Helicity and the Mañé critical value

GABRIEL P PATERNA

We establish a relationship between the helicity of a magnetic flow on a closed surface of genus \( \geq 2 \) and the Mañé critical value.

53C25, 53C21, 35J15

1 Results

Let \( N \) be a closed oriented 3–manifold with a volume form \( \Omega \). A vector field \( F \) on \( N \) that preserves \( \Omega \) is said to be null-homologous or exact if the closed 2–form \( i_F \Omega \) is exact. Given a volume preserving null-homologous vector field \( F \), the helicity \( \mathcal{H}(F) \) is defined by setting

\[
\mathcal{H}(F) := \int_N \tau \wedge d\tau = \int_N \tau(F) \Omega,
\]

where \( \tau \) is any primitive of \( i_F \Omega \). It is easy to check that this definition is independent of the choice of primitive \( \tau \). The helicity (also referred to as the asymptotic Hopf invariant) measures how much in average field lines wrap and coil around one another. The term “helicity” was introduced by K Moffatt [12] who also pointed out the topological nature of the invariant. We refer to Arnold and Khesin [2] for a complete account of this concept as well as its interpretation as an average self-linking number.

In this note we wish to study the following class of volume preserving flows. Let \( M \) be a closed oriented surface of genus \( \geq 2 \) and let \( g \) be a Riemannian metric on \( M \). The unit circle bundle \( SM \) determined by \( g \) is a closed 3–manifold with volume form \( \Omega = \alpha \wedge d\alpha \), where \( \alpha \) is the contact 1–form of the geodesic flow of \( g \). Let \( X \) be the geodesic vector field and let \( V \) denote the infinitesimal generator of the circle action on the fibres of \( SM \). Suppose we are given in addition a 2–form \( \sigma \) on \( M \) (automatically closed). We may write \( \sigma = f \mu_g \), where \( \mu_g \) is the area form of \( g \). The vector field \( F := X + f V \) preserves the volume form \( \Omega \) and in fact is null-homologous. Its helicity is easy to compute and one obtains

\[
\mathcal{H}(F) = 2\pi A + [\sigma]^2 / \chi,
\]

Published: 21 July 2009

DOI: 10.2140/agt.2009.9.1413
where $A$ is the area of $g$, $[\sigma]$ is the total integral of $\sigma$ and $\chi$ is the Euler characteristic of $M$ (we refer to Section 2 for proofs of these elementary facts). In fact formula (1) holds also for the 2–sphere, but in that case $\mathcal{H}(F) > 0$ always.

There is a well known symplectic interpretation for the vector field $F$. Denote by $\pi: TM \to M$ the canonical foot-point projection. Consider the twisted symplectic form $\omega := -d\alpha + \pi^*\sigma$ on $TM$ and the Hamiltonian $H: TM \to \mathbb{R}$ given by the kinetic energy of $g$, i.e., $H(x, v) := g_x(v, v)/2$. The Hamiltonian vector field $\xi_H$ of $H$ with respect to $\omega$ restricted to $SM$ is precisely $F$ ($\xi_H$ is defined by $i_{\xi_H} \omega = dH$). Note that $SM$ coincides with the energy hypersurface $H = 1/2$. The flow of $F$ is called a magnetic flow or a twisted geodesic flow. It is also well known, that considering $\xi_H$ restricted to a hypersurface $H = k$ is equivalent to considering the vector field $F_s$ on $SM$ determined by $s \sigma$, where $s = 1/\sqrt{2\kappa}$. We can think of $s$ as the intensity that modulates the magnetic field $\sigma$. The projections to $M$ of the orbits of $F_s$ model the motion of a particle with charge $s$ under the effect of the magnetic field $\sigma$. If we let $\omega_s := -d\alpha + s\pi^*\sigma$, then the pair $(SM, \omega_s)$ determines a Hamiltonian structure for every $s$.

This note is partially motivated by the following two open questions:

(A) For which values of $s \in (0, \infty)$ is the Hamiltonian structure $(SM, \omega_s)$ of contact type? (Recall that contact type means that there exists a smooth 1–form $\lambda$ on $SM$ such that $d\lambda = \omega_s$ and $\lambda(F_s)$ never vanishes.)

