Variational methods in the presence of symmetry

Jonathan M. Borwein*and Qiji J. Zhu†

May 15, 2013

Abstract

The purpose of this paper is to survey and to provide a unified framework to connect a diverse group of results, currently scattered in the literature, that can be usefully viewed as consequences of applying variational methods to problems involving symmetry. Here, *variational methods* refer to mathematical treatment by way of constructing an appropriate *action function* whose critical points—or saddle points—correspond to or contain the desired solutions.

1 Introduction

The purpose of this paper is to survey and to provide a unified framework to connect a diverse group of results, currently scattered in the literature, that can be usefully viewed as consequences of applying variational methods to problems involving symmetry. Here, *variational methods* refer to mathematical treatment by way of constructing an appropriate *action function* whose critical points correspond to or contain the desired solutions.

Variational methods can be viewed as a mathematical form of the *least action principle* in physics. Variational methods have been a powerful tool in both pure and applied mathematics ever since the systematic development of calculus of variations commenced over 300 years ago. With the discovery of modern variational principles and the development of nonsmooth analysis, the range of application of such techniques has been extended dramatically (see recent monographs [14, 19, 33, 40, 43]).

*Centre for Computer-assisted Research Mathematics and its Applications (CARMA), School of Mathematical and Physical Sciences, University of Newcastle, Callaghan, NSW 2308, Australia. Email: jonathan.borwein@newcastle.edu.au, jborwein@gmail.com

†Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA. Email: zhu@wmich.edu

Symmetry is ubiquitous in the real world and its modelling, and it also presents frequently in variational problems. Often when the action function is symmetric the solution also has a certain symmetry. Research on such symmetric variational problems is currently scattered in the literature and sometimes is treated only implicitly. Since Felix Klein introduced his *Erlangen Program* [29], symmetry is, in general, treated as invariance with respect group actions. This, however, proves to be inadequate in dealing with many variational problems involving symmetry.

For example, a key result leading to Adrian Lewis' celebrated representation of subdifferentials of spectral functions [31] is:

Proposition 1. *Let f be a convex permutation invariant function of several real variables. Then $y \in \partial f(x) = \{y \colon \langle y, z-x\rangle \le f(z) - f(x)\}$ if and only if*

$$y^{\downarrow} \in \partial f(x^{\downarrow}) \text{ and } \langle y, x\rangle = \langle y^{\downarrow}, x^{\downarrow}\rangle.$$

Here $x^{\downarrow}$ denotes the decreasing rearrangement of x.

We note that in Proposition 1 (a) the permutation invariance of the action function f is not preserved by the solution – its subdifferential and (b) the decreasing rearrangement is not invertible and, therefore, cannot be described by a group action. These are not the only special features of variational problems involving symmetry. Some other 'anomalies' include (c) there may be no compatible topology for the symmetrization process and (d) at times only approximate symmetries can be constructed. A typical example involving (c) and (d) is the existence of Schwarz-symmetric solutions to Laplace-type partial differential equations as discussed in Section 3.4.2.

There have been some prior efforts to deal systematically with variational problems involving symmetry. An early result is the *Palais principle of symmetric criticality* [35]. When the action group consists of differentiable isometries, this principle states that finding symmetric critical points of a smooth action function requires only handling its restriction to the invariant submanifold corresponding to the group action. Limiting the application of the Palais principle is its strong smoothness requirement on the action function and the restrictive isometric property it needs for the group action.

In [30], Ledyaev and Zhu adopted the use of nonsmooth functions on smooth manifolds as a framework to deal with symmetry in variational problems. Spectral functions provide examples that can benefit from such a framework (see [14, Chapter 7]). The research in [30] established a set of tools for nonsmooth variational problems on smooth manifolds and illustrated that linearity of the underlying space is not essential in dealing with many variational problems. However, this approach still relies on characterizing symmetry as invariance with respect to group actions. As a result the special features of variational problems involving symmetry alluded to above were not adequately addressed.

Recently, van Schaftingen [42] and Squassina [45] proposed versions of symmetric minimax and variational principles—in different levels of generality—that are stimulated by *polarization* approximations of Steiner symmetry [15]. Tailoring variational principles to a specific type of symmetry helps in fitting such variational principles to the targeted problem. Yet this also limits the application of such principles to other problems. A more practicable approach to understanding variational problems involving symmetry is by providing a few simple overarching principles which lead to the development of systematic approaches for dealing with symmetries of diverse nature.

We follow this path in the current work. First we lay out a few such simple general variational principles. To make the general principles flexible enough and appropriate to the variational approach we define symmetry as an invariance with respect to a *semigroup* and study action functions that are 'sub-invariant' under the given semigroup action. The key to successfully applying these principles then relies on finding appropriate semigroups and related symmetrizations.

We shall illustrate this with a suite of examples of contexts in which variational methods have already been used successfully in solving problems. As it is neither possible nor useful for our purpose to be comprehensive, we instead will illustrate our approach using selected examples that effectively illustrate the four special features, (a) through (d), highlighted above.

The organization of the remainder of the paper is as follows. We layout the general principles in Section 2 and then turn to their applications in Section 3. Section 4 is devoted to discussion of saddle points in the presence of symmetry, and we make various conclusions in Section 5.

2 Variational Principles in Presence of Symmetry

2.1 Invariance and symmetry

As indicated, we frame symmetry as invariance or sub-invariance with respect to the action of a prescribed semigroup G with identity (monoid) acting on the ambient space. Throughout the paper, unless stated otherwise, we shall assume each semigroup does possesses an identity. Let (X, d) be a complete metric space with the semigroup action $G \times X \mapsto X$

$$(g, x) \mapsto gx. \tag{1}$$

In the sequel we will always assume that, for any $g \in G$, mapping $x \mapsto gx$ is continuous. For any subset S of X the G-orbit of S is defined by

$$G \cdot S := \{gs : g \in G, s \in S\}.$$

Consider an extended valued *lsc* (lsc) function $f : X \mapsto R \cup \{+\infty\}$. We denote the (lower) *level set* of f at $a \in (-\infty, \infty]$ by

$$[f \leq a] := \{x \in X : f(x) \leq a\}.$$

Definition 2 (Invariance)**.** *We say that an extended valued function $f : X \mapsto R \cup \{+\infty\}$ is* sub-invariant *with respect to the semigroup action (1) if, for any $g \in G$ and $x \in X$,*

$$f(gx) \leq f(x). \tag{2}$$

If the inequality is strict for all $gx \neq x$, we say f is strictly sub-invariant*.*

We say that an extended valued function $f : X \mapsto R \cup \{-\infty\}$ is (strictly) super-invariant *with respect to the semigroup action (1) if $-f$ is (strictly) sub-invariant. If $f : X \mapsto R$ is both sub- and super-invariant then we say f is* invariant*.*

Clearly, when f is invariant its level set $[f = a] := \{x \in X : f(x) = a\}$ is invariant under the semigroup action, i.e.,

$$G \cdot [f = a] \subset [f = a].$$

This relationship is convenient when working with equality constraints but non-essential since

$$[f = a] = [f \leq a] \cap [-f \leq -a].$$

Thus, we will focus on sub-invariance below. We also observe that if G is a group then the concept of either sub-invariance or super-invariance coincides with invariance.

Although the invariance of the level set under the semigroup action is a type of symmetry property, in many situations the following stronger form of symmetry is more significant.

Definition 3 (Symmetrization)**.** *Let G be a semigroup and let f be a sub-invariant function. A map $S : X \mapsto X$ is a (G, f)-*symmetrization *if, for any $x \in X$,* *(i) for any $g \in G$, $S(gx) = gS(x) = S(x)$;*

(ii) for any x, $S^2(x) = S(x)$;

(iii) for any x, $f(S(x)) \leq f(x)$.

*In particular, if $S(x) \in cl(G \cdot s)$ then (iii) holds for any lsc function f. In this case we will simply call S a G-*symmetrization.

This framework relies on ideas in [42, 45]. The main difference is in condition (iii) in the definition of symmetrization. In our definition, the symmetrization is linked to a particular action function f. Verifying (iii) is often the key to application and is nontrivial. Such verification is, however, usually much easier than trying to show $S(x) \in \text{cl}\ (G \cdot s)$ which ensures that (iii) holds for every lower semicontinuous function f; an unnecessarily strong condition which fails in many cases.

An important known concrete example that fits this framework is establishing the existence of symmetric solutions of certain Dirichlet type problems. In those problems, as we shall see, often condition (iii) can be verified using the *Palais-Smale* property [36] of the action function. The framework herein is also more flexible.

2.2 Symmetric extremal principle

Many symmetry properties that one can derive using variational arguments are based on the following simple result.

Proposition 4 (Symmetric extremal principle)**.** *Let $f : X \mapsto R \cup \{+\infty\}$ be a* sub-invariant *extended valued function with respect to the semigroup action (1). Then*

$$G \cdot [f \leq a] \subset [f \leq a].$$

In particular, letting $a = \inf f$, we have

$$G \cdot \operatorname{argmin}(f) \subset \operatorname{argmin}(f).$$

Moreover, if f is strictly sub-invariant, then, for any $x \in \operatorname{argmin}(f)$ and any $g \in G$, $gx = x$.

Proof. For any $x \in [f \leq a]$ and any $g \in G$, $f(gx) \leq f(x) \leq a$ implies $gx \in [f \leq a]$. The two set inclusions follows directly. The last conclusion follows directly from the definition of strict sub-invariance. This suffices to complete the proof. □

Similarly to Proposition 4 we have

Proposition 5 (Symmetric minimization)**.** *Let* $f : X \mapsto R \cup \{+\infty\}$ *be an* sub-invariant *extended valued function with respect to the semigroup action (1) and* S *is a* $(G,F)-$*symmetrization. Then*

$$S(\operatorname{argmin}(f)) \subset \operatorname{argmin}(f).$$

Proof. It is impossible to lie strictly below the minimum! $\square$

2.3 Symmetric variational principles

The symmetric extremal principle only applies to a problem that attains its infimum—that is, has a minimum. When the existence of a minimum is not guaranteed we need symmetric versions of a 'variational principle' [12, 22]. Versions of symmetric variational principles related to Steiner symmetrization have been discussed in [45]. These results require additional structure assumptions to ensure the compatibility with the Steiner symmetrization and its approximation by polarizations. We present simpler versions with more flexibility of application. The trade-off is a loss of precision when dealing with a specific Steiner symmetrization.

