of the field: Taking for c the segment $[x, x+v]$, where v is a vector at x that one lets go to 0, generates at the limit a linear map $v \to \omega_x(v)$. This map, an element of the dual of T_x, is called a *covector* at x. A 1-form, therefore, can be conceived as a (smooth enough) field of covectors. In coordinates, covectors such as $v \to v^i$, where v^i is the ith component of v at point x, form a basis for covectors at x. (They are what is usually denoted by $\mathrm{d}x^i$; but d^i makes better notation, that should be used instead, on a par with ∂_i for basis vectors.)

[27] Same trick, with Riemann sums of the form $\sum_\mathrm{T} \vec{\mathbf{T}} \bullet \mathrm{J}(x_\mathrm{T})$. After (5.2) and (5.4), $\vec{\mathbf{T}} \bullet \mathrm{J} = L\vec{\mathbf{T}} \cdot L\mathrm{J} = L^a L\vec{\mathbf{T}} \cdot \mathrm{J} = |\det(L)|\vec{\mathrm{T}} \cdot \mathrm{J}$. Hence $\int_\Sigma \mathbf{n} \bullet \mathrm{J} = |\det(L)| \int_\Sigma n \cdot \mathrm{J}$.

8. The Stokes theorem

The Stokes "theorem" hardly deserves such a status in the present approach, for it reduces to a mere

DEFINITION 8.1. The exterior derivative $\mathrm{d}\omega$ of the $(p-1)$-form ω is the p-form $c \to \int_{\partial c} \omega$.

In plain words: To integrate $\mathrm{d}\omega$ over the p-chain c, integrate ω over its boundary ∂c. (This applies to straight or twisted chains and forms equally. Note that d is well defined, thanks to the continuity of ∂ from $\mathcal{C}_{p-1}$ to $\mathcal{C}_p$.) In symbols: $\int_{\partial c} \omega = \int_c \mathrm{d}\omega$, which is the common form of the theorem, or equivalently,

$$\langle \partial c; \omega \rangle = \langle c; \mathrm{d}\omega \rangle \quad \forall c \in \mathcal{C}_p \text{ and } \omega \in \mathcal{F}^{p-1} \tag{8.1}$$

(put tildes over $\mathcal{C}$ and $\mathcal{F}$ for twisted chains and forms), which better reveals what is going on: d is the *dual* of ∂ (YOSIDA [1980]). As a corollary of (4.1), one has

$$\mathrm{d} \circ \mathrm{d} = 0. \tag{8.2}$$

A form ω is *closed* if $\mathrm{d}\omega = 0$, and *exact* if $\omega = \mathrm{d}\alpha$ for some form α. (Synonyms, perhaps more mnemonic, are *cocycle* and *coboundary*. The integral of a cocycle over a boundary, or of a coboundary over a cycle, vanishes.)

REMARK 8.1. In A_n, all closed forms are exact: this is known as the *Poincaré Lemma* (see, e.g., SCHUTZ [1980, p. 140]). But closed forms need not be exact in general manifolds: this is the dual aspect of the "not all cycles bound" issue we discussed earlier. Studying forms, consequently, is another way, dual to homology, to investigate topology. The corresponding theory is called *cohomology* (JÄNICH [2001], MADSEN and TORNEHAVE [1997]).

In three dimensions, the d is the affine version of the classical differential operators, grad, rot, and div, which belong to the Euclidean structure. Let's review this.

First, the gradient: Given a smooth function φ, we define $\operatorname{grad}\varphi$ as the vector field such that, for any 1-cell c with unit tangent field τ,

$$\int_c \tau \cdot (\operatorname{grad}\varphi) = \int_{\partial c} \varphi, \tag{8.3}$$

the latter quantity being of course $\varphi(c(1)) - \varphi(c(0))$. By linearity, this extends to any 1-chain. One recognizes (8.1) there. The relation between gradient and d, therefore, is ${}^1(\operatorname{grad}\varphi) = \mathrm{d}\,{}^0\varphi \equiv \mathrm{d}\varphi$, the third term being what is called the *differential* of φ. (The zero superscript can be dropped, because there is only one way to turn a function into a 0-form, whatever the metric.) The vector field $\operatorname{grad}\varphi$ is a proxy for the 1-form $\mathrm{d}\varphi$.

Thus defined, $\operatorname{grad}\varphi$ depends on the metric. If the dot product is changed from "$\cdot$" to "$\bullet$", the vector field whose circulation equals the right-hand side of (8.3) is a different proxy, $\mathbf{grad}\,\varphi$, which relates to the first one, as one will see using (5.2), by $\operatorname{grad}\varphi = L^a L\,\mathbf{grad}\,\varphi$.

Up in degree, rot and div are defined in similar fashion. Thus, all in all,

$$^1(\operatorname{grad}\varphi) = \mathrm{d}\,^0\varphi, \quad ^2(\operatorname{rot} u) = \mathrm{d}\,^1 u, \quad ^3(\operatorname{div} v) = \mathrm{d}\,^2 v. \tag{8.4}$$

Be well aware that all forms here are *straight*. Yet their proxies may behave in confusing ways with respect to orientation, as we shall presently see.

About curl, (8.4) says that the curl of a smooth field u, denoted $\operatorname{rot} u$, is the vector field such that, for any inner oriented surface S,

$$\int_S n \cdot \operatorname{rot} u = \int_{\partial S} \tau \cdot u. \tag{8.5}$$

Here, τ corresponds to the induced orientation of ∂S, and n is obtained by the Ampère rule. So the ambient orientation is explicitly used. Changing it reverses the sign of $\operatorname{rot} u$. The curl behaves like the cross product in this respect. If, moreover, the dot product is changed, the bold curl and the meager one relate as follows:

PROPOSITION 8.1. *With $u \cdot v = Lu \cdot Lv$ and $\mathbf{Or} = \operatorname{sign}(\det(L))Or$, one has*

$$\mathbf{rot}\, u = \big(\det(L)\big)^{-1} \operatorname{rot}(L^a Lu). \tag{8.6}$$

PROOF. Because of the hybrid character of (8.5), with integration over an outer oriented surface on the left, and over an inner oriented line on the right, the computation is error prone, so let's be careful. On the one hand (Note 25), $\int_{\partial S} \boldsymbol{\tau} \cdot u = \int_{\partial S} \tau \cdot L^a Lu = \int_S n \cdot \operatorname{rot}(L^a Lu)$. On the other hand (Note 27), setting $\mathbf{J} = \mathbf{rot}\, u$, we know that $\int_S \mathbf{n} \cdot \mathbf{J} = |\det(L)| \int_S n \cdot \mathbf{J}$, hence … but wait! In Note 27, we had both normals n and $\mathbf{n}$ on the same side of the surface, but here (see Fig. 3.2, left), they may point to opposite directions if $\mathbf{Or} \neq Or$. The correct formula is thus $\int_S \mathbf{n} \cdot \mathbf{rot}\, u = \det(L) \int_S n \cdot \mathbf{rot}\, u \equiv \int_S n \cdot \operatorname{rot}(L^a Lu)$, hence (8.6). □

As for the divergence, (8.4) defines $\operatorname{div} v$ as the function such that, for any volume V with outgoing normal n on ∂V,

$$\int_V \operatorname{div} v = \int_{\partial V} n \cdot v. \tag{8.7}$$

No vagaries due to orientation this time, because both integrals represent the same kind of form (twisted). Moreover, $\mathbf{div}\, v = \operatorname{div} v$, because the same factor $|\det(L)|$ pops up on both sides of $\int_V \mathbf{div}\, v = \int_{\partial V} \mathbf{n} \cdot v$. (These integrals, as indicated by the boldface summation sign, are with respect to the "bold" measure. For the one on the left, it's the 3D measure $|\mathbf{vol}|$, and $\mathbf{vol} = \det(L)\,\mathrm{vol}$ after (5.2).)

REMARK 8.2. The invariance of div is consistent with its physical interpretation: if v is the vector field of a fluid mass, its divergence is the rate of change of the volume occupied by this mass, and though volumes depend on the metric, volume *ratios* do not, again after (5.2).

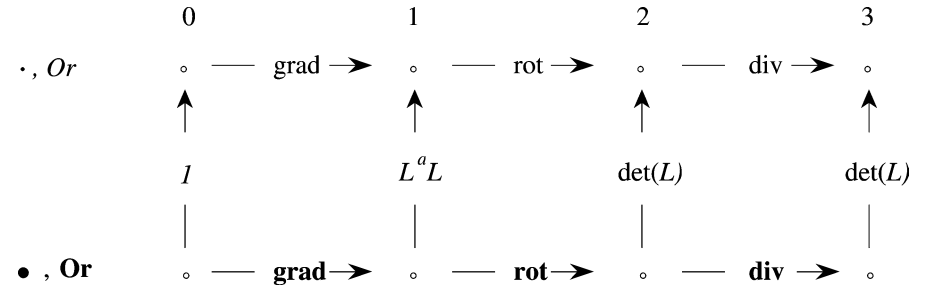

FIG. 8.1. Vertical arrows show how to relate vector or scalar proxies that correspond to the *same* straight form, of degree 0 to 3, in two different Euclidean structures. For *twisted* forms, use the same diagram, but with $|\det(L)|$ substituted for $\det(L)$.

For reference, Fig. 8.1 gathers and displays the previous results. This is a commutative diagram, from which transformation formulas about the differential operators can be read off.[28]

As an illustration of how such a diagram can be used, let us prove something the reader has probably anticipated: the invariance of Faraday's law with respect to a change of metric and orientation. Let two vector fields E and B be such that $\partial_t \mathrm{B} + \mathrm{rot}\,\mathrm{E} = 0$, and set $\mathbf{B} = \mathrm{B}/\det(L)$, $\mathbf{E} = (L^a L)^{-1}\mathrm{E}$, which represent the same differential forms (call them b and e) in the $\{\cdot, \mathbf{Or}\}$ framework, as B and E in the $\{\cdot, Or\}$ one. Then $\partial_t \mathbf{B} + \mathbf{rot}\,\mathbf{E} = 0$. We now turn to the significance of the single physical law underlying these two relations.

9. The magnetic field, as a 2-form

Electromagnetic forces on moving charges, i.e., currents, will now motivate the introduction of the magnetic field. Consider a current loop, I ampères strong, which is made to move – virtual move, again – so as to span a surface S (Fig. 9.1). The virtual work involved is then I times $\int_S n \cdot \mathrm{B}$ ("cut flux" rule), as explained in the caption. Experience establishes the linearity and continuity of the factor $\int_S n \cdot \mathrm{B}$, called the *induction flux*, as a function of S. Hence a 2-form, again the minimal description of the (empirical) magnetic field, which we denote by b and call *magnetic induction*.

In spite of the presence of n in the formula, b is not a twisted but a straight 2-form, as it should, since ambient orientation cannot influence the sign of the virtual work in any way. Indeed, what is relevant is the direction of the current along the loop, which inner-orients c, and the inner orientation of S is the one that matches the orientation of the chain $c' - c$ ("final position minus initial position" in the virtual move). The intervention of a normal field, therefore, appears as the result of the will to represent b with help of a vector, the traditional B such that $b = {}^2\mathrm{B}$. No surprise, then, if this vector

[28] It should be clear that L might depend on the spatial position x, so this diagram is more general than what we contracted for. It gives the correspondence between differential operators relative to different Riemannian structures on the same 3D manifold.

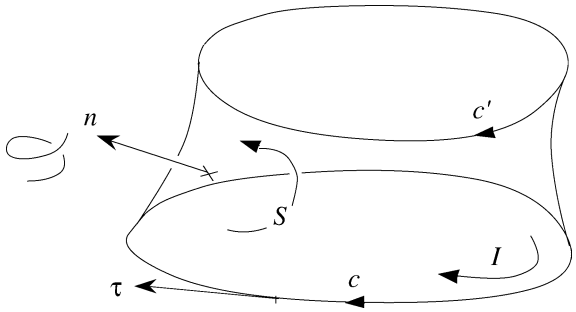

FIG. 9.1. Conventions for the virtual work due to B on a current loop, in a virtual move from position c to position c'. The normal n is the one associated, by Ampère's rule, with the inner orientation of S, a surface such that $\partial S = c' - c$. The virtual work of the J × B force, with $\mathrm{J} = I\tau$, is then I times the flux $\int_S n \cdot \mathrm{B}$.

Nature of the proxy	*for a* straight	*or* twisted	*DF of degree*
function	polar	axial	*0*
vector field	polar	axial	*1*
vector field	axial	polar	*2*
function	axial	polar	*3*

FIG. 9.2. Nature of the proxies in *non*-oriented 3D space with dot product.

proxy "changes sign" with ambient orientation! Actually, it cannot do its job, that is, represent b, without an ambient orientation.

If one insists on a proxy that can act to this effect in autonomy, this object has to carry on its back, so to speak, an orientation of ambient space, i.e., it must be a field of *axial* vectors. Even so, the dependence on metric is still there, so the benefit of using such objects is tiny. Yet, why not, if one is aware that (polar) vector field and axial vector field are just mathematical *tools*,[29] which may be more or less appropriate, depending on the background structures, to represent a given physical entity. In this respect, it may be useful to have a synoptic guide (Fig. 9.2).

We can fully appreciate, now, the difference between j and b, between current flow and magnetic flux. Current density, the twisted 2-form j, is meant to be integrated over surfaces Σ with crossing direction: its proxy J is independent of the ambient orientation. Magnetic induction, the straight 2-form b, is meant to be integrated over surfaces S with inner orientation: its proxy B changes sign if ambient orientation is changed. Current, clearly, flows through a surface, so intensity is one of these "through variables" of

[29] Thus axiality or polarity is by no means a property of the physical objects. But the way physicists write about it doesn't help clarify this. For instance (BAEZ and MUNIAIN [1994, p. 61]): "In physics, the electric field E is called a vector, while the magnetic field B is called an axial vector, because E changes sign under parity transformation, while B does not". Or else (ROSEN [1973]): "It is well known that under the space inversion transformation, $P : (x, y, z) \to (-x, -y, -z)$, the electric field transforms as a polar vector, while the magnetic field transforms as an axial vector, $P : \{E \to -E, B \to B\}$". This may foster confusion, as some passages in BALDOMIR and HAMMOND [1996] demonstrate.

Note 20. But thinking of the magnetic flux as going *through* S is misleading. Hence the expression used here, flux *embraced by* a surface.[30]

10. Faraday and Ampère

We are now ready to address Faraday's famous experiment: variations of the flux embraced by a conducting loop create an electromotive force. A mathematical statement meant to express this law with maximal economy will therefore establish a link between the integral of b over a fixed surface S and the integral of e over its boundary ∂S. Here it is: one has

$$\partial_t \int_S b + \int_{\partial S} e = 0 \quad \forall S \in \mathcal{C}_2, \tag{10.1}$$

i.e., for any straight 2-chain, and in particular, any inner oriented surface S. Numbers in (10.1) have dimension: webers for the first integral, and volts (i.e., Wb/s) for the second one. *Inner* orientation of ∂S (and hence, of S itself) makes lots of physical sense: it corresponds to selecting one of the two ways a galvanometer can be inserted in the circuit idealized by ∂S. Applying the Stokes theorem – or should we say, the definition of d – we find the local, infinitesimal version of the global, integral law (10.1), as this:

$$\partial_t b + \mathrm{d}e = 0, \tag{10.2}$$

the metric- and orientation-free version of $\partial_t \mathrm{B} + \mathrm{rot}\,\mathrm{E} = 0$.

As for Ampère's theorem, the expression is similar, except that twisted forms are now involved:

$$-\partial_t \int_\Sigma d + \int_{\partial \Sigma} h = \int_\Sigma j \quad \forall \Sigma \in \widetilde{\mathcal{C}}_2, \tag{10.3}$$

i.e., for any twisted 2-chain, and in particular, any outer oriented surface Σ. Its infinitesimal form is

$$-\partial_t d + \mathrm{d}h = j, \tag{10.4}$$

again the purely affine version of $-\partial_t \mathrm{D} + \mathrm{rot}\,\mathrm{H} = \mathrm{J}$. Since j is a twisted form, d must be one, and h as well,[31] which suggests that its proxy H will not behave like E under a change of the background Euclidean structure. Indeed, one has $\mathbf{H} = \mathrm{sign}(\det(L))(L^a L)^{-1}\mathrm{H}$ in the now familiar notation. In non-oriented space with metric, the proxy H would be an axial vector field, on a par with B. Vector proxies D and J would be polar, like E.

At this stage, we may announce the strategy that will lead to a discretized form of (10.1) and (10.3): Instead of requesting their validity for *all* chains S or Σ, we shall be

[30] This exposes the relative inadequacy of the "across vs. through" concept, notions which roughly match those of straight 1-form and twisted 2-form (BRANIN [1961]). Actually, between lines and surfaces on the one hand, and inner or outer orientation on the other hand, it's *four* different "vectorial" entities one may have to deal with, and the vocabulary may not be rich enough to cope.

[31] A *magnetomotive force* (m.m.f.), therefore, is a real value (in ampères) attached to an *outer* oriented line γ, namely the integral $\int_\gamma h$.

content with enforcing them for a *finite* family of chains, those generated by the 2-cells of an appropriate finite element mesh, hence a system of differential equations. But first, we must deal with the constitutive laws linking b and d to h and e.

11. The Hodge operator

For it seems a serious difficulty exists there: Since b and h, or d and e, are objects of different types, simple proportionality relations between them, such as $b = \mu h$ and $d = \varepsilon e$, won't make sense if μ and ε are mere scalar factors. To save this way of writing, as it is of course desirable, we must properly redefine μ and ε as *operators*, of type 1-*FORM* $\to$ 2-*FORM*, one of the forms twisted, the other one straight.

So let's try to see what it takes to go from e to d. It consists in being able to determine $\int_\Sigma d$ over any given outer oriented surface Σ, knowing two things: the form e on the one hand, i.e., the value $\int_c e$ for any inner oriented curve c, and the relation $\mathrm{D} = \varepsilon \mathrm{E}$ between the proxies, on the other hand. (Note that ε can depend on position. We shall assume it's piecewise smooth.) How can that be done?

The answer is almost obvious if Σ is a small[32] piece of plane. Build, then, a small segment c meeting Σ orthogonally at a point x where ε is smooth. Associate with c the vector $\vec{c}$ of same length that points along the crossing direction through Σ, and let this vector also serve to inner-orient c. Let $\vec{\Sigma}$ stand for the vectorial area of Σ, and take note that $\vec{\Sigma}/\operatorname{area}(\Sigma) = \vec{c}/\operatorname{length}(c)$. Now dot-multiply this equality by D on the left, εE on the right. The result is

$$\int_\Sigma d = \varepsilon(x) \frac{\operatorname{area}(\Sigma)}{\operatorname{length}(c)} \int_c e, \tag{11.1}$$

which does answer the question.

How to lift the restrictive hypothesis that Σ be small? Riemann sums, again, are the key. Divide Σ into small patches T, as above (Fig. 6.1, right), equip each of them with a small orthogonal segment c_T, meeting it at x_T, and such that $\vec{c}_\mathrm{T} = \vec{\mathrm{T}}$. Next, define $\int_\Sigma d$ as the limit of the Riemann sums[33] $\sum_\mathrm{T} \varepsilon(x_\mathrm{T}) \int_{c_\mathrm{T}} e$. One may then define the *operator* ε, with reuse of the symbol, as the map $e \to d$ just constructed, from $\mathcal{F}^1$ to $\widetilde{\mathcal{F}}^2$. A similar definition holds for μ, of type $\widetilde{\mathcal{F}}^1 \to \mathcal{F}^2$, and for the operators ε^{-1} and μ^{-1} going in the other direction. (Later, we shall substitute ν for μ^{-1}.)

REMARK 11.1. We leave aside the anisotropic case, with a (symmetric) tensor ε^{ij} instead of the scalar ε. In short: Among the variant "bold" metrics, there is one in which ε^{ij} reduces to unity. Then apply what precedes, with "orthogonality", "length", and "area" understood in the sense of this modified metric. (The latter may depend on position, however, so this stands a bit outside our present framework. Details are given in BOSSAVIT [2001b].)

[32] To make up for the lack of rigor which this word betrays, one should treat c and Σ as "p-vectors" ($p = 1$ and 2 respectively), which are the infinitesimal avatars of p-chains. See BOSSAVIT [1998b] for this approach.

[33] Singular points of ε, at which $\varepsilon(x_\mathrm{T})$ is not well defined, can always be avoided in such a process, unless Σ coincides with a surface of singularities, like a material interface. But then, move Σ a little, and extend d to such surfaces by continuity.

REMARK 11.2. When the scalar ε or μ equals 1, what has just been defined is the classical *Hodge operator* of differential geometry (BURKE [1985], SCHUTZ [1980]), usually denoted by $*$, which maps p-forms, straight or twisted, to $(n-p)$-forms of the other kind, with $** = \pm 1$, depending on n and p. In dimension $n = 3$, it's a simple exercise to show that the above construction then reduces to $*^1 u = {}^2\tilde{u}$, which prompts the following definition: $*^0\varphi = {}^3\tilde{\varphi}$, $*^1 u = {}^2\tilde{u}$, $*^2 u = {}^1\tilde{u}$, ${}^3\varphi = *^0\tilde{\varphi}$. Note that $** = 1$ for all p in 3D.

The metric structure has played an essential role in this definition: areas, lengths, and orthogonality depend on it. So we now distinguish, in the Maxwell equations, the two metric-free main ones,

$$\partial_t b + \mathrm{d}e = 0, \tag{10.2}$$

$$-\partial_t d + \mathrm{d}h = j, \tag{10.4}$$

and the metric-dependent constitutive laws

$$b = \mu h, \tag{11.2}$$

$$d = \varepsilon e, \tag{11.3}$$

where μ and ε are operators of the kind just described. To the extent that no metric element is present in these equations, except for the operators μ and ε, from which one can show the metric can be inferred (BOSSAVIT [2001b]), one may even adopt the radical point of view (DI CARLO and TIERO [1991]) that μ and ε *encode* the metric information.

12. The Maxwell equations: Discussion

With initial conditions on e and h at time $t = 0$, and conditions about the "energy" of the fields to which we soon return, the above system makes a well-posed problem. Yet a few loose ends must be tied.

First, recall that j is supposed to be known. But reintroducing Ohm's law at this stage would be no problem: replace j in (10.4) by $j^s + \sigma e$, where j^s is a given twisted 2-form (the source current), and σ a third Hodge-like operator on the model of ε and μ.

12.1. Boundary conditions, transmission conditions

Second, boundary conditions, if any. Leaving aside artificial "absorbing" boundary conditions (MITTRA, RAMAHI, KHEBIR, GORDON and KOUKI [1989]), not addressed here, there are essentially four basic ones, as follows.

Let's begin with "electric walls", i.e., boundaries of perfect conductors, inside which $\mathrm{E} = 0$, hence the standard $n \times \mathrm{E} = 0$ on the boundary. In terms of the form e, it means that $\int_c e = 0$ for all curves c contained in such a surface. This motivates the following definition, stated in dimension n for generality: S being an $(n-1)$-manifold, call $\mathcal{C}_p(S)$ the space of p-chains whose components are all supported in S; then,

DEFINITION 12.1. The trace $\mathrm{t}_S\omega$ of the p-form ω is the restriction of ω to $\mathcal{C}_p(S)$, i.e., the map $c \to \int_c \omega$ restricted to p-chains based on components which are contained in S.

Of course this requires $p < n$. So the boundary condition at an electric wall S^e is $\mathrm{t}_{S^e} e = 0$, which we shall rather write, for the sake of clarity, as "$\mathrm{t}e = 0$ on S^e". Symmetrically, the condition $\mathrm{t}h = 0$ on S^h corresponds to a magnetic wall S^h.

The Stokes theorem shows that d, and t, commute: $\mathrm{d}\mathrm{t}\omega = \mathrm{t}\mathrm{d}\omega$ for any ω of degree not higher than $n-2$. Therefore $\mathrm{t}e = 0$ implies $\mathrm{t}\mathrm{d}e = 0$, hence $\partial_t(\mathrm{t}b) = 0$ by (10.2), that is, $\mathrm{t}b = 0$ if one starts from null fields at time 0. For the physical interpretation of this, observe that $\mathrm{t}b = 0$ on S^b means $\int_S b = 0$ for any surface piece S belonging to S^b, or else, in terms of the vector proxy, $\int_S n \cdot \mathrm{B} = 0$, which implies $n \cdot \mathrm{B} = 0$ on all S^b: a "no-flux" surface, called a "magnetic barrier" by some. We just proved anew, in the present language, that electric walls are impervious to magnetic flux. One will see in the same manner that $\mathrm{t}j = 0$ corresponds to "insulating boundaries" ($n \cdot \mathrm{J} = 0$) and $\mathrm{t}d = 0$ to "dielectric barriers" ($n \cdot \mathrm{D} = 0$). If j is given with $\mathrm{t}j = 0$ at the boundary of the domain of interest (which is most often the case) then $\mathrm{t}h = 0$ on S^h implies $\mathrm{t}d = 0$ there. (In eddy current problems, where d is neglected, but j is only partially given, $\mathrm{t}h = 0$ on S^h implies $\mathrm{t}j = 0$, i.e., no current through the surface.)

Conditions $\mathrm{t}b = 0$ or $\mathrm{t}d = 0$ being thus weaker than $\mathrm{t}e = 0$ or $\mathrm{t}h = 0$, one may well want to enforce them independently. Many combinations are thereby possible. As a rule (but there are exceptions in non-trivial topologies, see BOSSAVIT [2000]), well-posedness in a domain D bounded by surface S obtains if S can be subdivided as $S = S^e \cup S^h \cup S^{eh}$, with $\mathrm{t}e = 0$ on S^e (electric wall), $\mathrm{t}h = 0$ on S^h (magnetic wall), and *both* conditions $\mathrm{t}\mathrm{d}e = 0$ and $\mathrm{t}\mathrm{d}h = 0$ on S^{eh}, which corresponds to $\mathrm{t}b = 0$ and $\mathrm{t}d = 0$ taken together (boundary which is both a magnetic and a dielectric barrier, or, in the case of eddy-current problems, an insulating interface).

REMARK 12.1. It may come as a surprise that the standard Dirichlet/Neumann opposition is not relevant here. It's because a Neumann condition is just a Dirichlet condition composed with the Hodge and the trace operators (BOSSAVIT [2001c]): Take for instance the standard $n \times \mu^{-1}\,\mathrm{rot}\,\mathrm{E} = 0$, which holds on magnetic walls in the E formulation. This is (up to an integration with respect to time) the proxy form of $\mathrm{t}h = 0$, i.e., of the *Dirichlet* condition $n \times \mathrm{H} = 0$. In short, Neumann conditions on e are Dirichlet conditions on h, and the other way round. They only become relevant when one eliminates either e or h in order to formulate the problem in terms of the other field exclusively, thus breaking the symmetry inherent in Maxwell's equations (which we have no intention to do unless forced to!).

