Optimization
Vol. 58, No. 6, August 2009, 685–712

On the stratification of a class of specially structured matrices

Peter Jonker, Georg Still and Frank Twilt*

Department of Applied Mathematics, University of Twente, Enschede, The Netherlands

(*Received 9 May 2006; final version received 26 March 2007*)

We consider specially structured matrices representing optimization problems with quadratic objective functions and (finitely many) affine linear equality constraints in an n-dimensional Euclidean space. The class of all such matrices will be subdivided into subsets ['strata'], reflecting the features of the underlying optimization problems. From a differential-topological point of view, this subdivision turns out to be very satisfactory: Our strata are smooth manifolds, constituting a so-called Whitney Regular Stratification, and their dimensions can be explicitly determined. We indicate how, due to Thom's Transversality Theory, this setting leads to some fundamental results on smooth one-parameter families of linear-quadratic optimization problems with (finitely many) equality and inequality constraints.

Keywords: parametric quadratic optimization problems; generalized critical points; Whitney regular stratification

AMS Subject Classifications: 90C30; 58A35; 90C31

1. Introduction

The structure of general non-linear (parametric) finite optimization problems has been studied intensively in the past (see, e.g. [3,4,5,10]). In this article, we deal with the important special subclass of (parametric) linear-quadratic programming problems. In order to introduce the subject of this article, let Q be the following linear–quadratic optimization problem with only equality constraints:

$$Q\begin{cases} \min\limits_{x\in\mathbb{R}^n} \dfrac{1}{2}x^TAx + a^Tx \quad \text{such that} \\ Bx + b = 0, \end{cases}$$

where A is a symmetric $n\times n$-matrix, B an $m\times n$-matrix; $a\in\mathbb{R}^n$, $b\in\mathbb{R}^m$ and $(\cdot)^T$ stands for transpose. Then, the Karush–Kuhn–Tucker equations (KKT) take the matrix form:

$$\text{KKT} \quad \begin{bmatrix} A & B^T \\ B & O \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} = 0,$$

*Corresponding author. Email: f.twilt@ewi.utwente.nl

ISSN 0233–1934 print/ISSN 1029–4945 online

DOI: 10.1080/02331930701763793
http://www.informaworld.com

where O is the $m \times m$-null matrix and $\lambda \in \mathbb{R}^m$. We put

$$M := \begin{bmatrix} A & B^T \\ B & O \end{bmatrix}; \quad c := \begin{bmatrix} a \\ b \end{bmatrix}$$

It is natural to represent the optimization problem Q by the $(m+n) \times (m+n+1)$ matrix $[M \vdots c]$, which is obtained from M by augmenting this matrix with c as $(m+n+1)$th column. In the sequel, the set of all matrices $[M \vdots c]$ will be referred to as the *representation* space $\mathcal{M}_{n,m}$. Note that $\mathcal{M}_{n,m}$ may be identified with $\mathbb{R}^N$, $N = (1/2)n(n+3) + m(n+1)$.

Now we are in the position to present our aim:

1.1. *Aim*

- To give a subdivision [stratification] of $\mathcal{M}_{n,m}$ into finitely many, pair wise disjoint sets – in the sequel called *strata* – such that:
  - Each stratum represents only optimization problems with the *same features*;
  - Each stratum attains a unique 'simplest element' *(normal form)*;
  - All strata are smooth manifolds, constituting a *Whitney regular stratification*;
  - The *dimension* of each stratum can explicitly be expressed in terms of the features of the underlying optimization problem Q.
- To motivate the above subdivision by indicating its relevance for 1-parameter families of linear-quadratic optimization problems in $\mathbb{R}^n$ with finitely many (in-) equality constraints.

This article is organized as follows. In the next section, we spell out the features of the optimization problems Q, underlying the stratification of the representation space $\mathcal{M}_{n,m}$, and derive a characterization of the strata. Moreover, in this section, we shall formulate our main results in terms of three theorems. In Section 3, we explain the relevance of these results for 1-parametric linear-quadratic optimization, and mention some other possible applications. The remaining sections are devoted to the proofs of our three theorems.

2. Results

Let Q be an optimization problem as introduced in Section 1, with A and B its associated matrices. Let S be a matrix with as columns a (linear independent) basis of $\ker(B)$, where $\ker(\cdot)$ stands for 'kernel'. Then the so called *restriction of A to* $\ker(B)$ is defined as $A_{|\ker B} := S^T AS$. The *inertia* of this restriction, denoted $\mathrm{In}(\cdot)$, is given by:

$$\mathrm{In}(A_{\big|\ker B}) := (\xi^+, \xi^-, \xi^0),$$

where ξ^+, ξ^- and ξ^0 are the numbers of positive, negative and zero eigenvalues of $S^T AS$ respectively. Note that, by Sylvester's Law, this inertia is independent of the ambiguity in the choice of S. In the sequel, we put rank $(B) = k$. Thus, S is an $n \times (n-k)$-matrix[1] and $\xi^+ + \xi^- + \xi^0 = n - k$.

The *feasible set* of Q is denoted by $\mathcal{F}_Q$, and let $\mathcal{K}_Q$ be the *Karush–Kuhn–Tucker set* of Q (i.e. the solution set of the equation KKT in Section 1). So, we have:

$$\mathcal{F}_Q := \{x \in \mathbb{R}^n \mid Bx + b = 0\}; \quad \mathcal{K}_Q := \{y \in \mathbb{R}^{n+m} \mid My + c = 0\}.$$

Now we are ready to list the kind of features of Q, which we have in mind.

2.1. *Q-Features*

- $\mathcal{F}_Q$ is empty/non-empty ($\Leftrightarrow$ rank$[B \vdots b] =$ rank$(B) + 1$/rank$[B \vdots b] =$ rank(B)).
- $\mathcal{K}_Q$ is empty/non-empty ($\Leftrightarrow$ rank$[M \vdots c] =$ rank$(M) + 1$/rank$[M \vdots c] =$ rank(M)).
- Feasible set equation is homogeneous/inhomogeneous ($\Leftrightarrow b = 0 \,/\, b \neq 0$).
- KKT-equation is homogeneous/inhomogeneous ($\Leftrightarrow c = 0 \,/\, c \neq 0$).
- Rank(B) $(= k)$.
- If $b = 0$: rank $[B^T \vdots a]$ $(= k/k + 1)$
- In $(A_{|\ker B})$ $(= (\xi^+, \xi^-, \xi^0))$
- Sign $[M \vdots c]$; $:=$ sign $(y^T c)$ for any $y \in \mathcal{K}_Q$, where sign $(\cdot)$ stands for signature.

Note that some of these features are not 'independent' from each other. For example, the condition $\mathcal{K}_Q \neq \emptyset$ makes only sense if $\mathcal{F}_Q \neq \emptyset$ in order to define sign $[M \vdots c]$ it is necessary that $\mathcal{K}_Q \neq \emptyset$. On the other hand, rank(B) and In$(A_{|\ker B})$ are – irrespectively of the other features – always well-defined.[2]

Special attention should be paid to the last feature: Let y' be another vector in $\mathcal{K}_Q$. Then $M(y - y') = 0$ and thus: [use that A and thus also M is symmetric]

$$(y - y')^T c = -(y^T M(y - y'))^T = 0.$$

Hence, sign $[M \vdots c]$ does not depend on the ambiguity in the choice of y.

We collect the above Q-features in groups that respect the possible dependencies:

(1) rank(B); In$(A_{|\ker B})$; $\mathcal{F}_Q = \emptyset$ [thus $b \neq 0$]
(2) rank(B); In$(A_{|\ker B})$; $b \neq 0$; $\mathcal{F}_Q \neq \emptyset$; $\mathcal{K}_Q = \emptyset$
(3) rank(B); In$(A_{|\ker B})$; $b \neq 0$; $\mathcal{F}_Q \neq \emptyset$; $\mathcal{K}_Q \neq \emptyset$ sign$[M \vdots c]$
(4) rank(B); In$(A_{|\ker B})$; $b = 0$ [thus $\mathcal{F}_Q \neq \emptyset$]; $\mathcal{K}_Q = \emptyset$ [thus $a \neq 0$]
(5) rank(B); In$(A_{|\ker B})$; $b = 0$; [thus $\mathcal{F}_Q \neq \emptyset$]; $a \neq 0$; rank$[B^T \vdots a] = k$ [thus $\mathcal{K}_Q \neq \emptyset$]
(6) rank(B); In$(A_{|\ker B})$; $b = 0$ [thus $\mathcal{F}_Q \neq \emptyset$]; rank$[B^T \vdots a] = k+1$ [thus $a \neq 0$]; $\mathcal{K}_Q \neq \emptyset$ sign$[M \vdots c]$
(7) rank(B); In$(A_{|\ker B})$; $b = 0$; $a = 0$ [thus $\mathcal{F}_Q \neq \emptyset$ $\mathcal{K}_Q \neq \emptyset$]

It is easily verified that this 'clustered set of features'[3] induces a partition of the representation set $\mathcal{M}_{n,m}$ into strata $V^{(\ell)}_{\xi,k,[\tau]}$, according to Table 1, where ℓ refers to the underlying cluster of Q-features, ξ to any vector (ξ^+, ξ^-, ξ^0) with non-negative components summing up to $(n-k)$, and $\tau = \pm 1$ or $= 0$ [only if $\ell = 3$ or $\ell = 6$].

Some of the strata are empty. This can be derived from the very definitions of $V^{(\ell)}_{\xi,k,[\tau]}$, but is also a direct consequence of the forthcoming Theorem 1.

We proceed by formulating our main results.

Table 1. The strata of $\mathcal{M}_{n,m}$.

$[M\vdots c] \in V^{(\ell)}_{\xi,k,[\tau]}$ iff		$(A_{\mid \ker B}) = \xi$, rank$(B) = k$, and moreover,	$V^{(\ell)}_{\xi,k,[\tau]} = \emptyset$ iff
$\ell =$			
1		rank$[B \vdots b] = k+1$	$k = m$
2	$b \neq 0$	rank$[B \vdots b] = k$, rank$[M \vdots c] = 1 + \text{rank}(M)$	$k = 0$, or $= n$, or $\xi^0 = 0$
3	$b \neq 0$	rank$[M \vdots c] = \text{rank}(M)$, sign$[M \vdots c] = \tau$	$k = 0$
4	$b = 0$	rank$[M \vdots c] = 1 + \text{rank}(M)$	$\xi^0 = 0$
5	$b = 0$, $a \neq 0$	rank$[B^T \vdots a] = k$	$k = 0$
6	$b = 0$	rank$[B^T \vdots a] = k+1$, rank$[M \vdots c] = \text{rank}(M)$, sign$[M \vdots c] = \tau$	$\xi^+ = \xi^- = 0$, or $\xi^- = 0$, $\tau = +1, 0$ or $\xi^+ = 0$, $\tau = -1$, 0
7	$b = 0$, $a = 0$		

Table 2. Specification of $\bar{n}$ in the normal forms $[\bar{N} \vdots \bar{n}]$.

$\ell =$	$\tau =$	a_1	a_{ξ^+}	a_{ξ^-}	a_{ξ^0}	b_1	b_2
1	–	0	0	0	0	0	e
2	–	0	0	0	e	e	0
3	+1	$-e$	0	0	0	e	0
3	−1	e	0	0	0	e	0
3	0	0	0	0	0	e	0
4	–	0	0	0	e	0	0
5	0	e	0	0	0	0	0
6	+1	0	0	e	0	0	0
6	−1	0	e	0	0	0	0
6	0	0	e	e	0	0	0
7	0	0	0	0	0	0	0

THEOREM 1 (Normal forms) *Each non-void stratum $V^{(\ell)}_{\xi,k,[\tau]}$ contains a unique matrix (normal form) of the type*:

$$[\overline{N} \vdots \overline{n}] := \begin{bmatrix} \overline{A} & \overline{B}^T & \vdots & \overline{a} \\ \overline{B} & O & \vdots & \overline{b} \end{bmatrix},$$

with

$$\overline{A} = \begin{bmatrix} O & O \\ O & J_\xi \end{bmatrix}, \quad J_\xi = \begin{bmatrix} I_{\xi^+} & O & O \\ O & -I_{\xi^-} & O \\ O & O & O \end{bmatrix}, \quad \overline{B} = \begin{bmatrix} I_k & O \\ O & O \end{bmatrix},$$

where I_{ξ^+} stand for the $\xi^+ \times \xi^+$ unit matrix etc., O for null matrices of appropriate dimensions, and $\overline{a}, \overline{b}$ are 'compositions in accordance with the block structure of $\overline{N}$ 'of vectors of the type $0 = (0, \ldots, 0)^T$ ['null'] and $e = (1, 0, \ldots, 0)^T$ ['unit'] of appropriate dimensions. In fact, if

$$\overline{a} = (a_1^T, a_2^T)^T \in \mathbb{R}^k \times \mathbb{R}^{n-k}, \text{ with } a_2 = (a_{\xi^+}^T, a_{\xi^-}^T, a_{\xi^0}^T)^T \in \mathbb{R}^{\xi^+} \times \mathbb{R}^{\xi^-} \times \mathbb{R}^{\xi^0}$$
$$\overline{b} = (b_1^T, b_2^T)^T \in \mathbb{R}^k \times \mathbb{R}^{m-k}$$

then $\overline{a}$ and $\overline{b}$ are as indicated in Table 2.

