Available online at www.sciencedirect.com
ScienceDirect
International Journal of Solids and Structures 44 (2007) 4723–4741
INTERNATIONAL JOURNAL OF SOLIDS and STRUCTURES
www.elsevier.com/locate/ijsolstr

Imperfect symmetry: A new approach to structural optima via group representation theory

P.L. Várkonyi, G. Domokos *

Department of Mechanics, Materials and Structures and Center for Applied Mathematics and Computational Physics, Budapest University of Technology and Economics, H-1111 Budapest, Muegyetem rkp 3, Hungary

Received 11 August 2006; received in revised form 15 November 2006
Available online 5 December 2006

Abstract

While *imperfect symmetry* (i.e., small deviation from symmetry) is abundant in Nature, it is not common among engineering structures. Here we investigate whether, and under which conditions the response of engineering structures with given symmetry group Γ may be *improved* by adding small perturbations x_i to the geometry. We will prove that a naturally emerging representation of Γ in the space of the variables $\{x_i\}$ plays a key role and, based on this representation, we formulate exact conditions of improvability, utilizing classical representation theory of finite groups. We also present various examples among which *optimal* structures with *imperfect symmetry* emerge, somewhat counter to the engineer's intuition.

Keywords: Imperfect symmetry; Optimisation; Representation theory

1. Introduction

Engineering structures are often built with some degree of symmetry, in fact, it is hard to find structures *without* local or global symmetry. The concept of symmetry not only simplifies the design and construction process; there is a deep-rooted belief among structural engineers that symmetrical geometry represents an optimum, at least locally. We will show that this intuition is not unfounded. It is remarkable that an extensive study on shape optimization (Sokolowski and Zolesio, 1992) lists exclusively examples with reflection symmetry. Nevertheless, bold designers abandon this concept once in a while (Tzonis, 1999), resulting in structures with fundamental asymmetry. Structures with *slight* asymmetry seem to appear as less desirable, "imperfect" solutions to engineering problems.

In contrast, Nature produces creatures with *imperfect symmetry* (i.e., slight asymmetry) in abundant quantities: the human body itself, while obeying a fundamental body plan of planar reflection symmetry, displays a remarkable collection of small asymmetries, ranging from the heart's location to the functional neural wiring

* Corresponding author. Fax: +3614631773.
E-mail addresses: vpeter@mit.bme.hu (P.L. Várkonyi), domokos@iit.bme.hu (G. Domokos).

0020-7683/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijsolstr.2006.11.047

of the brain. These asymmetries have all emerged in the process of evolution, seeking competitive optima (cf. Várkonyi et al., 2006).

Whether or to what extent slight asymmetry should be treated as *imperfect* (as opposed to *perfect* symmetry) is a philosophical question beyond the scope of this paper. As Weyl (1952) notes, Western art, similar to Nature, is inclined to relax rigorous symmetry. Nevertheless, it is rare that asymmetry means merely the lack of symmetry: asymmetric compositions are often *very close* to some symmetric pattern (cf. Fig. 1). Undoubtedly, both artistic and human beauty is often associated with *slight asymmetry* so it should not be viewed as inferior to intact, perfect symmetry.

Based on the appearance of slightly asymmetric optima in Nature one could hope to find their analogies in structural design as optimal solutions; this is one goal of the current paper. On the other hand, appearance of slight asymmetry in western art suggests that slightly asymmetrical structures may even be visually attractive. Nature produces slightly asymmetric optima in the process of evolution, i.e., time appears as an essential parameter (Várkonyi et al., 2006). In the structural analogy the study of *individual structures* may only reveal whether the given, symmetric configuration is locally a 'pessimum' or optimum, i.e., *whether or not it can be improved via small perturbations*. This is the question, which we will address in full generality. In order to locate *slightly asymmetric optima* we will follow the analogy to evolution and study *one-parameter families* of structures in search for *bifurcations of structural optima*; this rather complex issue will be addressed only at specific points in the paper.

Existence of structural optima with imperfect symmetry (i.e., bifurcations of structural optima) has been reported in Várkonyi (2006) for structures with reflection symmetry. However, Várkonyi (2006) predicts that if only *one* symmetry-breaking geometric variable (x_1) is admitted, the perfect configuration ($x_1 = 0$) is typically optimal, at the same time, structural optima with imperfect symmetry ($|x_1| \ll 1$) are atypical in a wide class of engineering optimization problems (see 'local' optimization criteria in Section 2). This confirms the engineer's fundamental intuition. Nevertheless, examples with different symmetry groups and higher number of variables suggest that this may not be always the case, so it is natural to ask the following, generalized questions:

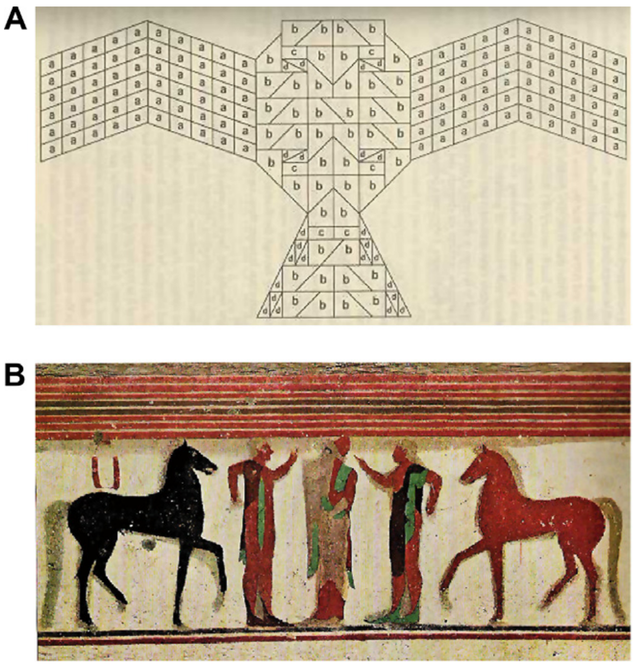

Fig. 1. (A) Falcon-shaped Vedic sacrificial altar made of brick (Joseph, 2000). (B) Wall-painting of the Etruscan 'Tomb of the Baron', Tarquinia.

1. What are exactly the conditions for the local optimality of structures with arbitrary *(finite) symmetry groups* Γ *and arbitrary number of symmetry-breaking variables* x_i?
2. What are exactly the conditions for the existence of *slightly asymmetric optima (i.e., bifurcations of optima)* in a one-parameter family of structures with *arbitrary (finite) symmetry group* Γ *and arbitrary number of symmetry-breaking variables* x_i?

In this paper we address primarily the first question and provide rigorous conditions for *typical* cases. The answer depends essentially on the *representation of the group* Γ *in the space of the variables* $\{x_i\}$, and classical group representation theory provides easy tools to determine the properties of this representation. Group-theory based analysis does not offer a full answer; however, by using this method one can quickly exclude the bulk of the cases where the structure may *not* be improved with the given variables. If improvability cannot be *excluded* based alone on group theory arguments, the structure will be called 'potentially improvable'. To verify *actual* improvability in such cases, further, more cumbersome structural analysis is required.

We will discuss the second question only partially in Section 6, in connection with the structural examples. Our results demonstrate that the relation of the answers to the first and to the second question is far from trivial; however, a general treatment of the second question is beyond the scope of this paper.

Section 2 introduces the basic concepts on a simple reflection-symmetric structure. We point out fundamental differences between *local* and *global* optimum criteria and between *one* and *two* symmetry-breaking variables. Section 3 discusses our main question on improvability in full generality, assuming an arbitrary, finite symmetry group and arbitrary number of perturbing variables. Section 4 illustrates the results on simple examples, Section 5 outlines exceptional cases, and results are summarized in Section 6. It is also discussed in this part how the main questions of the paper are related to the design of structural optima with imperfect symmetry. Finally, Appendices A.1–A.4 contain the proofs of several lemmas and theorems.

2. Structural optima and symmetry

We consider the simplest kind of optimization problem: a scalar "optimization" potential $U(p,x)$ will be associated with the structure and we seek local minima of U as optimal structural configurations. The variable x will refer to the *deviation* from the symmetric configuration, i.e., $x=0$ will be always associated with the symmetric problem. The parameter p will describe a *family* of structures, each of which possesses the same symmetry at $x=0$. As final goal, we seek *optimum-bifurcation diagrams* in the $[x,p]$ plane, describing how optima evolve as the parameter p is varied.

Fig. 2 presents a planar three-hinged structure subjected to vertical loading N at the internal hinge C, the length of the bars is denoted by l_i, the area of the cross sections by A_i, the bending stiffnesses by $\mathrm{EI}_i (i=1,2)$. The *horizontal* location of C is identified by the variable x, the *vertical* location of C by the *parameter* p. After obtaining the internal forces N_1, N_2 and the critical (Euler) loads N_1^{cr}, N_2^{cr}, the *risk against buckling* can be calculated in the individual bars as $f_1 = N_1/N_1^{\mathrm{cr}}$, $f_2 = N_2/N_2^{\mathrm{cr}}$, in more detail:

$$f_i(p,\mathbf{x}) = N_i(p,\mathbf{x})/N_i^{\mathrm{cr}}(p,\mathbf{x}) = \frac{N_i(p,\mathbf{x})l_i^2(p,\mathbf{x})}{EI_i\pi^2}, \tag{1}$$

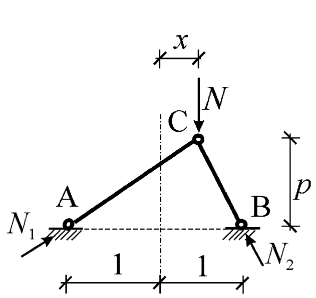

Fig. 2. A simple three-hinged model loaded by the concentrated force N.

We will now introduce the concept of *global* and *local* optimum criteria.