(B) For which values of $s \in (0, \infty)$ does $F_s$ have a closed orbit?

The literature on Question (B) is vast and it is impossible to do it justice in this brief note. Suffices to say that the study of the problem of existence of closed orbits for magnetic flows was initiated by V I Arnold [1] and S P Novikov [13] in the early 80’s with subsequent work by many others. We refer the reader to Ginzburg [8; 9] for a survey of some of these results, particularly for the case of surfaces discussed here.

Obviously (A) and (B) are related by Taubes’ proof of the Weinstein conjecture [16]. In the exact case $[\sigma] = 0$, both questions were solved by Contreras, Macarini and the author [6] with the help of Aubry–Mather theory. An important ingredient in [6] was the Mañé critical value whose definition we now recall. Let $p: \tilde{M} \to M$ denote the universal covering of $M$. Since we are assuming that $M$ has $\chi < 0$, $\tilde{M}$ is diffeomorphic to $\mathbb{R}^2$ and thus $p^*\sigma$ has a primitive. Set

$$c(g, \sigma) := \inf_{\theta} \sup_{x \in \tilde{M}} \frac{1}{2} \left| \theta x \right|^2.$$
where the infimum runs over all 1–forms $\theta$ with $d\theta = p^*\sigma$, and the norm of $\theta$ is taken with respect to the lifted Riemannian metric. It is easy to see that $c(g, \sigma) < \infty$ using the fact that on the upper-half plane with the hyperbolic metric, the primitive $-y^{-1}dx$ of the area form $y^{-2}dx \wedge dy$ is bounded. If $\sigma = \mu_g$, then $1/\sqrt{2c(g, \sigma)}$ coincides with Cheeger’s isoperimetric constant of the universal covering (cf Burns and Paternain [3]). The critical value $c = c(g, \sigma)$ is also relevant for us because it is known by work of the author [14] that for any $s \in (0, s_c)$, every nontrivial homotopy class of $M$ contains the projection of a closed orbit of $F_s$, where $s_c := 1/\sqrt{2c}$. A thorough discussion of the relevance of the Mañé critical value to the symplectic topology of hypersurfaces may be found in Contreras, Macarini and Paternain [5].

Let us return to the helicity now. In the nonexact case ($[\sigma] \neq 0$) an inspection of (1) tells us that there is a unique positive value of $s$ for which $H(F_s) = 0$ and is given by

$$s_h^2 := \frac{-2\pi \chi A}{|\sigma|^2}.$$ 

Since the helicity vanishes, $(SM, \omega_{s_h})$ cannot be of contact type. How does it relate with $s_c$? The answer is given by the following theorem:

**Theorem** For an arbitrary pair $(g, \sigma)$ on a closed surface of genus $\geq 2$ with $[\sigma] \neq 0$, we have $s_c \leq s_h$ with equality if and only if $g$ has constant Gaussian curvature and $\sigma$ is a constant multiple of the area form of $g$.

We note that if $g$ has constant curvature $-1$ and $\sigma = \mu_g$, then $s_c = s_h = 1$. The vector field $F_1$ is the horocycle flow, which is uniquely ergodic and has of course, zero helicity. The theorem is saying that unless we are in this well understood homogeneous situation, if we wish to answer Questions (A) and (B) above, we would need to wrestle with a nontrivial interval $[s_c, s_h]$ whose Hamiltonian structures are probably out of reach of current technology. Presumably, every $s \in [s_c, s_h]$ is not of contact type, but even for $s_c$ this is not known in full generality.

The proof of the theorem has two ingredients. One was already present in [3, Theorem B] (but its relation with the helicity was not exposed) and it will give fairly easily the inequality and the fact that, if equality holds, then $g$ must have constant curvature. However to show that $\sigma$ must be a constant multiple of the area form requires a new tool. This is provided by techniques closely related with the Selberg trace formula and the study of an appropriate Radon transform of 1–forms on geodesic circles of the Riemann surface $M$.

