TUTORIAL

Differential forms for plasma physics

R. S. MacKay†

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

Differential forms provide a coordinate-free way to express many quantities and relations in mathematical physics. In particular, they are useful in plasma physics. This tutorial gives a guide so that you can read the plasma physics literature that uses them and apply them yourself.

Key words: plasma confinement

---

1. Introduction

Differential forms allow one to go beyond what vector calculus can cope with. For example, they permit extension of the operator $B \cdot \nabla$ for a magnetic field B in a principled way to more general objects than scalar functions, powerful generalisations of Gauss' and Stokes' theorems, and generalisations of three-dimensional (3-D) results like that a curl-free vector field is locally a gradient and a divergence-free one is locally a curl. These extensions and generalisations are particularly useful in contexts like Hamiltonian dynamics, where the state space is usually not three-dimensional.

One can write everything about differential forms in index notation (as in Misner, Thorne & Wheeler (1973)) and index notation generalises to arbitrary tensors (differential forms are only the antisymmetric covariant tensors). But the concepts for differential forms are clearer without the reference to coordinate systems that index notation implies. In particular one can see which relations involve an assumed volume form or an assumed Riemannian metric and which are independent of these.

Differential forms were developed into a complete theory by Cartan in 1899. Although they have been used to good effect in electromagnetics (see Warnick & Russer (2014) for a recent review, including many pointers to the literature, of which I find Deschamps (1981) particularly good) and in plasma physics (e.g. Cary & Littlejohn (1983), who gave a tutorial appendix), many plasma physicists are unfamiliar with them. The aim of this tutorial is to make them accessible to plasma physicists.

We give a definition of differential forms on a general oriented smooth manifold M. Loosely speaking, a manifold is a topological space such that every point has a neighbourhood homeomorphic to either $\mathbb{R}^n$ or $\mathbb{R}_+ \times \mathbb{R}^{n-1}$ (the latter corresponding to boundary points of M). The integer n is called the dimension of the manifold. A smooth manifold is a manifold for which the concepts of tangent vectors and

† Email address for correspondence: R.S.MacKay@warwick.ac.uk

differentiable functions $f : M \to \mathbb{R}$ are defined and satisfy desirable properties. A more precise definition is given in appendix A. It is oriented if there is a continuous choice of set of n linearly independent tangent vectors at each point (called a frame). Examples of oriented smooth manifolds include $\mathbb{R}^n$, the circle, which I will consider as $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, the n-torus $\mathbb{T}^n$ (which is the product of n copies of the circle) and the n-sphere $\mathbb{S}^n$, which is the set of points in $\mathbb{R}^{n+1}$ such that $\sum_{i=0}^{n} x_i^2 = 1$ (note that, apart from $n = 1$, it is not a product of n copies of the circle).

For integer k between 0 and the dimension n of the manifold (including 0 and n), a differential k-form on M is an antisymmetric k-linear map from ordered k-tuples of tangent vectors at any point of M to the reals. Antisymmetric means it changes sign under interchange of any pair of its arguments. The case of a 0-form is just a scalar function from M to $\mathbb{R}$. The case $k = n$ is called a top form. If we want to indicate a differential form ω at a point $p \in M$, we write ω_p. I shall assume whatever degree of differentiability is required for the results I shall state.

A good introduction to differential forms in the context of Hamiltonian mechanics is chap. 7 of Arnol'd (1978). A lot more advanced material of relevance to hydrodynamics and magnetohydrodynamics is given in Arnol'd & Khesin (1998).

The paper starts by developing the main concepts for differential forms in the context of magnetic fields. It then goes on to use them in charged particle dynamics. It closes with a short discussion and an appendix on some additional topics.

2. Magnetic fields

A magnetic field B is a volume-preserving 3-D vector field.

There are two ways to view a vector field on a manifold. The first is as a field of tangent vectors. First define a differentiable parametrised curve to be a map γ from $[-1, +1] \to M$ such that in one or any local coordinate system $\phi : U \to \mathbb{R}^n$ with U a neighbourhood of the image of γ, $\phi \circ \gamma$ is a differentiable map from $[-1, +1] \to \mathbb{R}^n$. Then a tangent vector is an equivalence class of differentiable parametrised curves through a point, where two such curves γ_1, γ_2 are considered equivalent if $d(\gamma_1(t), \gamma_2(t)) = o(t)$ as $t \to 0$, using any metric d on M. Intuitively, this says that γ_1 and γ_2 pass through the same point at $t = 0$ and have the same velocity at $t = 0$. So a tangent vector is the velocity of a parametrised curve. A vector field is a continuous field of tangent vectors; think of it as the velocity field of a fluid flow.

The second is as a first-order differential operator ('derivation') on scalar functions on the manifold, i.e. a linear operator L on the set of smooth functions $f : M \to \mathbb{R}$ that satisfies Leibniz' product rule $L(fg) = (Lf)g + f(Lg)$.

The link is that a vector field v defined as the velocity of a flow induces a first-order operator on differentiable functions $f : M \to \mathbb{R}$ by $v(f) = Df(v)$, where the derivative Df_p at a point $p \in M$ is the linear map on tangent vectors v at p such that $f(p + \varepsilon v) - f(p) = \varepsilon Df_p(v) + o(\varepsilon)$ as $\varepsilon \to 0$ in any local coordinate system. It is common to write $Df(v)$ as $v \cdot \nabla f$, but that notation suggests it depends on the choice of a Riemannian metric g on M, because for vectors a, b, $a \cdot b$ represents $g(a, b)$ and ∇ is defined using a Riemannian metric. In fact, ∇ is defined by the relation $v \cdot \nabla f = Df(v)$ for all vectors v, so the dependence on the metric cancels out. Thus I prefer to avoid the $v \cdot \nabla f$ notation.

Every first-order differential operator L on functions is of the form $L(f) = Df(v)$ for some vector field v in the first sense. This can be proved using linearity and Leibniz' rule.

2.1. *Volume preservation and magnetic flux form*

The vector calculus way to write the volume-preservation condition is div $B = 0$. To explain how to write it in differential forms language will take several steps.

Firstly, we let Ω be the relevant volume form in three dimensions. A volume form on a manifold is a non-degenerate top form. For a top form, non-degenerate means that there exists an n-tuple on which it is non-zero. In the 3-D context, the interpretation is that $\Omega(\xi, \eta, \zeta)$ is the signed volume of the parallelepiped spanned by any ordered triple of tangent vectors (ξ, η, ζ). In vector calculus language the standard Euclidean volume is $\Omega(\xi, \eta, \zeta) = \xi \cdot (\eta \times \zeta)$.

The next step is to introduce the flux form $\beta = i_B \Omega$ associated with B. This is a 2-form giving the flux of B through any infinitesimal parallelogram. Specifically,

$$i_B \Omega(\xi, \eta) = \Omega(B, \xi, \eta). \tag{2.1}$$

For a general vector field X, the contraction operator i_X on an arbitrary differential k-form ω ($k > 0$) is the $(k-1)$-form defined by putting X as the first argument of ω. We extend to the case of 0-forms f by defining $i_X f = 0$. For future reference, non-degeneracy of a general k-form ω ($k > 0$) is defined by $i_X \omega = 0 \implies X = 0$.

For some purposes, in particular to work out the magnetic flux through a surface, β is the more natural view of a magnetic field than B. Given a 2-form β and a volume form Ω in three dimensions there is always an associated vector field B such that $i_B \Omega = \beta$ and it is unique. This is because Ω is assumed to be non-degenerate.

A k-form ω can be integrated over oriented submanifolds of dimension k. Heuristically, we subdivide the submanifold into small parallelepipeds of dimension k spanned by k tangent vectors $\xi_1, \ldots \xi_k$, sum up $\omega(\xi_1, \ldots \xi_k)$ and take the limit as their size goes to zero. For example $\int_S \beta$ gives the magnetic flux through a surface S and $\int_V \Omega$ gives the volume of a region V. Note that in this notation, no variable of integration is specified; this is because the formulation is coordinate free.

For any k-form ω, $0 \leqslant k < n$, there is a $(k+1)$-form $d\omega$ such that for any oriented $(k+1)$-submanifold V with boundary ∂V (with associated orientation),

$$\int_V d\omega = \int_{\partial V} \omega. \tag{2.2}$$

This was called Stokes' theorem by Cartan, because it encompasses the usual Stokes' theorem (we will get to the differential forms version of curl later), but it also includes Gauss' theorem and Green's theorem. The operator d, called exterior derivative, can be defined by

$$d\omega(\xi_1, \ldots \xi_{k+1}) = \lim_{\varepsilon \to 0} \varepsilon^{-k} \int_{\partial W_\varepsilon} \omega, \tag{2.3}$$

where W_ε is the parallelepiped spanned by $(\varepsilon\xi_1, \ldots \varepsilon\xi_{k+1})$ in some local coordinate system. This definition of d makes Stokes' theorem obvious. On 0-forms f (scalar functions), d is just the usual derivative: applied to a vector ξ it gives $df(\xi) = Df(\xi)$, the directional derivative along ξ. As already discussed, this is commonly written as $\xi \cdot \nabla f$. The point of the definition of d is to extend it to k-forms with $k > 0$. The definition of d can be extended to act on top forms by sending them all to 0.

