数理解析研究所講究録
第717巻 1990年 159-177

Group Symmetry and Sparsity in Matrix Computation

室田一雄　Kazuo MUROTA
(東大 工 計数　University of Tokyo)

池田清宏　Kiyohiro IKEDA
(長岡技科大 建設　Nagaoka University of Technology)

1 Introduction

Group-symmetry has been extensively employed to enhance efficiency in the numerical computation in many different fields of mathematical sciences. Sparsity of matrices is another important aspect to be taken into account for computational efficiency. In this paper we show how group-symmetry and sparsity can be utilized for efficient matrix computations. We discuss this issue in the context of numerical analysis of bifurcation behavior of axisymmetric structures based on a pair of recent papers [11], [15] of the authors.

We consider a static (or stationary, steady) bifurcation problem expressed by a system of equations

$$f(v,\lambda)=0$$

in a finite-dimensional vector v with a bifurcation parameter λ. We are concerned with the bifurcation diagram, i.e., the loci of (v,λ) satisfying the equations in a region of our interest.

As a typical problem that we have in mind, we mention the problem of determining the deformation of the spherical diamond shell of Figure 1 under a uniform vertical loading, where v is a 93 $(=3\times 31)$ dimensional vector representing displacements of 31 free nodes, and λ is a loading parameter.

In applications we often encounter bifurcation problems that have group-symmetry (in the sense to be made precise in §3). For instance, the system of equations describing the spherical diamond shell has a symmetry expressed by a dihedral group if all the physical properties are symmetric. Also finite dimensional approximations to group-symmetric bifurcation problems described by differential equations often fall into the category we shall consider.

A typical numerical procedure for obtaining the bifurcation diagram would involve the following two phases:

(i) extending and tracing a path (or locus), and

(ii) identifying bifurcation points as well as the directions of new branches.
The first phase (i) amounts to solving a system of nonlinear equations by, e.g., a predictor-corrector (Newton-Raphson) type method, whereas the second to determining the rank of the Jacobian matrix J of f with respect to v as well as its null space.

When the bifurcation problem has group-symmetry we can make use of this fact in various ways to enhance the computational efficiency of the numerical analysis [1], [2], [4], [5], [8], [9], [10]. In particular, a recent paper [9] by T. J. Healey illustrates to the full extent how we can reduce the computational effort in the path-tracing phase (i) by exploiting symmetry.

The main concern of this paper is to exploit symmetry in the second phase (ii) of singularity-detection, in which we are to see if the Jacobian matrix is singular at particular points. The basic idea of the proposed method is to find a suitable (unitary) matrix H, which is independent of particular points, such that the transformed Jacobian matrix H^*JH takes a block-diagonal form.

The advantage of this block-diagonalization is twofold. Obviously, the test for singularity of J is decomposed into a series of tests for singularity of diagonal blocks, which are much smaller in size and hence much easier to handle. Moreover, the diagonal blocks of H^*JH are not all distinct but contain several pairs (or tuples, depending on the symmetry) of identical matrices. Thus we have only to deal with distinct representatives among the diagonal blocks, and we can check for their singularities independently of one another, or in parallel if such a machinery is available.

The proposed method is based on the well-established mathematical facts about the (nonunique) decomposition of a group representation into a direct sum of irreducible representations (see §2). Such a transformation matrix H, as well as the block-diagonal form H^*JH, is not uniquely determined from the group representation theoretic point of view, which fact allows us to select a "good" one from the computational point of view.

For a symmetric structure like the one in Figure 1, the proposed choice of H is induced naturally from the underlying geometric symmetry, and is a block-diagonal matrix composing of many small diagonal blocks. In other words, the block-diagonal form H^*JH is obtained through a series of "local" coordinate transformations. This "locality" is computationally advantageous in that it requires small amount of computation and preserves, to some extent, the sparsity of the original Jacobian matrix J.

For the example problem of spherical diamond shell (Figure 1), the proposed method puts the Jacobian matrix J of size 93 into a block-diagonal matrix H^*JH with 8 diagonal blocks, among which are 6 distinct blocks of sizes 6, 7, 8, 10, 15 and 16.

This paper is organized as follows. Preliminaries from group representation theory are given in §2 with particular stress on dihedral groups. Sections 3 and 4 consider the block-diagonalization of the Jacobian matrix under a group symmetry from a computational point of view. The former affords general framework for bifurcation problems with symmetry expressed by a finite group G, and the latter specializes it to axisymmetric truss structures, the symmetry of which is represented by dihedral groups. We

derive estimates for the computational complexity of the proposed procedure to show its efficiency.

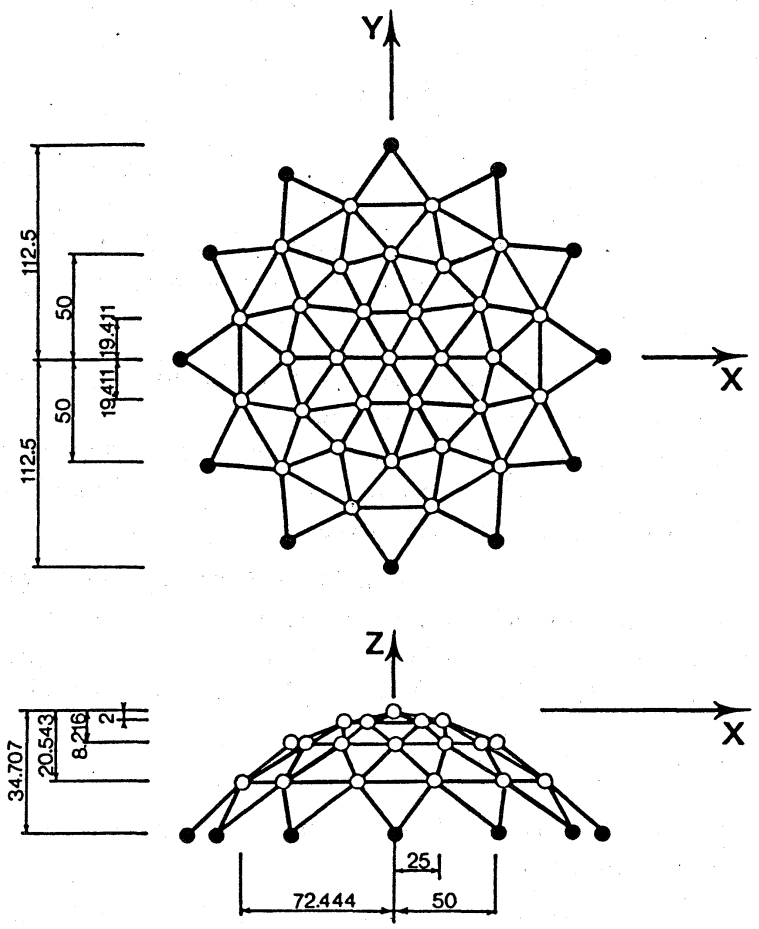


Fig.1. Spherical diamond shell

2 Preliminaries on Group Representation

This section is to introduce the notations as well as some fundamental facts about linear representations of finite groups to be used in this paper. For complete accounts, the reader is referred to textbooks such as [3], [7], [12], and [14].

