Proceedings of the Royal Society of Edinburgh, **135A**, 915–933, *2005*

Optimal coercivity inequalities in $W^{1,p}(\Omega)$

Giles Auchmuty
Department of Mathematics, University of Houston,
Houston, TX 77204-3008, USA (auchmuty@uh.edu)

(MS received 30 June 2004; accepted 24 February 2005)

This paper describes the characterization of optimal constants for some coercivity inequalities in $W^{1,p}(\Omega)$, $1 < p < \infty$. A general result involving inequalities of p-homogeneous forms on a reflexive Banach space is first proved. The constants are shown to be the least eigenvalues of certain eigenproblems with equality holding for the corresponding eigenfunctions. This result is applied to three different classes of coercivity results on $W^{1,p}(\Omega)$. The inequalities include very general versions of the Friedrichs and Poincaré inequalities. Scaling laws for the inequalities are also described.

1. Introduction

This paper describes the characterization of optimal constants, and corresponding optimal functions, for some inequalities satisfied by functions in the Sobolev spaces $W^{1,p}(\Omega)$, $1 < p < \infty$. Here Ω is a bounded, connected, open set in $\mathbb{R}^n$ which satisfies the assumptions 2.1 and 2.2 given below.

First, some general properties of p-homogeneous inequalities on a reflexive Banach space are derived in §§ 3–5. It is shown that, for certain classes of problems, the optimal constants may be found using either a constrained, or an associated unconstrained, variational principle. The extremality condition for the minimizer of the unconstrained problem is used to show that the optimal constant is the least eigenvalue of a related eigenproblem. The corresponding eigenfunctions will be optimal functions for the inequality.

These results are then applied to three different classes of inequalities on $W^{1,p}(\Omega)$. These include $W^{1,p}$-versions of Friedrichs-type inequalities in §§ 6 and 7, some different inequalities involving boundary integrals in §§ 8 and 9, and some generalized Poincaré-type inequalities in § 10. For each of the inequalities, an associated scaling law is described.

These inequalities generalize some well-known, and often used, results from the theory of Sobolev spaces. In particular, the inequalities are used to prove the equivalence of various norms on $W^{1,p}(\Omega)$. Such results appear in [14, ch. 7.4] and more recent results appear in [3, § 6.3.5]. Most of the published proofs are non-constructive. For numerical and other purposes it is of considerable interest to know how the constants depend on the geometry and size of the underlying region. The results are illustrated in § 11 by describing the constants in some special cases for rectangular boxes in the plane.

Here, three different classes of inequalities will be distinguished. They differ in only one of the functionals involved, but the associated extremality conditions lead to different types of eigenproblems at optimality. The inequalities have a variety of names in the literature. Currently the term *Poincaré inequality* is commonly used for inequalities of the form (2.5) below. Here we will follow the older usage [12, § 5.3] of calling results similar to (6.1) *Friedrichs' inequality*. Recently, there has also been interest in the discrete analogues of these inequalities (see [7] and the references therein).

In this paper we shall use various standard results from the calculus of variations and convex analysis. Background material on such methods may be found in [6, 15], both of which have discussions of the variational principles for the Dirichlet eigenvalues and eigenfunctions of second-order elliptic operators. The variational principles used here are variants of the principles described there and are analogous to those for the Laplacian described in [5, § 5]. Some different unconstrained variational principles for eigenvalue problems are described in [4].

2. Definitions and notation

Let Ω be a non-empty, bounded, connected, open subset of $\mathbb{R}^n$ with boundary $\partial\Omega$. Such a set Ω is called a *region*. Let $L^p(\Omega)$ be the usual real Lebesgue space of all functions $u : \Omega \to [-\infty, \infty]$ which are pth-power integrable with respect to Lebesgue measure on Ω, $1 < p < \infty$. Let σ and $\mathrm{d}\sigma$ represent the Hausdorff $(n-1)$-dimensional measure and integration with respect to this measure, respectively. The space $L^p(\partial\Omega, \mathrm{d}\sigma)$ is the space of all such pth-power integrable functions on $\partial\Omega$. The corresponding norms are $\|u\|_p$ and $\|u\|_{p,\partial\Omega}$ and are defined by

$$\|u\|_p^p := \int_\Omega |u|^p \,\mathrm{d}x \quad \text{and} \quad \|u\|_{p,\partial\Omega}^p := \int_{\partial\Omega} |u|^p \,\mathrm{d}\sigma. \tag{2.1}$$

All functions in this paper will take values in $\bar{\mathbb{R}} := [-\infty, \infty]$ and we shall write

$$\bar{u} := |\Omega|^{-1} \int_\Omega u \,\mathrm{d}x \quad \text{and} \quad \bar{u}_\partial := |\sigma(\partial\Omega)|^{-1} \int_{\partial\Omega} u \,\mathrm{d}\sigma$$

for the mean values of u over the region Ω and the boundary $\partial\Omega$, respectively. Also $p^* := p/(p-1)$ is the dual index to p.

When $u \in L^p(\Omega)$, its weak jth derivative is denoted by $\mathrm{D}_j u$. The Sobolev space $W^{1,p}(\Omega)$ is defined to be the space of all functions in $L^p(\Omega)$ whose weak first derivatives $\mathrm{D}_j u$, $1 \leqslant j \leqslant n$, are all in $L^p(\Omega)$. The standard norm on $W^{1,p}(\Omega)$ is denoted $\|u\|_{1,p}$ and is defined by

$$\|u\|_{1,p}^p := \int_\Omega \left[\sum_{j=1}^n |\mathrm{D}_j u|^p + |u|^p \right] \mathrm{d}x. \tag{2.2}$$

This space is a Banach space. When $p = 2$, this becomes the Hilbert space $H^1(\Omega)$ with the standard H^1-inner product

$$[u, v]_1 := \int_\Omega [u(x)v(x) + \nabla u(x) \cdot \nabla v(x)] \,\mathrm{d}x. \tag{2.3}$$

Here $\nabla u := (\mathrm{D}_1 u, \mathrm{D}_2 u, \ldots, \mathrm{D}_n u)$ is the gradient of the function u and we shall write

$$\|\nabla u\|_p^p := \int_\Omega \sum_{j=1}^n |\mathrm{D}_j u|^p \, \mathrm{d}x. \tag{2.4}$$

For our analysis we also require some mild regularity conditions on Ω and $\partial\Omega$. First we shall require that the Sobolev embedding theorem and the Rellich–Kondrachov theorem hold for $W^{1,p}(\Omega)$, as is stated below.

Assumption 2.1. The embedding $i : W^{1,p}(\Omega) \to C^0(\bar{\Omega})$ is compact when $p > n$ and $i : W^{1,p}(\Omega) \to L^q(\Omega)$ is compact for $1 \leqslant q < q_c$ when $p \leqslant n$ and $q_c = np/(n-p)$.

Criteria for this assumption are given in [1] and [8, ch. V]. In particular, when assumption 2.1 holds, then the *Poincaré inequality* follows. Namely, there is a $C_p = C_p(\Omega) > 0$, which depends on Ω and p only, such that

$$\|\nabla u\|_p \geqslant C_p \|u - \bar{u}\|_p \quad \text{for all } u \in W^{1,p}(\Omega). \tag{2.5}$$

For a (non-constructive) proof see [9, § 5.8]. A detailed analysis of criteria that guarantee inequalities such as this is found in [8, ch. V, §§ 4 and 5].

We also require a trace condition. Assume that $\partial\Omega$ has a finite surface σ-measure and is a finite union of disjoint Lipschitz surfaces. When this holds, there is an outward unit normal ν defined σ-almost everywhere (σ-a.e.) on $\partial\Omega$. For the definition of this and related terms, see [10, ch. 4]. Let Γ denote the boundary trace operator. We will then require the following assumption.

Assumption 2.2. The boundary trace operator $\Gamma : W^{1,p}(\Omega) \to L^p(\partial\Omega, \mathrm{d}\sigma)$ is continuous.