Table 3. Codimensions of the strata.

$\ell=$	codim $\mathrm{V}^{(\ell)}_{\xi,k,[\tau]}$
1	$(n-k)(m-k)+\frac{1}{2}\xi^0(\xi^0+1)$
2	$(n-k+1)(m-k)+\frac{1}{2}\xi^0(\xi^0+1)$
3	$(n-k+1)(m-k)+\frac{1}{2}\xi^0(\xi^0+3)+1-\tau^2$
4	$m+(n-k)(m-k)+\frac{1}{2}\xi^0(\xi^0+1)$
5	$m+(n-k)(m-k+1)+\frac{1}{2}\xi^0(\xi^0+1)$
6	$m+(n-k)(m-k)+\frac{1}{2}\xi^0(\xi^0+3)+1-\tau^2$
7	$m+n+(n-k)(m-k)+\frac{1}{2}\xi^0(\xi^0+1)$

Some of the blocks/vectors are possibly empty.

Theorem 2 (Manifolds)

- *Each non-void stratum* $V^{(\ell)}_{\xi,k,[\tau]}$ *is a smooth submanifold of* $\mathbb{R}^N$ (*i.e. locally diffeomorphic to an open set in* $\mathbb{R}^q$, *where q stands for* dim $(V^{(\ell)}_{\xi,k,[\tau]})$).
- *codim* $(V^{(\ell)}_{\xi,k,[\tau]})[=N-q]$ *can explicitly be expressed in terms of the parameters* n, m, k, ξ *and* $[\tau]$, *see Table* 3.

Recall that $\xi^+ + \xi^- + \xi^0 = n-k$.

A finite subdivision of $\mathbb{R}^N$ into pair wise disjoint smooth manifolds [strata] is called a *Whitney regular stratification* if neighbouring strata stick together in such a 'regular' way that the *local topological type*[4] remains constant along (the connected components) of each stratum. See [2,10] for a more analytical definition.

With this concept in mind, we have the following result:

Theorem 3 (Whitney regular stratification) *The subdivision of* $\mathcal{M}_{n,m}$ *into the sets* $V^{(\ell)}_{\xi,k,[\tau]}$ *is a Whitney regular stratification.*

We end up this section with a short comment on the relationship between the proofs of Theorems 1, 2 and 3. The proof of the latter theorem is the more sophisticated one since it relies upon some basic properties from algebraic geometry. In fact, let $\mathcal{A}$ be a locally finite partition of $\mathbb{R}^N$ into *semi algebraic* sets[5] (strata). Assume, moreover, that the *homogeneity property*[6] for each stratum of $\mathcal{A}$ holds. Then, basically due to the fact that any non-void semi algebraic set contains at least one regular[7] point (hence, all strata are smooth manifolds), and taking Whitney's Theorem on the 'bad set' of two semi-algebraic smooth manifolds into account, it follows that $\mathcal{A}$ is a Whitney regular stratification (cf [2]). So, we are done when we are able to prove that our strata $V^{(\ell)}_{\xi,k,[\tau]}$ fulfill both conditions.

The homogeneity property is a direct consequence of the proof of Theorem 1 (which is based on the observation that, given any $V^{(\ell)}_{\xi,k,[\tau]}$ and any $[M \vdots c]$ in this stratum, a diffeomorphism from $V^{(\ell)}_{\xi,k,[\tau]}$ onto itself exists which maps $[M \vdots c]$ to the normal form $[\overline{N} \vdots \overline{n}]$ in $V^{(\ell)}_{\xi,k,[\tau]}$). The semi algebraic character of the stratification is proved in Section 6, Corollary 3. We emphasize that the Whitney regularity of our stratification automatically yields the 'manifold statement' in Theorem 2. However, in order to derive the formulae for the codimensions (which are crucial for the applications we have in mind, cf Section 3), we must treat each stratum separately. In fact, using again the homogeneity property, it is sufficient to identify defining systems around the *normal forms* (cf Theorem 1) for *each* stratum. Doing so, we prove again – as a side result – that all our strata are smooth manifolds.

3. Motivation

In parametric optimization, one seeks to investigate how changes of the parameter do influence the characteristics of the optimization problem. In general, this influence will be quite unpredictable. However, in some cases it is possible to identify subclasses of optimization problems, which on one hand are large enough to cover 'almost all' problems under consideration, and on the other hand, are restrictive enough to allow relevant statements. For discussions on this 'genericity point of view', see e.g. the treatises [5] and [10].

In the case of linear-quadratic optimization, the results in Section 2 do play a crucial role in this genericity approach. In order to clarify this role, let $Q(t)$ be a smooth r-parameter family of linear-quadratic optimization problems, i.e. problems Q as introduced in Section 1 for which the entries of A, B, a and b are smooth functions of a parameter $t \in \mathbb{R}^r$. (Note that, if $r = 1$, such a family may be regarded as a *smooth curve* in the representation space $\mathcal{M}_{n,m}$).

Basically due to Thom's Transversality Theorem for Whitney Regular Stratifications (cf [6,10]) and our Theorem 3, we have the following 'genericity result':

- Generically[8] $Q(\mathbb{R}^r)$ only intersects strata $V^{(\ell)}_{\xi,k,[\tau]}$ of codimension $\leq r$, and the inverse image $Q^{-1}(\mathcal{M}_{n,m})$ is Whitney regular stratified with strata $Q^{-1}(V^{(\ell)}_{\xi,k,[\tau]})$. Moreover,
- If $Q^{-1}(V^{(\ell)}_{\xi,k,[\tau]}) \neq \emptyset$, then codim $Q^{-1}(V^{(\ell)}_{\xi,k,[\tau]}) =$ codim $V^{(\ell)}_{\xi,k,[\tau]}$.

In the case where $r = 1$, it follows that, generically, a 1-parameter family of optimization problems $Q(t)$, $t \in \mathbb{R}$, only intersects codim 0 or codim 1 strata of $\mathcal{M}_{n,m}$. Due to Theorem 2, we are able to give a list of all such strata, together with the corresponding Q-features (cf Tables 4 and 5). It turns out that all underlying $Q(t)$'s – if their feasible sets, respectively KKT-sets are non-empty – fulfill the conditions LICQ and ND (cf. endnote 2), with as the only exception the optimization problems represented by the codim 1 stratum $V^{(3)}_{0,n,[\pm 1]}$, $m = n + 1$. In this latter case it is easily verified that $\mathcal{F}_Q$ consists of a singleton, whereas dim $\mathcal{K}_Q = 1$.

We call x a *generalized critical* (g.c.) point for Q, whenever $x \in \mathcal{F}_Q$, and $Ax + a$ together with the columns of B^T, form a linear dependent set of vectors[9] (in $\mathbb{R}^n$). Note that a feasible x is g.c. point iff either LICQ is violated at x, or LICQ holds and $(x^T, \lambda^T)^T \in \mathcal{K}_Q$ for some λ in $\mathbb{R}^m$.

The above considerations lead to:

Lemma (Two types) *Let $Q(t)$ be a smooth* 1*-parameter family of optimization problems Q. Then, generically, at a g.c. point $Q(t)$ admits one of the following two possibilities*:

Either

- LICQ *and* ND *hold, and in this case*: $m \leq n$ *or*
- LICQ *is violated, and in this case*: $m = n + 1$.

Now, we broaden our scope and consider linear-quadratic optimization problems, say $\tilde{Q}$, with *m equality* and *p inequality* constraints, $m < n$, $p \geq 0$. Each combination

Table 4. Strata of codim 0.

$V^{(\ell)}_{\xi,k,[\tau]}$ of codim 0	Q-features
$m \geq n+1$; $\ell = 1$; $k=n$; ($\xi^0 = 0$)	$\mathcal{F}_Q = \emptyset$
$m=n$; $\ell=3$; $k=m$ ($\xi^0=0$); $\tau=\pm 1$	$\mathcal{K}_Q = \{\bar{y}\}$; LICQ, ND hold
$0 \leq m \leq n$; $\ell=3$; $k=m$; $\xi^0=0$; $\tau=\pm 1$	$\mathcal{K}_Q = \{\bar{y}\}$; LICQ, ND hold
$m=0$; $\ell=6$; ($k=0$); $\xi^0=0$; $\tau=\pm 1$	No constraints; $\mathcal{K}_Q=\{\bar{y}\}$; ND holds

Between parentheses: condition which is implied by the preceding conditions in this row.

of the equality constraints, together with any subset of (active) inequality constraints, yields a sub problem of type Q, giving rise to conditions LICQ and ND. An additional non-degeneracy condition comes up: The non-vanishing of all Lagrange parameters, associated to active inequality constraints (under LICQ) at the g.c. points (of the various sub problems). This condition is referred to as to the strict complementarity condition (SCC). For the sake of completeness, we quote in a rather rough form – a result that we obtained in [11]:

Theorem (Three types) *Let $\widetilde{Q}(t)$ be a smooth* 1*-parameter family of optimization problems $\widetilde{Q}$. Then, generically, $\widetilde{Q}(t)$ admits at a g.c. point, one of following alternatives, which (apart from some additional, technical conditions) may be characterized as:*

Either

- LICQ, ND *and* SCC *hold* (*regular g.c. point*), *or*
- LICQ *and* ND *hold, but* SSC *not, and precisely one of the Lagrange parameters to the active inequality constraints vanishes* (*degenerate g.c. point*), *or*
- LICQ *is violated and* # (*all active constraints*) $= n+1$ (*degenerate g.c. point*).

Moreover, the set of all *regular g.c. points* constitutes a smooth curve in $\mathbb{R}^n$ (*critical curve*) and the degenerate g.c. points are isolated points in the topological closure of this curve. For each regular g.c. point its nature with respect to the corresponding problem $\widetilde{Q}(t)$, (i.e. minimum, maximum. or saddle point) is determined by a *quadruple of indices*. Traversing the topological closure of the critical curve, these quadruples only change when a degenerate g.c. point is passed through (in which case the transitions of these quadruples can be described in terms of characteristics of this degenerate g.c. point).

In [11], we focused on the (rather tough) verification – in the generic case – of the (additional) conditions characterizing the three types of g.c. points in the above theorem, as well as on the relationship with the five (!) types of g.c. points (cf [7,8]) which are possible in the general case of smooth optimization problems with finitely many (in-)equality constraints. Although, the present Theorem 3 and Tables 4 and 5 constitute one of the cornerstones for [11], in this article we merely applied these results, without giving their proofs.[10]

This motivational section is concentrated on 1-parameter families $Q(t)$. However, we feel that our results enjoy also some intrinsic value; possibly they can also be used to develop genericity approaches to other problems that are characterized by

Table 5. Strata of codim 1.

$V^{(\ell)}_{\xi,k,[\tau]}$ of codim 1	Q-features
$m > n+1$; does not occur	– – – –
$m = n+1$; $\ell = 3$; $k = n$; ($\xi^0 = 0$); $\tau = \pm 1$	LICQ violated; $\mathcal{F}_Q = \{\bar{x}\}$, dim $\mathcal{K}_Q = 1$
$m = n$; $\ell = 1$; $k = n-1$; $\xi^0 = 0$	$\mathcal{F}_Q = \emptyset$
$m = n$; $\ell = 3$; $k = n$; ($\xi_0 = 0$); $\tau = 0$	$\mathcal{K}_Q = \{\bar{y}\}$; LICQ, ND hold
$m = n$; $\ell = 5$; $k = n = 1$; ($\xi^0 = 0$)	$\mathcal{K}_Q = \{\bar{y}\}$; LICQ, ND hold
$0 < m < n$; $\ell = 2$; $k = m$; $\xi^0 = 1$	$\mathcal{F}_Q \neq \emptyset$; $\mathcal{K}_Q = \emptyset$; LICQ holds
$0 < m < n$; $\ell = 3$; $k = m$; $\xi^0 = 0$; $\tau = 0$	$\mathcal{K}_Q = \{\bar{y}\}$; LICQ, ND hold
$0 < m < n$; $\ell = 6$; $k = m = 1$; $\xi^0 = 0$; $\tau = \pm 1$	$\mathcal{K}_Q = \{\bar{y}\}$; LICQ, ND hold
$m = 0$; $\ell = 4$; ($k = 0$); $\xi^0 = 1$;	$\mathcal{F}_Q \neq \emptyset$; $\mathcal{K}_Q = \emptyset$
$m = 0$; $n > 1$; $\ell = 6$; ($k = 0$); $\xi^0 = 0$; $\tau = 0$	$\mathcal{K}_Q = \{\bar{y}\}$; ND holds
$m = 0$; $n = 1$; $\ell = 7$; ($k = 0$); $\xi^0 = 0$;	$\mathcal{K}_Q = \{\bar{y}\}$; ND holds

Between parentheses: condition which is implied by the preceding condition in this row.

matrices, structured according to $[M \vdots c]$. For example, it is easily shown (by analyzing the strata $V^{(\ell)}_{\xi,k,[\tau]}$ of codim 2) that 2-parameter families $Q(t)$, with $0 < m < n$, generically admit only g.c. points for which either LICQ and ND hold, or LICQ holds but ND not. Another example, not related to Parametric Optimization, is the following: The so-called *extraneous singularities* for a gradient Newton flow are characterized by a rank condition on certain matrices of the form $[M \vdots c]$, with $B = \mathrm{O}$, $b = 0$, cf. [10, Ch. 9]. This leads to a generic description of the sets of extraneous singularities (at least for lower dimensions).

