Journal of Mathematical Physics

Some mathematics for quasi-symmetry

Cite as: J. Math. Phys. **61**, 093503 (2020); https://doi.org/10.1063/1.5142487
Submitted: 13 December 2019 . Accepted: 04 August 2020 . Published Online: 10 September 2020

J. W. Burby, N. Kallinikos, and R. S. MacKay

Journal of Mathematical Physics ARTICLE scitation.org/journal/jmp

Some mathematics for quasi-symmetry

Cite as: J. Math. Phys. 61, 093503 (2020); doi: 10.1063/1.5142487
Submitted: 13 December 2019 • Accepted: 4 August 2020 •
Published Online: 10 September 2020

J. W. Burby,[1] N. Kallinikos,[2,a)] and R. S. MacKay[2]

AFFILIATIONS

[1] Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
[2] Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom

[a)] **Author to whom correspondence should be addressed:** nikos.kallinikos@warwick.ac.uk

ABSTRACT

Quasi-symmetry of a steady magnetic field means integrability of first-order guiding-center motion by a spatial symmetry. Here, we derive many restrictions on the possibilities for a quasi-symmetry. We also derive an analog of the Grad–Shafranov equation for the flux function in a quasi-symmetric magnetohydrostatic field.

I. INTRODUCTION

The concept of quasi-symmetry was introduced by Boozer (1983) and then distilled into a design principle for stellarators by Nührenberg and Zille (1988). In its strongest sense, it means integrability of first-order guiding-center motion by a spatial symmetry. An excellent survey of the subject was provided by Helander (2014), assuming magnetohydrostatic (MHS) fields, that is, magnetohydrodynamic equilibrium with isotropic pressure and no mean flow.

A fundamental step was made by Burby and Qin (2013), who stated necessary and sufficient local conditions for integrability of guiding-center motion in terms of a continuous symmetry of three differential forms derived from the magnetic field and made clear that quasi-symmetry can be separated from the issue of whether the magnetic field is MHS or not.

Perturbative calculations of Garren and Boozer (1991), however, make it look very likely that the only possibility for exact quasi-symmetry for MHS fields with bounded magnetic surfaces is axisymmetry. Our paper gives first steps to deciding whether or not this is true.

In this paper, we prove many consequences of quasi-symmetry and thereby restrictions on possible quasi-symmetric fields. In the case of a quasi-symmetric MHS field, we derive a generalization of the axisymmetric Grad–Shafranov (GS) equation.

Burby and Qin (2013) built in an assumption that a quasi-symmetry must be a circle action. Here, we relax this requirement, though prove that under some mild conditions, it is actually a circle action.

We write many equations using differential forms. For those unfamiliar with differential forms, the book of Arnol'd (1978) (Chap. 7), is a classic, and there is a tutorial (MacKay, 2019) specifically for plasma physicists.

Throughout the paper, we will assume enough smoothness that the equations we write make sense, at least in a weak sense.

II. GUIDING-CENTER MOTION

We consider non-interacting charged particles in a steady, smooth (at least C^1) magnetic field B in 3D satisfying div $B = 0$, with $B \neq 0$ in the region of interest.

The (non-relativistic) motion of a particle of mass m, charge e, and position q in a magnetic field B on oriented Euclidean $\mathbb{R}^3$ has a formulation as a Hamiltonian system of three degrees of freedom (DoF),

$$i_V \omega = dH \tag{1}$$

for the vector field $V = (\dot{q}, \dot{p})$ on the cotangent bundle $T^*\mathbb{R}^3$, with Hamiltonian function and symplectic form (non-degenerate closed 2-form) given by

$$H(q,p) = \frac{|p|^2}{2m}, \tag{2}$$

$$\omega = -d\vartheta - e\pi^*\beta. \tag{3}$$

Here, p is a cotangent vector at $q \in \mathbb{R}^3$ (applied to a tangent vector ξ to $\mathbb{R}^3$, it produces $p(\xi) = p \cdot \xi$), $|p|$ is its Euclidean norm, ϑ is the tautological 1-form on $T^*\mathbb{R}^3$ defined by $\vartheta_{(q,p)}(\delta q, \delta p) = p(\delta q)$, $\pi : (q,p) \mapsto q$ is the natural map from $T^*\mathbb{R}^3$ to $\mathbb{R}^3$, π^* is the pullback by π, and $\beta = i_B\Omega$ for volume-form Ω on $\mathbb{R}^3$ corresponding to the Euclidean metric and chosen orientation. Note that $\operatorname{div} B = 0$ is equivalent to $d\beta = 0$.

One could allow time-dependent B, electric fields, arbitrary oriented Riemannian 3-manifold, and relativistic effects, but to focus ideas, we avoid all of these (the cases with electrostatic fields and relativity are treated in Appendix A).

If the perpendicular speed $v_\perp$ is less than $r_B|\Omega_B|$, where r_B is the radius of curvature of the field lines and $\Omega_B = -e|B|/m$ is the "gyrofrequency," then there is a locally unique "guiding center" X within r_B of q and "gyro-radius vector" ρ perpendicular to $B(X)$ and smaller than r_B such that

$$v = \frac{e}{m}B(X) \times \rho + v_\parallel b(X), \tag{4}$$

$$q = X + \rho, \tag{5}$$

where $v = \dot{q}$, $b = B/|B|$, and $v_\parallel = v \cdot b$. Indeed, the above formulas provide a local diffeomorphism from $(X, \rho, v_\parallel)$ to (q, v) for $|\rho| < r_B$.

If B varies slowly on the length-scales of ρ and $v_\parallel/\Omega_B$, then rotation of ρ about $B(X)$ is an approximate symmetry of the particle motion. There is a corresponding adiabatic invariant,

$$\mu = \frac{mv_\perp^2}{2|B(X)|} = \frac{1}{2}|e\Omega_B(X)||\rho|^2, \tag{6}$$

called the "magnetic moment."

If one neglects the variation of μ with time, one can reduce charged particle motion by gyro-rotation (Littlejohn, 1983) to obtain a Hamiltonian system of 2DoF with state $(X, v_\parallel)$ and

$$H = \frac{1}{2}mv_\parallel^2 + \mu|B(X)|, \tag{7}$$

$$\omega = -e\pi^*\beta - md(v_\parallel \pi^* b^\flat), \tag{8}$$

with π^* now being the pullback for the map $\pi(X, v_\parallel) = X$. The 2-form ω is non-degenerate iff $\widetilde{B}_\parallel \neq 0$ [to be introduced in (14)]. The equation $i_V\omega = dH$ for $V = (\dot{X}, \dot{v}_\parallel)$ implies

$$e\dot{X} \times \widetilde{B} = \mu\nabla|B| + m\dot{v}_\parallel b, \tag{9}$$

$$\dot{X} \cdot b = v_\parallel, \tag{10}$$

with the modified field

$$\widetilde{B} = B + \frac{m}{e}v_\parallel c, \text{ where } c = \operatorname{curl} b. \tag{11}$$

These can be rearranged to give

$$\dot{X} = \frac{1}{\widetilde{B}_\parallel}\left(v_\parallel \widetilde{B} + \frac{\mu}{e} b \times \nabla|B|\right), \tag{12}$$

$$\dot{v}_\parallel = -\frac{\mu}{m}\frac{\widetilde{B}}{\widetilde{B}_\parallel} \cdot \nabla|B|, \tag{13}$$

where

$$\widetilde{B}_{\|} = \widetilde{B} \cdot b. \tag{14}$$

We call (12) and (13) *first-order guiding-center motion* (FGCM); "first-order" because, as shown by Littlejohn (1983), it is possible to derive higher-order approximations, but we will restrict attention to first-order in this paper.

The Hamiltonian formulation (7) and (8) and drift equations (12) and (13) hold for an arbitrary oriented 3D Riemannian manifold, with $|\,|, \cdot, \times, \nabla$, div, and curl interpreted appropriately. Note that the above system is defined for $\widetilde{B}_{\|} \neq 0$, which is a reasonable assumption because the zeroth-order term in (14) is $|B| \neq 0$. In toroidal geometry, however, one can treat the degeneracy at $\widetilde{B}_{\|} = 0$ to avoid any arising inconsistencies in gyrokinetics and derive at the same time a canonical Hamiltonian structure for the purpose of symplectic integration (Burby and Ellison, 2017).

The zeroth-order approximation to FGCM (using $1/e$ as a convenient smallness parameter) is

$$\dot{X} = v_{\|} b, \tag{15}$$

$$\dot{v}_{\|} = -\frac{\mu}{m} b \cdot \nabla |B|. \tag{16}$$

We call this zeroth-order guiding-center motion (ZGCM).

Both FGCM and ZGCM conserve H of (7). We write E for the value of H.

In ZGCM, the guiding center moves along a fieldline. It may be *circulating*, meaning $v_{\|}$ has constant sign, or *bouncing*, meaning it is confined to an interval where $|B(X)| \leq E/\mu$ and $v_{\|}$ changes sign on reaching each end (the usual terminology for "bouncing" is "trapped," but this is inappropriate in a context where the whole point is to determine whether the particles are confined).

In FGCM, there are drifts of the guiding center across the field. These come from the modification $\widetilde{B}$, the $\nabla|B|$ term, and the $\widetilde{B}_{\|}$ denominator in (12). There are variants of FGCM, which agree to first order in $1/e$, but we choose the one above because it has a natural Hamiltonian formulation, which we believe is important and, in particular, allows us to discuss its integrability.

III. CONTINUOUS SYMMETRIES OF HAMILTONIAN SYSTEMS

Definition III.1. A continuous symmetry of a Hamiltonian system (M, H, ω) on a manifold M is a C^1 vector field U on M such that the Lie derivatives $L_U H$ and $L_U \omega$ are both zero.

It follows that there is a conserved quantity locally, and it can be extended globally under mild conditions. This is a Hamiltonian version of Noether's theorem.

Theorem III.2. *U is a continuous symmetry for a Hamiltonian system (M, H, ω) with vector field V iff there is a conserved local function K for V. If there are combinations $fU + gV$ of U and V with closed or recurrent trajectories realizing a basis of first homology $H_1(M)$, then K is global.*

Proof. $L_U \omega = 0$ and $d\omega = 0$ imply $d i_U \omega = 0$, so by Poincaré's lemma, $i_U \omega = dK$ for some local function K, and then,

$$i_V dK = i_V i_U \omega = -i_U dH = -L_U H = 0. \tag{17}$$

The converse is similar. If there is a combination $w = fU + gV$ of U and V with a closed trajectory γ, then

$$\int_\gamma i_U \omega = \int_0^T (f i_U + g i_V) i_U \omega \; dt, \tag{18}$$

where t is time along w and T is the period. The first term vanishes by antisymmetry of ω, and the second vanishes because of (17). For a recurrent trajectory, close it by a short arc and bound the error to obtain that the integral of $i_U \omega$ in its homology direction is zero [the concept of homology direction is described by Fried (1982)]. If $\int_\gamma i_U \omega = 0$ holds for γ representing a basis of $H_1(M)$, we deduce that K is global. □

Definition III.3. A 2DoF Hamiltonian system with vector field V is integrable if it has a continuous symmetry U with global conserved quantity K, and U, V are linearly independent almost everywhere (a.e.) (equivalently dK, dH are linearly independent a.e.).

Note that Definition III.1 implies that the symmetry U and the Hamiltonian vector field V commute because $i_{[U,V]}\omega = L_U i_V \omega - i_V L_U \omega = L_U dH = dL_U H = 0$ and ω is non-degenerate.

For an integrable 2DoF system, the bounded regular components of level sets of (K, H) are 2-tori, and there is a coordinate system in which U, V are both constant vector fields on each of them. [A component C of a level set of a C^1 function $F : M \to N$ is *regular* if DF

is surjective everywhere on C; in the present context, $F = (K, H)$ and $N = \mathbb{R}^2$.] This is a special case of the Arnol'd-Liouville (AL) theorem (Arnol'd, 1978). Here is a statement and proof.

Theorem III.4. *If U, V are commuting vector fields on a bounded surface S, independent everywhere on it, then S is a 2-torus, and there are coordinates $(\theta^1, \theta^2) \bmod 2\pi$ on it in which U, V are constant.*

Proof. Let ϕ^U be the flow of U and ϕ^V be the flow of V. For $t = (t_1, t_2) \in \mathbb{R}^2$, let $\phi_t = \phi^U_{t_1} \circ \phi^V_{t_2}$. Because the two commute and are independent, ϕ is a transitive action of the group $\mathbb{R}^2$ on S. Choose a point $0 \in S$, and let T be the set of $t \in \mathbb{R}^2$ such that $\phi_t(0) = 0$. It is a discrete subgroup of $\mathbb{R}^2$ (as a group under addition). Then, S is diffeomorphic to $\mathbb{R}^2/T$. Since S is bounded, T must be isomorphic to $\mathbb{Z}^2$. Thus, T is generated by a pair (T^1, T^2) of independent vectors in $\mathbb{R}^2$. Let A be the matrix with columns (T^1, T^2). Then, we obtain an action of $\mathbb{S}^1 \times \mathbb{S}^1$ (with $\mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$) on S by $(\theta, x) \mapsto \phi_{A\theta/2\pi}(x)$, where $\theta = (\theta^1, \theta^2) \in \mathbb{S}^1 \times \mathbb{S}^1$ and $x \in S$. Keeping $0 \in S$ fixed, this action defines a diffeomorphism $\mathbb{S}^1 \times \mathbb{S}^1 \to S$. In these coordinates, U is the first column of $2\pi A^{-1}$, and V is its second column, thus constant vector fields. □

Definition III.5. Coordinates $(\theta^1, \theta^2) \bmod 2\pi$ on a 2-torus in which commuting vector fields U, V on it are constant are called Arnol'd-Liouville (AL) coordinates.

