Minimization of magnetic forces on stellarator coils

Rémi Robin¹ and Francesco A. Volpe²,*

¹ Laboratoire Jacques-Louis Lions, Sorbonne Université, Paris, France
² Renaissance Fusion, Fontaine, France

E-mail: remi.robin@inria.fr and francesco.volpe@renfusion.eu

Received 12 April 2022, revised 31 May 2022
Accepted for publication 7 June 2022
Published 6 July 2022

Abstract
Magnetic confinement devices for nuclear fusion can be large and expensive. Compact stellarators are promising candidates for cost-reduction, but introduce new difficulties: confinement in smaller volumes requires higher magnetic field, which calls for higher coil-currents and ultimately causes higher Laplace forces on the coils—if everything else remains the same. This motivates the inclusion of force reduction in stellarator coil optimization. In the present paper we consider a coil winding surface, we prove that there is a natural and rigorous way to define the Laplace force (despite the magnetic field discontinuity across the current-sheet), we provide examples of cost associated (peak force, surface-integral of the forcesquared) and discuss easy generalizations to parallel and normal force-components, as these will be subject to different engineering constraints. Such costs can then be easily added to the figure of merit in any multi-objective stellarator coil optimization code. We demonstrate this for a generalization of the REGCOIL code (Landreman 2017 Nucl. Fusion 57 046003), which we rewrote in python, and provide numerical examples for the NCSX (Zarnstorff et al 2001 Plasma Phys. Control. Fusion 43 A237–49) (now QUASAR) design. We present results for various definitions of the cost function, including peak force reductions by up to 40%, and outline future work for further reduction.

Keywords: stellarator, high field, coil optimization, force minimization

(Some figures may appear in colour only in the online journal)

1. Introduction
Stellarators are non-axisymmetric toroidal devices that magnetically confine fusion plasmas [3]. Thanks to specially shaped coils they do not require a current in the plasma, hence are more stable and steady-state than tokamaks. However, they exhibit comparable confinement, hence tend to be about as large. Like tokamaks, confinement can be improved (and size reduced) by adopting stronger magnetic fields.

Fields as high as 8–12 T were only tested in two series of tokamak experiments at the MIT and ENEA, culminated respectively in Alcator C-mod [4] and FTU [5]. For comparison, ITER has a field of 5.3 T on axis [6]. Other high-field tokamaks were designed but not built [7, 8], although the new high-field SPARC tokamak has been designed, modeled and its construction is expected to start in 2021 [9].

For stellarators and heliotrons, there is broad agreement that power-plants will require at least 4–6 T [10], but fields as high as 8–12 T have only been proposed very recently [11]. Two private companies are working toward that goal³,⁴. Generating strong fields requires high currents and of course results in high forces on the coils (unless their design is modified, as we will argue in this paper). Up to 5 T, the issue can be resolved by adequately reinforced coil-support structures and coil-spacers [12]. However, a further increase to 10 T will result in 4× higher forces. This calls for including force-reduction in the coil design and optimization process, along with other criteria.

³ Type One Energy. https://typeoneenergy.com.
⁴ Renaissance Fusion. https://stellator.energy.
Such need was recognized earlier on for heliotrons, and spurred reduced force (so-called force-free) heliotron designs [13]. From a mathematical standpoint this is not too surprising, since helical fields in heliotrons resemble the eigenfunctions of the curl operator on a torus [14]: \( \nabla \times \mathbf{B} = \mathbf{A} \). This, combined with Maxwell–Ampere law, implies that \( \mathbf{B} \) and the current density \( j \) are parallel, and there are no Laplace forces on the coils.

Modular coils for advanced stellarators, on the other hand, are the result of numerical optimization. The most common optimization criterion is to reproduce the target magnetic field to within one part in \( 10^4 \) or \( 10^5 \). Typically this is solved on a 2D toroidal surface external to the plasma, called Coil Winding Surface (CWS). On that surface, numerical codes compute the current potential (thus, ultimately, the current pattern) that best reproduces the target plasma boundary, in a least-squares sense [1]. This is the principle of the seminal NESCOIL code [15, 16]. Further developments included engineering-constrained nonlinear optimizers [17] and the Tikhonov-regularized REGCOIL [1]. The latter includes the squared coil-current density in the objective function, which leads to more ‘gentle’, easier-to-build coil shapes. All these codes fix the CWS; more recently, a free-CWS 3D search method was developed [18].

In the present article, we generalize REGCOIL to include coil-force reduction. This is obtained by adding to the objective function a term quantifying the Laplace forces on the CWS. Several metrics are possible, for example the surface-current of density \( j \) on a surface-current of density \( j \), Section 2 and appendices A.1–A.4 offer an extensive mathematical derivation of such expression. Based on that, possible cost functions are proposed and briefly discussed in section 3. In section 4 we determine that, based on the Hilbert space where the current density is defined, the correct norm to use for its regularization is the \( H^1 \) norm. This is new w.r.t. REGCOIL and other codes adopting the \( L^2 \) norm, and has numerical implications. Finally section 4.3 illustrates the numerical results obtained with the two main cost-functions for the quasi-axisymmetric stellarator design formerly known as NCSX [2], then QUASAR, including a reduction of the peak force by up to 40%.

2. Laplace force on a surface

2.1. Notations

We start by introducing the following notations:

- \( \mathcal{S} \) is a smooth two-dimensional Riemannian submanifold of \( \mathbb{R}^3 \), diffeomorphic to the two-torus standing for the CWS.
- \( n \) is the unit vector field normal to \( \mathcal{S} \) and pointing outward.
- \( \mathcal{X}(\mathcal{S}) \) is the set of smooth vector fields on \( \mathcal{S} \).

- \( \langle X \cdot Y \rangle \) denote the scalar product (in \( \mathbb{R}^3 \)) between the vector fields \( X \) and \( Y \). When both vector are tangent to \( \mathcal{S} \), we sometime denote \( \langle X \cdot Y \rangle_{T\mathcal{S}} \) the scalar product at \( x \in \mathcal{S} \) (which coincides with the one in \( \mathbb{R}^3 \)).
- \( L^p(\mathcal{S}) \) and \( H^1(\mathcal{S}) \) are the Hilbert spaces defined as the completion of \( \mathcal{C}^\infty(\mathcal{S}) \) for the norms

\[
|f|_{L^p(\mathcal{S})} = \left( \int_{\mathcal{S}} |f|^p \, dS \right)^{1/p}
\]

\[
|f|_{H^1(\mathcal{S})}^2 = \int_{\mathcal{S}} \left( f^2 + \langle \nabla f \cdot \nabla f \rangle \right) \, dS.
\]

- \( \mathcal{X}(\mathcal{S}) \) and \( \mathcal{X}^{1,2}(\mathcal{S}) \) are the Hilbert spaces defined as the completion of \( \mathcal{X}(\mathcal{S}) \) for the norms

\[
|j|_{\mathcal{X}^{1,2}(\mathcal{S})} = \sqrt{|j_1|_{H^1(\mathcal{S})}^2 + |j_2|_{H^1(\mathcal{S})}^2 + |j_3|_{H^1(\mathcal{S})}^2},
\]

where \( j_1, j_2, \) and \( j_3 \) are the components of \( j \) in \( \mathbb{R}^3 \) for an arbitrary orthogonal basis.