2.3.1 Symmetric Ekeland variational principle

Theorem 6 (Symmetric variational principle)**.** *Let* (X, d) *be a complete metric space with a semigroup action* G*. Let* $f : X \mapsto R\cup\{+\infty\}$ *be a* G*-sub-invariant lsc function which is bounded from below. Suppose that*

$$f(z) < \inf_X f + \varepsilon.$$

Then, for any $g \in G$ *and* $\lambda > 0$ *there exists* y *such that*

(i) $d(y, gz) \le \lambda$*;*

(ii) $f(y) + (\varepsilon/\lambda)\, d(y, gz) \le f(z)$*; and*

(iii) $f(x) + (\varepsilon/\lambda)\, d(x, y) > f(y)$ *for* $x \ne y$*.*

Proof. Since f is sub-invariant we have $f(gz) \le f(z) < \inf_X f + \varepsilon$. Applying Ekeland's variational principle [14, 22] to gz suffices to complete the proof. $\square$

Remark 7. We note that for the trivial semigroup action $gx = x, \forall g \in G$, any function on X is invariant. Thus, Theorem 6 is a true, if easy, generalization of the Ekeland variational principle. $\diamondsuit$

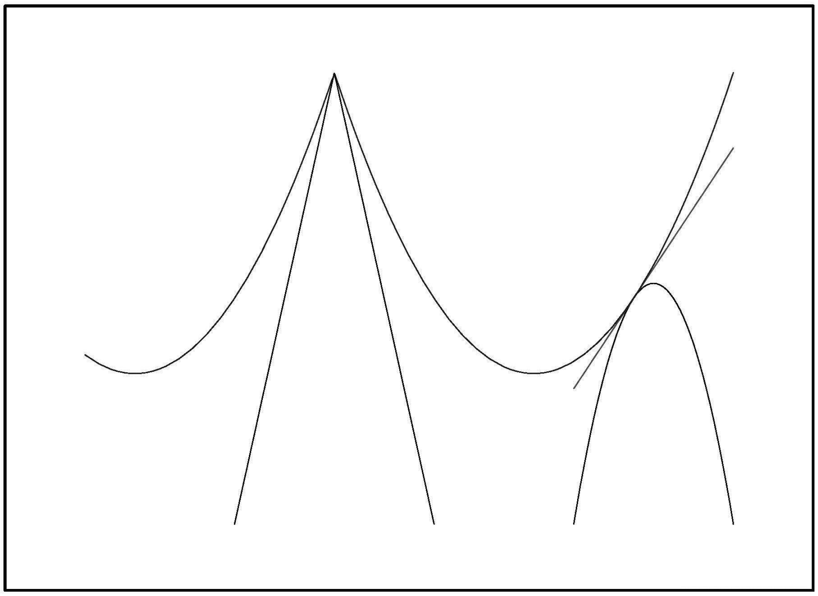

Figure 1: Ekeland (L) and Borwein-Preiss (R) variational principles.

2.3.2 Symmetric Borwein-Preiss variational principle

Similarly we have the following symmetric special case of the Borwein-Preiss variational principle. Recall that a Banach space is *super-reflexive* if it possess an equivalent uniformly convex norm or an equivalent uniformly smooth norm [12, 13]. Hilbert space and each abstract L^p space with $1 < p < \infty$ is super-reflexive.

Theorem 8 (Symmetric smooth variational principle)**.** *Let $(X, \|\cdot\|)$ be a super-reflexive Banach space with a semigroup action G. Let $f : X \mapsto R \cup \{+\infty\}$ be a G-sub-invariant lsc function which is bounded from below. Suppose that*

$$f(z) < \inf_X f + \varepsilon.$$

Then, for any $g \in G$ and $\lambda > 0, p \geq 1$ there exists y such that

(i) $f(gz) < \inf_X f(x) + \varepsilon$*,*

(ii) $\|y - gz\| \leq \lambda$*;*

(iii) $f(y) + (\varepsilon/\lambda^p)\, \|y - gz\|^p \leq f(gz)$*; and*

(iv) $f(x) + (\varepsilon/\lambda^p)\, \|x - y\|^p \geq f(y)$*.*

Proof. This is similar to the proof of the symmetric Ekeland variational principle except we use the super-reflexive Borwein-Preiss variational principle [12] as a starting point. $\square$

Figure 1 illustrates the difference in the two principles.

3 Applications

Applying the framework in Section 2 to concrete problems requires us to carefully determine the action function f, to find a suitable semigroup G with related symmetrization S, and to verify their compatibility. Although there are some general patterns, this process is largely problem-specific.

The main purpose of this section is to illustrate this process with several examples involving different forms of symmetry. We arrange our examples according to the special features alluded to in the introduction.

3.1 Symmetrization not compatible with topology

We start with a simple example:

Example 9 (Minimum of a symmetric function)**.** Consider the problem of minimizing a permutation invariant lsc convex function $f(x) : R^N \mapsto R \cup \{+\infty\}$. Assume that f is coercive (has compact lower level sets) and is bounded from below. Thus, f attains its minimum.

Let $G_1 := P(N)$ be the permutation group on $\{1, 2, \ldots, N\}$ and define $S_1(x) := \bar{x}\vec{1}$, where $\bar{x} := \sum_{n=1}^{N} x_n/N$ and the vector $\vec{1}$ has all components 1. We can directly verify that S_1 is a (G_1, f)-symmetrization. Thus, by the Symmetric Extremal Principle the minimizer of f must have the form $z = \bar{z}\vec{1}$.

This has reduces an N-dimensional optimization problem to to a one dimensional problem. We note that in this case $S_1(x) \notin \mathrm{cl} G_1 \cdot x$ unless $x = a\vec{1}$ for some $a \in R$. $\diamondsuit$

Since f need only be a lsc function, the method in Example 9 can easily be applied to minimization problems with invariant constraints, on using an indicator function (as defined below) to turn the constrained minimization problem into an unconstrained one. As an illustration, we use symmetry with respect to $P(N)$ to give a proof of the well known algebraic-geometric mean inequality.

Example 10 (Arithmetic-Geometric mean inequality)**.** Consider

$$\min\ f(x) := -\sum_{n=1}^{N} \log(x_n) + \iota_C(x),$$

where $C := \{x : \langle x, \vec{1}\rangle = K, x \geq 0\}$, and $\iota_C(x) = 0$ for $x \in C$ and is ∞ otherwise; represents the indicator function of the set C.

Then f satisfies all the conditions for the function in Example 9 and, therefore, has a minimum of the form $S_1(x) = \bar{x}\vec{1}$. The constraint $S_1(x) \in C$ forces $\bar{x} = K/N$ and the minimum is $-N\log(K/N)$. This now easily leads to the Arithmetic-Geometric mean inequality. $\diamondsuit$

Likewise, we prove the classical relative entropy inequality [9] using symmetry.

Example 11 (Relative entropy inequality)**.** Consider

$$\min\, f(p, q) := -\sum_{n=1}^{N} p_n \log(p_n/q_n) + \iota_C(p, q),$$

$C := \{(p, q) : \langle p, \vec{1}\rangle = \langle q, \vec{1}\rangle = 1, (p, q) \geq 0.\}$.

We shall show that $f(p, q) \geq f(\vec{1}, \vec{1}) = 0$. Now f is invariant with respect to the group action $G_2 : g(p, q) := (gp, gq), g \in P(N)$, and $S_2(p, q) := (\bar{p}\vec{1}, \bar{q}\vec{1})$ is a $(P(N), f)$-symmetrization. Again, f has a minimum $S_2(p, q) = (\bar{p}\vec{1}, \bar{q}\vec{1})$. The constraint $S_2(p, q) \in C$ forces $S_2(p, q) = (\vec{1}, \vec{1})$ and the minimum to be 0 as needed.

Note that, in general, $f(p, q) > f(S_2(p, q))$, that is, the invariance of f is not preserved by the symmetrization. Moreover, although f is defined on R^{2N} it is not $P(2N)$-invariant. Carefully choosing the semigroup G_2 is very important. $\diamondsuit$

Certainly, using convexity in these last two examples leads to shorter proofs. Nevertheless, the proofs here highlight the role of symmetry in such inequalities.

3.2 The role of sub-invariance

Much of the analysis of symmetrical properties using variational arguments follows the pattern in the three examples 9, 10, 11 described above. The hard work lies in verifying the conditions. Since the symmetric extremal principle is a consequence of the G-sub-invariance property of the given function f, it is unsurprising that the latter is often more powerful. We illustrate with two more subtle inequalities.

We start with the Muirhead inequality [34] which is described in terms of majorization (see e.g. [8]). Recall that, for vectors $x, y \in R^N$, we say x *majorizes* y, denoted by $x \prec y$, if

$$\sum_{n=1}^{k} x_n^{\downarrow} \geq \sum_{n=1}^{k} y_n^{\downarrow}, \text{ for } 1 \leq k < N \text{ and } \sum_{n=1}^{N} x_n^{\downarrow} = \sum_{n=1}^{N} y_n^{\downarrow}.$$

We will use a semigroup characterization of majorization that follows the exposition in [26].

We define, for $\delta > 0$, the δ-average operators a^δ_{ij} by

$$a^\delta_{ij} x := \begin{cases} x - \delta e^i + \delta e^j & \text{if } x_i - x_j > 2\delta, \\ x & \text{otherwise,} \end{cases}$$

and use G_3 to denote the semigroup of all the finite compositions of the δ-average operators. Then [26, Lemma 4] can be stated as

Proposition 12 (Characterization of majorization)**.** *We have $x \prec y$ if and only if $y = gx$ for some $g \in G_3$.*

The Muirhead inequality concerns the following function

$$T[y](x) := \sum_{\pi \in P(N)} x^{y_1}_{\pi(1)} x^{y_2}_{\pi(2)} ... x^{y_N}_{\pi(N)} \; y, x \in R^N_+.$$

Theorem 13 (Muirhead inequality)**.** *For any $x \in R^N_+$, the function $y \mapsto T[y](x)$ is G-sub-invariant.*

Proof. We need only to show that $a^\delta_{ij} y = z$ implies $T[y](x) \geq T[z](x)$. If $y_i - y_j < 2\delta$ then $y = z$ and $T[y](x) = T[z](x)$. We now consider the nontrivial case in which $y_i - y_j > 2\delta$ so that $z_i = y_i - \delta$ and $z_j = y_j + \delta$ and $z_k = y_k, k \neq i, j$. Without loss of generality assuming $i < j$, we calculate

$$\begin{aligned}
& T[y](x) - T[z](x) \qquad (3) \\
= & \sum_{\pi \in P(N)} x^{y_1}_{\pi(1)} ... x^{y_{i-1}}_{\pi(i-1)} x^{y_{i+1}}_{\pi(i+1)} ... x^{y_{j-1}}_{\pi(j-1)} x^{y_{j+1}}_{\pi(j+1)} ... x^{y_N}_{\pi(N)} x^{y_N}_{\pi(N)} \times \\
& \times [x^{y_i}_{\pi(i)} x^{y_j}_{\pi(j)} + x^{y_i}_{\pi(j)} x^{y_j}_{\pi(i)} - x^{y_i-\delta}_{\pi(i)} x^{y_j+\delta}_{\pi(j)} - x^{y_i-\delta}_{\pi(j)} x^{y_j+\delta}_{\pi(i)}] \\
= & \sum_{\pi \in P(N)} x^{y_1}_{\pi(1)} ... x^{y_{i-1}}_{\pi(i-1)} x^{y_{i+1}}_{\pi(i+1)} ... x^{y_{j-1}}_{\pi(j-1)} x^{y_{j+1}}_{\pi(j+1)} ... x^{y_N}_{\pi(N)} x^{y_N}_{\pi(N)} \times \\
& \times x^{y_j}_{\pi(i)} x^{y_j}_{\pi(j)} [x^{y_i - y_j}_{\pi(i)} + x^{y_i - y_j}_{\pi(j)} - x^{y_i - y_j - \delta}_{\pi(i)} x^{\delta}_{\pi(j)} - x^{y_i - y_j - \delta}_{\pi(j)} x^{\delta}_{\pi(i)}] \\
= & \sum_{\pi \in P(N)} x^{y_1}_{\pi(1)} ... x^{y_{i-1}}_{\pi(i-1)} x^{y_{i+1}}_{\pi(i+1)} ... x^{y_{j-1}}_{\pi(j-1)} x^{y_{j+1}}_{\pi(j+1)} ... x^{y_N}_{\pi(N)} x^{y_N}_{\pi(N)} \times \\
& \times x^{y_j}_{\pi(i)} x^{y_j}_{\pi(j)} (x^{y_i - y_j - \delta}_{\pi(i)} - x^{y_i - y_j - \delta}_{\pi(j)})(x^{\delta}_{\pi(i)} - x^{\delta}_{\pi(j)}).
\end{aligned}$$