The concepts of integrability and AL coordinates have higher dimensional analogs, but the 2DoF context suffices here.

Note that if K, H are C^3, then dK, dH independent a.e. implies that the union of non-regular level sets of (K, H) has measure zero. We have not found this result in the literature but are grateful to Vassili Gelfreich for providing a proof, attached here as Appendix B.

IV. QUASI-SYMMETRY

Definition IV.1. Given a magnetic field B on an oriented 3D Riemannian manifold Q, a vector field u on Q is a quasi-symmetry *of B if $U = (u, 0)$ is a continuous symmetry for FGCM for all values of magnetic moment μ.*

We assume B nowhere zero on Q in order for FGCM to make sense. Note that in contrast to most of the literature [e.g., Helander (2014)], we do not assume that B is MHS. Indeed, one might like to apply the concept of quasi-symmetry to magnetohydrodynamic equilibria with a mean flow [cf. Simakov and Helander (2011)] or with anisotropic pressure, for example. The concept of FGCM does not require an MHS field, so neither should quasi-symmetry.

A simple example of a quasi-symmetric magnetic field is any axisymmetric B in Euclidean space. Take $u = \partial_\phi = r\hat{\phi}$ in cylindrical coordinates (r, ϕ, z). Axisymmetry of B can be defined in various ways, e.g., $L_u B = 0$ or $L_u \beta = 0$ for this u. They are equivalent because $\operatorname{div} u = 0$ and

$$i_{[u,B]}\Omega = L_u\beta - (\operatorname{div} u)\beta. \tag{19}$$

Our first main theorem is as follows:

Theorem IV.2. *A vector field u is a quasi-symmetry of a magnetic field B iff*

$$L_u|B| = 0, \tag{20}$$

$$L_u\beta = 0, \tag{21}$$

$$L_u b^\flat = 0. \tag{22}$$

Proof. Recall the Hamiltonian and symplectic form for FGCM,

$$H = \frac{1}{2}mv_\parallel^2 + \mu|B(X)|, \tag{23}$$

$$\omega = -e\pi^*\beta - md(v_\parallel \pi^* b^\flat). \tag{24}$$

Then, $L_U H = \mu L_u|B|$, so $L_U H = 0$ for all μ iff $L_u|B| = 0$. Next,

$$L_U\omega = -eL_u\beta - md(v_\parallel L_u b^\flat) = -eL_u\beta - mdv_\parallel \wedge L_u b^\flat - mv_\parallel dL_u b^\flat. \tag{25}$$

Apply this to an arbitrary pair of tangents to Q and set $v_\parallel = 0$ to deduce that $L_U\omega = 0$ implies $L_u\beta = 0$. Apply it to an arbitrary tangent ξ to Q and the vector $(0, 1)$ tangent to $Q \times \mathbb{R}$ to deduce that $i_\xi L_u b^\flat = 0$, so $L_u b^\flat = 0$.

In the other direction, if $L_u\beta = 0$ and $L_u b^\flat = 0$, then $L_U\omega = 0$.

So we have proved that u is a quasi-symmetry of B iff $L_u|B| = 0$, $L_u\beta = 0$, and $L_u b^\flat = 0$. □

Remark IV.3. Actually if $(u, 0)$ is a continuous symmetry for FGCM for one set of non-zero e, m, μ then it is for all e, m, μ.

We write the three conditions of Theorem IV.2 in vector calculus for comparison (recall $c = \text{curl}\, b$):

$$u \cdot \nabla|B| = 0, \tag{26}$$

$$\text{curl}\,(B \times u) = 0, \tag{27}$$

$$c \times u + \nabla(u \cdot b) = 0. \tag{28}$$

Quasi-symmetry has the following significant consequences:

Theorem IV.4. *If u is a quasi-symmetry of a magnetic field B, then $L_u B^\flat = 0$, $L_u\Omega = 0$, and $L_u B = 0$.*

Proof. To prove $L_u B^\flat = 0$, use $B^\flat = |B| b^\flat$, so

$$L_u B^\flat = (L_u|B|) b^\flat + |B| L_u b^\flat. \tag{29}$$

By Theorem IV.2, $L_u|B| = 0$ and $L_u b^\flat = 0$. Thus, $L_u B^\flat = 0$.

To prove $L_u\Omega = 0$, note that $\beta \wedge b^\flat = |B|\Omega$. Applying L_u, we obtain

$$L_u\beta \wedge b^\flat + \beta \wedge L_u b^\flat = (L_u|B|)\Omega + |B| L_u\Omega. \tag{30}$$

According to Theorem IV.2, the first three terms of this are zero. As $|B| \neq 0$, we obtain $L_u\Omega = 0$.

To prove that $L_u B = 0$, note that it can alternatively be written as $[u, B] = 0$. Use the formula

$$i_{[u,B]}\Omega = L_u i_B \Omega - i_B L_u \Omega, \tag{31}$$

which holds for any pair of vector fields u, B and any differential form Ω, in particular, the volume-form. By Theorem IV.2, $L_u\beta = 0$, and we just proved that $L_u\Omega = 0$. So using Ω non-degenerate, we see that $[u, B] = 0$, cf. (19). □

Similarly to that for $L_u b^\flat$ in (28), the first result of Theorem IV.4 is written as

$$u \times J = \nabla(u \cdot B), \tag{32}$$

where $J = \text{curl}\, B$, because

$$L_u B^\flat = i_u dB^\flat + d i_u B^\flat = i_u i_J \Omega + d(u \cdot B). \tag{33}$$

The second says

$$\text{div}\, u = 0, \tag{34}$$

and for the third,

$$L_u B = [u, B] = u \cdot \nabla B - B \cdot \nabla u = \text{curl}\,(B \times u) + (\text{div}\, B)u - (\text{div}\, u)B. \tag{35}$$

Since $\text{div}\, B = 0$ and we already proved that $\text{div}\, u = 0$, then $[u, B] = 0$ can be written in this case as $\text{curl}\,(B \times u) = 0$.

Noting that some steps in the above proof are reversible, we can derive various alternative necessary and sufficient conditions for quasi-symmetry. The following theorem gives some examples, from which we shall frequently use (i) or (ii). Case (i) is a slight generalization of the formulation by Burby and Qin (2013).

Theorem IV.5. *A vector field u is a quasi-symmetry of a magnetic field B iff any of the following sets of conditions hold:*

i. $L_u|B| = 0$, $L_u\beta = 0$, $L_u B^\flat = 0$;
ii. $L_u\Omega = 0$, $L_u\beta = 0$, $L_u B^\flat = 0$;
iii. $L_u\Omega = 0$, $L_u B = 0$, $L_u B^\flat = 0$.

Proof. (i) To prove the first set, we use (29). Thus, under $L_u|B| = 0$ and $B \neq 0$, we obtain $L_u B^\flat = 0$ iff $L_u b^\flat = 0$, which converts Theorem IV.2 to (i).

(ii) The second comes from the first and (30).
(iii) The third comes from the second and (31).

□

Here are some additional consequences of quasi-symmetry.

Theorem IV.6. *If u is a quasi-symmetry of B, then*

i. $L_u(u\cdot B) = 0, L_u(u\cdot b) = 0,$
ii. $[u,b] = 0, [u, B/|B|^2] = 0, [u, u_\perp] = 0$ *(where $u_\perp$ is the component of u perpendicular to B), and*
iii. $[u,J] = 0$ *(where $J = \operatorname{curl} B$), $[u,[J,B]] = 0$, $L_u(J\cdot B) = 0$, and $L_J(u\cdot B) = 0$.*

Proof. (i) $u\cdot B = i_u B^\flat$ so $L_u(u\cdot B) = i_u L_u B^\flat + i_{[u,u]}B^\flat$, both of which are zero.
For $u\cdot b$, apply L_u to $u\cdot B = |B| u\cdot b$ and use the above plus $B \neq 0$ to deduce that $L_u(u\cdot b) = 0$.
(ii) For $[u,b]$, use $[u,B] = 0$, $B = |B|b$, and $L_u|B| = 0$ to obtain $[u,b] = 0$.
Similarly, $[u, B/|B|^2] = 0$ or indeed $[u, f(|B|)B] = 0$ for any function f.
$u_\perp = u - (u\cdot B)B/|B|^2$, and L_u on each of these terms is zero, so $L_u u_\perp = 0$.
(iii) $J = \operatorname{curl} B$ translates to $i_J\Omega = dB^\flat$. Apply L_u to each side. $L_u i_J\Omega = i_J L_u\Omega + i_{[u,J]}\Omega$ and $L_u dB^\flat = dL_u B^\flat = 0$. However, $L_u\Omega = 0$, so we deduce that $i_{[u,J]}\Omega = 0$. Ω is non-degenerate, so $[u,J] = 0$.
For $[u,[J,B]]$, we use the Jacobi identity $[u,[J,B]] + [J,[B,u]] + [B,[u,J]] = 0$. We already proved that $[B,u] = 0$ and $[u,J] = 0$. So $[u,[J,B]] = 0$.
$L_u(J\cdot B) = L_u i_J B^\flat = i_J L_u B^\flat = 0$, using $[u,J] = 0$.
$L_J(u\cdot B) = i_J d(u\cdot B) = i_J L_u B^\flat = 0$, using (33). □

If B is a vacuum field with quasi-symmetry u, note that Theorem IV.6(i) can be strengthened to $u\cdot B = \text{const.}$, using (33) [or (32)].

V. FLUX FUNCTION

The condition $L_u\beta = 0$ of Theorem IV.2 merits additional comment. We discuss it in a more general context than quasi-symmetry. Specifically, we require only $L_u\beta = 0$, $\operatorname{div} B = 0$, and $\operatorname{div} u = 0$.

Because $d\beta = 0$ and $\beta = i_B\Omega$, $L_u\beta = 0$ is equivalent to $di_u i_B\Omega = 0$. Thus, by Poincaré's lemma, $i_u i_B\Omega = d\psi$ for some function ψ locally (in vector calculus, $B\times u = \nabla\psi$), and both u and B are tangent to regular level sets of ψ.

An important question is whether ψ is global. It is global if there are combinations $fu + gB$ with closed or recurrent trajectories realizing a basis of $H_1(Q)$. For the case of Q being a solid torus with a circulating magnetic field, B has a closed trajectory realizing $H_1(Q)$, so ψ is global. For more complicated domains, it might fail.

Definition V.1. A flux function for a field B on Q is a globally defined function $\psi : Q \to \mathbb{R}$ with $i_B d\psi = 0$ and $d\psi \neq 0$ a.e.

Note that the existence of a flux function ψ is an assumption of the standard approach to quasi-symmetry [e.g., Helander (2014)], whereas here, we derived it as a consequence, at least as a local function. For many purposes, however, we will need to assume that ψ is global and has non-zero derivative a.e. (see Definition VI.1). Note that in the paper [Helander (2014)], ψ is chosen to be the toroidal flux enclosed by the level set, and the term "flux function" is used for any function of ψ.

If ψ is global, it follows from the classification of surfaces that bounded regular components of level sets of ψ are 2-tori because they support a nowhere-zero vector field (u or B).

Definition V.2. The bounded regular components of level sets of a flux function are called flux surfaces.

Furthermore, u and B are independent everywhere on such a 2-torus because $i_u i_B\Omega = d\psi$. $L_u\beta = 0$ with $\operatorname{div} u = 0$ implies that $[u,B] = 0$ [cf. (19)]. Using Theorem III.4, it follows that there are coordinates $(\theta^1,\theta^2) \bmod 2\pi$ on the 2-torus in which both u and B are constant vector fields. There are many works on "flux coordinates," including the book (D'haeseleer *et al.*, 1991) and the recent paper (Kruger and Greene, 2019), but we are not aware of any of them using this very natural AL approach. The closest we have seen is the paper of Hamada (1962).

We now derive a formula for the winding ratios of $X = u, B$ on a flux surface.

Definition V.3. The winding ratio ι_X of a vector field X on a 2-torus with coordinates $(\theta^1,\theta^2) \in \mathbb{S}^1\times\mathbb{S}^1$ is the limit as $t\to\infty$ of the ratio of the number of revolutions made in θ^1 along a trajectory of vector field X to that in θ^2. ι_X is considered as a point in the projective line $\mathbb{R}P^1$ to include the option of ∞ and to ignore the sign of X.

For vector fields X of "Poincaré type" (those having a cross section) on a 2-torus, the limit exists and is the same for all trajectories and for both signs of X. Since u and B are conjugate to non-zero constant vector fields on each flux surface, they are of Poincaré type.