- The spaces \( L^p(\mathcal{S}, \mathbb{R}^3) \) and \( H^{1,2}(\mathcal{S}, \mathbb{R}^3) \) are related to \( \mathcal{C}^\infty(\mathcal{S}, \mathbb{R}^3) \) in the same way as \( \mathcal{X}(\mathcal{S}) \) and \( \mathcal{X}^{1,2}(\mathcal{S}) \) are related to \( \mathcal{X}(\mathcal{S}) \).

- \( \pi \) is the projector on the tangent bundle. For any \( Y \in \mathcal{C}^\infty(\mathcal{S}, \mathbb{R}^3) \), we define

\[
\forall x \in \mathcal{S}, (\pi(Y))_x = Y_x - \langle Y_x \cdot n(x) \rangle n(x) \in T_x \mathcal{S}. \tag{1}
\]

Since \( \pi(Y) \) is clearly a tangent vector field on \( \mathcal{S} \), it belongs to \( \mathcal{X}(\mathcal{S}) \).

2.2. Limit definition of Laplace force exerted by a current-sheet on itself

Let \( j \) be a vector field on \( \mathcal{S} \), representing the surface-current density, i.e. the current per unit length (not per unit surface, as is usually the case for this notation). The Laplace force is the magnetic component of the Lorentz force; the Laplace force per unit surface (not per unit volume) is given by \( \mathbf{F} = j \times \mathbf{B} \), although here, quite often, it will simply be called ‘force’, for brevity.

A surface-current of density \( j \) causes a discontinuity in the component of the magnetic field \( \mathbf{B} \) tangential to the surface: \( \mathbf{n}_{12} \times (\mathbf{H}_2 - \mathbf{H}_1) = j \). The resulting jump in the tangential component of \( \mathbf{B} \) results in a normal force wherever \( j \neq 0 \). That force, proportional to \( |j|^2 \), tries to increase the thickness of the CWS. To ensure that force remains reasonably small, one can easily add a cost \( |j|_{L^2} \) or \( |j|_{\mathcal{X}} \), to the multi-objective figure-of-merit or optimize under a constraint on \( j \).

From now on, though, we will focus on the other contributions to the Laplace force. We define them in a location \( y \in \mathcal{S} \) as follows:

\[
\mathbf{F}(y) = \lim_{\varepsilon \to 0} \frac{1}{2} \{ [\mathbf{B}(y + \varepsilon \mathbf{n}(y)) + \mathbf{B}(y - \varepsilon \mathbf{n}(y))].
\]

Let us focus on the case where \( \mathbf{B} \) is only generated by currents on \( \mathcal{S} \); there are no permanent magnets nor magnetically
susceptible media. In any \( y \notin S \), the field is given by the Biot–Savart law in vacuo:

\[
B(y) = BS(j)(y) = \int_S J(x) \times \frac{y - x}{|y - x|^3} \, dS(x), \tag{2}
\]

where, to reduce the amount of notation, we dropped the \( \frac{\mu_0}{4\pi} \) factor in front of the integral. The notation \( BS(j) \) refers to the Biot–Savart operator, function of \( j \), that maps each \( y \notin S \) in the local field, \( B(y) \).

**Remark 1.** \( BS(j) \) cannot be defined on \( S \), unless \( j = 0 \). This is because \( j \) is not being integrable for \( y \in S \). The jump in the tangential component of the induced magnetic field mentioned above is thus caused by the discontinuity of \( BS(j)(y + \varepsilon n(y)) \) at \( \varepsilon = 0 \).

However, since \( BS(j) \) is well-defined in locations \( y \notin S \), we can define for any \( y \in S \) and any \( \varepsilon > 0 \) the bilinear map

\[
L_{\varepsilon}(j_1, j_2)(y) = \frac{1}{2} \left\{ J_1(y) \times [BS(j_2)(y + \varepsilon n(y)) + BS(j_2)(y - \varepsilon n(y))] \right\}. \tag{3}
\]

This describes the Laplace force that a surface-current of density \( j_1 \) exerts on another of density \( j_2 \), per unit surface. Since we are dealing with a stellarator, these currents are constant in time and there is no need to include induced fields and the associated forces.

The ‘average Laplace force’ that a current of density \( j \) exerts on point \( y \in S \) is thus \( L_{\varepsilon}(j)(y) = \lim_{\varepsilon \to 0} L_{\varepsilon}(j_1, j_2)(y) \).

This definition, however, raises several questions:

(a) Under which assumptions on \( j \) can we ensure that \( L_{\varepsilon}(j) \) is well defined (i.e., that the limit is well-defined)?

(b) Can we find an explicit expression of \( L_{\varepsilon}(j) \) (i.e., without a limit on \( \varepsilon \))?

(c) Which functional space does \( L_{\varepsilon}(j) \) belong to (for \( j \) in a given functional space)?

The first point is more theoretical, but is necessary to answer the second and third one, which have very practical consequences. Indeed, without an explicit expression for \( L_{\varepsilon}(j) \), the numerical computation of the Laplace force may be a complex matter. A typical approach would involve three different scales. From the smallest to the largest, these are the discretisation-length of \( S \), \( h \), the infinitesimal distance \( \varepsilon \), and the characteristic distance of variation of the magnetic field, \( d_S \).

An accurate computation of \( B(y + \varepsilon n(y)) \) requires \( S \) to be finely discretized, with the discretisation-length \( h \ll \varepsilon \). This is because \( \int_S [y + \varepsilon n(y)] - x \times dS(x) \) blows up when \( \varepsilon \to 0 \). Indeed when we replace the integral with a discrete sum (with \(|y_{i,j} - y_{i+1,j} | \approx |y_{i,j} - y_{i,j+1} | \approx h \)) and take the limit for small \( \varepsilon \),

\[
\tilde{B}(y_{i,j} + \varepsilon n(y_{i,j})) = \sum_{k,l} \{ j(x_{i,k}) \times \frac{y_{i,j} + \varepsilon n(y_{i,j}) - x_{i,j}}{|y_{i,j} + \varepsilon n(y_{i,j}) - x_{i,j}|^3} \} dS(x_{i,k}).
\]

Figure 1. Average norm of \( B \) as a function of the distance \( \varepsilon \) from the surface \( S \), for different grid sizes, more coarse (red) or fine (blue). \( h_{\text{min}} \) and \( h_{\text{max}} \) refer to the smallest and largest mesh size (the poloidal \times toroidal mesh being non-uniform in real space). The plot guides the selection of \( \varepsilon \); an excessively small value, \( \varepsilon \ll h \), results in the numerical artifact of a diverging field.

which for small \( \varepsilon \) diverges like \( \varepsilon^{-2} \), as shown in figure 1 for NCSX.

The semi-sum \( \tilde{B}(y + \varepsilon n(y)) + \tilde{B}(y - \varepsilon n(y)) \) is numerically more stable, but we still need \( h \ll \varepsilon \) (as it will be shown later in figure 3). Such fine mesh makes it costly to accurately compute \( L_{\varepsilon}(j)(y) \) as \( \lim_{\varepsilon \to 0} L_{\varepsilon}(j_1, j_2)(y) \).

The functional space of \( L_{\varepsilon}(j) \) is also important to understand what type of penalization can be applied to minimize this force, or a related metric.

