![TU/e Eindhoven University of Technology logo]

Decoupling of bidiagonal systems involving singular blocks

Document status and date:
Published: 01/01/1989

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

Download date: 16. Nov. 2023

Eindhoven University of Technology
Department of Mathematics and Computing Science

RANA 89-25
December 1989
DECOUPLING OF BIDIAGONAL
SYSTEMS INVOLVING
SINGULAR BLOCKS
by
R.M.M. Mattheij

DECOUPLING OF BIDIAGONAL SYSTEMS INVOLVING SINGULAR BLOCKS

R.M.M. Mattheij
Faculteit Wiskunde en Informatica
Technische Universiteit Eindhoven
Postbus 513
5600 MB Eindhoven
The Netherlands

Abstract

For one step difference equations, where the matrix coefficients may be singular, a stability analysis based on using fundamental solutions and their inverses does not apply. This paper shows how well-boundedness of the Green's function leads to a kind of dichotomy of the fundamental solution, including certain "parasitic solutions" (which arise because of the singularity of the fundamental solutions). This then is used to show how one can find a stable decoupling and thus a numerical algorithm for solving a discrete BVP. Several examples sustain the analysis.

1. Introduction

In analyzing numerical methods for boundary value problems, one often uses explicit representations or estimates of Green's functions. The notion of well-conditioning of such a BVP is intimately connected to a moderate bound of these Green's functions. As a consequence one can show cf. [6] that there exists a natural splitting of the solution space into nongrowing and nondecaying solutions. It is essential for the latter analysis that the fundamental solution is nonsingular everywhere. A discrete analogue of such an analysis is straight-forward (cf. [2]).

However, if the differential equations are implicit, and have an index ≥ 1, i.e. if in the (linearized) system, the coefficient matrix of the derivative term is allowed to be singular, this may lead to problems. These differential algebraic equations are in fact ODEs coupled with algebraic relations, cf. [5]. For certain numerical methods the rank deficiency is inherited in some of the matrix coefficients of the difference equation. This then explains our interest in rank deficient coefficients. Another example of one step recursions with such a rank deficiency arises from discretization of stiff problems; e.g. if we have

$$\frac{dx}{dt} = \lambda(t)x, \quad \lambda \in I\!R$$

and we use an implicit method, the trapezezoidal rule say, then the resulting recursion reads

$$(1 - h_i \lambda(t_{i+1}))x_{i+1} = (1 + h_i \lambda(t_i))x_i.$$

Clearly for $h_i \lambda(t_{i+1}) = 1$ or $h_i \lambda(t_i) = -1$ either one of the coefficients is zero. A generalisation to systems is obvious.

In the singular coefficient case, we no longer have a fundamental solution that is invertible everywhere. This then makes the more or less standard argumentation for relating well-conditioning to dichotomy no longer applicable. Indeed, apparently some "parasitic" solutions are generated (filling in for the trivial solutions in the Green's function); we have to show that these behave well in some sense too.

In this paper we shall exclusively deal with the discrete *rank deficient case*. We shall analyze the Green's function and draw conclusions with respect to growth of ("parasitic") solutions. As a major result we can thus show how and why a decoupling technique makes sense and is stable for such problems. The analysis is in particular applied to problems which involve block tridiagonal matrices as arise e.g. from collocation methods for BVP. However, they may also be useful for multiple shooting techniques and related methods.

In section 2 we generalise concepts and estimates for Green's functions for the singular case. In section 3 we then show what can be said about the growth of basis modes (as far as they exist). As a special application of this analysis we consider in section 4 tridiagonal matrices arising from a particular repartitioning in a staircase matrix, thus inducing singular blocks. The analysis finally enables us to derive a stable decoupling algorithm in section 5. The theory is illustrated by some

examples in section 6.

