Transformations into optimal parallelism in euclidean spaces (or: how to explain the shape of the electron-density distribution inside a crystal)

Ehrhard Behrends, Fritz Mädler

Abstract.

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ be vectors in the finite dimensional euclidean space $\mathbb{R}^d$. We investigate the problem of how one can find a U such that $\sum_i \langle U\mathbf{x}_i, \mathbf{y}_i \rangle$ is maximal when U runs through the orthogonal group or the special orthogonal group, i.e., we are looking for a U such that the $U\mathbf{x}_1, \ldots, U\mathbf{x}_n$ and the $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are "as parallel as possible".

For $d = 3$ this problem arises, for instance, from the data analysis of crystallographic diffraction experiments on orientationally disordered systems: the $\mathbf{x}$'s stand for the atom positions of fragments M_1 of the crystal structure, the $\mathbf{y}$'s are taken from the set M_2 of maxima of the electron-density distribution reconstructed from diffraction data, and one must know the above transformations U in order to determine the M_1-M_2-configurations of minimum distance since they are responsible for the shape of the density distribution.

It turns out that one can associate a $d \times d$-matrix G with this problem in such a way that the relevant U are precisely those for which the trace of GU is maximal. Using this transformation we are able to provide an explicit solution by means of the singular value decomposition of G.

Several further topics in connection with this problem are also discussed. In particular we investigate in which cases the optimal position is unique, and we study generalizations to nondiscrete situations and to the infinite dimensional setting.

AMS-classification: 15A18, 46C05, 47B10, 82D25.

keywords: Euclidean space, (special) orthogonal group, singular value decomposition, nuclear operator.

1. Introduction

We provide the finite dimensional $\mathbb{R}^d$ with the usual euclidean scalar product. Given families of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ we are interested in mappings of the form $\mathbf{x} \mapsto \mathbf{a} + U\mathbf{x}$ which transform the $\mathbf{x}_i$ such that the expression

$$\sum_i ||\mathbf{a} + U\mathbf{x}_i - \mathbf{y}_i||^2 \tag{1.1}$$

is as small as possible. Here $\mathbf{a}$ might be any vector and U an arbitrary orthogonal transformation. (Sometimes it is reasonable to restrict oneself to the orthogonal U with determinant one.)

This problem plays a role in crystallography for the following reasons. If molecules or complex ions are located in different orientations in the crystal structure, the investigation of the spatial electron-density distribution (EDD) can reveal their static or dynamic ambiguity. To give an example, fig. 1 depicts an isosurface of the EDD inside the unit cell of cubic KPF_6 as reconstructed from a high-pressure X-ray diffraction-experiment [7]. As an important step of the analysis, the orientational non-uniqueness of the octahedral PF_6 anion M_1 has to be solved, which in this case means to give an answer to the question: how can the regular PF_6 octahedron with the *six* F anions at its corners produce the *eight* cushion-like parts of the EDD shown in fig. 1 ? The answer is depicted in fig. 2. For physical reasons the F anions $\mathbf{x}_i$ must be as close as possible to a subset $M_2 \subset M$ of the eight EDD maxima $\mathbf{y}_i$ which form the vertices of a cube. There are four minimum orientations U of PF_6 which conjointly produce the EDD, and always three F anions of different U superpose to one maximum of their conjoint centroid. Furthermore, if the $\mathbf{y}_i$'s are taken from suitable subsets M_2 of the set M of the surrounding six nearest neighbour K^+ ions, then the M_1-M_2-configurations provide predictions which are in remarkable coincidence with the results derived from the experiments. (For details, cf. [4].)

In general, the assignment $\mathbf{x}_i \mapsto \mathbf{y}_i$ is unknown and becomes part of the problem solution. Then the minimum of the numbers in (1.1) for variable $\mathbf{a}$ and U can be found by Monte-Carlo methods and evolution strategies [5]. In the present paper we are going to provide explicit solutions of (1.1) for applications where the relation $\mathbf{x}_i \mapsto \mathbf{y}_i$ is either known beforehand from expertise or can be guessed.

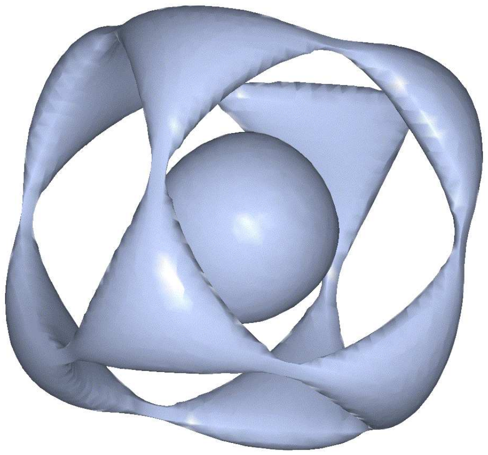


Figure 1: An equi-electron-density surface of hexafluorophosphate (PF_6) inside a cubic KPF_6 crystal. The question: how can the PF_6 octahedron with its *six* F anions produce the *eight* cushion-like parts of the electron density?

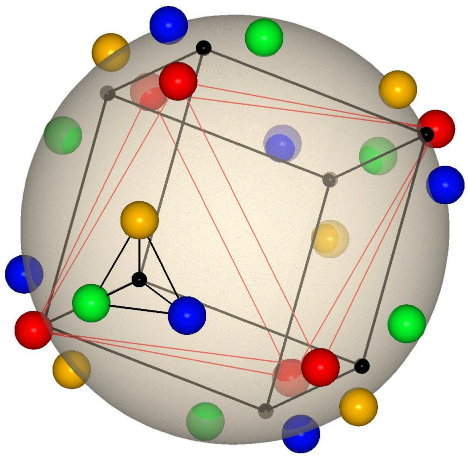


Figure 2: The answer: the density maxima form the vertices of a cube; there are four minimum orientations U for the PF_6 octahedron, and always three F anions of different U superpose to one maximum which is located in their conjoint centroid.

In [4] it was demonstrated how the optimal vector $\mathbf{a}$ can be found, and thus one may assume that $\mathbf{a} = \mathbf{0}$. Also, from

$$||U\mathbf{x} - \mathbf{y}||^2 = \langle U\mathbf{x} - \mathbf{y}, U\mathbf{x} - \mathbf{y}\rangle = ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - 2\langle U\mathbf{x}, \mathbf{y}\rangle$$

it is clear that a U minimizes the expression in (1.1) iff it maximizes

$$\sum_i \langle U\mathbf{x}_i, \mathbf{y}_i\rangle. \tag{1.2}$$

In this way we arrive at the following problem:

> Given $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ in $\mathbb{R}^d$, determine U in the orthogonal group $\mathcal{O}(d)$ (resp. in the special orthogonal group $\mathcal{SO}(d)$) such that $\sum_i \langle U\mathbf{x}_i, \mathbf{y}_i\rangle$ is maximal.

