Chapter 1

Maxwell in Chains

The purpose of this chapter is to give an overview of the subject along with a pedagogical introduction to the formalism used. The discussion will be informal. Readers well versed in exterior calculus may wish to read this chapter very cursorily or skip it altogether.

1 Atomistic picture: chains

All of the present chapter will be based on the reasonable *assumption* that the "continuous can be approximated by the discrete". We will thus take an atomistic point of view, replacing smooth configurations of the electromagnetic field by spiky ones, not because we believe that the fundamental constituents of the field are discrete, but because it serves the purpose of better visualization. Such an approach is legitimate whenever we can get arbitrarily close to the true configuration by refining the discretization. This, then, is what you must bear in mind when contemplating the pictures shown in this chapter and in much of the book: that they are to be refined and that ultimately a limit is to be taken to recover the continuum.

Having stated the rules of the game (approximation of the continuum by the discrete, with a continuum limit to be taken at the end), we will now introduce the discrete objects that allow to present electrodynamics and the laws governing it in an intuitive way.

1.1 Chains

Imagine some set of electric charges q_i $(i = 1, \ldots, N)$ located at points p_i in space. It will turn out to be useful (for book keeping and other purposes) to form the linear combination of the points, each weighted by its charge:

$$\rho = \sum_{i=1}^{N} q_i p_i \ .$$

Such a formal sum is called a *chain of degree* 0, or 0–chain for short. "Degree 0" signals that the objects being combined – the points – are zero-dimensional. The 0–chain ρ at hand is called the *electric charge density*.

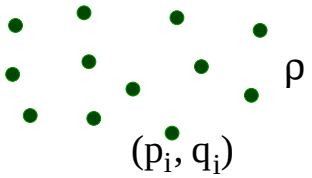

Figure 1.1: A set of point charges determine a 0–chain.

More generally, a 0–chain is any finite sum of points weighted by real numbers. 0–chains can be added and multiplied by numbers $\lambda \in \mathbb{R}$, so they form a real vector space (of infinite dimension).

Next, consider a stationary flow of electric current. In the spirit of the "atomistic picture", such a flow can be decomposed into lines or curves γ_i (the trajectories of the charged particles in the stationary flow), and each line carries an electric current I_i. The lines γ_i come with a sense of direction indicated by an arrow; see Fig. 1.2. If $I_i > 0$ the current associated with γ_i flows in the direction of the arrow on γ_i, if $I_i < 0$ it flows in the opposite direction. The formal sum of the flow lines weighted by their currents is called the *electric current density*, j:

$$j = \sum_{i=1}^{N} I_i \gamma_i \ .$$

It is an example of a 1–*chain*. More generally, a 1–chain is any finite linear combination of lines, or line segments, weighted by real numbers.

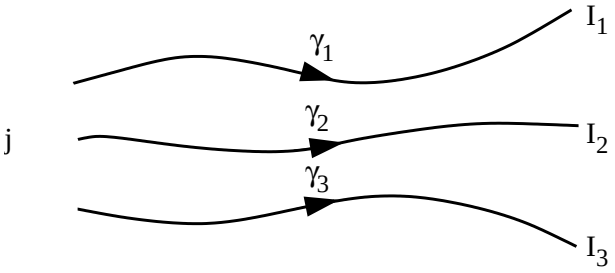

Figure 1.2: 1–chain model of a stationary current flow.

It should now be clear how to continue. A 2–*chain* is any finite linear combination of surfaces, or pieces of surface; a 3–*chain* is any finite linear combination of 3–dimensional domains in space; and so on. A k–chain is also called a chain of degree k. In d dimensions chains exist up to degree d. Chains with a fixed degree k can be added and multiplied by real numbers, so they form real vector spaces.

Before continuing, let me address one immediate objection you might have. You might say: "I thought the charge density was given by a scalar function, and I don't really see how my charge density function relates to your discrete model of the charge density as a 0–chain, or, for that matter, why you have any right to call that 0–chain $\sum_i q_i p_i$ the charge 'density' ρ !"

First of all, the 0–chain model for ρ is not just a discrete approximation but is in fact an accurate representation of reality within classical physics (which is of course modified by quantum mechanics). To the best of our knowledge, electric charge is concentrated in point particles carrying one fundamental unit of charge each. Second, even if the electric charge was distributed continuously, the charge density still wouldn't be a function! Under a space transformation, say ψ, a function $p \mapsto f(p)$ transforms into $p \mapsto f\big(\psi(p)\big)$. This is not the way the charge density "function" $p \mapsto \chi(p)$ transforms. The correct expression for the transformed χ involves the Jacobian $|\mathrm{Det}\, \psi'|$ of ψ, and is given by

$$p \mapsto \frac{\chi\big(\psi^{-1}(p)\big)}{|\mathrm{Det}\, \psi'\big(\psi^{-1}(p)\big)|} \,.$$

Indeed, if ψ is a scale transformation that contracts space, the charge density gets bigger (as the charges move closer together), and this is true even if the untransformed charge density is homogeneous, i.e. $\chi = \mathrm{const}$. On the other hand, a constant function f doesn't change at all under scale transformations: $f = f \circ \psi = \mathrm{const}$. The difference in the transformation behavior is of course due to the Jacobian being present in the transformation law for χ but not for f.

Thus the charge density "function" does not transform like a function, and in fact it isn't a function. As we shall see later on, the charge density "function" is the scalar determined by the *true* charge density and a fixed choice of volume form. The language "density" we are using is justified by the fact (yet to be demonstrated) that densities are canonically isomorphic to 0–chains in the discrete setting.

1.2 Inner and outer orientation

Recall that the flow lines of our discrete 1–chain model of the electric current density carry a *sense of direction*. Some other physical objects that can be modelled as 1–chains in 3–dimensional space are of a somewhat different type. For example, vortex lines in superconductors and superfluids do not possess a sense of direction but rather a *sense of circulation* – this is the sense in which the supercurrent flow circulates around them. Another example in this category are tornados (which we may roughly think of as lines extending from the earth's surface up to some height):

the most salient feature of a tornado is the circulation of the air around it.

We thus distinguish between two types of 1–chain: those with a sense of direction, and those with a sense of circulation. In mathematical language, what's different between the two types of 1–chain is the type of *orientation*. A 1–chain whose lines carry a sense of direction is called a 1–chain with *inner* orientation; one whose lines carry circulation is said to have *outer* orientation. An electromagnetic example of the latter kind is the magnetic field strength, $B = \sum \phi_i \gamma_i$. The lines γ_i of B are called *flux lines*, and the real weight factor ϕ_i is the *magnetic flux* carried by γ_i.

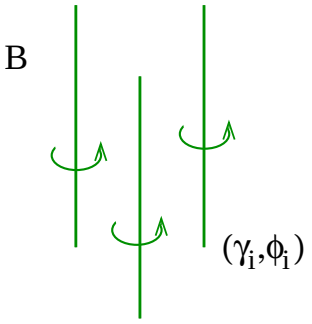

Figure 1.3: The magnetic field strength B is a 1–chain with outer orientation.

The same distinction – inner versus outer orientation – exists for chains of all degrees. Consider a 2–chain in 3–dimensional space. Its surface elements S_i may be oriented either by a sense of circulation (Fig. 1.4a), or by a transverse vector field defining a direction of passing through the surface (Fig. 1.4b). In the case of a 2–chain the first type of orientation is called inner, the second outer. Note that a surface with outer orientation (in 3–space) can be viewed as having two sides: an "in" side, and an "out" side; the transverse vectors point from the former to the latter. To give an example, the electric field strength $E = \sum V_i S_i$ is a 2–chain with outer orientation. In the static limit, the surface elements S_i organize into closed surfaces known as *equipotential surfaces*, and V_i is the voltage drop across the surface.

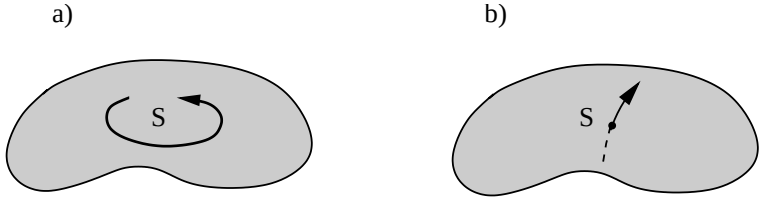

Figure 1.4: Surface with a) inner and b) outer orientation.

A 3–chain consists of domains in space. Giving an inner orientation to a 3–chain

in 3–space means to associate a right or left hand with each of its connected domains. Equivalently, we may assign to each domain an ordered set of linearly independent vectors $\{e_1, e_2, e_3\}$. A domain with outer orientation in 3–space is defined to be just a domain (with no extra attributes).

When every point of a 0–chain is equipped with a right or left hand, the result is a 0–chain with outer orientation. (Magnetic monopoles, if they existed, would have to be of that type.) Finally, to have complete systematics one decrees a point with inner orientation to be just a point.

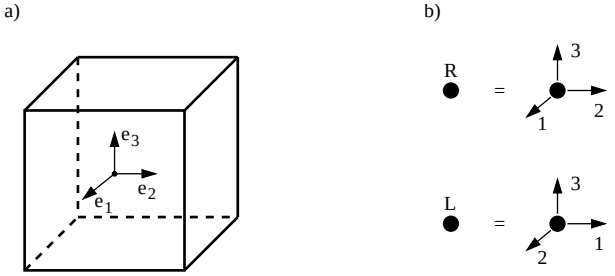

Figure 1.5: a) Cube with inner orientation, b) points with outer orientation.

Orientation is a dichotomy; i.e., fixing an orientation always means making a choice between two possibilities. The arrow on a line with inner orientation points one way, or the other way; the circulation of a line with outer orientation is either clockwise, or counterclockwise; an ordered set of linearly independent vectors $\{e_1, e_2, e_3\}$ forms a system which is either right–handed or left–handed. No third possibilities exist. Thus, fixing an orientation is invariably like choosing one member of the set $\{\pm 1\}$.

If c is an oriented chain (with inner or outer orientation), one adopts the reasonable convention that $-c$ is the same chain but with the opposite orientation. For example, if γ is a line with a sense of direction pointing one way, then $-\gamma$ is the same line but with the sense of direction pointing the opposite way. To see the rationale behind this convention, consider a single line γ carrying an electric current I. The electric current density is $j = I\gamma$ in that case. Now suppose the direction of the current flow is to be inverted. This can be implemented in the expression for j in two ways: we can either send I to $-I$, or we can invert the direction of the arrow on γ. Obviously, these two operations amount to the same physical operation of inverting the current. We also want $(-I) \times \gamma = I \times (-\gamma)$, or else there would be a conflict with the rules of vector calculus. Consistency then leaves us no choice but to say that $-\gamma$ is γ with the arrow pointing the other way.

The dichotomy of orientation also plays some role in a construction that relates

chains with inner and outer orientation. Let $\tilde{\gamma}$ be a line with outer orientation, and let γ be the same line but with an inner orientation. Then the pair $\gamma, \tilde{\gamma}$ determines an orientation of space; this is the orientation given by the hand the fingers of which curl according to the circulation of $\tilde{\gamma}$ when the thumb points in the direction of γ. Suppose the pair $\gamma, \tilde{\gamma}$ forms a right hand. We express this fact in formulas by the following notation:

$$\tilde{\gamma} = [\gamma, \text{right}] \, .$$

Yet, from the perspective of describing the chain with outer orientation $\tilde{\gamma}$, the sense of direction of γ is an irrelevant piece of data. There exists no good reason why we couldn't use $-\gamma$ instead of γ. With $\gamma, \tilde{\gamma}$ having been chosen to form a right hand, the new pair $-\gamma, \tilde{\gamma}$ will form a left hand, so by the same token as above,

$$\tilde{\gamma} = [-\gamma, \text{left}] \, .$$

Consistency therefore requires to make the identification

$$[\gamma, \text{right}] \equiv [-\gamma, \text{left}] \, .$$

This reveals the true meaning of the notation just introduced: if c is a chain with inner orientation, and $\text{Or} \in \{\text{right}, \text{left}\}$ an orientation of space, the notation $[c, \text{Or}]$ means an *equivalence class* consisting of the pair c, Or as well as the pair $-c, -\text{Or}$:

$$[c, \text{Or}] = [-c, -\text{Or}] \, .$$

Formally speaking, the group $\mathbb{Z}_2 = \{+1, -1\}$ acts on both factors of the pair, and we are forming equivalence classes by dividing by that group action. The operation of dividing by $\mathbb{Z}_2$ is distributive w.r.t. addition: $[c + c', \text{Or}] = [c, \text{Or}] + [c', \text{Or}]$, and linear w.r.t. multiplication: $[\lambda c, \text{Or}] = \lambda [c, \text{Or}]$ for $\lambda \in \mathbb{R}$.