### 2.3. Computing the Laplace force exerted by one current-sheet on another

Consider two linear densities of surface-currents \( j_1, j_2 \in X^{1,2}(S) \) and fix \( \varepsilon > 0 \).

Thanks to the well-known formula \( A \times (B \times C) = (A \cdot C)B - (A \cdot B)C \), we obtain from equations (2) and (3) that

\[
L_{\varepsilon}(j_1, j_2)(y) = \int_S \left\{ j_1(y) \cdot \left( \frac{y - x + \varepsilon n(y)}{2|y - x + \varepsilon n(y)|^3} \right) \right\} j_2(x) dx
\]

\[
\hspace{1cm} - \int_S \left\{ j_1(y) \cdot j_2(x) \right\} \left( \frac{y - x + \varepsilon n(y)}{2|y - x + \varepsilon n(y)|^3} \right) dx.
\]

The difficulty here is that \( \frac{1}{|y - x + \varepsilon n(y)|^3} \) is not integrable in two dimensions (remark 1). Hence, it does not make sense to take the limit for \( \varepsilon \to 0 \) directly inside the integral. Nevertheless, we can use the following equality:
followingequalities:

\[ \int_S \left< j_1(y) \cdot \frac{y - x + \varepsilon n(y)}{|y - x + \varepsilon n(y)|} \right> j_2(x) \, dx \]
\[ = \int_S \left< j_1(y) \cdot \nabla_x \frac{1}{|y - x + \varepsilon n(y)|} \right> j_2(x) \, dx, \quad (5) \]

where \( \nabla_x \) is the gradient in \( \mathbb{R}^3 \) with respect to the variable \( x \).

Furthermore, \( \varepsilon \rightarrow 0 \) for the first integral in equation (4) we have the following equalities:

\[ \int_S \left< j_1(y) \cdot \frac{y - x + \varepsilon n(y)}{|y - x + \varepsilon n(y)|} \right> j_2(x) \, dx \]
\[ = \int_S \left< j_1(y) \cdot \nabla_x \frac{1}{|y - x + \varepsilon n(y)|} \right> j_2(x) \, dx \]
\[ = \int_S \left< j_1(y) \cdot \frac{1}{|y - x + \varepsilon n(y)|} \right> j_2(x) \, dx \]
\[ + \int_S \left< j_1(y) \cdot \frac{y - x, n(x) + \varepsilon (n(y), n(x))}{|y - x + \varepsilon n(y)|^3} \right> j_2(x) \, dx. \quad (6) \]

Similarly, for the second integral in equation (4) we have:

\[ \int_S \left< j_1(y) \cdot j_2(x) \cdot \frac{y - x + \varepsilon n(y)}{|y - x + \varepsilon n(y)|} \right> dx \]
\[ = \int_S \left< j_1(y) \cdot j_2(x) \nabla_x \frac{1}{|y - x + \varepsilon n(y)|} \right> dx \]
\[ = \int_S \left< j_1(y) \cdot j_2(x) \nabla_y \frac{1}{|y - x + \varepsilon n(y)|} \right> dx \]
\[ + \int_S \left< j_1(y) \cdot j_2(x) \frac{y - x, n(x) + \varepsilon (n(y), n(x))}{|y - x + \varepsilon n(y)|^3} \right> n(x) \, dx. \quad (7) \]

Integrals (8) and (12), with integrands tangential to \( S \) (‘tangential terms’) are dealt with in appendices A.1 and A.2. Integrals (9) and (13), with integrands normal to \( S \) (‘normal terms’) are treated in appendices A.3 and A.4. Together, those appendices constitute proof of the following theorem.

**Theorem 1.** Let \( j_1, j_2 \in X^{1,2}(S) \). Then \( \hat{L}_e(j_1, j_2) \) has an \( \varepsilon \rightarrow 0 \) limit in \( L^p(S, \mathbb{R}^3) \), for any \( 1 \leq p < \infty \), denoted \( L(j_1, j_2) \). Furthermore, \( L \) is a continuous bilinear map \( X^{1,2}(S) \times X^{1,2}(S) \rightarrow L^p(S, \mathbb{R}^3) \) given by

\[ L(j_1, j_2)(y) = -\int_S \frac{1}{|y - x|} \text{div}_x(\pi_x j_1) \, dx \]
\[ + \pi_x j_1 \cdot \nabla_x j_2 \, dx \]
\[ + \int_S \left< j_1(y) \cdot n(x) \cdot \frac{y - x, n(x)}{|y - x|^3} \right> j_2(x) \, dx \]
\[ + \int_S \frac{1}{|y - x|} \left[ \int_S j_1(y) \cdot j_2(x) \text{div}_x(\pi_x) \right] \, dx. \quad (14) \]

\[ + \int_S \frac{1}{|y - x|} \left[ \int_S j_1(y) \cdot j_2(x) \right] \, dx \]
\[ = \int_S \frac{1}{|y - x + \varepsilon n(y)|} \left[ \int_S j_1(y) \cdot j_2(x) \right] \, dx \]
\[ - \int_S \frac{1}{|y - x + \varepsilon n(y)|} \left[ \int_S j_1(y) \cdot j_2(x) \right] \, dx. \quad (16) \]

\[ - \int_S \frac{1}{|y - x + \varepsilon n(y)|} \left[ \int_S j_1(y) \cdot j_2(x) \right] \, dx. \quad (17) \]

2.4. Expression in the zero-thickness limit

Recapitulating, the Laplace force has been initially defined as the \( \varepsilon \rightarrow 0 \) limit of the semi-sum of the magnetic field evaluated at a distance \( \varepsilon \) away from the CWS, \( S \), respectively inward and outward (equation (5)). This was shown to either be numerically costly or subject to numerical errors (figure 1).

An expression (equations (14)–(17)) has then been derived in theorem 1 for the Laplace force exerted by one current-sheet on another, per unit length. The special case \( j_1 = j_2 \) describes the self-interaction of a current-sheet.

Both treatments relied on an intrinsically 2D model for the currents on the CWS. A third approach is to treat the CWS as a 3D layer of infinitesimal thickness \( \varepsilon \). For some \( y \in S \) and if \( j \) is smooth enough, we could compute the \( \varepsilon \rightarrow 0 \) limit of

\[ \hat{L}_e(j_1)(y) = \int_{-\varepsilon/2}^{\varepsilon/2} [j_1(y + \varepsilon n(y)) \times B(y + \varepsilon n(y))] \, d\varepsilon_1. \]

Note that \( B(y + \varepsilon n(y)) \) is well-defined as we integrate on a 3D domain, and is given by:

\[ B(y + \varepsilon n(y)) = \int_{-\varepsilon/2}^{\varepsilon/2} \left[ j_1(x + \varepsilon n(x)) \times \frac{y - x + \varepsilon n(y) - \varepsilon n(x)}{|y - x + \varepsilon n(y) - \varepsilon n(x)|} \right] \, d\varepsilon_2. \]

In order to approximate the 3D volume with a 2D current-sheet, we suppose that \( \forall z \in S \) and \( \forall \varepsilon' \), we have \( j_1(z + \varepsilon' n(z)) = \frac{\varepsilon}{\varepsilon'} \). Thus,

\[ \hat{L}_e(j_1) = \int_{-\varepsilon/2}^{\varepsilon/2} \int_{-\varepsilon/2}^{\varepsilon/2} \left[ j_1(x) \times \left[ j_1(x) \times \frac{y - x + \varepsilon n(y) - \varepsilon n(x)}{|y - x + \varepsilon n(y) - \varepsilon n(x)|} \right] \, d\varepsilon_2 \right] \, d\varepsilon_1. \]

The quantity inside the brackets is very close to the one we got in theorem 1, starting with equations (2) and (3), except that we also have a contribution from \( \varepsilon n(x) \). It is possible to prove, using an argument similar to lemma 5, that replacing \( n(x) \) with \( n(y) \) does not change the limit. The intuition is that for \( x \) close to \( y \), \( n(x) \) is close to \( n(y) \). As a result, \( \hat{L}_e(j) \) has

\[ \text{(16)} \]

\[ \text{(17)} \]
the same limit as \( L_e(j) \) and the expression we found for the Laplace force (equations (14)–(17)) is consistent.