As a first step, we transform the problem slightly. With $\mathbf{u}_j$ = "the canonical j'th unit vector in $\mathbb{R}^d$" we denote by g_{jk} the number

$$g_{jk} := \sum_i \langle \mathbf{x}_i, \mathbf{u}_j\rangle \langle \mathbf{y}_i, \mathbf{u}_k\rangle \tag{1.3}$$

for $j, k = 1, \ldots, d$ and put $G := (g_{jk})$. Since $\mathbf{x} = \sum_j \langle \mathbf{x}, \mathbf{u}_j\rangle \mathbf{u}_j$ it follows for any $U = (u_{jk})$ that

$$\begin{aligned} \sum_i \langle U\mathbf{x}_i, \mathbf{y}_i\rangle &= \sum_{j,k} g_{jk} \langle U\mathbf{u}_j, \mathbf{u}_k\rangle \\ &= \sum_{j,k} g_{jk} u_{kj} \\ &= \mathrm{tr}(GU), \end{aligned}$$

where $\mathrm{tr}(M)$ denotes the *trace* of M for any square matrix M.

The entries g_{jk} of any $d \times d$-matrix G can be expressed as in (1.3) for suitable $\mathbf{x}_i$, $\mathbf{y}_i$, and hence it is natural to pass from the original problems to the following ones:

> Let G be an arbitrary real $d \times d$-matrix.
>
> *Problem 1:* Determine a U in $\mathcal{O}(d)$ such that $\mathrm{tr}(GU)$ is as large as possible.
>
> *Problem 2:* The same as problem 1, but with $\mathcal{SO}(d)$ instead of $\mathcal{O}(d)$.

We will solve these problems in *section 2* by using the singular value decomposition (*SVD*) of G. Whereas it is rather easy to determine the relevant U for problem 1, it turns out that problem 2 is more involved: if the determinant $\det(G)$ of G is negative it takes some care to show that the natural candidate $U \in \mathcal{SO}(d)$ in fact achieves the maximal $\mathrm{tr}(GU)$.

Sometimes there is no need to consider any U at all, i.e., $U = Id$ (the identity) already provides the maximal $\sum_i \langle U\mathbf{x}_i, \mathbf{y}_i\rangle$ resp. the maximal $\mathrm{tr}(GU)$. If this is the case, we say that the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$, $\mathbf{y}_1, \ldots, \mathbf{y}_n$ resp. the matrix G are *in optimal position.* In *Section 3* characterizations are derived, we also develop sufficient conditions which can easily be applied in cristallographic applications.

In *section 4* we study the connections between *symmetry assumptions* and the question to what extent the U with maximal $\mathrm{tr}(GU)$ are unique. Finally, *section 5* contains some *supplements*: a certain kind of "stability", possibilities to reduce the complexity of the problem, a short discussion of some generalizations, and a description of a random procedure to find an optimal position.

2. An explicit solution in terms of the *SVD*

In this section we fix a real $d \times d$-matrix G, by $G^\top$ we denote the transpose of G. Since $GG^\top$ is symmetric and nonnegative definite there are real numbers

$$s_1 \geq s_2 \geq \cdots \geq s_d \geq 0$$

such that the s_j^2 are the eigenvalues of $GG^\top$; the s_j are called the *singular values* of G. It is known that there are orthogonal matrices U_0 and V_0 such that

$$U_0 G V_0 = \begin{pmatrix} s_1 & 0 & 0 & \cdots & 0 \\ 0 & s_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_d \end{pmatrix}. \tag{2.1}$$

This fact can be expressed equivalently by saying that there are orthonormal bases $\mathbf{e}_1, \ldots, \mathbf{e}_d$ and $\mathbf{f}_1, \ldots, \mathbf{f}_d$ such that

$$G\mathbf{x} = \sum_j s_j \langle \mathbf{x}, \mathbf{e}_j \rangle \mathbf{f}_j \tag{2.2}$$

for every $\mathbf{x}$. A proof of the existence of this *singular value decomposition (SVD) of* G as well as methods how to calculate the matrices U_0, V_0 are contained in many standard texts on numerical analysis (see, e.g., [1] or [2]).

For our purposes the formula 2.2 is particularly useful. It implies that, for any $d \times d$-matrix U,

$$GU\mathbf{x} = \sum_j s_j \langle U\mathbf{x}, \mathbf{e}_j \rangle \mathbf{f}_j. \tag{2.3}$$

Consequently,

$$\begin{aligned} \mathrm{tr}(GU) &= \sum_k \langle \mathbf{f}_k, GU\mathbf{f}_k \rangle \\ &= \sum_{j,k} s_j \langle U\mathbf{f}_k, \mathbf{e}_j \rangle \langle \mathbf{f}_k, \mathbf{f}_j \rangle \\ &= \sum_j s_j \langle U\mathbf{f}_j, \mathbf{e}_j \rangle. \end{aligned}$$

It is now very simple to solve problem 1:

Theorem 2.1 *Let $\widehat{U} \in \mathcal{O}(d)$ be implicitly defined by $\widehat{U}\mathbf{f}_j := \mathbf{e}_j$. Then*

$$\mathrm{tr}(G\widehat{U}) = \max\{\mathrm{tr}(GU) \mid U \in \mathcal{O}(d)\},$$

and this number equals $s_1 + \cdots + s_d$.
(Alternatively, one may put $\widehat{U} := V_0 U_0$, with U_0, V_0 from the SVD *of G.)*

Proof. If U is an arbitrary orthogonal matrix, then – by the above calculations – the trace of GU is bounded from above by $s_1 + \cdots + s_d$, and in the case $U = \widehat{U}$ the trace obviously achieves this number.

For the proof of the second part denote by D the diagonal matrix with entries $s_1, \ldots, s_d$. Then, with $\widehat{U} := V_0 U_0$,

$$\begin{aligned} \mathrm{tr}(G\widehat{U}) &= \mathrm{tr}([GV_0]U_0) \\ &= \mathrm{tr}(U_0[GV_0]) \\ &= \mathrm{tr}(D) \\ &= s_1 + \cdots + s_d. \end{aligned}$$

Also, for an arbitrary orthonormal U, the trace of GU is the trace of $D\widetilde{U}$ with the orthonormal matrix $\widetilde{U} := V_0^{-1}UU_0^{-1}$ so that it is bounded by $s_1 + \cdots + s_d$. □

Note: Dirk Werner has observed that the preceding theorem can be thought of as a statement concerning a Banach space and its dual.

Let X be the real Banach space of all linear (and thus continuous) operators on $\mathbb{R}^d$, provided with the nuclear norm:

$$||G||_n := s_1 + \cdots + s_d.$$

Then, as is well-known, the dual space of X "is" the space of operators Y on the Hilbert space $\mathbb{R}^d$, provided with the usual operator norm: the linear functionals on X are precisely the mappings $\phi_H : G \mapsto \mathrm{tr}(GH)$, and ϕ_H has norm $||H||$.

The theorem asserts that $||G||_n$ is the maximum over the $\phi_U, U \in \mathcal{O}(d)$. With the help of the Hahn-Banach theorem and the "inverse" Krein-Milman theorem one may conclude that all extreme functionals must lie in $\mathcal{O}(d)$. On the other hand, it is easy to see that the $U \in \mathcal{O}(d)$ are in fact extreme points of Y: if $||U \pm A|| \leq 1$ holds, then $||Ux \pm Ax|| \leq 1$ for all normalized vectors, so that – since $||Ux|| = 1$ – the operator A must be trivial.

If it is known by some other method that $\mathcal{O}(d)$ is the extreme boundary of the unit ball of Y, then the theorem is a consequence of the Krein-Milman theorem.

