13 Exact forms, closed forms, loops and electromagnetism

Mathematics and physics help each other to understand and use different behavior of differential forms in simply connected and multiply connected domains.

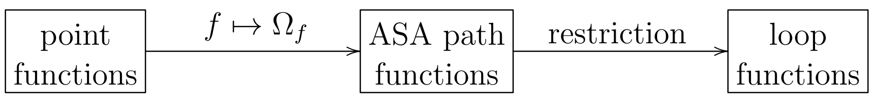
13a Path functions and loop functions

A closed path is also called a loop. By a loop function we mean a (real-valued) function on the set of all loops (in $\mathbb{R}^n$ or a given subset of $\mathbb{R}^n$). Every path function leads to a loop function (just restriction).

In particular, every 1-form leads to (a path function and) a loop function. For example, the winding number (recall Sect. 10d) is a loop function over $\mathbb{R}^2 \setminus \{0\}$; it corresponds to the 1-form $\frac{-y\,dx+x\,dy}{x^2+y^2}$ (and many others, as we'll see soon). Another interesting loop function (over $\mathbb{R}^2$) corresponds to the 1-form $-y\,dx + x\,dy$ (also discussed in Sect. 10d).

We restrict ourselves to additive stationary antisymmetric (ASA, for short) path functions.

Every function $f : \mathbb{R}^n \to \mathbb{R}$ (that is, point function) leads to an ASA path function Ω_f defined by $\Omega_f(\gamma) = f(\gamma(t_1)) - f(\gamma(t_0))$ whenever $\gamma : [t_0, t_1] \to \mathbb{R}^n$. If f is continuous, we may treat it as a 0-form and write $\Omega_f(\gamma) = \int_{\partial\gamma} f$. If $f \in C^1$ then $\Omega_f(\gamma) = \int_\gamma df$ (recall 11c3). But for now f is arbitrary. We have a diagram of linear mappings between vector spaces:



The composition of these two mappings is zero (just because $\gamma(t_1) = \gamma(t_0)$ implies $f(\gamma(t_1)) = f(\gamma(t_0))$). That is, the image of the first mapping is contained in the kernel of the second mapping. Interestingly, they are equal.

13a1 Lemma. An ASA path function Ω over $\mathbb{R}^n$ vanishes on all loops if and only if $\Omega = \Omega_f$ for some f.

Proof. "If": see above. "Only if": we define $f(x) = \Omega(\gamma_x)$ where $\gamma_x(t) = tx$ for $t \in [0,1]$. Given $\gamma : [t_0, t_1] \to \mathbb{R}^n$, $\gamma(t_0) = x$, $\gamma(t_1) = y$, the concatenation[1] $\gamma_x \cdot \gamma \cdot (\gamma_y)_{-1}$ is a loop, thus $0 = \Omega(\gamma_x \cdot \gamma \cdot (\gamma_y)_{-1}) = \Omega(\gamma_x) + \Omega(\gamma) - \Omega(\gamma_y)$, whence $\Omega(\gamma) = f(y) - f(x)$. □

Our choice of the paths γ_x does not really matter; any path γ_x from 0 to x works equally well, and gives the same result (think, why).

The same holds over a connected open subset of $\mathbb{R}^n$ (for instance, $\mathbb{R}^n \setminus \{0\}$).[2]

Given Ω, the function f is unique up to an additive constant. Proof: if $f(\gamma(t_1)) - f(\gamma(t_0)) = g(\gamma(t_1)) - g(\gamma(t_0))$ for all γ, then $f(\gamma(t_1)) - g(\gamma(t_1)) = f(\gamma(t_0)) - g(\gamma(t_0))$ for all γ, that is, $f - g = \text{const}$.

We specialize 13a1 to "good" path functions that correspond to 1-forms, $\Omega(\gamma) = \int_\gamma \omega$ (they all are ASA, of course).

13a2 Lemma. A 1-form ω on $\mathbb{R}^n$ satisfies $\int_\gamma \omega = 0$ for all loops γ if and only if $\omega = df$ for some $f \in C^1$.

Proof. "If": for $f \in C^1$ we have $\int_\gamma df = \Omega_f(\gamma) = 0$ for all loops γ.

"Only if": Lemma 13a1 applied to the path function $\Omega : \gamma \mapsto \int_\gamma \omega$ gives f such that $\int_\gamma \omega = \Omega_f(\gamma)$ for all paths γ. We take a straight path from x to $x+h$ (for arbitrary $x, h \in \mathbb{R}^n$) and get $f(x + \varepsilon h) - f(x) = \int_0^\varepsilon \omega(x + th, h)\, dt = \varepsilon\omega(x, h) + o(\varepsilon)$, that is, $(D_h f)_x = \omega(x, h)$, which shows that $f \in C^1$ and $df = \omega$. □

The same holds over an open subset of $\mathbb{R}^n$.

13b Exact forms and closed forms

13b1 Definition. A 1-form ω on an open set $G \subset \mathbb{R}^n$ is *exact,* if $\omega = df$ for some $f \in C^1(G)$.

Here is a reformulation of Lemma 13a2.

13b2 Corollary. A 1-form ω on G is exact if and only if $\int_\gamma \omega = 0$ for all loops γ in G.

[1] Take the stationarity into account...

[2] Also, over arbitrary open set, and moreover, arbitrary set; just choose a point in every path connected component. In this case f is unique up to a function constant on every component.

13b3 Exercise. (a) A 1-form $\omega = f\big(\sqrt{x^2+y^2+z^2}\big)(x\,dx + y\,dy + z\,dz)$ on $\mathbb{R}^3 \setminus \{(0,0,0)\}$ is exact for every continuous $f : (0,\infty) \to \mathbb{R}$. Prove it.

(b) Prove that

$$\int_\gamma \frac{x\,dx + y\,dy + z\,dz}{\sqrt{x^2+y^2+z^2}} = \int_\gamma (x\,dx + y\,dy + z\,dz)$$

for every $\gamma : [t_0, t_1] \to \mathbb{R}^3 \setminus \{(0,0,0)\}$ such that $\gamma(t_0) = \big(\frac{3}{10}, 0, -\frac{2}{5}\big)$ and $\gamma(t_1) = \big(0, \frac{6}{5}, \frac{9}{10}\big)$.