### 3. Examples of cost functions

After having rigorously defined the Laplace force-density \( L(j)(y) \) that a current-sheet of density \( j \) exerts on itself at location \( y \) (equations (14)–(17) for \( j_1 = j_2 = j \)), we now introduce some cost-functions to penalize high values of the force.

Two main options are possible, and considered here: (1) penalizing high cumulative (or, equivalently, surface-averaged) forces, or (2) penalizing or even forbidding excessively high locally maxima of the force. Further variants are possible for specific force-components (e.g. tangential or normal to the CWS) or a weighted combination of them, with higher weights assigned to the engineeringly more demanding component, depending on the specific stellarator design. Such variants go beyond the scope of the present paper, and are left for future work.

A natural choice from the functional analysis point of view is to use a penalization of the form

\[
|L(j)|_{L^p(S,R^3)} = \left( \int_S |L(j)(x)|^p \, dx \right)^{1/p}.
\]

(18)

The case \( p = 2 \) is well-known: it represents the cumulative (or, barring a factor, the surface-averaged) root-mean-square force. Higher values of \( p \) penalize more severely high values of the Laplace force (i.e., large oscillations around the average norm). By contrast, low values of \( p \) penalize the average norm of the Laplace force.

In principle it is also possible to use a \( L^\infty \) cost, \( \sup_j |L(j)| \), but the domain might be smaller than \( \mathcal{X}^{1,2}(S) \). However, such cost is not differentiable whenever the maximum is reached at multiple locations.

The second option is to introduce the cost

\[
C_e(j) = \int_S f_e(L(j)(x)) \, dx
\]

(19)
as the surface integral of the local cost

\[
f_e(w) = \frac{\max(w - c_0, 0)^2}{1 - \max(w - c_0, 0)}.
\]

(20)

The domain for this cost is not the entire space \( \mathcal{X}^{1,2}(S) \), but this cost captures more effectively the engineering constraints of building a high-field stellarator: the mechanical properties of support-structures and materials are such that forces below a threshold \( c_0 \) are negligible, forces higher and higher than \( c_0 \) should be penalized more and more, and forces above a second, ‘rupture’ threshold \( c_1 \) should be completely forbidden. Indeed, the local cost \( f_e \) evolves with the local force \( w \) as desired, as illustrated in figure 2.

**Remark 3.** It is unclear whether a minimizer exists in \( \mathcal{X}^{1,2}(S) \) for the costs discussed. As a consequence, a good practice is to add a regularizing term \( |j|_{\mathcal{X}^{1,2}(S)} \).

**Figure 2.** Plot of the local cost \( f_e \) as a function of the local force \( w \). Note that \( f_e \) diverges at \( c_1 \) and vanishes in \([0, c_0]\). In other words, the force is non-linearly optimized: small values are permitted, intermediate ones are increasingly, non-linearly penalized, and large ones are forbidden.

### 4. Numerical simulations

#### 4.1. Setup

To test our force-reduction method, we ran simulations for the NCSX stellarator equilibrium known as LI383 [2]: since our work is so closely related to REGCOIL, we adopted the same data-set, for ease of comparison with the original REGCOIL paper [1].

As mentioned in the introduction, the costs defined in section 3 are easily added to the cost-function in any stellarator coil optimization code. In our case such code was a new incarnation of REGCOIL, which we rewrote in python instead of fortran, and compiled with the Just In Time compiler Numba [19]. For the most part the new code is conceptually identical to REGCOIL, except that it uses equation (A5) of reference [1] in lieu of its normal, single-valued component (equation (A8) from the same paper). Equation (A5) would be numerically unstable if derivatives were taken by finite differences, but can be used here because we compute the derivatives explicitly. We compared results from the new code for LI383 and found them to agree with publicly available results from the original REGCOIL for the same case [1] to within 7 significant digits.

The surface-current \( j \) is divergence-free and thus taken in the form

\[
G \frac{\partial \Psi'}{\partial \theta} - I \frac{\partial \psi'}{\partial \zeta} + \frac{\partial \Phi'}{\partial \theta} \frac{\partial \psi'}{\partial \theta} - \frac{\partial \Psi'}{\partial \zeta} \frac{\partial \psi'}{\partial \theta}.
\]

(21)

Here \( \theta \) and \( \zeta \) are the poloidal and toroidal angle, \( G \) and \( I \) are optimization inputs (net poloidal and toroidal currents) and the current potential \( \Phi \) is decomposed in a 2D Fourier basis.

**Figure 3** illustrates how \( L_e \) converges to \( L \) (or, equivalently, the relative error vanishes) as \( \varepsilon \to 0 \). We recall that the numerical evaluation of \( L_e \) involves three characteristic distances \( h \) (discretisation length of the mesh), \( \varepsilon \), and \( d_B \) (characteristic distance of variation of the magnetic field). For reference we
computed $L$ on the same mesh (that is, for the same $h$), and obviously $d_B$ was also the same.

We observe that the error decreases with $\varepsilon$, as expected, but when $\varepsilon \lesssim h$ the convergence stops and the error grows again. This is a consequence of the calculation of $B$ not being accurate anymore, for $\varepsilon \lesssim h$ (see figure 1). Note that the reference value itself is an approximation. As we do not have an analytic expression, we also used a discretisation. The relative error is computed in $L^2$ norm.

In all simulations presented here, a period of both the CWS and the plasma surface were discretized as $64 \times 64$ meshes in the poloidal $\times$ toroidal direction. Half-periods of the two surfaces are rendered in 3D in figure 4.

The single-valued current potential $\Phi$ from which $j$ descends is represented by 8 or 12 harmonics in each direction. As we do not impose stellarator symmetry, we use as a basis the functions

$$\sin(k\theta + l\zeta), \cos(k\theta + l\zeta)$$

with $0 \leq k \leq N$ and $-N \leq l \leq N$. Since for $k = 0$ we can restrict to $0 < l$, the total number of degrees of freedom (DOF) is $2[(2N + 1)N + N]$.

Thus $N = 8$ harmonics in each direction correspond to 288 DOF, and $N = 12$ yields 624 DOF. Better results can be achieved with more harmonics, as shown in figure 5. However, a finer mesh is required, making the problem computationally more expensive.

The optimization is performed by conjugate gradient. With our implementation, a single evaluation of the gradient lasts approximately 2 min on a small cluster of 64 cores. The full optimization can last a few days.

