Universiteit Leiden
Opleiding Wiskunde

Geometric differential forms

Name: Jules Jacobs

Date: 28-09-2018

supervisor: dr. Roland van der Veen

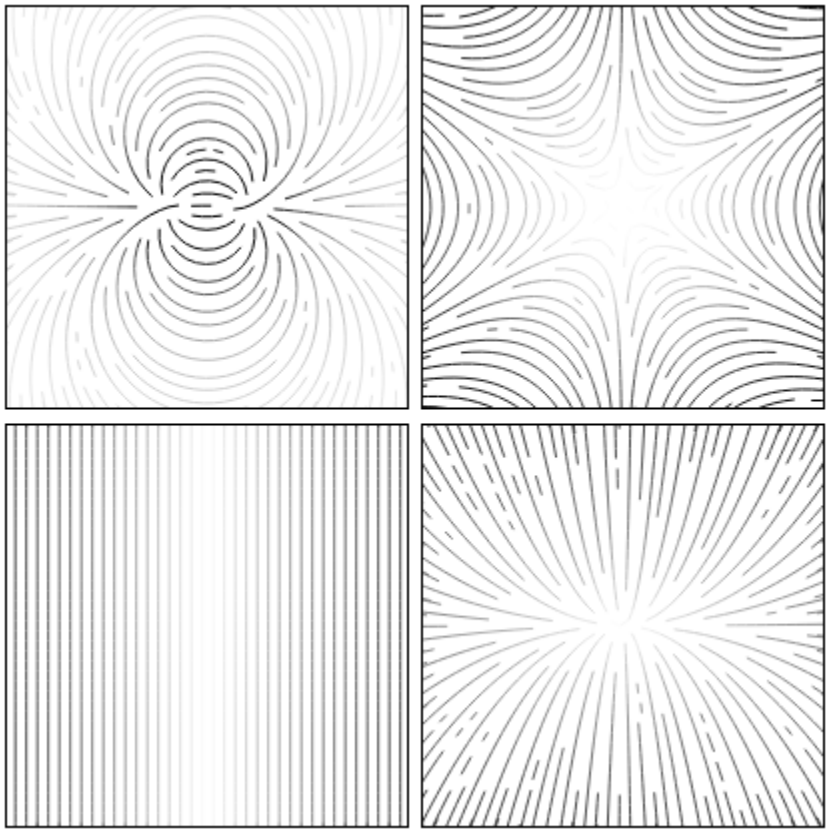
MASTER'S THESIS

Mathematical Institute
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

GEOMETRIC DIFFERENTIAL FORMS

JULES JACOBS

The book *Formes Differentielles* written by Henri Cartan, the son of Élie Cartan, describes differential forms as purely formal objects and does not give a geometric interpretation. In this thesis I will describe a geometric route to differential forms along the lines of Maxwell's geometric picture for the electric field. We will start with a geometric picture of what an n-form is, and and show that differential forms are the natural way to formalise these objects.



Contents

1. Introduction and Motivation

Mathematicians used to regard differentials dx as infinitessimally small quantities. The equation $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$ was seen as a relation between the small quantities df, dx, dy. The differential equation $\frac{dy}{dx} = -\frac{A}{B}$ could be rewritten as $Adx + Bdy = 0$ and this specified the direction of a curve in the x, y-plane, putting the variables x and y on equal footing. Given any point (x, y) we may use the relation $A(x, y)dx + B(x, y)dy = 0$ to find the local direction (dx, dy) of the integral curve through (x, y). Clairaut saw the connection and tried to express these integral curves as level sets $f(x, y) = c$ of a function f. Differentiating this relation we obtain $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = 0$. He noted that if $\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = Adx + Bdy$, i.e. $\frac{\partial f}{\partial x} = A$ and $\frac{\partial f}{\partial y} = B$, then f's level sets give integral curves of the differential equation. Clairaut saw that the symmetry of partial derivatives $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ meant that $\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}$ is a necessary condition for such an f to exist, and he also proved that this condition is sufficient. Clairaut generalised this idea to n variables and found $\sum A_i dx^i = df$ exists if $\frac{\partial A_i}{\partial x_j} - \frac{\partial A_j}{\partial x_i}$ for all i, j. We see the first glimpse of the exterior derivative.

Cauchy invented the notion of an integral $\int (A\frac{dx}{dt} + B\frac{dy}{dt})dt$ along a curve $x(t), y(t)$ to define integrals in the complex plane, and he noted that this integral is independent of the path γ if Clairauts condition $\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}$ is satisfied. This lead him to his integral theorem for holomorphic functions, to Green's theorem, and to the residue theorem. For Cauchy the integrand $A\frac{dx}{dt} + B\frac{dy}{dt}$ was intimately tied to the curve. Maxwell on the other hand wanted to put Faraday's geometric ideas about electric field lines on a solid mathematical footing, which lead him to the point of view that $\int (A\frac{dx}{dt} + B\frac{dy}{dt} + C\frac{dz}{dt})dt = \int_\gamma Adx + Bdy + Cdt$, and considered $Adx + Bdy + Cdz$ to be a separate object (namely, the electric field) that we can integrate along a curve γ. Surface integrals $\int_S Adxdy + Bdydz + Cdzdx$ appeared in the 19th century and Gauss' and Stokes' theorem were discovered. Poincaré wanted to extend Cauchy and Riemann's work to complex functions of n variables and was so led to his famous lemma. Frobenius was separately investigating integrability conditions for systems of first order homogeneous linear PDE's. Élie Cartan would eventually bring all of this together in his work on differential forms by formalising differential forms such as $Adxdy + Bdydz + Cdzdx$ as objects independent of the integral in which they appear. Stokes' theorem and Poincare's lemma are particularly simple and beautiful when stated in terms of differential forms.

Although Maxwell, following Faraday, had a clear geometric picture of electric field lines in space, the development of differential forms was primarily algebraic. The book *Formes Differentielles* [1] written by Henri Cartan, the son of Élie Cartan, describes differential forms as purely formal objects and does not give a geometric interpretation. In this thesis I will describe an alternate, geometric route to differential forms along the lines of Maxwell's geometric picture for the electric field. We will start with a geometric picture of what an n-form is, and and show that differential forms are the natural way to formalise these objects. The book *Gravitation* [6] contains a similar but different geometric interpretation of differential forms. The interpretation of 1-forms is the same, but the interpretation of higher forms is different. The article *A pictorial introduction to differential geometry, leading to*

Maxwell's equations as three pictures [3] contains an interpretation almost identical to ours, but differs in the interpretation of orientation and the description of electromagnetism.

2. The geometric picture

Faraday understood the electric field intuitively as lines in space emanating in all directions from positive and negative charges such that the lines start at a positive charge and end at a negative charge. Maxwell thought of this in terms of the motion of fluid: he imagined that positive charges were fluid sources and negative charges were fluid sinks. The path that a fluid particle takes from a source to a sink corresponds to one of Faraday's lines. Maxwell noted that the net fluid flow out of any closed surface in space is equal to the net amount of charge inside the surface. In terms of Faraday's intuition this means that the net number of lines piercing the surface is proportional to the net charge inside. Each line has an orientation and an intersection with the surface is counted positive or negative depending on the orientation. We may also ask how many lines pierce an open surface such as a small parallelogram in space. It is clear that the sign depends not only on the orientation of the lines but also on the orientation of the surface. If we rotate the parallelogram 180 degrees then the lines pierce it in precisely the opposite direction. One can imagine painting the parallelogram red on one side and blue on the other side, and count the intersection with a line positive or negative depending on whether the line goes from the blue to the red side or from the red to the blue side.

This picture generalises to n dimensional space. By analogy we can imagine a collection of negative and positive points, with lines emanating from the positive points and converging on the negative points. The net number of points inside an n-volume will be proportional to the net number of lines crossing the boundary of this n-volume, which is an (n-1)-dimensional submanifold. A point is a 0-dimensional object and a line is an 1-dimensional object. We therefore call a point density a 0-density and a line density a 1-density. We may measure the net intersection of a 1-density through a (n-1)-surface, and we may measure the net "intersection" of a 0-density with a n-surface, i.e. a volume. This suggests that we may generalise to measuring the net intersection of a k-density through a (n-k)-surface. By analogy with a line density, a k-density should then be a smoothly varying density of k-surfaces. A k-surface usually intersects a (n-k)-surface in a discrete set of points, which is something that we can count. In general a k-surface may lie parallel to a (n-k)-surface along one or more directions. This can already happen with a line density: the lines may lie parallel to the surface through which we're measuring the flux, and in that case those lines do not contribute to the net flux. It is therefore reasonable to state that the same should be true for the intersection of k-densities with a (n-k)-surface: if they lie parallel then the net intersection is zero. Let us enumerate the possibilities in three dimensions:

k	k-density	(3-k)-surface
0	point density	volume
1	line density	surface
2	surface density	line
3	volume density	point

Given a surface density and a curve we may count the net number of surfaces that the curve pierces through. This is similar to the intersection of a line density through a surface, except that we now have a single line and a density of surfaces instead of a single surface and a collection of lines. It is perhaps unclear what the intersection of a volume density with a single point is, but we can understand this by comparing it to the intersection of a point density with a single volume: instead of a single volume we now have a collection of volumes, and given any point we may count the number of those volumes that the point lies in. Just as we have positive and negative points, we also have positive and negative volumes. To understand more clearly the difference between a volume density and a point density consider the difference between the temperature distribution in space and a mass distribution in space. Given a mass distribution we may ask how much mass is in a given volume. It doesn't make sense to ask how much mass is at a point, if the mass is smoothly distributed (though it does make sense to ask about the mass density at a point). On the other hand, it does not make sense to ask how much temperature is inside a given volume, but it does make sense to ask what the temperature is at a given point. The difference can also be understood in terms of the transformation of space: if we stretch space by a factor 2 in all directions, then the temperature at a point remains the same, but the mass density at a point decreases by a factor of 8. The units of mass density are $\frac{kg}{m^3}$ whereas the units of temperature are just Kelvin.

Instead of $n = 3$ we may consider $n = 2$:

k	k-density	(2-k)-surface
0	point density	area
1	line density	line
2	surface density	point

The case $k = 1$ may be confusing: we compute the intersection of a line density with a line. This is a phenomenon that occurs for $k = \frac{n}{2}$ if n is even. In four space dimensions we intersect a 2-surface with a 2-surface density, as two 2-surfaces generally meet in a discrete set of points in 4d.

We can divide a (n-k)-surface into m pieces by cutting it up. Given a k-density we may determine the generalised flux through each of the m pieces separately. The flux through the whole surface will be the sum of the fluxes through the pieces. We conclude from this that k-densities are local in the sense that it is fully determined by measuring its flux through small (n-k)-surfaces.

2.1. **Probabilistic interpretation.** These k-densities also have a probabilistic interpretation. We may interpret a point density as a randomly chosen point, a line density as a randomly chosen curve from a distribution of curves, and in general a k-density as a randomly chosen k-surface from a distribution of k-surfaces. Given a fixed $(n-k)$-surface we may then ask what the *expected* number of intersections of the fixed (n-k)-surface with the random k-surface is. Suppose for simplicity that we have a finite collection of curves and a probability p_i for each curve. We can chop each individual curve up into m pieces and assign a probability p_i/m to each piece. By doing this we obtain a different probability distribution of smaller curves, but in terms of flux through surfaces this new distribution will behave in exactly the same way as the original curve distribution, except that the expected number of intersections with any fixed surface is diminised by a factor of n. To fix this issue we can work with non-normalised probability distributions, and assign

"probability" p_i to each piece, so that the chopped up density behaves exactly the same as the original. We conclude from this that the global connection structure of the k-density doesn't matter: a collection of long curves behaves the same as a (larger) collection of short lines.

2.2. **Smoothly varying densities.** We have to keep in mind that k-densities don't consist of a finite collection of k-surfaces, but of a collection that is smeared out continuously, just like a point density. If the density is smoothly varying then it will look like a constant density locally. Since we concluded that the behaviour of a density is fully determined by its local behaviour we hope to understand general densities by understanding constant densities. So what is a constant density? A point density is constant if given a volume, the number of points inside the volume doesn't change if we translate the volume in space. Similarly, a k-density is constant if given any (n-k)-surface, the flux through the surface doesn't change if we translate it in space. Furthermore, since we can approximate the surface by a collection of (n-k)-parallelepipeds, we can concentrate on measuring the generalised flux through (n-k)-dimensional parallelepipeds. So for a surface density we measure how many surfaces a straight line segment pierces through, for a line density we measure how many lines pierce through a parallelogram, and for a point density we measure how many points lie inside a 3-parallelepiped.

Operation	Type	Meaning	Example
$d\alpha$	α is a k density and $d\alpha$ is a $k-1$ density	Boundary	The boundary of a surface density is a line density, the boundary of a line density is a point density.
$\alpha \wedge \beta$	α is a k density, β is an l density, and $\alpha \wedge \beta$ is an $\alpha + \beta - n$ density	Intersection	The intersection of two surface densities in $\mathbb{R}^3$ is a line density.
$f^*\alpha$	$f : N \to M$ is a smooth function from an n manifold to a m manifold, and α is a k density on M	Intersect α with the image of f, then transport that density back to N	If α is a line density in $\mathbb{R}^3$ and f is a parametrisation of a surface, the intersection of α with the surface is a point density, and $f^*\alpha$ is that point density on the parameter space.
$\int_S \alpha$	α is a point density, and S is the surrounding space	Weighted number of points in α	If $\alpha$is the volume form on a sphere, it represents a constant point density, and $\int_S \alpha$ is the volume of the sphere.
$\int_M \alpha$, $M \subset S$	M is a $(n-k)$ submanifold of S, and $\alpha$is a k density on S	Intersection value of α with M	If $\alpha$is a line density and M is a surface, then $\int_M \alpha$ is the flux through the surface.
$\int_S \alpha \wedge \beta$	β is a $(n-k)$ density on S, and α is a k density on S	Intersection value of α with β	If α is a line density and β is a surface density, then $\int_S \alpha \wedge \beta$ is the flux through all the weighted surfaces of β.
$*a$	α is a k density and $*a$ is a $n-k$ density	Orthogonal complement of α	If α is a line density on $\mathbb{R}^3$, then $*a$ are surfaces orthogonal to the lines at every point.

TABLE 1. Operations on densities

Lemma/Theorem	Type	Meaning
$\alpha \wedge (\beta + \gamma) = \alpha \wedge \beta + \alpha \wedge \gamma$	α, β, γ arbitrary densities, with β and γ of the same degree	The intersection of α with the superposition of β and γ is equal to the superposition of the individual intersections.
$\int_M \alpha + \beta = \int_M \alpha + \int_M \beta$	α, β being k densities and M a $(n-k)$ submanifold	The intersection of the superposition of α and β with M is equal to the sum of the intersections.
$\int_{\partial M} \alpha = \int_M d\alpha$	α a line density on M	The flux of α through the boundary of M is equal to the weighted boundary points of α in M.
$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta$	α, β arbitrary densities	The boundary of the intersection of α and β occurs at the intersection of the boundary of α with β and vice versa.
$\int_{f(M)} \alpha = \int_M f^* \alpha$	α a k density and f a diffeomorphism between M and its image $f(M)$	The intersection of α with the image $f(M)$ is equal to the weighted number of points in αtransported back to M via f.

TABLE 2. Lemmas and theorems about densities

3. Exterior algebra

For comparison we now give a more traditional construction of the exterior algebra. In the following, let U be a finite dimensional vector space of dimension $N = \dim(U)$.