A functional $\mathcal{F} : W^{1,p}(\Omega) \to (-\infty, \infty]$ is said to be *homogeneous of degree p* (or *p-homogeneous*) provided that

$$\mathcal{F}(cu) = |c|^p \mathcal{F}(u) \quad \text{for all } c \in \mathbb{R},\ u \in W^{1,p}(\Omega).$$

We say the functional is *positive* if $\mathcal{F}(u) \geqslant 0$ for all u. It is said to be *G-differentiable* at $u \in W^{1,p}(\Omega)$ if there is a continuous linear functional $\mathrm{D}\mathcal{F}(u)$ such that

$$\lim_{t \to 0} t^{-1}[\mathcal{F}(u + th) - \mathcal{F}(u)] = D\mathcal{F}(u)(h) \quad \text{for all } h \in W^{1,p}(\Omega),$$

In this case, $\mathrm{D}\mathcal{F}(u)$ is called the *G-derivative* of $\mathcal{F}$ at u and this expression will also be denoted $\langle D\mathcal{F}(u), h \rangle$.

A real sequence $\{a_m : m \geqslant 1\}$ is said to be *decreasing* if $a_{m+1} \leqslant a_m$ and *strictly decreasing* for $a_{m+1} < a_m$, respectively, for all m. A function u is said to be *positive* on a set E, if $u(x) \geqslant 0$ and *strictly positive* if $u(x) > 0$, respectively, on E.

3. The *p*-homogeneous inequality

Our interest is in finding the optimal constant C_0, in inequalities of the form

$$\mathcal{F}(u) := \int_\Omega \sum_{j=1}^n |\mathrm{D}_j u|^p \, \mathrm{d}x + \mathcal{B}(u) \geqslant C_0 \mathcal{P}(u), \tag{3.1}$$

where $1 < p < \infty$, $\mathcal{B}$ and $\mathcal{P}$ are p-homogeneous functionals on $W^{1,p}(\Omega)$ and $C_0 > 0$. The constant C_0 in (3.1) is said to be *optimal* if it is the largest number such that (3.1) holds. A non-zero function $\hat{u}$ in $W^{1,p}(\Omega)$ *optimizes* (3.1) if the equality holds in (3.1) with the optimal choice of C_0. When $\hat{u}$ optimizes (3.1), so does any multiple of $\hat{u}$.

We are particularly interested in the case where $\mathcal{P} : W^{1,p}(\Omega) \to [0, \infty]$ is defined by

$$\mathcal{P}(u) := \int_\Omega \rho(x)|u(x)|^p \,\mathrm{d}x. \tag{3.2}$$

and $\rho : \Omega \to [0, \infty]$ satisfies the following assumption.

Assumption 3.1. The function ρ is positive with $\int_\Omega \rho \,\mathrm{d}x > 0$. When $1 < p \leqslant n$, ρ is in $L^q(\Omega)$ for some $q > q_0$ with $q_0 := n/p$. When $p > n$, ρ is in $L^1(\Omega)$.

Some properties of this functional may be summarized as follows.

Proposition 3.2. *Assume that Ω satisfies assumption 2.1, ρ satisfies assumption 3.1 and $\mathcal{P}$ is defined by (3.2). Then $\mathcal{P}$ is positive, bounded, convex, weakly continuous and G-differentiable on $W^{1,p}(\Omega)$ with*

$$\langle D\mathcal{P}(u), h\rangle = p \int_\Omega \rho|u|^{p-2}uh \,\mathrm{d}x \quad \textit{for all } u, h \in W^{1,p}(\Omega). \tag{3.3}$$

Proof. First consider the case $p > n$. From assumption 2.1, $u \in W^{1,p}(\Omega)$ implies that u is in $C^0(\bar{\Omega})$ and this embedding is compact. Use Hölder's inequality and assumption 3.1. Then

$$0 \leqslant \mathcal{P}(u) \leqslant \|\rho\|_1 \|u\|_\infty^p,$$

so $\mathcal{P}$ is continuous and bounded. $\mathcal{P}$ is convex as $|s|^p$ is convex on $\mathbb{R}$.

Assume $\{u_m : m \geqslant 1\}$ converges weakly to $\hat{u}$ in $W^{1,p}(\Omega)$. From assumption 2.1, it converges to $\hat{u}$ in the uniform norm on $C^0(\bar{\Omega})$. Thus

$$\rho(x)|u_m(x)|^p \to \rho(x)|\hat{u}(x)|^p \quad \text{pointwise on } \Omega.$$

The Lebesgue-dominated convergence theorem now implies that $\mathcal{P}$ is weakly continuous on $W^{1,p}(\Omega)$.

When $1 < p \leqslant n$ and assumption 3.1 holds, Hölder's inequality yields

$$0 \leqslant \mathcal{P}(u) \leqslant \|\rho\|_q \||u|^p\|_{q^*}. \tag{3.4}$$

Choose $r = pq/(q-1)$. Then, when assumption 2.1 holds, the embedding $i : W^{1,p}(\Omega) \to L^r(\Omega)$ is compact as $p < r < np/(p-1)$. Also

$$\||u|^p\|_{q^*} = \|u\|_r^{p/r},$$

so $\mathcal{P}$ is positive and bounded on $W^{1,p}(\Omega)$. It is convex as before.

If $\{u_m : m \geqslant 1\}$ converges weakly to $\hat{u}$ in $W^{1,p}(\Omega)$, then, from assumption 2.1, it converges strongly in $L^r(\Omega)$. Thus, there is a subsequence $\{u_{m_j} : j \geqslant 1\}$ which converges a.e. to $\hat{u}$ on Ω. In view of (3.4) and the previous equality, Lebesgue's dominated convergence theorem can be applied and $\mathcal{P}$ is weakly continuous as claimed.

The function $\psi(s) := |s|^p$ with $p > 1$ is continuously differentiable on $\mathbb{R}$ with $\psi'(s) = p|s|^{p-2}s$ for $s \neq 0$ and $\psi'(0) = 0$. Define

$$\Psi(t) := \mathcal{P}(u + th) = \int_{\Omega} \rho|u + th|^p \, \mathrm{d}x$$

with $u, h \in W^{1,p}(\Omega)$. Then the G-derivative of $\mathcal{P}$ at u is given by

$$\Psi'(0) = \langle D\mathcal{P}(u), h\rangle.$$

The conditions of [13, corollary 1.2.2, p. 124] hold in our case, so the t-derivative can be taken under the integral, and (3.3) follows. □

Note that this result implies that $\mathcal{P}(u)^{1/p}$ is a weakly continuous semi-norm on $W^{1,p}(\Omega)$.

The essential requirement for the functional $\mathcal{B} : W^{1,p}(\Omega) \to [0, \infty]$ is the following.

ASSUMPTION 3.3. The functional $\mathcal{B}$ is weakly lower semi-continuous (l. s. c.) on $W^{1,p}(\Omega)$.

We will describe inequalities based on three specific examples of this functional $\mathcal{B}$. The first is

$$\mathcal{B}_1(u) := \int_{\partial\Omega} b|\Gamma u|^p \, \mathrm{d}\sigma. \tag{3.5}$$

In the following, the trace operator Γ will often be omitted. We will require that the following holds.

ASSUMPTION 3.4. $b : \partial\Omega \to [0, \infty)$ is in $L^\infty(\partial\Omega, \mathrm{d}\sigma)$ and

$$\int_{\partial\Omega} b \, \mathrm{d}\sigma := b_0 > 0. \tag{3.6}$$

A second example is

$$\mathcal{B}_2(u) := \left| \int_{\partial\Omega} b\Gamma u \, \mathrm{d}\sigma \right|^p, \tag{3.7}$$

where we require that the following holds.

ASSUMPTION 3.5. $b : \partial\Omega \to [0, \infty]$ is in $L^{p^*}(\partial\Omega, \mathrm{d}\sigma)$ and (3.6) holds.

A third example is

$$\mathcal{B}_3(u) := \left| \int_{\Omega} c(x)u(x) \, \mathrm{d}x \right|^p. \tag{3.8}$$

We will require the following.