Let G be a finite group, whose order is denoted by $|G|$. Let V be a finite dimensional vector space over $F = \mathbf{R}$ or $\mathbf{C}$, where $\mathbf{R}$ is the field of real numbers and $\mathbf{C}$ the field of complex numbers. Denote by $GL(V)$ the group of all nonsingular linear transformations of V onto itself, and by $GL(N, F)$ the group of all nonsingular matrices over F of order N.

By a representation of G on V is meant a homomorphism $\hat{T} : G \to GL(V)$, i.e.,

$$\hat{T}(gh) = \hat{T}(g)\hat{T}(h), \qquad g, h \in G.$$

We call V the representation space and $N = \dim V$ the dimension or the degree of the representation. Fixing a basis $B = (u_1, \cdots, u_N)$ for V, we can associate with $\hat{T}(g)$ an $N \times N$ nonsingular matrix $T(g) = T_B(g) = (T_{ij}(g))_{1 \le i,j \le N}$ in a usual manner:

$$\hat{T}(g)u_j = \sum_{i=1}^{N} T_{ij}(g)u_i, \qquad g \in G.$$

The homomorphism $T : G \to GL(N, F)$ is called a matrix representation of G. A change of basis B to $C = (v_1, \cdots, v_N)$ where $v_j = \sum_{i=1}^{N} H_{ij}u_i$ with $H = (H_{ij}) \in GL(N, F)$ replaces the matrix T_B by another matrix $T_C = H^{-1}T_BH$.

We say that $\hat{T}$ (respectively T) is unitary if $\hat{T}(g)$ (respectively $T(g)$) is unitary for each $g \in G$.

Since we are interested in computational efficiency in this paper, we are concerned not only with representations in the abstract sense but also with their concrete matrix representations with respect to particular choices of bases. In fact, it is one of the central issues of this paper to choose "good" bases that yield matrix representations which are convenient to numerical computations.

Let P be a finite set, and assume that G acts on P through permutations $\pi(g)$, $g \in G$, of P. The permutation representation $\hat{T}$ is a representation of dimension $|P|$ defined as follows. Consider the vector space $V = F^P$, i.e., the linear space consisting of "formal" sums $\sum_{p \in P} f_p u_p$ ($f_p \in F$) with basis $(u_p \mid p \in P)$, and define $\hat{T} : G \to GL(V)$ by

$$\hat{T}(g)u_p = u_{\pi(g)p}.$$

A permutation representation is called the regular representation when $P = G$, and the unit representation when $|P| = 1$.

Let $\hat{T}_1$ and $\hat{T}_2$ be representations of G with respective representation spaces V_1 and V_2 over F. We say that $\hat{T}_1$ and $\hat{T}_2$ are equivalent if there exists a nonsingular linear map $\hat{H} : V_1 \to V_2$ such that

$$\hat{T}_1(g) = \hat{H}^{-1}\hat{T}_2(g)\hat{H}, \qquad g \in G.$$

We also say that two matrix representations T_1 and T_2 of G are equivalent if there exists a nonsingular matrix H such that

$$T_1(g) = H^{-1}T_2(g)H, \qquad g \in G.$$

The direct sum $\hat{T}_1 \oplus \hat{T}_2$ of $\hat{T}_1$ and $\hat{T}_2$ is a representation on the direct sum $V_1 \oplus_F V_2$ of the representation spaces V_1 and V_2, and is defined by

$$(\hat{T}_1 \oplus \hat{T}_2)(g) = \hat{T}_1(g) \oplus \hat{T}_2(g), \qquad g \in G.$$

The tensor product $\hat{T}_1 \otimes \hat{T}_2$ of $\hat{T}_1$ and $\hat{T}_2$ is a representation on the tensor product $V_1 \otimes_F V_2$ of V_1 and V_2, and is defined by

$$(\hat{T}_1 \otimes \hat{T}_2)(g) = \hat{T}_1(g) \otimes \hat{T}_2(g), \qquad g \in G.$$

A subspace W of V is said to be ($\hat{T}(G)$- or G-)invariant if $\hat{T}(g)w \in W$ for $w \in W$, $g \in G$. If W is invariant, the subrepresentation of $\hat{T}$ on W, denoted as $\hat{T}|W$, is defined by the restriction of $\hat{T}(g)$ to W for each $g \in G$. $\hat{T}$ is said to be irreducible if there exists no nontrivial invariant subspace W, where W is nontrivial if W is neither $\{0\}$ nor V. $\hat{T}$ is absolutely irreducible if it is irreducible as a representation over $\mathbf{C}$.

There exist a finite number of nonequivalent irreducible representations of G (over a fixed field F). We denote by

$$\{\hat{T}^\mu \mid \mu \in R(G)\}$$

a family of all nonequivalent irreducible representations of G, where $R(G)$ denotes an index set for the irreducible representations of G, and put

$$N^\mu = \dim \hat{T}^\mu.$$

The unit representation is indicated by $\mu = \mu_0$. We also fix a family

$$\{T^\mu \mid \mu \in R(G)\}$$

of nonequivalent irreducible matrix representations of G.

It is known as Maschke's theorem that any representation $\hat{T}$ of G on V can be expressed as a direct sum of irreducible representations. To be more precise, V is decomposed as

$$V = \bigoplus_{\mu \in R(G)} V^\mu \tag{2.1}$$

with

$$V^\mu = \bigoplus_{i=1}^{a^\mu} V_i^\mu, \tag{2.2}$$

where V_i^μ is invariant, and $\hat{T}|V_i^\mu$ is irreducible and equivalent to $\hat{T}^\mu$. Substituting (2.2) into (2.1), we obtain

$$V = \bigoplus_{\mu \in R(G)} \bigoplus_{i=1}^{a^\mu} V_i^\mu, \tag{2.3}$$

a direct sum decomposition into irreducible components.

The decomposition (2.1) is unique and called the isotypic (or standard) decomposition; each V^μ being called an isotypic (or homogeneous) component. On the other hand, the decomposition (2.2) is not unique though the multiplicity a^μ (≥ 0) is uniquely determined. Hence, the decomposition (2.3) into irreducible components is not unique either.