Definition 1. Let $T^n \subset U^n \to \mathbb{R}$ be the space of n-ary multilinear maps $U^n \to \mathbb{R}$, and define the tensor product $f \otimes g \in T^{n+m}$ of $f \in T^n$ and $g \in T^m$ to be $f \otimes g(u_1, \ldots, u_n, u_{n+1}, \ldots, u_{n+m}) = f(u_1, \ldots, u_n) g(u_{n+1}, \ldots, u_{m+n})$.

We are interested in alternating multilinear maps $\bigwedge^n \subset T^n$. Alternating multilinear maps will be what we integrate over a manifold. The input $(u_1, \ldots, u_n) \in U^n$ specifies a n-dimensional parallelepiped forming a piece of the manifold we are integrating over, and $f(u_1, \ldots, u_n)$ will be the value that the piece of the manifold contributes to the integral. If $(u_1, \ldots, u_n) = (\ldots, u, u, \ldots)$, that is, if two adjacent positions have the same value, then the parallelepiped is degenerate (has zero volume), so in this case we want f to be zero. This brings us to the following definition.

Definition 2. A map $f \in U^n \to \mathbb{R}$ is called alternating if it is multilinear and $f(\ldots, u, u, \ldots) = 0$ for all $u \in U$. That is, f is zero if two adjacent arguments have the same value.

This may look like it doesn't constrain the behaviour of f much, but we will see $f(\ldots, u, u, \ldots) = 0$ has great consequences if f is multilinear.

Lemma 3. *If f is alternating then $f(\ldots, u, v, \ldots) = -f(\ldots, v, u, \ldots)$ for $u, v \in U^n$*

Proof. Compute the value of $f(\dots, u+v, u+v, \dots)f$ by multilinearity:

$$\begin{aligned} f(\dots, u+v, u+v, \dots) =& f(\dots, u, u, \dots) + f(\dots, u, v, \dots) + \\ & f(\dots, v, u, \dots) + f(\dots, v, v, \dots) \end{aligned}$$

Since f is alternating,

$$\begin{aligned} f(\dots, u+v, u+v, \dots) &= 0 \\ f(\dots, u, u, \dots) &= 0 \\ f(\dots, v, v, \dots) &= 0 \end{aligned}$$

so the equation reads

$$f(\dots, u, v, \dots) + f(\dots, v, u, \dots) = 0$$

which implies the lemma $f(\dots, u, v, \dots) = -f(\dots, v, u, \dots)$ □

This means that swapping two arguments results in a minus sign. Since any permutation can be built from swapping adjacent positions, we can generalise this property to the following lemma.

Lemma. *If f is alternating, then $f \circ \sigma = sign(\sigma) f$ where $\sigma \in S_n$ is a permutation.*

Proof. Write $\sigma = s_1 \circ s_2 \circ \cdots \circ s_k$ as a sequence of adjacent swaps s_i. By the previous lemma, $f \circ s_i = -f = sign(s_i) f$. By induction, $f \circ \sigma = f \circ s_1 \circ \cdots \circ s_k = sign(s_1) \cdots sign(s_k) f = sign(\sigma) f$. □

For $n = 1$ an alternating multilinear map is simply a dual vector $f : U \to \mathbb{R}$, and for $n = 2$ this corresponds to antisymmetric maps $f(u, v) = -f(v, u)$, and for $n = \dim(U)$ we have $f = c \det(\cdot)$ for some $c \in \mathbb{R}$, since the determinant is characterised by its multilinear and alternating properties up to scaling. It will be useful to define a right action of permutations on multilinear maps.

Definition 4. A permutation $\sigma \in S_n$ acts from the right on $f \in T^n$ with $f\sigma = sign(\sigma) f \circ \sigma$.

We may unambiguously write $f\sigma\tau$ because $(f\sigma)\tau = sign(\sigma) sign(\tau) f \circ \sigma \circ \tau = sign(\sigma\tau) f \circ (\sigma\tau) = f(\sigma\tau)$. Using this notation we have that f is alternating iff $f\sigma = f$ because $sign(\sigma)^2 = 1$. We also define the right action of sets of permutations on f as the average value of the action of permutations in the set.

Definition 5. A set of permutations $A \subset S_n$ acts from the right on $f \in T^n$ with $fA = \frac{1}{|A|} \sum_{\sigma \in A} f\sigma$.

These actions are linear, so $(\sum f_i A) = (\sum f_i) A$ as if we were multiplying by a number.

Lemma 6. *Let $f \in \bigwedge^n$, then $fS_n = f$.*

Proof. $fS_n = \frac{1}{|S_n|} \sum_{\sigma \in S_n} f\sigma = \frac{1}{|S_n|} \sum_{\sigma \in S_n} f = f$. □

The following main lemma will be useful for establishing the remaining properties of alternating mutlilinear maps.

Lemma 7. *Let $A \subset S_n$ a set of permutations. Then $(fA)S_n = fS_n = (fS_n)A$.*

Proof. We first establish the lemma when A is a singleton σ. This follows from $\sigma S^n = S^n$:

$$\begin{aligned}(f\sigma)S_n &= \frac{1}{|S_n|}\sum_{\tau\in S_n} f\sigma\tau \\ &= \frac{1}{|S_n|}\sum_{\tau'\in\sigma S_n} f\tau' \\ &= fS_n\end{aligned}$$

and similarly $(fS_n)\sigma = fS_n$. Now

$$\begin{aligned}(fA)S_n &= \frac{1}{|A|}\sum_{\sigma\in A}(f\sigma)S_n \\ &= \frac{1}{|A|}\sum_{\sigma\in A} fS_n \\ &= fS_n\end{aligned}$$

and similarly $(fS_n)A = fS_n$. □

This lemma allows us to establish all the properties of alternating multilinear maps with ease.

Lemma 8. *Let $f \in T^n$, then $fS_n \in \bigwedge^n$.*

Proof. By the main lemma, $(fS_n)\sigma = fS_n$ for all permutations $\sigma \in S_n$. □

Together, this shows that the right action of S_n acts as a projection of T^n onto $\bigwedge^n$.

Definition 9. For $f \in \bigwedge^n$ and $g \in \bigwedge^m$, define the wedge product $f\wedge g = \frac{(n+m)!}{n!m!}(f\otimes g)S_{n+m}$.

The wedge product is bilinear because the tensor product is bilinear, and right action by S_{n+m} is linear.

Lemma 10. *The wedge product is associative, $(f\wedge g)\wedge h = f\wedge(g\wedge h)$ for $f \in \bigwedge^k, g\in\bigwedge^l, h\in\bigwedge^m$*

Proof. By the main lemma, $((f\otimes g)S_{k+l}\otimes h)S_{k+l+m} = (f\otimes g\otimes h)S_{k+l+m}$ because S_{k+l} acting on the first $k+l$ positions is a subset of S_{k+l+m}, and therefore

$$\begin{aligned}(f\wedge g)\wedge h &= \frac{(k+l)!}{k!l!}(f\otimes g)S_{k+l}\wedge h \\ &= \frac{(k+l)!}{k!l!}\frac{(k+l+m)!}{(k+l)!m!}((f\otimes g)S_{k+l}\otimes h)S_{k+l+m} \\ &= \frac{(k+l+m)!}{k!l!m!}(f\otimes g\otimes h)S_{k+l+m}\end{aligned}$$

Similarly,

$$\begin{aligned} f \wedge (g \wedge h) &= \frac{(l+m)!}{l!m!} f \wedge (g \otimes h) S_{l+m} \\ &= \frac{(l+m)!}{l!m!} \frac{(k+l+m)!}{k!(l+m)!} (f \otimes (g \otimes h) S_{l+m}) S_{k+l+m} \\ &= \frac{(k+l+m)!}{k!l!m!} (f \otimes g \otimes h) S_{k+l+m} \end{aligned}$$

□

Lemma 11. *The wedge product (anti)commutes with $g \wedge f = (-1)^{kl} f \wedge g$ for $f \in \bigwedge^k, g \in \bigwedge^l$.*

Proof. Let $\sigma_{k,l}$ be the permutation that swaps the first k positions with the remaining l positions. Using this permutation,

$$g \otimes f = (f \otimes g) \circ \sigma_{k,l} = sign(\sigma_{k,l})(f \otimes g)\sigma$$

We multiply both sides by $\frac{k!l!}{(k+l)!}$ and act with S_{k+l},

$$\frac{(k+l)!}{k!l!}(g \otimes f) S_{k+l} = sign(\sigma_{k,l}) \frac{(k+l)!}{k!l!} (f \otimes g) \sigma_{k,l} S_{k+l}$$

By the main lemma, $\sigma_{k,l}$ can be removed, and then this equation is nothing but

$$g \wedge f = sign(\sigma_{k,l}) f \wedge g$$

The proof is now complete, as $sign(\sigma_{k,l}) = (-1)^{kl}$ because this can be done with kl adjacent swaps. □

Corollary 12. *If n is odd, then $f \wedge f = 0$ for $f \in \bigwedge^n$.*

Proof. $f \wedge f = (-1)^{n \cdot n} f \wedge f = -f \wedge f$, so $f \wedge f = 0$. □

The following lemma shows the reason why we have the factorial prefactor, and also gives a way to define the wedge product when we are working over a ring instead of a field (so that we do not have $\frac{k!l!}{(k+l)!}$).

Lemma 13. *If $f \in \bigwedge^k, g \in \bigwedge^l$, then $f \wedge g = \sum_{\sigma \in S_{k,l}} (f \otimes g)\sigma$ where $S_{k,l}$ are all permutations that keep the first k elements in the same order and the remaining l elements in the same order, i.e. a permutation that interleaves the first k items with the remaining l items.*

Proof. We may write $\sum_{\sigma \in S_{k,l}} (f \otimes g)\sigma = |S_{k,l}|(f \otimes g) S_{k,l}$. Because f, g are alternating, $f = f S_k$ and $g = g S_l$. Therefore

$$\begin{aligned} \sum_{\sigma \in S_{k,l}} (f \otimes g)\sigma &= |S_{k,l}|(f \otimes g) S_{k,l} \\ &= |S_{k,l}|(f S_k \otimes g S_l) S_{k,l} \\ &= |S_{k,l}|(f \otimes g) S_{k+l} \\ &= f \wedge g \end{aligned}$$

The butlast step is justified because the set of permutations where we permute within the first k positions and within the remaining l positions, and then do a permutation that interleaves the k, l positions, gives the set of all permutations. For the last step we use $|S_{k,l}| = \frac{(k+l)!}{k!l!}$. □

Lemma 14. *If $f \in T^n$ is symmetric under any permutation σ with sign -1, then $fS_n = 0$.*

Proof. By the main lemma and $f \circ \sigma = f$, we have $fS_n = f\sigma S_n = sign(\sigma)(f \circ \sigma)S_n = -fS_n$, so $fS_n = 0$. □

Let us go back to the definition of alternating multilinear map: a map $f \in U^n \to \mathbb{R}$ is alternating if it is multilinear and $f(\dots, u, u, \dots) = 0$ for all $u \in U$. The n-tuple U^n of vectors in U represents an n-parallelepiped (it may be $n < \dim U$). Another way to specify a parallelepiped is with a linear map $V : \mathbb{R}^n \to U$. The parallelepiped is the image $V([0,1]^n)$ of the unit cube. Indeed $U^n \simeq \mathbb{R}^n \to U$. First a quick reminder about elementary row operations.

Definition 15. An elementary row operation $R_{ij}^c : \mathbb{R}^k \to \mathbb{R}^k$ is the linear map $R_{ij}^c = I + ce_ie_j^T$ that adds c times coordinate j to coordinate $i \neq j$.

Lemma 16. $(R_{ij}^c)^{-1} = R_{ij}^{-c}$ *and* $\det R_{ij}^c = 1$.

Proof. We calculate $R_{ij}^c R_{ij}^{-c} = (I + ce_ie_j^T)(I - ce_ie_j^T) = I^2 + ce_ie_j^T - ce_ie_j^T - c^2(e_ie_j^T)^2 = I$ since $i \neq j$. The matrix of R_{ij}^c is either lower or upper triangular with 1 on the diagonal, so by doing a cofactor expansion we find $\det R_{ij}^c = 1$. □

Lemma 17. *Left multiplication $R_{ij}^c A$ adds c times row j to row i and right multiplication AR_{ij}^c adds c times column i to column j.*

Proof. Calculate $R_{ij}^c A = (I + e_ie_j^T)A = A + e_iA_{j\cdot}$, so left multiplication adds c times row j to row i. For right multiplication, use $AR_{ij}^c = ((R_{ij}^c)^T A^T)^T = (R_{ji}^c A^T)^T$, which shows that right multiplication adds c times column i to column j. □

The conventional method is to have three types of elementary row operations, (1) the row operations R_{ij}^c (2) scaling a row by a nonzero constant (3) swapping two rows. Then it is proved that a matrix can be put in reduced row echelon form by left multiplying by elementary row operations. We take a different road so that only the row operations R_{ij}^c suffice, and we don't write a matrix in reduced row echelon form, but instead as a product EDE' where E and E' are a product of elementary row operations R_{ij}^c, and D is diagonal. The diagonal part is allowed to have zeroes, so this form works for singular matrices too.

Lemma 18. *Let a skew swap operation be the exchange of two rows while multiplying one of the rows by -1. We can build skew swap operations out of elementary row operations.*

Proof. To skew swap row i with row j, add row i to j, then subtract row j from i, then add row i to j. □

Lemma 19. *Any matrix $A \in M_{k,k}$ can be written as a product $A = EDE'$ where E and E' are a product of elementary row operations, and D is diagonal.*

Proof. We start with the matrix A and we compute the form EDE' with the following algorithm. Repeat the following steps for $i = 1$ to $i = k$: □

(1) If row i and column i of matrix A are zero, go to $i + 1$.
(2) If A_{ii} is zero, but some element in row or column i is nonzero, use a skew swap operation to introduce a nonzero in A_{ii}.

(3) Use row and column operations to make row i and column i zero, except for A_{ii}. Go to $i+1$.

At the end of the algorithm we are left with a diagonal matrix D. The algorithm only uses row operations and column operations, that is, it left and right multiplies A by some product of R^c_{ij}'s to reach diagonal form. Since $(R^c_{ij})^{-1} = R^c_{ij}$ we may invert those operations to transform D back into A, that is $EDE' = A$ where E and E' are products of row operations, and D is diagonal.

Lemma 20. *If f is alternating, then $f(v \circ A) = \det(A) f(v)$ for $v \in U^k$ and A a $k \times k$-matrix*

Proof. If $A = R^c_{ij}$ then $f(v \circ A) = f(v_1, \dots, v_i + cv_j, \dots, v_k) = f(v_1, \dots, v_i, \dots, v_k) + cf(v_1, \dots, v_j, \dots, v_k)$. The second term is zero, since v_j appears twice. So we have $f(v \circ A) = \det(A) f(v)$ since $\det(A) = 1$. If $A = D$ is diagonal, then $f(v \circ A) = (\prod_i D_i) f(v) = \det(A) f(v)$. Since any matrix can be written as $A = EDE'$, the lemma holds for any matrix. □

Corollary 21. *This lemma generalises properties that we've already encountered.*

(1) If $rank(A) < k$ then $f(v \circ A) = 0$. This is a generalisation of the defining property $f(\dots, u, u, \dots) = 0$: if the parallelepiped is degenerate in any way, the value of f is zero.
(2) If A is a permutation matrix, then $f(v \circ A) = \det(A) f(v)$ is the same as $f(v \circ \sigma) = sign(\sigma) f(v)$ for a permutation σ.
(3) If $A \in SL_k$ then $f(v \circ A) = f(v)$.

