ILLiad Request Printout

Transaction Number: 180584
Username: StuartHudson Name: Stuart Hudson Status: F - Faculty
ISSN/ISBN: 0375-9601
NotWantedAfter: 07/26/2005
Accept Non English: No
Accept Alternate Edition: No
Request Type: Article - ArticleExpress

Loan Information

LoanAuthor:
LoanTitle:
LoanPublisher:
LoanPlace:
LoanDate:
LoanEdition:
NotWantedAfter: 07/26/2005

Article Information

PhotoJournalTitle: Physics Letters A
PhotoJournalVolume: 194
PhotoJournalIssue:
Month:
Year: 1994
Pages: 49
Article Author: R.L.Dewar
Article Title: Almost invariant manifolds for divergence free fields.

Citation Information

Cited In:
Cited Title:
Cited Date:
Cited Volume:
Cited Pages:

OCLC Information

ILL Number:
OCLC Number:
Lending String:
Original Loan Author:
Original Loan Title:
Old Journal Title:
Call Number: Electronic Resource
Location: *ONLINE*

Notes

1/27/2005 10:04:18 AM jes Accession number: 4822346

ELSEVIER

24 October 1994

PHYSICS LETTERS A

Physics Letters A 194 (1994) 49–56

Almost invariant manifolds for divergence-free fields

R.L. Dewar [a], S.R. Hudson [a,1], P.F. Price [b,2]

[a] *Department of Theoretical Physics and Plasma Research Laboratory, Research School of Physical Sciences and Engineering, The Australian National University, Canberra, ACT 0200, Australia*

[b] *Computer Science Laboratory, Research School of Information Sciences and Engineering, The Australian National University, Canberra, ACT 0200, Australia*

Received 16 August 1994; accepted for publication 17 August 1994
Communicated by A.R. Bishop

Abstract

It is shown that toroidal surfaces that extremize a properly weighted surface integral of the squared normal component of a solenoidal three-vector field capture the local invariant dynamics, in that a field line that is anywhere tangential to the surface must be confined to the surface everywhere. In addition to an elementary three-vector calculus derivation, which relies on a curvilinear toroidal coordinate system, a coordinate-free geometric approach applicable to hypersurfaces (codimension-one submanifolds) of manifolds of arbitrary dimension is sketched.

1. Introduction

In order to compute accurately the equilibrium, stability and transport properties of nonaxisymmetric toroidal plasma containment devices such as stellarators, or tokamaks with field ripple, it is important to work in a curvilinear coordinate system such that the magnetic field lines are as close as possible to being tangent everywhere to the nested toroidal level surfaces of one of the coordinates. An exactly similar problem is that of constructing streamline coordinates for nonaxisymmetric, stationary vortex flows in incompressible fluids. If a surface can be found such that the velocity field is everywhere tangential, then the surface is *invariant* under the flow. If no suitable invariant surface can be found, then we seek an *almost invariant* surface. It is the purpose of this paper to present a definition of "almost invariant" which has the property of "capturing" sets that are invariant under the flow.

It can be shown [1–3] that, given an arbitrary curvilinear toroidal coordinate system, a "time"-dependent, one-degree-of-freedom Hamiltonian can always be constructed whose phase-space trajectories correspond to the field lines of an arbitrary magnetic field, with the generalized toroidal angle playing the role of "time". Thus the task of finding almost invariant surfaces may also be viewed as a Hamiltonian dynamics problem, though this paper will present the problem mainly in terms of magnetic fields.

An integrable time-periodic Hamiltonian is one whose phase-space trajectories all lie on invariant tori (regarding time as a toroidal angle) so the dynamics can be described using action–angle variables [4]. For a toroidal magnetic confinement system, integrability is equivalent to the existence of nested mag-

[1] Supported by an Australian National University Postgraduate Scholarship.
[2] Present address: Department of Defence, PO Box E33, Queen Victoria Terrace, Canberra, ACT 0200, Australia.

SSDI 0375-9601(94)00707-1

netic flux surfaces described by magnetic coordinate systems (see, e.g., Ref. [5]).

However, when such a system is perturbed from integrability, an attempt to reconstruct action–angle variables runs into the small resonant denominator problem of canonical perturbation theory [6]. No longer can one assume the existence of nested flux surfaces and the validity of magnetic coordinate systems is not guaranteed. Some comfort can be taken in the Kolmogorov–Arnol'd–Moser (KAM) theorem, which states that, under sufficiently small, smooth perturbations, some invariant tori will remain. However other tori with irrational rotation number (rotational transform in magnetic confinement parlance) will break up into invariant Cantor sets, and those with rational rotation number evaporate into islands and chaotic trajectories, leaving only a finite set of periodic orbits (closed field lines) as the invariant remnants. Thus, although the assumption of nested flux surfaces may be a good approximation in appropriate circumstances, a more generally satisfactory theory of magnetic coordinates must face up to the generic situation of nonintegrability.