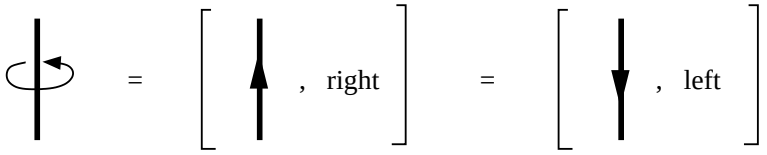

Figure 1.6: Graphical illustration of $\tilde{\gamma} = [\gamma, \text{right}] = [-\gamma, \text{left}]$.

In all of this we are implicitly assuming the ambient (or embedding) space and the chains in it to be *orientable*. To give a concrete example of where that assumption fails, consider the so-called Möbius strip, a surface made by twisting a rectangle and gluing opposite edges; see Fig. 1.7. Every line on a Möbius strip M can be given an inner orientation. Not so with outer orientation! Consider a closed line that

winds once (or an odd number of times) around the strip. Viewed as a 1–chain on the 2–manifold M such a closed line cannot be given a consistent outer orientation. Indeed, an outer orientation of a line on a surface amounts to a splitting into an "in" side and an "out" side. On a Möbius strip such a splitting does not exist globally: when you follow once around the winding curve, what you thought was the "in" side has turned into the "out" side! The origin of the difficulty is that the Möbius strip itself does not an admit an inner orientation (nor does it admit an outer orientation as a submanifold of 3–space). Transporting a sense of circulation continuously once around the strip, you find that the sense gets reversed.

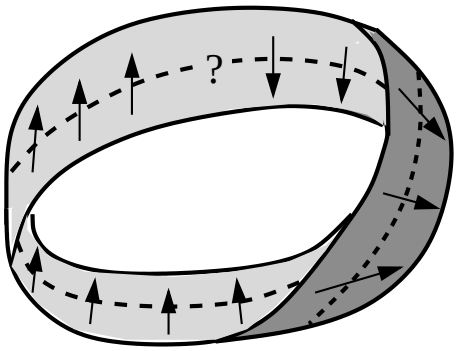

Figure 1.7: No outer orientation exists for a closed curve around a Möbius strip.

1.3 Boundary operator

To formulate Maxwell's equations in the chain picture, we shall need one differential-type operator. This is neither the curl, nor the divergence, but the *boundary operator*: a linear operator ∂, which transforms k–chains into $(k-1)$–chains by taking the boundary while preserving the type of orientation.

Let us go through the complete list of cases that occur in 3–dimensional space. To begin, consider a line γ with inner orientation. Its boundary $\partial\gamma$ is defined to be the final point, b, minus the initial point, a, of γ. In other words, $\partial\gamma$ is the 0–chain consisting of the two points a and b, with the former weighted by -1 and the latter by $+1$. Put in formulas: $\partial\gamma = -a + b$. If the end points coincide ($a = b$) the line γ is closed and has zero boundary ($\partial\gamma = 0$).

Now assume that the line γ carries an outer orientation by a sense of circulation (Fig. 1.8b). Then there is no invariant way to decide whether an end point of γ is an initial or final point. However, we can do the following instead. If $p \in \{a, b\}$ is an end point of γ, we pick some vector which is tangent to γ in p and points away from γ. In a natural manner, this "outward" tangent vector together with the

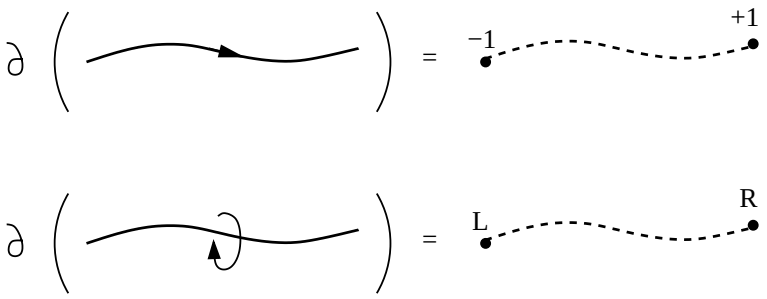

Figure 1.8: Boundary of a line with inner or outer orientation. `Need to label figures, lines, points.`

circulation of γ determines a system which is either right–handed or left–handed. Indeed, if the thumb points along the direction of the outward tangent vector, then it is either the fingers of a right hand, or those of a left hand, that curl around the thumb just like the circulation of γ curls around the vector. This right or left hand gets attributed to the end point p. The boundary of γ is defined to be the sum of the two end points with their respective outer orientations so determined. In the example shown in Fig. 1.8b,

$$\partial\gamma = [a, \text{right}] + [b, \text{left}] \ .$$

Of course we could equally well have written $\partial\gamma = [a - b, \text{right}]$.

Consider next a surface S with inner orientation. The boundary of S is a 1–chain, still with inner orientation, which may consist of zero, one, two, or more components: $\partial S = \sum_i \gamma_i$. Each component γ_i is a closed line. The sense of circulation of the surface S determines a sense of direction of each γ_i in the natural way: imagine a fluid circulating on S; the sense of direction of the boundary component γ_i is then given by the direction of the fluid flow near γ_i.

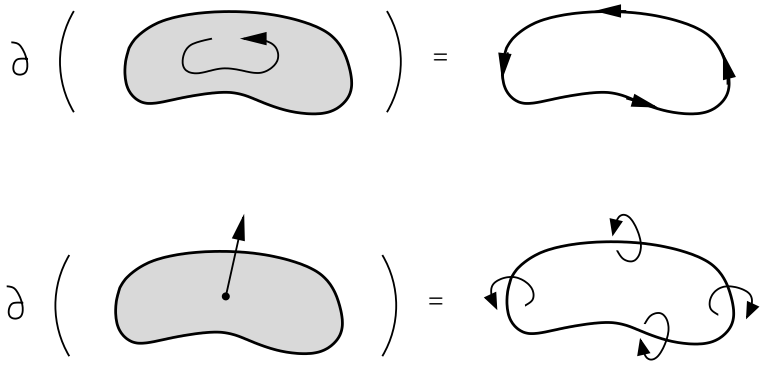

Figure 1.9: Boundary of a surface with inner or outer orientation. `Label surfaces and lines.`

When the type of orientation of the surface S gets changed from inner to outer,

the orientation type of the boundary components changes accordingly. Each piece of the boundary 1–chain is now given a sense of circulation. This sense is determined by first passing through the surface from the "in" side to the "out" side and then curving outward (i.e. away from the surface); see Fig. 1.9b.

Finally, let V be a connected domain in 3–dimensional space. The boundary ∂V is the closed surface of the domain V. If the orientation type of V is outer (which in the case of a domain in 3–space means no orientation at all) the "in" and "out" sides of ∂V are the sides facing the inside and outside of V, respectively.

For the case of inner orientation type the situation is slightly more complicated. Looking at the surface ∂V from the outside of V, we perceive the sense of circulation of ∂V as being counterclockwise (the mathematically positive sense) if the orientation of V is given by the right–hand rule. In the case of inner orientation by the left hand, what we see is the clockwise (or negative) circulation.

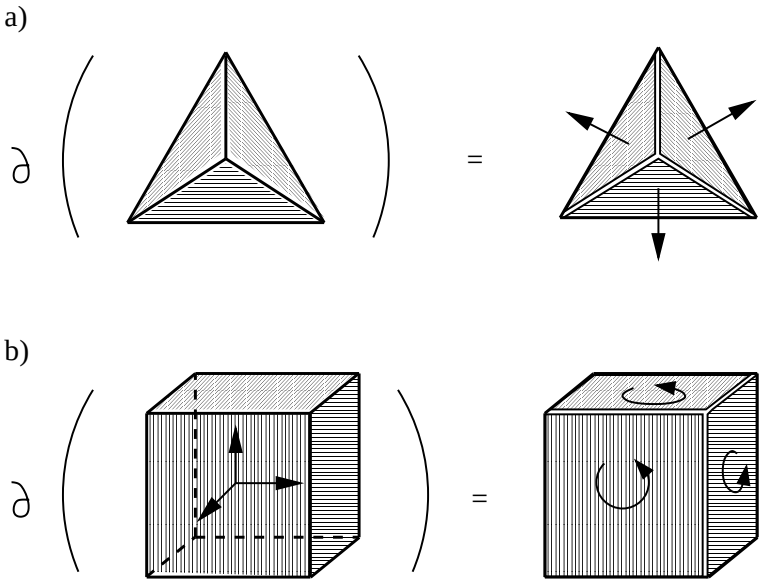

Figure 1.10: a) Boundary of a solid tetrahedron, b) boundary of an oriented cube.

In all of the cases we have discussed, the concept of boundary was explained at the example of elementary lines, surfaces, or domains. Nonetheless, since chains are linear combinations of these elementary objects and the boundary operator is defined to be linear, the above discussion completely specifies

$$\partial : k\text{–chains} \to (k-1)\text{–chains}.$$

The boundary of a 0–chain is trivially zero, as there are no chains of degree -1. This agrees with your intuition that points have no boundary.

A chain c with no boundary ($\partial c = 0$) is called *closed*. A closed chain is also called a *cycle*. We have remarked along the way that the boundary of any domain is a closed surface, and the boundary of any surface is a set of closed lines. It is intuitively obvious (although not so easy to prove) that the *boundary of a boundary always vanishes*:

$$\boxed{\partial \circ \partial = 0} \quad .$$

In abstract mathematics this property is often taken to be part of the system of axioms: if a degree-lowering operator ∂ does not satisfy $\partial \circ \partial = 0$, it simply isn't a boundary operator.

A more delicate issue is whether the converse holds true. Given some closed k–chain b we might ask whether there exists a $(k+1)$–chain c such that $b = \partial c$. In other words: is every cycle a boundary? The answer depends on the topology of space. If the space is $\mathbb{R}^n$, or an affine space modelled on $\mathbb{R}^n$, the answer is yes:

$$\boxed{\partial b = 0 \Rightarrow \exists c : b = \partial c} \qquad \text{(Poincaré lemma)}.$$

On the other hand, if the topology of space is non-trivial, the Poincaré lemma fails in general. For example, if U is $\mathbb{R}^3$ with one axis removed, there exist closed loops winding around that axis which cannot be expressed as the boundary of any surface in U; and if U is $\mathbb{R}^3$ with one point removed, there exists closed surfaces surrounding that point which cannot be expressed as the boundary of any domain in U.

1.4 Intersection pairing: integration

An indispensable tool in electrodynamics (as in any physical theory) is integration. The good objects to integrate are differential forms. In this book differential forms will not be introduced until Chapter [Ref], so for the time being we have to make do with a discrete version of integration.

The natural substitute for integration in the present discrete context is what will be called the *intersection pairing* of chains. It is an invariant bilinear pairing that exists between k–chains with inner orientation and $(3-k)$–chains with outer orientation. (Here 3 is the number of dimensions; in d dimensions the pairing would be between chains of degree k and $d-k$. For simplicity we will always be assuming $d = 3$ dimensions unless stated otherwise.)

To introduce the intersection pairing we start from the observation that, in d dimensions, the intersection between two subspaces of dimension k resp. l is generically of dimension $d-(d-k)-(d-l) = (k+l)-d$. Thus a k–chain c and a $(3-k)$–chain c' in 3 dimensions generically intersect (if they intersect at all) in a set

of points. The intersection pairing $P(c, c')$ is a real number computed by summing up weights of intersection points as follows.

