A Role of Symmetries in Evaluation of Fundamental Bounds

Miloslav Capek®, Senior Member, IEEE, Lukas Jelinek®, and Michal Masek®

Abstract—A problem of the erroneous duality gap caused by the presence of symmetries is solved in this article utilizing the point group theory. The optimization problems are first divided into two classes based on their predisposition to suffer from this deficiency. Then, the classical problem of Q-factor minimization is shown in an example where the erroneous duality gap is eliminated by combining solutions from orthogonal subspaces. The validity of this treatment is demonstrated in a series of subsequent examples of increasing complexity spanning the wide variety of optimization problems, namely, minimum Q-factor, maximum antenna gain, minimum total active reflection coefficient (TARC), or maximum radiation efficiency with self-resonant constraint. They involve problems with algebraic and geometric multiplicities of the eigenmodes and are completed by an example introducing the selective modification of modal currents falling into one of the symmetry-conformal subspaces. The entire treatment is accompanied by a discussion of finite numerical precision, and mesh grid imperfections and their influence on the results. Finally, the robust and unified algorithm is proposed and discussed, including advanced topics, such as the uniqueness of the optimal solutions, dependence on the number of constraints, or an interpretation of the qualitative difference between the two classes of the optimization problems.

Index Terms—Antenna theory, eigenvalues and eigenfunctions, electromagnetic modeling, method of moments, optimization.

I. INTRODUCTION

FUNDAMENTAL bounds expressed in terms of source quantities [1], [2] have shown their versatility and usefulness for a wide range of applications in antenna theory, microwaves, and optics. They delimit the performance of theoretically feasible structures that help to judge the performance of existent designs [3] and, in a few cases, lead to the conclusion that existing designs have already reached the bounds [4]. In addition, given that the bounds are far from the actual performance of the devices became the driving force to search for better designs [5]. However, despite recent success and a straightforward implementation, the problem with the presence of geometry symmetries remained open [1], [6], [7].

Under certain conditions, discussed in detail in this article, a large class of optimization problems experiences difficulties when symmetries are present. Although the problem is of a technical nature, it has a serious impact on the validity of the results since the degeneracy of eigenvalues introduces a duality gap, i.e., the difference between dual and primal solutions [8]. This duality gap is manifested by the fact that the current solution for a primal was not constructed correctly. The known empirical solutions to this issue utilize an ad hoc combination of the degenerated eigenvectors [6], [7]. This approach is difficult to apply inside a general solver dealing with a large class of problems and structures of arbitrary geometry. The main difficulty, however, arises with structures of higher order geometry degeneracies where the choice of modes to be combined is nontrivial. Since the shapes exhibiting symmetries are often used as initial designs, and since it is expected that the field of fundamental bounds will expand into a plethora of yet unsolved problems and researchers may face the problem again, a comprehensive and general treatment of this issue is of considerable importance.

The proposed solution adheres to the point group theory; namely, the von Neumann–Wigner theorem [9] is applied to a spectrum of eigenvalue traces given by the stationary points of the optimization problem. Consequently, the conditions under which the problem arises are discussed, including how the problem is always connected to an underlying (parameterized) eigenvalue problem introducing an erroneous duality gap. A simple procedure showing how to detect when the problem occurs and how to close the erroneous duality gap is given. The proposed recipe can also treat cases of realistic mesh grids, i.e., those not perfectly respecting the symmetry groups of the original object. The procedure was thoroughly tested on many canonical objects, such as a rectangular plate, square plate, metallic rim with the ground plane, in-parallel placed and crossed dipoles, and spherical shell.

This article is organized as follows. The situation is thoroughly analyzed in Section II. It is realized that the erroneous duality gap occurs only when the eigenvalue solution is required, i.e., for quadratically constrained quadratic programs (QCQPs) without linear terms. When the linear terms are present, this ambiguity vanishes as the solution does not use eigenvalue decomposition. The illustrative example of the erroneous duality gap is introduced in Section III and in Section IV with the help of point group theory. In Section V, some examples are explicitly treated, showing where and how the symmetries appeared and what is their influence on the problem. The properties of the method are discussed in Section VI. The uniqueness of the results (current density, port voltages, and so on) is investigated in light of the knowledge.
The spherical shell belongs to the O(3) symmetry group (see Appendix I and [10]). The arrangement embodies a Cs symmetry group. It is shown that the presence of symmetries may introduce additional degrees of freedom for the optimization, and it is specified where it is gained from the symmetry treatment. It is assumed that the algebraic properties of the operators are given later on). The symmetry group (see Appendix I and [10]). (b) Current density, Fig. 1. Various source quantities x to be optimized. (a) Port voltage, \( \mathbf{x} = \mathbf{V} \), of two in-parallel placed dipoles. The arrangement embodies a C\(_2\)v group (see Appendix I and [10]). (c) Current density, \( \mathbf{x} = \mathbf{a} \) on a spherical shell expanded in spherical harmonics. The spherical shell belongs to the O(3) symmetry group [10].

II. QCQP PROBLEMS

The evaluation of the source quantity-based fundamental bounds starts with a statement of the optimization problem. Two problems, denoted as \( P_1 \) and \( P_2 \), are shown below to distinguish when the problem with symmetries may (\( P_1 \)) or may not (\( P_2 \)) arise. After establishing the Lagrangians, the optimization problems are solved via dual formulation [8], the solution to which is here denoted as \( \langle \mathbf{x} \rangle \). The stationary points \( \mathbf{x} \) are solutions to dual problems solved via dual formulation [8], the solution to which is here denoted as \( \langle \mathbf{x} \rangle \). The stationary points \( \mathbf{x} \) are

A. Optimization Problem \( P_1 \)

Let us start with QCQP problem \( P_1 \) containing only quadratic terms

\[
\begin{align*}
\min & \quad \mathbf{x}^\mathsf{H}\mathbf{A}\mathbf{x} \\
\text{s.t.} & \quad \mathbf{x}^\mathsf{H}\mathbf{B}\mathbf{x} = 1 \\
& \quad \mathbf{x}^\mathsf{H}\mathbf{C}\mathbf{x} = 0
\end{align*}
\]

(1)

where \( \mathbf{x} \) is the optimized quantity, e.g., current density or port voltages in the source region (see Fig. 1), and \( \mathbf{A}, \mathbf{B}, \) and \( \mathbf{C} \) are integrodifferential operators represented in a basis defined by either piecewise or entire-domain basis functions \( \{\mathbf{\psi}_n\} \) [11]. It is assumed that the algebraic properties of the operators are compatible with the problem to be solved, e.g., \( \mathbf{B} \succ 0 \) and \( \mathbf{C} \) being generally indefinite (the true physical meaning of these operators is given later on).