An analogous problem occurs in the theory of area-preserving twist maps. Several approaches to defining almost invariant rotational curves have been proposed [7–10], with that of Meiss and Dewar [7,8] being the basis for the approach adopted in this paper. Meiss and Dewar define an almost invariant surface as one that extremizes the "quadratic flux", being the mean square vertical distance Δy between a trial curve C, say, and its image TC under the area-preserving twist diffeomorphism $T: \ (x,y) \mapsto (x',y')$. They show [7] from the Euler–Lagrange equation for extrema of the quadratic flux that, if C and TC are graphs over the coordinate x, then intersections of C and TC (where $\Delta y = 0$) correspond to orbits under the area preserving map. That is, if a point (x,y) is on both C and TC, then all its backward and forward iterates are on both C and TC. Since such orbits are invariant sets, this principle has the desirable property of capturing remnant local invariant dynamics, including KAM surfaces if the orbit is dense on C. The result can also be extended [8] to the case where C and TC are not graphs, where there is ambiguity in the meaning of "vertical distance", but this complication has no counterpart in the continuous-time case which is the subject of this paper.

The net flux, $\mathcal{F}$, through a surface Γ (taken to be a two-torus in $\mathbb{R}^3$) is defined by

$$\mathcal{F} \equiv \int_\Gamma \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S}, \tag{1}$$

where $\mathrm{d}\boldsymbol{S}$ denotes the vector surface element $\boldsymbol{n}\,\mathrm{d}S$, with $\boldsymbol{n}$ being the outward unit normal at a point on the surface Γ, and $\mathrm{d}S$ an element of surface area. In the absence of sources (magnetic monopoles) within the toroidal volume enclosed by Γ, $\mathcal{F}$ is identically zero. That is, the escape flux carried by field lines leaving the volume is exactly balanced by the return flux carried by field lines entering the volume. Thus we can define the escape flux, φ_1, as

$$\varphi_1 \equiv \tfrac{1}{2} \int_\Gamma |B_n|\,\mathrm{d}S, \tag{2}$$

where $B_n \equiv \boldsymbol{B} \cdot \boldsymbol{n}$.

Consider a deformation of Γ in a region where B_n is of one sign only, say positive, with the rest of Γ held fixed. Then the return flux, from regions where $B_n < 0$, cannot change, and therefore the escape flux cannot change either since they must balance. Thus φ_1 is left invariant by a wide class of variations of Γ and therefore does *not* form an appropriate objective functional for optimizing the choice of Γ by flux minimization. We are thus led, following Meiss and Dewar [7,8], to considering a second moment of $|B_n|$.

We define the *magnetic quadratic flux*, φ_2, by

$$\varphi_2 \equiv \tfrac{1}{2} \int_\Gamma w|B_n|^2\,\mathrm{d}S, \tag{3}$$

where w is a positive weight function whose choice will be found (see Eq. (16)) to be restricted by consideration of the Euler–Lagrange equation for surfaces that extremize φ_2. Note that the magnetic energy associated with the normal component of the magnetic field in a shell of infinitesimal thickness $w\,\mathrm{d}\rho$ (ρ being a generalized radial variable with appropriate dimensions) bounded by Γ is $\varphi_2\,\mathrm{d}n$, so φ_2 might perhaps be better thought of as a line density of the energy in the radial magnetic field rather than a flux.

We investigate the variational calculus of this functional using two methods. The first approach, somewhat more traditional from a plasma physics perspective, is to use a curvilinear coordinate system. This is developed in Section 2 where we also restrict the choice of the weight function by considering the existence of solutions of the Euler–Lagrange equation, and demonstrate the desirable property of solutions that "capture" invariant dynamics.

The second, more general, way developed in Section 3 is based on modern differential geometric techniques. Both methods derive the same result for the setting of Γ, a surface in ordinary three-space, but the second method actually is valid in the more general setting of Γ, a hypersurface of an arbitrary Riemannian manifold.

2. Three-vector analysis

In this approach we take the toroidal surface Γ to be a member of a family of nested tori labelled by a continuous parameter s. On each torus we set up generalized poloidal and toroidal angles θ and ζ, respectively, so that any point in space, $\boldsymbol{x}$, is given by $\boldsymbol{x} = \boldsymbol{r}(s,\theta,\zeta)$ and Γ is the surface s = const. Then the contravariant representation of any vector $\boldsymbol{A}$ is $A^s\boldsymbol{e}_s + A^\theta\boldsymbol{e}_\theta + A^\zeta\boldsymbol{e}_\zeta$, where $\boldsymbol{e}_s \equiv \partial_s\boldsymbol{r}$, $\boldsymbol{e}_\theta \equiv \partial_\theta\boldsymbol{r}$ and $\boldsymbol{e}_\zeta \equiv \partial_\zeta\boldsymbol{r}$, while the covariant representation is $A_s\nabla s + A_\theta\nabla\theta + A_\zeta\nabla\zeta$. The unit normal $\boldsymbol{n}$ is given by

$$\boldsymbol{n} = \frac{\nabla s}{|\nabla s|} = \frac{\boldsymbol{e}_\theta \times \boldsymbol{e}_\zeta}{|\boldsymbol{e}_\theta \times \boldsymbol{e}_\zeta|}, \tag{4}$$

and the vector surface area element by

$$\mathrm{d}\boldsymbol{S} = \boldsymbol{e}_\theta \times \boldsymbol{e}_\zeta \,\mathrm{d}\theta\,\mathrm{d}\zeta = \mathcal{J}\nabla s\,\mathrm{d}\theta\,\mathrm{d}\zeta, \tag{5}$$

where the Jacobian $\mathcal{J}$ is defined by

$$\mathcal{J} = (\nabla s\cdot\nabla\theta\times\nabla\zeta)^{-1} = \boldsymbol{e}_s\cdot\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta.$$