Let us start with the case $k = 1$ and $l = 2$. So let γ be a line, and let S be some surface intersecting γ transversally in a single point p. If the orientation type is inner for γ and outer for S, we can make an invariant comparison of orientations: the sense of direction of γ either agrees with that of the transverse vector field of S at the intersection point p, or it disagrees. In the former case the intersection pairing is defined to be $P(\gamma, S) = +1$, in the latter case $P(\gamma, S) = -1$. If γ and S do not intersect we set $P(\gamma, S) = 0$. It may also happen that γ and S intersect several times; in that case the value of $P(\gamma, S)$ is a sum over intersection points, where each point is counted as $+1$ or -1 depending on whether orientations do or do not agree.

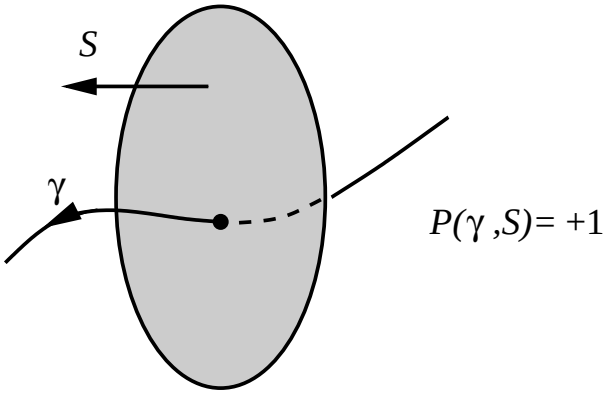

Figure 1.11: Intersection pairing of a line with a surface. `Label the point` p.

Since P is a bilinear function, this definition extends to any 1–chain $j = \sum_i I_i \gamma_i$ and 2–chain $E = \sum_i V_i S_i$ by

$$P(j, E) = \sum_{i,i'} I_i V_{i'} P(\gamma_i, S_{i'}) \; .$$

(If j is the electric current and E the electric field strength, the intersection pairing $P(j, E)$ has a physical meaning: it is the rate of energy transfer from the electric field to the charge carriers.)

The next case to consider is $k = 2$ and $l = 1$. The pairing now is between 2–chains with inner orientation and 1–chains with outer orientation. Everything works the same way as before, except that at each intersection point of a surface with a line, we now compare circulations. If the sense of circulation of the surface agrees with that of the line, the intersection point counts as $+1$, and if it doesn't, as -1. To give a physical example, if S is a surface with inner orientation and $B = \sum_i \phi_i \gamma_i$

the 1–chain of the magnetic field strength, the intersection pairing $P(S, B)$ is the magnetic flux through the surface.

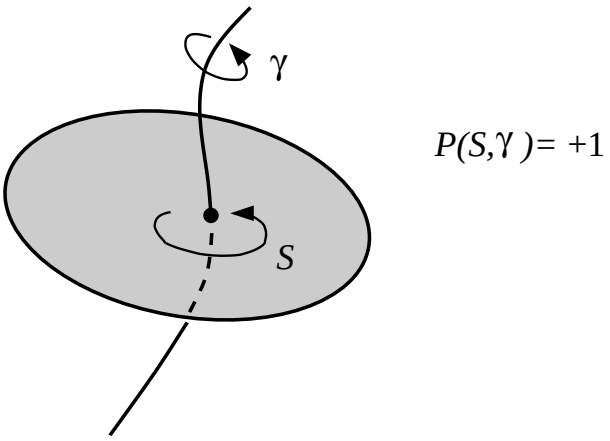

Figure 1.12: Intersection pairing of a surface with a line.

Turning to the case $k = 3$ and $l = 0$, let U be an oriented domain, and p a point with outer orientation. If p lies outside of U we set $P(U, p) = 0$. In the opposite case we compare orientations. If the orientations agree (i.e. both U and p carry the same hand – the right or left one), the intersection pairing is $P(U, p) = +1$. If they disagree, $P(U, p) = -1$. Again, this definition extends by linearity to the case of any 3–chain and 0–chain with inner resp. outer orientation.

The last case, $k = 0$ and $l = 3$, is the simplest one. There are no orientations to compare here. We just identify which points of the 0–chain lie in which regions of the 3–chain, and sum up the products of the weights. For example, if U is a domain with outer (i.e. no) orientation and $\rho = \sum q_i p_i$ is the 0–chain of the electric charge density, the intersection pairing

$$P(\rho, U) = \sum_{i:p_i \in U} q_i$$

is the total charge in U.

A word of warning is in order here. In order for the intersection pairing to be well defined, the chains have to intersect *transversally*. If chains are chosen at random, transversality will hold with probability one. Nevertheless, exceptions do exist. A curve γ might, for example, meet a surface S in some point a, then follow along the surface for some distance, and finally leave the surface again at point b; see Fig. 1.13. Such a curve has infinitely many points in common with the surface, but we cannot decide for any of these intersection points whether the orientations match or not. What's the intersection pairing $P(\gamma, S)$ for this degenerate situation?

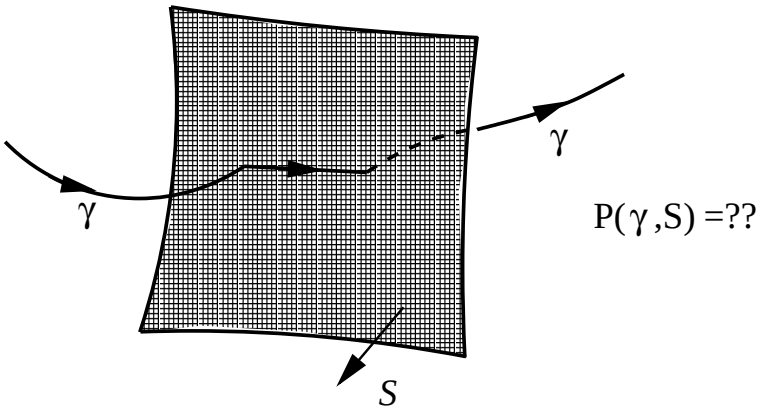

Figure 1.13: Non-transversal intersection of a curve with a surface.

Our discussion does not explain how to deal with exceptional cases like this one (at least not without some further thinking). Fortunately, the problems with such situations go away when one of the two chains is smeared out and turned into a smooth differential form (in which case the intersection pairing becomes the integral of the form over the chain). The plan of this book, however, is to avoid the use of differential forms in the first two chapters.

One way out for now is to tilt the curve by a very tiny amount, so that the degeneracy is lifted, the intersection becomes transversal, and we are back to a situation we know how to handle. This is the cheap solution we adopt here. Fixing the problem in a more satisfactory manner would require a certain amount of effort, and would torpedo the didactic purpose of the present chapter which is to give an easy and intuitive (if not always mathematically rigorous) introduction to the subject and the general structure of the formalism used.

1.5 Stokes' theorem

The intersection pairing satisfies an important relation involving the boundary operator ∂. If ∂c is a k–boundary with inner orientation which transversally intersects a $(3-k)$–chain ω with outer orientation, then

$$P\left(\partial c, \omega\right) = (-1)^{k+1} P\left(c, \partial\omega\right) \ .$$

Thus ∂ moves across inside the intersection pairing at the cost of a sign factor. The claimed relation will now be verified for a few basic cases.

i) $k = 0$. Let U be a connected domain, and let γ be a directed curve with boundary $\partial\gamma = b - a$. Suppose first that both end points a and b lie outside of U. Then the left–hand side vanishes trivially, and so does the right–hand side provided

that the curve c never enters U, as there are no points of intersection between γ and the boundary ∂U in that case. If the curve γ does enter U, the intersection pairing $P(\gamma, \partial U)$ still vanishes, since there is exactly one exit point for every entry point and their contributions cancel each other. Suppose next that both end points of γ lie inside of U. The value of $P(\gamma, \partial U)$ still vanishes, for essentially the same reasons as before, and so does the left–hand side: $P(\partial\gamma, U) = P(b, U) - P(a, U) = 1 - 1 = 0$. Suppose finally that one of the end points, say the starting point a, lies inside of U and the other one outside. Then $P(\partial\gamma, U) = -P(a, U) = -1$. On the right–hand side, there must now exist at least one point of intersection between γ and the boundary ∂U. If there exists exactly one such point, then $P(\gamma, \partial U) = +1$ since the orientations match (the curve goes from the inside to the outside, so it pierces the surface ∂U in the same sense as does the transverse vector field of ∂U). If there is more than one intersection point, the result still is $P(\gamma, \partial U) = 1$, as the additional intersection points come as pairs of entry and exit points and their contributions cancel. If the roles of a and b are exchanged, both sides of the equation change sign. Thus $P(\partial\gamma, U) = -P(\gamma, \partial U)$ in all cases.

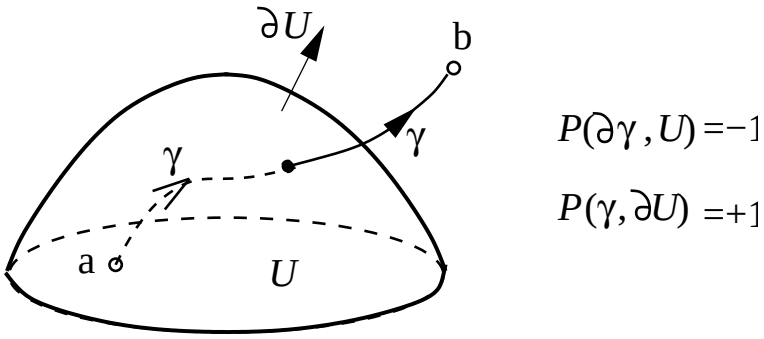

Figure 1.14: Stokes' theorem for $k = 0$.

ii) $k = 1$. Let S and Σ be two surfaces with inner and outer orientation respectively. Their boundaries are cycles equipped with a sense of direction resp. circulation. Now, it is a true fact from the topology of 3–dimensional Euclidean space that the intersection pairings $P(\partial S, \Sigma)$ and $P(S, \partial\Sigma)$ both vanish unless the two 1–chains ∂S and $\partial\Sigma$ are knotted w.r.t. each other. (Try a few examples to convince yourself!) Therefore let ∂S and $\partial\Sigma$ form a knot, and let that knot be simple as in Fig. 1.15. In the example shown there, both ∂S and $\partial\Sigma$ consist of a single closed curve. The surfaces S and Σ intersect in a curve with two end points, say a and b. The former is a point of intersection of S with $\partial\Sigma$, and the latter a point of intersection of Σ with ∂S. If the two intersection points are close to each other and

the surfaces S and Σ are smooth, it is obvious that the sense of direction of ∂S at b evolves into the sense of circulation of S at a just like the sense of transversality of Σ at b evolves into the sense of circulation of $\partial\Sigma$ at a. Therefore, the orientations either match at both intersection points, or they disagree at both points. Thus $P(\partial S, \Sigma) = P(S, \partial\Sigma)$. When the points a and b are not close to each other, or the surfaces S and Σ rough, this relation is still true by a continuity argument: pull the two intersection points apart; as long as the knot is kept intact, the two intersection pairings will remain unchanged.

These considerations provide some intuition and insight into why the claimed relation between intersection pairings holds. Of course they do not constitute a complete proof (of the nonetheless true fact) that it holds in all possible cases.