Third point, what about the apparently missing equations, $\mathrm{div}\,\mathrm{D} = \mathrm{Q}$ and $\mathrm{div}\,\mathrm{B} = 0$ in their classical form (Q is the density of electric charge)? These are not equations, actually, but relations implied by the Maxwell equations, or at best, constraints that initial conditions should satisfy, as we now show.

Let's first define q, the electric charge, of which the above Q is the proxy scalar field. Since j accounts for its flow, charge conservation implies $d_t \int_V q + \int_{\partial V} j = 0$ for all

volumes V, an integral law the infinitesimal form of which is

$$\partial_t q + \mathrm{d}j = 0. \tag{12.1}$$

Suppose both q and j were null before time $t = 0$. Later, then, $q(t) = -\int_0^t (\mathrm{d}j)(s)\,ds$. Note that q, like $\mathrm{d}j$, is a *twisted* 3-form, as should be the case for something that accounts for the density of a substance. (Twisted forms are often called "densities", by the way, as in BURKE [1985].)

Now, if one accepts the physical premise that no electromagnetic field exists until its sources (charges and their flow, i.e., q and j) depart from zero, all fields are null at $t = 0$, and in particular, after (10.4), $d(t) = d(0) + \int_0^t [(\mathrm{d}h)(s) - j(s)]\,ds$, hence, by using (8.2), $\mathrm{d}\,d(t) = -\int_0^t (\mathrm{d}j)(s)\,ds \equiv q(t)$, at all times, hence the derived relation $\mathrm{d}d = q$. As for b, the same computation shows that $\mathrm{d}b = 0$.

So-called "transmission conditions", classically $[n \times \mathrm{E}] = 0$, $[n \cdot \mathrm{B}] = 0$, etc., at material interfaces, can be evoked at this juncture, for these too are not equations, in the sense of additional constraints that the unknowns e, b, etc., would have to satisfy. They *are* satisfied from the outset, being a consequence of the very definition of differential forms (cf. Fig. 7.1).

12.2. Wedge product, energy

Fourth point, the notion of energy. The physical significance of such integrals as $\int \mathrm{B} \cdot \mathrm{H}$ or $\int \mathrm{J} \cdot \mathrm{E}$ is well known, and it's easy to show, using the relations displayed on Fig. 8.1, that both are metric-independent. So they should be expressible in non-metric terms. This is so, thanks to the notion of *wedge product*, an operation which creates a $(p+q)$-form $\omega \wedge \eta$ (straight when both factors are of the same kind, twisted otherwise) out of a p-form ω and a q-form η. We shall only describe this in detail in the case of a 2-form b and a 1-form h, respectively straight and twisted.

The result, a twisted 3-form $b \wedge h$, is known if integrals $\int_V b \wedge h$ are known for all volumes V. In quite the same way as with the Hodge map, the thing is easy when V is a small parallelepiped, as shown in Fig. 12.1. Observe that, if $b = {}^2\mathrm{B}$ and $h = {}^1\widetilde{\mathrm{H}}$, then $\int_V b \wedge h = \mathrm{B} \cdot \mathrm{H}\,\mathrm{vol}(V)$, if one follows the recipe of Fig. 12.1, confirming the soundness of the latter. The extension to finite-size volumes is made by constructing Riemann sums, as usual.

REMARK 12.2. Starting from the equality $\int b \wedge h' = \int \mathrm{B} \cdot \mathrm{H}'$, setting $b = \mu h$ yields $\int \mu\, h \wedge h' = \int \mu \mathrm{H} \cdot \mathrm{H}' = \int \mu \mathrm{H}' \cdot \mathrm{H} = \int \mu h' \wedge h$, a *symmetry* property of the Hodge operator to which we didn't pay attention till now. Note also that $\int \mu h \wedge h = \int \mu |\mathrm{H}|^2 > 0$, unless $h = 0$. Integrals such as $\int \mu h \wedge h'$, or $\int \nu b \wedge b'$, etc., can thus be understood as *scalar products* on spaces of forms, which can thereby be turned (after due completion) into Hilbert spaces. The corresponding norms, i.e., the square roots of $\int \mu h \wedge h$, of $\int \nu b \wedge b$, and other similar constructs on e or d, will be denoted by $|h|_\mu$, $|b|_\nu$, etc.

Other possible wedge products are ${}^0\varphi \wedge \omega = {}^0(\varphi\omega)$ (whatever the degree of ω), ${}^1u \wedge {}^1v = {}^2(u \times v)$, ${}^2u \wedge {}^1v = {}^3(u \cdot v)$. (If none or both factors are straight forms, the product is straight.) It's an instructive exercise to work out the exterior derivative of

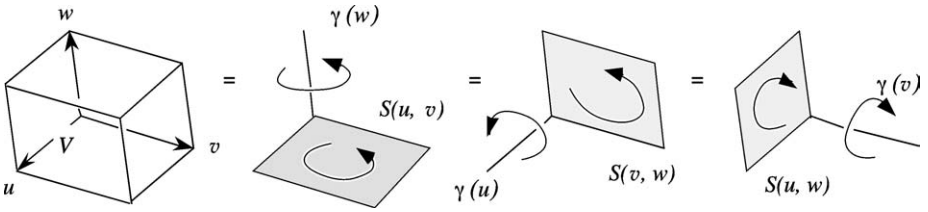

FIG. 12.1. There are three ways, as shown, to see volume V, built on u, v, w, as the extrusion of a surface S along a line segment γ. A natural definition of the integral of $b \wedge h$ is then $\int_V b \wedge h = (\int_{S(u,v)} b)(\int_{\gamma(w)} h) + (\int_{S(v,w)} b)(\int_{\gamma(u)} h) + (\int_{S(u,w)} b)(\int_{\gamma(v)} h)$. Note the simultaneous inner and outer orientations of S and γ, which should match (if the outer orientation of V is $+$, as assumed), but are otherwise arbitrary.

such products, using the Stokes theorem, and to look for the equivalents of the standard integration by parts formulas, such as

$$\int_\Omega (\mathrm{H} \cdot \operatorname{rot} \mathrm{E} - \mathrm{E} \cdot \operatorname{rot} \mathrm{H}) = \int_{\partial\Omega} n \cdot (\mathrm{E} \times \mathrm{H}),$$
$$\int_\Omega (\mathrm{D} \cdot \operatorname{grad} \Psi + \Psi \operatorname{div} \mathrm{D}) = \int_{\partial\Omega} \Psi n \cdot \mathrm{D}.$$

They are, respectively,

$$\int_\Omega (\mathrm{d}e \wedge h - e \wedge \mathrm{d}h) = \int_{\partial\Omega} e \wedge h, \tag{12.2}$$
$$\int_\Omega (\mathrm{d}\psi \wedge d + \psi\, \mathrm{d}d) = \int_{\partial\Omega} \psi\, d. \tag{12.3}$$

Now, let us consider a physically admissible field, that is, a quartet of forms b, h, e, d, which may or may not satisfy Maxwell's equations when taken together, but are each of the right degree and kind in this respect.

DEFINITION 12.2. The following quantities:

$$\frac{1}{2}\int \mu^{-1} b \wedge b, \quad \frac{1}{2}\int \mu h \wedge h, \quad \frac{1}{2}\int \varepsilon e \wedge e, \quad \frac{1}{2}\int \varepsilon^{-1} d \wedge d, \tag{12.4}$$

are called, respectively, *magnetic energy*, *magnetic coenergy*, *electric energy*, and *electric coenergy* of the field. The integral $\int j \wedge e$ is the *power* released by the field.

The latter definition, easily derived from the expression of the Lorentz force, is a statement about field–matter energy exchanges from which the use of the word "energy" could rigorously be justified, although we shall not attempt that here (cf. BOSSAVIT [1990a]). The definition entails the following relations:

$$\frac{1}{2}\int \mu^{-1} b \wedge b + \frac{1}{2}\int \mu h \wedge h \geqslant \int b \wedge h,$$
$$\frac{1}{2}\int \varepsilon^{-1} d \wedge d + \frac{1}{2}\int \varepsilon\, e \wedge e \geqslant \int d \wedge e,$$

with equality if and only if $b = \mu h$ and $d = \varepsilon e$. One may use this as a way to set up the constitutive laws.

REMARK 12.3. The well-posedness evoked earlier holds if one restricts the search to fields with finite energy. Otherwise, of course, nonzero solutions to (10.2), (10.4), (11.2), (11.3) with $j = 0$ do exist (such as, for instance, plane waves).

The integrals in (12.4) concern the whole space, or at least, the whole region of existence of the field. One may wish to integrate on some domain Ω only, and to account for the energy balance. This is again an easy exercise:

PROPOSITION 12.1 (Poynting's theorem). *If the field* $\{b, h, e, d\}$ *does satisfy the Maxwell equations* (10.2), (10.4), (11.2), (11.3), *one has*

$$d_t\left[\frac{1}{2}\int_\Omega \mu^{-1}b\wedge b + \frac{1}{2}\int_\Omega \varepsilon e\wedge e\right] + \int_{\partial\Omega} e\wedge h = -\int_\Omega j\wedge e$$

for any fixed domain Ω.

PROOF. "Wedge multiply" (10.2) and (10.4), from the right, by e and $-h$, add, use (12.2) and Stokes. □

As one sees, all equalities and inequalities on which a variational approach to Maxwell's theory can be based do have their counterparts with differential forms. We shall not follow this thread any further, since what comes ahead is not essentially based on variational methods. Let's rather close this section with a quick review of various differential forms in Maxwell's theory and how they relate.

12.3. The "Maxwell house"

To the field quartet and the source pair $\{q, j\}$, one may add the *electric potential* ψ and the *vector potential* a, a straight 0-form and 1-form respectively, such that $b = \mathrm{d}a$ and $e = -\partial_t a + \mathrm{d}\psi$. Also, the *magnetic potential* φ (twisted 0-form) and the twisted 1-form τ such that $h = \tau + \mathrm{d}\varphi$, whose proxy is the T of Carpenter's "T–Ω" method (CARPENTER [1977]). None of them is as fundamental as those in (10.2), (10.4), but each can be a useful auxiliary at times. The *magnetic current* k and *magnetic charge* m can be added to the list for the sake of symmetry (Fig. 12.2), although they don't seem to represent any real thing (GOLDHABER and TROWER [1990]).

For easier reference, Fig. 12.2 displays all these entities as an organized whole, each one "lodged" according to its degree and nature as a differential form. Since primitives in time may have to be considered, we can group the differential forms of electromagnetism in four similar categories, shown as vertical pillars on the figure. Each pillar symbolizes the structure made by spaces of forms of all degrees, linked together by the d operator. Straight forms are on the left and twisted forms on the right. Differentiation or integration with respect to time links each pair of pillars (the front one and the rear one) forming the sides of the structure. Horizontal beams symbolize constitutive laws.

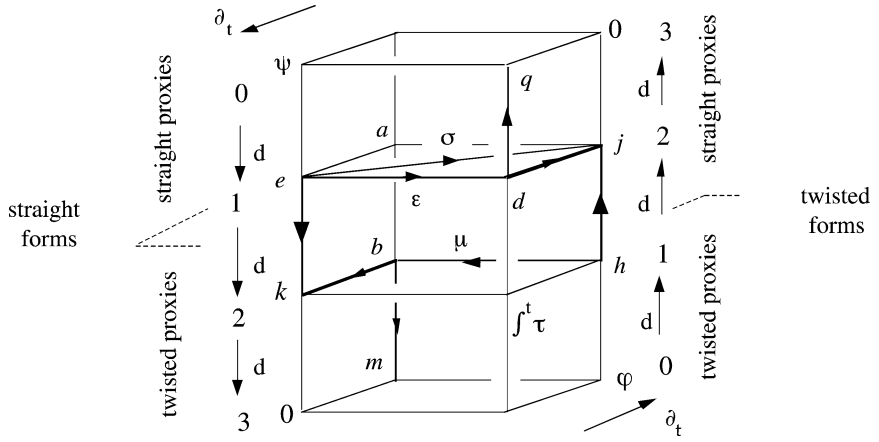

FIG. 12.2. Structures underlying the Maxwell system of equations. For more emphasis on their symmetry, Faraday's law is here taken to be $\partial_t b + \mathrm{d}e = -k$, with $k = 0$. (The straight 2-form k would stand for the flow of magnetic charge, if such a thing existed. Then, one would have $\mathrm{d}b = m$, where the straight 3-form m represents magnetic charge, linked with its current by the conservation law $\partial_t m + \mathrm{d}k = 0$.)

As one can see, each object has its own room in the building: b, a 2-form, at level 2 of the "straight" side, the 1-form a such that $b = \mathrm{d}a$ just above it, etc. Occasional asymmetries (e.g., the necessity to time-integrate τ before lodging it, the bizarre lay-out of Ohm's law ...) point to weaknesses which are less those of the diagram than those of the received nomenclature or (more ominously) to some hitch about Ohm's law (BOSSAVIT [1996]). Relations mentioned up to now can be directly read off from the diagram, up to sporadic sign inversions. An equation such as $\partial_t b + \mathrm{d}e = -k$, for instance, is obtained by gathering at the location of k the contributions of all adjacent niches, including k's, in the direction of the arrows. Note how the rules of Fig. 9.2, about which scalar- or vector-proxies must be twisted or straight, are in force.

But the most important thing is probably the neat separation, in the diagram, between "vertical" relations, of purely affine nature, and "horizontal" ones, which depend on metric. If this was not drawing too much on the metaphor, one could say that a change of metric, as encoded in ε and μ (due for instance to a change in their local values, because of a temperature modification or whatever) would shake the building horizontally but leave the vertical panels unscathed.

This suggests a method for *discretizing* the Maxwell equations: The orderly structure of Fig. 12.1 should be preserved, if at all possible, in numerical simulations. Hence in particular the search for finite elements *which fit differential forms*, which will be among our concerns in the sequel.

CHAPTER III

Discretizing

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, but heuristic considerations will be based on the much simpler case of static fields. The following example can illustrate both things, depending on whether the exciting current, source of the field, is transient or permanent, and lends itself to other useful variations.

13. A model problem

In a closed cavity with metallic walls (Fig. 13.1), which has been free from any electromagnetic activity 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 or magnetizable bodies inside the cavity, too, may scatter waves. Hence a complex evolution, which one may imagine simulating by numerical means. (How else?)

For the sake of generality, let's assume a symmetry plane, and a symmetrically distributed current. (In that case, the plane acts as a magnetic wall.) 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^e and Σ^h, as Fig. 13.1 shows, the two parts

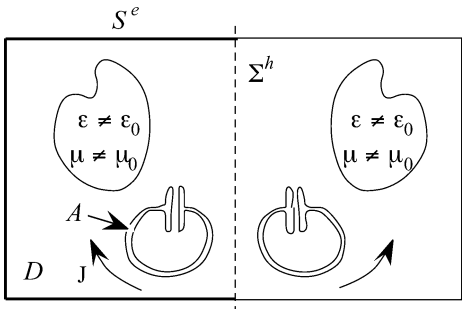

FIG. 13.1. Situation and notation (dimension 3). Region D is the left half of the cavity. Its boundary S has a part S^e in the conductive wall and a part Σ^h in the symmetry plane. Region A, the left "antenna", is the support of the given current density J (mirrored on the right), for which some generator, not represented and not included in the modelling, is responsible.

of its surface, an electric wall and a magnetic wall respectively, we write the relevant equations in D as

$$\begin{aligned} &\partial_t b + \mathrm{d}e = 0, && -\partial_t d + \mathrm{d}h = j, \\ &d = \varepsilon e, && b = \mu h, \\ &\mathrm{t}e = 0 \quad \text{on } S^e, && \mathrm{t}h = 0 \quad \text{on } \Sigma^h. \end{aligned} \tag{13.1}$$

The coefficients ε and μ which generate their Hodge namesakes are real, constant in time, but not necessarily equal to their vacuum values ε_0 and μ_0, and may therefore depend on x. (They could even be tensors, as observed earlier.) The current density j is given, and assumed to satisfy $j(t) = 0$ for $t \leqslant 0$. All fields, besides j, are supposed to be null before $t = 0$, hence initial conditions $e(0) = 0$ and $h(0) = 0$. Notice that $\mathrm{d}j = 0$ is *not* assumed: some electric charge may accumulate at places in the antenna, in accordance with the charge-conservation equation (12.1).

Proving this problem well-posed[34] is not our concern. Let's just recall that it is so, under reasonable conditions on j, when all fields e and h are constrained to have finite energy.

Two further examples will be useful. Suppose j has reached a steady value for so long that all fields are now time-independent. The magnetic part of the field, i.e., the pair $\{b, h\}$, can then be obtained by solving, in domain D,

$$\begin{aligned} &\mathrm{d}b = 0, && \mathrm{d}h = j, \\ &\qquad b = \mu h, \\ &\mathrm{t}b = 0 \quad \text{on } S^e, && \mathrm{t}h = 0 \quad \text{on } \Sigma^h. \end{aligned} \tag{13.2}$$

This is also a well-posed problem (magnetostatics), provided $\mathrm{d}j = 0$. As for the electric part of the field, which has no reason to be zero since the asymptotic charge density $q = q(\infty) = -\int_0^\infty \mathrm{d}j(t)\,dt$ does not vanish, as a rule, one will find it by solving

$$\begin{aligned} &\mathrm{d}d = q, && \mathrm{d}e = 0, \\ &\qquad d = \varepsilon e, \\ &\mathrm{t}e = 0 \quad \text{on } S^e, && \mathrm{t}d = 0 \quad \text{on } \Sigma^h \end{aligned} \tag{13.3}$$

(electrostatics). The easy task of justifying the boundary conditions in (13.2) and (13.3) is left to the reader. One should recognize in (13.3), thinly veiled behind the present notation, the most canonical example there is of elliptic boundary-value problem.[35]

Finally, let's give an example of eddy-current problem in harmonic regime, assuming a conductivity $\sigma \geqslant 0$ in D and $\sigma = 0$ in A. This time, all fields are of the form $u(t, x) =$

[34] Its physical relevance has been challenged (by SMYTH and SMYTH [1977]), on the grounds that assuming a given current density (which is routinely done in such problems) neglects the reaction of the antenna to its own radiated field. This is of course true – and there are other simplifications that one might discuss – but misses the point of what *modelling* is about. See UMAN [1977] and BOSSAVIT [1998b, p. 153], for a discussion of this issue.

[35] Mere changes of symbols would yield the stationary heat equation, the equation of steady flow in porous media, etc. Notice in particular how the steady current equation, with Ohm's law, can be written as $\mathrm{d}j = 0$, $j = \sigma e$, $\mathrm{d}e = 0$, plus boundary conditions (non-homogeneous, to include source terms).

Re[exp($i\omega t$) U(x)], with U complex-valued (SMALL CAPITALS will denote such fields). The given current in A, now denoted J^s (s for "source"), is solenoidal, displacement currents are neglected, and Ohm's law $\mathrm{J} = \sigma\,\mathrm{E} + \mathrm{J}^s$ is in force, where σ is of course understood as a Hodge-like operator, but positive semi-definite only. The problem is then, with the same boundary conditions as above,

$$\mathrm{dH} = \sigma\,\mathrm{E} + \mathrm{J}^s, \qquad \mathrm{H} = \nu\mathrm{B}, \qquad \mathrm{dE} = -i\omega\mathrm{B},$$

and B and H can be eliminated, hence a second-order equation in terms of E:

$$i\omega\sigma\,\mathrm{E} + \mathrm{d}\nu\,\mathrm{dE} = -i\omega\mathrm{J}^s, \tag{13.4}$$

with boundary conditions $\mathrm{tE} = 0$ on S^e and $\mathrm{t}\nu\,\mathrm{dE} = 0$ on Σ^h.

Nothing forbids σ and μ there to be complex-valued too. (Let's however request them to have Hermitian symmetry.) A complex μ can sometimes serve as a crude but effective way to model ferromagnetic hysteresis. And since the real σ can be replaced by $\sigma + i\omega\varepsilon$, we are not committed to drop out displacement currents, after all. Hence, (13.4) can well be construed as the general version of the Maxwell equations in harmonic regime, at angular frequency ω, with dissipative materials possibly present. In particular, (13.4) can serve as a model for the "microwave oven" problem. Note that what we have here is a Fredholm equation: Omitting the excitation term J^s and replacing σ by $i\omega\varepsilon$ gives the "resonant cavity problem" in D, namely, to find frequencies ω at which $\mathrm{d}\nu\,\mathrm{dE} = \omega^2\varepsilon\mathrm{E}$ has a nonzero solution E.

14. 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 is usual about some details. We pretend to work in n-dimensional Euclidean space E_n, but of course $n = 3$ is the case in point. The cells we use here are those introduced earlier [36] (Fig. 2.1), with the important caveat that they are all "open" cells, in the sense of Section 2, i.e., do not include their boundaries. (The only exception is for $p = 0$, nodes, which are both open and closed.) The corresponding closed cell will be denoted with an overbar (also used for the topological closure).

This being said, a *cellular paving* of some region R of space is a finite set of open 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. 14.1, left, c_3 is a face of c_4. If c is a face of c' which itself is a face of c'', then c is a face of c''. Cells in ambient dimension 3 or lower will be called *nodes*, *edges*, *facets*, and *volumes*, with symbols n, e, f, v to match.

We'll say we have a *closed* paving if R is closed. (Fig. 14.1, left, gives a two-dimensional example, where $R = \overline{D}$.) But it need not be so. Closed pavings are not

[36] Topologically simple *smooth* cells, therefore. But the latter condition is not strict and we shall relax it to *piecewise* smooth, in the sequel, without special warning.

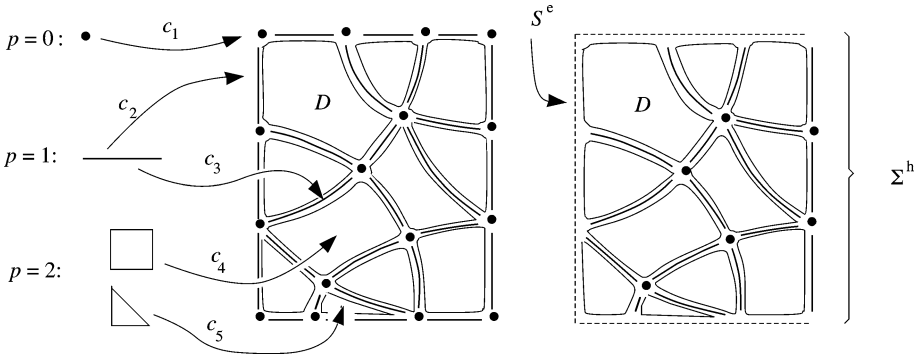

FIG. 14.1. Left: A few p-cells, contributing to a closed cellular paving of D. (This should be imagined in dimension 3.) Right: A culled paving, now "closed relative to" S^e. This is done in anticipation of the modelling we have in mind, in which cells of S^e would carry null degrees of freedom, so they won't be missed.

necessarily what is needed in practice, as one may rather wish to discard some cells in order to deal with boundary conditions. Hence the relevance of 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. The case we shall actually need, of a paving of $R = \overline{D} - S^e$ which is closed modulo S^e, is displayed on the right of Fig. 14.1. Informally said, "pave $\overline{D}$ first, then remove all cells from the electric boundary".

Each cell has its own inner orientation. These orientations are arbitrary and independent. In three dimensions, 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 these sets. (The general notation, rarely required, will be $\mathcal{S}_p$ for the set of p-cells and S_p for the number of such cells.)

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. As for the sign, recall that each cell orients its own boundary (Section 4), so this orientation may or may not coincide with the one attributed to σ. If orientations match, the sign is $+$, else it's $-$. Fig. 14.2 illustrates this point. (Also refer back to Fig. 4.1.)

Collecting these numbers in arrays, we obtain rectangular matrices **G**, **R**, **D**, called *incidence matrices* of the tessellation. For instance (Fig. 14.2), the incidence number for edge e and facet f is denoted $\mathbf{R}^e_f$, and makes one entry in matrix **R**, whose rows and columns are indexed over facets and edges, respectively. The entry $\mathbf{G}^n_e$ of **G** is -1 in the case displayed, because n, positively oriented, is at the start of edge e (cf. Fig. 3.4(c)). And so on. 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, will make cleaner notation in spatial

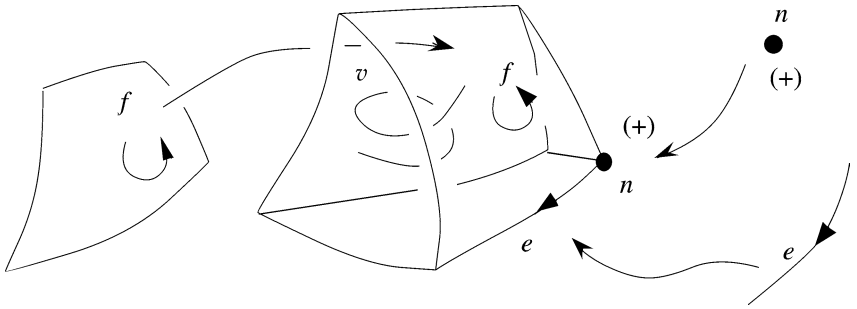

FIG. 14.2. Sides: Individual oriented cells. Middle: The same, plus a 3-cell, as part of a paving, showing respective orientations. Those of v and f match, those of f and e, or of e and n, don't. So $\mathbf{G}_e^n = -1$, $\mathbf{R}_f^e = -1$, and $\mathbf{D}_v^f = 1$.

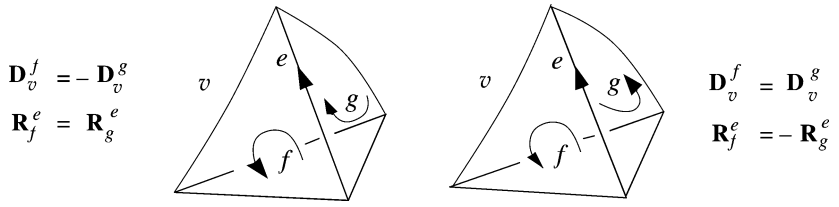

FIG. 14.3. Relation $\mathbf{DR} = 0$, and how it doesn't depend on the cells' individual orientations: In both cases, one has $\mathbf{D}_v^f \mathbf{R}_f^e + \mathbf{D}_v^g \mathbf{R}_g^e = 0$.

dimensions higher than 3, with $\mathbf{d}_0 = \mathbf{G}$, $\mathbf{d}_1 = \mathbf{R}$, $\mathbf{d}_2 = \mathbf{D}$. The mnemonic value of **G**, **R**, **D**, however, justifies keeping them in use.