ASSUMPTION 3.6. $c : \Omega \to [0, \infty]$ is in $L^{p^*}(\Omega)$ and

$$\int_{\Omega} c(x) \, \mathrm{d}x := \bar{c}|\Omega| > 0. \tag{3.9}$$

We shall treat the inequalities associated with each of these choices of $\mathcal{B}$ separately in later sections. In each of these examples, the functions ρ, b and/or c may be zero on sets of positive measure; in many important applications they will be characteristic functions of specific subsets.

Both $\mathcal{B}_2$ and $\mathcal{B}_3$ are functionals of the form

$$\mathcal{B}(u) := |b(u)|^p, \tag{3.10}$$

with b being a continuous linear functional on $W^{1,p}(\Omega)$. We will use the following general result for functionals of this form.

PROPOSITION 3.7. *Assume that b is a continuous linear functional on $W^{1,p}(\Omega)$, that $\mathcal{B}$ is defined by (3.10) and $1 < p < \infty$. Then $\mathcal{B}$ is continuous, convex and satisfies assumption 3.3. $\mathcal{B}$ is G-differentiable with $D\mathcal{B}(u) = 0$ when $b(u) = 0$ and*

$$\langle D\mathcal{B}(u), h\rangle = p|b(u)|^{p-2}b(u)b(h) \quad \textit{for all } u, h \in W^{1,p}(\Omega). \tag{3.11}$$

Proof. When $p > 1$, let $\psi(s) := |s|^p$, as above. ψ is convex and continuously differentiable on $\mathbb{R}$. Using standard results on compositions, $\mathcal{B}$ will be continuous and convex. Thus, it is weakly l. s. c. on $W^{1,p}(\Omega)$. Applying the chain rule for G-derivatives and the expression for ψ', (3.11) follows. □

4. Scaling of p-homogeneous inequalities

The inequalities studied here arise in the numerical analysis of elliptic equations, so it is of interest to know how they scale with the size of the domain. Given a reference region Ω, and $L > 0$, we define the scaled region $\Omega_L := \{Lx : x \in \Omega\}$.

Define the dilation operator $\mathcal{S}_L : W^{1,p}(\Omega) \to W^{1,p}(\Omega_L)$ by

$$\mathcal{S}_L u(y) := u_L(y) := u\left(\frac{y}{L}\right) \quad \text{for} \quad y \in \Omega_L. \tag{4.1}$$

$\mathcal{S}_L$ is a linear isomorphism and the change-of-variables rule yields that

$$\int_{\Omega_L} \sum_{j=1}^{n} |\mathrm{D}_j u_L(y)|^p \,\mathrm{d}y = L^{n-p} \int_{\Omega} \sum_{j=1}^{n} |\mathrm{D}_j u(x)|^p \,\mathrm{d}x. \tag{4.2}$$

For $L > 0$, define the functional $\mathcal{P}_L : W^{1,p}(\Omega_L) \to [0, \infty]$ by

$$\mathcal{P}_L(u_L) := \int_{\Omega_L} \rho_L(y)|u_L(y)|^p \,\mathrm{d}y \tag{4.3}$$

with $\rho_L := \mathcal{S}_L\rho$. This functional satisfies

$$\mathcal{P}_L(u_L) = L^n\mathcal{P}(u) \quad \text{for each } u \in W^{1,p}(\Omega). \tag{4.4}$$

Similarly, define the function $b_L : \partial\Omega_L \to [0, \infty]$ by $b_L(y) := b(y/L)$. Then the functionals $\mathcal{B}_{jL}$ defined by (3.5), (3.7) and (3.8) on $W^{1,p}(\Omega_L)$ scale according to

$$\mathcal{B}_{jL}(u_L) = L^{n-1}\mathcal{B}_j(u) \quad \text{for } j = 1, 2 \quad \text{and} \quad \mathcal{B}_{3L}(u_L) = L^n\mathcal{B}_3(u).$$

In the following sections we shall describe the scaled versions of the inequalities of the form (3.1). In each case the optimal functions will be dilations of the optimal function with $L = 1$.

5. Variational principles for optimal constants

In this section, we show that the problem of finding the optimal constants, and functions, in inequalities of the form (3.1) can be described using an unconstrained variational principle. This reformulation enables a much simpler description of the extremality conditions.

We now assume that X is a real reflexive Banach space. The following will also be used.

ASSUMPTION 5.1. $\mathcal{F}$ and $\mathcal{P}$ are continuous functionals on X which are homogeneous of degree $p > 1$.

ASSUMPTION 5.2. $\mathcal{F}$ is convex and there exists $c_0 > 0$ such that

$$\mathcal{F}(u) \geqslant c_0 \|u\|_X^p \quad \text{for all } u \in X. \tag{5.1}$$

ASSUMPTION 5.3. $\mathcal{P}$ is weakly continuous and bounded on X and there exists $v \in X$ such that $\mathcal{P}(v) > 0$.

Define

$$B := \{u \in X : \mathcal{F}(u) \leqslant 1\} \quad \text{and} \quad S := \{u \in X : \mathcal{F}(u) = 1\}. \tag{5.2}$$

Consider the variational problem of finding

$$\beta := \sup_{u \in S} \mathcal{P}(u). \tag{5.3}$$

When this β is finite, then the homogeneity condition (assumption 5.1) implies that

$$\beta \mathcal{F}(u) \geqslant \mathcal{P}(u) \quad \text{for all } u \in X. \tag{5.4}$$

Now assumptions 5.2 and 5.3 imply that $\beta > 0$, so we also have

$$\mathcal{F}(u) \geqslant \beta^{-1} \mathcal{P}(u) \quad \text{for all } u \in X. \tag{5.5}$$

THEOREM 5.4. *Assume that assumptions 5.1, 5.2 and 5.3 hold. Then β defined by (5.3) is finite and there is a $\hat{u} \in S$ with $\mathcal{P}(\hat{u}) = \beta$.*

Proof. First we shall show that $\beta := \sup_{u \in B} \mathcal{P}(u)$ is finite and that this supremum is attained. Then we prove this supremum is attained at a point in S which leads to the theorem.

The set B is closed, convex and bounded, so it is weakly compact, as $W^{1,p}(\Omega)$ is reflexive. $\mathcal{P}$ is weakly continuous from assumption 5.3, so there is a finite $\beta_0 > 0$ such that $\beta_0 = \sup_{u \in B} \mathcal{P}(u)$ and this infimum is attained, so there is a $\hat{u}$ in B with $\mathcal{P}(\hat{u}) = \beta_0$.

If $\mathcal{F}(\hat{u}) < 1$, then $\tau \hat{u} \in B$ for some $\tau > 1$. Then $\mathcal{P}(\tau \hat{u}) = |\tau|^p \beta_0 > \beta_0$. This contradicts the definition of β_0, so $\mathcal{F}(\hat{u}) = 1$. This implies that $\beta = \beta_0$, and the theorem follows. □

COROLLARY 5.5. *Assume that assumptions 5.1, 5.2 and 5.3 hold. Then (3.1) holds with $C_0 = \beta^{-1}$. Moreover, this constant is optimal and equality holds in (3.1) when u is any multiple of $\hat{u}$.*

Proof. Equation (3.1) follows from theorem 5.4 and equation (5.5). Substitution shows that equality holds in (5.5) whenever u is a multiple of any function $\hat{u}$ in S for which $\mathcal{P}(\hat{u}) = \beta$. The same is implied for (3.1). □

The variational principle described above provides the usual constrained variational characterization of the best constants in inequalities such as (3.1). Now, consider the functional $\mathcal{J} : X \to \mathbb{R}$ defined by

$$\mathcal{J}(u) := \tfrac{1}{2}\mathcal{F}(u)^2 - \mathcal{P}(u), \tag{5.6}$$

and the unconstrained problem of finding the infimum of $\mathcal{J}$ on X. The following theorem holds for this problem.