The Lagrangian of the problem \( P_1 \) is

\[
\mathcal{L}_1(\lambda_i, \mathbf{x}) = \mathbf{x}^\mathsf{H}\mathbf{H}(\lambda_i)\mathbf{x} + \lambda_1
\]

(2)

with its derivative

\[
\frac{\partial \mathcal{L}_1(\lambda_i, \mathbf{x})}{\partial \mathbf{x}^\mathsf{H}} = \mathbf{H}(\lambda_i)\mathbf{x}
\]

(3)

where \( \mathbf{H}(\lambda_i) = \nabla^2 \mathcal{L}_1 / \partial \mathbf{x}^\mathsf{H}\partial \mathbf{x} = \mathbf{A} - \lambda_1 \mathbf{B} - \lambda_2 \mathbf{C} \) is the Hessian matrix [12]. The stationary points \( \mathbf{x} \) are solutions to

\[
\frac{\partial \mathcal{L}_1(\lambda_i, \mathbf{x})}{\partial \mathbf{x}^\mathsf{H}} = 0
\]

(4)

or explicitly to

\[
\mathbf{A}\mathbf{x} - \lambda_2 \mathbf{C}\mathbf{x} = \lambda_1 \mathbf{B}\mathbf{x}.
\]

(5)

B. Optimization Problem \( P_2 \)

For the sake of completeness, the second optimization problem \( P_2 \) is defined as

\[
\begin{align*}
\min & \quad \mathbf{x}^\mathsf{H}\mathbf{A}\mathbf{x} \\
\text{s.t.} & \quad \mathbf{x}^\mathsf{H}\mathbf{B}\mathbf{x} = 1 \\
& \quad \mathbf{x}^\mathsf{H}\mathbf{C}\mathbf{x} = \text{Re}\{\mathbf{x}^\mathsf{H}\mathbf{c}\}
\end{align*}
\]

(6)

i.e., the second constraint contains a linear term in \( \mathbf{x} \) with \( \mathbf{c} \) being known column vector. Analogous to (2), the Lagrangian reads

\[
\mathcal{L}_2(\lambda_i, \mathbf{x}) = \mathbf{x}^\mathsf{H}\mathbf{H}(\lambda_i)\mathbf{x} + \lambda_2 \text{Re}\{\mathbf{x}^\mathsf{H}\mathbf{c}\} + \lambda_1.
\]

(7)

The derivative of the Lagrangian is

\[
\frac{\partial \mathcal{L}_2(\lambda_i, \mathbf{x})}{\partial \mathbf{x}^\mathsf{H}} = \mathbf{H}(\lambda_i)\mathbf{x} + \frac{\lambda_2}{2}\mathbf{c}.
\]

(8)

The stationary points \( \mathbf{x} \) are

\[
\mathbf{x} = -\frac{\lambda_1}{2}\mathbf{H}^{-1}(\lambda_i)\mathbf{c}
\]

(9)

with the demand that \( \mathbf{H}(\lambda_i) \succ 0 \).

C. Solution to Dual Problems

Primal problems \( P_1 \) and \( P_2 \) with stationary points (5) and (9) are generally nonconvex and are often approached using a dual function [8] defined as

\[
d_\rho(\lambda_i) = \inf_{\mathbf{x}} \{\mathcal{L}_\rho(\lambda_i, \mathbf{x})\}
\]

(10)

where \( \rho \in (1, 2) \). The supremum of the dual function

\[
d_\rho^* = \sup_{\lambda_i} \{d_\rho(\lambda_i)\}
\]

(11)

is a lower bound to the primal optimization problem [8], the solution to which is here denoted as \( p^* \). Since the dual function is convex [8], the solution to (11) can easily be found. Algebraic techniques reducing the computational burden behind the optimization of this type of problem are presented in [13].

Generally, the duality gap \( g^* \geq 0 \)

\[
g^* = p^* - d^*
\]

(12)

exists; nevertheless, problems involving the minimum Q-factor [6], maximum antenna gain [7], maximum radiation efficiency [14], [15], minimum total active reflection coefficient (TARC) [16], and their mutual tradeoffs [17], were shown to have no duality gap. Hence, to simplify the exposition, and without loss of generality, it is assumed, for the rest of this article, that there is no duality gap \( g^* \), i.e., \( p^* = d^* \) for both problems \( P_1 \) and \( P_2 \).
The typical workflow solving problem \( P_1 \) consists of an iterative evaluation of the generalized eigenvalue problem (5), taking the dominant eigenvalue \( \lambda_1 \) and setting the multiplier \( \lambda_2 \) so that \( \lambda_1 \) is maximized. On the contrary, the treatment of problem \( P_2 \) requires a repetitive solution to the system of linear equations. This is an important distinction between problems \( P_1 \) and \( P_2 \): issues with symmetries may occur in problem \( P_1 \), while they cannot appear for problems of type \( P_2 \).

III. ILLUSTRATIVE EXAMPLE—PROBLEM OF \( P_1 \)-TYPE

Let us demonstrate the effect of symmetries on a practical example of Q-factor minimization with a constraint on the self-resonance of the current, specifically

\[
\begin{align*}
\min \quad & \mathbf{1}^\mathbf{I^H} \mathbf{W} \\
\text{s.t.} \quad & \mathbf{1}^\mathbf{I^H R_0 I} = \frac{1}{2} \\
& \mathbf{1}^\mathbf{I^H X_0 I} = 0
\end{align*}
\]

(13)

where \( \mathbf{W} = \mathbf{A} = \omega \mathbf{\epsilon X_0}/\partial \omega \), \( \mathbf{R}_0 = \mathbf{B} \), and \( \mathbf{X}_0 = \mathbf{C} \) from (1), i.e., the problem of the minimum Q-factor falls into a class of \( P_1 \) problems, and \( \mathbf{Z}_0 = \mathbf{R}_0 + \mathbf{X}_0 \) is the impedance matrix for a scatterer made of a perfect electric conductor (PEC) (see [13] for the exact definition of all the matrix operators). The basis functions used are RWG functions [18], and the optimization variable \( \mathbf{x} = \mathbf{I} \) represents the surface current density as

\[
\mathbf{J}(\mathbf{r}) \approx \sum_n I_n \mathbf{\psi}_n(\mathbf{r}).
\]

(14)

All the operators were evaluated in the AToM package [19].

This problem has a long history starting with a seminal work of Chu [20] and has fully been described and solved in [6]. The solution to the dual problem (11) reads

\[
d^* = \max_{\lambda} \min_{m} \lambda_{1,m}
\]

(15)

with the eigenvalues \( \lambda_{1,m} \) defined by [see (5)]

\[
\frac{1}{2}(\mathbf{W} - \lambda_2 \mathbf{X}_0)\mathbf{I}_m = \lambda_{1,m} \mathbf{R}_0 \mathbf{I}_m.
\]

(16)

The definition of the Q-factor [21] can be rewritten as [6]

\[
Q(\mathbf{I}) = \max \left\{ \frac{\mathbf{1}^\mathbf{I^H X_m I} \mathbf{1}^\mathbf{I^H X_m I}}{\mathbf{1}^\mathbf{I^H R_0 I}} \right\}
\]

(17)

where

\[
\begin{align*}
\mathbf{X}_m &= \frac{1}{2}(\mathbf{W} + \mathbf{X}_0) \\
\mathbf{X}_e &= \frac{1}{2}(\mathbf{W} - \mathbf{X}_0).
\end{align*}
\]

(18a, b)

The formula (17) is valid for arbitrary current \( \mathbf{I} \) and can be used as a useful check of the duality gap \( g = Q(\lambda_2^*) - d^* \), where \( Q(\lambda_2^*) \) is a Q-factor evaluated via (17) with current \( \mathbf{I}_1 \) \((m = 1)\) found by (16) at \( \lambda_2 = \lambda_2^* \). When no duality gap occurs, we have

\[
Q^* = Q(\mathbf{I}_{\text{opt}}) = d^*.
\]

(19)

A solution to (15) is found here for two different shapes: an L-shape plate and a rectangular plate with a perfectly symmetric mesh grid [see Fig. 2(b)]. The effects of the nonregular mesh grid, as depicted in Fig. 2(c), are studied as well. The dual function and its maximum \( d^* \) at \( \lambda_2^* \) are shown in Fig. 3(a)–(c) corresponding to those of Fig. 2. Due to the large numerical dynamics in the bottom panes, the vicinity of the dual solutions are zoomed in the top panes of Fig. 3 with the traces for the actual value of Q-factor (17) added.