Just as $\mathrm{rot} \circ \mathrm{grad} = 0$ and $\mathrm{div} \circ \mathrm{rot} = 0$, one has $\mathbf{RG} = 0$ and $\mathbf{DR} = 0$. Indeed, for an edge e and a volume v, the $\{v, e\}$-entry of $\mathbf{DR}$ is $\sum_{f \in \mathcal{F}} \mathbf{D}_v^f \mathbf{R}_f^e$. Nonzero terms occur, in this sum over facets, only for those which both 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 meeting along e (Fig. 14.3), and hence two nonzero terms. As Fig. 14.3 shows, they have opposite signs, whatever the orientations of the individual cells, hence the result, $\mathbf{DR} = 0$. By a similar argument, $\mathbf{RG} = 0$, and more generally, $\mathbf{d}_{p+1}\mathbf{d}_p = 0$.

REMARK 14.1. The answer to the natural question, "then, is the kernel of **R** equal to the range of **G**?", is "yes" here, because $\overline{D} - S^e$ has simple topology. (See the remark at the end of Section 4 about homology. This time, going further would lead us into *co*homology.) For the same reason, $\ker(\mathbf{D}) = \mathrm{cod}(\mathbf{R})$. This will be important below.

It is no accident if this proof of $\mathbf{d} \circ \mathbf{d} = 0$ evokes the one about $\partial \circ \partial = 0$ in Section 4, and the caption of Fig. 4.1. The same basic observation, "the boundary of a boundary is zero" (TAYLOR and WHEELER [1992], KHEYFETS and WHEELER [1986]), underlies all proofs of this kind. In fact, the above incidence matrices can be used to find the boundaries, chainwise, of each cell. For instance, f being understood as the 2-chain

based on facet f with weight 1, one has $\partial f = \sum_{e\in\mathcal{E}} \mathbf{R}^e_f e$. So if S is the straight 2-chain $\sum_f w^f\, f$ with weights w^f (which we shall call a *primal 2-chain*, or "m-surface", using m as a mnemonic for the underlying mesh), its boundary[37] is the 1-chain

$$\partial S = \sum_{e\in\mathcal{E}} \sum_{f\in\mathcal{F}} \mathbf{R}^e_f w^f\, e.$$

More generally, let's write $\boldsymbol{\partial}_p$, boldface,[38] for the transpose of the above matrix $\mathbf{d}_{p-1}$. Then, if $c = \sum_{\sigma\in\mathcal{S}_p} w^\sigma \sigma$ is a p-chain, its boundary is $\partial c = \sum\{s \in \mathcal{S}_{p-1}\colon\ (\boldsymbol{\partial}_p \mathbf{w})^s s\}$, where **w** stands for the vector of weights. Thus, $\boldsymbol{\partial}$ is to ∂ what **d** is to d. Moreover, the duality between d and ∂ is matched by a similar duality between their finite-dimensional counterparts **d** and $\boldsymbol{\partial}$.

15. Dual mesh

A *dual* mesh, with respect to m, 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 an $(n-p)$-cell, called the *dual* of c and denoted $\tilde{c}$, which meets c at a single point x_c. (Ways to build $\tilde{c}$ will soon be indicated.) Hence a one-to-one 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 the tangent spaces at x_c to c and $\tilde{c}$ are complementary, the inner orientation of c provides an outer orientation for $\tilde{c}$ (Fig. 15.1). Incidence matrices $\widetilde{\mathbf{G}}$, $\widetilde{\mathbf{R}}$, $\widetilde{\mathbf{D}}$ can then be defined, as above, the sign of each nonzero entry depending on whether outer orientations match or not.

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 $\mathbf{R}^e_f = 1$, 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 $\widetilde{\mathbf{R}}^{\tilde{f}}_{\tilde{e}}$ is equal to $\mathbf{R}^e_f$. Same equality if $\mathbf{R}^e_f = -1$, and same reasoning for

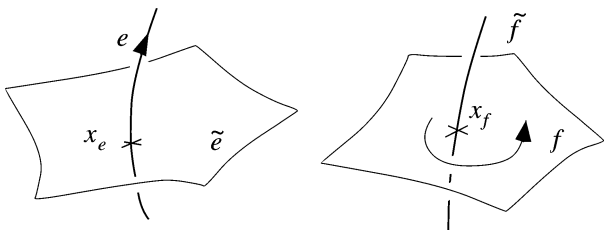

FIG. 15.1. Inner orientations of edge e and facet f, respectively, give crossing direction through $\tilde{e}$ and gyratory sense around $\tilde{f}$.

[37] More accurately, its boundary *relative to* Σ^h.

[38] Boldface, from now on, connotes mesh-related things, such as DoF arrays, etc.

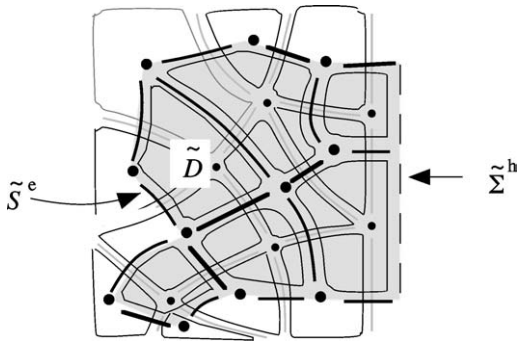

FIG. 15.2. A dual paving, overlaid on the primal one.

other kinds of cells, from which we conclude that the would-be dual incidence matrices $\widetilde{\mathbf{G}}$, $\widetilde{\mathbf{R}}$, $\widetilde{\mathbf{D}}$ 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 $\widetilde{R}$, the interior $\widetilde{D}$ of which is nearly D, but not quite (Fig. 15.2). A part of its boundary is paved by dual cells: We name it $\widetilde{S}^e$, owing to its kinship with S^e (not so obvious on our coarse drawing! but the finer the mesh, the closer $\widetilde{S}^e$ and S^e will get). The other part is denoted $\widetilde{\Sigma}^h$. So the cellular paving we now have is closed modulo $\widetilde{\Sigma}^h$, whereas the primal one was closed modulo S^e.

Given the mesh m, all its conceivable 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 two slightly more general procedures, the "star construction" and the "orthogonal construction" of meshes in duality. For this we shall consider only *polyhedral* meshes (those with polyhedral 3-cells), which is not overly restrictive in practice.

The orthogonal construction consists in having each dual cell orthogonal to its primal partner. (Cf. Figs. 15.3 and 15.5, left.) A particular case is the Voronoi–Delaunay tessellation (DIRICHLET [1850]), under the condition that dual nodes should be inside primal volumes. Alas, as Fig. 15.4 shows, orthogonality can be impossible to enforce, if the primal mesh is imposed. If one starts from a simplicial primal for which all circumscribed spheres have their center inside the tetrahedron, and all facets are acute triangles, all goes well. (One then takes these circumcenters as dual nodes.) But this property, desirable on many accounts, is not so easily obtained, and certainly not warranted by common mesh generators.

Hence the usefulness of the star construction, more general, because it applies to any primal mesh with star-shaped cells. A part A of A_n is *star-shaped* if it contains 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, pick such a center in each primal cell (the center of a primal node is itself), and join it to centers of all faces of the 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,

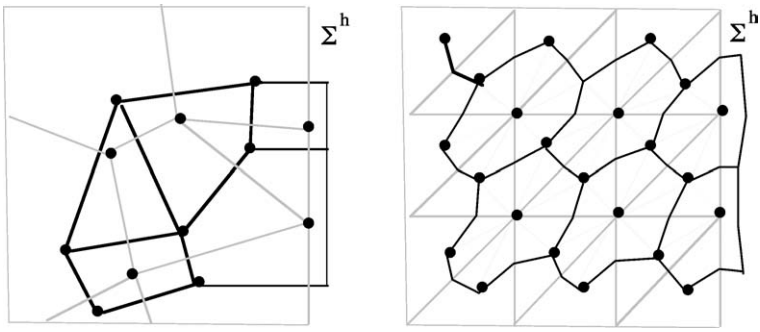

FIG. 15.3. Left: Orthogonal dual mesh. (Same graphic conventions as in Fig. 15.2, slightly simplified.) Right: Star construction of a dual mesh (close enough, here, to a barycentric mesh, but not quite the same). Notice the isolated dual edge, and the arbitrariness in shaping dual cells beyond Σ^h.

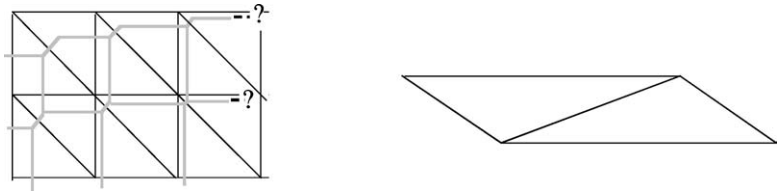

FIG. 15.4. Left: How hopeless the orthogonal construction can become, even with a fairly regular primal mesh. Right: Likely the simplest example of a 2D mesh without any orthogonal dual.

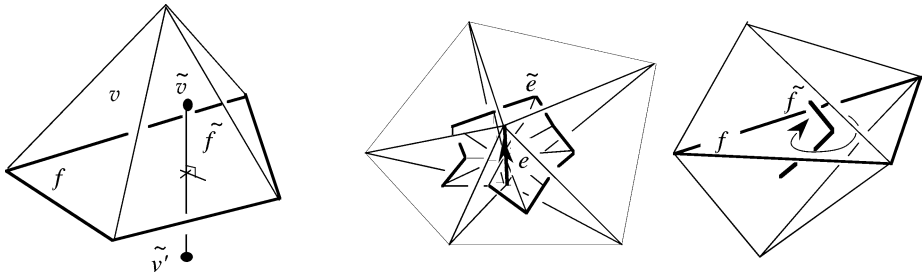

FIG. 15.5. Left: A facet f and its dual edge $\tilde{f}$ in the orthogonal construction ($\tilde{v}$ and $\tilde{v}'$ are the dual nodes which lie inside the volumes v and v' just above and just below f). From $\tilde{v}$, all boundary facets of v can directly be seen at right angle, but we don't require more: $\tilde{v}$ is neither v's barycenter nor the center of its circumscribed sphere, if there is such a sphere. Right: A dual facet and a dual edge, in the case of a simplicial primal mesh and of its barycentric dual. Observe the orientations.

$k \leqslant n - p$, which have one of their vertices at c's center, and other vertices at centers of cells incident on c. Figs. 15.3 and 15.5, right, give the idea. If all primal cells are simplices to start with, taking the barycenters of their faces as centers will give the *barycentric* dual mesh evoked a bit earlier.

REMARK 15.1. The recipe is imprecise about cells dual to those of Σ^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 centers of cells. It's all part of the unavoidable approximation error, which can be reduced at will by refinement.[39]

REMARK 15.2. If, as suggested above ("pave $\overline{D}$ first ..."), the primal mesh has been obtained by culling from a closed one, subcells built from the latter form a refinement of *both* the primal mesh and the dual mesh. The existence of this common "underlying simplicial complex" will be an asset when designing finite elements.

16. A discretization kit

We are ready, now, to apply the afore-mentioned strategy: Satisfy the balance equations (10.1) and (10.3) 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 inner 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 array of such "degrees of freedom" (DoF), $\{\mathbf{b} = \mathbf{b}_f\colon\ f \in \mathcal{F}\}$, 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, indirectly, the *flux* of b, embraced by the loop of a small enough magnetic probe, by reading off the induced e.m.f. 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. (Anyway, we'll soon see how to compute an approximation of the flux for any surface, knowing the $\mathbf{b}_f$'s, hence an approximation of b.) So one may be content with a method that would yield the four meaningful arrays of degrees of freedom, listing

- the edge e.m.f.'s, $\mathbf{e} = \{\mathbf{e}_e\colon\ e \in \mathcal{E}\}$,
- the facet fluxes, $\mathbf{b} = \{\mathbf{b}_f\colon\ f \in \mathcal{F}\}$,
- the dual-edge m.m.f.'s, $\mathbf{h} = \{\mathbf{h}_f\colon\ f \in \mathcal{F}\}$,
- and the dual-facet displacement currents, $\mathbf{d} = \{\mathbf{d}_e\colon\ e \in \mathcal{E}\}$,

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

In this respect, considering the integral form (10.1) and (10.3) of the basic equations will prove much easier than dealing with so-called "weak forms" of the infinitesimal equations (10.2) and (10.4). In fact, this simple shift of emphasis (which is the gist of Weiland's "finite integration theory", WEILAND [1992], and of Tonti's "cell method", TONTI [2001], MATTIUSSI [2000]) will so to speak *force on us* the right and unique discretization, as follows.

[39] A *refinement* of a paving is another paving of the same region, which restricts to a proper cellular paving of each original cell.

16.1. Network equations, discrete Hodge operator

Suppose the chain S in (10.1) is the simplest possible in the present context, that is, a *single* primal facet, f. The integral of e along ∂f is the sum of its integrals along edges that make ∂f, with proper signs, which are precisely the signs of the incidence numbers, by their very definition. Therefore, Eq. (10.1) applied to f yields

$$\partial_t \mathbf{b}_f + \sum_{e\in\mathcal{E}} \mathbf{R}^e_f \mathbf{e}_e = 0.$$

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. Taken together, in matrix form,

$$\partial_t \mathbf{b} + \mathbf{R}\mathbf{e} = 0, \tag{16.1a}$$

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

The same reasoning about each dual facet $\tilde{e}$ (the simplest possible outer-oriented surface that Σ in (10.3) can be) yields

$$-\partial_t \mathbf{d}_e + \sum_{f\in\mathcal{F}} \mathbf{R}^e_f \mathbf{h}_f = \mathbf{j}_e,$$

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

$$-\partial_t \mathbf{d} + \mathbf{R}^t \mathbf{h} = \mathbf{j}, \tag{16.1b}$$

the second group of network equations.

To complete this system, we need discrete counterparts to $b = \mu h$ and $d = \varepsilon e$, i.e., *network constitutive laws*, of the form

$$\mathbf{b} = \boldsymbol{\mu}\mathbf{h}, \qquad \mathbf{d} = \boldsymbol{\varepsilon}\mathbf{e}, \tag{16.2}$$

where $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$ are appropriate square symmetric matrices. Understanding how such matrices can be built is our next task. It should be clear that no *canonical* construction can exist – for sure, nothing comparable to the straightforward passage from (10.1), (10.3) to (16.1a), (16.1b) – because the metric of both meshes must intervene (Eq. (11.1) gives a clue in this respect). Indeed, the exact equivalent of (16.1), up to notational details, can be found in most published algorithms (including those based on the Galerkin method, see, e.g., LEE and SACKS [1995]), whereas a large variety of proposals exist as regards $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$. These "discrete Hodge operators" are the real issue. Constructing "good" ones, in a sense we still have to discover, is the central problem.

Our approach will be as follows: First – just not to leave the matter dangling too long – we shall give *one* solution, especially simple, to this problem, which makes $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$ *diagonal*, a feature the advantages of which we shall appreciate by working out a few examples. Later (in Section 20), a generic error analysis method will be sketched, from which a *criterion* as to what makes a good $\boldsymbol{\varepsilon}$–$\boldsymbol{\mu}$ pair will emerge. Finite elements will enter the stage during this process, and help find other solutions to the problem, conforming to the criterion.

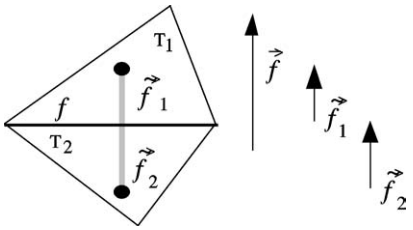

FIG. 16.1. The case of a discontinuous permeability (μ_1 and μ_2 in primal volumes T_1 and T_2, separated by facet f). We denote by $\vec{f}$ the vectorial area of f and by $\overrightarrow{f_1}$, $\overrightarrow{f_2}$, the vectors along both parts of $\tilde{f}$. Let u and v be arbitrary vectors, respectively normal and tangent to f, and let $\mathbf{H}_1 = u + v$ in T_1. Transmission conditions across f determine a unique uniform field $\mathrm{B}_2 = \mu_1 u + \mu_2 v$ in T_2. Then $\mathbf{b}_f = \mu_1 \vec{f} \cdot u$ and $\mu_2 \mathbf{h}_f = \mu_2 \overrightarrow{f_1} \cdot u + \mu_1 \overrightarrow{f_2} \cdot u$. As $\vec{f}$, $\overrightarrow{f_1}$, and $\overrightarrow{f_2}$ are collinear, u disappears from the quotient $\mathbf{b}_f / \mathbf{h}_f$, yielding (16.4).

The simple solution is available if one has been successful in building a dual mesh by the orthogonal construction (Figs. 15.3 and 15.5, left). Then, in the case when ε and μ are uniform,[40] one sets $\boldsymbol{\varepsilon}^{ee'} = 0$ if $e \neq e'$, $\boldsymbol{\mu}^{ff'} = 0$ if $f \neq f'$, and (cf. (11.1))

$$\boldsymbol{\varepsilon}^{ee} = \varepsilon \frac{\text{area}(\tilde{e})}{\text{length}(e)}, \qquad \boldsymbol{\mu}^{ff} = \mu \frac{\text{area}(f)}{\text{length}(\tilde{f})}, \tag{16.3}$$

which does provide diagonal matrices $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$. (The inverse of $\boldsymbol{\mu}$ will be denoted by $\boldsymbol{\nu}$.) The heuristic justification (TONTI [2001]) is that *if* the various fields happened to be piecewise constant (relative to the primal mesh), formulas (16.3) would exactly correspond to the very definition (11.1) of the Hodge operator. (Section 20 will present a stronger argument.) In the case of non-uniform coefficients, formulas such as

$$\boldsymbol{\mu}^{ff} = \frac{\mu_1 \mu_2 \ \text{area}(f)}{\mu_2 \,\text{length}(\tilde{f}_1) + \mu_1 \,\text{length}(\tilde{f}_2)}, \tag{16.4}$$

where $\tilde{f}_1$ and $\tilde{f}_2$ are the parts of $\tilde{f}$ belonging to the two volumes adjacent to f, apply instead (Fig. 16.1). Observe the obvious intervention of metric elements (lengths, areas, angles) in these constructions.

REMARK 16.1. Later, when edge elements w^e and facet elements w^f will enrich the toolkit, we shall consider another solution, that consists in setting $\boldsymbol{\varepsilon}^{ee'} = \int_D \varepsilon w^e \wedge w^{e'}$ and $\boldsymbol{\nu}^{ff'} = \int_D \mu^{-1} w^f \wedge w^{f'}$. For reference, let's call this the "Galerkin approach" to the problem. We shall use loose expressions such as "the Galerkin $\boldsymbol{\varepsilon}$", or "the diagonal hodge", to refer to various brands of discrete Hodge operators.

16.2. The toolkit

At this stage, we have obtained discrete counterparts (Fig. 16.2) to most features of the "Maxwell building" of Fig. 12.2, but time differentiation and wedge product still miss

[40] I'll use "uniform" and "steady" for "constant in space" and "constant in time", respectively.

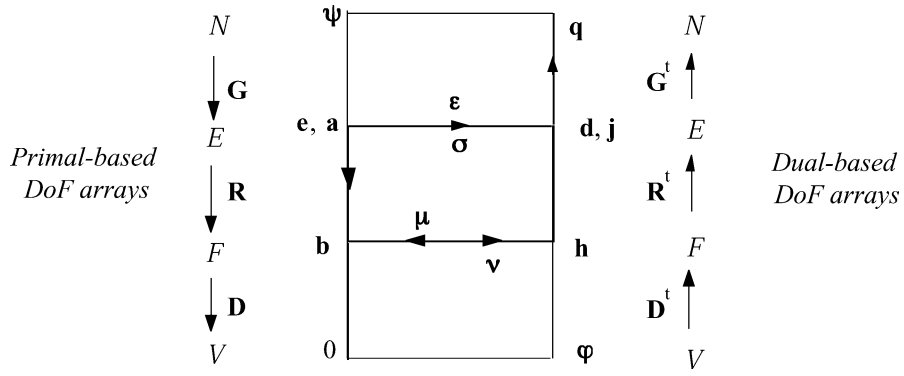

FIG. 16.2. A "discretization toolkit" for Maxwell's equations.

theirs. Some thought about how the previous ideas would apply in four dimensions should quickly suggest the way to deal with time derivatives: δt being the time step, call $\mathbf{b}^k$, $\mathbf{h}^k$, the values of $\mathbf{b}$, $\mathbf{h}$ at time $k\delta t$, for $k = 0, 1, \ldots$, call $\mathbf{j}^{k+1/2}$, $\mathbf{d}^{k+1/2}$, $\mathbf{e}^{k+1/2}$ those of $\mathbf{j}$, $\mathbf{d}$, $\mathbf{e}$ at time $(k + 1/2)\delta t$, and approximate $\partial_t \mathbf{b}$, at time $(k + 1/2)\delta t$, by $(\mathbf{b}^{k+1} - \mathbf{b}^k)/\delta t$, and similarly, $\partial_t \mathbf{d}$, now at time $k\delta t$, by $(\mathbf{d}^{k+1/2} - \mathbf{d}^{k-1/2})/\delta t$.

As for the wedge product, to $\int_D b \wedge h$ corresponds the sum $\sum_{f\in\mathcal{F}} \mathbf{b}_f \mathbf{h}_f$, which we shall denote by $(\mathbf{b}, \mathbf{h})$, with bold parentheses. Similarly, $\int_D d \wedge e$ corresponds to $\sum_{e\in\mathcal{E}} \mathbf{d}_e \mathbf{e}_e$, also denoted $(\mathbf{d}, \mathbf{e})$. Hence we may define "discrete energy" quadratic forms, $1/2(\boldsymbol{\nu}\mathbf{b}, \mathbf{b})$, $1/2(\boldsymbol{\mu}\mathbf{h}, \mathbf{h})$, $1/2(\boldsymbol{\varepsilon}\mathbf{e}, \mathbf{e})$, and $1/2(\boldsymbol{\varepsilon}^{-1}\mathbf{d}, \mathbf{d})$, all quantities with, indeed, the physical dimension of energy (but be aware that $(\mathbf{j}, \mathbf{e})$ is a power instead, like $\int_D j \wedge e$). Some notational shortcuts: Square roots such as $(\boldsymbol{\nu}\mathbf{b}, \mathbf{b})^{1/2}$, or $(\boldsymbol{\varepsilon}\mathbf{e}, \mathbf{e})^{1/2}$, etc., will be denoted by $|\mathbf{b}|_\nu$, or $|\mathbf{e}|_\varepsilon$, in analogy with the above $|b|_\nu$, or $|e|_\varepsilon$, and serve as various, physically meaningful *norms* on the vector spaces of DoF arrays. We'll say the "ν-norm", the "ε-norm", etc., for brevity.

PROPOSITION 16.1. *If Eqs.* (16.1)–(16.2) *are satisfied, one has*

$$d_t\left[\frac{1}{2}(\boldsymbol{\nu}\mathbf{b}, \mathbf{b}) + \frac{1}{2}(\boldsymbol{\varepsilon}\mathbf{e}, \mathbf{e})\right] = -(\mathbf{j}, \mathbf{e}). \tag{16.5}$$

PROOF. Take the bold scalar product of (16.1a) and (16.1b) by $\mathbf{h}$ and $-\mathbf{e}$, add, and use the equality $(\mathbf{Re}, \mathbf{h}) = (\mathbf{e}, \mathbf{R}^t\mathbf{h})$. □

REMARK 16.2. The analogue of $\int_S h \wedge e$, when S is some m-surface, is

$$\sum_{f\in\mathcal{F}(S),\, e\in\mathcal{E}} \mathbf{R}^e_f \mathbf{h}_f \mathbf{e}_e,$$

where $\mathcal{F}(S)$ stands for the subset of facets which compose S. (Note how this sum vanishes if S is the domain's boundary.) By exploiting this, the reader will easily modify (16.5) in analogy with the Poynting theorem. In spite of such formal correspondences,

energy and discrete energy have, a priori, no relation. To establish one, we shall need "interpolants", such as finite elements, enabling us to pass from degrees of freedoms to fields. For instance, facet elements will generate a mapping $\mathbf{b} \to b$, with $b = \sum_f \mathbf{b}_f w^f$. If $\boldsymbol{\nu}$ is the Galerkin hodge, then $\int_D \nu b \wedge b = (\boldsymbol{\nu}\mathbf{b}, \mathbf{b})$. Such built-in equality between energy and discrete energy is an exception, a distinctive feature of the Ritz–Galerkin approach. With other discrete hodges, even *convergence* of discrete energy, as the mesh is refined, towards the true one, should not be expected.

17. Playing with the kit: Full Maxwell

Now we have enough to discretize any model connected with Maxwell's equations. Replacing, in (13.1), rot by $\mathbf{R}$ or $\mathbf{R}^t$, ε and μ by $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$, and ∂_t by the integral or half-integral differential quotient, depending on the straight or twisted nature of the differential form in consideration, we obtain this:

$$\frac{\mathbf{b}^{k+1} - \mathbf{b}^k}{\delta t} + \mathbf{R}\mathbf{e}^{k+1/2} = 0, \qquad -\boldsymbol{\varepsilon}\frac{\mathbf{e}^{k+1/2} - \mathbf{e}^{k-1/2}}{\delta t} + \mathbf{R}^t \boldsymbol{\nu}\mathbf{b}^k = \mathbf{j}^k \tag{17.1}$$

(where $\mathbf{j}^k$ is the array of intensities through dual facets, at time[41] $k\delta t$), with initial conditions

$$\mathbf{b}^0 = 0, \qquad \mathbf{e}^{-1/2} = 0. \tag{17.2}$$

In the simplest case where the primal and dual mesh are plain rectangular staggered grids, (17.1) and (17.2) is the well known Yee scheme (YEE [1966]). So what we have here is the closest thing to Yee's scheme in the case of *cellular* meshes.