4. Proof of Theorem 1 (Normal forms)

We consider block matrices U of the following type:

$$U = \begin{bmatrix} V & K \\ O & W \end{bmatrix}, \quad \text{where}$$

V and W are regular matrices of dimensions $n \times n$ and $m \times m$ respectively, K is an arbitrary $n \times m$-matrix, and O the $m \times n$-null matrix.

Apparently, the set $\mathcal{G}$ of all matrices U is a *group* (w.r.t. to matrix multiplication).

For all $U \in \mathcal{G}$ and all $\eta \in \mathbb{R}\backslash\{0\}$, we introduce the map $\Psi_{U,\eta} : [M \vdots c] \mapsto [UMU^T \vdots \eta Uc]$.

By direct verification one finds:

$$[UMU^T \vdots \eta Uc] = \begin{bmatrix} A' & B'^T & \vdots & a' \\ B' & O & \vdots & b' \end{bmatrix}, \quad \text{with} \quad \begin{aligned} A' &= VAV^T + KBV^T + VB^TK^T \\ B' &= WBV^T; \\ O &= m \times m - \text{null matrix} \\ a' &= \eta(Va + Kb);\ b' = \eta Wb \end{aligned} \tag{1}$$

By (1) we have: (recall that $\mathcal{M}_{n,m} \equiv \mathbb{R}^N$)

LEMMA 1

- $\Psi_{U,\eta}$ *is a diffeomorphism of* $\mathcal{M}_{n,m}$ *onto itself, i.e.* $\Psi_{U,\eta} \in \mathrm{Diff}(\mathcal{M}_{n,m})$;
- $(U,\eta) \mapsto \Psi_{U,\eta}$ *is a morphism* $\mathcal{G} \times \mathbb{R}\backslash\{0\} \mapsto \mathrm{Diff}(\mathcal{M}_{n,m})$.

The next lemma is crucial:

Lemma 2 *The map $\Psi_{U,\eta}$ respects all Q-features mentioned in Section* 2.

Proof We extend the notations as introduced in (1). So, $\Psi_{U,\eta}([M \vdots c]) = [UMU^T \vdots \eta Uc] = [M' \vdots c']$, the optimization problem represented by $[M' \vdots c']$ is denoted Q', and matrix S' has as columns a basis for ker (B'), ...

Moreover, V^{-T} stands for $(V^T)^{-1} = (V^{-1})^T$, etc. Then, from (1) it follows:

$$x \in \mathcal{F}_Q \text{ iff } \eta V^{-T}x \in \mathcal{F}_{Q'} \quad \text{and} \quad y \in K_Q \text{ iff } \eta U^{-T}y \in \mathcal{K}_{Q'}. \tag{2}$$

Hence, the emptiness/non-emptiness of $\mathcal{F}_Q$ and $\mathcal{K}_Q$ are respected by $\Psi_{U,\eta}$.

For $y' \in \mathcal{K}_{Q'}$, from (1) and (2) it follows: $(y')^T c' = (\eta U^{-T}y)^T \eta Uc = \eta^2 y^T c$, and thus $\tau = \tau'$.

By (1), we also have: $z \in \ker(B)$ iff $V^{-T}z \in \ker(B')$. So, we may choose $S' = V^{-T}S$. Since $BS = 0$ and using (1) we find:
$\mathrm{In}(A'_{|\ker B'}) = \mathrm{In}(S'^T A' S') = \mathrm{In}(S^T AS + S^T V^{-1} K[BS] + [S^T B^T] K^T V^{-T} S = \mathrm{In}(S^T AS) = \mathrm{In}(A_{|\ker\ B})$.

The properties '$b = 0$ iff $b' = 0$', '$\mathrm{rank}(B) = \mathrm{rank}(B')$', and '$a = b = 0$ iff $a' = b' = 0$' are trivial consequences of (1). Finally, if $b = b' = 0$ we find by (1):
'z fulfils $B^T z + a = 0$ iff $z' = \eta W^{-T} z$ fulfils $B'^T z' + a' = 0$'. Hence, $\mathrm{rank}[B^T \vdots a] = \mathrm{rank}[B'^T \vdots a']$. ■

Now we are going to prove the first proposition of Theorem 1:

Lemma 3 (Normal forms for M) *Let $[M \vdots c]$ be an arbitrary, but fixed matrix in $V^{(\ell)}_{\xi,k,[\tau]}$. Then, a matrix $U \in \mathcal{G}$ exists such that*:

$$UMU^T = \overline{N},$$

where $\overline{N}$ is the normal form as described in the statement of Theorem 1.

Proof It is well-known that there exist regular matrices V_1, W_1 such that:

$$W_1 B V_1^T = \begin{bmatrix} I_k & O \\ O & O \end{bmatrix} (= \overline{B})$$

Now we choose a matrix, say U_1, from $\mathcal{G}$ by $V := V_1$, $W := W_1$ and $K := \mathrm{O}$. Then:

$$U_1 M U_1^T (= M') = \begin{bmatrix} A' & \overline{B}^T \\ \overline{B} & O \end{bmatrix}, \text{ with } A' = \begin{bmatrix} A'_{k,k} & A'_{k,n-k} \\ A'_{n-k,k} & A'_{n-k,n-k} \end{bmatrix} \text{(block matrix)}$$

Here, and in the sequel, a pair of subscripts of a block indicates its dimensions. For example, $A'_{k,n-k}$ is a $k \times (n-k)$-matrix; note that, by construction, the matrices $A'_{k,k}$ and $A'_{n-k,n-k}$ are symmetric, and $A'_{n-k,k} = (A'_{k,\ n-k})^T$.

Next, we choose a matrix, say U_2, from $\mathcal{G}$ by $V := I_n$, $W := I_m$, and

$$K := \begin{bmatrix} -\frac{1}{2} A'_{k,k} & O \\ -A'_{n-k,k} & O \end{bmatrix}.$$

By a simple calculation, we derive from (1):

$$U_2M'U_2^T = \begin{bmatrix} A'' & \overline{B}^T \\ \overline{B} & O \end{bmatrix}, \text{ with } A'' = \begin{bmatrix} O_{k,k} & O_{k,n-k} \\ O_{n-k,k} & A''_{n-k,n-k} \end{bmatrix}$$

Successive application of Lemma 2 yields: $\mathrm{In}(A_{|\ker B}) = \mathrm{In}(A''_{|\ker \overline{B}}) = \mathrm{In}(A''_{n-k,n-k})$, where the last equality is obtained by direct verification; note that $A''_{n-k,n-k}$ is symmetric.

Now, a regular $(n-k)\times(n-k)$-matrix Z exists such that $Z(A''_{n-k,n-k})Z^T = J_\xi$. Finally, we choose $U_3 \in \mathcal{G}$, by

$$V := \begin{bmatrix} I_k & O \\ O & Z \end{bmatrix}, \quad W := I_m, \quad \text{and } K := O.$$

Now, we define $U := U_3U_2U_1 (\in \mathcal{G})$. Then, due to the construction of U, and by Lemma 2, we have:

$$\Psi_{U,\eta}([M \,\vdots\, c]) = [\overline{N} \,\vdots\, \eta U c] \in V^{(\ell)}_{\xi,k,[\tau]},$$

in particular: $UMU^T = \overline{N}$. ■

Corollary 1 (Empty strata) *The list of empty strata in Table 1 follows directly from the special structure of $\overline{N}$.*

The following relationship between the inertia's of M and $A_{|\ker B}$ is well-known from literature, cf [10].

Corollary 2 (Inertia theorem) $\mathit{In}(M) = \mathit{In}(A_{|\ker B}) + (k, k, m-k)$.

Proof Let $\{e_1, \ldots, e_{n+m}\}$ be the standard basis for $\mathbb{R}^{n+m}$. By inspection, one finds that $\overline{N}$ admits the eigenvectors: $(e_1 \pm e_{n+1}), \ldots, (e_k \pm e_{n+k})$ [eigenvalues ± 1 according to $\pm$], $e_{k+1}, \ldots, e_n$ [eigenvalues $+1$, -1 or 0, distributed according to $\mathrm{In}(J_\xi)$], and $e_{n+k+1}, \ldots, e_{n+m}$ [eigenvalues 0]. These vectors constitute a basis for $\mathbb{R}^{n+m}$. Hence, the assertion follows from Lemmas 2, 3. (Note that $\overline{N}_{|\ker \overline{B}} = J_\xi$). ■

Before turning over to the remaining part of the proof of Theorem 1, we need some technical lemmas, concerning classes of matrices $U \in \mathcal{G}$ such that:

$$\Psi_{U,\eta}[\overline{N} \,\vdots\, c] = [\overline{N} \,\vdots\, \eta U c],$$

and, moreover, by appropriate choices of U, at least two of the sub vectors a_1, a_2, b_1, b_2 of c remain invariant under the mapping $c \mapsto Uc$.

Lemma 4 *Let $U \in \mathcal{G}$ be given by*:

$$V := \begin{bmatrix} P & O \\ O & I_{n-k} \end{bmatrix}, \quad W := \begin{bmatrix} P^{-T} & O \\ O & I_{m-k} \end{bmatrix},$$

with P a regular $k \times k$-matrix, and $K := \mathrm{O}$.

Then,

$$\Psi_{U,\eta}[\overline{N} \vdots c] = [\overline{N} \vdots c'], \quad \textit{with } c' = \eta((Pa_1)^T, a_2^T, (P^{-T}b_1)^T, b_2^T)^T.$$

Proof By inspection. (Note that the lemma also holds if $k=0$ [$V=I_n$; $W=I_m$; P, a_1 and b_1 non-existent], if $k=m$; [$W=P^{-T}$; b_2 non-existent] and if $k=n$ [$V=P$; J_ξ and a_2 non-existent]). ■

LEMMA 5 *Let $U \in \mathcal{G}$ be given by:*

$$V = \begin{bmatrix} I_k & P \\ O & I_{n-k} \end{bmatrix}, \quad K = \begin{bmatrix} L & O \\ R & O \end{bmatrix}, \quad W = I_m,$$

where P is a $k \times (n-k)$-matrix, L a $k \times k$-matrix, and R an $(n-k) \times k$-matrix. Then:

$$\bullet \quad U\overline{N}U^T = \overline{N} \quad \textit{iff} \quad \begin{cases} L^T + L + PJ_\xi P^T & = O \\ R + J_\xi P^T & = O \end{cases}$$

$$\bullet \quad \text{if } U\overline{N}U^T = \overline{N}, \textit{ then}: \Psi_{U,\eta}[\overline{N} \vdots c] = \Big[\overline{N} \vdots \eta((a_1 + Pa_2 + Lb_1)^T, (a_2 + Rb_1)^T, b_1^T, b_2^T)^T\Big]. \tag{3}$$

Proof By inspection (Note that due to condition (3), matrix R and to some extent also L, is determined by P, whereas P itself may be arbitrarily chosen). ■

LEMMA 6 *Given: H a regular $(n-k) \times (n-k)$-matrix, $k \geq 0$, and $U \in \mathcal{G}$ with:*

$$V = \begin{bmatrix} I_k & O \\ O & H \end{bmatrix}, \quad K = O, \ W = I_m.$$

Then,

$$\bullet \quad U\overline{N}U^T = \overline{N} \quad \textit{iff} \quad HJ_\xi H^T = J_\xi$$

$$\bullet \quad \textit{if } HJ_\xi H^T = J_\xi \textit{ then } \Psi_{U,\eta}[\overline{N} \vdots c] = \Big[\overline{N} \vdots \eta(a_1^T, (Ha_2)^T, b_1^T, b_2^T)^T\Big].$$

(*If $k=0$, then V reduces to H, and c to $(a_2^T, b_2^T)^T$; if $k=m$, then b_2 is non-existent*).