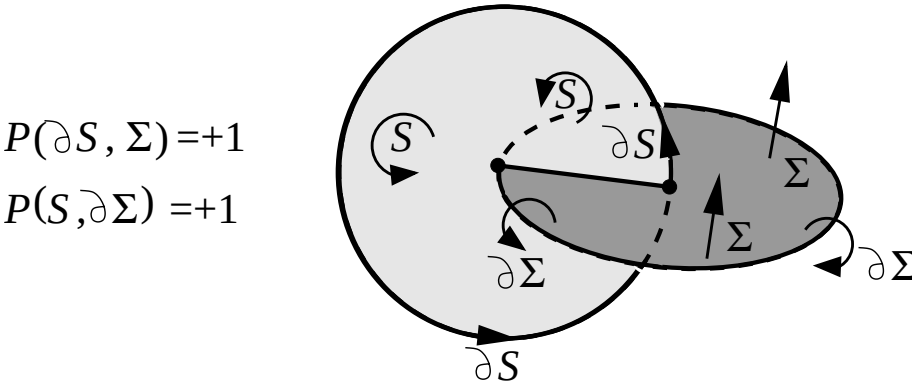

Figure 1.15: Stokes' theorem for $k = 1$.

iii) $k = 2$. Let U be an oriented domain and γ a curve with outer orientation. The discussion of this case is essentially identical to that for $k = 0$, except that we now have to compare orientations at *all* intersection points, including those of $\partial\gamma$ with U. It is not difficult to see that $P(\partial U, \gamma) = -P(U, \partial\gamma)$ holds in all cases.

In summary, the relation $P(\partial c, \omega) = (-1)^{\deg(\omega)} P(c, \partial\omega)$ always holds, as long as intersections are transversal and pairings well-defined. We have already (and perhaps mysteriously) pronounced the intersection pairing to be a discrete version of the integral. Let us now implement this fact in our notation, so as to build up familiarity with the standard formalism to be installed later on. For a k–chain c and a $(3-k)$–chain ω (with different types of orientation) we write

$$P(c, \omega) =: \int_c \omega \ .$$

For the time being, this may appear to be just a peculiar notation, which the reader is asked to accept in good faith. As we develop the formalism more systematically

it will transpire that, when the $(3-k)$–chain ω is converted into a smooth k–form (definitions later!) by the process of smearing, the intersection pairing $P(c,\omega)$ in fact becomes the standard integral, $\int_c \omega$, of the k–form ω over the k–chain c.

There is another piece of notation that will become of increasing importance as we go along. Let $\mathrm{d} : k\text{–chains} \to (k-1)\text{–chains}$ be the same as the boundary operator ∂ but for an alternating sign:

$$\mathrm{d}\omega := (-1)^{\deg(\omega)}\partial\omega \ .$$

A distinctive property of the operator d is that it is adjoint to ∂ with respect to the intersection pairing:

$$P(\partial c, \omega) = P(c, \mathrm{d}\omega) \ .$$

In integral notation, the relation between intersection pairings now takes the form

$$\boxed{\int\limits_{\partial c} \omega = \int\limits_{c} \mathrm{d}\omega} \qquad \text{(Stokes' theorem).}$$

(The modified boundary operator d turns into the exterior derivative on passing from chains to forms, and our discrete relation then becomes what is known as the Stokes' theorem of exterior calculus.)

The way we have stated it, Stokes' theorem assumes the types of orientation of c and ω to be inner resp. outer. However, instead of $P(c,\omega) = \int_c \omega$ we could as well have written $P(c,\omega) = \int_\omega c$, as the dependence of P on its arguments is symmetric. Therefore orientation types can be swapped, and Stokes' theorem still holds if the orientation is outer for c, and inner for ω.

2 Homogeneous Maxwell equations

The electric and magnetic field strengths are subject to a set of equations known as the homogeneous Maxwell equations. In the vector calculus used in traditional textbooks on electrodynamics one writes them as $\mathrm{div}\vec{B} = 0$ and $\mathrm{rot}\vec{E} = -\frac{\partial}{\partial t}\vec{B}$. The operator "div" is the divergence operator; to define it one has to make a choice of volume form for space. The definition of the vector differential operator "rot" requires making a choice of metric and orientation of space.

On the other hand, the true *physics* of the homogeneous Maxwell equations depends on no such information as the metric or the orientation. When written in their proper form, these equations in fact enjoy the property of general covariance, which is to say that they remain unchanged under any smooth coordinate transformation or diffeomorphism of space and time.

The simple chain picture developed in the present chapter does not make full use of the differentiable structure of space-time. Nonetheless, it allows us to present Maxwell's equations in a form free of contamination by the metric or other extraneous information.

2.1 What's a force field?

In traditional physics teaching, the physical term "force" is commonly introduced as a vector (or vector field). The evidence cited in favor of the vector nature of forces is Newton's second law:

$$\text{mass} \times \text{acceleration} = \text{force} \, .$$

Mass is a scalar and acceleration doubtless is a vector, so consistency leads one to conclude that force, too, must be a vector.

From a fundamental point of view, this conclusion is not entirely accurate. The first hint that there is more to the story comes from the observation that Newton's second law is not universal but holds only in its restricted domain of validity. According to Einstein's theory of general relativity, modifications occur when the geometry of space is non-Euclidean.

Another hint comes from measurement: to measure a conservative force, what you do is you displace a test body subject to the force and measure the resulting change in potential energy. In an ideal measurement, you measure the infinitesimal change in potential energy due to an infinitesimal displacement of the test body. In formulas,

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \big(W(a) - W(a + \epsilon u) \big) =: F_a(u) \, .$$

Here a is the location of the test body, u is the displacement vector, and $W(a)$ is the potential energy of the body at the point a. The same operative definition applies to non-conservative forces if $W(a) - W(a + \epsilon u)$ is replaced by the work needed to displace the test body by the infinitesimal translation vector from $a + \epsilon u$ to a.

For every choice of u, the measurement performed at a outputs a real number $F_a(u)$. It is an experimental fact that the dependence of $F_a(u)$ on u is linear. Thus the collection of data from the measurement determines a linear function $u \mapsto F_a(u)$. This function is called the *force* (or the force field at a) in the proper sense. Thus force properly understood is not a vector but a *linear function on vectors*,

$$\text{force} : \text{vectors} \to \mathbb{R} \, .$$

Considering the linear functions F_a at all points a in space at once, we get what is

called the *force field*. In mathematical terminology, a smooth assignment such as $F : a \mapsto F_a$ is called a *differential form of degree one*, or 1–form for short.

Thus a force field is not a vector field but a 1–form. Having appreciated this fact, we rewrite Newton's second law as follows:

$$\text{mass} \times \big\langle \text{acceleration}, \bullet \big\rangle = \text{force} \, .$$

Here $\langle \bullet, \bullet \rangle$ denotes the standard scalar product of vectors in Euclidean space. Note that the equation makes good sense: the right-hand side is a linear function on vectors, and so is the left-hand side – its value on the velocity vector v is the rate of change of the kinetic energy of the body, $m\langle \dot{v}, v \rangle = \frac{d}{dt}\left(\frac{m}{2}|v|^2\right)$. The second form of Newton's law is more revealing than the first one, as it clearly displays an assumption implicit to Newton's theory: that the geometry of space is Euclidean.

The use of differential forms has been outlawed in the first two chapters of this book by its author, in order to protect the student from getting absorbed in the details of exterior calculus before having acquired some solid physical intuition of the subject. We therefore need to make do with an alternative description of force fields for the moment. The primary task is to devise a discrete model of the *work functional* $W[\gamma]$: the function(al) that assigns to every directed curve γ the work done when moving the test body along γ. (In continuum language this would be minus the line integral of the force field along the curve γ.)

The work functionals W to be modelled are required to have all the properties of a linear function: if we divide the directed curve γ into two segments γ_1 and γ_2, the work must be additive: $W[\gamma_1 + \gamma_2] = W[\gamma_1] + W[\gamma_2]$; if we repeat the measurement, moving the test body twice along γ, the work must double: $W[2\gamma] = 2W[\gamma]$; and so on. Thus the objects to be described are linear functions W from the space of 1–chains (with inner orientation) into the real numbers.

Within the framework of the chain picture developed here, the only reasonable and invariant construction of such functions is provided by the intersection pairing: if K is some 2–chain with outer orientation, we can assign to any directed curve γ the real number $P(\gamma, K)$, which depends linearly on γ (as well as on K).

Hence in the atomistic view of electrodynamics taken in this chapter, we are going to model force fields as 2–chains $K = \sum w_i S_i$ with outer orientation, defining the work functional associated with K by

$$W[\gamma] = -P(\gamma, K) = -\int_\gamma K \, .$$

(This meets your expectation that $-W[\gamma]$ ought to be the line integral of the force field K along γ.) Recall how $\int_\gamma K$ is computed in our discrete picture: we look for

all the points of intersection of γ with the surface elements S_i of K, and we sum up all the relevant weights w_i, counting them as positive when orientations agree and negative when they disagree. The weight $-w_i$, of course, is to be interpreted as the work needed to advance the test body across the surface S_i. Our definition via the intersection pairing then makes good sense: it simply states that the total work $W[\gamma]$ sums up all the little amounts of work done along the way of γ.

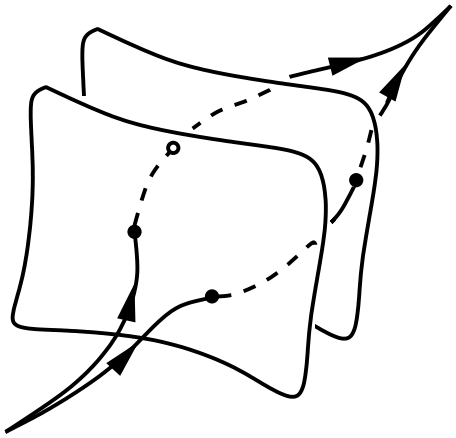

Figure 1.16: Path independence of the work functional for a conservative force field.

In chain language it is easy to state what it means for a force field K to be conservative: in that special case K must be *closed*, i.e. the 2–chain model of K has to consist of closed surfaces. (Actually, the true condition is that K be a boundary: $K = -\mathrm{d}\Phi$; but in Euclidean 3–space this amounts to the same, by the Poincaré lemma.) Indeed, if all surfaces of K are closed, the work $W[\gamma]$ depends only on the *end points* of the curve γ; for if we push the inner part of γ around while holding the end points fixed, the points of intersection of γ with K move around but the sum over weights of intersection points remains the same (Fig. 1.16). Thus if K is closed the work functional $W[\gamma]$ does not depend on the path taken by γ and, as one learns in elementary mechanics, path independence is the very criterion by which one decides whether a force field is conservative or not.

On the other hand, if a force field K is non-conservative, then the surface elements of K do *not* piece together to form closed global surfaces. Stretching our imagination a little bit, we then need to visualize an array of little pieces of surface, with appropriate weights attached to all the pieces and the array arranged in such a way as to approximate a given (continuous) force field.

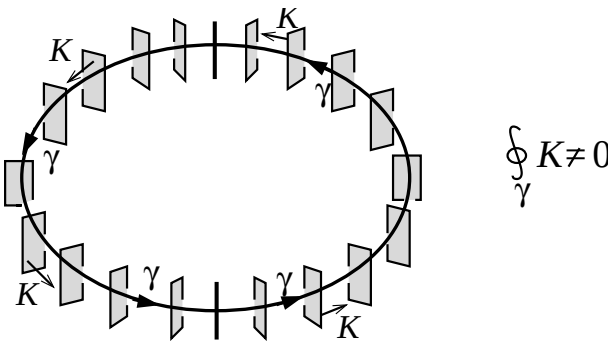

Figure 1.17: Non-conservative force field modelled by a 2–chain.

2.2 Electric field strength E

It is a basic experimental fact that electrically charged bodies exert forces upon one another. Indeed, the process of pushing around a charged body in the presence of other charges is observed to take up energy, or release energy, as the case may be. The force that acts on a charged body by virtue of its carrying electrical charge is called the *electrical force*. Its main characteristic is that it depends linearly on the body's charge, and is independent of other properties of the body such as its mass or velocity. The agent underlying the electrical force is a variable local property of space called the *electric field strength*, E.

To make a measurement of E, one picks a small test body with negligible mass and any nonzero charge Q and translates it along some curve γ. In such a process some work $W^{(\mathrm{e})}[\gamma]$ has to be done (the work will be positive or negative depending on whether one is pushing against the force field or along with it.) Division of the negative work $-W^{(\mathrm{e})}[\gamma]$ by the charge Q gives a number which we denote by $V[\gamma]$. Since the electrical work is proportional to the charge, the dependence on the charge of the test body drops out of the quotient $-W^{(\mathrm{e})}[\gamma]/Q = V[\gamma]$. The latter is an intrinsic (though variable) feature of space called the *voltage* along γ.