THEOREM 5.6. *Assume that assumptions 5.1, 5.2 and 5.3 hold with β and $\mathcal{J}$ defined by (5.3) and (5.6). Then $\mathcal{J}$ is weakly* l. s. c. *and coercive on X and attains its infimum at points $\pm\tilde{u}$ in X, where $\tilde{u} = \beta^{1/p}\hat{u}$ and $\hat{u}$ is a maximizer of $\mathcal{P}$ on S.*

Proof. When $\mathcal{F}$ is convex and continuous on X, so is $\mathcal{F}^2$, because it is positive. Thus, $\mathcal{F}^2$ and $\mathcal{J}$ are weakly l. s. c. on X. Define $c_1 := \sup_{u\in S} |\mathcal{P}(u)|$. This is finite from assumption 5.3, and homogeneity implies that $\mathcal{P}(u) \leqslant c_1\|u\|_X^p$ for all $u \in X$. Thus,

$$\mathcal{J}(u) \geqslant c_0^2\|u\|_X^{2p} - C_1\|u\|_X^p$$

for all $u \in X$. This shows that $\mathcal{J}$ is coercive on X. Hence, $\mathcal{J}$ is bounded below on X and attains its infimum. If $\tilde{u}$ is a minimizer, so is $-\tilde{u}$ as $\mathcal{J}$ is even.

Given $u \in X$, $u \neq 0$, $p > 1$, consider the ray $\Gamma_u := \{su : s > 0\}$. Let z be the unique point on this ray satisfying $\mathcal{F}(z) = 1$. Then $u = tz$, where $t^p = \mathcal{F}(u)$. Now

$$\mathcal{J}(sz) = \tfrac{1}{2}s^{2p} - s^p\mathcal{P}(z).$$

If $\mathcal{P}(z) \leqslant 0$, this expression is minimized at $s = 0$ and the infimum of $\mathcal{J}$ along the ray Γ_u is zero. If $\mathcal{P}(z) > 0$, this expression is minimized at $\tilde{s}$, where

$$\tilde{s}^p = \mathcal{P}(z) \quad \text{and} \quad \inf_{s>0} \mathcal{J}(sz) = -\tfrac{1}{2}\mathcal{P}(z)^2. \tag{5.7}$$

For $z \in S$, define $\mathcal{P}_+(z) := \max(0, \mathcal{P}(z))$. Then

$$\inf_{u\in X} \mathcal{J}(u) = \inf_{z\in S}\inf_{s>0} \mathcal{J}(sz) = \inf_{z\in S}[-\tfrac{1}{2}\mathcal{P}_+(z)^2] = -\tfrac{1}{2}\beta^2. \tag{5.8}$$

That is, the minimizers of $\mathcal{J}$ on X are $\tilde{u} = \pm\beta^{1/p}\hat{u}$, where $\hat{u}$ is a maximizer of $\mathcal{P}$ on S. Moreover, at these minimizers

$$\mathcal{F}(\tilde{u}) = \beta, \quad \mathcal{P}(\tilde{u}) = \beta^2 \quad \text{and} \quad \mathcal{J}(\tilde{u}) = -\tfrac{1}{2}\beta^2. \tag{5.9}$$

□

To describe the conditions satisfied by the solutions of these problems, we shall require the following.

ASSUMPTION 5.7. $\mathcal{F}, \mathcal{P}$ are G-differentiable on X.

When a functional $\mathcal{P}$ is p-homogeneous and G-differentiable at $u \in X$, the differentiation of $\mathcal{P}(tu)$ at $t = 1$ yields *Euler's rule*:

$$\langle \mathrm{D}\mathcal{P}(u), u\rangle = p\mathcal{P}(u). \tag{5.10}$$

Corollary 5.8. *Assume that assumptions 5.1, 5.2, 5.3 and 5.7 hold and that $\tilde{u}$ minimizes $\mathcal{J}$ on X. Then $\tilde{u}$ is a solution of*

$$\mathrm{D}\mathcal{F}(u) = \beta^{-1}\mathrm{D}\mathcal{P}(u). \tag{5.11}$$

Proof. Apply the chain rule to (5.6). Then

$$\mathrm{D}\mathcal{J}(u) = \mathcal{F}(u)\,\mathrm{D}\mathcal{F}(u) - \mathrm{D}\mathcal{P}(u).$$

From (5.9) the minimizers of $\mathcal{J}$ on X have $\mathcal{F}(\tilde{u}) = \beta$, so (5.11) follows. □

Consider now the general eigenvalue problem of solving

$$\mathrm{D}\mathcal{F}(u) = \mu\,\mathrm{D}\mathcal{P}(u). \tag{5.12}$$

That is, we wish to find those $(\mu, u) \in \mathbb{R} \times X$, with $u \neq 0$, which solve (5.12). This will be interpreted in the weak form

$$\langle \mathrm{D}\mathcal{F}(u), h\rangle = \mu\langle \mathrm{D}\mathcal{P}(u), h\rangle \quad \text{for all } h \in X. \tag{5.13}$$

Theorem 5.9. *Assume that assumptions 5.1, 5.2, 5.3 and 5.7 hold, and that β is defined by (5.3). Then β^{-1} is the least value of μ such that (5.13) has a non-zero solution in X.*

Proof. Let $\hat{v}$ be a non-zero solution of (5.13). Put $u = h = \hat{v}$ in (5.13). Then, by Euler's rule, (5.10) and (5.4) yield that

$$0 = p[\mathcal{F}(\hat{v}) - \mu\mathcal{P}(\hat{v})] \geqslant p(1 - \mu\beta)\mathcal{F}(\hat{v}).$$

From assumption 5.2, $\mathcal{F}(\hat{v}) > 0$, so $\mu \geqslant \beta^{-1}$. Moreover, v is a non-zero solution of (5.13) with $\mu = \beta^{-1}$, so the result follows. □

6. *p*-versions of Friedrichs' inequality

K. O. Friedrichs is credited with H^1-coercivity inequalities of the form

$$\int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^2\,\mathrm{d}x + \int_{\partial\Omega} |u|^2\,\mathrm{d}\sigma \geqslant C_0 \int_\Omega |u|^2\,\mathrm{d}x. \tag{6.1}$$

See [14, theorem 1.9] and [12, § 5.3] for proofs of this result and references to the earlier literature.

Here we describe the p-analogue of (6.1) with $1 < p < \infty$ and allow weights in the last two terms. That is, we prove an inequality of the form (3.1) with $\mathcal{B}_1$ given by (3.5) in place of $\mathcal{B}$ and $\mathcal{P}$ defined by (3.2). The intent is to identify the optimal constant C_{F} and corresponding optimal functions for the inequality

$$\mathcal{F}_1(u) := \int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^p\,\mathrm{d}x + \int_{\partial\Omega} b|u|^p\,\mathrm{d}\sigma \geqslant C_{\mathrm{F}} \int_\Omega \rho|u|^p\,\mathrm{d}x \tag{6.2}$$

for all $u \in W^{1,p}(\Omega)$. Here ρ and b obey assumptions 3.1 and 3.4, respectively. The value of C_F depends on the region Ω, the value of p and the functions b and ρ. It will be called *Friedrichs' constant* and some variational characterizations of it will be developed.

We need some basic properties of the functional $\mathcal{B}_1$.

PROPOSITION 6.1. *Assume that assumptions 2.2 and 3.4 hold and that $\mathcal{B}_1$ is defined by (3.5). Then $\mathcal{B}_1$ is convex, positive and continuous on $W^{1,p}(\Omega)$. Assumption 3.3 also holds and $\mathcal{B}_1$ is G-differentiable with*

$$\langle \mathrm{D}\mathcal{B}_1(u), h\rangle = p\int_{\partial\Omega} b|u|^{p-2}uh\,\mathrm{d}\sigma \quad \text{for all } u, h \in W^{1,p}(\Omega). \tag{6.3}$$

Proof. The integrand $b(x,s) := b(x)|s|^p$ is positive on $\Omega\times\mathbb{R}$, and $b(x,\cdot)$ is convex on $\mathbb{R}$, so $\mathcal{B}_1$ is positive and convex.