The nonsymmetrical case (a) causes no problems and \( Q^* = d^* \) for \( \lambda_2^* \), i.e., there is no duality gap. On the other hand, case (b) seemingly embodies a duality gap \( g = Q(\lambda_2^*) - d^* \). This “erroneous” duality gap is caused by the eigenvalue crossing (two eigensolutions to (16) are degenerate at \( \lambda_2^* \)). Neither of the degenerated eigenmodes satisfies the last constraint of (13), which is manifested by the immediate increase in the value of corresponding Q-factor \( Q(\mathbf{I}_m) \). It is shown later on that the degenerate solutions to (16) must properly be combined to satisfy this last constraint (to secure the self-resonance of the optimal current) and to close the gap. The last case (c) has no duality gap thanks to the slightly nonsymmetrical mesh grid.

This introductory example raises a series of questions.

1) When can problems with symmetries be expected?

2) How can the problem be detected?

3) How can the erroneous duality gap \( g \) caused by the presence of symmetries be fixed?

4) How can the robustness of the treatment for a numerical evaluation be improved?

These questions are addressed in the following text after a brief review of the elements of point group theory.

IV. PRESENCE OF SYMMETRIES

The point group theory\(^1\) constitutes the framework, both for the theoretical understanding and practical treatment of the issues related to symmetries.

\(^1\)Only the crucial parts essential for this work are reviewed here. The reader is referred to, e.g., [10] and references therein for a comprehensive explanation.
that object \( \Omega \) invariant to a set of point symmetries (e.g., rotation and reflection). Imagine further that object \( \Omega \) is discretized, and basis functions \( \{ \psi_n(r) \} \) are applied. It can be shown [10] that any operator, say \( \mathbf{A} \), represented in a basis \( \{ \psi_n(r) \} \) and preserving the symmetries, can be block-diagonalized as

\[
\tilde{\mathbf{A}} = \mathbf{\Gamma}^T \mathbf{A} \mathbf{\Gamma} = \begin{bmatrix}
A_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & A_G
\end{bmatrix}
\]  

(20)

where matrix \( \mathbf{\Gamma} \) is called a symmetry-adapted basis [10] and its construction for piecewise basis functions is shown, e.g., in [23]. Each block \( A_g \) in (20) belongs to a unique irreducible representation of the point group [10], briefly denoted hereinafter as “irreps” (see Appendix I for some notable examples relevant to this work).

An important consequence of relation (20) is that the eigenvalue decomposition of operator \( \mathbf{A} \) on a symmetrical structure is also separable into irreps, i.e., each eigenvector belongs to a particular irrep, and eigenvectors from different irreps are orthogonal to each other even with respect to any operator. A central observation pertaining to the spectrum of the operator \( \tilde{\mathbf{A}} \), attributed to von Neumann and Wigner [9], then states that, if operator \( \mathbf{A} \) is dependent on a certain parameter, such as frequency or Lagrange’s multiplier [see (16)], the traces of eigenvalues (abbreviated in this article as “eigentraces”) belonging to the same irreps cannot cross each other [23], [24]. Applying this theorem to Fig. 3(b), the blue and red traces must belong to modes from different irreps, as only traces of modes from different irreps can cross.2 Applying this theorem to Fig. 3(a), no problems with degeneracies occur since no symmetries are present, i.e., all modes belong to only one irrep (see Table III). In this case, therefore, there should be no crossing of eigentraces.

V. VARIOUS ASPECTS OF THE SYMMETRY PRESENCE

Several problems of various complexities are solved and interpreted in this section in terms of the point group theory. The necessity of combining two modes from different irreps to remove the erroneous duality gap is shown in Section V-A. When geometry multiplicities appear, more than one solution exists and modes can freely be combined, as shown in Section V-B. The study of how an imperfect mesh grid affects the symmetry treatment is conducted in Section V-C. Since the mesh grid is often made from rectangular or triangular elements, not all objects are perfectly represented, e.g., a spherical shell with triangular discretization elements [25]. The subsequent example in Section V-D shows that the theory introduced in this article is generally valid for the arbitrary representation of the unknown (source) quantities (see Fig. 1), by employing port-mode representation [26] to minimize the TARC [16], [27] of the metallic rim. The spectrum of the spherical shell is evaluated analytically and compared with the numerical solution in Section V-E. The last example in Section V-F deals with an academic, yet highly relevant, technique manipulating the eigenvalue traces of the isolated irreps.

A. Algebraic Multiplicity of Eigenvalues (Rectangular Plate)

The erroneous duality gap shown in Section III for a rectangular plate [see Fig. 2(b) and the results in Fig. 3(b)] is eliminated here by the proper combination of degenerate eigenvectors.

The optimization problem (13) is solved with (16) by separately utilizing (20) for irreps \( B_2 \) and \( A_2 \), i.e., two traces with a crossing at \( \lambda^* = 0.662 \) in Fig. 3(b). At the crossing point, the corresponding eigenvectors can be linearly combined without a change of dual function value \( g^* = \alpha (\lambda_2^*) \). Taking dominant modes from irreps as \( \mathbf{I}_a \in B_2, \mathbf{I}_b \in A_2 \) (see Fig. 4), we get

\[
\mathbf{I}_{\text{opt}} = \mathbf{I}_a + \alpha \mathbf{I}_b.
\]  

(21)

The erroneous duality gap in Fig. 3(b) (top pane) is a manifestation of the constraint’s violation in (13). Therefore, constant \( \alpha \) is found to fulfill

\[
\mathbf{p}_a^T \mathbf{X}_a \mathbf{I}_{\text{opt}} = 0.
\]  

(22)

Since modes \( \mathbf{I}_a \) and \( \mathbf{I}_b \) belong to different irreps, we have

\[
\mathbf{p}_a^T \mathbf{X}_a \mathbf{I}_b = 0, \text{ and}
\]

\[
\alpha = \sqrt{\frac{\mathbf{p}_a^T \mathbf{X}_a \mathbf{I}_a}{\mathbf{p}_a^T \mathbf{X}_a \mathbf{I}_b}}.
\]  

(23)

2Modes from different irreps have no physical interaction, which is the reason for the allowed crossing of their traces.

![Fig. 3](image-url)

(a) Nonsymmetrical geometry. (b) Symmetrical geometry. (c) Symmetrical geometry and nonsymmetrical mesh grid. Eigenvalues (Lagrange’s multipliers) \( \lambda_1 \) from (16) as functions of Lagrange’s multiplier \( \lambda_2 \) (bottom panes) and the corresponding Q-factors (17) (top panes, the green solid curves). The top panes show details in the vicinity of the optimal value of the multiplicator \( \lambda^* \). The two lowest eigenvalues \( \lambda_{1m} \) are depicted. For case (b), the lowest eigenvalue from \( B_2 \) (the blue line), \( A_2 \) (the red line), and \( B_1 \) (the orange line) irreducible representations [10] are shown (to be discussed in Section IV). The structures depicted in Fig. 2 are employed with the physical setting described in the caption of Fig. 2.
Fig. 4. Current densities associated with the first two modes of the eigenvalue problem (16) evaluated for the rectangular shape depicted in Fig. 2(b). (a) Capacitive mode belonging to irrep B2 [the blue line in Fig. 3(b)] with Q-factor $Q(I_1) = 44.57$. (b) Inductive mode belonging to irrep A2 [the red line in Fig. 3(b)] with Q-factor $Q(I_2) = 206.4$. (c) Combination of currents from subfigures (a) and (b) with the mixing coefficient $\alpha = 4.232$. The resulting current $I_{\text{opt}}$ is self-resonant and $Q(I_{\text{opt}}) = 38.02$.

with $\varphi \in [0, 2\pi)$ and the assumption that the square root is real. Except of nonphysical cases when the underlying mesh grid does not support inductive modes, this is always the case as for capacitive/inductive current $I_1$ we have to choose inductive/capacitive current $I_0$ so that the reactive powers have opposite sign. Combining the degenerated modes with (21) and (23), we get the optimal current seen in Fig. 4(c), fulfilling all constraints and $d^* = Q^*$.

