Technische Universität München

Zentrum Mathematik

Algebraic Methods for Convexity Notions in the Calculus of Variations

(Algebraische Methoden für Konvexitätsbegriffe in der Variationsrechnung)

Diplomarbeit

von Carl Friedrich Kreiner

Aufgabensteller:	Prof. Dr. Martin Brokate
Betreuer:	Dr. Johannes Zimmer
Letzter Abgabetermin:	15. Oktober 2003

Eigenständigkeitserklärung

Ich erkläre hiermit, dass ich die Diplomarbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

München, den 12. September 2003

1 Introduction

One of the central problems in the calculus of variations is the minimization of functionals of the form

$$I(u) = \int_\Omega W(x, u(x), \nabla u(x))\,\mathrm{d}x \qquad (\Omega \subset \mathbb{R}^d \text{ open, bounded})$$

among all functions $u\colon \mathbb{R}^n \to \mathbb{R}$ which lie in a suitable function space (usually a Sobolev space) and are subject to certain boundary conditions. Such problems arise in many different situations. Originally motivated by geometrical questions, important applications come now from mathematical modelling in the sciences.

Our work has been inspired by models for microstructures which develop during a solid to solid phase transition of the material. The structure of the minimizers u has contributed to understand the shape-memory effect that is observed in alloys like CuAlNi and NiTi. The solid crystalline structure is different at high temperatures (austenite) and low temperatures (martensite) and changes abruptly at a certain critical temperature. The austenite lattice has usually more symmetry than the martensite lattice. A comprehensive introduction to the theory of martensites can be found in the book [8].

In such models we investigate the behavior of the respective materials by minimizing the total stored energy, represented by the functional

$$E(u) = \int_\Omega W(\nabla u(x))\,\mathrm{d}x \tag{1.1}$$

where u is the elastic deformation of a body $\Omega \subset \mathbb{R}^n$, subject to a boundary condition, and W is the (nonnegative) energy density. The mathematical difficulty of this problem comes from the fact that the functional E is often not weakly lower semicontinuous on $W^{1,p}(\Omega;\mathbb{R}^m)$, with the consequence that there exists in general no minimizing Sobolev function u. The lack of weak lower semicontinuity is caused by the energy density W which fails to be convex—a sufficient condition—in these cases. In fact, W has for martensitic materials typically a multi-well structure, in particular a disconnected zero set. Therefore minimizing sequences may develop oscillations and converge to a weak limit that does not need to minimize E. The oscillations and the limit functions correspond to experimentally observed microstructures. We refer to [4, 8] for details.

The weak lower semicontinuity of (1.1) is the main assumption needed for the proof of existence of minimizers [14]. It is equivalent to the *quasiconvexity* of the energy density W, i.e., $E(u)$ is weakly lower semicontinuous on $W^{1,p}(\Omega;\mathbb{R}^m)$ if and only if $W\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ satisfies for every bounded domain $\Omega \subset \mathbb{R}^{m\times n}$,

$F \in \mathbb{R}^{m\times n}$ and for every test function $\varphi \in \mathscr{C}_0^\infty(\Omega, \mathbb{R}^m)$

$$W(F) \;\le\; \frac{1}{|\Omega|}\int_\Omega W(F+\nabla\varphi(x))\,\mathrm{d}x.$$

Obviously every convex function is quasiconvex. This notation of quasiconvexity was introduced by Morrey [29] but there is not yet any practically useful characterization. A rather negative result in this direction is that a local characterization cannot exist [23]. It is generally hard to verify for a given function directly whether it is quasiconvex.

To make the given functional (1.1) weakly lower semicontinuous, it is possible to *relax* the problem, that is, to consider the functional

$$\tilde{E}(u) = \int_\Omega W^{qc}(\nabla u)\,\mathrm{d}x \tag{1.2}$$

where W^{qc} is the *quasiconvex envelope* of W, defined by

$$W^{qc}(F) := \sup\{f(F) : f \text{ quasiconvex }, f \le W\}. \tag{1.3}$$

It is well known that the infima of E and $\tilde{E}$ coincide [14]. The transition from (1.1) to (1.2) corresponds to the homogenization of the microstructure and yields an accurate description of the macroscopic behavior without oscillations. The minimization problem is also closely related to the *quasiconvex hull* of $\mathcal{Z} := \{F \in \mathbb{R}^{m\times n} : W(F) = 0\}$. Let us prescribe a linear boundary condition

$$u(x) = F(x) \qquad (x \in \partial\Omega)$$

and consider (1.1) on all $u \in W^{1,p}$ that satisfy this boundary condition. Then the infimum of the functional (1.1) is zero if, and only if, F belongs to the quasiconvex hull of $\mathcal{Z}$ [39] which is defined as

$$\mathcal{Z}^{qc} := \left\{ G \in \mathbb{R}^{m\times n} : f(G) \le \sup_{Z\in\mathcal{Z}} f(Z) \text{ for all quasiconvex } f\colon \mathbb{R}^{m\times n} \to \mathbb{R} \right\}. \tag{1.4}$$

We also note that the level sets $L_{c,f} := \{X \in \mathbb{R}^{m\times n} : f(X) \le c\}$ of quasiconvex functions f satisfy $L_{c,f}^{qc} = L_{c,f}$.

Quasiconvexity itself being a rather inaccessible notion, several authors have introduced other notions. A function $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is called *polyconvex* if $f(X)$ can be expressed as a convex function of all minors (subdeterminants) of X. A function f is *rank-one convex* if $f(\lambda A + (1-\lambda)B) \le \lambda f(A) + (1-\lambda)f(B)$ whenever $\operatorname{rank}(A-B) = 1$. A less widespread notion is that f is called *separately convex* if $f(\lambda A+(1-\lambda)B) \le \lambda f(A)+(1-\lambda)f(B)$ whenever $(A-B)$ has only one nonzero entry.

With these definitions we have for a function $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ the following implications [14]

$$\left.\begin{array}{r} f \text{ convex} \implies f \text{ polyconvex} \implies f \text{ quasiconvex} \implies \\ \implies f \text{ rank-one convex} \implies f \text{ separately convex.} \end{array}\right\} \tag{1.5}$$

If $m = 1$ or $n = 1$ then all definitions coincide because, in this case, every rank-one convex function is convex. For $m, n \geq 2$ the converse of the first two implications is false [14]. The question whether quasiconvexity and rank-one convexity are equivalent, has long been an open problem. The answer is negative if we have $m \geq 3$ or $n \geq 3$ [38], positive for the space of diagonal 2×2-matrices [31], but still open for general 2×2-matrices.
The polyconvex, the rank-one convex and the separately convex hull (of a compact set K) are defined similarly to the quasiconvex hull—just substitute in (1.4) the word "quasiconvex" by polyconvex, rank-one convex—and denoted by K^{pc}, K^{rc}, and K^{sc}. If K^{co} is the usual convex hull we have the inclusions

$$K^{co} \supseteq K^{pc} \supseteq K^{qc} \supseteq K^{rc} \supseteq K^{sc}.$$

It is worth noting that in most cases where the quasiconvex hull is known explicitly, this was shown by proving that $K^{pc} = K^{rc}$ [4, 9].
The same can be carried out for the envelopes with the obvious modifications of (1.3).

In this thesis we consider some questions from this very active area of research. We develop new algorithms for the practical computation of separately convex and rank-one convex hulls of finite sets. They constitute approximations for the quasiconvex hull from below. Since the level sets of quasiconvex functions are quasiconvex this yields also algorithms for a discrete approximation from above of the quasiconvex envelope of a function f as well.
In Chapter 2 we recall a framework that encompasses both separate and rank-one convexity. We introduce relevant notation and mention some facts that reflect the complexity of the problem.
Chapter 3 discusses the case of separate convexity. We present a graph-theoretical algorithm, first for the case of a two-dimensional space, then for the general case. It is to our knowledge the first graph-theoretical algorithm in this area. Generalizations of our algorithm to infinite sets or other forms of directional convexity remain for now an area of further research.
Chapter 4 collects various facts and algorithms from algebraic geometry from the literature and illustrates them with examples. The content of this chapter is prerequisite for Chapter 5 where we use these algebraic methods for the detection of T_4-configurations. Such configurations have been of particular interest in the construction of solutions of certain differenial equations with

special properties [35, 30, 40] and as example like in our following Proposition 2.7. We pursue apparently a completely new approach and hopefully only the first step to a general algorithm for the computation of rank-one convex hulls that avoids critical discretizations wherever possible.
But already this rather special problem exhibits mathematically interesting features. We find more indications that the case of $\mathbb{R}^{2\times 2}$ (where the relation between quasiconvexity and rank-one convexity is unknown) shows a different behavior from higher dimensions. Even more importantly, by our implementation we give a tool for the exact computation of certain rank-one convex hulls, probably for the first time.

2 Directional Convexity

2.1 Definitions

Definition 2.1 *Let $\mathcal{D}$ be a set of vectors in $\mathbb{R}^d$. A function $f\colon \mathbb{R}^d \to \mathbb{R}$ is called a* $\mathcal{D}$-convex function *(or* $\mathcal{D}$-directionally convex*) if the one-variable functions $t \mapsto f(M + tD)$ are convex for all fixed $M \in \mathbb{R}^d$ and $D \in \mathcal{D}$.*
We will always assume that $\mathcal{D}$ spans $\mathbb{R}^d$.

This means that we speak of a $\mathcal{D}$-convex function if its restriction to each line parallel to a nonzero element of $\mathcal{D}$ is convex in the usual sense. The requirement that $\mathcal{D}$ spans $\mathbb{R}^d$ is not always found in the literature. However without this assumption $\mathcal{D}$-convex functions need not be continuous (see Lemma 2.4 and the following example). In all applications that we are interested in, $\mathcal{D}$ will satisfy this hypothesis.
Definition 2.1 contains several special cases.

- For $\mathcal{D} = \mathbb{R}^d$ we get the usual *convexity*. The same holds already if, e.g., $\mathcal{D}$ contains the $(d-1)$-dimensional sphere.
- For $\mathcal{D} = (\mathbb{R}^m \times \{0\}^n) \cup (\{0\}^m \times \mathbb{R}^n)$ with $m+n = d$ we get the definition of *bi-convexity* that was studied in [3].
- If $d = mn$ and if we identify $\mathbb{R}^d$ and $\mathbb{R}^{m\times n}$ then *rank-one convexity* is given by $\mathcal{D} := \{X \in \mathbb{R}^{m\times n} : \det(X) \le 1\}$.
- If $\mathcal{D}$ is an orthonormal basis of $\mathbb{R}^d$ we speak of *separate convexity*. Figure 2.1 refers to this example.

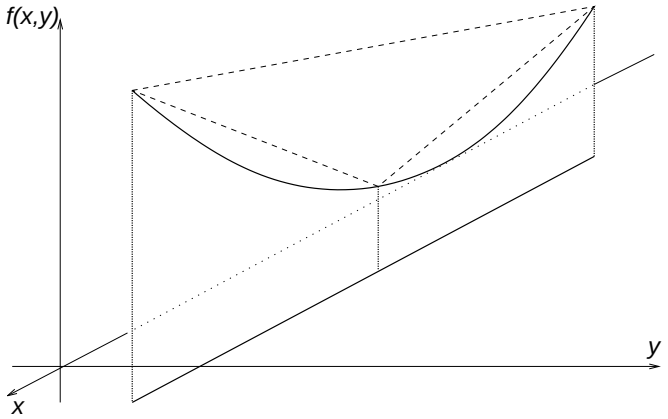


Figure 2.1: Section through the graph of a $\mathcal{D}$-convex function $f\colon \mathbb{R}^2 \to \mathbb{R}$ with $\mathcal{D} = \{\binom{1}{0}, \binom{0}{1}\}$

We say that $A, B \in \mathbb{R}^d$ are $\mathcal{D}$-*connected* if they lie on a $\mathcal{D}$-*line*, that means, if there is a $\gamma \in \mathbb{R}$ such that $\gamma(A - B) \in \mathcal{D} \setminus \{0\}$. We write $[A, B]$ for the line segment that connects the points A and B.

Now we generalize the notion of convex sets and convex hulls of sets. We start with the hulls. We will introduce two notions of $\mathcal{D}$-convex hulls. In the case of usual convexity both definitions coincide as we will see later (Prop. 2.6).

Definition 2.2 *Let $\mathcal{M} \subset \mathbb{R}^d$ be compact.*
(i) The functional $\mathcal{D}$-convex hull *of $\mathcal{M}$ is defined as*

$$\mathcal{M}^{\mathcal{D}c} := \left\{ X \in \mathbb{R}^d : f(X) \leq \sup_{M \in \mathcal{M}} f(M) \text{ for all } \mathcal{D}\text{-convex } f \colon \mathbb{R}^d \to \mathbb{R} \right\}.$$

(ii) The geometrical $\mathcal{D}$-convex hull *of $\mathcal{M}$ is defined as the smallest superset $\mathcal{N}$ of $\mathcal{M}$ with the following property: For any two points $A, B \in \mathcal{N}$ that are $\mathcal{D}$-connected we have $[A, B] \subset \mathcal{N}$.*
The geometrical $\mathcal{D}$-convex hull will be denoted by $\mathcal{M}_{\mathcal{D}c}$.

The definition of the geometrical hull is the natural generalization of the definition of the usual convex hull. However, this approach is for the applications like those mentioned in the introduction too narrow. Instead the functional $\mathcal{D}$-convex hull is needed which is in general larger (Lemma 2.6 and Prop. 2.7). For example, level sets of $\mathcal{D}$-convex functions are functionally $\mathcal{D}$-convex; this explains the interest in the functional $\mathcal{D}$-convex hull in the context of the $\mathcal{D}$-convexification of some function.
As there are two hulls, there are two notions of $\mathcal{D}$-convex sets as well.

Definition 2.3 *Let $\mathcal{M} \subset \mathbb{R}^d$ be compact.*
(i) A set $\mathcal{M} \subset \mathbb{R}^d$ is called a functionally $\mathcal{D}$-convex set *if $\mathcal{M} = \mathcal{M}^{\mathcal{D}c}$.*
(ii) A set $\mathcal{M} \subset \mathbb{R}^d$ is called a geometrically $\mathcal{D}$-convex set *if $\mathcal{M} = \mathcal{M}_{\mathcal{D}c}$.*

In Definition 2.1 we have considered only functions defined on the whole space $\mathbb{R}^d$; later we will only be interested in such functions. For completeness, we note that geometrically $\mathcal{D}$-convex sets are natural domains for $\mathcal{D}$-convex functions. Definition 2.1 makes sense if, and only if, for every fixed pair $(M, D) \in \mathbb{R}^d \times \mathcal{D}$ the domain of the function $t \mapsto f(M+tD)$ is an interval in $\mathbb{R}$. This is equivalent to the $\mathcal{D}$-convexity of the domain of f.

2.2 Basic properties

Lemma 2.4 *Let $f \colon \mathbb{R}^d \to \mathbb{R}$ be $\mathcal{D}$-convex. Then f is continuous, and even locally Lipschitz-continuous.*

Proof. See [26], Observation 2.3. □

This lemma uses our assumption from Definition 2.1 that $\mathcal{D}$ spans $\mathbb{R}^d$. We give an example to show that this assumption is essential.

Example. For this example only we drop the requirement that $\mathcal{D}$ spans $\mathbb{R}^d$. We consider $\mathcal{D} := \{\binom{1}{0}\} \subset \mathbb{R}^2$. Let $g\colon \mathbb{R} \to \mathbb{R}$ be an arbitrary discontinuous function and set

$$f(x, y) := g(y).$$

Then $f\colon \mathbb{R}^2 \to \mathbb{R}$ is constant on all lines of the form $M + t\binom{1}{0}$ hence $\mathcal{D}$-convex. But since g was not continuous f cannot be continuous either. ■

The next lemma characterizes the functional and the geometrical $\mathcal{D}$-convex hull. For the functional hull, it is sufficient to consider only a specific class of $\mathcal{D}$-convex functions. The geometrical hull may be computed by successive *lamination:* Take the union of all line segments that are parallel to a nonzero vector in $\mathcal{D}$ and that start and end in $\mathcal{M}$ and iterate this procedure.

Lemma 2.5 *Let $\mathcal{M} \subset \mathbb{R}^d$ be compact.*
(i) The functional $\mathcal{D}$-convex hull of $\mathcal{M}$ is the intersection of the zero sets of all nonnegative $\mathcal{D}$-convex functions that vanish on $\mathcal{M}$. In other words,

$$\mathcal{M}^{\mathcal{D}c} = \left\{X \in \mathbb{R}^d : f(X) = 0 \text{ for all } \mathcal{D}\text{-convex } f\colon \mathbb{R}^d \to [0, \infty) \text{ with } f|_{\mathcal{M}} \equiv 0\right\}.$$

(ii) Define $\mathcal{M}_0 := \mathcal{M}$ and inductively

$$\mathcal{M}_{j+1} := \Big\{[A, B] : A, B \in \mathcal{M}_j, A - \gamma B \in \mathcal{D} \text{ for some } \gamma \in \mathbb{R}\Big\}.$$

Then the geometrical $\mathcal{D}$-convex hull is $\mathcal{M}_{\mathcal{D}c} = \bigcup_{j=1}^{\infty} \mathcal{M}_j$.

Proof. (i) The inclusion "$\subseteq$" is clear. To prove the converse, let X be an element of the set on the right-hand side and assume there exists a $\mathcal{D}$-convex $g\colon \mathbb{R}^d \to \mathbb{R}$ with $g(X) > \sup_{M\in\mathcal{M}} g(M)$. Consider the function $f\colon \mathbb{R}^d \to \mathbb{R}$ defined by

$$f(Y) = \begin{cases} g(Y) - \sup\limits_{M\in\mathcal{M}} g(M) & \text{if } g(Y) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

This is the maximum of two $\mathcal{D}$-convex functions and therefore again $\mathcal{D}$-convex (this follows from the corresponding property of convex functions). We have $f \geq 0$, $f|_{\mathcal{M}} \equiv 0$ but $f(X) > 0$ in contradiction to the choice of X.
(ii) See [26], Observation 2.1. □

For general $\mathcal{D}$ however, we have only that the geometrical hull is contained in the functional hull. But as we will see in the next example (Proposition 2.7), this inclusion may be strict.

Lemma 2.6
(i) For all compact sets $\mathcal{M} \subset \mathbb{R}^d$ we have $\mathcal{M}_{\mathcal{D}c} \subseteq \mathcal{M}^{\mathcal{D}c}$.
(ii) Consider $\mathcal{D}_0 := \mathbb{R}^d$, i.e., the case of usual convexity. Then $\mathcal{M}_{\mathcal{D}_0 c} = \mathcal{M}^{\mathcal{D}_0 c}$.

Proof. (i) See [26], Observation 2.2.
(ii) The converse inclusion to (i) follows from the separation theorem. Since $\mathcal{M}_{\mathcal{D}_0 c}$ is convex, there exists, for every $P \notin \mathcal{M}_{\mathcal{D}_0 c}$ a linear functional f_P with $f_P(P) = 1$ and $f_P|_{\mathcal{M}} \equiv 0$. As linear functions are convex this shows $P \notin \mathcal{M}^{\mathcal{D}_0 c}$.
□

The following example for a four-point set with trivial geometrical but nontrivial functional $\mathcal{D}$-convex hull was discovered independently by several authors in different contexts [41, 3, 35], see also [26]. It has been presented, e.g., in the context of rank-one convexity on $\mathbb{R}^{2\times 2}_{\text{diag}}$ or separate convexity.

Proposition 2.7 (Tartar square) *Consider*

$$\mathcal{M} = \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\} \quad \textit{and} \quad \mathcal{D} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Then we have $\mathcal{M}_{\mathcal{D}c} = \mathcal{M}$ but $\mathcal{M}^{\mathcal{D}c}$ equals the set depicted in Figure 2.2, that is, the (usual) convex hull of $\{W, X, Y, Z\}$ (the square) and the line segments $[A, X], [B, Y], [C, Z], [D, W]$.