The solution to problem 2 necessitates a preparation: we need a generalization of the well-known fact that – for symmetric matrices A – the largest eigenvalue is the maximum of the numbers $\langle A\mathbf{e}, \mathbf{e}\rangle$, where $\mathbf{e}$ runs through the vectors of (euclidean) norm one.

Lemma 2.2 *Let A be a real symmetric $d \times d$-matrix, by $\lambda_1 \geq \cdots \geq \lambda_d$ we denote its eigenvalues (multiple eigenvalues are counted according to their multiplicity). Then, for $1 \leq d' \leq d$,*

$$\max\{\sum_{i=1}^{d'} \langle A\mathbf{e}_i, \mathbf{e}_i\rangle \mid \mathbf{e}_1, \ldots, \mathbf{e}_{d'} \text{ is an orthonormal system}\} = \lambda_1 + \cdots + \lambda_{d'}.$$

Similarly, the minimum over the $\sum_{i=1}^{d'} \langle A\mathbf{e}_i, \mathbf{e}_i\rangle$, with $\mathbf{e}_1, \ldots, \mathbf{e}_{d'}$ an orthonormal system, equals $\lambda_{d-d'+1} + \cdots + \lambda_d$.

Note: We are sure that the assertion is no surprise to the specialists, the cases $d' = 1$ and $d' = d$ correspond to standard results which are contained in many books. However, we have found no reference for general d' in the literature.

Proof. Since the second part follows easily by passing from A to $-A$ it suffices to treat the maximum. There the inequality "$\geq$" is obviously satisfied, it suffices to choose the $\mathbf{e}_i$ as suitable eigenvectors. Thus it remains to check "$\leq$".

We may (and will) assume that A is diagonal with entries $\lambda_1, \ldots, \lambda_d$, then the claim can be restated as follows:

> Let, for $i = 1, \ldots, d'$, $\mathbf{e}_i = (x_{1i}, \ldots, x_{di})^\top$ be vectors in $\mathbb{R}^d$ such that $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$. Then, with $a_j := \sum_{i=1}^{d'} (x_{ji})^2$,
>
> $$\sum_{j=1}^{d} \lambda_j a_j \leq \lambda_1 + \cdots + \lambda_{d'}. \tag{2.4}$$

For the proof of this assertion choose $\mathbf{e}_{d'+1}, \ldots, \mathbf{e}_d$ such that the $\mathbf{e}_1, \ldots, \mathbf{e}_d$ are an orthonormal basis of $\mathbb{R}^d$. If we continue to write $\mathbf{e}_i = (x_{1i}, \ldots x_{di})^\top$ for *all* i, then the matrix $U := (x_{ji})$ has orthonormal columns. Hence it also has orthonormal rows, and we may conclude that the a_j are bounded by 1; note also that they sum up to d'.

To finish the proof, let Δ be the collection of all $(a_1, \ldots, a_d)$ such that $a_j \in [0, 1]$ for every j, and $a_1 + \cdots + a_d = d'$. (2.4) is just the statement that

$$\Phi : (a_1, \ldots, a_d) \mapsto (\lambda_1 + \cdots + \lambda_{d'}) - \sum_{j=1}^{d} \lambda_j a_j$$

is a nonnegative function. But Φ is affine, and it is surely nonnegative on the extreme points of Δ: these consist of the $(a_1, \ldots, a_d)^\top$ for which precisely d' entries are one and the others vanish.

> As a variant one may fix $(a_1, \ldots, a_d) \in \Delta$ and consider the set Λ of all $(\lambda_1, \ldots, \lambda_d)$ such that $1 \geq \lambda_1 \geq \cdots \geq \lambda_d \geq 0$. The mapping Φ is nonnegative on Λ since it has this property on the extreme points $(0, \ldots, 0)$, $(1, 0, \ldots, 0)$, $(1, 1, 0, \ldots, 0)$,... of Λ, and this is (essentially) the claim to be proved.

□

Here is our solution to problem 2:

Theorem 2.3 *The maximal possible* $\mathrm{tr}(GU)$, $U \in \mathcal{SO}(d)$, *depends on the determinant* $\det(G)$ *of* G *as follows: the maximum equals*

$$s_1 + \cdots + s_d$$

if $\det(G) \geq 0$ *and*

$$s_1 + \cdots + s_{d-1} - s_d$$

in case $\det(G) \leq 0$.

The number $\mathrm{tr}(GU)$ *is maximal for* $U = \widetilde{U}$, *where* $\widetilde{U}$ *is defined by* $\widetilde{U}\mathbf{f}_j := \mathbf{e}_j$ *for* $j = 1, \ldots, d-1$ *and* $\widetilde{U}\mathbf{f}_d := z\mathbf{e}_d$, *where* $z \in \{-1, +1\}$ *is chosen such that* $\widetilde{U} \in \mathcal{SO}(d)$.

Proof. Suppose first that $\det(G)$ is strictly positive. Then it follows, with the notation of (2.1), that $s_i > 0$ for all i and that $\det(U_0)\det(V_0) = 1$. Hence the matrix $\widehat{U}$ from theorem 2.1 lies not only in $\mathcal{O}(d)$ but in $\mathcal{SO}(d)$.

If $\det(G) = 0$, then necessarily $s_d = 0$. Thus we may change the signs in the last column of $\widehat{U}$ without affecting the value of $\mathrm{tr}(G\widehat{U})$. Either the original $\widehat{U}$ or the new one will be in $\mathcal{SO}(d)$.

Let us turn to the more interesting case when $\det(G) < 0$. Then $\widehat{U}$ from theorem 2.1 has determinant -1 so that in the definition of the matrix $\widetilde{U}$ one has to work with $z = -1$ (otherwise we would have $\widehat{U} = \widetilde{U}$). It is obvious from (2.2) that $\mathrm{tr}(G\widetilde{U}) = s_1 + \cdots + s_{d-1} - s_d$, and it remains to prove that no $U \in \mathcal{SO}(d)$ gives rise to a better result. We pass from G to $G\widetilde{U}$, with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_d$ this matrix is diagonal with entries $s_1, \ldots, s_{d-1}, -s_d$. Thus, since $\mathcal{SO}(d)$ is a group, our claim is equivalent with the assertion that the trace of

$$\begin{pmatrix} s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{d-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & -s_d \end{pmatrix} U \tag{2.5}$$

is bounded by $s_1 + \cdots + s_{d-1} - s_d$ for every $U \in \mathcal{SO}(d)$.

The idea is to describe the $U \in \mathcal{SO}(d)$ explicitly and to reduce the claim with the help of the preceding lemma to the following elementary observation:

A function of the form

$$(\phi_1, \ldots, \phi_r) \mapsto \sum_{\rho=1}^{r} a_\rho \cos \phi_\rho + b_\rho \sin \phi_\rho + c \tag{2.6}$$

assumes its maximum at $(0, \ldots, 0)$ iff all b_ρ vanish and all a_ρ are nonnegative.