By construction, the voltage is a linear function on directed curves or, more generally, on 1–chains with inner orientation. If the curve γ is cut up into segments $\gamma_1, \ldots, \gamma_n$, the voltages add: $V[\sum_{i=1}^n \gamma_i] = \sum_{i=1}^n V[\gamma_i]$. This means that the voltage along γ results from summing (or integrating) over something; by definition, that something is the electric field strength E:

$$\boxed{V[\gamma] = \int_\gamma E} \qquad \text{(definition of } E\text{)}.$$

Thus by measuring the voltage along γ one is measuring the line integral of E along γ. A single such measurement does not determine E. However, by making many

voltage measurements, for all kinds of curves γ, we can get as much information on E as we desire. (E may change with time. If so, the voltage measurements have to be done at a rate faster than the rate of change of E.)

It is implicit in the operative definition of E that the electric force field acting on a charge q is

$$K^{(\mathrm{e})} = qE \ .$$

The explanations given in Section 2.1 then tell us to model E by a 2–chain with outer orientation. This modelling of course fits with the line integral definition of E: as we have seen, doing the line integral $\int_\gamma E$ in the discrete picture actually means to compute an intersection pairing: $\int_\gamma E = P(\gamma, E)$; and since the directed curve γ is a 1–chain with inner orientation, the intersection pairing $P(\gamma, E)$ makes immediate sense if and only if E is taken to be a 2–chain with outer orientation.

Thus a discrete approximation to E is a sum over surface elements S_i with a transverse sense of direction,

$$E = \sum_i V_i S_i \ ,$$

where the coefficients V_i are the voltages (or voltage drops) across the surfaces S_i. From this expression as well as the operative definition given above, it is clear that the electric field strength E carries the physical dimension of voltage, or energy per charge. (This is what is called the *absolute* dimension of E. Once we expand E with respect to a Cartesian basis, the components E_x, E_y, E_z will be seen to have the *relative* physical dimension of voltage per length.)

2.3 Magnetic field strength B

The force exerted upon a charged particle by the electric field strength does not depend on the velocity of the particle. It turns out that this electric force alone does not suffice – in combination with Newton's second law – to explain the motion of charged particles. For one thing, if a wire carrying an electrical current is placed in the vicinity of a permanent magnet, it experiences a force which is independent of its charge state but grows linearly with its current. For another, the trajectories of charged particles near a magnet are not straight but bent. These observations point to the existence of a velocity-dependent force called the *Lorentz force*. Its local cause is a physical quantity referred to as the *magnetic field strength*, B. Paralleling the electric case, we give an operative definition of B via the notion of work.

The ideal equipment to use for our gedanken-measurement of B is a piece of wire made of an perfect conductor – this means that the resistance of the wire is

very small, or even zero as in a superconductor. We are going to assume that the ideal wire is flexible, electrically neutral, and very thin as we are going to treat it as one-dimensional. Let such a wire be connected to an external source, say a battery, forcing upon the wire a current which is maintained at constant level $I \neq 0$ throughout the duration of the measurement. To start the experiment, pin down the wire in some configuration. As soon as the magnetic field is turned on (or the wire moved into the field region), magnetic forces start pulling on the segments of the wire. Then, keeping two points a and b of the wire fixed, change the directed intermediate path of the flexible wire from γ to γ' (see Fig. 1.18). In the process, the wire sweeps out an oriented surface S with boundary $\partial S = \gamma' - \gamma$.

Changing the configuration of the wire requires doing some work against the Lorentz force. This "magnetic" work $W^{(\mathrm{m})}[S]$ is proportional to the current, so we normalize by the current: $-W^{(\mathrm{m})}[S]/I =: \Phi[S]$. The quotient $\Phi[S]$ so obtained is called the *magnetic flux* through S. Note that its physical dimension is energy per current.

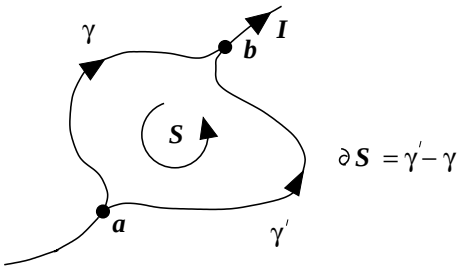

Figure 1.18: Using a test current to measure the magnetic flux through a surface S.

It is important to appreciate that the type of orientation to choose for the surface S is dictated by the nature of the measurement process. What we have done is to endow S with an inner orientation; it's the inner orientation given by the sense of circulation which agrees with that of the final path γ' and is opposite to that of the initial path γ. Note that there exists no other way of making an invariant choice. In particular, there exists no distinguished way of assigning to S an outer orientation by a transverse sense of direction, as we do not know how to tell the two sides of S apart without injecting extraneous information such as the right-hand or left-hand rule. Thus the surface orientation determined uniquely and invariantly by the measurement process is of the inner type. Conversely, given some surface S with inner orientation we know how to sweep the flexible wire over S and do the measurement of $\Phi[S] = -W^{(\mathrm{m})}[S]/I$.

The magnetic flux function $S \mapsto \Phi[S]$ is a linear function on 2–chains with inner orientation. Indeed, if we slice the surface S into pieces $S_1, \ldots, S_n$, the flux through the sum is the same as the sum of the fluxes: $\Phi[\sum_{i=1}^n S_i] = \sum_{i=1}^n \Phi[S_i]$. This cannot but mean that the magnetic flux comes from integrating something; that something is the magnetic field strength B:

$$\boxed{\Phi[S] = \int\limits_S B} \qquad \text{(definition of } B\text{)}.$$

As always in the present chapter, the integral sign stands for pairing by intersection: $\int_S B = P(S, B)$. Since S is a 2–chain with inner orientation, B must be a $(3-2)$–chain with outer orientation in order for the intersection pairing to make sense. Thus the chain model of the magnetic field strength is built from lines γ_i with a sense of circulation; we write

$$B = \sum_i \phi_i \gamma_i \ ,$$

where the coefficient ϕ_i is called the magnetic flux carried by the line γ_i.

From its definition, the magnetic field strength inherits the absolute physical dimension of magnetic flux, or energy per current, or action per charge. The last way of expressing the dimension acquires special significance in quantum mechanics, which provides a natural unit of magnetic flux by h/e, the ratio of Planck's constant over the electron charge.

The operative definition we have given for B implies the Lorentz force law for charges moving in a magnetic field. Unfortunately, to formulate the Lorentz force in direct and concise terms we need an operation (the interior product) not yet at our disposal. We shall therefore return to this point later, in Section 5.2.

2.4 No sources for B

Let γ and γ' be two directed curves with a common set of end points: $\partial\gamma = \partial\gamma'$. There are many ways of continuously deforming γ into γ' or, put differently, there exist many different surfaces S that connect γ with γ' in the sense of $\partial S = \gamma' - \gamma$; an example is shown in Fig. 1.19. If S_1 and S_2 are two such surfaces, how do their magnetic fluxes compare?

According to the definition of magnetic flux we have given, the question is answered by taking a line current (with the current I held fixed by an external source), pushing it along S_1 and S_2, and comparing the amounts of work done. If you do that experiment, you discover that no matter what surface you choose, the work is always the same!

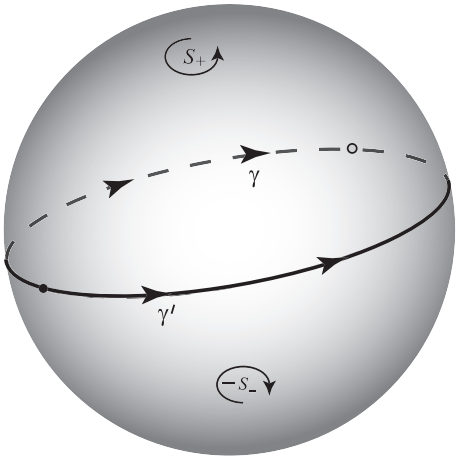

Figure 1.19: On a sphere, there exist two ways of connecting a given point a on the equator with its antipode b by following along the equator: these are given by the two directed curves γ and γ' in the figure. The 1–chain $\gamma' - \gamma$, the equator, bounds the upper hemisphere S_+ as well as the negative of the lower hemisphere, $-\mathrm{S}_-$. If the sphere is embedded in Euclidean 3–space, there exist many other surfaces that have the equator for their boundary. `Enter points` a, b.

To rephrase that finding, recall what it means for a force field to be conservative: if a test particle is pushed along a path γ in a conservative force field, the work done is *path*-independent; there is a dependence on the end points of the path but not the path itself. The magnetic force field under consideration is "conservative" in a similar sense: the work done while pushing a line current along a surface in a magnetic field is *surface*-independent: it depends on the initial and final configurations of the line current (or the boundary of the surface) but not the surface itself.

Thus if S and S' are two surfaces with inner orientation and the same boundary, the magnetic fluxes through them are experimentally found to be the same:

$$\partial S = \partial S' \Longrightarrow \Phi[S] = \Phi[S'] \; .$$

What conclusion about B can we draw from this observational fact?

If $\partial S = \partial S'$, then $\partial(S' - S) = 0$, and the Poincaré lemma ensures the existence of an oriented domain U such that $\partial U = S' - S$. Therefore $\Phi[S] = \Phi[S']$ leads to

$$0 = \Phi[S' - S] = \Phi[\partial U] = \int\limits_{\partial U} B \; ,$$

by the definition of the magnetic field strength. Conversely, if U is any oriented domain in space, its boundary can always be decomposed into pieces and written as a difference of two surfaces, $\partial U = S' - S$, where the boundaries of S and S' coincide. Hence for every oriented domain U,

$$\boxed{\int_{\partial U} B = 0} \qquad (B \text{ is source free}).$$

By Stokes' theorem $\int_{\partial U} B = \int_U \mathrm{d}B$ (which is often called Gauss' theorem in the present context) this is equivalent to

$$\mathrm{d}B = 0 \, .$$

It is clear how to interpret this law in the chain picture. Since $\mathrm{d}B = -\partial B$, the equation $\mathrm{d}B = 0$ says that the 1–chain of B has no boundary, or is closed. This means that approximating B by a 1–chain, $B = \sum \phi_i \gamma_i$, we can always arrange for the lines γ_i (the "magnetic flux lines") to be cycles. We express this fact by saying that magnetic flux lines have no beginning and no end. Alternatively we say that there exist no sources for B, or B is source free.

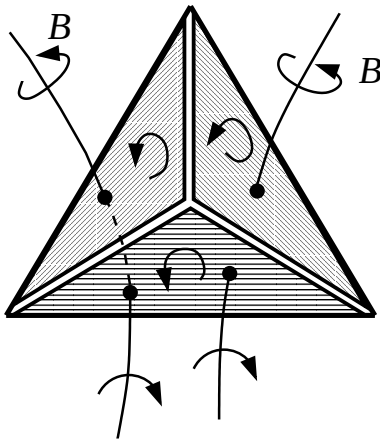

Figure 1.20: There is always as much magnetic flux leaving a closed surface as there is entering it.

Notice the following important aspect: determining the magnetic flux $\int_S B$ in the chain picture is essentially a counting exercise. Indeed, discretizing B by magnetic flux lines that all carry the same flux, say ϕ, we just need to count the points of intersection between the surface S and the flux lines: if N_+ (N_-) is the number of intersection points where the orientation of B matches that of S (resp. doesn't match), then

$$\int_S B = (N_+ - N_-)\phi \, .$$

Note in particular that we never need to make any length or angle measurement. All that matters for the value of the magnetic flux integral $\int_S B$ is whether the lines of B do or do not intersect the surface S; the angle of intersection is irrelevant. Thus no metric information ever enters here.