Now it is easy to see that all the summands are nonnegative and, therefore, $T[y](x) \geq T[z](x)$ as asserted. □

For fixed $x \in R_+^N$, $T[y](x)$ attains its minimum on any compact simplex $\{y \in R_+^N : \langle y, \vec{1}\rangle = k\}$. Also, it is easy to check that $S_3(y) = \bar{y}\vec{1}$ is a G_3-symmetry. Thus, by the symmetric extremal principle, for any $y, x \in R_+^N$,

$$T[y](x) \geq T[\bar{y}\vec{1}](x). \tag{4}$$

Relation (4) is very potent. Here are some attractive special cases:

Example 14. For $y = e^1$ we have $T[e^1](x) \geq T[\frac{1}{N}\vec{1}](x)$ or

$$(N-1)! \sum_{n=1}^{N} x_n \geq N! \, (x_1 ... x_N)^{1/N}.$$

Dividing both sides by $N!$ we get the AG-inequality. ♢

Example 15. For $y = e^1 + e^2$ we have $T[e^1 + e^2](x) \geq T[\frac{2}{N}\vec{1}](x)$ which simplifies to

$$\sqrt{\frac{\sum_{n \neq m} x_n x_m}{N(N-1)}} \geq (x_1 ... x_N)^{1/N}.$$

♢

Example 16. In general $T[e^1 + e^2 + ... + e^k](x) \geq T\left[\frac{k}{N}\vec{1}\right](x)$ gives us

$$\left(\frac{\sum_{1 \leq n_1 < ... < n_k \leq N} x_{n_1} ... x_{n_k}}{\binom{N}{k}}\right)^{1/k} \geq (x_1 ... x_N)^{1/N}.$$

♢

We point out that since (4) is a consequence of the Muirhead inequality, the Muirhead inequality itself is more powerful. What follows are two simple illustrations.

Example 17. It is easy to check that $a_{12}^1(2e^1) = e^1 + e^2$. Thus, by Muirhead inequality we have $T[e^1 + e^2](x) \leq T[2e^1](x)$. Explicitly this is

$$N \sum_{n=1}^{N} x_n^2 \geq \sum_{n \neq m} x_n x_m.$$

♢

Note that neither Example 16 nor Example 17 is a consequence of the general inequality (4).

Using the semigroup G_3 we can also state the Karamata inequality [28] as

Theorem 18 (Karamata inequality)**.** *Let f be an extended convex function on R and define $F : R^N \mapsto R \cup \{+\infty\}$ by*

$$F(x) := \sum_{n=1}^{N} f(x_n).$$

Then F is G_3-sub-invariant.

Proof. This is a direct consequence of Lemma 12 and [26, Theorem 1]. However, the semigroup characterization of majorization given in Proposition 12 leads to the following simple argument which better represents the nature of the inequality.

By the definition of G_3 we need only to show that $F(x) \geq F(a_{ij}^{\delta} x)$. When $a_{ij}^{\delta} x \neq x$, this is equivalent to, for $x_i - \delta \geq x_j + \delta$,

$$f(x_i) + f(x_j) - f(x_i - \delta) - f(x_j + \delta) \geq 0$$

or

$$\frac{f(x_i) - f(x_i - \delta)}{x_i - (x_i - \delta)} \geq \frac{f(x_j + \delta) - f(x_j)}{(x_j + \delta) - x_j},$$

which follows directly from the convexity of f. $\square$

Again, we could apply the symmetric extremal principle to $F(x)$ but this only gives us

$$\sum_{n=1}^{N} f(x_n) \geq N f\left(\frac{\sum_{n=1}^{N} x_n}{N}\right)$$

a weak form of the convexity of f. That said, the Karamata inequality itself has many rather useful applications as shown in [26]. This is yet another illustration that the sub-invariance property itself often captures more information than is encapsulated by the symmetric extremal principle.

3.3 Invariance mismatch

Next we consider examples in which the invariance properties of the action function and the solution are at odds.

Let $u_{ij} : R^N \mapsto R^N$ be a map such that $u_{ij}x$ *switches* the components x_i, x_j of x when

$$(x_i - x_j)(i - j) < 0.$$

Proposition 19. *Let G_4 be the semigroup of all the finite compositions of u_{ij}, and let $S_4(x) := x^{\downarrow}$ be a rearrangement of the components of x in non-increasing order. Then S_4 is a G_4-symmetrization.*

Proof. Define

$$f(x) := Nx_1 + (N-1)x_2 + \ldots + 2x_{N-1} + x_N.$$

It is easy to check that f is a strict G_4-sub-invariant function. Since for any $x \in R^N$, $G_4 \cdot x$ is compact, f attains its minimum at some y on $G_4 \cdot x$. By Proposition 4, for any $g \in G_4$, $gy = y$ and, therefore, $y = y^{\downarrow}$. Since $y \in G_4 \cdot x$, y and x have the same components so that $y = x^{\downarrow} = S_4(x)$. It is easy to directly check that $S_4(gx) = S_4(x)$ for any $g \in G_4$ and $S_4^2(x) = S_4(x)$. $\square$

This symmetry will help us in calculating the subdifferential of convex rearrangement-invariant functions. A general result [31] is:

Proposition 20 (Subdifferential of convex rearrangement invariant functions)**.** *Let $f : R^N \mapsto R \cup \{+\infty\}$ be a convex lsc rearrangement (i.e., $P(N)$-invariant) function. Then $y \in \partial f(x)$ if and only if*

$$y^{\downarrow} \in \partial f(x^{\downarrow}) \text{ and } \langle x, y\rangle = \langle x^{\downarrow}, y^{\downarrow}\rangle.$$

Proof. We can apply a finite number of u_{ij} operations to y consecutively to change y to $y^{\downarrow}$. Thus, there is an element $g \in G_4 \subset P(N)$ such that $gy = y^{\downarrow}$. Then $y \in \partial f(x)$ implies that the function

$$h(z) := f(z) - \langle y^{\downarrow}, z\rangle = f(g^{-1}z) - \langle y, g^{-1}z\rangle$$

attains its minimum at some $z = gx$. (Here g^{-1} is the inverse of g in the permutation group $P(N)$; note that G_4 is *not* a group.)

Since f is $P(N)$-invariant and, therefore, G_4-invariant, we see that h is G_4-sub-invariant. By Propositions 5 and 19, h attains its minimum at $(gx)^{\downarrow} = x^{\downarrow}$. That is, $y^{\downarrow} \in \partial f(x^{\downarrow})$. Moreover,

$$f(x) - \langle y, x\rangle = f(gx) - \langle y^{\downarrow}, gx\rangle = f(x^{\downarrow}) - \langle y^{\downarrow}, x^{\downarrow}\rangle,$$

implies $\langle x, y\rangle = \langle x^{\downarrow}, y^{\downarrow}\rangle$. The converse is evident from the inequalities

$$f(z) - \langle y, z\rangle = f(gz) - \langle y^{\downarrow}, gz\rangle \geq f(x^{\downarrow}) - \langle y^{\downarrow}, x^{\downarrow}\rangle = f(x) - \langle y, x\rangle,$$

and we are done. $\square$

Remark 21. One of the keys in the proof is that we can take the inverse g^{-1} in $P(N)$ without changing f. Thus, we needed to require f to be invariant with respect to the larger group $P(N)$. This is a stronger invariance than the G_4-invariance we ultimately derived for the subdifferential of f. $\diamondsuit$

Remark 22. Semigroup G_4 provides us a clear path to iteratively deriving $x^{\downarrow}$ from x. For proving Proposition 20, however, the semigroup G_4' generated by S_4 and the identity mapping suffices. $\diamondsuit$

Remark 23. Proposition 20 directly assisted in calculating the subdifferential of

$$f(x) := \max\{x_n : n = 1, 2, \dots, N\}$$

but only helps indirectly in dealing with a min function or more generally with the kth order statistics [14].

Easy generalizations to non-convex functions is not to be expected since we have a simple counter-example in $f(x_1, x_2) := x_1 x_2$ whose derivative $f'(x_1, x_2) = (x_2, x_1)$ certainly does not have the symmetry property of Proposition 20. $\diamondsuit$

When f is smooth, Proposition 20 characterizes the so called Schur convexity of f [44]. Since Schur convexity is related to many classical inequalities—such as the AG-mean inequality and Muirhead inequality, see [18]—we may expect semigroup G_4 also to help in proving such inequalities. Here is an illustration.

As with G_4 we define G_5 to be the finite composition of maps $v_{ij} : R^N \mapsto R^N$ such that each $v_{ij}x$ switches the components x_i, x_j of x when

$$(x_i - x_j)(i - j) > 0.$$

We can also verify that $S_5(x) := x^{\uparrow}$, the rearrangement of the components of x in non-decreasing order, is a G_5-symmetrization.

The semigroups G_4 and G_5 also enable us to give a more honest symmetry proof of the arithmetic-geometry mean inequality than that of Example 10.