2. Green's functions and their consequences

Let us consider a general implicit one step recursion for $\{x_i\}_{i=1}^N$

(2.1a) $A_i x_i + B_i x_{i+1} = f_i, \quad i = 1, \ldots, N-1$,

where $A_i, B_i \in I\!\!R^{n \times n}$, $f_i \in I\!\!R^n$. Assume we have a boundary condition (BC) for $\{x_i\}$ given by

(2.1b) $M_1 x_1 + M_N x_N = b$,

$M_1, M_N \in I\!\!R^{n \times n}$, $b \in I\!\!R^n$.
The problem (2.1) can be written as an $(n \times N)^2$ matrix equation

(2.2a) $A x = b$,

where

(2.2b) $$A := \begin{bmatrix} A_1 & B_1 & & & \\ & A_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & A_{N-1} & B_{N-1} \\ M_1 & & & & M_N \end{bmatrix}, \quad x := \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ \vdots \\ x_N \end{bmatrix}, \quad b := \begin{bmatrix} f_1 \\ \vdots \\ \vdots \\ f_{N-1} \\ b \end{bmatrix}.$$

The conditioning (with respect to absolute errors) is then characterized by the *conditioning constant* κ, defined as (cf. [2])

(2.2c) $\|A^{-1}\| =: \kappa$.

We assume (2.1) to be well posed, i.e. A is invertible. Then there exists a fundamental solution $\{\Phi_i\}_{i=0}^N$ with

(2.3a) $A_i \Phi_i + B_i \Phi_{i+1} = 0$

(2.3b) $M_1 \Phi_1 + M_N \Phi_N = I$.

Also, there exists a Green's function $\{G_{ij}\}_{i=1}^N$, $j = 1, \ldots, N-1$ with

(2.4a) $A_i G_{ij} + B_i G_{i+1,j} = \delta_{ij} I$,

(2.4b) $M_1 G_{ij} + M_N G_{Nj} = 0$.

Explicit expressions for $\{G_{ij}\}$ when $A_i \equiv I$ and B_i is nonsingular for all i can be found e.g. in [7]. In this section we shall derive, more generally, expressions for the potentially rank deficient case, i.e. we assume that for at least one index j a matrix A_j or B_j is singular, i.e.

(2.5) $A_j \Phi_j = -B_j \Phi_{j+1}$

is singular.

If we denote by A^+ the pseudo-inverse of a matrix A (cf. [4]), then

(2.6) $\qquad B_j \Phi_{j+1}(B_j \Phi_{j+1})^+ =: P_j$

defines an orthogonal projection $\mathbb{R}^n \to \text{Range}\,(B_j \Phi_{j+1})$. In order to construct the Green's function we make the Ansatz

(2.7a) $\qquad G_{ij} = \Phi_i K_j + F_{ij}, \quad i \le j \quad K_j \in \mathbb{R}^{n^2}$

(2.7b) $\qquad G_{ij} = \Phi_i L_j + F_{ij}, \quad i \le j \quad L_j \in \mathbb{R}^{n^2}\,,$

where $(1 - P_j)\,F_{ij} = 0$. (Note that $F_{ij} = 0$ is the nonsingular case.) From (2.4a) and (2.7) we see

(2.8a) $\qquad A_j \Phi_j K_j + B_j \Phi_{j+1} L_j = P_j$

(2.8b) $\qquad A_j F_{jj} + B_j F_{j,j+1} = I - P_j.$

Using (2.5) we obtain

(2.9) $\qquad B_j \Phi_{j+1}(L_j - K_j) = P_j$

and take

(2.10) $\qquad (L_j - K_j) = (B_j \Phi_{j+1})^+.$

Now we find from (2.4b) (N.B. $M_1 \Phi_1 = 1 - M_N \Phi_N$)

(2.11) $\qquad M_N \Phi_N(L_j - K_j) = -K_j\,,$

whence

(2.12) $\qquad K_j = -M_N \Phi_N(B_j \Phi_{j+1})^+.$

A similar derivation can be given for L_j. Summarizing we have

Property 2.13. Define $\bar{G}_{ij} = G_{ij} P_j$, then

(2.13a) $\qquad \bar{G}_{ij} = \Phi_i M_N \Phi_N(A_j \Phi_j)^+, \quad i \le j$

(2.13b) $\qquad \bar{G}_{ij} = \Phi_i M_1 \Phi_1(B_j \Phi_{j+1})^+, \quad i > j.$

Property 2.14.