Proof By inspection. ■

Now, we are ready to present the remaining part of the proof of Theorem 1:

Proof of Theorem 1 Due to Lemma 3, we may start with a matrix $[\overline{N} \vdots c] \in V^{(\ell)}_{\xi,k,[\tau]}$ and seek pairs $(U, \eta) \in \mathcal{G} \times \mathbb{R}\backslash\{0\}$ such that

$$U\overline{N}U^T = \overline{N}, \quad \eta Uc = \overline{n}.$$

Here, the 'normal vectors' $\overline{n}$ depend on the value of ℓ see (the statement of) Theorem 1. In order to keep our treatment relatively simple, we assume:

$$0 < k < m < n.$$

This assumption does not affect our final results: Some cases, which are excluded correspond to empty strata (Table 1), whereas in the remaining situations a simple adaptation to the general line of reasoning is possible.

We distinguish between the various values of ℓ:
$[\overline{N} \vdots c] \in V^{(1)}_{\xi,k,[\tau]}$:
Taking into account the structure of $\overline{N}$, we find $b_2 \neq 0$. We extend b_2 to a basis for $\mathbb{R}^{m-k}$. Consider the regular $(m-k)\times(m-k)$-matrix, say P, with this basis as columns. Then $Pe = b_2$, where $e = (1, 0, \ldots, 0)^T \in \mathbb{R}^{m-k}$, and thus $P^{-1}b_2 = e$. Put:

$$U_1 := \begin{bmatrix} I_{n+k} & D \\ O & P^{-1} \end{bmatrix}, \text{ with } D = \begin{bmatrix} -a & \vdots & O_{n,m-k-1} \\ -b_1 & \vdots & O_{k,m-k-1} \end{bmatrix} P^{-1};$$

$$\text{if } m-k = 1, \text{ then: } D = \begin{bmatrix} -a \\ -b_1 \end{bmatrix} P^{-1}.$$

It is easily verified that $U_1 \in \mathcal{G}$ [with $V := I_n$, the first n rows of D constituting K, etc.].

Apparently, we have:

$$U_1 \overline{N} U_1^T = \overline{N},\ \ U_1 c = \overline{n},\ \text{ where } \overline{n} \text{ is specified as in row } \ell = 1 \text{ of Table 2.}$$

Redefining $U := U_1(\in \mathcal{G})$ yields, in view of Lemma 1: $\Psi_{U,1}[\overline{N} \vdots c] = [\overline{N} \vdots \overline{n}] \in V^{(1)}_{\xi,k,[\tau]}$.
$[\overline{N} \vdots c] \in V^{(2)}_{\xi,k,[\tau]}$: Due to the structure of $\overline{N}$ and the characteristics of $V^{(2)}_{\xi,k,[\tau]}$ we have:

$$b_1 \neq 0,\ \ b_2 = 0,\ \ \xi^0 > 0,\ \text{ and } a_{\xi^0} \neq 0.$$

The vector $b_1 (\neq 0)$ can be completed to a basis for $\mathbb{R}^k$. Let P be the (regular) $k \times k$-matrix with this basis as rows, and put $e = (1, 0, \ldots, 0)^T \in \mathbb{R}^k$. Then, $P^T e = b_1$ and thus $P^{-T} b_1 = e$. With respect to this P, we choose $U_1 \in \mathcal{G}$ according to Lemma 4. We find:

$$U_1 \overline{N} U_1^T = \overline{N};\ \ U_1 c = ((Pa_1)^T, a_2^T, e^T, 0^T)^T.$$

Now, we are going to apply Lemma 6 with

$$H = \begin{bmatrix} I_r & L \\ O & E \end{bmatrix}, r = \xi^+ + \xi^-; \qquad L = \begin{bmatrix} -a_{\xi^+} & \vdots & \\ & \vdots & O_{r,\xi^0-1} \\ -a_{\xi^-} & \vdots & \end{bmatrix} E,$$

where E is a regular $\xi^0 \times \xi^0$-matrix with $Ea_{\xi^0} = e \in \mathbb{R}^{\xi^0}$. (Note: such E exists because $a_{\xi^0} \neq 0$; compare the construction of matrix P in the preceding case $\ell = 1$.) For this choice of H, we have $HJ_\xi\ H^T = J_\xi$. So, Lemma 6 yields a matrix $U_2 \in \mathcal{G}$ with $U_2 \overline{N} U_2^T = \overline{N}$, and

$$U_2((Pa_1)^T, a_2^T, e^T, 0^T)^T = ((Pa_1)^T, \overline{n}_2^T, e^T, 0^T)^T, \text{with } \overline{n}_2 = \begin{bmatrix} 0 \\ 0 \\ e \end{bmatrix} \in \mathbb{R}^{\xi^+} \times \mathbb{R}^{\xi^-} \times \mathbb{R}^{\xi^0}.$$

Next, we consider a $k \times (n-k)$-matrix, say F, with as first $(\xi^+ + \xi^-)$ columns 0, as $(\xi^+ + \xi^- + 1)$th-column the vector $-Pa_1$ and all (possible) other columns arbitrarily chosen; note that, since $\xi^0 \geq 1$, such matrices do exist. Apparently, we have $J_\xi F^T = O$.

We apply Lemma 5, with $P = F$, $R = O$ and $L = O$, and obtain a matrix $U_3 \in \mathcal{G}$ for which:

$$U_3 \overline{N} U_3^T = \overline{N}, \quad U_3((Pa_1)^T, \overline{n}_2^T, e^T, 0^T)^T = (0^T, \overline{n}_2^T, e^T, 0^T)^T.$$

Finally, we put $U := U_3 U_2 U_1$. Due to Lemma 1, we find

$$\Psi_{U,1}[\overline{N} \vdots c] = [\overline{N} \vdots \overline{n}] \in V^{(2)}_{\xi,k,[\tau]}.$$

$[\overline{N} \vdots c] \in V^{(3)}_{\xi,k,[\tau]}$:
Due to the structure of $\overline{N}$ and the characteristics of $V^{(3)}_{\xi,k,[\tau]}$, we have:

$$b_1 \neq 0, \; b_2 = 0.$$

As in the preceding case $(\ell = 2)$, the condition $b_1 \neq 0$ yields a matrix $U_1 \in \mathcal{G}$ such that:

$$U_1 \overline{N} U_1^T = \overline{N}; \; U_1 c = ((Pa_1)^T, a_2^T, e^T, 0^T)^T.$$

We proceed by proving:

$$J_\xi^2 a_2 = a_2.$$

First, we note: $a_2 = (a_{\xi^+}^T, a_{\xi^-}^T, a_{\xi^0}^T)^T$ with $a_{\xi^0} = 0$ [if $\xi^0 = 0$, this is trivial ('a_{ξ^0} empty'); if $\xi^0 > 0$, it follows from 'rank$[\overline{N} \vdots c] = $ rank$(\overline{N})$']. Moreover, it is easily verified that:

$$J_\xi^2 = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}, r = \xi^+ + \xi^-, \; [\text{if } \xi^0 = 0 \text{ then } J_\xi^2 = I_{n-k}, \; \text{if } \xi^0 = n-k, \text{then } a_2 = 0].$$

So, we may conclude that $J_\xi^2 a_2 = a_2$.

Now, we apply Lemma 5: Let P be a $k \times (n-k)$-matrix with as 1st row $[J_\xi \, a_2]^T$, $R := -J_\xi P^T$, and $L := -(1/2) P J_\xi P^T$. This yields a matrix, say $U_4 \in \mathcal{G}$, such that: $U_4 \overline{N} U_4^T = \overline{N}$, and

$$U_4((Pa_1)^T, a_2^T, e^T, 0^T)^T = ((a_1')^T, 0^T, e^T, 0^T)^T \in \mathbb{R}^k \times \mathbb{R}^{n-k} \times \mathbb{R}^k \times \mathbb{R}^{m-k}$$

(because $a_2 + \text{Re} = a_2 - J_\xi P^T e = a_2 - J_\xi^2 a_2 = a_2 - a_2 = 0$).

Put $c' := ((a'_1)^T, \; 0^T, e^T, 0^T)^T$. Apparently, vector $y := (-e^T, 0^T, -(a'_1)^T, 0^T)^T$ belongs to the KKT-set of the optimization problem represented by $[\overline{N} \vdots c']$, and $y^T c' = -2e^T a_1'$. If $q = -e^T a_1'$, then – due to Lemma 2 – we have

$$\tau (= \text{sign}[\overline{N} \vdots c']) = \text{sign}(q).$$

We distinguish between the cases $q \neq 0$ and $q = 0$.

If $q \neq 0$:

Apply Lemma 4 again, but now with $P := (1/\sqrt{|q|})[e \vdots P^\sim]^T$, where $P^\sim$ is a $k \times (k-1)$-matrix with as columns a basis for the orthogonal complement of a'_1, and put $\eta := \eta_5 = 1/\sqrt{|q|}$. Note that P is regular, then:

$$\eta P a_1' = \frac{-q}{|q|} e = -\tau e; \quad \eta P^{-T} e = e.$$

With these choices of P and η [if $k=1$, put $P := [1/\sqrt{|q|}]$ and $q = -a'_1$], Lemma 4 yields a pair $(U_5, \eta_5) \in \mathcal{G} \times \mathbb{R} \backslash \{0\}$ such that:

$$U_5 \overline{N} U_5^T = \overline{N}, \quad \text{and } \eta_5 U_5(c') = \overline{n},$$

where $\overline{n}$ is the normal form as indicated in Table 2, $\ell = 3, \tau = \pm 1$.

If $q = 0$:
From the very definition of q it follows that the 1st component of a'_1 vanishes.
So, if $k = 1$, we are done.
In case $k > 1$, we write $a_1' = (0, d^T)^T$, with $d \in \mathbb{R}^{k-1}$. Then Lemma 5, with

$$P := O_{k,n-k},\ R := O_{n-k,k}, \text{ and } L := \begin{bmatrix} 0 & \vdots & d^T \\ -d & \vdots & O_{k-1,k-1} \end{bmatrix} \quad (\text{hence, } L + L^T = O),$$

yields a matrix in $\mathcal{G}$, again denoted by U_5, such that

$$U_5 \overline{N} U_5^T = \overline{N}, \quad \text{and} \quad U_5(c') = \overline{n}$$

where $\overline{n}$ is the normal form as in Table 2, case $\ell = 3, \tau = 0$.

Altogether, if we redefine $U := U_5 U_4 U_1$ and put $\eta := \eta_5$ (if $q \neq 0$) or $\eta := 1$ (if $q = 0$), then the maps $\Psi_{U,\eta}$ meet our objectives.

$[\overline{N} \vdots c] \in V^{(4)}_{\xi,k,[\tau]}$:

Due to the structure of $\overline{N}$ and the characteristics of $V^{(4)}_{\xi,k,[\tau]}$, we have:

$$b = 0, \xi^0 > 0, \text{and } a_{\xi^0} \neq 0.$$

As in the case $\ell = 2$, the conditions '$\xi^0 > 0$', and '$a_{\xi^0} \neq 0$' give rise (by Lemma 6) to a matrix in $\mathcal{G}$, again denoted by U_2, such that

$$U_2 \overline{N} U_2^T = \overline{N}, \text{ and } U_2 c = (a_1', \overline{n}_2^T, 0^T, 0^T)^T,$$
$$\text{with } \overline{n}_2 = \begin{bmatrix} 0 \\ 0 \\ e \end{bmatrix} \in \mathbb{R}^{\xi^+} \times \mathbb{R}^{\xi^-} \times \mathbb{R}^{\xi^0}.$$

Next, we apply Lemma 5 to $[\overline{N} \vdots U_2 c]$, by choosing for P a $k \times (n-k)$-matrix with $-a_1'$ as its $(r+1)$th column, and all other columns arbitrary,[11] $R := -J_\xi P^T$, and $L := -(1/2) P J_\xi P^T$. We obtain a matrix, say $U_5 \in \mathcal{G}$, such that

$$U_5 \overline{N} U_5^T = \overline{N}, \text{ and } U_5(U_2 c) = \overline{n},$$

where $\overline{n}$ is the normal form as in Table 2, $\ell = 4$.
Redefining $U := U_5 U_2$, we find: $\Psi_{U,1}$ is the map we are looking for.

$[\overline{N} \vdots c] \in V^{(5)}_{\xi,k,[\tau]}$:

Due to the structure of $\overline{N}$ and the characteristics of $V^{(5)}_{\xi,k,[\tau]}$ we have:

$$a_1 \neq 0, \text{ and } a_2 = b_1 = b_2 = 0.$$

We extend a_1 to a basis for $\mathbb{R}^k$. Let P^{-1} be the $k \times k$-matrix with as columns this basis. Then, $P^{-1}e = a_1$, and thus $Pa_1 = e$. Now, Lemma 4, with this P, yields a matrix $U \in \mathcal{G}$, such that

$$\Psi_{U,1}[\overline{N} \vdots c] = [\overline{N} \vdots \overline{n}],$$

where $\overline{n}$ is as indicated in Table 2, $\ell = 5$.