Let $\mathbf{e}_1, \ldots, \mathbf{e}_d$ be an arbitrary orthonormal basis, r an integer such that $2r \leq d$, and $\phi_1, \ldots, \phi_r \in \mathbb{R}$. If $\mathbf{x} = \sum_i a_i \mathbf{e}_i$ is any vector, we put

$$\begin{aligned} U_{\phi_1,\ldots,\phi_r}(\mathbf{x}) &:= \sum_{\rho=1}^{r} (a_{2\rho-1} \cos \phi_\rho - a_{2\rho} \sin \phi_\rho)\mathbf{e}_{2\rho-1} + \\ &\qquad + (a_{2\rho-1} \sin \phi_\rho + a_{2\rho} \cos \phi_\rho)\mathbf{e}_{2\rho} + \sum_{k=2r+1}^{d} a_k \mathbf{e}_k. \end{aligned}$$

(This just means that, with respect to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_d$, the transformation $U_{\phi_1,\ldots,\phi_r}$ is represented by a matrix which consists of 2×2-rotations and – possibly – some 1's on the diagonal.)

Clearly, $U_{\phi_1,\ldots,\phi_r}$ is in $\mathcal{SO}(d)$, and it is known that every $U \in \mathcal{SO}(d)$ has this form for suitable $\mathbf{e}_1, \ldots, \mathbf{e}_d$ and $\phi_1, \ldots, \phi_r$. It has to be shown that

$$\operatorname{tr}(GU_{\phi_1,\ldots,\phi_r}) \leq s_1 + \cdots + s_{d-1} - s_d.$$

For the proof it will be convenient to define numbers s_i' by $s_i' := s_i$ for $i = 1, \ldots, d-1$, and $s_d' := -s_d$. Denote as before by $\mathbf{u}_1, \ldots, \mathbf{u}_d$ the canonical basis of $\mathbb{R}^d$ and put $\alpha_{ji} := \langle \mathbf{u}_j, \mathbf{e}_i \rangle$. Then $\mathbf{e}_i = \sum_j \alpha_{ji} \mathbf{u}_j$ for every i, and for any $\mathbf{x}$ such that $\mathbf{x} = \sum_j a_j \mathbf{u}_j$ it follows that $G\mathbf{x} = \sum_j s_j' a_j \mathbf{u}_j$.

Thus, if we fix any i, then

$$\begin{aligned} U_{\phi_1,\ldots,\phi_r}\mathbf{u}_i &= \sum_{\rho=1}^{r} (\alpha_{i,2\rho-1} \cos \phi_\rho - \alpha_{i,2\rho} \sin \phi_\rho)\mathbf{e}_{2\rho-1} + \\ &\qquad + (\alpha_{i,2\rho-1} \sin \phi_\rho + \alpha_{i,2\rho} \cos \phi_\rho)\mathbf{e}_{2\rho} + \sum_{k=2r+1}^{d} \alpha_{i,k}\mathbf{e}_k \\ &= \sum_{\rho=1}^{r} \sum_{j=1}^{d} (\alpha_{i,2\rho-1} \cos \phi_\rho - \alpha_{i,2\rho} \sin \phi_\rho)\alpha_{j,2\rho-1}\mathbf{u}_j + \\ &\qquad + (\alpha_{i,2\rho-1} \sin \phi_\rho + \alpha_{i,2\rho} \cos \phi_\rho)\alpha_{j,2\rho}\mathbf{u}_j + \sum_{k=2r+1}^{d} \sum_{j=1}^{d} \alpha_{i,k}\alpha_{j,k}\mathbf{u}_j. \end{aligned}$$

Consequently,

$$\begin{aligned} GU_{\phi_1,\ldots,\phi_r}\mathbf{u}_i &= \sum_{j=1}^{d} s'_j \mathbf{u}_j [\sum_{\rho=1}^{r}(\alpha_{i,2\rho-1}\cos\phi_\rho - \alpha_{i,2\rho}\sin\phi_\rho)\alpha_{j,2\rho-1} + \\ &\quad +(\alpha_{i,2\rho-1}\sin\phi_\rho + \alpha_{i,2\rho}\cos\phi_\rho)\alpha_{j,2\rho} + \sum_{k=2r+1}^{d}\alpha_{ik}\alpha_{jk}], \end{aligned}$$

and it follows that

$$\begin{aligned} \mathrm{tr}(GU_{\phi_1,\ldots,\phi_r}) &= \sum_i \langle \mathbf{u}_i, GU_{\phi_1,\ldots,\phi_r}\mathbf{u}_i\rangle \\ &= \sum_{\rho=1}^{r}[\sum_i s'_i(\alpha^2_{i,2\rho-1} + \alpha^2_{i,2\rho})\cos\phi_\rho] + c \\ &= \sum_{\rho=1}^{r}[\langle G\mathbf{e}_{2\rho-1}, \mathbf{e}_{2\rho-1}\rangle + \langle G\mathbf{e}_{2\rho}, \mathbf{e}_{2\rho}\rangle]\cos\phi_\rho + c \end{aligned}$$

with a number c not depending on the ϕ_ρ. By lemma 2.2 we have

$$\langle G\mathbf{e}_{2\rho-1}, \mathbf{e}_{2\rho-1}\rangle + \langle G\mathbf{e}_{2\rho}, \mathbf{e}_{2\rho}\rangle \geq s_{d-1} - s_d \geq 0,$$

and thus, since $\mathrm{tr}(GU_{\phi_1,\ldots,\phi_r})$ is of the form (2.6), the maximal value of this trace is achieved when all ϕ_ρ vanish. But

$$\mathrm{tr}(GU_{0\cdots 0}) = \mathrm{tr}(G) = s_1 + \cdots + s_{d-1} - s_d,$$

and this completes the proof. □

3. Families of vectors and matrices in optimal position

Our starting point was to determine $U \in \mathcal{O}(d)$ (or $U \in \mathcal{SO}(d)$) such that it maximizes $\sum_i \langle U\mathbf{x}_i, \mathbf{y}_i\rangle$ resp. $\mathrm{tr}(GU)$. Sometimes there is no need to calculate such a U since $U = Id$ (= the identity) already provides the best possible value. It is desirable to have conditions in order to decide when this is the case: in the applications of our results to cristallography one sometimes has a guess of what the optimal position should look like, and one wants to know whether this is in fact the case.

Definition 3.1 Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$, $\mathbf{y}_1, \ldots, \mathbf{y}_n$ and G be as above.

(i) The $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are called to be in *optimal position* (resp. in *1-optimal position*) if

$$\sum_i \langle \mathbf{x}_i, \mathbf{y}_i \rangle \geq \sum_i \langle U\mathbf{x}_i, \mathbf{y}_i \rangle$$

for every $U \in \mathcal{O}(d)$ (resp. every $U \in \mathcal{SO}(d)$).

(ii) Similarly, we say that G is in *optimal position* (resp. in *1-optimal position*) if the trace of G bounds the traces of all GU with $U \in \mathcal{O}(d)$ (resp. $U \in \mathcal{SO}(d)$).

Proposition 3.2 *Let G be any $d \times d$-matrix.*

(i) *G is in optimal position iff G is symmetric and all eigenvalues are nonnegative (i.e., if G is nonnegative definite).*

(ii) *G is in 1-optimal position iff it is symmetric and $\lambda_i + \lambda_j \geq 0$ for $i \neq j$ (as usual, $\lambda_1 \geq \cdots \geq \lambda_d$ denote the eigenvalues of G).*

Thus, if we know in advance that G is symmetric, then G is in optimal position (resp. in 1-optimal position) iff $\lambda_d \geq 0$ (resp. $\lambda_{d-1} + \lambda_d \geq 0$).