The last property shows that if we reparameterise the parallelepiped by a $\det A = 1$ matrix, then the value of f doesn't change. This implies that the shape of the parallelepiped does not matter; only its size and the subspace it is in matters. For instance, if we have a parallelogram in $\mathbb{R}^3$, then rotating or skewing that parallelogram in its own plane doesn't change the value of f. In other words, alternating maps remain well defined if we quotient the space U^k by SL_k.

Note that there is a different property $f(A \circ v) = \det(A) f(v)$ when $n = \dim(U)$, where $A : U \to U$. This property is about transforming the vector space U and only holds when $n = \dim(U)$. The lemma above is about transforming the parameter space, that is $A : \mathbb{R}^k \to \mathbb{R}^k$ and we have $v \circ A$ instead of $A \circ v$, and it holds for any n. This lemma is a generalisation of the property $f(v \circ \sigma) = sign(\sigma) f(v)$, which changes the parameter/index space, and also holds for any n.

Corollary 22. *If $f : (\mathbb{R}^k \to U) \to \mathbb{R}$ is alternating, then $g(v) = f(v \circ A)$ is alternating for $A : \mathbb{R}^k \to \mathbb{R}^k$*

Proof. Since $g(v) = \det(A) f(v)$, the corrolary immediately follows. In particular, g is multilinear and alternating. □

The following lemma shows that if f is only multilinear, then $g(v) = f(v \circ A)$ is still multilinear, even though $g(v) = \det(A) f(v)$ no longer holds.

Lemma 23. *If $f : (\mathbb{R}^k \to U) \to \mathbb{R}$ is multilinear, then $g(v) = f(v \circ A)$ is multilinear for $A : \mathbb{R}^k \to \mathbb{R}^k$*

Proof. Using the isomorphism $\mathbb{R}^k \to U \simeq U^k$ we have, in coordinates $g(v_1, \dots, v_n) = f(\sum_{i=1}^k v_i A_{i,1}, \sum_{i=1}^k v_i A_{i,2}, \dots, \sum_{i=1}^k v_i A_{i,k})$, from which it is clear that g is multilinear because g as a function of a single v_i holding the other $v_{i'}$ fixed is a composition of linear functions. □

The lemma and the corollary show that the following definitions are well defined.

Definition 24. Let V be a k-dimensional vector space. A map $f : (V \to U) \to \mathbb{R}$ is called V-multilinear over a vector space U if $g : (\mathbb{R}^k \to U) \to \mathbb{R}$ given by $g(v) = f(v \circ A)$ is k-multilinear where $A : \mathbb{R}^k \simeq V$ is a basis for V.

and

Definition 25. Let V be a k-dimensional vector space. A map $f : (V \to U) \to \mathbb{R}$ is called V-alternating over a vector space U if $g : (\mathbb{R}^k \to U) \to \mathbb{R}$ given by $g(v) = f(v \circ A)$ is k-alternating where $A : \mathbb{R}^k \simeq V$ is a basis for V.

The lemma and corollary ensure that it does not matter which basis we pick. If the map is multilinear resp. alternating for one basis A, then it is multilinear resp. alternating for any other basis A' because if $g(v) = f(v \circ A)$ is multilinear resp. alternating, then $g(v \circ (A' \circ A^{-1})) = f(v \circ A')$ is multilinear resp. alternating. This makes multilinear and alternating maps functors covariant in V and contravariant in U.

4. Formalising constant k-densities

This chapter will show that constant k-densities are naturally formalised as (n-k)-linear alternating forms. We will start with $n-k=1$ and show that $n-1$ densities are naturally formalised as dual vectors, and then we'll generalise to arbitrary k.

4.1. Constant $n-1$ densities.

We first need a clearer picture of what a constant $n-1$ density is. A constant $n-1$ density is, intuitively, a density of parallel hyperplanes that is uniformly spaced out through space. Because this is a vague notion we will work with sets of hyperplanes with uniform finite spacing in between, and then take the limit as the spacing goes to zero. Given a vector $v \in \mathbb{R}^n$ we wish to define the intersection value of v with the set of hyperplanes. If the spacing between the planes goes to zero then the number of intersections with v goes to infinity, so we must count each intersection with correspondingly lower value as the spacing goes to zero so that the intersection value converges.

Definition 26. Let S be a $n-1$ dimensional subspace of $\mathbb{R}^n$ and let $v \in \mathbb{R}^n, v \notin S$. We define the ϵ-density of hyperplanes to be the set $D_{S,v}(\epsilon) = \{S + n\epsilon v : n \in \mathbb{Z}\}$.

The density of the planes in $D_{S,v}(\epsilon)$ goes up as $\epsilon \to 0$. Note that $D_{S,v}$ and $D_{S,v'}$ are equivalent if $v - v' \in S$.

Definition 27. We define the oriented intersection value $I(D_{S,v}(\epsilon), w)$ with a vector $w \in \mathbb{R}^n$ as follows. First compute the unoriented intersection value $n = |D_{S,v}(\epsilon) \cap \{tw : t \in [0,1]\}|$ and then put $I(D_{S,v}(\epsilon), w) = \epsilon n$ if v and w lie on the same side of S and $I(D_{S,v}(\epsilon), w) = -\epsilon n$ if v and w lie on opposite sides of S, and $I(D_{S,v}(\epsilon), w) = 0$ if w lies in S.

Each intersection of w with $D_{S,v}(\epsilon)$ is counted with weight ϵ, and positively or negatively according to whether v and w pierce through S in the same direction or in the opposite direction.

Definition 28. For $w \in \mathbb{R}^n$ define $I(D_{S,v}, w) = \lim_{\epsilon \to 0} I(D_{S,v}(\epsilon), w)$.

Lemma 29. *The limit exists and $I(D_{S,v}, w)$ is linear in w.*

Proof. Since $v \notin S$ the $n-1$ dimensional subspace S together with v spans $\mathbb{R}^n$, and any vector w may be uniquely written as a linear combination $w = w_S + rv$ where $w_S \in S$ and $r \in \mathbb{R}$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be the linear map defined by $f(w_S + rv) = r$. Clearly $\ker f = S$ and $f^{-1}(r) = S + rv$, and therefore

$$D_{S,v}(\epsilon) = \{f^{-1}(n\epsilon) \,:\, n \in \mathbb{Z}\}$$

Therefore, for $w \in \mathbb{R}^n$,

$$\begin{aligned} n &= |D_{S,v}(\epsilon) \cap \{tw \,:\, t \in [0,1]\}| \\ &= |\{f^{-1}(n\epsilon) \,:\, n \in \mathbb{Z}\} \cap \{tw \,:\, t \in [0,1]\}| \\ &= \epsilon \left\lfloor \frac{1}{\epsilon} |f(w)| \right\rfloor \end{aligned}$$

If we take the limit $\epsilon \to 0$ then $\epsilon \left\lfloor \frac{1}{\epsilon}|f(w)| \right\rfloor \to |f(w)|$ so $|I(D_{S,v}, w)| = |f(w)|$. It remains to be shown that the signs also match. There are three cases:

If w lies in S then $f(w) = 0$ and $I(D_{S,v}, w) = 0$.

If w lies on the same side of S as v then $w = w_S + rv$ with $r > 0$ so $f(w) = r > 0$ and $I(D_{S,v}, w) > 0$.

If w lies on the opposite side of S as v then $w = w_S + rv$ with $r < 0$ so $f(w) = r < 0$ and $I(D_{S,v}, w) < 0$.

Therefore $I(D_{S,v}, w) = f(w)$. □

We have shown that the intersection value $I(D_{S,v}, w)$ of a constant $n-1$ density is just a linear map $f : \mathbb{R}^n \to \mathbb{R}$ with $f(v) = 1$ and $f(S) = 0$. Conversely, given any $f \neq 0$ we may put $S = \ker f$ and v any vector such that $f(v) = 1$ and then f represents the $n-1$ density of planes parallel to S with orientation and density such that the vector v has intersection value 1 with the density. If $f = 0$ then $f(w) = 0$ for any w so this represents the zero density: there are no planes in this density so that intersection value with any vector is zero.

This brings us to the following definition:

Definition 30. A $n-1$ density on $\mathbb{R}^n$ is a dual vector $f : \mathbb{R}^n \to \mathbb{R}$ and for a vector $w \in \mathbb{R}^n$ we call the value $f(w)$ the intersection value of f with w.

4.1.1. *The parallelogram rule for covectors.* The dual vectors form a vector space, so we would like to interpret those vector space operations geometrically. Scalar multiplication af represents a density of hyperplanes that is increased by a factor of a. The case $a = -1$ corresponds to reversing the orientation of the hyperplanes, so that each intersection counts as the negative. Interpreting the addition $h = f + g$ is more tricky if f and g are not multiples of each other, so let's start in $\mathbb{R}^2$. A covector is a uniform density of oriented straight lines, but let us go back to discrete sets of straight lines f_k and g_k, and let us construct $h_k = f_k + g_k$. The lines of f_k and g_k form a grid of parallelograms. Take any side of one of the parallelograms and walk along that side. If the side belongs to a line of $\ker f$, then we accumulate no intersections with f when walking along that side, but when we walk the whole side we accumulate $1/k$ intersections with g, or $-1/k$ intersections with g, depending on the direction we walk in. Label the sides with arrows so that the contribution is $+1/k$. When we walk on such a side of a parallelogram we walk from one g_k line to the next, while staying on the same f_k line, or vice versa. If we walk all the way around a parallelogram we end up where we started, so the net intersections are zero, so somewhere along our walk we had first a $+1/k$ contribution and then

a $-1/k$ contribution on two consecutive sides, when we walked from one corner of the parallelogram to the opposite corner. Therefore the line connecting those opposite corners has to be part of $\ker h$. We therefore construct h_k as the system of lines formed by drawing straight lines through all those opposite corners. Then we've at least got the direction of h_k lines right. To see that this also gives the correct density, consider walking one of the sides of the parallelogram. This gives a contribution of $\pm 1/k$, but it also walks from one of the h_k lines to the next, so the contribution is also $\pm 1/k$. By choosing the right orientation for h_k we get the correct result.

In higher dimensions it is essentially the same, except that the hyperplanes of f_k and g_k now intersect in whole hyperplanes instead of points. In $\mathbb{R}^3$ two planes intersect in a line, for instance. Instead of drawing a line through two opposite points of a parallelogram, we draw a hyperplane through two of those intersection lines. The situation can be visually understood in terms of the 2D case by imagining that all the extra dimensions lie orthogonal to the plane of the drawing. This is analogous to understanding the ordinary parallelogram rule for vector addition in $\mathbb{R}^n$ by using the ordinary drawing for $u+v$ by considering the subspace spanned by u and v, and imagining that all the other dimensions lie orthogonal to the drawing.

4.2. **Constant k densities.** A constant k density is, intuitively, a density of parallel affine subspaces of dimension k that is uniformly spaced out through space. Because this is a vague notion we will work with sets of affine subspaces with uniform finite spacing in between, and then take the limit as the spacing goes to zero. We want to define the oriented intersection value of a $n-k$ parallelepiped with the k density, so we'll take a quick look at oriented k parallelepipeds.

We can specify a k-parallelepiped in a vector space U by giving k vectors v_i so that the parallelepiped is the set $P = \{\sum_{i=1}^{k} x_i v_i \,:\, x_i \in [0,1]\}$. More compactly, we take a linear map $V : \mathbb{R}^k \to U$ which is in general given by

$$V(x) = x_1 v_1 + x_2 v_2 + \cdots + x_k v_k$$

and set $P = V([0,1]^k)$, the image of the map under the unit k-cube. For the purpose of computing intersections with densities it does not matter what the shape of the parallelepiped is: only the volume and the direction of the parallelepiped matters. If we scale one axis by 2 and another axis by $\frac{1}{2}$ then the total number of intersections will stay the same. By reparameterising by a $\det A = 1$ map we preserve the volume and direction, but change the shape and scale in the subspace in which the parallelepiped lies. This motivates the following definition:

Definition 31. A pure k-vector on a vector space U is a linear map $V : \mathbb{R}^k \to U$. We call two pure k-vectors V_1, V_2 equivalent if $V_1 = V_2 \circ A$ where $A : \mathbb{R}^k \to \mathbb{R}^k$ with $A \in SL(k)$.

We will use k-vectors to measure the intersection value with a $(n-k)$-density. Note that this definition gives ordinary vectors when $k = 1$: the map V is then of type $\mathbb{R} \to U$ which is of the form $V(r) = rv$ with $v \in U$, and $V([0,1])$ gives the content of the parallelepiped, which is a line segment when $k = 1$. Unlike the set $V([0,1])$, this definition of k-vector carries an orientation with it: the 1-vector $V(r) = rv$ is not equivalent to the 1-vector $V(r) = r(-v)$ because there is no $A \in SL(1)$ that transforms the one into the other. This is true generally: the set $V([0,1]^k)$ does not carry an orientation, but the map V, up to reparameterisation

a $A \in SL(k)$, does. To reverse the orientation of V, we pick any $A \in GL(k)$ with $\det A = -1$ and form $V' = V \circ A$. The map A reverses the orientation of the parameter space, which reverses the orientation of the parallelepiped. We will see later that reversing the orientation of a k-vector negates its intersection value with any $(n-k)$-density.

Given two parallelepipeds $P_1 \subset U$ and $P_2 \subset U$ we can create a bigger parallelepiped $P_3 = \{u_1 + u_2 : u_1 \in P_1, u_2 \in P_2\}$. We can define a similar operation on k-vectors.

Definition 32. Given a k-vector V_1 and a l-vector V_2 on U we define the $(k+l)$-vector $V_3 = V_1 \wedge V_2$ as $V_3(x,y) = V_1(x) + V_2(y)$ where $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^l$.

For instance, if V_1 and V_2 are 1-vectors, then $V_1 \wedge V_2$ is the 2-vector representing the parallelogram made by V_1 and V_2.

Lemma 33. *The $\wedge$-product on k-vectors is associative.*

Proof. Trivial. □

Definition 34. Let $L : U_1 \to U_2$ be a linear map between vector spaces. The pushforward of a k-vector V on U_1 is a k-vector L_*V on U_2 given by $L_*V = L \circ V$.

Just as a k-vector is given by k vectors $v_i \in U$ or as a map $V : \mathbb{R}^k \to U$, a $(n-k)$-density is given by k covectors $f_i \in U^*$ or as a map $F : U \to \mathbb{R}^k$. The k-vector represents the parallelepiped formed by the v_i, and the $(n-k)$-density represents the $(n-k)$-dimensional intersections of the hyperplane densities specified by the f_i. From a given f_i we may select one of its hyperplanes with the equation $f_i(u) = x_i$ with $x_i \in \mathbb{R}$. We select a different hyperplane as we vary x_i. By selecting a value x_i for $i \in \{1..k\}$ we select a hyperplane for each of the f_i. These chosen hyperplanes will in general intersect in a $(n-k)$-dimensional affine subspace, which corresponds to solving the k equations $f_i(u) = x_i$ for u simultaneously. The $(n-k)$-density built from the f_i represents the intersection of the hyperplane densities f_i.

In the setting of $F : U \to \mathbb{R}^k$ we have

$$F(u) = (f_1(u), f_2(u), \ldots, f_k(u))$$

and selecting x_i for $i \in \{1..k\}$ corresponds to selecting a point $x \in \mathbb{R}^k$, and the in general $(n-k)$-dimensional affine subspace is $F^{-1}(x)$. This brings us to the following definition:

Definition 35. A pure constant $(n-k)$-density on a vector space U is a linear map $F : U \to \mathbb{R}^k$. We call two pure $(n-k)$-densities F_1, F_2 equivalent if $F_1 = A \circ F_2$ where $A : \mathbb{R}^k \to \mathbb{R}^k$ with $A \in SL(k)$.