$[\overline{N} \vdots c] \in V^{(6)}_{\xi,k,[\tau]}$:

Due to the structure of $\overline{N}$ and the characteristics of $V^{(6)}_{\xi,k,[\tau]}$, we have:

$$b_1 = b_2 = 0, \text{ and } a_2 \neq 0.$$

Moreover, the condition 'rank$[\overline{N} \vdots c] =$ rank$[\overline{N}]$' yields:[12]

$$\xi^0 < n - k \text{ and, if } \xi^0 > 0, \text{we have } a_{\xi^0} = 0.$$

We are going to apply Lemma 5. From $a_2 \neq 0$, it follows that at least one component of a_2, say the jth component, has a value $q \neq 0$. Choose for P a $k \times (n-k)$-matrix with as its jth column $-(1/q)a_1$, and all other columns zero. Moreover we put $R := -J_\xi P^T$, and $L := -(1/2)PJ_\xi P^T$. With these choices we obtain, by Lemma 5, a matrix $U_6 \in \mathcal{G}$, such that

$$U_6\overline{N}U_6^T = \overline{N}, \text{ and } c'(:= U_6 c) = (0^T, a_2^T, 0^T, 0^T)^T,$$
$$a_2 = \begin{bmatrix} a_{\xi^+} \\ a_{\xi^-} \\ 0 \end{bmatrix} \in \mathbb{R}^{\xi^+} \times \mathbb{R}^{\xi^-} \times \mathbb{R}^{\xi^0}.$$

Apparently, the vector $y = (0^T, y_2^T, 0^T, 0^T)^T$, with $y_2^T = (-a_{\xi^+}^T, a_{\xi^-}^T, 0^T)^T$, belongs to the KKT-set of the optimization problem represented by $[\overline{N} \vdots c']$. Hence,

$$\tau(= \text{sign}([\overline{N} \vdots c']) = \text{sign}(-\|a_{\xi^+}\|^2 + \|a_{\xi^-}\|^2),$$

where $\|\cdot\|$, stands for Euclidean norm.[13]

Write $\alpha_+ = \|a_{\xi^+}\|$ (if $\xi^+ = 0$, then $\alpha_+ = 0$), $\alpha_- = \|a_{\xi^-}\|$ (if $\xi^- = 0$, then $\alpha_- = 0$), and note: $\alpha_+ + \alpha_- > 0$.

In order to apply Lemma 6, let H be a block matrix of the following structure:

$$H := \begin{bmatrix} H_{\xi^+} & O & O \\ O & H_{\xi^-} & O \\ O & O & I_{\xi^0} \end{bmatrix}$$

where $H.$ stands for *orthogonal* matrices, with dimensions as indicated in the subscript, and O for zero matrices with appropriate dimensions; some of these blocks may be empty, e.g. if $\xi^+ = 0$, or $\xi^- = 0$. Note that H is *regular*. By the special form of J_ξ we have:

$$HJ_\xi H^T = J_\xi.$$

We make suitable choices for H_{ξ^+} and H_{ξ^-}:
If $a_{\xi^+} \neq 0$, we extent $(1/\alpha_+)a_{\xi^+}$ to an orthonormal basis of $\mathbb{R}^{\xi^+}$, with this vector in first position. Now, H_{ξ^+} is the matrix with this basis as rows. If $a_{\xi^+} = 0$, we put $H_{\xi^+} = I_{\xi^+}$. We obviously have: $H_{\xi^+} a_{\xi^+} = \alpha_+ e$, with e in $\mathbb{R}^{\xi^+}$.

With respect to a_{ξ^-}, matrix H_{ξ^-} is defined similarly, and $H_{\xi^-} a_{\xi^-} = \alpha_- e$, with e in $\mathbb{R}^{\xi^-}$.

With this choice of H, Lemma 6 gives rise to a matrix $U_7 \in \mathcal{G}$, such that:

$$U_7 \overline{N} U_7^T = \overline{N}, \text{ and } U_7 c' = (0^T, a_2'^T, 0^T, 0^T)^T,$$

$$a_2' = \begin{bmatrix} \alpha_+ e \\ \alpha_- e \\ 0 \end{bmatrix} \in \mathbb{R}^{\xi^+} \times \mathbb{R}^{\xi^-} \times \mathbb{R}^{\xi^0}.$$

We distinguish between the cases $\tau = 0$, $\tau = 1$, and $\tau = -1$.

Case $\tau = 0$ In this case we have $(\alpha_+)^2 = (\alpha_-)^2$, and thus $\alpha_+ = \alpha_- > 0$ (use $\alpha_+ + \alpha_- > 0$, and $\alpha_+, \alpha_- \geq 0$). So, if we choose $\eta := (1/\alpha_+)$, and redefine $U := U_7 U_6$, the desired map is $\Psi_{U,\eta}$.

Case $\tau = -1$ In this case we have $\alpha_+ > \alpha_-$, and thus $\alpha_+ > 0$.

If $\alpha_- = 0$, we may proceed as in Case $\tau = 0$; with $\eta := (1/\alpha_+)$, we find the normal form described in Table 2, $\ell = 6$, $\tau = -1$.

If $\alpha_+ > \alpha_- > 0$, we apply Lemma 6 once again, but now with a matrix H of the form:

$$H := \begin{bmatrix} \cosh\varphi & 0^T & \sinh\varphi & 0^T \\ 0 & I_{(\xi^+)-1} & 0 & O \\ \sinh\varphi & 0^T & \cosh\varphi & 0^T \\ O & O & 0 & I_{(\xi^-)+(\xi^0)-1} \end{bmatrix}$$

where $\cosh(\cdot)$ stands for hyperbolic cosine, etc. and $0(O)$ are zero vectors (matrices) of appropriate dimensions.

It is easily verified that H is regular, and $HJ_\xi H^{\mathrm{T}} = J_\xi$. Thus, Lemma 6 yields a $U_8 \in \mathcal{G}$, such that

$$U_8 \overline{N} U_8^T = \overline{N}, \text{ and } U_8 U_7 c' = (0^T, a_2''^T, 0^T, 0^T)^T,$$
$$a_2'' = \begin{bmatrix} (\alpha_+ \cosh\varphi + \alpha_- \sinh\varphi)e \\ (\alpha_+ \sinh\varphi + \alpha_- \cosh\varphi)e \\ 0 \end{bmatrix}$$

We choose[14] a φ, say φ_0, with $(\alpha_+ \sinh\varphi_0 + \alpha_- \cosh\varphi_0) = 0$, and thus: $a_2'' = ((\alpha_+^2 - \alpha_-^2)^{(1/2)} e^T, 0^T, 0^T)^T$. Finally, by choosing $\eta := 1/(\alpha_+^2 - \alpha_-^2)^{(1/2)}$ and $U := U_8 U_7 U_6$, we find the desired normal form.

Case $\tau = +1$:
Similar to the preceding Case $\tau = -1$.
$[\overline{N} \vdots c] \in V^{(7)}_{\xi,k,[\tau]}$:
Trivial. ∎

5. Proof of Theorem 2 (Manifolds)

5.1. *Smooth manifold; defining system*

The subset $\mathcal{M}$ of $\mathcal{M}_{n,m}(\equiv \mathbb{R}^N)$ is called a smooth manifold of dimension q, if each $Z \in \mathcal{M}$ has an open[15] neighbourhood in $\mathcal{M}$, which is diffeomorphic to an open subset of $\mathbb{R}^q$.

Now, let $V^{(\ell)}_{\xi,k,[\tau]}$ be one of the non-void strata of $\mathcal{M}_{n,m}$, and $Z = [M \vdots c]$ a matrix in this stratum. Then, due to Lemmas 1, 2 and Theorem 1, an open neighbourhood in $V^{(\ell)}_{\xi,k,[\tau]}$ around Z, may be diffeomorpically mapped (by suitably chosen $\Psi_{U,\eta}$) onto an open $V^{(\ell)}_{\xi,k,[\tau]}$–neighbourhood, say $\mathcal{M}_0$, of the normal form $[\overline{N} \vdots \overline{n}]$ in $V^{(\ell)}_{\xi,k,[\tau]}$. So, we are done if we are able to show the existence of such $\mathcal{M}_0$, which is diffeomorphic to an open set in $\mathbb{R}^q$. Due to the Implicit Function Theorem, the latter assertion is equivalent to the existence of a so called *defining system around* $[\overline{N} \vdots \overline{n}]$, i.e. a set of smooth functions, say h_i, $i = 1, \ldots, N - q(=$ 'codimension'), on an $\mathbb{R}^N$-open neighbourhood $\mathcal{M}_1$ of $[\overline{N} \vdots \overline{n}]$ such that:

- $V^{(\ell)}_{\xi,k,[\tau]} \cap \mathcal{M}_1 = \{Z \in \mathcal{M}_1 \mid h_1(Z) = 0, \ldots, h_{N-q}(Z) = 0\}$
- On $V^{(\ell)}_{\xi,k,[\tau]} \cap \mathcal{M}_1$ the gradients of h_i, $i = 1, \ldots, N - q$, are linearly independent.

In the sequel, we shall need the following technical lemma:

LEMMA 7 *Let H be a matrix with the following 2×2-block structure*:

$$H = \begin{bmatrix} H_1 & H_2 \\ H_3 & H_4 \end{bmatrix},$$

where H_1 is a regular $p \times p$-matrix. Then:

$$\mathrm{Rank}(H) = p \quad \textit{iff} \quad H_4 = H_3(H_1)^{-1}H_2.$$

Proof See e.g. [10]. ∎

Now we are ready to present:

Proof of Theorem 2 From the observations above, we know that all we have to do, is to find defining systems for $V^{(\ell)}_{\xi,k,[\tau]}$ around its normal form $[\overline{N} \vdots \overline{n}]$. Later on, we shall distinguish between the various values for ℓ. However, since the conditions $\mathrm{rank}(B) = k$, and $\mathrm{In}(A_{|\ker B}) = \xi$ do not depend on a specific value of ℓ, we start with deriving equations which characterize these two conditions. Let A and B be the sub matrices of M as introduced in Section 1. We provide these matrices with the following block structures:

$$A = \begin{bmatrix} A_{11} & A_{21}^T & A_{31}^T \\ A_{21} & A_{22} & A_{32}^T \\ A_{31} & A_{32} & A_{33} \end{bmatrix}; \qquad B = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix}$$

Here,[16] the symmetric matrices A_{11}, A_{22}, and A_{33} have dimensions $(k \times k)$, $(r \times r)$ with $r = \xi^+ + \xi^-$, and $(\xi^0 \times \xi^0)$ respectively (inducing the dimensions of the other blocks in A); B_{11} is a $k \times k$-matrix, B_{21} a $(m\text{-}k) \times k$-matrix and B_{12} is of dimension $(k \times r)$, etc.

Now, we choose M close to the normal form $\overline{N}$, i.e.

$$A_{22} \approx \begin{bmatrix} I_{\xi^+} & O \\ O & -I_{\xi^-} \end{bmatrix}, \quad B_{11} \approx I_k, \text{ and all other blocks of } A \text{ and } B \approx O$$

In particular, we have: A_{22}, and B_{11} are regular.