As X conserves Ω and ψ, it also conserves an area-form on flux surfaces. Indeed, let

$$\mathcal{A} = i_n\Omega, \text{ where } n = \frac{\nabla\psi}{|\nabla\psi|^2}. \tag{36}$$

In vector calculus, $\mathcal{A}(\xi,\eta) = n\cdot(\xi\times\eta)$. Then, the restriction $\mathcal{A}_C$ of $\mathcal{A}$ to a regular component C of a level set of ψ is non-degenerate and conserved by X. To prove the conservation, it is enough to work out $L_X\mathcal{A}$ on (u,B), which form a basis of tangents to C. Using $[u,B]=0$, we have

$$i_u i_B L_X\mathcal{A} = L_X i_u i_B i_n\Omega = L_X i_n d\psi = L_X 1 = 0. \tag{37}$$

Theorem V.4. *If $\operatorname{div} u = \operatorname{div} B = 0$ and $i_u i_B\Omega = d\psi$, then for $X = u$ or B on a bounded regular component C of a level set of ψ,*

$$\iota_X = -\frac{\int_{\gamma_1} i_X\mathcal{A}_C}{\int_{\gamma_2} i_X\mathcal{A}_C}, \tag{38}$$

where γ_j is any closed loop on C making one turn in θ^j and none in the other.

Proof. As $L_X\mathcal{A}_C = 0$, we deduce that $i_X\mathcal{A}_C$ is closed, so its integral round a closed loop γ depends on only the homology class $[\gamma]$ of the loop. Take a long piece of trajectory of X on C and close it by a short arc on C, making a closed loop γ. It has homology class close to $N([\gamma_1]+\iota_X[\gamma_2])$ for some large integer N. $i_X\mathcal{A}_C(\dot\gamma)$ is zero except on the short arc. Taking the limit, we obtain

$$\int_{[\gamma_1]+\iota_X[\gamma_2]} i_X\mathcal{A}_C = 0, \tag{39}$$

hence the formula of the theorem. □

The same formula applies to the current density J for an MHS field with p constant on flux surfaces.

VI. THE INVARIANT TORI OF FGCM

Let us compute the conserved quantity K of FGCM resulting from quasi-symmetry. Recall from Theorem III.2 that K results from $i_U\omega = dK$. Recall from (8) that $\omega = -e\pi^*\beta - md(v_\parallel\pi^* b^\flat)$ and from Definition IV.1 that $U = (u,0)$. So

$$i_U\omega = -ei_u\beta - mL_U(v_\parallel\pi^* b^\flat) + mdi_u(v_\parallel b^\flat) = dK, \tag{40}$$

with

$$K = -e\psi + mv_\parallel u\cdot b, \tag{41}$$

using $L_U(v_\parallel\pi^* b^\flat) = v_\parallel L_u b^\flat = 0$. In particular, we see that K is global iff ψ is global.

Definition VI.1. A magnetic field B is quasi-symmetric if it has a quasi-symmetry u, and the associated flux function ψ is global with $d\psi \neq 0$ a.e.

Theorem VI.2. *If B is quasi-symmetric, then FGCM is integrable.*

Proof. By Definition IV.1, $(u,0)$ is a continuous symmetry for FGCM for all μ. From (41), the associated local conserved quantity K is global if ψ is global. To complete the verification of integrability (see Definition III.3), we must check that dH, dK are independent a.e. The derivatives dH and dK are independent at $(X,v_\parallel)$ iff $rdH + sdK = 0$, $r,s\in\mathbb{R}$ implies $r=s=0$. Now,

$$rdH + sdK = r(mv_\parallel dv_\parallel + \mu d|B|) + s(-ed\psi + m\,u\cdot b\,dv_\parallel + mv_\parallel d(u\cdot b)). \tag{42}$$

If this is zero and $(r,s)\neq(0,0)$, then the coefficient $rmv_\parallel + smu\cdot b$ of $dv_\parallel$ is zero, and so

$$\mu\,u\cdot b\,d|B| + v_\parallel(ed\psi - mv_\parallel d(u\cdot b)) = 0. \tag{43}$$

However, this is quadratic in $v_\parallel$, so if $d\psi\neq 0$ at X, it is zero for at most two values of $v_\parallel$ (typically for none because the coefficients are 1-forms in 3D). $d\psi\neq 0$ a.e. in X, so dH, dK are independent a.e. in $(X,v_\parallel)$. □

K governs how far particles move from a flux surface. Using conservation of K, we see that

$$v_{\parallel} = \frac{e\psi + K}{m\, u \cdot b}, \tag{44}$$

as long as $u \cdot b \neq 0$. Hence, by conservation of H in (7), the tori for FGCM are given in projection to guiding-center position by

$$\frac{1}{2}\frac{(e\psi + K)^2}{m(u \cdot b)^2} + \mu|B| = E, \tag{45}$$

with parallel velocity recovered by (44). This can be written as

$$\psi = -\frac{K}{e} \pm \frac{u \cdot b}{e}\sqrt{2m(E - \mu|B|)}. \tag{46}$$

We see the same division of motion into circulating and bouncing, as for ZGCM. Note that by $L_u|B| = 0$, the set of X where $|B(X)| > E/\mu$ is a set of u-lines.

Remark VI.3. The converse of Theorem VI.2 is not obvious. Perhaps there could be magnetic fields for which FGCM has a velocity-dependent symmetry. After all, gyro-rotation is a velocity-dependent symmetry for charged particle motion in a uniform field.

VII. EFFECT ON THE METRIC

Next, we examine the relation of a quasi-symmetry u to the Riemannian metric g. The conjecture of Garren and Boozer (1991) suggests that u is a Killing field for Euclidean metric g because an isometry with bounded orbits for Euclidean metric has to be a rotation.

Definition VII.1. A vector field u is a Killing field for a Riemannian metric g if $L_u g = 0$.

In Euclidean space, $L_u g = 0$ can be written as $\nabla u + (\nabla u)^T = 0$.

We have not managed to prove or disprove that a quasi-symmetry is a Killing field yet, but the following theorem gets two-thirds of the way (by showing that the subspace of possibilities for $L_u g$ at a point is constrained to a codimension-4 subspace of the 6D space of symmetric 3×3 matrices). It applies to an arbitrary Riemannian metric g.

Theorem VII.2. *Let u be a quasi-symmetry for magnetic field B, with flux function ψ. Wherever u, B are independent, then $d\psi \neq 0$ and (B, u, n) is a basis, with $n = \nabla\psi/|\nabla\psi|^2$. With respect to this basis, $L_u g$ has matrix*

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & L_u|u|^2 & i_n L_u u^\flat \\ 0 & u \cdot [n, u] & L_u|n|^2 \end{bmatrix}, \tag{47}$$

and the diagonal terms are related by

$$L_u|n|^2 = -|B|^2|n|^4 L_u|u|^2. \tag{48}$$

Proof. $d\psi \neq 0$ where u, B are independent because $i_u i_B \Omega = d\psi$ and Ω is non-degenerate. $\Omega(B, u, n) = i_n i_u i_B \Omega = i_n d\psi = 1 \neq 0$, so (B, u, n) is a basis.

The calculation of $L_u g$ makes use of Lemma VII.5 to follow, which says that the standard commutation relation $i_X L_u = L_u i_X - i_{[u,X]}$ applies not only to differential forms but also to any covariant 2-tensor, thus including the case of a metric tensor. We apply this to the metric tensor g for $X = B, u, n$ in turn.

For $X = B$, we obtain $i_B L_u g = 0$, so the first row of $L_u g$ is zero and also the first column by symmetry of g.

For $X = u$, we obtain $i_u L_u g = L_u i_u g = L_u u^\flat$. The diagonal component is obtained by contracting this with u: $i_u L_u u^\flat = L_u i_u u^\flat = L_u|u|^2$.

For $X = n$, we obtain

$$i_n L_u g = L_u i_n g - i_{[u,n]} g. \tag{49}$$

For the off-diagonal term, we contract (49) with u. First, $i_u L_u i_n g = L_u i_u i_n g = L_u(u \cdot n)$ but $u \cdot n = 0$ from $i_u d\psi = 0$. Second, $i_u i_{[u,n]} g = u \cdot [u, n]$. For the diagonal term, we contract (49) with n. First,

$$i_n L_u i_n g = i_n L_u n^\flat = i_n L_u (d\psi/|\nabla\psi|^2) = |\nabla\psi|^{-2} i_n L_u d\psi + (L_u|\nabla\psi|^{-2}) i_n d\psi = L_u|\nabla\psi|^{-2}, \tag{50}$$

using $L_u\psi = 0$ and $i_n d\psi = 1$. Second,

$$i_n i_{[u,n]} g = i_{[u,n]} n^\flat = L_u i_n n^\flat - i_n L_u n^\flat = L_u|\nabla\psi|^{-2} - |\nabla\psi|^{-2} i_n L_u d\psi - (L_u|\nabla\psi|^{-2}) i_n d\psi = 0. \tag{51}$$

Finally, we prove the indicated relation between the diagonal terms. Using $B \times u = \nabla\psi$,

$$|\nabla\psi|^2 = |B|^2|u|^2 - (u \cdot B)^2. \tag{52}$$

We proved $L_u(u \cdot B) = 0$ and $L_u|B| = 0$, so applying L_u to the above equation, we obtain $L_u|\nabla\psi|^2 = |B|^2 L_u|u|^2$, hence the result. □

Because g is symmetric, $L_u g$ is symmetric, but we give the two alternative expressions for the off-diagonal components in (47), and one can check that they are equal ($u \cdot [n,u] = i_{[n,u]}u^\flat = i_n L_u u^\flat - L_u i_n u^\flat$ but $n \cdot u = 0$).

One could choose other normalizations of $\nabla\psi$, but an advantage of the chosen one is that $n \cdot [u,n] = 0$, as proved in (51), besides $\det g = 1$ for (B,u,n).

Relation (48) can alternatively be obtained by using $\operatorname{div} u = 0$ and the local expression $\Omega = \pm\sqrt{\det g}\, dx^1 \wedge dx^2 \wedge dx^3$, with $\det g$ being the determinant of the matrix representing g in coordinate system (x^1, x^2, x^3) [i.e., $g(\xi,\eta) = g_{ij}\xi^i\eta^j$]. A calculation shows that $\operatorname{div} u = \frac{1}{2}\operatorname{tr}(g^{-1}L_u g)$ and hence (48).

The previous theorem highlights the importance of $L_u u^\flat$ not just for the off-diagonal term but also because $L_u|u|^2 = i_u L_u u^\flat$. So

Theorem VII.3. *A quasi-symmetry u is a Killing field iff* $L_u u^\flat = 0$.

In vector calculus, $L_u u^\flat = 0$ can be written as $w = 0$, where

$$w = v \times u + \nabla|u|^2, \tag{53}$$

with $v = \operatorname{curl} u$. The relation between the two is $w^\flat = L_u u^\flat$. In terms of w, (47) combined with (48) can be written as

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & u \cdot w & n \cdot w \\ 0 & n \cdot w & -|B|^2|n|^4 u \cdot w \end{bmatrix}. \tag{54}$$

Either from (47), using (48) and the aforementioned symmetry of $L_u g$, or (54), we see that for $|B|, |\nabla\psi| \neq 0$, the (1,1)-minor of $L_u g$ vanishes iff both the diagonal and off-diagonal terms vanish.

Corollary VII.4. If u is a quasi-symmetry, then $L_u g$ is degenerate. If $L_u g$ is non-zero, it has rank 2.

We conclude this section with the required lemma.

Lemma VII.5. For any covariant 2-tensor g and vector fields u, X,

$$i_{[u,X]}g = L_u i_X g - i_X L_u g. \tag{55}$$

Proof. For arbitrary vector fields u, X and Y, and covariant 2-tensor g,

$$(L_u g)(X,Y) = L_u(g(X,Y)) - g(L_u X, Y) - g(X, L_u Y). \tag{56}$$

This says

$$i_Y i_X L_u g = L_u i_Y i_X g - i_Y i_{[u,X]} g - i_{[u,Y]} i_X g. \tag{57}$$

Now, $i_X g$ is a differential form, so the usual commutation relation

$$L_u i_Y i_X g = i_Y L_u i_X g + i_{[u,Y]} i_X g \tag{58}$$

can be employed for the first term on the right. It results that

$$i_Y i_X L_u g = i_Y L_u i_X g - i_Y i_{[u,X]} g. \tag{59}$$

This is true for all Y, hence the result. □

VIII. CIRCLE ACTION

In the case of axisymmetry, the trajectories of u are all closed and have a common period (single points on the axis of symmetry and circles elsewhere, of period 2π). We say the flow of u generates a circle action.

Definition VIII.1. A circle action on a manifold M is a differentiable mapping $\Phi : \mathbb{S}^1 \times M \to M$, $(\theta, x) \mapsto \Phi_\theta(x)$ such that $\Phi_0(x) = x$ and $\Phi_{\theta+\theta'}(x) = \Phi_\theta(\Phi_{\theta'}(x))$ for all $\theta, \theta' \in \mathbb{S}^1$ and $x \in M$.

The orbit of a point x under a circle action is either a point or diffeomorphic to a circle.

Burby and Qin (2013) formulated quasi-symmetry in terms of a circle action preserving FGCM. Given a circle action, one can obtain a vector field $u = \partial_\theta \Phi_\theta|_{\theta=0}$. It is a quasi-symmetry if Φ preserves FGCM. In our treatment of quasi-symmetry here, we do not require the trajectories of a quasi-symmetry to be circles, but we prove now that under mild conditions any quasi-symmetry does generate a circle action locally.