As with k-vectors, we may reparameterise $(n-k)$-densities using $A \in SL(k)$, and if we take $A \in GL(k)$ with $\det A = -1$ then $F' = A \circ F$ reverses the orientation of F.

Definition 36. Given a pure constant $(n-k)$ and $(n-l)$ density F_1 and F_2 on U, we define the $(n-k-l)$ density $F_3 = F_1 \wedge F_2$as $F_3(u) = (F_1(u), F_2(u))$, where $u \in U$.

If F_1 and F_2 are k-densities, then $F_1 \wedge F_2$ will represent their intersection.

Definition 37. Let $L : U_1 \to U_2$ be a linear map between vector spaces. The pullback of a pure constant $(n-k)$-density F on U_2 is a pure constant $(n-k)$-density on L^*F on U_1 given by $L_*F = F \circ L$.

4.2.1. *The intersection value of a $(n-k)$-density with a k-vector.* If F and V are in general position then an affine subspace $F^{-1}(x)$ will intersect $V([0,1]^k)$ in zero or one points. Consider a grid of points with spacing $\epsilon > 0$ in $\epsilon\mathbb{Z}^k \subset \mathbb{R}^k$. The preimage $F^{-1}(\epsilon\mathbb{Z}^k)$ consists of an infinite number of translations of $F^{-1}(0)$ with uniform spacing in between. Each of these intersects $V([0,1]^k)$ in zero or one points, and in general position only a finite number of these intersect $V([0,1]^k)$, say $n(\epsilon) = |V(C) \cap F^{-1}(\epsilon\mathbb{Z}^k)|$. The limit of $\epsilon^k n(\epsilon)$ is the number of weighted intersections we wish to count. We have

$$\epsilon^k n(\epsilon) = \epsilon^k |F(V(C) \cap F^{-1}(\epsilon\mathbb{Z}^k))| = \epsilon^k |F(V(C)) \cap \epsilon\mathbb{Z}^k|$$

As the grid spacing becomes finer, the number of grid points $\epsilon\mathbb{Z}^k$ in $F(V([0,1]^k))$ approaches $\epsilon^{-k} Vol(F(V([0,1]^k)) = \epsilon^{-k} \det(F \circ V)$. So as $\epsilon \to 0$, the number of weighted intersections $\epsilon^k n(\epsilon) \to \det(F \circ V)$. So $\det(F \circ V)$ is equal to the weighted number of intersections of F with V. If F and V are not in general position then $\det(F \circ V)$ is zero. This can be understood geometrically as saying that V lies parallel to F, and by an arbitrarily small translation V can be positioned in between the affine subspace of F, making the intersection zero.

The intersection value of a k-vector V and a k-density F can be computed by $\det(F \circ V)$. Given a k-density F we have the map $V \mapsto \det(F \circ V)$ that computes the intersection value for any V. We take linear combinations of such maps to get general k-densities:

Definition 38. The set of constant $(n-k)$-densities on a vector space U is $\bigwedge^k U = span\{V \mapsto \det(F \circ V) : F : U \to \mathbb{R}^k\}$ with $V : \mathbb{R}^k \to U$.

We extend the wedge product and pullback to (non-pure) constant k-densities as follows.

Definition 39. For constant $(n-k)$-density f given by a pure constant $(n-k)$-density F, we define the pullback $L^* f$ as the constant density given by the pure constant density $L^* F$. For linear combinations of pure densities, we let L^* act linearly.

That is, the pullback acts as usual on pure densities, and we extend it to linear combinations by acting linearly.

Definition 40. For constant $(n-k)$ and $(n-l)$ densities f_1, f_2 given by pure constant densities F_1, F_2, we define the wedge product $f_1 \wedge f_2$ as the pure constant density given by $F_1 \wedge F_2$. For linear combinations of pure densities, we let $\wedge$ act bilinearly.

Lemma 41. *$f \in \bigwedge^k U$ satisfies $f(V \circ A) = \det(A) f(V)$.*

Proof. The formula holds for $f = V \mapsto \det(F \circ V)$ because det is multiplicative, and therefore it also holds for linear combinations. □

Corollary. *For f, g be constant $(n-k)$ and $(n-l)$ densities respectively, it holds that $f \wedge g = (-1)^{kl} g \wedge f$.*

Proof. $(f \wedge g)(V) = (g \wedge f)(V \circ A)$ where A is the permutation matrix that swaps the first k with the remaining l components. By the lemma, $(g \wedge f)(V \circ A) = \det(A)(g \wedge f)(V)$, and $\det(A) = (-1)^{kl}$. Therefore $f \wedge g = (-1)^{kl} g \wedge f$ □

Compare this to the traditional definition of f as a multilinear alternating map $f : U^k \to \mathbb{R}$. A k-tuple of vectors $v \in U^k$ gives rise to a map $V \in \mathbb{R}^k \to U$ via $V(e_i) = v_i$, and vice versa we may put $v_i = V(e_i)$ to go in the other direction, so k-tuples of vectors U^k are in bijection with maps $\mathbb{R}^k \to U$. That $f : U^k \to \mathbb{R}$ is alternating means that $f(\sigma v) = sign(\sigma)f(v)$. This corresponds to taking A to be the permutation matrix associated with σ, and then $sign(\sigma) = \det(A)$.

5. Formalising varying k-densities

We're in position to tackle varying k-densities now that we have found a mathematical representation of constant k-densities. If the k-density is slowly varying in the sense that the direction and density of the k-planes is slowly varying, we may represent a varying k-density as an assignment of a constant k-density at each point. So a varying k-density is a map $\alpha : M \to \bigwedge^k M$ where $M = \mathbb{R}^n$ is the space we're working on (we'll deal with manifolds later). If we want the density to be slowly varying then we insist that α is continuous or even smooth. This is well defined because $\bigwedge^k \mathbb{R}^n$ is a finite dimensional vector space. Given such a map α and a (n-k)-vector v attached to point p, we can measure the local flux through v as $\alpha(p)(v)$, which is the flux through v assuming that α is the constant density $\alpha(p)$. This is accurate if α is slowly varying and v is sufficiently small.

5.1. **Flux through curved surfaces.** We wish to determine the flux of a k-density α through a curved (n-k)-surface S, which we denote as $\int_S \alpha$. To do this we approximate S by small (n-k)-vector segments and add up all the fluxes through those. Let us first deal with the case $k = n - 1$, that is, we have a varying hyperplane density α and a curve $S = \{\gamma(t) : t \in [0, 1]\}$ and we wish to calculate how many surfaces of α the curve pierces through. We take n points t_i to get points on the curve $p_i = \gamma(t_i)$ and approximate the curve with the straight segments $\Delta p_i = p_{i+1} - p_i$ going from p_i to p_{i+1}. If the segment is sufficiently small we may approximate the intersection value as $\alpha(p_i)(\Delta p_i)$, and in addition $\Delta p_i = p_{i+1} - p_i = \gamma(t_{i+1}) - \gamma(t_i) \approx \gamma'(t_i)\Delta t_i$, so the intersection value of segment i is approximately $\alpha(\gamma(t_i))(\gamma'(t_i))\Delta t_i$. If we add those all up we get:

$$\int_S \alpha \approx \sum_{i=1}^{n} \alpha(\gamma(t_i))(\gamma'(t_i))\Delta t_i$$

By taking the limit $n \to \infty$ this becomes the Riemann sum for the integral $\int \alpha(\gamma(t))(\gamma'(t))dt$, so we take this as the definition of $\int_S \alpha$:

$$\int_S \alpha = \int \alpha(\gamma(t))(\gamma'(t))dt$$

For general k we have a (n-k)-surface $S = \{\gamma(t) : t \in U\}$ where the parameterisation domain $U \subset \mathbb{R}^{n-k}$. The map γ transforms a small rectangle (or higher dimensional analogues) in the domain into a small parallelogram in $\mathbb{R}^n$. We therefore cut up the domain U into small rectangles, apply γ to get parallelograms in $\mathbb{R}^n$, and then sum the flux through those parallelograms. This leads us to

$$\int_S \alpha \approx \sum_{i=1}^{n} \alpha(\gamma(t_i))(\gamma'(\Delta_1 t_i \wedge \cdots \wedge \Delta_{n-k} t_i))$$

This is a Riemann sum an integral over U, so we define:

$$\int_S \alpha = \int_U \alpha(\gamma(t))(D\gamma(t)(e_1 \wedge \cdots \wedge e_{n-k}))dt_1 \dots dt_{n-k}$$

This definition agrees with the $k = n - 1$ definition.

5.2. **Wedge product.** We define the wedge product of two densities α and β as the pointwise wedge product of their locally constant densities: $(\alpha \wedge \beta)(p) = \alpha(p) \wedge \beta(p)$. This product inherits the algebraic properties of the wedge product on constant densities. It also inherits a geometric interpretation: the wedge product of locally constant densities is their intersection, so the wedge product of varying densities is also their intersection, because it is the intersection locally. Examples:

(1) We have the vertical planes dx and horizontal planes dy, and the intersection is the uniform point density $dx \wedge dy$.
(2) Given functions $f(x, y)$ and $g(x, y)$ we have vertical planes $f dx$ that vary in density and another density of vertical planes $g dx$ that vary in density. These planes don't intersect because they both go horizontally, and indeed $(f dx) \wedge (g dx) = fg(dx \wedge dx) = 0$.
(3) Given functions $f(x, y)$ and $g(x, y)$ we have vertical planes $f dx$ that vary in density and horizontal planes $g dy$ that vary in density. These planes do intersect because $f dx$ goes horizontally whereas $g dy$ goes vertically, and indeed $(f dx) \wedge (g dx) = fg(dx \wedge dy)$, which is a density of points with local density $f(x, y)g(x, y)$.

5.3. **Exterior derivative and Stokes' theorem.** The exterior derivative of a k-density represents the boundary of the density. If we imagine the density being made up of a finite number of submanifolds then the exterior derivative is the collection of boundaries of those submanifold. Consider the case $k = 1$, namely a line density. The boundary of a line segment is two points, so the boundary of a 1-density α is a density of points, that is, a 0-density $d\alpha$. A 0-density $d\alpha$ allows us to measure how many points are inside a small box around a point p by doing $d\alpha(p)(v)$ where v is an n-vector that represents the box. The boundary points inside the box are connected to lines of α. Each line in α starts at a negative point and ends at a positive point. The intersection value of lines through the sides of the box is equal to the value of points inside the box. We assume that α is of the form $f dx^2 \wedge \cdots \wedge dx^n$ and later extend the results to arbitrary $n - 1$ forms by linearity. The form $\alpha$is smooth so around a point p we may take a linear approximation and write $\alpha = cx^2 \wedge \cdots \wedge dx^n + \sum_i a_i x^i dx^2 \wedge \cdots \wedge dx^n$. The constant density $cx^2 \wedge \cdots \wedge dx^n$ has no boundary so we may also assume $c = 0$. The question we now wish to answer is *how much boundary value does α have inside a box?*

The lines of α are going in the x^1 direction. Any variation in the density of α perpendicular to x^1 contributes no boundary points. Only the variation in line density in the direction of the lines themselves results in lines starting and ending. Therefore the boundary of α is only dependent on the term $a_1 x^1 dx^2 \wedge \cdots \wedge dx^n$ and the other terms do not contribute to the boundary. The density of the lines is increasing linearly with coefficient a_1 and the number of boundary points is equal to the number of lines that start as we move in the x^1 direction. Therefore $d\alpha = a_1 dx^1 \wedge \cdots \wedge dx^n$, so in general if $\alpha = \sum_i f_i dx^I$ where I is a multi-index that

includes all but i,

$$d\alpha = \sum_i (-1)^i \frac{\partial f_i}{\partial x_i} dx^1 \wedge \cdots \wedge dx^n$$

5.4. **Pullback.** Given a smooth map $f : A \to B$ where A, B are open subsets of $\mathbb{R}^n, \mathbb{R}^m$ and a form α on $\mathbb{R}^m$ we define the pullback as the pointwise pullback for constant densities along the linear approximation to f:

$$f^*\alpha(p) = Df(p)^*\alpha(f(p))$$

This means that

$$f^*\alpha(p)(v) = \alpha(f(p))(Df(p)v)$$

So given a parallelepiped v at p we measure the intersection number with $f^*\alpha$ by transporting it forward to $Df(p)v$ at $f(p)$ and measuring that intersection number with α.

How do we visualise this? Consider the image of f in B and support this is a k-surface in B. The form α intersects with this k-surface. We transport the result of this intersection back to A, and that's the density $f^*\alpha$, which may in general be a density of different dimension than α, although its codimension stays the same.

5.4.1. *Change of variables.* The following proof is a slight modification of the proof by Lax [5].

Theorem 42. *Let N, M compact n-manifolds and α an n-form on M, and $f : N \to M$ an orientation preserving diffeomorphism. Then $\int_N f^*\alpha = \int_M \alpha$.*

Proof. By induction on n.

If $n = 0$ then N and M are a finite set of points a_i, b_i and $b_i = f(a_i)$ and the corresponding points have the same orientation $\sigma_i \in \{-1, 1\}$. Then

$$\int_B \alpha = \sum_i \sigma_i \alpha(b_i) = \sum_i \sigma_i \alpha(f(a_i))$$

. For the left hand side, $f^*\alpha(x) = \alpha(f(x))$, so

$$\int_A f^*\alpha = \sum_i \sigma_i \alpha(f(a_i))$$

. So the theorem holds for $n = 0$.

If $n > 0$ we may cut N, M into simply connected regions using a partition of unity, so it suffices to show the theorem in the case that N, M are simply connected. Because α is an n-form on an n-manifold, $d\alpha = 0$, and because M is simply connected we may apply Poincaré's lemma to find A such that $dA = \alpha$. Apply Stokes' theorem to the left hand side:

$$\int_N f^*\alpha = \int_N f^* dA = \int_N df^* A = \int_{\partial N} f^* A$$

and to the right hand side:

$$\int_M \alpha = \int_M dA = \int_{\partial M} A$$

By the inductive hypothesis, $\int_{\partial N} f^* A = \int_{\partial M} A$ because A is an $(n-1)$-form. So the theorem holds for $n > 0$. □

5.5. **k-densities on manifolds.** By analogy with the pullback we may think of a density on the k-surface as given by a density on A. That this doesn't depend on how the surface is parameterised is because all differential form operations respect the pullback.