Example 24. Consider the constrained minimization problem

$$\begin{aligned} \text{minimize } f(x) &:= x_1 + x_2 + \ldots + x_N \\ \text{subject to } g(x) &:= x_1 x_2 \ldots x_N \geq 1, \\ & \quad\; x_n \geq 0, n = 1, 2, \dots, N. \end{aligned} \tag{5}$$

Since f is linear the minimum is attained on the part of the boundary of the convex feasible set defined by $x_1 x_2 \ldots x_N = 1$. Let z be a minimal point. Since both f and g are invariant with respect to the group action $P(N)$ which contains both G_4 and G_5, the minimum is also attained at both $z^{\downarrow}$ and $z^{\uparrow}$. Since f and the feasible set are both convex, we have that the convex hull $[z^{\downarrow}, z^{\uparrow}] \subset$ argmin. This implies that $z^{\downarrow} = z^{\uparrow}$ because $x_1 x_2 \ldots x_N = 1$ does not contain any non-degenerate line segment. Again we conclude that all components of z are the same and deduce that the minimum N is attained at $z_1 = z_2 = \ldots = z_N = 1$, which implies

$$\frac{x_1 + x_2 + \ldots + x_N}{N} \geq (x_1 x_2 \ldots x_N)^{1/N}.$$

This is again the classical algebraic-geometric mean inequality. $\diamondsuit$

Proposition 20 is a special case of a more general formula for subdifferentials of convex spectral functions. Let $O(N)$ be the group of $N \times N$ orthogonal matrices and $S(N)$ be the $N \times N$ symmetric matrices endowed with the norm induced by the matrix inner product

$$\langle A, B\rangle = \mathrm{tr}(A^\top B).$$

For $U \in O(N)$ and $A \in S(N)$, we define a group action by

$$U \cdot A := UAU^\top.$$

Define $S_4^*(A) := \mathrm{diag}\ \lambda(A)$, where $\lambda(A)$ is the vector of eigenvalues of A in a non-increasing order and $\mathrm{diag}(x)$ signifies the diagonal matrix generated by vector x. Then by virtue of the von Neumann-Theobald inequality [9, 14]

$$\langle A, B\rangle \leq \langle S_4^*(A), S_4^*(B)\rangle. \tag{6}$$

Correspondingly to Proposition 20 we have:

Proposition 25 (Orthogonal invariance)**.** *Let $f : S(N) \mapsto R \cup \{+\infty\}$ be a convex lsc function invariant with respect to the orthogonal transformations. Then $Y \in \partial f(X)$ if and only if*

$$\mathrm{diag}\,\lambda(Y) \in \partial f(\mathrm{diag}\,\lambda(X)) \text{ and } \langle X, Y\rangle = \langle \lambda(X), \lambda(Y)\rangle.$$

Proof. Define G_4^* to be the semigroup generated by S_4^* and the identity mapping. We now verify that $Z \mapsto f(Z) - \langle S_4^*(Y), Z\rangle$ is sub-invariant with respect to G_4^*. The rest of the proof is like that of Proposition 20—as simplified in Remark 22. □

Remark 26. Define $\phi(x) := f(\mathrm{diag}\ x)$. It is an easy matter to check that when f is orthogonally invariant then ϕ is permutation invariant and $y \in \partial\phi(x)$ if and only if $\mathrm{diag}\ y \in \partial f(\mathrm{diag}\ x)$. Thus, Proposition 25 actually reduces the problem of computing the subdifferential of f on $S(N)$ to one of computing the subdifferential of ϕ on the much smaller dimensional space R^N. □

It is not hard to verify that, for any $w \in R^N$,

$$A \mapsto \langle w^\downarrow, \lambda\rangle(A) := \langle w^\downarrow, \lambda(A)\rangle$$

is a convex $O(N)$-invariant function. As a corollary of Proposition 25 we have

Corollary 27. *For any vectors $w, x \in R^N$,*

$$\mathrm{diag}\ w^\downarrow \in \partial\langle w^\downarrow, \lambda\rangle(\mathrm{diag}\ x^\downarrow). \tag{7}$$

Formula (7) is one of the keys in building the representation theorem for subdifferentials of the spectral functions (see [14, 31]).

3.4 Approximation of symmetrization

3.4.1 Rearrangement of infinite series

Let us first consider how to extend Theorem 20 from finite dimensions to l^2 or another sequence space. Although the idea is similar to that of the finite dimensional case, two technical difficulties need to be addressed. First, using decreasing or increasing rearrangements of the components as a target symmetry no longer works directly. Moreover, in order to generalize the Theobald inequality of (6) to l^2 we may need to strip away some of the zeros components to avoid the pathological situation identified in [14, Exercise 7.3.8].

In [10, 11] this is achieved through a particular rearrangement which sandwiches zeros in between positive and negative components. Here, we adopt a different approach that is more natural both in using the symmetry and in being closer to the argument for the case of finite dimensional problem. One main difference in comparison to the finite dimensional arguments is that we must resort to approximations to help prove that the symmetry is compatible with the semigroup and the action function.

We fix notation first. We represent l^2 bilaterally as

$$l^2 = \{x = \sum_{n=-\infty}^{\infty} x_n e^n : \sum_{n=-\infty}^{\infty} x_n^2 < \infty\},$$

where e^n is the standard base in which the nth component is 1 and all the other components are 0. We will use the *left* and *right shift* operators defined by

$$R_S x := \sum_{n=-\infty}^{\infty} x_{n-1} e^n \quad \text{and} \quad L_S x := \sum_{n=-\infty}^{\infty} x_{n+1} e^n.$$

The inner product and the *Hamilton product* are defined by

$$\langle x, y \rangle := \sum_{n=-\infty}^{\infty} x_n y_n \tag{8}$$

and

$$x \circ y := \sum_{n=-\infty}^{\infty} x_n y_n e^n, \tag{9}$$

respectively. For $k < l$, we denote

$$1_k^l := \sum_{n=k}^{l} e^n,$$

and we will allow $k = -\infty$ and $l = \infty$.

For any $x \in l^2$, define $S_6(x) := x^*$ to be a rearrangement of the components—with the possibility of deleting or adding an arbitrary number of zeros—such that all the positive components have nonnegative indices and are arranged in non-increasing order, followed by zeros if necessary, and all the negative components have negative indices arranged in non-increasing order, preceded by zeros as necessary. For example, if

$$x = (...... - 2, 3, -1, -5, -4, \mathbf{7}, 4, 5, 2, 0, 0, ...)$$

then

$$x^* = (......0, -1, -2, -4, -5, \mathbf{7}, 5, 4, 3, 2, 0, ...)$$

where the boldfaced component $\mathbf{7}$ corresponding to index 0.

Our next goal is to define a semigroup action for which $S_6 = *$ is the natural symmetry. We need two basic operations: 'switch' and 'move' as defined next. The *switch* operator s_{nm} switches components x_n and x_m if $n < m < 0$ or $0 \le n < m$ and when doing so brings x closer to x^* in norm. More precisely, if $n < m < 0$ or $0 \le n < m$,

$$s_{nm}x := x - x_n e^n - x_m e^m + \max(x_n, x_m)e^n + \min(x_n, x_m)e^m.$$

The *move* operator moves positive components to the right of $n = 0$ (inclusive) and negative components to the left of $n = -1$. In doing so we must shift some of the components to make room and we must also make sure the move brings x closer to x^*. The precise definition follows:

$$m_n x := \begin{cases} x \circ 1_{-\infty}^{k-1} - x_n e^n + x_n e^k + R_S(x \circ 1_k^{\infty}) & \text{if } n < 0 \text{ and } x_n > 0 \\ x \circ 1_{l+1}^{\infty} - x_n e^n + x_n e^l + L_S(x \circ 1_{-\infty}^{l}) & \text{if } n \ge 0 \text{ and } x_n < 0 \\ x & \text{otherwise,} \end{cases}$$

where $k := \min\{m \ge 0 : \sup_{i \ge m} |x_i| < x_n\}$ and $l := \max\{m < 0 : \sup_{i \le m} |x_i| < -x_n\}$. The most important property of these two operators is that when applied to x they increase the inner product $\langle y^*, x\rangle$ with respect to any y^*.

Lemma 28. *Let $x, y \in l^2$ and assume that $y = y^*$. Then*

$$\langle y, x\rangle \le \langle y, s_{nm}x\rangle,$$

and

$$\langle y, x\rangle \le \langle y, m_n x\rangle.$$

Proof. The first inequality is obvious. We turn to the second. When $n < 0$ and $x_n > 0$, the inequality follows from the following estimates

$$\begin{aligned}
& \langle y, m_n x\rangle - \langle y, x\rangle \\
= & \sum_{i=-\infty}^{k-1} y_i x_i - y_n x_n + x_n y_k + \sum_{i=k+1}^{\infty} x_{i-1} y_i - \sum_{i=-\infty}^{\infty} x_i y_i \\
= & -y_n x_n + x_n y_k - x_k y_k + \sum_{i=k+1}^{\infty} (x_{i-1} - x_i) y_i \\
= & -y_n x_n + x_n y_k - \sum_{i=k}^{\infty} x_i (y_i - y_{i+1}) \text{ (using Abel's formula)} \\
\geq & \; x_n y_k - \sum_{i=k}^{\infty} |x_i| (y_i - y_{i+1}) \text{ (since } - x_n y_n \geq 0 \text{ and } y_i - y_{i+1} \geq 0) \\
\geq & \; x_n (y_k - \sum_{i=k}^{\infty} (y_i - y_{i+1})) = 0 \text{ (since } x_n \geq |x_i|, i \geq k).
\end{aligned}$$

The case when $n \geq 0$ and $x_n < 0$ is analogous. $\square$

For a natural number N we define $G^N :=$ {all finite composition of s_{nm} and m_n where $|n|, |m| \leq N$} and set $G_6 := \cup_{N=1}^{\infty} G^N$. It is easy to verify that G^N and G_6 are semigroups and S_6 is a G_6-symmetrization.

Now, define H to be the semigroup of all self-mappings of l^2 which add or delete an arbitrary number of zeros and then perform a permutation of the components. We can check that for any $y \in l^2$ there exists $h_y, h^y \in H$ such that $h_y y^* = y$ and $y^* = h^y y$. Next we observe, on applying Lemma 28, that the function $\varphi(x) := -\langle y^*, x\rangle$ is sub-invariant under the semigroup action G_6. Moreover, for $x \in l^2$ and $h \in H$, if the components of $x^* \circ 1_k^l$ are a subset of $\{(hx)_n, |n| \leq N\}$, then $\varphi(x)$ attains a minimum on $G^N(hx)$ at some x_h^N and $x^* \circ 1_k^l = x_h^N \circ 1_k^l$.

Letting $k \to -\infty$ and $l \to \infty$ we see that $x_h^N \to x^*$ as $N \to \infty$. Now we can show the following generalization of Proposition 20 holds:

Proposition 29 (Subgradients on Hilbert space)**.** *Let $f : l^2 \mapsto R \cup \{+\infty\}$ be a convex lsc function invariant under the semigroup action H. Then $y \in \partial f(x)$ if and only if*

$$y^* \in \partial f(x^*) \text{ and } \langle x, y\rangle = \langle x^*, y^*\rangle.$$

Proof. Let $y \in \partial f(x)$. Then, for all $z \in l^2$,

$$\begin{aligned} f(z) - \langle y^*, z\rangle &= f(h_y z) - \langle h_y y^*, h_y z\rangle \\ &= f(h_y z) - \langle y, h_y z\rangle \\ &\geq f(x) - \langle y, x\rangle \\ &= f(h^y x) - \langle y^*, h^y x\rangle. \end{aligned} \tag{10}$$

Since f is invariant with respect to the action H, the minimum of $f(h^y x) - \langle y^*, h^y x\rangle$ is the same as that of the sub-invariant function $-\langle y^*, h^y x\rangle$ on $G^N(h^y x)$ and is $x^N_{h^y}$. Thus, for all N,

$$f(z) - \langle y^*, z\rangle \geq f(x^N_{h^y}) - \langle y^*, x^N_{h^y}\rangle.$$

Taking the limit as $N \to \infty$ we have

$$f(z) - \langle y^*, z\rangle \geq f(x^*) - \langle y^*, x^*\rangle,$$

or $y^* \in \partial f(x^*)$.

Setting $z = x^*$ in (10) we have $f(x^*) - \langle y^*, x^*\rangle \geq f(x) - \langle y, x\rangle$. Since $f(x^*) = f(x)$ we have $\langle y^*, x^*\rangle \leq \langle y, x\rangle$. But the opposite inequality always holds, so we have $\langle y^*, x^*\rangle = \langle y, x\rangle$.