13b4 Definition. A 1-form ω of class C^1 on an open set $G \subset \mathbb{R}^n$ is *closed,* if $d\omega = 0$.

Every exact 1-form (of class C^1) is closed, since $d(df) = 0$ by (11e4).

The form $\frac{-y\,dx + x\,dy}{x^2+y^2}$ is closed (since its restriction to U_i is exact for $i = 1, 2, 3, 4$, recall 10d) but not exact (by 13b2).

13b5 Exercise. [1] Let $\omega = f_1\,dx + f_2\,dy + f_3\,dz$ be a closed 1-form on $\mathbb{R}^3 \setminus \{(0,0,0)\}$, whose coefficient functions f_1, f_2, f_3 are homogeneous of degree $k \ne -1$ (that is, $f_i(tx, ty, tz) = t^k f(x,y,z)$). Prove that $\omega = dg$ where $g = \frac{1}{k+1}(xf_1 + yf_2 + zf_3)$.[2]

13b6 Lemma. A 1-form ω of class C^1 on G is closed if and only if $\int_{\partial\Gamma} \omega = 0$ for all singular 2-boxes Γ in G.

Proof. "Only if": $\int_{\partial\Gamma} \omega = \int_\Gamma d\omega = 0$; "if": $\int_\Gamma d\omega = \int_{\partial\Gamma} \omega = 0$ for all Γ, therefore $d\omega = 0$ by the argument on page 179: $\frac{1}{\varepsilon^2}\int_{\Gamma_\varepsilon} d\omega \to (d\omega)(x,h,k)$. $\square$

We have again a diagram of linear mappings between vector spaces:

$$\boxed{\begin{matrix}\text{0-forms}\\ \text{of class } C^2\end{matrix}} \xrightarrow{\quad d \quad} \boxed{\begin{matrix}\text{1-forms}\\ \text{of class } C^1\end{matrix}} \xrightarrow{\quad d \quad} \boxed{\begin{matrix}\text{2-forms}\\ \text{of class } C^0\end{matrix}}$$

The composition of these two mappings is zero. That is, the image of the first mapping (exact forms) is contained in the kernel of the second mapping (closed forms). They are equal for some G, not for all G.

Homotopy

It was noted (in Sect. 10e) that a singular 2-box may be thought of as a path in the space of paths. And now we need a path in the space of loops.

[1] Shurman, Ex. 9.11.1.

[2] Hint: first check the dx term of dg, remembering that ω is closed; a homogeneous function must satisfy Euler's identity $xD_1f + yD_2f + zD_3f = kf$.

13b7 Definition. Let $\gamma_1, \gamma_2 \in C^1([t_0, t_1] \to G)$ be loops in an open set $G \subset \mathbb{R}^n$.

(a) A *homotopy*[1] between γ_1 and γ_2 (in G) is a mapping $\Gamma \in C^1([t_0, t_1] \times [0, 1] \to G)$ such that

$$\Gamma(t, 0) = \gamma_1(t)\,, \quad \Gamma(t, 1) = \gamma_2(t) \quad \text{for all } t \in [t_0, t_1]\,;$$
$$\Gamma(t_0, u) = \Gamma(t_1, u) \quad \text{for all } u \in [0, 1]\,;$$

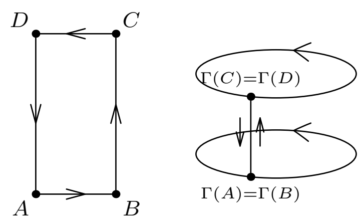
(b) γ_1 and γ_2 are *homotopic,* if there exists a homotopy between them;

(c) γ_1 is *null homotopic,* if it is homotopic to a trivial γ_2 (that is, $\gamma_2(\cdot) = \text{const}$).

13b8 Exercise. Prove that "homotopic" is an equivalence relation.[2]

13b9 Proposition. If loops γ_1, γ_2 in an open set $G \subset \mathbb{R}^n$ are homotopic in G then $\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$ for all closed 1-forms ω on G.

Proof. We take a homotopy Γ between γ_1 and γ_2; $\int_{\partial\Gamma} \omega = 0$ by 13b6. It remains to note that $\int_{\partial\Gamma} \omega = \int_{\gamma_1} \omega - \int_{\gamma_2} \omega$, since $\Gamma|_{[t_0,t_1]\times\{0\}} \sim \gamma_1$, $\Gamma|_{[t_0,t_1]\times\{1\}} \sim \gamma_2$, and $\Gamma|_{\{t_0\}\times[0,1]} \sim \Gamma|_{\{t_1\}\times[0,1]}$. $\square$



13b10 Exercise. Let $G = \mathbb{C} \setminus \{0\}$ be the punctured complex plane. Prove that

(a) every loop $\gamma \in C^1([0, 1] \to G)$ may be written as $\gamma(t) = r(t)\mathrm{e}^{\mathrm{i}\theta(t)}$, $r \in C^1([0, 1] \to (0, \infty))$, $\varphi \in C^1([0, 1] \to \mathbb{R})$;[3]

(b) the loop $t \mapsto r(t)\mathrm{e}^{\mathrm{i}\theta(t)}$ is homotopic (in G) to the loop $t \mapsto \mathrm{e}^{\mathrm{i}\theta(t)}$;

(c) the loop $t \mapsto \mathrm{e}^{\mathrm{i}\theta(t)}$ is homotopic (in G) to the loop $t \mapsto \mathrm{e}^{2\pi\mathrm{i}Nt}$, $N = (\theta(1) - \theta(0))/(2\pi)$;

(c) two loops $t \mapsto \mathrm{e}^{2\pi\mathrm{i}N_1 t}$, $t \mapsto \mathrm{e}^{2\pi\mathrm{i}N_2 t}$ are homotopic if and only if $N_1 = N_2$.

13b11 Exercise. Let G be $\mathbb{R}^3$ without the (union of the) three coordinate

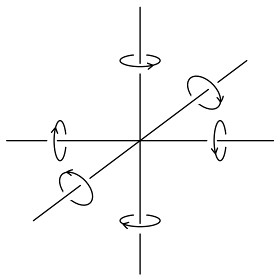
[1] Namely, a homotopy of class C^1.