The direct sum decomposition (2.3) means that the matrix representation is put into a block-diagonal form

$$T(g) = \bigoplus_{\mu \in R(G)} \bigoplus_{i=1}^{a^\mu} T_i^\mu(g), \qquad g \in G \tag{2.4}$$

with T_i^μ being irreducible, if we choose a basis of V consistent with the decomposition (2.3), i.e., if we first choose a basis $(h_{ij}^\mu \mid j = 1, \cdots, N^\mu)$ for each V_i^μ and adopt their union as a basis of V. In this paper, we take advantage of the fact that we can impose a further condition

$$T_i^\mu(g) = T^\mu(g), \qquad g \in G, \quad 1 \leq i \leq a^\mu. \tag{2.5}$$

Lemma (Schur). Let T_1 and T_2 be irreducible matrix representations of G over F, and J be a matrix over F. Assume that

$$T_1(g)J = JT_2(g), \qquad g \in G.$$

(i) If T_1 and T_2 are not equivalent, then $J = O$.
(ii) If T_1 and T_2 are equivalent, then $J = O$ or J is nonsingular.
(iii) If $T_1(g) = T_2(g)$, $g \in G$, and T_1 is absolutely irreducible, then $J = \alpha I$ for some $\alpha \in F$. □

The dihedral groups and their representations play substantial roles in this paper. The dihedral group D_n of order $2n$ is defined as

$$D_n = \{1, r, \cdots, r^{n-1}; s, sr, \cdots, sr^{n-1}\},$$

where

$$r^n = s^2 = (sr)^2 = 1.$$

A matrix representation T of D_n is determined by $T(r)$ and $T(s)$.

We index the family of nonequivalent irreducible representations of D_n by

$$\begin{aligned} R(D_n) &= \{\mathrm{s}j \mid j = 1,2,3,4\} \cup \{\mathrm{d}j \mid j = 1, \cdots, (n-2)/2\} \quad \text{for } n \text{ even;} \\ R(D_n) &= \{\mathrm{s}j \mid j = 1,2\} \cup \{\mathrm{d}j \mid j = 1, \cdots, (n-1)/2\} \quad \text{for } n \text{ odd,} \end{aligned}$$

where the first component $\delta =$ "s" or "d" of the index $\mu = \delta j$ indicates the dimension of the representation; "s" for one-dimensional representations and "d" for two-dimensional ones.

The one-dimensional irreducible representations $T^{\mathrm{s}j}$ of D_n are uniquely determined and given by

$$\begin{aligned} T^{\mathrm{s}1}(r) &= 1, & T^{\mathrm{s}1}(s) &= 1; \\ T^{\mathrm{s}2}(r) &= 1, & T^{\mathrm{s}2}(s) &= -1; \\ T^{\mathrm{s}3}(r) &= -1, & T^{\mathrm{s}3}(s) &= 1; \\ T^{\mathrm{s}4}(r) &= -1, & T^{\mathrm{s}4}(s) &= -1. \end{aligned}$$

The two-dimensional irreducible matrix representations $T^{\mathrm{d}j}$ are not unique but the following choice seems to be convenient:

$$T^{\mathrm{d}j}(r) = R^j, \qquad T^{\mathrm{d}j}(s) = S,$$

where

$$R = \begin{pmatrix} \cos(2\pi/n) & -\sin(2\pi/n) \\ \sin(2\pi/n) & \cos(2\pi/n) \end{pmatrix}, \qquad S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.6}$$

These irreducible representations over $\mathbf{R}$ are absolutely irreducible.

3 Exploiting General Group Symmetry

In this section we propose a computational method for the numerical bifurcation analysis of structures with group symmetry in general. The method will be applied to axisymmetric structures (with D_n-symmetry) in the next section.

3.1 Group equivariance

We consider a system of equations

$$f_i(v,\lambda) = 0, \qquad i = 1,\cdots,N, \tag{3.1}$$

in a finite-dimensional real vector $v = (v_j \mid j = 1,\cdots,N) \in \mathbf{R}^N = V$ with a bifurcation parameter $\lambda \in \mathbf{R}$. Suppose further that this problem is equivariant (cf. [6], [16]) to a finite group G through a unitary matrix representation T of G on V:

$$T(g)f(v,\lambda) = f(T(g)v,\lambda), \qquad g \in G,$$

where $f = (f_i \mid i = 1,\cdots,N)$. Then the Jacobian matrix $J = (J_{ij})$, $J_{ij} = \partial f_i/\partial v_j$, commutes with the group action:

$$T(g)J(v,\lambda) = J(v,\lambda)T(g), \qquad g \in G, \tag{3.2}$$

if v is G-symmetric, i.e., if $T(g)v = v$, $g \in G$.

A typical numerical procedure for finding the bifurcation diagram would involve the following two phases:

(i) extending and tracing a path (or locus), and

(ii) identifying bifurcation points as well as the directions of new branches.

The first phase amounts to solving a system of nonlinear equations by, e.g., a predictor-corrector type method, whereas the second to determining the rank of J as well as its null space.

We shall explain in the following subsection how we can make use of the nonunique decomposition (2.3), which is a refinement of the isotypic decomposition, for efficient computation in the phase (ii) of singularity detection.

3.2 Block-diagonalization

Since T is assumed to be unitary, we can find (cf. §2) a unitary (orthogonal) matrix H such that

$$H^*T(g)H = \bigoplus_{\mu\in R(G)} \bigoplus_{i=1}^{a^\mu} T_i^\mu(g) = \bigoplus_{\mu\in R(G)} \bigoplus_{i=1}^{a^\mu} T^\mu(g), \qquad g \in G, \tag{3.3}$$

where (2.5) is used. Note that $T^\mu(g)$ is a square matrix of order N^μ. If we accordingly partition the column set of H as

$$H = (H_i^\mu \mid \mu \in R(G),\ i = 1,\cdots,a^\mu) \tag{3.4}$$

with H_i^μ being an $N \times N^\mu$ matrix, we have

$$T(g)H_i^\mu = H_i^\mu T^\mu(g). \tag{3.5}$$

With reference to (3.4) we put

$$H^\mu = (H_i^\mu \mid i = 1, \cdots, a^\mu), \quad \mu \in R(G);$$

then

$$H = (H^\mu \mid \mu \in R(G)). \tag{3.6}$$

A block-diagonal matrix $\tilde{J}$ is obtained by transforming J with H:

$$\tilde{J} = H^* J H.$$

If we define

$$\tilde{J}_{ij}^{\mu\nu} = (H_i^\mu)^* J H_j^\nu,$$

$\tilde{J}$ is partitioned consistently with (3.3):

$$\tilde{J} = (\tilde{J}_{ij}^{\mu\nu} \mid \mu, \nu \in R(G);\ i = 1, \cdots, a^\mu; j = 1, \cdots, a^\nu).$$

Then Schur's lemma (i) implies that

$$\tilde{J}_{ij}^{\mu\nu} = O \qquad \text{if } \mu \neq \nu, \tag{3.7}$$

since $(H^*T(g)H)\tilde{J} = \tilde{J}(H^*T(g)H)$ by (3.2). It follows that

$$\tilde{J} = \bigoplus_{\mu \in R(G)} \tilde{J}^\mu, \tag{3.8}$$

where

$$\tilde{J}^\mu = (\tilde{J}_{ij}^{\mu\mu} \mid i, j = 1, \cdots, a^\mu), \qquad \mu \in R(G),$$

is a square matrix of order $a^\mu N^\mu$.