Notice that the mixing coefficient $\alpha$ has the same form as in [28], where two dominant characteristic modes (capacitive and inductive) were combined to get a minimum Q-factor.

B. Geometry Multiplicity of Eigenvalues (Square Plate)

This example attempts to highlight the difference between a degeneracy across irreps (Section V-A) and the higher dimension of a single irrep, a situation where the symmetries introduce additional degrees of freedom (this section).

Let us consider a setting depicted in Fig. 5, which shows a square plate made of a perfectly conducting material, discretized with a symmetric mesh grid, and centered with respect to the coordinate system. Antenna gain $G$ in a direction $\hat{\varphi}$ and polarization $\hat{e}$ for a self-resonant current is to be maximized as

$$\min \quad -I^H U(\hat{e}, \hat{\varphi}) I$$

$$\text{s.t.} \quad I^H (R_0 + R_\rho - \nu X_0) I = 1$$

$$I^H X_0 I = 0$$

(24)

where $F(\hat{e}, \hat{\varphi}) = [F_1^T(\hat{\varphi}) F_2^T(\hat{\varphi})]^T$ and $R_\rho$ is a material matrix defined in [2]. The optimization problem (24) is solved according to the procedure from Section II-A by combining constraints, as proposed in [7]. The solution reads

$$G^* = d^* = -4\pi \max_m \frac{\lambda_{1,m}}{v}$$

(26)

where

$$-\nu I_m = \lambda_{1,m} (R_0 + R_\rho - \nu X_0) I_m$$

(27)

with $\nu = -\lambda_2/\lambda_{1,m} \in [v_{\text{min}}, v_{\text{max}}]$ being picked so that the matrix on the right-hand side of (27) is positive definite [7]. A further acceleration of the formula (27) is possible (see [7] for details).

The optimization problem (24) differs from (13) in two respects. First, matrix $U$ has rank 2, which means that only two eigenvalues from (27) differ from zero. Second, matrix $U$ explicitly depends on the observation coordinate, which also must be taken into account when considering the symmetries of the problem. Notice that, for a general observation coordinate $\hat{\varphi}$, the physical problem is not symmetric although the antenna geometry is.

For the purpose of this example, let us assume that the direction for radiation intensity maximization has been set to $\hat{\varphi} = \hat{z}$ and that the electrical size is $ka = 1/2$. The material parameters were set to be equivalent to copper at frequency $f = 1$ GHz. No restrictions were imposed on polarization $\hat{e}$, meaning that the solution can equally be formed by polarization pointing into $\hat{\varphi}$- and $\hat{\varphi}$-directions (or their combination). With these settings and the mesh grid from Fig. 5, the optimization problem complies with symmetries of the $C_{4v}$ point group (see Table VI).

The solution to (27) is depicted in Fig. 6 with an immediate observation of twice degenerated eigenetraces. These traces belong to irrep E (the only 2-D irrep of point group $C_{4v}$). Since there is no other eigentrace crossing these two at $v^*$ (all other eigenvalues are zero), there is no need to combine modes to fulfill the third constraint as in Section V-A. Instead, both solutions are valid on their own. They are geometry
multiplicities because, for \( \hat{r} = \hat{z} \), the two rank-one matrices \( \mathbf{F}_\theta \) and \( \mathbf{F}_\varphi \) forming operator \( \mathbf{U} \) are linearly independent

\[
\mathbf{F}_\phi = \mathbf{F}_\gamma \mathbf{C}_4
\]

where \( \mathbf{C}_4 \in \mathbb{R}^{N \times N} \) is the (unitary) rotation matrix by angle \( \varphi = \pi/2 \) around \( \hat{z} \)-axis, \( \mathbf{C}_4^4 = \mathbf{I} \), represented in basis \( \{ \psi_n(r) \} \); therefore

\[
\mathbf{F} = \begin{bmatrix} \mathbf{F}_\theta \\ \mathbf{F}_\varphi \end{bmatrix} \mathbf{C}_4
\]

yields twice degenerated eigenvalue \( \lambda_1 \) in (27) since, according to (25)

\[
\mathbf{F}^\dagger \mathbf{F} = \mathbf{F}_\theta^\dagger \mathbf{F}_\theta + \mathbf{C}_4^\dagger \mathbf{F}_\varphi \mathbf{F}_\gamma \mathbf{C}_4.
\]

Adapting the knowledge gained in this section on an example of minimal Q-factor optimization from Section V-A with a shape from, e.g., the \( C_{4v} \) point group (a square plate), a problem originates where two modes out of three degeneracies have to be combined as (21) to fulfill the third constraint (22). In such a case, these two modes have to be from different irreps, specifically \( \mathbf{I}_A \in \mathcal{I}_A, \mathbf{I}_B \in \mathcal{I}_B \), and \( A \neq B \) so that

\[
\text{sign} \left( \mathbf{t}^A \mathbf{x}_0 \mathbf{L}_r \right) = -\text{sign} \left( \mathbf{t}^B \mathbf{x}_0 \mathbf{L}_r \right).
\]

Otherwise, the erroneous duality gap cannot be eliminated.

### C. Imperfections of the Mesh Grid

The understanding gained in the previous sections will be exploited here on an example of mesh grid imperfection, where the point group rules are obscured by the fact that all computations are made with finite numerical precision.

Two structures of different point groups are assumed: a square plate (\( C_{4v} \)) and a rectangular plate (\( C_{2v} \)). The optimization of the Q-factor introduced in Section IV and solved in Section V-A is considered. The discretization grids are made of square pixels (see the insets on the left of Fig. 7) or compressed both horizontally and vertically (see the insets on the right of Fig. 7). Assuming that the mesh grid lies in the \( xy \) plane with the bottom-left corner at the origin, the compression is provided via transformation

\[
\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x/L_x \xi \\ y/L_y \xi \end{bmatrix}
\]

applied on every grid node, where \( L_x, L_y \) are side lengths of the square or rectangle and \( \xi \in (1, \infty) \).

The smooth distortion of the symmetric mesh grid enables the evolution of an erroneous duality gap to be seen, depicted as a normalized quantity in Fig. 7. For \( \xi = 1 \), the mesh grids preserve the symmetry of the object, and an erroneous duality gap exists (see the left part of Fig. 7 highlighted by the gray background color). The error given by the difference between the primal and dual solution attains \( 34% \) for the square plate and about \( 17% \) for the rectangular plate, respectively. For a reasonably large value of \( \xi \), say \( \xi > 1 + 10^{-2} \), the nonsymmetry of the grid is significant enough that no special treatment is required (duality gap is zero) (see the right part of Fig. 7, highlighted by the green background color). The most challenging cases lie between these two regions, highlighted by the red background color in Fig. 7 and, often, occur in practice due to rounding errors and other numerical imperfections. This region deserves further attention because the symmetry treatments from Sections V-A and V-B have to be properly adapted.