Now consider $d\beta$ for $\beta = i_B \Omega$. It is a 3-form in three dimensions. The space of top forms at a point is one-dimensional, and Ω is non-degenerate, so $d\beta$ is a multiple of Ω. div B is defined to be that multiple:

$$d\beta = \text{div} B \ \Omega. \tag{2.4}$$

Thus, in differential forms, the volume-preserving condition is written as $d\beta = 0$. Any differential form whose exterior derivative is zero is called closed. Thus the flux form for a magnetic field is closed.

Note how the Cartan version of Stokes' theorem includes Gauss' theorem: for any vector field X, volume form Ω and top-dimensional volume V, apply it to $i_X\Omega$ and use the definition of div; it says $\int_V \text{div} X \; \Omega = \int_{\partial V} i_X\Omega$.

2.2. *Magnetic 1-form*

In addition to the magnetic field B and its flux form $\beta = i_B\Omega$, it is useful to consider the associated 1-form $B^\flat$. This is defined by

$$B^\flat(\xi) = B \cdot \xi \tag{2.5}$$

on any tangent vector ξ. Here $\cdot$ denotes the inner product of two vectors. For a general Riemannian metric g, this is defined by $a \cdot b = g(a, b)$. Equivalently we can write $B^\flat = i_B g$, by extending the contraction notation from differential forms to any covariant tensor, such as a metric tensor (a covariant tensor is a multilinear map from the tangent space to the reals, without the antisymmetry condition).

The inverse operation to $^\flat$ is denoted $^\sharp$, taking a 1-form to a vector field. It is well defined because any Riemannian metric is assumed to be non-degenerate (in fact positive definite, but one can extend both operations to Lorentzian metrics too).

It is also convenient to define $|B| = \sqrt{g(B, B)}$. Using the above, one can write $|B|^2 = i_B B^\flat$.

Finally, we revisit the relation $df(v) = v \cdot \nabla f$ for a scalar function f. Using the above notation we see we can write

$$(\nabla f)^\flat = df, \quad \nabla f = (df)^\sharp. \tag{2.6a,b}$$

3. Cross-product

The cross-product of vectors enters many formulae involving the magnetic field, e.g. the Lorentz force[1] $F = J \times B$, where J is the current-density vector field. The cross-product is specific to three dimensions. There are two ways of writing it in differential forms, which are equivalent if the assumed volume form Ω is natural for the Riemannian metric g, i.e. applied to any orthonormal ordered triple, Ω produces ± 1 according to the orientation of the triple.

The first way is that $J \times B$ is the vector such that

$$(J \times B)^\flat = i_B i_J \Omega. \tag{3.1}$$

Note the reversal of order and the explicit dependence of the cross-product on a volume form Ω and a Riemannian metric (through $^\flat$). One could write $J \times B = (i_B i_J \Omega)^\sharp$, but it is natural to consider forces as covectors rather than vectors, because the work done by a force vector F moving through a displacement ξ is $F^\flat(\xi)$. So I prefer the first formulation. Similarly, momenta are naturally covectors p, with Newton's law $\dot{p} = F^\flat$ and $p(\dot{q})$ tells how fast the action integral changes for velocity $\dot{q}$ of configuration.

The second way is that

$$i_{J \times B}\Omega = J^\flat \wedge B^\flat, \tag{3.2}$$

---

[1] Some call it Laplace force.

where the wedge product of a k-form ω and an m-form η is defined to be the $(k+m)$-form

$$(\omega \wedge \eta)(\xi_1, \ldots \xi_{k+m}) = \sum_{\sigma \in Sh(k,m)} \text{sgn}(\sigma) \ \omega(\xi_{\sigma(1)}, \ldots \xi_{\sigma(k)}) \ \eta(\xi_{\sigma(k+1)}, \ldots \xi_{\sigma(k+m)}), \quad (3.3)$$

with $Sh(k, m)$ being the set of permutations σ of $\{1, \ldots k + m\}$ such that $\sigma(1) < \cdots \sigma(k)$ and $\sigma(k+1) < \cdots \sigma(k+m)$ (shuffles). This definition of $J \times B$ again depends explicitly on a volume form and a Riemannian metric.

It is worth mentioning at this stage how d and i_B act on wedge products,

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-)^k \omega \wedge d\eta, \quad (3.4)$$

where k is the degree of ω, and the same for i_B. Note the case of the wedge product with a 0-form f is just multiplication: $f \wedge \eta = f\eta = \eta \wedge f$.

4. Curl and de Rham cohomology

Written in differential forms, the curl of a 3-D vector field B is the vector field J such that

$$i_J \Omega = dB^\flat. \quad (4.1)$$

So it depends on both the volume form Ω and the Riemannian metric (via $\flat$). We see that the flux form $j = i_J \Omega$ is the natural object here.

Note that the above expresses the relation between a current density J and a magnetic field B in units for which the magnetic permeability $\mu_0 = 1$. We shall take that convention throughout.

A fundamental ingredient of the theory of differential forms is that $d^2 = 0$. One way to see this is if α is an arbitrary $(k-1)$-form, then let $\beta = d\alpha$. Applying Stokes' theorem twice starting with an arbitrary $(k+1)$-submanifold V gives $\int_V d\beta = \int_{\partial V} \beta = \int_{\partial \partial V} \alpha = 0$ because the boundary of a boundary is empty. So $d^2\alpha = d\beta = 0$.

So for example, the flux form j defined above is automatically closed ($dj = 0$). This expresses that div of a curl is zero. I used the notation B for an arbitrary 3-D vector field, which might suggest that I was assuming it is volume preserving, but that was not used above.

As a second example, if B is a gradient vector field ∇f then its curl is automatically zero. This is because $(\nabla f)^\flat = df$, so $j = d^2 f = 0$.

Now the question arises for a differential k-form β, if $d\beta = 0$ then is $\beta = d\alpha$ for some $(k-1)$-form α? In the positive case, β is called exact. In a contractible subset of a manifold this is true (Poincaré's lemma). It generalises the standard results that (i) if curl $v = 0$ then v is locally the gradient of some function ϕ, and (ii) if div $B = 0$ then B is locally curl A for some vector field A. The latter is usually derived as a global result in $\mathbb{R}^3$ via Helmholtz' theorem, but Poincaré's lemma can be proved in a bounded contractible subset, analogous to the standard way for proving (i) (see appendix A).

There can be global topological obstructions, however. For example, for the flux form $\beta = i_B \Omega$ for a volume-preserving field B we have β is closed. So locally, there is a 1-form a such that $\beta = da$; $A = a^\sharp$ is known as a vector potential[2] for B, so the relation is that $\beta = dA^\flat$. But if there is a (non-contractible) closed surface S over

---

[2]It would be more natural to regard the 1-form a as a potential for B, and this is done in some of the literature and is then often denoted by A, but I stick with plasma physics convention.

which $\int_S \beta \neq 0$ then there cannot be a globally defined vector potential for B. In practice for physical magnetic fields this appears never to occur, though maybe in a big enough piece of the universe we will find it does. Note that the vector potential for B is not unique: one can add any gradient to A.

As a second example, however, suppose B is a vacuum magnetic field, i.e. curl $B = 0$, equivalently $dB^\flat = 0$. Then $B^\flat$ is locally the derivative of a 0-form (i.e. scalar function) ϕ, i.e. $B^\flat = d\phi$ or equivalently $B = \nabla\phi$. But it may well fail to be a gradient globally. For example, if the region in which B is a vacuum field is a solid torus, as would arise by generation of B by currents in coils surrounding it, then for any closed curve γ going the non-contractible way in the solid torus, $\int_\gamma B^\flat = \int_{\mathcal{A}} j$ where $\mathcal{A}$ is a disc in the surrounding 3-D space spanning γ, and this is the total external current I_{ext} through the hole in the solid torus, whereas if $B^\flat = d\phi$ then its integral along any closed curve is zero. In this example there is an easy solution: $B = \nabla\phi$ in the solid torus for a multivalued function ϕ, which increases by I_{ext} for each revolution around the solid torus.