First we intend to *minimize the total mass*, i.e., we have $U(p,x) = A_1 l_1 + A_2 l_2$ with equal, given $f_1 = f_2 = r$ risk of buckling in each bar (i.e., $A_1 \neq A_2$ is determined according to this constraint). The resulting optimum-bifurcation diagram (Fig. 3A) is not surprising. Since U is smooth and symmetrical ($U(p,x) = U(p,-x)$), $x = 0$ is either at a local minimum or maximum of U, as predicted by elementary catastrophe theory (Poston and Stewart, 1978), and the type of $x = 0$ changes typically at pitchfork bifurcations.

Now we modify our potential to obtain an example for *local* optimum criteria.

Engineers often prefer to apply identical structural elements. By adopting this concept we now prescribe $A_1 = A_2$, $I_1 = I_2$ resulting in $f_1 \neq f_2$. The task will be to *minimize the risk of buckling*, i.e., we consider the *highest risk* among the bars. The potential is defined as $U(p,x) = \max\{f_1(p,x), f_2(p,x)\}$. We call this criterion "local" because it considers the local, *weak* points (in this case, bars) of the structure individually; this approach is perhaps more natural for engineers than the global criterion. The optimum-bifurcation diagram features an unusual 'X-bifurcation' (Fig. 3B). Notice that $x = 0$ remains a local optimum on both sides of the bifurcation point. The reason of the unexpected results is the special type of potential U. Observe that $f_1(p,x) = f_2(p,-x)$, and $x = 0$ is typically a *non-smooth* local optimum (Fig. 4), which does not vanish at bifurcation points. The bifurcation patterns of similar optimum diagrams have been studied in Várkonyi and Domokos (2006). It has also been demonstrated there that one-parameter families of such examples typically do not contain critical points where symmetrical optima bifurcate, i.e., *slightly asymmetrical optima do not exist in typical cases*. The aim of the present paper is to generalize these results to *arbitrary* number of symmetry-breaking variables x_i (instead of one) and *arbitrary, finite symmetry groups* (instead of planar reflection symmetry).

Before we enter the general and rigorous discussion, we would like to illustrate that increasing the number of variables *by one* already has radical effects. The previous variable will be denoted by $x = x_1$ and we introduce x_2 as half of the *vertical distance* between the support hinges. As before, $x_1 = x_2 = 0$ refers to the

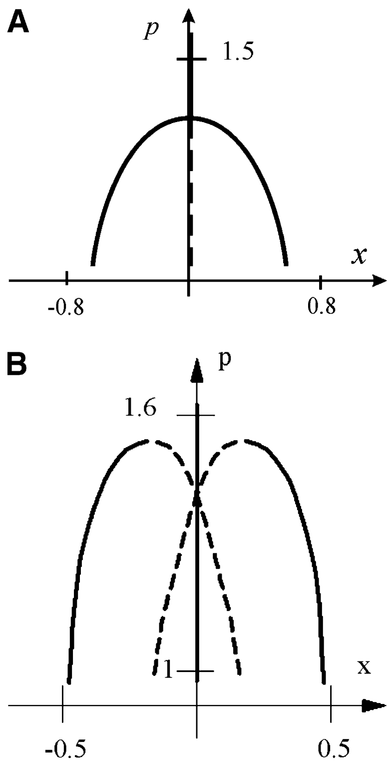

Fig. 3. (A) Optimization diagram of total mass with prescribed safety against buckling in both bars. (B) Optimization diagram of safety against buckling if the cross sections are equal and prescribed (continuous line: optimum, dashed line: pessimum).

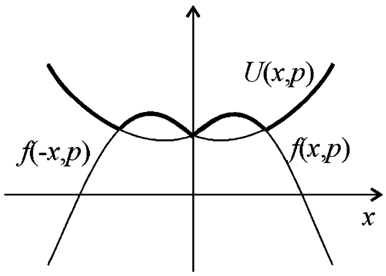

Fig. 4. The potential $U(x,p)$ at $p = 0.5$, generated from (1).

symmetric configuration (Fig. 5A). The only modification needed in the previous arguments is to replace x by the vector $\mathbf{x} = [x_1 x_2]^T$, however, this minor modification results in radical change. The potential U is again the *maximal risk among individual members* criterion. In Fig. 5B and C we can observe that $\mathbf{x} = 0$ need not be a local optimum any more, and the bifurcation patterns of the optimum diagram (Fig. 6) become similar to those of smooth functions (such as Fig. 3A). Still, similar to the single-variable case, a *typical* random perturbation increases the potential (i.e., makes the structure worse), however a *special, adequate* one might improve it (Fig. 5D). Thus, the 2-variable case suggests that improvement of symmetrical structures by small perturbations is *difficult, however possible* in case of 'local' optimization criteria. With other words, spontaneous imperfection (e.g., due to errors during the construction) has typically a bad effect, however, calculated geometrical imperfectness of a symmetrical structure may – under some circumstances – be of advantage. In the next section we proceed to define these circumstances rigorously for arbitrary number of variables and arbitrary, finite symmetry groups.

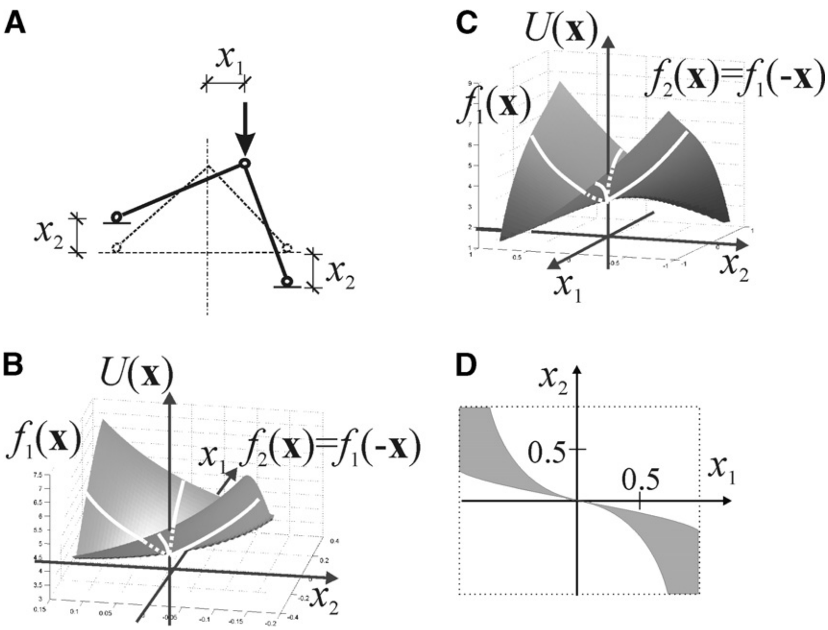

Fig. 5. (A) A simple three-hinged model with two perturbing variables. (B) Optimization potential of the structure if $p = 0.15$. The symmetrical configuration is partially smooth optimum. (C) Optimization potential if $p = 2$. The symmetrical configuration is a partially smooth saddle, i.e., not optimum. Notice in (B) and (C) that optimization with only one of the variables (white sections of the surfaces) would result in non-smooth optima. (D) The grey domain indicates values of x, for which $U(p,x) < U(p,0)$ if $p = 2$. Notice that a randomly chosen small ($|x| \ll 1$) value of x is typically out of this range. Thus, a small, random perturbation of the symmetry typically spoils the structure.

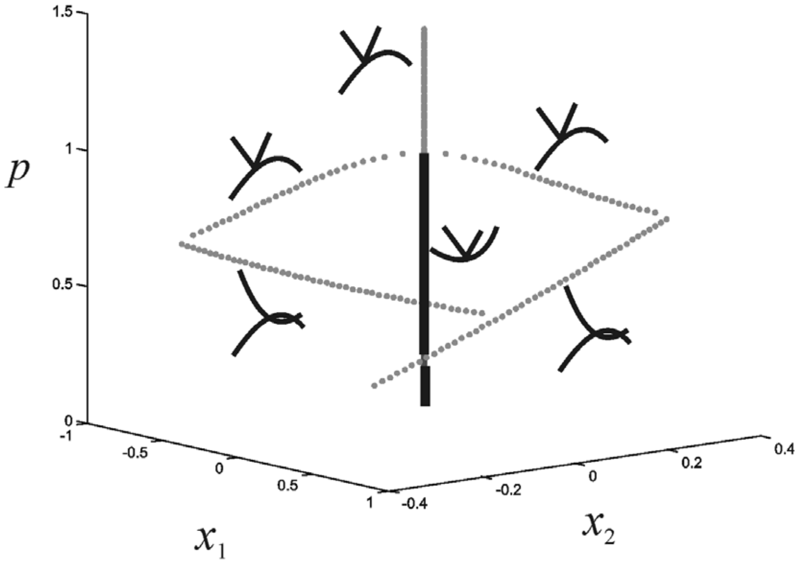

Fig. 6. Optimum diagram of the example of Fig. 5A. Thick, black lines denote local minima, thin grey lines denote maxima, and saddle points of the function $U(x_1, x_2)$. Small pictograms indicate the local shape of U (partially smooth or smooth) at all branches.

3. The general problem

3.1. Definition

In this section, we discuss the natural generalization (arbitrary symmetry group Γ, d variables) of our previous observations on optimisation problems with *local* optimum criteria. The corresponding results of Section 2 suggest that symmetrical structural forms tend to be local optima. If the number of variables is low, then the symmetrical configuration corresponds to a non-smooth, "robust" optimum – in this case there is little hope to improve the structure by local perturbations. However, if the number of variables is sufficiently high then *smooth submanifolds* will pass through the point associated with the symmetric structure. In the subspace of these manifolds one can hope to improve the symmetric configuration, however actual improvability depends on nonlinear terms in f_i (Fig. 5B vs. C). In this section we take a rigorous approach to find sufficient, necessary, as well as sufficient *and* necessary conditions for the existence of the smooth submanifolds, i.e., for the *potential improvability* of the symmetric structure. The goal of these criteria is to determine *without detailed example-specific analysis* the number and type of perturbing variables, which make a structure potentially improvable.