On the other hand, if

$$y^* \in \partial f(x^*) \text{ and } \langle x, y\rangle = \langle x^*, y^*\rangle,$$

then, for any $z \in l^2$,

$$\begin{aligned} f(z) - \langle y, z\rangle &= f(h^y z) - \langle y^*, h^y z\rangle \\ &\geq f(x^*) - \langle y^*, x^*\rangle \\ &= f(x) - \langle y, x\rangle. \end{aligned} \tag{11}$$

That is, $y \in \partial f(x)$, as asserted. □

Remark 30. (a) Again the invariance property of the action function f (H-invariance) is different from that of the subdifferential (G_6-invariance, where G_6 is a proper subset of H). (b) The requirement of f being invariant with respect to semigroup action H is stronger than rearrangement invariance. However, most commonly occurring spectral functions such as norms, the log barrier, sup etc, do satisfy this requirement. (c) To check that $*$ is a (H, f)-symmetrization using approximation from G^N seems crucial. ♢

3.4.2 Rearrangement of measurable functions

Similar ideas can also be implemented in continuous infinite dimensional spaces. One motivation is the search for symmetric solutions to Laplace's equation

$$\Delta u = f \text{ in } \Omega, \qquad u|_{\partial\Omega} = 0. \tag{12}$$

Solutions of equation (12) correspond to critical points of the action function

$$F(u) := \int_{\Omega} \left(\frac{|\nabla u|^2}{2} + fu \right) \mu(dx), \tag{13}$$

in the Sobolev space $H_0^1(\Omega)$.

Arguing directly that F attains its infimum is more than a bit technical. Variational principles help here since F is bounded from below, and so as a consequence of the Ekeland variational principle, there is a sequence u_n such that

$$F'(u_n) \to 0 \textit{ and } F(u_n) \to c = \inf F. \tag{14}$$

It turns out that F satisfies the *Palais-Smale condition* [36], i.e., convergence of sequences in (14) implies that u_n has a convergent subsequence $H_0^1(\Omega)$ whose limit attains the minimum of F and thus the solution to (12).

We turn to describing an appropriate concept of symmetry. Consider the Lebesgue measure space $(R^n, \mathcal{M}, \mu^n)$. We write μ when the dimension is clear. A set transformation $T : \mathcal{M} \mapsto \mathcal{M}$ is called a *rearrangement* if T is *monotone*: $A \subset B$ implies $T(A) \subset T(B)$ and *measure preserving*: $\mu(T(A)) = \mu(A)$ for all $A \in \mathcal{M}$. A rearrangement of sets induce a rearrangement of functions in the class of *symmetrizable* functions $\mathcal{S} := \{u : \mu(u > \inf u) < \infty\}$ by

$$Tu(x) := \sup\{c > \inf u : x \in T(\{u > c\})\}, \qquad \forall x \in R^n.$$

Clearly, all the rearrangements form a semigroup. Rearrangement is nonexpansive as established in Theorem 3, Corollary 1 of [20]:

Theorem 31. *Let T be a rearrangement and let j be a strictly convex function with $j(0) = 0$. Then*

$$\int_{R^n} j(|Tu - Tv|)\, d\mu \le \int_{R^n} j(|u - v|\,)d\mu, \qquad \forall u, v \in \mathcal{S}, \tag{15}$$

as soon as either of the integrals in (15) is finite.

For $j(t) := t^2$, this leads to $\|Tu - Tv\|_2 \leq \|u - v\|_2$ and, as a consequence, we obtain the following *Hardy-Littlewood inequality*

$$\int_{R^n} u\, v\, d\mu \leq \int_{R^n} Tu\, Tv\, d\mu, \qquad \forall u, v \in L^2_+(R^n). \tag{16}$$

Moreover, any measure preserving rearrangement, satisfies *Cavalieri's principle* [15, Eqn. (3.7)], for any continuous function f,

$$\int_{R^N} f(u)\, u(d\mu) = \int_{R^N} f(Tv)\, v(d\mu). \tag{17}$$

3.4.3 Steiner symmetrization

We now turn to *Steiner symmetrization.* Let $\Sigma \subset R^n$ be a k-dimensional plane. We say that $S : \mathcal{M} \mapsto \mathcal{M}$ is a (k,n)-*Steiner symmetry induced by* Σ if, for every $x \in \Sigma$ and $M \in \mathcal{M}$,

$$S(M) \cap (x + \Sigma^\perp) := B_r(x) \cap (x + \Sigma^\perp)$$

with $\mu^k(B_r(x) \cap (x + \Sigma^\perp)) = \mu^k(M \cap (x + \Sigma^\perp))$. In particular, for $k = 0$ we get as a special case the *Schwarz symmetry* or decreasing rearrangement symmetry defined by $S_7(M) = B_r(0)$ for r such that $\mu(B_r(0)) = \mu(M)$. Clearly, Schwarz symmetry induces a rearrangement. The induced rearrangement for functions u in $\mathcal{S}$ will be denoted by $u^* := S_7(u)$.

We wish to show that when f and Ω are Schwarz symmetric so is the solution u to equation (12). For this we need to find a semigroup G_7 such that the action function F is G_7-sub-invariant and the Schwarz symmetry $S_7 := *$ is a (G_7, F)-symmetrization. It turns out again that we need to approximate $*$. A polarization first introduced by Wolontis for plane sets [50] and later extended to functions by Baenstein and Taylor [7] suits this purpose.

Let X_0 be a hyperplane in R^N that does not contain the origin and so dividing R^N into two closed half-spaces. Denote by X_+ the half-space contains 0 and X_- the closed half-space complementary to X^+. Let σ be the reflection exchanging the two half-spaces. The *polarization* of a function f at X_0 is

$$f^\sigma(x) := \begin{cases} \max\{f(x), f(\sigma x)\} & x \in X_+, \\ \min\{f(x), f(\sigma x)\} & x \in X_-, \\ f(x) & x \in X_0. \end{cases}$$

It turns out the semigroup G_7 of all finite compositions of polarizations is exactly what we need as shown in the following theorem summarizing relevant results by Brock and Solynin [15] and in Schaftingen [41].

Theorem 32 (Properties of polarization)**.** *Let G_7 denote the semigroup all finite compositions of polarizations. Then we have*

(1) Hardy-Littlewood inequality*:*

$$\int fg \leq \int f^\sigma g^\sigma, \quad \forall \sigma \in G_7. \tag{18}$$

(2) Decreasing L^p norm*:*

$$\|f - g\|_p \geq \|f^\sigma - g^\sigma\|_p, p \geq 1, \forall \sigma \in G_7. \tag{19}$$

(3) Universal strong approximation of Schwarz symmetrization in L^p*: There exists a sequence $g_k \in G_7$ such that, for any $f \in L^P$, $\|g_k f - f^*\|_p \to 0$.*

(4) Weak approximation of Schwarz symmetrization in $W^{1,p}$*: Let $f \in W^{1,p}$. Then there exists a sequence $g_k \in G_7$ such that $g_k f \to f^*$ weakly in $W^{1,p}$.*

(5) Characterization of $*$*: $f^* = f$ if and only if $f^\sigma = f$ for all $\sigma \in G_7$.*

(6) Preservation of the norm*: $\|f^\sigma\|_{H^1} = \|f\|_{H^1}$ for all $\sigma \in G_7$.*

Now we can establish the symmetry of a solution to equation (12) when its data is symmetric.

Theorem 33 (Symmetric Laplace solution)**.** *Suppose that both f and Ω are Schwarz symmetric then so is every solution u to equation (12).*

Proof. Since the action function F defined in (13) is convex, critical points of F are all minima. Let G_7 be the semigroup of finite compositions of polarizations. Then properties (1) and (6) of the polarization implies that, the function F defined in (13) is G_7-sub-invariant when $f = f^*$. Property (3), (4) and (5) of the polarization implies that the Schwarz symmetry $*$ is a (G_7, F)-symmetrization. Thus, if $\Omega = \Omega^*$ and $f = f^*$, then every solution of the PDE (12) is Schwarz symmetric, that is $u = u^*$. □

Remark 34. (a) The use of approximate polarization is essential and non-trivial. (b) Using symmetrization helped but did not make the work easy. (c) Although we have a general framework, interesting concrete problems have to be dealt with symmetry by symmetry. ♢

For instance, we do not know whether the framework is applicable to the periodical solutions of planar motion of two bodies. Mathematically, this can be formulated as minimization of the action functional

$$F(x) := \int_0^P \left[\frac{\|x'(t)\|^2}{2} + \frac{1}{\|x(t)\|} \right] dt$$

in the space of periodic orbits $\{x \in H^1([0,P],R^2) : x(0) = x(P)\}$ [24]. Clearly F is rotation-invariant and Kepler 'showed' the solution is a circle. Thus, both the action function and the solution are rotation-invariant.

Can one find a (semi)group G and a (G,F)-symmetrization to fit this problem into the above framework? We return to this problem in Example 49 using a different approach.

3.4.4 The isoperimetric problem

The isoperimetric problem has a long history. Legend says that Queen Dido of Carthage already knew that among all the shapes with the same perimeter a disk has the maximum area.

Rigorously, we can formulate the result as

$$\mathrm{Per}(A) \geq \mathrm{Per}(A^*),$$

where $A \subset R^2$ and 'Per' denotes the perimeter. However, rigorous solution is rather recent. Here we discuss Steiner's solution using symmetry [46]. The method may have been known earlier to Gergonne (see [23, 37]). Following [16], we will assume the more recent knowledge from geometric measure theory that (a) for any $A \subset R^2$, $\mathrm{Per}(A) \geq \mathrm{Per}(\mathrm{co}A)$ so that we need only consider convex sets; and that (b) among all convex sets with equal area contained in a fixed bounded region, a perimeter minimizing set exists.

Labeling the standard coordinate system of R^2 by 0, we use θ to denote the coordinate system by rotating 0 counter-clockwise θ degrees. Let A be a bounded convex set in R^2 and let u_θ and l_θ be the functions representing the upper and lower boundaries of A in θ so that

$$A = \{(x,y) : a_\theta \leq x \leq b_\theta, l_\theta(x) \leq y \leq u_\theta(x)\}.$$

We define the *θ-symmetrization*: of A by

$$s_\theta A := \{(x,y) : a_\theta \leq x \leq b_\theta, (l_\theta(x) - u_\theta(x))/2 \leq y \leq (u_\theta(x) - l_\theta(x))/2\},$$

and we define the semigroup G_8 to be the collection of finite compositions of θ-symmetrizations. It is not hard to check that the Schwarz symmetrization $*$ is then a G_8-symmetrization.

Also, the semigroup action on convex sets preserves the area and reduces the perimeter as is made precise in the following lemma.

Lemma 35. *Let A be a bounded convex set in R^2. Then, for any $g \in G_8$,*

$$\mathrm{Area}(A) = \mathrm{Area}(gA) \text{ and } \mathrm{Per}(A) \geq \mathrm{Per}(gA).$$

Moreover, the inequality is strict if $gA \neq A$.