[2] Beware of C^1 when proving transitivity; try
$\Gamma(t, u) = \begin{cases} \Gamma_1(t, 1 - (1 - 2u)^2) & \text{for } u \le 1/2, \\ \Gamma_2(t, (2u - 1)^2) & \text{for } u \ge 1/2. \end{cases}$.

[3] Hint: $\theta(t) - \theta(0) = \int_{\gamma|_{[0,t]}} \frac{-y\,dx + x\,dy}{x^2 + y^2}$.

axes, and ω a closed 1-form on G. Prove that

$$\int_{\gamma_1} \omega + \dots + \int_{\gamma_6} \omega = 0$$



where $\gamma_1, \dots, \gamma_6$ are the circles shown on the picture.[1]

13b12 Exercise. Let us define a circle (in $\mathbb{R}^3$) as such a path: $\gamma : [0, 2\pi] \to \mathbb{R}^3$, $\gamma(t) = a + b \cos t + c \sin t$, where $a, b, c \in \mathbb{R}^3$, $|b| = |c| > 0$ and $\langle b, c \rangle = 0$. Let G be $\mathbb{R}^3$ without the three coordinate axes (as in the previous exercise). Classify circles in G up to homotopy (in G). (Intuitive explanation is enough; no need to prove it.)

Answer: $2 \cdot (3^3 - 1) + 1 = 53$ homotopy classes.[2]

13b13 Exercise. (a) Let $G \subset \mathbb{R}^2$ be such that $\mathbb{R}^2 \setminus G$ is a finite set, of m points. Consider *simple* loops in G and their homotopy classes. Count these classes. (Intuitive explanation is enough; no need to prove it.)
Answer: $2^{m+1} - 1$ homotopy classes.

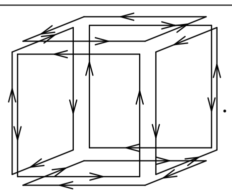
(b) The same for $G \subset S^2$.
Answer: $2^m - 1$ homotopy classes (for $m > 0$).

(c) The same for $G \subset \mathbb{R}^3$ such that $\mathbb{R}^3 \setminus G$ is the union of m (pairwise different) rays from the origin.
Answer: $2^m - 1$ homotopy classes (for $m > 0$).

13b14 Corollary. (to 13b9) If γ is null homotopic in G then $\int_\gamma \omega = 0$ for all closed 1-forms ω on G.

13b15 Definition. A connected open set $G \subset \mathbb{R}^n$ is *simply connected,* if every loop (of class C^1) in G is null homotopic.

---

[1]Hint:  .

[2]Hint: maybe it is easier to do the next exercise first; also, 3^3 means: $\{x\text{-axis}, y\text{-axis}, z\text{-axis}\} \to \{\text{negative part}, \text{positive part}, \text{neither}\}$.

13b16 Proposition. Every closed 1-form ω on a simply connected G is exact.

Proof. By 13b14, $\int_\gamma \omega = 0$ for all loops γ in G; by 13b2, ω is exact. $\square$

If G is convex, then it is simply connected (think, why).

Simple connectivity is preserved by diffeomorphisms[1] (think, why).

The punctured space $\mathbb{R}^n \setminus \{0\}$ is simply connected for $n > 2$ (since a punctured sphere is diffeomorphic to a vector space) but not for $n = 2$ (by 13b14).

A 1-form (of class C^1) on G is closed if and only if it is locally exact (since every point of G has a convex neighborhood in G).

Homology and cohomology

Let $G \subset \mathbb{R}^n$ be an open set. Two closed 1-forms on G are called *cohomologous,* if their difference is exact.

All closed 1-forms on G are a vector space; all exact 1-forms (of class C^1) are its subspace; the quotient space (closed)/(exact) consists of all equivalence classes $[\omega] = \{\omega + \alpha : \alpha \text{ exact}\} = \{\omega + df : f \in C^1(G)\}$; these classes are called *cohomology classes*[2] of G.

A 1-chain in G is called 1-*boundary* (in G), if it is the boundary of some 2-chain (in G).

A 1-*cycle* in G is, by definition, a 1-chain in G whose boundary is zero.

Every 1-boundary is a 1-cycle by (11d1).

Again, a diagram of linear mappings between vector spaces:

$$\boxed{\text{2-chains}} \xrightarrow{\quad\partial\quad} \boxed{\text{1-chains}} \xrightarrow{\quad\partial\quad} \boxed{\text{0-chains}}$$

The composition of these two mappings is zero. That is, the image of the first mapping (1-boundaries) is contained in the kernel of the second mapping (1-cycles). They are equal for some G, not for all G.

Two 1-cycles on G are called *homologous,* if their difference is 1-boundary.

All 1-cycles in G are a vector space; all 1-boundaries are its subspace; the quotient space (cycles)/(boundaries) consists of all equivalence classes $[C] = \{C + B : B \text{ boundary}\}$; these classes are called *homology classes*[3] of G.

If ω is closed, α is exact, C is a cycle and B is a boundary then

$$\int_{C+B} \omega + \alpha = \int_C \omega + \int_C \alpha + \int_B \omega + \int_B \alpha = \int_C \omega \,,$$

[1] Also by homeomorphisms; I do not prove it.

[2] De Rham cohomology classes of dimension 1.

[3] Singular homology classes of dimension 1.

since $\int_C \alpha = \int_C df = \int_{\partial C} f = 0$, and $\int_B \omega = 0$ by 13b6 (and $\int_B \alpha = 0$ by both reasons). Thus, $\int_{[C]} [\omega]$ is well-defined, and is a bilinear function of a homology class $[C]$ and a cohomology class $[\omega]$.