Each diagonal block $\tilde{J}^\mu$ can have a further structure by virtue of our choice (2.5) and Schur's lemma (iii). That is, if T^μ is absolutely irreducible, then

$$\tilde{J}_{ij}^{\mu\mu} = \alpha_{ij}^\mu I_{N^\mu}, \qquad i, j = 1, \cdots, a^\mu,$$

for some $\alpha_{ij}^\mu \in F$. If we define

$$A^\mu = (\alpha_{ij}^\mu \mid i, j = 1, \cdots, a^\mu), \tag{3.9}$$

this means that $\tilde{J}^\mu$ is in a block-diagonal form (after suitable arrangement of rows and columns) which has A^μ as its N^μ identical diagonal blocks of size a^μ:

$$\tilde{J}^\mu = \bigoplus_{l=1}^{N^\mu} A^\mu.$$

To sum up, we obtain

$$\tilde{J} = H^* J H = \Big(\bigoplus_{\mu \in R_a(G)} \bigoplus_{l=1}^{N^\mu} A^\mu \Big) \oplus \Big(\bigoplus_{\mu \in R(G) - R_a(G)} \tilde{J}^\mu \Big), \tag{3.10}$$

where $R_a(G)$ denotes the subsets of $R(G)$ consisting of absolutely irreducible representations of G. This type of block-diagonalization, as well as the one of (3.8) corresponding to the isotypic decomposition (2.1), is well known in many application areas. In bifurcation theory, as shown by [5], (3.8) plays an essential role in the analysis of bifurcation at simple points. This principle is also used in quantum mechanics [7], [14] in deriving "selection rules".

Obviously, J is singular iff A^μ or $\tilde{J}^\mu$ is singular for some $\mu \in R(G)$. Hence, instead of testing for the singularity of J, of size N, we may test for the singularity of $|R(G)|$ smaller matrices: A^μ, of size a^μ, for $\mu \in R_a(G)$ and $\tilde{J}^\mu$, of size $a^\mu N^\mu$, for $\mu \in R(G) - R_a(G)$. This procedure involves an extra work of computing the matrices A^μ and $\tilde{J}^\mu$; the amount of computation for it depends on the choice of H. The following subsection shows a natural choice of H, which results in an overall improvement in efficiency.

3.3 "Good" basis

This subsection suggests a choice of the matrix H in (3.3) and (3.10) for a class of bifurcation problems including those arising from truss structures.

Let us recall the spherical diamond shell of Figure 1, and suppose that the steady state of this structure is described in terms of displacements of the nodes. We denote by P the set of the free nodes; $|P| = 31$. As a consequence of the geometric symmetry of this structure, the bifurcation problem is equivariant to the dihedral group D_6, if the physical characteristics and the loading pattern have also D_6 symmetry; $r \in D_6$ corresponds to the anti-clockwise rotation around Z-axis at angle $\pi/3$ and $s \in D_6$ to the reflection with respect to XZ-plane. The space V of the displacements can be regarded as the tensor product of $\mathbf{R}^P$ and the 3 dimensional Euclidean space $\mathbf{R}^3$, and the action of D_6 on V conforms with this tensor product. That is, the representation T of D_6 on V is the tensor product of a permutation representation T_P on P (or on $\mathbf{R}^P$) and a representation T_E on $\mathbf{R}^3$.

On the basis of this observation, we shall discuss in a rather general framework in this subsection. The following section will give a more concrete description for truss structures with D_n-symmetry.

Now we shall assume for (3.1) in general that

$$V = \mathbf{R}^P \otimes \mathbf{R}^E \tag{3.11}$$

for some finite sets P and E, and that

$$T = T_P \otimes T_E,$$

where T_P is a permutation representation of G on P and T_E a representation of G on $\mathbf{R}^E$. Note that the matrix T is of size $N = |P| \times |E|$ and each row (or column) set is indexed by $P \times E$.

Let

$$P = \bigcup_{k \in K} P_k \tag{3.12}$$

be the decomposition of P into disjoint orbits with respect to the action of G. Obviously, $\mathbf{R}^{P_k}$ is a G-invariant subspace for each k. Also let

$$E = \bigcup_{l \in L} E_l$$

be a decomposition of E into disjoint subsets such that $\mathbf{R}^{E_l}$ is a G-invariant subspace for each l. Then $P \times E$ is partitioned into disjoint sets as

$$P \times E = \bigcup_{k \in K} \bigcup_{l \in L} P_k \times E_l = \bigcup_{\kappa} (P \times E)_\kappa. \tag{3.13}$$

Here and henceforth, we put

$$\kappa = (k, l), \qquad (P \times E)_\kappa = P_k \times E_l,$$

and the summation over κ will mean that for $k \in K$ and $l \in L$. Put

$$N_\kappa = |(P \times E)_\kappa| = |P_k| \times |E_l|.$$

Since $V_\kappa = \mathbf{R}^{P_k} \otimes \mathbf{R}^{E_l}$ is G-invariant, we have a block-diagonal decomposition of T:

$$T(g) = \bigoplus_{\kappa} T_\kappa(g), \qquad g \in G.$$

Note that this decomposition is of combinatorial nature, in that it is obtained only through partition of the set of coordinates without involving rotation of coordinate axes.

Each diagonal block T_κ of T is to be decomposed further into irreducible components through a suitable "local" change of the basis of V_κ. If we denote by H_κ the unitary matrices of size N_κ representing such "local" transformations, we have (cf. (3.3), (3.4), (3.5))

$$(H_\kappa)^* T_\kappa(g) H_\kappa = \bigoplus_{\mu \in R(G)} \bigoplus_{i=1}^{a_\kappa^\mu} T^\mu(g), \qquad g \in G.$$

The whole transformation matrix H is a block-diagonal matrix:

$$H = \bigoplus_{\kappa} H_\kappa \tag{3.14}$$

and

$$H^* T(g) H = \bigoplus_{\kappa} \bigoplus_{\mu \in R(G)} \bigoplus_{i=1}^{a_\kappa^\mu} T^\mu(g) = \bigoplus_{\mu \in R(G)} \bigoplus_{i=1}^{a^\mu} T^\mu(g), \qquad g \in G,$$

where

$$a^\mu = \sum_{\kappa} a_\kappa^\mu.$$

Then, as explained in the previous subsection, $\tilde{J} = H^* J H$ is in a block-diagonal form (3.10) with $|R(G)|$ distinct blocks; A^μ is of size a^μ, and $\tilde{J}^\mu$ is of size $a^\mu N^\mu$.