---

1 More generally, there is a Mañé critical value associated to any covering of $M$ on which $\sigma$ becomes exact. The main result in [6] asserts that if $\sigma$ is exact on $M$, a hypersurface with energy $k$ is of contact type if and only if $k > c_0$, where $c_0$ is the Mañé critical value of the abelian covering. This value coincides with the minimum of Mather’s alpha function. Moreover, every energy level has a closed orbit.
Caveat on terminology  In magnetohydrodynamics (MHD) the role of \( F \) is played by a magnetic field \( B \), frozen into a fluid of infinite conductivity filling \( N \). What we call here magnetic field is the 2–form \( \sigma \) and the two “fields” should not be confused.

2 Proofs

Let \( M \) be a closed oriented surface with genus \( \geq 2 \) and \( g \) a Riemannian metric. As above \( SM \) is the unit circle bundle. We consider a 2–form \( \sigma \) on \( M \) with total integral \( \int \sigma \). The vector field of the magnetic flow is given by \( F := X + f V \) where \( f \) is defined by \( \sigma = f \mu_g \), and \( \mu_g \) is the area form of \( g \) with total integral \( A \). There is a coframe of 1–forms \( \{ \alpha, \gamma, \psi \} \) in \( SM \) related by the structure equations \( d\alpha = \psi \wedge \gamma \), \( d\gamma = -\psi \wedge \alpha \) and \( d\psi = -K \alpha \wedge \gamma \), where \( K \) is the Gaussian curvature of \( g \) and \( \alpha \) is the contact 1–form dual to the geodesic vector field \( X \). The coframe \( \{ \alpha, \gamma, \psi \} \) is dual to \( \{ X, H, V \} \), where \( H = [V, X] \).

Let \( \Omega := \alpha \wedge d\alpha \) be the Sasaki volume form on \( SM \). A calculation using the structure equations shows that \( i_F \Omega = da - f \alpha \wedge \gamma = da - f \pi^* \mu_g \). In other words \( i_F \Omega = da - \pi^* \sigma \). It is easy to find a primitive for \( \pi^* \sigma \). Write \( \sigma = -aK \mu_g + d\beta \), where \( a \) satisfies \( |\sigma| = -a 2\pi \chi \) and \( \beta \) is a 1–form on \( M \). Then \( \pi^* \sigma = a d\psi + d \pi^* \beta \) and thus \( i_F \Omega = d\tau \), where \( \tau := \alpha - a \psi - \pi^* \beta \). This shows that \( F \) is null-homologous and

\[
\tau(F)(x, v) = 1 - a f(x) - \beta_x(v).
\]

Since the function \((x, v) \mapsto \beta_x(v)\) is odd with respect to the flip \( v \mapsto -v \) we have

\[
\int_{SM} \beta_x(v) \Omega = 0.
\]

It follows that the helicity of \( F \) is given by

\[
\mathcal{H}(F) = 2\pi A - a 2\pi |\sigma| = 2\pi A + |\sigma|^2 / \chi,
\]

which proves (1).

Remark 2.1  The calculation of the helicity for an arbitrary magnetic flow was also carried out in [14, Equation (4)]. Up to the factor \(-2\pi A\), the helicity is precisely what I called in [14] the action of the Liouville measure. The Proposition in [14] could be rephrased by saying that if \( F_{sh} \) has no conjugate points, then \( g \) must have constant curvature, \( \sigma \) is a constant multiple of the area form and \( F_{sh} \) is a horocycle flow. In fact, the proof of the Proposition in [14] shows that for any \( s \in (s_h, \infty) \), \( F_s \) has conjugate points.

Algebraic & Geometric Topology, Volume 9 (2009)
The helicity for geodesic and horocycle flows in the constant curvature case is also computed in [2, Proposition 4.9] and [17, Example 2.2.1].

**Remark 2.2** If we replace $\sigma$ by $s\sigma$ in the argument above we obtain a primitive $\tau_s := \alpha - as \psi - s\pi^* \beta$ of $i_{F_s} \Omega = -\omega_s$ such that

$$
\tau_s(F_s) = 1 - a s^2 f(x) - s \beta_x(v).
$$

This shows right away that if $f$ does not vanish (ie $\sigma$ is symplectic), then there is $s_0$ such that for any $s > s_0$, $(SM, \omega_s)$ is of contact type and therefore it has a closed orbit.