Theorem VIII.2. *If u is a quasi-symmetry for B, ψ is global, S is a bounded regular component of a level set of ψ (flux surface), and S contains a regular component T of a joint level set of either*

i. $(|B|, \psi)$ *with* $u \cdot B \neq 0$ *or*
ii. $(u \cdot B, \psi)$,

then T is a circle and a closed u-line, and all u-lines on S are closed and of the same period. Furthermore, all nearby flux surfaces have the same properties and the same period. The circles are non-contractible on the flux surfaces, and all have the same rational winding ratio on this interval of flux surfaces.

Proof. The vector field u is nowhere zero on S because $i_u i_B \Omega = d\psi \neq 0$. A bounded regular component T of a level set of two functions in 3D is a circle. $L_u \psi = 0$. In case (i), $L_u|B| = 0$ and $d|B|, d\psi$ are independent, so T is a u-line. In case (ii), $L_u(u \cdot B) = 0$ and $d(u \cdot B), d\psi$ are independent, so T is a u-line.

Now, $[u, B] = 0$, so by Theorem III.4, u is conjugate to a constant vector field on S. Because one u-line on S is closed, it follows that all u-lines on S are closed and have the same period.

Independence of $d\psi$ and $d|B|$ [respectively, $d(u \cdot B)$] on T implies the same for all nearby components of level sets of $(\psi, |B|)$ [respectively, $(\psi, u \cdot B)$]. So we obtain the same result for all nearby flux surfaces.

To prove that the period of the u-lines is the same for nearby flux surfaces, we treat the two cases separately.

In case (i), let the function $f(x) = 2\pi/\tau(x)$, where $\tau(x)$ is the period of the u-line through the given point x and define the vector field $R = u/f$. Then, R generates a circle action Φ. Define the circle-average $\langle \omega \rangle$ of any differential form or vector field ω by

$$\langle \omega \rangle = \frac{1}{2\pi} \int_0^{2\pi} \Phi_\theta^* \omega \; d\theta. \tag{60}$$

Note that the circle-average of $L_R \omega$ is zero for any ω. From $u = fR$ follows

$$0 = L_u B^\flat = (B \cdot R) \; df + f L_R B^\flat. \tag{61}$$

Take the circle-average of this equation to obtain

$$0 = (B \cdot R) \; df \tag{62}$$

because f and $B \cdot R$ are constant along Φ-orbits. So if $u \cdot B \neq 0$, then $df = 0$. This applies on a neighborhood of S, so f is locally constant.

In case (ii), $L_u B^\flat = 0$ implies that $i_u i_J \Omega = -d(u \cdot B)$ [Eq. (33)] and $[u, J] = 0$ [Theorem IV.6(iii)]. So u, J are commuting vector fields on each level set of $u \cdot B$. Now, $d(u \cdot B) \neq 0$, so the u-lines on the level set of $u \cdot B$ are closed and have the same period (using Theorem III.4). By independence of $d(u \cdot B)$ and $d\psi$, this gives us a u-line on each nearby flux surface, and it has the same period. Thus, they all have the same period.

The u-lines are non-contractible on the flux surfaces because of the conjugacy to a constant vector field. The winding ratio of u as a function of ψ is continuous and rational, so it is constant. □

Remark VIII.3. There is also a partial converse to the above theorem. Namely, if u is a quasi-symmetry that is equal to the infinitesimal generator of a circle action Φ_θ, then there is a globally defined ψ such that $i_u \beta = d\psi$. This may be seen by directly computing the exterior derivative of $\psi = i_u \langle A^\flat \rangle$, where A is the vector potential.

Note that case (ii) can not occur for an MHS field with p constant on flux surfaces because of Theorem IX.2 to follow, but it might be useful in other situations.

The rational winding ratio $m : n$ of the u-lines is called the *type* of the quasi-symmetry. With θ^1 poloidal and θ^2 toroidal, $m : n = 0 : 1$ is called quasi-axisymmetric (QA), $1 : 0$ is called quasi-poloidal (QP), and anything else is called quasi-helical (QH).

Note that both quasi-symmetry and circle action allow the possibility of short fibres, for example, a region of $m : n$ QH may shrink onto a closed u-line (which will be also a B-line) around which the rest have the winding ratio $m : n$ (a "Seifert fibration"). If the u-period of the main u-lines is 2π, then the period of the short fibre is $2\pi/n$.

Note also that the construction in case (ii) of tori with $u \cdot B$ constant supporting commuting vector fields u, J applies for a general quasi-symmetry u, without requiring the derivatives of $u \cdot B$ and ψ to be independent. The formula of Theorem V.4 for winding ratios extends to those for u and J on these tori, with n replaced by $\nabla(u \cdot B)/|\nabla(u \cdot B)|^2$.

IX. RELATION TO STANDARD TREATMENTS

The standard approach to quasi-symmetry, as exemplified by Helander (2014), assumes a magnetohydrostatic field B from the start.

Definition IX.1. A magnetic field B is magnetohydrostatic (MHS) if $J \times B = \nabla p$ for some function p, where $J = \mathrm{curl}\, B$.

It then assumes a flux function ψ, i.e., a function such that $B \cdot \nabla\psi = 0$ and $\nabla\psi \neq 0$ a.e. Given the MHS assumption, this is not a great restriction because p satisfies $B \cdot \nabla p = 0$; the only catch is that dp might not be non-zero a.e. [in particular, in the surrounding vacuum region but also because dp must vanish at all rational surfaces where the resonant Fourier harmonic of $1/|B|^2$ does not vanish (Boozer, 1981)]. Then, the level sets of ψ are assumed to be bounded in the region of interest and hence 2-tori. Actually, Helander (2014) takes ψ to be the toroidal flux bounded by the level set of ψ, but that is not essential. The standard approach also assumes that p is constant on flux surfaces, which is automatic if the field has density of irrational flux surfaces but might otherwise be a restriction. Then, it is proved that there are "Boozer" coordinates (Boozer, 1981), which, in particular, make the magnetic field lines straight. Guiding-center motion is formulated in the Boozer angles as a Lagrangian system and seen to have an ignorable linear combination of the angles if the field strength is constant along a family of straight lines on each flux surface (not, in general, the same straight lines as the field lines), and so guiding-center motion is integrable. The field is said to be quasi-symmetric if this is the case.

An alternative approach is due to Hamada (1962) but requires $dp \neq 0$ a.e. It constructs a different coordinate system on flux surfaces but with similar properties, and quasi-symmetry is identified as the result of an ignorable coordinate again. Helander (2014) identifies a whole class of coordinate systems that will do as well.

Here, we explain how our approach connects to these. In our definition and analysis of quasi-symmetry, we have not required the field to be MHS, but if it is MHS and has quasi-symmetry u in our sense with a global flux function ψ and if p is constant on flux surfaces, it turns out that $u \cdot B$ is also. The latter in this case will be denoted by C. We showed that quasi-symmetry implies that $[u, B] = 0$ and $[u, B/|B|^2] = 0$. We claim that the resulting AL coordinates on flux surfaces augmented by ψ give Hamada coordinates in the first case and Boozer coordinates in the second one. Because of $L_u|B| = 0$, it follows that $|B|$ is constant along the u-lines, which are straight in either case. So under the assumptions of MHS with p constant on flux surfaces, quasi-symmetry in our sense implies quasi-symmetric in the standard sense.

We will now prove these statements. Afterward, we will address the converse question.

Theorem IX.2. *If u is a quasi-symmetry for an MHS field B and p is constant on flux surfaces, then $u \cdot B$ is constant on flux surfaces.*

Proof. We already have $L_u(u \cdot B) = 0$ (Theorem IV.6). Now, $L_B(u \cdot B) = L_B i_u B^\flat = i_u L_B B^\flat$ because $[u, B] = 0$. Translated to differential forms, the MHS equation says $i_B dB^\flat = dp$. So $L_B B^\flat = dp + d|B|^2$. Thus, $L_B(u \cdot B) = i_u dp + i_u d|B|^2$. The first term is zero because $i_u d\psi = 0$ and p is constant on flux surfaces. The second is zero from $L_u|B| = 0$. u and B are independent and tangent to flux surfaces. Combining these two results, $u \cdot B$ is constant on flux surfaces. □

Thus, in the MHS case, $u \cdot B$ is a function $C(\psi)$, and so together with $B \times u = \nabla\psi$, we can describe the magnetic field in terms of u and ψ. [Technically, $u \cdot B = C(\psi)$ is not quite correct because if a level set of ψ has more than one regular component, then $u \cdot B$ could take different values on the different components, but we use the formulation $u \cdot B = C(\psi)$ with this understanding. Similarly, we will write p being constant on flux surfaces as $p = p(\psi)$.] We can do the same for J. The results are stated in the following theorem:

Theorem IX.3. *If u is a quasi-symmetry for an MHS field B with p constant on flux surfaces, then B has the form*

$$B = \frac{1}{|u|^2}(C(\psi)u + u \times \nabla\psi). \tag{63}$$

Assume C and p are absolutely continuous functions of ψ. Then,

$$J = -p'(\psi)u - C'(\psi)B. \tag{64}$$

Proof. Equation (63) comes straight from $B \times u = \nabla\psi$ if we cross with u and use $u \cdot B = C(\psi)$. Absolute continuity of a function is enough for its derivative to exist a.e. and for the fundamental theorem of calculus to hold. The MHS condition implies that J is tangent to flux surfaces, so it is a linear combination of u and B, say $J = \kappa u + \lambda B$ for some functions κ, λ. Then, from $\nabla p = J \times B = \kappa u \times B = -\kappa\nabla\psi$, we find $\kappa = -p'$. Likewise, from (32), we have $\nabla C = u \times J = \lambda u \times B = -\lambda\nabla\psi$, and so $\lambda = -C'$. □

Remark IX.4. (a) The function C has an interpretation as a current. More precisely, if γ_1, γ_2 are u-lines with values ψ_1, ψ_2 of flux function and their (common) period is denoted τ (Sec. VIII), then the current I across any surface S spanning γ_1 and γ_2 is $I = \int_S J^\flat = \int_{\gamma_1 - \gamma_2} B^\flat = \int_{\gamma_1 - \gamma_2} B \cdot u ds = (C(\psi_1) - (\psi_2))\,\tau$, where s is time along u.

(b) For future reference, it is also useful to express the quasi-symmetry u in terms of B and ψ. So now, if we cross $B \times u = \nabla\psi$ with B, we have

$$u = \frac{1}{|B|^2}(C(\psi)B - B \times \nabla\psi), \tag{65}$$

using $u \cdot B = C(\psi)$. An alternative expression can be obtained by crossing $B \times u = \nabla\psi$ with $\nabla|B|$ and using $u \cdot \nabla|B| = 0$. In this way, the quasi-symmetry can be written as

$$u = \frac{\nabla\psi \times \nabla|B|}{B\cdot\nabla|B|}. \quad (66)$$

We emphasize that the two expressions are not equivalent as the former assumes MHS fields in using $u\cdot B = C(\psi)$, while the latter does not but uses $L_u|B| = 0$ instead.

Definition IX.5. Given a magnetic field B with a flux function ψ, coordinates (θ^1,θ^2,ψ), $\theta^j \in \mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$, are magnetic coordinates *if the B-lines are straight in them.*

Note that this remains true under any linear transformation on θ^j in $SL(2,\mathbb{Z})$, so it is conventional to take θ^1 in the poloidal direction on a flux surface and θ^2 in the toroidal direction, defined (up to orientation) by its embedding in $\mathbb{R}^3$. For topologists, a poloidal loop is a "meridian," and a toroidal one is a "longitude."

Definition IX.6. A set of magnetic coordinates is called Hamada coordinates *if the current density takes the form $J = \nabla I\times\nabla\theta^1 + \nabla G\times\nabla\theta^2$ for some functions I, G of ψ only (in differential forms, the current flux-form $j = i_J\Omega = dI\wedge d\theta^1 + dG\wedge d\theta^2$).*

Definition IX.7. A set of magnetic coordinates is called Boozer coordinates *if the magnetic field takes the form $B = I\nabla\theta^1 + G\nabla\theta^2 + K\nabla\psi$ with I, G functions of ψ only (in differential forms, $B^\flat = Id\theta^1 + Gd\theta^2 + Kd\psi$).*

Note that in each case, there is freedom to choose where to put the origin of (θ^1,θ^2) on each flux surface.

Theorem IX.8. *If u is a quasi-symmetry for an MHS field B, then AL coordinates for $[u,B] = 0$ give Hamada coordinates, and $|B|$ is constant along a system of straight lines in these coordinates.*

Proof. Quasi-symmetry implies $[u,B] = 0$ (Theorem IV.4). Construct AL coordinates (θ^1,θ^2) for this commutation relation and augment to 3D by ψ. Then, both u and B are constant vector fields on each flux surface in these coordinates, i.e., $u = u^1\partial_1 + u^2\partial_2$ and $B = B^1\partial_1 + B^2\partial_2$, with the coefficients being functions of ψ only, where ∂_i is short for ∂_{θ^i}. In particular, the B-lines are straight, so (θ^1,θ^2) are magnetic coordinates.