Proof. We need only show the conclusion for $g = s_\theta$. The property of preserving the volume follows directly from computation:

$$\begin{aligned}\mathrm{Area}(A) &= \int_{a_\theta}^{b_\theta} (u_\theta(x) - l_\theta(x))\, dx \\ &= \int_{a_\theta}^{b_\theta} \left(\frac{u_\theta(x) - l_\theta(x)}{2} - \frac{l_\theta(x) - u_\theta(x)}{2} \right) dx = \mathrm{Area}(s_\theta A).\end{aligned}$$

For the reduction of the perimeter we use the convexity of the function $t \mapsto \sqrt{1+t^2}$ and the representation

$$\mathrm{Per}(A) = (u_\theta(a) - l_\theta(a)) + (u_\theta(b) - l_\theta(b)) + \int_{a_\theta}^{b_\theta} \left(\sqrt{1 + u_\theta'^2(x)} + \sqrt{1 + l_\theta'^2(x)} \right) dx$$

to finish the proof. □

We are now ready to prove

Theorem 36 (Isoperimetric inequality)**.** *If $A \subset R^2$ is convex and has finite perimeter, then*

$$\mathrm{Per}(A) \geq \mathrm{Per}(A^*).$$

Proof. We know that Per attains a minimum at $B \in G_8 \cdot A$ and Per is strictly G_8-sub-invariant by Lemma 35. Thus, by Proposition 4 we must have $gB = B$ for any $g \in G_8$. This means B is a disc centered at $(0,0)$. Moreover, $B \in G_8 \cdot A$ implies that Area(B)=Area(A). Thus, $B = A^*$ and the inequality follows. □

4 Saddle Points

So far all the examples are about extrema of the action function. But saddle points are also important. We illustrate two methods. The first uses Ambrosetti and Rabinowitz's notion of a mountain pass [4], and the second introduces Palais' principle of symmetric criticality [35].

4.1 Mountain passes

The mountain pass theorem due to Ambrosetti and Rabinowitz is an important tool in dealing with solutions of both ordinary and partial differential equations that correspond to saddle points of certain action functions. One can understand this theorem intuitively as its name suggests. Consider a basin surrounded by mountains as illustrated in Figure 2.

To travel from a village in the basin to a city outside these surrounding mountains one must follow a mountain pass crossing the mountain ridge. We

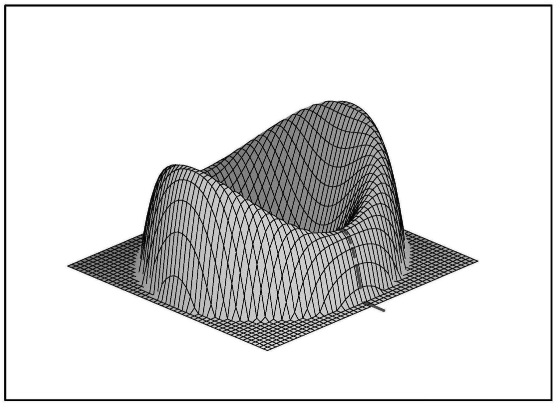

Figure 2: A typical mountain pass.

can see that for each such mountain pass, there will be a point with the highest elevation. Assuming appropriate compactness, among all the possible mountain passes, there must be a pass with a lowest highest elevation—which it seems clear will give us a saddle-type critical point of the terrain. Following [4], we now make this intuitive description precise.

Definition 37 (Separation)**.** *Let X be a Banach space and let S be a closed subset of X. We say that S* separates *two points a and b in X provided that a and b belong to disjoint connected components of $X \backslash S$.*

We can now precisely state the mountain pass theorem [4].

Theorem 38 (Mountain pass)**.** *Let X be a Banach space, let $a, b \in X$ and let $f : X \mapsto R$ be a continuous and Gâteaux differentiable function. Define*

$$c := \inf_{x \in \Gamma(a,b)} \max_{t \in [0,1]} f(x(t)),$$

where $\Gamma(a, b) := \{x \in C([0, 1], X) \mid x(0) = a, x(1) = b\}$.

Suppose that S is a closed subset of X such that $S \subset \{x \in X \mid f(x) \geq c\}$ and S separates a and b. Suppose also that f satisfies the Palais-Smale condition of Section 3.4.2. Then there exists a point $\bar{x} \in S$ such that $f(\bar{x}) = c$ and $f'(\bar{x}) = 0$.

The Palais-Smale condition [36] ensures the existence of an exact critical point. Without it, in general, one can only derive an approximate version of the mountain pass theorem. Suppose in addition that f has a symmetric property in the form of sub-invariance with respect to a semigroup action G

on the space X. Then we can also expect the approximate critical point is close to a symmetric point.

We present such an approximate version of the symmetric mountain pass theorem below and then illustrate how to use it to derive the existence of symmetric critical points. For this purpose we need the following definition.

Definition 39 (Compatible metric)**.** *Let (X, d) be a complete metric space with a semigroup action G. We say that the metric d is* compatible *with G if, for any $x \in X$ and $g \in G$,*

$$d(x, y) \geq d(gx, gy).$$

Note that when G is a group, d compatible to G is equivalent to: *for any $x \in X$ and $g \in G$, $d(x, y) = d(gx, gy)$*, that is to say, g is an isometry on X.

Theorem 40 (Approximate symmetric mountain pass)**.** *Let X be a Banach space with a semigroup action G, and suppose the norm on X is compatible with G. Let $f\colon X \mapsto \mathbb{R}$ be a continuous and Gâteaux differentiable G-sub-invariant function. Suppose that $a, b \in X$ are invariant under the semigroup action G, that is, for any $g \in G$, $a = ga$ and $b = gb$. Define*

$$c := \inf_{x \in \Gamma(a,b)} \max_{t \in [0,1]} f(x(t)).$$

Suppose that S is a closed subset of X such that $S \subset \{x \in X \mid f(x) \geq c\}$ and S separates a and b. Then, for any $g \in G$ satisfying $gS = S$, there exist points $x_\varepsilon, z_\varepsilon \in X$ such that

(i) $\|x_\varepsilon - gz_\varepsilon\| < \frac{1}{2}\varepsilon$,

(ii) $c < f(x_\varepsilon) < c + \frac{5}{4}\varepsilon^2$,

(iii) $d(S; x_\varepsilon) < \frac{3}{2}\varepsilon$ *and*

(iv) $\|f'(x_\varepsilon)\| < \frac{3}{2}\varepsilon$.

Proof. Since S separates a and b we can find two disjoint open sets U and V such that $X \backslash S = U \cup V$ and $a \in U$ while $b \in V$. Fix ε so that $0 < \varepsilon < \frac{1}{2}\min(1, d(S; a), d(S; b))$. Let $z \in \Gamma(a, b)$ satisfy

$$\max\{f(z(t)) \mid t \in [0,1]\} < c + \frac{\varepsilon^2}{4}. \tag{20}$$

Set $h(x) := \varepsilon \max(0, \varepsilon - d(S; x))$, and define a function $\varphi : \Gamma(a, b) \mapsto \mathbb{R}$ by

$$\varphi(x) := \max\{f(x(t)) + h(x(t)) \mid t \in [0,1]\}.$$

Note that for any $x \in \Gamma(a,b)$ $x([0,1]) \cap S \neq \emptyset$, since $x(0) = a \in U$, $x(1) = b \in V$ and $X \backslash S = U \cup V$. It follows that for any $x \in \Gamma(a,b)$

$$\varphi(x) \geq \max\{f(x(t)) + h(x(t)) \mid t \in [0,1] \text{ and } x(t) \in S\} \geq c + \varepsilon^2$$

so that

$$\inf_{\Gamma(a,b)} \varphi \geq c + \varepsilon^2. \tag{21}$$

Moreover, we have

$$\varphi(z) \leq \max\{f(z(t)) + h(z(t)) \mid t \in [0,1]\} \leq \big(c + \frac{\varepsilon^2}{4}\big) + \varepsilon^2 < \inf_{\Gamma(a,b)} \varphi + \frac{\varepsilon^2}{4}. \tag{22}$$

On the other hand, the mappings $(gx)(t) = gx(t)$, for $x \in \Gamma(a,b)$ together with the identity mapping generate a semigroup on $\Gamma(a,b)$ which we denote G_9. It is easy to see that φ is G_9 sub-invariant. Applying the symmetric Ekeland variational principle of Theorem 6 to φ on $\Gamma(a,b)$ we can find a path $y \in \Gamma(a,b)$ such that

$$\varphi(y) \leq \varphi(gz), \tag{23}$$

$$\|y - gz\| \leq \varepsilon/2, \tag{24}$$

and

$$\varphi(x) + \frac{\varepsilon}{2}\|x - y\| \geq \varphi(y) \qquad \forall x \in \Gamma(a,b). \tag{25}$$

Now let M be the subset of $[0,1]$ consisting of all points where $(f+h) \circ y$ attains its maximum on $[0,1]$. We prove first that there exists $\bar{t} \in M$ such that $\|f'(y(\bar{t}))\| \leq \frac{3}{2}\varepsilon$.

Indeed, first note (25) shows that for any $\eta \in C([0,1];X)$ with $\eta(0) = \eta(1) = 0$,

$$-\frac{\varepsilon}{2}\|\eta\| \leq \liminf_{s \to 0+} \frac{\varphi(y + s\eta) - \varphi(y)}{s}.$$

Using the definition of the Gâteaux differential of f and the fact that h has Lipschitz constant ε, it follows that the last inequality is dominated by

$$\liminf_{s \to 0+} \frac{1}{s}\Big[\max_{t \in [0,1]}((f+h)(y(t)) + s\langle f'(y(t)), \eta(t)\rangle) - \max_{t \in [0,1]}((f+h)(y(t))\Big] + \varepsilon\|\eta\|.$$

Hence

$$-\frac{3\varepsilon}{2}\|\eta\| \leq \liminf_{s \to 0+} \frac{m(k + sl) - m(k)}{s}, \tag{26}$$

where $k = (f + h) \circ y$, $l = \langle f'(y), \eta \rangle$ and m is the continuous convex function on $C([0,1]; X)$ defined by $m(x) := \max\{x(t) \mid t \in [0,1]\}$.

Recall that the convex subdifferential of m has the following representation [25] $\partial m(x) = \{\mu \mid \mu$ is a Radon probability measure supported in $M(x)\}$ where $M(x) := \{t \in [0,1] \mid x(t) = m(x)\}$. It follows from (26) and the 'max-formula' for the convex subdifferential (see e.g. [14, Theorem 4.2.7]) that

$$\begin{aligned} -\frac{3\varepsilon}{2}\|\eta\| &\leq \liminf_{s\to 0+} \frac{m(k+sl) - m(k)}{s} \\ &\leq \max\{\langle l, \mu \rangle \mid \mu \in \partial m(k)\} \\ &= \max\{\int \langle f'(y), \eta \rangle \, d\mu \mid \mu \in \partial m(k)\}. \end{aligned}$$

By a standard minimax theorem (e.g., [5, Theorem 6.2.7]) we have

$$\begin{aligned} -\frac{3\varepsilon}{2} &= \inf_{\eta} \max_{\mu} \Big\{ \int \langle f'(y), \eta \rangle \, d\mu \,\Big|\, \mu \in \partial m(k), \|\eta\| \leq 1, \eta(0) = \eta(1) = 0 \Big\} \\ &= \max_{\mu} \inf_{\eta} \Big\{ \int \langle f'(y), \eta \rangle \, d\mu \,\Big|\, \mu \in \partial m(k), \|\eta\| \leq 1, \eta(0) = \eta(1) = 0 \Big\} \\ &= \max_{\mu} \Big\{ -\int \|f'(y)\| \, d\mu \,\Big|\, \mu \in \partial m(k) \Big\} \\ &\leq -\min\left\{ \|f'(y(t))\| \mid t \in M(k) \right\}. \end{aligned}$$

Combining (20) and (21) we can verify that

$$M(k) \cap \{0,1\} = \emptyset. \tag{27}$$

Therefore, there exists $\bar{t} \in M(k) = M$ such that $\|f'(y(\bar{t}))\| \leq 3\varepsilon/2$. It remains to show that points $x_\varepsilon = y(\bar{t})$ and $z_\varepsilon = z(\bar{t})$ satisfy (i), (ii) and (iii).