We now consider variations in the surface Γ, holding the spatial dependence of the vector field $\boldsymbol{B}(\boldsymbol{x})$ fixed. In the inverse representation in terms of (s,θ,ζ), we vary the function $\boldsymbol{r}(s,\theta,\zeta)$, so that the variation in $\boldsymbol{n}$ is given by

$$\delta\boldsymbol{n} = \frac{(\partial_\theta\delta\boldsymbol{r})\times\boldsymbol{e}_\zeta + \boldsymbol{e}_\theta\times(\partial_\zeta\delta\boldsymbol{r})}{|\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta|} - \frac{\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta\delta|\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta|}{|\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta|^2}, \tag{6}$$

and the variation in $|\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta| = (|\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta|^2)^{1/2}$ is given by

$$\delta|\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta| = [(\partial_\theta\delta\boldsymbol{r})\times\boldsymbol{e}_\zeta + \boldsymbol{e}_\theta\times(\partial_\zeta\delta\boldsymbol{r})]\cdot(\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta)/|\boldsymbol{e}_\theta\times\boldsymbol{e}_\zeta|. \tag{7}$$

Using Eqs. (5)–(7) and integrating by parts with respect to θ and ζ we can now prove the lemmas, for arbitrary scalar and vector point functions f and $\boldsymbol{f}$, respectively,

$$\int_\Gamma f\delta\,\mathrm{d}S = -\int_\Gamma \mathrm{d}S\,\delta\boldsymbol{r}\cdot(\boldsymbol{n}\times\nabla)\times(f\boldsymbol{n}), \tag{8}$$

and

$$\int_\Gamma \mathrm{d}S\,\boldsymbol{f}\cdot\delta\boldsymbol{n} = -\int_\Gamma \mathrm{d}S\,\delta\boldsymbol{r}\cdot(\boldsymbol{n}\times\nabla)\times\boldsymbol{f}_s, \tag{9}$$

where $\boldsymbol{f}_s$ is the projection of $\boldsymbol{f}$ in the tangent plane to Γ at the point $\boldsymbol{x} = \boldsymbol{r}(s,\theta,\zeta)$,

$$\boldsymbol{f}_s \equiv (\mathbf{I} - \boldsymbol{nn})\cdot\boldsymbol{f}. \tag{10}$$

Here $\mathbf{I}$ is the idemfactor or unit tensor. (See Fig. 1 for the case $\boldsymbol{f} = \boldsymbol{B}$.) Note that, although we used a curvilinear coordinate system to derive Eqs. (8) and (9), the equations themselves do not involve coordinates, and allow us to calculate the variation of φ_2 in a coordinate-free manner.

It is actually easiest to consider first the more general functional

$$\varphi = \int_\Gamma f(\boldsymbol{x},\boldsymbol{n})\;\mathrm{d}S. \tag{11}$$

On using the two lemmas above and the identities $(\nabla\boldsymbol{n})\cdot\boldsymbol{n} \equiv 0$, $(\nabla\boldsymbol{f}_s)\cdot\boldsymbol{n} \equiv -(\nabla\boldsymbol{n})\cdot\boldsymbol{f}$ and $\nabla f = \partial f/\partial\boldsymbol{x} + (\nabla\boldsymbol{n})\cdot(\partial f/\partial\boldsymbol{n})$ we find the first variation

$$\delta\varphi = \int_\Gamma \mathrm{d}S\,\boldsymbol{n}\cdot\delta\boldsymbol{r}\nabla\cdot\left[\left(\frac{\partial f}{\partial\boldsymbol{n}}\right)_s + f\boldsymbol{n}\right]. \tag{12}$$

This depends only on $\boldsymbol{n}\cdot\delta\boldsymbol{r}$, as it should since variations of $\boldsymbol{r}$ within the surface correspond to relabelling of the θ and ζ lines and leave the actual surface invariant. The Euler–Lagrange equation which makes φ stationary is

$$\nabla\cdot\left[\left(\frac{\partial f}{\partial\boldsymbol{n}}\right)_s + f\boldsymbol{n}\right] = 0, \tag{13}$$