4.2. Adding force minimization and improving regularization in REGCOIL

We propose to integrate the costs introduced above in the same optimization scheme as NESCOIL [16] and REGCOIL [1].

As a reminder, NESCOIL seeks the current of density $j$, on a fixed $S$, that maximizes magnetic field accuracy on the plasma boundary $S_P$ (hence, indirectly, in the plasma). It does so by minimizing the ‘plasma-shape objective’ or ‘field accuracy objective’

$$\chi^2_B = \int_{S_P} (B(x) \cdot n(x))^2 \, dS(x).$$

REGCOIL, instead, compromises between field accuracy and coil simplicity by minimizing $\chi^2_B + \lambda \chi^2_j$, where $\lambda$ is a weight and the ‘current-density objective’ or ‘regularizing term’ $\chi^2_j$ is a penalty on high values of $j$, in the sense of the $L^2$ norm:

$$\chi^2_j = \int_S |j|^2 \, dS.$$  \hspace{1cm} (23)

Heavier weighting makes $\Phi$ (hence $j$, hence the coils) more regular, but at the expense of reduced field accuracy. Such cost is identical to $\chi^2_K$ of reference [1], but is renamed $\chi^2_j$ for consistency of notation with another regularizing term that we need to introduce:

$$\chi^2_{Vj} = \int_S (|\nabla j_x|^2 + |\nabla j_y|^2 + |\nabla j_z|^2) \, dS.$$  \hspace{1cm} (24)

This new term is motivated by theorem 1: as the Laplace force can only be defined for $j \in X^{1,2}(S)$, it is natural to add a penalization on the gradient of $j$ and not just on $j$. Basically we are replacing the $L^2$ norm of $j$ with the $H^1$ norm of $j$.

Here we propose to further generalize the REGCOIL cost function to
where $\chi^2$ is a ‘force objective’ that penalizes strong forces on the current-sheet, i.e. among the coils. Per the discussion in section 3, possible definitions include:

$$\chi^2 = [L(\mathbf{j})]_{L=0(R,S)}^2 = \int_S |L(\mathbf{j})|^2 dS$$

where $F_{\gamma} = C_{\gamma} = \int_S f_{\gamma}(L(\mathbf{j})) dS$

(26)

with $f_{\gamma}$ defined as in equation (20) and plotted in figure 2. As stress limits, here we set $c_0 = 5 \times 10^6$ Pa and $c_1 = 10^7$ Pa.

4.3. Numerical results

There is obviously a trade-off between conflicting objectives in equation (25), or special cases of that equation. Special cases include the REGCOIL-like minimization of $\chi^2 = \chi_B^2 + \lambda_1 \chi^2_2$ and force-minimization without regularization ($\chi^2 = \chi_B^2 + \gamma \chi^2_2$).

In the REGCOIL-like case (curves in figure 5) we fixed $\lambda_2 = \gamma = 0$ and minimized $\chi_B^2 + \lambda_1 \chi^2_2$ for various choices of $\lambda_1$. By this scan we re-obtained the well-known trade-off between $\chi_B^2$ and $\chi^2_2$ (or, equivalently, field-accuracy and coil-simplicity) [1], but do not plot it for brevity. Interestingly, we also found a trade-off between $\chi_B^2$ and $\gamma \chi^2_2$, even though $\chi^2_2$ was not part of the $\chi_B^2 + \lambda_1 \chi^2_2$ minimization. In other words, more accurate fields come at the expense of higher forces, even when forces are not accounted in the minimization. The trade-off between these global quantities is plotted in figure 5(a), and a trade-off between related, local quantities is plotted in figure 5(b). This can be explained as follows. Accumulation of currents (high $\chi^2_2$), e.g. due to complicated patterns, typically leads to accumulation of forces (high $\chi^2_2$) because closer current-filaments exert stronger forces onto each other. This correlation between $\chi^2_2$ and $\chi^2_2$, combined with the well-known anti-correlation between $\chi^2_2$ and $\chi_B^2$ [1] implies that $\chi^2_F$ anti-correlates with $\chi^2_B$.

In the force-minimization case, instead, we fixed $\lambda_1 = \lambda_2 = 0$ and minimized $\chi_B^2 + \gamma \chi^2_2$ for various choices of $\gamma$. Not surprisingly, we found a trade-off between $\chi_B^2$ and $\chi^2_2$ (symbols in figure 5). Interestingly, $\chi_B^2$ also exhibits a trade-off with $\chi^2_2$, in spite of the latter not being part of the minimization. This suggests that $\chi_B^2$ has a regularizing effect on $\mathbf{j}$, as it will become apparent in figures 7 and 8.

Finally, figure 5 confirms that a higher number of Fourier harmonics and hence of DOF reproduces the magnetic field more accurately. This is why for the remainder of the article we adopt the higher number of DOF; 624.

Also, we no longer scan the weights, but fix them to yield reasonable compromises between field accuracy, current regularization and/or force minimization. In particular, calculations were performed for the following four choices of weights and $\chi^2_F$ in equation (25):

$$\begin{align*}
\text{Case} & & \lambda_1 & & \lambda_2 & & \gamma & & \chi^2_F \\
1 & & 1.5 \times 10^{-16} & & 0 & & 0 & & 0 \\
2 & & 0 & & 0 & & 10^{-17} & & |L(\mathbf{j})|^2_{L=0(R,S)} \\
3 & & 0 & & 0 & & 10^{-16} & & C_{\gamma} \\
4 & & 10^{-19} & & 10^{-19} & & 10^{-16} & & C_{\gamma} \\
\end{align*}$$

(28)

Case 1 is basically REGCOIL, whereas case 2 and 3 are effectively NESCOIL but with minimized forces, according to two different force metrics. Finally, case 4 explicitly combines force minimization with regularization, but in a broader sense compared to REGCOIL, as discussed in connection with equation (24). The force metric for case 2 penalizes high root mean squared surface-averaged forces (equation (26)), whereas the metric for cases 3 and 4 non-linearly penalizes high local forces (equation (27)).

The results for these four cases are plotted in figure 6 (circles) and compared with REGCOIL results (curve). In particular figure 6(a) refers to surface-integrated, ‘global’ objectives, and figure 6 to ‘local’ maxima. Note the logarithmic plots. As expected, case 1 agrees with REGCOIL. Case 2 (defined in terms of the ‘global’ $|L(\mathbf{j})|^2_{L=0(R,S)}$) overperforms in the ‘global’ figure 6(a), as expected. Actually, it performs better than REGCOIL even in terms of local metrics (figure 6(b)). Compared to REGCOIL, peak-forces are reduced in cases 3
and 4 (figure 6(b)), and remain lower than the chosen $c_1$, as is expected from the definition of $C_e$ and $f_e$ (equations (19) and (20)) However, this happens at the expense of higher cumulative forces (figure 6(a)).

Details on the four cases are presented in figures 7 and 8. Columns from left to right refer to cases from 1 to 4. From top to bottom, the rows in figure 7 present contours of (a) The norm of $j$, related to $\chi_j^2$ and $\chi_j^{\mathcal{L}_2}$. (b) The magnetic field normal to $S$, related to $\chi_B^2$, and (c) The norm of the Laplace force per unit surface, related to $\chi_F^2$.