A similar numerical behavior can therefore be expected. Indeed,

PROPOSITION 17.1. *The scheme* (17.1) *and* (17.2) *is stable for δt small enough, provided both $\boldsymbol{\varepsilon}$ and $\boldsymbol{\nu}$ are symmetric positive definite.*

PROOF. For such a proof, one may assume $\mathbf{j} = 0$ and nonzero initial values in (17.2), satisfying $\mathbf{D}\mathbf{b}^0 = 0$. Eliminating $\mathbf{e}$ from (17.1), one finds that

$$\mathbf{b}^{k+1} - 2\mathbf{b}^k + \mathbf{b}^{k-1} + (\delta t)^2 \mathbf{R}\boldsymbol{\varepsilon}^{-1}\mathbf{R}^t \boldsymbol{\nu}\mathbf{b}^k = 0. \tag{17.3}$$

Since $\mathbf{D}\mathbf{R} = 0$, the "loop invariant" $\mathbf{D}\mathbf{b}^k = 0$ holds, so one may work in the corresponding subspace, $\ker(\mathbf{D})$. Let's introduce the (generalized) eigenvectors $\mathbf{v}_i$ such that $\mathbf{R}\boldsymbol{\varepsilon}^{-1}\mathbf{R}^t\mathbf{v}_i = \lambda_i \boldsymbol{\mu}\mathbf{v}_i$, which satisfy $(\boldsymbol{\mu}\mathbf{v}_i, \mathbf{v}_j) = 1$ if $i = j$, 0 if $i \neq j$. In this "μ-orthogonal" basis, $\mathbf{b}^k = \boldsymbol{\mu}\Sigma_i \eta_i^k \mathbf{v}_i$, and (17.3) becomes

$$\eta_i^{k+1} - \big(2 - \lambda_i(\delta t)^2\big)\eta_i^k + \eta_i^{k-1} = 0$$

for all i. The η_i^ks, and hence the $\mathbf{b}^k$s, stay bounded if the characteristic equation of each of these recurrences has imaginary roots, which happens (Fig. 17.1) if $0 < \lambda_j \delta t < 2$ for all j. □

[41] For easier handling of Ohm's law, $\mathbf{j}(k\delta t)$ may be replaced by $(\mathbf{j}^{k+1/2} + \mathbf{j}^{k-1/2})/2$.

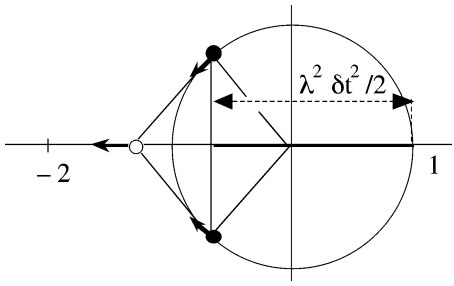

FIG. 17.1. The white spot lies at the sum of roots of the characteristic equation $r^2 - (2 - \lambda_i(\delta t)^2)r + 1 = 0$. Stability is lost if it leaves the interval $[-2, 2]$.

In the case of the original Yee scheme, eigenvalues could explicitly be found, hence the well-known relation (YEE [1966]) between the maximum possible value of δt and the lengths of the cell sides. For general grids, we have no explicit formulas, but the thumbrule is the same: δt should be small enough for a signal travelling at the speed of light (in the medium under study) not to cross more than one cell during this lapse of time.

This stringent stability condition makes the scheme unattractive if not fully explicit, or nearly so: $\boldsymbol{\varepsilon}$ should be *diagonal*, or at the very least, block-diagonal with most blocks of size 1 and a few small-size ones, and $\boldsymbol{\nu}$ should be sparse. If so is the case, each time step will only consist in a few matrix–vector products plus, perhaps, the resolution of a few small linear systems, which makes up for the large number of time steps. Both conditions are trivially satisfied with the orthogonal construction (cf. (16.3), (16.4)), but we have already noticed the problems this raises. Hence the sustained interest for so-called "mass-lumping" procedures, which aim at replacing the Galerkin $\boldsymbol{\varepsilon}$ by a diagonal matrix without compromising convergence: see COHEN, JOLY and TORDJMAN [1993], ELMKIES and JOLY [1997], HAUGAZEAU and LACOSTE [1993] (a coordinate-free reinterpretation of which can be found in BOSSAVIT and KETTUNEN [1999]).

REMARK 17.1. Obviously, there is another version of the scheme, in **h** and **d**, for which what is relevant is sparsity of $\boldsymbol{\varepsilon}^{-1}$ and diagonality of $\boldsymbol{\mu}$, i.e., of $\boldsymbol{\nu}$. Unfortunately, the diagonal lumping procedure that worked for edge elements fails when applied to the Galerkin $\boldsymbol{\nu}$, i.e., to the mass-matrix of facet elements (BOSSAVIT and KETTUNEN [1999]).

There are of course other issues than stability to consider, but we shall not dwell on them right now. For *convergence* (to be treated in detail later, but only in statics), cf. MONK and SÜLI [1994], NICOLAIDES and WANG [1998], BOSSAVIT and KETTUNEN [1999]. On *dispersion* properties, little can be said unless the meshes have some translational symmetry, at least locally, and this is beyond our scope. As for *conservation* of some quantities, it would be nice to be able to say, in the case when $\mathbf{j} = 0$, that "total discrete energy is conserved", but this is only almost true. Conserved quantities, as one will easily verify, are $\frac{1}{2}(\boldsymbol{\mu}\mathbf{h}^{k+1}, \mathbf{h}^k) + \frac{1}{2}(\boldsymbol{\varepsilon}\mathbf{e}^{k+1/2}, \mathbf{e}^{k+1/2})$ and

$\frac{1}{2}(\boldsymbol{\mu}\mathbf{h}^k, \mathbf{h}^k) + \frac{1}{2}(\boldsymbol{\varepsilon}\mathbf{e}^{k-1/2}, \mathbf{e}^{k+1/2})$, both independent of k. So their half-sum, which can suggestively be written as

$$W_k = \tfrac{1}{2}(\boldsymbol{\mu}\mathbf{h}^{k+1/2}, \mathbf{h}^k) + \tfrac{1}{2}(\boldsymbol{\varepsilon}\mathbf{e}^k, \mathbf{e}^{k+1/2}),$$

if one agrees on $\mathbf{h}^{k+1/2}$ and $\mathbf{e}^k$ as shorthands for $[\mathbf{h}^k + \mathbf{h}^{k+1}]/2$ and $[\mathbf{e}^{k-1/2} + \mathbf{e}^{k+1/2}]/2$, is conserved: *Not* the discrete energy, definitely, however close.

18. Playing with the kit: Statics

Various discrete models can be derived from (17.1) by the usual maneuvers (neglect the displacement current term $\boldsymbol{\varepsilon}\mathbf{e}$, omit time-derivatives in static situations), but it may be more instructive to obtain them from scratch. Take the magnetostatic model (13.2), for instance: Replace forms b and h by the DoF arrays $\mathbf{b}$ and $\mathbf{h}$, the d by the appropriate matrix, as read off from Fig. 16.2, and obtain

$$\mathbf{D}\mathbf{b} = 0, \qquad \mathbf{h} = \boldsymbol{\nu}\mathbf{b}, \qquad \mathbf{R}^t\mathbf{h} = \mathbf{j}, \tag{18.1}$$

which automatically includes the boundary conditions, thanks for having discarded[42] "passive" boundary cells. Observe that $\mathbf{G}^t\mathbf{j} = 0$ must hold for a solution to exist: But this is the discrete counterpart, as Fig. 16.2 shows, of $\mathrm{d}j = 0$, i.e., of $\operatorname{div}\mathrm{J} = 0$ in vector notation.

In the next section, we shall study the convergence of (18.1). When it holds, all schemes equivalent to (18.1) that can be obtained by algebraic manipulations are thereby equally valid – and there are lots of them. First, let $\mathbf{h}^{\mathbf{j}}$ be one of the facet-based arrays[43] such that $\mathbf{R}^t\mathbf{h}^{\mathbf{j}} = \mathbf{j}$. Then $\mathbf{h}$ in (18.1) must be of the form $\mathbf{h} = \mathbf{h}^{\mathbf{j}} + \mathbf{D}^t\boldsymbol{\varphi}$. Hence (18.1) becomes

$$\mathbf{D}\boldsymbol{\mu}\mathbf{D}^t\boldsymbol{\varphi} = -\mathbf{D}\boldsymbol{\mu}\mathbf{h}^{\mathbf{j}}. \tag{18.2}$$

This, which corresponds to $-\operatorname{div}(\mu(\operatorname{grad}\Phi + \mathrm{H}^j)) = 0$, the scalar potential formulation of magnetostatics, is not interesting unless $\boldsymbol{\nu}$ is diagonal, or nearly so, since $\boldsymbol{\mu}$ is full otherwise. So it requires the orthogonal construction, and is not an option in the case of the Galerkin $\boldsymbol{\nu}$. It's a well-studied scheme (cf. BANK and ROSE [1987], COURBET and CROISILLE [1998], GALLOUET and VILA [1991], HEINRICH [1987], HUANG and XI [1998], SÜLI [1991]), called "block-centered" in other sectors of numerical engineering (KAASSCHIETER and HUIJBEN [1992], WEISER and WHEELER [1988]), because degrees of freedom, assigned to the *dual* nodes, appear as lying inside the primal volumes,

[42] Alternatively (and this is how non-homogeneous boundary conditions can be handled), one may work with enlarged incidence matrices $\mathbf{R}$ and $\mathbf{D}$ and enlarged DoF arrays, taking all cells into account, then assign boundary values to passive cells, and keep only active DoFs on the left-hand side.

[43] There are such arrays, owing to $\mathbf{G}^t\mathbf{j} = 0$, because $\ker(\mathbf{G}^t) = \operatorname{cod}(\mathbf{R}^t)$, by transposition of $\operatorname{cod}(\mathbf{G}) = \ker(\mathbf{R})$, in the simple situation we consider. Finding one is an easy task, which does not require solving a linear system. Also by transposition of $\operatorname{cod}(\mathbf{R}) = \ker(\mathbf{D})$, one has $\ker(\mathbf{R}^t) = \operatorname{cod}(\mathbf{D}^t)$, and hence $\mathbf{R}^t(\mathbf{h} - \mathbf{h}^{\mathbf{j}}) = 0$ implies $\mathbf{h} = \mathbf{h}^{\mathbf{j}} + \mathbf{D}^t\boldsymbol{\varphi}$.

or "blocks". Uniqueness of $\boldsymbol{\varphi}$ is easily proved,[44] which implies the uniqueness – not so obvious, a priori – of $\mathbf{h}$ and $\mathbf{b}$ in (18.1).

Symmetrically, there is a scheme corresponding to the vector potential formulation (i.e., $\mathrm{rot}(\nu\,\mathrm{rot}\,\mathrm{A}) = \mathrm{J}$):

$$\mathbf{R}^t\boldsymbol{\nu}\mathbf{R}\mathbf{a} = \mathbf{j}, \tag{18.3}$$

obtained by setting $\mathbf{b} = \mathbf{R}\mathbf{a}$, where the DoF array $\mathbf{a}$ is indexed over (active) edges. (If $\boldsymbol{\nu}$ is the Galerkin hodge, (18.3) is what one obtains when using edge elements to represent the vector potential.) Existence in (18.3) stems from $\mathbf{G}^t\mathbf{j} = 0$. No uniqueness this time, because $\ker(\mathbf{R})$ does not reduce to 0, but all solutions $\mathbf{a}$ give the same $\mathbf{b}$, and hence the same $\mathbf{h} = \boldsymbol{\nu}\mathbf{b}$.

REMARK 18.1. Whether to "gauge" $\mathbf{a}$ in this method, that is, to impose a condition that would select a unique solution, such as $\mathbf{G}^t\boldsymbol{\varepsilon}\mathbf{a} = 0$ for instance, remains to these days a contentious issue. It depends on which method is used to solve (18.3), and on how well the necessary condition $\mathbf{G}^t\mathbf{j} = 0$ is implemented. With iterative methods such as the conjugate gradient and its variants, and if one takes care to use $\mathbf{R}^t\mathbf{h}^{\mathbf{j}}$ instead of $\mathbf{j}$ in (18.3), then it's better *not* to gauge (REN [1996]).

This is not all. If we refrain to eliminate $\mathbf{h}$ in the reduction from (18.1) to (18.3), but still set $\mathbf{b} = \mathbf{R}\mathbf{a}$, we get an intermediate two-equation system,

$$\begin{pmatrix} -\boldsymbol{\mu} & \mathbf{R} \\ \mathbf{R}^t & 0 \end{pmatrix}\begin{pmatrix} \mathbf{h} \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{j} \end{pmatrix}, \tag{18.4}$$

often called a *mixed* algebraic system (ARNOLD and BREZZI [1985]). (Again, little interest if $\boldsymbol{\mu}$ is full, i.e., unless $\boldsymbol{\nu}$ was diagonal from the outset.) The same manipulation in the other direction (eliminating $\mathbf{h}$ by $\mathbf{h} = \mathbf{h}^{\mathbf{j}} + \mathbf{D}^t\boldsymbol{\varphi}$, but keeping $\mathbf{b}$) gives

$$\begin{pmatrix} -\boldsymbol{\nu} & \mathbf{D}^t \\ \mathbf{D} & 0 \end{pmatrix}\begin{pmatrix} \mathbf{b} \\ \boldsymbol{\varphi} \end{pmatrix} = \begin{pmatrix} -\mathbf{h}^{\mathbf{j}} \\ 0 \end{pmatrix}. \tag{18.5}$$

We are not yet through. There is an interesting variation on (18.5), known as the mixed-hybrid approach. It's a kind of "maximal domain decomposition", in the sense that all volumes are made independent by "doubling" the degrees of freedom of $\mathbf{b}$ and $\mathbf{h}$ (two distinct values on sides of each facet not in Σ^h). Let's redefine the enlarged arrays and matrices accordingly, and call them $\overline{\mathbf{b}}$, $\overline{\mathbf{h}}$, $\overline{\nu}$, $\overline{\mathbf{D}}$, $\overline{\mathbf{R}}$. Constraints on $\overline{\mathbf{b}}$ (equality of up- and downstream fluxes) can be expressed as $\mathbf{N}\overline{\mathbf{b}} = 0$, where $\mathbf{N}$ has very simple structure (one 1×2 block, with entries 1 and -1, for each facet). Now, introduce an array $\boldsymbol{\lambda}$ of facet-based Lagrange multipliers, and add $(\boldsymbol{\lambda}, \mathbf{N}\overline{\mathbf{b}})$ to the underlying Lagrangian of (18.5). This gives a new discrete formulation (still equivalent to (18.1), if one derives $\mathbf{b}$

[44] It stems from $\ker(\mathbf{D}^t) = 0$. Indeed, $\mathbf{D}^t\boldsymbol{\psi} = 0$ means that $\Sigma_v \mathbf{D}_v^f \boldsymbol{\psi}_v = 0$ for all primal facets f. For some facets (those in Σ^h), there is but *one* volume v such that $\mathbf{D}_v^f \neq 0$, which forces $\boldsymbol{\psi}_v = 0$ for this v. Remove all such volumes v, and repeat the reasoning and the process, thus spreading the value 0 to all $\boldsymbol{\psi}_v$s.

and $\mathbf{h}$ from $\overline{\mathbf{b}}$ and $\overline{\mathbf{h}}$ the obvious way):

$$\begin{pmatrix} -\overline{\nu} & \overline{\mathbf{D}}^t & \mathbf{N}^t \\ \overline{\mathbf{D}} & 0 & 0 \\ \mathbf{N} & 0 & 0 \end{pmatrix} \begin{pmatrix} \overline{\mathbf{b}} \\ \boldsymbol{\varphi} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} -\overline{\mathbf{h}}^{\mathbf{j}} \\ 0 \\ 0 \end{pmatrix}.$$

Remark that the enlarged $\overline{\nu}$ is block-diagonal (as well as its inverse $\overline{\mu}$), hence easy elimination of $\overline{\mathbf{b}}$. What then remains is a symmetric system in $\boldsymbol{\varphi}$ and $\boldsymbol{\lambda}$:

$$\begin{pmatrix} \overline{\mathbf{D}}\,\overline{\mu}\,\overline{\mathbf{D}}^t & \overline{\mathbf{D}}\,\overline{\mu}\,\mathbf{N}^t \\ \mathbf{N}\,\overline{\mu}\,\overline{\mathbf{D}}^t & \mathbf{N}\,\overline{\mu}\,\mathbf{N}^t \end{pmatrix} \begin{pmatrix} \boldsymbol{\varphi} \\ \boldsymbol{\lambda} \end{pmatrix} = -\begin{pmatrix} \overline{\mathbf{D}}\,\overline{\mu}\,\overline{\mathbf{h}}^{\mathbf{j}} \\ \mathbf{N}\,\overline{\mu}\,\overline{\mathbf{h}}^{\mathbf{j}} \end{pmatrix}.$$

The point of this manipulation is that $\overline{\mathbf{D}}\overline{\mu}\overline{\mathbf{D}}^t$ is *diagonal*, equal to $\mathbf{K}$, say. So we may again eliminate $\boldsymbol{\varphi}$, which leads to a system in terms of only $\boldsymbol{\lambda}$:

$$\mathbf{N}[\overline{\mu} - \overline{\mu}\,\overline{\mathbf{D}}^t\mathbf{K}^{-1}\overline{\mathbf{D}}\,\overline{\mu}]\mathbf{N}^t\boldsymbol{\lambda} = \mathbf{N}[\overline{\mu}\,\overline{\mathbf{D}}^t\mathbf{K}^{-1}\overline{\mathbf{D}}\,\overline{\mu} - \overline{\mu}]\overline{\mathbf{h}}^{\mathbf{j}}. \tag{18.6}$$

Contrived as it may look, (18.6) is a quite manageable system, with a sparse symmetric matrix. (The bracketed term on the left is block-diagonal, like $\overline{\mu}$.)

REMARK 18.2. In $(\boldsymbol{\lambda}, \mathbf{N}\overline{\mathbf{b}})$, each $\boldsymbol{\lambda}_f$ multiplies a term $(\mathbf{N}\overline{\mathbf{b}})_f$ which is akin to a magnetic charge. Hence the $\boldsymbol{\lambda}_f$s should be interpreted as facet-DoFs of a magnetic potential, which assumes the values necessary to reestablish the equality between fluxes that has been provisionally abandoned when passing from $\mathbf{b}$ to the enlarged (double size) flux vector $\overline{\mathbf{b}}$. This suggests a way to "complementarity" (obtaining bilateral estimates of some quantities) which is explored in BOSSAVIT [2003].

There is a dual mixed-hybrid approach, starting from (18.4), where *dual* volumes are made independent, hence (in the case of a simplicial primal mesh) three DoFs per facet, for both $\mathbf{b}$ and $\mathbf{h}$, and two Lagrange multipliers to enforce their equality. This leads to a system similar to (18.6) – but with twice as many unknowns, which doesn't make it attractive.

Systems (18.2), (18.3), (18.4), (18.5) and (18.6) all give the same solution pair $\{\mathbf{b}, \mathbf{h}\}$ – the solution of (18.1). Which one effectively to solve, therefore, is uniquely a matter of algorithmics, in which size, sparsity, and effective conditioning should be considered. The serious contenders are the one-matrix semi-definite systems, i.e., (18.2), (18.3), and (18.6). An enumeration of the number of off-diagonal terms (which is a fair figure of merit when using conjugate gradient methods on such matrices), shows that (18.6) rates better than (18.3), as a rule. The block-centered scheme (18.2) outperforms both (18.3) and (18.6), but is not available[45] with the Galerkin hodge. Hence the enduring interest (CHAVENT and ROBERTS [1991], KAASSCHIETER and HUIJBEN [1992], MOSÉ, SIEGEL, ACKERER and CHAVENT [1994], HAMOUDA, BANDELIER and RIOUX-DAMIDAU [2001]) for the "mixed-hybrid" method (18.6).

Each of the above schemes could be presented as the independent discretization of a specific mixed or mixed-hybrid variational formulation, and the literature is replete

[45] Unless one messes up with the computation of the terms of the mass-matrix, by using ad-hoc approximate integration formulas. This is precisely one of the devices used in mass-lumping.

with sophisticated analyses of this kind. Let's reemphasize that all these schemes are *algebraically* equivalent, as regards **b** and **h**. Therefore, an error analysis of one of them applies to all: For instance, if $\boldsymbol{\nu}$ is the Galerkin hodge, the standard variational convergence proof for (18.3), or if $\boldsymbol{\mu}$ is the diagonal hodge of (16.4), the error analysis we shall perform next section, on the symmetrical system (18.1).

19. Playing with the kit: Miscellanies

The advantage of working at the discrete level from the outset is confirmed by most examples one may tackle. For instance, the discrete version of the eddy-current problem (13.4) is, without much ado, found to be

$$i\omega\boldsymbol{\sigma}\mathbf{E} + \mathbf{R}^t\boldsymbol{\nu}\mathbf{R}\mathbf{E} = -i\omega\mathbf{J}^s. \tag{19.1}$$

As a rule, σ vanishes outside of a closed region $C = D - \Delta$ of the domain, C for "conductor". (Assume, then, that A, which is $\mathrm{supp}(J^s)$, is contained in Δ.) The system matrix then has a non-trivial null space, $\ker(\boldsymbol{\sigma}) \cap \ker(\mathbf{R})$, and uniqueness of $\mathbf{E}$ is lost. It can be restored by enforcing the constraint $\mathbf{G}^t\boldsymbol{\varepsilon}_\Delta\mathbf{E} = 0$, where $\boldsymbol{\varepsilon}_\Delta$ is derived from $\boldsymbol{\varepsilon}$ by setting to zero all rows and columns which correspond to edges borne by C. Physically, this amounts to assume a zero electric charge density outside the conductive region $C = \mathrm{supp}(\sigma)$. (Beware, the electric field obtained this way can be seriously wrong about A, where this assumption is not warranted, in general. However, the electric field in C is correct.) Mathematically, the effect is to limit the span of the unknown $\mathbf{E}$ to a subspace over which $i\omega\boldsymbol{\sigma} + \mathbf{R}^t\boldsymbol{\nu}\mathbf{R}$ is regular.

In some applications, however, the conductivity is nonzero in all D, but may assume values of highly different magnitudes, and the above matrix, though regular, is ill-conditioned. One then will find in the kit the right tools to "regularize" such a "stiff" problem. See CLEMENS and WEILAND [1999] for an example of the procedure, some aspects of which are studied in BOSSAVIT [2001a]. Briefly, it consists in adding to the left-hand side of (19.1) a term, function of $\mathbf{E}$, that vanishes when $\mathbf{E}$ is one of the solutions of (19.1), which supplements the $\mathbf{R}^t\boldsymbol{\nu}\mathbf{R}$ matrix by, so to speak, what it takes to make it regular (and hence, to make the whole system matrix well conditioned, however small σ can be at places). The modified system is

$$i\omega\boldsymbol{\sigma}\mathbf{E} + \mathbf{R}^t\boldsymbol{\nu}\mathbf{R}\mathbf{E} + \boldsymbol{\sigma}\mathbf{G}\boldsymbol{\delta}\mathbf{G}^t\boldsymbol{\sigma}\mathbf{E} = -i\omega\mathbf{J}^s, \tag{19.2}$$

where $\boldsymbol{\delta}$ is a Hodge-like matrix, node based, diagonal, whose entries are $\boldsymbol{\delta}^{nn} = \int_{\tilde{n}} 1/\mu\sigma^2$. A rationale for this can be found in BOSSAVIT [2001a]: In a nutshell, the idea is to "load the null space" of $\mathbf{R}^t\boldsymbol{\nu}\mathbf{R}$, and dimensional considerations motivate the above choice of $\boldsymbol{\delta}$. Our sole purpose here is to insist that all this can be done at the discrete level.

REMARK 19.1. One *might* motivate this procedure by starting from the following equation, here derived from (19.2) by simply using the toolkit in the other direction ("discrete" to "continuous"):

$$i\omega\sigma\,\mathrm{E} + \mathrm{rot}(\nu\,\mathrm{rot}\,\mathrm{E}) - \sigma\,\mathrm{grad}\left(\frac{1}{\mu\sigma^2}\,\mathrm{div}(\sigma\,\mathrm{E})\right) = -i\omega\mathrm{J}^s, \tag{19.3}$$

but which can be seen as a natural regularization of (13.4). (We revert to vector proxies here to call attention on the use of a variant of the $-\Delta = \mathrm{rot} \circ \mathrm{rot} - \mathrm{grad} \circ \mathrm{div}$ formula, which is relevant when both μ and σ are uniform in (19.3).) This is a time-honored idea (LEIS [1968]). Part of its present popularity may stem from its allowing standard discretization via *node-based* vector-valued elements (the discrete form is then of course quite different[46] from (19.2)), because E in (19.3) has more a priori regularity than E in (13.4). Even if one has reasons to prefer using such elements, the advantage is only apparent, because the discrete solution may converge towards something else than the solution of (13.4) in some cases (e.g., reentrant corners, cf. COSTABEL and DAUGE [1997]), where the solution of (19.3) has *too much* regularity to satisfy (13.4). This should make one wary of this approach.

Many consider the nullspace of $\mathbf{R}^t \boldsymbol{\nu} \mathbf{R}$ as a matter of concern, too, as regards the eigenmode problem,

$$\mathbf{R}^t \boldsymbol{\nu} \mathbf{R} \mathbf{E} = \omega^2 \boldsymbol{\varepsilon} \mathbf{E}, \tag{19.4}$$

because $\omega = 0$ is an eigenvalue of multiplicity N (the number of active nodes). Whether the concern is justified is debatable, but again, there are tools in the kit to address it. First, regularization, as above:

$$[\mathbf{R}^t \boldsymbol{\nu} \mathbf{R} + \boldsymbol{\varepsilon} \mathbf{G} \boldsymbol{\delta} \mathbf{G}^t \boldsymbol{\varepsilon}] \mathbf{E} = \omega^2 \boldsymbol{\varepsilon} \mathbf{E}, \tag{19.5}$$

with $\boldsymbol{\delta}^{nn} = \int_{\tilde{n}} 1/\mu\varepsilon^2$ this time. Zero is not an eigenvalue any longer, but new eigenmodes appear, those of $\boldsymbol{\varepsilon} \mathbf{G} \boldsymbol{\delta} \mathbf{G}^t \boldsymbol{\varepsilon} \mathbf{E} = \omega^2 \boldsymbol{\varepsilon} \mathbf{E}$ under the restriction $\mathbf{E} = \mathbf{G} \boldsymbol{\psi}$. As remarked by WHITE and KONING [2000], we have here (again, assuming uniform coefficients) a phenomenon of "spectral complementarity" between the operators $\mathrm{rot} \circ \mathrm{rot}$ and $-\mathrm{grad} \circ \mathrm{div}$. The new modes, or "ghost modes" as they are called in WEILAND [1985], have to be sifted out, which is in principle easy[47] (evaluate the norm $|\mathbf{G}^t \boldsymbol{\varepsilon} \mathbf{E}|_\delta$), or "swept to the right" by inserting an appropriate scalar factor in front of the regularizing term. Second solution (TRAPP, MUNTEANU, SCHUHMANN, WEILAND and IOAN [2002]): Restrict the search of $\mathbf{E}$ to a complement of $\ker(\mathbf{R}^t \boldsymbol{\nu} \mathbf{R})$, which one can do by so-called "tree-cotree" techniques (ALBANESE and RUBINACCI [1988], MUNTEANU [2002]). This verges on the issue of *discrete Helmholtz decompositions*, another important tool in the kit, which cannot be given adequate treatment here (see RAPETTI, DUBOIS and BOSSAVIT [2002]).