Proof. (i) Let G be symmetric with nonnegative eigenvalues λ_i. Then, with respect to a suitable orthonormal basis, we have $\mathrm{tr}(GU) = \sum_i \lambda_i \langle \mathbf{e}_i, U\mathbf{e}_i \rangle$ for every U, and from this it is immediately clear that G is in optimal position.

Conversely, let us start with a $G = (g_{ij})$ in optimal position, first we will show that G is symmetric. Fix different indices i_0, j_0 and consider a rotation U_ϕ by the angle ϕ in the subspace generated by $\mathbf{u}_{i_0}$ and $\mathbf{u}_{j_0}$. It is easy to verify that

$$\mathrm{tr}(GU_\phi) = (g_{i_0 i_0} + g_{j_0 j_0}) \cos \phi + (g_{i_0 j_0} - g_{j_0 i_0}) \sin \phi + c$$

with a constant c independent of ϕ. By assumption, $\phi = 0$ leads to the maximal value, and – since we deal with a function of the form (2.6) – we conclude that $g_{i_0 j_0} - g_{j_0 i_0} = 0$.

Thus G is symmetric, and we may assume that it is diagonal. If one of the entries were negative it would be obvious how to produce a larger trace by a suitable reflection. This completes the proof of (i).

(ii) If G is symmetric with $\lambda_{d-1} - \lambda_d \geq 0$, then we know from lemma 2.2 that

$$\langle G\mathbf{e}, \mathbf{e} \rangle + \langle G\mathbf{f}, \mathbf{f} \rangle \geq 0$$

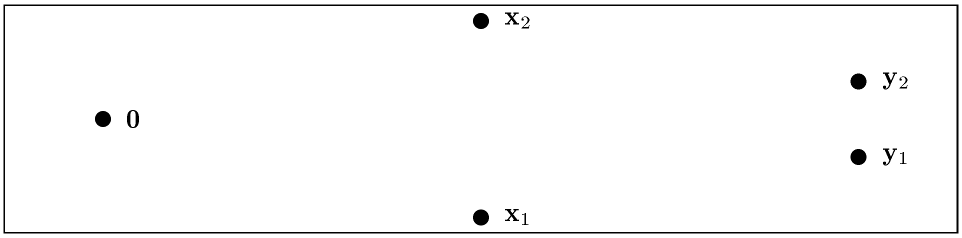
for orthonormal $\mathbf{e}$, $\mathbf{f}$. Hence, by the calculations from the proof of theorem 2.3, we have

$$\mathrm{tr}(GU_{\phi_1,\ldots,\phi_r}) = \sum_{\rho=1}^{r} a_\rho \cos\phi_\rho + c$$

with *nonnegative* a_ρ for arbitrary $U_{\phi_1,\ldots,\phi_r}$, and therefore we cannot do better than choose $\phi_1 = \cdots = \phi_r = 0$. This means that G is in 1-optimal position.

If, conversely, G is in 1-optimal position, then it follows as in the proof of (i) that G is symmetric. If $\lambda_{d-1} + \lambda_d < 0$ were true, then a rotation by the angle π in the last two coordinates – an element of $\mathcal{SO}(d)$ – would give rise to a larger trace. Contradiction. □

We now turn to families $\mathbf{x}_1,\ldots,\mathbf{x}_n$, $\mathbf{y}_1,\ldots,\mathbf{y}_n$. Consider for example the following two-dimensional situation:



It is "somehow obvious" that $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{y}_1, \mathbf{y}_2$ are in optimal position, they are already "as parallel as possible". Similar two- and three-dimensional examples abound in cristallography, here is a result by which one can often verify in a simple way that the guess is correct:

Proposition 3.3 *Let* $\mathbf{x}_1,\ldots,\mathbf{x}_n$, $\mathbf{y}_1,\ldots,\mathbf{y}_n$ *be given such that there exist* $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ *and* $\hat{\mathbf{x}}_i$, $\hat{\mathbf{y}}_i$ *such that*

- (i) $\hat{\mathbf{x}}$ *and* $\hat{\mathbf{y}}$ *are* codirectional, *i.e., one of these vectors is a positive multiple of the other (or, equivalently, in the Cauchy-Schwarz inequality one has equality).*
- (ii) $\hat{\mathbf{x}}_i$ *is codirectional with* $\hat{\mathbf{y}}_i$ *for every* i.
- (iii) $\sum_i \hat{\mathbf{x}}_i = \sum_i \hat{\mathbf{y}}_i = \mathbf{0}$.
- (iv) $\mathbf{x}_i = \hat{\mathbf{x}} + \hat{\mathbf{x}}_i$, *and* $\mathbf{y}_i = \hat{\mathbf{y}} + \hat{\mathbf{y}}_i$.

Then the $\mathbf{x}_1,\ldots,\mathbf{x}_n$, $\mathbf{y}_1,\ldots,\mathbf{y}_n$ *are in optimal position.*

Note: As an illustration consider the above example. There the obvious choice is

$$\hat{\mathbf{x}} = (\mathbf{x}_1 + \mathbf{x}_2)/2 \text{ and } \hat{\mathbf{y}} = (\mathbf{y}_1 + \mathbf{y}_2)/2.$$

Proof. Let $U \in \mathcal{O}(d)$ be arbitrary. The number

$$\begin{aligned}\sum_i \langle U\mathbf{x}_i, \mathbf{y}_i\rangle &= \sum_i \langle U(\hat{\mathbf{x}} + \hat{\mathbf{x}}_i), \hat{\mathbf{y}} + \hat{\mathbf{y}}_i\rangle \\ &= d\langle U\hat{\mathbf{x}}, \hat{\mathbf{y}}\rangle + \sum_i \langle U\hat{\mathbf{x}}_i, \hat{\mathbf{y}}_i\rangle.\end{aligned}$$

is bounded by

$$||\hat{\mathbf{x}}||||\hat{\mathbf{y}}|| + \sum_i ||\hat{\mathbf{x}}_i||||\hat{\mathbf{y}}_i||,$$

and for $U = Id$ this value is achieved. □

4. SYMMETRY

How unique is the optimal position? We put

$$\begin{aligned}\mathcal{M}_G &:= \{U \in \mathcal{O}(d) \mid \mathrm{tr}(GU) = \max_{V \in \mathcal{O}(d)} \mathrm{tr}(GV)\}, \\ \mathcal{M}_G^1 &:= \{U \in \mathcal{SO}(d) \mid \mathrm{tr}(GU) = \max_{V \in \mathcal{SO}(d)} \mathrm{tr}(GV)\}, \\ \mathcal{S}_G &:= \{U \in \mathcal{O}(d) \mid G = GU\}, \\ \mathcal{S}_G^1 &:= \{U \in \mathcal{SO}(d) \mid G = GU\}.\end{aligned}$$

Our notation is as before, in particular the numbers $s_1, \ldots, s_d$ denote the singular values in decreasing order.

Fix any $U_0 \in \mathcal{M}_G$ or in $\mathcal{M}_G^1$. Then $GU = GU_0(U_0^{-1}U)$ so that, in order to describe all U with a certain property, we may pass from G to GU_0, and hence we may assume that G is in (1-)optimal position. Similarly we can pass from G to $V_0GV_0^{-1}$ since $V_0GV_0^{-1}U$ has the same trace as $GV_0^{-1}UV_0$, and – for orthogonal V_0 – the matrix U is in $\mathcal{SO}(d)$ iff $V_0^{-1}UV_0$ is. Thus it suffices to treat the problem for diagonal G in (1-)optimal position.