The dual solution to the example of the square plate and Q-factor minimization, as depicted in Fig. 7 (top pane), is...
repeated in Fig. 8. The close vicinity around $\lambda^*$ point is studied for $\zeta = \{1, 1 + 10^{-5}, 1 + 10^{-2}\}$, i.e., for three various representatives of different regions in Fig. 7. It is seen that crossings of eigentrajectories for the symmetric case evolve into the crossing avoidance scenario initially described in [9] (English transcription) and recalled in [24]. A problematic case appears in Fig. 8(b) where the values of $\lambda_1$ for $\lambda^*$ are very close to each other; in this particular case, they are the same up to six significant digits, yet not separated into irreps. The assumption $I^I CI_b = 0$ secured by (20) is, therefore, not valid anymore, and one has to solve (22) via (21) using

$$|a|^2 + 2 \text{Re} \left\{ a \frac{I^I X_0 I_b}{I^I X_0 I_b} \right\} + \frac{I^I X_0 I_a}{I^I X_0 I_b} = 0.$$  \hfill (33)

Formula (33) is a generalization of (23) for slightly nonsymmetrical structures or for perturbed nonsymmetrical mesh grids of symmetric structures. The only difference is the selection of suitable modes $I_a$ and $I_b$ to be combined. This fact is discussed further on the algorithmic level in Section VI-A. Notice that the scenario shown in Fig. 8(c) contains one eigentrajectory (the blue curve), which is significantly separated from the others to represent the true solution to the problem on its own.

The study in this section suggests that one may avoid issues related to the symmetries when nonsymmetrical mesh grids are utilized; therefore, the symmetries within the operators are broken. However, as revealed by Fig. 7, the threshold over which the mesh can be considered nonsymmetrical enough to disable the erroneous duality gap is not sharp and cannot be used for the general treatment of this problem.

D. Change of Basis (TARC of a Lossy Metallic Rim)

It is shown in this example that the presence of symmetries strongly affects the physics even when the basis (14) is changed, i.e., the operators are represented in another basis, which is still compatible with the point group of the studied object. A prominent example of this behavior is a port modes representation [26], which advantageously reduces the size of the problem. Another advantage is that, since the unknowns are the terminal voltages [see Fig. 1(b)], the optimal solution is directly realizable.

A metallic rim placed over a parasitic ground plane is shown in Fig. 9. The size of the ground plane is 150 mm $\times$ 75 mm, the height of the rim is 2.5 mm, and the height over the ground plane is 2.5 mm (the dimensions are adjusted to be equivalent to a smartphone chassis). The material of the chassis is copper. The discretization grid was generated to accommodate port modes and admittances $B_{L,i}$, and the matching susceptances $B_{L,i}$ (see [16]) for the detailed optimization procedure.

The position of the ports is specified in Fig. 9, with the polarization of the delta gaps pointing toward $+\hat{y}$-direction. The port admittance matrix $y$ is of $4 \times 4$ size and complies with the symmetries of the $C_{2v}$ point group. The port voltages and admittances $(R_{0,i}^{-1} - jB_{L,i})$ enforcing simultaneously zero reflections on all ports are obtained as solutions to an eigenvalue problem [16]

$$y v_i = (R_{0,i}^{-1} - jB_{L,i}) v_i,$$  \hfill (34)

and are depicted in Table I one by one as belonging to different irreps. When properly normalized, they evoke the character table for the $C_{2v}$ point group (see Table V). If port $P_1$ is taken as the initial port, port $P_3$ is identified as its rotation by $\pi$, port $P_2$ as reflection through $xz$ plane, and port $P_4$ as reflection through $yz$ plane. Knowing this, the voltage solutions can be assigned to the irreps that they represent.

The initial values of matching $B_L$ and loading $R_0$ given by (34) can further be optimized, as described in [16]. TARC values $T^*$ for all excitation schemes are summarized in the last column of Table I, concluding that the feeding scheme $v_3$, i.e., with the voltage orientation along the loop formed by the rim, gives the minimum TARC. This excitation scheme belongs to...
the B\textsubscript{1} irrep and dominates up to frequency $f \approx 750$ MHz. Around that frequency, the best performing irrep switches to another one.

One notable implication of the symmetries is that the voltage schemes from Table V are identical in amplitude, which simplifies the feeding circuitry (see [29] for a detailed study).

The conclusions drawn in this section, i.e., that the effects of the symmetries remain the same with a proper change of basis, apply for many practical applications. For example, the entire domain basis of characteristic modes [30] suffers from the necessity of eigentrace tracking [23], [24]. On the other hand, the proper use of symmetries introduces additional degrees of freedom, e.g., for MIMO antenna design [31], [32]. Another notable example involves reduction with the Schur complement [6].

E. Analytically Solvable Problem (A Spherical Shell)

The next optimization problem is solved analytically. A minimal dissipation factor $\delta$ [15] is found with the optimal current being self-resonant [5]. A spherical shell of radius $a$ and electrical size $ka$ is considered. Explicitly, the optimization problem reads [5], [17]

$$\begin{align*}
\min \quad P_{\text{lost}} \\
\text{s.t.} \quad P_{\text{rad}} &= 1 \\
\quad P_{\text{react}} &= 0
\end{align*}$$  \hspace{1cm} (35)

where the values of lost power $P_{\text{lost}}$, radiated power $P_{\text{rad}}$, and reactive power $P_{\text{react}} = 2\omega(W_{\text{m}} - W_{\text{c}})$ are given by quadratic forms as before. The optimal dissipation factor is evaluated as $\delta = P_{\text{lost}} / P_{\text{rad}}$ [15].

Let us start with a proper representation of the operators, here, in an entire domain basis of regular spherical waves $u_p(kr)$ [33]

$$J(r) = \sum_p a_p u_p(kr).$$  \hspace{1cm} (36)

The operators are given elementwise as

$$R_{0,pq} = \left\langle u_p, R_0 u_q \right\rangle = Z_0 k \int \int \frac{U_{pq}}{4\pi R} \sin(kR) \, dV \, dV'$$  \hspace{1cm} (37)

with

$$U_{pq} = u^*_p(kr) \cdot u_q(kr') - \frac{1}{k^2} \nabla \cdot u^*_p(kr) \nabla' \cdot u_q(kr')$$  \hspace{1cm} (38)

TABLE I

| $P_1$ | $P_2$ | $P_3$ | $P_4$ | irrep | 1/$R_{01}$ | $R_0$ | $T^*$ |
|-------|-------|-------|-------|-------|-----------|-------|-------|
| v\textsubscript{1} | +1 | +1 | +1 | +1 | A\textsubscript{1} | 451.3 | 24800 | 0.3064 |
| v\textsubscript{2} | +1 | -1 | +1 | -1 | A\textsubscript{2} | -313.8 | 156700 | 0.4212 |
| v\textsubscript{3} | +1 | -1 | -1 | +1 | B\textsubscript{1} | 28.64 | 98.52 | 0.2374 |
| v\textsubscript{4} | +1 | -1 | +1 | -1 | B\textsubscript{2} | 21.23 | 149.7 | 0.3302 |

[Fig. 10. Eigenvalue (Lagrange's multipliers) $\lambda_1$ from (41) as functions of Lagrange's multiplier $\lambda_2$. The TM modes are presented by the red eigentraces (since the characteristic number is negative, $\lambda_2 < 0$, the curves are increasing). The TE modes are represented by the blue eigentraces (since $\lambda_2 > 0$). Both curves are three times degenerated (in correspondence with the geometrical multiplicity of dominant TM and TE modes).