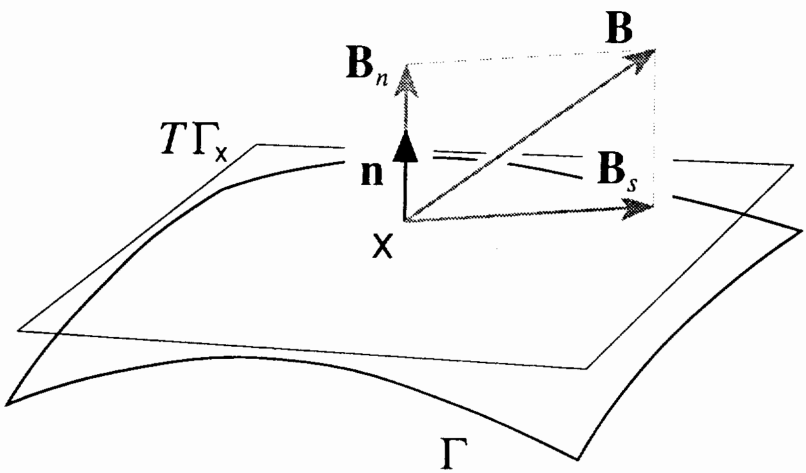

Fig. 1. Illustration of the decomposition of $\boldsymbol{B}$ at the point x on the surface Γ into a component $\boldsymbol{B}_n$ in the subspace spanned by the unit normal, $\boldsymbol{n}$ and the projection $\boldsymbol{B}_s$ in the tangent plane, $T\Gamma_x$. In this section $x = \boldsymbol{r}(s,\theta,\zeta)$, but the figure also illustrates schematically the decomposition used for higher dimensional manifolds.

where $(\ )_s$ denotes the surface projection defined by Eq. (10).

If we take $f = \frac{1}{2}B_n^2$, corresponding to choosing $w = 1$ in Eq. (3), Eq. (13) gives, on using $\nabla\cdot\boldsymbol{B} = 0$,

$$\boldsymbol{B}_s\cdot\nabla B_n = \tfrac{1}{2}B_n^2\nabla\cdot\boldsymbol{n}. \tag{14}$$

On attempting to solve Eq. (14) by the method of characteristics, i.e. by integrating along a "pseudo field line" defined as a curve on Γ such that $\boldsymbol{B}_s$ is everywhere tangential to it, we immediately observe a fatal problem – the right-hand side is $\geqslant 0$ (except perhaps for tori with very small aspect ratio), so that B_n must increase monotonically along the pseudo field line. Thus, if we start at a point where B_n is positive, it will blow up as we integrate further along the line, while in the case of negative B_n blowup will occur in the reverse direction along the line. We must therefore conclude that no stationary surfaces exist in general for the quadratic flux defined with unit weight function.

One possibility for rectifying this situation might be to use as w an $\boldsymbol{x}$-dependent scalar (e.g. B^2), but this does not have the desired property of leading to a driving term in the Euler–Lagrange equation that vanishes or manifestly satisfies the solubility condition [11] for inverting the $\boldsymbol{B}_s\cdot\nabla$ operator. Instead we are led to introducing an auxiliary vector field $\boldsymbol{C}$, say, obeying the conditions

$$\nabla\cdot\boldsymbol{C} = 0, \quad C_n \equiv \boldsymbol{n}\cdot\boldsymbol{C} > 0 \tag{15}$$

at least in a neighbourhood of the surface making φ_2 stationary. The choice of $\boldsymbol{C}$ will be discussed later.

We now take as weight function in Eq. (3)

$$w = C_n^{-1}. \tag{16}$$

Inserting $f = B_n^2/2C_n$ in Eq. (13) and using the divergence-free properties of $\boldsymbol{B}$ and $\boldsymbol{C}$, we find the Euler–Lagrange equation

$$\left(\boldsymbol{B}_s - \frac{B_n}{C_n}\boldsymbol{C}_s\right)\cdot\nabla\left(\frac{B_n}{C_n}\right) = 0. \tag{17}$$

Since the right-hand side now vanishes, there is now no problem with solvability of the partial differential equation. Indeed we can immediately solve it by integrating along the characteristic pseudo field lines defined by the dynamical system on Γ,

$$\dot{\boldsymbol{x}} = \boldsymbol{B}_s - \frac{B_n}{C_n}\boldsymbol{C}_s, \tag{18}$$

with $\dot{\boldsymbol{x}}$ being the derivative of $\boldsymbol{x}$ with respect to a "time" variable with dimensions of length divided by magnetic field. We assume that the vector field on the right-hand side of Eq. (18) never vanishes, so that there are no fixed points and the characteristics cannot cross. The solution of Eq. (17) is just

$$\frac{B_n}{C_n} = \text{const} \tag{19}$$

on a pseudo field line defined by Eq. (18).

Note that, if $B_n = 0$ at some point, then it will be zero on all points of the pseudo field line passing through that point. Furthermore, this "pseudo" field line will in fact be a *real* magnetic field line (orbit of the dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{B}$). This is the direct analogue of Theorem 1 of Dewar and Meiss [8]. An immediate corollary is the *invariance capture property*: The quadratic-flux-extremizing surfaces with $w = 1/C_n$ "stick to" any invariant structures such as closed field lines (periodic orbits of $\dot{\boldsymbol{x}} = \boldsymbol{B}$) and magnetic surfaces (KAM tori – quasiperiodic orbits of $\dot{\boldsymbol{x}} = \boldsymbol{B}$ covering a surface ergodically) that they touch.

Consider the case of a quadratic-flux-extremizing surface containing a hyperbolic closed field line X corresponding to the X-point of a magnetic island in the return map defined on the Poincaré section $\zeta =$ const. Now follow the sign of B_n as we move along a curve on Γ cutting the pseudo field lines transversely and intersecting X twice. Since $B_n = 0$ on X, and Eqs. (15) and (19) show that the sign of B_n is conserved on a pseudo field line, we see that the sign of B_n must change in the same sense on both crossings of X. However, this means that B_n must also change sign *somewhere in between* the crossings of X. Hence there must be at least one other closed field line O, say, interleaved with X. Assuming that the island whose (chaotic) separatrix is defined by the stable and unstable manifolds of X contains a single elliptic closed field line (O-point in the Poincaré section) we identify O with this periodic orbit. An exactly analogous result was also found in the area-preserving map case [7,8].