Due to Lemma 7, we have: (for any B close enough to the normal form $\overline{B}$)

$$\operatorname{rank}(B) = k \text{ iff } [B_{22}B_{23}] = B_{21}B_{11}^{-1}[B_{12}B_{13}], \text{ i.e. } \begin{matrix} B_{22} = B_{21}B_{11}^{-1}B_{12} \\ B_{23} = B_{21}B_{11}^{-1}B_{13} \end{matrix} \tag{4}$$

Next, we consider the $n \times (n-k)$-matrix S, with the following block structure:

$$S = \begin{bmatrix} -B_{11}^{-1}B_{12} & -B_{11}^{-1}B_{13} \\ I_r & O \\ O & I_{n-k-r} \end{bmatrix}.$$

Obviously, we have: $\operatorname{rank}(S) = n - k$ and –under condition (4)– $BS = O$. Thus the columns of S form a basis for $\ker(B)$. By straightforward calculation:

$$S^T AS = \begin{bmatrix} D_{11} & D_{21}^T \\ D_{21} & D_{22} \end{bmatrix},$$

with

$$\left.\begin{aligned} D_{11} &= A_{22} + B_{12}^T B_{11}^{-T} A_{11} B_{11}^{-1} B_{12} - B_{12}^T B_{11}^{-T} A_{21}^T - A_{21} B_{11}^{-1} B_{12} \\ D_{21} &= A_{32} + B_{13}^T B_{11}^{-T} A_{11} B_{11}^{-1} B_{12} - B_{13}^T B_{11}^{-T} A_{21}^T - A_{31} B_{11}^{-1} B_{12} \\ D_{22} &= A_{33} + B_{13}^T B_{11}^{-T} A_{11} B_{11}^{-1} B_{13} - B_{13}^T B_{11}^{-T} A_{31}^T - A_{31} B_{11}^{-1} B_{13} \end{aligned}\right\} (\dagger)$$

So, by choosing M close enough to $\overline{N}$, we have: $D_{11} \approx A_{22}$, thus D_{11} is regular, and D_{21}, $D_{22} \approx$ O. From Lemma 7, it follows:

$$\operatorname{rank}(S^T AS) = r (= \xi^+ + \xi^-) \text{ iff } D_{22} = D_{21} D_{11}^{-1} D_{21}^T .$$

The latter condition, combined with the above expressions (†) yields:

$$A_{33} = -B_{13}^T B_{11}^{-T} A_{11} B_{11}^{-1} B_{13} + B_{13}^T B_{11}^{-T} A_{31}^T + A_{31} B_{11}^{-1} B_{13} + D_{21} D_{11}^{-1} D_{21}^T \tag{5}$$

Altogether, we have shown for matrices M close enough[17] to $\overline{N}$:

$$\operatorname{rank}(B) = k \wedge (\operatorname{In} A_{|\ker B}) = \xi \quad \text{iff} \quad (4) \wedge (5) \text{ holds.}$$

Taking into account the symmetry of A_{11}, the combined Condition (4) ∧ (5) gives rise to a set of $(n-k)(m-k) + (1/2)\xi^0(\xi^0 + 1)$ scalar equations, with on the left hand sides only variables corresponding to the (independent) entries of B_{22}, B_{23} and A_{33}, whereas the right hand sides may be considered – due to (†) – as smooth functions on the entries of the other sub matrices of M. Clearly, the gradients of these equations w.r.t. the (independent) entries of B_{22}, B_{23}, A_{33} are linearly independent. Hence,[18]

$$(4) \wedge (5) \text{ is a defining system of } \{M|\text{rank}(B) = k, \text{In}(A_{|\ker B}) = \xi\} \text{ around } \overline{N} \qquad (*)$$

By the Implicit Function Theorem, there exists an open neighbourhood, say $\mathcal{N}$, in $\{M|\ \text{rank}(B) = k,\ \text{In}(A_{|\ker B}) = \xi\}$ around $\overline{N}$, such that on $\mathcal{N}$ the condition (4) $\wedge$ (5) holds (as an independent set of equations).

In the sequel, we consider open neighbourhoods in $\mathcal{N} \times \mathbb{R}^{n+m}$ around $[\overline{N} \vdots \overline{n}]$, denoted $V^{(\ell)}$, where ℓ refers to the fact that $[\overline{N} \vdots \overline{n}]$ is supposed to be the normal form of $V^{(\ell)}_{\xi,k,[\tau]}$.

We are looking for necessary and sufficient conditions in order that matrices $[M \vdots c]$ in $V^{(\ell)}$, and close to $[\overline{N} \vdots \overline{n}]$, are contained in the stratum $V^{(\ell)}_{\xi,k,[\tau]}$. Doing so, we assume:[19] $0 < k < m < n$; $0 < r(= \xi^+ + \xi^-) < n - k$.

We proceed by distinguishing between the various ℓ-values.

If $\ell = 1$:

Apart from the features assured on $V^{(1)}$ by (4) $\wedge$ (5), only one additional condition is needed to characterize $V^{(1)}_{\xi,k,[\tau]}$: rank$[B \vdots b] = k + 1$; cf Table 1.

From rank $[\overline{B} \vdots \overline{b}] = k + 1$ and by a continuity argument it follows: For $[M \vdots c] \in V^{(1)}$ close enough to $[\overline{N} \vdots \overline{n}]$, and thus $[B \vdots b]$ close enough to $[\overline{B} \vdots \overline{b}]$, we have: rank $[B \vdots b] = k + 1$. So, apart from (4) and (5), no other equations are needed to give a local characterization of $V^{(1)}_{\xi,k,[\tau]}$ around $[\overline{N} \vdots \overline{n}]$. Hence, see (*), we find that (4) $\wedge$ (5) gives rise to a defining system for $V^{(1)}_{\xi,k,[\tau]}$ around $[\overline{N} \vdots \overline{n}]$. Thus $V^{(1)}_{\xi,k,[\tau]}$ is a smooth manifold with codimension as indicated in Table 3.

If $\ell = 2$:

Apart from the features assured on $V^{(2)}$ by (4) $\wedge$ (5), three additional conditions are used to characterize $V^{(2)}_{\xi,k,[\tau]}$: $b \neq 0$, rank$[M \vdots c] = \text{rank}(M) + 1$ and rank$[B \vdots b] = k$; cf Table 1.

For $[M \vdots c] \in V^{(2)}$ close enough to $[\overline{N} \vdots \overline{n}]$ the first two conditions remain automatically valid (and therefore, they do not contribute to the defining system we are looking for). The persistency of the inequality $b \neq 0$ follows directly from a continuity argument. In case of the second condition, we use Corollary 2: rank$(M) = 2[\text{rank}(B)] + r$. Due to (4) $\wedge$ (5), rank$(B)(= k)$ and $\xi^+ + \xi^-(= r)$ remain constant on $V^{(2)}$. Since rank$[\overline{N} \vdots \overline{n}] = 2k + r + 1$, the persistency of 'rank$[M \vdots c] >$ rank(M)' follows from a continuity argument applied to $[M \vdots c]$.

So, we focus on the third additional condition: rank$[B \vdots b] = k$. With the notations as introduced before, we consider the condition

$$\text{rank}[B \vdots b] = \text{rank}\begin{bmatrix} B_{11} & B_{12} & B_{13} & \vdots & b_1 \\ B_{21} & B_{22} & B_{23} & \vdots & b_2 \end{bmatrix} = k.$$

Put $[B \vdots b]$ sufficiently close to $[\overline{B} \vdots \overline{b}]$. Then: $B_{11} \approx I_k$, $b_1 \approx e(\neq 0)$, and all other blocks of $B \approx O$.

In view of Lemma 7: 'rank $[B \vdots b] = k$' iff '$[B_{22} B_{23} \vdots b_2] = B_{21} B_{11}^{-1} [B_{12} B_{13} \vdots b_1]$'.

Since (4) is already fulfilled on $V^{(2)}$, the right hand side of this equivalency reduces to:

$$b_2 = B_{21} B_{11}^{-1} b_1 \qquad (6)$$

Note that (6) gives rise to $(m-k)$ scalar equations. Moreover, on its left hand side, only the entries of b_2 occur, showing up neither in the right-hand side of (6), nor in (4) $\wedge$ (5). From this, it follows that (4) $\wedge$ (5) $\wedge$ (6) forms an independent system; compare also the (similar) proof of (*). Altogether (4) $\wedge$ (5) $\wedge$ (6) yields a defining system for $V^{(2)}_{\xi,k,[\tau]}$ around $[\overline{N} \vdots \overline{n}]$.

So, $V^{(2)}_{\xi,k,[\tau]}$ is a smooth manifold of codimension as specified in Table 3.

If $\ell = 3$:

Apart from the features assured on $V^{(3)}$ by (4) and (5), three additional conditions are needed to characterize $V^{(3)}_{\xi,k,[\tau]}$: $b \neq 0$, rank $[M \vdots c] = \text{rank}(M)$, $\text{sign}[M \vdots c] = \tau$; cf = Table 3.

The persistency of the inequality $b \neq 0$, for $[M \vdots c]$ sufficiently near to $[\overline{N} \vdots \overline{n}]$, follows (again) by a continuity argument.

In order to treat the condition $\text{rank}[M \vdots c] = \text{rank}(M)$, we partition c into sub vectors $a_1, a_2^{\pm}, a_2^0, b_1, b_2$, according to the block structures of A and B. We consider the matrices

$$[M \vdots c] = \begin{bmatrix} A_{11} & A_{21}^T & A_{31}^T & B_{11}^T & B_{21}^T & \vdots & a_1 \\ A_{21} & A_{22} & A_{32}^T & B_{12}^T & B_{22}^T & \vdots & a_2^{\pm} \\ A_{31} & A_{32} & A_{33} & B_{13}^T & B_{23}^T & \vdots & a_2^0 \\ B_{11} & B_{12} & B_{13} & O & O & \vdots & b_1 \\ B_{21} & B_{22} & B_{23} & O & O & \vdots & b_2 \end{bmatrix}$$

and

$$[F \vdots c'] = \begin{bmatrix} A_{11} & A_{21}^T & B_{11}^T & A_{31}^T & B_{21}^T & \vdots & a_1 \\ A_{21} & A_{22} & B_{12}^T & A_{32}^T & B_{22}^T & \vdots & a_2^{\pm} \\ B_{11} & B_{12} & O & B_{13} & O & \vdots & b_1 \\ A_{31} & A_{32} & B_{13}^T & A_{33} & B_{23}^T & \vdots & a_2^0 \\ B_{21} & B_{22} & O & B_{23} & O & \vdots & b_2 \end{bmatrix}$$

where the latter matrix is obtained from the first one, by interchanging the 3rd and 4th block columns, and subsequently the 3rd and 4th block rows.

We put:

$$Z = \begin{bmatrix} A_{11} & A_{21}^T & B_{11}^T \\ A_{21} & A_{22} & B_{12}^T \\ B_{11} & B_{12} & O \end{bmatrix} \text{ and } I^{\pm} = \begin{bmatrix} O & O & I_k \\ O & J^{\pm} & O \\ I_k & O & O \end{bmatrix}, \text{ where } J^{\pm} = \begin{bmatrix} I_{\xi^+} & O \\ O & -I_{\xi^-} \end{bmatrix}.$$

Let $[M \vdots c]$ be sufficiently near to $[\overline{N} \vdots \overline{n}]$. Then: $Z(\approx I^{\pm})$ is a regular $(2k+r) \times (2k+r)$-matrix. By construction of $V^{(3)}$, and using Corollary 2, we find: rank(M) (= rank $(F)) = 2k + r$, and thus rank $[M \vdots c]$ (= rank $[F \vdots c']) = 2k + r$. Now, applying Lemma 7 to $[F \vdots c']$, with Z in the role of H_1, yields the following necessary & sufficient condition for $\text{rank}[M \vdots c] = \text{rank}(M)$:

$$\begin{bmatrix} A_{33} & B_{23}^T & \vdots & a_2^0 \\ B_{23} & O & \vdots & b_2 \end{bmatrix} = \begin{bmatrix} A_{31} & A_{32} & B_{13}^T \\ B_{21} & B_{22} & O \end{bmatrix} Z^{-1} \begin{bmatrix} A_{31}^T & B_{21}^T & \vdots & a_1 \\ A_{32}^T & B_{22}^T & \vdots & a_2^{\pm} \\ B_{13} & O & \vdots & b_1 \end{bmatrix}.$$

In principle, this gives rise to six matrix equations. However, the equations corresponding with the blocks A_{33}, B_{23}^T, B_{23}, and O on the left-hand side are already fulfilled. This follows from Lemma 7, applied to matrix F with Z in the role of H_1. So, the above equations reduce to:

$$a_2^0 = [A_{31}\ A_{32}\ B_{13}^T] Z^{-1} (a_1^T, (a_2^{\pm})^T, b_1^T)^T \tag{7}$$

$$b_2 = [B_{21}\ B_{22}\ O] Z^{-1} (a_1^T, (a_2^{\pm})^T, b_1^T)^T \tag{8}$$

The latter relation is nothing else than Relation (6). In fact, we note that if again $[M \vdots c] \approx [\overline{N} \vdots \overline{n}]$, thus $\text{rank}(Z) = 2k + r$, a unique triple $(u, v, w) \in \mathbb{R}^k \times \mathbb{R}^r \times \mathbb{R}^k$ exists such that:

$$\begin{bmatrix} A_{11} & A_{21}^T & B_{11}^T \\ A_{21} & A_{22} & B_{12}^T \\ B_{11} & B_{12} & O \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2^{\pm} \\ b_1 \end{bmatrix}, \text{ or } \begin{bmatrix} u \\ v \\ w \end{bmatrix} = Z^{-1} \begin{bmatrix} a_1 \\ a_2^{\pm} \\ b_1 \end{bmatrix} \tag{**}$$

Hence, by substituting (**) in (8) and using (4):

$$b_2 = B_{21}u + B_{22}v = B_{21}B_{11}^{-1}[B_{11}u + B_{12}v] = B_{21}B_{11}^{-1}b_1.$$

Altogether, for $[M \vdots c] \in V^{(3)}$ close enough to $[\overline{N} \vdots \overline{n}]$:

$$\text{rank}[M \vdots c] = \text{rank}(M) \text{ iff } (6) \wedge (7) \text{ holds.}$$

Note that in the left-hand side of (6) $\wedge$ (7) only entries appear, which are neither present in its right hand side, nor in (4) $\wedge$ (5).