By 13b2,

$$[\omega] = [0] \quad \Longleftrightarrow \quad \forall [C] \;\; \int_{[C]} [\omega] = 0$$

(think, why). The numbers $\int_{[C]} [\omega]$ are traditionally called periods of ω. We see that a closed 1-form is exact if and only if all its periods are zero. This is a special case of the first part of De Rham theorem.[1] Its second part states, in particular, that

$$[C] = [0] \quad \Longleftrightarrow \quad \forall [\omega] \;\; \int_{[C]} [\omega] = 0 \,.$$

For example, the punctured plane $G = \mathbb{R}^2 \setminus \{0\}$ has a one-dimensional space of homology classes and one-dimensional space of cohomology classes. Deleting m points from the plane we get m-dimensional spaces of homologies and cohomologies. But for $\mathbb{R}^3 \setminus \{0\}$ they are trivial (0-dimensional).

Exact 2-forms and loop functions

13b17 Definition. A 2-form ω on an open set $G \subset \mathbb{R}^n$ is *exact,* if $\omega = d\alpha$ for some 1-form α of class C^1 on G.

13b18 Exercise. If α is a closed 1-form and β is an exact 1-form then the 2-form $\alpha \wedge \beta$ is exact.

Prove it.

If ω is exact then $\int_\Gamma \omega$ depends only on $\partial\Gamma$; that is, $\int_{\Gamma_1} \omega = \int_{\Gamma_2} \omega$ whenever singular 2-boxes Γ_1, Γ_2 satisfy $\partial\Gamma_1 = \partial\Gamma_2$. And more generally, $\int_{C_1} \omega = \int_{C_2} \omega$ whenever 2-chaines C_1, C_2 satisfy $\partial C_1 = \partial C_2$. The proof is immediate: $\int_{C_1} \omega = \int_{C_1} d\alpha = \int_{\partial C_1} \alpha = \int_{\partial C_2} \alpha = \int_{C_2} d\alpha = \int_{C_2} \omega$.[2]

Every 1-form ω on G leads to a loop function $\gamma \mapsto \int_\gamma \omega$. By 13b2, this loop function is trivial (zero) if and only if ω is exact. Thus, each equivalence class $\{\omega + \alpha : \alpha \text{ exact}\} = \{\omega + df : f \in C^1(G)\}$ leads to a loop function. (Not a cohomology class, unless ω is closed.)

If two loops γ_1, γ_2 are homotopic in G, then the 1-chain $\gamma_1 - \gamma_2$ is equivalent to some 1-boundary $\partial\Gamma$ by the argument of the proof of 13b9. Thus, the difference $\int_{\gamma_1} \omega - \int_{\gamma_2} \omega = \int_{\partial\Gamma} \omega = \int_\Gamma d\omega$ is uniquely determined by the exact

[1] If you want to know more about this deep theorem, see Nicolaescu, Wikipedia.

[2] The converse holds by (the first part of) De Rham theorem for 2-forms.

2-form $d\omega$. In particular, $d\omega$ determines uniquely the loop function on all null homotopic loops. We see that an exact 2-form leads to a loop function, provided that G is simply connected; here is a formal statement.

13b19 Proposition. If ω is an exact 2-form on a simply connected open set $G \subset \mathbb{R}^n$, then for every loop γ in G, $\int_\gamma \alpha$ does not depend on the choice of α such that $d\alpha = \omega$.

A different situation appears when G is not simply connected. For example, the 1-form $\omega = \frac{1}{2\pi} \frac{-y\,dx + x\,dy}{x^2+y^2}$ on $G = \mathbb{R}^2 \setminus \{0\}$ leads to the loop function called the winding number (recall 10d). In this case $d\omega = 0$; accordingly, homotopic loops have equal winding numbers, and null homotopic loops have winding number zero. In order to know the whole loop function we need also the period of ω.

In this example $G = G_1 \cup G_2 = (U_1 \cup U_2) \cup (U_3 \cup U_4)$ (U_i being as in 10d); restrictions $\omega|_{G_1}$, $\omega|_{G_2}$ are exact, but $\omega|_{G_1 \cup G_2} = \omega$ is not. The corresponding loop function (winding number) is trivial on G_1 and G_2 but nontrivial on $G_1 \cup G_2$ (which never happens to usual functions).

13c Electrostatics

Greatness of the electromagnetic theory cannot be overestimated. It unites many seemingly unrelated phenomena, such as these:

* amber attracts lightweight particles;
* magnetic compass points to the north;
* solid body stiffness (and Lorentz contraction…);
* light;
* radio waves (radio, TV, WiFi, …);
* X-rays (computed tomography…).

Still, mechanics and optics are (more or less) separate branches of physics; and moreover, electrostatics, magnetostatics and electromagnetic waves are branches of electromagnetism, which is quite practical, especially for engineers. However, the situation is changing; engineers reconsider such notions as potential difference and electromotive force[1] and try the language of differential forms;[2] philosophers discuss the ontological status of loop functions.[3]

We begin with electrostatics.

[1]Kirk T. McDonald (2012) "Voltage drop, potential difference and $\mathcal{EMF}$".
[2]G.A. Deschamps (1981) "Electromagnetics and differential forms".
[3]A. Afriat (2013) "Is the world made of loops?"

Electrostatics is mathematically very similar to Newtonian gravitation treated in Sect. 9g. By Coulomb's law, the electrostatic force exerted by a particle of charge q at point ξ on a particle of charge q_0 at point x is $\frac{q_0 q}{\varepsilon_0} E_\xi(x)$, and $\frac{q}{\varepsilon_0} E_\xi(\cdot)$ is the electric field generated by q,

$$E_\xi(x) = E_0(x-\xi) = \frac{1}{4\pi}\frac{x-\xi}{|x-\xi|^3} = -\frac{1}{4\pi}\nabla U_0(x-\xi)\,; \tag{13c1}$$

here the function $U_0 : x \mapsto \frac{1}{|x|}$ is proportional to the electrostatic potential (energy), and ε_0 is the electric constant.[1] For a charge distribution with continuous density $\rho(\xi)$ the potential is

$$U_\rho(x) = \int \frac{\rho(\xi)\,\mathrm{d}\xi}{|\xi - x|}\,.$$

For the homogeneous charge distribution (that is, $\rho(\xi) = 1$) within the ball B_R of radius R centered at the origin, the potential is

$$U(x) = \int_{B_R} \frac{\mathrm{d}\xi}{|x-\xi|} = \begin{cases} \frac{4\pi R^3}{3|x|} & \text{for } |x| \ge R, \\ \frac{2\pi}{3}(3R^2 - |x|^2) & \text{for } |x| \le R; \end{cases} \tag{13c2}$$

$4\pi R^3/3$ is the total charge of the ball B_R. Once again, the potential, and hence the force exerted by the homogeneous ball on a particle is the same as if the whole charge of the ball were concentrated at its center, if the point is outside the ball. Also, the potential of the homogeneous sphere does not depend on the point x when x is inside the sphere.