The two rows in figure 8 present the force components normal and tangential to $S$. All quantities are plotted as functions of the poloidal and toroidal angles.

As anticipated, case 2 is as regular as case 1, in spite of its $\chi_j^2$ not containing a regularizing objective. By contrast, case 3 reproduces the field with high accuracy and exhibits reduced peak forces, as expected from the definition of $C_e$, but with a complicated current-pattern. That is ameliorated by adding some regularization: case 4 is the best compromise between coil simplicity (first row in figure 8), field accuracy (second row) and reduced forces (figure 9 and third row of figure 8).

Incidentally all cases, including REGCOIL (case 1) and the magnetically most accurate case 3, exhibit residual field errors of up to 60 mT. Lower errors can be achieved by adopting a higher number of DOF, as is intuitive and suggested by figure 5, but this is computationally more intensive and beyond the scope of the present paper.

From the point of view of the surface-integrated or surface-averaged forces, the best result in figure 7 is a modest reduction by 5% for case 2, relative to REGCOIL. From the point of view of peak forces, however, the best result in figure 7 is a reduction by 40% for case 4, relative to REGCOIL. Correspondingly, the peak tangential force is reduced by 50% and the peak normal force by 20% (figure 8). Note that maxima for different components occur at different toroidal and poloidal locations.

More dramatic reductions were obtained in figure 5, especially in peak forces. However, they were obtained for low-accuracy cases on the top left of figure 5(b): a stellarator with those characteristics would suffer from very low coil-forces, but it would also be a poor match of the target field.

There is some arbitrariness in how to discretize the continuous current-distributions of figure 7. Figure 10 illustrates possible filamentations for cases 1 and 4. Note the accumulation of current filaments, i.e. coils, in regions of high forces (the color contours in the background). This reflects the fact that, by definition, current filaments tend to ‘crowd’ in regions of high j, and this proximity results in high forces. However, for different discretizations they are subject to different forces. This offers additional DOF for force-minimization, which are left for future work.

5. Summary, conclusions and future work

To summarize, force-minimization is an important aspect of stellarator coil-optimization, especially for future high-field stellarators. In the present article we rigorously proved in section 2.3 that the Laplace force exerted by a surface-current onto one another can be written as in equations (14)–(17). From that, one can calculate the auto-interaction $L(j)$ of a current-distribution with itself, and distill that information in a single scalar. Possible metrics were discussed in section 3, and two of them were used for detailed numerical calculations: two possible ‘force objectives’ (equations (26) and (27)) were added to the cost function of the well-known REGCOIL code [1]. In addition, the $L^2$ norm of $j$ was replaced with the $H^1$ norm of $j$, for reasons explained in sections 2.1 and 4.2.

This approach permitted to simultaneously optimize the coils of the NCSX stellarator for magnetic fidelity, regularity and low forces. For instance, 40% lower peak-forces were obtained compared to REGCOIL, for similar plasma-shape accuracy and current regularity. These results were presented as case 1 and 4 in figures 6–8. Force reduction is an important criterion in stellarator optimization, and future high-field designs might benefit from our approach.

Unfortunately force minimization made the new approach (case 4) significantly slower than REGCOIL (case 1). This motivated the adoption of a low number of Fourier harmonics and thus of DOF in case 4 and, for consistency, in all cases.
Figure 7. Results of minimizing equation (25) for NCSX, for four different cases (four different choices of weights in the equation, as summarized by the table in equation (28)). Each column refers to a different case; its title is color-coded like the corresponding data-point in figure 6. From top to bottom the three rows refer respectively to the results for simultaneous (1) current regularization (if any), (2) field accuracy and (3) force-minimization (if any). Case 4 (last column) demonstrates that it is possible to simultaneously optimize these three competing objectives without excessively penalizing any of them with respect to established codes. On the contrary, case 4 actually exhibits higher field accuracy and lower peak forces compared to REGCOIL (first column). Shown in the legends are the root mean square (RMS) surface-averages and local maxima of the quantities plotted, as well as the $H^1$ norm of $j$ and $C_e$ force metric (equation (27)). The quantities actually minimized are marked in purple. 624 DOF are used for $j$ in every simulation. It is well-known from [1] and figure 5 that a higher number of DOF will lower all individual metrics $\chi_2^B$, $\chi_2^j$, $\chi_2^\nabla j$, $\chi_2^F$ and find better compromises among them. Correspondingly, all contours presented here will improve, for all four cases, and by the same proportion. However, this will require more computational resources, and is left as future work.

The resulting field inaccuracies are high, but they are just as high with REGCOIL, under the same circumstances (figure 7). Fortunately, reference [1] and figure 5 indicate that such inaccuracies rapidly disappear by adopting more DOF, and just as rapidly in our code as for REGCOIL. At the same time, more DOF lead to more coil-simplicity [1] and lower coil-forces (figure 5). In the future, optimizing the code for speed and/or running it on a super-computer will allow to retain a higher number of DOF.

In the present paper we introduced (figure 2) and successfully demonstrated (figure 5) the non-linear optimization of the coil-forces: we introduced constraints $c_0$ and $c_1$ to allow stresses below $c_0$, increasingly penalize stresses in the $[c_0, c_1]$ range, and forbid stresses above $c_1$. Future work could impose stricter constraints and tailor them differently for normal and longitudinal forces, as they tend to differ (figure 8) and they obey to different engineering and material constraints. In addition, we could non-linearly optimize other quantities. For example we could allow field inaccuracies of one part in $10^5$, penalize inaccuracies up to one part in $10^4$, and attribute infinite cost to larger discrepancies.

Finally, in the present work the CWS was fixed. Future shape optimization of the CWS, inspired by references [20, 21], is expected to further reduce the coil-forces.

Acknowledgments

The authors thank Mario Sigalotti for the fruitful discussions and for carefully reading the manuscript. The authors express their gratitude to M. Landreman and the NCSX group for making the original REGCOIL code and the LI383 dataset publicly available. This work has been partly supported by Inria’s Action Exploratoire StellaCage.
Appendix A

A.1. First tangential term

Integration by parts on a compact manifold $\mathcal{M}$ without boundary is given by the following formula. Let $f \in C^\infty(\mathcal{M})$ and $\mathbf{X}$ a smooth vector field on $\mathcal{M}$, then

$$\int_{\mathcal{M}} \text{div}(f \mathbf{X}) = 0 = X f + \int_{\mathcal{M}} f \text{ div } \mathbf{X}. \quad (29)$$

We also recall that $X f = \langle X \cdot \nabla f \rangle$ in Euclidean coordinates.