The description of $\mathcal{M}_G$ is easy:

Proposition 4.1 *Let G be diagonal with entries $s_1, \ldots, s_d$. Then*

(i) *$\mathcal{M}_G = \mathcal{S}_G$;*

(ii) *$\mathcal{S}_G$ is isomorphic with $\mathcal{O}(r)$, where r denotes the number of vanishing s_i.*

Proof. (i) That the U with $GU = G$ are in $\mathcal{M}_G$ is obvious. Conversely, let $U \in \mathcal{M}_G$ be given. With the notation of (2.3) this means that

$$\mathrm{tr}(GU) = \sum_i s_i \langle U\mathbf{e}_i, \mathbf{e}_i\rangle = \sum_i s_i,$$

and this is possible only if $U\mathbf{e}_i = \mathbf{e}_i$ for all i such that $s_i > 0$. (2.3) implies that $GU = G$.
(ii) By the preceding proof there are as many $U \in \mathcal{M}_G$ as there are orthogonal matrices on the span of $\{\mathbf{e}_i \mid s_i = 0\}$. □

The situation for the special orthogonal group is more complicated. If G is diagonal with entries $s_1, \ldots, s_d$, then the results parallel the preceding ones: $\mathcal{M}_G^1 = \mathcal{S}_G^1$, and this group is isomorphic with $\mathcal{SO}(r)$, where r is as above.
In the case $\det(G) < 0$ we have to argue more subtly:

Proposition 4.2 *Let G be a diagonal $d \times d$-matrix with entries $s_1, \ldots, s_{d-1}$, $-s_d$, where $s_1 \geq \ldots \geq s_d > 0$.*

(i) *$\mathcal{S}_G^1$ consists only of the identity.*
(ii) *$\mathcal{M}_G^1$ might contain $\mathcal{S}_G^1$ properly. In fact, in contrast to the preceding cases $\mathcal{M}_G^1$ need not be a subgroup of $\mathcal{SO}(d)$.*
(iii) *Suppose that $s_{d-1} > s_d$. Then $\mathcal{M}_G^1$ contains precisely the matrix Id.*
(iv) *If $s_{d-1} = s_d$ holds, define a number $w \leq d-1$ such that $s_{w-1} > s_w = \cdots = s_d$; if all s_i are identical, we put $w = 1$.*
Let a normalized vector $\mathbf{e}$ in the linear span V of $\mathbf{u}_w, \ldots, \mathbf{u}_{d-1}$ and $\phi \in \mathbb{R}$ be given. Denote by $U_{\mathbf{e},\phi}$ the rotation by the angle ϕ in the plane spanned by $\mathbf{e}$ and $\mathbf{u}_d$ (with $\mathbf{u}_1, \ldots, \mathbf{u}_d$= the canonical unit vectors of $\mathbb{R}^d$).
Then $\mathcal{M}_G^1$ is precisely the collection of all these $U_{\mathbf{e},\phi}$, and consequently $\mathcal{M}_G^1$ is a group iff $w = d-1$.

Proof. (i) is obvious. For the proof of (ii) consider the following matrices:

$$G := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \; U_1 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \; U_2 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

U_1 and U_2 lie in $\mathcal{M}_G^1$, but $U_1 U_2$ does not.
(iii) We adopt the notation and the results from the proof of theorem 2.3.

Let $\mathbf{e}_1, \ldots, \mathbf{e}_{2r}$ be an orthonormal family, $\phi_1, \ldots, \phi_r$ real numbers and $U_{\phi_1,\ldots,\phi_r}$ the associated element of $\mathcal{SO}(d)$. Then

$$\operatorname{tr}(GU_{\phi_1,\ldots,\phi_r}) = \sum_{\rho=1}^{r} [\langle G\mathbf{e}_{2\rho-1}, \mathbf{e}_{2\rho-1}\rangle + \langle G\mathbf{e}_{2\rho}, \mathbf{e}_{2\rho}\rangle] \cos\phi_\rho + c$$

holds.

As a consequence of lemma 2.2 we know that all

$$\langle G\mathbf{e}_{2\rho-1}, \mathbf{e}_{2\rho-1}\rangle + \langle G\mathbf{e}_{2\rho}, \mathbf{e}_{2\rho}\rangle$$

are bounded from below by $s_{d-1} - s_d$, and thus they are strictly positive. It follows that $U_{\phi_1,\ldots,\phi_r}$ can be in $\mathcal{M}_G^1$ only if $\phi_1 = \cdots = \phi_r = 0$, i.e., if $U_{\phi_1,\ldots,\phi_r} = Id$.
(iv) Let $\mathbf{e} \in V$ be given. Then $\langle G\mathbf{e}, \mathbf{e}\rangle = s_{d-1}$ so that

$$\langle G\mathbf{e}, \mathbf{e}\rangle + \langle G\mathbf{u}_d, \mathbf{u}_d\rangle = 0.$$

Therefore

$$\mathrm{tr}(GU_{\mathbf{e},\phi}) = \langle G\mathbf{e}, \mathbf{e}\rangle + \langle G\mathbf{u}_d, \mathbf{u}_d\rangle \cos\phi + c$$

is a constant which equals $s_1 + \cdots + s_{d-1} - s_d$, and hence $U_{\mathbf{e},\phi} \in \mathcal{M}_G^1$ as claimed.

For the reverse inclusion let $U_{\phi_1,\ldots,\phi_r}$ be as in the preceding proof of (iii). We suppose that $U_{\phi_1,\ldots,\phi_r}$ lies in $\mathcal{M}_G^1$, that it is not the identity, and that the ϕ_ρ are not in $2\pi\mathbb{Z}$.
Since $\mathrm{tr}(GU_{\phi_1,\ldots,\phi_r})$ assumes its maximal value at $\phi_1 = \cdots = \phi_r = 0$ it follows that all

$$\langle G\mathbf{e}_{2\rho-1}, \mathbf{e}_{2\rho-1}\rangle + \langle G\mathbf{e}_{2\rho}, \mathbf{e}_{2\rho}\rangle,\ \rho = 1, \ldots, r$$

vanish. From lemma 2.2 we conclude that

$$\begin{aligned} 0 \ &= \ \sum_\rho \langle G\mathbf{e}_{2\rho-1}, \mathbf{e}_{2\rho-1}\rangle + \langle G\mathbf{e}_{2\rho}, \mathbf{e}_{2\rho}\rangle \\ &\geq \ s_{d-2r+1} + s_{d-2r+2} + \cdots + s_{d-1} - s_d, \end{aligned}$$

and this implies that $r = 1$.

It remains to be shown that the span of $\mathbf{e}_1$, $\mathbf{e}_2$ coincides with the span of $\mathbf{e}$, $\mathbf{u}_d$ for a suitable $\mathbf{e} \in V$ with $w \leq i \leq d-1$: then $U_{\phi_1,\ldots,\phi_r} = U_{\mathbf{e},\phi'}$ with a properly chosen ϕ'.