[46] When $\boldsymbol{\sigma}$ and $\boldsymbol{\nu}$ are the Galerkin hodges, (19.2) corresponds to the edge-element discretization of (19.3).

[47] These ghost modes are *not* the (in)famous "spurious modes" which were such a nuisance before the advent of edge elements (cf. BOSSAVIT [1990b]). Spurious modes occur when one solves the eigenmode problem $\mathrm{rot}(\nu\,\mathrm{rot}\,\mathrm{E}) = \omega^2 \varepsilon \mathrm{E}$ by using *nodal vectorial* elements. Then (barring exceptional boundary conditions) the $\mathrm{rot}(\nu\,\mathrm{rot})$ matrix is regular (because the approximation space does not contain gradients, contrary to what happens with edge elements), but also – and for the same reason, as explained in BOSSAVIT [1998a] – poorly conditioned, which is the root of the evil. It would be wise *not* to take "ghost modes" and "spurious modes" as synonyms, in order to avoid confusion on this tricky point.

CHAPTER IV

Finite Elements

We now tackle the convergence analysis of the discrete version of problem (13.2), magnetostatics:

$$\mathbf{Db} = 0, \qquad \mathbf{h} = \boldsymbol{\nu}\mathbf{b}, \qquad \mathbf{R}^t\mathbf{h} = \mathbf{j}. \tag{18.1}$$

A preliminary comment on what that means is in order.

A few notational points before: The mesh is denoted m, the dual mesh is $\widetilde{m}$, and we shall subscript by m, when necessary, all mesh-related entities. For instance, the largest diameter of all p-cells, $p \geqslant 1$, primal and dual, will be denoted γ_m (with a mild abuse, since it also depends on the metric of the dual mesh), and called the "grain" of the pair of meshes. The computed solution $\{\mathbf{b}, \mathbf{h}\}$ will be $\{\mathbf{b}_m, \mathbf{h}_m\}$ when we wish to stress its dependence on the mesh-pair. And so on.

A first statement of our purpose is "study $\{\mathbf{b}_m, \mathbf{h}_m\}$ when γ_m tends to 0". Alas, this lacks definiteness, because how the *shapes* of the cells change in the process does matter a lot. In the case of triangular 2D meshes, for instance, there are well-known counter-examples (BABUŠKA and AZIZ [1976]) showing that, if one tolerates too much "flattening" of the triangles as the grain tends to 0, convergence may fail to 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 m or $\widetilde{m}$ of the family is the transform by similarity of one of them. The notation "$m \to 0$" will then refer to a sequence of meshes, all belonging to some definite uniform family, and such that their γ_ms tend to zero. Now we redefine our objective: Show that the error incurred by taking $\{\mathbf{b}_m, \mathbf{h}_m\}$ as a substitute for the real field $\{b, h\}$ tends to zero when $m \to 0$.

The practical implications of achieving this are well known. If, for a given m, the computed solution $\{\mathbf{b}_m, \mathbf{h}_m\}$ 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 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.

Fortunately, such refinement rules do exist (this is an active area of research: BÄNSCH [1991], BEY [1995], DE COUGNY and SHEPHARD [1999], MAUBACH [1995]). 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. 20.1. As one can see, at most five different shapes

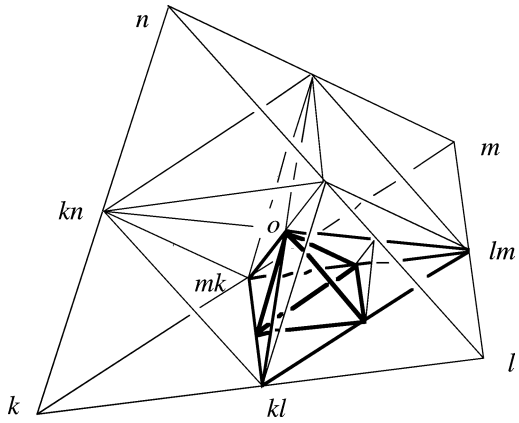

FIG. 20.1. Subdivision rule for a tetrahedron $\mathrm{T} = \{k, l, m, n\}$. (Mid-edges are denoted kl, lm, etc., and o is the barycenter.) A first halving of edges generates four small tetrahedra and a core octahedron, which itself can be divided into eight "octants" such as $\mathrm{O} = \{o, kl, lm, mk\}$, 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.

can occur, for each tetrahedral shape present in the original coarse mesh. In practice, not all volumes get refined simultaneously, so junction dissection schemes are needed, which enlarges the catalogue of shapes, but the latter is bounded nonetheless.)

For these reasons, we shall feel authorized to assume uniformity in this sense. We shall also posit that the hodge entries, whichever way they are built, only depend (up to a multiplicative factor) on the *shapes* of the cells contributing to them. Although stronger than necessary, these assumptions will make some proofs easier, and thus help focus on the main ideas.

20. Consistency

Back to the comparison between $\{\mathbf{b}_m, \mathbf{h}_m\}$ and $\{b, h\}$, a natural idea is to compare the computed DoF arrays, $\mathbf{b}_m$ and $\mathbf{h}_m$, with arrays of the same kind, $r_m b = \{\int_f b\colon\ f \in \mathcal{F}\}$ and $r_m h = \{\int_{\tilde{f}} h\colon\ f \in \mathcal{F}\}$, composed of the fluxes and m.m.f.'s of the (unknown) solution $\{b, h\}$ of the original problem (13.2). This implicitly defines two operators with the same name, r_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 m.m.f.'s. (No risk of confusion, since the name of the operand, b or h, reveals its nature.)

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

$$\mathbf{D} r_m b = 0, \qquad \mathbf{R}^t r_m h = \mathbf{j}, \tag{20.1}$$

since the entries of $\mathbf{j}$ are precisely the intensities across dual facets. Comparing with (18.1), we obtain

$$\mathbf{D}(\mathbf{b}_m - r_m b) = 0, \qquad \mathbf{R}^t(\mathbf{h}_m - r_m h) = 0, \tag{20.2}$$

and

$$(\mathbf{h}_m - r_m h) - \boldsymbol{\nu}(\mathbf{b}_m - r_m b) = (\boldsymbol{\nu} r_m - r_m \nu) b \equiv \boldsymbol{\nu}(r_m \mu - \boldsymbol{\mu} r_m) h. \tag{20.3}$$

Let us compute the μ-norm of both sides of (20.3). (For this piece of algebra, we shall use the notation announced in last chapter: $(\mathbf{b}, \mathbf{h})$ for a sum such as $\sum_{f\in\mathcal{F}} \mathbf{b}_f \mathbf{h}_f$, and $|\mathbf{h}|_\mu$ for $(\boldsymbol{\mu}\mathbf{h}, \mathbf{h})^{1/2}$, the μ-norm of $\mathbf{h}$, and other similar constructs.)

As this is done, "square" and "rectangle" terms appear. The rectangle term for the left-hand side is $-2(\mathbf{b}_m - r_m b, \mathbf{h}_m - r_m h)$, but since $\mathbf{D}(\mathbf{b}_m - r_m b) = 0$ implies the existence of some $\mathbf{a}$ such that $\mathbf{b}_m - r_m b = \mathbf{R}\mathbf{a}$, we have

$$(\mathbf{b}_m - r_m b, \mathbf{h}_m - r_m h) = (\mathbf{R}\mathbf{a}, \mathbf{h}_m - r_m h) = \big(\mathbf{a}, \mathbf{R}^t(\mathbf{h}_m - r_m h)\big) = 0,$$

after (20.2). Only square terms remain, and we get

$$\begin{aligned} &|\mathbf{h}_m - r_m h|_\mu^2 + |\mathbf{b}_m - r_m b|_\nu^2 \\ &\quad = \big|(\boldsymbol{\nu} r_m - r_m \nu) b\big|_\mu^2 \equiv \big|(\boldsymbol{\mu} r_m - r_m \mu) h\big|_\nu^2 \equiv (\boldsymbol{\nu} r_m b - r_m h, r_m b - \boldsymbol{\mu} r_m h). \end{aligned} \tag{20.4}$$

On the left-hand side, which has the dimension of an energy, we spot two plausible estimators for the error incurred by taking $\{\mathbf{b}_m, \mathbf{h}_m\}$ as a substitute for the real field $\{b, h\}$: the "error in (discrete) energy" [respectively coenergy], as regards $\mathbf{b}_m - r_m b$ [respectively $\mathbf{h}_m - r_m h$]. Components of $\mathbf{b}_m - r_m b$ are what can be called the "residual fluxes" $\mathbf{b}_f - \int_f b$, i.e., the difference between the computed flux embraced by facet f and the genuine (but unknown) flux $\int_f b$. Parallel considerations apply to h, with m.m.f.'s along $\tilde{f}$ instead of fluxes. It makes sense to try and *bound* these error terms by some function of γ_m. So let us focus on the right-hand side of (20.4), for instance on its second expression, the one in terms of h.

By definition of r_m, the f-component of $r_m(\mu h)$ is the flux of $b = \mu h$ embraced by f. On the other hand, the flux array $\boldsymbol{\mu} r_m h$ is the result of applying the discrete Hodge operator to the m.m.f. array $r_m h$, so the compound operators $r_m \mu$ and $\boldsymbol{\mu} r_m$ will not be equal: they give different fluxes when applied to a generic h. This contrasts with the equalities $(\mathbf{D} r_m - r_m \mathrm{d}) b = 0$ and $(\mathbf{R}^t r_m - r_m \mathrm{d}) h = 0$, which stem from the Stokes theorem. The mathematical word to express such equalities is "conjugacy": $\mathbf{D}$ and d are conjugate via r_m, and so are $\mathbf{R}^t$ and d, too. Thus, μ and $\boldsymbol{\mu}$ are *not* conjugate via r_m – and this is, of course, the reason why discretizing entails some error.

Yet, it may happen that $r_m \mu$ and $\boldsymbol{\mu} r_m$ *do* coincide for *some* hs. This is so, for instance, with piecewise constant fields, when $\boldsymbol{\mu}$ is the diagonal hodge of (16.3) and (16.4): in fact, these formulas were motivated by the desire to achieve this coincidence for such fields. Also, as we shall prove later, $r_m \nu$ and $\boldsymbol{\nu} r_m$ coincide on facet-element approximations of b, i.e., on divergence-free fields of the form $\sum_{f\in\mathcal{F}} \mathbf{b}_f w^f$ (which are meshwise constant), when $\boldsymbol{\nu}$ is the Galerkin hodge. Since all piecewise smooth fields differ from such special fields by some small residual, and the finer the mesh the smaller, we may

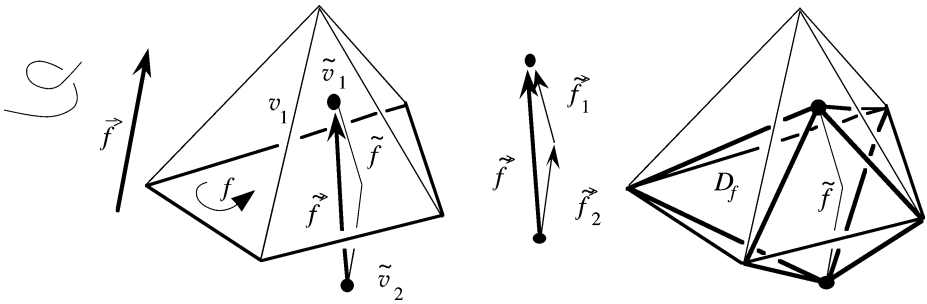

FIG. 20.2. As in Fig. 16.1, $\vec{f}$ denotes the vectorial area of facet f: the vector of magnitude area(f), normal to f, that points away from f in the direction derived from f's inner orientation by Ampère's rule. By $\vec{\tilde{f}}$ we denote the vector that joins the end points of the associated dual edge $\tilde{f}$. (An ambient orientation is assumed here. One could do without it by treating both $\vec{f}$ and $\vec{\tilde{f}}$ as axial vectors.) In case ν is not the same on both sides of f, understand $\nu\,\vec{\tilde{f}}$ as $\nu_2\,\vec{\tilde{f}}_2 + \nu_1\,\vec{\tilde{f}}_1$, where $\vec{\tilde{f}}_2$ and $\vec{\tilde{f}}_1$ are as suggested. Region D_f is the volume enclosed by the "tent" determined by the extremities of $\tilde{f}$ and the boundary of f. Note that $\vec{f}$ and $\nu\,\vec{\tilde{f}}$ always cross f in the same direction, but only in the orthogonal construction are they parallel (cf. Fig. 16.1): In that case, (20.6) can be satisfied by a *diagonal* hodge – cf. (16.3) and (16.4).

in such cases expect "asymptotic conjugacy", in the sense that the right-hand side of (20.4) will tend to 0 with m, for a piecewise smooth b or h. This property, which we rewrite informally but suggestively as

$$\boldsymbol{\nu} r_m - r_m \nu \to 0 \quad \text{when } m \to 0, \qquad \boldsymbol{\mu} r_m - r_m \mu \to 0 \quad \text{when } m \to 0 \tag{20.5}$$

(two equivalent statements), is called *consistency* of the approximation of μ and ν by $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$. Consistency, thus, implies asymptotic vanishing of the error in (discrete) energy, after (20.4).

Let's now take a heuristic step. (We revert to vector proxies for this. Fig. 20.2 explains about $\vec{f}$ and $\vec{\tilde{f}}$, and n and τ are normal and tangent unit vector fields, as earlier. The norm of an ordinary vector is $|\,|$.) Remark that the right-hand side of (20.4) is, according to its rightmost avatar, a sum of terms, one for each f, of the form

$$\left[\sum_{f'} \boldsymbol{\nu}^{ff'} \int_{f'} n \cdot \mathrm{B} - \int_{\tilde{f}} \nu\tau \cdot \mathrm{B}\right]\left[\int_f \mu n \cdot \mathrm{H} - \sum_{f''} \boldsymbol{\mu}^{ff''} \int_{\tilde{f}''} \tau \cdot \mathrm{H}\right],$$

which we'll abbreviate as $[\mathrm{B}, f][\mathrm{H}, f]$. Each should be made as small as possible for the sum to tend to 0. Suppose ν is uniform, and that boundary conditions are such that B and H are uniform. Then $[\mathrm{B}, f] = \mathrm{B} \cdot (\sum_{f'} \boldsymbol{\nu}^{ff'} \vec{f}' - \nu\,\vec{\tilde{f}})$. This term vanishes if

$$\sum_{f' \in \mathcal{F}} \boldsymbol{\nu}^{ff'} \vec{f}' = \nu\,\vec{\tilde{f}}\,. \tag{20.6}$$

(This implies $\sum_{f' \in \mathcal{F}} \boldsymbol{\mu}^{ff'} \nu\,\vec{\tilde{f}}' = \vec{f}$, and hence, cancellation of $[\mathrm{H}, f]$, too.) We therefore adopt this geometric compatibility condition as a *criterion* about $\boldsymbol{\nu}$. Clearly, the

diagonal hodge of (16.4) passes this test. But on the other hand, no diagonal $\boldsymbol{\nu}$ can satisfy (20.6) unless $\vec{f}$ and $\nu \overset{\rightsquigarrow}{f}$ are collinear.

PROPOSITION 20.1. *If $\boldsymbol{\nu}$ is diagonal, with $\boldsymbol{\nu}^{ff}\vec{f} = \nu\overset{\rightsquigarrow}{f}$, as required by the criterion, there is consistency.*

PROOF. (All C's, from now on, denote constants, not necessarily the same each time, possibly depending on the solution, but not on the mesh.) This time, the solution proxy B is only piecewise smooth, and possibly discontinuous if ν is not uniform, but its component parallel to $\vec{f}$, say $\mathcal{B}$, satisfies $|\mathcal{B}(x) - \mathcal{B}(y)| \leqslant C|x - y|$ in the region D_f of Fig. 20.2. One has[48] $\int_f n \cdot \mathrm{B} = \mathrm{area}(f)\mathcal{B}(x_f)$ and $\int_{\tilde{f}} \nu\tau \cdot \mathrm{B} = \mathrm{length}(\nu\overset{\rightsquigarrow}{f})\mathcal{B}(x_{\tilde{f}})$, for some averaging points x_f and $x_{\tilde{f}}$, the distance of which doesn't exceed γ_m, hence $[\mathrm{B}, f] \leqslant C\gamma_m \boldsymbol{\nu}^{ff}\,\mathrm{area}(f)$, by factoring out $\boldsymbol{\nu}^{ff}\,\mathrm{area}(f) \equiv \mathrm{length}(\nu\overset{\rightsquigarrow}{f})$, and similarly, $[\mathrm{H}, f] \leqslant C\gamma_m \boldsymbol{\mu}^{ff}\,\mathrm{length}(\nu\overset{\rightsquigarrow}{f})$. Noticing that $\mathrm{area}(f)\,\mathrm{length}(\nu\overset{\rightsquigarrow}{f}) = 3\int_{D_f}\nu$, and summing up with respect to f, one finds that

$$|\mathbf{h}_m - r_m h|_\mu^2 + |\mathbf{b}_m - r_m b|_\nu^2 \leqslant C\gamma_m^2, \tag{20.7}$$

the consistency result. □

Going back to (20.4), we conclude that both the ν-norm of the residual flux array and the μ-norm of the residual m.m.f. array tend to 0 as fast as γ_m, or faster,[49] a result we shall exploit next.

One may wonder whether the proof can be carried out in the case of a non-diagonal hodge, assuming (20.6). The author has not been able to do so on the basis of (20.6) only. The result is true under stronger hypotheses (stronger than necessary, perhaps): When the construction of $\boldsymbol{\nu}$ is a local one, i.e., $\boldsymbol{\nu}^{ff'} = 0$ unless facets f and f' belong to a common volume, and when the *infimum* δ_m of all cell diameters verifies $\delta_m \geqslant \beta\gamma_m$, with β independent of m. Then $\boldsymbol{\nu}$ has a band structure, and its terms behave in γ_m^{-1}, which makes it easy to prove that $[\mathrm{B}, f]$ is in $\mathrm{O}(\gamma_m^2)$. Handling $[\mathrm{H}, f]$ is more difficult, because $\boldsymbol{\mu}$ is full, and the key argument about averaging points not being farther apart than γ_m breaks down. On the other hand, owing to the band structure of $\boldsymbol{\nu}$, and its positive-definite character, $\boldsymbol{\mu}^{ff'}$ is small for distant f and f', which allows one to also bound $[\mathrm{H}, f]$ by $C\gamma_m^2$. The number of faces being in γ_m^{-3}, consistency ensues.

There is some way to go to turn such an argument into a proof, but this is enough to confirm (20.6) in its status of criterion as regards $\boldsymbol{\nu}$, a criterion which is satisfied, by construction (Fig. 16.1), in FIT (WEILAND [1996]) and in the cell method (TONTI

[48] In case ν is not the same on both sides of f, understand $\mathrm{length}(\nu\overset{\rightsquigarrow}{f})$ as $\nu_1\,\mathrm{length}(\overset{\rightsquigarrow}{f_1}) + \nu_2\,\mathrm{length}(\overset{\rightsquigarrow}{f_2})$. The underlying measure of lengths is not the Euclidean one, but the one associated with the metric induced by the Hodge operator ν.

[49] Convergence in γ_m^2 is in fact often observed when the meshes have some regularity, such as crystal-like symmetries, which may cancel out some terms in the Taylor expansions implicit in the above proof. For instance, the distance between points x_f and $x_{\tilde{f}}$ may well be in γ_m^2 rather than γ_m. This kind of phenomenon is commonplace in Numerical Analysis (SCHATZ, SLOAN and WAHLBIN [1996]).

[2001]), but allows a much larger choice. We'll see in a moment how and why it is satisfied in the Galerkin approach.

21. Stability

So, the left-hand side of (20.4) tends to 0. Although this is considered by many as sufficient in practice, one cannot be satisfied with such "discrete energy" estimates. Fields should be compared with fields. To really prove convergence, one should build from the DoF arrays $\mathbf{b}_m$ and $\mathbf{h}_m$ an approximation $\{b_m, h_m\}$ of the pair of differential forms $\{b, h\}$, and show that the discrepancies $b_m - b$ and $h_m - h$ tend to 0 with m in some definite sense. So we are after some map, that we shall denote by p_m, that would transform a flux array $\mathbf{b}$ into a 2-form $p_m\mathbf{b}$ and an m.m.f. array $\mathbf{h}$ into a twisted 1-form $p_m\mathbf{h}$. The map should behave naturally with respect to r_m, i.e.,

$$r_m p_m \mathbf{b} = \mathbf{b}, \qquad r_m p_m \mathbf{h} = \mathbf{h}, \tag{21.1}$$

as well as

$$|p_m r_m b - b|_\nu \to 0 \quad \text{and} \quad |p_m r_m h - h|_\mu \to 0 \quad \text{when } m \to 0 \tag{21.2}$$

(asymptotic vanishing of the "truncation error" $p_m r_m - 1$). A satisfactory result, then, would be that both $|b - p_m\mathbf{b}_m|_\nu$ and $|h - p_m\mathbf{h}_m|_\mu$ tend to 0 with m (convergence "in energy"). As will now be proved, this is warranted by the following inequalities:

$$\alpha |p_m \mathbf{b}|_\nu \leqslant |\mathbf{b}|_\nu, \qquad \alpha |p_m \mathbf{h}|_\mu \leqslant |\mathbf{h}|_\mu \tag{21.3}$$

for all $\mathbf{b}$ and $\mathbf{h}$, where the constant $\alpha > 0$ does not depend on m. Since $|\mathbf{b}|_\nu$ and $|\mathbf{h}|_\mu$ depend on the discrete hodge, (21.3) is a property of the approximation scheme, called *stability*.

PROPOSITION 21.1. *Consistency* (20.5) *being assumed,* (21.2) *and* (21.3) *entail convergence.*

PROOF. By consistency, the right-hand side of (20.4) tends to 0, whence $|\mathbf{b}_m - r_m b|_\nu \to 0$, and $|p_m\mathbf{b}_m - p_m r_m b|_\nu \to 0$ by (21.3). Therefore $p_m\mathbf{b}_m \to b$, "in energy", thanks to (21.2). Same argument about h. □

This is Lax's celebrated folk theorem (LAX and RICHTMYER [1956]): *consistency + stability = convergence.*

Below, we shall find a systematic way to construct p_m, the so-called *Whitney map.* But if we don't insist right now on generality, there is an easy way to find a suitable such map in the case of a simplicial primal mesh and of DoF arrays $\mathbf{b}$ that satisfy $\mathbf{Db} = 0$ (luckily, only these do matter in magnetostatics). The idea is to find a vector proxy $\overline{\mathrm{B}}$ uniform inside each tetrahedron with facet fluxes $\overline{\mathrm{B}} \cdot \vec{f}$ equal to $\mathbf{b}_f$. (Then, $\operatorname{div} \overline{\mathrm{B}} = 0$ 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 $\overline{\mathrm{B}}$) to be determined from four fluxes linked by one linear relation, $\sum_f \mathbf{D}_v^f \mathbf{b}_f = 0$, so the problem has a solution, which we take as $p_m\mathbf{b}$.

Then,[50] $p_m r_m b \to b$. As for the stability condition (21.3), one has $|p_m \mathbf{b}|_\nu^2 = \int_D \nu |\overline{\mathrm{B}}|^2$, a quadratic form with respect to the facet fluxes, which we may therefore denote by $(\mathbf{b}, \mathbf{Nb})$, with $\mathbf{N}$ some positive definite matrix. Now, suppose first a *single* tetrahedron in the mesh m, and consider the Rayleigh-like quotient $(\mathbf{b}, \boldsymbol{\nu}\mathbf{b})/(\mathbf{b}, \mathbf{Nb})$. Its lower bound, strictly positive, depends only on the *shape* of the tetrahedron, not on its size. Then, uniformity of the family of meshes, and of the construction of $\boldsymbol{\nu}$, allows us to take for α in (21.3) the smallest of these lower bounds, which is strictly positive and independent of m. We may thereby conclude that $p_m \mathbf{b}_m$ converges towards b in energy.

No similar construction on the side of h is available, but this is not such a handicap: if $p_m \mathbf{b}_m \to b$, then $\nu p_m \mathbf{b}_m \to h$. This amounts to setting p_m on the dual side equal to $\nu p_m \boldsymbol{\mu}$. The problem with that is, $p_m \mathbf{h}$ 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. But never mind, since this "non-conformal" approximation converges in energy.

Yet, we need a more encompassing p_m map, if only because $\mathbf{Db} = 0$ was just a happy accident. Before turning to that, which will be laborious, let's briefly discuss convergence in the full Maxwell case.

22. The time-dependent case

Here is a sketch of the convergence proof for the generalized Yee scheme (17.1) and (17.2) of last chapter.

First, linear interpolation in time between the values of the DoF arrays, as output by the scheme, provides DoF-array-valued functions of time which converge, when δt tends to zero, towards the solution of the "spatially discretized" equations (16.1) and (16.2). This is not difficult.

Next, linearity of the equations allows one to pass from the time domain to the frequency domain, via a Laplace transformation. Instead of studying (16.1) and (16.2), therefore, one may examine the behavior of the solution of

$$-p\mathbf{D} + \mathbf{R}^t \mathbf{H} = \mathbf{J}, \qquad p\mathbf{B} + \mathbf{R}\mathbf{E} = 0, \tag{22.1}$$

$$\mathbf{D} = \boldsymbol{\varepsilon}\mathbf{E}, \qquad \mathbf{B} = \boldsymbol{\mu}\mathbf{H}, \tag{22.2}$$

when $m \to 0$. 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 (16.1) and (16.2) will ensue, by inverse Laplace transformation. The main problem, therefore, is to compare $\mathbf{E}$, $\mathbf{B}$, $\mathbf{H}$, $\mathbf{D}$, as given by (22.1) and (22.2), with $r_m \mathrm{E}$, $r_m \mathrm{B}$, $r_m \mathrm{H}$, $r_m \mathrm{D}$, where small capitals, again, denote Laplace transforms, but of differential forms this time.