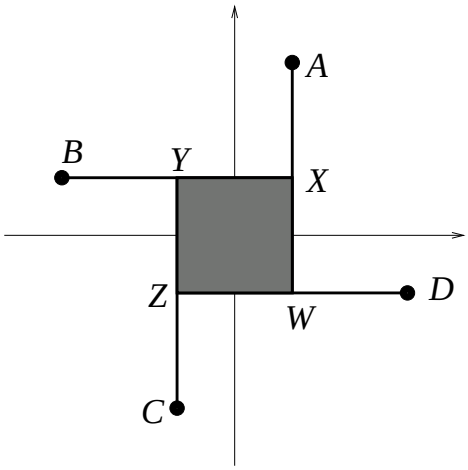


Figure 2.2: Tartar square

Proof. The first statement is clear because the elements of $\mathcal{M}$ are pairwise not $\mathcal{D}$-connected.
In order to see that $W, X, Y, Z \in \mathcal{M}^{\mathcal{D}c}$ (with the notation from Figure 2.2) we assume that there exists a $\mathcal{D}$-convex function $f\colon \mathbb{R}^2 \to [0, \infty)$ with $f|_{\mathcal{M}} \equiv 0$ and $f(X) > 0$. By convexity of f on the line $A + t\binom{0}{1}$ we get $f(W) > f(X)$ and, by convexity of f on the line $D + t\binom{1}{0}$, $f(Z) > f(W)$. Going on this way, we arrive at the contradiction

$$f(X) > f(Y) > f(Z) > f(W) > f(X)$$

hence $X \in \mathcal{M}^{\mathcal{D}c}$ and, similarly, $W, Y, Z \in \mathcal{M}^{\mathcal{D}c}$. With these four points the line segments $[A, W], [B, X], [C, Y], [D, Z]$ and then the interior of the square $WXYZ$ must belong to $\mathcal{M}^{\mathcal{D}c}$.
This is all of $\mathcal{M}^{\mathcal{D}c}$ because every other point in $\mathbb{R}^2$ can separated from $\mathcal{M}$ by a $\mathcal{D}$-convex function: Consider the function

$$f\left(\begin{smallmatrix} p_1 \\ p_2 \end{smallmatrix}\right) := \begin{cases} p_1 p_2 & \text{if } p_1 > 0 \text{ and } p_2 > 0 \\ 0 & \text{otherwise} \end{cases}$$

This function is $\mathcal{D}$-convex and nonzero exactly on the open first quadrant. The construction works similarly for every open quadrant, and for every translated open quadrant. Every point outside the square $WXYZ$ and the line segments $[A, W], [B, X], [C, Y], [D, Z]$ lies in some translated open quadrant that does not contain an element of $\mathcal{M}$. □

Informally speaking, the reason for the difference between functional and geometrical $\mathcal{D}$-convex hull in the previous example is that the latter cannot generate the polygon that is present in this example. The iterative method from Lemma 2.5 (ii) adds only interior points of a line segment to the kth iterate $\mathcal{M}_k$ if both endpoints already belong to $\mathcal{M}_{k-1}$.

We give one more characterization of the functional $\mathcal{D}$-convex hull. The concept of envelopes of functions is very important in the calculus of variations.

Definition 2.8 *Let $g\colon \mathbb{R}^d \to \mathbb{R}$ be a function.*
The $\mathcal{D}$-convex envelope $C_{\mathcal{D}}g\colon \mathbb{R}^d \to \mathbb{R}$ is defined by

$$C_{\mathcal{D}}g(X) := \sup\{f(X) : f\colon \mathbb{R}^d \to \mathbb{R} \ \mathcal{D}\text{-convex with } f(Y) \leq g(Y) \,\forall\, Y \in \mathbb{R}^d\},$$

i.e., the $\mathcal{D}$-convex envelope is the pointwise supremum of all $\mathcal{D}$-convex functions f satisfying $f(Y) \leq g(Y)$ for all $Y \in \mathbb{R}^d$.

We note that, as supremum of $\mathcal{D}$-convex functions, $C_{\mathcal{D}}g$ is $\mathcal{D}$-convex as well. This follows from corresponding properties of convex functions (in the usual sense) that are not lost due to the restriction of convexity to certain directions.

Lemma 2.9 *Let $\mathcal{M} \subset \mathbb{R}^d$ be compact and $g\colon \mathbb{R}^d \to \mathbb{R}$ denote the associated distance function, i.e.,*

$$g(X) := \min_{M \in \mathcal{M}} \|X - M\|.$$

Then then functional $\mathcal{D}$-convex hull $\mathcal{M}^{\mathcal{D}c}$ is the zero set of the $\mathcal{D}$-convex envelope $C_{\mathcal{D}}g$ of the distance function.

Proof. See [25], Theorem 3.1. □

In the area of computational convexity, extremal points are the crucial notion for theoretical and practical aspects. The most important result is the theorem of Krein-Milman which states that every convex set is the convex hull of its extremal points.
We now turn to corresponding generalizations for $\mathcal{D}$-convexity. Again, there are two approaches.

Definition 2.10 *Let $\mathcal{M} \subset \mathbb{R}^d$ be compact.*
(i) A point $E \in \mathcal{M}$ is called a geometrically $\mathcal{D}$-extremal point *of $\mathcal{M}$ if there exists no line segment $[A, B] \subset \mathcal{M}$ parallel to some nonzero vector in $\mathcal{D}$ that contains E als interior point.*
(ii) A point $E \in \mathcal{M}$ is called a $\mathcal{D}$-Choquet point *of $\mathcal{M}$ if the Dirac measure δ_E is the only probability measure on $\mathcal{M}$ that represents E, that is, if we have the following implication:*

$$\mu \in \mathrm{Prob}(\mathcal{M}) \quad \textit{and for all } \mathcal{D}\textit{-convex } f\colon \mathbb{R}^d \to \mathbb{R} : f(E) \leq \int_{\mathcal{M}} f(X)\ \mathrm{d}\mu(X)$$

then $\mu = \delta_E$.

The first notion seems more intuitive and is much easier to use for computational issues but, as we will see in 2.12 and 2.13, it is not appropriate for functionally $\mathcal{D}$-convex sets. This is not surprising because the functional $\mathcal{D}$-convexity is defined by duality and not by purely geometrical properties. In the case of usual convexity both notions coincide ([33], Prop.1.4).

Lemma 2.11 *Let E be a $\mathcal{D}$-Choquet point of a compact set $\mathcal{M} \subset \mathbb{R}^d$. Then E is a geometrically $\mathcal{D}$-extremal point as well.*

Proof. Suppose E is not geometrically $\mathcal{D}$-extreme. Then there exists a $\mathcal{D}$-line segment in $\mathcal{M}$ which contains E as an interior point, that is, there are $A, B, \in \mathcal{M}$ such that $A - \gamma B \in \mathcal{D} \setminus \{0\}$ for some $\gamma \in \mathbb{R}$ and $E = \lambda A + (1 - \lambda B)$ for some $\lambda \in (0, 1)$. Then we have with $\mu := \lambda\delta_A + (1-\lambda)\delta_B$ for every $\mathcal{D}$-convex function $f\colon \mathbb{R}^d \to \mathbb{R}$ by definition and by the fundamental property of the Dirac measure

$$\begin{aligned} f(E) &\leq \lambda f(A) + (1-\lambda) f(B) \\ &= \lambda \int_{\mathcal{M}} f(X)\ \mathrm{d}\delta_A(X) + (1-\lambda) \int_{\mathcal{M}} f(X)\ \mathrm{d}\delta_B(X) = \int_{\mathcal{M}} f(X)\ \mathrm{d}\mu(X) \end{aligned}$$

hence E is not a $\mathcal{D}$-Choquet point and this is a contradiction. □

The converse of Lemma 2.11 is not true; we give a counterexample.

Proposition 2.12 (Kirchheim star) *Consider* $\mathbb{R}^{2\times 2}_{\mathrm{sym}}$ *(which can be identified with* $\mathbb{R}^3$*) and* $\mathcal{D} = \{X \in \mathbb{R}^{2\times 2}_{\mathrm{diag}} : \det(X) = 0\}$ *(rank-one convexity). Then the functional* $\mathcal{D}$*-convex hull of*

$$\mathcal{M} := \left\{ \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$$

is $\mathcal{M}^{\mathcal{D}c} = \{t \cdot M : t \in [0,1], M \in \mathcal{M}\}$*. The zero matrix is a geometrically* $\mathcal{D}$*-extreme point but not a* $\mathcal{D}$*-Choquet point.*

Proof. This is a special case of [22], Corollary 4.19. □

Now we come to the generalization of the theorem of Krein-Milman for $\mathcal{D}$-convexity.

Theorem 2.13 (Kružík) *Let* $\mathcal{M} \subset \mathbb{R}^d$ *be compact and* $\mathcal{E}(\mathcal{M})$ *be the set of all* $\mathcal{D}$*-Choquet points of* $\mathcal{M}$*. Then*

$$\mathcal{M}^{\mathcal{D}c} = \mathcal{E}(\mathcal{M})^{\mathcal{D}c}.$$

In particular, every (compact) functionally $\mathcal{D}$*-convex set is the functional* $\mathcal{D}$*-convex hull of its* $\mathcal{D}$*-Choquet extremal points.*

Proof. See [24]. □

The previous theorem shows that the notion of $\mathcal{D}$-Choquet points is the appropriate one for functional $\mathcal{D}$-convexity. However, this is not very helpful for the computation of functional $\mathcal{D}$-convex hulls because—at present knowledge—it does not yield an implementable algorithm.

3 Separate convexity

We have already mentioned the notion of separate convexity as $\mathcal{D}$-convexity with $\mathcal{D}$ being the canonical basis of $\mathbb{R}^d$. Throughout the chapter, the canonical basis of $\mathbb{R}^d$ will be denoted by $\{e_1, e_2, \dots, e_d\}$.

Definition 3.1 *A function $f\colon \mathbb{R}^d \to \mathbb{R}$ is called* separately convex *if it is $\mathcal{D}$-convex for $\mathcal{D} = \{e_1, e_2, \dots, e_d\}$, i.e., if the functions $t \mapsto f(x + te_j)$ are convex for all fixed $x \in \mathbb{R}^n$ and $1 \leq j \leq d$.*
For a set $\mathcal{M} \subset \mathbb{R}^d$ we define the (functional) separately convex hull *to be the functional $\mathcal{D}$-convex hull for $\mathcal{D} = \{e_1, e_2, \dots, e_d\}$. It is denoted by $\mathcal{M}^{sc}$.*

Since we are considering functional separately convex hulls only in this chapter, we shall mostly drop the word "functional".
The term "separate convexity" is used in the literature for different special cases of $\mathcal{D}$-convexity, such as $\mathcal{D} = (\mathbb{R}^2 \times \{0\}) \cup (\{(0,0)\} \times \mathbb{R}) \subset \mathbb{R}^3$ in [30].
Separate convexity on $\mathbb{R}^d$ can be interpreted as restriction of rank-one convexity on $\mathbb{R}^{d\times d}$ to the subspace of diagonal matrices, denoted by $\mathbb{R}^{d\times d}_{\text{diag}}$. To see this, we identify the vector $(v_1, v_2, \dots, v_d)^T \in \mathbb{R}^d$ with the matrix

$$\begin{pmatrix} v_1 & & & 0 \\ & v_2 & & \\ & & \ddots & \\ 0 & & & v_d \end{pmatrix} \in \mathbb{R}^{d\times d}_{\text{diag}}.$$

Obviously a diagonal matrix has rank one if and only if it has only one nonzero entry, and these matrices correspond to multiples of the canonical basis vectors $e_j \in \mathbb{R}^d$.
We now come to some facts that distinguish separate convexity from general $\mathcal{D}$-convexity, in particular from rank-one convexity.

3.1 Separately convex hulls of finite sets

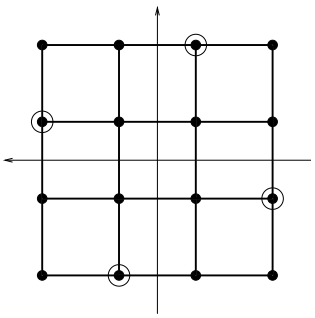
The main focus of interest of this thesis is the computation of special $\mathcal{D}$-convex hulls of finite sets. In the case of separate convexity we have the significant advantage that we may restrict ourselves to a suitably defined grid. The main result is Theorem 3.4 which appeared first in [26]. It was used to construct an algorithm relying on geometrically extremal points. We will take a different approach but start from the same result.

Definition 3.2 *Let $\mathcal{M} \subset \mathbb{R}^d$ be finite and denote for a vector $v \in \mathbb{R}^d$ its jth component by $x_j(v)$. Then we call the set*

$$\operatorname{grid}(\mathcal{M}) := \{x_1(v) : v \in \mathcal{M}\} \times \{x_2(v) : v \in \mathcal{M}\} \times \cdots \times \{x_d(v) : v \in \mathcal{M}\}.$$

the grid associated to $\mathcal{M}$.
For a point $v \in \mathrm{grid}(\mathcal{M})$ we denote by v^{j+} (resp. v^{j-}) the point in $\mathrm{grid}(\mathcal{M})$ *whose coordinates, except the jth one, coincide with those of v, and whose jth coordinate is the successor (resp. predecessor) of $x_j(v)$ in $\{x_j(w) : w \in \mathcal{M}\}$ if there exists a successor (resp. predecessor).*

Example. For $\mathcal{M} = \left\{\binom{1}{3}, \binom{-3}{1}, \binom{-1}{-3}, \binom{3}{-1}\right\}$ (cf. Proposition 2.7), we have $\mathrm{grid}(M) = \{-3,-1,1,3\} \times \{-3,-1,1,3\}$.
For the vector $v = \binom{1}{1}$ we have $v^{1-} = \binom{-1}{1}$ and $v^{j+} = \binom{3}{1}$. For the vector $w = \binom{1}{3}$, a "boundary point" of the grid, w^{2+} does not exist. ■



The following definition generalizes separate convexity to discrete functions defined on such grids. We will require convexity on the grid lines which are parallel to the coordinate axes. Instead of a discrete condition we may think of the function f being linearly interpolated on these interesting lines; then the discrete condition translates to usual convexity for a one-dimensional function.

Definition 3.3 *Let $\mathcal{M} \subset \mathbb{R}^d$ be finite. A function $f\colon \mathrm{grid}(\mathcal{M}) \to \mathbb{R}$ is called* grid-separately convex *if f satisfies the following convexity condition on every tripel (v_{j-}, v, v^{j+}) in* $\mathrm{grid}(\mathcal{M})$

$$f(v) \le f(v^{j-})\frac{x_j(v) - x_j(v^{j-})}{x_{j+}(v) - x_j(v^{j-})} + f(v^{j+})\frac{x_j(v^{j+}) - x_j(v)}{x_{j+}(v) - x_j(v^{j-})}.$$

(ii) The (functional) grid-separately convex hull *of a set $S \subset \mathrm{grid}(\mathcal{M})$ is*

$$\left\{x \in \mathrm{grid}(\mathcal{M}) : f(x) \le \max_{s\in S} f(s) \ \forall \text{ grid-separately convex } f\colon \mathrm{grid}(\mathcal{M}) \to \mathbb{R}\right\}.$$

We note that the characterization from Lemma 2.5 (i) holds also for the grid-separately convex hull of a set S: It is the zero set of all nonnegative grid-separately convex functions that vanish on S.
Now we come to the main theorem of this section.

Theorem 3.4 (Extension of grid-separately convex functions)
Let $\mathcal{M} \subset \mathbb{R}^d$ be finite. A function $f\colon \mathrm{grid}(\mathcal{M}) \to \mathbb{R}$ can be extended to a global separately convex function $\tilde{f}\colon \mathbb{R}^d \to \mathbb{R}$ if (and only if) it is grid-separately convex.

Proof. The proof for the 'if' part is based on multilinear interpolation; for details see [26], Thm. 3.1. The 'only-if' part is clear because every tripel (v^{j-}, v, v^{j+}) lies on a line parallel to a coordinate axis. □

Definition 3.5 *Let $\mathcal{M} \subset \mathbb{R}^d$ be finite and $S \subset \text{grid}(\mathcal{M})$. The* box complex of *$S$, denoted $\mathbf{B}(S)$, is the union of all sets of the form $I_1 \times I_2 \times \cdots \times I_d$ where the I_j have either the form $\{a_j\}$ for some $a_j \in \{x_j(v) : v \in \mathcal{M}\}$ or the form $[x_j(v), x_j(v^{j+})]$ for some $v \in \text{grid}(\mathcal{M})$.*

Indeed, the box complex equals the separately convex hull of S.

Proposition 3.6 *Let $S \subset \text{grid}(\mathcal{M})$ for some finite set $\mathcal{M} \subset \mathbb{R}^d$.*
Then $S^{sc} = \mathbf{B}(S)$.

Proof. The statement follows by induction on the dimension of the elementary boxes $I_1 \times I_2 \times \cdots \times I_d$ from Lemma 2.5 (ii). □

Corollary 3.7 *Let $\mathcal{M} \subset \mathbb{R}^d$ be finite and S its grid-separately convex hull.*
Then $\mathcal{M}^{sc} = \mathbf{B}(S)$.

Proof. See [26], Cor. 3.2. □

We can now formulate the algorithm of Matoušek and Plecháč ([26], Section 3.2).

Algorithm 3.8
Input: $\mathcal{M} \subset \mathbb{R}^d$ finite.
Procedure:

1. Compute $\text{grid}(\mathcal{M})$.
2. Initialize $S = \text{grid}(\mathcal{M})$.
3. Find a point $v \in S \setminus \mathcal{M}$ such that there exists no line segment $[g, h]$ with $g, h \in S$ which contains v as interior point. If no such point exists go to step 5.
4. Set $S := S \setminus \{v\}$ (i.e., remove v from S) and repeat Step 3.
5. Compute the box complex of S.

Output: $\mathcal{M}^{sc} = \mathbf{B}(S)$.

This algorithm has running time $O(n^d)$. This is optimal with respect to complexity among all known algorithms for this task. For more details see [26] and for applications of this algorithm see [25].

3.2 Graph-theoretical algorithm for $\mathbb{R}^2$

The next two sections will present a graph-theoretical algorithm for the computation of the separately convex hull of a finite set. The main observation is that the grid from Section 3.1 can be interpreted as a graph with orientation induced by the convexity requirement on grid-separately convex functions. We

will treat first the case of $\mathbb{R}^2$ and postpone the slightly more difficult general case to the next section.
We recall the definition of an oriented graph.

Definition 3.9 *An* oriented graph G *is a pair* (N, E) *where* N *is a finite set of* nodes *and* E *is a subset of* $N \times N$*. An element* (n_1, n_2) *of* E *will be called an* (oriented) edge *from* n_1 *to* n_2*, and we write* $n_1 \to n_2$*.*
A cycle *in an oriented graph* G *is a finite sequence of edges of the form*

$$n_1 \to n_2 \to n_3 \to \cdots \to n_{k-1} \to n_k = n_1.$$

Now let $\mathcal{M} \subset \mathbb{R}^2$ be a finite set. We are going to associate an oriented graph to grid($\mathcal{M}$). To do this we observe that on each grid line there is a point of $\mathcal{M}$. We want to identify those grid points which belong to the grid-separately convex hull of $\mathcal{M}$. With the remark after Definition 3.3 these are the points x with $f(x) = 0$ for all grid-separately convex $f \geq 0$ with $f|_{\mathcal{M}} \equiv 0$.
The set N of nodes of our graph is the set of grid points. The edges connect neighboring grid points g and g^{j+} (or g^{j-}). They shall be oriented such that they point away from the grid points that belong to $\mathcal{M}$. More precisely, we define for $j \in \{1, 2\}$ and for all $g \in \text{grid}(\mathcal{M})$ such that g^{j+} exists

$$\begin{aligned} g \to g^{j+} &:\Longleftrightarrow \exists\, v \in \mathcal{M} : x_j(v) \leq x_j(g) \text{ and } x_k(v) = x_k(g)\ \forall\ j \neq k, \\ g^{j+} \to g &:\Longleftrightarrow \exists\, v \in \mathcal{M} : x_j(v) > x_j(g) \text{ and } x_k(v) = x_k(g)\ \forall\ j \neq k. \end{aligned} \tag{3.1}$$

By the definition of the grid, at least one of these two conditions on the right hand side must be satisfied. In the case that both conditions are satisfied at the same time, we allow two edges connecting g and g^{j+}. We will speak of a double edge and write $g \leftrightarrows g^{j+}$.
An example for a graph (without double edge) is shown in Figure 3.1.