In order to analyze the last additional condition ($\text{sign}[M \vdots c] = \tau$), we note that (6) $\wedge$ (7) and (**) imply that $y = -(u^T, v^T, 0^T, w^T, 0^T)^T$ is a solution for $My + c = 0$. Hence,

$$c^T y = -(a_1^T, (a_2^{\pm})^T, b_1^T) Z^{-1} (a_1^T, (a_2^{\pm})^T, b_1^T)^T.$$

Suppose $\tau = \pm 1$:

Then, at $[\overline{N} \vdots \overline{n}]$ we have $c^T y \neq 0$. This inequality remains valid for $[M \vdots c]$ close enough to $[\overline{N} \vdots \overline{n}]$ (continuity argument). So, in these cases no additional conditions are needed.

Suppose $\tau = 0$:

Then, at $[\overline{N} \vdots \overline{n}]$ we have $c^T y = 0$. In this case, we have to cope with an additional condition:

$$(a_1^T, (a_2^{\pm})^T, b_1^T) Z^{-1} (a_1^T, (a_2^{\pm})^T, b_1^T)^T = 0. \tag{9}$$

For $[M \vdots c] \approx [\overline{N} \vdots \overline{n}]$, the left-hand side of (9) is a smooth function on entries of $[M \vdots c]$, but not on the entries of B_{22}, B_{23}, A_{33}, a_2^0, and b_2, cf. (4), (5), (6), (7).

If $\tau = \pm 1$, the independency of (4) $\wedge$ (5) $\wedge$ (6) $\wedge$ (7) is proved as for (4) $\wedge$ (5) in (*).

If $\tau = 0$, the extension to a system (4) $\wedge$ (5) $\wedge$ (6) $\wedge$ (7) $\wedge$ (9) does not hurt the independency. This is so, because the partial derivative at $[\overline{N} \vdots \overline{n}]$, cf. Table 2, $\ell = 3$, $\tau = 0$, of the left-hand side of (9) with respect to the first component of $(a_1^T, (a_2^{\pm})^T, b_1^T)$ does not vanish. In fact, for this partial derivative we find:

$$2[1, 0, \ldots, 0] \begin{bmatrix} O & O & I_k \\ O & J^{\pm} & O \\ I_k & O & O \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ e \end{bmatrix} = 2$$

Summarizing, the defining systems that, we seek, are (4) $\wedge$ (5) $\wedge$ (6) $\wedge$ (7) [if $\tau = \pm 1$], and (4) $\wedge$ (5) $\wedge$ (6) $\wedge$ (7) $\wedge$ (9) [if $\tau = 0$]. Taking the dimensions of B_{22}, B_{23} , A_{33} [symmetric!], a_2^0, and b_2 into account, we find the codimensions as indicated in Table 2, $\ell = 3$.

If $\ell = 4$:

Apart from the features, assured on $V^{(4)}$ by (4) and (5), two additional conditions are needed to characterize $V^{(4)}_{\xi,k,[\tau]}$: $\mathrm{rank}[M \vdots c] = \mathrm{rank}(M) + 1$, $b = 0$, cf Table 1. The first (rank) condition does not contribute to a (possible) defining system; see case $\ell = 2$ where we encountered a similar situation.

The second additional condition yields:

$$b_1 = 0, b_2 = 0 \tag{10}$$

The entries of the left-hand side of (10) do not occur in (4) $\wedge$ (5). So, as in the situation described in (*), the conditions (4) $\wedge$ (5) $\wedge$ (10) constitute a defining system for $V^{(4)}_{\xi,k,[\tau]}$ around $[\overline{N} \vdots \overline{n}]$, with co-dimension as indicated in Table 3.

If $\ell = 5$:

Apart from the conditions assured on $V^{(5)}$ by (4) and (5) three additional conditions are needed to characterize $V^{(5)}_{\xi,k,[\tau]}$: $\mathrm{rank}[B^T \vdots a] = \mathrm{rank}(B^T)(= k)$, $b = 0$, $a \neq 0$; cf. Table 1. For $[M \vdots c]$ sufficiently near to $[\overline{N} \vdots \overline{n}]$, we have: $B_{11}^T (\approx I_k)$ is regular. Applying Lemma 7 to $[B^T \vdots a] (\approx [\overline{B}^T \vdots \overline{a}])$, with B_{11}^T in the role of H_1, and taking (4) into account, we find the following necessary & sufficient condition for $\mathrm{rank}[B^T \vdots a] = \mathrm{rank}(B^T)$:

$$\begin{aligned} a_2^{\pm} &= B_{12}^T B_{11}^{-T} a_1 \\ a_2^0 &= B_{13}^T B_{11}^{-T} a_1 \end{aligned} \tag{11}$$

The second additional condition is just (10), whereas the last condition does not affect a defining system (since, by a continuity argument, $a \neq 0$ remains valid if $[M \vdots c] \approx [\overline{N} \vdots \overline{n}]$). Similar to the preceding cases, we find as a desired defining system: (4) $\wedge$ (5) $\wedge$ (10) $\wedge$ (11), with codimension according to Table 3, $\ell = 5$.

If $\ell = 6$:

Now, in addition to the features assured on $V^{(6)}$ by (4) and (5), four extra conditions must be fulfilled in order to characterize $V^{(6)}_{\xi,k,[\tau]}$:

$$\text{rank}[B^T \vdots a] = k + 1, b = 0, \text{rank}[M \vdots c] = \text{rank}(M), \text{sign}[M \vdots c] = \tau.$$

The inequality $\text{rank}[B^T \vdots a] > k$ remaining valid for all $[M \vdots c] \in V^{(6)}$ sufficiently close to $[\overline{N} \vdots \overline{n}]$ – cf the preceding cases – we focus on the latter three conditions. We adopt the notations, used in case $\ell = 3$.

Again the second additional condition is just (10). Due to (10), we find for $[M \vdots c] (\approx [\overline{N} \vdots \overline{n}])$ in $V^{(6)}$ as necessary & sufficient condition for $\text{rank}[M \vdots c] = \text{rank}(M)$: (compare case $\ell = 3$).

$$a_2^0 = [A_{31}\ A_{32}\ B_{13}^T] Z^{-1} (a_1^T, (a_2^{\pm})^T, 0^T)^T \tag{12}$$

Now, we turn over to the last additional condition (sign $[M \vdots c] = \tau$).

As in case $\ell = 3$, we have for $[M \vdots c] (\approx [\overline{N} \vdots \overline{n}])$ in $V^{(6)}$:

$$\text{sign}[M \vdots c] = -\text{sign}((a_1^T, (a_2^{\pm})^T, 0^T)^T Z^{-1} (a_1^T, (a_2^{\pm})^T, 0^T)^T).$$

If $\tau = \pm 1$, then sign $[M \vdots c] \neq 0$ remains valid for $[M \vdots c]$ close enough to $[\overline{N} \vdots \overline{n}]$, and no extra equations in the desired defining system are needed.

If $\tau = 0$, we introduce:

$$(a_1^T, (a_2^{\pm})^T, 0^T) Z^{-1} (a_1^T, (a_2^{\pm})^T, 0^T)^T = 0 \tag{13}$$

A straightforward calculation learns:

The partial derivative at $[\overline{N} \vdots \overline{n}]$, with $\overline{n}$ as in Table 1, $\ell = 6$, $\tau = 0$, of the left hand side of (13) w.r.t. the jth component of vector $(a_1^T, (a_2^{\pm})^T)$, equals -2 (if $j = k + 1$) and $+2$ (if $j = k + \xi^+ + 1$). Note that sub vector $\overline{n}_2^{\pm} = (e^T, e^T)^T$ is not 'empty' if $V^{(6)}_{\xi,k,[0]}$ is not 'empty'.

Altogether, we find: The desired defining system is (4) $\wedge$ (5) $\wedge$ (10) $\wedge$ (12) , if $\tau = \pm\ 1$, and (4) $\wedge$ (5) $\wedge$ (10) $\wedge$ (12) $\wedge$ (13), if $\tau = 0$. The proof of the independency (being similar to the proof given in case $\ell = 3$, see also (*)), will be deleted.

If $\ell = 7$: Similar to the preceding cases, we find as a defining system: (4) $\wedge$ (5) $\wedge$ $a = 0 \wedge b = 0$. See also endnote 13. ■

6. Proof of Theorem 3 (Whitney regular stratification)

The proof of Theorem 3 relies upon some well-known results from Algebraic Geometry / Stratification Theory:

A subset of $\mathbb{R}^N$ is called *semi-algebraic* if it is generated – in the Boolean sense – by finitely many polynomial equalities and inequalities. Consider a partition of $\mathbb{R}^N$ into semi algebraic strata. Moreover, let us assume that this partition fulfils the so-called *homogeneity property* (cf. Section 2, endnote 6). Then this partition is a Whitney regular stratification. (Compare the comment at the end of Section 2).

In the present situation, Lemma 1, Lemma 2 and (the proof of) Theorem 1 obviously assure the homogeneity property. So, we are done if we are able to prove that our partition of $\mathbb{M}_{n,m}$ into the strata $V^{(\ell)}_{\xi,k,[\tau]}$ is a semi-algebraic stratification. In fact, this can be shown. See forthcoming Corollary 3. In proving Corollary 3, the following complication arises: All conditions for $V^{(\ell)}_{\xi,k,[\tau]}$ see Table 1, take the form of polynomial (in-)equalities ['rank conditions'], with the exception of '$\tau = \pm$ 1, 0' and '$\mathrm{In}(\mathrm{A}_{|\ker \mathrm{B}}) = \xi$' which are – at first sight – not of polynomialnature.[20] In case '$\tau = \pm$ 1, 0', we overcome this problem by introducing an equivalent global condition which is polynomial (see forthcoming Lemma 8).

For '$\mathrm{In}(\mathrm{A}_{|\ker \mathrm{B}}) = \xi$', we will use (see forthcoming Corollary 3) another basic result from Algebraic Geometry, namely: The stratification induced by the connected components of a semi algebraic set is again semi algebraic (cf [2]). ■

We finish this section by presenting the missing links in the above proof of Theorem 3.

LEMMA 8 *Let* $k + \xi^+ + \xi^- \geq 0$. *Then there exists a polynomial on* $\mathbb{R}^N$, *say* $\Pi_{\xi,k}$, *such that*:

$$\mathrm{sign}[M \vdots c] = \mathrm{sign}\Pi_{\xi,k}([M \vdots c]).$$

for all[21] $[M \vdots c] \in V^{(\ell)}_{\xi,k,[\tau]}$, $\ell = 3, 6$.

Proof We put $p = 2k + \xi^+ + \xi^-$ and thus $0 < p \leq n + k \leq n + m$.

Let $\omega = m_{j1}, \ldots, m_{jp}\}$, $1 \leq j_1 < \cdots < j_p \leq n + m$ be an ordered set of columns of M.

For any vector d in $\mathbb{R}^{n+m}$, we introduce $\omega(d) := (d^T m_{j_1}, \ldots, d^T m_{j_p})^T \in \mathbb{R}^p$. The determinant of the matrix $[\omega(m_{ji})]$ is just the *Gramian* of ω, denoted G_ω. As it is well-known, the Gramian of a set of vectors is always non-negative and only vanishes in the case of linear dependency (cf [1]). Define $G_\omega(i)$ as the determinant of the matrix obtained from $[\omega_{ji})]$ by changing the ith column of the latter matrix into $\omega(c)$. Then, the desired polynomial is:

$$\Pi_{\xi,k}([M \vdots c]) = \sum_\omega -G_\omega \sum_{i=1}^p c_{j_i} G_\omega(i),$$

where the summation takes place over *all possible* sets ω of p columns in M.

To see this, let $[M \vdots c] \in V^{(\ell)}_{\xi,k,[\tau]}$, $\ell = 3,6$. Note that rank $(M) = \mathrm{rank}[M \vdots c]$. Due to Corollary 2 we have $p = \mathrm{rank}(M)$. Thus among the sets ω of columns of M, there exist ω's which are linearly independent. Given any such ω, we obviously have:

$$c = \sum_{i=1}^p \mu_i m_{j_i}, \tag{14}$$

where the μ_i's are uniquely determined. Now, let y be a vector in $\mathbb{R}^{n+m}$ with the j_ith components equal to $-\mu_i$, and the other components equal to zero. Then $My + c = 0$, and thus τ ($=$sign $[M \vdots c]$)$=$sign$(y^T c)$. Moreover, we find:

$$y^T c = -\sum_{i=1}^{p} \mu_i c_{j_i}, \tag{15}$$

where c_{j_i} stands for the j_ith component of c. From (14), it follows:

$$\omega(c) = \sum_{i=1}^{p} \mu_i \omega(m_{j_i}).$$

The latter relation may be viewed to as to a system of linear equations in the p unknown's μ_i. Applying Cramer's rule yields:

$$\mu_i = \frac{G_\omega(i)}{G_\omega}, \quad i = 1, \ldots, p,$$

Substituting this result into (15) gives:

$$y^T c = -\sum_{i=1}^{p} \frac{G_\omega(i)}{G_\omega} c_{j_i}.$$

Multiplication with $(G_\omega)^2$ and noting that $G_\omega > 0$, yields:

$$\tau = \text{sign}[M \vdots c] = \text{sign}(y^T c) = \text{sign}\left[-G_\omega \sum_{i=1}^{p} c_{j_i} G_\omega(i)\right]$$

Summation over all possible[22] ω's yields our assertion; use that sign$[M \vdots c]$ does not depend on the ambiguity in the choice of y (cf. Section 2). ∎

LEMMA 9 *The set* $\mathtt{C}$ *of all* $p \times p$*-matrices* C, *with det* $(C) > 0$, *is path wise connected* (*and thus* $\mathtt{C}$ *is a connected set in* $\mathbb{R}^{p^2}$).