If $\{u_m : m \geqslant 1\}$ converges to $\hat{u}$ in $W^{1,p}(\Omega)$, then assumption 2.2 implies that $\{\Gamma u_m\}$ converges strongly to $\Gamma\hat{u}$ in $L^p(\partial\Omega, \mathrm{d}\sigma)$. Thus, there is a subsequence $\{\Gamma u_{m_j} : j \geqslant 1\}$ which converges σ-a.e. to $\Gamma\hat{u}$ on $\partial\Omega$, so

$$b(x, u_{m_j}(x)) \to b(x, \hat{u}(x)) \quad \sigma\text{-a.e. on } \partial\Omega \text{ as } j\to\infty.$$

When assumption 3.4 holds, the Lebesgue-dominated convergence theorem shows that $\mathcal{B}_1$ is continuous, as claimed. Since $\mathcal{B}_1$ is continuous and convex, assumption 3.3 holds.

The proof that $\mathcal{B}_1$ is G-differentiable and (6.3) holds parallels that of the corresponding part of the proof of proposition 3.2. □

A consequence of this result is the observation that $\mathcal{B}_1(u)^{1/p}$ is a continuous semi-norm on $W^{1,p}(\Omega)$.

Consider the variational problem of minimizing $\mathcal{F}_1$ on the set

$$S_1 := \{u \in W^{1,p}(\Omega) : \mathcal{P}(u) = 1\}. \tag{6.4}$$

When ρ satisfies assumption 3.1, proposition 3.2 shows that S_1 will be a weakly closed unbounded subset of $W^{1,p}(\Omega)$. Define

$$\alpha := \inf_{u\in S_1} \mathcal{F}_1(u). \tag{6.5}$$

Let $W^{1,p}_m(\Omega)$ be the subspace of $W^{1,p}(\Omega)$ of all functions with mean value zero. Then each $u \in W^{1,p}(\Omega)$ has a unique decomposition of the form

$$u = \bar{u} + v \quad \text{with } v \in W^{1,p}_m(\Omega). \tag{6.6}$$

The triangle inequality for norms yields

$$\|u\|_{1,p} \leqslant |\bar{u}||\Omega|^{1/p} + \|v\|_{1,p}. \tag{6.7}$$

THEOREM 6.2. *Assume that assumptions 2.1, 2.2, 3.1 and 3.4 hold and that $1 < p < \infty$. There is an optimal constant $C_F > 0$ which has corresponding optimal functions for (6.2).*

Proof. To prove this we shall show that the infimum of $\mathcal{F}_1$ on S_1 is attained, and then (6.2) holds with the optimal $C_{\mathrm{F}} = \alpha > 0$. Since $\mathcal{F}_1(u) \geqslant 0$ we have $\alpha \geqslant 0$. Let $\{u_m : m \geqslant 1\}$ be a decreasing, minimizing sequence for $\mathcal{F}_1$ on S_1. Then

$$\|\nabla u_m\|_p^p \leqslant \mathcal{F}_1(u_1) \quad \text{for all } m \geqslant 1.$$

For each m, write $u_m = \bar{u}_m + v_m$ as in (6.6). Then

$$\|\nabla u_m\|_p = \|\nabla v_m\|_p,$$

so $\|v_m\|_{1,p}$ is bounded from the Poincaré inequality (2.5).

When $\mathcal{B}_1$ is defined by (3.5) and assumption 3.4 holds, since $\mathcal{B}_1^{1/p}$ is a semi-norm, we obtain

$$\begin{aligned}|\bar{u}_m| {b_0}^{1/p} &= \mathcal{B}_1(\bar{u}_m)^{1/p} \\ &\leqslant \mathcal{B}_1(u_m)^{1/p} + \mathcal{B}_1(v_m)^{1/p} \\ &\leqslant \mathcal{F}_1(u_1)^{1/p} + C_2\|v_m\|_{1,p}.\end{aligned}$$

In the last inequality, the fact that Γ is continuous from assumption 2.2 was used. This yields the result that $\{\bar{u}_m : m \geqslant 1\}$ is bounded. Then (6.7) implies that $\{u_m : m \geqslant 1\}$ is bounded in $W^{1,p}(\Omega)$. This sequence has a weak limit $\hat{u}$, as $W^{1,p}(\Omega)$ is reflexive when $1 < p < \infty$, and $\hat{u}$ is in S_1, as S_1 is weakly closed. The functional $\mathcal{F}_1$ is weakly l. s. c. on $W^{1,p}(\Omega)$ from proposition 6.1. Thus, $\mathcal{F}_1(\hat{u}) = \alpha$, as the sequence $\{u_m : m \geqslant 1\}$ is minimizing. Hence, the infimum in (6.5) is attained.

If $\alpha = 0$, then $\|\nabla \hat{u}\|_p = 0$. Thus, $\hat{u}$ is constant on Ω, since Ω is connected. From (3.7), this constant must be 0, which is impossible if $\hat{u} \in S_1$. Hence, $\alpha > 0$. By homogeneity, (6.2) follows, with $C_{\mathrm{F}} := \alpha$, and any multiple of $\hat{u}$ is an optimal function for (6.2). □

Theorem 6.2 implies that the expression

$$\|u\|_{b,p} := \mathcal{F}_1(u)^{1/p} \tag{6.8}$$

is a norm on $W^{1,p}(\Omega)$. Moreover, we have the following corollary.

COROLLARY 6.3. *Assume that assumptions 2.1, 2.2 and 3.4 hold and* $1 < p < \infty$. *Then the* (b, p) *norm defined by (6.8) is an equivalent norm to the standard norm on* $W^{1,p}(\Omega)$.

Proof. When assumption 3.4 holds and $u \in W^{1,p}(\Omega)$,

$$\mathcal{F}_1(u) \leqslant \|\nabla u\|_p^p + \|b\|_{\infty,\partial\Omega}\|\Gamma u\|_{p,\partial\Omega}^p.$$

Since Γ is continuous, from assumption 2.2, this yields

$$\|u\|_{b,p} \leqslant (1 + C)\|u\|_{1,p} \quad \text{for some positive } C.$$

Take $\rho \equiv 1$ in theorem 6.2. Then, since $b \geqslant 0$ and (6.2) holds,

$$2\mathcal{F}_1(u) \geqslant \|\nabla u\|_p^p + C_{\mathrm{F}}\|u\|_p^p \geqslant \min(1, C_{\mathrm{F}})\|u\|_{1,p}^p.$$

These two inequalities imply the equivalence of these norms on $W^{1,p}(\Omega)$. □

To describe the scaling of these inequalities, let Ω, Ω_L be as in § 4. Define ρ_L, b_L as before. When (6.2) holds on Ω, we multiply through by L^{n-p} and use the formulae of § 4 to find that

$$\int_{\Omega_L} \sum_{j=1}^{n} |\mathrm{D}_j u|^p \,\mathrm{d}y + L^{1-p} \int_{\partial\Omega_L} b_L |u|^p \,\mathrm{d}\sigma \geqslant C_{\mathrm{F}} L^{-p} \int_{\Omega_L} \rho_L |u|^p \,\mathrm{d}y \tag{6.9}$$

for all $L > 0$, $u \in W^{1,p}(\Omega_L)$. This is the general scale-dependent version of (6.2) and equality will hold here for some functions in $W^{1,p}(\Omega_L)$.

7. Friedrichs' constant as an eigenvalue

In this section we show that the optimal constant C_{F} in (6.2), and also the optimal functions, can be described as the least eigenvalue, and associated eigenfunctions, respectively, of a p-Laplacian eigenproblem on Ω.

This will be done by using the unconstrained variational formulation introduced in § 5. Take $X = W^{1,p}(\Omega)$, $\mathcal{F}_1$ in place of $\mathcal{F}$, and assume that $\mathcal{P}$ is defined by (3.2). When assumptions 2.1 and 3.1 hold for ρ, proposition 3.2 shows that $\mathcal{P}$ satisfies assumptions 5.1, 5.3 and 5.7. Similarly, when assumptions 2.1 and 2.2 hold, $\mathcal{F}_1$ satisfies assumptions 5.1 and 5.3. The characterization of C_{F} in (6.2) may be compared with (5.5) and theorem 5.9 to show that C_{F} is the least eigenvalue of an eigenproblem of the form (5.13).