The auxiliary divergence-free field $\boldsymbol{C}$ is somewhat arbitrary, but if we are given poloidal and toroidal angles $\theta(\boldsymbol{x})$ and $\zeta(\boldsymbol{x})$ then a natural choice is $\boldsymbol{C} = \nabla\theta \times \nabla\zeta$, so that $\mathrm{d}S = \mathrm{d}\theta\,\mathrm{d}\zeta/C_n$. It is then appropriate to ask whether the choice of θ and ζ can be restricted by requiring that φ_2 be stationary under variations in θ and ζ. Varying θ and ζ in Eq. (3) with $w = 1/C_n$ while holding Γ fixed, we find

$$\delta\varphi_2 = \int_0^{2\pi}\!\!\int_0^{2\pi} \mathrm{d}\theta\,\mathrm{d}\zeta\,\frac{B_n}{C_n}\left(\delta\theta\frac{\partial}{\partial\theta} + \delta\zeta\frac{\partial}{\partial\zeta}\right)\frac{B_n}{C_n}. \tag{20}$$

Thus φ_2 cannot be stationary under variations of the angles unless B_n/C_n is constant on Γ, which is impossible (unless $B_n \equiv 0$, in which case $\varphi_2 \equiv 0$). Thus we conclude that the choice of angles must be made on grounds other than pure quadratic flux minimization. Other criteria might be numerical convenience, the spectral condensation optimization of Hirshman and Meier [12], or physical and mathematical grounds such as those leading to Boozer coordinates [13].

3. General geometrical method

We now indicate how to generalize the above result to manifolds of arbitrary dimension in an elegant, coordinate free way by using the methods of modern differential geometry. Thus we consider an m-dimensional hypersurface, Γ, of a general $(m+1)$-dimensional Riemannian manifold, N. (For the plasma containment problem we take m equal to 2 and N simply a flat three-space.) We also require a vector field $\boldsymbol{B}$ to be defined in a neighbourhood (at least) of Γ within N and to be divergence free, i.e. $\nabla\cdot\boldsymbol{B} = 0$.

An example of $\boldsymbol{B}$ in higher dimension might be the vector field of a nondissipative dynamical system, specifically a nonautonomous Hamiltonian system, with N being the phase space and time, and with a suitable metric chosen to provide the Riemannian structure required for the present theory. Indeed the magnetic field problem of the previous section provides an example if we treat it using the magnetic field-line Hamiltonian approach [1–3]. Then the manifold N has the structure of a two-torus (with coordinates being a poloidal angle as generalized coordinate and a toroidal angle as "time") crossed with the positive real line (corresponding to the toroidal flux function as generalized momentum). In this case there is a natural metric, that giving the length in physical three-space (although even here there is a certain arbitrariness due to the necessity to introduce the auxiliary field $\boldsymbol{C}$).

The hypersurface Γ inherits a Riemannian structure from that of N and the unit normal $\boldsymbol{n}$ and volume element $\mathrm{d}S$ are well defined in this structure. Thus the quadratic flux φ_2 is still well-defined by Eq. (3) above.

We wish to derive the conditions (Euler–Lagrange

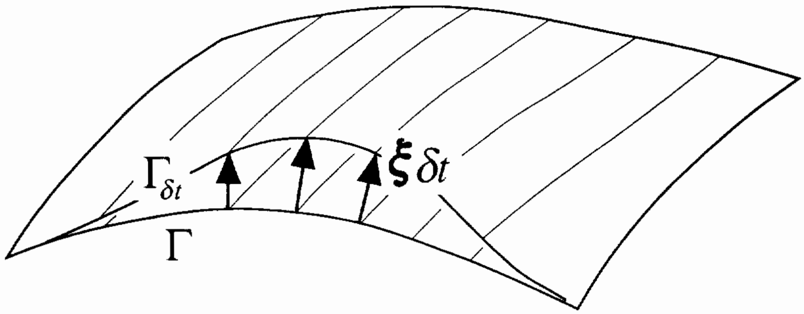

Fig. 2. Infinitesimal variation in surface Γ produced by the vector field $\boldsymbol{\xi}$ during the interval in the parameter t, 0 to $0+\delta t$. It is seen that $\boldsymbol{\xi}\delta t$ forms the generalization of the variation $\delta \boldsymbol{r}$ used in the three-vector formulation.

equations) under which φ_2 will be stationary for all smooth one-parameter families of (compactly supported) variations of Γ. We can generate such a family, Γ_t, by acting on Γ with the local flow, $\Phi_t \in N$, produced by a vector field $\boldsymbol{\xi}(\mathrm{x})$ defined for all points $\mathrm{x} \in N$ in a neigbourhood of Γ. That is, $\Gamma_t = \Phi_t(\Gamma)$, with $\partial_t \Phi_t(\mathrm{x}) = \boldsymbol{\xi}(\Phi_t(\mathrm{x}))$ and with $\Phi_0(q\mathrm{x}) = \mathrm{x}$, as illustrated in Fig. 2. The vector field $\boldsymbol{\xi}$ is in the tangent space of N but we require the transversality condition that it be not (at least not entirely) in the tangent space of Γ when x is on Γ. We can now define a one-parameter family of quadratic fluxes

$$\varphi_2(t) \equiv \tfrac{1}{2} \int_{\Gamma_t} B_n^2 \, \mathrm{d}S_t, \tag{21}$$

so that the stationarity condition is $\varphi_2'(0) = 0$. (We have taken the weight function to be unity for simplicity, though the same reasons for needing to generalize this as were found in Section 2 will be found in the general case as well.)