### 2.1 Mañé’s critical value and helicity

By the conformal equivalence theorem there exists a unique positive scalar $C^\infty$ function $\rho$ such that the metric $\rho^2 g$ has constant negative curvature and the same area as $g$. Let $\rho_g$ be the *conformality coefficient* given by

$$
\rho_g := \frac{1}{A} \int_M \rho \mu_g.
$$

By the Cauchy–Schwartz inequality, $\rho_g \leq 1$ and equality holds if and only if $g$ itself is a metric of constant negative curvature. In [3, Theorem B] it is shown that

$$
(2) \quad c(g, \sigma) \geq \frac{[\sigma]^2}{4\pi \chi A \rho_g^2}.
$$

In order to make this note self-contained we will give a proof of (2), which is actually a little simpler than the one in [3] and it will naturally lead us to the proof of the Theorem. The key idea comes from a similar estimate of Katok [10] of the Cheeger isoperimetric constant. Without loss of generality we may suppose that $g$ has area $A = -2\pi \chi$ and hence $g_0 := \rho^2 g$ has constant negative curvature $-1$.

We lift everything to the universal covering $p: \mathbb{H}^2 \rightarrow M$ and we consider the Lagrangian $L(x, v) = g_x(v, v)/2 - \theta_x(v)$, where $\theta$ is a primitive of $p^* \sigma$. The critical value $c(g, \sigma)$ can also be characterised as the infimum of the values of $k \in \mathbb{R}$ such that the action

$$
A_{L+k}(\gamma) := \int_0^T (L + k)(\gamma(t), \dot{\gamma}(t)) \, dt \geq 0
$$

for all absolutely continuous closed curves $\gamma: [0, T] \rightarrow \mathbb{H}^2$ and any $T > 0$. This was Mañé’s original approach to the critical value [11] and in the setting of nonexact magnetic fields, a proof of the equivalence between these two ways of characterising $c$ may be found in [3].

*Algebraic & Geometric Topology*, Volume 9 (2009)
Consider a geodesic circle $C_r$ of $g_0$ of radius $r$. It has $g_0$–length $2\pi \sinh r$ and encloses a disk $D_r$ of $g_0$–area $2\pi(\cosh r - 1)$. Its $g$–length is given by
\[
\ell_g(C_r) = \int_0^{2\pi \sinh r} \rho^{-1}(\gamma(t)) \, dt,
\]
where $\gamma : [0, 2\pi \sinh r] \to \mathbb{H}^2$ is a parametrisation of $C_r$ with speed one with respect to $g_0$. Now parametrise $C_r$ to have speed $\sqrt{2c}$ with respect to $g$. We must have
\[
A_{L+c}(C_r) \geq 0
\]
for all $r > 0$. In other words, using Stokes theorem and the definitions,
\[
\sqrt{2c} \ell_g(C_r) - \int_{D_r} p^* \sigma \geq 0
\]
for all $r > 0$. We can write $\sigma = a \mu_{g_0} + d\beta$, where $\mu_{g_0}$ is the area form of $g_0$. Clearly $[\sigma] = -a 2\pi \chi$. Thus
\[
(3) \quad \sqrt{2c} \ell_g(C_r) - a 2\pi(\cosh r - 1) - \int_{C_r} p^* \beta \geq 0
\]
for all $r > 0$. The key observation now is that the projection to $M$ of a circle $C_r$ in $\mathbb{H}^2$ converges to a horocycle when the radius goes to infinity, and the projection to the unit sphere bundle of $(M, g_0)$ of the normalised arc length measure weakly converges to an invariant probability measure for the horocycle flow. But the only invariant probability measure for the horocycle flow is the Liouville measure $\nu$ of $g_0$ [7].

If we now divide (3) by $2\pi \sinh r$, let $r$ go to infinity and use the definition of weak convergence we derive
\[
\sqrt{2c} \int_M \rho^{-1} \mu_{g_0} - a A \geq 0,
\]
since the integral of $\beta$ (regarded as a function on $TM$) over the unit circle bundle of $g_0$ with respect to $\nu$ must vanish. Equivalently, using that $[\sigma] = a A = -a 2\pi \chi$ we have
\[
\sqrt{2c} \geq \frac{[\sigma]}{\int_M \rho^{-1} \mu_{g_0}}.
\]
But
\[
-2\pi \chi \rho_g = \int_M \rho \mu_g = \int_M \rho^{-1} \mu_{g_0}.
\]
Hence
\[
\sqrt{2c} \geq \frac{[\sigma]}{-2\pi \chi \rho_g}
\]
which proves (2) when $A = -2\pi \chi$ (note that $c(g, \sigma) = c(g, -\sigma)$).