Let $\Gamma \equiv \{\gamma_i\, i = 1, 2, \ldots, r\}$ denote the symmetry group associated with the perfect structure (r is the order of Γ). In practical engineering problems, Γ is usually a cyclic or dihedral group. It is required that the loads, the internal forces, and, in fact, any external condition, which has an effect on the optimization process, support the 'Γ-symmetry' of the structure.

The symmetry of the structure will be perturbed via symmetry-breaking variables collected in vector $\mathbf{x} = [x_1\ x_2 \cdots x_d]^{\mathrm{T}}$, $\mathbf{x} \in \mathbf{R}^d$. These variables represent the set of structures, which are considered as candidate solutions of the optimization problem. The structure corresponding to $\mathbf{x} = \mathbf{x}_0$ is referred to as $\mathbf{S}(\mathbf{x}_0)$. We apply the following three restrictions on $\mathbf{x}$:

(i) $\mathbf{S}(\mathbf{x})$ is Γ-invariant if and only if $\mathbf{x} = 0$.
(ii) the set $\{\mathbf{S}(\mathbf{x}), \mathbf{x} \in \mathbf{R}^d\}$ is Γ-invariant.
(iii) $\mathbf{S}(\mathbf{x}_1) \equiv \mathbf{S}(\mathbf{x}_2)$ if and only if $\mathbf{x}_1 = \mathbf{x}_2$.

Condition (i) is a natural consequence of the fact that our investigation relies on the optimality of the Γ-symmetrical configuration $\mathbf{x} = 0$ *compared to non-symmetrical ones*, i.e., disturbed configurations should not be Γ-symmetrical. According to condition (ii), if an asymmetrical configuration $\mathbf{S}(\mathbf{x})$ is a potential solution then the transformed configuration $\gamma_i(\mathbf{S}(\mathbf{x}))$, $\gamma_i \in \Gamma$ is also potential solution. Since the set of possible solutions is limited primarily by external conditions, which should not break the Γ-symmetry, (ii) is a *natural symmetry condition*. Finally condition (iii) is purely technical: it states that the symmetry-breaking variables are 'independent' of each other in the sense that non-identical perturbations yield indeed non-identical structures.

According to conditions (ii), (iii) there are $d_i : \mathbf{R}^d \to \mathbf{R}^d i = 1, 2, \ldots, r$ transformations satisfying $\gamma_i(\mathbf{S}(\mathbf{x})) = \mathbf{S}(d_i(\mathbf{x}))$. We restrict ourselves to the case when d_i are *linear transformations*, (which includes all examples of practical interest according to the authors' experience). Since $\gamma_i(\mathbf{S}(0)) = \mathbf{S}(d_i(0)) = \mathbf{S}(0)$ yields $d_i(0) = 0$ by condition (iii), the symmetry transformations now correspond to simple matrix multiplications in the space of the variables:

$$\gamma_i(\mathbf{S}(\mathbf{x})) = \mathbf{S}(\mathbf{D}_i\mathbf{x}), \tag{2}$$

Due to Eq. (2) and condition (iii), the set of matrices $D \equiv \{\mathbf{D}_i i = 1, 2, \ldots, r\}$ mimic the group structure of Γ: if $\gamma_i, \gamma_j, \gamma_k \in \Gamma$ satisfy $\gamma_i\gamma_j = \gamma_k$, we also have $\mathbf{D}_i\mathbf{D}_j = \mathbf{D}_k$. Thus, D is a *linear representation* of group Γ (Jones, 1998); it will be referred to in this paper as the 'induced representation' of Γ' (Note that the concept of induced representations is used in group representation theory in a different sense, see e.g., Barut and Raczka (1986). It is also worth mentioning that Γ has another representation in the *physical space* of the structure, since the elements γ_i correspond to matrix transformations in an adequate physical coordinate system. The latter representation of Γ will not gain importance during the following investigations).

In case of local optimization criteria, the potential of the structures is of the form

$$U(\mathbf{x}) = \max_i f_i(\mathbf{x}) \quad i = 1, 2, \ldots, k, \tag{3}$$

where the functions $f_i(\mathbf{x})$ are local goodness measures corresponding to 'weak' elements/points of the structure. We are interested in local properties of $U(\mathbf{x})$ at $\mathbf{x} = 0$, thus we only need to consider the weakest points of the perfect configuration $\mathbf{S}(0)$, i.e., those $f_i(\mathbf{x})$ functions for which $f_i(0) = U(0)$. (At the same time, $\mathbf{S}(0)$ has usually more than one 'weakest' points due to its symmetry.) The functions $f_i(\mathbf{x})$ are supposed to be analytic, which allows approximating them via linearization: $f_i(\mathbf{x}) = U(\mathbf{0}) + \mathbf{g}_i^{\mathrm{T}}\mathbf{x} + \mathrm{o}(|\mathbf{x}|^2)$. Thus, Eq. (3) yields

$$U(\mathbf{x}) = U(\mathbf{0}) + \max_i(\mathbf{g}_i^{\mathrm{T}}\mathbf{x}) + \mathrm{o}(|\mathbf{x}|^2) \quad i = 1, 2, \ldots, k \quad \text{if } |\mathbf{x}| \ll 1. \tag{4}$$

We have already seen this type of potential in the first example, where $x = 0$ was not only a local optimum, but it was a 'robust' one, i.e., for $|x| \ll 1$, we had $U(x) - U(0) \approx c|x|$($c > 0$ is a constant), while at smooth optima we would have typically $U(x) - U(0) \approx c|x|^2$. This kind of non-smooth optimum is a characteristic property of similar examples. Before going into details, we give an exact definition of robust optima, which applies for problems with arbitrary number of variables:

Definition 1. The point $\mathbf{x} = 0$ is a robust local optimum (or minimum) of the scalar function U($\mathbf{x}$), $\mathbf{x} \in \mathbf{R}^n$ if there exist real scalars $\delta, \varepsilon > 0$ such that $|\mathbf{x}| < \delta$ yields $U(\mathbf{x}) - U(0) \geqslant \varepsilon \cdot |\mathbf{x}|$. ($|\mathbf{x}|$ denotes the l_2-norm of the vector $\mathbf{x}$.)

Based on Definition 1, functions of type (3) can be classified according to the following, simple scheme:

(A) $\mathbf{x} = 0$ is a robust minimum of $U(\mathbf{x})$. In this case $\mathbf{S}(0)$ is *not improvable* via small perturbations.
(B) $\mathbf{x} = 0$ is a singular point (minimum saddle or maximum) of $U(\mathbf{x})$, however it is not a robust minimum. In this case, $\mathbf{S}(0)$ is called '*potentially improvable*', because we have two possibilities according to *nonlinear terms* of the generating $f_i(\mathbf{x})$ functions:
   (B1) $\mathbf{x} = 0$ is local minimum: $\mathbf{S}(0)$ *cannot* be improved via small perturbations.
   (B2) $\mathbf{x} = 0$ is not local minimum: $\mathbf{S}(0)$ *can* be improved via small perturbations.
(C) $\mathbf{x} = 0$ is not singular point of $U(\mathbf{x})$. In this case, $\mathbf{S}(0)$ is improvable via small perturbations.

As we will show, $U(\mathbf{x})$ cannot be of type C if conditions (i)–(iii) are satisfied (Lemma 1). At the same time, one can decide whether $\mathbf{S}(0)$ belongs to (A) or (B) *without* computing structural behaviour, solely based on the symmetry group Γ and the variables x_i. Our goal is to describe this algorithm (Section 3.2) together with some interesting additional results (Section 3.3) and also, to formulate sufficient as well as necessary criteria for $\mathbf{S}(0)$ belonging either to (A) or to (B) (Section 3.4).

If $\mathbf{S}(0)$ belongs to (B), the question of *actual* improvability (class B1 vs. B2) depends on nonlinear terms of the functions $f_i(\mathbf{x})$, and finding the answer requires detailed computations on the structure under investigation. Although we will provide specific examples of such computations in Section 4, we can not give any general method to distinguish between structures in (B1) and (B2), so we call all structures in category (B) "*potentially improvable*".

3.2. Exact conditions of potential improvability

In this section, conditions of potential improvability of $\mathbf{S}(0)$ are derived. After presenting a general, however, somewhat cumbersome condition in Section 3.2.1, we develop an easy-to-check condition for *typical* cases in Section 3.2.2.

Throughout the rest of the paper, we use the following notations (in accordance with previous ones): vector $\mathbf{x}$ for the perturbing variables and d for the number of variables, $\mathbf{S}(\mathbf{x})$ for the corresponding structures, $\Gamma \equiv \{\gamma_1, \gamma_2, \ldots, \gamma_r\}$ for the symmetry group of $\mathbf{S}(0)$, r for the order of Γ, $D \equiv \{\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_r\}$ for the elements of the induced representation, $P_i (i = 1, 2, \ldots, k)$ for the weakest points of $\mathbf{S}(0)$, $f_i(\mathbf{x})$ for the potentials associated with P_i and finally $\mathbf{g}_i = \text{grad } f_i(\mathbf{x})|_{\mathbf{x}=0}$.

3.2.1. Example-specific classification of problems

Considering vectors $\mathbf{g}_i$ in R^d, we can characterize the local configuration of U at $\mathbf{x} = 0$ as stated in

Theorem 1. *Based on the classification of Definition* 3.1, $\mathbf{S}(0)$ *belongs to*
a: *class 'A' iff the point* $\mathbf{x} = 0$ *is inside the convex hull of the endpoints of vectors* $\mathbf{g}_i$,
b: *class 'B' iff it is at the border of the convex hull and*
c: *class 'C' iff it is outside the convex hull.*

(For the proof see Appendix A.1). To apply Theorem 1, all gradients $\mathbf{g}_i$ have to be computed via structural analysis.