13c3 Exercise. For a radial vector field F on $\mathbb{R}^n$,

$$F(x) = f(|x|)x\,, \quad f \in C^1[0,\infty)\,, \quad f'(0) = 0\,,$$

prove that $F \in C^1(\mathbb{R}^n \to \mathbb{R}^n)$ and

$$\operatorname{div} F(x) = f'(|x|)|x| + nf(|x|)\,;$$

here

$$\operatorname{div} F(x) = D_1 F_1 + \dots + D_n F_n \quad \text{for } F = (F_1, \dots, F_n)\,.$$

13c4 Exercise. For a radial function $g : \mathbb{R}^n \ni x \mapsto f(|x|) \in \mathbb{R}$, $f \in C^2[0,\infty)$, $f'(0) = 0$, prove that $g \in C^2(\mathbb{R}^n)$ and

$$\operatorname{div} \nabla g(x) = f''(|x|) + \frac{n-1}{|x|} f'(|x|)\,.$$

13c5 Exercise. For U of (13c2) check that

$$\operatorname{div} \nabla U(x) = \begin{cases} 0 & \text{for } |x| > R, \\ -4\pi & \text{for } |x| < R. \end{cases}$$

[1] $\varepsilon_0 \approx 8.854 \cdot 10^{-12} \, \frac{\mathrm{s^4 A^2}}{\mathrm{m^3 kg}}$.

13d Magnetostatics, and linking number

Consider a steady electric current I that flows along a loop γ in $\mathbb{R}^3$ (maybe a wire). By the Biot-Savart law, this current generates a magnetic field B,

$$B(x) = \mu_0 I B_\gamma(x)\,, \quad B_\gamma(x) = \frac{1}{4\pi}\int_{t_0}^{t_1} \frac{\gamma'(t)\times(x-\gamma(t))}{|x-\gamma(t)|^3}\,\mathrm{d}t\,; \tag{13d1}$$

here μ_0 is the magnetic constant.[1]

13d2 Exercise. Consider a loop γ_R in the x,z plane, consisting of a straight path from $(0,0,-R)$ to $(0,0,R)$ and half of a circle $x^2+z^2=R^2$. Prove that[2]

$$B_{\gamma_R}(x,y,z) \to B_{\gamma_\infty}(x,y,z) = \frac{1}{2\pi}\Big(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\Big) \quad \text{as } R\to\infty\,.$$

In the limit we have a current on the whole z axis; not really a loop, of course, but anyway, the integral converges, and gives a well-known vector field (recall 12b3, 12c7 and the paragraph after 12d3); its curl vanishes outside the z axis. And its circulation around a loop is the winding number of the projection of the loop to the x,y plane, known also as the linking number of the loop and the axis.

What about a non-closed path? This case is beyond magnetostatics; it may seem steady, but it is not: charges accumulate at the endpoints. Nevertheless, let us try a half $\gamma_{-\infty,0}$ of the z axis, from $-\infty$ to 0.

13d3 Exercise. Check that

$$B_{\gamma_{-\infty,0}}(x,y,z) = \frac{1}{4\pi}\Big(1-\frac{z}{\sqrt{x^2+y^2+z^2}}\Big)\Big(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\Big)\,.$$

We wonder, does the curl vanish outside the axis? The circulation of $B_{\gamma_{-\infty,0}}$ around a circle $x^2+y^2=r^2$, $z=\text{const}$ is easy to calculate, it is $\frac12\big(1-\frac{z}{\sqrt{z^2+r^2}}\big)$. As $r\to 0$, the circulation converges to 1 for $z<0$ and to 0 for $z>0$. This is natural; but if the curl vanishes outside the axis then the circulation must be constant, and it is not!

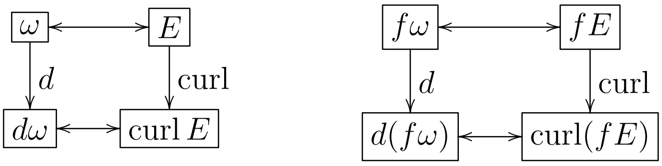
Let us calculate the curl. To this end we need an equality

$$\operatorname{curl}(fE) = f\operatorname{curl}E + \nabla f\times E \tag{13d4}$$

[1] $\mu_0 = 4\pi\cdot 10^{-7}\,\frac{\mathrm{m\cdot kg}}{\mathrm{s^2\cdot A^2}}$. In fact, $\varepsilon_0\mu_0 c^2 = 1$, where c is the speed of light!

[2] Hint: $(z^2+a^2)^{3/2} = a^{-2}\frac{\mathrm{d}}{\mathrm{d}z}\frac{z}{\sqrt{z^2+a^2}}$.

for all $f \in C^1(\mathbb{R}^3)$ and $E \in C^1(\mathbb{R}^3 \to \mathbb{R}^3)$; this is the equality (11e5) $d(f\omega) = df \wedge \omega + f\,d\omega$ translated into the language of vector fields,



since $df \wedge \omega$ corresponds to $\nabla f \times E$ by 12a1(a), and df corresponds to ∇f, of course.

We apply (13d4) to

$$f(x,y,z) = \frac{z}{\sqrt{x^2+y^2+z^2}}\,, \quad E(x,y,z) = \Big(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0 \Big)\,;$$

here $\operatorname{curl}(fE) = \nabla f \times E$ (outside the z axis), since $\operatorname{curl} E = 0$.

13d5 Exercise. Check that

$$\nabla f(x,y,z) = \frac{(-xz, -yz, x^2+y^2)}{(x^2+y^2+z^2)^{3/2}}\,,$$
$$(\nabla f \times E)(x,y,z) = -\frac{(x,y,z)}{(x^2+y^2+z^2)^{3/2}}\,.$$

That is, $(\nabla f \times E)(x) = -\frac{x}{|x|^3}$, and we get

$$\operatorname{curl}(B_{\gamma_{-\infty,0}})(x) = \frac{1}{4\pi} \frac{x}{|x|^3}$$

outside the z axis.