(i) $\qquad \| \Phi_i M_N \Phi_N \Phi_j^+ \| \le \kappa \| A_j \|, \quad i \le j$

(ii) $\qquad \| \Phi_i M_1 \Phi_1 \Phi_{j+1}^+ \| \le \kappa \| B_j \|, \quad i \ge j + 1.$

Proof. First we show that if $A_j \Phi_j y = 0$ for some nontrivial vector y, then $\Phi_{j+1} y = 0$. Suppose the latter is not true; without restriction assume $\|y\|_2 = 1$. Let Y be an orthogonal matrix with y as its first column and $Y(\alpha)$ equal to Y but for its first column which is αy, $\alpha \in \mathbb{R}$. Define another matrix solution $\{\Psi_i(\alpha)\}$, where $\Psi_i(\alpha) = \Phi_i Y$, $i \le j$ and $\Psi_i(\alpha) = \Phi_i Y(\alpha)$, $i \ge j + 1$. (Note that in particular $A_j \Psi_j(\alpha) = -B_j \Psi_{j+1}(\alpha)$.) We have

$$M_1 \Psi_1(\alpha) + M_N \Psi_N(\alpha) = M_1 \Phi_1 Y + M_N \Phi_N Y(\alpha)$$

$$= Y + (\alpha - 1) M_N \Phi_N [y \mid 0 \cdots 0].$$

Hence there clearly exists a choice for $\alpha \neq 1$ such that $M_1 \Psi_1(\alpha) + M_N \Psi_N(\alpha)$ is nonsingular, showing that $\{\Psi_i(\alpha)\}$ is another fundamental solution. However, since $\Psi_1(\alpha) Y^{-1} = \Phi_1$ we can't have uniqueness unless $\Phi_i y = 0$, $i = j+1, \ldots, N$.

So for the rest of the proof we can therefore restrict ourselves to Range $(A_j \Phi_j (A_j \Phi_j)^+) = \text{Range}(B_j \Phi_{j+1} (B_j \Phi_{j+1})^+) =: S_j$, i.e. for (ii) we deduce from the foregoing that if $u \in \text{range}(\Phi_{j+1} \Phi_{j+1}^+) \Rightarrow B_j \Phi_{j+1} u \neq 0$. We can then proceed as follows

$$\|\Phi_i M_1 \Phi_1 (B_j \Phi_{j+1})^+\| = \max_{z \neq 0} \frac{\|\Phi_i M_1 \Phi_1 (B_j \Phi_{j+1})^+ z\|}{\|z\|}.$$

It is not restrictive to assume $B_j \Phi_{j+1} (B_j \Phi_{j+1})^+ z = z$, so with $(B_j \Phi_{j+1})^+ z =: y$, i.e. $B_j \Phi_{j+1} y = z$, we find

$$\|\Phi_i M_1 \Phi_1 (B_j \Phi_{j+1})^+\| = \max_{y \neq 0} \frac{\|\Phi_i M_1 \Phi_1 y\|}{\|B_j \Phi_{j+1} y\|} \geq \max_{y \neq 0} \frac{\|\Phi_i M_1 \Phi_1 y\|}{\|B_j\| \, \|\Phi_{j+1} y\|}.$$

Since $\Phi_{j+1} y =: u \neq 0$, we thus see that the last expression is underestimated by

$$\max_{u \neq 0} \frac{\|\Phi_i M_1 \Phi_1 \Phi_{j+1}^+ u\|}{\|B_j\| \, \|u\|} = \frac{\|\Phi_i M_1 \Phi_1 \Phi_{j+1}^+\|}{\|B_j\|},$$

from which (ii) follows. Estimate (i) goes similarly. $\quad\square$

Clearly, if all A_j and B_j are nonsingular, the expressions in (2.13) coincide with the standard form Green's function.

Next, we consider the complementary parts (see (2.7) and (2.8b)).

Property 2.15.

(i) $\{F_{ij}\}_{i \leq j}$ and $\{F_{ij}\}_{i \geq i+1}$ satisfy the homogeneous recursion of (2.1).