It should be clear that the partition (3.13) above does not put $\tilde{J}$ into a block-diagonal form. That is, the submatrix $\tilde{J}_{\kappa\lambda}$ of $\tilde{J}$ with row-set $(P \times E)_\kappa$ and column-set $(P \times E)_\lambda$ does not vanish in general even if $\kappa \neq \lambda$, since T_κ and T_λ may have some irreducible components in common.

Nevertheless, $\tilde{J}_{\kappa\lambda}$ is likely to vanish for many pairs of (κ, λ) in practical applications. This is not because of the group symmetry but because of the sparsity of J and the locality of our transformation H. In fact, if $J_{\kappa\lambda} = O$ (where $J_{\kappa\lambda}$ is the submatrix of J with the same row/column subsets as $\tilde{J}_{\kappa\lambda}$), then $\tilde{J}_{\kappa\lambda} = (H_\kappa)^* J_{\kappa\lambda} H_\lambda = O$. This means that each A^μ or $\tilde{J}^\mu$ in (3.10) inherits the sparsity of J in the sense that its (κ, λ) block vanishes if $J_{\kappa\lambda} = O$.

3.4 Computational cost

We now give a rough estimate of the computational efficiency of the above procedure when each irreducible representation of G is absolutely irreducible.

The computational cost for obtaining the nonzero elements of $\tilde{J}$ from J is estimated as follows. For each (κ, λ), the distinct nonzero elements of $\tilde{J}_{\kappa\lambda} = (H_\kappa)^* J_{\kappa\lambda} H_\lambda$ can be computed with

$$\sum_{\mu \in R(G)} (a_\kappa^\mu N_\kappa N_\lambda + a_\kappa^\mu N_\lambda a_\lambda^\mu)$$

multiplications and additions. Hence all the elements of $\tilde{J}$ can be computed with

$$\begin{aligned} &\sum_\kappa \sum_\lambda \sum_{\mu \in R(G)} (a_\kappa^\mu N_\kappa N_\lambda + a_\kappa^\mu N_\lambda a_\lambda^\mu) \\ &= \sum_{\mu \in R(G)} \left((N + a^\mu)(\sum_\kappa a_\kappa^\mu N_\kappa) \right) \end{aligned} \tag{3.15}$$

operations, where

$$\sum_\kappa N_\kappa = N, \qquad \sum_\kappa a_\kappa^\mu = a^\mu$$

are used.

Assuming the Gaussian elimination for testing for singularity of a matrix, the overall computational cost of the proposed procedure is estimated by

$$\sum_{\mu \in R(G)} \left((N + a^\mu)(\sum_\kappa a_\kappa^\mu N_\kappa) + \frac{1}{3}(a^\mu)^3 \right). \tag{3.16}$$

In the case of D_n-symmetric structures, we can derive an explicit form for this expression. Note also that if J is known to be a symmetric matrix, then the above estimate reduces to its half.

The cost (3.15) for computing $\tilde{J}$ is of the order of N^2 (if $|G|$ and $|E|$ are of constant order). To see this, first observe

$$\sum_{\mu} a_{\kappa}^{\mu} \leq \sum_{\mu} a_{\kappa}^{\mu} N^{\mu} = N_{\kappa} = |P_k||E_l| \leq |G||E|, \quad \kappa = (k,l).$$

It then follows that

$$\begin{aligned} \sum_{\mu}(N + a^{\mu})(\sum_{\kappa} a_{\kappa}^{\mu} N_{\kappa}) &\leq 2N \sum_{\mu}\sum_{\kappa} a_{\kappa}^{\mu} N_{\kappa} = 2N \sum_{\kappa}(\sum_{\mu} a_{\kappa}^{\mu}) N_{\kappa} \\ &\leq 2N|G||E| \sum_{\kappa} N_{\kappa} = 2|G||E|N^2. \end{aligned}$$

It is to be emphasized that this bound is due to our choice of "local" transformation H; if H were a dense matrix, the computation of $\tilde{J}$ alone would involve $O(N^3)$ operations.

In the above argument we have assumed that all the entries of J are computed first. In many applications (e.g., truss structures, finite element method, or other discretization schemes), however, the matrix J is often sparse, being an assembly of $O(N)$ elementary matrices. Then the computation of $\tilde{J}$ can be done more efficiently, compatibly with such assembly process, as follows.

Let $\bar{J}^e$ be the elementary matrix, say, of size N^e, corresponding to element e; we denote by J^e the $N \times N$ matrix obtained by augmenting $\bar{J}^e$ with zeros. (In the case of truss structures we have $N^e \leq 6$.) We assume that J is expressed as

$$J = \sum_{e=1}^{M} J^e$$

with M elementary matrices. Referring to H_{κ} in (3.14), we denote by $\hat{H}_{\kappa}$ the $N \times N$ matrix obtained by augmenting H_{κ} with zeros; then (3.14) is written as $H = \sum_{\kappa} \hat{H}_{\kappa}$. We partition the column set of H_{κ} as

$$H_{\kappa} = (H_{\kappa}^{\mu} \mid \mu \in R(G)),$$

where H_{κ}^{μ} denotes the $N \times a_{\kappa}^{\mu} N^{\mu}$ submatrix corresponding to the representation μ. Accordingly we express $\hat{H}_{\kappa}$ as

$$\hat{H}_{\kappa} = (\hat{H}_{\kappa}^{\mu} \mid \mu \in R(G))$$

with $N \times a^{\mu} N^{\mu}$ submatrices $\hat{H}_{\kappa}^{\mu}$, which is obtained by augmenting H_{κ}^{μ} with zeros. Using these expressions we have

$$\begin{aligned} \tilde{J} &= H^* J H \\ &= \sum_{e=1}^{M}\sum_{\kappa}\sum_{\kappa'}(\hat{H}_{\kappa})^* J^e \hat{H}_{\kappa'} \\ &= \sum_{e=1}^{M}\sum_{\kappa}\sum_{\kappa'}((\hat{H}_{\kappa}^{\mu})^* J^e \hat{H}_{\kappa'}^{\mu'} \mid \mu, \mu' \in R(G)) \\ &= \bigoplus_{\mu \in R(G)} (\sum_{e=1}^{M}\sum_{\kappa}\sum_{\kappa'}(\hat{H}_{\kappa}^{\mu})^* J^e \hat{H}_{\kappa'}^{\mu}). \end{aligned} \tag{3.17}$$

The last equality follows from (3.7); note also that $(\hat{H}^{\mu}_{\kappa})^* J^e \hat{H}^{\mu'}_{\kappa'} \neq O$ in general. Since $\bar{J}^e$ is small in size, there are only small number (at most $(N^e)^2$) of pairs (κ, κ') that contribute to this sum. For each $(e, \mu, \kappa, \kappa')$, the contribution of the matrix product $(\hat{H}^{\mu}_{\kappa})^* J^e \hat{H}^{\mu}_{\kappa'}$ to A^{μ} can be computed with $a^{\mu}_{\kappa} N^e (N^e + a^{\mu}_{\kappa'})$ operations. Thus, if a^{μ}_{κ} and N^e are small, being independent of N and $|G|$ (which is often the case), we can compute $\tilde{J}$ with $O(M|G|)$ operations. This implies that the first term in (3.16) could be replaced by a term of $O(N|G|)$ if $M = O(N)$. See [1], [11] and [15].