Compare this to what is done in traditional text books viewing the magnetic field strength as a vector, $\vec{B}$. In order to compute the magnetic flux through a surface, $\int_S \vec{B} \cdot d\vec{S}$, one picks the component of $\vec{B}$ normal to the surface, and integrates it against the surface density induced on S by the Euclidean structure of space. In the process, the Euclidean metric is used several times: once to evaluate the scalar product between $\vec{B}$ and the normal vector of S, and then again to determine the density to be used for integration. Thus one makes repeated use of the metric en route to an answer that's totally independent of the metric!

2.5 Faraday's law of induction

In all of the above, the test current was kept at a constant level I. This may not be the best way of doing the experiment in practice, and it also raises a question: since we believe energy to be a conserved quantity, and work had to be done to push the line current into its new configuration, where did the energy go?

The only reasonable answer is that energy was exchanged with the external source maintaining the constant ideal current flow. Indeed, pushing the line current makes the charge carriers experience a force – that's the Lorentz force – which slows them down or speeds them up as the case may be; so to keep the charge carriers at constant speed and maintain the current, energy must be drawn from the external source or transferred to it (if it isn't dissipated as heat by the non-ideal conditions of a real experiment).

There exists a variant of the previous experiment which allows a rigid wire to be used instead of a flexible one. The idea is to bend the ideally conducting wire into a closed loop γ ($\partial\gamma = 0$). Then, leaving the shape of the loop unchanged we pull it to another location, γ'. If an external source maintains the current at a constant level I, the work done is still $-I \int_S B$ where S is any surface with boundary $\partial S = \gamma' - \gamma$.

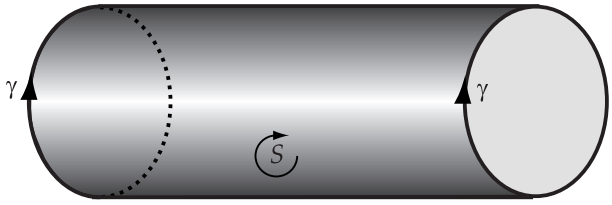

Figure 1.21: Variant of the magnetic flux experiment: a rigid and ideally conducting loop is pulled to a new location. `Left loop should be` γ'.

Yet another variant of the experiment is to disconnect the ideally conducting loop from the external source. If we now push or pull the loop in a magnetic field, its current will change as there is no longer anything maintaining it. In particular, the

final current will in general be non-vanishing even if the initial current was zero. The reason is that by pulling the charge carriers in a magnetic field we subject them to the Lorentz force and thereby make them start moving in a transverse direction (and in a truly ideal conductor there is nothing that brings them back to rest).

Let us now alter the experimental strategy altogether. So far we've been assuming the magnetic field strength B to be frozen in time (if B changes with time the experiment still works in principle, but the measurements must then be done at a rate faster than the rate of change of B); given a fixed 1–chain of magnetic flux lines, we pushed around a line current and learned that its moving charges experienced a conservative force, the Lorentz force. Now we change the scheme and imagine moving the lines of B, while keeping the conducting wire in fixed position. This may be viewed as a mere change of reference frame, and by the principle of relativity we expect that the charge carriers of the conductor again experience a force.

A concrete setup is to use the loop geometry, and start with zero initial current in the conducting loop γ. If a magnet is then moved past the loop, the resulting forces set the mobile charge carriers of the conductor in motion, causing a transient current to circulate in the loop γ. By Ohm's law for a *real* conductor (current = voltage / resistance) that current translates back to a voltage.

Experiments of a similar type were carried out by Faraday in the early 1830's. What he discovered can be formulated in the following concise terms: *the voltage around a loop γ equals the negative rate of magnetic flux cutting across γ*:

$$\boxed{\int\limits_{\gamma} E = -\tfrac{\partial}{\partial t} \int\limits_{S} B} \qquad \text{(Faraday's law of induction)},$$

where S is any oriented surface bounded by γ. Note that the choice of surface S is arbitrary, as the magnetic flux is the same for all S with boundary $\partial S = \gamma$.

Note further that Faraday's law of induction a priori has nothing to do with conductors, or the voltage or current induced in a conductor; it is simply a statement relating the magnetic and electric field strengths. Conductors and induced currents are mentioned in connection with Faraday's law only because they serve the purpose of measurement and verification of the law.

We can get a slightly different perspective by taking instead of the loop γ the surface S as the given data. Using Stokes' theorem with $\gamma = \partial S$ on the left-hand side, we obtain $\int_S \mathrm{d}E = -\frac{\partial}{\partial t} \int_S B$. Since Faraday's law is true for any surface S, we deduce

$$\mathrm{d}E = -\dot{B} \, .$$

In words: the boundary of the 2–chain of E equals the negative time derivative of the 1–chain of B. In this formulation it is especially evident that Faraday's law makes no informative statement about the closed part of E. There may always exist closed surfaces (for example the equipotential surfaces of a static background electric field) in the 2–chain of E; they drop out of the expression $\mathrm{d}E = \partial E$ on the left-hand side. What Faraday's law states is that *non-closed* surface elements of E appear whenever and wherever B is time dependent.

How should we visualize that statement in the chain picture? The answer is that Faraday's law cannot be pictured quite so easily, as it involves the operation of taking a derivative, a notion tied to the continuum. This difficulty also alerts us to a problem with our naive chain formalism: differentiating a time dependent chain with respect to time, one gets an answer which in general is no longer a chain but a singular limit of chains (in fact, one gets what is called a "deRham current"). Strictly speaking, therefore, the equation $\mathrm{d}E = -\dot{B}$ is mathematically inconsistent as an equation for chains.

However, having already discretized in space, let us now go all the way and make time discrete, too, replacing the derivative $\dot{B}$ by a quotient of differences:

$$\dot{B}(t) \approx \frac{B(t + \Delta t/2) - B(t - \Delta t/2)}{\Delta t} \ .$$

If B is a chain, then so is the discrete-time approximation to $\dot{B}$, although the $\Delta t \to 0$ limit $\dot{B}$ no longer is. Thus, approximating the time derivative with a finite time step Δt gets us around the mathematical inconsistency.

Moreover, we can now draw a picture. If, at time t, there exists some surface element S of E which is not closed but has a boundary, then according to Faraday's law a change in B must be occurring near S in the time interval from $t - \Delta t/2$ to $t + \Delta t/2$. Let's say that S has the shape of a rectangle. Then a likely explanation is that some magnetic flux line with a kink in it (see Fig. 1.22) straddles the surface S by its two flux configurations at the times $t - \Delta t/2$ and $t + \Delta t/2$; in other words, the kink of the magnetic flux line relocates from one corner of S to the opposite corner during the said time interval. (Such a process may, for example, occur for a Josephson vortex in a layered superconductor when the magnetic field axis is tilted by a small angle with respect to the plane of the layers.) If the flux line carries flux ϕ, the discrete approximation to Faraday's law,

$$\Delta t \cdot \mathrm{d}E(t) = B(t - \Delta t/2) - B(t + \Delta t/2) \ ,$$

is satisfied if we assign to S the voltage $V = \phi/\Delta t$. Note that the outer orientation

of S is determined by the sense of direction in which the surface element S is pierced by the circulation of the magnetic flux line at the *earlier* time $t - \Delta t/2$.

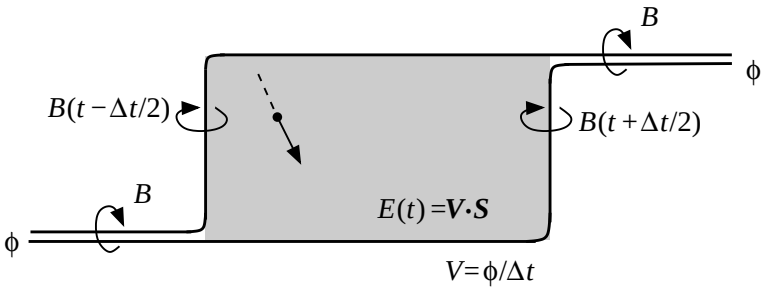

Figure 1.22: If the kink of a magnetic flux line carrying flux ϕ swishes over a surface S in time Δt, an electric field strength $E = (\phi/\Delta t)S$ is induced. In the example of the figure, the outer orientation of S points from back to front. `Label` S `is missing.`

We thus arrive at the intuitive notion that the world surfaces swept out by moving flux lines are instantaneous surface elements of E. For the pedagogical purpose of even better visualization we might imagine that the vacuum is a fluorescent medium – what a ridiculous concept! – and as magnetic flux lines move through it, the medium lights up, and what shines (only to return instantly to darkness) are the surface elements of E.

We now understand the "geometric" meaning of the statement (which is actually of purely differential, i.e., non-metric character) that "the voltage around a loop γ equals the negative rate of magnetic flux cutting across γ": every magnetic flux line crossing γ sweeps out a surface element pierced by γ; that surface element is part of the 2–chain of E; since the surface element is pierced by γ, it contributes to the intersection pairing $P(\gamma, E) = \int_\gamma E$, which is the voltage around γ.

The equations $\mathrm{d}B = 0$ and $\dot{B} + \mathrm{d}E = 0$ are called the *homogeneous Maxwell equations*. We shall take a closer look at them later and will then recognize that they are two different manifestations of a single fundamental principle.

2.6 Caveat: non-uniqueness of the chain picture

The current subsection is *not* addressed to second-year students of electrodynamics. Rather, it is injected for the benefit of those readers who already have some experience trying to visualize the electromagnetic field by "field lines, or gear wheels" [Feynman], and have concluded that you invariably run into some kind of contradiction or inconsistency, unless you adopt the abstract field idea...

3 Inhomogeneous Maxwell equations

There are two distinct aspects to the electromagnetic field and its interaction with matter: the influence of the field on the charges, and the reverse influence of the charges on the field. As was discussed in Section 2, the electromagnetic field exerts forces upon electric charges and electric currents. This first aspect is quantitatively expressed by the field strengths, E and B.

Our next theme is that electric charges and electric currents in turn give rise to electromagnetic fields. One should imagine that charges and currents in some sense "excite" the space surrounding them. This second aspect is captured and made quantitative by a second set of fields, the excitations D and H.

The excitations D and H are not independent quantities but can be expressed in terms of the field strengths E and B, and vice versa. That relation, however, is not universal! Rather, it depends on the geometry of space and time (it changes when going from flat space to the gravitationally curved space-time near a black hole; it is also modified in the formulation of an effective theory of electromagnetism in macroscopic media). We therefore see it fit to first introduce the excitations (and the equations obeyed by them) in their own right, postponing the details of their relation to the field strengths until later.

To establish the physical reality of the excitations D and H, we shall give ad hoc descriptions of how they can be measured, exploiting the static screening properties of metals and superconductors. For the time being we do not explain the physical basis of these measurements, and why they work and give the correct answer under suitable experimental conditions. This will be reviewed later, when all foundational issues have been settled.

Basic to the equations obeyed by D and H is the law of charge conservation, so that's where we must start.

3.1 Continuity equation

Recall the discrete model of the electric charge density ρ as a 0–chain $\rho = \sum q_i p_i$ consisting of point charges q_i at positions p_i. One is often interested in the total charge in a space domain U. That charge, $Q[U]$, is determined in the chain formalism by intersecting the 0–chain ρ with the 3–chain U:

$$\boxed{Q[U] = \int\limits_U \rho} \quad .$$

By definition, the intersection pairing $\int_U \rho \equiv P(\rho, U)$ is computed by identifying the points of ρ contained in U and simply summing up their charges.

A further simplification occurs in reality. Observations made in the microscopic world of atoms and subatomic particles show that electric charge is "quantized": it always comes as some integer multiple of an elementary unit of charge, $e = 1.602 \times 10^{-19}$ Coulomb; this is the magnitude of the charge of both the proton and the electron – the two constituents of the hydrogen atom. Given charge quantization, we can evaluate $Q[U]$ by just counting the number of positive and negative elementary charges in U, taking the difference, and multiplying by the charge quantum e.