$$\mathbf{R}_0 = [R_{0,pq}], \text{ and similarly for } \mathbf{X}_0 \text{ and } \mathbf{R}_p,$$  \hspace{1cm} (39)

where $p = 2(l^2 + l - 1 + (-1)^m) + \tau$ and $l = \{0, 1, \ldots, L\}$ denotes the order of spherical harmonics, $m = \{-l, \ldots, l\}$ denotes the parity, $r = \{1, 2\}$ denotes the used radial profile (consult [33], [34] for details), $Z_0$ is the impedance of a vacuum, $k$ is the wavenumber, and $R = |r - r'|$. Importantly, the choice of spherical waves for spherical object leads to diagonal matrices $\mathbf{R}_p$, $\mathbf{R}_0$, and $\mathbf{X}_0$. Consequently, the eigenvalue problem (5) for the problem (35) reads

$$(\mathbf{R}_p - \lambda_2 \mathbf{X}_0) \alpha = \lambda_1 \mathbf{R}_0 \alpha$$  \hspace{1cm} (40)

which can further be separated into individual equations for each spherical wave

$$\frac{R_{pp}}{R_{op}} - \lambda_2 X_{qp} \quad R_{op} = \delta_p - \lambda_2 \lambda_p = \lambda_{1q}$$  \hspace{1cm} (41)

where $\delta_p$ is the dissipation factor [35] and $\lambda_p$ is a characteristic number,\textsuperscript{3} both being evaluated for dominant spherical waves in [36]. Since the dissipation factors $\delta_p$ are positive and characteristic numbers $\lambda_p$ are indefinite, (41) generates straight lines increasing (decreasing) with multiplicator $\lambda_2$ for capacitive, $\lambda_p < 0$ (inductive, $\lambda_p > 0$) modes (see Fig. 10).

To solve the dual problem (11), two modes, say the qth and the rth spherical waves, have to be chosen so that their traces intersect with the lowest value of $\lambda_1$. This task is accomplished by taking the dominant TM and TE modes, $q \in \{2, 4, 6\}$ and $r \in \{1, 3, 5\}$, respectively, with

$$\lambda_2^* = \frac{\delta_q - \delta_p}{\lambda_p - \lambda_q}.$$  \hspace{1cm} (42)

Substituting (42) into (41) for $p = q$ or $p = r$ yields $\lambda_2^* = \delta^*$ (see Fig. 10). The solution (42), however, does not secure the fulfillment of the self-resonant constraint. This constraint is

\textsuperscript{3}Spherical harmonics are the characteristic modes of a spherical shell [25].
that the traces $\lambda_{1,1}(\ell_2), \lambda_{1,2}(\ell_2)$ were untouched. In order to manipulate only with traces of irrep $B_1$, the reactance matrix of the rim is modified (by the inclusion of frequency independent tuning reactances) as

$$\tilde{X}_0 = X_0 + \Gamma_{B_1} \tilde{X}_1 \Gamma_{B_1}^T$$  \hspace{1cm} (45)$$

where

$$\tilde{X}_L = \begin{bmatrix} X_{L,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_{L,N} \end{bmatrix}$$  \hspace{1cm} (46)$$

is a matrix of tuning coefficients. Notice that matrices $\Gamma_b$ of the symmetry adapted basis belonging to irrep $B_1$ were used in the opposite direction than in relation (20). This composition guarantees that, irrespective of the matrix $\tilde{X}_L$, only properties of $X_0$ attached to irrep $B_1$ will be modified.

The use of the reactance matrix (45) in (16) instead of matrix $X_0$ generates the eigenvalue problem

$$\frac{1}{2}(W - \ell_2 \tilde{X}_0)I_m = \lambda_{1,m} R_0 L_m$$  \hspace{1cm} (47)$$

the results of which are shown in Fig. 12 for the optimal Lagrange’s multiplier $\ell_2 = \lambda_2^* \approx 0.6577$ and for a single nonzero parameter $X_{L,i} = \tilde{X}_L_i$. Notice that the particular index $i$ of a selected tuning parameter is free to choose and is a function of basis functions ordering.

Orthogonal properties of symmetry adapted bases belonging to different irreps (20) can further be employed to simplify and speed up the evaluation of Fig. 12. It has already been mentioned that (45) cannot change eigentrajectories belonging to irreps different than $B_1$. The blue and red traces in Fig. 12 are, therefore, independent of the tuning parameter, and there is no need to recalculate them (they attain the same value as in Fig. 11). To that point, relation (47) is left multiplied by $\Gamma_{B_1}^T$, and $I_m = \Gamma_{B_1} \tilde{X}_m$ is substituted, which leads to

$$\frac{1}{2} (\tilde{W}_{B_1} - \tilde{\lambda}_2^* \tilde{X}_0_{B_1} - \tilde{\lambda}_2^* \tilde{X}_L) \Gamma_{B_1} = \lambda_{1,m} \tilde{R}_{0,B_1} \Gamma_{B_1}.$$  \hspace{1cm} (48)$$

The eigenvalue problem (48) generates only those eigensolutions that belong to irrep $B_1$ (orange eigentrajectory in Fig. 12).

It is seen in Fig. 11(b) that three eigentrajectories are crossing each other at $\lambda_2^*$. This means that at least two solutions compliant with (21) are possible. These solutions are depicted in Fig. 13 in terms of the resulting surface currents. Case (a) is the classical solution known for $C_{2v}$ combining $B_2$ and $A_2$ irreps [see Fig. 4(c)]. Case (b) in Fig. 13 is similar to case (a) in shape, but the maximum current density appears on the shorter side. This is possible due to added reactive matching elements that effectively elongate the side. As seen in this context, cases (a) and (b) are close to two geometrically degenerated solutions, normally appearing on the square rim ($C_4v$ point group). Finally, combining solely $B$ irreps results in case (c).

In order to reduce the Q-factor in the third irrep, an existence of frequency-independent reactance with no energy accumulation was assumed, which is unphysical. A physically more acceptable possibility would be to manipulate the first two modes (from irrep $A_2, B_2$), so they become equal to the third mode. This will increase the Q-factor value, but
the gained benefit may be the equality of three eigenvalues (more degrees of freedom). Another possibility is a selective manipulation with a specific subset of characteristic modes. The same attempt was already undertaken with geometry manipulations, preserving the symmetries [37], with selective excitation [38], or with reactive tuning [39]. With the technique introduced above, one characteristic mode from each irrep can be modified, so they all have the same eigenvalue at arbitrary $ka$. This is possible with simple reactive matching, and the procedure above offers a simple recipe for how to do it.

VI. DISCUSSION

The determination of fundamental bounds in the presence of symmetries raised several interesting points to be discussed in this section.

A. Robust Algorithm to Eliminate Erroneous Duality Gaps

The procedure capable of dealing with all possible scenarios related to the presence of symmetries is depicted in Fig. 14.

Its robustness was tested against various examples, involving both crossing (mesh grid preserving the symmetries) and near crossing avoidance (slightly unsymmetrical mesh grids), and including shapes from all point groups depicted in Table II.