First we show that $\mathbf{e}_1$ and $\mathbf{e}_2$ lie in the span of V and $\mathbf{u}_d$. Let G' be the diagonal $d \times d$-matrix with entries $s_{d-1}, \ldots, s_{d-1}, -s_d$. It follows from lemma 2.2 that

$$\begin{aligned} s_{d-1} - s_d \ &\leq \ \langle G'\mathbf{e}_1, \mathbf{e}_1\rangle + \langle G'\mathbf{e}_2, \mathbf{e}_2\rangle \\ &\leq \ \langle G\mathbf{e}_1, \mathbf{e}_1\rangle + \langle G\mathbf{e}_2, \mathbf{e}_2\rangle \\ &= \ s_{d-1} - s_d. \end{aligned}$$

Therefore

$$\langle G'\mathbf{e}_1, \mathbf{e}_1\rangle + \langle G'\mathbf{e}_2, \mathbf{e}_2\rangle = \langle G\mathbf{e}_1, \mathbf{e}_1\rangle + \langle G\mathbf{e}_2, \mathbf{e}_2\rangle,$$

and consequently neither $\mathbf{e}_1$ nor $\mathbf{e}_2$ can have a nontrivial component at positions $1, \ldots, w-1$.

We write $\mathbf{e}_j = \mathbf{f}_j + a_j\mathbf{u}_d$, with $\mathbf{f}_j \in V$ and $a_j \in \mathbb{R}$ $(j = 1, 2)$. Then, since $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthonormal, we have

$$\begin{aligned} 1 &= ||\mathbf{f}_1||^2 + a_1^2 \\ 1 &= ||\mathbf{f}_2||^2 + a_2^2. \end{aligned}$$

We also know that

$$\begin{aligned} 0 &= \langle G\mathbf{e}_1, \mathbf{e}_1\rangle + \langle G\mathbf{e}_2, \mathbf{e}_2\rangle \\ &= s_d[(||\mathbf{f}_1||^2 + ||\mathbf{f}_2||^2) - (a_1^2 + a_2^2)], \end{aligned}$$

and therefore

$$||\mathbf{f}_1||^2 + ||\mathbf{f}_2||^2 = a_1^2 + a_2^2$$

holds. It follows that $||\mathbf{f}_2||^2 = a_1^2$ and $||\mathbf{f}_1||^2 = a_2^2$.

But $\mathbf{e}_1$, $\mathbf{e}_2$ are orthogonal so that $a_1a_2 = -\langle\mathbf{f}_1, \mathbf{f}_2\rangle$. The preceding calculations yield

$$|\langle\mathbf{f}_1, \mathbf{f}_2\rangle|^2 = ||\mathbf{f}_1||^2||\mathbf{f}_2||^2,$$

i.e., $\mathbf{f}_1$ and $\mathbf{f}_2$ lie in the same one-dimensional subspace of V. This completes the proof. □

5. SUPPLEMENTS

STABILITY OF THE OPTIMAL POSITION

Let the $d\times d$-matrix G be in 1-optimal position. For orthonormal vectors $\mathbf{e}$ and $\mathbf{f}$ we denote by $U_{\mathbf{e},\mathbf{f},\phi}$ the rotation by the angle ϕ in the plane generated by $\mathbf{e}$, $\mathbf{f}$. Then, as we already know,

$$\operatorname{tr}(GU_{\mathbf{e},\mathbf{f},\phi}) = (\langle G\mathbf{e}, \mathbf{e}\rangle + \langle G\mathbf{f}, \mathbf{f}\rangle)\cos\phi + c.$$

The assumption implies that $\langle G\mathbf{e}, \mathbf{e}\rangle + \langle G\mathbf{f}, \mathbf{f}\rangle$ is nonnegative. The absolute size of this number measures how fast the trace will decrease with ϕ if G is replaced by $GU_{\mathbf{e},\mathbf{f},\phi}$. Therefore it is natural to introduce

$$S_G := \inf\ (\langle G\mathbf{e}, \mathbf{e}\rangle + \langle G\mathbf{f}, \mathbf{f}\rangle),$$

where the infimum is taken over all pairs of orthonormal vectors: S_G measures something like the *stability of the 1-optimal position.*

With the above results we can be explicit: by lemma 2.2, S_G equals $s_{d-1}+s_d$ or $s_{d-1}-s_d$ if $\det(G)$ is positive or negative, respectively.

REDUCTION OF THE PROBLEM

Suppose that, in the original problem, the $\mathbf{x}_1,\ldots,\mathbf{x}_n$ as well as the $\mathbf{y}_1,\ldots,\mathbf{y}_n$ lie in an r-dimensional subspace V of $\mathbb{R}^d$. Without loss of generality we may assume that V is the span of the first r canonical unit vectors, and this implies that the associated matrix $G=(g_{jk})$ has zero entries for $j>r$ or $k>r$. Consequently the $\mathbf{e}_j$, $\mathbf{f}_j$ from the *SVD* which are associated with nonzero s_j lie in V, and therefore one may find a $U\in\mathcal{O}(d)$ such that $\operatorname{tr}(GU)$ is maximal and U leaves the orthogonal complement of V invariant. To phrase it differently: a U with maximal $\sum_i\langle U\mathbf{x}_i,\mathbf{y}_i\rangle$ can already be found in $\mathcal{O}(r)$.

Similarly one can argue if the $\mathbf{x}_1,\ldots,\mathbf{x}_n$ resp. the $\mathbf{y}_1,\ldots,\mathbf{y}_n$ lie in r-dimensional subspaces V_1 resp. V_2, where not necessarily $V_1=V_2$. Transforming the $\mathbf{x}_i$ in a first step by an orthogonal U to V_2 and applying the preceding argument we conclude: there exists $U\in\mathcal{O}(d)$ such that $g(U)=\sum_i\langle U\mathbf{x}_i,\mathbf{y}_i\rangle$ is maximal and all $U\mathbf{x}_i$ lie in V_2.

We note that there might be U with maximal $g(U)$ where the $U\mathbf{x}_i$ are not necessarily in V_2. As a simple example consider a situation where the associated matrix G has only zero entries. Then *every* U gives rise to a maximal $g(U)$.

Moral of the story: if r is smaller than d, then one may reduce the complexity of the calculations considerably.

For the case of the *special* orthogonal group one has to argue a little more subtly. As a simple example consider

$$\mathbf{x}_1=\mathbf{y}_2=\mathbf{u}_1,\quad \mathbf{x}_2=\mathbf{y}_1=-\mathbf{u}_1$$

in $\mathbb{R}^2$ (with $\mathbf{u}_1=$ "the first unit vector"). Considered as elements of the one-dimensional $\mathbb{R}\mathbf{u}_1$ these vectors are in 1-optimal position, but in two dimensions there is the possibility of rotating the $\mathbf{x}_i$ which gives rise to a strictly larger $g(U)$. Thus one can only hope to reduce the problem to a $(r+1)$-dimensional one if the $\mathbf{x}_i$ and the $\mathbf{y}_i$ lie in r-dimensional subspaces V_1 and V_2, respectively. (Nevertheless, in this case it is true that there exists a $U_0\in\mathcal{SO}(d)$ such that $\sum_i\langle U_0\mathbf{x}_i,\mathbf{y}_i\rangle=\max_{U\in\mathcal{SO}(d)}\sum_i\langle U\mathbf{x}_i,\mathbf{y}_i\rangle$, and all $U_0\mathbf{x}_i$ lie in V_2. This is left to the reader.)