Algorithm 3.10
Input: $\mathcal{M} \subset \mathbb{R}^2$ finite.
Procedure:

1. Compute grid($\mathcal{M}$) and set up the graph G as described above. Initialize $S := \mathcal{M}$.
2. Add all nodes g, g^{j+} to S that are connected by a double edge.
3. Search for a cycle in G. If no cycle is found, go to Step 6.
4. If a cycle is found then add all nodes on the cycle to S. Add all nodes on paths to S that lead into the cycle.
5. Repeat from Step 3 (but only for cycles that are not yet detected).
6. Compute the box complex of S.

Output: $\mathcal{M}^{sc} = \mathbf{B}(S)$.

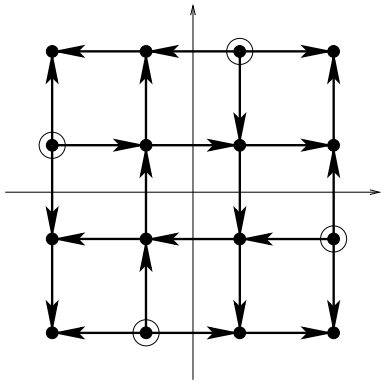


Figure 3.1: Graph associated to $\mathcal{M} = \left\{\binom{1}{3}, \binom{-3}{1}, \binom{-1}{-3}, \binom{3}{-1}\right\}$ (elements of $\mathcal{M}$ circled)

We demonstrate Algorithm 3.10 with two examples and then prove its correctness. The examples are simple but show already the main ideas of the proof.

Example. We consider again $\mathcal{M} = \left\{\binom{1}{3}, \binom{-3}{1}, \binom{-1}{-3}, \binom{3}{-1}\right\}$ (cf. Prop. 2.7 and the example after Def. 3.2). The associated graph is shown in Figure 3.1. Since no two points of $\mathcal{M}$ are connected by a line parallel to a coordinate axis (this was the reason why this example was introduced in the beginning) there is no double edge.
There is a single cycle, around the square in the center. After adding the involved nodes, S has eight elements (Figure 3.2). This is the grid-separately convex hull and its box complex is the Tartar square from Figure 2.2. ■

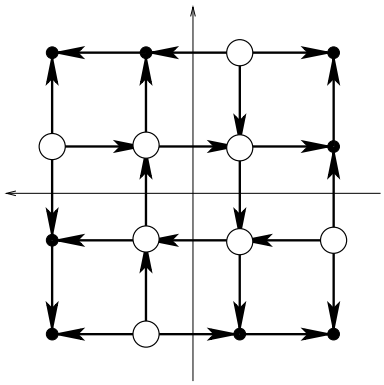


Figure 3.2: Grid-separately convex hull of $\mathcal{M} = \left\{\binom{1}{3}, \binom{-3}{1}, \binom{-1}{-3}, \binom{3}{-1}\right\}$ in the associated graph

Example. We consider $\mathcal{M} := \left\{\binom{1}{3}, \binom{-1}{1}, \binom{-3}{-3}, \binom{3}{-1}\right\}$. Figure 3.3 (a) shows

the associated grid. It contains neither double edges nor cycles. The algorithm yields $S = \mathcal{M}$. Since the box complex of S coincides with S we obtain that $\mathcal{M}^{sc} = \mathcal{M}$.
To verify this we need according to Theorem 3.4 for every grid point g a nonnegative grid-separately convex function f_g with $f|_{\mathcal{M}} \equiv 0$ and $f_g(g) > 0$. Figure 3.3 (b) shows an example of such a function for the grid point $g = \binom{1}{-1}$. The nodes where f_g vanishes are in the shaded region. ■

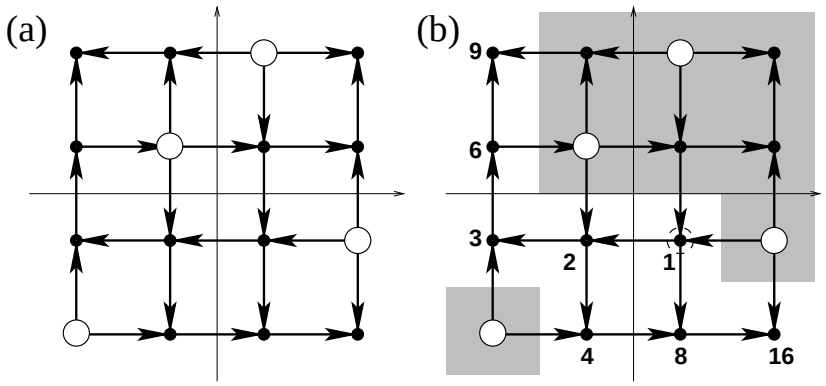


Figure 3.3: $\mathcal{M} := \left\{\binom{1}{3}, \binom{-1}{1}, \binom{-3}{-3}, \binom{3}{-1}\right\}$. (a) Associated graph. (b) Values of a grid-separately convex function that separates $\mathcal{M}$ and the marked point.

Now we show the correctness of Algorithm 3.10.

Theorem 3.11 *Algorithm 3.10 computes the separately convex hull for arbitrary finite sets $\mathcal{M} \subset \mathbb{R}^2$.*

Proof. According to Corollary 3.7 it suffices to prove that S is the grid-separately convex hull of $\mathcal{M}$. Denote the latter by $\mathcal{M}^H$.
The inclusion $S \subseteq \mathcal{M}^H$ follows from the construction of S and G.
Two nodes g, g^{j+} with $g \to g^{j+}$ and $g^{j+} \to g$ must lie on a line parallel to a coordinate axes between two elements of $\mathcal{M}$. The convexity condition for grid-separately convex functions implies $g, g^{j+} \in \mathcal{M}^H$.
If the graph contains a cycle $g_0, g_1, \ldots, g_m$ then we have for every separately convex function $f \geq 0$ with $f|_A \equiv 0$

$$f(g_0) \leq f(g_1) \leq f(g_2) \leq \ldots \leq f(g_{m-1}) \leq f(g_m) \leq f(g_0),$$

so we have equality everywhere in this chain. Suppose that $f(g_0) = C > 0$ and that g_0, g_1 and $a \in \mathcal{M}$ lie on a line with $x_j(a) \leq x_j(g_0) \leq x_j(g_1)$. Then we have

$$C = f(g_0) \;\leq\; f(a)\frac{x_j(g_0) - x_j(a)}{x_j(g_1) - x_j(a)} + f(g_1)\frac{x_j(g_1) - x_j(g_0)}{x_j(g_1) - x_j(a)} \;<\; 0 + C,$$

which is a contradiction. Hence $f(g_0) = 0$ and the cycle belongs to $\mathcal{M}^H$.
If there is a path $[a, g_1, \dots g_m, h]$ with $a \in \mathcal{M}, h \in \mathcal{M}^H$ then we have

$$f(a) \le f(g_1) \le f(g_2) \le \ldots \le f(g_{m-1}) \le f(g_m) \le f(h),$$

hence $f(g_k) = 0$ for $k = 1 \dots m$ and $g_k \in \mathcal{M}^H$.
There are no other possibilites how a node may be added to the output of the algorithm so we have shown $\mathcal{M}^H \supseteq S$.
To prove the converse inclusion $\mathcal{M}^H \subseteq S$, let $g_0 \in \mathrm{grid}(\mathcal{M}) \setminus S$. We construct a nonnegative grid-separately convex function f on $\mathrm{grid}(\mathcal{M})$ with $f(g_0) > 0$. This will show $g_0 \notin \mathcal{M}^H$.
Let $P_1 := [g_0, g_{1,1}, g_{1,2}, \dots, g_{1,m_1}], P_2, \dots P_r$ denote all paths which start at g_0 and end somewhere at the boundary of the grid and let Z denote the set of all grid points which do not belong to any path P_k. Note that all P_k are finite; otherwise some path would contain a loop which would imply that $g_0 \in S$.
We define now inductively a sequence of functions with grid-separately convex limit. We start with $f^0 \colon \mathrm{grid}(\mathcal{M}) \to \mathbb{R}$ by $f^0(g_0) := 1$ and $f^0(g) := 0$ for all $g \in \mathrm{grid}(\mathcal{M}) \setminus \{g_0\}$.
Assume now that f^{j-1} is already constructed and let J_j be the set of all nodes which are the $(j+1)$th in some path, i.e., $J_j = \{g_{h,j} \in P_h : h = 1 \dots n\}$. For every $g_{h,j} \in J_j$ choose a $c(g_{h,j})$ such that

$$\frac{c(g_{h,j}) - f^{j-1}(g_{h,j-k})}{\|g_{h,j} - g_{h,j-k}\|} > \frac{f^{j-1}(g_{h,j-l}) - f^{j-1}(g_{h,j-m})}{\|g_{h,j-l} - g_{h,j-m}\|} \quad \forall \quad 1 \le k, l, m \le j-1$$
$$\text{and} \quad c(g_{h,j}) \frac{\|g_{h,j} - g_{h,j-k}\|}{\|g_{h,j} - g\|} > f^{j-1}(g_{h,j-k}) \qquad \forall \quad 1 \le k \le j-1,\ g \in M.$$

These are finitely many conditions of the form $c(g_{h,j}) >$ const., hence this choice is always possible. Note that, if $g_{h_1,j} = g_{h_2,j}$ for some $h_1 \neq h_2$, then $c(g_{h_1,j})$ may be different from $c(g_{h_2,j})$. Now set for $g \in \mathrm{grid}(\mathcal{M})$

$$f^j(g) := \begin{cases} \max\{f^{j-1}(g), c(g_{h_1,j}), \dots, c(g_{h_s,j})\} & \text{if } g = g_{h_1,j} = \ldots = g_{h_s,j}, \\ f^{j-1}(g) & \text{otherwise.} \end{cases}$$

Thus $\mathrm{supp}(f^j) = \bigcup_{l=1}^{j} J_l$ and f^j is by construction grid-separately convex on $\mathrm{supp}(f^j) \cup Z$.
Since $J_j = \emptyset$ for $j > t := \max\{m_h : h = 1, \dots r\}$, this construction terminates. Then $f := f^t$ is grid-separately convex and $f(g_0) > 0$, hence we may conclude $g_0 \notin \mathcal{M}^H$ and the proof is finished. □

3.3 Graph-theoretical algorithm for $\mathbb{R}^d$

Algorithm 3.10 does not generalize immediately to higher dimensions.

One impediment in dimension $d \geq 3$ is that, if the grid is defined as in Section 3.1, not every grid line contains a point of $\mathcal{M}$. This occurs for example in the following example.

Example. Consider the set

$$\mathcal{M} := \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

as depicted in Figure 3.4. According to Definition 3.2 we have $\text{grid}(\mathcal{M}) = \{0,1\}^3$, i.e., $\text{grid}(\mathcal{M})$ equals the vertices of the unit cube. We attempt to set up the graph as in the previous section: The grid points become the nodes, and the edges get oriented following (3.1).
This is impossible for the three dashed edges. None of the two alternatives in (3.1) holds on them because they do not contain a point from $\mathcal{M}$. ■

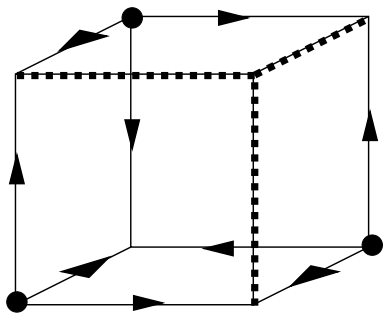


Figure 3.4: Unoriented edges in a three point configuration.

This example shows that there may exist pairs (g, g^{i+}) such that none of the two alternatives in (3.1) holds. In order to transfer the results from the last section we need a new framework.

Definition 3.12 *A* partially oriented graph *G is a tripel (N, E, F) where N is a finite set of* nodes *and E, F are disjoint subsets of $N \times N$.*
An element (n_1, n_2) of E will be called an oriented edge *from n_1 to n_2, and we write $n_1 \to n_2$. An element (n_1, n_2) of F will be called an* unoriented edge *from n_1 to n_2, and we write $n_1 - n_2$.*

In graph theory this mix of an oriented and an unoriented graph is normally avoided (see, e.g. [10, 20]). Unoriented edges $(-)$ are replaced by oriented double edges $(\leftrightarrows)$.
We cannot do this in our situation: Our oriented edges $g \to g^{j+}$ correspond to the monotone increment of certain separately convex functions on the line

segment $[g, g^{j+}]$. An oriented double edge $g \leftrightarrows g^{j+}$ means that all separately convex functions under consideration are constant on $[g, g^{j+}]$.
We have to face the situation that it is not enough to search for cycles in the graph which is partially oriented according to (3.1).

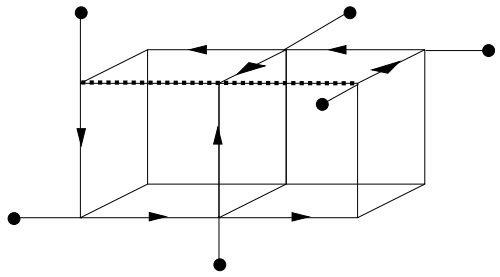


Figure 3.5: Partial grid to a six point configuration.

Example. Consider the six point configuration in Figure 3.5. The dashed line cannot be oriented without closing a cycle: If oriented leftwards the front face of the left cube is bounded by a cycle; if oriented rightwards this happens to the top face of the right cube. The only orientation for the two indicated edges that would not close a cycle would require the left edge pointing to the right and vice versa. But since oriented edges correspond to monotone growth of every nonnegative separately-convex function vanishing on the given six-point set this orientation would indicate a maximum of every such function, in contradiction to its convexity on the dashed line. ■

Another feature that results ultimately from unoriented edges as well appears in the following example.

Example. Consider the five-point set

$$\mathcal{M} := \left\{ \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

The first four points form the familiar Tartar square, the fifth point lies above its center (see Figure 3.6).
It is easy to see that $\mathcal{M}^{sc}$ consists of the Tartar square (embedded in $\mathbb{R}^3$) and the line segment on the x_3-axis that connects the fifth point and the origin. This example shows that the algorithm will need some sort of iteration because the line segment would not be detected in any algorithm involving only the initial graph. But note that the origin is a grid point from the very beginning on. ■

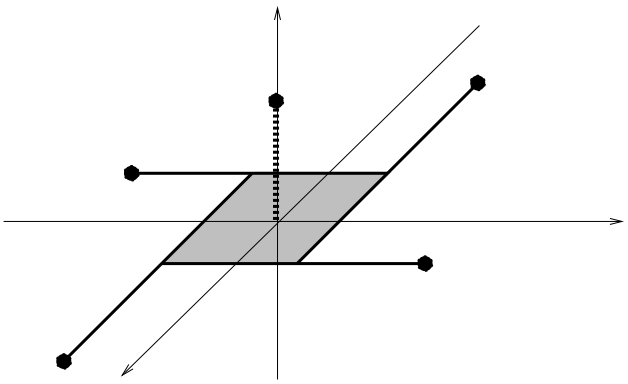


Figure 3.6: Example for the need of an iterative algorithm

Therefore the strategy in the case of general $\mathbb{R}^d$ is slightly different from the case of $\mathbb{R}^2$. For every grid point g there, we tried to find a path from g into some cycle in order to see that $g \in \mathcal{M}^{sc}$. Now we will rather try to find an orientation of the grid such that all paths starting at g end at the border of the grid and not in a cycle; this would show $g \notin \mathcal{M}^{sc}$.
As for the iterative component of the algorithm, note that, as soon as a grid point $s \in \text{grid}(\mathcal{M})$ is identified to lie in the grid-separately convex hull of $\mathcal{M}$ we may assign orientations to some more edges:

$$\begin{array}{lll} g \to g^{j+} & :\Longleftrightarrow & x_j(s) \leq x_j(g) \text{ and } x_k(s) = x_k(g) \ \forall\ j \neq k \\ g^{j+} \to g & :\Longleftrightarrow & x_j(s) > x_j(g) \text{ and } x_k(s) = x_k(g) \ \forall\ j \neq k \end{array} \tag{3.2}$$

This is possible because we know $f(s) = 0$ for all nonnegative grid-separately convex functions with $f|_{\mathcal{M}} \equiv 0$ and because we can therefore treat s as if it were already an element of $\mathcal{M}$. This permits an update of the graph. It does not require the computation of a new grid because s was already a grid point.

With the considerations before and the reasoning in the proof of Theorem 3.11 we arrive at the following algorithms.

Algorithm 3.13
Input: $\mathcal{M} \subset \mathbb{R}^d$ finite ($d \geq 3$).
Procedure:

1. Compute grid($\mathcal{M}$) and set up the partially oriented graph following the rules (3.1).
2. Initialize $S := \mathcal{M}$ (to become the grid-separately convex hull), $R := \emptyset$ (grid points not in the grid-separately convex hull) and $U := \text{grid}(\mathcal{M}) \backslash \mathcal{M}$ (undecided nodes).
3. Choose $g \in U$ and conduct Algorithm 3.14 for S, R, U, g and an auxiliary set $H = \{g\}$.

4. Perform the grid update according to (3.2) for all $s \in S$.
5. If $U \neq \emptyset$ then repeat from 3.
6. Compute the box complex of S.

Output: $\mathcal{M}^{sc} = \mathbf{B}(S)$.

Algorithm 3.14 [1]

Input: $\text{grid}(\mathcal{M}) \subset \mathbb{R}^d$ finite, $S, R, U \subset \text{grid}(M)$ with $\text{grid}(M) = S \cup R \cup U$, a grid point g, an auxiliary set H.

Procedure:

1. If $g \in S$ then return "found".
2. If there is an edge $g \to h$ then we found a cycle. Return "found".
3. For all edges $g \to h$ with $h \notin R$ conduct Algorithm this 3.14 for S, R, U, h and $H := H \cup \{h\}$.
4. For all edges $g - g^{j+}$ such that g^{j+}, g^{j-} exist and are not contained in R conduct this Algorithm 3.14 for S, R, U, g^{j+} and $H := H \cup \{g^{j+}\}$ and for S, R, U, g^{j-} and $H := H \cup \{g^{j-}\}$.
5. If any of the calls in Step 3 returns "found" then $S := S \cup \{h\}$, $U := U \setminus \{h\}$ and return "found".
6. If all calls in Step 3 return "fail" then $R := R \cup \{h\}$, $U := U \setminus \{h\}$ and return "fail".
7. If any pair of calls in Step 4 returns "found" for both g^{j+} and g^{j-} then $S := S \cup \{g^{j+}, g^{j-}\}$, $U := U \setminus \{g^{j+}, g^{j-}\}$ and return "found".
8. If all calls in Step 3 return "fail" then $R := R \cup \{h\}$, $U := U \setminus \{h\}$.
9. If no pair of calls in Step 4 returns "found" for both g^{j+} and g^{j-} then $R := R \cup \{g^{j+}, g^{j-}\}$, $U := U \setminus \{g^{j+}, g^{j-}\}$.
10. Return "fail".

Example. We resume the six point configuration from Figure 3.5. It is given by

$$\mathcal{M} := \left\{ \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

The grid is given by $\{-2,-1,0,1,2\} \times \{-1,0,1,2\} \times \{-1,0,1,2\}$, and we have $S = \mathcal{M}$ and $U = \text{grid}(\mathcal{M}) \setminus \mathcal{M}$.

We choose $g = (0,0,0)^T \in U$ and enter Algorithm 3.14. There are two oriented edges leaving from g, namely in the e_1 and e_3 directions. So there are in Step 3 two recursive calls of Algorithm 3.14 with $h_1 = (0,0,1)^T$ and $h_2 = (1,0,0)^T$.

[1] We follow the practice of many programming languages to keep the same identifier if a variable is updated, e.g. $S := S \cup \{g\}$ means that we add the element g to the old set S and call the new set again S.