The workflow is as follows. It is assumed that the problems belonging to class $P_1$ are solved with a dedicated solver (steps 1 and 2 in Fig. 14). While the optimal multiplicator $\lambda^*_2$ is found (step 3), identify multiplicity $|M|$ of eigenvalue $\lambda_1$ (step 4), consider that they may vary up to relative error $\varepsilon$ due to numerical errors and mesh imperfections. According to $|M|$, decide whether the eigenmodes have to be combined (step 5). Notice that the decision shall not be based on constraint fulfillment as a true duality gap might exist. When degeneracies do not appear, follow the standard procedure (steps 6 and 7), i.e., determine the value of primal problem $p^*$ (or verify that constraints are fulfilled). When the necessity of mode combination is detected, a special routine replacing step 6 is called for (steps 6A–6D). First, evaluate projections for a matrix generating one of the constraints (step 6A, matrix $C_I$). Identify block-diagonal matrices within $C_I$ and assign them with different irreps (step 6B). Pick one mode from two different irreps; if not possible, pick two modes arbitrarily (step 6C) and find a value of parameter $\alpha$ (step 6D).
B. Distinction Between $P_1$-Type and $P_2$-Type Problems

We have seen that the presence of symmetries has serious consequences for the correct evaluation of problems from class $P_1$, defined by (1). Conversely, problems from class $P_2$, defined by (6), remain untouched. The reason is the presence of a linear term in the constraints, which is typically a consequence of a prescribed or, in other words, uncontrollable field quantity. A good example is a prescription for complex power balance, heavily utilized in [13]

$$I^H Z I = I^H V$$

(49)

where $V$ is a vector of excitation coefficients of the incident electric field intensity. Analogous to (49), all the linear terms with a current as the unknown couple the optimized quantity to the (external) field. This type of constraint makes the bounds sharper since it connects the optimized quantities and their excitation together.

C. Uniqueness of the Optimal Solution

The explicit solution to problems $P_1$ and $P_2$ enlighten the uniqueness of the solutions. In order to simplify the discussion, let us assume that matrices $A$ and $C$ in (1) and (6) have full rank, all the matrices are fixed, and the optimized quantity is properly representable in a basis (14), i.e., the basis is chosen so that it respects the nature of the optimized problem.

The solution to problem $P_2$ is unique. The solution to non-symmetric problem $P_1$ is nonunique only with respect to the phase of the optimal current. For problem $P_1$ with algebraic multiplicities, as shown in Section V-A, there is only one value of mixing coefficient $|\alpha|$, and only the phase of the mixed current may be chosen arbitrarily. Finally, when geometrical multiplicities occur, as shown in Section V-B, the optimal current further contains an arbitrary linear combination of geometrically degenerated eigenmodes (see Table II).

The uniqueness of the optimal currents $I_{opt}$ generating fundamental bounds implies that these bounds are not feasible except for the rare case in which the region used for the optimization, $\Omega$, is already an optimal solution to the shape synthesis problem. Moreover, the excitation used has to be in accordance with the optimal current found, that is

$$V = Z I_{opt}.$$  

(50)

If this is not the case, and once the initial shape $\Omega$ has to be perturbed to meet the condition (50) (see [40]), the removals of the degrees of freedom immediately cause the deterioration of the fundamental bounds, which consequently indicates that the original bound was not feasible.

D. More Than Two Constraints

The existence of a duality gap is not a function of symmetries. Furthermore, the number of constraints and the number of degeneracies are not related in any way. Irrespective of the number of constraints, and considering that there is no duality gap $g^*$, the “erroneous” duality gap introduced by the presence of symmetries is always eliminated by the proper choice of just one constant, $\alpha$. This statement is explained as follows.

Let us consider a problem from class $P_1$ with multiple constraints leading to the formula for the stationary points $I$ in a form

$$AI - \lambda_2 CI - \lambda_3 DI + \cdots = \lambda_1 BI.$$  

(51)

The dual problem is solved by determining the set of optimal multipliers $\{\lambda_1^*, \lambda_2^*, \lambda_3^*, \ldots\}$ (see step 3 in Fig. 14). When degeneracies are detected (step 5 in Fig. 14), a combination of modes from different irreps has to be used (steps 6A–6D in Fig. 14), introducing an additional degree of freedom, parameter $\alpha$. Notice that the values of the multipliers, $\{\lambda_1^*, \lambda_2^*, \lambda_3^*, \ldots\}$, are not changed by combining degenerated modal currents. Since we know from the beginning that the problem has no duality gap, i.e., the solution to the primal problem is equal to the solution of the dual problem, we know that all the constraints

$$I^H B I_1 + |\alpha|^2 I^H B I_2 = 0$$

$$I^H C I + |\alpha|^2 I^H C I_2 = 1$$

$$I^H D I + |\alpha|^2 I^H D I_2 = d$$

$$\vdots = \vdots$$

(52)

can be fulfilled by properly setting just one parameter $\alpha$. This means that we can pick only one of the constraints (52) and determine the proper value of the parameter $\alpha$.

The multiobjective optimization problems of fundamental bounds (see [7], [13], [17], [41]) are often formulated with additional constraints (52) representing additional objectives. Such formulations, therefore, suffer from the erroneous duality gap issues once they fit into the $P_1$ class.
E. Reduction of Computational Complexity

A useful side product of (20) is the possibility to compute the eigenvalue decomposition of each block (irrep) separately, which leads to the acceleration of the optimization. The ratio \( S \) between computation time of generalized eigenvalue decomposition utilizing symmetries and computational time of a standard decomposition is obtained by analyzing eigenvalue decomposition of (20).

In order to present only the most salient features, it is assumed that that one eigenvalue is to be found within each irrep. The potential speed-up in such a case is

\[
S \propto \frac{1}{c N^q} \sum_{g=1}^{G} N_g^q
\]

(53)

where \( G \) is the number of irreps [see (20)], \( N_g = \dim(A_g) \), \( q \) indicates asymptotic complexity of used eigenvalue algorithm, and \( c \) indicates how many times the algorithm has been used, e.g., \( q \approx 2 \) and \( c = G \) for implicitly restarted Arnoldi method [42], and \( q \approx 3 \) and \( c = 1 \) for generalized Schur decomposition [43].

In order to provide a specific example, the \( C_{2v} \) group used in this article is considered for evaluation of speed-up \( S \). The dominant solution in each of the four 1-D irreps is demanded, and the number of modes in each irrep is the same, \( N_g = N/4 \). Relation (53) is simplified to

\[
S \propto \frac{4}{c N^q} \left( \frac{N}{4} \right)^q = \frac{1}{4^2} = \frac{1}{16}
\]

(54)

both for implicitly restarted Arnoldi method and generalized Schur decomposition.

An extreme case is shown in [44] for a body of revolution, where the system of basis functions forms a reducible system so that inverse of the resulting matrix is directly possible.

VII. Conclusion

It has been shown that one entire class of optimization problems generating fundamental bounds in electromagnetism is encumbered with potential issues induced by symmetries. When no linear constraints are present, care must be taken with the investigation of the primal solution. This applies to structures with an imperfect discretization mesh grid as well, where the elimination of an erroneous duality gap might be even more problematic since the separation into irreducible representations is not possible.

A heuristic, yet general and point group theory-based technique to remove the erroneous duality gap has been presented and tested for various examples of varying complexity. The formula was tested for all canonical bodies (rectangular and square plates, a triangular shape, a spherical shell, and so on).

This work helps to understand the role of symmetries in establishing source quantity-based bounds. The challenges related to the presence of symmetries, when properly treated, introduce additional degrees of freedom. All conclusions apply not only to optimal, yet abstract and usually nonrealizable, currents but also to optimal port mode excitation and other feasible representations of integrodifferential operators, both for their surface and volumetric formulations.