The discussion above introduces the fascinating topic of de Rham cohomology. The kth de Rham cohomology group H^k of a manifold M is defined to be the quotient of the set of all closed k-forms on M by the set of all exact ones, i.e. consider two closed forms to be equivalent if they differ by an exact one. It is a group under addition. Actually it is a real vector space. Its dimension $\beta_k(M)$ is called the kth Betti number; H^1 will play a role when we come to the Hamiltonian version of Noether's theorem.

Now that we have learnt about the wedge product and Poincaré's lemma we can formulate the Clebsch representation of a volume-preserving vector field in differential forms, namely $d\beta = 0$, for β the flux 2-form, implies $\beta = da$ locally, for some 1-form a, as above. Every 1-form a can be written locally as fdg for some functions f and g. So $\beta = d(fdg) = df \wedge dg$, which is the differential forms version of $B = \nabla f \times \nabla g$ in three dimensions. There are usually global topological obstructions, however.

5. Lie derivative

The Lie derivative L_B of differential forms (or more general tensors, to be defined later in this section) along a vector field B is one of the most useful concepts, describing how the differential form changes as viewed along the flow of the vector field. I call the vector field B, but no 3-D or volume-preserving conditions are assumed.

We begin with the simple case of a 0-form, i.e. scalar function f. Then $L_B f$ is defined by

$$L_B f = df(B) = i_B df = Df(B) = B \cdot \nabla f. \tag{5.1}$$

So L_B is just the first-order operator view of the vector field B. It is so important in plasma physics, however, that my PhD supervisor John Greene chose car registration $B \cdot GRAD$ (using a screw for the dot). Equations of the form $L_B f = g$ with vector field B and function g given are called magnetic differential equations.

The real power of the Lie derivative is to describe how fast other objects like k-forms with $k > 0$ or vector fields change along a vector field B. In Euclidean space one can get away with writing expressions like $B \cdot \nabla B$ but they require careful interpretation (e.g. in curvilinear coordinates) and do not always behave how you might expect. The nice way to define how fast a general tensor field varies along a vector field is via the flow of the vector field.

Given a smooth vector field B on M, it defines a flow $\phi : \mathbb{R} \times M \to M$, $(t, x) \mapsto \phi_t(x)$ (at least locally around $t = 0$) by

$$\partial_t \phi_t(x) = B(\phi_t(x)), \tag{5.2}$$

with $\phi_0(x) = x$. As a shorthand, write the derivative $D\phi_t$ with respect to $x \in M$ as ϕ_{t*} (note the asterisk is subscripted, indicating that it pushes tangent vectors forward). Define the pullback operator ϕ_t^* of the differentiable map ϕ_t on differential forms (note superscript), e.g. k-form ω, by

$$(\phi_t^* \omega)_p(\xi_1, \ldots \xi_k) = \omega_{\phi_t(p)}(\phi_{t*}\xi_1, \ldots \phi_{t*}\xi_k). \tag{5.3}$$

Define

$$L_B \omega = \partial_t \phi_t^* \omega|_{t=0}. \tag{5.4}$$

Cartan proved the 'magic formula'

$$L_B = i_B d + d i_B \tag{5.5}$$

for L_B on differential forms. We recover the above definition on 0-forms, $L_B = i_B d$, because of the convention that $i_B = 0$ on functions. For top forms, we see that $L_B = d i_B$ because of the convention that $d = 0$ on top forms. Thus an alternative way to write the definition of div with respect to a chosen volume-form Ω is

$$L_B \Omega = \text{div} B \ \Omega. \tag{5.6}$$

One of the main points of the Lie derivative is to articulate how the integral of a form over a surface moving with the flow of a vector field B evolves. The answer is

$$\partial_t \int_{\phi_t S} \omega = \int_{\phi_t S} L_B \omega. \tag{5.7}$$

The two terms of Cartan's magic formula capture how ω changes as viewed along the flow and the effect of how the boundary of $\phi_t S$ changes.

Some nice properties of the Lie derivative on forms are,

$$d L_B = L_B d, \tag{5.8}$$
$$L_B(\alpha \wedge \beta) = (L_B \alpha) \wedge \beta + \alpha \wedge L_B \beta, \tag{5.9}$$
$$L_{\lambda B} \omega = \lambda L_B \omega + d\lambda \wedge i_B \omega. \tag{5.10}$$

Note the case of (5.9) where α is a 0-form f gives

$$L_u(f\beta) = (L_u f)\beta + f L_u \beta, \tag{5.11}$$

and the case of (5.10) where ω is a 0-form f gives

$$L_{\lambda B} f = \lambda L_B f. \tag{5.12}$$

The definition (5.4) extends to all covariant tensors, defined to be multilinear maps on ordered sets of vectors. An example that is not a differential form is a metric tensor, because it is symmetric rather than antisymmetric. The magic formula does not work, however, for general covariant tensors.

It is convenient to extend the definition of the Lie derivative to contravariant tensors (multilinear maps on ordered sets of covectors) and mixed tensors (multilinear maps on ordered sets of vectors and covectors) too. The main example that will concern us is just vector fields, so we restrict attention to that case. A formula to compute the Lie derivative of a general tensor is given in appendix A. The pullback of a vector field Y under the flow ϕ_t of vector field B is

$$\phi_t^* Y|_p = \phi_{-t*} Y|_{\phi_t(p)}. \tag{5.13}$$

Then we define

$$L_B Y = \partial_t \phi_t^* Y|_{t=0}. \tag{5.14}$$

An alternative way to write $L_B Y$ is the commutator $[B, Y]$. To understand this, think of a vector field X as the associated first-order operator $i_X d$ on functions. Then $[B, Y] = BY - YB$ is interpreted as $i_B d i_Y d - i_Y d i_B d$. Although this looks second order, the second-derivative terms cancel and it is actually first order and equals $L_B Y$. An alternative way to think of the commutator $[B, Y]$ is to flow from a point p for time t along B followed by time s along Y and compare the result with flowing from p along Y for time s followed by B for time t. The difference in any local coordinate system is of order st and if one takes the limit of the quotient as $s, t \to 0$ one obtains a vector at p, which we call $[B, Y]$. In vector calculus language, $[B, Y] = B \cdot \nabla Y - Y \cdot \nabla B$. The commutator is antisymmetric.

A useful relation is that on differential forms

$$i_{[J,B]} = i_J L_B - L_B i_J = L_J i_B - i_B L_J, \tag{5.15}$$

which we shall use shortly.

We give some examples of how the Lie derivative shows up in plasma physics. I say a magnetohydrostatic (MHS) field[3] is a 3-D volume-preserving field B such that $J \times B = \nabla p$ for some scalar function p, where $J = \operatorname{curl} B$. In differential forms these equations are written as $i_B i_J \Omega = dp$ and $i_J \Omega = dB^\flat$. One can eliminate mention of J to obtain $i_B dB^\flat = dp$. Now $d i_B B^\flat = d|B|^2$, so it can also be written as

$$L_B B^\flat = d(p + |B|^2). \tag{5.16}$$

Continuing further, apply (5.15) to Ω to obtain

$$i_{[J,B]} \Omega = i_J L_B \Omega - L_B i_J \Omega. \tag{5.17}$$

Now $L_B \Omega = 0$ by div $B = 0$. Also $i_J \Omega = dB^\flat$ and d commutes with L_B on forms, so

$$i_{[J,B]} \Omega = -d L_B B^\flat = -d^2(p + |B|^2) = 0 \tag{5.18}$$

for a MHS field. As Ω is non-degenerate, we deduce that $[J, B] = 0$.

Conversely, for a 3-D volume-preserving vector field B that commutes with its curl, J, there is a possibly multivalued scalar function p such that $J \times B = \nabla p$. This is because the hypotheses imply $d L_B B^\flat = 0$ so $L_B B^\flat$ is locally df for some function f and then we can set $p = f - |B|^2$. One can interpret solutions of $[J, B] = 0$ as MHS fields subject to a possible additional force round each non-contractible loop, like an electromotive force for a charged plasma.

---

[3]This is not universal terminology. The concept of MHS field lies between a magnetostatic field, being a magnetic field resulting from a prescribed steady current distribution (curl $B = J$, div $B = 0$), and a magnetohydrodynamic equilibrium, which allows steady mean flow of plasma and anisotropic pressure.