6. Examples and elementary applications

6.1. **Solid of revolution.** We can specific a volume in three dimensions by using an indicator function in cylindrical coordinates: $V : \mathbb{R}^3 \to \{0, 1\}$ such that a point (x, r, θ) is in the volume if $V(x, r, \theta) = 1$. The volume is given by

$$Vol(V) = \int_{\mathbb{R}\times[0,\infty)\times[0,2\pi]} V(x, r, \theta) r dr \wedge d\theta \wedge dx$$

If the volume is axisymmetric then I does not depend on θ and the integral becomes

$$\begin{aligned} Vol(V) &= \int_{\mathbb{R}\times[0,\infty)\times[0,2\pi]} V(x, r) r dr \wedge d\theta \wedge dx \\ &= \int_{\mathbb{R}\times[0,\infty)} 2\pi V(x, r) r dr \wedge dx \end{aligned}$$

Choosing S as the integration domain, we get

$$Vol(V) = \int_V 2\pi r dr \wedge dx$$

The integrand $2\pi r dr \wedge dx$ may be interpreted as the volume of a ring of radius r and thickness dr and width dx, such that each little square of the two dimensional $V(x, r)$ contributes $2\pi r dr \wedge dx$ to the volume when rotated around the x-axis. We can find the following antiderivatives of $2\pi r dr \wedge dx$:

$$d(-2\pi x r dr) = -2\pi r dr \wedge dx - 2\pi x dr \wedge dr = 2\pi r dr \wedge dx$$

and

$$d(\pi r^2 dx) = 2\pi r dr \wedge dx$$

Using Stokes' theorem we can express the volume as a line integral:

$$\begin{aligned} Vol(V) &= \int_V 2\pi r dr \wedge dx \\ &= \int_V d(-2\pi x r dr) \\ &= \int_{\partial V} -2\pi x r dr \end{aligned}$$

and

$$\begin{aligned} Vol(V) &= \int_V 2\pi r dr \wedge dx \\ &= \int_V d(\pi r^2 dx) \\ &= \int_{\partial V} \pi r^2 dx \end{aligned}$$

Compare this to the traditional calculus formulas for surfaces of revolution when V is bounded by two functions. When the surface is bounded between $x = f(r)$ and $x = g(r)$:

$$Vol(V) = 2\pi \int_a^b r|f(r) - g(r)|dr$$

When the surface is bounded between $r = f(x)$ and $r = g(x)$:

$$Vol(V) = \pi \int_a^b |f(x)^2 - g(x)^2|dx$$

The integral $\int_{\partial V}$ picks up both f and g terms along the boundary ∂V.

6.2. **Alternative to Lagrange multipliers.** Suppose we want to minimise $f(x, y) = ax + by$ along the constraint $g(x, y) = x^2 + y^2 - R^2 = 0$. According to the method of Lagrange multipliers we form the Lagrangian $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ and set its partial derivatives to zero:

$$\begin{aligned} \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} &= 0 \\ \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} &= 0 \\ g(x, y) &= 0 \end{aligned}$$

For the example this gives

$$\begin{aligned} a + \lambda 2x &= 0 \\ b + \lambda 2y &= 0 \\ x^2 + y^2 &= R^2 \end{aligned}$$

We eliminate λ from the first equation to find $\lambda = -a/2x$. When substituting this into the equation equation we get $b + (-a/2x)2y = 0$, or $ay = bx$. We may find the solution by using the constraint $x^2 + y^2 = R^2$.

We will now solve this problem in a slightly different way. The first two equations say that $\nabla f = -\lambda \nabla g$, so the gradient of f is a multiple of the gradient of g, that is, the gradient of f and the gradient of g are in the same direction. Geometrically this is the same as saying that the contours of f and g are parallel. We may express this with the wedge product as $df \wedge dg = 0$. For the example $f(x, y) = ax + by$ along the constraint $g(x, y) = x^2 + y^2 - R^2 = 0$ this gives the following equation:

$$(adx + bdy) \wedge (2xdx + 2ydy) = (2ay - 2bx)dx \wedge dy = 0$$

We immediately get the condition $ay = bx$. So instead of introducing Lagrange multipliers we may also solve the system

$$\begin{aligned} df \wedge dg &= 0 \\ g(x, y) &= 0 \end{aligned}$$

Which is the same as having eliminated λ from the Lagrange multiplier system.

7. Thermodynamics

Thermodynamics is concerned with the macroscopic behaviour of systems that are microscopically very complex. The quintessential example is a container with a given quantity of gas and a piston that can be used to compress the gas. A lot of variables are required to descibe the all the positions and velocities of the gas molecules, but the macroscopic state can be characterised by global quantities such as the temperature, pressure, and position of the piston (or equivalently, the volume to which the gas is compressed). In fact, two variables are sufficient: given the temperature T and the volume V, the pressure P is determined by the equation $PV = nRT$ where nR is a constant, so we may specify the state by (P, V), (V, T), or (T, P). The state can be manipulated by moving the piston and by heating and cooling the gas. We move the system from state S to S' via path γ by doing the movement of the piston and the heating/cooling continuously. It turns out that the total amount of heat Q required to move from S to S' depends on the path γ, as does the total amount of work W that was used to move the piston, but the sum $Q + W$ only depends on the initial and final states. We can express the heat as the integral of a one form q on the state space so that $Q = \int_\gamma q$ and similarly $W = \int_\gamma w$. The statement that $Q + W$ doesn't depend on the path is equivalent to the statement that $q + w$ is closed. This is the first law of thermodynamics, and it can be stated in differential and integral form:

$$d(q + w) = 0$$
$$\oint_\gamma q + w = 0$$

Because $q + w$ is closed and the state space is $\mathbb{R}^2$ we can find a potential E (the energy) such that $dE = q + w$. This leads to two more equivalent statements of the first law:

$$dE = q + w$$
$$\int_\gamma q + w = E(\gamma_1) - E(\gamma_0)$$

The second law of thermodynamics is similar: the form $\frac{q}{T}$ is closed:

$$d(\frac{q}{T}) = 0$$
$$\oint_\gamma \frac{q}{T} = 0$$

We can find a potential S (the entropy) such that $dS = \frac{q}{T}$. This leads to two more equivalent statements of the second law:

$$dS = \frac{q}{T}$$
$$\int_\gamma \frac{q}{T} = S(\gamma_1) - S(\gamma_0)$$

Each of these may be derived without differential forms, but the formalism makes this particularly effortless. For the gas example we have $w = PdV$ and differential

forms make it easy to show that the work done by a Carnot cycle is the area enclosed in the P, V plane:

$$\int_{\partial A} w = \int_A dw = \int_A d(PdV) = \int_A dP \wedge dV$$

We can also use differential forms to derive the Maxwell relations. Consider the differential $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$ and wedge both sides with dy to get $df \wedge dy = \frac{\partial f}{\partial x}dx \wedge dy$. We may "divide" this equation by $dx \wedge dy$ to get our formula

$$(\frac{\partial f}{\partial x})_y = \frac{df \wedge dy}{dx \wedge dy}$$

where $(\frac{\partial f}{\partial x})_y$ indicates that we are taking the derivative of f with respect to x keeping y constant. If dividing forms seems questionable, note that the space of n-forms in n dimensions is one dimensional, so for any two n-forms α, β we have $\alpha = k\beta$ for some $k \in \mathbb{R}$, so we put $\frac{\alpha}{\beta} = k$.

We substitute the definition of work $w = PdV$ and the second law $q = TdS$ into the first law $d(q+w) = 0$ to get $dT \wedge dS = -dP \wedge dV$. We can easily obtain Maxwell relations by dividing this equation by various 2-forms, for instance by $dV \wedge dS$:

$$\frac{dT \wedge dS}{dV \wedge dS} = -\frac{dP \wedge dV}{dV \wedge dS} = \frac{dP \wedge dV}{dS \wedge dV}$$

We have found $(\frac{\partial T}{\partial V})_S = (\frac{\partial P}{\partial S})_V$, or in words "The temperature change with volume at constant entropy is equal to the pressure change with entropy at constant volume". We obtain the three other Maxwell relations by dividing by $dT \wedge dP$, $dT \wedge dV$, and $dS \wedge dP$. As another application, consider Euler's triple product rule for three coordinates on a two dimensional surface:

$$(\frac{\partial x}{\partial y})_z(\frac{\partial y}{\partial z})_x(\frac{\partial z}{\partial x})_y = -1$$

This is easily proved using $(\frac{\partial f}{\partial x})_y = \frac{df \wedge dy}{dx \wedge dy}$:

$$\frac{dx \wedge dz}{dy \wedge dz} \cdot \frac{dy \wedge dx}{dz \wedge dx} \cdot \frac{dz \wedge dy}{dx \wedge dy} = \frac{dx \wedge dz}{dz \wedge dx} \cdot \frac{dy \wedge dx}{dx \wedge dy} \cdot \frac{dz \wedge dy}{dy \wedge dz} = (-1)^3 = -1$$

The same method generates various generalisations relating the partial derivatives of n coordinates on a $m < n$ dimensional surface. For instance, adding $\wedge dw$ to each term gives

$$(\frac{\partial x}{\partial y})_{z,w}(\frac{\partial y}{\partial z})_{x,w}(\frac{\partial z}{\partial x})_{y,w} = -1$$

on a three dimensional surface, which is arguably not much of a generalisation because the w just goes along for the ride. However, on a three dimensional surface we also get

$$(\frac{\partial x}{\partial y})_{z,w}(\frac{\partial y}{\partial z})_{w,x}(\frac{\partial z}{\partial w})_{x,y}(\frac{\partial w}{\partial x})_{y,z} = 1$$

which permutes the coordinates cyclically like Euler's triple product rule. For $m = n - 1$ we have the following cyclic product identity

$$\prod_{i=0}^{n}(\frac{\partial x_i}{\partial x_{i+1}})_{x_{i+2},\ldots,x_{i+n}} = (-1)^n$$

where the index in x_i is taken modulo n. For $n = 2$ this is simply

$$\frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial x} = 1$$

The cyclic identity is easily generalised to arbitrary $m < n$. It might be fun to investigate the precise set of identities that this method can generate. For instance, is the set of identities generated by the cyclic product identity plus variables like w that go along for the ride?

8. Liouville's theorem

In Hamiltonian mechanics the time evolution is governed by the Hamiltonian function $H(p, q)$ which is a function of n positions p_i and n momenta q_i. The space $(p, q) \in \mathbb{R}^{2n}$ is called phase space. The equations of motion are

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}$$
$$\dot{q}_i = +\frac{\partial H}{\partial p_i}$$

These equations determine the trajectory of a point in phase space. Liouville's theorem is concerned with the time evolution of a random initial state: suppose the initial state is chosen from some probability distribution on phase space, what does the probability distribution look like at time t? The interpretation of volume forms as point densities helps: a probability distribution is just a $2n$-form α on phase space. If (p, q) is a point in phase space and v is a $2n$-vector on phase space, then $\alpha(p, q)(v)$ measures the probability mass or the number of points in the volume v at (p, q). Let $\phi_t : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be the flow that advances a state forward by time t, then the map ϕ^*_{-t} advances the probability distribution by time t. That this is so can be seen in the following way: let (p, q) be a point in phase space and v be a $2n$-vector, i.e. a small volume element in phase space. The number of points that lie inside this small volume at time t can be found by evolving this small volume backwards in time to $t = 0$ and then measuring the number of points that lie in that small volume: $\alpha(\phi_{-t}(p, q))(D\phi_{-t}(p, q)(v)) = \phi^*_{-t}\alpha(p, q)(v)$. We define $\omega = dp_1 \wedge dq_1 \wedge \cdots \wedge dp_n \wedge dq_n$ to be the phase space volume form, and write $\alpha = \rho\omega$. We get $\phi^*_{-t}\alpha = \phi^*_{-t}(\rho)\phi^*_{-t}(\omega)$. We therefore put $\omega_t = \phi^*_t(\omega)$ and study its behaviour.

Theorem 43. *(Lioville)* $\omega_t = \omega$.

Proof. We have

$$\begin{aligned}
\phi^*_t(\omega) &= \phi^*_t(dp_1 \wedge dq_1 \wedge \cdots \wedge dp_n \wedge dq_n) \\
&= d(\phi^*_t p_1) \wedge d(\phi^*_t q_1) \wedge \cdots \wedge d\phi^*_t(p_n) \wedge d\phi^*_t(q_n) \\
&= dp_1(t) \wedge dq_1(t) \wedge \cdots \wedge dp_n(t) \wedge dq_n(t)
\end{aligned}$$

We now show that the derivative is zero. Suppose first that $n = 1$ so that we have $\omega_t = dp(t) \wedge dq(t)$. Compute $\dot{\omega}_t$ using the product rule:

$$\begin{aligned}\dot{\omega}_t &= d\dot{p} \wedge dq + dp \wedge d\dot{q} \\ &= d(-\frac{\partial H}{\partial q}) \wedge dq + dp \wedge d(\frac{\partial H}{\partial p}) \\ &= -\frac{\partial^2 H}{\partial p \partial q} dp \wedge dq + dp \wedge dq \frac{\partial^2 H}{\partial p \partial q} \\ &= 0\end{aligned}$$

If $n > 1$ we still have $\dot{\omega}_t = 0$ because now the product rule gives $2n$ terms, each (p, q) pair of which cancels by the same calculation as for a single pair. Since ω_t is consant, we find $\omega_t = \omega$ for all t. □

This theorem implies that the probability density ρ evolves in a very simple way: $\rho_t = \phi_{-t}^* \rho = \rho(p(-t), q(-t))$. That is, to find the density near a point (p, q) at time t, we simply evolve (p, q) backwards to $t = 0$ and find the density of ρ there. In particular, probability densities that are invariant under time evolution must be conserved quantities. If the conserved quantities of the Hamiltonian H are known to be some finite set, then the time translation invariant probability densities must be a function of those quantities only, and not of all the p_i, q_i. This is useful in statistical mechanics: if we have a large number of gas particles in a box governed by some Hamiltonian we will usually have only one conserved quantity, namely the energy H. Lioville's theorem then tells us that any equilibrium state is of the form $\rho(p, q) = f(H(p, q))$ where $f : \mathbb{R} \to \mathbb{R}$, which is a lot simpler than an arbitrary $\rho : \mathbb{R}^{2n} \to \mathbb{R}$.

9. Currents

9.1. **Distributions.** We've seen that in order to represent manifolds by forms we need Dirac delta functions. The theory of distributions is the formalisation of Dirac's original idea. The space of distributions $D(\mathbb{R}^n)$ on $\mathbb{R}^n$ is defined to be the dual space of smooth functions $S(\mathbb{R}^n) \subset \mathbb{R}^n \to \mathbb{R}$. Given a smooth test function $\phi : \mathbb{R}^n \to \mathbb{R}^n$ and a distribution $f \in D(\mathbb{R}^n)$ we have the pairing $\langle f, \phi \rangle = f(\phi) \in \mathbb{R}$. Given a measurable function g we can define the associated distribution f_g as

$$\langle f_g, \phi \rangle \quad = \quad \int_{\mathbb{R}^n} g(x)\phi(x)dx$$

The test functions ϕ can be seen as spread out points: ordinary functions $g : \mathbb{R}^n \to \mathbb{R}$ are evaluated in a point, whereas a distribution $f_g : S(\mathbb{R}^n) \to \mathbb{R}^n$ is evaluated in a spread out point. By forgoing the ability to evaluate in an exact point we gain a larger space of generalised functions that includes the Dirac delta function:

$$\langle \delta(x), \phi \rangle \quad = \quad \phi(x)$$

Given any distribution we can define its derivative

$$\langle Df, \phi \rangle \quad = \quad -\langle f, D\phi \rangle$$

One can check that this definition of the derivative corresponds to the ordinary notion of derivative for distributions that come from differentiable functions using

integration by parts. Define the Heaviside step function as

$$\theta(x) \quad = \quad \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

The derivative $D\theta = \delta(0)$. We can build up distributions on $\mathbb{R}^n$ as products of distributions on $\mathbb{R}^1$, and in that case $\delta_n(x) = \delta_1(x_1)\delta_1(x_2)\cdots\delta_1(x_n)$. Since any distribution has a derivative, so does the Dirac delta:

$$\langle D\delta(x), \phi\rangle \quad = \quad -\langle \delta(x), D\phi\rangle = -D\phi(x)$$

where D is some derivative operator that satisfies the partial integration formula. For instance, for $D = \frac{\partial}{\partial x}$ we have

$$\left\langle \frac{\partial}{\partial x}\delta(x), \phi \right\rangle \quad = \quad -\frac{\partial}{\partial x}\phi(x)$$

We write $\frac{\partial}{\partial x}f = \partial_x f = f_x$, and in 1 dimension f'. For the multidimensional delta function expressed as a product $\delta(p) = \delta((x,y,z)) = \delta(x)\delta(y)\delta(z)$, we have $\partial_x\delta(p) = \delta_x(p) = \delta'(x)\delta(y)\delta(z)$, and similarly for partial derivatives with respect to y and z.
To establish properties of the Dirac delta function we use that it is the derivative of the Heaviside function.