It is intuitively clear (and a standard result in the area of geometric variations of geometric integrals, see, e.g., Ref. [14]) that the first variation $\varphi_2'(0)$ will only depend on the vector field $\boldsymbol{\xi}$ along Γ and not on the extension of $\boldsymbol{\xi}$ in a neighbourhood of Γ, i.e. only on the tangent vectors to the curves $\Phi_t(\mathrm{x})$ at $t = 0$, $\mathrm{x} \in \Gamma$. It is also standard that the variation vector field $\boldsymbol{\xi}$ can be taken to be normal to Γ. This is because any tangential component simply results in a reparametrization of Γ_t and leaves the quadratic flux φ_2 invariant.

Thus

$$\begin{aligned} \varphi_2'(0) &= \tfrac{1}{2}\left(\int_\Gamma \partial_t(B_n^2)\,\mathrm{d}S + \int_\Gamma B_n^2 \partial_t(\mathrm{d}S) \right)_{t=0} \\ &= \int_\Gamma \boldsymbol{B}_n \cdot \nabla_{\boldsymbol{\xi}} \boldsymbol{B}_n \, \mathrm{d}S - \tfrac{1}{2} \int_\Gamma B_n^2 \boldsymbol{H} \cdot \boldsymbol{\xi} \, \mathrm{d}S, \end{aligned} \tag{22}$$

where $\boldsymbol{B}_n \equiv B_n \boldsymbol{n}$ is the projection of $\boldsymbol{B}$ in the normal direction (so $B_n^2 = \boldsymbol{B}_n \cdot \boldsymbol{B}_n$) and $\boldsymbol{H}$ is the mean curvature of the hypersurface Γ in N, related to the unit normal $\boldsymbol{n}$ by

$$\boldsymbol{H} = -(\boldsymbol{\nabla} \cdot \boldsymbol{n})\boldsymbol{n}. \tag{23}$$

Here we have extended $\boldsymbol{n}$ into a neighbourhood of Γ in N by, for example, defining $\boldsymbol{n}(\mathrm{x})$ to be the unit normal to Γ_t with t chosen so x is on Γ_t.

The fact that the derivative of the surface element is given by $-(\boldsymbol{H} \cdot \boldsymbol{\xi})\,\mathrm{d}S$ is a standard result of the geometric calculus of variations. It is, for example, the starting point for the theory of "minimal surfaces". The first term in Eq. (22) is a statement of one of the defining relationships of Riemannian geometry, namely that

$$\partial_X(\boldsymbol{Y} \cdot \boldsymbol{Z}) = (\nabla_X \boldsymbol{Y}) \cdot \boldsymbol{Z} + \boldsymbol{Y} \cdot (\nabla_X \boldsymbol{Z}), \tag{24}$$

for all vector fields $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ in N. Here ∇_X means the covariant derivative in the direction $\boldsymbol{X}$ (in $\mathbb{R}^n$ it is simply $\boldsymbol{X} \cdot \boldsymbol{\nabla}$) and the directional derivative ∂_X is defined by

$$(\partial_X f)|_{\mathrm{x}_0} = \left. \frac{\mathrm{d}f(\mathrm{x}(t))}{\mathrm{d}t} \right|_{t=0}, \tag{25}$$

where $\mathrm{x}(t)$ is any smooth curve whose tangent vector $\mathrm{x}'(0)$ at $\mathrm{x}(0) = \mathrm{x}_0$ is $\boldsymbol{X}(\mathrm{x}_0)$. Again, in $\mathbb{R}^n$, ∂_X is simply $\boldsymbol{X}\cdot\boldsymbol{\nabla}$.

Since $\boldsymbol{B}_n = (\boldsymbol{B}_n\cdot\boldsymbol{n})\boldsymbol{n}$ and $\boldsymbol{\xi} = (\boldsymbol{\xi}_n\cdot\boldsymbol{n})\boldsymbol{n}$,

$$\boldsymbol{B}_n\cdot\nabla_{\boldsymbol{\xi}}\boldsymbol{B}_n = (\boldsymbol{B}_n\cdot\boldsymbol{n})(\boldsymbol{\xi}\cdot\boldsymbol{n})\boldsymbol{n}\cdot\nabla_{\boldsymbol{n}}\boldsymbol{B}_n. \tag{26}$$

Using $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$ we calculate $\boldsymbol{n}\cdot\nabla_{\boldsymbol{n}}\boldsymbol{B}_n$. We calculate using a locally defined set of orthogonal unit vectors $(\{\boldsymbol{e}_i\}_{i=1,\dots,m}, \boldsymbol{n})$, where the $\{\boldsymbol{e}_i\}_{i=1,\dots,m}$ are tangential to Γ. The tangential and normal components of $\boldsymbol{B}$ are denoted $\boldsymbol{B}_s$ and $\boldsymbol{B}_n$, respectively, and are given by $\boldsymbol{B}_s = (\boldsymbol{B}\cdot\boldsymbol{e}_i)\boldsymbol{e}_i$ and $\boldsymbol{B}_n = (\boldsymbol{B}\cdot\boldsymbol{n})\boldsymbol{n}$. Then $\boldsymbol{\nabla}\cdot\dots = \boldsymbol{e}_i\cdot(\nabla_{\boldsymbol{e}_i}\dots) + \boldsymbol{n}\cdot(\nabla_{\boldsymbol{n}}\dots)$ (summation convention).