Another basic fact of nature is *conservation of electric charge*. It has never been seen that electric charges materialize all of a sudden out of nothing or disappear into nothing. Although there exist fundamental processes by which N_i elementary particles can be converted into a different number $N_\mathrm{f} \neq N_\mathrm{i}$ of elementary particles, this always happens in such a way that equal amounts of positive and negative charge are created or annihilated. [see Figure]

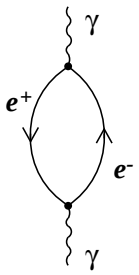

Figure 1.23: In quantum electrodynamics pairs of electrons (e^-) and positrons (e^+) may be created from a photon (γ), or may annihilate to form a photon. Electric charge is conserved in the process.

What does charge conservation imply for $Q[U]$, the total charge in U? The obvious answer for a fixed space domain U is that $Q[U]$ can change with time only by one kind of mechanism: by charges entering or leaving U through the surface ∂U. Consequently, there exists a physical quantity $I[\partial U]$, called the *electric current* through ∂U, such that

$$\dot{Q}[U] = -I[\partial U] \; .$$

The minus sign accounts for the fact that a positive charge crossing the surface ∂U from the inside of U to the outside – that's the direction of the canonical outer orientation of ∂U – makes a positive contribution to $I[\partial U]$, but a negative contribution to the rate of change of $Q[U]$.

The total current through ∂U results from summing up the charge displacements through all the parts of the surface ∂U. Thus $I[\partial U]$ is the result of integrating a local physical quantity, which is called the *electric current density* j. In the chain formalism, j is modelled by a 1–chain with inner orientation. (The line elements of j are obtained by a limit procedure which constructs the velocity vectors tangent to the trajectories of the charged particles and weights them by their charges). By definition of the 1–chain of j, the total electric current $I[S]$ through any surface S with outer orientation is given by the intersection pairing $P(j, S) = \int_S j$ between j and S:

$$\boxed{I[S] = \int\limits_S j} \qquad \text{(definition of } j\text{)}.$$

Thus the current $I[S]$ is determined by those line elements of j that cross S. That's precisely how it should be on physical grounds. Note, in particular, that all that matters for the intersection pairing $\int_S j$ is whether a line element of j does or does not cross S. In fact, the angle of the crossing (which enters in the vector calculus expression $I[S] = \int_S \vec{j} \cdot d\vec{S}$) has no bearing on the value of $I[S]$.

Recall that $\dot{Q}[U] = -I[\partial U]$, by the law of charge conservation. On inserting the integral expressions $Q[U] = \int_U \rho$ and $I[S] = \int_S j$, this conservation law takes the form

$$\boxed{\tfrac{\partial}{\partial t} \int\limits_U \rho = - \int\limits_{\partial U} j} \qquad \text{(continuity equation)}.$$

In words: *the rate of change of the electric charge in U equals the electric current through $-\partial U$*. This statement holds for any fixed (i.e. time independent) domain U. By Stokes' theorem (actually Gauss' theorem) $\int_{\partial U} j = \int_U \mathrm{d}j$, the continuity equation can be recast in the "differential" form

$$\dot{\rho} = -\mathrm{d}j \ ,$$

which in turn is equivalent to $\partial j = \dot{\rho}$.

The qualification made for the chain version of Faraday's law $\partial E = -\dot{B}$ in Section 2.5 applies equally well to the continuity equation $\partial j = \dot{\rho}$: strictly speaking, this equation makes sense as an equation for chains only after discretization of the time derivative by a finite time step Δt:

$$\Delta t \cdot \partial j(t) = \rho(t + \Delta t/2) - \rho(t - \Delta t/2) \ .$$

When written in this discrete form, the physical meaning of the continuity equation becomes particularly transparent: the line elements of $j(t)$ are the trajectories of the

point charges of ρ in the time interval $[t-\Delta t/2, t+\Delta t/2]$, weighted by $(\Delta t)^{-1}$; since these trajectories by continuity connect the points of $\rho(t-\Delta t/2)$ with the points of $\rho(t+\Delta t/2)$, the boundary of $\Delta t \cdot j(t)$ is the 0–chain $\rho(t+\Delta t/2) - \rho(t-\Delta t/2)$.

In the time-continuum limit $\Delta t \to 0$, the 1–chain model of j becomes singular: the spatial extensions of its line elements ($\sim \Delta t$) go to zero while their weights ($\sim (\Delta t)^{-1}$) go to infinity. In this limit the 1–chain of j becomes a so-called deRham current of degree one, see Section [Ref]. (As a matter of fact, it was the physical example of the electric current j that motivated deRham to give the name "current" to the mathematical objects now named after him.)

Later [Ref] we will take a look at the transformation behavior of j in the limit of continuous time and space, and will see that j transforms like a 2–form (of odd type), which is distinct in a subtle manner from the transformation behavior of a vector field.

3.2 Electric excitation D

A canonical method to measure static (or low-frequency) electric fields is an experimental setup known as "Maxwell's double plates", which will now be described.

Assuming that you are given two very thin metallic plates of identical size and shape, do the following. First of all, by connecting both plates with the ground, remove any excess charges from them to ensure electrical neutrality. Then disconnect the plates from the ground, and establish electrical contact *between* the two plates by placing them right on top of each other. Now move the two plates (while maintaining mutual contact) to the place where you want to make a measurement of the electric field. Once the plates have been installed in the desired location, separate them uniformly by a small distance. The metallic plates are now found to be charged; one of them carries charge q, the other charge $-q$. Roughly speaking, q is big if the electric field at the measurement location is strong, and is small if the electric field is weak.

To make this more quantitative, denote the surface occupied by the plates right before separation by S. Give the surface S an outer orientation by picking either one of the two transverse senses of direction. (It doesn't matter which you choose, but you must make some choice for book keeping purposes.) The surface S now has an "in" side and an "out" side. Correspondingly, the choice of orientation allows you to distinguish between the "in" and the "out" metallic plate. Measure the total charge on the "out" plate. Denoting the result by $Q'[S]$, you have determined an

assignment $S \mapsto Q'[S]$.

It is clear that such an experiment can in principle be done for any surface S with a reasonable shape, where "reasonable" means that one should be able to manufacture two identical, metallic, stackable, approximately two-dimensional bodies with the size and shape of S. Thus Maxwell's double plate experiment determines a function $S \mapsto Q'[S]$ on a large set of surfaces oriented by a transverse vector.

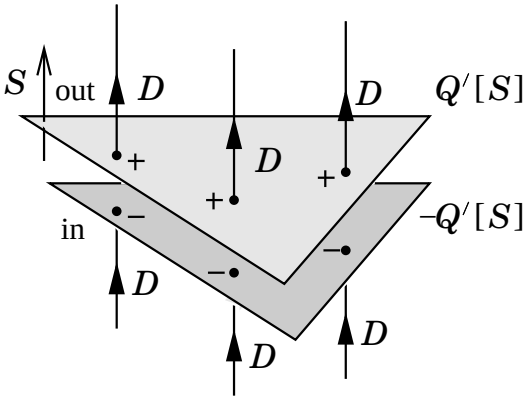

Figure 1.24: Maxwell's double plate experiment measures the electric flux $Q'[S]$ through a surface S.

In view of the experimental setup using metallic bodies, one might expect some dependence on the type of metal used, its preparation history, the temperature etc. However, this is not so; the outcome $Q'[S]$ of the experiment does not depend on any material properties (at least not in an ideal setup with very thin metallic plates and adiabatic conditions for the charge carriers in the metal). Rather, the result $Q'[S]$ of the measurement is an intrinsic property of the empty space at the surface S and the time of the measurement; it is called the *electric flux* through S.

Experimentally the electric flux function $S \mapsto Q'[S]$ turns out to be linear: if you divide the measuring surface into pieces, $S = \sum S_i$, you find that the charges add up: $Q'[S] = \sum Q'[S_i]$. Thus mathematically speaking $Q'[S]$ extends to a linear function(al) on the vector space of 2–chains with outer orientation.

Linearity tells us that the electric flux $Q'[S]$ arises from evaluating the integral over S of something or, in our discrete approximation, from computing the intersection pairing between S and some chain. We shall refer to that chain as the *electric excitation*, D:

$$\boxed{Q'[S] = \int_S D} \qquad \text{(definition of } D\text{)}.$$

Maxwell called D the electric displacement field. In order for the intersection pairing $\int_S D = P(D, S)$ to make sense and be invariantly defined, D must be chain modelled by a 1–chain with inner orientation. Clearly, D carries the absolute physical dimension of electric charge. Note that the definition of D does not depend on which orientation you chose for S: if the orientation of S is reversed, both $Q'[U]$ and $\int_S D$ change sign while D remains invariant.

Thus we distinguish between two manifestations of the electric field: the field strength E, and the excitation D. The electric field strength determines the voltage $V[\gamma] = \int_\gamma E$, while the electric excitation determines the electric flux $Q'[S] = \int_S D$. As will be explained later, D is related to E by the metric structure of space and a fundamental constant, the dielectric polarizability of the vacuum.

3.3 Gauss' law

We are now in a position to state another fundamental law of nature constraining the electric field. In Maxwell's double plate experiment the choice of surface S was quite arbitrary (it was only constrained by practicability of the experiment). In particular, the surface S might have been with or without boundary. For a closed surface, of course, doing the measurement in a single step may be difficult if not possible. (How would you first attach and then detach two concentric metallic spheres, say?) However, having established linearity of the electric flux we can always cut a closed surface S into two pieces S_+ and S_-, do the Maxwell double plate measurement separately on the pieces, and add the charges: $Q'[S] = Q'[S_+] + Q\prime[S_-]$ or, equivalently, $\int_S D = \int_{S_+} D + \int_{S_-} D$.

Let the surface S now be closed. In that case S is the boundary of some domain in space, U. As usual, the surface $S = \partial U$ gets an outer orientation by the transverse vector pointing from the inside of U to the outside. Thus when doing the Maxwell double plate measurement for the closed surface $S = \partial U$, the final state charge $Q'[S]$ is understood to be the charge on the *outer* one of the two metallic bodies.

By definition and claim, the charge $Q'[S]$ on the outer plate measures the electric flux $\int_S D$. It is natural to compare that charge with the total amount of electric charge contained in the domain U, which is the integral of the charge density, $\int_U \rho = Q[U]$. The fundamental law to be stated, namely Gauss' law, says that these two charges are always equal: $Q'[\partial U] = Q[U]$, or

$$\boxed{\int\limits_{\partial U} D = \int\limits_{U} \rho} \qquad \text{(Gauss' law).}$$

In words: *the electric flux through a closed surface $S = \partial U$ equals the total electric charge enclosed by S.*

Since $\int_{\partial U} D = \int_U \mathrm{d}D$ by Stokes' theorem (actually, Gauss' theorem) and since the choice of domain U is arbitrary, we can recast Gauss' law as

$$\mathrm{d}D = \rho \, .$$

An equivalent statement is $\partial D = -\rho$, which says that the boundary of the 1–chain of D is the negative of the 0–chain of ρ. That statement is interpreted as follows.

The electric excitation D, being modelled as a 1–chain with inner orientation, is a superposition $D = \sum q_i \gamma_i$ of directed curves γ_i, each carrying an electric flux q_i. We may assume all q_i to be positive (if $q_i < 0$ we just reverse the direction, $\gamma_i \to -\gamma_i$, and use $q_i \gamma_i = (-q_i) \cdot (-\gamma_i)$). Gauss' law is compatible with some curves of D being cycles ($\partial \gamma_i = 0$). More informatively, if a curve γ_i carrying electric flux q_i is not a cycle, Gauss' law $-\rho = \partial D = \sum q_i \partial \gamma_i$ dictates that γ_i must begin on a positive charge q_i and end on a negative charge $-q_i$. (Recall that the boundary of a directed curve is given by the final point minus the initial point.)

Notice that Gauss' law by itself does not determine the electric excitation D. Given some charge density ρ, there exist infinitely many choices of D that satisfy $\partial D = -\rho$. Figure [Fig] shows a very simple configuration of an electric flux line and two charges consistent with Gauss' law. Note that if the two charges are static, the figure does not reflect physical reality. (Many configurations of an electric flux line must be superimposed to build up a static electric dipole field.)