There are two calls in Step 4 as well because there is a pair of unoriented edges as required, leading to $(0, \pm 1, 0)^T$.
The call of Algorithm 3.14 with h_1 is the most interesting and we do not consider the other calls here any further. There are again two oriented edges leaving from h_1, in the $-e_2$ and the e_3 direction. Both lead to the border of the grid and Algorithm 3.14 will eventually yield "fail". There is a pair of undirected edges in the $\pm e_1$-direction. As we see in Figure 3.5, both lead into a cycle. We may conclude that the grid-separately convex hull of $\mathcal{M}$ contains

$$\mathcal{M} \cup \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

After testing the remaining undecided points we would see that this is indeed the full grid-separately convex hull. $\mathcal{M}^{sc}$ is constructed as its box-complex of the above set. ■

4 Tools from Algebraic Geometry

This chapter collects definitions and basic facts from algebraic geometry that we will use in the following chapter on rank-one convexity. For a comprehensive exposition of related results we refer to standard textbooks like [12, 13]. Readers familiar with the relevant notation can skip the chapter.

4.1 Ideals and Varieties

In the next chapter we will consider certain systems of polynomial equations. The framework for the structure of the solution sets of such systems is provided by the theory of real varieties. We assume the reader to be familiar with the notion of an ideal that is introduced in many textbooks on linear algebra like [7, 17].

Definition 4.1 *Consider the real polynomial ring $R := \mathbb{R}[x_1, x_2, \dots, x_d]$ in d indeterminates and a list $L := \{p_1, \dots, p_s\} \subset R$ of polynomials. The (affine)* real variety $\mathscr{V}_{\mathbb{R}}(L)$ *associated to L is the set of real solutions of the system of polynomial equations*

$$p_1(x_1, x_2, \dots, x_d) = 0, \quad p_2(x_1, x_2, \dots, x_d) = 0, \quad \dots, \quad p_s(x_1, x_2, \dots, x_d) = 0.$$

Instead of L we might as well consider the ideal generated by the polynomials in L.

Lemma 4.2 *Let $L \subset \mathbb{R}[x_1, x_2, \dots, x_d]$ be a list of polynomials, I the ideal generated by the polynomials in L, in symbols $I = \langle p_1, \dots p_s \rangle \subseteq \mathbb{R}[x_1, x_2, \dots, x_d]$. Denote $\mathscr{V}_{\mathbb{R}}(I)$ the* variety associated to *I, that is,*

$$\mathscr{V}_{\mathbb{R}}(I) := \{x = (x_1, x_2, \dots x_d) \in \mathbb{R}^d : p(x) = 0 \text{ for all } p \in I\}.$$

Then $\mathscr{V}_{\mathbb{R}}(I) = \mathscr{V}_{\mathbb{R}}(L)$.

Proof. See [12], Chapter 2, Sect. 5, Prop. 9. □

There are extremely close connections between the theory of ideals and the theory of varieties. A famous result is Hilbert's Nullstellensatz ([12], Chapter 4, Sect. 1, Thm. 1 and 2) but since it has no direct applications to this thesis there is no need to quote this very general theorem here. But we will see that useful results are often formulated in terms of rings although we are actually interested in the corresponding varieties.

Lemma 4.3 *Let $\mathscr{V}$ and $\mathscr{W}$ be varieties. Then $\mathcal{V} \cap \mathcal{W}$ and $\mathcal{V} \cup \mathcal{W}$ are varieties as well.*

Proof. See [12], Chapter 1, Sect. 2, Lemma 2. □

We proceed now towards the definition of Gröbner bases, a very important notion in the theory of ideals. First, we need the monomial orderings.

Definition 4.4
(i) A monomial ordering *on* $\mathbb{R}[x_1, x_2, \ldots, x_d]$ *is any relation* $>$ *on the set of monomials* $M := \{x^\alpha : x = (x_1, x_2, \ldots, x_d), \alpha \in \mathbb{N}_0^d\}$ *(where* $x^\alpha = \prod x_j^{\alpha_j}$*) such that*

- $>$ *is a total ordering on* M*.*
- *If* $x^\alpha > x^\beta$ *and* $\gamma \in \mathbb{N}_0^d$ *then* $x^{\alpha+\gamma} > x^{\beta+\gamma}$*.*
- $>$ *is a well ordering on* M*.*

(ii) Let $>$ *be a monomial order and* $p = \sum_{\alpha \in A} \lambda_\alpha x^\alpha \in \mathbb{R}[x_1, x_2, \ldots, x_d]$ *with* $\lambda_\alpha \in \mathbb{R} \setminus \{0\}$*. The* leading monomial $\mathrm{LT}(p)$ *of* p *is the maximum of the finite set* $\{x^\alpha : \alpha \in A\}$ *with respect to the order* $>$*.*

Example. Consider the real polynomial ring $R := \mathbb{R}[x, y]$ in two variables. The set of monomials is $M = \{x^r y^s : r, s \in \mathbb{N}_0\}$. The lexicographic order is defined by

$$x^{r_1} y^{s_1} > x^{r_2} y^{s_2} \qquad :\Longleftrightarrow \qquad r_1 > r_2 \quad \text{or} \quad r_1 = r_2, s_1 > s_2,$$

and this is a monomial order ([12], Chapter 2, Sect. 2). We have, for example, $x^2 > y^3$, $x^4 y > x y^4$ and $x^3 y^3 > x^3 y^2$. The leading monomial of the polynomial $x^2 - y^3 + x^4 y - x y^4 + x^3 y^3 - x^3 y^2$ is $x^4 y$. With respect to the lexicographic order, the absolute degree $(r + s)$ of the monomials is not taken into account. ■

Now we are ready for Gröbner bases.

Definition 4.5 *Let* $>$ *be a monomial order and* $I \subset \mathbb{R}[x_1, x_2, \ldots, x_d]$ *an ideal other than* $\{0\}$*. A* Gröbner basis with respect to $>$ *is a finite subset* $G = \{g_1, g_2, \ldots g_t\} \subset I$ *such that*

$$\langle \mathrm{LT}(g_1), \mathrm{LT}(g_2), \ldots, \mathrm{LT}(g_t) \rangle = \langle \{\mathrm{LT}(p) : p \in I\} \rangle.$$

There is a Gröbner basis for any monomial order and any nontrivial ideal (see Lemma 4.6 below). This allows a broad variety of applications of the theory of Gröbner bases and, for computational issues, the choice of an advantageous monomial order.

Example. We continue the previous example $R = \mathbb{R}[x, y]$ with the lexicographic order and consider the ideal $I = \langle xy^2 + 1, x^2 - 1 \rangle$. We use the software tool *Macaulay 2* [18] for the computation of a Gröbner basis of I. We set up the ring, the monomial order and the ideal with the following commands:

```
i1 : R = QQ[x,y,MonomialOrder=>Lex];

i2 : I = ideal ( x*y^2+1, x^2-1);
o2 : Ideal of R
```

The Gröbner basis is computed by

```
i3 : gb I
o3 = | y4-1 x+y2 |
o3 : GroebnerBasis
```

and this should be read as $G = \{y^4 - 1, x + y^2\}$. Indeed, these two polynomials are contained in I. We have, e.g,

```
i4 : (x*y^2-1) * ( x*y^2+1 ) - y^4 * ( x^2-1 )
      4
o4 = y  - 1
```

We abstain from the technical verification that G satisfies Definition 4.5. ■

Lemma 4.6 *Let $>$ be a monomial order and $I \subset \mathbb{R}[x_1, x_2, \ldots, x_d]$ an ideal other than $\{0\}$.*
(i) Let G be a Gröber basis for I with respect to $>$. Then I is generated by G. (This justifies the use of the word "basis".)
(ii) There exists a Gröbner basis G for I with respect to $>$.

Proof. (i) See [12], Chapter 2, Sect. 5, Thm. 4.
(ii) See [12], Chapter 2, Sect. 5, Cor. 6. □

For a given monomial order the Gröbner basis is usually not unique. In fact, every finite superset of a Gröbner basis is again a Gröbner basis. Therefore the term of a *reduced Gröbner basis* has been introduced. The reduced Gröbner basis contains a minimal number of polynomials, and it is unique.

In the ring of univariate polynomials $\mathbb{R}[x]$ there is a well-known division algorithm to decompose a polynomial p with respect to a given polynomial f to a sum $p = qf + r$ where r is of degree less than f. This leads to an easy answer for the membership problem, i.e., to decide whether some polynomial p belongs to a given ideal I or not. In the univariate ring $\mathbb{R}[x]$, every ideal I can be generated by a single polynomial f (a ring with this property is called a principal ideal ring) [7, 17]. The remainder after the division of p by f is zero if and only if $p \in I$.
With a given monomial order on $\mathbb{R}[x_1, x_2, \ldots x_d]$ the division algorithm can be generalized to multivariate polynomials. We do not describe in detail how

a polynomial is divided by an ordered finite list of polynomials. However, not every list B of polynomials has the property that the division of a polynomial p by the B yields the remainder zero if p lies in the ideal generated by B. This makes it impossible to describe an ideal by an arbitrarily chosen generating system.
The remedy are Gröbner bases:

Theorem 4.7 *Let $p \in \mathbb{R}[x_1, \ldots, x_d]$, $I \subset \mathbb{R}[x_1, \ldots, x_d]$ an ideal and G be a Gröbner basis for I with respect to an arbitrary monomial order $>$ on $\mathbb{R}[x_1, \ldots, x_d]$. Then there exists a unique $r_> \in \mathbb{R}[x_1, \ldots, x_d]$ with the following properties:*

- *There is a $g \in I$ with $p = g + r_>$.*
- *No term of $r_>$ is divisible by a leading term of G, i.e., by an element of $\{\mathrm{LT}(g_j) : g_j \in G\}$.*

The remainder is zero if and only if $p \in I$.

Proof. See [12], Chapter 2, Sect. 6, Prop. 1. □

Since the particular choice of the polynomial order does not matter for the properties of the resulting Gröbner basis we will often speak of "a Gröbner basis" meaning the Gröbner basis corresponding to some monomial order.
The previous theorem has far-reaching consequences for various computational aspects. Several properties of an ideal or its associated variety can be simply read off the Gröbner basis with respect to any fixed monomial order.
As an example we mention a criterion how to decide whether a variety $\mathcal{V}_{\mathbb{R}}(I)$ consists of at most finitely many points.

Theorem 4.8 *Let $I \subsetneq \mathbb{R}[x_1, x_2, \ldots, x_d]$ be an ideal and $\mathcal{V}_{\mathbb{R}}(I)$ its associated real variety. Assume that one of the following conditions is satisfied:*

- *The $\mathbb{R}$-vector space $\mathbb{R}[x_1, x_2, \ldots, x_d]/I$ is finite-dimensional.*
- *Let G be a Gröbner basis for I. Then for each $1 \leq j \leq n$ there exists some $m_k \in \mathbb{N}_0$ such that $x_j^{m_j}$ is the leading term for some $g \in G$.*

Then $\mathcal{V}_{\mathbb{R}}(I)$ consists of at most finitely many points.
We say that I and $\mathcal{V}_{\mathbb{R}}(I)$ are zero-dimensional.

Proof. See [12], Chapter 5, Sect. 3, Thm. 6. □

Example. We continue the last example with $I = \langle xy^2 + 1, x^2 - 1\rangle \subset \mathbb{R}[x, y]$. To verify the first condition in Theorem 4.8 we compute a basis for the vector space $\mathbb{R}[x, y]/I$

```
i5 : basis (R/I)
o5 = | 1 y y2 y3 |
```

and we see that this vector space is four-dimensional, in particular finite-dimensional. As for the second condition, we have already found that $\{y^4 - 1, x + y^2\}$ is a Gröbner basis of I with respect to the lexicographic order. The leading terms of these two polynomials are y^4 and x, respectively, hence the second condition is satisfied. We conclude that I is zero-dimensional. ■

The drawback in the use of Gröbner bases is their high computational complexity [27]. There have been great efforts to create efficient implementations. In the applications that do not require a specific monomial order a very important aspect is the choice of a suitable order. The *graded reverse lexicographic order* (grevlex-order, see [12], Chapter 2, Sect. 2) is particularly favorable [5]. In fact, it is the default monomial order in the software package*Macaulay 2* [18] and recommended in the package *Singular* [19]. Both packages are intended for research in algebraic geometry and intricate computations related to it. The efficient implementation has been particularly important for these packages. We use *Macaulay 2* for our computations in this thesis.

4.2 Eliminant method and Sturm sequences

In Section 5.3 we will be faced with the following situation. We have a zero-dimensional ideal I (cf. Theorem 4.8) and we are interested in points of the associated real variety $\mathscr{V}_{\mathbb{R}}(I)$ which satisfy certain polynomial inequalities. This section introduces a method to compute the real numbers which may appear as coordinates of the points in $\mathscr{V}_{\mathbb{R}}(I)$.

Let I be zero-dimensional. Consider the quotient ring $Q := \mathbb{R}[x_1, \ldots, x_d]/I = \{p + I : p \in \mathbb{R}[x_1, \ldots, x_d]\}$ with its induced addition and multiplication. By Theorem 4.8, Q is a finite dimensional vector space over $\mathbb{R}$.
Every $f \in Q$ acts on Q by multiplication, i.e., we consider the mapping $m_f : q \mapsto f \cdot q$. It is well-defined because multiplication in the quotient ring Q is well-defined, and linear due to the distributive law

$$m_f(q + r) = f \cdot (q + r) = f \cdot q + f \cdot r = m_f(q) + m_f(q).$$

Hence m_f is an endomorphism of the finite-dimensional vector space. The minimal monomial of m_f is called the *eliminant corresponding to f*. By the theorem of Cayley-Hamiltion, the eliminant divides the characteristical polynomial of m_f and the latter will give us all information we will be interested in. In the algorithm we will choose a basis of Q, compute a matrix representation T_f of m_f and then the characteristic polynomial of T_f.

Theorem 4.9 *Let $I \subset \mathbb{R}[x_1, \ldots, x_d]$ be a zero-dimensional ideal and $f \in \mathbb{R}[x_1, \ldots, x_d]$ a polynomial (e.g., a projection $f(x_1, \ldots, x_d) = x_j$, $1 \leq j \leq d$). Let $\lambda \in \{f(x) : x \in \mathscr{V}_{\mathbb{R}}(I)\}$ be a value of f on $\mathscr{V}_{\mathbb{R}}(I)$. Then:*

- *λ is a root of the eliminant corresponding to f, i.e., of the minimal polynomial of the endomorphism m_f.*
- *λ is an eigenvalue of m_f.*

Proof. See [13], Chapter 2, Theorem 4.5. □

There is a stronger version of Theorem 4.9 if we take into account the *complex variety* $\mathscr{V}_{\mathbb{C}}(I) := \{x \in \mathbb{C}^d : p(x) = 0 \,\forall\, p \in I\}$. The theorem of Stickelberger [36] states that there is a one-to-one correspondence between eigenvectors v_x of m_f and points $x \in \mathscr{V}_{\mathbb{C}}(I)$, and the value of $f(x)$ equals the eigenvalue corresponding to v_x. This shows in particular that the dimension of the algebra Q is greater or equal the number of points in $\mathscr{V}_{\mathbb{C}}(I)$.
When we compare this result with Theorem 4.9 it becomes clear why the following *eliminant method* (Alg. 4.10) will give us only necessary but not sufficient criteria. It will be an algorithm to check whether, given a zero-dimensional ideal I and a single polynomial inequality $p(x) > 0$, it is *possible* that there exists a point $x \in \mathscr{V}_{\mathbb{R}}(I)$ that satisfies this inequality. For a sufficient criterion and for several polynomial inequalities we will discuss the computationally expensive BKR algorithm in the next section.

Algorithm 4.10 (Eliminant method)
Input: $p \in R := \mathbb{R}[x_1, \ldots, x_d]$, a zero-dimensional $I \subset R$.
Procedure:

1. Compute a basis B of $Q := R/I$ as $\mathbb{R}$-algebra (this involves computing a Gröbner basis for I).
2. Compute the matrix T_p that represents m_p with respect to the basis B.
3. Compute the characteristic polynomial χ_p of T_p. (The eliminant corresponding to p has the same roots as this polynomial.)
4. If χ_p has *no* positive roots then there is *no* $x \in \mathscr{V}_{\mathbb{R}}(I)$ with $p(x) > 0$.

The computation of eliminants is implemented in the extension `realroots.m2` to *Macaulay 2* by Sottile and Grayson [36]. We demonstrate the algorithm with our recurring example.

Example. We consider again $R = \mathbb{R}[x, y]$ and $I = \langle xy^2+1, x^2-1\rangle$. The associated complex variety consists of the four points $(1, i), (1, -i), (-1, 1), (-1, -1)$. From now on we will use the more efficient grevlex-order. A Gröbner basis of I with respect to this monomial order is $\{y^2 + x, x^2 - 1\}$ and a basis of the $\mathbb{R}$-algebra $Q := R/I$ is given by $\{1, x, xy, y\}$. We consider the coordinate functions $p(x, y) = x$ and $q(x, y) = y$. The multiplication endomorphisms m_p and

m_q are with respect to the given basis of Q represented by the matrices

$$T_p = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad T_q = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

For example, the third column of T_p means that $(x+I)(xy+I) = y+I$. Indeed, since $y(x^2-1) \in I$ we have $x^2y + I = y + I$.
The characteristic polynomials of these matrices are

$$\chi_p(t) = t^4 - 2t^2 + 1 = (t^2-1)^2 \quad \text{and} \quad \chi_q(t) = t^4 - 1 = (t^2-1)(t^2+1).$$

We conclude that the points in $\mathscr{V}_{\mathbb{R}}(I)$ must have both x- and y-coordinates in $\{\pm 1\}$ because these are the only real roots to χ_p and χ_q. Observe however that there is *no* point in $\mathscr{V}_{\mathbb{R}}(I)$ with x-coordinate $+1$ as the real variety consists of the two points $\{(-1,1),(-1,-1)\}$ only!
The above computations were done with the following sequence of *Macaulay 2* commands.

```
i1 : R = QQ[x,y];

i2 : I = ideal (x*y^2+1,x^2-1);
o2 : Ideal of R

i3 : gb I
o3 = | y2+x x2-1 |
o3 : GroebnerBasis

i4 : Q = R/I
o4 = Q
o4 : QuotientRing

i5 : basis Q
o5 = | 1 x xy y |
             1       4
o5 : Matrix Q  <--- Q

i6 : p=x;q=y;

i8 : load "realroots.m2"
--loaded realroots.m2
```

```
i9 : regularRep p
o9 = | 0 1 0 0 |
     | 1 0 0 0 |
     | 0 0 0 1 |
     | 0 0 1 0 |
               4        4
o9 : Matrix QQ  <--- QQ

i10 : regularRep q
o10 = | 0 0 -1 0  |
      | 0 0 0  -1 |
      | 0 1 0  0  |
      | 1 0 0  0  |
               4        4
o10 : Matrix QQ  <--- QQ

i11 : charPoly(p,Z)
       4     2
o11 = Z  - 2Z  + 1
o11 : QQ [Z]

i12 : charPoly(q,Z)
       4
o12 = Z  - 1
o12 : QQ [Z]
```

We note for completeness, that the actual eliminant corresponding to p is t^2-1 because T_p^2 is the identity matrix and this is therefore the minimal polynomial of T_p. ■

In this simple case we saw the roots of the characteristic polynomials immediately. In general however, the question whether a one-variable polynomial has positive roots (Step 4) is best carried out with Sturm sequences. This tool to locate real roots of a polynomial equation $f(t) = 0$ dates back to the early 19^{th} century.

Algorithm 4.11 (Sturm sequence)
Input: $f \in \mathbb{R}[t]$, $a, b \in \mathbb{R}$ with $a < b$ and $f(a), f(b) \neq 0$.
Procedure:

1. Set $f_0 := 0$, $f_1 := f'$ (the derivative) and for $k \geq 1$ recursively f_{k+1} to be the negative of the remainder of the division of f_{k-1} by f_k, i.e., if $f_{k-1} = q_k f_k + r$ with $q_k, r_k \in \mathbb{R}[t]$ and $\deg(r) < \deg(f_k)$ then $f_{k+1} := -r_k$.

If f has a multiple root this sequence terminates at zero, otherwise with a nonzero constant.

2. Count the sign changes in the sequence $A := (f_0(a), f_1(a), \ldots, f_s(a))$ and $B := (f_0(b), f_1(b), \ldots, f_s(b))$. (In counting sign changes, zeros are ignored.)

Output: The difference of sign changes in the sequences A and B is the number of real solutions of $f(x) = 0$ in the interval (a, b).

In recent textbooks this algorithm is normally demonstrated only for polynomials f with no multiple roots [13, 36] and for this situation implemented in the *Macaulay 2* extension package `realroots.m2` [36]. A careful analysis of the proof shows that this assumption is not necessary, see [11].