GENERALIZATIONS I: THE NONDISCRETE CASE

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $f, g : \Omega \to \mathbb{R}^d$ two measurable functions. Which U in $\mathcal{O}(d)$ resp. in $\mathcal{SO}(d)$ provides the maximal

$$\int_\Omega \langle Uf(\omega), g(\omega)\rangle d\mu(\omega)?$$

The problem we started from corresponds to the case of the counting measure on a finite set.

First, one needs to know that the integrals under consideration do exist. A natural assumption is that both f and g are in the Bochner space $L^2(\Omega, \mathbb{R}^d)$. (In this way we can, e.g., treat the case of sequences $\mathbf{x}_1, \mathbf{x}_2, \ldots$ and $\mathbf{y}_1, \mathbf{y}_2, \ldots$ provided that $\sum ||\mathbf{x}_i||^2$ and $\sum ||\mathbf{y}_i||^2$ are finite.) And then one can argue similarly to the above approach: the optimal U is precisely that matrix for which the trace of GU is maximal, where $G = (g_{jk})$ is defined by

$$g_{jk} := \int_\Omega \langle f(\omega), \mathbf{u}_j\rangle \langle g(\omega), \mathbf{u}_k\rangle.$$

GENERALIZATIONS II: THE INFINITE-DIMENSIONAL CASE

Let H be any separable real Hilbert space, we will replace the matrix G from section 2 by a linear continuous operator $T : H \to H$. We will assume that T is a *nuclear operator*. Then T and all TU with unitary U will have a well-defined trace[1], and we can ask how large the maximum of the numbers

$$\mathrm{tr}(TU)$$

will be.

It is possible to solve this problem with the same ideas as in the finite-dimensional case. First one has to know that nuclear operators admit a singular value decomposition: there are orthonormal bases $\mathbf{e}_1, \mathbf{e}_2, \ldots$ and $\mathbf{f}_1, \mathbf{f}_2, \ldots$ of H as well as numbers $s_1 \geq s_2 \geq \cdots \geq 0$ with $s_1 + s_2 + \cdots < \infty$ such that

$$T\mathbf{x} = \sum s_i \langle \mathbf{x}, \mathbf{e}_i\rangle \mathbf{f}_i$$

for every $\mathbf{x}$. It can easily be shown as in section 2 that the unitary operator $\widehat{U}$ which is defined by $\mathbf{f}_i \mapsto \mathbf{e}_i$ then gives rise to the value maximally possible for $\mathrm{tr}(TU)$ and that this maximum is $s_1 + s_2 + \cdots$, the *nuclear norm* of T.

[1] For terminology cf. [6].

What about the subtleties in connection with the special orthogonal group in this case? There exists no reasonable definition of the determinant for an arbitrary unitary operator, but one could restrict oneself to operators of the form "identity plus finite dimensional". For such operators a notion of "determinant" exists (see [6]), the unitary U of this form for which the determinant is one are our candidates to replace the elements of the special orthogonal group.

It turns out that in infinite dimensional spaces this variant does not give rise to interesting results. The maximum of the $\mathrm{tr}(TU)$ with unitary U is always the supremum of the $\mathrm{tr}(TU)$ with "special unitary" U (this "special" supremum, however, is not necessarily attained). This is an easy consequence of the fact that the s-numbers $s_1, s_2, \ldots$ tend to zero.

Similar remarks apply to questions in connection with "optimal position" and "symmetry".

A RANDOM PROCEDURE TO DETERMINE A 1-OPTIMAL POSITION

Let G be any $d \times d$-matrix, we know how to find a $\widetilde{U} \in \mathcal{M}_G^1$ with the help of the SVD of G. There is another possibility: one might use the following random procedure which is easy to implement and which provides an approximate solution surprisingly fast.

To describe the simple idea put $G_1 := G$ and choose orthonormal random vectors $\mathbf{e}, \mathbf{f}$. These can be generated, e.g., by producing two vectors $\mathbf{x}, \mathbf{y}$ the components of which are equidistributed in $[-1, +1]$; then put $\mathbf{e} := \mathbf{x}/||\mathbf{x}||$ and $\mathbf{f} := (\mathbf{y} - \langle \mathbf{e}, \mathbf{y} \rangle \mathbf{e})/||\mathbf{y} - \langle \mathbf{e}, \mathbf{y} \rangle \mathbf{e}||$. Denote by $U_{\mathbf{e},\mathbf{f},\phi} \in \mathcal{SO}(d)$ the rotation by the angle ϕ in the $\mathbf{e}$-$\mathbf{f}$-plane. Then, as can easily be shown,

$$\mathrm{tr}(GU_{\mathbf{e},\mathbf{f},\phi}) = (\langle G\mathbf{e}, \mathbf{e} \rangle + \langle G\mathbf{f}, \mathbf{f} \rangle) \cos \phi + (\langle G\mathbf{e}, \mathbf{f} \rangle - \langle G\mathbf{f}, \mathbf{e} \rangle) \sin \phi + c.$$

Choose – with the help of elementary calculus – a ϕ where this number is as large as possible and replace G by $G_2 := GU_{\mathbf{e},\mathbf{f},\phi}$. Repeat this procedure with G_2 instead of G and continue this way. Then, trivially, the $\mathrm{tr}(G_1), \mathrm{tr}(G_2), \ldots$ increase, and computer experiments indicate that these numbers converge remarkably fast to the largest possible $\mathrm{tr}(GU)$, $U \in \mathcal{SO}(d)$. (For example, if $d = 3$, one obtains 8 relevant digits of the maximal trace already after roughly fifty cycles.)

The procedure works since every $U \in \mathcal{SO}(d)$ is a product of at most $[d/2]$ rotations of the form $U_{\mathbf{e},\mathbf{f},\phi} \in \mathcal{SO}(d)$. An error analysis could be given, but since we have an exact method at our disposal we omit to describe the clumsy details.

References

[1] P. Deuflhard, A. Hohmann. *Numerical Analysis.* de Gruyter, Berlin, New York, 1995.

[2] G.H. Golub, Ch.V. van Loan. *Matrix Computations (second edition).* The John Hopkins University Press. Baltimore and London, 1989.

[3] E. Hewitt, K. Ross. *Abstract Harmonic Analysis II.* Springer-Verlag, Berlin-Heidelberg-New York, 1970.

[4] F. Mädler, E. Behrends, K. Knorr. *A geometric centroid principle and its applications.* submitted to Acta Cryst. Sec. A.

[5] K. Knorr, F. Mädler. *The application of evolution strategies to disordered structures.* J. Appl. Cryst. **32** (1999), 902–910.

[6] A. Pietsch. *Eigenvalues and s-Numbers.* Cambridge studies in advanced mathematics 13. Cambridge University Press, Cambridge 1987

[7] H. Sowa, K. Knorr, F. Mädler, H. Ahsbahs, G. Kutoglu. *The high-pressure behaviour of the cubic phase of KPF_6 – An X-ray single-crystal and entropy-maximization study.* Z. Kristallogr. **214** (1999), 542–546.

I. Mathematisches Institut, Freie Universität Berlin, Arnimallee 2–6,
D-14 195 Berlin, Germany;
e-mail: behrends@math.fu-berlin.de

Hahn-Meitner-Institut, Glienicker Straße 100,
D-14 109 Berlin, Germany;
e-mail: maedler@hmi.de