Using these expressions we have (strongly using the orthogonality of the basis vectors)

$$\begin{aligned}\boldsymbol{\nabla}\cdot\boldsymbol{B} &= \boldsymbol{e}_i\cdot\nabla_{\boldsymbol{e}_i}\boldsymbol{B}_s + (\boldsymbol{B}_n\cdot\boldsymbol{n})\boldsymbol{e}_i\cdot\nabla_{\boldsymbol{e}_i}\boldsymbol{n}\\ &\quad + (\boldsymbol{B}\cdot\boldsymbol{e}_i)\boldsymbol{n}\cdot\nabla_{\boldsymbol{n}}\boldsymbol{e}_i + \boldsymbol{n}\cdot\nabla_{\boldsymbol{n}}\boldsymbol{B}_n.\end{aligned} \tag{27}$$

Then, using Eq. (23),

$$\begin{aligned}\boldsymbol{n}\cdot\nabla_{\boldsymbol{n}}\boldsymbol{B}_n &= -\boldsymbol{e}_i\cdot\nabla_{\boldsymbol{e}_i}\boldsymbol{B}_s + (\boldsymbol{B}_n\cdot\boldsymbol{n})\boldsymbol{H}\cdot\boldsymbol{n}\\ &\quad - (\boldsymbol{B}\cdot\boldsymbol{e}_i)\boldsymbol{n}\cdot\nabla_{\boldsymbol{n}}\boldsymbol{e}_i.\end{aligned} \tag{28}$$

Now

$$\boldsymbol{n}\cdot(\nabla_{\boldsymbol{n}}\boldsymbol{e}_i) = (\boldsymbol{\xi}\cdot\boldsymbol{n})^{-1}\boldsymbol{n}\cdot\nabla_{\boldsymbol{\xi}}\boldsymbol{e}_i = (\boldsymbol{\xi}\cdot\boldsymbol{n})^{-1}\boldsymbol{n}\cdot(\nabla_{\boldsymbol{e}_i}\boldsymbol{\xi}), \tag{29}$$

where, for computational convenience, we have extended the tangential basis $\{\boldsymbol{e}_i\}_{i=1,\dots,m}$ so that they remain tangential to Γ_t for small t. Then $\nabla_{\boldsymbol{\xi}}\boldsymbol{e}_i - \nabla_{\boldsymbol{e}_i}\boldsymbol{\xi} = [\boldsymbol{\xi},\boldsymbol{e}_i]$ is tangential.

Putting together Eqs. (26)–(28) we have

$$\begin{aligned}\boldsymbol{B}_n\cdot\nabla_{\boldsymbol{\xi}}\boldsymbol{B}_n &= (\boldsymbol{B}_n\cdot\boldsymbol{n})(\boldsymbol{\xi}\cdot\boldsymbol{n})(\boldsymbol{B}_n\cdot\boldsymbol{n})\boldsymbol{H}\cdot\boldsymbol{n}\\ &\quad - (\boldsymbol{B}_n\cdot\boldsymbol{n})(\boldsymbol{\xi}\cdot\boldsymbol{n})(\boldsymbol{B}_n\cdot\boldsymbol{e}_i)(\boldsymbol{\xi}\cdot\boldsymbol{n})^{-1}\boldsymbol{n}\cdot\nabla_{\boldsymbol{e}_i}\boldsymbol{\xi}\\ &\quad - (\boldsymbol{B}_n\cdot\boldsymbol{n})(\boldsymbol{\xi}\cdot\boldsymbol{n})\boldsymbol{e}_i\cdot\nabla_{\boldsymbol{e}_i}\boldsymbol{B}_s\\ &= |\boldsymbol{B}_n|^2\boldsymbol{H}\cdot\boldsymbol{\xi} - (\nabla_{\boldsymbol{B}_s}\boldsymbol{\xi})\cdot\boldsymbol{B}_n - (\boldsymbol{e}_i\cdot\nabla_{\boldsymbol{e}_i}\boldsymbol{B}_s)(\boldsymbol{B}_n\cdot\boldsymbol{\xi}).\end{aligned} \tag{30}$$