Example. Consider the polynomial $f(t) := t^4 - 2t^3 - 5t^2 + 7t - 2$, plotted in Figure 4.1. We are interested in the number of roots of f in the interval $(0, 1)$. The Sturm sequence reads

$$\begin{aligned} f_0(t) &= t^4 - 2t^3 - 5t^2 + 7t - 2, \\ f_1(t) &= 4t^3 - 6t^2 - 10t + 7, \\ f_2(t) &= \frac{13}{4}t^2 - 4t + \frac{9}{8}, \\ f_3(t) &= \frac{2148}{169}t - \frac{1246}{169}, \end{aligned}$$

and the polynomial $f_4(t)$ is a positive constant $c > 0$. For $t = 0$ we get the sequence $(-2, 7, \frac{9}{8}, -\frac{1246}{169}, c)$. The signs in this sequence are $(-, +, +, -, +)$ hence there are three sign changes. The sign sequence at $t = 1$ is $(-1, -1, 1, 1, 1)$ and contains one sign change. We conclude that there f has two real roots in the interval $(0, 1)$. ■

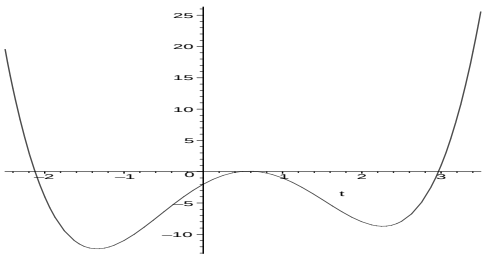


Figure 4.1: Plot of $f(t) = t^4 - 2t^3 - 5t^2 + 7t - 2$

4.3 The BKR algorithm

Let $R = \mathbb{R}[x_1, \ldots, x_d]$ be the ring of real polynomials in d indeterminates and $I \subset R$ a zero-dimensional ideal. Then, according to Theorem 4.8, the associated real variety $\mathcal{V}_{\mathbb{R}}(I) \subset \mathbb{R}^d$ consists of at most finitely many points. We take a finite list of polynomials $\{p_1, \ldots, p_s\} \subset R$ and consider for sign conditions $\diamondsuit_j \in \{<, =, >\}$ constraint regions of the form

$$G(\diamondsuit_1, \diamondsuit_2, \ldots, \diamondsuit_d) := \{x = (x_1, \ldots x_d) \in \mathbb{R}^d : p_1(x) \diamondsuit_1 0,\ p_2(x) \diamondsuit_2 0, \ldots,\ p_s(x) \diamondsuit_s 0\},$$

i.e., for possible combination of s sign conditions the region in $\mathbb{R}^d$ where each polynomial on the list satisfies the given sign condition.
The BKR algorithm addresses the following problem: Count, simultaneously for all sign conditions $\diamondsuit_j$, the number $c(p_1, \diamondsuit_1; p_2, \diamondsuit_2; \ldots; p_s, \diamondsuit_s)$ of points in the intersections $\mathcal{V}(I) \cap G(\diamondsuit_1, \ldots, \diamondsuit_s)$. It is described in full generality in [32] for the present multivariate case. The original paper [6] by Ben-Or, Kozen and Reif (hence the abbreviation BKR) and the detailed description [34] refer to the case $d = 1$. The multivariate case can be reduced to this univariate case using the Shape Lemma [37]; this method is however much less efficient. The case of a polynomial list with a single element ($s = 1$) can be treated with other methods, see [12, 36].
We now proceed to the definition of the important *trace form*. In the previous section we have introduced the multiplication mapping m_f on the finite-dimensional algebra $Q := R/I$, defined by $q \mapsto f \cdot q$. As endomorphism of a finite-dimensional vector space m_f has a trace $\operatorname{tr} m_f$ which can be computed as the trace of the matrix T_f that represents m_f with respect to some basis of Q.
In this way we may define, for any $g \in Q$, a bilinear form $B_g \colon Q \times Q \to \mathbb{R}$ by

$$B_g(q, r) := \operatorname{tr}(m_{g \cdot q \cdot r}).$$

This is obviously well-defined and symmetric, and bilinearity follows from

$$B_g(q, r_1 + r_2) = \operatorname{tr}(m_{gq(r_1+r_2)}) = \operatorname{tr}(m_{gqr_1} + m_{gqr_2}) = B_g(q, r_1) + B_g(q, r_2).$$

With respect to some basis of Q, B_g can be represented by a symmetric matrix S_g. For $g \in R$, we mean by S_g the matrix associated to the equivalence class of g in $Q = R/I$. The signature of the matrix S_g, i.e., the difference between the numbers of positive and negative eigenvalues, will be denoted by $\sigma(S_g)$
The key result is now the following theorem.

Theorem 4.12 *With notation and definitions from above, we have for all* $g \in R$

$$\sigma(S_g) = |\{x \in \mathcal{V}_{\mathbb{R}}(I) : g(x) > 0\}| - |\{x \in \mathcal{V}_{\mathbb{R}}(I) : g(x) < 0\}|. \tag{4.1}$$

Proof. See [32], Theorem 2.1. □

From this theorem we get directly the following linear system

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} c(h,>) \\ c(h,<) \\ c(h,=) \end{pmatrix} = \begin{pmatrix} \sigma(S_g) \\ \sigma(S_{g^2}) \\ \sigma(S_1) \end{pmatrix} \tag{4.2}$$

The first equation restates (4.1), the second one uses that $\{x : g^2(x) < 0\} = \emptyset$ and $\{x : g^2(x) > 0\} = \{x : g(x) > 0\} \cup \{x : g(x) < 0\}$. Since for the constant polynomial $g = 1$, (4.1) states $\sigma(S_1) = |\mathcal{V}_{\mathbb{R}}(I)|$ the third line follows from the fact that the three cases $<, >$ and $=$ are exhaustive. We denote the matrix on the left-hand side of (4.2) by A.
Rather than explaining the details of the algorithm in full and abstract generality, we present an example that shows most features of the BKR algorithm. All computations were conducted with *Macaulay 2* [18]. We have implemented the BKR algorithm in its full generality in an extension package `bkr.m2`. This program is not yet published and may be found in Appendix B.

Example. We consider $R = \mathbb{R}[x, y]$ with the ideal $I = \langle x^3 - x - y, y^2 - 4y + x \rangle$ and the polynomial list $\{x, y, x^2 - y - 1\}$. Figure 4.2 shows the geometric situation.

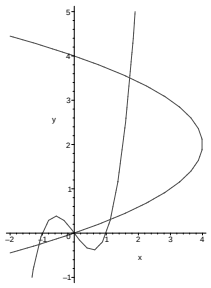


Figure 4.2: The variety $\mathcal{V}_{\mathbb{R}}(I)$: consists of the intersection points of the two curves $y = x^3 - y$ and $x = 4y - y^2$.

A reduced Groebner basis of I with respect to the graded reverse lexicographic order is given by the generating polynomials $\{x^3 - x - y, 2y^2 + x - 4y\}$. I is zero-dimensional and has degree 6. A basis of the 6-dimensional $\mathbb{R}$-algebra Q is given by $\{1, x, x^2, x^2y, xy, y\}$.
In order to set up the linear system we need to determine the matrices S_x, S_{x^2} and S_1. We start with S_x and show for two entries where they come from. The

$(1,1)$-entry of S_x is the trace of $m_{x\cdot 1\cdot 1}$; the $(3,4)$-entry of S_x is the trace of $m_{x\cdot x^2\cdot x^2 y}$. We find that m_x and m_{x^5y} are represented by the matrices

$$T_x = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 & 0 & 0 \end{pmatrix} \quad \text{and} \quad T_{x^5y} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & -4 & -3 & -9 & -17 & -7 \\ 0 & -2 & -4 & -17 & -7 & -1 \\ 4 & 1 & 7 & 24 & 2 & 16 \\ 1 & 7 & 17 & 65 & 24 & 2 \\ 3 & 16 & 2 & 1 & 63 & 8 \end{pmatrix},$$

respectively, since for example (first column of T_{x^5y}) $1\cdot x^5y \equiv 4x^2y+xy-2x+3y \mod I$. We have $\operatorname{tr} T_x = 0$ and $\operatorname{tr} T_{x^5y} = 48$. Going on this way, we get

$$S_x = \begin{pmatrix} 0 & 4 & 12 & 48 & 3 & 0 \\ 4 & 12 & 4 & -1 & 48 & 3 \\ 12 & 4 & 15 & 48 & -1 & 48 \\ 48 & -1 & 48 & 140 & -19 & 188 \\ 3 & 48 & -1 & -19 & 188 & 0 \\ 0 & 3 & 48 & 188 & 0 & -4 \end{pmatrix}$$

with characteristic polynomial $\chi_x(t) = t^6 - 351t^5 - 11181t^4 + 8191780t^3 + 9460777t^2 - 107337245t$. From this, we find with Descartes' rule that the signature $\sigma(S_x) = 1$.

Similarly, we obtain $\sigma(S_{x^2}) = 3$ and $\sigma(S_1) = 4$. The fact that the signature of S_1 equals 4 tells us that there are precisely 4 points in the real variety whereas we know from $\dim Q = 6$ and Theorem 5.3.6 of [12] that the complex variety $\mathcal{V}_{\mathbb{C}}(I)$ contains 6 points. Solving the resulting linear system (4.2) we find $c(x,>) = 2, c(x,=) = 1$ and $c(x,<) = 1$.

Now comes the crucial step for the treatment of several simultaneous polynomial constraints. Two identities like (4.2) may be combined with the help of the matrix tensor product (Kronecker product). In our example we get

$$\left(\begin{array}{ccc|ccc|ccc} 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right) \left(\begin{array}{c} c(x,>;y,>) \\ c(x,>;y,<) \\ c(x,>;y,=) \\ \hline c(x,<;y,>) \\ c(x,<;y,<) \\ c(x,<;y,=) \\ \hline c(x,=;y,>) \\ c(x,=;y,<) \\ c(x,=;y,=) \end{array}\right) = \left(\begin{array}{c} \sigma(S_{x\cdot y}) \\ \sigma(S_{x\cdot y^2}) \\ \sigma(S_{x\cdot 1}) \\ \hline \sigma(S_{x^2\cdot y}) \\ \sigma(S_{x^2\cdot y^2}) \\ \sigma(S_{x^2\cdot 1}) \\ \hline \sigma(S_{1\cdot y}) \\ \sigma(S_{1\cdot y^2}) \\ \sigma(S_{1\cdot 1}) \end{array}\right). \tag{4.3}$$

The matrix on the left-hand side is $A \otimes A$. We may speak of a computation tree which is shown in Figure 4.3.

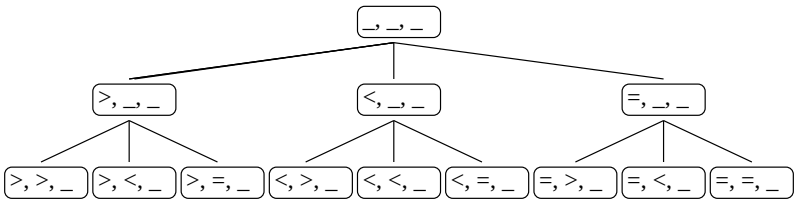


Figure 4.3: Computation tree after considering the first two constraints

As before, we compute the signatures on the right-hand side; it turns out to be the vector $(3, 1, 1, 1, 3, 3, 1, 3, 4)^T$. By solving the linear system we get the vector $(2, 0, 0, 0, 1, 0, 0, 0, 1)^T$, that is, we have $c(x, >; y, >) = 2$, $c(x, <; y, <) = 1$, $c(x, =; y, =) = 1$ and $c(x, \diamond, y, \diamond) = 0$ for all other sign combinations. This means that $\mathcal{V}_{\mathbb{R}}(I)$ consists of two points in the first quadrant, one point in the negative quadrant and the origin.

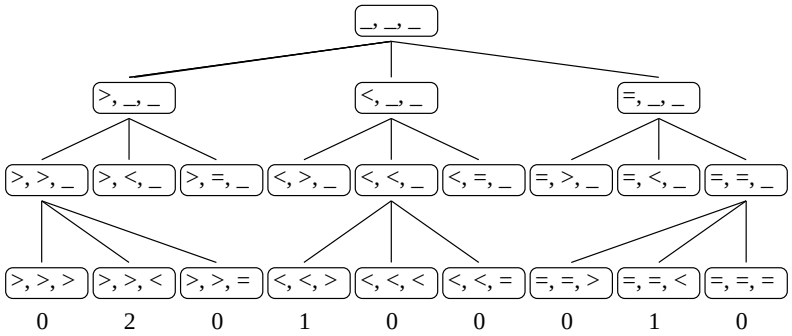


Figure 4.4: Computation tree after considering all three constraints. The branches of the second level that led to no variety points were pruned.

The second crucial idea of the algorithm is the following "pruning of the computation tree" (Figure 4.4). Writing (4.3) without the vanishing $c(x, \diamond; y, \diamond)$ and deleting the corresponding columns of $A \otimes A$ we get

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} c(x, >; y, >) \\ c(x, <; y, <) \\ c(x, =; y, =) \end{pmatrix} = \begin{pmatrix} \sigma(S_{x \cdot y}) \\ \sigma(S_{x \cdot y^2}) \\ \sigma(S_{x \cdot 1}) \\ \sigma(S_{x^2 \cdot y}) \\ \sigma(S_{x^2 \cdot y^2}) \\ \sigma(S_{x^2 \cdot 1}) \\ \sigma(S_{1 \cdot y}) \\ \sigma(S_{1 \cdot y^2}) \\ \sigma(S_{1 \cdot 1}) \end{pmatrix}.$$

We extract a full-rank submatrix on the left-hand side (first, second and last

row) and take the corresponding entries from the signature vector; this yields

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} c(x,>;y,>) \\ c(x,<;y,<) \\ c(x,=;y,=) \end{pmatrix} = \begin{pmatrix} \sigma(S_{x\cdot y}) \\ \sigma(S_{x\cdot y^2}) \\ \sigma(S_{1\cdot 1}) \end{pmatrix}.$$

This reduced system may be tensored with the original system 4.2 the same way we did to obtain 4.3. This leads to the system

$$\left(\begin{array}{c|c|c} A & A & 0 \\ \hline A & -A & 0 \\ \hline A & A & A \end{array}\right) \begin{pmatrix} c(x,>;y,>;(x^2-y-1),>) \\ c(x,>;y,>;(x^2-y-1),<) \\ c(x,>;y,>;(x^2-y-1),=) \\ c(x,<;y,<;(x^2-y-1),>) \\ c(x,<;y,<;(x^2-y-1),<) \\ c(x,<;y,<;(x^2-y-1),=) \\ c(x,=;y,=;(x^2-y-1),>) \\ c(x,=;y,=;(x^2-y-1),<) \\ c(x,=;y,=;(x^2-y-1),=) \end{pmatrix} = \begin{pmatrix} \sigma(S_{x\cdot y\cdot(x^2-y-1)}) \\ \sigma(S_{x\cdot y\cdot(x^2-y-1)^2}) \\ \sigma(S_{x\cdot y\cdot 1}) \\ \sigma(S_{x\cdot y^2\cdot(x^2-y-1)}) \\ \sigma(S_{x\cdot y^2\cdot(x^2-y-1)^2}) \\ \sigma(S_{x\cdot y^2\cdot 1}) \\ \sigma(S_{1\cdot 1\cdot(x^2-y-1)}) \\ \sigma(S_{1\cdot 1\cdot(x^2-y-1)^2}) \\ \sigma(S_{1\cdot 1\cdot 1}) \end{pmatrix}.$$

The right-hand side turns out to be $(-1,3,3,-3,1,1,-2,4,4)^T$. Solving the linear system we obtain the expected solution

$$\begin{aligned} c(x,>;y,>;(x^2-y-1),<) &= 2 \\ c(x,<;y,<;(x^2-y-1),>) &= 1 \\ c(x,=;y,=;(x^2-y-1),<) &= 1 \end{aligned}$$

and $c(x,\diamond;y,\diamond;(x^2-y-1),\diamond)=0$ for all other sign combinations.
If done with `bkr.m2` the commands are

```
i1 : R = QQ[x,y];

i2 : I = ideal(x^3-x-y,y^2-4*y+x);
o2 : Ideal of R

i3 : load "bkr.m2"
--loaded bkr.m2

i4 : bkr ( {x, y, x^2-y-1}, R, I)
o4 = {{<<>, 1}, {>><, 2}, {==<, 1}}
o4 : List
```

The whole computation took 3.7 seconds. ■

5 Rank-one convexity

With this chapter we turn to the most important special case of directional convexity. The introduction has outlined some of the applications of rank-one convexity. Here is the formal definition.

Definition 5.1
(i) A function $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ *is called* rank-one convex *if it is convex on all rank-one lines, i.e., if the functions* $t \mapsto f(M + tX)$ *are convex for all fixed* $M, X \in \mathbb{R}^{m\times n}$ *with* $\operatorname{rank}(X) = 1$. *Equivalently,* f *satisfies*

$$f(\lambda A + (1-\lambda)B) \le \lambda f(A) + (1-\lambda) f(B)$$

for all $A, B \in \mathbb{R}^{m\times n}$ *with* $\operatorname{rank}(A - B) = 1$ *and* $\lambda \in [0, 1]$.
(ii) For a compact set $\mathcal{M} \subset \mathbb{R}^{m\times n}$ *we define its* (functional) rank-one convex hull *by*

$$\mathcal{M}^{rc} := \{X \in \mathbb{R}^{m\times n} : f(X) \le \sup_{M\in\mathcal{M}} f(M) \text{ for all rank-one convex } f\colon \mathbb{R}^{m\times n} \to \mathbb{R}\}.$$

(iii) For a function $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ *we define its* rank-one convex envelope $\mathrm{R}h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ *by*

$$\mathrm{R}h(X) := \sup\{f(X) : f \text{ rank-one convex with } f(Y) \le h(Y)\ \forall\, Y \in \mathbb{R}^{m\times n}\},$$

i.e., as pointwise supremum of all rank-one convex functions $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ *satisfying* $f(Y) \le h(Y)$ *for all* $Y \in \mathbb{R}^{m\times n}$.

Rank-one convexity makes sense only on matrix spaces. Therefore we write now about $\mathbb{R}^{m\times n}$. Whenever necessary, we will identify $\mathbb{R}^{m\times n}$ with $\mathbb{R}^{mn}$ without switching symbols.
Like in the Chapter 3 we shall drop the term "functional" because we are concerned about functional rank-one convex hulls only.

5.1 Computation of rank-one convex hulls and envelopes

The computation of rank-one convex envelopes of functions is a task that has been investigated previously. The rank-one convex envelope gives an upper bound for the quasiconvex envelope of a function, and the latter is of interest in applications. The computation of the rank-one convex hull can be reduced to the computation of the rank-one convex envelope of a function, namely, the distance function (Lemma 2.9).
Apparently, the most general algorithm was developed by Dolzmann [16, 15]. It computes the rank-one convex envelope of a given function g on a ball $B_r(0)$

and is based on discretization. The values of Rg are computed on a uniform grid $\mathcal{G}_{h,r} := \{hM : M \in \mathbb{Z}^{m\times n}\} \cap B_r(0)$ with mesh size h. The set of all rank-one lines is replaced by the finite set

$$\mathcal{R}_{h,r} := \left\{ hX : X = ab^T, a \in \mathbb{Z}^m, b \in \mathbb{Z}^n, |a|, |b| \leq \sqrt{\frac{2r}{h}} \right\}.$$

We point out that it is necessary to discretize two different spaces, namely the domain of g and the set of all rank-one lines where convexity is required.

Algorithm 5.2 (Dolzmann)
Input: $g\colon B_r(0) \to \mathbb{R}$.
Procedure:

1. Initialize $f_0 := g|_{\mathcal{G}_{h,r}}$ and $i := 0$.
2. For all $P \in \mathcal{G}_{h,r}$ and for all $Q \in \mathcal{R}_{h,r}$ convexify the restriction of f_{i+1} to the grid points on $\{P + tQ : t \in \mathbb{R}\}$
3. If $\|f_{i+1} - f_i\| > \varepsilon$ for some prescribed tolerance ε then set $i := i + 1$ and repeat from 2. Otherwise $f := f_{i+1}$.