6. So what?

A nice consequence of the equation $i_B i_J \Omega = dp$ for MHS fields and the resulting commutation relation $[J, B] = 0$ is that bounded regular components of level sets of p are tori, B and J are tangent to them and the flows of B and J are simultaneously conjugate to rotations of different winding ratios on them. A component of a level set of p is regular if $dp \neq 0$ everywhere on it. The flow ϕ^B of B on a 2-torus T is conjugate to a rotation if there exists $h: T \to \mathbb{R}^2/\mathbb{Z}^2$ such that $h(\phi_t^B(x)) = h(x) + rt$ for all $x \in T$, for some $r \in \mathbb{R}^2$, called the rotation vector. The map h is called a conjugacy. Simultaneous conjugacy for B and J means they use the same conjugacy but may have different rotation vectors. These results are known, but the following derivation is slick.

Firstly, a regular component of a level set of any scalar function in three dimensions is a 2-D submanifold. Secondly, applying i_B or i_J to $i_B i_J \Omega = dp$ produces $i_B dp = 0$ and $i_J dp = 0$ because Ω is antisymmetric, so B and J are tangent to the regular level sets of p (this is the same as taking the inner product of $J \times B = \nabla p$ with B or J, but gets rid of the irrelevant apparent dependency on a Riemannian metric). Thirdly, if $dp \neq 0$ everywhere on a submanifold and $i_B i_J \Omega = dp$ then J and B are independent everywhere on it. Finally, the Arnol'd–Liouville theorem (e.g. Arnol'd (1978)) implies that for a bounded 2-D submanifold invariant under the flows of two commuting vector fields that are independent everywhere on it, the submanifold is a torus and the flows are simultaneously conjugate to rotations of different winding ratios on them. The latter conclusion is along the lines of the construction of magnetic coordinates such as Boozer and Hamada coordinates but makes it more natural. Indeed, $[J, B] = 0$ leads directly to Hamada coordinates.

The Arnol'd–Liouville theorem is usually presented for integrable Hamiltonian systems, but the above key step applies more generally. Note that many people deduce the submanifold is a torus just from its being bounded, supporting a nowhere-zero vector field (both J and B satisfy this as it is a consequence of their being independent), and orientable (which is a consequence of being a regular component of a level set of a function in an orientable space). One just computes the Euler characteristic of the submanifold to be zero by Poincaré's index theorem and then uses the classification of surfaces. That proof does not, however, give the additional information that J and B are simultaneously conjugate to rotations.

The winding ratio ι of the magnetic field on a flux surface can be expressed nicely using differential forms. It is defined as the long-time limit of the ratio of the number of turns a fieldline makes in the poloidal direction to the number in the toroidal direction. So it is the ratio of the components of the rotation vector. In differential forms,

$$\iota = -\frac{\displaystyle\int_{\gamma_t} i_B i_n \Omega}{\displaystyle\int_{\gamma_p} i_B i_n \Omega}, \tag{6.1}$$

where $n = \nabla \psi / |\nabla \psi|^2$, over any cycles γ_j on the flux surface such that γ_t makes one toroidal turn and no poloidal ones and *vice versa* for γ_p.

7. Applying diffeomorphisms

If B is a vector field on a manifold M, or ω a differential form on it, and φ is a diffeomorphism from M to another manifold N (or possibly the same one), there is a

natural way to produce a corresponding vector field and differential form on N. Firstly, let φ_* denote the derivative $D\varphi$ of φ, as in § 5. Then $\tilde{B} = \varphi_* B$ is a vector field on N. As an example, h a conjugacy of the flow of B on a torus T to a rotation says $h_* B = r$, constant on $\mathbb{R}^2/\mathbb{Z}^2$. One might think that φ_* could be used to make new magnetic fields from old ones, but the volume-preservation condition might not be satisfied. If we suppose B preserves volume form Ω on M then $di_B\Omega = 0$. Applying φ_* to this we obtain $di_{\tilde{B}}\varphi_*\Omega$, where the definition of φ_* has been extended to the pushforward on k-forms ω by

$$(\varphi_*\omega)_p(\xi_1, \ldots \xi_k) = \Omega_{\varphi^{-1}(p)}(\varphi_*^{-1}\xi_1, \ldots \varphi_*^{-1}\xi_k), \tag{7.1}$$

and we note that φ_* passes through d and through contractions (in the sense that $\varphi_* i_B\Omega = i_{\varphi_* B}\varphi_*\Omega$). Thus $\tilde{B}$ preserves a volume form on N, but perhaps not the one you had in mind. If one has a prior choice $\tilde{\Omega}$ of volume form on N then $\tilde{B}$ preserves it if $\varphi_*\Omega = \tilde{\Omega}$.

Let us consider instead the action of φ_* on the magnetic flux-form $\beta = i_B\Omega$. We know that div $B = 0$ if $d\beta = 0$. We let $\tilde{\beta} = \varphi_*\beta$ on N. Applying φ_* to $d\beta = 0$ we deduce that $d\tilde{\beta} = 0$. So φ_* takes a magnetic flux form to one on N, without any further conditions. This is one of the reasons to consider the magnetic flux form to be more fundamental than the magnetic field.

Such diffeomorphisms φ (with $N = M$) arise as the flow ϕ_t of the velocity field v of a perfectly conducting fluid (which can be time dependent). The induction equation $\partial_t B = \text{curl}(v \times B)$ implies $\partial_t \beta = -L_v \beta$ for the magnetic flux form. The definition of L_v on forms used the pullback, but one can equally use the pushforward with change of sign. So the flow applies ϕ_{t*} to β. This is Alfvén's frozen flux theorem.

It is common to restrict attention to incompressible flows, in which case the induction equation implies that $\partial_t B = [v, B]$ so ϕ_{t*} then also gives the evolution of B. This is often referred to as preservation of the topology of the magnetic field, but is much stronger than that term suggests, because of Alfvén's theorem. In general the induction equation for B is $\partial_t B = -[v, B] - (\text{div } v)B$ so there is an additional stretching effect on B from convergence of v, beyond that implied by ϕ_{t*}.

8. Coordinates

The point of differential forms is to be coordinate free. If you are desperate to connect to index notation, however, here is the correspondence.

A coordinate system or chart on an open subset U of a manifold M of dimension n is a set of n functions $x^1, \ldots x^n : U \to \mathbb{R}$ such that the map $p \mapsto (x^1(p), \ldots x^n(p))$ is a diffeomorphism from U to its image in $\mathbb{R}^n$. It is conventional to use superscripts for the coordinate functions.

The linear operators $\partial_i = (\partial/\partial x^i)$ (keeping the other x^j fixed) form a basis for the tangent space T_pM at a point $p \in U$ (in the differential operator view of vectors). It is conventional to use subscripts for these partial derivative operators. So a vector B at p can be expanded as $B = \sum_{i=1}^n B^i \partial_i$. It is conventional to use superscripts for the components of a vector. It is convenient to adopt the summation convention that in an expression with an index appearing once as a subscript and once as a superscript, there is an implied sum over the possible values of that index, so $B = B^i \partial_i$, for example.

The 1-forms dx^i at a point p form a basis of the cotangent space T_p^*M, so $B^\flat$ can be expanded as $B^\flat = B_i dx^i$, using subscripts for its components, and summation convention.

It is common to think of a vector field B and its associated 1-form $B^\flat$ as being the same object and refer to the B^i as being its contravariant components and the B_i as its covariant components. They are related, as in the definition of $^\flat$, by an assumed Riemannian metric. A metric tensor g can be expanded as $g(\xi, \eta) = g_{ij}\xi^i\eta^j$ on pairs of vectors ξ, η with respect to their components in a coordinate system. Then the relation between B and $B^\flat$ can be written $B_i = g_{ij}B^j$. The notation $^\flat$ corresponds to the operation of lowering indices in index notation. The operation $^\sharp$ corresponds to raising them. If α is a 1-form $\alpha_i dx^i$ then $\alpha^\sharp$ is a vector field $\alpha^i\partial_i$ with $\alpha^i = g^{ij}\alpha_j$, the matrix g^{ij} being the inverse of the matrix g_{ij}.

A minor caution: many plasma physicists use the term 'contravariant representation' of a vector field B to mean the flux form β, rather than the vector field itself. They agree, however, that the 'covariant representation' of a vector field B is the 1-form $B^\flat$.

A basis for 2-forms at a point is given by the $dx^i \wedge dx^j$ with $i < j$ to avoid duplication. It can be tidier to think of the basis elements as being $dx^i \wedge dx^j - dx^j \wedge dx^i$ with $i < j$, which are twice the preceding ones. So one can expand a 2-form like β as $\beta = \beta_{ij}dx^i \wedge dx^j$, with β_{ij} antisymmetric. Similarly, a volume form Ω in three dimensions can be written as $\Omega = \Omega_{ijk}dx^i \wedge dx^j \wedge dx^k$ with Ω_{ijk} being completely antisymmetric, although of course there is only one independent 3-form in three dimensions, say $dx^1 \wedge dx^2 \wedge dx^3$ so Ω_{ijk} is a multiple $\mathcal{J}^{-1}$ of the usual ε_{ijk} ($\mathcal{J}$ is often called the Jacobian of the coordinate system).