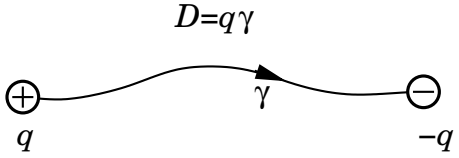

Figure 1.25: "Charges are sources of electric excitation".

3.4 Magnetic excitation H

We have met two incarnations of the electric field: one as a 2–chain with outer orientation (E), and another one as a 1–chain with inner orientation (D). On the magnetic side we have only seen the magnetic field strength B so far; the setting will now be completed by establishing a magnetic partner for it. Since B is modelled

by a 1–chain with outer orientation, symmetry makes us expect its partner to be a 2–chain with inner orientation. This expectation is in fact borne out. The magnetic partner of B is called the *magnetic excitation* H. It can in principle be measured by a procedure close in spirit to Maxwell's double plate experiment, as follows.

The basic piece of equipment to use is a long cylinder or rod (which we refer to as the "device" for short) cooled down to a very low temperature, where the material becomes superconducting and exhibits the Meissner effect. The precise shape of the device does not matter. What is important is that it be approximately one-dimensional, i.e. the cylinder or rod should have a small cross section.

We are not going to explain here what a superconductor in the Meissner phase is, except to say that it expels magnetic flux from its interior for energetic reasons (which have to do with the coherent nature of its macroscopically ordered quantum ground state). The flux expulsion is effected by dissipationless electric currents circulating around the surface of the superconductor.

To prepare the device, let its temperature rise above the point where it goes normal-conducting. Then take the device to a location devoid of magnetic field and, there, cool it down into the superconducting state (this is just a precautionary step to get rid of trapped flux and long-lived supercurrents that might otherwise corrupt the outcome of the measurement).

Start the experiment by moving the superconducting device to the location where you want to make a measurement of the magnetic field. If a magnetic field is present there, electric (super)currents will now be flowing near the surface of the superconductor. Measure the total current circulating around the device. (This may not be so easy to do directly, but should be possible via the so-called Coulomb drag effect. In any case, our "experiment" still makes good sense as a gedanken-experiment.) The measured current provides the information we are seeking.

Let us put this in succinct mathematical terms. Because the device is approximately one-dimensional, it will suffice to idealize it as a curve. When placed at the location of the measurement, the device occupies some specific curve, say γ. To do the book keeping, let γ be given an outer orientation by some sense of circulation. The surface supercurrent then counts as positive if it circulates in accordance with the circulation of γ, else as negative. In this way the gedanken-measurement associates with the oriented curve γ a well-defined current $I'[\gamma]$. Although the measurement involves a superconductor, the measured current $I'[\gamma]$ actually does *not* depend (under ideal circumstances) on the specific properties of the superconducting material used, but reflects an intrinsic property of space.

Repeating the measurement for all kinds of curves determines a function $\gamma \mapsto I'[\gamma]$ on 1–chains with outer orientation. Because electromagnetism is a linear theory, this function once again turns out to be linear. Apart from an overall sign, it is the integral of the magnetic excitation H along γ:

$$\boxed{I'[\gamma] = -\int_\gamma H} \qquad \text{(definition of } H\text{)}.$$

By this definition H carries the absolute physical dimension of electric current. As always, the integral is to be interpreted as an intersection pairing: $\int_\gamma H = P(H, \gamma)$. The pairing is well-defined in general, as 2–chains H with outer orientation and 1–chains γ with inner orientation generically intersect (in three space dimensions) in a finite number of points. The definition of H, once again, does not depend on the choice of orientation for the curve γ: reversing the orientation of γ changes the sign of both $I'[\gamma]$ and $\int_\gamma H$ while preserving H.

In summary, there exist two distinct manifestations of the magnetic field: the field strength B, and the excitation H. As we shall see, they are related to each other by the metric structure of space and a fundamental constant, the magnetic permeability of the vacuum.

3.5 Ampere's law

Spurred by Oersted's observation (1820) that a magnetic needle is deflected by a varying electric current, Ampere carried out a flurry of experiments on currents and the forces acting between them. This research led to the formulation of what is now known as Ampere's law. To cut a long story short, we shall motivate Ampere's law in the context of the gedanken-experiment described in the previous section.

While the choice of curve γ used in our definition of the magnetic excitation H was quite arbitrary, we now take γ to be closed. So let γ be some fixed loop in space and, as before, let γ be given an outer orientation by some sense of circulation. We can imagine doing two different things. On the one hand, we can manufacture a superconductor of suitable size and shape, place it at the loop γ, and measure the circulating supercurrent, $I'[\gamma]$. On the other hand, putting the superconductor aside and picking some surface S with outer orientation and boundary $\partial S = \gamma$, we can measure the current $I[S]$ flowing through S. Thus there are two currents, which we can compare: the supercurrent circulation $I'[\gamma]$ around the loop $\gamma = \partial S$, and the (normal) current $I[S]$ crossing the surface S. Ampere's law states that one

current is the negative of the other:

$$-I'[\partial S] = \int_{\partial S} H = \int_S j = I[S] \qquad \text{(Ampere's law)}.$$

Thus $I[S] + I'[\partial S] = 0$ or, in other words, the superconducting current flow exactly "screens" (or compensates) the ordinary current flow.

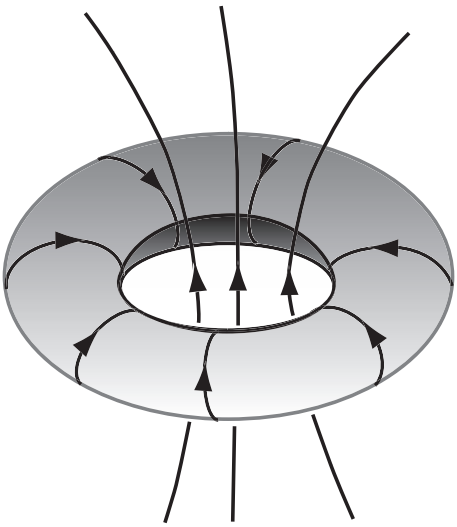

Figure 1.26: An electrical current flowing through the "hole" of a superconducting torus is compensated by supercurrents circulating around the torus.

We should emphasize, however, that the superconductor is not what Ampere's law is about (in fact, superconductors were not discovered until 90 years after Ampere's work). While screening by supercurrents is a convenient tool to *measure* the magnetic excitation, H already exists in the absence of the measuring device. Thus Ampere's law is really just meant as a statement that relates the electric current flow through S to the ring integral of the magnetic excitation around ∂S.

Ampere's law makes sense in the static limit but, as was pointed out by Maxwell, it becomes inconsistent for more general, time dependent situations. Indeed, in the latter cases the currents through two surfaces S_1 and S_2 will in general be different, $I[S_1] \neq I[S_2]$, even if the boundaries $\partial S_1 = \partial S_2 = \gamma$ are the same. We then face a dilemma: the left-hand side of Ampere's law depends only on the boundary of S, whereas the right-hand side depends on S itself; a clear contradiction!

If the lines of the electric current j are closed, however, the problem of surface dependence goes away. In that case, deforming the surface S (while keeping the boundary $\partial S = \gamma$ fixed) just shifts the points of intersection of j and S without changing the intersection pairing $\int_S j = P(j, S)$. The lines of j are closed in the static limit, as follows from $\partial j = \dot{\rho} = 0$. This is the limit where Ampere's law makes

sense and, as a matter of fact, holds true.

3.6 Ampere-Maxwell law

It was Maxwell who had the crucial insight (ca 1861) that, in order to correct Ampere's law and get an equation that makes sense in dynamical situations, the integrand j on the right-hand side must be replaced by $j + \dot{D}$. This modification removes the surface dependence of the integral. Indeed, the sum $j + \dot{D}$ is closed by Gauss' law and the continuity equation:

$$\mathrm{d}(j + \dot{D}) = \mathrm{d}j + \frac{\partial}{\partial t}\mathrm{d}D = -\dot{\rho} + \dot{\rho} = 0 \ .$$

Therefore, if S_1 and S_2 are two surfaces with a common boundary $\partial S_1 = \partial S_2 = \gamma$, which guarantees the existence of a domain U with $\partial U = S_1 - S_2$ by the Poincare lemma, Stokes' theorem gives

$$\int\limits_{S_1 - S_2} (j + \dot{D}) = \int\limits_{\partial U} (j + \dot{D}) = \int\limits_{U} \mathrm{d}(j + \dot{D}) = 0 \ ,$$

and the desired surface independence $\int_{S_1}(j + \dot{D}) = \int_{S_2}(j + \dot{D})$ follows.

Notice that forming the sum $j + \dot{D}$ makes mathematical and physical sense, as both j and $\dot{D}$ are (singular) 1–chains with inner orientation and the physical dimension of current. Maxwell called $\dot{D}$ the displacement current.

His modification of Ampere's law in integral form reads:

$$\boxed{\int\limits_{\partial S} H = \int\limits_{S} (j + \dot{D})} \qquad \text{(Ampere-Maxwell law)}.$$

The equivalent "differential" version again follows on rewriting the left-hand side by Stokes' law, $\int_{\partial S} H = \int_S \mathrm{d}H$, and exploiting the arbitrariness of S:

$$\mathrm{d}H = j + \dot{D} \ .$$

In the language of chains, this says that the boundary of the 2–chain of the magnetic excitation H agrees with the sum of the 1–chains of the electric current j and the displacement current $\dot{D}$.

To draw a picture, let us again make time discrete, approximate $\dot{D}$ by a quotient of differences, and multiply the equation by the time step Δt:

$$\Delta t \cdot \partial H(t) = \Delta t \cdot j(t) + D(t + \Delta t/2) - D(t - \Delta t/2) \ .$$

Then, suppose the 2–chain of H at time t contains a non-closed surface element S. If S has the shape of a parallelogram or similar and carries weight $I = q/\Delta t$,

the following may be happening. In the time interval from $t - \Delta/2$ to $t + \Delta t/2$ an electric flux line with flux q is sweeping out the surface S [see Figure]. To satisfy Gauss' law, charges $\pm q$ are attached to the ends of the flux line. The positive charge contributes to j a line element with inner orientation pointing in the direction of motion of the flux line, while the negative charge contributes a line element with the opposite orientation. Reversing the direction of the flux line at the earlier time (to account for the minus sign in the discretization of $\dot{D}$), we see that the process at hand contributes a cycle $q\partial S$ to the 1–chain $\Delta t \cdot j(t) + D(t + \Delta t/2) - D(t - \Delta t/2)$. That cycle is the boundary of the surface element $\Delta t \cdot (q/\Delta t)S$ in $\Delta t \cdot H(t)$, in agreement with the time-discrete form of the Ampere-Maxwell law.

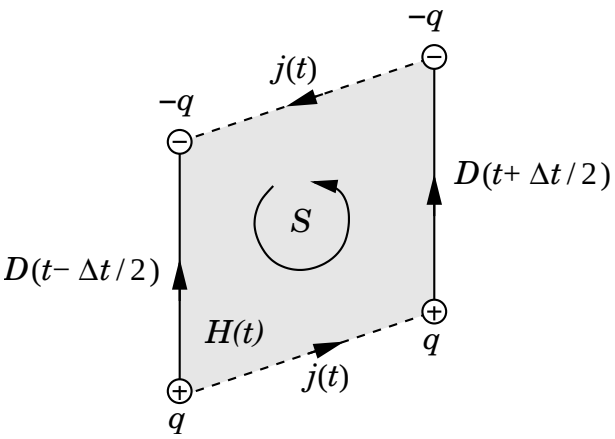

Figure 1.27: Illustration of the Ampere-Maxwell law: a surface S swept out by a flux line of D is an instantaneous surface of the magnetic excitation H.

Gauss' law $\mathrm{d}D = \rho$ and the Ampere-Maxwell law $\mathrm{d}H = j + \dot{D}$ are called the *inhomogeneous Maxwell equations.*