*Algebraic & Geometric Topology, Volume 9 (2009)*
Next we observe that the inequality $s_c^2 \leq s_h^2$ is equivalent to $c \geq [\sigma]^2 / (-4\pi \chi A)$ which follows immediately from (2) since $\rho_g \leq 1$. Moreover, if equality holds then $\rho_g = 1$ and $g$ must have constant curvature. What remains to prove in the Theorem from the introduction is that if $s_c = s_h$ then $f$ must also be constant. Before we proceed any further we would like to record inequality (2) in the following form:

**Proposition 2.3** For any pair $(g, \sigma)$ on a surface with genus $\geq 2$ we have

$$2c(g, \sigma) \geq 2\rho_g^2 c(g, \sigma) \geq 1 - \frac{\mathcal{H}(F)}{2\pi A}.$$  

As above, and without loss of generality we shall assume that $g$ has constant curvature $-1$. Since $s_c = s_h$ we can write $\sigma = a\mu_g + d\beta$, where $[\sigma] = -a 2\pi \chi$ and $2c = a^2$. We may suppose that in fact $a = \sqrt{2c}$. Using (3) we obtain

$$2\pi \sqrt{2c} (1 + \sinh r - \cosh r) - \int_{C_r} p^* \beta \geq 0$$

which implies

$$\int_{C_r} p^* \beta \leq 2\pi \sqrt{2c} (1 - e^{-r}) \leq 2\pi \sqrt{2c}$$

for all $r > 0$. In the next subsection we explain how use the bound (4) to show that in fact $\beta$ must be closed, and consequently $\sigma$ is a constant multiple of $\mu_g = \mu_g^0$.

### 2.2 A Radon transform

Let $h: M \to \mathbb{R}$ be a smooth function with $\int_M h(x) \, dx = 0$. As before $M = \mathbb{H}^2 / \Gamma$ and $p: \mathbb{H}^2 \to M$ the quotient map. We consider the Radon transform $\hat{h}_r$ of $h$ on geodesic disks defined as follows. Given $x \in M$, let $\bar{x}$ be a lift of $x$ and let $D(\bar{x}, r)$ be the disk with center $\bar{x}$ and radius $r$. We set

$$\hat{h}_r(x) := \int_{D(\bar{x}, r)} h \circ p(y) \, dy.$$  

It is easy to check that this definition is independent of the lift of $x$.

Let $\varphi_0, \varphi_1, \ldots$ denote a complete orthonormal sequence of real eigenfunctions of the Laplacian of $M$ corresponding to eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$ $\to \infty$. Write $h = \sum_j a_j \varphi_j$. Since $h$ has zero average over $M$, $a_0 = 0$.

**Lemma 2.4** $\hat{h}_r(x) = \sum_j a_j q_r(s_j) \varphi_j(x)$, where $q_r$ is the function $(s \in \mathbb{C})$

$$q_r(s) = 4\sqrt{2} \int_0^r \cos su (\cosh r - \cosh u)^{1/2} \, du$$

and $s_j$ is any of the roots of $\frac{1}{4} + s_j^2 = \lambda_j$. 

*Algebraic & Geometric Topology, Volume 9 (2009)*
Proof This is the application of the techniques connected with the Selberg trace formula [15] and is fully explained in Section 2 of Randol’s chapter in Chavel’s book [4]. It goes as follows. We let \( k_r(x, y) \) be the function on \( \mathbb{H}^2 \times \mathbb{H}^2 \) such that \( k_r(x, y) = 1 \) if \( y \in D(x, r) \) and \( k_r(x, y) = 0 \) otherwise. Set \( K_r(x, y) := \sum_{y \in \Gamma} k_r(x, y y) \). Then it is easy to check that

\[
\hat{h}_r(x) = \int_M h(y) K_r(x, y) \, dy.
\]