The contraction operator i_B is easily written in index notation, e.g. $i_B\Omega$ is the 2-form with components $B^i\Omega_{ijk}$. The fact that Ω is a multiple of ε is a possible reason for plasma physicists to refer to β as the contravariant representation of B, but the Jacobian must be taken into account too.

The condition for a volume form Ω to be natural for a Riemannian metric g can be written in coordinates as $\Omega = \pm\sqrt{\det g}\, dx^1 \wedge \ldots dx^n$, where $\det g$ is the determinant of the matrix for the covariant representation of g in coordinates $x^1, \ldots x^n$. The two possible signs for Ω correspond to the two possible orientations of a connected orientable manifold.

For a simple k-form fdx^I, where I is an ordered set of k distinct indices $i_1, \ldots i_k$ and $dx^I = dx^{i_1} \wedge \ldots \wedge dx^{i_k}$, the exterior derivative has the formula

$$d(fdx^I) = (\partial_i f)dx^i \wedge dx^I. \tag{8.1}$$

A general k-form is a linear combination of simple ones.

On a k-form ω,

$$(L_B\omega)_{i_1\cdots i_k} = B^c\partial_c\omega_{i_1\cdots i_k} + (\partial_{i_1}B^c)\omega_{ci_2\cdots i_k} + \cdots + (\partial_{i_k}B^c)\omega_{i_1\cdots i_{k-1}c}. \tag{8.2}$$

On a vector field Y,

$$L_BY = (B^j\partial_j Y^i - Y^j\partial_j B^i)\partial_i. \tag{8.3}$$

9. Charged particle motion in a magnetic field

We now turn to the dynamics of a charged particle in a magnetic field B. Denote its mass by m and its charge by e. The equation of motion in Euclidean space is

$$m\ddot{q} = e\dot{q} \times B(q). \tag{9.1}$$

It is fruitful to put this into Hamiltonian form. The standard way is to choose a vector potential A for B and introduce a momentum variable $p = m\dot{q} + eA(q, t)$ and let the

Hamiltonian be $H(q, p, t) = (1/2m)|p - eA(q, t)|^2$. The canonical Hamilton equations $\dot{q} = \partial_p H$, $\dot{p} = -\partial_q H$, reproduce the right equations of motion.

It is better, however, to abandon the canonical view of Hamiltonian systems. Instead, a Hamiltonian system $\dot{x} = X(x, t)$ (possibly time dependent) on a manifold M is defined by

$$i_X \omega = dH, \tag{9.2}$$

for a function $H : M \times \mathbb{R} \to \mathbb{R}$ called the Hamiltonian, and symplectic form ω on M. A symplectic form is a non-degenerate closed 2-form. This equation defines (existence and uniqueness) the vector field X because ω is non-degenerate. Note that non-degeneracy of ω requires M to have even dimension; denoting it by $2n$, we call n the number of degrees of freedom (DoF).

The standard example is that M is the cotangent bundle T^*Q of a manifold Q, i.e. the set of covectors to Q. One can write a covector as (q, p) where $q \in Q$ and p is a covector at q, i.e. $p : T_q Q \to \mathbb{R}$ and is linear. T^*Q has a natural symplectic form, as follows. Define the natural 1-form α on T^*Q by $\alpha_{(q,p)}(\delta q, \delta p) = p(\delta q)$. In a local coordinate system q^i on Q we define associated coordinates p_i so that $p(\delta q) = p_i \delta q^i$ (with summation convention). Then $\alpha = p_i dq^i$. Finally, let $\omega = -d\alpha$. It is a closed (indeed exact), non-degenerate 2-form. In the above coordinates $\omega = dq^i \wedge dp_i$.

A simple mechanical system on T^*Q is defined by this symplectic form ω and $H(q, p) = \frac{1}{2} p^T M^{-1} p + V(q)$ for some positive definite 'mass' matrix M, which takes vectors to covectors, and 'potential' $V : Q \to \mathbb{R}$. Solving $i_X \omega = dH$ for $X = (\dot{q}, \dot{p})$ gives

$$\dot{q} = M^{-1} p, \tag{9.3}$$
$$\dot{p} = -dV_q, \tag{9.4}$$

so $M\ddot{q} = -dV_q$. One can allow M to depend on q, which is important to treat mechanical linkages for example, but it adds extra terms to the equations of motion, analogous to centrifugal and Coriolis forces. An equivalent way to put this is that for a linkage with configuration space Q the kinetic energy is half the norm squared of the momentum covector for some Riemannian metric g on Q, $|p|_q^2 = g^{ij}(q)p_i p_j$ in local coordinates.

Let us treat the motion of a charged particle in a magnetic field. We will do it on a general oriented 3-D manifold Q with Riemannian metric g. The Riemannian metric gives $|v|^2 = g_q(v, v) = g_{ij}(q)v^i v^j$ for a vector v, and for a covector p the length squared is defined to be $|p|^2 = g^{ij}(q)p_i p_j$, as above. The dynamics is on the cotangent bundle T^*Q. We take $H = (1/2m)|p|^2$ and[4]

$$\omega = -d\alpha - e\pi^* \beta, \tag{9.5}$$

where α is the natural 1-form on T^*Q, $\beta = i_B \Omega$, $\pi : T^*Q \to Q$ is $\pi(q, p) = q$ and π^* is its pullback from Q to T^*Q (recall (5.3)). So the dynamics $(\dot{q}, \dot{p})$ is given by solving

$$\omega((\dot{q}, \dot{p}), (\xi_q, \xi_p)) = dH(\xi_q, \xi_p) \ \forall \xi. \tag{9.6}$$

Specialising to the case of Euclidean metric, this gives $\dot{q} = p/m$, so $p = m\dot{q}$, the ordinary kinetic momentum, and $-\dot{p}(\xi_q) - e\beta(\dot{q}, \xi_q) = 0$ for all ξ_q. Now $\beta(\dot{q}, \xi_q) = \Omega(B, \dot{q}, \xi_q)$, so this says that $\dot{p} = -eB \times \dot{q}$, which indeed recovers the right equations of motion.

Advantages of the differential forms formulation $i_X \omega = dH$ of Hamiltonian dynamics are that it makes it easy to see that:

---

[4]If B has a vector potential A, or better a 1-form potential a, this can be written as $\omega = -d(\alpha + e\pi^* a)$.

(i) If H is time independent then H is conserved along X: $i_X dH = i_X i_X \omega = 0$ by antisymmetry (in the time-dependent case, along solutions we obtain $dH/dt = i_X dH + \partial_t H$, so $dH/dt = \partial_t H$).

(ii) ω is conserved along X: $L_X \omega = i_X d\omega + d i_X \omega = 0$ because $d\omega = 0$ and $d^2 H = 0$.

(iii) It automatically takes care of acceleration in arbitrary coordinate systems (which otherwise requires introducing the Levi-Civita connection into Newton's equations, e.g. centrifugal and Coriolis forces).

(iv) It allows to obtain conservation laws and reductions from continuous symmetries.

Let us expand on the latter point. We say a vector field U on M is a continuous symmetry of (H, ω) if $L_U H = 0$ and $L_U \omega = 0$. Then the second equation shows that $d i_U \omega = 0$, thus U is locally Hamiltonian, i.e. $i_U \omega = dK$ for some function $K : M \to \mathbb{R}$ locally. Furthermore, $i_X dK = i_X i_U \omega = -i_U dH = 0$ from the first condition of a symmetry. So K is conserved by X. This is a version of Noether's theorem.

There remains the question whether K is globally defined. Let $\tilde{M}$ be the universal cover of M, i.e. the set of equivalence classes of curves from a chosen base point, under the equivalence relation of continuous deformation fixing the ends of the curve. Then the dynamics can be lifted to $\tilde{M}$ and K is globally defined (up to a constant) on $\tilde{M}$. Next, note that $[X, U] = 0$ because

$$i_{[X,U]}\omega = i_X L_U \omega - L_U i_X \omega = 0, \tag{9.7}$$

using $L_U \omega = 0$, $i_X \omega = dH$ and $L_U H = 0$. As ω is non-degenerate, it follows that $[X, U] = 0$. Thus X and U are commuting vector fields on each level set of K. If the first homology group $H_1(M)$ (1-D cycles modulo boundaries of 2-D surfaces) is spanned by closed trajectories of the set of vector fields of the form $aU + bX$ for arbitrary choices of functions a, b then the change in K around any non-contractible loop is zero and so K is well defined on M. To see this, let γ be a closed orbit of $aU + bX$, then

$$\int_\gamma i_U \omega = \int i_U \omega(\dot{\gamma}) \, dt = \int i_U \omega(aU + bX) \, dt = \int \omega(U, aU) + \omega(U, bX) \, dt = 0, \tag{9.8}$$

because of antisymmetry for the first term and the second is $b \, dH(U) = 0$. It follows that the change in K round such a loop is zero. It is not even necessary to find closed trajectories. Asymptotic cycles in the sense of Schwartzmann will do (Fried 1982).