The G-differentiability of $\mathcal{F}_1$ is directly verified. Equation (5.13) then becomes, in this case, the problem of finding (μ, u) in $\mathbb{R} \times W^{1,p}(\Omega)$ with $u \neq 0$ that solve

$$\int_{\Omega} \left[\sum_{j=1}^{n} |\mathrm{D}_j u|^{p-2} \mathrm{D}_j u \mathrm{D}_j h - \mu\rho |u|^{p-2} u h \right] \mathrm{d}x + \int_{\partial\Omega} b |u|^{p-2} u h \,\mathrm{d}\sigma = 0 \tag{7.1}$$

for all $h \in W^{1,p}(\Omega)$.

When $p = 2$, this is the weak form of a linear eigenvalue problem for the Laplacian on Ω: namely, to find non-trivial solutions (μ, u) of

$$\int_{\Omega} \left[\sum_{j=1}^{n} \mathrm{D}_j u \mathrm{D}_j h - \mu\rho u h \right] \mathrm{d}x + \int_{\partial\Omega} b u h \,\mathrm{d}\sigma = 0 \quad \text{for all } h \in H^1(\Omega). \tag{7.2}$$

This is the weak form of the eigenproblem

$$-\Delta u = \mu\rho u \quad \text{in } \Omega, \tag{7.3}$$

$$(\nabla u) \cdot \nu + b u = 0 \quad \text{on } \partial\Omega. \tag{7.4}$$

This boundary condition is of Robin type, when b is strictly positive and of Neumann type on any subset where $b = 0$. In § 12, we shall determine the value of C_{F} for rectangles in the plane by direct solution of this problem.

For general $p > 1$, (7.1) is the weak form of an eigenvalue problem for the p-Laplacian on Ω. Namely one seeks non-zero solutions of the problem:

$$-\Delta_p u = -\sum_{j=1}^{n} \mathrm{D}_j (|\mathrm{D}_j u|^{p-2} \mathrm{D}_j u) = \mu\rho |u|^{p-2} u \quad \text{in } \Omega, \tag{7.5}$$

$$\sum_{j=1}^{n} |\mathrm{D}_j u|^{p-2}(\mathrm{D}_j u)\nu_j + b|u|^{p-2}u = 0 \quad \text{on } \partial\Omega. \tag{7.6}$$

Friedrichs' constant may now be characterized as follows.

THEOREM 7.1. *Assume that assumptions 2.1, 2.2, 3.1 and 3.4 hold with $1 < p < \infty$. Then the optimal constant $C_{\mathrm{F}} > 0$ in (6.2) is the least eigenvalue μ_1 of (7.1). Equality holds in (6.2) if and only if u is an eigenfunction of (7.1) corresponding to the least eigenvalue μ_1.*

Proof. The fact that C_{F} is the least eigenvalue of (7.1) follows from theorem 5.9. If u_1 is a corresponding eigenfunction of (7.1), then we put $u = h = u_1$ in (7.1) to see that equality holds in (6.2).

Conversely, if $\tilde{u}$ is a non-zero function for which the equality holds in (6.2), then it is a multiple of a function which maximizes $\mathcal{P}$ on the unit sphere in $W^{1,p}(\Omega)$ with the (b,p)-norm. Hence, it is a multiple of a minimizer of the associated $\mathcal{J}$ on $W^{1,p}(\Omega)$. Corollary 5.8 now yields the result when one observes that any multiple of a solution of (7.1) is again a solution of the equation. □

8. Coercivity inequalities with boundary integrals

The analysis of finite-element methods for linear elliptic operators uses coercivity inequalities of the form

$$\int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^2 \,\mathrm{d}x + \left(\int_{\partial\Omega} bu \,\mathrm{d}\sigma\right)^2 \geqslant C_1 \int_\Omega |u|^2 \,\mathrm{d}x \tag{8.1}$$

for all $u \in H^1(\Omega)$ (see [2,7] for discussions of this). This inequality differs from (6.2) and the author is not aware of a published proof of this result which includes an estimate of the constant C_1. A slightly different inequality is given in [3, example 6.3.16].

A more general form of this is that, for $1 < p < \infty$, there exists a constant $C_B > 0$ such that

$$\mathcal{F}_2(u) := \int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^p \,\mathrm{d}x + \left|\int_{\partial\Omega} bu \,\mathrm{d}\sigma\right|^p \geqslant C_B \int_\Omega \rho|u|^p \,\mathrm{d}x \tag{8.2}$$

for all $u \in W^{1,p}(\Omega)$. Here b and ρ satisfy assumptions 3.5 and 3.1. The value of C_B will depend on Ω, p, b and ρ. Here we shall characterize the optimal constant C_B, and the optimizing functions in this inequality, via some variational principles. In particular we will show that C_B is the least eigenvalue of a eigenproblem with integro-differential boundary conditions.

The analysis of this inequality parallels that for Friedrichs' inequality. We first consider the variational problem of minimizing $\mathcal{F}_2$ on the set S_1 defined by (6.4). Write

$$\alpha_2 := \inf_{u \in S_1} \mathcal{F}_2(u). \tag{8.3}$$

The proof of the following theorem shows that the minimizers of this problem exist and are functions for which equality holds in (8.2) with the value $\alpha_2 = C_B$.

THEOREM 8.1. *Assume that assumptions 2.1, 2.2, 3.1 and 3.5 hold and that* $1 < p < \infty$. *There is an optimal constant* $C_B > 0$ *which has corresponding optimal functions for (8.2).*

Proof. First, note that $\alpha_2 > 0$. Let $\{u_m : m \geqslant 1\}$ be a decreasing, minimizing sequence for $\mathcal{F}_2$ on S_1. Then

$$\|\nabla u_m\|_p^p \leqslant \mathcal{F}_2(u_1) \quad \text{for all } m \geqslant 1.$$

For each m, write $u_m = \bar{u}_m + v_m$ as in (6.6). Then $\|\nabla u_m\|_p = \|\nabla v_m\|_p$, so $\|v_m\|_{1,p}$ is bounded from the Poincaré inequality (2.5).

Define the linear functional b on $W^{1,p}(\Omega)$ by

$$b(u) := \int_{\partial\Omega} b\Gamma u \, \mathrm{d}\sigma. \tag{8.4}$$

Then $b(u_m) = b_0\bar{u}_m + b(v_m)$. When assumptions 2.2 and 3.5 hold, b will be continuous, so

$$b_0|\bar{u}_m| \leqslant |b(u_m)| + C\|v_m\|_{1,p}.$$

Now $|b(u_m)|$ is bounded, as we have a descent sequence for $\mathcal{F}_2$; thus, $|\bar{u}_m|$ is uniformly bounded. Hence, the sequence $\{u_m : m \geqslant 1\}$ is bounded in $W^{1,p}(\Omega)$ from (6.7). The concluding arguments in the proof of theorem 6.2 now apply, and complete the proof of this theorem. □

This result implies that the expression

$$\|u\|_{\partial,p} := \mathcal{F}_2(u)^{1/p} \tag{8.5}$$

defines a norm on $W^{1,p}(\Omega)$. This may be strengthened to the following corollary.

COROLLARY 8.2. *Assume that assumptions 2.1, 2.2 and 3.5 hold and that* $1 < p < \infty$. *Then the* (∂, p) *norm defined by (8.5) is an equivalent norm to the standard norm on* $W^{1,p}(\Omega)$.