[50] This is an exercise, for which the following hints should suffice. Start from b, piecewise smooth, such that $\mathrm{d}b = 0$, set $\mathbf{b} = r_m b$, get $\overline{\mathrm{B}}$ as above, and aim at finding an upper bound for $|\mathrm{B} - \overline{\mathrm{B}}|$, where B is the proxy of b, over a tetrahedron T. For this, evaluate $\nabla\lambda \cdot \int_\mathrm{T}(\mathrm{B} - \overline{\mathrm{B}})$, where λ is an affine function such that $|\nabla\lambda| = 1$. Integrate by parts, remark that $\int_f \lambda n \cdot \overline{\mathrm{B}} = \lambda(x_f)\mathbf{b}_f$, where x_f is the barycenter of f. Taylor-expand $n \cdot \mathrm{B}$ about x_f, hence a bound in $C\gamma_m^4$ for $\int_{\partial\mathrm{T}} \lambda n \cdot (\mathrm{B} - \overline{\mathrm{B}})$, from which stems $|\int_\mathrm{T}(\mathrm{B} - \overline{\mathrm{B}})| \leqslant C\gamma_m^4$. Use uniformity to conclude that $|\mathrm{B} - \overline{\mathrm{B}}| \leqslant C\gamma_m$.

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

$$p\boldsymbol{\mu}(\mathbf{H} - r_m \mathrm{H}) + \mathbf{R}(\mathbf{E} - r_m \mathrm{E}) = p(r_m \mu - \boldsymbol{\mu} r_m)\mathrm{H}, \tag{22.3}$$

$$-p\boldsymbol{\varepsilon}(\mathbf{E} - r_m \mathrm{E}) + \mathbf{R}^t(\mathbf{H} - r_m \mathrm{H}) = -p(r_m \varepsilon - \boldsymbol{\varepsilon} r_m)\mathrm{E}. \tag{22.4}$$

Then, right-multiply (22.3) (in the sense of $(\ ,\ )$) by $(\mathbf{H} - r_m \mathrm{H})^*$ and the complex conjugate of (22.4) by $-(\mathbf{E} - r_m \mathrm{E})$, 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 (20.3), on the one hand, and (22.3), (22.4), 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 discrete hodge in statics, therefore, is a good one in transient situations. Let's tentatively promote this remark to the rank of 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.*

At this stage, we may feel more confident about the quality of the toolkit: If the discrete hodges and the meshes are compatible in the sense of (20.6), so that consistency can be achieved, if there is a way to pass from DoFs to fields which binds energy and discrete energy tightly enough for stability (21.3) to hold, then convergence will ensue. So we need the p_m operator. We would need it, anyway, to determine fluxes, e.m.f.'s, etc., at a finer scale than what the mesh provides – motivation enough to search for interpolants, but not the most compelling reason to do so: Field reconstruction from the DoFs is needed, basically, *to assess stability*, in the above sense, and thereby, the validity of the method. Whitney forms, which will now enter the scene, provide this mechanism.

23. Whitney forms

Let's summarize the requirements about the generic map p_m. It should map each kind of DoF array to a differential form of the appropriate kind: $p_m\mathbf{e}$, starting from an edge-based DoF array $\mathbf{e}$, should be a 1-form; $p_m\mathbf{b}$, obtained from a facet-based $\mathbf{b}$, should be a 2-form, and so forth. Properties (21.1) and (21.2) should hold for all kinds, too, so in short,

$$r_m p_m = 1, \quad p_m r_m \to 1 \quad \text{when } m \to 0. \tag{23.1}$$

The stability property (21.3) will automatically be satisfied in the case of a uniform family of meshes. Moreover, we expect $\mathrm{d}b = 0$ when $\mathbf{Db} = 0$, $\mathrm{d}e = 0$ when $\mathbf{Re} = 0$, etc. More generally, $\mathbf{Ra} = \mathbf{b}$ should entail $\mathrm{d}a = b$, and so forth. These are desirable features of the toolkit. The neatest way to secure them is to enforce the structural property

$$\mathrm{d}p_m = p_m\mathbf{d}, \tag{23.2}$$

at all levels (Fig. 23.1): d and $\mathbf{d}$ should be conjugate, via p_m, or said differently, Fig. 23.1 should be a *commutative diagram*. Remarkably, these prescriptions will prove sufficient to generate interpolants in an essentially unique way. Such interpolants are known as *Whitney forms* (WHITNEY [1957]), and we shall refer to the structure they constitute as the *Whitney complex*.

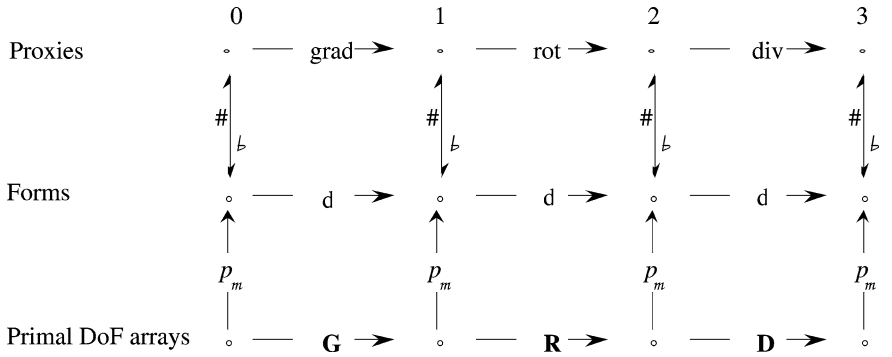

FIG. 23.1. Diagrammatic rendering of (23.2), with part of Fig. 8.1 added. Flat and sharp symbols represent the isomorphism between differential forms and their scalar or vector proxies.

23.1. Whitney forms as a device to approximate manifolds

We address the question by taking a detour, to see things from a viewpoint consistent with our earlier definition of differential forms as maps from manifolds to numbers. A differential form, say, for definiteness, b, maps a p-manifold S to the number $\int_S b$, with $p = 2$ here. Suppose we are able to approximate S by a p-chain, i.e., here, a chain based on facets, $p_m^t S = \sum_{f \in \mathcal{F}} \mathbf{c}^f f$. Then a natural approximation to $\int_S b$ is $\int_{p_m^t S} b$. But this number we know, by linearity: since $\int_f r_m b = \mathbf{b}_f$, it equals the sum $\sum_f \mathbf{c}^f \mathbf{b}_f$, that we shall denote $\langle \mathbf{c}; \mathbf{b} \rangle$ (with boldface brackets). Hence an approximate knowledge of the field b, i.e., of all its measurable attributes – the fluxes – from the DoF array $\mathbf{b}$. In particular, fluxes embraced by *small* surfaces (small with respect to the grain of the mesh) will be computable from $\mathbf{b}$, which meets our expectations about interpolating to local values of b. The question has thus become "how best to represent S by a 2-chain?". Fig. 23.2 (where $p = 1$, so a curve c replaces S) gives the idea.

Once we know about the manifold-to-chain map p_m^t, we know about Whitney forms: For instance, the one associated with facet f is, like the field b itself, a map from surfaces to numbers, namely the map $S \to \mathbf{c}^f$ that assigns to S its weight with respect to f. We denote this map by w^f and its value at S by $\int_S w^f$ or by $\langle S; w^f \rangle$ as we have done earlier. (The notational redundancy will prove useful.) Note that $\langle p_m^t S; \mathbf{b} \rangle = \int_S \sum_f \mathbf{b}_f w^f = \int_S p_m \mathbf{b} \equiv \langle S; p_m \mathbf{b} \rangle$, which justifies the "$p_m^t$" notation: A transposition is indeed involved.

23.2. A generating formula

Now, let's enter the hard core of it. A simplicial primal mesh will be assumed until further notice. (We shall see later how to lift this restriction.) Results will hold for any spatial dimension n and all simplicial dimensions $p \leqslant n$, but will be stated as if n was 3 and $p = 1$ or 2 (edge and facet elements). So we shall also write proofs, even recursive ones that are supposed to move from p to $p + 1$ (see, e.g., Proposition 23.1), as if p had a specific value (1 or 2), and thereby prefer $\mathbf{R}, \mathbf{D}$, or $\mathbf{R}^t, \mathbf{D}^t$, to $\mathbf{d}$ or $\boldsymbol{\partial}$. That the proof has general validity notwithstanding should be obvious each time.

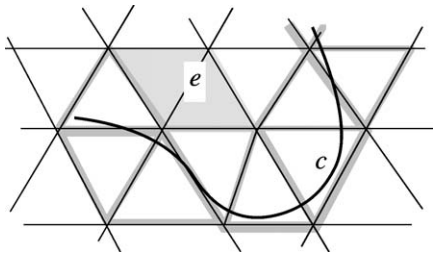

FIG. 23.2. Representing curve c by a weighted sum of mesh-edges, i.e., by a 1-chain. Graded thicknesses of the edges are meant to suggest the respective weights assigned to them. Edges such as e, whose "control domain" (shaded) doesn't intersect c, have zero weight. (A weight can be negative, if the edge is oriented backwards with respect to c.) Which weights thus to assign is the central issue in our approach to Whitney forms.

We use $\lambda^n(x)$ for the barycentric weight of point x with respect to node n, when x belongs to one of the tetrahedra which share node n (otherwise, $\lambda^n(x) = 0$). We'll soon see that $w^n = \lambda^n$ is the natural choice for nodal 0-forms, and again this dual notation will make some formulas more readable. We define $\lambda^e = \lambda^m + \lambda^n$, when edge $e = \{m, n\}$, as well as $\lambda^f = \lambda^l + \lambda^m + \lambda^n$ for facet $f = \{l, m, n\}$, etc. When $e = \{m, n\}$ and $f = \{l, m, n\}$, we denote node l by $f - e$. Thus λ^{f-e} refers to (in that case) λ^l, and equals $\lambda^f - \lambda^e$. The oriented segment from point x to point y is xy, the oriented triangle formed by points x, y, z, in this order, is xyz. And although node n and its location x_n should not be confused, we shall indulge in writing, for instance, ijx for the triangle based on points x_i, x_j, and x, when i and j are node labels.

The weights in the case of a "small manifold", such as a point, a segment, etc.,[51] will now be constructed, and what to use for non-small ones, i.e., the maps w^e, w^f, etc., from lines, surfaces, etc., to reals, will follow by linearity. The principle of this construction is to enforce the following commutative diagram property:

$$\boldsymbol{\partial}\, p^t_m = p^t_m \partial, \tag{23.3}$$

which implies, by transposition, $\mathrm{d} p_m = p_m \mathbf{d}$, the required structural property (23.2).[52] We shall not endeavor to prove, step by step, that our construction does satisfy (23.3), although that would be an option. Rather, we shall let (23.3) inspire the definition that follows, and then, directly establish that $\mathrm{d} p_m = p_m \mathbf{d}$. This in turn will give (23.3) by transposition.

DEFINITION 23.1. Starting from $w^n = \lambda^n$, the simplicial Whitney forms are

$$w^e = \sum_{n \in \mathcal{N}} \mathbf{G}^n_e \lambda^{e-n}\, \mathrm{d}w^n, \quad w^f = \sum_{e \in \mathcal{E}} \mathbf{R}^e_f \lambda^{f-e}\, \mathrm{d}w^e, \quad w^v = \sum_{f \in \mathcal{F}} \mathbf{D}^f_v \lambda^{v-f}\, \mathrm{d}w^f \tag{23.4}$$

(and so on, recursively, to higher dimensions).

[51] The proper underlying concept, not used here, is that of *multivector* at point x.

[52] If moreover $\ker(\boldsymbol{\partial}_p) = \mathrm{cod}(\boldsymbol{\partial}_{p+1})$, i.e., in the case of a trivial topology, then $\ker(\mathrm{d}_p) = \mathrm{cod}(\mathrm{d}_{p-1})$, just as, by transposition, $\ker(\mathbf{d}_p) = \mathrm{cod}(\mathbf{d}_{p-1})$. One says the Whitney spaces of forms, as linked by the d_p, form an *exact sequence*.

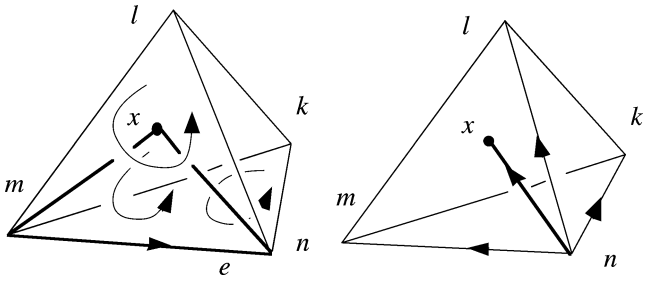

FIG. 23.3. Left: With edge $e = \{m, n\}$ and facets $\{m, n, k\}$ and $\{m, n, l\}$ oriented as shown, the 2-chain to associate with the "join" $x \vee e$, alias mnx, can only be $\lambda^k(x)mnk + \lambda^l(x)mnl$. This is what (23.5) says. Right: Same relation between the join $x \vee n$ and the 1-chain $\lambda^k(x)nk + \lambda^l(x)nl + \lambda^m(x)nm$.

Let us justify this statement, by showing how compliance with (23.3) suggests these formulas. The starting point comes from finite element interpolation theory, which in our present stand consists in expressing a point x as a weighted sum of nodes, the weights $w^n(x)$ being the barycentric ones, $\lambda^n(x)$. (Note how the standard p_m for nodal DoFs, $p_m\boldsymbol{\varphi} = \sum_n \boldsymbol{\varphi}_n w^n$, comes from $p^t_m x = \sum_n w^n(x)n$ by transposition.) Recursively, suppose we know the proper weights for a segment yz, i.e., the bracketed terms in the sum $p^t_m yz = \sum_e \langle yz; w^e\rangle e$, and let us try to find $p^t_m xyz$. By linearity, $p^t_m xyz = \sum_e \langle yz; w^e\rangle p^t_m(x \vee e)$, where the "join" $x \vee e$ is the triangle displayed in Fig. 23.3, left. So the question is: which 2-chain best represents $x \vee e$? As suggested by Fig. 23.3, the only answer consistent with (23.3) is

$$p^t_m(x \vee e) = \sum_{f\in\mathcal{F}} \mathbf{R}^e_f \lambda^{f-e}(x)\, f. \tag{23.5}$$

Indeed, this formula expresses $x \vee e$ as the average of mnk and mnl (the only two facets f for which $\mathbf{R}^e_f \neq 0$), with weights that depend on the relative proximity of x to them. So $p^t_m xyz = \sum_{e,f} \mathbf{R}^e_f \lambda^{f-e}(x)\langle yz; w^e\rangle f \equiv \sum_f \langle xyz; w^f\rangle f$, hence

$$\langle xyz; w^f\rangle = \sum_e \mathbf{R}^e_f \lambda^{f-e}(x)\langle yz; w^e\rangle. \tag{23.6}$$

On the other hand, since a degenerate triangle such as xzx should get zero weights, we expect $0 = \langle xzx; w^f\rangle = \sum_e \mathbf{R}^e_f \lambda^{f-e}(x)\langle zx; w^e\rangle$, and the same for $\langle xxy; w^f\rangle$. From this (which will come out true after Proposition 23.1 below), we get

$$\begin{aligned}\langle xyz; w^f\rangle &= \sum_e \mathbf{R}^e_f \lambda^{f-e}(x)\langle yz + zx + xy; w^e\rangle \\ &= \sum_e \mathbf{R}^e_f \lambda^{f-e}(x)\big\langle \partial(xyz); w^e\big\rangle = \sum_e \mathbf{R}^e_f\, \lambda^{f-e}(x)\langle xyz; \mathrm{d}w^e\rangle\end{aligned}$$

for any small triangle xyz, by Stokes, and hence $w^f = \sum_e \mathbf{R}^e_f \lambda^{f-e}\,\mathrm{d}w^e$.

Thus, formulas (23.4) – which one should conceive as the unfolding of a unique formula – are forced on us, as soon as we accept (23.5) as the right way, amply suggested by Fig. 23.3, to pass from the weights for a simplex s to those for the join $x \vee s$. The

reader will easily check that (23.4) describes the Whitney forms as they are more widely known, that is, on a tetrahedron $\{k, l, m, n\}$,

$$w^n = \lambda^n$$

for node n,

$$w^e = \lambda^m \, \mathrm{d}\lambda^n - \lambda^n \, \mathrm{d}\lambda^m$$

for edge $e = \{m, n\}$,

$$w^f = 2(\lambda^l \, \mathrm{d}\lambda^m \wedge \mathrm{d}\lambda^n + \lambda^m \, \mathrm{d}\lambda^n \wedge \mathrm{d}\lambda^l + \lambda^n \, \mathrm{d}\lambda^l \wedge \mathrm{d}\lambda^m)$$

for facet $f = \{l, m, n\}$, and

$$w^v = 6(\lambda^k \, \mathrm{d}\lambda^l \wedge \mathrm{d}\lambda^m \wedge \mathrm{d}\lambda^n + \lambda^l \, \mathrm{d}\lambda^m \wedge \mathrm{d}\lambda^n \wedge \mathrm{d}\lambda^k + \lambda^m \, \mathrm{d}\lambda^n \wedge \mathrm{d}\lambda^k \wedge \mathrm{d}\lambda^l + \lambda^n \, \mathrm{d}\lambda^k \wedge \mathrm{d}\lambda^l \wedge \mathrm{d}\lambda^m)$$

for volume $v = \{k, l, m, n\}$. In higher dimensions (WHITNEY [1957]), the Whitney form of a p-simplex $s = \{n_0, n_1, \ldots, n_p\}$, with inner orientation implied by the order of the nodes, is

$$w^s = p! \sum_{i=0,\ldots,p} (-1)^i w^{n_i} \, \mathrm{d}w^{n_0} \wedge \cdots \langle i \rangle \cdots \wedge \mathrm{d}w^{n_p},$$

where the $\langle i \rangle$ means "omit the term $\mathrm{d}w^{n_i}$".

From now on, we denote by W^p the finite-dimensional subspaces of $\mathcal{F}^p$ generated by these basic forms.

REMARK 23.1. To find the vector proxies of w^e and w^f, substitute ∇ and $\times$ to d and $\wedge$. The scalar proxy of w^v is simply the function equal to $1/\mathrm{vol}(v)$ on v, 0 elsewhere. The reader is invited to establish the following formulas:

$$w^{mn}(x) = (kl \times kx)/6\,\mathrm{vol}(klmn), \qquad w^{mnk}(x) = xl/3\,\mathrm{vol}(v),$$

very useful when it comes to actual coding. (Other handy formulas, at this stage, are $\mathrm{rot}(x \to v \times ox) = 2v$ and $\mathrm{div}(x \to ox) = 3$, where o is some origin point and v a fixed vector. As an exercise, one may use this to check on Proposition 23.3 below.)

REMARK 23.2. One may recognize in (23.6) the development of the 3×3 determinant of the array of barycentric coordinates of points x, y, z, with respect to nodes l, m, n, hence the geometrical interpretation of the weights displayed in Fig. 23.4.

23.3. Properties of Whitney forms

Thus in possession of a rationale for (23.4), we now derive from it a few formulas, for their own sake and as a preparation for the proof of the all important $\mathrm{d}p_m = p_m \mathbf{d}$ result, Proposition 23.3 below.

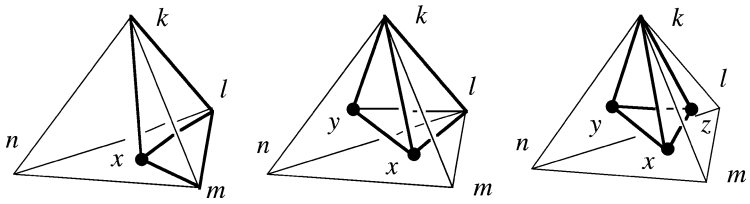

FIG. 23.4. Just as the barycentric weight of point x with respect to node n is vol$(klmx)$, if one takes vol$(klmn)$ as unit, the weight of the segment xy with respect to edge $\{m,n\}$ is vol$(klxy)$, and the weight of the triangle xyz with respect to facet $\{l,m,n\}$ is vol$(kxyz)$.

PROPOSITION 23.1. *For each p-simplex, there is one linear relation between Whitney forms associated with $(p-1)$-faces of this simplex. For instance, for each f,*

$$\sum_{e\in\mathcal{E}} \mathbf{R}_f^e \lambda^{f-e} w^e = 0.$$

PROOF. By (23.4), $\sum_e \mathbf{R}_f^e \lambda^{f-e} w^e = \sum_{e,n} \lambda^{f-e}\lambda^{e-n}\mathbf{R}_f^e \mathbf{G}_e^n w^n = 0$, thanks to the relation $\mathbf{RG}=0$, because $\lambda^{f-e}\lambda^{e-n}$, which is the same for all e in ∂f, can be factored out. □

As a corollary, and by using $\mathrm{d}(\lambda\omega) = \mathrm{d}\lambda\wedge\omega + \lambda\,\mathrm{d}\omega$, we have

$$w^f = -\sum_{e\in\mathcal{E}} \mathbf{R}_f^e \,\mathrm{d}\lambda^{f-e}\wedge w^e,$$

and other similar alternatives to (23.4).

PROPOSITION 23.2. *For each p-simplex s, one has*

$$\text{(i)}\quad \lambda^s\,\mathrm{d}w^s = (p+1)\,\mathrm{d}\lambda^s\wedge w^s, \qquad \text{(ii)}\quad \mathrm{d}\lambda^s\wedge \mathrm{d}w^s = 0. \tag{23.7}$$

PROOF. This is true for $p=0$. Assume it for $p=1$. Then

$$\mathrm{d}w^f = \sum_e \mathbf{R}_f^e\,\mathrm{d}\lambda^{f-e}\wedge\mathrm{d}w^e = \sum_e \mathbf{R}_f^e\,\mathrm{d}\lambda^f\wedge\mathrm{d}w^e \equiv \mathrm{d}\lambda^f\wedge\sum_e \mathbf{R}_f^e\,\mathrm{d}w^e$$

by (ii), hence $\mathrm{d}\lambda^f\wedge\mathrm{d}w^f=0$. Next,

$$\begin{aligned}\lambda^f\,\mathrm{d}w^f &= \lambda^f\left(\sum_e \mathbf{R}_f^e\,\mathrm{d}\lambda^f\wedge\mathrm{d}w^e\right) = \mathrm{d}\lambda^f\wedge\left(\sum_e \mathbf{R}_f^e\lambda^f\,\mathrm{d}w^e\right)\\ &= \mathrm{d}\lambda^f\wedge\left(w^f + \sum_e \mathbf{R}_f^e\lambda^e\,\mathrm{d}w^e\right),\end{aligned}$$

which thanks to (i) equals

$$\begin{aligned}\mathrm{d}\lambda^f\wedge\left(w^f + 2\sum_e \mathbf{R}_f^e\,\mathrm{d}\lambda^e\wedge w^e\right) &= \mathrm{d}\lambda^f\wedge w^f - 2\mathrm{d}\lambda^f\wedge\sum_e \mathbf{R}_f^e\,\mathrm{d}\lambda^{f-e}\wedge w^e\\ &= 3\mathrm{d}\lambda^f\wedge w^f,\end{aligned}$$

which proves (i) for $p=2$. Hence (ii) for $p=2$ by taking the d. □

Next, yet another variant of (23.4), but without summation this time. For any edge e such that $\mathbf{R}^e_f \neq 0$, one has

$$\mathbf{R}^e_f w^f = \lambda^{f-e}\,\mathrm{d}w^e - 2\,\mathrm{d}\lambda^{f-e} \wedge w^e. \tag{23.8}$$

This is proved by recursion, using $\mathbf{G}^n_{e'} w^{e'} = \lambda^{e'-n}\,\mathrm{d}w^n - \mathrm{d}\lambda^{e'-n} w^n$, where $n = e \cap e'$, and the identity $\mathbf{G}^n_{e'}\mathbf{G}^n_e = -\mathbf{R}^{e'}_f \mathbf{R}^e_f$. We may now conclude with the main result about structural properties (cf. Fig. 23.1):

PROPOSITION 23.3. *One has*

$$\mathrm{d}w^e = \sum_{f\in\mathcal{F}} \mathbf{R}^e_f w^f,$$

and hence, by linearity, $\mathrm{d}p_m = p_m \mathbf{d}$.

PROOF. Since both sides vanish out of the "star" of e, i.e., the union $\mathrm{st}(e)$ of volumes containing it, one may do as if $\mathrm{st}(e)$ were the whole meshed region. Note that $\sum_f \mathbf{R}^e_f \lambda^f = 1 - \lambda^e$ on $\mathrm{st}(e)$. Then,

$$\begin{aligned}\sum_f \mathbf{R}^e_f w^f &= \sum_f [\lambda^{f-e}\,\mathrm{d}w^e - 2\,\mathrm{d}\lambda^{f-e} \wedge w^e] = (1-\lambda^e)\,\mathrm{d}w^e - 2\,\mathrm{d}(1-\lambda^e) \wedge w^e \\ &= (1-\lambda^e)\,\mathrm{d}w^e + \lambda^e \wedge \mathrm{d}w^e \equiv \mathrm{d}w^e,\end{aligned}$$

by using (i). Now, $\mathrm{d}(p_m \mathbf{a}) = \mathrm{d}(\sum_e \mathbf{a}_e w^e) = \sum_{e,f} \mathbf{R}^e_f \mathbf{a}_e w^f = \sum_f (\mathbf{R}\mathbf{a})_f w^f = p_m(\mathbf{d}\mathbf{a})$. □

As a corollary, $\mathrm{d}W^{p-1} \subset W^p$, and if $\ker(\mathbf{d}_p) = \mathrm{cod}(\mathbf{d}_{p-1})$, then $\ker(\mathrm{d}; W^p) = \mathrm{d}W^{p-1}$, the *exact sequence* property of Whitney spaces in case of trivial topology.

23.4. "Partition of unity"

For what comes now, we revert to the standard vector analysis framework, where w^f denotes the proxy vector field (i.e., $2(\lambda^l \nabla\lambda^m \times \nabla\lambda^n + \cdots)$) of the Whitney form w^f.

Recall that barycentric functions sum to 1, thus forming a "partition of unity": $\sum_{n\in\mathcal{N}} w^n = 1$. We shall drop the ugly arrows in what follows, and use symbol f not only as a label, but also for the vectorial area of f (Fig. 20.2). Same dual use of $\tilde{f}$. Same convention for xyz, to be understood as a triangle or as its vectorial area, according to the context.