Noting that

$$\begin{aligned}\boldsymbol{e}_i\cdot\nabla_{\boldsymbol{e}_i}(\boldsymbol{\xi}\cdot\boldsymbol{B}_n)\boldsymbol{B}_s &= (\boldsymbol{\xi}\cdot\boldsymbol{B}_n)\boldsymbol{e}_i\cdot\nabla_{\boldsymbol{e}_i}\boldsymbol{B}_s\\ &\quad + (\nabla_{\boldsymbol{B}_s}\boldsymbol{\xi})\cdot\boldsymbol{B}_n + \boldsymbol{\xi}\cdot\nabla_{\boldsymbol{B}_s}\boldsymbol{B}_n\end{aligned} \tag{31}$$

we have

$$\begin{aligned}\boldsymbol{B}_n\cdot\nabla_{\boldsymbol{\xi}}\boldsymbol{B}_n &= |\boldsymbol{B}_n|^2\boldsymbol{H}\cdot\boldsymbol{\xi} + (\nabla_{\boldsymbol{B}_s}\boldsymbol{B}_n)\cdot\boldsymbol{\xi}\\ &\quad - \boldsymbol{e}_i\cdot\nabla_{\boldsymbol{e}_i}(\boldsymbol{\xi}\cdot\boldsymbol{B}_n)\boldsymbol{B}_s.\end{aligned} \tag{32}$$

Putting this into Eq. (22) we get

$$\varphi_2'(0) = \int_{\Gamma} (\tfrac{1}{2}|\boldsymbol{B}_n|^2\boldsymbol{H} + \nabla_{\boldsymbol{B}_s}\boldsymbol{B}_n)\cdot\boldsymbol{\xi}\,\mathrm{d}S, \tag{33}$$

since the third term on the right-hand side of Eq. (32) is the divergence on Γ of $(\boldsymbol{\xi}\cdot\boldsymbol{B}_n)\boldsymbol{B}_s$ and integrates to zero, by the divergence theorem, as $\boldsymbol{\xi}$ is of compact support.

For stationarity of φ_2, the right-hand side of Eq. (33) must vanish for arbitrary normal $\boldsymbol{\xi}$, and this implies the Euler–Lagrange equation

$$(\nabla_{\boldsymbol{B}_s}\boldsymbol{B}_n + \tfrac{1}{2}|\boldsymbol{B}_n|^2\boldsymbol{H})\cdot\boldsymbol{n} = 0. \tag{34}$$

Eq. (34) represents the generalization to hypersurfaces of arbitrary dimension of the result Eq. (14) derived for two-tori in $\mathbb{R}^3$. The weight function correction can also be generalized in a straightforward manner.

4. Conclusion

We have derived Euler–Lagrange equations for hypersurfaces that extremize quadratic flux functionals. For divergence-free flows in $\mathbb{R}^3$, a class of weight functions has been found such that the surfaces retain the invariant structures, such as closed flow lines, that survive after perturbation away from integrability. In higher dimension an obvious application is to Hamiltonian flows, which are divergence free owing to the symplectic nature of the dynamical system. For instance, one might use the principle to define a surface over which to measure the flux due to the breaking of an adiabatic invariant.

However the dimension of the phase space of an n-dimensional Hamiltonian system is $2n$, while the dimension of constant-action surfaces in integrable systems, and invariant tori (if they exist) in nonintegrable systems is n, so that the theory as developed so far forms a possible basis for generalizing action–angle representations to nonintegrable systems only in the case $n = 1$ (which is nonintegrable if the Hamiltonian is time-dependent). It would thus be of interest to generalize the concept of quadratic flux to lower dimensional submanifolds than hypersurfaces.

Acknowledgment

One of us (RLD) thanks Dr. S. Hyde for pointing out a simplification in the derivation of Eq. (8).

References

[1] A.H. Boozer, Phys. Fluids 26 (1983) 1288.
[2] J.R. Cary and R.G. Littlejohn, Ann. Phys. 151 (1983) 1.
[3] Z. Yoshida, Phys. Plasmas 1 (1994) 208.
[4] H. Goldstein, Classical mechanics, 2nd Ed. (Addison-Wesley, Reading, MA, 1980).
[5] W.D. D'haeseleer, W.N.G. Hitchon, J.D. Callen and J.L. Shohet, Flux coordinates and magnetic field structure (Springer, Berlin, 1991).
[6] A.J. Lichtenberg and M.A. Lieberman, Applied mathematical sciences, no. 38. Regular and chaotic dynamics, 2nd Ed. (Springer, Berlin, 1992); 1st Ed. entitled Regular and stochastic motion (1983).
[7] J.D. Meiss and R.L. Dewar, in: Proc. Miniconference on Chaos and order, Centre for Mathematical Analysis, Australian National University, Canberra, Australia, 1991, eds. N. Joshi and R.L. Dewar (World Scientific, Singapore) pp. 97–103.
[8] R.L. Dewar and J.D. Meiss, Physica D 57 (1992) 476.
[9] J.K. Moser, An unusual variational problem connected with Mather's theory for monotone twist mappings, in: Proc. Euler Inst. Dynamical systems semester, Saint Petersburg, 1991, to appear.
[10] R.S. Mackay and M.R. Muldoon, Phys. Lett. A 178 (1993) 245.
[11] W.A. Newcomb, Phys. Fluids 2 (1959) 362.
[12] S.P. Hirshman and H.K. Meier, Phys. Fluids 28 (1985) 1387.
[13] A.H. Boozer, Phys. Fluids 24 (1981) 1999.
[14] L. Simon, Lectures on geometric measure theory, in: Proc. Centre for Mathematical Analysis, Vol. 3, Australian National University, Canberra, 1983.