Output: An approximation f for the rank-one convex envelope Rg.

For a more detailed description and convergence results of this algorithm see [16]. Numerical examples and some error estimates are documented in [15].
A similar approach and more applications are presented in [1]. The main difference is the way how the space of rank-one lines is discretized. The task is rewritten as global nonlinear optimization problem but admittedly there is no way to estimate the quality of the approximation.
The paper [2] is interested in applications to mathematical models for martensitic materials. These applications suggest a special discretization of the space of rank-one directions. The algorithm yields good results for the specific class of problems it is designed for.
The mentioned algorithms have in common that they have high computational complexity and therefore need a long running time. The quality of the results depends highly on the chosen discretization, especially of the set of rank-one lines. It is very hard to find a good discretization unless it is possible to exploit favorable properties of the particular problem.
High precision is required despite the high costs for it. An example for numerical instability is given in [25]. Even worse for the computation of the rank-one convex hull of a set via the distance function, a discretization-based algorithm may fail at all. Consider, e.g., the case of a two-point set $\{A, B\}$ with $\operatorname{rank}(B - A) = 1$. Algorithm 5.2 will fail if either of the points A, B is not a grid point. The same happens if both are grid points but the rank-one line $\{A + t(B - A) : t \in \mathbb{R}\}$ is not contained in the discretized set $\mathcal{R}_{h,r}$ of rank-one lines.

We will take a new and different approach for the computation of rank-one convex hulls which avoids the discretization of the space of rank-one lines. Instead, we will use the underlying algebraical structure. Since these methods are still to be developed we will keep to finite sets.

5.2 T_k-configurations

We will concentrate on finite sets without rank-one connections but nontrivial rank-one convex hull. We say that a set $\mathcal{M}$ does not contain a rank-one connection if $\operatorname{rank}(A-B) \geq 2$ for all $A, B \in \mathcal{M}$ with $A \neq B$. We already know an example of such a set.

Example. Recall from Proposition 2.7 the *Tartar square*, formed by the four diagonal matrices

$$\begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}.$$

The space $\mathbb{R}^{2\times 2}_{\text{diag}}$ of real diagonal matrices can be identified with the real plane $\mathbb{R}^2$ as done in Figure 2.2; the rank-one directions correspond in this model to the coordinate axes. ■

There are some natural generalizations of the Tartar square. If the four corners are not planar the interior of the tetragon is no longer part of the hull but the "frame" may remain intact. Furthermore we need not restrict ourselves to tetragons. This leads us to the notion of T_k-configurations. We will see examples later.

Definition 5.3 *A finite set $\mathcal{M} = \{M_1, M_2, \dots M_k\} \subset \mathbb{R}^{m\times n}$ of k matrices is called a* T_k-configuration *if there exist a permutation $\sigma \in \operatorname{Sym}(k)$, rank-one matrices $C_1, C_2, \dots C_k$, scalars $\kappa_1, \kappa_2, \dots \kappa_k$ with $\kappa_j > 0$ and matrices $X_1, X_2, \dots X_k \in \mathbb{R}^{m\times n}$ such that the relations*

$$X_{j+1} - X_j = C_j, \qquad M_{\sigma(j)} - X_{j+1} = \kappa_j C_j \tag{5.1}$$

hold, where the index j is counted modulo k.

Unlike the definition in [21] we prefer to define the term "T_k-configuration" rather for sets instead of ordered tuples. We want to avoid the situation that some matrices, improperly indexed, do not "form a T_k-configuration" while they do after a permutation of indices. This requires the use of the permutation σ in the preceeding definition.
The following lemma states that the rank-one convex hull of a T_k-configuration is indeed non-trivial. In particular, the line segments drawn in Figure 5.1 are part of the rank-one convex hull.

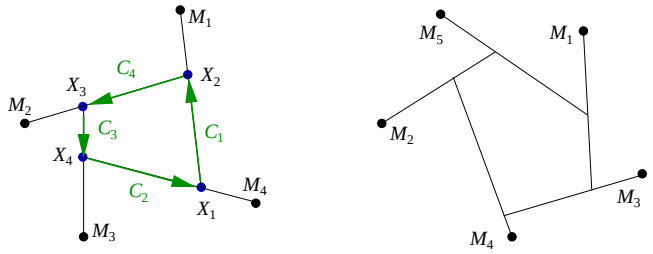


Figure 5.1: General T_4-configuration and T_5 configuration

Lemma 5.4 *Let $\mathcal{M} = \{M_1, M_2, \dots M_k\}$ be a T_k-configuration where the M_i are indexed such that $\sigma = \mathrm{id}$. Then*

$$\bigcup_{j=1}^{k} [M_j, X_j] \;\subseteq\; \mathcal{M}^{rc}.$$

Proof. The reasoning of the proof to Proposition 2.7 carries over. It suffices to prove $X_1, X_2, \dots X_k \in \mathcal{M}^{rc}$. Let $f\colon \mathbb{R}^{2\times 2} \to \mathbb{R}$ be a nonnegative rank-one convex function with $f|_{\mathcal{M}} \equiv 0$. (It is enough to consider such functions due to Lemma 2.5 (i).) Then we have by the convexity of f on the rank-one lines $\{M_j + tC_j : t \in \mathbb{R}\}$

$$f(X_1) \geq f(X_2) \geq \ldots \geq f(X_k) \geq f(X_1)$$

hence $f(X_j) = \text{const.}$ for all j. Since $f(X_1) > 0$ would imply

$$f(X_2) \leq f(M_1)\frac{\|M_1 - X_2\|}{\|M_1 - X_1\|} + f(X_1)\frac{\|X_2 - X_1\|}{\|M_1 - X_1\|} < 0 + f(X_1) = f(X_2)$$

in contradiction to the rank-one convexity of f on $\{M_4 + tC_4 : t \in \mathbb{R}\}$, we may conclude $X_1, X_2, \dots X_k \in \mathcal{M}^{rc}$. □

We remark that all T_2- and T_3-configurations contain necessarily rank-one connections. This is obvious for T_2. As for T_3-configurations, it can be verified by elementary linear algebra that, whenever three points X_1, X_2, X_3 are pairwise rank-one connected, they must lie in a plane that consists of rank-one lines. We postpone this to Corollary 5.13.

Example. (i) The Tartar square from Example 5.2 and Proposition 2.7 is a T_4-configuration.
(ii) An example for a nonplanar T_4-configuration where the interior of the cycle does not belong to the rank-one convex hull is, e.g.,

$$M_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 5 & 1 \\ 2 & 1 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}.$$

In this example, we have $\sigma = \mathrm{id}$,

$$X_1 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad X_4 = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix},$$

$$C_1 = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad C_3 = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}, \quad C_4 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

and $\kappa_1 = \kappa_2 = \kappa_3 = \kappa_4 = 1$. ■

We introduce a special configuration which arises as degenerated limit of T_k-configurations with $k = 4$.

Definition 5.5 *A four-element set $\mathcal{M} = \{M_1, M_2, \dots M_4\} \subset \mathbb{R}^{m\times n}$ is called a* Kirchheim star *(or a* degenerated T_k-configuration*) if there exists a matrix $X \in \mathbb{R}^{m\times n}$ such that*

$$\operatorname{rank}(M_j - X) = 1 \quad (j = 1, 2, 3, 4) \quad \textit{and} \quad X \in \mathcal{M}^{co}, \tag{5.2}$$

where $\mathcal{M}^{co}$ denotes the usual convex hull of $\mathcal{M}$.

An example for a Kirchheim star was already given in Proposition 2.12.

Lemma 5.6 *Let $\mathcal{M}$ be a Kirchheim star. Then we have*

$$\bigcup_{j=1}^{4} [X, M_j] \ \subseteq \ \mathcal{M}^{rc}.$$

Proof. See [22], Example 4.18. □

In the special case $\mathbb{R}^{2\times 2}$, the T_4-configurations are in some sense the universal example for finite sets with nontrivial rank-one convex hull. This is due to the following theorem.

Theorem 5.7 (Székelyhidi, '03) *Let $\mathcal{M} \subset \mathbb{R}^{2\times 2}$ be a compact set without rank-one connections but $\mathcal{M}^{rc} \neq \mathcal{M}$. Then $\mathcal{M}$ contains a T_4-configuration or a Kirchheim star.*

Proof. See [40], Theorem 2. □

For the space of diagonal 2×2-matrices this was already proved by Tartar [41]. The conclusion does not hold in matrix spaces of higher dimensions. There is a counterexample of six points in the space

$$\left\{ \begin{pmatrix} x & y & 0 \\ 0 & 0 & z \end{pmatrix} : x, y, z \in \mathbb{R} \right\} \subset \mathbb{R}^{2\times 3};$$

for details see [21], Proposition 22.

5.3 An algorithm for detection of T_4-configurations

We address the following problem.

Problem 5.8 *Let four matrices $M_1, \dots, M_4 \in \mathbb{R}^{m\times n}$ without rank-one connections (i.e.,* $\operatorname{rank}(M_i - M_j) \geq 2$ *for $i \neq j$) be given.*
Do they form a T_4-configuration?

With our algorithm to solve this problem we are giving a partial answer to a question raised in [21] about the efficient computation of general rank-one convex hulls and in particular for the detection of T_k-configurations. The algorithm we develop in the following generalizes easily to any $k > 4$. The practical usability however is for $k > 4$ limited by the high costs of computations in polynomial rings with many variables.
By Definition 5.3, a T_4-configuration in $\mathbb{R}^{m\times n}$ satisfies (5.1). Since we do not know σ (but without loss of generality we may assume $\sigma(1) = 1$) we have to consider $3! = 6$ systems of equations. For fixed σ and known M_j (5.1) is a system in the variables C_j, κ_j and X_j, i.e., in $8mn + 4$ variables. The system consists of $4\binom{m}{2}\binom{n}{2}+8mn = mn(mn-m-n+9)$ quadratic and linear equations. In $\mathbb{R}^{2\times 2}$, the lowest-dimensional case of interest, this means 36 equations in 36 variables.
The direct approach to check for each of these 6 systems whether it has a solution turns out to be very inefficient. It involves computing a Gröbner basis for an ideal with $O(m^2n^2)$ generators in a polynomial ring in $8mn$ indeterminates. The computer algebra package *Macaulay 2* [18] which uses very efficient algorithms for this task could not solve this problem in reasonable time even for simple examples. For results concerning the exponential complexity of the computation of Gröbner bases see [5].
A much more efficient algorithm can be found by using the underlying algebraic and geometric structures. We will derive several necessary or sufficient conditions for the presence of a T_4-configuration. First, we need a new definition.

Definition 5.9 *The* rank-one cone *with respect to a matrix $M \in \mathbb{R}^{m\times n}$ is the set*

$$\mathscr{R}_1(M) := \{X \in \mathbb{R}^{m\times n} : \operatorname{rank}(X - M) \leq 1\},$$

i.e., the set of all matrices that are rank-one connected to M.

We have obviously $\mathscr{R}_1(M) = M + \mathscr{R}_1(0)$ where 0 represents the zero matrix. In order to describe $\mathscr{R}_1(0)$ algebraically we use the following notation: Let $X = (x_{ij})$ be an $m \times n$-matrix of the indeterminates $x_{11}, x_{12}, \dots, x_{1n}, x_{21}, \dots, x_{mn}$. We denote the real polynomial ring in these indeterminates by $\mathbb{R}[X]$. Whenever necessary, we will silently identify $\mathbb{R}^{mn}$ and $\mathbb{R}^{m\times n}$ without switching symbols.

A matrix has rank one if and only if all 2×2-minors vanish: a nonzero 2×2-minor would indicate two linearly independent rows and therefore a rank ≥ 2. This observation means that $\mathscr{R}_1(0)$ is the common zero set of the $\frac{1}{4}m(m-1)n(n-1)$ quadratic polynomials

$$\det \begin{pmatrix} x_{rs} & x_{ru} \\ x_{ts} & x_{tu} \end{pmatrix}, \qquad 1 \leq r < t \leq m, \quad 1 \leq s < u \leq n. \tag{5.3}$$

These polynomials generate an ideal in $\mathbb{R}[X]$, and $\mathscr{R}_1(0)$ is the algebraic variety associated to this ideal. As every $Y \in \mathscr{R}_1(0)$ can be written as $Y = vw^T$ with $v \in \mathbb{R}^m, w \in \mathbb{R}^n$, uniquely up to multiplication of v and w with reciprocal scalars, the dimension of $\mathscr{R}_1(0) \subset \mathbb{R}^{m \times n}$ equals $m + n - 1$.
Let us return to the T_4 and let us assume for the moment that the matrices M_1, M_2, M_3, M_4 form a T_4-configuration with $\sigma = \mathrm{id}$. A first important observation is that the corners of the inner tetragon lie in the intersections of rank-one cones, i.e., we have

$$X_i \in \mathscr{J}_j := \mathscr{R}_1(M_j) \cap \mathscr{R}_1(M_{j-1}).$$

As intersection of varieties $\mathscr{J}_j$ is again a variety (Lemma 4.3). We find that if $m, n \geq 3$ then $\mathscr{J}_j$ is generically empty. For $m = 2$ or $n = 2$, the variety $\mathscr{J}_j$ is generically a two-dimensional surface. In the case of $\mathbb{R}^{2\times 2}$, an elementary but tedious computation shows that this intersection surface is a one-sheeted hyperboloid unless M_j and M_{j-1} are rank-one connected (see Theorem 5.14). We omit the proof for this special structure because we will not exploit this particular special geometric structure.
The variety $\mathscr{J}_j$ is described by the ideal $I_j \subset \mathbb{R}[X_j]$ generated by the 2×2-minors of $(X_j - M_j)$ and $(X_j - M_{j-1})$, i.e., by the quadratic polynomials

$$\det \begin{pmatrix} x_{j,rs} - M_{j,rs} & x_{j,ru} - M_{j,ru} \\ x_{j,ts} - M_{j,ts} & x_{j,tu} - M_{j,tu} \end{pmatrix}, \det \begin{pmatrix} x_{j,rs} - M_{(j-1),rs} & x_{j,ru} - M_{(j-1),ru} \\ x_{j,ts} - M_{(j-1),ts} & x_{j,tu} - M_{(j-1),tu} \end{pmatrix},$$

$$1 \leq r < t \leq m, \quad 1 \leq s < u \leq n \tag{5.4}$$

where $\mathbb{R}[X_j]$ is the polynomial ring in the mn indeterminates of the above polynomials (just as was $\mathbb{R}[X]$) and $M_{j,rs}$ denotes the rs entry in matrix M_j. The other condition we need is that, for each j, the matrices M_j, X_{j+1} and X_j lie on a line, and in this particular order. This yields the equations and inequalities

$$\lambda_j M_j + (1 - \lambda_j) X_j = X_{j+1}, \qquad 0 < \lambda_j < 1, \qquad \text{for } 1 \leq j \leq 4. \tag{5.5}$$

In order to describe this in terms of varieties we introduce, the same way as $\mathbb{R}[X_j]$ above, the polynomial ring $\mathcal{P} := \mathbb{R}[X_1, X_2, X_3, X_4, \lambda_1, \lambda_2, \lambda_3, \lambda_4]$ in

$4mn+4$ indeterminates. Then we obtain from (5.5) naturally the polynomials

$$\lambda_j M_{j,rs} + (1-\lambda_j)x_{j,rs} - x_{(j+1),rs} \qquad \text{for } 1 \le j \le 4,\ 1 \le r \le m,\ 1 \le s \le n. \tag{5.6}$$

These $4mn$ polynomials and the polynomials in (5.4), the latter taken for all $1 \le j \le 4$, generate an ideal $I \subset \mathcal{P}$. For $\sigma \in \mathrm{Sym}(4)$ we mean by I_σ the ideal generated similarly but where in (5.4) and (5.6) the M_j are substituted by $M_{\sigma^{-1}(j)}$. The real variety associated to I_σ will be denoted by $\mathscr{V}_\sigma \subset \mathbb{R}^{4mn+4}$. The following proposition summarizes the situation.

Proposition 5.10 *Let $\mathcal{M} = \{M_1, M_2, M_3, M_4\} \subset \mathbb{R}^{m\times n}$. Then $\mathcal{M}$ is a T_4-configuration if and only if there is a $\sigma \in \mathrm{Sym}(4)$ such that $\mathscr{V}_\sigma \subset \mathbb{R}^{4mn+4}$ contains a point $(X_1, X_2, X_3, X_4, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$ with $\lambda_j \in (0,1)$ for $1 \le j \le 4$.*

This leads to the following algorithm, commented below.

Algorithm 5.11
Input: $\mathcal{M} = \{M_1, M_2, M_3, M_4\} \subset \mathbb{R}^{m\times n}$ without rank-one connections.
Procedure: For all $\sigma \in \mathrm{Sym}(4)$ conduct the following test. $(*)$

1. For $j = 1, 2, 3, 4$ compute a Gröbner basis for the ideal $I_{\sigma,j}$ generated by the 2×2-minors of $(X_j - M_{\sigma^{-1}(j)})$ and the 2×2-minors of $(X_j - M_{\sigma^{-1}(j-1)})$.
   If $I_{\sigma,j} = \mathbb{R}[X_j]$ for some j then test failed: return to $(*)$ and continue with the next σ.
2. Compute a Gröbner basis for the ideal generated by the polynomials in (5.6) where the M_j are substituted by $M_{\sigma^{-1}(j)}$.
3. Compute a Gröbner basis for the ideal I_σ generated by the union of the ideals in step 1 and 2.
   If $I_\sigma = \mathcal{P}$ then test failed: return to $(*)$ and continue with the next σ.
4. Apply the BKR algorithm to the quotient ring $\mathcal{P}/I_\sigma$, if applicable, and the polynomials $p_j := \lambda_j(1-\lambda_j)$ in order to determine if there is a point $(X_1, X_2, X_3, X_4, \lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathscr{V}_\sigma$ with $\lambda_j \in (0,1)$ for all $1 \le j \le 4$.
   If not, test failed: return to $(*)$ and continue with the next σ.
   If yes, END: this is a T_4.

Output: $\mathcal{M}$ is a T_4-configuration if this was found in Step 4. $\mathcal{M}$ is not a T_4-configuration if the test failed for all $\sigma \in \mathrm{Sym}(4)$.

Step 1 corresponds to (5.4). If any of the ideals $I_{\sigma,j}$ equals the whole ring $\mathcal{P}$ we may conclude that $\mathscr{V}_\sigma$ is empty and therefore by Proposition 5.10 that this σ cannot be used to satisfy (5.1).
Step 2 corresponds to (5.6) and Step 3 takes the union of the ideals arising from (5.4) and (5.6). Here we may find again that $\mathscr{V}_\sigma$ is empty and abort the test.

Step 4 is applicable only if the ideal I_σ is zero-dimensional in $\mathcal{P}$. This is true in all examples we conducted, including all degenerated cases. However, we could not prove that it can never happen that I_σ fails the assumptions in Theorem 4.8.
The BKR algorithm is described in Section 4.3 and the references therein. It yields a lot more information than we actually need as we only want to know whether the number of points of the variety $\mathscr{V}_\sigma \subset \mathbb{R}^{4mn+4}$ in the region given by $0 < \lambda_j < 1$ for $j = 1, 2, 3, 4$ or, equivalently, in the region

$$\left\{(X_1, X_2, X_3, X_4, \lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^{4mn+4} : \lambda_j(1 - \lambda_j) > 0 \text{ for } j = 1, 2, 3, 4\right\} \tag{5.7}$$

is zero or strictly greater than zero. We use therefore a simplified version of the BKR algorithm that is tailored to our problem and runs much more efficient. The last step of the algorithm merely restates Proposition 5.10.
We emphasize that this algorithm answers only the binary question whether a given set forms a T_4-configuration or not. To find explicitly the inner quadrangle, or more precisely the X_j, numerical procedures are the adequate means. Newton's method can provide a good approximation. Numerical algorithms may fail since the computation of rank-one convex hulls is not numerically stable.