(ii) Range$(B_j F_{j,j+1}) \cap \text{Range}(A_j F_{jj}) = \{0\}$.

(iii) There exists an orthogonal projection Q_j such that

$$B_j F_{j,j+1} = (I - P_j) Q_j$$

$$A_j F_{jj} = (I - P_j)(I - Q_j).$$

Proof. (ii): If Range$(B_j F_{j,j+1}) \cap \text{Range}(A_j F_{jj}) \neq \{0\}$, then there exist vectors a and b such that $A_j F_{jj} a + B_j F_{j,j+1} b = 0$ with $A_j F_{jj} a \neq 0$. Now define $\{f_i\}$ by

$$f_i = F_{ij}\, a\,, \quad i \le j$$

$$f_i = F_{ij}\, b\,, \quad i \ge j+1.$$

Then clearly $\{f_i\}$ satisfies the homogeneous recursion for $i = j$ and on account of (i) also for $i \ne j$. So there exists a vector c such that $f_i = \Phi_i\, c$ for all i, and in particular for $i = j$. Hence $A_j \Phi_j c = A_j F_{jj} a \ne 0$. Since $A_j \Phi_j c \in \mathrm{Range}(P_j)$ and $(A_j F_{jj}) \in \mathrm{Range}(1 - P_j)$ this is clearly a contradiction

(iii) Define Q_j by $(I - P_j)\, Q_j = B_j F_{j,j+1} (B_j F_{j,j+1})^+$ (cf. (2.6)). $\Box$

Property 2.16. The splitting for G_{ij} in (2.7) is unique and hence $\bar{G}_{ij}$ is uniquely defined by (2.13a,b).

Proof. Clearly $\bar{G}_{ij}$ is uniquely defined. Now suppose $\{H_{ij}\}$ is defined as $\{F_{ij}\}$ in (2.7). Then $\{F_{ij} - H_{ij}\}$ satisfies the homogeneous recursion. Hence for some matrix S we have

$$F_{ij} - H_{ij} = \Phi_i\, S.$$

Since we must have $M_1 F_{1j} + M_N F_{Nj} = M_1 H_{1j} + M_N H_{Nj} = 0$ (see (2.4b)) we deduce $(M_1 \Phi_1 + M_N \Phi_N) S = 0 \Rightarrow S = 0$, so $F_{ij} = H_{ij}$. $\Box$

We find immediately:

Corollary 2.17. $\|\bar{G}_{ij}\|_2\,,\ \|F_{ij}\|_2 \le \|G_{ij}\|_2$.

3. Estimates for the fundamental solutions

In section 2 we derived expressions for the Green's function in terms of the fundamental solution $\{\Phi_i\}$ and certain *parasitic solutions* $\{F_{ij}\}_{j\,\text{fixed}}$. The latter kind of "solutions" are effectively one-sided solutions, i.e. built up from components which satisfy the homogeneous recursion for either $i > j$ or $i \leq j$; this is in contrast to the *fundamental solution* which exists globally (albe it with possible zero components on a range of consecutive indices).

We shall now show that well conditioning of the BVP (2.1), (2.2) implies a kind of dichotomy of $\{\Phi_i\}$ and a kind of one sided stability for $\{F_{ij}\}$. To do this we first have to generalise some results of [6]; not so much because our problem is discrete, but rather because it is singular.

First we examine the structure of $\{\Phi_i\}$ more in detail. Let $\text{rank}(\Phi_1) = n - l$ and $\text{rank}(\Phi_N) = n - m$. (N.B. $0 < l + m \leq n$.) Then there clearly exists an orthogonal V such that

$$(3.1a) \qquad \Phi_1 V = [* \mid \underset{\underset{l}{\longleftrightarrow}}{\varnothing}] \,,$$

i.e. the last l columns of $\Phi_1 V$ consist of zeros; consequently the last columns of $\Phi_N V$ must have full rank. However if $m > 0$, we may postmultiply V by an orthogonal matrix W,

$$(3.1b) \qquad W = \begin{bmatrix} \hat{W} & 0 \\ 0 & I_l \end{bmatrix}, \quad \hat{W} \in \mathbb{R}^{(n-l)^2} \,,$$

such that

$$(3.1c) \qquad \Phi_N V W = [\underset{\underset{m}{\longleftrightarrow}}{\varnothing} \mid *] \,.$$