We can improve Theorem 1 since the gradients $\mathbf{g}_i$ are not independent of each other. Pick one arbitrary weak point (P_1) of $\mathbf{S}(0)$ (such as the left bar of the example of Fig. 2.). Due to the Γ-symmetry of $\mathbf{S}(0)$, each point $\gamma_i(P_1)$ is associated with the same potential as P_1 if $\mathbf{x} = 0$ (in the introductory example $\gamma_1(P_1) = P_1$ and $\gamma_2(P_1) = P_2$ are the left and the right bar, respectively). It is possible that some additional points have the same potential, however such a coincidence can be considered as atypical (unless the potential of the structure possesses some hidden constraints, which possibility is discussed in Section 5). The local potentials of the weak points $\gamma_i(P_1) = P_i$ are $f_i(\mathbf{x}) = f_1(\mathbf{D}_i\mathbf{x})$, yielding $\mathbf{g}_i = \text{grad } (f_1(\mathbf{D}_i\mathbf{x}))|_{x=0} = \mathbf{D}_i^{\mathrm{T}}$ grad $(f_1(\mathbf{x}))|_{x=0} = \mathbf{D}_i^{\mathrm{T}}\mathbf{g}_1$. Thus, the resultant potential $U(\mathbf{x})$ of the structure is typically of the form (cf. Eq. (4)):

$$U(\mathbf{x}) = U(0) + \max_i((\mathbf{D}_i^{\mathrm{T}}\mathbf{g}_1)^{\mathrm{T}}\mathbf{x}) + \mathrm{o}(|\mathbf{x}|^2) = U(0) + \max_i\left((\mathbf{D}_i^{\mathrm{T}}\mathbf{g}_1)^{\mathrm{T}}\bar{\mathbf{x}}\right) \cdot |\mathbf{x}| + \mathrm{o}(|\mathbf{x}|^2) \tag{5}$$

where $\bar{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$. In Eq. (5), '>' might emerge instead of '=' in atypical cases.

Based on Eq. (5), we derive *typical* conditions for the classification of optimization problems, which are simpler to check than Theorem 1. The set of vectors $\mathbf{D}_i^{\mathrm{T}}\mathbf{g}_1$ in (5) is called the orbit of $\mathbf{g}_1$, with respect to the representation $D^{\mathrm{T}} \equiv \{\mathbf{D}_i^{\mathrm{T}}\}$, which we denote by $\mathrm{orb}_{D^{\mathrm{T}}}\mathbf{g}_1$. Let us introduce the following concepts:

Definition 2. An n-dimensional representation R is 'cyclic' if there exits a vector $\mathbf{v} \in R^n$ so that $\dim(\mathrm{orb}_R\,\mathbf{v}) = n$.

Definition 3. If R is an n-dimensional (cyclic) representation, $\mathbf{v} \in R^n$, and $\dim(\mathrm{orb}_R\mathbf{v}) = n$, $\mathbf{v}$ is 'cyclic vector' of R.

The above definitions are used mostly in the theory of the representations of inifinite groups (Barut and Raczka, 1986), where properties of complicated representations are studied via their decompositions to the direct sum of *cyclic representations*. At the same time, representations D of finite groups Γ can be further decomposed:

an adequate similarity transformation applied on all elements of a D transforms them to block-diagonal matrices (each with the same block-sizes), in which all blocks correspond to lower dimensional representations of Γ, called *subrepresentations* of D. Representations, which cannot be further decomposed are called *irreducible*. The representation theory of finite groups strongly relies on the fact that decomposition to the direct sum of irreducible representations is unique, thus it is an effective tool of analyzing complex representations. Due to the above reason, the notion of cyclic representations is not widely used when finite groups are considered. However, in this paper the notion of cyclic representations/vectors becomes essential by

Lemma 1. *$x = 0$ is a robust, local optimum* (i.e., $\mathbf{S}(0)$ *is in class* A) *iff* g_1 *is a cyclic vector of* D^{T}. *Otherwise it is typically potentially improvable* (*type* B).

Proof of Lemma 1. Let $\mathbf{D}_k$ be an arbitrary element of D. The sum of the orbit of $\mathbf{g}_1$ is invariant to multiplication by $\mathbf{D}_k^T$, because the matrices $\mathbf{D}_i\mathbf{D}_k$, $i = 1, 2, \ldots, r$ are a permutation of $\mathbf{D}_i$, $i = 1, 2, \ldots, r$, i.e.,

$$\mathbf{D}_k^{\mathrm{T}} \sum_{i=1}^{r} \mathbf{D}_i^{\mathrm{T}} \mathbf{g}_1 = \sum_{i=1}^{r} (\mathbf{D}_i \mathbf{D}_k)^{\mathrm{T}} \mathbf{g}_1 = \sum_{i=1}^{r} \mathbf{D}_i^{\mathrm{T}} \mathbf{g}_1. \tag{6}$$

Eq. (6) yields

$$\sum_{i=1}^{r} \mathbf{D}_i^{\mathrm{T}} \mathbf{g}_1 = 0, \tag{7}$$

since $\mathbf{x} = 0$ is the only invariant point of D^{T} by condition (i) and Eq. (2). According to (7), $\mathbf{x} = 0$ is a convex combination of vectors $\mathbf{D}_i^{\mathrm{T}} \mathbf{g}_1$ (all coefficients are $1/r$), i.e., it is inside their convex hull, unless the convex hull is degenerate. The latter case occurs if $\mathbf{g}_1$ is not a cyclic vector of D^{T}. The convex hull has then only borders, and $\mathbf{x} = 0$ is one of them. Thus, Theorem 1 yields Lemma 1. □

Notice that the second part of Lemma 1 is only *typically* true. The origin of this fact has been pointed out after Eq. (5). Furthermore, we remark that Lemma 1 states the emptiness of class C as a consequence of condition (i). One could, however, analyze problems violating condition (i). In that case, (7) does not follow from (6), and such problems would belong typically to class C.

Now, in order to classify the problem, we only have to decide whether $\dim(\mathrm{orb}_{D^{\mathrm{T}}} \mathbf{g}_1) = d$ or not. This property depends both on D and on $\mathbf{g}_1$. The former one represents the type of the symmetry-breaking variables, while the latter one is completely example-specific. In the forthcoming section, we derive typical conditions for $\dim(\mathrm{orb}_{D^{\mathrm{T}}} \mathbf{g}_1) = d$, which rely *only* on D. These conditions can be applied even if the specific form of U is unknown.

3.2.2. General classification of problems

No we improve Lemma 1 in such a way that $\mathbf{g}_1$ will no more be needed as input data. In fact, whether $\mathbf{g}_1$ is cyclic vector of D or not, depends primarily on D, and only secondarily on $\mathbf{g}_1$. Clearly, if D is not cyclic, it has no cyclic vectors by Definition 2, however if D is cyclic, an arbitrary vector is not necessarily cyclic vector of D. (The zero vector is a trivial example of non-cyclic vectors.). However

Lemma 2. *If D is cyclic, a randomly chosen, non-zero vector is typically cyclic vector of D.*

See the proof in Appendix A.2. According to Lemma 2 it is sufficient to decide whether D is cyclic or not, if we want to determine the potential improvability of $\mathbf{S}(\mathbf{0})$ in *typical* cases. Thus we have derived a 'typical' condition, which is not example-specific.

Moreover, determining whether a representation of a finite group is cyclic or not, is a simple task by using

Lemma 3. *A representation of a finite group Γ is cyclic if and only if it is subrepresentation of the regular representation of Γ.*

See the proof in Appendix A.3. We mention that the *regular representation* (R_Γ) *of group* Γ is a representation of special interest. Its decomposition contains d_i examples of each irreducible representation I_i of Γ, where $d_i = \dim(I_i)$. Moreover, $\dim(R_\Gamma)$ is equal to the order of (the number of elements in) Γ according to the Dimensionality theorem (Jones, 1998).

Now, we can summarize Lemmas 1–3 in the main result of the paper:

Theorem 2. $\mathbf{S}(0)$ is typically not potentially improvable if D is a subrepresentation of the regular representation of Γ; otherwise it is typically potentially improvable.

Deciding whether D^{T} (or, equivalently, D) is a subrepresentation of R_Γ or not, is a simple standard task, the solution technique can be found in textbooks of representation theory (Jones, 1998). It has the following steps:

(1) Determining the 'character' of D, i.e., the r-dimensional vector $\mathrm{char}(D) = [\mathrm{trace}(\mathbf{D}_j) j = 1, 2, \ldots, r]^{\mathrm{T}}$.
(2) Determining the list I_i $i = 1, 2, \ldots, l$ of the irreducible representations of Γ and the character of each of them.
(3) Solving the

$$\mathrm{char}(D) = \sum (m_i \mathrm{char}(I_i)) \tag{8}$$

linear vector equation. It can be shown that the solution always exists and it is unique, furthermore, $m_i \in N$ (non-negative integers). The constants m_i represent the number of I_i components in D.
(4) As mentioned after Lemma 3, the number of I_i components in R_Γ is $\dim(I_i)$, thus D is a subrepresentation of R_Γ iff $m_i \leqslant \dim(I_i)$ for every i.

The main step of the process is the solution of a simple system of linear equations, provided that I_i are known. In case of simple groups (such as the symmetry groups associated with engineering structures), the irreducible representations I_i are known, they can be found in text books.

3.3. Further properties of optimization problems

Here we show two further results on the optimization problems, which are based on the type of D and follow from results of representation theory.

3.3.1. Verification of condition (i)

Conditions (ii) and (iii) must be checked at the beginning of the analysis (otherwise D makes no sense). At the same time condition (i) need not be satisfied to perform the analysis. It has already been mentioned that such problems typically belong to class C. Now, we show how the violation of condition (i) is indicated by the results of the analysis of the induced representation.