Under the MHS condition with p constant on flux surfaces, J is given by (64). Thus, the current 2-form is $j = i_J\Omega = -p' i_u\Omega - C' i_B\Omega$. Now, Ω can be written as $\Omega = \mathcal{J} d\psi\wedge d\theta^1\wedge d\theta^2$, where $\mathcal{J}$ is the Jacobian of the coordinate system. The Jacobian in these coordinates is a function of ψ only because $\operatorname{div} u = \operatorname{div} B = 0$ implies $\operatorname{div}\partial_i = 0$ for $i = 1,2$, so $0 = di_{\partial_1}\Omega = -d\mathcal{J}\wedge d\psi\wedge d\theta^2$ and $0 = di_{\partial_2}\Omega = d\mathcal{J}\wedge d\psi\wedge d\theta^1$. Thus, $d\mathcal{J}\wedge d\psi = 0$, and hence, $\mathcal{J}$ is a function of ψ. Finally,

$$j = \mathcal{J}(-p'(u^2 d\psi\wedge d\theta^1 - u^1 d\psi\wedge d\theta^2) - C'(B^2 d\psi\wedge d\theta^1 - B^1 d\psi\wedge d\theta^2)), \quad (67)$$

and so has the form $j_1 d\psi\wedge d\theta^1 + j_2 d\psi\wedge d\theta^2$ for some $j_i(\psi)$, which (together with being magnetic coordinates) is the defining condition of Hamada coordinates. Because u is constant in these coordinates on a flux surface, it follows from $L_u|B| = 0$ that $|B|$ is constant along a system of straight lines. □

Note that J is also constant in Hamada coordinates on each flux surface because we proved in (64) that it is a linear combination (by functions of ψ) of u and B, which are constant on each flux surface. Thus, instead of making AL coordinates for $[u,B] = 0$, we could equally well make AL coordinates for $[J,B] = 0$ or for $[J,u] = 0$, the first of which holds for any MHS field and the second of which holds for any quasi-symmetric field. To prove that $[J,B] = 0$ for any MHS field, note simply that

$$i_{[J,B]}\Omega = i_J L_B\Omega - L_B i_J\Omega = -L_B dB^\flat = -dL_B B^\flat = 0 \quad (68)$$

because $L_B B^\flat = d(p + |B|^2)$. $[J,B] = 0$ gives Hamada coordinates, and $|B|$ is constant along a system of straight lines if the field is quasi-symmetric.

Theorem IX.9. *If u is a quasi-symmetry for an MHS field B, then AL coordinates for $[u, B/|B|^2] = 0$ give Boozer coordinates, and $|B|$ is constant along a system of straight lines in these coordinates.*

Proof. Quasi-symmetry implies $[u, B/|B|^2] = 0$ (Theorem IV.6). Construct AL coordinates (θ^1,θ^2) for this commutation relation and augment by ψ. Then, both u and $B/|B|^2$ are constant vector fields on each flux surface in these coordinates. In particular, the B-lines are straight, even if $|B|$ might vary along them. So these are magnetic coordinates again.

Under the MHS condition with p constant on flux surfaces, we see that the AL coordinates have the additional property that lines of $\nabla\psi\times B$ are straight too. This is because $u\cdot B$ is constant on flux surfaces, so $\nabla\psi\times B/|B|^2 = u - CB/|B|^2$ [coming from (65)] is a constant vector field on each flux surface. This is a property of Boozer coordinates, but the defining property of Boozer coordinates (beyond being magnetic coordinates) is that $B^\flat = Id\theta^1 + Gd\theta^2 + Kd\psi$ for some functions I, G, K with I, G functions of ψ only. Now, we prove that this holds for our AL coordinates. Without the constraints on I and G, $B^\flat$ has the above form because the coordinate 1-forms form a basis for the cotangent space. As before, we proved that $C = u \cdot B$ is constant on flux surfaces, so $C = i_u B^\flat = Iu^1 + Gu^2$. In addition, $1 = i_{B/|B|^2} B^\flat = Ih^1 + Gh^2$, where $h^j = B^j/|B|^2$. The coefficients (u^j, h^j) in these two equations are functions of ψ only, so it follows that I and G are too. Thus, the AL coordinates for $[u, B/|B|^2] = 0$ in an MHS quasi-symmetric field are Boozer coordinates. Because u is constant in these coordinates and $|B|$ is constant along u-lines, it follows that $|B|$ is constant along a system of straight lines on each flux surface. □

Again, we note that the same coordinates are obtained if we start from the commutation relation $[u, \nabla\psi \times B/|B|^2] = 0$, which holds in quasi-symmetry, or from $[B/|B|^2, \nabla\psi \times B/|B|^2] = 0$, whose significance comes from the next result.

Theorem IX.10. *If B is MHS with flux function ψ and p constant on flux surfaces, then $[B/|B|^2, \nabla\psi \times B/|B|^2] = 0$.*

Proof. Write $h = B/|B|^2$ and $k = \nabla\psi \times h$. We apply

$$i_{[h,k]} = L_h i_k - i_k L_h \tag{69}$$

to $d\psi, B^\flat, k^\flat$, in turn, using the operator

$$L_h = |B|^{-2} L_B + d|B|^{-2} \wedge i_B, \tag{70}$$

which reduces to $|B|^{-2}L_B$ on 0-forms. In the first case, we get immediately $i_{[h,k]}d\psi = 0$, since $L_h d\psi = 0$ and $i_k d\psi = 0$. Next, inserting the MHS condition in the form $L_B B^\flat = d(p + |B|^2)$, we have

$$i_{[h,k]} B^\flat = -i_k L_h B^\flat = -i_k(|B|^{-2} d|B|^2 + |B|^2 d|B|^{-2}) = 0, \tag{71}$$

as well. For the last case, note first that

$$i_{[h,k]} k^\flat = L_h |k|^2 - i_k L_h k^\flat = |B|^{-2} L_B i_k k^\flat - |B|^{-2} i_k L_B k^\flat = |B|^{-2} i_{[B,k]} k^\flat \tag{72}$$

since $i_B k^\flat = 0$. However, using $i_k\Omega = d\psi \wedge h^\flat$, we also see that

$$i_B i_{[B,k]}\Omega = i_B(L_B i_k \Omega - i_k L_B \Omega) = i_B(d\psi \wedge L_B h^\flat) = -d\psi \wedge L_B 1 = 0, \tag{73}$$

since $L_B\psi = 0$, $L_B\Omega = 0$, and $i_B d\psi = 0$. Therefore, the vector field $[B, k]$ is parallel to B, and so $i_{[h,k]}k^\flat = 0$ too. Since $(\nabla\psi, B, k)$ form a basis for the tangent space, we deduce that $[h, k] = 0$. □

Thus, starting with an MHS field B with a flux function ψ and p constant on flux surfaces, one can construct Boozer coordinates as AL coordinates for the commutation relation of Theorem IX.10. If B is in addition quasi-symmetric, then $|B|$ is constant along a system of straight lines, namely, the u-lines, because $L_u|B| = 0$ and u, given by (65), is a constant vector field on each flux surface.

The standard approaches to quasi-symmetry are based on a symmetry (ignorable coordinate) of the gyro-averaged Lagrangian in Boozer or Hamada coordinates. There are two guiding-center Lagrangians that appear in the quasi-symmetry literature. One of these is given by

$$L(X, \dot{X}) = \frac{1}{2} m(b \cdot \dot{X})^2 + eA \cdot \dot{X} - \mu|B(X)|, \tag{74}$$

with μ being a positive real parameter, which bears a resemblance to a well-known Lagrangian for the (non-gyro-averaged) Lorentz force, namely, $L_{LL}(q, \dot{q}) = m|\dot{q}|^2/2 + eA \cdot \dot{q}$. However, the resemblance is misleading. In contrast to L_{LL}, the Lagrangian (74) is degenerate, which means that its associated Legendre transformation is not invertible. A consequence of this degeneracy is that the Euler–Lagrange equations derived from (74) do not uniquely specify a second-order system of ordinary differential equations on TQ. Nevertheless, the Lagrangian (74) is still a valid Lagrangian for FGCM in a weaker sense. Indeed, if X is the Q-component of a solution of the system (12) and (13), then X must be a critical point of the action functional defined by integrating (74) with respect to time. That said, the fundamental role of (74) in the theory of charged particle dynamics in strong magnetic fields is unclear. While (74) is a good Lagrangian for FGCM, it is not a good Lagrangian for higher-order guiding-center dynamics. In fact, there are no known extensions of (74) to higher-order guiding-center theory.

The other guiding-center Lagrangian that appears in the quasi-symmetry literature is

$$L(X, v_\parallel, \dot{X}, \dot{v}_\parallel) = (eA + mv_\parallel b) \cdot \dot{X} - (\frac{1}{2} mv_\parallel^2 + \mu|B|), \tag{75}$$

which is due to Littlejohn (1983). This Lagrangian bears resemblance to another Lagrangian for the Lorentz force, $L_{PP}(q, v, \dot{q}, \dot{v}) = (eA + mv) \cdot \dot{q} - m|v|^2/2$. In this case, the resemblance is more than superficial. After first writing the variational principle for the Lorentz force in Poincaré–Cartan form in the extended phase space (q, v, t) as $\delta \int \alpha_{PP} = 0$, with $\alpha_{PP} = (eA + mv) \cdot dq - (m|v|^2/2)dt$ [see Arnol'd (1978)], Littlejohn used sequences of near-identity transformations in phase space and asymptotic symplectic reduction to systematically derive (75). In fact, Littlejohn's derivation naturally produces generalizations of (75) that account for all higher-order effects in guiding-center theory. In this sense, it is clear that the Lagrangian (75) plays a fundamental role in the theory of strongly magnetized particle dynamics, in contrast to the Lagrangian (74).

In a similar way, the variational principle for (75) in extended state space $(X, v_\parallel, t)$ is expressed as $\delta \int_\gamma \alpha = 0$ over variations of compact support of paths γ with

$$\alpha = eA^\flat + mv_\parallel b^\flat - (\frac{1}{2} m v_\parallel^2 + \mu|B|)\, dt. \tag{76}$$

The vector field V resulting from such a variational principle is given by the unique choice $(V, 1)$ in the kernel of $d\alpha$ (which is one-dimensional because the extended state space is odd-dimensional and B is assumed non-zero). Then, a continuous symmetry is defined to be a vector field U such that $L_U \alpha = dZ$ for some function Z. This is because flowing with such a vector field does not change the variational principle (actually, one could replace dZ by any closed 1-form, but it is conventional to take it exact). It follows, as in the standard Noether theorem, that $i_U d\alpha = d(Z - i_U\alpha)$ and so $i_{(V,1)} d(Z - i_U\alpha) = -i_U i_{(V,1)} d\alpha = 0$, so $K = Z - i_U\alpha$ is conserved by $(V, 1)$. If $U = (u, 0)$, then

$$i_U \alpha = e\, i_u A^\flat + mv_\parallel i_u b^\flat = e\, u \cdot A + mv_\parallel u \cdot b. \tag{77}$$

Finally, we address the converse question: given a quasi-symmetric system in the standard sense, identify the quasi-symmetry in our sense.

Theorem IX.11. *If the magnetic field B has a flux function ψ and is MHS with p constant on flux surfaces, density of irrational surfaces, $p'(\psi) \neq 0$ a.e. and $|B|$ is constant along a family of straight lines in Hamada coordinates, then it is quasi-symmetric with $u = -(J + C'B)/p'$ for a function C of ψ such that $C = u \cdot B$.*

Proof. For an MHS field, $[J, B] = 0$. If $dp \neq 0$, then J, B are independent. Take AL coordinates for this commutation relation. By the discussion after Theorem IX.8, they are Hamada coordinates. If $|B|$ is constant along a family of straight lines, then that implies $|B|$ is constant along

$$u = \kappa J + \lambda B \tag{78}$$

for some functions κ, λ of p or equivalently of ψ if a different flux function has been chosen with the same level sets. Their ratio is determined by the lines of constant $|B|$, but their magnitudes are otherwise free.

Now, $i_u i_B \Omega = \kappa i_J i_B \Omega = -\kappa dp = -\kappa p' d\psi$. Therefore, $L_u \beta = 0$. Choose, in particular, $\kappa = -1/p'$ to obtain $i_u i_B \Omega = d\psi$.

It remains to prove that $L_u B^\flat = 0$. From (78), $i_u i_J \Omega = \lambda i_B i_J \Omega = \lambda dp = \lambda p' d\psi$, by the MHS equation. Also $[u, B] = 0$ since $[J, B] = 0$. From the MHS condition again $L_B B^\flat = d(p + |B|^2)$, and $[u, B] = 0$, we have $L_B(u \cdot B) = i_u L_B B^\flat = i_u d(p + |B|^2) = 0$ because both p and $|B|$ are constant along u. Then, density of irrational surfaces implies, assuming continuity, that $u \cdot B$ is constant on flux surfaces. In other words, $u \cdot B$ is a function C of ψ. Therefore,

$$L_u B^\flat = i_u dB^\flat + d i_u B^\flat = i_u i_J \Omega + d(u \cdot B) = (\lambda p' + C') d\psi. \tag{79}$$

Thus, choosing $\lambda = -C'/p'$, we satisfy the last condition for quasi-symmetry. □

Remark IX.12. In the previous proof, we can show that $L_u B^\flat = 0$ using circle averaging instead, as follows. First, we note that if $|B|$ is not constant on a flux surface, then the u-lines are closed. If, exceptionally, $|B|$ is constant on a flux surface, then the ratio is undetermined, but we can choose it to make the u-lines closed. The period τ of the u-lines is constant on flux surfaces but, in general, varies with ψ. We are free to simultaneously scale κ and λ by any function of ψ, however. Thus, we can scale them to make $\tau = 2\pi$. With this choice, we can now apply circle averaging [defined by (60)] to (79). The average of the left-hand side is zero, being L_u of something. All of λ, p, and C are constant along u. Thus, the average of the right-hand side is just itself. Consequently, $0 = (\lambda p' + C')d\psi$. It follows that $L_u B^\flat = 0$.