- We determine $\delta(ax)$. Let $a > 0$, then $\theta(ax) = \theta(x)$. If we differentiate both sides we get $a\delta(ax) = \delta(x)$. Let $a < 0$, then $\theta(ax) = 1 - \theta(x)$. If we differentiate both sides we get $-a\delta(ax) = \delta(x)$. We combine these to $|a|\delta(ax) = \delta(x)$, or

  $$\delta(ax) \quad = \quad \frac{1}{|a|}\delta(x)$$

- We determine $\delta(f(x))$. Let f be a smooth function, then $\theta(f(x))' = \theta'(f(x))f'(x) = \delta(f(x))f'(x)$. Turning this around we get

  $$\delta(f(x)) \quad = \quad \frac{\theta(f(x))'}{f'(x)}$$

  The $\theta(f(x))$ term is constant 1 on intervals where $f(x) > 0$, so $\theta(f(x)) = \sum_i I_{[a_i,b_i]}(x)$ where $[a_i, b_i]$ are the intervals where $f(x) > 0$. Taking the derivative we get

  $$\theta(f(x))' \quad = \quad \sum_i \delta(x - a_i) - \delta(x - b_i)$$

  so we get

  $$\begin{aligned} \delta(f(x)) &= \frac{1}{f'(x)}\sum_i \delta(x-a_i) - \delta(x-b_i) \\ &= \sum_i \frac{1}{f'(x)}\delta(x-a_i) - \frac{1}{f'(x)}\delta(x-b_i) \\ &= \sum_i \frac{1}{f'(a_i)}\delta(x-a_i) + \frac{1}{-f'(b_i)}\delta(x-b_i) \end{aligned}$$

We have $f'(a_i) > 0$ and $f'(b_i) < 0$ because those a_i are the points where f crosses 0 upwards and b_i are the points where f crosses 0 downwards. Therefore we have

$$\delta(f(x)) = \sum_i \frac{1}{|f'(a_i)|}\delta(x-a_i) + \frac{1}{|f'(b_i)|}\delta(x-b_i)$$

wherefore we can write this as

$$\delta(f(x)) = \sum_i \frac{1}{|f'(p_i)|}\delta(x-p_i)$$

where p_i are the points where f crosses 0 in either direction.

9.2. **Currents.** Differential forms with distributions rather than smooth functions as coefficients are called currents. Currents can have mass concentrated in a point, or more generally, in a submanifold. In this section we will make precise the idea that an k-dimensional submanifold N of M can be viewed as a differential k-form ω that has all its mass concentrated on that submanifold, and such that $\int_N \alpha = \int_M \omega \wedge \alpha$ for all $n-k$ forms α.

Definition 44. A k-current is a continuous linear functional on the space of C^∞ k-forms.

Examples:

(1) Given a k-submanifold N we define a current ω given by $\omega(\alpha) = \int_N \alpha$.
(2) Given an $n-k$ form β we define a current ω given by $\omega(\alpha) = \int_M \beta \wedge \alpha$.
(3) Given a k-vector field u we define a current ω given by $\omega(\alpha) = \int_M u \vee \alpha$.
(4) Given a point $x \in M$ and a k-vector $u \in \bigwedge^k T_x M$ we define a current ω given by $\omega(\alpha) = \alpha(x)(u)$.

In each of these cases we may use the notation $\omega = N$, $\omega = \beta$, $\omega = u$ for convenience.

Definition 45. The boundary of a current ω is defined by $d\omega(\alpha) = \omega(d\alpha)$.

This definition agrees with the boundary of a submanifold. Let N be a submanifold of M, then

$$N(d\alpha) = \int_N d\alpha = \int_{\partial N} \alpha = \partial N(\alpha)$$

It also agrees with the exterior derivative of a form, up to sign. Let β be a $n-k$ form, then $d(\beta \wedge \alpha) = d\beta \wedge \alpha + (-1)^{p(n-p)}\beta \wedge d\alpha$. Therefore

$$\begin{aligned}
\beta(d\alpha) &= \int_M \beta \wedge d\alpha \\
&= (-1)^{p(n-p)}\left(\int_M d(\beta \wedge \alpha) - d\beta \wedge \alpha\right) \\
&= (-1)^{p(n-p)}\left(\int_M d(\beta \wedge \alpha) - \int_M d\beta \wedge \alpha\right) \\
&= (-1)^{p(n-p)}\left(\int_{\partial M} \beta \wedge \alpha - \int_M d\beta \wedge \alpha\right) \\
&= (-1)^{p(n-p)+1}\int_M d\beta \wedge \alpha \\
&= (-1)^{p(n-p)+1}(d\beta)(\alpha)
\end{aligned}$$

This definition satisfies all the usual rules of d on forms, e.g. $dd\omega(\alpha) = \omega(dd\alpha) = \omega(0) = 0$.

Definition 46. Given a current ω and a form β we define $\omega \wedge \beta$ as $(\omega \wedge \beta)(\alpha) = \omega(\beta \wedge \alpha)$, and similarly $(\beta \wedge \omega)(\alpha) = \omega(\alpha \wedge \beta)$.

These two definitions are coherent, because

$$\begin{aligned}(\omega \wedge \beta)(\alpha) &= \omega(\beta \wedge \alpha) \\ &= \omega((-1)^{pq}\alpha \wedge \beta) \\ &= (-1)^{pq}\omega(\alpha \wedge \beta) \\ &= (-1)^{pq}(\beta \wedge \omega)(\alpha)\end{aligned}$$

So indeed we have $\omega \wedge \beta = (-1)^{pq}\beta \wedge \omega$ as for the wedge product of two forms. Furthermore, when ω itself is a current coming from a form, then

$$\begin{aligned}(\omega \wedge \beta)(\alpha) &= \omega(\beta \wedge \alpha) \\ &= \int_M \omega \wedge (\beta \wedge \alpha) \\ &= \int_M (\omega \wedge \beta) \wedge \alpha \\ &= (\omega \wedge \beta)(\alpha)\end{aligned}$$

Where the left hand size is the wedge product of a current and a form, and the right hand side is the wedge product of two forms. This wedge product satisfies the usual rules on currents, e.g. $d(\omega \wedge \beta) = d\omega \wedge \beta + (-1)^p \omega \wedge d\beta$, because

$$\begin{aligned}d(\omega \wedge \beta)(\alpha) &= (\omega \wedge \beta)(d\alpha) \\ &= \omega(\beta \wedge d\alpha) \\ &= \omega(d(\beta \wedge \alpha) + (-1)^p d\beta \wedge \alpha) \\ &= \omega(d(\beta \wedge \alpha)) + \omega((-1)^p d\beta \wedge \alpha) \\ &= (d\omega)(\beta \wedge \alpha) + \omega((-1)^p d\beta \wedge \alpha) \\ &= (d\omega \wedge \beta)(\alpha) + (-1)^p(\omega \wedge d\beta)(\alpha)\end{aligned}$$

We use the notation $\int \omega = \omega(1)$ where 1 is the constant 1 function, when ω is an 0-current. This gives $\int_N \alpha = \int N \wedge \alpha$ for a submanifold N.

For a map $f : N \to M$ we define $f\omega$ by $f\omega(\alpha) = \omega(f^*\alpha)$. This allows us to transport a current forward from N to M.

9.3. **Poincaré lemma for currents.** The Poincaré lemma for forms states that any form a with $da = 0$ has a local antiderivative A with $dA = a$. We wish to extend this lemma to currents.

Lemma 47. *Let ω be a current with $d\omega = 0$ on a domain U for which the ordinary Poincaré lemma has been proven, i.e. $d\omega(\alpha) = \omega(d\alpha) = 0$ for all α with support in U. Then there exists a current Ω such that $d\Omega = \omega$ on U.*

Proof. We are looking for a current Ω with $d\Omega = \omega$ locally, i.e. $d\Omega(\alpha) = \omega(\alpha)$ with α having support in U. By the rule for differentiation we also have $d\Omega(\alpha) = \Omega(d\alpha)$, so we need $\Omega(d\alpha) = \omega(\alpha)$. By the proof of the ordinary Poincaré lemma for U, there exists a linear antiderivative map L with $dL(\beta) = \beta$ for all β with support on U and $d\beta = 0$. Take $\Omega(\beta) = \omega(L\beta)$. Then $d\Omega(\beta) = \omega(Ld\beta) = \omega(\beta)$ because $dLd\beta = d\beta$ so $Ld\beta$ and β differ by an exact form, and ω is zero on exact forms. □

10. Linking number

10.1. **Winding number.** Winding number of a curve γ around 0 is n,

$$\int_\gamma \alpha \quad = \quad 2\pi n$$

where

$$\alpha \quad = \quad \frac{xdy - ydx}{x^2 + y^2}$$

If we integrate this form around a closed loop we always get a multiple of 2π. Currents can explain this intuitively: $d\alpha = 2\pi\delta_0 dx \wedge dy$. We see that $d\alpha$ is 2π times the Dirac delta at 0. If $\gamma = \partial A$, then $\int_\gamma a = \int_{dA} \alpha = \int_A d\alpha = \int_A 2\pi\delta_0 dx \wedge dy = 2\pi n$. Intuitively, a loop that winds around the origin multiple times is the boundary of an area that has an integer weight.

10.2. **Intersection Theory.** We can define the oriented intersection of transversal submanifolds N, M. This corresponds to the wedge product of their currents, if that is well defined (this requires transversality). We can define the intersection number $I(N, M)$ of two submanifolds of dimention $k+l = N$, and this corresponds to $\int N \wedge M$. We immediately get the whole theory for free, e.g. $I(M, N) = (-1)^{pq} I(N, M)$, $I(\partial M, N) = (-1)^p I(M, \partial N)$, etc.

10.3. **Linking Number.**

Definition 48. The linking number $L(a, b)$ of two closed 1-vectors in $\mathbb{R}^3$ is defined as $\int_{\mathbb{R}^3} A \wedge b$ where $dA = a$.

This definition does not depend on the choice of A, for let $dB = b$, then $\int A \wedge b = \int A \wedge dB = \int dA \wedge B = \int a \wedge B$. Hence we also see that the definition is symmetric: $L(a, b) = L(b, a)$. Geometrically the linking number can be interpreted as follows: take a disk A whose boundary is a, and a disk B whose boundary is b, then the linking number is the intersection number $I(A, b) = I(a, B)$.

If the 1-vectors are 1-submanifolds then the linking number is also the degree of the map $\Gamma : \mathbb{T}^2 \to \mathbb{S}^2$ given by

$$\Gamma(t_1, t_2) \quad = \quad \frac{a(t_1) - b(t_2)}{||a(t_1) - b(t_2)||}$$

Proof. Let γ_1, γ_2 be closed 1-submanifolds of $\mathbb{R}^3$. By a homotopy we may assume that γ_1 lies in the xy-plane. To calculate the degree of Γ we may pick a form $\omega \in \bigwedge^2 \mathbb{S}^2$ with $\int_{\mathbb{S}^2} \omega = 1$, then $\deg \Gamma = \int_{\mathbb{T}^2} \Gamma^* \omega$. We pick ω concentrated on the xy-plane. Then the $\int_{\mathbb{T}^2} \Gamma^* \omega$ integral only picks up points where $\gamma_1(t_1) - \gamma_2(t_1)$ lies in the xy-plane. Since γ_1 already lies in the xy-plane, these are points where γ_2 lies in the xy-plane as well, and there the integral becomes $\sum_{t_2 \in T_2} \int_{\mathbb{S}^1} \frac{\gamma_1(t_1) - \gamma_2(t_2)}{||\gamma_1(t_1) - \gamma_2(t_2)||} dt_1$ where T_2 are all the t_2 such that $\gamma_2(t_2)$ lies in the xy-plane. The inner integral is the winding number of γ_1 around $\gamma_2(t_2)$, which is equal to the intersection number of a point with the surface A with $\partial A = \gamma_1$. □

10.4. **Sketch of the proof of Gauss' linking integral.** Currents give us a short proof of Gauss' linking integral. Define

$$G_x(y) = \frac{x-y}{\Omega_n |x-y|^n}$$

Where Ω_n is the surface of an n-ball. The boundary of the vector field $dG_x = \delta_x$ where δ_x is the diagonal Dirac delta current given by $\delta_x(f) = f(x)$. We also have $A(x) = \int \delta_x \wedge A = \int_{y\in\mathbb{R}^n} \delta_x(y) \wedge A(y)$.

$$\begin{aligned}
L(a,b) &= \int_{\mathbb{R}^n} A \wedge b \\
&= \int_{x\in\mathbb{R}^n} A(x) \wedge b(x) \\
&= \int_{x\in\mathbb{R}^n} \int_{y\in\mathbb{R}^n} \delta_x(y) \wedge A(y) \wedge b(x) \\
&= \int_{x\in\mathbb{R}^n} \int_{y\in\mathbb{R}^n} \delta_y(x) \wedge A(y) \wedge b(x) \\
&= \int_{x\in\mathbb{R}^n} \int_{y\in\mathbb{R}^n} d_x G_y(x) \wedge A(y) \wedge b(x) \\
&= \int_{x\in\mathbb{R}^n} \int_{y\in\mathbb{R}^n} d_x G_y(x) \wedge A(y) \wedge b(x) \\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} G \wedge d(\bar{A} \wedge \bar{b}) \\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} G \wedge (d\bar{A} \wedge \bar{b} + (-1)^p \bar{A} \wedge d\bar{b}) \\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} G \wedge d\bar{A} \wedge \bar{b} \\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} G \wedge \bar{a} \wedge \bar{b} \\
&= \int_{a\times b} G
\end{aligned}$$

The exchange of derivative and integral can be formally justified using the dominated convergence theorem.

10.5. **Index of a vector field.** Given a vector field $v : \mathbb{R}^n \to \mathbb{R}^n$ we can define $w : \mathbb{R}^n \to \mathbb{S}^{n-1}$ as $w(x) = \frac{x}{||x||}$. Let ω be the volume form of $\mathbb{S}^{n-1}$, then we can define $\alpha = w^*\omega \in \bigwedge^{n-1}$. We define the index in an area M as

$$Index(v, M) \quad = \quad \int_M d\alpha$$

By Stokes' theorem we have $Index(v, M) = \int_{\partial M} \alpha$. This fact requires an elaborate proof without currents. The form α has singularities at $v(x) = 0$ and this causes trouble for $d\alpha$ but not with currents. Note that if B is a small sphere, then $\deg w|_B = Index(v, B)$.

11. Electrodynamics

11.1. **Electrostatics.** In electrostatics we have a charge distribution $\rho : \mathbb{R}^3 \to \mathbb{R}$ and wish to find the associated electric field $E : \mathbb{R}^3 \to \mathbb{R}^3$. With differential forms the charge distribution becomes a 3-form and the electric field becomes a 2-form. The equations are

$$dE = \rho$$
$$d * E = 0$$

Because $d * E = 0$ we can find a potential V for $*E$ that satisfies $dV = *E$. Using this potential the first equation becomes

$$d * dV = \rho$$

and the second equation is satified automatically because $d^2 = 0$.