Each group has a one-dimensional *trivial representation* I_1, in which all group elements are represented by 1 (one-dimensional unit matrix). Clearly, D may not be the trivial representation by condition (i), because all points are invariant points of the trivial representation. Furthermore

Lemma 4. *Condition* (*i*) *is satisfied iff D* (*or the equivalent representation* D^{T}) *has no trivial component*.

See the proof in Appendix A.4. According to Lemma 4, the process of deciding whether D is cyclic also indicates whether condition (i) is met or not.

3.3.2. Irreducible representations

As already mentioned, Lemma 2 and the consequent results are typical but not exact. One main reason of this restriction is that cyclic representations have non-cyclic vectors (e.g., the 0 vector). However, we have

Lemma 5. *If D is irreducible*, *all non-zero vectors are cyclic vectors of D*.

The proof of Lemma 6 can be found in Barut and Raczka, 1986, pp. 146. Furthermore, the representation D emerging in structural optimization problems is always real-valued (since U is real-valued function). Lemma 5 can be improved for real representations as

Lemma 5A. If D is irreducible among real-valued representations, all non-zero real vectors are cyclic vectors of D.

The proof of Lemma 5A is essentially the same as that of Lemma 5.

3.4. Weaker criteria of improvability

Theorem 2 proved to be an effective tool for classifying optimization problems without performing extensive structural computations. At the same time, there are cases where even simpler criteria can be applied to answer the same question. In Section 3.2.2 we summarized how to determine whether a representation is cyclic or not. This summary shows that irreducible representations are always cyclic. Since one-dimensional representations are always irreducible, we have

Theorem 3. $d \geqslant 2$ is typically a necessary condition of potential improvability.

We mention without proof that this statement can be improved in case of special groups, e.g., in case of cyclic groups of odd order, $d \geqslant 4$ is also a necessary condition.

At the same time, cyclic representations are subrepresentations of the regular representation R_Γ (Theorem 2) and they may not contain the trivial component I_1 (Lemma 4). Since $\dim(R_\Gamma) = r$, $\dim(I_1) = 1$, and R_Γ has one trivial component, cyclic representations are at most $r - 1$ dimensional. Thus,

Theorem 4. $d \geqslant r$ is a sufficient condition of potential improvability.

If $2 \leqslant d \leqslant r - 1$, the more precise condition of Theorem 2 should be applied.

4. Simple examples

In this Section, six simple optimization examples illustrate our group theory-based method (see Figs. 7 and 8), and we also analyse the examples of Section 2 (referred to as 1A, B). While these ones have D_1 (reflection-) symmetry, four of the new Examples (2A–D, see Fig. 7) have D_2 symmetry, and two of them (3A, B and Fig. 8) have D_3. All structures consist of straight elastic bars of equal cross sections, connected by hinges. In all cases, the *risk of buckling* for the total structure is minimised. From among the individual bars, the worst one determines the global risk of the structure. Thus, $U(\mathbf{x})$ is of type (3), where the potential of the ith bar is of the form (1). We introduce a geometrical parameter p (vertical size of the structures) in all examples, thus numerical computations will result in bifurcation diagrams of optima (see Fig. 9).

The steps of the analysis follow the description in Section 3.2.2. We show these steps for Example 2A only.

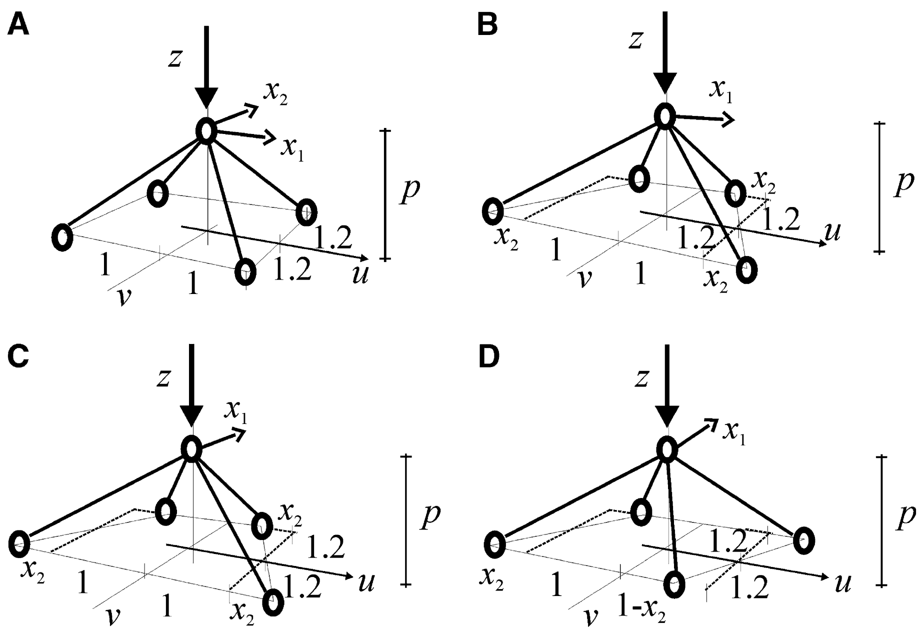

Fig. 7. Example 2 with four different perturbations.

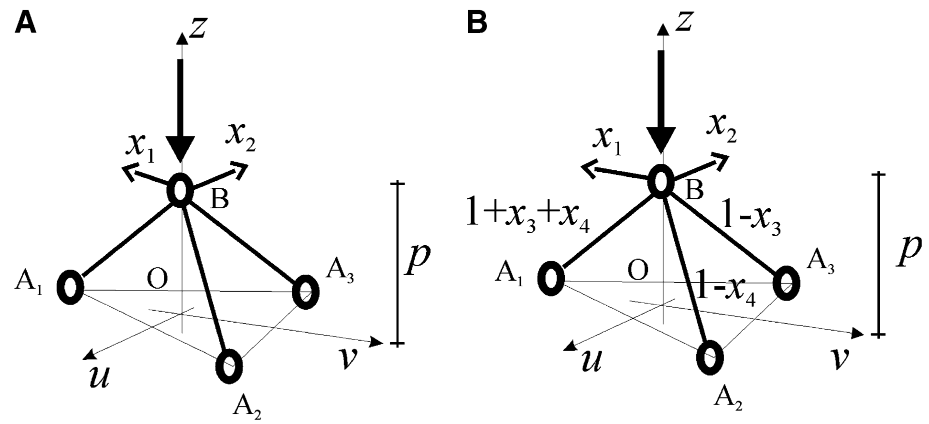

Fig. 8. Example 3 with two different perturbations. $A_1A_2A_3$ is a regular triangle. There are two variables at A, while at B, the number of perturbations is 4; x_3 and x_4 refer to the cross-sectional areas of the bars.

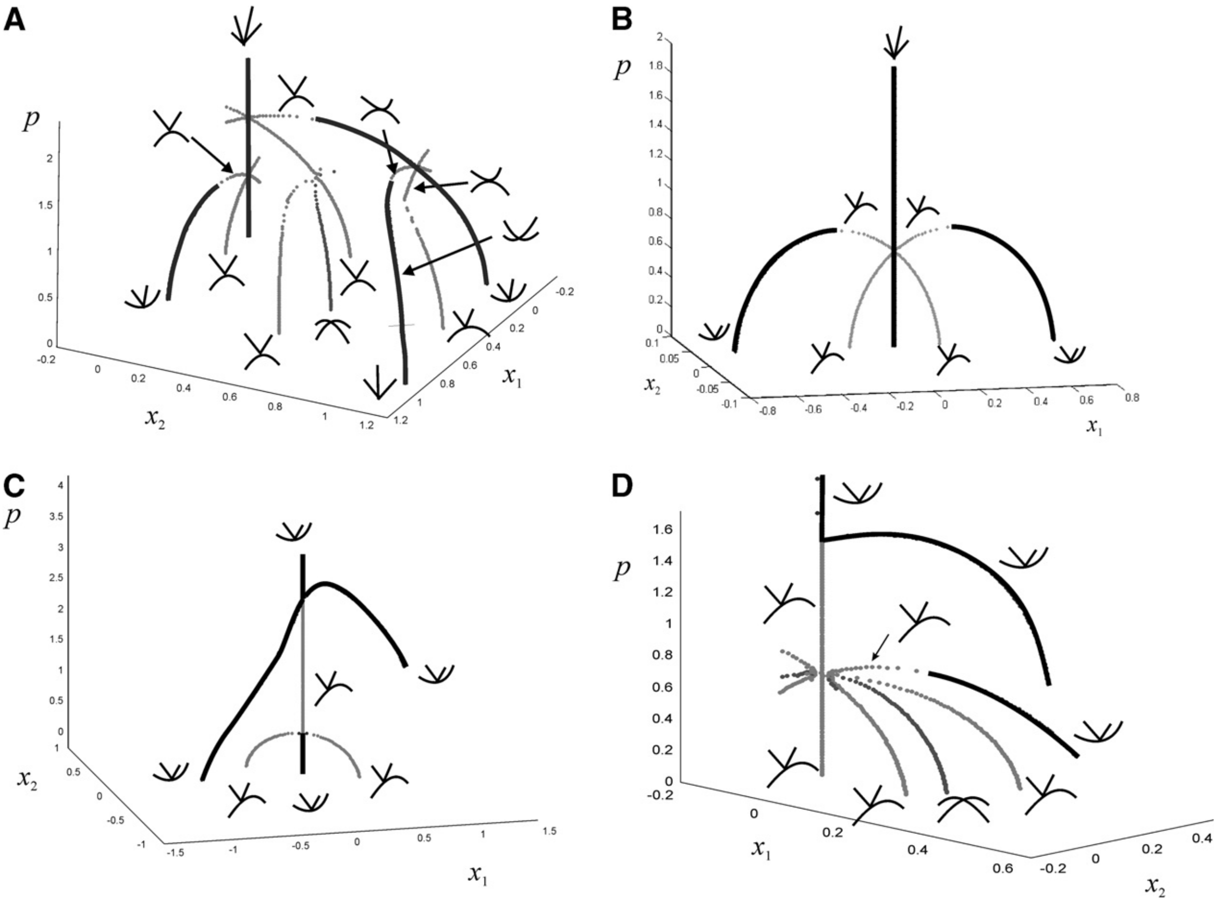

Fig. 9. Optimum-bifurcation diagrams for Examples 2A–D computed by the simplex method. Thick, black lines denote local minima, thin grey lines denote maxima, and saddle points of the functions $U(x_1, x_2)$. Small pictograms indicate the local shape of U (sharp, wedge-like or smooth) at all branches. Notice that sharp critical points are always robust optima and wedge-like points are optima or saddles. For clarity, only the domain $x_1, x_2 \geqslant 0$ has been plotted for 2A and 2D.