Clearly (i) follows directly from (24). For (ii) combine (21), (22) and (23) to get

$$c + \varepsilon^2 \leq \inf_{\Gamma(a,b)} \varphi \leq f(y(\bar{t})) + h(y(\bar{t})) = \varphi(y) \leq \varphi(z) \leq c + \frac{5\varepsilon^2}{4}.$$

Since $0 \leq h \leq \varepsilon^2$ we obtain $c \leq f(x_\varepsilon) \leq c + 5\varepsilon^2/4$. For (iii) we combine $f(gz(\bar{t})) + h(gz(\bar{t})) \geq \varphi(gz) \geq \varphi(y) \geq c + \varepsilon^2$ and (20) to conclude $h(gz(\bar{t})) > 3\varepsilon^2/4 > 0$. This implies that $d(S; gz_\varepsilon) = d(S; gz(\bar{t})) < \varepsilon$. This combined with (24) gives that $d(S; x_\varepsilon) = d(S; y(\bar{t})) \leq 3\varepsilon/2$, and we are done. □

Example 41. Consider the function $F(x,y) := x^2 - y^2$. Define $S_{10}(x,y) = (0,y)$ and define G_{10} to be the semigroup generated by S_{10} and the identity mapping. Define the space of paths

$$\Gamma := \{\gamma \in C([0,1], R^2) : \gamma(0) = a := (0,1), \gamma(1) = b := (0,-1)\}$$

and set $S = \{(0, y) \mid y \in R\}$. We can check that

$$0 = \inf_{(x,y)(t)\in\Gamma} \max_{t\in[0,1]} F(x(t), y(t)),$$

and that S separates $a = S_{10}(a)$ and $b = S_{10}(b)$. Moreover, $g = S_{10}$ maps Γ to itself and $gS = S$.

Thus, applying Theorem 40, there exist sequences (x_i, y_i) and $(0, z_i)$ such that $\|(x_i, y_i) - (0, z_i)\| \to 0$, $F(x_i, y_i) \to 0$, and $F'(x_i, y_i) \to 0$. Taking limits of this sequence of approximate critical points identifies $(0, 0)$ as a symmetric critical point for F. $\diamondsuit$

4.2 Symmetric solutions of a semilinear elliptic PDE

We now consider using the method of Section 4.1 to derive the existence of Schwarz symmetric solutions for semilinear elliptic partial differential equations of the form

$$\begin{aligned} -\Delta x &= F'(x) & y &\in \Omega, \\ x &= 0 & y &\in \partial\Omega. \end{aligned} \tag{28}$$

Let X again be the Sobolev space $H^1_0(\Omega)$. Then solutions of (28) correspond to critical points of the functional

$$f(x) := \frac{1}{2}\|x\|^2 - \int_\Omega F(x(y))\, dy.$$

Example 42. Consider the case when $\Omega = \Omega^*$ and $F(x) = |x|^p$, where $2 < p < 2^* := (2N-2)/(N-2)$. Clearly, the Dirichlet problem (28) has the trivial Schwarz (S_7) symmetric solution $x(y) = 0$. We will show that there also exits a nontrivial Schwarz symmetric solution by analyzing the related action function defined on the Sobolev space $H^1_0(\Omega)$. We refer the readers to Adams' classical book [1] for properties of this Sobolev space. Since $2 < p < 2^* := (2N-2)/(N-2)$, $X = H^1_0(\Omega)$ is compactly imbedded in $L^p(\Omega)$, i.e., the imbedding $X \mapsto L^p(\Omega)$ maps bounded closed subsets of X to compact subsets of $L^p(\Omega)$. Thus, by the Sobolev inequality, $f(x) \geq r$ for some $r > 0$ on the unit sphere S_X of X. Clearly, $f(0) = 0$. Moreover, fixing $x = x^* \neq 0$ we have,

$$f(tx) = \frac{1}{2}t^2\|x\|^2 - t^p \int_\Omega |x|^p\, dy \to -\infty$$

as $t \to +\infty$. Thus, there exists $b = tx = b^*$ such that $f(b) \leq 0$.

It is tempting to apply the approximate symmetric mountain pass Theorem 40 with $g = *$ the Schwarz symmetrization. This does not work as it is known

[2] that, for a continuous path $u(t) \in \Gamma(a,b)$, $(u(t))^*$ may not be continuous. Using the universal approximation property of Theorem 32, van Schaftingen has provided a delicate version of the approximate symmetric mountain pass lemma that does work [42]. The following is a simplified version.

Theorem 43 (Approximate symmetric minimax)**.** *Let $X := H_0^1(\Omega)$ where $\Omega = \Omega^*$. Let $f\colon X \mapsto \mathbb{R}$ be a continuous and Gâteaux differentiable G_7-sub-invariant function. Suppose that $a^*, b^* \in X$. Define*

$$c := \inf_{x\in\Gamma(a^*,b^*)} \max_{t\in[0,1]} f(x(t)).$$

Suppose that S is a closed subset of X such that $S \subset \{x \in X \mid f(x) \geq c\}$ and S separates a^* and b^*. Then, there exist points $x_\varepsilon, z_\varepsilon^* \in X$ such that

(i) $\|x_\varepsilon - z_\varepsilon^*\|_{L^2(\Omega)} < \varepsilon$,

(ii) $c < f(x_\varepsilon) < c + \frac{5}{4}\varepsilon^2$,

(iii) $d(S; x_\varepsilon) < \frac{3}{2}\varepsilon$ *and*

(iv) $\|f'(x_\varepsilon)\| < \frac{3}{2}\varepsilon$.

Sketch of the proof. We need only to modify the proof of Theorem 40 slightly. For the pass $z(t)$ defined in the proof of Theorem 40, define $\tilde{z}(t) \in \Gamma(a^*,b^*) \cap G_7 \cdot z(t)$ such that $\|\tilde{z}(t) - (z(t))^*\|_{L^2(\Omega)} < \varepsilon/2$ by [42, Proposition 3.1]. Then use $\tilde{z}(t)$ to replace $gz(t)$ in the proof of Theorem 40. $\square$

We now continue our discussion of Example 42. For $a^* = 0$ and $b^* = tx^*$, applying Theorem 43 we can find sequences x_k, z_k such that $\|x_k - z_k^*\|_{L^2} \to 0$, $\lim_{k\to\infty} f(x_k) \geq r$ and $f'(x_k) \to 0$. It is well known that f satisfies the Palais–Smith condition (see e.g. [14, p.279]). Thus, without loss of generality we may assume x_k converges to $\bar{x}$ in H_0^1. It follows that z_k^* converges to $\bar{x}$ in L^2, and so $\bar{x} = \bar{x}^*$ is a Schwarz symmetric critical point of f. $\diamondsuit$

Remark 44. The symmetric minimax theorem in [42] is more general and also applies to several other symmetrizations that can be approximated by polarizations. Squassina [45] proposed related versions of variational principles that makes the proof of the above symmetric minimax theorem more efficient besides various other applications. $\diamondsuit$

Remark 45. We note that *symmetric mountain pass theorems* have been used in work [4, 27] among others in a different sense where symmetry of the action function is used to derive the existence of infinitely many (not necessarily symmetric) critical points. $\diamondsuit$

4.3 Principle of symmetric criticality

Palais [35] proposed a powerful principle of symmetric criticality. Here we provide a simplified version to illustrate the idea:

Theorem 46 (Principle of symmetric criticality)**.** *Let X be a Hilbert space with an isometric linear group action G and let $F \in C^1(X)$ be G-invariant. Denote*

$$\Sigma := \{x \in X : gx = x, \ \forall g \in G\}.$$

Then any critical point of $F|_\Sigma$ is also a critical point for F.

Proof. For any $g \in G$, and a vector v in the tangent space of Σ at $x \in \Sigma$: $T\Sigma|_x$, $F \circ g = F$ implies that

$$dF_x(v) = dF_{gx}(g(v)).$$

Since g is an isometry

$$\langle g\nabla F(x), g(v)\rangle = \langle \nabla F(x), v\rangle = dF_x(v).$$

On the other hand $gx = x$ implies

$$dF_{gx}(g(v)) = \langle \nabla F(gx), g(v)\rangle = \langle \nabla F(x), g(v)\rangle.$$

Thus,

$$\langle g\nabla F(x), g(v)\rangle = \langle \nabla F(x), g(v)\rangle.$$

or

$$g\nabla F(x) = \nabla F(x).$$

It follows that $\nabla F(x) \in T\Sigma|_x$ and x is a critical point of F. $\square$

Example 47. Consider the function $F(x, y) := x^2 - y^2$. The reflection

$$r(x, y) := (-x, y),$$

is a linear isometry with the invariant set

$$\Sigma := \{(0, y) : y \in R\}.$$

We can see that $F(x, y) := x^2 - y^2$ is invariant with respect to r and $(0, 0)$ is a critical point of $F(x, y)|_\Sigma = y^2$. Hence $(0, 0)$ is a critical point of F by Theorem 46. $\diamondsuit$

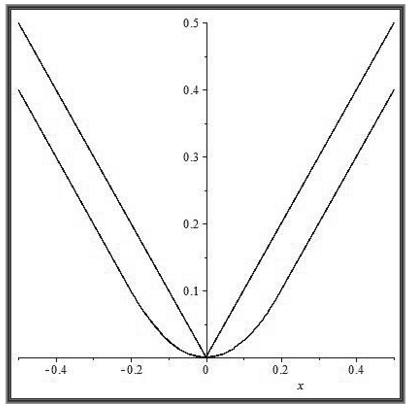

Figure 3: $e_{0.2}|x|$.

Remark 48. The behavior of $g(x, y) = |x| - |y|$ is similar to that of $F(x, y) = x^2 - y^2$. The Palais principle of symmetric criticality does not apply to g due to the lack of smoothness. This restriction seems to be largely a technical issue. In fact, the Moreau envelop (see [39, 40]) of the absolute function

$$e_\lambda|x| := \begin{cases} x - \frac{\lambda}{2} & x > \lambda \\ \frac{x^2}{2\lambda} & x \in [-\lambda, \lambda] \\ -x - \frac{\lambda}{2} & x < -\lambda \end{cases}$$

is a smooth function and approximates $|x|$ as $\lambda \to 0$ (see Fig. 3).

Thus, $g_\lambda(x, y) := e_\lambda|x| - e_\lambda|y|$ is a family of smooth functions that approximates $g(x, y)$. Moreover, we can check that g_λ inherits the symmetry of g. Thus, we can apply the Palais principle of symmetric criticality to g_λ and then taking limits as $\lambda \to 0$ to conclude that $(0, 0)$ is a generalized critical point of g in the sense that 0 belongs to the *derivative container* of Warga [48] at $(0, 0)$.