PROPOSITION 23.4. *At all points x, for all vectors v,*

$$\sum_{f\in\mathcal{F}} \big(w^f(x)\cdot v\big) f = v. \tag{23.9}$$

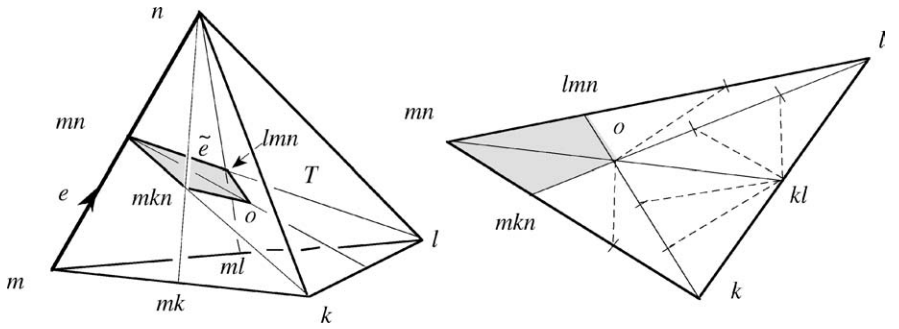

FIG. 23.5. Why $\int_T w^e = \tilde{e}$ in the barycentric construction of the dual mesh. First, the length of the altitude from n is $1/|\nabla w^n|$, therefore $\int_T \nabla w^n = klm/3$. Next, the average of w^n or w^m is $1/4$. So $\int_T w^e \equiv \int_T [w^m \nabla w^n - w^n \nabla w^m]$ is a vector equal to $(klm/3 + kln/3)/4$. As the figure shows (all twelve triangles on the right have the same area), this is precisely the vectorial area of $\tilde{e}$.

This is a case of something true of all simplices, and a consequence of the above construction in which the weights $\langle xyz; w^f(x)\rangle$ were assigned in order to have $xyz = \sum_f \langle xyz; w^f(x)\rangle f$. Replacing there w^f by its proxy, and xyz and f by their vectorial areas, we do find (23.9). As a corollary (replace f by g, v by $vw^f(x)$, and integrate in x), the entries $\boldsymbol{\nu}^{fg}$ of the Galerkin facet elements mass matrix satisfy

$$\sum_{g\in\mathcal{F}} \boldsymbol{\nu}^{fg} g = v\tilde{f},$$

where $v\tilde{f}$ is as explained on Fig. 20.2, but with the important specification that here, we are dealing with the *barycentric* dual mesh. That $\int vw^f = v\tilde{f}$ is an exercise in elementary geometry, and a similar formula holds for all Whitney forms (Fig. 23.5). Now, compare this with (20.6), the compatibility condition that was brought to light by the convergence analysis: We have proved, at last, that the Galerkin hodges do satisfy it.

24. Higher-degree forms

Let's sum up: Whitney forms were built in such a way that the partition of unity property (23.9) ensues. This property makes the mass matrix $\boldsymbol{\nu}$ of facet elements satisfy, with respect to the mesh and its barycentric dual, a compatibility criterion, (20.6), which we earlier recognized as a requisite for consistency. Therefore, we may assert that *Whitney forms of higher polynomial degree*, too, *should satisfy* (23.9), and take this as heuristic guide in the derivation of such forms.

Being a priori more numerous, higher-degree forms will make a finer partition. But we have a way to refine the partition (23.9): Multiply it by the λ^ns, which themselves form a partition of unity. This results in

$$\sum_{f\in\mathcal{F}, n\in\mathcal{N}} \big(\lambda^n w^f(x)\cdot v\big) f = v,$$

hence the recipe: Attach to edges, facets, etc., the products $\lambda^n w^e$, $\lambda^n w^f$, etc., where n spans $\mathcal{N}$. Instead of the usual Whitney spaces W^p, with forms of polynomial degree

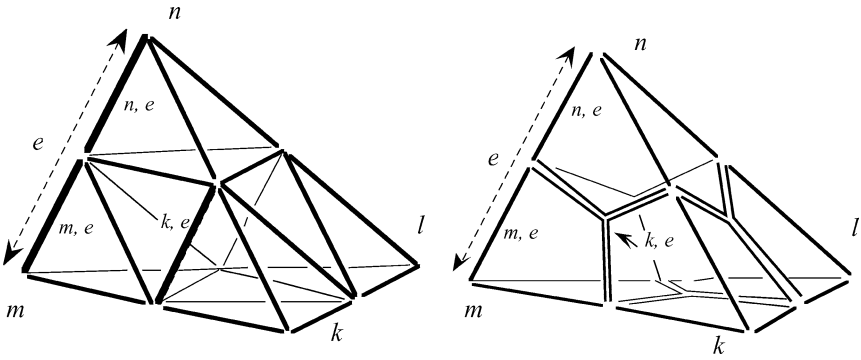

FIG. 24.1. Left: "Small" edges, in one-to-one correspondence with the forms $\lambda^n w^e$, and how they are labelled. Right: A variant where some small edges, such as $\{k, e\}$, are broken lines. These three crooked small edges, with proper signs, add up to the null chain, hence the compatibility condition of Note 53 is built in.

1 at most, we thus obtain larger spaces W_2^p, with forms of polynomial degree 2 at most. (For consistency, W^p may now be denoted W_1^p.) As we shall prove in a moment (under the assumption of trivial topology, but this is no serious restriction), the complex they constitute enjoys the exact sequence property: If for instance $b = \sum_{n,f} \mathbf{b}_{nf} \lambda^n w^f$ satisfies $\mathrm{d}b = 0$ (which means it has a divergence-free proxy) then there are DoFs $\mathbf{a}_{ne}$ such that $b = \mathrm{d}(\sum_{n,e} \mathbf{a}_{ne} \lambda^n w^e)$. (How to define W_k^p, for polynomial degrees $k = 3, \ldots,$ should now be obvious.)

Note however that, because of Proposition 23.1, these new forms are not linearly independent. For instance, the span of the $\lambda^n w^e$s, over a tetrahedron, has dimension 20 instead of the apparent 24, because Proposition 23.1 imposes one linear relation per facet. Over the whole mesh, with N nodes, E edges, F facets, the two products $\lambda^m w^e$ and $\lambda^n w^e$ for each edge $e = \{m, n\}$, and the three products $\lambda^{f-e} w^e$ for each facet f, make a total of $2E + 3F$ generators for W_2^1. But with one relation per facet, the dimension of W_2^1 is only $2(E + F)$. (The spans of the $\lambda^n w^n$s, the $\lambda^n w^f$s, and the $\lambda^n w^v$s, have respective dimensions $N + E$, $3(F + V)$, and $4V$. The general formula is $\dim(W_2^p) = (p + 1)(S_p + S_{p+1})$, where S_p is the number of p-simplices. Note that $\sum_p (-1)^p \dim(W_2^p) = \sum_p (-1)^p S_p \equiv \chi$, the Euler–Poincaré constant of the meshed domain.)

Owing to this redundancy, the main problem with these forms is, how to interpret the DoFs. With standard edge elements, the DoF $\mathbf{a}_{e'}$ is the integral of the 1-form $a = \sum_e \mathbf{a}_e w^e$ over edge e'. In different words, the square matrix of the circulations $\langle e'; w^e \rangle$ is the identity matrix: edges and edge elements are *in duality* in this precise sense (just like the basis vectors and covectors ∂_i and d^j of Note 26). Here, we cannot expect to find a family of 1-chains in such duality with the $\lambda^n w^e$s. The most likely candidates in this respect, the "small edges" denoted $\{n, e\}$, etc., on Fig. 24.1, left, don't pass, because the matrix of the $\langle \{n', e'\}; \lambda^n w^e \rangle$ is not the identity matrix. If at least this matrix was regular, finding chains in duality with the basis forms, or the other way round, would be straightforward. But regular it is not, because of the relations of Proposition 23.1. We might just omit one small edge out of three on each facet, but this is an ugly solution. Better to reason in terms of *blocks* of DoF of various dimensions, and to be content

with a rearrangement of chains that makes the matrix block-diagonal: Blocks of size 1 for small edges which are part of the "large" ones, blocks of size three for small edges inside the facets. Each of these 3-blocks corresponds to a subspace of dimension *two*, owing to Proposition 23.1, be it the subspace of forms or of chains. The triple of degrees of freedom, therefore, is up to an additive constant. Yet, the circulations[53] do determine the *form*, if not the DoF, uniquely ("unisolvence" property).

The reader will easily guess about "small facets" (16 of them on a single tetrahedron, for a space of dimension $3(F+T) = 3(4+1) = 15$) and "small volumes" (four), in both variants.

Which leaves us with the task of proving the exact sequence property, that is to say, the validity of Poincaré's Lemma in the complex of the W_2^p: Show that $\mathrm{d}b = 0$ for $b \in W_2^p$ implies the existence, locally at least, of $a \in W_2^{p-1}$ such that $b = \mathrm{d}a$. We'll treat the very case this notation suggests, i.e., $p = 2$, and assume trivial topology ("contractible" meshed domain), which does no harm since only a local result is aimed at. We use rot and div rather than d for more clarity. First, two technical points:

LEMMA 24.1. *If $\sum_{n\in\mathcal{N}} \beta_n \lambda^n(x) = \beta_0$ for all x, where the βs are real numbers, then $\beta_n = \beta_0$ for all nodes $n \in \mathcal{N}$.*

PROOF. Clear, since $\sum_n \lambda^n = 1$ is the only relation linking the $\lambda^n(x)$s. □

LEMMA 24.2. *If $a \in W^1$, then $2\,\mathrm{rot}(\lambda^n a) - 3\lambda^n \,\mathrm{rot}\, a \in W^2$.*

PROOF. If $a = w^e$ and $n = f - e$, this results from (23.8). If n is one of the end points of e, e.g., $e = \{m, n\}$, a direct computation, inelegant as it may be, will do: $2\,\mathrm{d}\lambda^n \wedge (\lambda^m\,\mathrm{d}\lambda^n - \lambda^n\,\mathrm{d}\lambda^m) = -2\lambda^n\,\mathrm{d}\lambda^n \wedge \mathrm{d}\lambda^m = \lambda^n\,\mathrm{d}w^e$. □

Now,

PROPOSITION 24.1. *If the W_1^p sequence is exact, the W_2^p sequence is exact.*

PROOF (at level $p = 2$). Suppose $b = b_0 + \sum_{n\in\mathcal{N}} \lambda^n b_n$, with b_0 and all the b_n in W^2, and $\mathrm{div}\, b = 0$. Taking the divergence of the sum and applying Lemma 24.1 in each volume, one sees that $\mathrm{div}\, b_n$ is the same field for all n. So there is some common $\bar{b}$ in W^2 such that $\mathrm{div}(b_n - \bar{b}) = 0$ for all n, and since the W^p complex is exact, there is an a_n in W^1 such that $b_n = \bar{b} + \mathrm{rot}\, a_n$. Hence, $b = b_0 + \bar{b} + \sum_n \lambda^n \,\mathrm{rot}\, a_n$. By Lemma 24.2, there is therefore some $\hat{b}$ in W^2 such that $b = \hat{b} + \frac{2}{3}\,\mathrm{rot}(\sum_n \lambda^n a_n)$. Since $\mathrm{div}\,\hat{b} = 0$, the solenoidal b in W_2^2 we started from is indeed the curl of some element of W_2^1. □

Very little is needed to phrase the proof in such a way that the contractibility assumption becomes moot. Actually, the complexes W_1^p and W_2^p have *the same cohomology*,

[53] Since the matrix has no maximal rank, small-edge circulations must satisfy compatibility conditions for the form to exist. (Indeed, one will easily check that any element of W_2^1 has a null circulation along the chain made by the boundary of a facet minus four times the boundary of the small facet inside it.) This raises a minor problem with the r_m map, whose images need not satisfy this condition. The problem is avoided with a slightly different definition of the small edges (KAMEARI [1999]), as suggested on the right of Fig. 24.1.

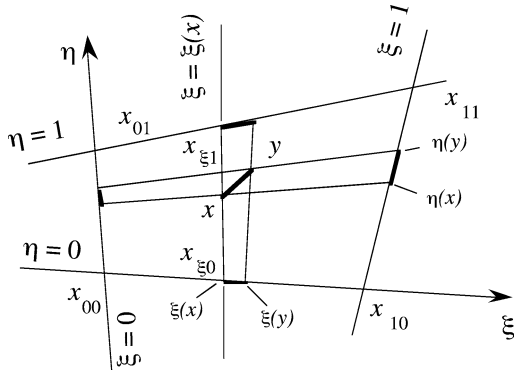

FIG. 25.1. The system of projections, in dimension 2.

whatever the topology of the domain and the culling of passive simplices (i.e., those bearing a null DoF) implied by the boundary conditions.

25. Whitney forms for other shapes than simplices

This simple idea, *approximate p-manifolds by p-chains based on p-cells of the mesh*, is highly productive, as we presently see.

25.1. Hexahedra

First example, the well-known isoparametric element (ERGATOUDIS, IRONS and ZIENKIEWICZ [1968]) on hexahedra can thus be understood. A 2D explanation (Fig. 25.1) will suffice, the generalization being easy. Let us take a convex quadrangle based on points $x_{00}, x_{10}, x_{01}, x_{11}$, and wonder about which weights $w^n(x)$ should be assigned to them (label n designates the generic node) in order to have $x = \sum_{n\in\{00,10,10,11\}} w^n(x)x_n$ in a sensible way. The weights are obvious if x lies on the boundary. For instance, if $x = (1-\xi)x_{00}+\xi x_{10}$, a point we shall denote by $x_{\xi 0}$, weights are $\{1-\xi, \xi, 0, 0\}$. Were it $x \equiv x_{\xi 1} = (1-\xi)x_{01}+\xi x_{11}$, we would take $\{0, 0, 1-\xi, \xi\}$. Now, each x is part of some segment $[x_{\xi 0}x_{\xi 1}]$, for a *unique* value $\xi(x)$ of the weight ξ, in which case $x = (1-\eta)x_{\xi 0} + \eta x_{\xi 1}$, for some $\eta = \eta(x)$, hence it seems natural to distribute the previous weights in the same proportion:

$$\begin{aligned} x = {} & \big(1-\eta(x)\big)\big(1-\xi(x)\big)x_{00} + \big(1-\eta(x)\big)\xi(x)x_{10} \\ & + \eta(x)\big(1-\xi(x)\big)x_{01} + \eta(x)\xi(x)\,x_{11}, \end{aligned} \tag{25.1}$$

and we are staring at the basis functions. They form, obviously, a partition of unity.

Looking at what we have done, and generalizing to dimension 3 or higher, we notice a *system of projections*, associated with a trilinear [54] *chart*, $x \to \{\xi(x), \eta(x), \zeta(x)\}$, from

[54] Thus called because ξ, η, and ζ, though cubic polynomials in terms of the Cartesian coordinates of x, are affine functions of each of them, taken separately.

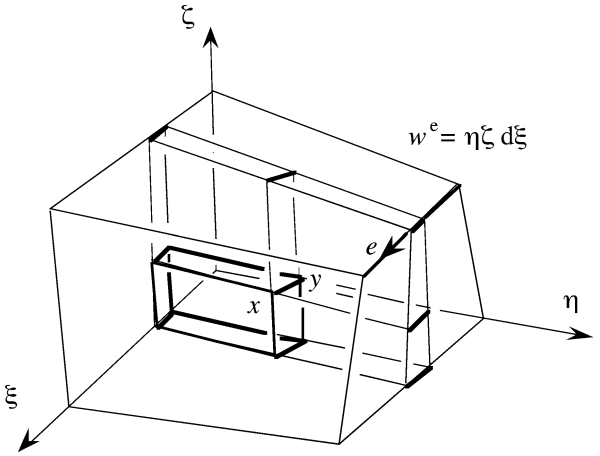

FIG. 25.2. Weight $w^e(xy)$ is the $\xi\eta\zeta$-volume of the "hinder region" of xy with respect to edge e.

a hexahedron to the unit cube in $\xi\eta\zeta$-space. The successive projections (which can be performed in any order) map a point $x \equiv x_{\xi\eta\zeta}$ to its images $x_{0\eta\zeta}$ and $x_{1\eta\zeta}$ on opposite facets[55] $\xi = 0$ and $\xi = 1$, then, recursively, send these images to points on opposite edges, etc., until eventually a node n is reached. In the process, the weight $\langle x; w^n\rangle$ of x is recursively determined by formulas such as (assuming for the sake of the example that n belongs to the facet $\xi = 0$)

$$\langle x_{\xi\eta\zeta}; w^n\rangle = (1-\xi)\langle x_{0\eta\zeta}; w^n\rangle.$$

The final weight of x with respect to n is thus the product of factors, such as here $(1-\xi)$, collected during the projection process. (They measure the relative proximity of each projection to the face towards which next projection will be done.) The last factor in this product is 1, obtained when the projection reaches n. Observe the fact, essential of course, that whatever the sequence of projections, the partial weights encountered along the way are the same, only differently ordered, and hence the weight of x with respect to node n is a well-defined quantity.

The viewpoint thus adopted makes the next move obvious. Now, instead of a point x, we deal with a vector v at x, small enough for the segment xy (where $y = x + v$) to be contained in a single hexahedron. The above projections send x and y to facets, edges, etc. Ending the downward recursion one step higher than previously, at the level of edges, we get projections $x_e y_e$ of xy onto all edges e. The weight $\langle xy; w^e\rangle$ is the product of weights of x collected along the way, but the last factor is now the algebraic ratio $x_e y_e/e$ (which makes obvious sense) instead of 1. Hence the analytical expression of the corresponding Whitney form, for instance, in the case of Fig. 25.2, $w^e = \eta\zeta\, d\xi$. (Notice the built-in "partition of unity" property: $xy = \Sigma_e\langle xy; w^e\rangle e$.) The proxies, $w^e = \eta\zeta\nabla\xi$ in this example, were proposed as edge elements for hexahedra by VAN WELIJ [1985].

[55] Be aware that p-faces need not be "flat", i.e., lie within an affine p-subspace for $p > 1$, in dimension higher than 2. To avoid problems this would raise, we assume here a mesh generation which enforces this extra requirement.

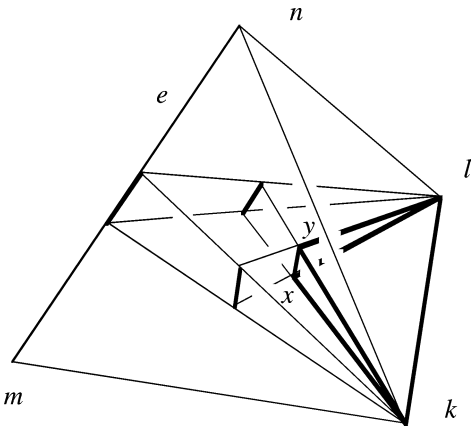

FIG. 25.3. There too, weight $w^e(xy)$ is the relative volume of the hinder region.

One may wonder whether weights such as $\langle xy; w^e\rangle$ have a geometric interpretation there too. They do: $\langle xy; w^e\rangle$ is the relative volume, in the *reference hexahedron*[56] $H = \{\xi, \eta, \zeta\}$: $0 \leqslant \xi \leqslant 1$, $0 \leqslant \eta \leqslant 1$, $0 \leqslant \zeta \leqslant 1$, of the "hinder region" of Fig. 25.2, made of points "behind" xy with respect to edge e. This may seem fairly different from the situation in Fig. 23.4, middle, but a suitable reinterpretation of the system of projections in the tetrahedron (Fig. 25.3) shows the analogy.

A similar reasoning gives facet elements: the last weight, for a small triangle xyz, is $x_f y_f z_f / f$, which again makes sense: Take the ratio of the areas (an affine notion) of the images of these surfaces in the reference cube, with sign $+$ if orientations of $x_f y_f z_f$ and f match, $-$ otherwise. Whitney forms such as $w^f = \xi\, \mathrm{d}\eta\, \mathrm{d}\zeta$ (when f is the facet $\xi = 1$) result. The proxy of that particular one is $\xi \nabla\eta \times \nabla\zeta$.

25.2. Prisms

So, Cartesian coordinates and barycentric coordinates provide two systems of projections which make obvious the weight allocation. These systems can be mixed: one of them in use for $p < n$ dimensions, the other one for the $n - p$ remaining dimensions. In dimension 3, this gives only one new possibility, the prism (Fig. 25.4).

Such a variety of shapes makes the mesh generation more flexible (DULAR, HODY, NICOLET, GENON and LEGROS [1994]). Yet, do the elements of a given degree, edge elements say, fit together properly when one mixes tetrahedra, hexahedra, and prisms? Yes, because of the recursivity of the weight allocation: If a segment xy lies entirely in the facet common to two volumes of different kind, say a tetrahedron and a prism, the weights $\langle xy; w^e\rangle$ for edges belonging to this facet only depend on what happens in the facet, i.e., they are the same as evaluated with both formulas for w^e, the one valid in the tetrahedron, the one valid in the prism. This is enough to guarantee the *tangential continuity* of such composite edge elements.

[56]Recall that all tetrahedra are affine equivalent, which is why we had no need for a reference one. The situation is different with hexahedra, which form several orbits under the action of the affine group.

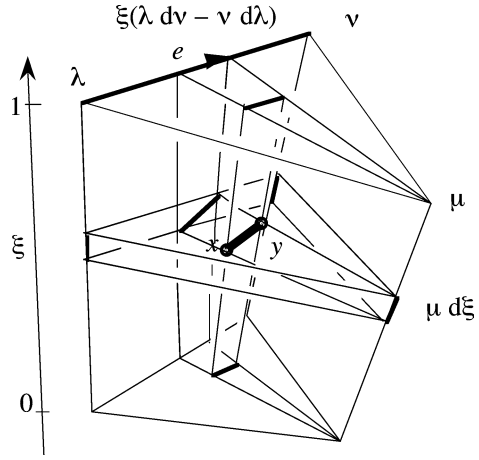

FIG. 25.4. Projective system and edge elements for a prism. Observe the commutativity of the projections.

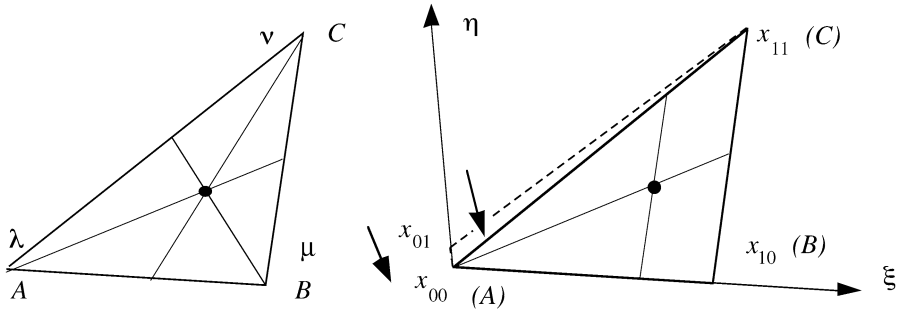

FIG. 25.5. Projective systems for the same triangle, in the barycentric coordinates on the left, and by degeneracy of the quadrilateral system on the right.

25.3. "Degeneracies"

Yet one may yearn for even more flexibility, and edge elements for *pyramids* have been proposed (COULOMB, ZGAINSKI and MARÉCHAL [1997], GRADINARU and HIPTMAIR [1999]). A systematic way to proceed, in this respect, is to recourse to "degenerate" versions of the hexahedron or the prism, obtained by fusion of one or more pair of nodes and or edges.

To grasp the idea, let's begin with the case of the degenerated quadrilateral, in two dimensions (Fig. 25.5). With the notations of the figure, where $\{\lambda, \mu, \nu\}$ are the barycentric coordinates in the left triangle, the map $\{\mu, \nu\} \to \{\eta, \xi\}$, where $\eta = \nu/(\mu + \nu)$ and $\xi = \mu + \nu$, sends the interior of the triangle to the interior of the right quadrilateral. When, by deformation of the latter, x_{10} merges with x_{00}, the projective system of the quadrilateral generates a new projective system on the triangle.

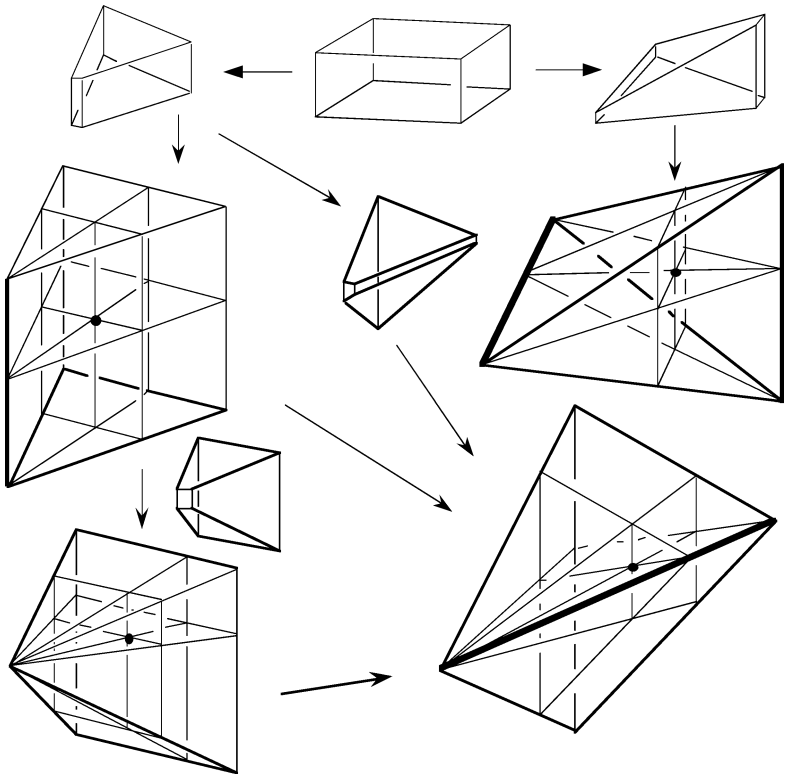

FIG. 25.6. Projective systems in four degenerations of the hexahedron. Thick lines indicate the merged edges.

The weights assigned to the nodes, and hence the nodal elements, are the same in both systems, for $\xi\eta = \nu$ for point C (cf. (25.1)), $\xi(1-\eta) = \mu$ for B, and the sum $(1-\xi)(1-\eta) + (1-\xi)\eta$, attributed to A by adding the loads of x_{00} and x_{01}, does equal λ. But the edge elements differ: For AC, $\eta\,\mathrm{d}\xi \equiv -(1-\lambda)^{-1}\mu\,\mathrm{d}\lambda$ on the right instead of $\lambda\,\mathrm{d}\nu - \nu\,\mathrm{d}\lambda$ on the left, $-(1-\lambda)^{-1}\mu\,\mathrm{d}\lambda$ for AB, and $\mathrm{d}\nu + (1-\lambda)^{-1}\nu\,\mathrm{d}\lambda$ for BC. (The singularity of shape functions at point A is never a problem, because integrals where they appear always converge.)

In dimension 3, the principle is the same: When two edges merge, by degeneration of a hexahedron or of a prism, the Whitney form of the merger is the sum of the Whitney forms of the two contributors, which one may wish to rewrite in a coordinate system adapted to the degenerate solid. Figs. 25.6 and 25.7 show seven degeneracies, all those that one can obtain from a hexahedron or a prism with plane facets under the constraint of not creating curved facets in the process. As one sees, the only novel shape is the pyramid, while the prism is retrieved once and the tetrahedron four times.