The homology group $H_1(M)$ is dual to the de Rham cohomology group $H^1(M)$ in the sense that given $\gamma \in H_1$ and $\alpha \in H^1$, there is a natural scalar $\langle \alpha, \gamma \rangle = \int_\gamma \alpha$. It is well defined despite the freedom to deform γ and add any exact 1-form to α.

This version of Noether's theorem is more sophisticated than the usual one for Lagrangian systems, where the symmetry is restricted to being on configuration space. Hence the need for the additional step of checking whether K is global.

Note the trivial case of $U = X$, which achieves nothing. If U, X are independent almost everywhere, however, then so are dK, dH, so we obtain a genuine reduction of the dynamics by one dimension by restricting to level sets of K.

Actually, one can reduce by two dimensions, by also quotienting by the flow ϕ of U if its orbit space is a manifold. The resulting vector field of X on $K^{-1}(k)/\phi$ is Hamiltonian with respect to the reduced symplectic form ω and Hamiltonian H, which we denote by the same symbols. Note that the symplectic form indeed reduces to the quotient because $\omega(u, \xi) = 0$ for all $\xi \in \ker dK$.

Does the symplectic form have physical manifestations? The answer is yes, e.g. (MacKay 1992). One manifestation is that Liouville volume $\Omega = \omega^{\wedge n}/n!$ is conserved by a system of n DoF (here $\omega^{\wedge n}$ is the wedge product of ω with itself n times, and the factor $n!$ is conventional). Since H is also conserved, we deduce conservation of energy-surface volume μ_E on $H^{-1}(E)$ defined uniquely by taking any $(2n - 1)$-form μ such that $\mu \wedge dH = \Omega$ and restricting it to $H^{-1}(E)$. One can write $\mu = i_n\Omega$, with $n = \nabla H/|\nabla H|^2$ for any Riemannian metric. This is the basis for the theory of entropy for classical mechanical systems. Another is the action $S = \int_A \omega$ for any area A spanning a closed curve γ, which is also conserved under the flow of X. Perhaps the action of a closed curve is not considered physical, but bear in mind that if γ is a periodic orbit of a Hamiltonian system then it is generically part of a family of such, parametrised by the value E of H, and the period $T = dS/dE$. Furthermore, if the Hamiltonian has slow time dependence then a periodic orbit drifts in energy in such a way as to preserve its action to high order of approximation (an 'adiabatic invariant').

10. Charged particle in an axisymmetric magnetic field

We give an example of application of Noether's theorem to charged particle motion in a magnetic field in Euclidean space. Recall that it is Hamiltonian with $H = (1/2m)|p|^2$ and $\omega = -d\alpha - e\pi^*\beta$ on $T^*\mathbb{R}^3$, where α is the natural 1-form for a cotangent bundle and $\beta = i_B\Omega$. Let vector field $u = \partial_\phi$ on $T^*\mathbb{R}^3$ with respect to cylindrical coordinates (r, ϕ, z) on $\mathbb{R}^3$. This uses the first-order operator view of a vector field, but we could write it as a velocity field $r\hat{\phi} = (0, r, 0, 0, 0, 0)$ in $(r, \phi, x, p_r, p_\phi, p_z)$. The momentum coordinates are defined so that the natural 1-form $\alpha = p_i dq^i$. Then $H = (1/2m)(p_r^2 + r^{-2}p_\phi^2 + p_z^2)$.

Say B is axisymmetric if $L_u\beta = 0$ (an alternative definition is $[u, B] = 0$, but since B and the chosen u are volume preserving, this reduces to $L_u\beta = 0$). Then $di_u\beta = 0$ because $d\beta = 0$. So $i_u\beta = d\psi$ for some function ψ locally, called a flux function. In fact, ψ is global because $\mathbb{R}^3$ is contractible. But if for some reason the field was defined or axisymmetric only on some axisymmetric solid torus, for example, then ψ would still be global, because $i_u d\psi = i_u i_u \beta = 0$ by antisymmetry so ψ is independent of ϕ.

An alternative approach to deriving a flux function for an axisymmetric magnetic field is to use a vector potential A for B and define axisymmetry by $L_u A^\flat = 0$. Then $i_u\beta = i_u dA^\flat = -di_u A^\flat$, so we get a global flux function $\psi = -i_u A^\flat = -A \cdot u = -rA_\phi$.

The relation $i_u\beta = d\psi$ implies $i_u i_B\Omega = d\psi$. This is written in vector calculus as $B \times u = \nabla\psi$. Since $u = r\hat{\phi}$ we deduce that

$$B_z = -\frac{1}{r}\frac{\partial\psi}{\partial r}, \tag{10.1}$$

$$B_r = \frac{1}{r}\frac{\partial\psi}{\partial z}. \tag{10.2}$$

Note this is a 1 DoF Hamiltonian system for the reduction of fieldline flow by axisymmetry: $H = \psi(r, z)$, $\omega = r dz \wedge dr$. To get the full fieldline flow one just adds $\dot{\phi} = (1/r)B_\phi$.

The axisymmetry of B makes the charged particle dynamics axisymmetric too: $L_u\omega = 0$, $L_u H = 0$. We determine the resulting constant of the motion by noting that

$$i_u\omega = dp_\phi - ed\psi. \tag{10.3}$$

Thus the motion conserves $L = p_\phi - e\psi$, a modified angular momentum. As usual, $p_\phi = mr^2\dot{\phi}$. So we deduce that $\dot{\phi} = (1/mr^2)(L + e\psi(r, z))$. We can reduce the system by axisymmetry to a family of systems on (r, z, p_r, p_z) with $p_\phi - e\psi = L$ constant,

$$H = \frac{1}{2m}\left(p_r^2 + p_z^2 + \frac{(L + e\psi)^2}{r^2}\right) \tag{10.4}$$

$$\omega = dr \wedge dp_r + dz \wedge dp_z - eB_\phi(r, z)dz \wedge dr. \tag{10.5}$$

This is the basic Hamiltonian system for the motion of charged particles in a tokamak. To proceed further, however, it is good to notice a further approximate symmetry: gyro-phase rotation. We will tackle that in the next section, for fields that are not necessarily axisymmetric.

11. Magnetic moment and guiding-centre motion

If $B(q(t), t)$ seen by a charged particle changes slowly on the time scale of one gyroperiod $2\pi/\Omega$ (gyrofrequency $\Omega = -e|B|/m$), then the magnetic moment $\mu = m|v_\perp|^2/2|B|$ is an adiabatic invariant. This means that it remains close to its initial value for an exceedingly long time (depending on the smoothness of the variation of B).

This is a consequence of the Hamiltonian structure of the dynamics and is most simply revealed using the differential forms approach (as was shown by Littlejohn (1983) though he used a variational formulation on extended state space). We treat the time-independent case, but almost no change is required for the time-dependent case.

The idea is to change coordinates from (q, p) to $(X, \rho, p_\parallel)$ with guiding centre $X \in \mathbb{R}^3$, gyroradius vector ρ perpendicular to $B(X)$ and parallel momentum $p_\parallel \in \mathbb{R}$, given by solving the following system of equations:

$$p = eB(X) \times \rho + p_\parallel b(X) \tag{11.1}$$

$$q = X + \rho, \tag{11.2}$$

where $b = B/|B|$. For $e|B|$ large, there is a unique solution with ρ small (approximately $p \times B/e|B|^2$ evaluated at q), by the implicit function theorem. More precisely, this works if $|\rho|$ is less than the radius of curvature of the fieldlines.