The sparsity of J is partially inherited to $\tilde{J}$ (or to A^{μ}), as has been mentioned at the end of §3.3. There are many computational techniques to exploit sparsity in computing determinants and hence the second term $(a^{\mu})^3/3$ in (3.16) could be reduced considerably.

4 Axisymmetric Truss Structures

The method discussed in §3 is applied to axisymmetric truss structures, which enjoy both equivariance to D_n and sparsity. It is noteworthy in connection to the block-diagonalization (3.10) of J that each irreducible representation of D_n is absolutely irreducible, i.e., $R_a(D_n) = R(D_n)$.

We are mainly interested in the comparison of the computational efficiency of the proposed (block-diagonalization) method and that of the conventional (direct) method for computing determinants or for solving linearized equations. We assume that the Jacobian matrix J, of size N, is symmetric. As will be explained below, the computational efficiency of the proposed method, when applied to truss structures, stems from the following properties:

(P1) The transformation matrix H can be chosen as a block-diagonal matrix, representing a "local" coordinate change.
(P2) The Jacobian matrix J is often sparse, being an assembly of element stiffness matrices.
(P3) The band structure of J is inherited to each block $\tilde{J}^{\mu}$, again due to the "locality" of H.

The dominant computational cost in the direct method is the cost (a) to sweep out J by the Cholesky method. If we denote by B the half band width of J, we may estimate this cost by

$$B^2(N - B) + \frac{1}{6}B^3 \tag{4.1}$$

in terms of the number of arithmetic operations.

The major computational cost of the proposed method consists of the cost (b) to compute the diagonal blocks of $\tilde{J}$ of (3.10), and the cost (c) to sweep out $\tilde{J}^{\mu}$ by the Cholesky method for $\mu \in R(D_n)$. As has been observed at the end of §3.3 for the general case, we can reduce the cost (b) significantly by taking advantage of sparsity (P2) of J as follows.

Let us denote by P the set of free nodes, the number of which equals $N/3$, and divide it into $|K|$ disjoint orbits P_k, $k \in K$, as in §3.3. That is, each P_k consists of

those nodes which can be transformed to one another by the geometric transformations represented by the elements of D_n. We classify the orbits into four types labeled 0, $1M$, $1V$, and 2, as depicted in Figure 2. We partition J as

$$J = (J_{pq} \mid p, q \in P), \tag{4.2}$$

where J_{pq} is a 3×3 matrix corresponding to a pair (p, q) of nodes. For convenience, we refer to the set of fixed nodes also as an orbit.

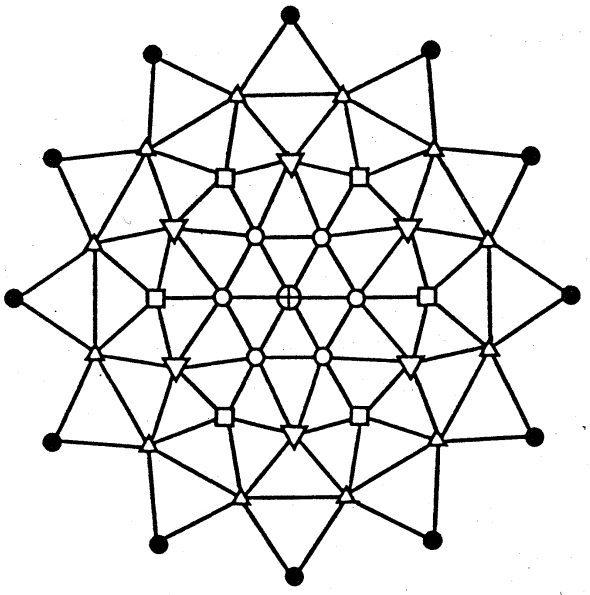


Fig.2. Orbital decomposition of the spherical diamond shell

Suppose that the truss structure is composed of M elements (members). Consider an element e that connects a free node $\bar{p}(e)$ and another node $\bar{q}(e)$ (free or fixed) which belong to orbits $\bar{k}(e)$ and $\bar{l}(e)$, respectively. We sometimes use the short hand notations: $\bar{p} = \bar{p}(e)$, $\bar{q} = \bar{q}(e)$, $\bar{k} = \bar{k}(e)$, and $\bar{l} = \bar{l}(e)$. Let J^e denote the $N \times N$ matrix that corresponds to the member stiffness matrix $\bar{J}^e$; J^e is partitioned as

$$J^e = (J^e_{pq} \mid p, q \in P) \tag{4.3}$$

in accordance with (4.2). If nodes $\bar{p}(e)$ and $\bar{q}(e)$ are free, J^e_{pq} is given by

$$J^e_{pq} = \begin{cases} \hat{J}^e & \text{if } (p,q) = (\bar{p}(e), \bar{p}(e)) \text{ or } (\bar{q}(e), \bar{q}(e)) \\ -\hat{J}^e & \text{if } (p,q) = (\bar{p}(e), \bar{q}(e)) \text{ or } (\bar{q}(e), \bar{p}(e)) \\ O & \text{otherwise} \end{cases}$$

for some 3×3 symmetric matrix $\hat{J}^e$; and if nodes $\bar{q}(e)$ is fixed (and $\bar{p}(e)$ is free)

$$J^e_{pq} = \begin{cases} \hat{J}^e & \text{if } (p,q) = (\bar{p}(e), \bar{p}(e)) \\ O & \text{otherwise.} \end{cases}$$

The whole Jacobian matrix J is expressed as

$$J = \sum_{e=1}^{M} J^e. \tag{4.4}$$

Substituting (4.4) and (3.6) into (3.10) we see

$$\tilde{J}^\mu = \sum_{e=1}^{M} \tilde{J}^{\mu e},$$

where

$$\tilde{J}^{\mu e} = (H^\mu)^* J^e H^\mu. \tag{4.5}$$

If we partition H^μ, of size $N \times N^\mu$, into blocks as

$$H^\mu = (H^\mu_{pk} \mid p \in P,\ k \in K) \tag{4.6}$$

with each block H^μ_{pk}, of size $3 \times N^\mu_k$, corresponding to node p and orbit k, our choice of H suggested in §3.3 (see also [15]) enjoys the "locality" in the sense that

$$H^\mu_{pk} = O \quad \text{if } p \notin P_k.$$

Substituting (4.3) and (4.6) into (4.5), we obtain the expression

$$\tilde{J}^{\mu e} = (\tilde{J}^{\mu e}_{kl} \mid k, l \in K)$$

with

$$\tilde{J}^{\mu e}_{kl} = \sum_{p=\bar{p}(e),\bar{q}(e)} \sum_{q=\bar{p}(e),\bar{q}(e)} (H^\mu_{pk})^* J^e_{pq} H^\mu_{ql}. \tag{4.7}$$