Let us start with the first tangential term (equation (8)):

$$\int_{S} \left\langle j_1(y) \cdot \nabla_{S} \left( \frac{1}{|y-x|} \right) \right\rangle_{\mathbb{R}^3} j_2(x) dx$$

$$= \int_{S} \left\langle \pi_x j_1(y) \cdot \nabla_{S} \left( \frac{1}{|y-x|} \right) \right\rangle_{T_x S} j_2(x) dx$$

$$= \int_{S} \left\langle \pi_x j_1(y) \cdot \nabla_{S} \left( \frac{1}{|y-x|} \right) \right\rangle_{T_x S} dx \quad (30)$$

$$= - \int_{S} \left\langle \frac{1}{|y-x|} \right\rangle_{T_x S} \text{div}_{x} (j_2(x) \pi_x j_1(y)) dx \quad (31)$$
Thus the term in equation (5) is equal to:

$$\sum_{j=1}^{3} \left( \int_{|y-x|+\varepsilon n(y)}^{1} \frac{1}{|y-x|+\varepsilon n(y)} \left[ f_{j}(x) \text{div}_{x}(\pi_{x}j_{1}(y)) \right] \right. \left. + \langle \pi_{x}j_{1}(y) \cdot \nabla f_{j}(x) \rangle \right) \, dx.$$ 

(32)

Note that the two integrals (respectively with the sign + or − at denominator) converge to the same limit. Their sum yields the integral on the right-hand-side of equation (14).

A.2. Second tangential term

Now, let us tackle the term in equation (12). We start by computing the $i$th component of that integral, i.e. its projection on $e_{i}$, then follow a derivation similar to equations (30)–(32):

$$\int_{S} \langle j_{1}(y) \cdot j_{2}(x) \rangle \left( \frac{1}{|y-x|+\varepsilon n(y)} \right) \, dx$$

$$= - \int_{S} \langle j_{1}(y) \cdot j_{2}(x) \rangle \left( \frac{1}{|y-x|+\varepsilon n(y)} \right) \, dx$$

$$= \left( \int_{S} \langle j_{1}(y) \cdot j_{2}(x) \rangle \langle \pi_{x}e_{i} \cdot \nabla_{x} \rangle \, dx \right) + \left( \langle \pi_{x}e_{i} \cdot \nabla_{x} \rangle \langle j_{1}(y) \cdot j_{2}(x) \rangle \right) \, dx.$$ 

Using the notation of remark 2, we find the vector form of the integral:

$$- \int_{S} \frac{1}{|y-x|+\varepsilon n(y)} \left( \langle j_{1}(y) \cdot j_{2}(x) \rangle \text{div}_{x}(\pi_{x}) \right. \left. + \nabla_{x} \langle j_{1}(y) \cdot j_{2}(x) \rangle \right) \, dx.$$ 

Due to the same arguments invoked in equations (33)–(35), both integrals, with the sign + and − at denominator, converge in $X^{0}(S)$ to the same limit,

$$- \int_{S} \frac{1}{|y-x|+\varepsilon n(y)} \left[ \langle j_{1}(y) \cdot j_{2}(x) \rangle \text{div}_{x}(\pi_{x}) + \nabla_{x} \langle j_{1}(y) \cdot j_{2}(x) \rangle \right] \, dx.$$ 

Their sum yields integral (16).

A.3. First normal term

Let us now focus on the normal component of 4, namely equation (9). This is in effect the sum of two integrals, which we will discuss separately:

$$\int_{S} \langle j_{1}(y) \cdot n(x) \rangle \left( \frac{y-x+n(x)}{|y-x|+\varepsilon n(x)} \right) j_{2}(x) \, dx$$

(36)

$$\int_{S} \langle j_{1}(y) \cdot n(x) \rangle \left( \frac{\pm \varepsilon n(x)}{|y-x|+\varepsilon n(x)} \right) j_{2}(x) \, dx.$$ 

(37)

First notice that we have the following estimate:

Lemma 3. \(\exists C > 0, \forall x \neq y \in S, \frac{|y-x \cdot n(x)|}{|y-x|} \leq C.\)

Proof of lemma 3. Let us suppose there exist two sequences $(x_{n})$, $(y_{n})$ in $S$ such that $x_{n} \neq y_{n}$ and $\frac{|y_{n} - x_{n} \cdot n(x_{n})|}{|y_{n} - x_{n}|} \to \infty$. Up to an extraction, we can suppose that $x_{n} \to x_{0} \in S$. If $y_{n}$ does not converge to $x_{0}$, we can extract a subsequence such that $\frac{|y_{n} - x_{n} \cdot n(x_{n})|}{|y_{n} - x_{n}|}$ does not diverge. This is a contradiction, hence both $x_{n}$ and $y_{n}$ converge to $x_{0} \in S$.

Let $\Gamma(x,y) = (y - x \cdot n(x))$. As $S$ is smooth, so is $\Gamma$. Its partial derivatives are...
Thus, at the point $(x_0, y_0)$, both first derivatives vanish. As a consequence, for $n$ large enough there exists $C > 0$ such that
\[ \Gamma(x_0, y_0) = C|x_0 - y_0|^2, \]
contradiction. □

Now, we need to find a minoration of $|y - x + \varepsilon n(y)|$.

**Lemma 4.** For $\varepsilon$ small enough, for all $\mu > 0$,\n\[
|x - y + \varepsilon n(y)|^\mu \geq \left( \frac{1}{\sqrt{2}} |x - y| \right)^\mu, \varepsilon^\mu.
\]

**Proof of lemma 4.**\n\[
|y - x + \varepsilon n(y)|^2 = |y - x|^2 + \varepsilon(y - x, n(y)) + \varepsilon^2
\geq |y - x|^2 - C \varepsilon |y - x|^2 + \varepsilon^2 \quad \text{by lemma 3.}
\]
Thus for $\varepsilon \leq 1/(2C)$, we have, $\forall \mu > 0$,
\[
|x - y + \varepsilon n(y)|^\mu \geq \left( \frac{1}{\sqrt{2}} |x - y| \right)^\mu, \varepsilon^\mu.
\]
□

Using lemma 3 and 4, for some constant $C_m$, \[ \frac{|y - x + \varepsilon n(y)|}{|y - x + \varepsilon n(y)|^2} \]
is dominated by $C \frac{1}{|y - x|^2}$, which is integrable. By dominated convergence
\[
\int_S \left( j_1(y) \cdot n(x) \right) \frac{\langle y - x, n(x) \rangle}{|y - x + \varepsilon n(y)|^3} j_2(x)dx
\]
\[
\times^{S^2} \int_S \left( j_1(y) \cdot n(x) \right) \frac{\langle y - x, n(x) \rangle}{|y - x|^3} j_2(x)dx, \quad (38)
\]
i.e. we obtained integral (15).

Now we have to deal with $\frac{\varepsilon n(y)}{|y - x + \varepsilon n(y)|^2}$, but we will show their net contribution to converge to zero.

To begin with, we could use the smallness of the term $\langle j_1(y) \cdot n(x) \rangle$ to ensure integrability. Instead, we will prove the following lemma which will also be useful later. Let $\Delta = \{(z, z) \in S \cap S \} \subset S^2$.

**Lemma 5.** Let $f_\varepsilon : S^2 \setminus \Delta \ni (x, y) \mapsto \frac{1}{|y - x + \varepsilon n(y)|^2} - \frac{\varepsilon n(y)}{|y - x + \varepsilon n(y)|^3} dx$. Then $\forall \eta > 0, \exists M > 0, \forall \varepsilon < \eta, \forall (x, y), |\varepsilon^\nu f_\varepsilon(x, y)| \leq M |\varepsilon^{1/2} - 1|.