But, as was predictible from the 2-dimensional case, it's *new* Whitney forms, on these solids, that are produced by the merging, because the projection systems are different. In particular, we have now *five* distinct projective systems on the tetrahedron (and two on the pyramid and the prism), and the equality of traces is not automatic any longer. One

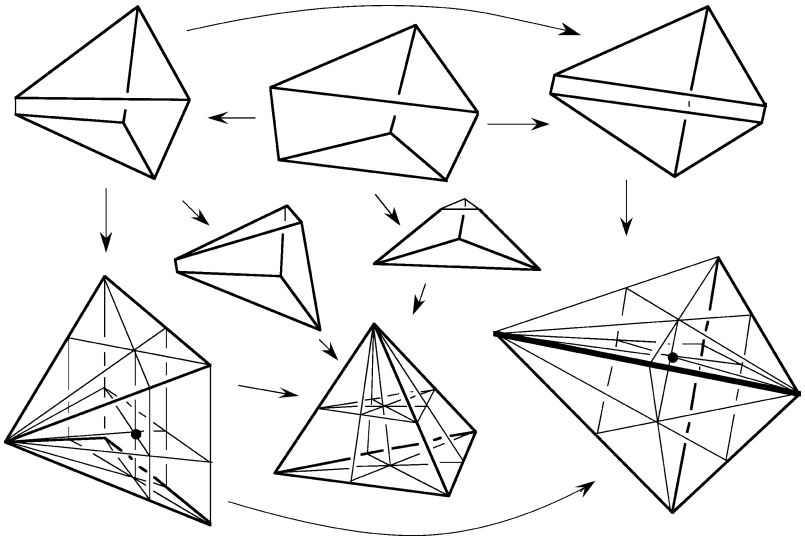

FIG. 25.7. Projective systems in three degenerations of the prism. Note how the pyramid has two ways to degenerate towards the tetrahedron.

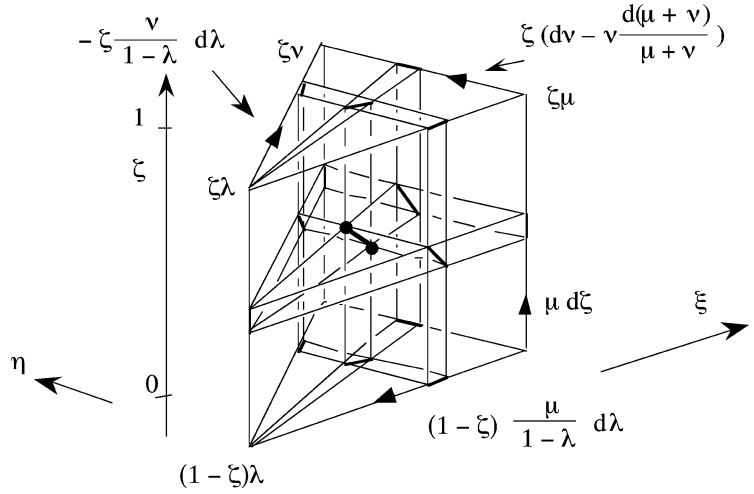

FIG. 25.8. Nodal and edge elements for the projective system of Fig. 25.5. One passes from the previous coordinate system $\{\xi, \eta, \zeta\}$ to the prism-adapted $\{\zeta, \lambda, \mu, \nu\}$ system by the formulas $\xi = \mu + \nu,\ \eta = \nu/(\mu + \nu)$, with $\lambda + \mu + \nu = 1$.

must therefore care about correct assembly, in order to get the same projection system on each facet.

The advantage of having the pyramid available is thus marred by the necessity of an extended shape-functions catalogue (on at least two triangular facets of a pyramid, the projection system cannot match the tetrahedron's one), and by the existence of cum-

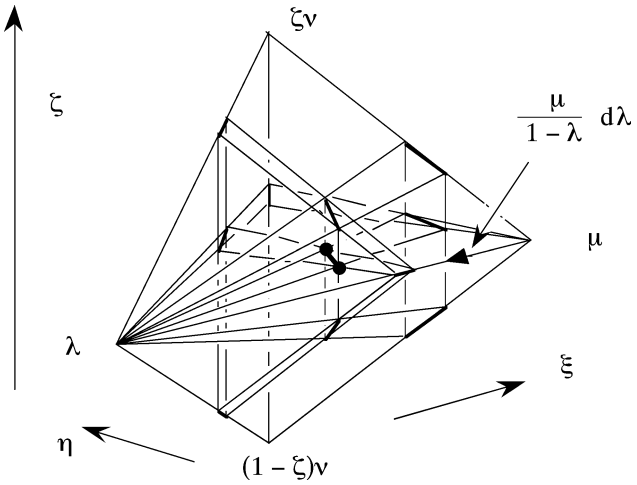

FIG. 25.9. Degeneration of the prism of Fig. 25.8. Two edges disappear, and a new edge element, $\mu(1-\lambda)^{-1}\,\mathrm{d}\lambda$ is created by the merging. The coordinate system is the same here as in Fig. 25.8, so $\{\lambda, \mu, \nu\}$ should not be confused with barycentric coordinates of this tetrahedron. Denoting the latter by $\{\bar\kappa, \bar\lambda, \bar\mu, \bar\nu\}$, and using the formulas $\nu = \bar\nu + \bar\kappa$ and $\zeta = \bar\nu/(\bar\nu + \bar\kappa)$, one has $\xi = \bar\mu + \bar\nu + \bar\kappa = 1 - \bar\lambda$, $\eta = (\bar\nu + \bar\kappa)/(1 - \bar\lambda)$. Thus, for instance, the shape function $\mu(1-\lambda)^{-1}\,\mathrm{d}\lambda$ rewrites as $\bar\mu(1-\bar\lambda)^{-1}\,\mathrm{d}\bar\lambda$ in barycentric coordinates.

bersome assembly rules. Yet, finding the new shape-functions is not too difficult, as exemplified by Figs. 25.8 and 25.9.

25.4. Star-shaped cells, dual cells

Let's end all this by an indication on how to build Whitney forms on any star-shaped polyhedron.

Suppose each p-cell of the mesh m, for all p, has been provided with a "center", in the precise sense of Section 15, i.e., a point with respect to which the cell is star-shaped. Then, join the centers in order to obtain a simplicial refinement, $\overline{m}$ say, where the new sets of p-simplices are $\overline{\mathcal{S}}_p$, the old sets of cells being $\mathcal{S}_p$. In similar style, let $\mathbf{u}$ and $\overline{\mathbf{u}}$ stand for DoF arrays indexed over $\mathcal{S}_p$ and $\overline{\mathcal{S}}_p$ respectively, with the compatibility relation $\mathbf{u}_s = \Sigma_{s'} \pm \overline{\mathbf{u}}_{s'}$ for all s in $\mathcal{S}_p$, the sum running over all small simplices in the refinement of cell s, and the signs taking care of relative orientations. To define $p_m\mathbf{u}$, knowing what $p_{\bar m}\overline{\mathbf{u}}$ is, we just take the *smallest*, in the energy norm, of the $p_{\bar m}\overline{\mathbf{u}}$'s, with respect to all $\overline{\mathbf{u}}$'s compatible with $\mathbf{u}$.

The family of interpolants thus obtained is to the cellular mesh, for all purposes, what Whitney forms were to a simplicial mesh. Whether they deserve to be called "Whitney forms" is debatable, however, because they are metric-dependent, unlike the standard Whitney forms. The same construction on the dual side provides similar pseudo-Whitney forms on the dual mesh. (More precisely, there is, as we have observed at the end of Section 15, a common simplicial refinement of both m and $\widetilde{m}$. The process just defined constructs forms on both, but it's easy to check that the pseudo-Whitneys on the

primal mesh are just the Whitney forms.) This fills a drawer in the toolkit, the emptiness of which we took some pain to hide until now, although it was conspicuous at places, on Fig. 23.1, for instance.

References

ALBANESE, R., RUBINACCI, G. (1988). Integral formulations for 3-D eddy-currents computation using edge-elements. *IEE Proc. A* **135**, 457–462.

ARMSTRONG, M.A. (1979). *Basic Topology* (McGraw-Hill, London).

ARNOLD, D.N., BREZZI, F. (1985). Mixed and non-conforming finite element methods: implementation, postprocessing and error estimates. M^2AN **19**, 7–32.

BABUŠKA, I., AZIZ, A.K. (1976). On the angle condition in the finite element method. *SIAM J. Numer. Anal.* **13**, 214–226.

BAEZ, J., MUNIAIN, J.P. (1994). *Gauge Fields, Knots and Gravity* (World Scientific, Singapore).

BALDOMIR, D., HAMMOND, P. (1996). *Geometry of Electromagnetic Systems* (Oxford Univ. Press, Oxford).

BANK, R.E., ROSE, D.J. (1987). Some error estimates for the box method. *SIAM J. Numer. Anal.* **24**, 777–787.

BÄNSCH, E. (1991). Local mesh refinement in 2 and 3 dimensions. *Impact Comput. Sci. Engrg.* **3**, 181–191.

BEY, J. (1995). Tetrahedral grid refinement. *Computing* **55**, 355–378.

BOSSAVIT, A. (1990a). Eddy-currents and forces in deformable conductors. In: Hsieh, R.K.T. (ed.), *Mechanical Modellings of New Electromagnetic Materials: Proc. IUTAM Symp., Stockholm, April 1990* (Elsevier, Amsterdam), pp. 235–242.

BOSSAVIT, A. (1990b). Solving Maxwell's equations in a closed cavity, and the question of spurious modes. *IEEE Trans. Magn.* **26**, 702–705.

BOSSAVIT, A. (1996). A puzzle. *ICS Newsletter* **3** (2), 7; **3** (3), 14; **4** (1) (1997) 17–18.

BOSSAVIT, A. (1998a). *Computational Electromagnetism* (Academic Press, Boston).

BOSSAVIT, A. (1998b). Computational electromagnetism and geometry. *J. Japan Soc. Appl. Electromagn. Mech.* **6**, 17–28, 114–123, 233–240, 318–326; **7** (1999) 150–159, 249–301, 401–408; **8** (2000) 102–109, 203–209, 372–377.

BOSSAVIT, A. (1999). On axial and polar vectors. *ICS Newsletter* **6**, 12–14.

BOSSAVIT, A. (2000). Most general 'non-local' boundary conditions for the Maxwell equations in a bounded region. *COMPEL* **19**, 239–245.

BOSSAVIT, A. (2001a). 'Stiff' problems in eddy-current theory and the regularization of Maxwell's equations. *IEEE Trans. Magn.* **37**, 3542–3545.

BOSSAVIT, A. (2001b). On the notion of anisotropy of constitutive laws: some implications of the 'Hodge implies metric' result. *COMPEL* **20**, 233–239.

BOSSAVIT, A. (2001c). On the representation of differential forms by potentials in dimension 3. In: van Rienen, U., Günther, M., Hecht, D. (eds.), *Scientific Computing in Electrical Engineering* (Springer-Verlag, Berlin), pp. 97–104.

BOSSAVIT, A. (2003). Mixed-hybrid methods in magnetostatics: complementarity in one stroke. *IEEE Trans. Magn.* **39**, 1099–1102.

BOSSAVIT, A., KETTUNEN, L. (1999). Yee-like schemes on a tetrahedral mesh, with diagonal lumping. *Int. J. Numer. Modelling* **12**, 129–142.

BRANIN JR., F.H. (1961). An abstract mathematical basis for network analogies and its significance in physics and engineering. *Matrix and Tensor Quarterly* **12**, 31–49.

BURKE, W.L. (1985). *Applied Differential Geometry* (Cambridge Univ. Press, Cambridge).

DI CARLO, A., TIERO, A. (1991). The geometry of linear heat conduction. In: Schneider, W., Troger, H., Ziegler, F. (eds.), *Trends in Applications of Mathematics to Mechanics* (Longman, Harlow), pp. 281–287.

CARPENTER, C.J. (1977). Comparison of alternative formulations of 3-dimensional magnetic-field and eddy-current problems at power frequencies. *Proc. IEE* **124**, 1026–1034.

CHAVENT, G., ROBERTS, J.E. (1991). A unified presentation of mixed, mixed-hybrid finite elements and standard finite difference approximations for the determination of velocities in waterflow problems. *Adv. Water Resources* **14**, 329–348.

CLEMENS, M., WEILAND, T. (1999). Transient eddy-current calculation with the FI-method. *IEEE Trans. Magn.* **35**, 1163–1166.

COHEN, G., JOLY, P., TORDJMAN, N. (1993). Construction and analysis of higher order elements with mass-lumping for the wave equation. In: *Mathematical Aspects of Wave Propagation Phenomena* (SIAM, Philadelphia), pp. 152–160.

COSTABEL, M., DAUGE, M. (1997). Singularités des équations de Maxwell dans un polyèdre. *C. R. Acad. Sci. Paris I* **324**, 1005–1010.

DE COUGNY, H.L., SHEPHARD, M.S. (1999). Parallel refinement and coarsening of tetrahedral meshes. *Int. J. Numer. Meth. Engng.* **46**, 1101–1125.

COULOMB, J.L., ZGAINSKI, F.X., MARÉCHAL, Y. (1997). A pyramidal element to link hexahedral, prismatic and tetrahedral edge finite elements. *IEEE Trans. Magn.* **33**, 1362–1365.

COURBET, B., CROISILLE, J.P. (1998). Finite volume box schemes on triangular meshes. M^2AN **32**, 631–649.

VAN DANTZIG, D. (1934). The fundamental equations of electromagnetism, independent of metrical geometry. *Proc. Cambridge Phil. Soc.* **30**, 421–427.

VAN DANTZIG, D. (1954). On the geometrical representation of elementary physical objects and the relations between geometry and physics. *Nieuw. Archief vor Wiskunde* **3**, 73–89.

DIRICHLET, G.L. (1850). Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. *J. Reine Angew. Math.* **40**, 209.

DULAR, P., HODY, J.-Y., NICOLET, A., GENON, A., LEGROS, W. (1994). Mixed finite elements associated with a collection of tetrahedra, hexahedra and prisms. *IEEE Trans. Magn.* **30**, 2980–2983.

EBELING, F., KLATT, R., KRAWCZYK, F., LAWINSKY, E., WEILAND, T., WIPF, S.G., STEFFEN, B., BARTS, T., BROWMAN, M.J., COOPER, R.K., DEAVEN, H., RODENZ, G. (1989). The 3-D MAFIA group of electromagnetic codes. *IEEE Trans. Magn.* **25**, 2962–2964.

ECKMANN, B. (1999). Topology, algebra, analysis – relations and missing links. *Notices AMS* **46**, 520–527.

ELMKIES, A., JOLY, P. (1997). Éléments finis d'arête et condensation de masse pour les équations de Maxwell: le cas de dimension 3. *C. R. Acad. Sci. Paris Sér. I* **325**, 1217–1222.

ERGATOUDIS, J.G., IRONS, B.M., ZIENKIEWICZ, O.C. (1968). Curved, isoparametric, 'quadrilateral' elements for finite element analysis. *Int. J. Solids Struct.* **4**, 31–42.

FIRESTONE, F.A. (1933). A new analogy between mechanical and electrical systems. *J. Acoust. Soc. Am.* **0**, 249–267.

GALLOUET, T., VILA, J.P. (1991). Finite volume element scheme for conservation laws of mixed type. *SIAM J. Numer. Anal.* **28**, 1548–1573.

GELBAUM, B.R., OLMSTED, J.M.H. (1964). *Counterexamples in Analysis* (Holden-Day, San Francisco).

GOLDHABER, A.S., TROWER, W.P. (1990). Resource letter MM-1: Magnetic monopoles. *Am. J. Phys.* **58**, 429–439.

GRADINARU, V., HIPTMAIR, R. (1999). Whitney elements on pyramids. *ETNA* **8**, 154–168.

HALMOS, P.R. (1950). *Measure Theory* (Van Nostrand, Princeton).

HAMOUDA, L., BANDELIER, B., RIOUX-DAMIDAU, F. (2001). Mixed formulation for magnetostatics. In: *Proc. Compumag* (paper PE4–11).

HARRISON, J. (1998). Continuity of the integral as a function of the domain. *J. Geometric Anal.* **8**, 769–795.

HAUGAZEAU, Y., LACOSTE, P. (1993). Condensation de la matrice masse pour les éléments finis mixtes de $H(\mathrm{rot})$. *C. R. Acad. Sci. Paris I* **316**, 509–512.

HEINRICH, B. (1987). *Finite Difference Methods on Irregular Networks* (Akademie-Verlag, Berlin).

HENLE, A. (1994). *A Combinatorial Introduction to Topology* (Dover, New York).

HILTON, P.J., WYLIE, S. (1965). *Homology Theory, An Introduction to Algebraic Topology* (Cambridge Univ. Press, Cambridge).

HUANG, J., XI, S. (1998). On the finite volume element method for general self-adjoint elliptic problems. *SIAM J. Numer. Anal.* **35**, 1762–1774.

HYMAN, J.M., SHASHKOV, M. (1997). Natural discretizations for the divergence, gradient, and curl on logically rectangular grids. *Comput. Math. Appl.* **33**, 81–104.

JÄNICH, K. (2001). *Vector Analysis* (Springer, New York).

KAASSCHIETER, E.F., HUIJBEN, A.J.M. (1992). Mixed-hybrid finite elements and streamline computation for the potential flow problem. *Numer. Meth. PDE* **8**, 221–266.

KAMEARI, A. (1999). Symmetric second order edge elements for triangles and tetrahedra. *IEEE Trans. Magn.* **35**, 1394–1397.

KHEYFETS, A., WHEELER, J.A. (1986). Boundary of a boundary and geometric structure of field theories. *Int. J. Theor. Phys.* **25**, 573–580.

KOENIG, H.E., BLACKWELL, W.A. (1960). Linear graph theory: a fundamental engineering discipline. *IRE Trans. Edu.* **3**, 42–49.

KOTTLER, F. (1922). Maxwell'sche Gleichungen und Metrik. *Sitzungber. Akad. Wien IIa* **131**, 119–146.

LAX, P.D., RICHTMYER, R.D. (1956). Survey of the stability of linear finite difference equations. *Comm. Pure Appl. Math.* **9**, 267–293.

LEE, J.-F., SACKS, Z. (1995). Whitney elements time domain (WETD) methods. *IEEE Trans. Magn.* **31**, 1325–1329.

LEIS, R. (1968). Zur Theorie elektromagnetischer Schwingungen in anisotropen inhomogenen Medien. *Math. Z.* **106**, 213–224.

MADSEN, I., TORNEHAVE, J. (1997). *From Calculus to Cohomology* (Cambridge Univ. Press, Cambridge).

MATTIUSSI, C. (2000). The finite volume, finite element, and finite difference methods as numerical methods for physical field problems. *Adv. Imag. Electron Phys.* **113**, 1–146.

MAUBACH, J.M. (1995). Local bisection refinement for N-simplicial grids generated by reflection. *SIAM J. Sci. Stat.* **16**, 210–227.

MITTRA, R., RAMAHI, O., KHEBIR, A., GORDON, R., KOUKI, A. (1989). A review of absorbing boundary conditions for two and three-dimensional electromagnetic scattering problems. *IEEE Trans. Magn.* **25**, 3034–3038.

MONK, P., SÜLI, E. (1994). A convergence analysis of Yee's scheme on nonuniform grids. *SIAM J. Numer. Anal.* **31**, 393–412.

MOSÉ, R., SIEGEL, P., ACKERER, P., CHAVENT, G. (1994). Application of the mixed hybrid finite element approximation in a groundwater flow model: Luxury or necessity?. *Water Resources Res.* **30**, 3001–3012.

MUNTEANU, I. (2002). Tree-cotree condensation properties. *ICS Newsletter* **9**, 10–14.

NICOLAIDES, R., WANG, D.-Q. (1998). Convergence analysis of a covolume scheme for Maxwell's equations in three dimensions. *Math. Comp.* **67**, 947–963.

POST, E.J. (1972). The constitutive map and some of its ramifications. *Ann. Phys.* **71**, 497–518.

RAPETTI, F., DUBOIS, F., BOSSAVIT, A. (2002). Integer matrix factorization for mesh defect detection. *C. R. Acad. Sci. Paris* **334**, 717–720.

REN, Z. (1996). 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)*, pp. 119–124.

DE RHAM, G. (1936). Relations entre la topologie et la théorie des intégrales multiples. *L'Enseignement Math.* **35**, 213–228.

DE RHAM, G. (1960). *Variétés différentiables* (Hermann, Paris).

ROSEN, J. (1973). Transformation properties of electromagnetic quantities under space inversion, time reversal, and charge conjugation. *Am. J. Phys.* **41**, 586–588.

RUDIN, W. (1973). *Functional Analysis* (McGraw-Hill, New York).

SCHATZ, A.H., SLOAN, I.H., WAHLBIN, L.B. (1996). Superconvergence in finite element methods and meshes that are locally symmetric with respect to a point. *SIAM J. Numer. Anal.* **33**, 505–521.

SCHOUTEN, J.A. (1989). *Tensor Analysis for Physicists* (Dover, New York).

SCHUTZ, B. (1980). *Geometrical Methods of Mathematical Physics* (Cambridge Univ. Press, Cambridge).

SEIFERT, H., THRELFALL, W. (1980). *A Textbook of Topology* (Academic Press, Orlando) (first German ed., 1934).

SHAW, R., YEADON, F.J. (1989). On $(a \times b) \times c$. *Am. Math. Monthly* **96**, 623–629.

SMYTH, J.B., SMYTH, D.C. (1977). Critique of the paper 'The electromagnetic radiation from a finite antenna'. *Am. J. Phys.* **45**, 581–582.

SORKIN, R. (1975). The electromagnetic field on a simplicial net. *J. Math. Phys.* **16**, 2432–2440.

SÜLI, E. (1991). Convergence of finite volume schemes for Poisson's equation on nonuniform meshes. *SIAM J. Numer. Anal.* **28**, 1419–1430.
TAYLOR, E.F., WHEELER, J.A. (1992). *Spacetime Physics* (Freeman, New York).
TEIXEIRA, F.L., CHEW, W.C. (1999). Lattice electromagnetic theory from a topological viewpoint. *J. Math. Phys.* **40**, 169–187.
TONTI, E. (1996). On the geometrical structure of electromagnetism. In: Ferrarese, G. (ed.), *Gravitation, Electromagnetism and Geometrical Structures* (Pitagora, Bologna), pp. 281–308.
TONTI, E. (2001). A direct formulation of field laws: the cell method. *CMES* **2**, 237–258.
TRAPP, B., MUNTEANU, I., SCHUHMANN, R., WEILAND, T., IOAN, D. (2002). Eigenvalue computation by means of a tree–cotree filtering technique. *IEEE Trans. Magn.* **38**, 445–448.
UMAN, M.A. (1977). Reply to Smyth and Smyth. *Am. J. Phys.* **45**, 582.
VEBLEN, O., WHITEHEAD, J.H.C. (1932). *The Foundations of Differential Geometry* (Cambridge Univ. Press, Cambridge).
WEILAND, T. (1992). Maxwell's grid equations. In: *Proc. URSI Int. Symp. Electromagnetic Theory, Sydney*, pp. 37–39.
WEILAND, T. (1996). Time domain electromagnetic field computation with finite difference methods. *Int. J. Numer. Modelling* **9**, 295–319.
WEILAND, T. (1985). Three dimensional resonator mode computation by finite difference methods. *IEEE Trans. Magn.* **21**, 2340–2343.
WEISER, A., WHEELER, M.F. (1988). On convergence of block-centered finite differences for elliptic problems. *SIAM J. Numer. Anal.* **25**, 351–375.
VAN WELIJ, J.S. (1985). Calculation of eddy currents in terms of H on hexahedra. *IEEE Trans. Magn.* **21**, 2239–2241.
WHITE, D.A., KONING, J.M. (2000). A novel approach for computing solenoidal eigenmodes of the vector Helmholtz equation. CEFC'00 (p. 328 of the "digest of abstracts").
WHITNEY, H. (1957). *Geometric Integration Theory* (Princeton Univ. Press, Princeton).
YEE, K.S. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans. AP* **14**, 302–307.
YOSIDA, K. (1980). *Functional Analysis* (Springer-Verlag, Berlin) (first ed., 1965).

Further reading

DODZIUK, J. (1976). Finite-difference approach to the Hodge theory of harmonic forms. *Amer. J. Math.* **98**, 79–104.
FRANKEL, T. (1997). *The Geometry of Physics, An Introduction* (Cambridge Univ. Press, Cambridge).
HIPTMAIR, R. (2001). Discrete Hodge operators. *Progress in Electromagnetics Research* **32**, 122–150.
KOTIUGA, P.R. (1984). Hodge decompositions and computational electromagnetics. Thesis (Department of Electrical Engng., McGill University, Montréal).
MAXWELL, J.C. (1864). On reciprocal figures and diagrams of forces. *Phil. Mag. Ser.* **4**, 250–261.
VON MISES, R. (1952). On network methods in conformal mapping and in related problems. In: Appl. Math. Series **18** (US Department of Commerce, NBS), pp. 1–6.
MÜLLER, W. (1978). Analytic torsion and R-torsion of Riemannian manifolds. *Adv. Math.* **28**, 233–305.
NEDELEC, J.C. (1980). Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.* **35**, 315–341.
POST, E.J. (1979). Kottler–Cartan–van Dantzig (KCD) and noninertial systems. *Found. Phys.* **9**, 619–640.
POST, E.J. (1984). The metric dependence of four-dimensional formulations of electromagnetism. *J. Math. Phys.* **25**, 612–613.
REN, Z., IDA, N. (2002). High-order elements of complete and incomplete bases in electromagnetic field computation. *IEE Proc. Science, Measurement and Technology* **149**, 147–151.
SILVESTER, P., CHARI, M.V.K. (1970). Finite element solution of saturable magnetic field problems. *IEEE Trans. PAS* **89**, 1642–1651.
TAFLOVE, A. (1995). *Computational Electromagnetics: The Finite-Difference Time-Domain Method* (Artech House, Boston).

XIANG, YOU-QING, ZHOU, KE-DING, LI, LANG-RU (1989). A new network-field model for numerical analysis of electromagnetic field. In: Shunnian, Ding (ed.), *Electromagnetic Fields in Electrical Engineering: Proc. BISEF'88, October 19–21, Beijing* (Pergamon Press, Oxford), pp. 391–398.

YIOULTSIS, T.V., TSIBOUKIS, T.D. (1997). Development and implementation of second and third order vector finite elements. *IEEE Trans. Magn.* **33**, 1812–1815.