Interestingly, the curl of this magnetic field is proportional to the electric field E_0 (recall (13c1)) of a charge at 0,

$$\operatorname{curl}(B_{\gamma_{-\infty,0}}) = E_0\,. \tag{13d6}$$

This is not a coincidence but a manifestation of an important relation between electric and magnetic fields in dynamics (rather than statics); we'll return to it later.

It is worth to try a short non-closed straight path, since an arbitrary path may be thought of as consisting of infinitesimal elements of this kind. It is sufficient to consider an interval $[0,\varepsilon]$ of the z axis, since all operations of vector analysis are invariant under shifts and rotations (that is, are well-defined on a 3-dimensional affine Euclidean space rather than just $\mathbb{R}^3$). Here is why. First, operations on differential forms are invariant under all diffeomorphisms (recall 11f5, 11f6). Second, the correspondence between differential forms and vector fields is invariant under shifts and rotations (recall Sect. 12a).

By shift, $\operatorname{curl}(B_{\gamma_{-\infty,\varepsilon}})(x,y,z) = E_0(x,y,z-\varepsilon)$, thus,

$$\operatorname{curl}(B_{\gamma_{0,\varepsilon}})(x,y,z) = E_0(x,y,z-\varepsilon) - E_0(x,y,z) = -(D_3E_0)_{(x,y,z)}\varepsilon + o(\varepsilon)$$

(since $B_{\gamma_{0,\varepsilon}} = B_{\gamma_{-\infty,\varepsilon}} - B_{\gamma_{-\infty,0}}$). On the other hand,

$$B_{\gamma_{0,\varepsilon}} = \frac{1}{4\pi}\int_0^\varepsilon \frac{(0,0,1)\times(x-(0,0,t))}{|x-(0,0,t)|^3}\,\mathrm{d}t = \frac{\varepsilon}{4\pi}\frac{(0,0,1)\times x}{|x|^3} + o(\varepsilon)\,;$$

we see that

$$\text{(13d7)} \qquad \frac{1}{4\pi}\operatorname{curl}_x \frac{(0,0,1)\times x}{|x|^3} = -(D_3E_0)_x\,.$$

Our argument fails on the z axis, but (13d7) holds on $\mathbb{R}^3\setminus\{0\}$, since both sides are continuous on $\mathbb{R}^3\setminus\{0\}$. By the invariance under shifts and rotations,

$$\text{(13d8)} \qquad \frac{1}{4\pi}\operatorname{curl}_x \frac{h\times(x-\xi)}{|x-\xi|^3} = -(D_hE_0)_{x-\xi}$$

for all $\xi,h\in\mathbb{R}^3$. Thus,

$$\begin{aligned}\operatorname{curl} B_\gamma(x) &= \frac{1}{4\pi}\operatorname{curl}\int_{t_0}^{t_1}\frac{\gamma'(t)\times(x-\gamma(t))}{|x-\gamma(t)|^3}\,\mathrm{d}t = \\ &= \frac{1}{4\pi}\int_{t_0}^{t_1}\operatorname{curl}_x\frac{\gamma'(t)\times(x-\gamma(t))}{|x-\gamma(t)|^3}\,\mathrm{d}t = -\int_{t_0}^{t_1}\big(D_{\gamma'(t)}E_0\big)_{x-\gamma(t)}\,\mathrm{d}t = \\ &= \int_{t_0}^{t_1}\frac{\mathrm{d}}{\mathrm{d}t}E_0(x-\gamma(t))\,\mathrm{d}t = E_0(x-\gamma(t_1)) - E_0(x-\gamma(t_0))\end{aligned}$$

for all $x\in\mathbb{R}^3\setminus\gamma([t_0,t_1])$.

13d9 Exercise. Prove that the curl of the integral (above) is indeed equal to the integral of the curl.[1]

Thus, for every loop γ,

$$\operatorname{curl} B_\gamma = 0 \quad \text{on } \mathbb{R}^3\setminus\gamma([t_0,t_1])\,.$$

[1] Hint: use Theorem 7e1.

Let $\gamma : [t_0, t_1] \to \mathbb{R}^3$ be a loop, $t_0 < 0 < t_1$, $\gamma(0) = (0,0,0)$, $\gamma'(0) = (0,0,1)$, and $\gamma(t) \ne (0,0,0)$ for $t \ne 0$. Then it appears that

$$\varepsilon B_\gamma(\varepsilon x, \varepsilon y, \varepsilon z) \to \frac{1}{2\pi}\Big(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \Big) \quad \text{as } \varepsilon \to 0+$$

and the circulation of B_γ around a circle $x^2 + y^2 = \varepsilon^2$, $z = 0$ converges to 1 as $\varepsilon \to 0+$. (I do not prove these facts.) It follows that the circulation equals 1 for all ε small enough (such that γ crosses the closed ε-disk only once).

For two loops γ_1, γ_2 that do not cross themselves, nor one another, the circulation of B_{γ_1} around γ_2 is always an integer, the famous linking number, given by the Gauss linking integral (Gauss 1833)[1]

$$\mathrm{Lk}(\gamma_1, \gamma_2) = \frac{1}{4\pi} \iint \frac{\det(\gamma_1'(s), \gamma_2'(t), \gamma_1(s) - \gamma_2(t))}{|\gamma_1(s) - \gamma_2(t)|^3} \mathrm{d}s\mathrm{d}t \,.$$

13e Electrodynamics

> *From a long view of the history of mankind – seen from, say, ten thousand years from now – there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics. The American Civil War will pale into provincial insignificance in comparison with this important scientific event of the same decade.*
>
> *Richard Feynman.*[2]

Electrodynamics describes electromagnetism by a pair of vector fields, E (electric) and B (magnetic), on $\mathbb{R}^3$, depending also on time,

$$E : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 \,, \quad B : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 \,,$$

satisfying the famous Maxwell equations:[3]

$$\operatorname{div} E = \frac{1}{\varepsilon_0}\rho \,, \qquad \operatorname{div} B = 0 \,,$$
$$\operatorname{curl} E = -\frac{\partial B}{\partial t} \,, \qquad \operatorname{curl} B = \mu_0 j + \underbrace{\varepsilon_0 \mu_0}_{1/c^2} \frac{\partial E}{\partial t} \,;$$

[1]See also Wikipedia,
De Zela, "Linking Maxwell, Helmholtz and Gauss through the Linking Integral",
Ricca, Nipoti "Gauss' linking number revisited".