(1) The perfect structure has D_2 symmetry, with the following symmetry transformations:$\{\gamma_1 \equiv \textit{identity}$; $\gamma_2 \equiv$ rotation by π around axis z; $\gamma_3 \equiv$ reflection to plane $[vz]$; $\gamma_4 \equiv$ reflection to plane $[uz]\}$. These correspond to the $\mathbf{x} \to \mathbf{D}_i\mathbf{x}$ $(i = 1,2,3,4)$ transformations, where

$$\mathbf{D}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{D}_2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{D}_3 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{D}_4 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{9}$$

The character of $D \equiv \{\mathbf{D}_i\}$ is $[2 \quad -2 \quad 0 \quad 0]^{\mathrm{T}}$.

(2) The characters of the irreducible representations of group D_2 are taken from Jones (1998) and are collected in the top panel of Table 2.

(3) The unique solution of (8) is $m_1 = m_2 = 0$, $m_3 = m_4 = 1$.

(4) Since $m_i \leqslant \dim(I_i^{(D_2)})$ for all i, D is subrepresentation of the regular representation of D_2, the perfect structure is not improvable by Theorem 2.

The results of all examples are summarized in Tables 1–3. They show that 1A, and 2C are potentially improvable, the rest is not improvable.

The numerically obtained bifurcation diagrams of Examples 1A, B (Figs. 3B and 6) support the theoretical predictions. The diagrams of the novel examples are collected in Fig. 9, and Fig. 10. Since 3B has $d = 4$ variables in addition to the parameter p, its optimum diagram would be 5 dimensional. Instead of this one, we plotted the optimum diagram of a 'restricted version' of 3B, in which $x_2 = 0$, and $x_3 = x_4$, thus there are only 2 free variables. Most plots are in harmony with the predictions, however not all of them. The restricted plot

Table 1
Representation theoretical analysis of Examples 1A, B

Irreducible representations of D_1		$I_1^{(D_1)}$	$I_2^{(D_1)}$
Dimension		1	1
Character		$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$
Example 1A	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	0	1
Example 1B	$\begin{bmatrix} 2 \\ -2 \end{bmatrix}$	0	2
	char(D)	m_1	m_2

$I_i^{(D_1)}$ denotes the irreducible representations of the symmetry group D_1 ($I_1^{(D_1)}$ is the trivial representation). Neither of the induced representation contains the I_1 component, i.e., condition (i) is satisfied in both cases (cf. Lemma 4). 1A is cyclic, but 1B is not ($2I_2$ components).

Table 2
Analysis of Examples 2A–D

Irreducible representations of D_2		$I_1^{(D_2)}$	$I_2^{(D_2)}$	$I_3^{(D_2)}$	$I_4^{(D_2)}$
Dimension		1	1	1	1
Character		$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$
Example 2A	$[2 \quad -2 \quad 0 \quad 0]^{\mathrm{T}}$	0	0	1	1
Example 2B	$[2 \quad -2 \quad 0 \quad 0]^{\mathrm{T}}$	0	0	1	1
Example 2C	$[2 \quad -2 \quad 2 \quad -2]^{\mathrm{T}}$	0	0	0	2
Example 2D	$[2 \quad 0 \quad 0 \quad -2]^{\mathrm{T}}$	0	1	0	1
	char(D)	m_1	m_2	m_3	m_4

$I_i^{(D_2)}$ denotes the irreducible representations of the symmetry group D_2 ($I_1^{(D_2)}$ is the trivial representation). None of the induced representations contains the I_1 component, i.e., condition (i) is satisfied in both cases. 2A, B, D are cyclic, but 2C is not ($2I_4$ components).

Table 3
Analysis of Examples 3A, B

Irreducible representations of D_3		$I_1^{(D_3)}$	$I_2^{(D_3)}$	$I_3^{(D_3)}$
Dimension		1	2	1
Character		$\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix}$	$\begin{bmatrix}2\\-1\\-1\\0\\0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\\1\\-1\\-1\\-1\end{bmatrix}$
Example 3A	$[2\ \ -1\ \ -1\ \ 0\ \ 0\ \ 0]^T$	0	1	0
Example 3B	$[4\ \ -2\ \ -2\ \ 0\ \ 0\ \ 0]^T$	0	2	0
	char(D)	m_1	m_2	m_3

$I_i^{(D_3)}$ denotes the irreducible representations of the symmetry group D_3 ($I_1^{(D_3)}$ is the trivial representation). Neither of the induced representation contains I_1 component; both are cyclic.

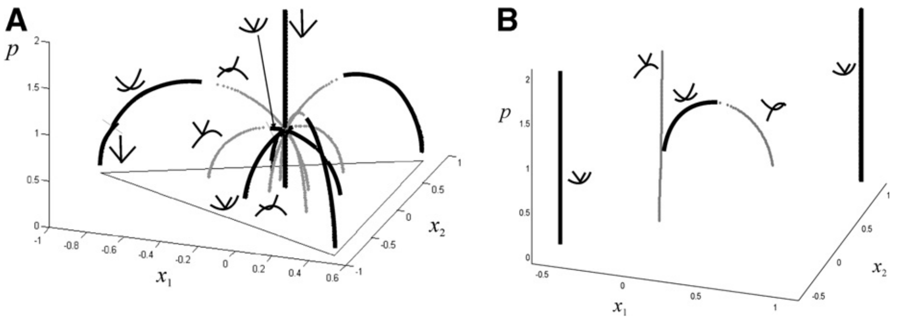

Fig. 10. Optimum-bifurcation diagrams for Example 3A and a restricted version of 3B. Thick, black lines denote local minima, thin grey lines denote maxima, and saddle points of the functions $U(x_1, x_2)$. Small pictograms indicate the local shape of U (robust, wedge-like or smooth) at all branches.

of 3B shows that this example cannot belong to class A, since it is improvable in some range of p even with the restrictions on the variables. The plot of 2B also shows unexpected behaviour.

Finding exceptions is not impossible since the main results of this paper are only 'typical', with several possibilities of exceptional behaviour. At the same time, these simple examples suggest that we should be cautious when declaring unexpected behaviour simply as atypical. In the next section we collect the potential sources of exceptions, and also find the reasons, why these two examples yield surprising results.

5. Exceptional cases

At two instances in Section 3, while deriving the main results of the paper, we utilized statements which were true *typically*, however, not *always*:

(1) In Section 3.2.1, we generated the set of weak points of $\mathbf{S}(\mathbf{0})$ from one such point P_1 as $\gamma_i(P_1)$, $i = 1, 2, \ldots, r$. However it is possible that there are more weakest points in $\mathbf{S}(\mathbf{0})$, because there is a point $P' \neq \gamma_i(P_1)$ for which accidentally $f_1(0) = f_{P'}(0)$. In this case, a problem might fall into class A instead of class B.
(2) Cyclic representations have non-cyclic vectors such as the 0 vector. If $\mathbf{g}_1$ happens to be a non-cyclic vector, the problem falls into class B instead of class A.

Clearly, our simple examples cannot be type (1) exceptions, since the whole set of their weak points can be generated from one such point P_1 as $\gamma_i(P_1)$. Class (1) contains only more complex structures with higher number of weak points. However, we can observe type (2) exceptions at several examples: some optimum diagrams contain bifurcation points at $\mathbf{x} = 0$, $p = p_0$. In fact, these are exactly the points where $\mathbf{g}_1$ is non-cyclic. Since bifurcations occur at isolated values of p, this phenomenon is atypical. At the same time, the vectors $\mathbf{g}_1$ associated with Examples 2B and 3B seem to be non-cyclic for *arbitrary values of p*. Below we explore the reason of this phenomenon.

While looking at Example 2B, it turned out that an example-specific constraint $\partial f/\partial x_2 = 0$ emerges in the potential $f(\mathbf{x})$ of one of the bars (see Várkonyi (2006) for details), and vectors of the form $\mathbf{g} = [* \ 0]^T$ happen to be non-cyclic vectors of the induced representation of Example 2B (which is the same as (Várkonyi and Domokos, 2006)). The identity $\partial f/\partial x_2 = 0$ cannot be detected without detailed analysis of the problem, i.e., the fast analysis necessarily fails at this point. Nevertheless, 2B would show standard behaviour (in this case no improvability) if we chose another type of optimization potential, such as risk against exceeding maximum stress in the bars.

In Example 3B (and also 3A) we again find $\partial f/\partial x_2 = 0$ if f denotes the potential of bar A_3B. However, this constraint is now far from example-specific: in Section 3.2.1, we generated the set of weak points from one of them as $\gamma_i(P_1)$, $\gamma \in \Gamma$. This method results in 6 weakest points in case of D_3 symmetry. However, Example 3 has only 3 bars, i.e., only three weakest points. Three pairs of the points $\gamma_i(P_1)$ coincide, because P_1 *is invariant under an element* γ_i *of* Γ (e.g., bar $A_3B \equiv P_1$ is invariant under reflection to plane *[vz]*). Consequently, $f(x_1, x_2, x_3, x_4) = f(x_1, -x_2, x_3, x_4)$, yielding $\partial f/\partial x_2 = 0$. This type of exception cannot be considered as atypical, since many symmetrical structures have their weakest points in special positions. At the same time, this exception is robust in the sense that the special behaviour is preserved if a different potential U is applied to the same structure.