APPENDIX I

Character Tables

This appendix lists character tables of point groups used in this article. Each table also contains symmetry operations available for the group (those are grouped into conjugacy classes and enumerated in the first row of the table) and the corresponding irreps (enumerated in the first column of the table). The table entries consist of group characters (numbers of the table) and denote the traces of the matrix representations for a corresponding class and irrep. The number of irreps corresponds to the number of classes [10]; all rows and columns of the character table are orthogonal.

The symmetry operations used in Tables III–VI are follows:

- E—the identical operation; 
- \( \sigma_x \)—a reflection (t is a place-holder for a specific type of reflection); 
- \( C_n(u) \)—a rotation by \( 2\pi/n \) around the \( u \)-axis; and 
- \( \sigma_{uv} \)—a reflection through plane \( uv \).

The character corresponding to identity operation E indicates the dimension of the irrep and the geometric multiplicity of the eigenvalues within that irrep. For example, current solutions falling into the E irrep of \( C_{4v} \) group is twice degenerated (see Table VI). This applies to the solutions of the problem (27) (see Fig. 6).

Nonsymmetric objects belong to point group \( C_1 \) (see Table III). Objects with one reflection plane, often classified as having odd and even solutions, belong to point group \( C_s \) (see...
Table VI

| Character Table for Point Group C_{4v}, a Square Plate Belonging To |
|-------------------|-------------------|-------------------|-------------------|
| C_{4v} E | 2C_{4}(z) | C_{2}(x) | 2σ_{v} | 2σ_{d} |
| A_{1} +1 | +1 | +1 | +1 | +1 |
| A_{2} +1 | +1 | +1 | −1 | −1 |
| B_{1} +1 | −1 | +1 | +1 | −1 |
| B_{2} +1 | −1 | +1 | −1 | +1 |
| E +2 | 0 | 0 | 0 | 0 |

Table IV). The remaining two groups mentioned here are C_{2v} (e.g., rectangular plate) in Table V and C_{4v} (e.g., square plate) in Table VI.

REFERENCES

[1] M. Gustafsson, D. Tyllé, C. Ehrenborg, M. Cismasu, and S. Norbedo, "Antenna current optimization using MATLAB and CVX," in Proc. FERMAT, vol. 15, May-Jun. 2016, pp. 1–29.

[2] L. Jelinek and M. Capek, "Optimal currents on arbitrarily shaped surfaces," IEEE Trans. Antennas Propag., vol. 65, no. 1, pp. 329–341, Jan. 2017.

[3] S. R. Best and D. L. Hanna, "A performance comparison of fundamental small-antenna designs," IEEE Trans. Antennas Propag. Mag., vol. 52, no. 1, pp. 47–70, Feb. 2010.

[4] M. Capek et al., "Optimal planar electric dipole antennas: Searching for antennas reaching the fundamental bounds on selected metrics," IEEE Antennas Propag. Mag., vol. 61, no. 4, pp. 19–29, Aug. 2019.

[5] L. Jelinek, K. Schab, and M. Capek, "Radiation efficiency cost of resonance tuning," IEEE Trans. Antennas Propag., vol. 66, no. 12, pp. 6716–6723, Dec. 2018.

[6] M. Capek, M. Gustafsson, and K. Schab, "Minimization of antenna quality factor," IEEE Trans. Antennas Propag., vol. 65, no. 8, pp. 4115–4123, Aug. 2017.

[7] M. Gustafsson and M. Capek, "Maximum gain, effective area, and directivity," IEEE Trans. Antennas Propag., vol. 67, no. 8, pp. 5282–5293, Aug. 2019.

[8] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[9] J. von Neumann and E. Wigner, In the Works of Einstein in Adiabatic Processes. Singapore: World Scientific, 2000.

[10] R. McWeeny, Symmetry: An Introduction to Group Theory and Its Applications. London, U.K.: Pergamon Press, 1963.

[11] A. F. Peterson, S. L. Ray, and R. Mittra, Computational Methods for Electromagnetics. Hoboken, NJ, USA: Wiley, 1998.

[12] J. Nocedal and S. Wright, Numerical Optimization. New York, NY, USA: Springer, 2006.

[13] M. Gustafsson, K. Schab, L. Jelinek, and M. Capek, "Upper bounds on absorption and scattering," New J. Phys., vol. 22, no. 7, Sep. 2020, Art. no. 073013.

[14] M. Uzsöky and L. Solymár, "Theory of super-directive linear arrays," IEEE Trans. Antennas Propag., vol. 65, no. 8, pp. 4134–4145, Aug. 2017.

[15] M. Capek, V. Losenicky, L. Jelinek, and M. Gustafsson, "Validating the characteristic modes solvers," IEEE Trans. Antennas Propag., vol. 65, no. 8, pp. 3914–3915, Aug. 2017.

[16] R. Harrington, "Reactively controlled directive arrays," IEEE Trans. Antennas Propag., vol. 26, no. 3, pp. 390–395, May 1978.

[17] M. Manteghi and Y. Rahmat-Samii, "Multiport characteristics of a wide-band cavity backed annular patch antenna for multipolarization operations," IEEE Trans. Antennas Propag., vol. 53, no. 1, pp. 466–474, Jan. 2005.

[18] M. Capek and L. Jelinek, "Optimal composition of modal currents for minimal quality factor Q," IEEE Trans. Antennas Propag., vol. 64, no. 12, pp. 5230–5242, Dec. 2016.

[19] M. Capek, L. Jelinek, and M. Gustafsson, "Excitation of orthogonal radiation states," 2020, arXiv:2003.09378. [Online]. Available: http://arxiv.org/abs/2003.09378.

[20] R. Harrington and J. Mautz, "Theory of characteristic modes for conducting bodies," IEEE Trans. Antennas Propag., vol. 19, no. 5, pp. 622–628, Sep. 1971.

[21] N. Peitzmeier and D. Manteuffel, "Upper bounds and design guidelines for realizing uncorrelated ports on multimode antennas based on symmetry analysis of characteristic modes," IEEE Trans. Antennas Propag., vol. 67, no. 6, pp. 3902–3914, Jun. 2019.

[22] J. Mautz and R. Harrington, "A combined-source solution for radiation and scattering from a perfectly conducting body," IEEE Trans. Antennas Propag., vol. 27, no. 4, pp. 445–454, Jul. 1979.

[23] J. E. Hansen, Ed., Numerical Methods for Large Eigenvalue Problems. Philadelphia, PA, USA: SIAM, 2003.
Miloslav Capek (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree from Czech Technical University in Prague, Prague, Czech Republic, in 2009 and 2014, respectively.

He was appointed as an Associate Professor at Czech Technical University in Prague, in 2017. He leads the development of the Antenna Toolbox for MATLAB (AToM) package. He has authored or coauthored more than 100 journal and conference papers. His research interests include electromagnetic theory, electrically small antennas, numerical techniques, fractal geometry, and optimization.

Dr. Capek is a member of the Radioengineering Society, a Regional Delegate of European Association on Antennas and Propagation (EurAAP), and an Associate Editor of *IET Microwaves, Antennas and Propagation*.

Lukas Jelinek received the Ph.D. degree from Czech Technical University in Prague, Prague, Czech Republic, in 2006.

In 2015, he was appointed as an Associate Professor at the Department of Electromagnetic Field, Czech Technical University in Prague. His research interests include wave propagation in complex media, general field theory, and numerical techniques and optimization.

Michal Masek received the M.Sc. degree in electrical engineering from Czech Technical University in Prague, Prague, Czech Republic, in 2015, where he is currently pursuing the Ph.D. degree in the area of modal tracking and characteristic modes.

He is a member of the team developing the Antenna Toolbox for MATLAB (AToM).