Proof Let $C \in \mathtt{C}$ be arbitrary, but fixed. We are done, if we can find a continuous path, say C_t, $t \in [0,1]$, such that det $(C_t) > 0$, $t \in [0,1]$, $C_0 = C$, and $C_1 = I_p$. It is well-known, cf. [1], that we have the following polar decomposition of C:

$$C = PY,$$

where P is a positive-definite, and Y an orthogonal matrix. (In fact, $P = (CC^T)^{1/2}$ and, since $\det(C) > 0$, also $\det(Y) > 0$). Moreover, Y being orthogonal, we may write:

$$Y = X \,\text{diag}(I_s, R(\varphi_1), \ldots, R(\varphi_d))X^{-1},$$

where X is an orthogonal matrix and diag$(\cdot)$ a diagonal block-matrix, with on its diagonal: $I_s = s \times s$-identity matrix and $R(\varphi)$ are 2×2-matrices of the form:

$$R(\varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}, 0 < \varphi \le \pi,\ \ s + 2d = p.$$

Now, the continuous path, we are seeking, is defined as follows:

$$C_t = \begin{cases} ((1-2t)P + 2tI_p)Y & \text{if } t \in \left[0, \frac{1}{2}\right] \\ X \operatorname{diag}(I_s, R(2(1-t)\varphi_1), \ldots, R(2(1-t)\varphi_d))X^{-1}, & \text{if } t \in \left[\frac{1}{2}, 1\right] \end{cases}$$

(Note that $((1-2t)P + 2tI_p)$ is positive-definite for all $t \in [0, 1/2]$). ■

We consider the sets

$$\sum_{\xi,k} = \{M \mid \operatorname{rank}(B) = k; \operatorname{In}(A_{|\ker B}) = \xi\}.$$

Apparently, $\sum_{\xi,k}$ may be viewed to as to a subset (even a smooth manifold, cf. endnote 13) of $\mathbb{R}^{N-m-n}$.

LEMMA 10 *All sets* $\sum_{\xi,k}$, *with* $k < n$, *are path wise connected, and thus connected.*[23]

Proof Let $M \in \sum_{\xi,k}$ be arbitrary, but fixed, and $\overline{N}$ the normal form for M as introduced in Theorem 1. In Lemma 3 we proved the existence of matrices $U \in \mathcal{G}$:

$$\left. \begin{aligned} U = \begin{bmatrix} V & K \\ O & W \end{bmatrix}, \quad \det(V) \neq 0,\ \det(W) \neq 0 \\ \text{such that } UMU^T = \overline{N}. \end{aligned} \right\} \qquad (***)$$

We are done if we are able to prove that we can choose U such that $\det(V) > 0$, $\det(W) > 0$. For then, due to Lemmas 2, 9, we can define a continuous path $U_E M U_E^T$ from $\overline{N}$ to M in $\sum_{\xi,k}$, where

$$U_t = \begin{bmatrix} V_t & (1-t)K \\ O & W_t \end{bmatrix}.$$

Suppose $\det(W) < 0$. We choose a matrix of the form U, sayU', where U' is the diagonal matrix with all diagonal entries equal to $+1$ with the exception of the 1st and$(n+1)$th ones which are equal to -1. Then $(U'U)M(U'U)^T = \overline{N}$ (for $k > 0$ use Lemma 4), and

$$U'U = \begin{bmatrix} V' & K' \\ O & W' \end{bmatrix}, \quad \det(V') \neq 0,\ \det(W') > 0.$$

If $\det(V') > 0$, we are done. If not, we choose a matrix of the form U, say U'', as introduced in Lemma 6 with H a diagonal matrix with all diagonal entries equal to $+1$, with the exception of the first one which equals -1 (since $k < n$ this is possible). Then $(U''U'U)M(U''U'U)^T = \overline{N}$, and

$$U''U'U = \begin{bmatrix} V'' & K'' \\ O & W' \end{bmatrix}, \quad \det(V'') > 0,\ \det(W') > 0.$$

■

COROLLARY 3 *The strata* $V^{(\ell)}_{\xi,k,[\tau]}$ *constitute a semi algebraic stratification for* $\mathbb{M}_{n,m}$.

Proof For any pair (k, r), with $0 \le k \le \min(n, m)$ and $0 \le r \le n-k$ we define

$$W_{(k,r)} = \{M \mid \text{rank}(B) = k, \text{rank}(M) = 2k + r\}.$$

In view of Corollary 2, the set $W_{(k,r)}$ is the union of sets of the type $\sum_{\xi,k}$, $\xi=(\xi^+, \xi^-, \xi^0)$ with $\xi^+ + \xi^- = r$. Here, the union is taken over all $(r+1)$ possible sets of this type. If $k=n$, then $r=0$ and $W_{(k,r)}$ reduces to $W_{(n,0)} = \{M|\text{rank}(B)=n\}$, which is semi algebraic. If $k < n$, then the sets $\sum_{\xi,k}$, $\xi^+ + \xi^- = r$, are connected [use Lemma 10], *open and closed* sets in $W_{(k,r)}$. Hence, the sets $\sum_{\xi,k}$, $\xi^+ + \xi^- = r$, are the connected components of $W_{(k,r)}$. The set $W_{(k,r)}$ is semi algebraic, and so are its connected components $\sum_{\xi,k}$.

Finally, we emphasize that the (conditions for) the strata $V^{(\ell)}_{\xi,k,[\tau]}$ are obtained from (those for) the sets $\sum_{\xi,k}$ by taking into account additional conditions, which (Table 1 and Lemma 8) have the form of polynomial (in-)equalities. Thus, the strata $V^{(\ell)}_{\xi,k,[\tau]}$ are semi-algebraic subsets of $\mathbb{M}_{n,m}$. ∎

Acknowledgements

The authors thank the two referees. Their comments were very helpfull to improve this paper considerably. We also thank Dini Heres for preparing the layout of the manuscript.

Notes

1. If $k=n$ $(\le m)$, then S is not defined, but we formally put $(\xi^+, \xi^-, \xi^0)=(0, 0, 0)$ in this case. In order to focus on the general line of reasoning, here and in the sequel we shall not dwell on these types of 'degeneracies'; see, however, the forthcoming Tables 1, 2, 4, 5 and related comments.
2. Note that 'rank$(B)=m$'(thus $0 \le m \le n$) means: the linear independency constraint qualification (LICQ) holds at all (possible) feasible points for Q. Moreover, if – under LICQ – we have $\mathcal{K}_Q \ne \emptyset$, then $\xi^0=0$ yields a well-known second-order non-degeneracy condition on the 'restricted Hessian of the Lagrange function' at a critical point $\bar{x}$ for Q. The condition $\xi^0=0$, together with LICQ, will be referred to as to ND. Note that if rank$(B)=m$ and $\xi^0=0$, we have: rank $M=n+m$ and thus, $\mathcal{K}_Q$ consists of the singleton $(x^T \bar{\lambda}^T)^T$. See, e.g. the forthcoming Corollary 2 (Inertia Theorem).
3. One could expect that sign$[M \,\vdots\, c]$ not only plays a role in Clusters 3 and 6 but also in Clusters 5 and 7. This is not the case, because under the features in this clusters, we always have: sign$[M \,\vdots\, c]=0$.
4. Two points, say x and y, in a connected component of a stratum are said to 'have the same topological type with respect to the stratification' whenever there exist two neighbourhoods of x respectively y, and a homeomorphism between these neighbourhoods that preserves the local stratification around these points.
5. For a definition of *semi-algebraic (partition)*, see Section 6.
6. That is, given any two points in the same stratum, there exists a diffeomorphism of an $\mathbb{R}^N$-neighbourhood of one point onto an $\mathbb{R}^N$-neighbourhood of the other point which preserves strata.
7. That is, locally around such a regular point the semi algebraic set is a smooth submanifold.
8. That is, restricted to the C^1-open and-dense set of mappings $Q(\cdot)$, which are transversal to each stratum $V^{(\ell)}_{\xi,k,[\tau]}$.
9. Thus, the gradients at x of the object function for Q, together with the gradients of the (active) constraint functions form a linear dependent set. See e.g., [7,11].

10. In fact, up till now, only an informal, hand written manuscript with these proofs is available.
11. Note that such P exists since $\xi^0 \geq 1$ and $\xi^+ + \xi^- = r$ implies: $(r+1) \leq (n-k)$.
12. So, $\xi^k = \xi^- = 0$ does not occur (cf. list of empty strata in Table 1).
13. As a consequence of this expression for τ we have: combinations '$\xi^- = 0$, with either $\tau = 0$ or $\tau = 1$' and '$\xi^+ = 0$, with either $\tau = 0$ or $\tau = -1$' do not occur; compare also the list of empty strata in Table 1.
14. Such φ_0 exists and is uniquely determined, in fact: $\varphi_0 = \tanh^{-1}(-\alpha_-/\alpha_+)$; note that $|-\alpha_-/\alpha_+| < 1$.
15. Open with respect to the relative topology on $\mathcal{M}$, induced by the standard topology on $\mathbb{R}^N$.
16. In contradistinction to Section 4, here a pair of sub indices of a block does *not* indicate its dimension.
17. Use alo the continuous dependency of the eigenvalues of a symmetric matrix on its entries.
18. In fact, due to Lemmas 1–3, this proves that the set $\{M | \text{rank}(B) = k, \text{In}(A_{|ker\ B}) = \xi\}$ is a smooth manifold in $\mathbb{R}^{N-n-m}$ of codimension $(n-k)(m-k) + (1/2)\xi^0(\xi^0 + 1)$.
19. This assumption does not imply a restriction to our final results.
20. Note that in the proof of Theorem 2 (cases $\ell = 3,6$) we characterized these conditions – essentially – by means of polynomials. However these characterizations are of a 'local nature', whereas for our present aim global characterizations are needed.
21. Note that violation of the condition $k + \xi^+ + \xi^- > 0$ yields in the cases $\ell = 3,6$ empty strata, cf Table 1.
22. Here it is not necessary to restrict ourselves to sets ω which are linearly independent since in case of linear dependency of ω we have $G_\omega = 0$.
23. It is not difficult to show that this also holds for all other sets $\sum_{\xi,k}$ with the exception of $\sum_{0,k=m=n}$, which admits precisely two connected components.
24. With respect to relative topology on $W_{(k,r)}$.

References

[1] J.R. Gantmacher, *The Theory of Matrices*, Vol. 1 and Vol. 2, Chelsea Publ. Company, New York, 1960.

[2] C.G. Gibson, K. Wirthmüller, A.A. Du Plessis, et al., *Topological Stability of Smooth Mappings*, Lecture Notes in Math., Vol. 552, Springer-Verlag, Berlin, 1976.

[3] J. Guddat, et al., *Parametric optimization and related topics III*, in *Series Approximation and Optimization*, Peter Lang Verlag, Frankfurt, 1993.

[4] J. Guddat, et al., *Parametric optimization and related topics IV*, in *Series Approximation and Optimization*, Peter Lang Verlag, Frankfurt, 1997.

[5] J. Guddat, F. Guerra, and H.Th. Jongen, *Parametric Optimization: Singularities, Pathfollowing and Jumps*, Teubner and John Wiley, Chichester, 1990.

[6] M.W. Hirsch, *Differential Topology*, Springer-Verlag, New York, 1976.

[7] H.Th. Jongen, P. Jonker, and F. Twilt, Critical sets in parametric optimization, Math. Programming 34 (1986), pp. 333–353.

[8] H.Th. Jongen, P. Jonker, and F. Twilt, *One-parameter families of optimization problems: equality constraints*, J. Optim. Theory Appl. 48 (1986), pp. 141–161.

[9] H.Th. Jongen, J. Möbert, J. Rückmann, et al., *On inertia and Schur complement in optimization*, Lin. Algebra Appl. 95 (1987), pp. 97–109.

[10] H.Th. Jongen, P. Jonker, and F. Twilt, *Nonlinear Optimization* in *Finite Dimension: Morse Theory, Chebyshev Approximation, Transversality, Flows, Parametric Aspects*, Kluwer Ac. Publisher, Dordrecht, Boston, 2000.

[11] P. Jonker, G. Still, and F. Twilt, *One parametric linear-quadratic optimization problems*, Ann. Oper. Res. 101 (2001), pp. 221–253.