The geometric interpretation of the differential form E is the same as the geometric interpretation of the usual vector field E: a field of curved lines that go from negative charges to positive charges, that is, the boundary of the lines is the charge density: $dE = \rho$. The equation $d * E = 0$ means that the planes orthogonal to the lines have no net boundary, that is, they close back onto themselves. For instance, for the field of a single electron the planes orthogonal to the electric field lines form spherical shells. These spherical shells are the surfaces of constant electric potential: $dV = *E$.

11.2. **Magnetostatics.** Magnetostatics is similar except we have a current density $I : \mathbb{R}^3 \to \mathbb{R}^3$ and the magnetic field $B : \mathbb{R}^3 \to \mathbb{R}^3$. With differential forms the current density becomes a 2-form and the magnetic field becomes a 1-form. The equations are

$$dB = I$$
$$d * B = 0$$

Because $d * B = 0$ we can find a vector potential M for $*B$ that satisfies $dM = *B$. Using this potential the first equation becomes

$$d * dM = I$$

and the second equation is satified automatically because $d^2 = 0$.

The geometric interpretation of the differential form B is a bit different than the magnetic field B, because the form B is a 1-form, and thus represents a density of surfaces. These surfaces are orthogonal to the conventional magnetic field: the conventional magnetic field is $*B$. The reason we define B as a one form rather than as the two form $*B$ is to put the equations in the same form as for electrostatics: $dB = I$ is analogous to $dE = \rho$, except that now we have surfaces B whose boundary is the current I, which is a line density. Another reason is that the equation does not depend on a right hand rule when written in this form: going from the right hand rule to the left hand rule corresponds to taking a minus sign whenever the Hodge star is applied, but $d * B = 0$ remains the same with or without a minus sign, so the equations are manifestly not dependent on handedness. With vector calculus we use the cross product to define B, which depends on handedness, but when computing the Lorentz force we use another cross product. If we change handedness the direction of the magnetic field reverses, but the sign of the magnetic part of the Lorentz force reverses as well, so the physics does not depend on handedness.

The equation $d*B = 0$ is interpreted as saying that the boundary of the conventional magnetic field is zero: the magnetic field lines close back onto themselves. The magnetic potential M is a surface density whose boundary is the conventional magnetic field $*B$, which is possible to find because $*B$ has no boundary. It is intuitively clear that there are many ways to choose a surface density such that its boundary is a given magnetic field $*B$, in the same way that there are many surfaces whose boundary is a given loop. For the electric field the freedom to choose V was much more limited: any two potentials differ only by a scalar. Any two magnetic potentials differ by a form N with $dN = 0$, because then and only then is $d * d(M + N) = d * dM = I$ still satisfied. Geometrically, we may add any boundaryless surface density N to the magnetic potential M. This has the effect of shifting the surfaces in M without shifting their boundaries.

11.3. **Maxwell's equations.** In Minkowski spacetime we can form the 4 dimensional current density

$$J = \rho + I \wedge dt$$

and the Maxwell tensor

$$F = E + B \wedge dt$$

Maxwell's equations are

$$dF = J$$
$$d * F = 0$$

Note that the Hodge star is with respect to the Minkowski metric. Because $d*F = 0$ we can find a 4-potential A for $*F$ that satisfies $dA = *F$. Using this potential the first equation becomes

$$d * dA = J$$

and the second equation is satified automatically because $d^2 = 0$.

This shows another advantage of using B rather than $*B$: we can easily combine $E + B \wedge dt$ into a single 2-form whose boundary is the spacetime current density J. Geometrically, the spacetime current density J may be seen as the world lines of charges through spacetime. The orientation of the charges depends on whether they are a positive charge or a negative charge. A negative charge is the same as a positive charge traveling backward in time. A loop in spacetime may be sliced in slices of constant t, and this represents the creation of a positive+negative charge pair that travel away from each other and then travel back together and annihilate. The boundary of the Maxwell tensor F is that current density J, and as with electrostatics and magnetostatics, its coboundary is zero.

It is conventional to use $*F$ in place of F, but to retain the symmetry with electrostatics and magnetostatics we don't do this, and thus have $dF = J$ rather than $d*F = J$. Since F is a two form in 4d, so is $*F$, and both represent a 2-surface density in 4d. The spacetime potential A represents a 3-surface in 4d.

If we take the pullback of F to a 3d slice of constant t we get the electric field. That is, the intersection of the 2-surface density$F$ with a surface of constant t gives the line density E.

11.4. **The Lorentz force.** The Maxwell tensor bilinear antisymmetric. We may turn this into a linear endomorphism using the metric isomorphism with the dual space. This linear map is skew symmetric with respect to the Minkowski inner product, so it is an element of the Lie algebra of the Lorentz group. In other words, the Maxwell tensor is an infinitessimal rotation or Lorentz boost at each point. The Lorentz force on a particle with 4-velocity v is $F \cdot v$. This force tends to rotate or boost the particle in the direction of the infinitessimal rotation or boost given by F. Geometrically, F is a 2-surface in $\mathbb{R}^4$, and at a point p the rotation/boost $F(p)$ happens in the direction of the surface $F(p)$. For instance, if the surface at $F(p)$ lies in the xy-plane, then it is a magnetic field at p that tends to rotate the trajectory of a particle in the xy plane (a moving electron in a uniform magnetic field goes in circles). If the surface $F(p)$ lies in the xt-plane, then it is an electric field that tends to boost (i.e. accelerate) the particle in the x direction (an electron in a uniform electric field is uniformly accelerated).

11.5. **The Lagrangian of a charged particle.** The relativistic Lagrangian of a charged particle traveling with trajectory $x(\tau) \in \mathbb{R}^4$ is $L(\tau) = |x'(\tau)| + A(x(\tau))(x'(\tau))$ where A is the four potential and the norm is with respect to the Minkowski metric. The action is the integral of the Lagrangian over the trajectory:

$$\begin{aligned} S[x] &= \int |x'(\tau)| + A(x(\tau))(x'(\tau))d\tau \\ &= \int |x'(\tau)|d\tau + \int A(x(\tau))(x'(\tau))d\tau \\ &= |x| + \int A \wedge x \end{aligned}$$

That is, the action is the spacetime length of the trajectory plus the intersection value of the four potential with the trajectory. The principle of least action states that the trajectory that a particle takes extremises the action. We can derive the Lorentz force law from the Lagrangian using the Euler-Lagrange equations.

12. Computing the Hodge decomposition

Let a be an n-form on M. Poincaré's lemma tells us that we can always find a local antiderivative b with $db = a$ on some open set around a point. The topology of the manifold may prevent us from finding a global antiderivative. The same applies to the coderivative: we can always find a local c with $\delta c = a$, but we cannot always find a global one. The db and δc are orthogonal with respect to the inner product on forms induced by the metric because d and δ are adjoint:

$$\langle db, \delta c \rangle = \left\langle d^2 b, c \right\rangle = \langle 0, c \rangle = 0$$

Or equivalently

$$\langle db, \delta c \rangle = \left\langle b, \delta^2 c \right\rangle = \langle b, 0 \rangle = 0$$

This shows that the component of a for which we can find an antiderivative with respect to d is orthogonal to the component of a for which we can find an antiderivative with respect to δ, which suggests that perhaps we can split $a = a_1 + a_2$ such that $a_1 = db$ has an antiderivative with respect to b and $a_2 = \delta c$ has an antiderivative with respect to δ. It turns out that this is not enough, even for the 1 dimensional case. Consider the 1-form $d\theta$ on the circle. This form does not have an

antiderivative despite the suggestive notation $d\theta$, because θ is a multi-valued function. The form $d\theta$ is well defined because the derivative coincides for all branches. Intuitively we view $d\theta$ as a force field going around the circle, and then we'd like to find a potential for this force field. A single-valued potential does not exist, but the form $d\theta$ is harmonic, that is, it's derivative and coderivative are both zero. This is the final component of the decomposition of a:

Theorem 49. *Hodge decomposition.*

Given a k-form a on a manifold M we can decompose a into three orthogonal components $a = a_1 + a_2 + a_3$ such that $a_1 = db$, $a_2 = \delta c$ and a_3 is harmonic ($da_3 = 0$ and $\delta a_3 = 0$).

Corollary 50. *Helmholz decomposition.*

Given a vector field F on $\mathbb{R}^3$ we may find a potential $\phi : \mathbb{R}^3 \to \mathbb{R}$ and a vector potential $A : \phi : \mathbb{R}^3 \to \mathbb{R}^3$ such that $F = \nabla\phi + \nabla \times A$.

Proof. This is the special case of the Hodge decomposition for $M = \mathbb{R}^3$ and $k = 1$. The potential ϕ is a 0-form and the gradient ∇ corresponds to d on 0-forms. The vector potential A is a 2-form and the curl $\nabla\times$ corresponds to the codifferential on 2-forms. We do not have a harmonic component because of the topology of $\mathbb{R}^3$: the 1 cohomology of $\mathbb{R}^3$ is trivial hence the only harmonic 1-form on $\mathbb{R}^3$ is zero. □

That the harmonic component a_3 is orthogonal to a_1 and a_2 follows from similar reasoning as why a_1 and a_2 are orthogonal:

$$\langle a_1, a_3\rangle = \langle db, a_3\rangle = \langle b, \delta a_3\rangle = \langle b, 0\rangle = 0$$
$$\langle a_2, a_3\rangle = \langle \delta c, a_3\rangle = \langle c, da_3\rangle = \langle c, 0\rangle = 0$$

Where we've used that a harmonic form satisfies $da_3 = 0$ and $\delta a_3 = 0$.

There is another way to view the decomposition, and that is via the decomposition from linear algebra that is sometimes called the fundamental theorem of linear algebra: given an operator A from V to W we can decompose $V = \ker A + im\, A^T$ or if A is an operator from W to V we can decompose $V = im\, A + \ker A^T$. In our case we have two operators $d_k : \bigwedge^k M \to \bigwedge^{k+1} M$ and $\delta_k : \bigwedge^k M \to \bigwedge^{k-1} M$. If the spaces were finite dimensional we could apply the decomposition twice, once to d and once to δ. We get a decomposition of $\bigwedge^k$ into four components:

$$\ker d_k \cap \ker \delta_k, \ker d_k \cap im\, \delta_k^T$$
$$im\, d_k^T \cap \ker \delta_k, im\, d_k^T \cap im\, \delta_k^T$$

In our case $d_n^T = \delta_{n+1}$ and $\delta_n^T = d_{n-1}$, so the subspaces are

$$\ker d_k \cap \ker \delta_k, \ker d_k \cap im\, d_{k-1}$$
$$im\, \delta_{k+1} \cap \ker \delta_k, im\, \delta_{k+1} \cap im\, d_{k-1}$$

The space $\ker d_k \cap \ker \delta_k$ are the harmonic forms. The space $\ker d_k \cap im\, d_{k-1}$ is associated with $a_1 = db$ and the space $im\, \delta_{k+1} \cap \ker \delta_k$ is associated with $a_2 = \delta c$. The space $im\, \delta_{k+1} \cap im\, d_{k-1}$ is zero, since a form a in this space can be written both as $a = db$ and $a = \delta c$. Appling d and δ to a we find $da = 0$ and $\delta a = 0$, so a is harmonic, so $im\, \delta_{k+1} \cap im\, d_{k-1} \subset \ker d_k \cap \ker \delta_k$. However, by construction all these spaces are orthogonal, so we find $im\, \delta_{k+1} \cap im\, d_{k-1} = \{0\}$.

This argument is not rigorous because we are working in a Hilbert space not a finite dimensional space, but perhaps this argument can be made rigorous. Given

a Hilbert space H and a subspace X we may decompose H into X and X^T if X is closed. So if we could show that the kernel of d_k and δ_k are closed, we are done.

12.1. **Computing the Hodge decomposition numerically.** To compute the Hodge decomposition $a = a_1 + a_2 + a_3 = db + \delta c + a_3$ numerically we use

$$b = argmin_b \, ||a - db||^2$$
$$c = argmin_c \, ||a - \delta c||^2$$

This works because the spaces are orthogonal: we may extract the component in the image of d by minimising the remaining norm $||a - db||^2$, and similarly for δ: we may extract the componenet in the image of δ by minimising the remaining norm $||a - \delta c||^2$. We rewrite these norms using the inner product:

$$\begin{aligned} ||a - db||^2 &= \langle a - db, a - db \rangle \\ &= \langle a, a \rangle + \langle a, -db \rangle + \langle -db, a \rangle + \langle -db, -db \rangle \\ &= \langle a, a \rangle - 2 \langle a, db \rangle + \langle db, db \rangle \\ &= \langle a, a \rangle - 2 \langle a, db \rangle + \langle b, \delta db \rangle \end{aligned}$$

The constant term $\langle a, a \rangle$ does not matter for the minimisation so we may ignore it. To minimise the remainder $\langle b, \delta db \rangle - 2 \langle a, db \rangle$ we pick a direction ν and insist that the derivative in this direction is zero:

$$\begin{aligned} &\frac{\partial}{\partial \epsilon} [\langle b + \epsilon \nu, \delta d(b + \epsilon \nu) \rangle - 2 \langle a, d(b + \epsilon \nu) \rangle] \\ =&\frac{\partial}{\partial \epsilon} [\langle b + \epsilon \nu, \delta d(b + \epsilon \nu) \rangle - 2 \langle a, d(b + \epsilon \nu) \rangle] \\ =&\frac{\partial}{\partial \epsilon} [\langle b, \delta db \rangle + \epsilon \langle \nu, \delta db \rangle + \epsilon \langle b, \delta d\nu \rangle + \epsilon^2 \langle \nu, \delta d\nu \rangle - 2 \langle a, db \rangle - 2\epsilon \langle a, d\nu \rangle] \\ =& \langle \nu, \delta db \rangle + \langle b, \delta d\nu \rangle + 2\epsilon \langle \nu, \delta d\nu \rangle - 2 \langle a, d\nu \rangle \end{aligned}$$

We now set $\epsilon = 0$ to find the equation

$$\langle \nu, \delta db \rangle + \langle b, \delta d\nu \rangle - 2 \langle a, d\nu \rangle = 0$$

We may rewrite $\langle \nu, \delta db \rangle = \langle d\nu, db \rangle = \langle db, d\nu \rangle = \langle b, \delta d\nu \rangle$ so the equation is $2 \langle \nu, \delta db \rangle - 2 \langle a, d\nu \rangle = 0$. We may also rewrite $\langle a, d\nu \rangle = \langle \delta a, \nu \rangle = \langle \nu, \delta a \rangle$. We conclude that for all directions ν we have

$$\langle \nu, \delta db \rangle = \langle \nu, \delta a \rangle$$

at the minimum. Since equality holds for all ν we have $\delta db = \delta a$. In other words, if we define the operator $A = \delta b$ and the right hand side $f = \delta a$ we have to solve the equation $Ax = f$ and our solution $b = x$. Since we cannot solve infinite dimensional problems numerically we will instead pick a finite dimensional subspace spanned by basis vectors ϕ_i. Then we wish to find $b = \sum_{i=1}^{n} x_i \phi_i$ in that subspace with coefficients $x_i \in \mathbb{R}$. We cannot satisfy an infinite dimensional system with only a finite set of parameters a_i so we do not require that $\langle \nu, \delta db \rangle = \langle \nu, \delta a \rangle$ for all ν, but only for those ν in the same subspace. So we get n equations for $\nu = \phi_j$:

$$\langle \phi_j, \delta db \rangle = \langle \phi_j, \delta a \rangle$$

Substitute $b = \sum_{i=1}^n x_i \phi_i$, then by linearity

$$\left\langle \phi_j, \delta d(\sum_{i=1}^n x_i \phi_i) \right\rangle = \sum_{i=1}^n x_i \langle \phi_j, \delta d \phi_i \rangle$$

So the equations are

$$\sum_{i=1}^n x_i \langle \phi_j, \delta d \phi_i \rangle = \langle \phi_j, \delta a \rangle$$

Define the vector $f_j = \langle \phi_j, \delta a \rangle$ and matrix $A_{ij} = \sum_{i=1}^n x_j \langle \phi_i, \delta d \phi_i \rangle$, then the equations are

$$\sum_{i=1}^n x_i A_{ij} = f_j$$

for $j \in \{1, \dots, n\}$, or simply

$$Ax = f$$

It remains to choose appropriate basis functions ϕ_i and to compute $A_{ij} = \sum_{i=1}^n x_j \langle \phi_i, \delta d \phi_i \rangle$ and $f_j = \langle \phi_j, \delta a \rangle$. The same can be applied for c instead of b if we flip d and δ in the derivation above.