Proof. When assumption 3.5 holds and $u \in W^{1,p}(\Omega)$, we obtain

$$\mathcal{F}_2(u) \leqslant \|\nabla u\|_p^p + \|b\|_{p^*,\partial\Omega}^p \|\Gamma u\|_{p,\partial\Omega}^p.$$

Since Γ is continuous, from assumption 2.2 this yields

$$\|u\|_{\partial,p} \leqslant (1 + C)\|u\|_{1,p} \quad \text{for some positive } C.$$

Take $\rho \equiv 1$ in theorem 8.1. Then, since $b \geqslant 0$ and (8.2) holds,

$$2\mathcal{F}_2(u) \geqslant \|\nabla u\|_p^p + C_B\|u\|_p^p \geqslant \min(1, C_B)\|u\|_{1,p}^p.$$

These two inequalities imply the equivalence of these norms on $W^{1,p}(\Omega)$. □

When Ω_L is defined as in § 4, the scaled version of (8.2) is obtained by multiplying by L^{n-p} and using the formulae in § 4. This yields

$$\int_{\Omega_L} \sum_{j=1}^{n} |\mathrm{D}_j u|^p \, \mathrm{d}y + L^{n(1-p)} \left| \int_{\partial\Omega_L} b_L u \, \mathrm{d}\sigma \right|^p \geqslant C_B L^{-p} \int_{\Omega_L} \rho_L |u|^p \, \mathrm{d}y \tag{8.6}$$

for all $L > 0$, $u \in W^{1,p}(\Omega_L)$. Moreover, there are functions in $W^{1,p}(\Omega_L)$ for which equality holds here.

9. The optimal constant C_B

In the last section, the constant C_B was identified as the value of a variational problem. This problem has the form described in § 5, with X replaced by $W^{1,p}(\Omega)$, $\mathcal{F}$ by $\mathcal{F}_2$, and β by α_2.

Just as in § 5, an unconstrained variational principle for this problem may be introduced, and we find that the minimizers of our problem are given by the eigenfunctions of the analogue of equation (5.13) corresponding to the least eigenvalue. It is straightforward to complete the verification that $\mathcal{F}_2$ is G-differentiable on $W^{1,p}(\Omega)$. In this case, equation (5.13) becomes the problem of finding the non-zero solutions (μ, u) in $\mathbb{R} \times W^{1,p}(\Omega)$ of

$$\int_\Omega \left[\sum_{j=1}^n |\mathrm{D}_j u|^{p-2}\, \mathrm{D}_j u\, \mathrm{D}_j h - \mu\rho|u|^{p-2}uh\right] \mathrm{d}x + |b(u)|^{p-2}b(u)b(h) = 0 \qquad (9.1)$$

for all $h \in W^{1,p}(\Omega)$. When $p = 2$, this reduces to

$$\int_\Omega \left[\sum_{j=1}^n \mathrm{D}_j u\, \mathrm{D}_j h - \mu\rho uh\right] \mathrm{d}x + b(u)b(h) = 0 \quad \text{for all } h \in H^1(\Omega). \qquad (9.2)$$

This is the weak form of the eigenproblem

$$-\Delta u = \mu\rho u \quad \text{in } \Omega, \qquad (9.3)$$

$$(\nabla u)\cdot\nu + b(u)b = 0 \quad \text{on } \partial\Omega. \qquad (9.4)$$

This boundary condition is an integro-differential equation. Nevertheless, standard elliptic spectral theory applies to this eigenproblem with only minimal changes.

For general $p > 1$, (9.1) is the weak form of an eigenvalue problem for the p-Laplacian on Ω. Namely, one seeks non-zero solutions of (7.5) subject to the integro-differential boundary condition

$$\sum_{j=1}^n |\mathrm{D}_j u|^{p-2}(\mathrm{D}_j u)\nu_j + b|b(u)|^{p-2}b(u) = 0 \quad \text{on } \partial\Omega. \qquad (9.5)$$

The optimal constant C_B may now be characterized in a similar way to that of the Friedrich constant in § 7. The proof of the following is essentially the same as that of theorem 7.1.

THEOREM 9.1. *Assume that assumptions 2.1, 2.2, 3.1 and 3.5 hold with* $1 < p < \infty$. *Then the optimal constant* $C_B > 0$ *in (8.2) is the least eigenvalue* μ_1 *of (9.1). Equality holds in (8.2) if and only if* u *is an eigenfunction of (9.1) corresponding to the least eigenvalue* μ_1.

10. Generalized Poincaré inequalities

The name *Poincaré inequality* is attached to a number of different results. In [11, § 7.9] and [8, ch. V, § 3], inequalities of the form

$$\int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^p \, \mathrm{d}x \geqslant C_p \int_\Omega |u - \bar{u}|^p \, \mathrm{d}x \tag{10.1}$$

are described with specific simple formulae for (lower bounds on) C_p.

Here we shall consider the question of finding the optimal constant C_P in the inequality

$$\mathcal{F}_3(u) := \int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^p \, \mathrm{d}x + \left| \int_\Omega cu \, \mathrm{d}x \right|^p \geqslant C_P \int_\Omega \rho |u|^p \, \mathrm{d}x \tag{10.2}$$

for all $u \in W^{1,p}(\Omega)$. Here c and ρ satisfy assumptions 3.6 and 3.1. When $p = 2$ and the functions c, ρ are constants, this is one of the forms given in [14, ch. 1]. The value of C_P will depend on Ω, p, c and ρ.

Here we shall characterize the optimal constant C_P and the optimizing functions in this inequality via some variational principles. In particular, we will show that C_P is the least eigenvalue of a Neumann eigenproblem for an integro-differential operator on Ω.

Just as before, consider the problem of minimizing the functional $\mathcal{F}_3$ on the set S_1 defined by (6.4). Write

$$\alpha_3 := \inf_{u \in S_1} \mathcal{F}_3(u). \tag{10.3}$$

The proof of the following theorem shows that the minimizers of this problem exist, there are functions for which equality holds in (10.2) and the value $\alpha_3 = C_P$.

THEOREM 10.1. *Assume that assumptions 2.1, 3.1 and 3.6 hold and that $1 < p < \infty$. There is an optimal constant $C_P > 0$ which has corresponding optimal functions for (10.2).*

Proof. First, note that $\alpha_3 > 0$. Let $\{u_m : m \geqslant 1\}$ be a decreasing, minimizing sequence for $\mathcal{F}_3$ on S_1. Then

$$\|\nabla u_m\|_p^p \leqslant \mathcal{F}_3(u_1) \quad \text{for all } m \geqslant 1.$$

For each m, write $u_m = \bar{u}_m + v_m$ as in (6.6). Then $\|\nabla u_m\|_p = \|\nabla v_m\|_p$, so $\|v_m\|_{1,p}$ is bounded from the Poincaré inequality (2.5).

Define the linear functional c on $W^{1,p}(\Omega)$ by

$$c(u) := \int_\Omega cu \, \mathrm{d}x. \tag{10.4}$$

Then $c(u_m) = \bar{c}|\Omega|\bar{u}_m + c(v_m)$. When assumption 3.6 holds, c will be continuous, so the fact that $|c(u_m)|$ is uniformly bounded implies that $|\bar{u}_m|$ is uniformly bounded. Hence, the sequence $\{u_m : m \geqslant 1\}$ is bounded in $W^{1,p}(\Omega)$ from (6.7). The concluding arguments in the proof of theorem 6.2 apply again here, to yield the proof of this theorem. □

This result implies that the expression

$$\|u\|_{c,p} := \mathcal{F}_3(u)^{1/p} \tag{10.5}$$

defines a norm on $W^{1,p}(\Omega)$. This may be strengthened to the following corollary.

COROLLARY 10.2. *Assume that assumptions 2.1, 3.1 and 3.6 hold and that* $1 < p < \infty$. *Then the* (c,p) *norm defined by (10.5) is an equivalent norm to the standard norm on* $W^{1,p}(\Omega)$.

Proof. When assumption 3.6 holds and $u \in W^{1,p}(\Omega)$,

$$\mathcal{F}_3(u) \leqslant \|\nabla u\|_p^p + \|c\|_{p^*}^p \|u\|_p^p \leqslant C\|u\|_{1,p}^p$$

with $C > 0$. Take $\rho \equiv 1$ in theorem 10.1. Then, since $c \geqslant 0$ and (10.2) holds,

$$2\mathcal{F}_2(u) \geqslant \|\nabla u\|_p^p + C_P\|u\|_p^p \geqslant \min(1, C_P)\|u\|_{1,p}^p.$$

These two inequalities imply the equivalence of these norms on $W^{1,p}(\Omega)$. □

When Ω_L is defined as in § 4, the scaled version of (10.2) is obtained by using the formulae in § 4 and multiplying by L^{n-p}. This yields

$$\int_{\Omega_L} \sum_{j=1}^{n} |\mathrm{D}_j u|^p \,\mathrm{d}y + L^{n-p(n+1)} \left| \int_{\Omega_L} c_L u \,\mathrm{d}y \right|^p \geqslant C_P L^{-p} \int_{\Omega_L} \rho_L |u|^p \,\mathrm{d}y. \tag{10.6}$$

for all $L > 0$, $u \in W^{1,p}(\Omega_L)$. Equality holds here for some functions in $W^{1,p}(\Omega_L)$.