As we see, type (2) exceptions have two subclasses:

(a) $\mathbf{g}_1$ becomes a non-cyclic vector atypically, or due to an example-specific constraint.
(b) the weak point P_1 is invariant under a subgroup Γ' of Γ, which automatically results in a symmetry-constraint on $\mathbf{g}_1$.

Examples of the latter type (2*b*) behave as if they had only weaker symmetry (cf. the modified analysis of Example 3 in Table 4, where the assumption of C_3-symmetry helps to generate the set of weakest points correctly as $P_i \equiv \gamma_i(P_1)$). In the extreme case $\Gamma \equiv \Gamma'$, U becomes smooth and all problems are potentially improvable, regardless of the number and type of perturbing variables.

The listed, various possibilities of exceptional behaviour indicate that the presented results should be applied with caution, this section provided some hints in this respect. Exceptions of type 2*b* occur in numerous cases, however, these cases can be readily recognised if the position of the weakest point is determined. Exceptions of type 1, and 2*a* are much rarer, however, declaring that such cases are atypical is questionable; engi-

Table 4
Modified analysis of Examples 3A, B, assuming C_3 instead of D_3 symmetry

irreducible representations of C_3		$I_1^{(C3)}$	$I_2^{(C3)}$	$I_3^{(C3)}$
Dimension		1	1	1
Character		$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}1\\(-1)^{\frac{2}{3}}\\(-1)^{\frac{4}{3}}\end{bmatrix}$	$\begin{bmatrix}1\\(-1)^{\frac{4}{3}}\\(-1)^{\frac{2}{3}}\end{bmatrix}$
Example 3A	$[2 \quad -1 \quad -1]^T$	0	1	0
Example 3B	$[4 \quad -2 \quad -2]^T$	0	2	0
	char(D)	m_1	m_2	m_3

3A is cyclic, but 3B is not ($2I_2$ components).

neering designs may often contain hidden constraints. Exceptions of these types cannot be recognised without detailed analysis of the problems.

6. Summary and connection to slightly asymmetrical optima

In this paper we analysed whether and under which conditions engineering structures with symmetry may be improved via small geometric perturbations. Our main motivation was the observation of many creatures in Nature with secondary *slight asymmetry*. Since those creatures evolved as competitive optima, we looked for similar solutions among structures, despite the engineer's fundamental intuition that symmetry is locally optimal. Our analysis showed that in several cases of practical interest the engineer's intuition is correct, and symmetrical structures may not be locally improved. On the other hand, there are still many cases where the opposite is true and our goal was to identify those cases and also to indicate, exactly how the local perturbations have to be introduced to achieve better structural response.

We realised that one has to distinguish between *global* and *local* optimum criteria. The former refer to some overall scalar quantity (potential) U associated with the structure as a whole (e.g., total mass), the latter derive the optimisation potential as the supremum of local potentials U_i, each of which is associated with a *weak point* (e.g., individual bar of a truss, individual fibre of a beam) of the structure. We found that on one hand structures with *global* criteria are relatively easy to improve; on the other hand this formulation is less adapted to real engineering problems. Consequently, our paper focused on the problem how to improve symmetrical structures with *local* optimum criteria.

Whether or not such a structure with symmetry group Γ can be improved, depends essentially on the properties of the *induced representation* D, i.e., the representation of Γ in the space of the symmetry-breaking variables. On one hand, this is a strong result because structural symmetry is mostly described by simple (cyclic, dihedral) groups and the representations of those groups are easy to obtain. On the other hand, the result appears weaker because we can identify only *potential* improvability by this methods, to prove *actual* improvability more complex, problem-specific structural analysis s needed where the gradient $\mathbf{g}_1$ of $U(\mathbf{0})$ has to be considered. We provided examples of such an analysis.

Specifically, *improvability* can be predicted for *typical cases* by classifying the problem according to the simple scheme in Section 3.1. Briefly, classes A, B and C refer to *not improvable, potentially improvable* and *improvable* cases. Based on the induced representation D the symmetrical structure $\mathbf{S}(\mathbf{0})$ can be classified according to the following conditions:

- if D has a trivial component, condition (i) of Section 3.1 is not satisfied (Lemma 5), the perturbing variables are badly chosen (the problem belongs typically to class C, exceptionally to class B or A). if D has no trivial component, and
- D is a subrepresentation of R_Γ, $\mathbf{S}(\mathbf{0})$ is typically not improvable, i.e., of class A.
- D is not a subrepresentation of R_Γ, $\mathbf{S}(\mathbf{0})$ is typically potentially improvable (class B)

In case of a one-parameter *family* of structures $\mathbf{S}(p, \mathbf{x})$ (where for each value of p criteria (i)–(iii) are satisfied) one can look for bifurcations of optima resulting in ''slightly asymmetrical optima'' (i.e., optima arbitrarily close to symmetry). The existence of these was the subject of question 2 in the Introduction. Despite the formal similarity of questions 1 and 2, the answers are remarkably different. On one hand, the structural examples of Section 4 illustrate that potential improvability often leads to the emergence of slightly asymmetric optima in one-parameter families of structures (cf. the bifurcations in Fig. 7C and D). On the other hand, one has the natural intuition that the lack of improvability (local optimality) of perfect symmetry precludes the emergence of slightly asymmetrical optima in a one-parameter structural family. This intuition is true for simple reflection-symmetry (one-parameter families of class A structures typically contain no bifurcating optima according to the bifurcation analysis of Várkonyi and Domokos (2006)) but fails in the general case: Fig. 10 is an example of bifurcating asymmetrical optima in a family of non-improvable structures. A deeper analysis of this example in Várkonyi (2006) suggests that similar bifurcations may occur in many other cases.

According to the above arguments, local optimality of perfect symmetry (question 1) has much in common with slightly asymmetrical optima (question 2), however transforming the results of this paper directly to statements about slight asymmetry in structural optimization is anything but trivial. There is still much to do in exploring the role of imperfect symmetry in structural optimization and we are convinced that learning from Nature, which was the motivation of our paper, repeatedly offers new aspects in structural design.

Acknowledgement

The authors thank Gert van der Heijden and Lajos Rónyai for several helpful comments. The financial support of OTKA Grant TS49885 is gratefully acknowledged.

Appendix A

A.1. Proof of Theorem 1

We only prove Theorem 1a, the other two statements are straightforward consequences.

Proof. *'if' part of Theorem 1a.* Assume that $\mathbf{x} = 0$ is inside the convex hull. In this case it can be written as a convex combination of the nodes of the convex hull:

$$0 = \sum_{i=1}^{k} c_i \mathbf{g}_i \text{ where } c_i > 0, \quad \sum_{i=1}^{k} c_i = 1. \tag{10}$$

Transposing both sides of Eq. (10) and multiplying them by a unit vector ($\mathbf{v} \in R^d$,$|\mathbf{v} = 1$) yields

$$0 = \sum_{i=1}^{k} c_i \mathbf{g}_i^{\mathrm{T}} \ \mathbf{v}. \tag{11}$$

In the above sum, either all components are 0 or some of the components are positive. If there exists a $\mathbf{v} = \mathbf{v}_0$, for which all terms are zero, then $\mathbf{v}_0$ is orthogonal to the vectors $\mathbf{g}_i$, i.e., the vectors do not span R^d, which means that their convex hull is degenerate: it has no internal point at all in R^d. This is in contradiction with the initial assumption. Thus, there must be a positive component in (11) for arbitrary $\mathbf{v}$:

$$\max_{1 \leqslant j \leqslant k}(\mathbf{g}_j^{\mathrm{T}} \mathbf{v}) > 0 \text{ for arbitrary } |\mathbf{v}| = 1, \tag{12}$$

The function on the left side of Eq. (12) is continuous in $\mathbf{v}$ and the set $\{\mathbf{v} \in R^d, |\mathbf{v}| = 1\}$ is compact. According to the Extreme Value Theorem (see e.g., Malik, 1992), such functions have a global minimum, which is positive, due to Eq. (12)

$$m = \min_{|v|=1}(\max_{1 \leqslant j \leqslant r}(\mathbf{g}_j^{\mathrm{T}} \mathbf{v})) > 0, \tag{13}$$

Eq. (4) can be rearranged as

$$U(\mathbf{x}) = U(\mathbf{0}) + \max_{1 \leqslant i \leqslant r}((\mathbf{g}_i^{\mathrm{T}} \bar{\mathbf{x}}) \cdot |\mathbf{x}| + \mathrm{o}(|\mathbf{x}|^2)), \tag{14}$$

where $\bar{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ and $|\bar{\mathbf{x}}| = 1$. From (13) and (14) we have

$$U(\mathbf{x}) \geqslant U(\mathbf{0}) + m \cdot |\mathbf{x}| + \mathrm{o}(|\mathbf{x}|^2), \tag{15}$$

which means that $\mathbf{x}$=0 is robust minimum of $U(\mathbf{x})$. □

Proof. *Indirect proof of the 'only if' part of Theorem 1a.* Suppose that $\mathbf{x} = 0$ is outside the convex hull, but U has robust minimum at $\mathbf{x} = 0$. Let vector $\mathbf{g}^*$ point to the closest point of the convex hull to $\mathbf{x} = 0$. Because of the robustness of the local minimum and Eq. (4), there exits k for an arbitrary vector $\mathbf{v}$ satisfying $\mathbf{g}_k^{\mathrm{T}}\mathbf{v} > 0$. Substituting $\mathbf{v} = -\mathbf{g}^*$ yields

$$\mathbf{g}^{*\mathrm{T}}\mathbf{g}_k < 0, \tag{16}$$

Consider $1 \gg \varepsilon > 0$ constant and vector $\mathbf{g}^{**} = (1-\varepsilon)\mathbf{g}^* + \varepsilon\mathbf{g}_k$, which should be inside the convex hull, since it is convex combination of $\mathbf{g}^*$ and $\mathbf{g}_k$. In this case

$$|g^{**}| = g^{**\mathrm{T}}g^{**} = (1-\varepsilon)^2 g^{*\mathrm{T}}g^* + 2\varepsilon(1-\varepsilon)g^{*\mathrm{T}}g_k + \varepsilon^2 g_k^{\mathrm{T}}g_k, \tag{17}$$