[2]"The Feynman Lectures on Physics" (1964) Volume II, Sect. 1-6.

[3]Coefficients in Maxwell equations depend on the system of units; the form given here fits SI.

here ε_0 is the electric constant (see 13c), μ_0 is the magnetic constant (see 13d), $\rho : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ is the charge density, and $j : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ is the current density. Significantly, $\varepsilon_0 \mu_0 c^2 = 1$, where c is the speed of light.

Surely, this monumental physical law could not be discovered at once in an insight of genius. See Fitzpatrick[1] for its instructive history and meaning. Here are few remarks.

The equality $\operatorname{div} E = \frac{1}{\varepsilon_0}\rho$ (in a special case) was observed in 13c5.

The equation $\operatorname{curl} B = \mu_0 j + \varepsilon_0 \mu_0 \frac{\partial E}{\partial t}$ means that the circulation of B around a loop must be equal to the flux of $\mu_0 j + \varepsilon_0 \mu_0 \frac{\partial E}{\partial t}$ through any surface bounded by this loop. In magnetostatics $\frac{\partial E}{\partial t} = 0$; the flux of $\mu_0 j$ remains. And indeed, such equality was observed in Sect. 13d for a current along the z axis; the circulation of $B_\gamma = \frac{1}{\mu_0 I} B$ is the linking number of the loop and the axis.

Beyond magnetostatics, a current along a half of the z axis, from $-\infty$ to 0, was also treated in Sect. 13d. In this case a charge $q(t) = It$ accumulates at the origin, and generates[2] the electric field $E(x,t) = \frac{1}{\varepsilon_0} It E_0(x)$ (recall (13c1)). Thus, $\varepsilon_0 \mu_0 \frac{\partial E}{\partial t} = \mu_0 I E_0$; this is indeed the curl of $B = \mu_0 I B_{\gamma_{-\infty,0}}$ observed in (13d6).

Now you may wonder, why $\operatorname{div} B = 0$ and why $\operatorname{curl} E = -\frac{\partial B}{\partial t}$. But do you wonder, why at all the 6 functions (on $\mathbb{R}^3 \times \mathbb{R}$)? Because $6 = 3 + 3$, really? But why just two vector fields? Why not one, or three, or four? Why not one scalar field and one vector field?

You may say: these are questions to the Great Architect of the Universe... Well, but He/She "begins to appear as a pure mathematician".[3] And indeed, mathematics answers:

$$6 = \binom{4}{2} .$$

It means: not a pair of 3-dimensional vector fields, but a 2-form in the 4-dimensional space-time!

You may say: but does it help to understand, why $\operatorname{div} B = 0$ and why $\operatorname{curl} E = -\frac{\partial B}{\partial t}$? Oh yes, it does! Here is the 2-form on $\mathbb{R}^4 = \{(x_1, x_2, x_3, t)\}$:

$$\omega = (E_1\,dx_1 + E_2\,dx_2 + E_3\,dx_3) \wedge dt + B_1\,dx_2 \wedge dx_3 + B_2\,dx_3 \wedge dx_1 + B_3\,dx_1 \wedge dx_2 .$$

And here is the answer: since ω is exact! That is,

$$\omega = dA$$

[1] Fitzpatrick, "Classical electromagnetism".

[2] Really, "generates"? Well, at least, this charge and this field are compatible...

[3] "... from the intrinsic evidence of his creation, the Great Architect of the Universe now begins to appear as a pure mathematician." James Hopwood Jeans, in his book "The Mysterious Universe".

for some 1-form A on $\mathbb{R}^4 = \{(x_1, x_2, x_3, t)\}$. We have

$$A = A_1\,dx_1 + A_2\,dx_2 + A_3\,dx_3 + A_0\,dt\,;$$

$$dA = \underbrace{(D_1A_2 - D_2A_1)}_{B_3}\,dx_1\wedge dx_2 + \underbrace{(D_2A_3 - D_3A_2)}_{B_1}\,dx_2\wedge dx_3 + \underbrace{(D_3A_1 - D_1A_3)}_{B_2}\,dx_3\wedge dx_1 + \underbrace{(D_1A_0 - D_0A_1)}_{E_1}\,dx_1\wedge dt + \underbrace{(D_2A_0 - D_0A_2)}_{E_2}\,dx_2\wedge dt + \underbrace{(D_3A_0 - D_0A_3)}_{E_3}\,dx_3\wedge dt\,;$$

$$\operatorname{div} B = D_1B_1 + D_2B_2 + D_3B_3 = (D_2D_3 - D_3D_2)A_1 + (D_3D_1 - D_1D_3)A_2 + (D_1D_2 - D_2D_1)A_3 = 0\,;$$

$$\begin{aligned}\operatorname{curl} E &= (D_2E_3 - D_3E_2, D_3E_1 - D_1E_3, D_1E_2 - D_2E_1) = \\ &= (D_2D_3A_0 - D_2D_0A_3 - D_3D_2A_0 + D_3D_0A_2, \dots, \dots) = \\ &= \big(D_0(D_3A_2 - D_2A_3), \dots, \dots\big) = -D_0B\,.\end{aligned}$$

Magically, all questions are answered! In addition, the next exercise explains why $\varepsilon_0\mu_0c^2 = 1$, where c is the speed of light.