12.2. **Explicit calculation for the flat 2-torus.** Suppose $M = \mathbb{T}^2$ with the flat metric. If we are interested in decomposing an 1-form a into $db + \delta c + a_3$ where b is a 0-form and c is a 2-form. For b need to calculate $A_{ij} = \sum_{i=1}^n x_j \langle \phi_i, \delta d \phi_i \rangle$ and $f_j = \langle \phi_j, \delta a \rangle$ where ϕ_i are 0-forms, i.e. $\mathbb{R}$-valued functions on $\mathbb{T}^2$.

12.2.1. *Calculation of A_{ij}.* We can rewrite $\langle \phi_j, \delta d \phi_i \rangle = \langle d\phi_j, d\phi_i \rangle$, and d is the gradient and the inner product is the standard inner product in the flat metric:

$$\langle \phi_j, \delta d \phi_i \rangle = \int_{\mathbb{T}^2} [\frac{\partial \phi_i}{\partial x} \cdot \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \cdot \frac{\partial \phi_j}{\partial y}] dx \wedge dy$$

where x and y are the standard coordinates. We subdivide $\mathbb{T}^2$ into triangles Δ_i and choose ϕ_i linear on each triangle: $\phi_i = \alpha_i x + \beta_i y + \gamma_i$, so $\frac{\partial \phi_i}{\partial x} \cdot \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \cdot \frac{\partial \phi_j}{\partial y} = \alpha_i \alpha_j + \beta_i \beta_j$. So on a triangle Δ_k we have

$$\int_{\Delta_k} [\frac{\partial \phi_i}{\partial x} \cdot \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \cdot \frac{\partial \phi_j}{\partial y}] dx \wedge dy = |\Delta_k| (\alpha_i \alpha_j + \beta_i \beta_j)$$

where $|\Delta_k|$ is the area of the triangle. In total we have

$$\begin{aligned} A_{ij} &= \langle \phi_i, \delta d \phi_i \rangle \\ &= \int_{\mathbb{T}^2} [\frac{\partial \phi_i}{\partial x} \cdot \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \cdot \frac{\partial \phi_j}{\partial y}] dx \wedge dy \\ &= \sum_{k \in S_{ij}} \int_{\Delta_k} [\frac{\partial \phi_i}{\partial x} \cdot \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \cdot \frac{\partial \phi_j}{\partial y}] dx \wedge dy \\ &= \sum_{k \in S_{ij}} |\Delta_k| (\alpha_i \alpha_j + \beta_i \beta_j) \end{aligned}$$

Where the sum $k \in S_{ij}$ runs over the triangles for which ϕ_i and ϕ_j have support (or equivalently we could choose α_i and β_i to be zero on those triangles not in the support). In our case we choose ϕ_i such that on exactly one grid point $\phi_i(x_p) = 0$,

and $\phi_i(x_{p'}) = 0$ for all other grid points, and linearly interpolated in the interior of each triangle. So if x_p are the grid points then $\phi_i(x_p) = \delta_{ip}$ where δ_{ip} is the Kronecker delta.

12.2.2. *Calculation of f_j.* If $a = gdx + hdy$ then $\delta a = \frac{\partial g}{\partial y} - \frac{\partial h}{\partial x}$, a 0-form. The inner product is

$$\begin{aligned} f_j &= \langle \phi_j, \delta a \rangle \\ &= \int_{\mathbb{T}^2} \phi_j (\frac{\partial g}{\partial y} - \frac{\partial h}{\partial x}) dx \wedge dy \\ &= \sum_k \int_{\Delta_k} \phi_j (\frac{\partial g}{\partial y} - \frac{\partial h}{\partial x}) dx \wedge dy \end{aligned}$$

We approximate the integral using the Newton-Cotes rule

$$\begin{aligned} \int_{\Delta_k} \phi_j (\frac{\partial g}{\partial y} - \frac{\partial h}{\partial x}) dx \wedge dy &\approx |\Delta_k| \sum_p \phi_j(x_p) (\frac{\partial g}{\partial y}(x_p) - \frac{\partial h}{\partial x}(x_p)) \\ &= |\Delta_k| \sum_p \delta_{jp} (\frac{\partial g}{\partial y}(x_p) - \frac{\partial h}{\partial x}(x_p)) \\ &= |\Delta_k| (\frac{\partial g}{\partial y}(x_j) - \frac{\partial h}{\partial x}(x_j)) \end{aligned}$$

where x_p are the corners of the triangle Δ_k and δ_{jp} is the Kronecker delta. So we may approximate

$$f_j \approx \sum_k |\Delta_k| (\frac{\partial g}{\partial y}(x_j) - \frac{\partial h}{\partial x}(x_j))$$

12.3. **Explicit calculation for the general case.** Suppose M is an N-dimensional manifold with arbitrary metric, and we are interested in decomposing a k-form a into $db + \delta c + a_3$. For b need to calculate $A_{ij} = \sum_{i=1}^n x_j \langle \phi_i, \delta d\phi_i \rangle$ and $f_j = \langle \phi_j, \delta a \rangle$ where ϕ_i are $(k-1)$-forms on M.

12.3.1. *Calculation of A_{ij}.* We can rewrite $\langle \phi_j, \delta d\phi_i \rangle = \langle d\phi_j, d\phi_i \rangle$, and the inner product is the standard inner product induced by the metric:

$$\langle \phi_j, \delta d\phi_i \rangle = \int_M d\phi_j \wedge \star d\phi_i$$

where x and y are the standard coordinates. We subdivide M into simplices Δ_i and choose ϕ_i linear on each simplex: $\phi_i = \sum_I \sum_{n=0}^N \alpha_n^I x_n dx^I + \sum_I \gamma_i^I dx^I$ (note

that this depends on the coordinate system!), so

$$\begin{aligned}
&(\sum_I \sum_{n=0}^N \alpha_n^I x_n dx^I + \sum_I \gamma_i^I dx^I) \wedge \star(\sum_I \sum_{n=0}^N \alpha_n^I x_n dx^I + \sum_I \gamma_i^I dx^I) \\
=&(\sum_I \sum_{n=0}^N \alpha_n^I x_n dx^I) \wedge \star(\sum_I \sum_{n=0}^N \alpha_n^I x_n dx^I) + \text{lower order} \\
=&\sum_{I,I'} \sum_{n,n'=0}^N \alpha_n^I x_n dx^I \wedge \star \alpha_{n'}^{I'} x_{n'} dx^{I'} \\
=&\sum_{I,I'} (\sum_{n,n'=0}^N \alpha_n^I x_n \alpha_{n'}^{I'} x_{n'}) dx^I \wedge \star dx^{I'}
\end{aligned}$$

. On a simplex Δ_k we have

$$\int_{\Delta_k} [\frac{\partial \phi_i}{\partial x} \cdot \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \cdot \frac{\partial \phi_j}{\partial y}] dx \wedge dy = |\Delta_k|(\alpha_i \alpha_j + \beta_i \beta_j)$$

where $|\Delta_k|$ is the volume of the simplex. In total we have

$$\begin{aligned}
A_{ij} &= \langle \phi_i, \delta d\phi_i \rangle \\
&= \int_M [\frac{\partial \phi_i}{\partial x} \cdot \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \cdot \frac{\partial \phi_j}{\partial y}] dx \wedge dy \\
&= \sum_{k \in S_{ij}} \int_{\Delta_k} [\frac{\partial \phi_i}{\partial x} \cdot \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \cdot \frac{\partial \phi_j}{\partial y}] dx \wedge dy \\
&= \sum_{k \in S_{ij}} |\Delta_k|(\alpha_i \alpha_j + \beta_i \beta_j)
\end{aligned}$$

Where the sum $k \in S_{ij}$ runs over the triangles for which ϕ_i and ϕ_j have support (or equivalently we could choose α_i and β_i to be zero on those triangles not in the support). In our case we choose ϕ_i such that on exactly one grid point $\phi_i(x_p) = 0$, and $\phi_i(x_{p'}) = 0$ for all other grid points, and linearly interpolated in the interior of each triangle. So if x_p are the grid points then $\phi_i(x_p) = \delta_{ip}$ where δ_{ip} is the Kronecker delta.

12.3.2. *Calculation of f_j.* If $a = gdx + hdy$ then $\delta a = \frac{\partial g}{\partial y} - \frac{\partial h}{\partial x}$, a 0-form. The inner product is

$$\begin{aligned}
f_j &= \langle \phi_j, \delta a \rangle \\
&= \int_M \phi_j (\frac{\partial g}{\partial y} - \frac{\partial h}{\partial x}) dx \wedge dy \\
&= \sum_k \int_{\Delta_k} \phi_j (\frac{\partial g}{\partial y} - \frac{\partial h}{\partial x}) dx \wedge dy
\end{aligned}$$

We approximate the integral using the Newton-Cotes rule

$$\begin{aligned}\int_{\Delta_k} \phi_j(\frac{\partial g}{\partial y} - \frac{\partial h}{\partial x})dx \wedge dy &\approx |\Delta_k| \sum_p \phi_j(x_p)(\frac{\partial g}{\partial y}(x_p) - \frac{\partial h}{\partial x}(x_p)) \\ &= |\Delta_k| \sum_p \delta_{jp}(\frac{\partial g}{\partial y}(x_p) - \frac{\partial h}{\partial x}(x_p)) \\ &= |\Delta_k|(\frac{\partial g}{\partial y}(x_j) - \frac{\partial h}{\partial x}(x_j))\end{aligned}$$

where x_p are the corners of the triangle Δ_k and δ_{jp} is the Kronecker delta. So we may approximate

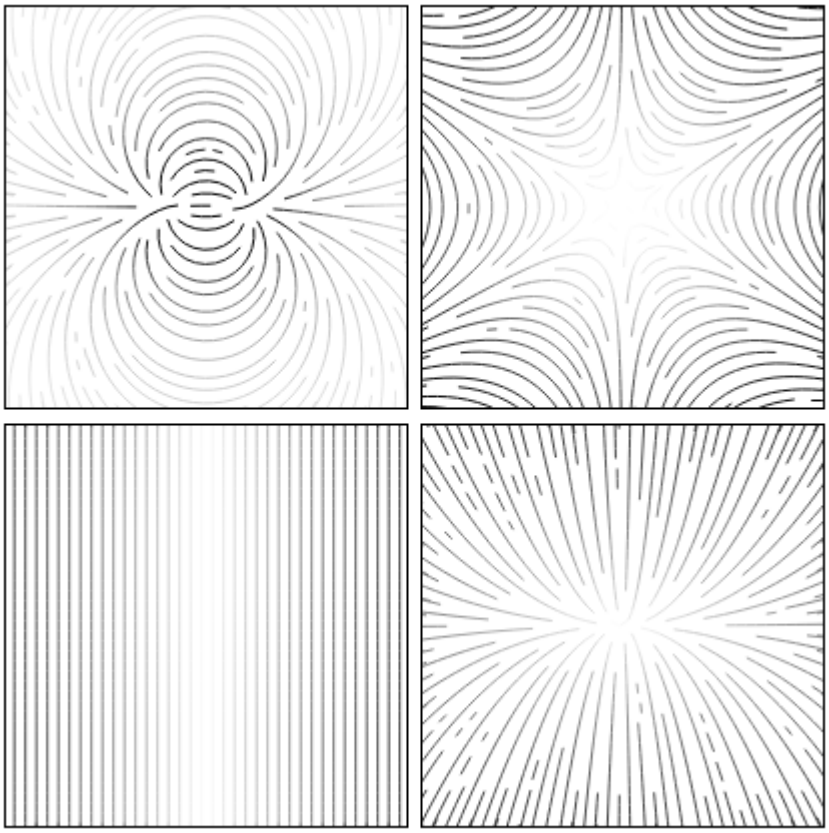
$$f_j \approx \sum_k |\Delta_k|(\frac{\partial g}{\partial y}(x_j) - \frac{\partial h}{\partial x}(x_j))$$

13. Drawing 1-forms on $\mathbb{R}^2$

1-forms on $\mathbb{R}^2$ are geometrically interpreted as a density of lines. In order to draw this we first draw a field of lines of rougly constant density, and then change the opacity of the lines to match the density of the form. The algorithm to draw 1-forms α works as follows:

(1) Pick a spacing distance $\epsilon > 0$ between the lines.
(2) Pick a random starting point and put it on a stack.
(3) Repeat these steps until the stack is empty:
    (a) Pick the point on the top of the stack.
    (b) Follow the line forward and backward using a numerical ODE method.
    (c) Draw the lines connecting the points with opacity dictated by the value of $|\alpha(x)|$.
    (d) For each point x that is computed by the ODE method, put $x + \epsilon v$ and $x - \epsilon v$ on the stack, where v is a unit normal to the current line at point x.

This method produces these results:



14. Conclusion

In this thesis we have described a method for visualising differential forms and given a geometric interpretation of the operations on differential forms such as the exterior derivative, the integral, the wedge product, the Hodge star, and the pull-back. In order to make this more precise we have given an alternative treatment of alternating linear maps that makes the connection to the geometric interpretation more concrete. We have also seen how this geometric interpretation can be applied to electromagnetism, thermodynamics, classical mechanics, linking numbers, and the Hodge decomposition. We've also described a method for numerically computing the Hodge decomposition and a method for drawing 1-forms on $\mathbb{R}^2$.

References

[1] H. Cartan. *Differential forms*. Dover Books on Mathematics Series. Dover Publications, 2006.

[2] S.S. Chern, F.R. Smith, and G. de Rham. *Differentiable Manifolds: Forms, Currents, Harmonic Forms*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2012.

[3] J. Gratus. A pictorial introduction to differential geometry, leading to Maxwell's equations as three pictures. *ArXiv e-prints*, September 2017.

[4] Victor J. Katz. The history of differential forms from clairaut to poincarÃ©. *Historia Mathematica*, 8(2):161 – 188, 1981.

[5] Peter D. Lax. Change of variables in multiple integrals ii. *The American Mathematical Monthly*, 108(2):115–119, 2001.
[6] C.W. Misner, K.S. Thorne, and J.A. Wheeler. *Gravitation*. Number pt. 3 in Gravitation. W. H. Freeman, 1973.
[7] R.S. Strichartz. *A Guide to Distribution Theory and Fourier Transforms*. World Scientific, 2003.