For this case the analogue of (5.13) is to find non-zero solutions (μ, u) in $\mathbb{R} \times W^{1,p}(\Omega)$ of

$$\int_{\Omega} \left[\sum_{j=1}^{n} |\mathrm{D}_j u|^{p-2} \,\mathrm{D}_j u \,\mathrm{D}_j\, h - \mu\rho|u|^{p-2}uh \right] \mathrm{d}x + |c(u)|^{p-2}c(u)c(h) = 0 \tag{10.7}$$

for all $h \in W^{1,p}(\Omega)$. When $p = 2$, this reduces to

$$\int_{\Omega} \left[\sum_{j=1}^{n} \mathrm{D}_j u \,\mathrm{D}_j\, h - \mu\rho uh \right] \mathrm{d}x + c(u)c(h) = 0 \quad \text{for all } h \in H^1(\Omega). \tag{10.8}$$

This is the weak form of the eigenproblem

$$-\Delta u + c(u)c = \mu\rho u \quad \text{in } \Omega, \tag{10.9}$$

$$(\nabla u) \cdot \nu = 0 \qquad \text{on } \partial\Omega. \tag{10.10}$$

This is an eigenvalue problem for an integro-differential equation on Ω. It is worth noting that when the function c here is itself an eigenfunction of the Neumann problem for equation (7.3), then the eigenfunctions of this problem are precisely the eigenfunctions of the Neumann problem for (7.3), and only one eigenvalue is different.

For general $p > 1$, (10.7) is the weak form of an eigenvalue problem for the p-Laplacian on Ω. Namely, one seeks non-zero solutions of the system

$$-\sum_{j=1}^{n} \mathrm{D}_j(|\mathrm{D}_j u|^{p-2}\,\mathrm{D}_j u) + |c(u)|^{p-2}c(u)c = \mu\rho|u|^{p-2}u \quad \text{in } \Omega, \tag{10.11}$$

$$\sum_{j=1}^{n} |\mathrm{D}_j u|^{p-2}(\mathrm{D}_j u)\nu_j = 0 \quad \text{on } \partial\Omega. \tag{10.12}$$

11. Optimal inequalities for boxes

Brenner [7] describes discrete analogues of the boundary inequality (11.2) and the Poincaré inequality (11.3) below for two-dimensional polygons and three-dimensional polyhedra. To illustrate the preceding analysis, and for comparison with the results in [7], we will find explicit formulae for the optimal constants when the region Ω is taken to be a rectangle. Take $p = 2$ and $\Omega := (0, \pi) \times (0, h)$ with $h > 0$ the height of the rectangle and let all the coefficient functions be identically 1.

The three different inequalities may be written as

$$\int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^2 \,\mathrm{d}x + \int_{\partial\Omega} |u|^2 \,\mathrm{d}s \geqslant C_{\mathrm{F}}(h) \int_\Omega |u|^2 \,\mathrm{d}x, \tag{11.1}$$

$$\int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^2 \,\mathrm{d}x + \left| \int_{\partial\Omega} u \,\mathrm{d}s \right|^2 \geqslant C_B(h) \int_\Omega |u|^2 \,\mathrm{d}x, \tag{11.2}$$

$$\int_\Omega \sum_{j=1}^{n} |\mathrm{D}_j u|^2 \,\mathrm{d}x + \left| \int_{\Omega} u \,\mathrm{d}x \right|^2 \geqslant C_P(h) \int_\Omega |u|^2 \,\mathrm{d}x. \tag{11.3}$$

Here ds replaces dσ, as it represents arc length on the boundary $\partial\Omega$.

From (7.3) and (7.4), the value of $C_{\mathrm{F}}(h)$ in (11.1) is the least eigenvalue of the Robin–Laplace problem

$$-\Delta u = \mu u \quad \text{in } \Omega \quad \text{and} \quad (\nabla u)\cdot\nu + u = 0 \quad \text{on } \partial\Omega. \tag{11.4}$$

Similarly, the value of $C_B(h)$ in (11.2) is the least eigenvalue of the Laplacian eigenproblem

$$-\Delta u = \mu u \quad \text{in } \Omega \quad \text{and} \quad (\nabla u)\cdot\nu + \int_{\partial\Omega} u \,\mathrm{d}s = 0 \quad \text{on } \partial\Omega. \tag{11.5}$$

Finally, the value of $C_P(h)$ in (11.3) is the least eigenvalue of the modified Laplacian eigenproblem

$$-\Delta u + \left(\int_\Omega u \,\mathrm{d}x \right) = \mu u \quad \text{in } \Omega \quad \text{and} \quad (\nabla u)\cdot\nu = 0 \quad \text{on } \partial\Omega. \tag{11.6}$$

The eigenfunctions of (11.6) are precisely the eigenfunctions of the Neumann Laplacian on Ω. The first non-zero eigenvalue of the Neumann Laplacian is $\lambda_1^{(\mathrm{N})} = \min(1, \pi/h^2)$. Thus, the optimal constant in (11.3) for this rectangle is found to be

$$C_P(h) = \min(h\pi, 1, \pi/h^2). \tag{11.7}$$

A careful analysis of a family of constrained variational principles for $C_B(h)$ leads to the result that

$$C_B(h) = \min(1, \pi/h^2). \tag{11.8}$$

The first eigenvalue, and the corresponding eigenfunction of (11.4) may also be found explicitly. The least eigenvalue is

$$C_{\mathrm{F}}(h) = k_0 + (k_1(h))^2, \tag{11.9}$$

where $k_0 = 0.407\,42$ and $0.5(\pi/h) < k_1(h) < (\pi/h)$. In fact, k_0 and k_1 are the smallest positive solutions, respectively, of

$$\tan k\pi = \frac{-2k}{1-k^2} \quad \text{and} \quad \tan kh = \frac{-2k}{1-k^2}.$$

References

1 R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, 2nd edn (Academic, 2003).

2 D. N. Arnold, F. Brezzi, B. Cockburn and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Analysis* **39** (2002), 1749–1779.

3 K. Atkinson and W. Han. *Theoretical numerical analysis* (Springer, 2001).

4 G. Auchmuty. Variational principles for self-adjoint elliptic eigenproblems. In *Nonsmooth/nonconvex mechanics: modeling, analysis, and numerical methods* (ed. D. Y. Gao, R. W. Ogden and G. E. Stavroulakis), pp. 15–42 (Kluwer, 2001).

5 G. Auchmuty. The main inequality of vector analysis. *Math. Models Meth. Appl. Sci.* **14** (2004), 1–25.

6 P. Blanchard and E. Brüning. *Variational methods in mathematical physics* (Springer, 1992).

7 S. C. Brenner. Poincaré–Friedrichs inequalities for piecewise H^1 functions. *SIAM J. Numer. Analysis* **41** (2003), 306–324.

8 D. E. Edmunds and W. D. Evans. *Spectral theory and differential operators* (Oxford University Press, 1987).

9 L. C. Evans. *Partial differential equations* (Providence, RI: American Mathematical Society, 1999).

10 L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions* (Boca Raton, FL: CRC Press, 1992).

11 D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order* (Springer, 1977).

12 G. Hellwig. *Differential operators of mathematical physics* (Addison-Wesley, 1967).

13 J. Jost and X. Li-Jost. *Calculus of variations* (Cambridge University Press, 1998).

14 J. Necas. *Les méthodes directes en théorie des équations elliptiques* (Paris: Masson, 1967).

15 E. Zeidler. *Nonlinear functional analysis and its applications. III. Variational methods and optimization* (Springer, 1985).

(*Issued 14 October 2005*)