Theorem IX.13. *If magnetic field B has a flux function ψ and is MHS with p constant on flux surfaces and density of irrational surfaces, and $|B|$ is constant along a family of straight lines in Boozer coordinates, then it is quasi-symmetric with $u = (CB + \nabla\psi \times B)/|B|^2$ for a function C of ψ such that $(C^2/2)' = -|u|^2 p' - u \cdot J$.*

Proof. For an MHS field with a flux function ψ, we proved that $[B/|B|^2, \nabla\psi \times B/|B|^2] = 0$ (Theorem IX.10), and the corresponding AL coordinates are Boozer (discussion after Theorem IX.10). If $|B|$ is constant along a family of straight lines in these coordinates, then there are functions C, λ of ψ only such that $|B|$ is constant along

$$u = (CB + \lambda \nabla\psi \times B)/|B|^2. \tag{80}$$

The ratio of C, λ is determined by the lines of constant $|B|$, but their magnitudes are otherwise free. We see that $C = u \cdot B$; hence, $u \cdot B$ is constant on flux surfaces.

Now, $B \times u = \lambda \nabla\psi$, so we obtain $i_u i_B \Omega = \lambda d\psi$ and $L_u \beta = 0$ accordingly. Let us choose $\lambda = 1$ to obtain $i_u i_B \Omega = d\psi$.

It remains to prove that $L_u B^\flat = 0$. The MHS equation can be written as $i_B i_J \Omega = dp$. Thus, from the above expression for u, we deduce that

$$i_u i_J \Omega = \frac{1}{|B|^2}(C i_B i_J \Omega - i_J(d\psi \wedge B^\flat) = \frac{1}{|B|^2}(C dp + (i_J B^\flat) d\psi) = \kappa d\psi \tag{81}$$

since $i_J d\psi = 0$, with

$$\kappa = (Cp' + J \cdot B)/|B|^2 = (|u|^2 p' + u \cdot J)/C. \tag{82}$$

Therefore,

$$L_u B^\flat = i_u i_J \Omega + d(u \cdot B) = (\kappa + C') d\psi. \tag{83}$$

Next, note that $[u, B/|B|^2] = 0$ because u is given by (80), $B \cdot \nabla C = 0$, and $[B/|B|^2, \nabla\psi \times B/|B|^2] = 0$. However, $|B|$ is constant along u, so $[u, B] = 0$. Then, apply L_B to (81). We have $[u, B] = 0$ and $[J, B] = 0$ from the MHS condition, and $L_B \Omega = 0$. Hence, $L_B \kappa = 0$. Using density of irrational surfaces, it follows, assuming continuity, that κ is constant on flux surfaces, i.e., $\kappa = \kappa(\psi)$. Then, choose $\kappa = -C'$ to obtain $L_u B^\flat = 0$. Inserting this in (82) proves the result for $(C^2/2)'$. □

Remark IX.14. Alternatively, in the previous proof, we can show that $L_u B^\flat = 0$ using circle-averaging, as follows. If $|B|$ is not constant on a flux surface, then the u-lines are closed. If it is constant, then we can choose the ratio $C : \lambda$ to make the u-lines closed. In either case, the period of the u-lines is constant on a flux surface. We can scale C, λ simultaneously by a function of ψ to make the period 2π. Apply circle averaging to (83) and use κ, C constant along u to obtain $0 = (\kappa + C') d\psi$. Thus, $L_u B^\flat = 0$.

Another common treatment of quasi-symmetry for an MHS field with flux function and p constant on flux surfaces [e.g., Simakov and Helander (2011)] is based on the relation

$$B \times \nabla\psi = E\,\nabla\psi \times \nabla|B| + FB, \tag{84}$$

where $E = -|B|^2/B \cdot \nabla|B|$ and $F = B \times \nabla\psi \cdot \nabla|B|/(B \cdot \nabla|B|)$. Quasi-symmetry in this approach is then formulated as F being constant on flux surfaces,

$$F = F(\psi). \tag{85}$$

This setup fits in our framework because Eq. (84) is none other than $i_u i_B \Omega = d\psi$ and $L_u|B| = 0$ combined together, and condition (85) says that $u \cdot B$ is constant on flux surfaces in the MHS case. To see the first one, cross $B \times u = \nabla\psi$ with B to obtain $B \times \nabla\psi = (u \cdot B)B - |B|^2 u$ and insert (66) for u. The second one follows from Theorem IX.2 since the function F is precisely $u \cdot B$, i.e., $F = C$ for MHS fields.

Finally, we show that $L_u b^\flat = 0$ implies that $\int_c dl$ is constant when the curve c is drawn from any continuous family of field line segments within a given flux surface and with fixed endpoint values of $|B|$. Let $\gamma : [s_0, s_1] \to \mathbb{R}^3$ be the restriction of a field line to an interval $[s_0, s_1]$ such that $|B|(\gamma(s_0)) = k_0$ and $|B|(\gamma(s_1)) = k_1$, where $k_0, k_1 \in \mathbb{R}_+$. If ϕ_λ is the u-flow, then $\gamma_\lambda = \phi_\lambda \circ \gamma$ is a field line segment contained in the same flux surface as γ for each λ. In addition, the integral $I_\lambda = \int_{\gamma_\lambda} dl$ is independent of λ because

$$\frac{d}{d\lambda}\int_{\phi_\lambda \circ \gamma} dl = \frac{d}{d\lambda}\int_{\phi_\lambda \circ \gamma} b^\flat = \frac{d}{d\lambda}\int_\gamma \phi_\lambda^* b^\flat = \int_\gamma L_u b^\flat = 0. \tag{86}$$

Therefore, the arc lengths of the field line segments γ_λ are all the same. Moreover, because $L_u|B| = 0$, the endpoint values of $|B|$ for γ_λ are independent of λ, i.e., $|B|(\gamma_\lambda(s_0)) = k_0$ and $|B|(\gamma_\lambda(s_1)) = k_1$ for each λ. The desired result now follows upon noting that any continuous family of field line segments in a given flux surface with fixed endpoint values of $|B|$ can be generated by flowing some field line segment along u.

X. QUASI-SYMMETRIC GRAD–SHAFRANOV EQUATION

In the axisymmetric case, magnetohydrostatics is reduced to a nonlinear elliptic partial differential equation called the Grad–Shafranov (GS) equation (Grad and Rubin, 1958; Shafranov, 1958) [but previous versions were published by Lüst and Schlüter in 1957 and by Chandrasekhar and Prendergast in 1956, and in the fluids context, the analogous equation was published by Hicks in 1899]. It takes as input two functions $p(\psi)$ and $C(\psi)$ and is an equation for $\psi(r, z)$ in cylindrical polar coordinates. The GS equation has a nice variational principle

(Berestycki and Brézis, 1980) (but probably there are earlier references), reasonable existence theory for solutions (Berestycki and Brézis, 1980; Ambrosetti and Mancini, 1980), and well developed codes for its numerical solution, e.g, Jardin (2010), Chap. 4.

Here, we generalize the GS equation to magnetohydrostatics with a general quasi-symmetry u.

First, we derive a pre-GS equation, which does not assume magnetohydrostatics. Furthermore, the only part of quasi-symmetry that it uses is $i_u i_B \Omega = d\psi$.

Theorem X.1. *If $i_u i_B \Omega = d\psi$, then*

$$\Delta\psi - \frac{u \times v}{|u|^2} \cdot \nabla\psi + \frac{u \cdot v}{|u|^2}\, u \cdot B - u \cdot J = 0, \tag{87}$$

where $\Delta = div\,\nabla$, $v = curl\, u$, and $J = \operatorname{curl} B$.

Proof. We have $B \times u = \nabla\psi$, and so

$$B^\flat \wedge u^\flat = i_{B\times u}\Omega = i_{\nabla\psi}\Omega. \tag{88}$$

Applying d to the above equation, we obtain

$$\Delta\psi\ \Omega = d(B^\flat \wedge u^\flat) = dB^\flat \wedge u^\flat - B^\flat \wedge du^\flat = i_J\Omega \wedge u^\flat - B^\flat \wedge i_v\Omega = (u \cdot J - B \cdot v)\,\Omega. \tag{89}$$

Hence, by non-degeneracy of Ω,

$$\Delta\psi = u \cdot J - B \cdot v. \tag{90}$$

For the last term of (90), contract (88) with v and u to find

$$v \times u \cdot \nabla\psi = i_u i_v (B^\flat \wedge u^\flat) = i_u((B \cdot v)u^\flat - (u \cdot v)B^\flat) = (B \cdot v)|u|^2 - (u \cdot v)(u \cdot B). \tag{91}$$

Hence,

$$B \cdot v = (v \times u \cdot \nabla\psi + (u \cdot v)(u \cdot B))/|u|^2. \tag{92}$$

Substituting this into (90), we deduce Eq. (87). □

An immediate consequence of $i_u i_B \Omega = d\psi$ is also the additional equation $u \cdot \nabla\psi = 0$, which comes by contracting with u.

Next, we derive two further conditions, which follow from $\operatorname{div} B = 0$ and $L_u B^\flat = 0$, assuming, in addition, the properties $\operatorname{div} u = 0$ and $L_u(u \cdot B) = 0$ of a quasi-symmetry.

Theorem X.2. *If $i_u i_B \Omega = d\psi$, $\operatorname{div} u = 0$, and $L_u(u \cdot B) = 0$, then $div\, B = 0$ iff $B \cdot w = 0$, where $w^\flat = L_u u^\flat$.*

Proof. Recalling $\beta = i_B\Omega$, we have

$$|u|^2 d\beta = (i_u d\beta) \wedge u^\flat = (L_u\beta - d i_u\beta) \wedge u^\flat = (i_{[u,B]}\Omega + i_B L_u\Omega) \wedge u^\flat = (i_{[u,B]}u^\flat)\,\Omega \tag{93}$$

$$= (L_u i_B u^\flat - i_B L_u u^\flat)\,\Omega = -(B \cdot w)\,\Omega \tag{94}$$

because $d i_u \beta$, $L_u\Omega$, and $L_u i_B u^\flat$ are zero. $d\beta = (\operatorname{div} B)\Omega$, hence the result. □

Theorem X.3. *Under the assumptions of the previous theorem, $L_u B^\flat = 0$ iff $B \times w = [u, \nabla\psi]$.*

Proof. Note first that since $i_u L_u B^\flat = L_u i_u B^\flat = 0$, we have $i_u(L_u B^\flat \wedge u^\flat) = -|u|^2 L_u B^\flat$. Thus, $L_u B^\flat$ vanishes iff $L_u B^\flat \wedge u^\flat$ does. Applying then L_u to (88), we obtain

$$(L_u B^\flat) \wedge u^\flat = L_u i_{\nabla\psi}\Omega - B^\flat \wedge L_u u^\flat = i_{[u,\nabla\psi]}\Omega - i_{B\times w}\Omega, \tag{95}$$

since $L_u\Omega = 0$. Ω is non-degenerate, hence the result. □

Thus, the pre-GS equation requires extra conditions to guarantee that $\operatorname{div} B = 0$ and $L_u B^\flat = 0$. Both of them are automatic in the case of isometries, that is, if u is a Killing field. To see this, write $w^\flat = L_u u^\flat = i_u L_u g$ and $[u, \nabla\psi]^\flat = L_u(d\psi) - i_{\nabla\psi} L_u g$, recalling Lemma VII.5, and take into account the first condition, $L_u\psi = 0$. In the general case, however, they appear to be non-trivial additional conditions.

The condition of Theorem X.2 is also automatic if $w = 0$. For a quasi-symmetry, the latter was precisely the condition for u to be a Killing field (Theorem VII.3). The next result shows that this is also true if instead u satisfies the first and third supplementary conditions.

Theorem X.4. *If* $\operatorname{div} u = 0$, $L_u\psi = 0$, *and* $B \times w = [u, \nabla\psi]$, *then* $w = 0$ *iff* $L_u g = 0$.

Proof. If $L_u g = 0$, then straightforwardly $w^\flat = L_u u^\flat = i_u L_u g = 0$.
Let $w = 0$. Then, $[u, \nabla\psi] = 0$. Using $L_u\Omega = 0$ and $L_u d\psi = dL_u\psi = 0$,

$$i_{[u,u\times\nabla\psi]}\Omega = L_u i_{u\times\nabla\psi}\Omega - i_{u\times\nabla\psi}L_u\Omega = L_u(u^\flat \wedge d\psi) = L_u u^\flat \wedge d\psi = 0. \tag{96}$$

Thus, $[u, u\times\nabla\psi] = 0$ too, since Ω is non-degenerate.
Consider then the basis $(u, \nabla\psi, u\times\nabla\psi)$. $i_u L_u g = L_u u^\flat = 0$. Furthermore, since u commutes with $\nabla\psi$ and $u\times\nabla\psi$,

$$i_{\nabla\psi}L_u g = L_u i_{\nabla\psi}g = L_u d\psi = dL_u\psi = 0, \tag{97}$$

$$i_{u\times\nabla\psi}L_u g = L_u i_{u\times\nabla\psi}g = L_u(u\times\nabla\psi)^\flat = L_u i_{\nabla\psi}i_u\Omega = i_{\nabla\psi}i_u L_u\Omega = 0, \tag{98}$$

using $L_u\psi = 0$ and $L_u\Omega = 0$ again. Hence, $L_u g = 0$. □

We now turn to combining quasi-symmetry with magnetohydrostatics, obtaining another of our main results.