5.4 Improvements and generalizations

An implementation of Algorithm 5.11 in the software package *Macaulay 2* [18] can be found in Appendix A. It contains some improvements that contribute to its efficiency.
Since the BKR algorithm tends to be expensive (see [34] for comments on complexity)—even the simplified version of it—we should look for possibilities to avoid it. A necessary condition for all four conditions in (5.7) to be satisfied is of course that each of them is fulfilled. One way to check whether one polynomial sign condition can be satisfied is the eliminant method (see Algorithm 4.10 in Section 4.2).
Algorithm 4.10 is used four times in the procedure after Step 3 with the polynomials $p_j := \lambda_j(1 - \lambda_j)$ $(j = 1, 2, 3, 4)$ and the ideal $I_\sigma \subset \mathcal{P}$. If we find that, for some j, there is no point $y \in \mathscr{V}_\sigma$ where the polynomial p_j takes a positive value, then the criterion in Proposition 5.10 cannot be satisfied, i.e., we may abort the current test in Algorithm 5.11, skip Step 4, and return to $(*)$.

The eliminant method allows us to check for a Kirchheim stars as well. Definition 5.5 translates to the algebraic relations

$$\det \begin{pmatrix} x_{rs} - M_{j,rs} & x_{ru} - M_{j,ru} \\ x_{ts} - M_{j,ts} & x_{tu} - M_{j,tu} \end{pmatrix} = 0, \qquad 1 \le j \le 4, \quad \begin{matrix} 1 \le r < t \le m, \\ 1 \le s < u \le n, \end{matrix}$$

$$X = \sum_{j=1}^{4} \mu_j M_j, \qquad \sum_{j=1}^{4} \mu_j = 1, \qquad \mu_j > 0 \ \text{ for } \ 1 \leq j \leq 4 \tag{5.8}$$

in the $mn + 4$ indeterminates $X = (x_{rs})$ and μ_j.
If the eliminant method finds 0 as common root of the four characteristic polynomials related to the coordinates λ_j then it may be possible that there is a point in $\mathscr{V}_\sigma$ with $\lambda_j = 0$ for all $1 \leq j \leq 4$, a necessary condition for a Kirchheim star. In this case the BKR method should be used to check whether the variety associated to (5.8) is empty or not. In the latter case we may conclude that we have a Kirchheim star.

5.5 Statistical experiments and results

As immediate application of our algorithm, we have conducted extensive test series on random matrices in $\mathbb{R}^{2\times 2}$ and $\mathbb{R}^{3\times 3}$. There were several reasons for this. One was to check the usability and efficiency of the implementation with more than only the known, manually constructed examples. Secondly we were looking for new examples for T_4-configurations, Kirchheim stars and other interesting cases such as T_4-configurations which admit solutions to (5.1) for more than one $\sigma \in \mathrm{Sym}(4)$. Furthermore we tried to get an impression how frequent T_4-configurations are, i.e., how often we may expect a given four-element set $\mathcal{M}$ of matrices to form a T_4 configuration.
The implementation of Algorithm 5.11 with the enhancements described in Section 5.4 in the computer algebra package *Macaulay 2* [18] is found in Appendix A. All computations were done on a Dual Pentium III with 1GHz pulse frequence. For every experiment, we had *Macaulay 2* generate four random matrices $\mathcal{M} = \{M_1, M_2, M_3, M_4\}$ with integer entries in $[0, R]$ for $R = 20, 30, 50, 150$. First, the set $\mathcal{M}$ was translated such that M_4 coincided with the zero matrix and checked for possible rank-one connections (in which case the experiment was aborted). Then the algorithm, including measures listed in Section 5.4, was applied to the translated set.
After every 2500 experiments, *Macaulay 2* was restarted. This had purely technical reasons. The package gets significantly slower during a lengthy computation, probably because of a suboptimal implementation of the garbage collector that clears the memory.
Table 1 shows some results. In the case of $\mathbb{R}^{2\times 2}$ we found many examples for T_4-configurations. Apparently the set of all T_4-configurations, considered as a subset of $(\mathbb{R}^{2\times 2})^4$, has positive measure as about 9% of all random four-element sets were found to form a T_4. This had not been expected in the beginning since T_4-configurations often seem to be very special. We found even quite a

	$\mathbb{R}^{2\times2}$			$\mathbb{R}^{4\times2}$	$\mathbb{R}^{3\times3}$
Range R	30	50	150	50	20
Number of experiments	5000	40000	40000	20000	60000
with a rank-one connection	748	607	110	0	0
T_4 configurations	368	3487	3554	0	0
thereof sixfold T_4 configurations	2	88	65	0	0
Kirchheim stars	0	0	0	0	0
not a T_4-configuration	3884	35906	36336	20000	60000
Average time per experiment	n/a	8.77 s	9.70 s	3.71 s	0.42 s

Table 1: Overview over results of our statistical experiments

few *sixfold T_4-configurations*. By this term, we mean a set $\mathcal{M}$ such that the equations (5.1) can be satisfied for *every* $\sigma \in \mathrm{Sym}(4)$. An example for this constellation is

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 5 & 8 \\ 1 & 5 \end{pmatrix}, \quad \begin{pmatrix} 3 & -9 \\ -2 & 2 \end{pmatrix}, \quad \begin{pmatrix} -4 & -3 \\ -6 & -3 \end{pmatrix}.$$

As predicted for $\mathbb{R}^{2\times2}$ by Székelyhidi ([40], Theorem 3) which we got to know during the execution of our experiments, we found no twofold or threefold T_4-configurations. A T_4-configuration admits a real solution for (5.1) either for only one σ (up to a rotation) or for all $\sigma \in \mathrm{Sym}(4)$. A larger range of entries in the matrices leads obviously to less configurations with rank-one connections. In the cases of $\mathbb{R}^{3\times3}$ and $\mathbb{R}^{4\times2}$ no random set of matrices turned out to be a T_4-configuration. In $\mathbb{R}^{3\times3}$, no random configuration yielded four nonempty intersections $\mathscr{J}_j$ of the respective rank-one cones. It was already a rare exception (ca. 0.1% of experiments) to get just one nonempty intersection. The generically empty intersections were mentioned above. The rank-one cones are only five-dimensional objects in a nine-dimensional space; therefore the fact is intuitively not surprising. In $\mathbb{R}^{4\times2}$ however, the $\mathscr{J}_j$ are two-dimensional, but the ideal I_σ equaled the whole ring $\mathcal{P}$ in all experiments.

5.6 Rank-one convexity in $\mathbb{R}^{2\times2}$

We conclude this chapter with several facts about the geometrical structure of rank-one cones in the particularly interesting case of $\mathbb{R}^{2\times2}$.
We start for later reference with an observation from basic linear algebra.

Lemma 5.12 *Let $A, B \in \mathbb{R}^{2\times2}$ with $\mathrm{rank}(A), \mathrm{rank}(B) \leq 1$ and $\mathrm{rank}(A-B) \leq 1$. Then $\mathrm{rank}(\lambda A + \mu B) \leq 1$ for all $\lambda, \mu \in \mathbb{R}$.*

(In this situation we say that A and B lie in a rank-one plane*, i.e., in a plane consisting of rank-one matrices.)*

Proof. In $\mathbb{R}^{2\times 2}$, a matrix A has rank less or equal than one if, and only if, $\det(A) = 0$. For every $A, B \in \mathbb{R}^{2\times 2}$ and $\lambda, \mu \in \mathbb{R}$ we have

$$\det(\lambda A + \mu B) = \lambda^2 \det(A) + \mu^2 \det(B) + \lambda\mu(a_{11}b_{22} + a_{22}b_{11} - a_{12}b_{21} - a_{21}b_{12}),$$

hence under the assumption $\det(A) = \det(B) = 0$ we obtain $\det(\lambda A + \mu B) = \lambda\mu(a_{11}b_{22}+a_{22}b_{11}-a_{12}b_{21}-a_{21}b_{12})$. In particular, we get $\det(A-B) = -a_{11}b_{22}-a_{22}b_{11} + a_{12}b_{21} + a_{21}b_{12}$. If we suppose that $\operatorname{rank}(A - B) \leq 1$; this implies that the right-hand side is zero, hence the assertion. □

Now we see that there are no nontrivial T_3-configurations.

Corollary 5.13 *Let $\mathcal{M}$ be a T_3-configuration. Then the elements of $\mathcal{M}$ are pairwise rank-one connected. In particular, $\mathcal{M}^{rc}$ equals the usual convex hull of $\mathcal{M}$.*

Proof. Let $\mathcal{M}$ be a T_3-configuration. In particular, there exist rank-one matrices C_1, C_2, C_3 with $C_1 - C_2 = C_3$ that satisfy Definition 5.3. The same property holds, for every choice of $j, k \in \{1, \dots m\}$ and $r, s \in \{1, \dots n\}$, for the 2×2-submatrices of the C_i consisting of the entries at the positions jr, js, kr, ks. Since all 2×2-minors vanish for a rank-one matrix Lemma 5.12 shows that the affine plane which is spanned by $\mathcal{M}$ consists of rank-one lines. In particular, the elements of $\mathcal{M}$ are pairwise rank-one connected. □

We turn to the structure of the intersection surface of the rank-one cones with respect to two different matrices.

Theorem 5.14 *Let A and B be distinct matrices in $\mathbb{R}^{2\times 2}$ and $\mathscr{J} := \mathscr{R}_1(A) \cap \mathscr{R}_1(B)$. Then we have the following two cases:*

- *If $\operatorname{rank}(A - B) \leq 1$ then $\mathscr{J}$ consists of two intersecting planes.*
- *If A, B are not rank-one connected then $\mathscr{J}$ is a one-sheeted hyperboloid.*

Proof. The theorem is shown by elementary analytic geometry. Without loss of generality we may assume B to be the zero matrix. Denote

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$$

then the conditions for $X \in \mathscr{J} := \mathscr{R}_1(A) \cap \mathscr{R}_1(B)$ read

$$\begin{aligned} wz - xy &= 0 \\ (a - w)(d - z) - (b - x)(c - y) &= 0 \end{aligned} \tag{5.9}$$

where we can simplify the second equation by removing the brackets and substituting the first equation, and we obtain

$$(ad - bc) - az - dw + by + cx = 0. \tag{5.10}$$

We assume now first $a \neq 0$. Then we obtain by multiplying (5.9) with a and inserting (5.10)

$$\begin{aligned} & w(by + cx - dw + \det(A)) - axy &=& \ 0 \\ \Leftrightarrow \qquad & dw^2 - bwy - cwx + axy &=& \ w\det(A) \end{aligned}$$

or in matrix form

$$\begin{pmatrix} w & x & y \end{pmatrix} \begin{pmatrix} d & -\frac{c}{2} & -\frac{b}{2} \\ -\frac{c}{2} & 0 & \frac{a}{2} \\ -\frac{b}{2} & \frac{a}{2} & 0 \end{pmatrix} \begin{pmatrix} w \\ x \\ y \end{pmatrix} = w\det(A). \tag{5.11}$$

We denote the symmetric matrix by S. We can write $\mathscr{J}$ then as

$$\mathscr{J} = \left\{ \begin{pmatrix} w & x \\ y & z \end{pmatrix} : (w, x, y) \text{ solves (5.11) and } z = \frac{1}{a}(\det(A) - dw + by + cx) \right\}.$$

As z can be expressed by the other three variables we can identify $\mathscr{J}$ with the solution set of Equation (5.11) which defines a quadric in $\mathbb{R}^3$. There is a complete classification of the 16 geometric objects that may occur as solution set to a quadratic system of the form $v^T S v + s^T x + c = 0$ with $S \in \mathbb{R}^{3\times 3}$, $s \in \mathbb{R}^3$, $c \in \mathbb{R}$ and the the unknown $v \in \mathbb{R}^3$, see, e.g., [28]. We now proceed to identifying which ones qualify as solutions to (5.11).
S has the characteristic polynomial

$$\chi_S(t) = t^3 - dt^2 - \frac{1}{4}t(a^2 + b^2 + c^2) + \det(A).$$

This polynomial has three real roots $\lambda_1 \geq \lambda_2 \geq \lambda_3$ (as S is symmetric). Between any two roots there must be a root of the derivative $\chi_S'(t) = 3t^2 - 2dt - \frac{1}{4}(a^2 + b^2 + c^2)$. The roots of this derivative are

$$r_\pm = \frac{2d \pm \sqrt{4d^2 + 3(a^2 + b^2 + c^2)}}{6}$$

and since $a^2 + b^2 + c^2 \geq a^2 > 0$ by assumption we have $r_+ > 0 > r_-$, hence S has at least one positive eigenvalue λ_1 and at least one negative eigenvalue λ_3. The second property of S that we can see from the characteristic polynomial is that 0 is an eigenvalue if and only if $\det(A) = 0$, and because of $a \neq 0$ we have in this case $\operatorname{rank}(S) = 2$. From the classification we may conclude that $\mathscr{J}$ consists of two intersecting planes.

The possibilites for the case $\det(A) \neq 0$, i.e., the case that A is not a rank-one matrix, are a one-sheeted hyperboloid, a two-sheeted hyperboloid and a cone. We will return to this point after treating the case $a = 0$.
If $a = 0$ we will first assume $d \neq 0$. Similarly we eliminate w and arrive at the equations for $\mathscr{J}$

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 0 & \frac{d}{2} & -\frac{c}{2} \\ \frac{d}{2} & 0 & -\frac{b}{2} \\ -\frac{c}{2} & -\frac{b}{2} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \det(A)$$
$$dw - cx - by = \det(A).$$

The analysis of this equation carries over and we draw the same conclusions. For $a = d = 0$ and $b \neq 0$ we find

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{b}{2} \\ 0 & c & 0 \\ \frac{b}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \det(A)$$
$$-cx - by = \det(A).$$

The eigenvalues of the matrix are $c, \frac{b}{2}$ and $-\frac{b}{2}$ hence there exist a positive and a negative eigenvalue, and only in the case of $\det(A) = 0$, that is $c = 0$, zero is an eigenvalue and we have two intersecting planes.
The case $a = b = d = 0$ and $c \neq 0$ corresponds to a rank-one matrix A and yields two intersecting planes.

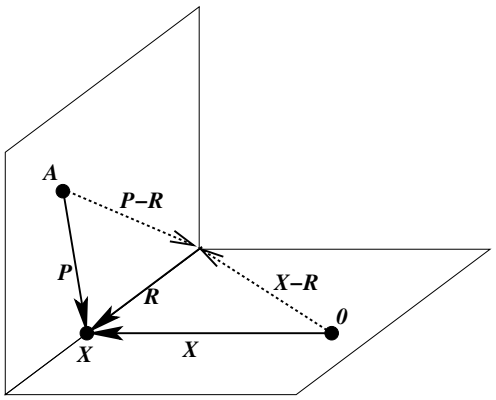


Figure 5.2: Notation for the last part of proof 5.14

It remains to show that, in the case of $\operatorname{rank}(A) = 2$, a cone or a two-sheeted hyperboloid cannot occur. To see this we observe that every point $X \in \mathscr{J}$ lies on precisely two rank-one lines that are contained in $\mathscr{J}$. With $P = A - X$ and $Q = B - X$, we need to show that there exist R_1, R_2 with $\det(R_j) = \det(P - R_j) = \det(Q - R_j) = 0$. Figure 5.2 illustrates the situation. We will

not distinguish explicitely between elements of $\mathbb{R}^4$ or $\mathbb{R}^{2\times 2}$ and their tangent spaces.
The three conditions read

$$\begin{aligned} p_{11}p_{22} + r_{11}r_{22} - p_{11}r_{22} - p_{22}r_{11} - p_{12}p_{21} - r_{12}r_{21} + p_{12}r_{21} + p_{21}r_{12} &= 0 \\ x_{11}x_{22} + r_{11}r_{22} - x_{11}r_{22} - x_{22}r_{11} - x_{12}x_{21} - r_{12}r_{21} + x_{12}r_{21} + x_{21}r_{12} &= 0 \\ r_{11}r_{22} - r_{12}r_{21} &= 0 \end{aligned}$$

and because of $\det(A - X) = \det(X) = 0$ this yields

$$\begin{aligned} -p_{11}r_{22} - p_{22}r_{11} + p_{12}r_{21} + p_{21}r_{12} &= 0 \\ -x_{11}r_{22} - x_{22}r_{11} + x_{12}r_{21} + x_{21}r_{12} &= 0 \\ r_{11}r_{22} - r_{12}r_{21} &= 0. \end{aligned}$$

The first two equations are linearly independent since $\operatorname{rank}(A) = 2$ hence they define a two-dimensional plane in $\mathbb{R}^4$ containing the origin. This plane intersects $\mathscr{R}_1(0)$ in two intersecting lines, so we find, up to multiplication with a real constant, two solutions R.
This concludes the proof because out of the three objects in question for $\mathscr{J}$, namely the cone, the two-sheeted hyperboloid and the one-sheeted hyperboloid, only the last one has the property that each element $X \in \mathscr{J}$ is contained in two lines that lie in $\mathscr{J}$. □

Acknowledgements

First of all I would like to thank my advisor Johannes Zimmer for idea to and his active support for this interesting thesis project and for arranging the necessary permissions for my stay at the California Institute of Technology. He was always accessible for advice (also advice not strictly related to mathematics) and open for suggestions regarding the particular direction of the research.
Further thanks go to Kaushik Bhattacharya and Martin Brokate, professors at the California Institute of Technology and the Technische Universität München, respectively. They made it possible for me to write this thesis under the supervision of Johannes Zimmer at the California Institute of Technology, the former by accepting me as a guest in his group, the latter by being "Aufgabensteller der Diplomarbeit" and accepting my thesis for evaluation. The Freistaat Bayern supported my stay financially.
For discussions on the topics in or related to my work, I would like to thank (additionally and alphabetically) Isaac Chenchiah, Stefan Müller, Bernd Sturmfels, László Székelyhidi, and Thorsten Theobald. For various help during the time of writing up the thesis I express my gratitude to Patrick Dondl, Florian Rupp, and (last but not least) my parents.