Then the dynamics has approximate rotation symmetry of ρ about B. In particular,

$$H = \frac{|p|^2}{2m} = \frac{1}{2m}(p_\parallel^2 + e^2|B(X)|^2|\rho|^2) \tag{11.3}$$

is exactly rotation invariant in ρ. The symplectic form requires more work

$$\omega = -d\alpha - e\beta, \tag{11.4}$$

where I have dropped the π^* in front of β because it is obvious what is meant. In these $(X, \rho, p_\parallel)$ coordinates and choosing perpendicular components (ρ_1, ρ_2) for ρ and denoting the parallel component of X by X_3,

$$\alpha = e|B|\rho_2(dX_1 + d\rho_1) - e|B|\rho_1(dX_2 + d\rho_2) + p_\parallel dX_3. \tag{11.5}$$

The proposed symmetry vector field is $u = (0, 0, 0, -\rho_2, \rho_1, 0)$ in $(X, \rho, p_\parallel)$ coordinates. We will compute $L_u\alpha$ and $L_u\beta$ and then combine them to deduce that $L_u\omega$ is small.

Using Cartan's formula for L_u on differential forms,

$$L_u \rho_2 d\rho_1 = i_u d\rho_2 \wedge d\rho_1 + d(-\rho_2^2) = \rho_1 d\rho_1 + \rho_2 d\rho_2 - d(\rho_2^2) \tag{11.6}$$

$$L_u(-\rho_1 d\rho_2) = -i_u d\rho_1 \wedge d\rho_2 - d(\rho_1^2) = \rho_2 d\rho_2 + \rho_1 d\rho_1 - d(\rho_1^2). \tag{11.7}$$

They sum to zero. Also $L_u(p_\parallel dX_3) = 0$ and $L_u|B(X)| = 0$. So we are left with

$$\begin{aligned}
L_u \alpha &= e|B|L_u(\rho_2 dX_1 - \rho_1 dX_2) \\
&= e|B|i_u(d\rho_2 \wedge dX_1 - d\rho_1 \wedge dX_2) = e|B|(\rho_1 dX_1 + \rho_2 dX_2).
\end{aligned} \tag{11.8}$$

Thus

$$L_u(-d\alpha) = -e|B(X)|(d\rho_1 \wedge dX_1 + d\rho_2 \wedge dX_2). \tag{11.9}$$

For $L_u\beta$, we have to take into account that $\beta = |B(q)|dq_1 \wedge dq_2$ with $q = X + \rho$. So

$$e\beta = e|B(X + \rho)|(dX_1 \wedge dX_2 + dX_1 \wedge d\rho_2 + d\rho_1 \wedge dX_2 + d\rho_1 \wedge d\rho_2). \tag{11.10}$$

Recalling that β is closed,

$$\begin{aligned}
L_u(-e\beta) &= -edi_u\beta = ed\left(|B(q)|(\rho_1 dX_1 + \rho_2 dX_2 + \rho_2 d\rho_2 + \rho_1 d\rho_1)\right) \\
&= e|B(q)|(d\rho_1 \wedge dX_1 + d\rho_2 \wedge dX_2) \\
&\quad + ed|B(q)| \wedge (\rho_1 dX_1 + \rho_2 dX_2 + \rho_2 d\rho_2 + \rho_1 d\rho_1).
\end{aligned} \tag{11.11}$$

The first term almost cancels $L_u(-d\alpha)$, the difference being just due to where $|B|$ is evaluated, but ρ is small and B varies slowly in space. The second term is small because it is proportional to $d|B|$ and B varies slowly in space. Specifically it is

$$e\left((\rho_2 \partial_{q_1}|B| - \rho_1 \partial_{q_2}|B|)\ dq_1 \wedge dq_2 + \partial_{q_3}|B|\ dX_3 \wedge (\rho_1 dq_1 + \rho_2 dq_2)\right). \tag{11.12}$$

Hence $L_u\omega$ is small and u is an approximate symmetry of the dynamics.

It follows that $i_u\omega$ is approximately d of some function and that function is approximately conserved. It is conventionally written as $-m\mu/e$ where the magnetic moment

$$\mu = \frac{e^2}{2m}|B||\rho|^2 = \frac{m|v_\perp|^2}{2|B|}. \tag{11.13}$$

The proof is that, ignoring the dependence of $|B|$ on position,

$$i_u\omega = -e|B|(\rho_1 d\rho_1 + \rho_2 d\rho_2) = -\frac{e|B|}{2}d|\rho|^2 = -\frac{m}{e}d\mu. \tag{11.14}$$

Reducing by the approximate symmetry u produces a 2 DoF system called first-order guiding-centre motion, on the space of $(X, p_\parallel) \in \mathbb{R}^3 \times \mathbb{R}$,

$$H = \frac{1}{2m}p_\parallel^2 + \mu|B(X)| \tag{11.15}$$

$$\omega = -d(p_\parallel b^\flat) - e\beta_X = b^\flat \wedge dp_\parallel - p_\parallel db^\flat - ei_B\Omega. \tag{11.16}$$

Let $c = \operatorname{curl} b$, so $i_c\Omega = db^\flat$. Then ω can be written

$$\omega = b^\flat \wedge dp_\parallel - ei_{\widetilde{B}}\Omega, \tag{11.17}$$

with the modified magnetic field

$$\widetilde{B} = B + \frac{p_\parallel}{e} c. \tag{11.18}$$

The 2-form ω is closed; it is non-degenerate except where $\widetilde{B} \cdot b = 0$. Then the equation of motion is given by solving

$$i_{(\dot{X}, \dot{p}_\parallel)} \omega = dH. \tag{11.19}$$

Applying to a vector of the form $(0, \delta p_\parallel)$, we deduce that

$$\dot{X}_\parallel = p_\parallel / m. \tag{11.20}$$

Applying to a vector of the form $(\xi, 0)$:

$$e(\widetilde{B} \times \dot{X}) \cdot \xi = \xi \cdot (\mu \boldsymbol{\nabla} |B| + \dot{p}_\parallel b). \tag{11.21}$$

So

$$e\widetilde{B} \times \dot{X} = \mu \boldsymbol{\nabla} |B| + \dot{p}_\parallel b. \tag{11.22}$$

Taking the inner product with $\widetilde{B}$ yields

$$\dot{p}_\parallel = -\mu \frac{\widetilde{B} \cdot \boldsymbol{\nabla} |B|}{\widetilde{B} \cdot b}. \tag{11.23}$$

Finally, taking the cross-product of (11.22) with b and using (11.20) we obtain

$$\dot{X} = \frac{1}{\widetilde{B} \cdot b} \left(\frac{\mu}{e} b \times \boldsymbol{\nabla} |B| + \frac{p_\parallel}{m} \widetilde{B} \right). \tag{11.24}$$

Equations (11.23) and (11.24) are the guiding-centre equations.

One sees there is a repulsion from increasing $|B|$ along b (which corresponds exactly to conservation of H in (11.15)), and there is a perpendicular drift driven by $\mu \boldsymbol{\nabla}_\perp |B|$ and $(p_\parallel^2/me) b \cdot \boldsymbol{\nabla} b$ (coming from the curl b term of $\widetilde{B}$).

There are alternative forms for the guiding-centre equations, agreeing to first order in $1/e$, but the advantage of the above approach is that the resulting drift equations retain Hamiltonian structure, allowing deeper understanding of their behaviour and extension to gyrokinetic theory (e.g. Qin *et al.* 2007).

12. Discussion

We have shown how differential forms can make some results in plasma physics simpler and more intuitive. They also open up the possibility of new discoveries.

An example is the interaction of two charged particles in a magnetic field. We found a new constant of the motion in the case where the particles have the same e/m (including sign), corresponding to a novel symmetry that I call 'locomotive coupling rod symmetry' (MacKay & Pinheiro 2006, 2008). This is an important case because it includes the interaction of two electrons or of two protons.

The topic in plasma physics where I have found differential forms most useful is that of quasi-symmetry. Quasi-symmetry just means integrability of guiding-centre motion, but without assuming the symmetry corresponds to an isometry. It is a good principle for stellarator design (Helander 2014). Using differential forms, we have gone considerably beyond what was known before (Burby, Kallinikos & MacKay 2019), though we have not yet a complete understanding of quasi-symmetry.

Acknowledgements

Appendix A

A.1. *Smooth manifolds*

Here is a slightly more formal introduction to smooth manifolds. A topological manifold of dimension n is a second countable Hausdorff topological space such that every point has a neighbourhood homeomorphic to $\mathbb{R}^n$ or $\mathbb{R}_+ \times \mathbb{R}^{n-1}$. A neighbourhood of a point is an open set containing that point. Second countable means there is a countable set of open sets such that any open set can be obtained by finite intersections and arbitrary unions of them. Hausdorff means that for any two points x, y there are disjoint neighbourhoods U_x, U_y of x, y. A differentiable manifold M is a topological manifold with an open cover $U_\alpha, \alpha \in A$, and maps $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ which are homeomorphisms to their images such that the change-of-chart maps $\phi_\alpha \circ \phi_\beta^{-1}$ between the subsets of $\mathbb{R}^n$ on which they are defined are diffeomorphisms. The open cover is called an atlas and the maps ϕ_α are called charts. It is C^r if the change-of-chart maps are C^r.