Theorem X.5. *If MHS field B is quasi-symmetric with quasi-symmetry u, flux function ψ, and p a function of ψ, then*

$$\Delta\psi - \frac{u\times v}{|u|^2}\cdot\nabla\psi + \frac{u\cdot v}{|u|^2}C(\psi) + CC'(\psi) + |u|^2 p'(\psi) = 0, \tag{99}$$

where $v = \operatorname{curl} u$ *and* $C = u\cdot B$.

Proof. In the MHS case with p constant on flux surfaces, $u\cdot B$ is a function $C(\psi)$ from Theorem IX.2, and $u\cdot J = -CC' - |u|^2 p'$ from (64). Substitute these into the pre-GS equation (87) to obtain (99). □

Equation (99) is our quasi-symmetric Grad–Shafranov equation. For given u, $C(\psi)$, and $p(\psi)$, it comprises a semilinear elliptic PDE for the dependent variable ψ. Solutions of (99), however, do not necessarily give MHS fields. There are several additional conditions that are required.

First of all, Eq. (99) needs supplementing by the condition $L_u\psi = 0$, equivalently

$$u\cdot\nabla\psi = 0, \tag{100}$$

which restricts (99) to u-invariant solutions, reducing it effectively to 2D. For the special case of axisymmetry, $u = r\hat{\phi}$, then $|u| = r$, $v = 2\hat{z}$, $u\cdot v = 0$, and $u\times v = 2r\hat{r}$, so the usual GS equation is recovered in cylindrical polar coordinates. Similarly, for the case of helical symmetry, $u = r\hat{\phi} + l\hat{z}$, where l is a constant, then $|u| = \sqrt{r^2 + l^2}$, $v = 2\hat{z}$, $u\cdot v = 2l$, and $u\times v = 2r\hat{r}$, so the helical GS equation (Johnson *et al.*, 1958) is obtained.

Recalling Theorem IX.3, the magnetic field B can be obtained by formula (63). If $w = 0$, as for axisymmetry and helical symmetry, no further conditions beyond (99) and (100) are required for MHS fields.

If $w \neq 0$, however, it is not automatic from (63) that $\operatorname{div} B = 0$ nor that $L_u B^\flat = 0$. Thus, in general, one must add the conditions of Theorems X.2 and X.3 to ensure them. The first one reads

$$(u\times w)\cdot\nabla\psi - (u\cdot w)C(\psi) = 0 \tag{101}$$

as we can see by replacing B from (63) into $B\cdot w = 0$. For the non-isometry case, the second one can be reduced to

$$[v\times w - 2(w\cdot\nabla)u]\cdot\nabla\psi = 0, \tag{102}$$

$$[(u\cdot v)w + 2((u\times w)\cdot\nabla)u]\cdot\nabla\psi + |w|^2 C(\psi) = 0, \tag{103}$$

as the next result shows. It is worth noting that, owing to $[u, \nabla\psi]$, the original condition in this case involves second-order partial differential equations, but they can be reduced to first-order ones by making use of (100), as described in (104) below. Thus, in the end, the second-order quasi-symmetric Grad–Shafranov equation is augmented by four first-order quasilinear partial differential equations given by (100)–(103).

Theorem X.6. *Let B be of the form (63) with $C \neq 0$. If u is not locally a Killing field, then $B \times w = [u, \nabla\psi]$ reduces to (102)* and *(103) under (100)* and *(101).*

Proof. First of all, for any vector field X, we have

$$[u,\nabla\psi]\cdot X = i_{[u,\nabla\psi]}X^\flat = (L_u i_{\nabla\psi} - i_{\nabla\psi}L_u)X^\flat = L_u i_X d\psi - i_{\nabla\psi}L_u X^\flat$$
$$= (i_{[u,X]} + i_X L_u)d\psi - i_{\nabla\psi}L_u X^\flat = i_{\nabla\psi}([u,X]^\flat - L_u X^\flat), \tag{104}$$

using $L_u d\psi = dL_u\psi = 0$ from (100). Moreover, switching to vector calculus,

$$[u,X]^\flat - L_u X^\flat = [(u\cdot\nabla)X - (X\cdot\nabla)u]^\flat - i_u dX^\flat - d(i_u X^\flat)$$
$$= [(u\cdot\nabla)X - (X\cdot\nabla)u + u\times\operatorname{curl}X - \nabla(u\cdot X)]^\flat$$
$$= [v\times X - 2(X\cdot\nabla)u]^\flat. \tag{105}$$

Now, if $u\times w = 0$, then $u\cdot w = 0$ from (101) for $C\neq 0$, and so $w = 0$; hence, $L_u g = 0$ from Theorem X.4. Therefore, if u is not locally a Killing field, $(u, w, u\times w)$ is a basis. Then, project $B\times w = [u,\nabla\psi]$ to the directions $X = u, w, u\times w$ and use (104) and (105).

For $X = u$, we see directly from (104) that the projection of $B\times w = [u,\nabla\psi]$ to u is trivially satisfied,

$$0 = u\cdot(B\times w - [u,\nabla\psi]) = w\cdot(u\times B) + i_{\nabla\psi}L_u u^\flat = -w\cdot\nabla\psi + w\cdot\nabla\psi. \tag{106}$$

For $X = w$, we obtain (102),

$$0 = w\cdot(B\times w - [u,\nabla\psi]) = -i_{\nabla\psi}[v\times w - 2(w\cdot\nabla)u]^\flat. \tag{107}$$

For $X = u\times w$, using (100) and (101) in its original form $B\cdot w = 0$, we arrive at (103),

$$0 = u\times w\cdot(B\times w - [u,\nabla\psi])$$
$$= [(B\times w)\times u]\cdot w - i_{\nabla\psi}[v\times(u\times w) - 2((u\times w)\cdot\nabla)u]^\flat$$
$$= [Cw - (w\cdot u)B]\cdot w - i_{\nabla\psi}[(v\cdot w)u - (v\cdot u)w - 2((u\times w)\cdot\nabla)u]^\flat$$
$$= C|w|^2 + i_{\nabla\psi}[(v\cdot u)w + 2((u\times w)\cdot\nabla)u]^\flat. \tag{108}$$

□

The four first-order partial differential equations (100)–(103) in 3D imply a linear dependence among them. Without going into the tedious details, we comment that the latter though is equivalent to the degeneracy of $L_u g$, which we saw earlier in Theorem VII.2.

Finally, the current density J can be found from (64). However, $J = \operatorname{curl} B$ is not automatic either. Still for solutions of (87) and therefore (99), this amounts to $L_u B^\flat = 0$ again as the next result shows.

Theorem X.7. *Let B, J be of the form (63) and (64). On the set of solutions of the pre-GS equation, $L_u B^\flat = 0$ iff $J = curl\,B$.*

Proof. Write $L_u B^\flat = i_u dB^\flat + d(i_u B^\flat)$ and $J = \operatorname{curl} B$ as $s = 0$, where $s = dB^\flat - i_J\Omega$. Using $u\cdot B = C$ and $i_u i_J\Omega = -C' d\psi$, we see that

$$L_u B^\flat = i_u s. \tag{109}$$

For the converse, note that $|u|^2 s = i_u(u^\flat\wedge s) + u^\flat\wedge i_u s$ and

$$u^\flat\wedge s = -d(u^\flat\wedge B^\flat) + i_v\Omega\wedge B^\flat - (i_J u^\flat)\Omega = (\Delta\psi + B\cdot v - J\cdot u)\,\Omega. \tag{110}$$

Thus, in light of (90), we deduce that on solutions of (87),

$$s = -|u|^{-2}u^\flat\wedge L_u B^\flat, \tag{111}$$

which completes the proof. □

In conclusion, the system of Eqs. (99)–(103) describes the conditions that a quasi-symmetry u and the corresponding flux function ψ must satisfy in magnetohydrostatics.

XI. VARIATIONAL PRINCIPLE FOR THE QUASI-SYMMETRIC GS EQUATION

Two questions arise: (i) does (99) has solutions, and (ii) how do we incorporate the supplementary conditions (100)–(103)?

In this section, we principally address the first question.

There are a number of relevant results on the existence theory for semilinear elliptic PDEs, for example, Theorem 15.12 in the book of Gilbarg and Trudinger (2001) and Theorem 9.12 in the book of Amann (1976). However, even for the axisymmetric GS equation, there are regimes with no solutions (Ambrosetti and Mancini, 1980) and regimes with more than one solution. We have to do more work to reach definitive conclusions.

In the meantime, however, we address here the question whether (99) has a variational principle because it would be one useful route to prove existence of a minimizer or some other critical point by variational means, cf. Berestycki and Brézis (1980), and to understand the set of solutions. To that end, we resort to the Helmholtz conditions, as formulated in Olver (1993).

For a variational problem of the form $D\mathcal{L}_\psi = 0$ (often written $\delta\mathcal{L}[\psi] = 0$) with $\mathcal{L}[\psi]= \int_Q L(\psi, \nabla\psi)\ dV$ (where L is called the *Lagrangian* and may involve more derivatives) on smooth functions $\psi : Q \to \mathbb{R}$, the *Euler–Lagrange operator* E on smooth functions ψ is defined by writing

$$D\mathcal{L}_\psi v = -(E[\psi], v) \tag{112}$$

for all $v : Q \to \mathbb{R}$ (often written $\delta\psi$) satisfying suitable boundary conditions, where $(f, g) = \int_Q fg dV$ is the standard inner product on $L^2(Q, \mathbb{R})$. So the Euler–Lagrange equations are $E[\psi] = 0$.

The Helmholtz conditions say that a differential operator E on functions $\psi : Q \to \mathbb{R}$ is the Euler–Lagrange operator for some variational problem iff $(f, DEg) = (DEf, g)$ for all functions f, g for which both sides are defined. We write this as $DE^* = DE$, where DE^* is the adjoint operator defined wherever $(DEf, g) = (f, DE^* g)$ makes sense, and refer to such operators as self-adjoint, but ignoring the question of equality of domains that is part of the standard definition.

A catch with applying the Helmholtz conditions is that the variational property is not preserved between equivalent equations, not even, for example, $\lambda E[\psi] = 0$ and $E[\psi] = 0$, where λ is some non-zero function on Q. In fact, the Helmholtz criterion for the left-hand side of (99) as it stands implies the highly restrictive case $u \times v = 0$ since Δ is self-adjoint. However, the axisymmetric and helical cases suggest the use of the factor $\lambda = |u|^{-2}$, so let us consider

$$E[\psi] = |u|^{-2}\Delta\psi - |u|^{-4}u \times v \cdot \nabla\psi + |u|^{-4}u \cdot v\, C(\psi) + |u|^{-2}CC'(\psi) + p'(\psi). \tag{113}$$

Theorem XI.1. *E given by (113) is the Euler–Lagrange operator for some variational problem $\mathcal{L}$ iff $L_u u^\flat = 0$. In this case,*

$$\mathcal{L}[\psi] = \int \left(\frac{1}{2|u|^2}(|\nabla\psi|^2 - C(\psi)^2) + C(\psi)Y \cdot \nabla\psi - p(\psi)\right) dV, \tag{114}$$

where Y is a vector field such that div $Y = u \cdot v/|u|^4$.

Proof. Note first that the last three terms of E are functions of ψ only. Therefore, their derivative is a multiplication operator, which is always self-adjoint. Thus, the problem is reduced to just $E[\psi] = |u|^{-2}\Delta\psi - F \cdot \nabla\psi$, where $F = u \times v/|u|^4$. Now E is a linear differential polynomial, and so $DE = E$, i.e.,

$$DE = |u|^{-2}\Delta - F \cdot \nabla. \tag{115}$$

Integration by parts shows that $DE^* g = \Delta(g|u|^{-2}) + \operatorname{div}(gF)$ for the formal adjoint of DE. Using the identities $\Delta(g|u|^{-2}) = \operatorname{div}\nabla(g|u|^{-2}) = |u|^{-2}\Delta g + 2\nabla|u|^{-2} \cdot \nabla g + g\Delta|u|^{-2}$ and $\operatorname{div}(gF) = g \operatorname{div} F + F \cdot \nabla g$, we arrive at

$$DE^* = |u|^{-2}\Delta - (2G + F) \cdot \nabla - \operatorname{div} G, \tag{116}$$

where $G = -\nabla|u|^{-2} - F$. In other words, $G^\flat = (d|u|^2 + i_u du^\flat)/|u|^4 = L_u u^\flat/|u|^4$. Hence, $DE^* = DE$ iff $G = 0$, i.e., $L_u u^\flat = 0$.