13e1 Exercise. [1] Consider such a special case:

$$A = A_2(x_1 - ct)\,dx_2 + A_3(x_1 - ct)\,dx_3\,.$$

Check that in this case

$$E = \begin{pmatrix} 0 \\ cA_2'(x_1 - ct) \\ cA_3'(x_1 - ct) \end{pmatrix}, \qquad B = \begin{pmatrix} 0 \\ -A_3'(x_1 - ct) \\ A_2'(x_1 - ct) \end{pmatrix},$$

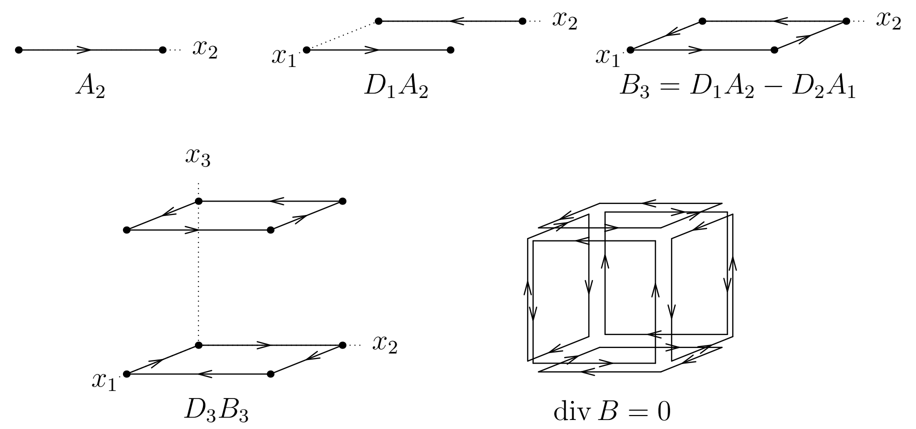
$$\operatorname{div} E = 0\,, \qquad \operatorname{curl} B = \frac{1}{c^2}\frac{\partial E}{\partial t}\,.$$

Such solutions are called electromagnetic waves. Explain, why. In what direction do these waves travel, and how fast?

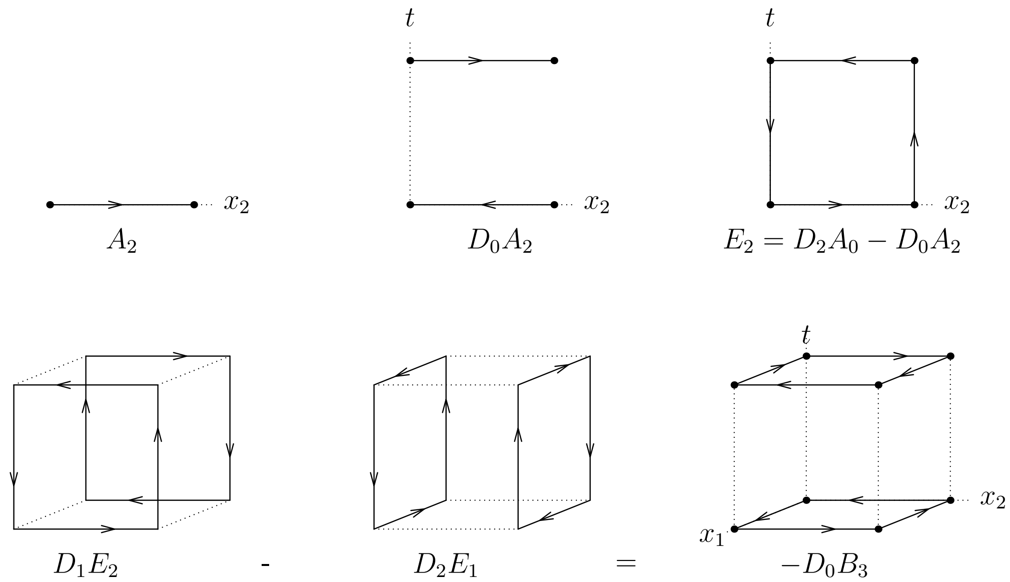
13e2 Exercise. Explain the pictures below; insert the missing ε, $1/\varepsilon$, $1/\varepsilon^2$ and $\lim_{\varepsilon\to 0+}$ as needed.

[1] Sjamaar, Ex. 2.18(v).

Magnetic field as a loop function in space



Electric field as a loop function in space-time



Should we treat A as a physical field that underlies E and B? No, we should not, since A cannot be measured (neither in practice nor in principle). According to the so-called gauge field theory,[1] all measurable quantities are invariant under the gauge transformation

$$A \mapsto A + df\,, \quad f \in C^1(\mathbb{R}^4)\,.$$

[1]See also Wikipedia (English, Hebrew).

Thus, A cannot be measured; only its equivalence class $\{A + df : f \in C^1(\mathbb{R}^4)\}$ can.

We have several mathematically equivalent descriptions of the same physical object:

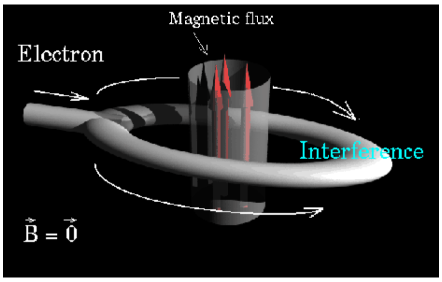
∗ a pair of vector fields;

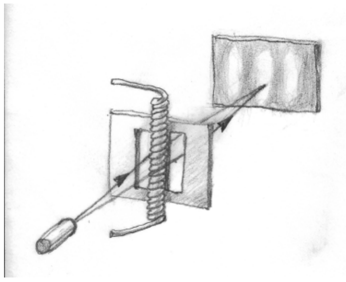
∗ an exact 2-form;

∗ an equivalence class of 1-forms;

∗ a loop function.

The latter is equivalent to others, since $\mathbb{R}^4$ is simply connected (recall 13b19). We cannot try a different, multiply connected Universe; but we can try a multiply connected region of the given space-time. Can it happen that electromagnetic field is absent in two regions but present in their union?[1] Classical physics gives negative answer; charged particles interact only locally with the electromagnetic field. Amazingly, quantum physics gives affirmative answer; true, the interaction is local, but a particle has its wave function, spread in space! The relevant physical phenomenon is the famous Aharonov-Bohm effect.[2] [3]





[1]Recall the end of Sect. 13b.

[2]Images from: Edward Sternin, Bartosz Milewski.

[3]See also Wikipedia (English, Hebrew).

Index