Using the expansion \( h = \sum_j a_j \varphi_j \) we obtain

\[
\hat{h}_r(x) = \sum_j a_j \int_M \varphi_j(y) K_r(x, y) \, dy.
\]

But it is shown in [4, Chapter X, Theorem 1 and page 277] that

\[
\int_M \varphi_j(y) K_r(x, y) \, dy = q_r(s_j) \varphi_j(x),
\]

where \( q_r(s) \) is calculated in [4, page 275] yielding the formula in the lemma.

\[\square\]

Remark 2.5 If \( s_j \) is real, then

\[
q_r(s_j) = 4 \sqrt{2} \int_0^r \cos s_j u \left( \cosh r - \cosh u \right)^{1/2} \, du.
\]

If there are small eigenvalues, then \( s_j \) would be purely imaginary and if we let \( \alpha_j = |s_j| \), then

\[
q_r(s_j) = 4 \sqrt{2} \int_0^r \cosh \alpha_j u \left( \cosh r - \cosh u \right)^{1/2} \, du.
\]

Lemma 2.6 For every \( j \), there is \( r_n \to \infty \) such that \( q_{r_n}(s_j) \to \infty \).

Proof On account of Remark 2.5 it suffices to prove the lemma when \( s_j \) is real and positive since if \( s_j \) is purely imaginary or zero \( \cosh \alpha_j u \geq 1 \) for all \( u \) which implies that \( q_r(s_j) \geq q_r(s) \) where \( s \) is any real number.

Suppose then that \( s_j \) is real and positive and note

\[
\frac{q_r(s_j)}{4 \sqrt{2}} = \frac{1}{(\cosh r)^{1/2}} \int_0^r \cos s_j u \left( \cosh r - \cosh u \right)^{1/2} \, du
\]

\[
\geq \frac{1}{(\cosh r)^{1/2}} \int_0^r \cos s_j u \left( \cosh r - \cosh u \right) \, du
\]

\[
= \frac{1}{(\cosh r)^{1/2} s_j (1 + s_j^2)} \left( \cosh r \sin s_j r - s_j \sinh r \cos s_j r \right).
\]
Thus, if we take \( r_n = \pi(2n + 1/2)/s_j \) we derive
\[
\frac{q_{r_n}(s_j)}{4\sqrt{2}} \geq \frac{(\cosh r_n)^{1/2}}{s_j(1+s_j^2)}
\]
which proves the lemma. \( \square \)

**Lemma 2.7** Suppose that \( \hat{h}_r(x) \leq C \) for all \( x \in M \) and all \( r > 0 \). Then \( h \equiv 0 \).

**Proof** We will show that all the Fourier coefficients \( a_j \) vanish. Suppose \( a_k \geq 0 \) for some \( k \). By compactness there is a constant \( B_k \) such that \( \varphi_k + B_k \) is a positive function. Multiply both sides of the inequality
\[
\sum_j a_j q_r(s_j) \varphi_j(x) \leq C
\]
by \( \varphi_k(x) + B_k \) and integrate with respect to \( x \) to obtain \( (a_0 = 0) \)
\[
a_k q_r(s_k) \leq C B_k A
\]
for all \( r > 0 \). By Lemma 2.6 this can only happen if \( a_k = 0 \).

If \( a_k \leq 0 \) we proceed in a similar way by considering a constant \( B_k \) such that \( \varphi_k + B_k \) is a negative function. In any case we obtain \( a_k = 0 \) as desired. \( \square \)

### 2.3 End of the proof of the Theorem

Write \( d\beta = h \mu_g \) where \( h \) has zero average over \( M \). Inequality (4) is saying that \( \hat{h}_r(x) \leq 2\pi \sqrt{2}c \) for all \( x \in M \) and \( r > 0 \). By Lemma 2.7, \( h \) vanishes identically and \( \beta \) must be closed.

### References


Department of Pure Mathematics and Mathematical Statistics
University of Cambridge, Cambridge CB3 0WB, UK
g.p.paternain@dpmms.cam.ac.uk

Received: 5 March 2009

Algebraic & Geometric Topology, Volume 9 (2009)