Whether such a scheme can be systematically applied to, say, the class of *difference convex* (or DC) functions [6] is an interesting question. $\diamondsuit$

Example 49 (Two body problem revisited)**.** Let us revisit the two body problem in which we wish to minimize the action function

$$F(x) := \int_0^P \left[\frac{\|x'(t)\|^2}{2} + \frac{1}{\|x(t)\|} \right] dt$$

in the space of periodic orbits $\{x \in H^1([0, P], R^2) : x(0) = x(P)\}$. We noted before that our invariance tools did not comport well. Let

$$G := \{\text{rotations around the origin}\}.$$

Then G is a group of isometries. It is easy to check that the Lagrange action function F is G-invariant. Thus, the Principle of symmetric criticality applies to the two body problem. In other words to find the critical point of F we need only to look for critical point of $F(x)$ on

$$\Sigma := \{x \in H^1([0,P],R^2) : x(0) = x(P), gx = x, g \in G\}$$

which is the set of all P-periodic H^1 cyclic trajectories. Thus, $x \in \Sigma$ has the form

$$x(t) := a\left(\cos\frac{2\pi t}{P}, \sin\frac{2\pi t}{P}\right),$$

where a is a parameter. Minimizing

$$F(a) = F(x) = \int_0^P \left[\frac{(2\pi a)^2}{2P^2} + \frac{1}{a}\right] dt$$

yields $a = (P/2\pi)^{2/3}$. Hence the orbit is

$$x(t) = \left(\frac{P}{2\pi}\right)^{\frac{2}{3}} \left(\cos\frac{2\pi t}{P}, \sin\frac{2\pi t}{P}\right),$$

as is well known [24]. $\diamondsuit$

5 Conclusion

Variational problems involving symmetry present new challenges that cannot be adequately addressed by traditional views of symmetry as invariance with respect to group actions. We have proposed a framework that expands the description of symmetry to include sub-invariance with respect to semigroup actions. As illustrated by the many examples discussed in this paper, the framework in Section 2 provides alternative perspectives on diverse variational problems involving symmetry in both finite and infinite dimensional spaces.

This exercise show us that the spirit of *Erlangen program* is very much applicable to variational problems involving symmetry: the key in understanding such a problem is to understand the underlying semigroup. Our effort here is only a first step in this direction. It certainly invites more questions then we have resolved. A few directly related to our discussion here are listed below as examples of challenges ahead.

1. The symmetric variational principles discussed in Section 2.3 represent only one possible way of extending variational principles to deal with problems involving symmetry. The emphasis in Section 2.3 is to make the

symmetric variational principles simple and general so that they have the possibility of application to many different situations. The downside is the loss of precision in dealing with any given problems. Different forms of symmetric variational principles were discussed in [45], motivated by and specifically designed for the polarization approximation of the Schwarz symmetry. There are of course many other possibilities. Where is the right point of compromise? Or do we intrinsically need many different types of symmetric variational principle?

2. It was shown in Example 49 that, using the Palais principle of symmetric criticality [35], the two body problem has a periodic solution. However, it is also clear that this periodic solution is, in fact, a minimum of the action function. It is natural to ask whether it is possible to use the framework in Section 2 to deal with it.

3. Many examples in Sections 3.1 and 3.2 show the close relationship between the symmetric variational principles and inequalities. Much can be done in this direction. Adapting the proof of the Muirhead inequality of Theorem 13 for other functions with symmetry is an interesting prospect. For example, whether this method can be helpful in dealing with permanents and the permanent inequality and extensions [18, 32] is a particularly interest question to explore.

4. Likewise we would like to see a nonsmooth extension of Palais principle of symmetric criticality, at least to continuous difference convex functions as discussed in Remark 47.

These are only a few of the many possible interesting directions for further research. We hope that they will bring about the attention of researchers to this interesting area and stimulate further research in this direction.

References

[1] Adams, R. A., *Sobolev spaces*, Pure and Applied Mathematics, **65**, New York-London: Academic Press, 1975.

[2] Almgren, F.J. Jr. and Lieb, E.H., The (non) continuity of symmetric ecreasing rearrangement (Cortona, 1988), Symposia Mathematica, Vol. XXX pp 89–102, Academic Press, London 1989.

[3] Ambrosetti, A., Variational methods and nonlinear problems: classical results and recent advances. In M. Matzeu and A. Vignolli, editors, Topological nonlinear analysis, Vol. 15 of Progr. Nonlinear Differential Equations Appl. pp. 1-36, Birkhauser, 1995.

[4] Ambrosetti, A. and Rabinowitz, P.H., Dual variational methods in critical point theory and applications. J. Functional Analysis, **14** (1973) 349–381.

[5] Aubin, J.-P. and Ekeland, I., *Applied Nonlinear Nnalysis*, John Wiley & Sons Inc., New York, 1984.

[6] Bačák, M and Borwein, J. M., On Difference Convexity of Locally Lipschitz Functions. *Optimization*, **60** (2011), 961–978.

[7] Baernstain II, A. and Taylor, B. A., Spherical rearrangements, subharmonic functions, and $*-$functions in n-space. Duke Math. J., **43** (1976), 245-268.

[8] Bhatia, R, *Matrix Analysis*, Springer-Verlag, New York, 1997.

[9] Borwein, J. M. & Lewis, A. S., *Convex Analysis and Nonlinear Optimization: Theory and Examples*, CMS Springer-Verlag Books, Springer-Verlag, New York, 2000.

[10] Borwein, J. M., Read, J., Lewis, A. S. and Zhu, Q. J. Convex spectral functions of compact operators, J. Nonlinear Convex Anal.**1** (2000) 17–35.

[11] Borwein, J. M., Lewis, A. S. and Zhu, Q. J. Convex spectral functions of compact operators, Part II: Lower semicontinuity and rearragement invariance, In Proceedings of the Optimization Miniconference VI, pp. 1-18, Kluwer Academic Publ. 2000.

[12] Borwein, J. M. and Preiss, D., A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, Trans. Amer. Math. Soc. **303** (1987) 517-527.

[13] Borwein, J. M. and Vanderwerff J. G., *Convex Functions: Constructions, Characterizations and Counterexamples*, Encyclopedia of Mathematics and Applications, **109**, Cambridge University Press, 2010.

[14] Borwein J. M. and Zhu Q. J., *Techniques of Variational Analysis*, CMS Springer-Verlag Books, Springer-Verlag, New York, 2005.

[15] Brock, F. & Solynin, A. Yu., An approach to symmetrization via polarizatization, Trans. AMS, **352** (1999) 1759–1796.

[16] Burchard, A., *A Short Course on Rearrangement Inequalities*, Lecture notes, IMDEA Winter School, Madrid, January 2009.

[17] van Brunt, B., *The Calculus of Variations*, Springer-Verlag, New York, 2004.

[18] Bullen, P. S., *A dictionary of inequalities*, Longmann, 1998.

[19] Clarke, F. H., Ledyaev, Yu. S., Stern, J. R. and Wolenski, R. P., *Nonsmooth Analysis and Control Theory*, Graduate Text in mathematics, Vol. 178. Springer-Verlag, New York, 1998.

[20] Crowe, J.A. and Zweibel, J. A. and Rosenbloom, P. C., Rearrangement of functions, J. Funct. Anal., **66** (1986) 432–438.

[21] Eggleston, H.G., *Convexity*, Cambridge University Press, Cambridge 1958.

[22] Ekeland I., On the variational principle, J. Math. Anal. Appl., **47** (1974), 324–353.

[23] Gergonne, J.D., Géométrie. Recherche de la surface plane de moindre surface, entre tous ceux de même volume, Annales de Gergonne, **4** (1813-1814) 338-343.

[24] Gordon, W.B., A minimizing property of keplerian orbits, American J. Math., **99** (1977), 961–971.

[25] Ioffe, A.D. and Levin, V. L. Subdifferentials of convex functions, Trudy Moskov. Mat. Obšč, **26** (1972), 3-73.

[26] Kadelburg, Z. Dukić, D., Lukić, M.& Matić, I. Inequalities of Karamata, Schur and Muirhead, and some other applications, The Teaching of Mathematics, **8** (2005), 31–45.

[27] Kajikiya, R. A critical point theorem related to the symmetric mountain pass lemma and its applications to elliptic equations, J. Functional Analysis, **225** (2005), 352-370.

[28] Karamata, J. Sur une inégalité relative aux fonctions convexes, Pulb. Math. Univ. Belgrade, **1** (1932), 145-148.

[29] Klein, F., 1872. "Vergleichende Betrachtungen ber neuere geometrische Forschungen" ('A comparative review of recent researches in geometry'), Mathematische Annalen, **43** (1893) 63–100.

[30] Ledyaev, Yu. S. & Zhu, Q. J. Nonsmooth analysis on smooth manifolds, Trans. Amer. Math. Soc., **359** (2007), 3687–3732.

[31] Lewis, A. S. Nonsmooth analysis of eigenvalues. Math. Program. **84** (1999) 1-24.

[32] Minc, H., *Permanents*, Encyclopedia of Mathematics and its Applications, Vol. 6, Addison-Wesley, New York, 1978.

[33] Mordukhovich, B. S., *Variational Analysis and Generalized Differentiation* I, II, Springer-Verlag, 2006.

[34] Muirhead, R.F. Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters, Proc. Edinburgh Math. Soc., **21** (1903), 144-157.

[35] Palais, R. S. The principle of symmetric criticality, Comm. Math. Physics, **69** (1979) 19–30.

[36] Palais, R. S. and Smale, S., A generalized Morse theory, Bull. Amer. Math. Soc., **70** (1964) 165-172.

[37] Piñeyro, P. J. H., Gergonne; the isoperimetric problem and the Steiner's symmetrization. to appear in The Amer. Math. Monthly.

[38] Rabinowitz, P. H., Critical points theory and applications to differential equations: a survey. In M. Matzeu and A. Vignolli, editors, Topological nonlinear analysis, Vol. 15 of Progr. Nonlinear Differential Equations Appl. pp. 464–513, Birkhauser, 1995.

[39] Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, N. J. 1970.

[40] Rockafellar, R. T. and Wets, R. J. B., *Variational Analysis*, Vol. 317 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1998.

[41] Van Schaftingen, J., Universal approximation of symmetrization by polarizatization, Proc. AMS, **134** (2005) 177–186.

[42] Van Schaftingen, J., Symmetrization and minimax principles, Comm. in Contemporary Mathematics, **7** (2005) 463–481.

[43] Schirotzek, W., *Nonsmooth Anlaysis*, Springer, 2007.

[44] Schur, I., Uber eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie, Sitzunsber. Berlin. Math. Ges. **22** (1923) 9-20.

[45] Squassina, M., Symmetry in variational principles and applications, to appear in Journal of London Math Soc.

[46] Steiner, J., Einfache bewiese der isoperimetrischen Hauptsätze, J. Reine Angew. Math. **18** (1938) 689-788.

[47] Valentine, F. A., *Convex Sets*, McGraw Hill, Boston, 1964

[48] Warga, J., Derivate containers, inverse functions, and controllability, in Calculus of Variations and Control Theory, D.L. Russell, Ed., Academic Press, New York, 1976.

[49] Weyl, H., *Symmetry*, Princeton University Press, 1952

[50] Wolontis, V., Properties of conformal invariants, Amer. J. Math. **74** (1952) 587-606.