A function $f : M \to \mathbb{R}$ is differentiable if the maps $f \circ \phi_\alpha^{-1} : \mathbb{R}^n \to \mathbb{R}$ are differentiable where defined.

A.2. *Poincaré's lemma*

This is the statement that if a k-form β is closed on a contractible subset U of a manifold then $\beta = d\alpha$ for some $(k-1)$-form on U. Here is a proof. U contractible means there is a vector field X on U with forward flow ϕ that maps U into itself and the image $\phi_t U$ contracts to a point as $t \to \infty$. Define $(k-1)$-form α on U

$$\alpha = -\int_0^\infty i_X \phi_t^* \beta \, \mathrm{d}t. \tag{A 1}$$

Then

$$d\alpha = -\int_0^\infty \mathrm{d}i_X \phi_t^* \beta \, \mathrm{d}t = -\int_0^\infty L_X \phi_t^* \beta + i_X \, \mathrm{d}\phi_t^* \beta, \tag{A 2}$$

using $L_X = di_X + i_X d$ on forms. The second integrand can be written as $i_X \phi_t^* d\beta$ so is zero by the hypothesis that β is closed. The first integrand can be written as $\partial_t \phi_t^* \beta$, thus $d\alpha$ is minus the change in $\phi_t^* \beta$ from $t = 0$ to ∞. But $\phi_t^* \beta \to 0$ as $t \to \infty$ because of the contraction to a point, and $\phi_0^* \beta = \beta$. Thus $d\alpha = \beta$.

A.3. *Hodge star*

Although I have chosen not to use it, you may come across the Hodge star operator $\star$ in your reading, so here is a little introduction to it. In an oriented Riemannian manifold of dimension n it gives a natural bijection between k-forms and $(n-k)$-forms. Let us specialise to $n = 3$. Then any 3-form ω is a scalar multiple of the Riemannian volume form Ω, say $\omega = f\Omega$, because the space of top-forms at a point is one-dimensional and Ω is non-degenerate. The relation is denoted by $\omega = \star f$, equivalently $f = \star\omega$. Thus for example one could write div $B = \star d\beta$ (with $\beta = i_B \Omega$).

Similarly, we have seen that a vector field B (not necessarily volume preserving) produces both a 2-form $\beta = i_B \Omega$ and a 1-form $B^\flat$. The relation between the two is denoted by $\beta = \star B^\flat$, $B^\flat = \star \beta$. Thus one can write $J = \mathrm{curl}\, B$ as $J^\flat = \star dB^\flat$ and $\mathrm{div}\, B = \star d \star B^\flat$. The combination $\delta = \star d \star$ is called the co-differential, so we could write $\mathrm{div}\, B = \delta B^\flat$. In particular the Laplacian of a function f is $\Delta f = \delta df$. The Laplacian can be extended to k-forms by $\Delta = \delta d + d\delta$.

I will not give the general definition of $\star$ here, but it has the nice feature that on odd-dimensional Riemannian manifolds, $\star^2 = 1$, as you have seen in the 3-D case. Beware, however, that in even dimension or for Lorentzian manifolds, $\star^2 = \pm 1$ depending on the degree of the forms involved.

A.4. *Electromagnetism in space–time*

A beautiful observation is that in a 4-D Lorentzian manifold, two of Maxwell's equations can be expressed as the Faraday tensor being a closed 2-form. Given electric field E and magnetic field B, the Faraday tensor is the 2-form

$$F = B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy + E_x dx \wedge dt + E_y dy \wedge dt + E_z dz \wedge dt, \quad \text{(A 3)}$$

in a local Minkowski coordinate system (i.e. such that the metric is $ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2$ to leading order). Then Faraday's law ($\partial B / \partial t = -\mathrm{curl}\, E$) and $\mathrm{div}\, B = 0$ are equivalent to the single statement $dF = 0$. This is because

$$F = \pi^* \beta + (\pi^* E^\flat) \wedge dt, \quad \text{(A 4)}$$

where π maps space–time to space, so using (8.1), $dF = 0$ if $d\beta = 0$ and $\partial \beta / \partial t + dE^\flat = 0$.

The electromagnetic force on a charge e with 4-velocity U in space–time is the 1-form $f = -e i_U F$. One can do a relativistic treatment of the reduction of charged particle motion to guiding-centre motion.

Using the Hodge star, the two remaining Maxwell equations can be expressed as

$$d \star F = \star J, \quad \text{(A 5)}$$

where J is a 1-form on space–time representing densities of charge and current. Applying d to (A 5) yields $d \star J = 0$, which expresses charge conservation. The beauty of the formulation is that it applies to arbitrary Lorentzian manifolds, taking care automatically of the effects of curvature.

A.5. *Helicity*

The helicity integral $\mathcal{H} = \int A \cdot B \, dV$, where A is a vector potential for B, plays an important role in Taylor's theory of relaxation of plasma in a magnetic field and also can be interpreted as an average rate of winding of field lines around each other (Arnol'd & Khesin 1998). It has a nice representation in differential forms, namely

$$\mathcal{H} = \int a \wedge \mathrm{d}a, \quad \text{(A 6)}$$

where $a = A^\flat$ is the 1-form potential for B. This integrand is a baby version of the famous Chern–Simons form on 3-manifolds, where the concept of differential form is extended to take values in a Lie algebra instead of just $\mathbb{R}$ (Arnol'd & Khesin 1998).

A.6. *Lie derivative on general tensors*

A tensor T of type (p, q) is a multilinear map from the space of p covectors and q vectors at a point to $\mathbb{R}$. Its Lie derivative along a vector field Y can be computed by

$$
\begin{aligned}
(L_Y T)(\alpha_1, \ldots, \alpha_p, X_1, \ldots X_q) = {}& Y(T(\alpha_1, \ldots, \alpha_p, X_1, \ldots X_q)) \\
& - T(L_Y \alpha_1, \alpha_2, \ldots X_q) - T(\alpha_1, L_Y \alpha_2, \ldots, X_q) - \cdots \\
& - T(\alpha_1, \ldots \alpha_p, L_Y X_1, X_2, \ldots X_q) - \cdots - T(\alpha_1, \ldots X_{q-1}, L_Y X_q),
\end{aligned}
\tag{A 7}
$$

for any 1-forms α_i and vector fields X_j.

REFERENCES

ARNOL'D, V. I. 1978 *Mathematical Methods of Classical Mechanics*. Springer.

ARNOL'D, V. I. & KHESIN, B. A. 1998 *Topological Methods in Hydrodynamics*. Springer.

BURBY, J. W., KALLINIKOS, N. & MACKAY, R. S. Some mathematics for quasi-symmetry. arXiv:1912.06468.

CARY, J. R. & LITTLEJOHN, R. G. 1983 Noncanonical Hamiltonian mechanics and its application to magnetic field line flow. *Ann. Phys.* **151**, 1–34.

DESCHAMPS, G. A. 1981 Electromagnetics and differential forms. *Proc. IEEE* **69** (6), 676–687.

FRIED, D. 1982 The geometry of cross sections to flows. *Topology* **21**, 353–371.

HELANDER, P. 2014 Theory of plasma confinement in non-axisymmetric magnetic fields. *Rep. Prog. Phys.* **77**, 087001.

LITTLEJOHN, R. G. 1983 Variational principles of guiding centre motion. *J. Plasma Phys.* **29**, 111–125.

MACKAY, R. S. 1992 Some aspects of the dynamics and numerics of Hamiltonian systems. In *The Dynamics of Numerics and Numerics of Dynamics* (ed. D. S. Broomhead & A. Iserles), IMA Conf. Proc., pp. 137–193. Oxford.

MACKAY, R. S. & PINHEIRO, D. 2006 Interaction of two charges in a uniform magnetic field: I. Planar problem. *Nonlinearity* **19**, 1713–1745.

MACKAY, R. S. & PINHEIRO, D. 2008 Interaction of two charges in a uniform magnetic field: II spatial case. *J. Nonlinear Sci.* **18**, 615–666.

MISNER, C. W., THORNE, K. S. & WHEELER, J. A. 1973 *Gravitation*. Freeman.

QIN, H., COHEN, R. H., NEVINS, W. M. & XU, X. Q. 2007 Geometric gyrokinetic theory for edge plasmas. *Phys. Plasmas* **14**, 056110.

WARNICK, K. F. & RUSSER, P. H. 2014 Differential forms and electromagnetic field theory. *Prog. Electromagnetic Res.* **148**, 83–112.