If $\bar{p}(e)$ and $\bar{q}(e)$ are free nodes and belong to different orbits $\bar{k}(e)$ and $\bar{l}(e)$ (i.e., $\bar{k}(e) \neq \bar{l}(e)$), (4.7) is evaluated to

$$\tilde{J}^{\mu e}_{kl} = \begin{cases} (H^\mu_{\bar{p}\bar{k}})^* \hat{J}^e H^\mu_{\bar{p}\bar{k}} & \text{if } (k,l) = (\bar{k}(e), \bar{k}(e)) \\ -(H^\mu_{\bar{p}\bar{k}})^* \hat{J}^e H^\mu_{\bar{q}\bar{l}} & \text{if } (k,l) = (\bar{k}(e), \bar{l}(e)) \\ -(H^\mu_{\bar{q}\bar{l}})^* \hat{J}^e H^\mu_{\bar{p}\bar{k}} & \text{if } (k,l) = (\bar{l}(e), \bar{k}(e)) \\ (H^\mu_{\bar{q}\bar{l}})^* \hat{J}^e H^\mu_{\bar{q}\bar{l}} & \text{if } (k,l) = (\bar{l}(e), \bar{l}(e)) \\ O & \text{otherwise.} \end{cases} \tag{4.8}$$

If both nodes are free and belong to the same orbit (i.e., $\bar{k}(e) = \bar{l}(e)$),

$$\tilde{J}^{\mu e}_{kl} = \begin{cases} (H^{\mu}_{\bar{p}\bar{k}} - H^{\mu}_{\bar{q}\bar{k}})^* \hat{J}^e (H^{\mu}_{\bar{p}\bar{k}} - H^{\mu}_{\bar{q}\bar{k}}) & \text{if } (k,l) = (\bar{k}(e), \bar{k}(e)) \\ O & \text{otherwise.} \end{cases} \tag{4.9}$$

If node $\bar{q}(e)$ is fixed (and $\bar{p}(e)$ is free),

$$\tilde{J}^{\mu e}_{kl} = \begin{cases} (H^{\mu}_{\bar{p}\bar{k}})^* \hat{J}^e H^{\mu}_{\bar{p}\bar{k}} & \text{if } (k,l) = (\bar{k}(e), \bar{k}(e)) \\ O & \text{otherwise.} \end{cases} \tag{4.10}$$

The cost (b) for computing all $\tilde{J}^{\mu}$ for $\mu \in R(D_n)$ is expressed as

$$\sum_{e=1}^{M} \sum_{\mu \in R(D_n)} C^{\mu e} \tag{4.11}$$

where $C^{\mu e}$ is the cost for computing $\tilde{J}^{\mu e}$ for the member e according to (4.8) to (4.10). If the nodes $\bar{p}(e)$ and $\bar{q}(e)$ are free and belong to different orbits $\bar{k}(e)$ and $\bar{l}(e)$, $C^{\mu e}$ is estimated by

$$C^{\mu e} = 3\{\gamma^{\mu}_{k}(N^{\mu}_{k})^2 + \gamma^{\mu}_{l}(N^{\mu}_{l})^2 + \min(\gamma^{\mu}_{k}, \gamma^{\mu}_{l}) N^{\mu}_{k} N^{\mu}_{l} + 3\gamma^{\mu}_{k} N^{\mu}_{k} + 3\gamma^{\mu}_{l} N^{\mu}_{l}\}, \tag{4.12}$$

where N^{μ}_{k} denotes the number of the columns of H^{μ}_{pk}, γ^{μ}_{k} indicates the density of H^{μ}_{pk}, i.e., the ratio of the number of non-zero entries of H^{μ}_{pk} to $3N^{\mu}_{k}$. If nodes $\bar{p}(e)$ and $\bar{q}(e)$ belong to the same orbit $\bar{k}(e)$ or if node $\bar{q}(e)$ is fixed, we have

$$C^{\mu e} = 3\{\gamma^{\mu}_{\bar{k}}(N^{\mu}_{\bar{k}})^2 + 3\gamma^{\mu}_{\bar{k}} N^{\mu}_{\bar{k}}\}.$$

Tables 1 and 2 list the values of N^{μ}_{k} and γ^{μ}_{k} obtained by [15]. Table 3 lists the computational cost for $\tilde{J}^{\mu e}$ estimated by (4.11) with Tables 1 and 2.

As for the cost (c) we have the following estimate, similar to (4.1):

$$\sum_{\mu \in R(D_n)} \{(B^{\mu})^2 (N^{\mu} - B^{\mu}) + \frac{1}{6}(B^{\mu})^3\}, \tag{4.13}$$

where B^{μ} denotes the half band width of $\tilde{J}^{\mu}$. Note that $\tilde{J}^{\mu}$ is of size

$$N^{\mu} = \sum_{k \in K} N^{\mu}_{k}.$$

It is remembered that the use of parallel computers will further reduce the costs for (b) and (c) to a great extent, thereby enhancing the efficiency of the proposed method. With the aid of more than $n/2$ computing units, these blocks are computed in parallel and the cost (c) will be reduced up to approximately $2/n$ times as what is presented above. Moreover, one can even split the computation of each block into parts and reduce the cost (b) further. Supported with infinite parallel units, the cost (b) would be almost nullified.

Table 1 Column-size N_k^μ of H_{pk}^μ.

Type of Orbit k	μ					
	s1	s2	s3	s4	d1	dj ($j \geq 2$)
0	1	0	0	0	1	0
$1M$	2	1	1	2	3	3
$1V$	2	1	2	1	3	3
2	3	3	3	3	6	6

Table 2 Density γ_k^μ of H_{pk}^μ.

Type of Orbit k	μ					
	s1	s2	s3	s4	d1	dj ($j \geq 2$)
0	1/3	0	0	0	2/3	0
$1M$	1/2	2/3	2/3	1/2	5/9	5/9
$1V$	1/2	2/3	1/2	2/3	5/9	5/9
2	5/9	5/9	5/9	5/9	5/9	5/9

Table 3 Computational cost for a truss member.