Appendix A. *Macaulay 2* implementation for the detection of T_4-configurations

```
_

load "bkr.m2"
load "kirchheim.m2"
-----------------------------------------------------------------------
-- Twodtest receives 2x2-matrices A,B,C as input and tests this ordered
tuple for being a T4.
-- The fourth matrix is always the zero matrix, in the comments denoted
with D.
--
twodtest = (Aa,Bb,Cc,nn) -> (
-----------------------------------------------------------------------
-- Initializing the ring: X = intersection(hyperplane,ROcone(A),ROcone(B))
--                        Y = intersection(hyperplane,ROcone(C),ROcone(D))
--                        lam,mue,rho,sig: Parameters.
--
R := QQ[lam1,lam2,lam3,lam4,x1,x2,x3,x4,y1,y2,y3,y4,w1,w2,w3,w4,z1,z2,z3,z4,
  MonomialOrder=>Eliminate 4];
--
-----------------------------------------------------------------------
-- Technical preparations of little importance
--
X := matrix{{x1,x2},{x3,x4}};
Y := matrix{{y1,y2},{y3,y4}};
W := matrix{{w1,w2},{w3,w4}};
Z := matrix{{z1,z2},{z3,z4}};

A := matrix Aa ** R;
B := matrix Bb ** R;
C := matrix Cc ** R;
n := matrix nn ** R;
x := submatrix(X,{0},) | submatrix(X,{1},);
y := submatrix(Y,{0},) | submatrix(Y,{1},);
w := submatrix(W,{0},) | submatrix(W,{1},);
z := submatrix(Z,{0},) | submatrix(Z,{1},);
--
-----------------------------------------------------------------------
--Computation of the normal vector n to the hyperplane by exterior product
```

```
--
-------------------------------------------------------------------------
-- Iab: Conditions for X = intersection(hyperplane,ROcone(A),ROcone(B))
-- Icd: Conditions for Y = intersection(hyperplane,ROcone(C),ROcone(D))
--
Iab := ideal( det(x*transpose n), det(X-A), det(X-B));
GBab := transpose (gens (gb Iab));
Icd := ideal( det(y*transpose n), det(Y-C), det(Y));
GBcd  := transpose (gens (gb Icd));
--
Ibc := ideal( det(w*transpose n), det(W-B), det(W-C));
GBbc := transpose (gens (gb Ibc));
Ida := ideal( det(z*transpose n), det(Z-A), det(Z));
GBda := transpose (gens (gb Ida));
--
-------------------------------------------------------------------------
I1 := ideal ( A_(0,0) + lam1 * (Z-A)_(0,0) - X_(0,0),
A_(0,1) + lam1 * (Z-A)_(0,1) - X_(0,1),
A_(1,0) + lam1 * (Z-A)_(1,0) - X_(1,0),
A_(1,1) + lam1 * (Z-A)_(1,1) - X_(1,1) );
GBi1 := transpose (gens (gb I1));
-------------------------------------------------------------------------
I2 := ideal ( B_(0,0) + lam2 * (X-B)_(0,0) - W_(0,0),
B_(0,1) + lam2 * (X-B)_(0,1) - W_(0,1),
B_(1,0) + lam2 * (X-B)_(1,0) - W_(1,0),
B_(1,1) + lam2 * (X-B)_(1,1) - W_(1,1) );
GBi2 := transpose (gens (gb I2));
-------------------------------------------------------------------------
I3 := ideal ( C_(0,0) + lam3 * (W-C)_(0,0) - Y_(0,0),
C_(0,1) + lam3 * (W-C)_(0,1) - Y_(0,1),
C_(1,0) + lam3 * (W-C)_(1,0) - Y_(1,0),
C_(1,1) + lam3 * (W-C)_(1,1) - Y_(1,1) );
GBi3 := transpose (gens (gb I3));
-------------------------------------------------------------------------
I4 := ideal ( lam4 * Y_(0,0) - Z_(0,0),
lam4 * Y_(0,1) - Z_(0,1),
lam4 * Y_(1,0) - Z_(1,0),
lam4 * Y_(1,1) - Z_(1,1) );
GBi4 := transpose (gens (gb I4));
-------------------------------------------------------------------------
-- Definition of the actual ideal I which describes essentially the
```

```
--        conditions for a T4 configuration.
--
-- NOTE: lam, mu, rho, sig are subject to the constraints
--     lam >= 0, mu >=0, lam+mu  <= 1
--     rho >= 0, sig>=0, rho+sig <= 1
--
-- A nonempty V(I) is not necessarily associated to a T4 configuration.
--
I := I1 + I2 + I3 + I4 + Iab + Ibc + Icd + Ida;
GBi := transpose (gens (gb I));
--
-------------------------------------------------------------------------
-- Criterion 1) We are done if I=R (no T4, empty variety V(I))
-- Criterion 2) This algorithm fails if V(I) is not zero-dimensional.
--
if dim I == -1 then return ("fail", "empty var.(I)");
if dim I >0 then return ("fail", "not generic (dim I>0)");
--
-------------------------------------------------------------------------
-- Counting lam, mu, rho, sig that fulfil at least one of the bound.cond.
-- exam lists the eliminants and the number of their roots between 0 and 1.
-- Followed by various warnings which may indicate a special case.
--
-- A,B,C,D do not form T4 configuration if there is at least one entry =0.
-- Otherwise, all combinations must be checked carefully.
--
QR := R/I;
epimo = map(QR,R,gens R);
--
el1 := eliminant(epimo(lam1),QQ[s]);
el2 := eliminant(epimo(lam2),QQ[s]);
el3 := eliminant(epimo(lam3),QQ[s]);
el4 := eliminant(epimo(lam4),QQ[s]);

if (signAtOne(el1)==0 and signAtOne(el2)==0 and signAtOne(el3)==0
   and signAtOne(el4)==0) then (
      kh:=kirchheim(A,B,C); print kh;
      if (kh=="T4") then return ("kirchheim","Kirchheim star")
);

if num01Roots(el1)*num01Roots(el2)*num01Roots(el3)*num01Roots(el4)==0
```

```
  then return ("fail","failed elim.test");
--print exam;

ct := bkr ({lam1*(1-lam1),lam2*(1-lam2),lam3*(1-lam3),lam4*(1-lam4)},R,I);

print ct_1;
return ct

)



--------------------------------------------------------------------------
--------------------------------------------------------------------------
--Twodproc calls twodtest for all possible permutations of (A,B,C)

twodproc = (A,B,C) -> (
      if (det(A)==0 or det(B)==0 or det(C)==0 or det(A-B)==0
                    or det(A-C)==0 or det(B-C)==0) then (
            print "contains a rank-one connection";
            backup = try (get "twodfile") else " ";
            "twodfile" << backup << endl << "A=" | (net A)^1 | " B=" |
                (net B)^1 | " C=" | (net C)^1 | "contains a rank-one
                connection" << endl  << close;
            return null
      );

      a := submatrix(A,{0},) | submatrix(A,{1},);
      b := submatrix(B,{0},) | submatrix(B,{1},);
      c := submatrix(C,{0},) | submatrix(C,{1},);
      G = a || b || c;
      n := matrix{{det(submatrix(G,,{1,2,3})) ,
                  (-1)*det(submatrix(G,,{0,2,3})) ,
                  det(submatrix(G,,{0,1,3})) ,
                  (-1)*det(submatrix(G,,{0,1,2})) }};
      if n == matrix{{0,0,0,0}} then (
            print "Not a generic configuration, A,B,C,D lie in a plane
                  (not a hyperplane).";
            "DABC" << endl << "A = " << toString A << endl << "B = " <<
                  toString B << endl << "C = " << toString C << endl
                  <<close ;
```

```
            error "Error: A,B,C,D lie in a plane."
      );

      succ:=0;
print "Test 1";
      t1:=twodtest(A,B,C,n) ;
      if t1 _ 0 =="T4" then succ= 1;
print "Test 2";
      t2:=twodtest(A,C,B,n);
      if t2 _ 0 =="T4" then succ= 1;
print "Test 3";
      t3:=twodtest(B,A,C,n);
      if t3 _ 0 =="T4" then succ= 1;
print "Test 4";
      t4:=twodtest(B,C,A,n);
      if t4 _ 0 =="T4" then succ= 1;
print "Test 5";
      t5:=twodtest(C,A,B,n);
      if t5 _ 0 =="T4" then succ= 1;
print "Test 6";
      t6:=twodtest(C,B,A,n);
      if t6 _ 0 =="T4" then succ= 1;

      backup = try (get "twodfile") else " ";
      sucput:="";
      if succ == 1 then sucput="is a T4" else sucput="rejected";
      "twodfile" << backup << endl << "A=" | (net A)^1 | " B=" | (net B)^1
           | " C=" | (net C)^1 | "      " | sucput << endl << net( { t1_1,
           "--", t2_1, "--", t3_1}) << endl << net ( {t4_1, "--", t5_1,
           "--", t6_1 } ) << endl << close;

      print { t1_1, "--", t2_1, "--", t3_1 };
      print { t4_1, "--", t5_1, "--", t6_1 }
)

-----------------------------------------------------------------------
-- Twodrazor: Utility to reset the variables. Necessary for Macaulay2 only.

twodrazor = () -> (
erase symbol lam1; erase symbol lam2; erase symbol lam3; erase symbol lam4;
erase symbol x1;erase symbol x2;erase symbol x3;erase symbol x4;
```

```
erase symbol y1;erase symbol y2;erase symbol y3;erase symbol y4;
erase symbol w1;erase symbol w2;erase symbol w3;erase symbol w4;
erase symbol z1;erase symbol z2;erase symbol z3;erase symbol z4;
return null
)

-----------------------------------------------------------------------------
-- Twodseries tests a series of random matrices.
-- Twodseries expects the file randommatrix.m2 to be already loaded.

twodseries = i -> (
      for j from 1 to i do (
            print j;
            twodrnd ();
            var:=  timing twodproc (A,B,C) ;--else "A=" | (net A)^1 | " B="
                    | (net B)^1 | " C=" | (net C)^1 ;
            backup:= try get "twodfile";
            "twodfile" << backup << j << var <<endl << close;
      )
);

-----------------------------------------------------------------------------
setRandomString = seed -> setRandomSeed fold((i,j) -> 101*i + j, 0, ascii
       seed)
-----------------------------------------------------------------------------
randommatrix = () -> (
      D := matrix {{random 150,random 150},{random 150,random 150}};
      A = matrix {{random 150,random 150},{random 150,random 150}};
      B = matrix {{random 150,random 150},{random 150,random 150}};
      C = matrix {{random 150,random 150},{random 150,random 150}};

      print  A ; print " ";
      print B; print " ";
      print C; print " ";
      print D;

      A = A-D; B = B-D; C= C-D; return null
)

--
```

Appendix B. Implementation of the full BKR algorithm in *Macaulay 2*

```
-- bkr.m2: Implementation of the BenOr-Kozen-Reif algorithm
load "realroots.m2"
-----------------------------------------------------------------------
--bkr: input: list of polynomials (constraints), Ring, Ideal

bkr = (polynomialList,R,I) -> (
  ListOfRows := {0,1,2};
  ListOfRegions := {0,1,2};
  M0 := matrix {{1,-1,0},{1,1,0},{1,1,1}};
  M := M0;
  k:= 0;
  while k < #polynomialList do (             -- tensoring
       if k>0 then (
          ListOfRegions = ListOfRegions | apply(ListOfRegions, i-> i+3^k)
          | apply(ListOfRegions, i-> i + 2*3^k);
          ListOfRows = ListOfRows | apply(ListOfRows, i-> i+3^k)
          | apply(ListOfRows, i-> i + 2*3^k);
          M = M0 ** M
      );
      s := transpose assembleRHS ( ListOfRows, k, polynomialList, R, I);
      sizes = numgens target s;
      print (M | s);                         -- solving
      c := toList( apply( sizes, i -> cramer (i,M,s) ) );
        relativeListOfRegions := {};
      for i from 0 to #c-1 do                -- pruning
           if c_i != 0 then
                relativeListOfRegions = relativeListOfRegions | {i};
      k = k+1;
      M = submatrix(M,,relativeListOfRegions);
      if (numgens source M != numgens target M) then (
           seqq := extractFullRank M;
           M = seqq _ 0;
           relativeListOfRows = seqq _ 1
      ) else relativeListOfRows = toList (0.. (numgens target M)-1);
-- preparing next iteration
      ListOfRegions = takeIndices( ListOfRegions, relativeListOfRegions);
```

```
     ListOfRows = takeIndices( ListOfRows, relativeListOfRows);
     if ListOfRegions=={} then return ("fail",{{}});
     print k; print ListOfRows; print ListOfRegions; print c;
  );
  detM:= det M;
  ct:= select ( c, i-> i!=0);
  apply (#ListOfRegions,
     i -> {uncoding (ListOfRegions_i,#polynomialList) , ct_i})
)
---------------------------------
assembleRHS = (listind,k,listpol, R, I) -> (
     QR := R/I;
     epimorph := map(QR,R,gens R);
     stmp:= {};
     bd:= apply(listind, i->ternaryDecomp i);
     for j from 0 to #bd-1 do (
          bdj:=bd_j;
          p:=1;
          for l from 0 to k do (
               bdjl:= try bdj_l else 0;
               if bdjl == 0 then p = p*listpol_l
                    else if bdjl == 1 then p = p*listpol_l*listpol_l
          );
          stmp = stmp | {signatureWrtPol (epimorph(p))};
     );
     matrix {stmp}
)
---------------------------------
signatureWrtPol = h -> (
     A := ring h;
     assert( dim A == 0 );
     assert( char A == 0 );
     S := QQ[zt];
     TrF := traceForm(h) ** S;
     IdZ := zt * id_(S^(numgens source TrF));
     f := det(TrF - IdZ);
     variations (descartes f) - variations (descartes minusf f)
)
---------------------------------
ternaryDecomp = n -> (
     S:=ZZ/3;
```

```
     if n < 0 then error "Ternary Decomposition expects positive integer.";
     if n == 0 then return {0};
     loc:= {};
     while n>0 do (
          if n_S == 0 then (
               loc = loc | {0};
               n = n /3
          ) else if n_S == 1 then (
               loc = loc | {1};
               n = (n-1)/3
          ) else if n_S == 2 then (
               loc = loc | {2};
               n = (n-2)/3
          ) else error "Error in ternary decomposition."
     );
     loc
)
---------------------------------
cramer = (i,M,s) -> (
     loc:=halfcramer (i,M,s);
     try loc = loc / det M
       else error "Can't use Cramer rule (No division in this ring)";
     loc
)
---------------------------------
halfcramer = (i, M, s) -> (
        colnum= numgens source M;
        sizes = numgens target s;
        assert (colnum == sizes);
        assert (colnum == numgens target M);
        assert (i >= 0 and i < colnum);
        if i==0 then Cr = s | submatrix(M, , toList(1..colnum-1))
             else if i==colnum-1 then
           Cr = submatrix(M, , toList(0.. colnum-2)) | s
             else Cr = submatrix(M, , toList(0..i-1)) | s
                       | submatrix(M, , toList(i+1..colnum-1));
        det Cr
)
---------------------------------
extractFullRank = M -> (
     colnum := numgens source M;
```

```
     rownum := numgens target M;
     if colnum > rownum then error "expected more rows than columns";
     assist=subsets(rownum,colnum);
     for j from 0 to #assist-1 do (
          if det submatrix(M,assist_j,) != 0
               then return (submatrix(M,assist_j,),assist_j)
     );
     error "no submatrix with full rank"
)

---------------------------------
takeIndices = (biglist,smallist) -> (
     tmplist:={};
     for i from 0 to #biglist -1 do
          if member(i,smallist) then tmplist = tmplist | {biglist#i};
     tmplist
)

---------------------------------
uncoding = (n,k) -> (
     tem:=ternaryDecomp n;
     signseq:="";
     thissign:=" ";
     for j from 0 to #tem-1 do (
          if tem_j == 0 then thissign=">";
          if tem_j == 1 then thissign="<";
          if tem_j == 2 then thissign="=";
          signseq = signseq | thissign
     );
     for l from #signseq+1 to k do  signseq = signseq | ">" ;
     signseq
)
--
```

References

[1] Ernesto Aranda and Pablo Pedregal. On the computation of the rank-one convex hull of a function. *SIAM J. Sci. Comput.*, 22(5):1772–1790 (electronic), 2000.

[2] S. Aubry, M. Fago, and M. Ortiz. A constrained sequential-lamination algorithm for the simulation of sub-grid microstructure in martensitic materials. *Computer Methods in Applied Mechanics and Engineering*, 192(26-27):2823–2843, 2003.

[3] Robert J. Aumann and Sergiu Hart. Bi-convexity and bi-martingales. *Israel J. Math.*, 54(2):159–180, 1986.

[4] J. M. Ball and R. D. James. Proposed experimental tests of a theory of fine microstructure and the two-well problem. *Phil. Trans. R. Soc. London (A)*, 338(1650):389–450, feb 15 1992.

[5] David Bayer and Michael Stillman. A theorem on refining division orders by the reverse lexicographic order. *Duke Math. J.*, 55(2):321–328, 1987.

[6] Michael Ben-Or, Dexter Kozen, and John Reif. The complexity of elementary algebra and geometry. *J. Comput. System Sci.*, 32(2):251–264, 1986. 16th annual ACM-SIGACT symposium on the theory of computing (Washington, D.C., 1984).

[7] Sterling K. Berberian. *Linear algebra*. Oxford United Press, Oxford, 1992.

[8] Kaushik Bhattacharya. *Martensitic Phase Transformations*. Oxford University Press, 2002.

[9] Kaushik Bhattacharya and Georg Dolzmann. Relaxation of some multi-well problems. *Proc. Roy. Soc. Edinburgh Sect. A*, 131(2):279–320, 2001.

[10] Béla Bollobás. *Graph theory*, volume 63 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979. An introductory course.

[11] William Snow Burnside. *The theory of equations*. Hodges, Figgins & Co., Dublin, 1881.

[12] David Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1997. An introduction to computational algebraic geometry and commutative algebra.

[13] David Cox, John Little, and Donal O'Shea. *Using algebraic geometry*, volume 185 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[14] Bernard Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 1989.

[15] G. Dolzmann and N. J. Walkington. Estimates for numerical approximations of rank one convex envelopes. *Numer. Math.*, 85(4):647–663, 2000.

[16] Georg Dolzmann. Numerical computation of rank-one convex envelopes. *SIAM J. Numer. Anal.*, 36(5):1621–1635 (electronic), 1999.

[17] Gerd Fischer. *Lineare Algebra*, volume 17 of *Grundkurs Mathematik*. Vieweg-Verlag, Braunschweig, 9 edition, 1989.

[18] Daniel R. Grayson and Michael E. Stillman. Macaulay 2, a software system for research in algebraic geometry. Available at `http://www.math.uiuc.edu/Macaulay2/`.

[19] G.-M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 2.0. A Computer Algebra System for Polynomial Computations, Centre for Computer Algebra, University of Kaiserslautern, 2001. `http://www.singular.uni-kl.de`.

[20] Jonathan Gross and Jay Yellen. *Graph Theory and its Applications*. Discrete Mathematics and its Applications. CRC Press, Boca Raton, Fla., USA, 1999.

[21] B. Kirchheim, S. Müller, and V. Šverák. Studying nonlinear pde by geometry in matrix space. Preprint 49/2002, Max Planck Institute for Mathematics in the Sciences, Leipzig, 2002.

[22] Bernd Kirchheim. Rigidity and geometry of microstructures. Lecture notes 16/2003, Max Planck Institute for Mathematics in the Sciences, Leipzig, 2003.

[23] Jan Kristensen. On the non-locality of quasiconvexity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 16(1):1–13, 1999.

[24] Martin Kružík. Bauer's maximum principle and hulls of sets. *Calc. Var. Partial Differential Equations*, 11(3):321–332, 2000.

[25] J. Matoušek. On directional convexity. *Discrete Comput. Geom.*, 25(3):389–403, 2001.

[26] J. Matoušek and P. Plecháč. On functional separately convex hulls. *Discrete Comput. Geom.*, 19(1):105–130, 1998.

[27] Ernst Mayr. Membership in polynomial ideals over $\mathbb{Q}$ is exponential space complete. In *STACS 89 (Paderborn, 1989)*, volume 349 of *Lecture Notes in Comput. Sci.*, pages 400–406. Springer, Berlin, 1989.

[28] Kurt Meyberg and Peter Vachenauer. *Höhere Mathematik 1*. Springer-Verlag, Heidelberg, 1989.

[29] Charles B. Morrey, Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.

[30] S. Müller and V. Šverák. Convex integration for lipschitz mappings and counterexamples to regularity. Preprint 26/1999, Max Planck Institute for Mathematics in the Sciences, Leipzig, 1999.

[31] Stefan Müller. Rank-one convexity implies quasiconvexity on diagonal matrices. *Internat. Math. Res. Notices*, (20):1087–1095, 1999.

[32] P. Pedersen, M.-F. Roy, and A. Szpirglas. Counting real zeros in the multivariate case. In *Computational algebraic geometry (Nice, 1992)*, volume 109 of *Progr. Math.*, pages 203–224. Birkhäuser Boston, Boston, MA, 1993.

[33] Robert R. Phelps. *Lectures on Choquet's theorem*, volume 1757 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2001.

[34] Marie-Françoise Roy and Aviva Szpirglas. Complexity of computation on real algebraic numbers. *J. Symbolic Comput.*, 10(1):39–51, 1990.

[35] V. Scheffer. *Regularity and irregularity of solutions to nonlinear second order elliptic systems of partial differential equations and inequalities*. PhD thesis, Princeton University, 1974.

[36] Frank Sottile. From enumerative geometry to solving systems of polynomial equations. In David Eisenbud, Daniel R. Grayson, and Michael Stillman, editors, *Computations in algebraic geometry with Macaulay 2*, volume 8 of *Algorithms and Computation in Mathematics*, pages 1–30. Springer-Verlag, Berlin, 2002.

[37] Bernd Sturmfels. Solving systems of polynomial equations. Preprint on homepage, to appear as book, 2003.

[38] Vladimír Šverák. Rank-one convexity does not imply quasiconvexity. *Proc. Roy. Soc. Edinburgh Sect. A*, 120(1-2):185–189, 1992.

[39] Vladimír Šverák. On lower-semicontinuity of variational integrals. *Tatra Mt. Math. Publ.*, 4:217–220, 1994. Equadiff 8 (Bratislava, 1993).

[40] László Székelyhidi. *Elliptic regularity versus rank-one convexity.* PhD thesis, Universität Leipzig, 2003.

[41] Luc Tartar. Some remarks on separately convex functions. In *Microstructure and phase transition*, volume 54 of *IMA Vol. Math. Appl.*, pages 191–204. Springer, New York, 1993.

Contents