**Proof of lemma 5.**
\[
f_\varepsilon(x, y) = \frac{|y - x - \varepsilon n(y)|^3 - |y - x + \varepsilon n(y)|^3}{|y - x + \varepsilon n(y)|^3 |y - x - \varepsilon n(y)|^3}
\]
\[
= \frac{|y - x - \varepsilon n(x)| - |y - x + \varepsilon n(\varepsilon)|}{|y - x + \varepsilon n(\varepsilon)|}
\times \frac{(|y - x + \varepsilon n(y)|^2 + |y - x + \varepsilon n(y)| |y - x - \varepsilon n(y)| + |y - x - \varepsilon n(y)|^2)}{|y - x + \varepsilon n(y)|^3 |y - x - \varepsilon n(y)|^3}.
\]
Using the fact that square root is $1/2$-Hölder continuous ($a \geq b \geq 0, \sqrt{a} - \sqrt{b} \leq \sqrt{a - b}$) and lemma 3, there exists $C > 0$ such that
\[
|y - x + \varepsilon n(y)| - |y - x - \varepsilon n(y)| \leq 2 \varepsilon |(x, y, n(y))| \leq C \varepsilon |x - y|.
\]
Now we use the minoration of the denominator from lemma 4. Up to a global multiplicative constant $M$, we get, for any $0 \leq \nu \leq 4$,\n\[
|f_\varepsilon(x, y)| \leq 4C \varepsilon |x - y|^{1 - \nu} \varepsilon^\nu
\leq M \frac{1}{\varepsilon^{\frac{3}{2} - \alpha}} |x - y|^\nu,
\]
for any $-0.5 \leq \alpha \leq 3.5$

Thanks to lemma 5 with any $\alpha \in (1/2, 1)$, there exists $C > 0$ such that
\[
\int_S \left( j_1(y) \cdot n(x) \right) \frac{\varepsilon n(y)}{|y - x + \varepsilon n(y)|^3} j_2(x)dx \times^{S^2} \int_S \left( j_1(y) \cdot n(x) \right) \frac{\varepsilon n(y)}{|y - x - \varepsilon n(y)|^3} j_2(x)dx = 0. \quad (39)
\]
In summary, equation (9) is the sum of two integrals converging respectively as in equations (38) and (39). Ultimately the ‘first normal term’ converges to equation (15).

### A.4. Second normal term

The same reasoning just applied to integral (9) also applies to integral (13),

\[
\int_{S} (j_1(y) \cdot j_2(x)) \frac{y - x \pm \varepsilon n(y)}{|y - x|} \, dx,
\]

which converges to

\[
\int_{S} (j_1(y) \cdot j_2(x)) \frac{y - x, n(x)}{|y - x|^3} \, dx,
\]

i.e. to equation (17).

This concludes the proof of theorem 1: one by one, in appendices A.1–A.4, we have obtained all terms in integral (13),

\[
\left| \left| \left| \int_{S} \left( j \frac{y - x}{|y - x|^3} \right) \cdot n \right| \right| \right| \leq C |j|_{L^2(S)}.
\]

**Remark 4.** We do not expect \(L(j_1, j_2)\) to be in \(L^\infty(S, \mathbb{R}^3)\). Indeed, \(H^1(S)\) is not embedded in \(L^\infty(S)\) for manifolds of dimension 2. For example, there is no constant \(C > 0\) such that

\[
\left| \int_{S} \frac{1}{|y - x|^3} \, dx \right| \leq C |j|_{L^2(S)}.
\]

**ORCID iDs**

Rémi Robin  
https://orcid.org/0000-0001-5609-1659

Francesco A. Volpe  
https://orcid.org/0000-0002-7193-7090

**References**

[1] Landreman M. 2017 An improved current potential method for fast computation of stellarator coil shapes *Nucl. Fusion* 57 046003

[2] Zarnstorff M.C. et al 2001 Physics of the compact advanced stellarator NCSX *Plasma Phys. Control. Fusion* 43 A237–49

[3] Helander P. 2014 Theory of plasma confinement in non-axisymmetric magnetic fields *Rep. Prog. Phys.* 77 087001

[4] Marmar E.S. et al 2015 Alcator C-mod: research in support of ITER and steps beyond *Nucl. Fusion* 55 104020

[5] Pucella G. et al 2015 Overview of the FTU results *Nucl. Fusion* 55 104005

[6] ITER Technical Basis 2002 Number 24 in ITER EDA Documentation Series (Vienna: International Atomic Energy Agency)

[7] Coppi B. et al 2013 New developments, plasma physics regimes and issues for the ignitor experiment *Nucl. Fusion* 53 104013

[8] Meade D et al 2002 Fire, exploring the frontiers of burning plasma science *J. Plasma Fusion Res. Ser.* 5 143–8

[9] Creely A.J. et al 2020 Overview of the SPARC tokamak *J. Plasma Phys.* 86 865860502

[10] Sagara A., Igitkhanov Y. and Najmabadi F. 2010 Review of stellarator/helicon design issues towards MFE DEMO *Fusion Eng. Des.* 85 1336–41

[11] Queral V., Volpe F.A., Spong D., Cabrera S. and Tabarés F. 2018 Initial exploration of high-field pulsed stellarator approach to ignition experiments *J. Fusion Energy* 37 275

[12] Schauer F., Egorov K. and Bykov V. 2013 Helias 5-b magnet system structure and maintenance concept *Fusion Eng. Des.* 88 1619–22

[13] Imagawa S., Sagara A., Watanabe K., Satow T. and Motojima O. 2002 Magnetic field and force of helical coils for force free helical reactor (FFHR) *J. Plasma Fusion Res. Ser.* 5 537–40

[14] Alonso-Rodríguez A., Camaño J., Rodríguez R., Valli A. and Venegas P. 2018 Finite element approximation of the spectrum of the curl operator in a multiply connected domain *Found. Comput. Math.* 18 1493–533

[15] Merkel P. 1986 An integral equation technique for the exterior and interior Neumann problem in toroidal regions *J. Comput. Phys.* 66 83–98

[16] Merkel P. 1987 Solution of stellarator boundary value problems with external currents *Nucl. Fusion* 27 867–71

[17] Strickler D.J., Hirshman S.P., Spong D.A., Cole M.J., Lyon J.F., Nelson B.E., Williamson D.E. and Ware A.S. 2004 Development of a robust quasi-poloidal compact stellarator *Fusion Sci. Technol.* 45 15–26

[18] Zhu C., Hudson S.R., Song Y. and Wan Y. 2018 New method to design stellarator coils without the winding surface *Nucl. Fusion* 58 016008

[19] Lam S.K., Pitrou A. and Seibt S. 2015 Numba: a LLVM-based python JIT compiler *Proc. 2nd Workshop on the LLVM Compiler Infrastructure in HPC, LLVM’15* (New York: Association for Computing Machinery)

[20] Paul E. 2020 Adjoint methods for stellarator shape optimization and sensitivity analysis *PhD Dissertation* University of Maryland

[21] Privat Y., Robin R and Sigalotti M 2022 Optimal shape of stellarators for magnetic confinement fusion *J. Math. Pures.* 163 231–64

[22] Han Y. and Zhu M. 2016 Hardy–Littlewood–Sobolev inequalities on compact Riemannian manifolds and applications *J. Differ. Equ.* 260 1–25

[23] Hebey E. 2000 *Nonlinear Analysis on Manifolds: Sobolev Spaces and Inequalities* (Courant Lecture Notes in Mathematics) (New York: Courant Institute of Mathematical Sciences)