(a) $\bar{k}(e) = \bar{l}(e)$ or $\bar{q}(e)$ is fixed

Type of $\bar{k}(e)$	$C^{\mu e}$						$\sum_\mu C^{\mu e}$	
	s1	s2	s3	s4	d1	dj ($j \geq 2$)	n =even	n =odd
0	4	0	0	0	8	0	12	12
$1M$	15	8	8	15	30	30	$15n$+16	$15n$+8
$1V$	15	8	15	8	30	30	$15n$+16	$15n$+8
2	30	30	30	30	90	90	$45n$+30	$45n$+15

(b) $\bar{k}(e) \neq \bar{l}(e)$

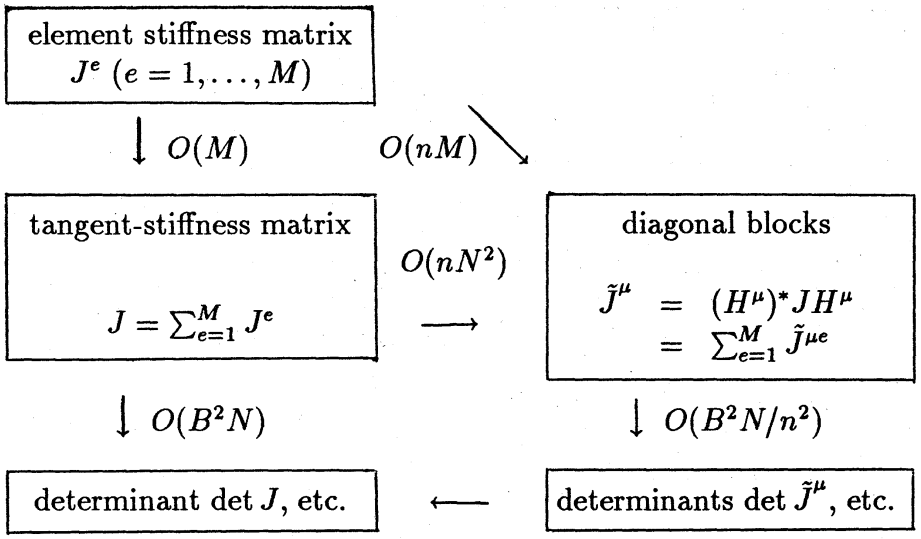
Type of $(\bar{k}(e), \bar{l}(e))$	$C^{\mu e}$						$\sum_\mu C^{\mu e}$	
	s1	s2	s3	s4	d1	dj ($j \geq 2$)	n =even	n =odd
(0,0)	9	0	0	0	18	0	27	27
$(1M, 1M)$	36	18	18	36	75	75	(75/2)n+33	(75/2)n+33/2
$(1V, 1V)$	36	18	36	18	75	75	(75/2)n+33	(75/2)n+33/2
(2,2)	75	75	75	75	240	240	120n+60	120n+30
$(0, 1M)$	21	8	8	15	43	30	15n+35	15n+27
$(0, 1V)$	21	8	15	8	43	30	15n+35	15n+27
(0,2)	37	30	30	30	108	90	45n+55	45n+40
$(1M, 1V)$	36	18	26	26	75	75	(75/2)n+31	(75/2)n+33/2
$(1M, 2)$	54	43	43	54	150	150	75n+44	75n+22
$(1V, 2)$	54	43	54	43	150	150	75n+44	75n+22

s3 and s4 apply only for n =even.

Besides the efficiency in computation time the procedure above achieves the economy in memory space. It works with the element stiffness-matrices $\bar{J}^e$ (consisting of $\hat{J}^e$ of size 3) and the diagonal blocks $\tilde{J}^\mu$ of size a^μ ($\mu \in R(D_n)$), thereby avoiding

the explicit construction of the matrix J of size N. Moreover, $\tilde{J}^{\mu}$ inherits the band structure and can be stored as such.

As the conclusion we schematize the computational flow below.



References

[1] A. Bossavit: Symmetry, groups, and boundary value problems — A progressive introduction to noncommutative harmonic analysis of partial differential equations in domains with geometrical symmetry, *Computer Methods in Applied Mechanics and Engineering*, Vol. 56 (1986), 167–215.

[2] H.-C. Chen and A. H. Sameh: A matrix decomposition method for orthotropic elasticity problems, *SIAM Journal on Matrix Analysis and Applications*, Vol. 10 (1989), 39–64.

[3] C. W. Curtis and I. Reiner: *Representation Theory of Finite Groups and Associative Algebras*, John Wiley (Interscience), 1962.

[4] M. Dellnitz and B. Werner: Computational methods for bifurcation problems with symmetries — with special attention to steady state and Hopf bifurcation points, *Journal of Computational and Applied Mathematics*, Vol. 26 (1989), 97–123.

[5] H. Fujii and M. Yamaguti: Structure of singularities and its numerical realization in nonlinear elasticity, *Journal of Mathematics of Kyoto University*, Vol. 20 (1980), 489–590.

[6] M. Golubitsky, I. Stewart and D. G. Schaeffer: *Singularities and Groups in Bifurcation Theory*, Vol. 2, Springer, 1988.

[7] M. Hammermesh: *Group Theory and Its Application to Physical Problems*, Addison-Wesley, 1962.

[8] T. J. Healey: *Symmetry, Bifurcation and Computational Methods in Nonlinear Structural Mechanics*, Ph. D. Thesis, Univ. Illinois, 1985.

[9] T. J. Healey: A group-theoretic approach to computational bifurcation problems with symmetry, *Computer Methods in Applied Mechanics and Engineering*, Vol. 67 (1988), 257–295.

[10] T. J. Healey: Numerical bifurcation with symmetry: diagnosis & computation of singular points, *Proceedings of the International Conference on Bifurcation Theory and Its Numerical Analysis*, L. Kaitai, J. Marsden, M. Golubitsky and G. Iooss, eds., Xi'an Jiaotong University Press, 1989.

[11] K. Ikeda and K. Murota: Bifurcation analysis of axisymmetric structures using block-diagonalization, Technical Report METR 89-15, Department of Mathematical Engineering, University of Tokyo, 1989.

[12] N. Jacobson: *Basic Algebra II*, Freeman, San Francisco, 1980.

[13] D. J. Klein, C. H. Carlisle and F. A. Matsen: Symmetry adaptation to sequences of finite groups, in *Advances in Quantum Chemistry* (P. Löwdin, ed.), Vol. 5 (1970), Academic, 219–260.

[14] W. Miller, Jr.: *Symmetry Groups and Their Applications*, Academic, 1972.

[15] K. Murota and K. Ikeda: Computational use of group theory in bifurcation analysis of symmetric structures, *SIAM Journal on Scientific and Statistical Computing*, to appear.

[16] D. H. Sattinger: *Group Theoretic Methods in Bifurcation Theory*, Lecture Notes in Mathematics, 762, Springer, 1979.