Substituting (16) into (17) and neglecting nonlinear terms in ε, we get

$$|g^{**}| < |g^*| - 2\varepsilon|g^*| + o(\varepsilon^2) < |g^*|, \tag{18}$$

which is contradiction, since $\mathbf{g}^*$ was supposed to be the closest point of the convex hull to $\mathbf{x} = 0$. □

A.2. Proof of Lemma 2

Let D be a cyclic representation. Consider the vector-matrix function $\mathbf{M}(\mathbf{v}) = [\mathbf{D}_1\mathbf{v}\ \mathbf{D}_2\mathbf{v}\cdots\mathbf{D}_r\mathbf{v}]$. $\mathbf{M}$ is of size $d \times r$. D has a cyclic vector $\mathbf{v}_0$, for which $\mathrm{rank}(\mathbf{M}) = d$, i.e., $\mathbf{M}(\mathbf{v}_0)$ has a non-zero subdeterminant $\mathbf{S}(\mathbf{M}(\mathbf{v}_0)) \neq 0$ of order d. $\mathbf{S}(\mathbf{M}(\mathbf{v}))$ is a polynomial of $\mathbf{v}$, which is either $\mathbf{S}(\mathbf{M}(\mathbf{v})) \equiv 0$, or $\mathbf{S}(\mathbf{M}(\mathbf{v})) \neq 0$ for typical $\mathbf{v}$. Now we have the latter case, since $\mathbf{S}(\mathbf{M}(\mathbf{v}_0)) \neq 0$, thus the orbit of $\mathbf{v}$ is typically d-dimensional.

A.3. Proof of Lemma 3

We divide the proof into 3 steps:

– Step 1.: if D is not subrepresentation of the regular representation of Γ, it is not cyclic.
– Step 2.: the regular representation is cyclic.
– Step 3.: subrepresentations of a cyclic representation are cyclic.

Throughout the proof, we apply the notation $\mathbf{M} = \mathrm{diag}(\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_n)$ for block-diagonal matrices, and $*$ for arbitrary matrices, vectors or scalars.

Proof of Step 1: Assume that D is not subrepresentation of the regular representation R of Γ. An adequate unitary transformation decomposes the elements $\mathbf{D}_i$ of D to the direct sum of irreducible representations (i.e., to block-diagonal form). Consider the block diagonal form of representation D. According to the *Dimensionality theorem* (Jones, 1998), each irreducible representation $I \equiv \{\mathbf{I}_j, j = 1, 2, \ldots, r\}$ of Γ occurs $\dim(I)$ times in R, thus there must be an I, which occurs at least $\dim(I) + 1$ times in D. Consider now the following unitary form of the elements $\mathbf{D}_j \in D$:

$$\mathbf{D}_j = \mathrm{diag}(\overbrace{\mathbf{I}_j, \mathbf{I}_j, \ldots, \mathbf{I}_j}^{\dim(I)+1}, *) \tag{19}$$

Consider an arbitrary d-dimensional vector $\mathbf{v}$ in the form

$$\mathbf{v} = [\mathbf{v}_1^{\mathrm{T}}\mathbf{v}_2^{\mathrm{T}} \cdots \mathbf{v}_{\dim(I)+1}^{\mathrm{T}} * *]^{\mathrm{T}}, \tag{20}$$

where the $\mathbf{v}_j$-s are $\dim(I)$-dimensional vectors. Since $\dim(I) + 1$ $\dim(I)$-dimensional vectors are linearly dependent, they have a non-trivial zero linear combination:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots c_{d_1+1}\mathbf{v}_{d_1+1} = 0, \tag{21}$$

Multiplying both sides by $\mathbf{I}_j$, we obtain

$$c_1\mathbf{I}_j\mathbf{v}_1 + c_2\mathbf{I}_j\mathbf{v}_2 + \ldots c_{\dim(I)+1}\mathbf{I}_j\mathbf{v}_{\dim(I)+1} = 0 \quad \text{for any arbitrary } j, \tag{22}$$

The elements of the orbit of $\mathbf{v}$ are

$$\mathbf{D}_j\mathbf{v} = \begin{bmatrix} \mathbf{I}_j & & & & \\ & \ddots & & & \\ & & \mathbf{I}_j & & \\ & & & * & * \\ & & & * & * \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_{\dim(I)+1} \\ * \\ * \end{bmatrix} = \begin{bmatrix} \mathbf{I}_j\mathbf{v}_1 \\ \vdots \\ \mathbf{I}_j\mathbf{v}_{\dim(I)+1} \\ * \\ * \end{bmatrix}. \tag{23}$$

Eq. (22) determines a subspace of C^d, which contains all elements of the orbit in case of arbitrary $\mathbf{v}$, thus D is not cyclic.

Proof of step 2: The regular representation, by definition, consists of permutation matrices, and the orbit of vector $[1 \quad 0 \quad 0\ldots0]^{\mathrm{T}}$ is the set of vectors $[1 \quad 0 \quad 0\ldots0]^{\mathrm{T}}$, $[0 \quad 1 \quad 0\ldots0]^{\mathrm{T}}$, $\ldots$, $[0 \quad 0\ldots0 \quad 1]^{\mathrm{T}}$.We have found a cyclic vector of R, thus R is cyclic.

Proof of step 3: Let $D' \equiv \{\mathbf{D}'_i\}$ denote a d'-dimensional subrepresentation of the d-dimensional representation $D \equiv \{\mathbf{D}_i\}$, and assume that the latter one is cyclic, i.e., it has a cyclic vector $\mathbf{v}$. Assume further that D has been transformed to the block-diagonal form $\mathbf{D}_i = \mathrm{diag}(\mathbf{D}_{i'}, *)$. Separate $\mathbf{v}$ as $\mathbf{v} = [\mathbf{v}'^{\mathrm{T}} *]^{\mathrm{T}}$, where $\dim(\mathbf{v}') = d'$. The elements of the orbit of $\mathbf{v}$ are now of the form $\mathbf{D}_i\mathbf{v} = [(\mathbf{D}_{i'}\mathbf{v}')^{\mathrm{T}} *]^{\mathrm{T}}$. If vectors $\mathbf{D}_i\mathbf{v}$, $i = 1,2,\ldots,r$ span R^d, vectors $\mathbf{D}_{i'}\mathbf{v}'$, $i = 1,2,\ldots,r$ also span $R^{d'}$. Thus we have found that $\mathbf{v}'$ is cyclic vector of D', i.e., D' is cyclic.

A.4. Proof of Lemma 4

Assume that D has a trivial component. We can consider the following block-diagonal form of the induced representation: $\mathbf{T}\mathbf{D}_i\mathbf{T}^{-1} = \mathrm{diag}(1, *)$. Consider the vector $\mathbf{x} = \mathbf{T}^{-1}[1 \quad 0 \quad 0\ldots0]^{\mathrm{T}}$! In this case $\mathbf{D}_i\mathbf{x} = \mathbf{T}^{-1}\mathbf{T}\mathbf{D}_i\mathbf{T}^{-1}\mathbf{T}\mathbf{x} = \mathbf{T}^{-1}\cdot\mathrm{diag}(1, *)\cdot[1 \quad 0 \quad 0\ldots0]^{\mathrm{T}} = \mathbf{T}^{-1}[1 \quad 0 \quad 0\ldots0]^{\mathrm{T}} = \mathbf{x}$, i.e., $\mathbf{x}$ is an invariant point of D. Thus, we have

$$\gamma_i(\mathbf{S}(\mathbf{x})) = \mathbf{S}(\mathbf{D}_i\mathbf{x}) = \mathbf{S}(\mathbf{x}), \tag{24}$$

which means that $\mathbf{S}(\mathbf{x})$ is invariant under Γ, contradicting condition (i). The converse statement can be proven in the same steps in reversed order; Eq. (24) follows from condition (iii).

References

Barut, A.O., Raczka, R., 1986. Theory of Group Representations and Applications. World Scientific, Singapore.

Jones, H.F., 1998. Groups, Representations and Physics, second ed. IOP Publishing Ltd, London.

Joseph, G.G., 2000. The Crest of the Peacock, non-European roots of Mathematics. Princeton University Press, Princeton and Oxford.

Malik, S.C., 1992. Mathematical Analysis, second ed. Wiley, New York.

Poston, T., Stewart, J., 1978. Catastrophe Theory and its Applications, Pitman, London. Weyl, H., 1952. Symmetry. Princeton University Press, Princeton.

Sokolowski, J., Zolesio, J.P., 1992. Introduction to Shape Optimization, Shape sensitivity analysis. Springer Series in Computational Mathematics, Springer, Berlin.

Tzonis, A., 1999. Santiago Calatrava: The Poetics of Movement. Universe Publishing, New York.

Várkonyi, P.L., 2006. Ph.D. Thesis, Budapest Univ. of Technology and Economics. <http://www.szt.bme.hu/munkatrs/Varkonyi/PhD.pdf/>.

Várkonyi, P.L., Domokos, G., 2006. Symmetry, optima and bifurcations in structural design. Nonlinear Dynamics 43, 47–58.

Várkonyi, P.L., Meszéna, G., Domokos, G., 2006. Emergence of asymmetry in evolution. Theoretical Population Biology 70, 63–75.

Weyl, H., 1952. Symmetry. Princeton University Press, Princeton.