The first term of (114) introduces $-|u|^{-2}\Delta\psi - \nabla|u|^{-2} \cdot \nabla\psi$ into the Euler–Lagrange operator. For $L_u u^\flat = 0$, $\nabla|u|^{-2}$ reduces to $-F$, and so we recover the first two terms of (113). The third term of the Lagrangian yields $-C\operatorname{div} Y = -Cu \cdot v/|u|^4$, and the remaining terms easily restore the rest of (113). □

Note that existence of a vector field Y such that div $Y = u \cdot v/|u|^4$ may look difficult to satisfy, but it can be expressed equivalently as saying that $a = u^\flat \wedge du^\flat/|u|^4$ is exact, since $du^\flat = i_v\Omega$ and so $u^\flat \wedge du^\flat = (i_v u^\flat)\Omega$ and $a = di_Y\Omega$ accordingly. In this way, we see that it is not much of a restriction because a is automatically closed, being a top-form.

In the case of axisymmetry where $u \cdot v = 0$, the variational functional (114) for $Y = 0$ recovers the Lagrangian (Berestycki and Brézis, 1980) for the usual GS equation in cylindrical polar coordinates.

In the case of helical symmetry, $u \cdot v/|u|^4 = 2l/(r^2 + l^2)^2$, then $Y = -l/(r(r^2 + l^2))\hat{r}$, so we derive a Lagrangian for the helical GS equation.

Recall, however, Theorem X.4 for u satisfying the supplementary conditions. In this case, unfortunately, $L_u u^\flat = 0$ is the condition for u to be a Killing field and thus in Euclidean space holds only if u generates an orbit of $SE(3)$. Then, we are back to the axisymmetric GS equation (rejecting translations and the helical case because they do not have bounded flux surfaces). This is where the second question comes in. Can the extra conditions be incorporated as constraints in a variational principle? Perhaps, as in Sec. X, one should view the problem as a simultaneous system of equations for ψ and u.

An alternative approach to extending the GS equation to general quasi-symmetry is to circle-average the MHS equation in the form $i_B dB^\flat = dp$ and derive an equation for ψ with respect to the circle-averaged Riemannian metric. Strangely, the resulting GS equation always has a variational principle. However, as in our analysis here, there are extra conditions that must be satisfied, and it is not clear that they can. This will be written in a separate publication.

XII. PERSPECTIVES

Is every quasi-symmetry a Killing field? At least in the Euclidean case? Or at least if one requires magnetohydrostatics? Or might there be some "Kovalevskaya" examples? (Recall that Kovalevskaya found non-axisymmetric integrable cases for the dynamics of a top.) It is not even clear whether these questions are global or local in nature. The isometry condition $L_u g = 0$ is certainly a local one, and this at least hints that the questions may be local. If this is indeed the case, a prolongation analysis based on the Cartan–Kuranishi theorem may be sufficient to provide definitive answers. We will report on such an analysis in a future publication.

The main point of stellarators is to achieve confined guiding-center trajectories without significant toroidal current. Is quasi-symmetry compatible with this goal? The toroidal current enclosed by a flux surface is just $\int_\gamma B^\flat$ round any poloidal loop on the flux surface. It is conventionally written as $2\pi I(\psi)$. Not surprisingly, $\int_{pol} B^\flat = 0$ iff the winding ratio for the current is $\iota_J = 0$. We see no incompatibility between this and quasi-symmetry, but it depends on there being some non-axisymmetric quasi-symmetries.

Quasi-symmetry may be too strong an ideal to aim for. Weaker conditions would suffice for single-particle confinement. Omnigenity (Helander, 2014) is one such concept, which just requires flux surfaces and the average drift in flux function for a guiding center to be zero to leading approximation. This is automatic for circulating particles of ZGCM on irrational flux surfaces but requires a condition for all bouncing particles (Landreman and Catto, 2012) and for circulating particles on rational surfaces (this is not usually recognized but follows by the same arguments as for bouncing particles). Quasi-symmetry implies omnigenity, but perhaps not vice versa, so omnigenity would allow a bit more scope (Landreman and Catto, 2012).

More generally, one invariant torus for FGCM at each value of energy and magnetic moment will confine all those inside it. This might be too weak an approach, however, because particle interactions would lead to exchange of energy and magnetic moment and drive them across the confining tori.

Alternatively, approximate quasi-symmetry may be enough; after all first-order guiding-center motion is only approximate. This can be achieved to some order by near-axis expansions (Landreman *et al.*, 2019), and we intend to address it in more detail.

Quasi-symmetry may also be too weak an ideal to aim for. Even if quasi-symmetric field configurations do exist, they may do a poor job of confining the hot alpha particles generated by thermonuclear burn. The issue is such particles have much larger gyro-radii than bulk plasma particles. In the best case, confinement properties of alpha particles might be well-captured by guiding-center theory with higher-order corrections, in which case it would be interesting to study possible hidden symmetries of these high-order terms. In the worst case, guiding-center theory is useless for describing alpha particle orbits, and other approaches, perhaps based on a more brute-force optimization approach, should be pursued.

Finally, we have restricted attention to symmetries of FGCM of the form $U = (u, 0)$, with u being a vector field on guiding-center position, not involving the parallel velocity. Might there be parallel-velocity-dependent symmetries that render FGCM integrable? Likewise, we have concentrated here entirely on the context of Hamiltonian symmetries. Might there be relevant non-Hamiltonian symmetries? This work is in progress.

ACKNOWLEDGMENTS

This work was supported mainly by a grant from the Simons Foundation (601970, RSM) and partly by the National Science Foundation under Grant No. DMS-1440140 while the first and third authors were in residence at the Mathematical Sciences Research Institute in Berkeley, CA, during the Fall 2018 semester. Research presented in this article was also supported by the Los Alamos National Laboratory LDRD program under Project No. 20180756PRD4.

We are grateful to other members of the MSRI semester for useful comments and of the Simons collaboration team, in particular, to Adam Golab for researching the literature on existence results for solutions of the GS equation and discussion on the supplementary conditions for its quasi-symmetric generalization. We would also like to thank Vassili Gelfreich for Appendix B and Miles Wheeler for helpful suggestions about existence of solutions of the quasi-symmetric GS equation.

APPENDIX A: ADDITIONS OF ELECTROSTATIC POTENTIAL AND RELATIVITY

To add the effect of an electrostatic potential Φ to the theory of this paper, add $e\Phi(q)$ to H in both equations (2) and (7). The drift equations (12) and (13) gain additional terms $b \times \nabla\Phi/\widetilde{B}_\parallel$ and $-e\nabla\Phi/\widetilde{B}_\parallel$, respectively. The conditions of Theorems IV.2 and IV.5 need augmenting by $L_u\Phi = 0$.

To add relativistic effects, one simply replaces $p = mv$ by $p = \gamma mv$ with Lorentz factor $\gamma = (1 - |v|^2/c^2)^{-1/2}$ and the kinetic energy in H by $c\,(m^2c^2 + |p|^2)^{1/2}$. Likewise, the magnetic moment changes to $\mu = p_\perp^2/(2m|B|)$, and the kinetic part of the guiding-center Hamiltonian changes to $c\,(m^2c^2 + p_\parallel^2 + 2m\mu|B|)^{1/2}$. The conditions for quasi-symmetry are unchanged.

Alternatively, taking a fully space–time view and allowing general time-dependent electric and magnetic fields, the equation of motion is

$$\frac{dp}{d\tau} = -\,e\, i_U F, \tag{A1}$$

where F is the Faraday 2-form, U is the contravariant 4-velocity, $p = mU^\flat$ is the covariant 4-momentum, and τ is proper time for the particle. In a time–space coordinate system (t, x, y, z) with locally Minkowski metric $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$,

$$F = B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy + E_x dx \wedge dt + E_y dy \wedge dt + E_z dz \wedge dt \tag{A2}$$

and $U = \gamma(c, v^x, v^y, v^z)$. Noting that two of Maxwell's equations are equivalent to saying F is closed, the motion of the charged particle can be written in Hamiltonian form with

$$H = \frac{|p|_*^2}{2m}, \tag{A3}$$

$$\omega = -\,d\vartheta - eF \tag{A4}$$

on the level set $H = mc^2/2$, where ϑ is the tautological 1-form on T^*Q for space–time Q and $|p|_*$ is the induced metric on cotangents. The number of degrees of freedom is now 4. Guiding-center reduction can still be performed resulting in a 3DoF system. Integrability would now require two further integrals beyond the Hamiltonian. It would be interesting to find out whether time-translation symmetry can be replaced.

APPENDIX B: REGULARITY OF INVARIANTS

Theorem B.1. *(V. Gelfreich) If (i) $F : M^n \longrightarrow N^k$ between manifolds of dimensions $n \geq k$ is C^{n-k+1} and (ii) the set $X \subset M^n$ where dF is not of rank k has measure $\mu_n(X) = 0$, then $\mu_n(Z) = 0$, where $Z = F^{-1}(Y)$ and $Y = F(X)$.*

Proof. Under the C^{n-k+1} condition, Sard's theorem implies $\mu_n(Y) = 0$. Every point x in the complement of X is regular for F so has a neighborhood V_x where F can be used for k components of a smooth coordinate system. Then, $\mu_n(Z \cap V_x) = 0$. Since $\mu_n(X) = 0$, for every $\epsilon > 0$, there exists an open set $U \subset M$ such that $X \subset U$ and $\mu_n(X) < \epsilon$. The complement of U is closed and so can be covered by a countable number of the neighborhoods V_x. Therefore, $\mu_n(Z) < \epsilon$. Hence, $\mu_n(Z) = 0$. □

Corollary B.2. If $F : M^4 \longrightarrow \mathbb{R}^2$, $F = (H, K)$ is C^3, then the union of the non-regular sets has measure zero.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

Amann, H., "Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces," SIAM Rev. **18**, 620–709 (1976).

Ambrosetti, A. and Mancini, G., "A free boundary problem and a related semilinear equation," Nonlinear Anal. **4**, 909–915 (1980).

Arnol'd, V. I., *Mathematical Methods of Classical Mechanics*, Graduate Texts in Mathematics Vol. 60 (Springer, 1978).

Berestycki, H. and Brézis, H., "On a free boundary problem arising in plasma physics," Nonlinear Anal. **4**, 415–436 (1980).

Boozer, A. H., "Plasma equilibrium with rational magnetic surfaces," Phys. Fluids **24**, 1999–2003 (1981).

Boozer, A. H., "Transport and isomorphic equilibria," Phys. Fluids **26**, 496–499 (1983).

Burby, J. W. and Ellison, C. L., "Toroidal regularization of the guiding center Lagrangian," Phys. Plasmas **24**, 110703 (2017).

Burby, J. W. and Qin, H., "Toroidal precession as a geometric phase," Phys. Plasmas **20**, 012511 (2013).

D'haeseleer, W. D., Hitchon, W. N. G., Callen, J. D., and Shohet, J. L., *Flux Coordinates and Magnetic Field Structure* (Springer, 1991).

Fried, D., "The geometry of cross-sections to flows," Topology **21**, 353–371 (1982).
Garren, D. A. and Boozer, A. H., "Existence of quasihelically symmetric stellarators," Phys. Fluids B **3**, 2822–2834 (1991).
Gilbarg, D. and Trudinger, N. S., *Elliptic Partial Differential Equations of Second Order*, Classics in Mathematics (Springer, 2001).
Grad, H. and Rubin, H., "Hydromagnetic equilibria and force-free fields," in *Proceedings of 2nd United Nations Conference on Peaceful Uses of Atomic Energy* (United Nations, Geneva, 1958), Vol. 31, pp. 190–197.
Hamada, S., "Hydromagnetic equilibria and their proper coordinates," Nucl. Fusion **2**, 23–37 (1962).
Helander, P., "Theory of plasma confinement in non-axisymmetric magnetic fields," Rep. Prog. Phys. **77**, 087001 (2014).
Jardin, S., *Computational Methods in Plasma Physics*, Computational Science Series (CRC Press, 2010).
Johnson, J. L., Oberman, C. R., Kulsrud, R. M., and Frieman, E. A., "Some stable hydromagnetic equilibria," Phys. Fluids **1**, 281–296 (1958).
Kruger, S. E. and Greene, J. M., "The relationship between flux coordinates and equilibrium-based frames of reference in fusion theory," Phys. Plasmas **26**, 082506 (2019).
Landreman, M. and Catto, P. J., "Omnigenity as generalised quasisymmetry," Phys. Plasmas **19**, 056103 (2012).
Landreman, M., Sengupta, W., and Plunk, G. G., "Direct construction of optimised stellarator shapes. Part 2. Numerical quasisymmetric solutions," J. Plasma Phys. **85**, 905850103 (2019).
Littlejohn, R. G., "Variational principles of guiding centre motion," J. Plasma Phys. **29**, 111–125 (1983).
MacKay, R. S., "Tutorial on differential forms in plasma physics," J. Plasma Phys. **86**, 925860101 (2020).
Nührenberg, J. and Zille, R., "Quasihelically symmetric toroidal stellarators," Phys. Lett. A **129**, 113–117 (1988).
Olver, P. J., *Applications of Lie Groups to Differential Equations*, Graduate Texts in Mathematics Vol. 107 2nd ed. (Springer, 1993).
Shafranov, V. D., "On magnetohydrodynamical equilibrium configurations," Sov. Phys. JETP **6**, 545–554 (1958).
Simakov, A. N. and Helander, P., "Plasma rotation in a quasisymmetric stellarator," Plasma Phys. Control. Fusion **53**, 024005 (2011).