It is therefore not restrictive to identify V and $V W$. Next, let U_1 be an orthogonal matrix such that

$$(3.2a) \qquad U_1^T \Phi_1 V = \begin{bmatrix} \Box & \varnothing \\ \varnothing & \varnothing \\ \varnothing & \varnothing & \varnothing \end{bmatrix} \begin{matrix} \updownarrow m \\ \\ \updownarrow l \end{matrix} =: \Psi_1$$

(i.e. Ψ_1 is upper triangular) and U_N an orthogonal matrix such that

$$(3.2b) \qquad U_N^T \Phi_N V = \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ \varnothing & & \varnothing \\ \varnothing & \Box & \end{bmatrix} \begin{matrix} \updownarrow m \\ \\ \updownarrow l \end{matrix} =: \Psi_N$$

(i.e. Ψ_N is lower triangular). Let Φ_1 be a special initial value defined by (cf. (3.2a))

$$(3.3a) \qquad \Phi_1 := \left\{ \Psi_1 + \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & I_l \end{bmatrix} \right\} V^T.$$

It is simple to verify that

$$(3.3b) \qquad \Phi_N \overline{\Phi}_1^{-1} = U_N^T \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ \varnothing & A & \varnothing \\ \varnothing & B & C \end{bmatrix} ,$$

for some $A \in I\!R^{(n-m-l)^2}$, $B \in I\!R^{(n-m)\times l}$ and $C \in I\!R^{l\times l}$, A and C being nonsingular. As a following-
ing step let

$$(3.4a) \qquad A =: \bar{U} \bar{\Sigma} \bar{V}^T ,$$

be the SVD of A, where $\bar{\Sigma}$ has positive elements $\sigma_1, \ldots, \sigma_{n-m-l}$. Assume $0 < \sigma_1, \ldots, \sigma_k \leq 1$ and
$\sigma_{k+1}, \ldots, \sigma_{n-m-l} > 1$, for some $k \in I\!N$. Like in [6] we split $\bar{\Sigma}$ in two parts (which effectively
corresponds to a separation of non-increasing and increasing modes *coupled* by the boundary
condition); so we introduce

$$(3.4b) \qquad D_2 = \text{diag}(\sigma_{k+1}, \ldots, \sigma_{n-m-l})$$

$$(3.4c) \qquad \bar{D} = \begin{bmatrix} I_k & \varnothing \\ \varnothing & D_2 \end{bmatrix} .$$

Inspired by this partitioning, let a projection P be given by

$$(3.5) \qquad P := \begin{bmatrix} I_{m+k} & \varnothing \\ \varnothing & \varnothing \end{bmatrix} .$$

Finally, let

$$(3.6) \qquad K := P + (I - P) \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ \varnothing & \bar{D} & \varnothing \\ \varnothing & B\bar{V} & C \end{bmatrix}$$

(note that K is nonsingular).
Then we define a fundamental solution $\{\bar{\Phi}_i\}$ by

$$(3.7a) \qquad \bar{\Phi}_i := \Phi_i \overline{\Phi}_1^{-1} X K^{-1}$$

where

$$(3.7b) \qquad X := \begin{bmatrix} I_m & \varnothing & \varnothing \\ \varnothing & \bar{V} & \varnothing \\ \varnothing & \varnothing & I_l \end{bmatrix} ,$$

and *separated* BC $\bar{M}_1, \bar{M}_N$ by

$$(3.8a) \qquad \bar{M}_1 := P X^T U_1^T$$

$$(3.8b) \qquad \bar{M}_N := (I - P) Y^T U_N^T$$

where

$$(3.8c) \qquad Y := \begin{bmatrix} I_m & \varnothing & \varnothing \\ \varnothing & \bar{U} & \varnothing \\ \varnothing & \varnothing & I_l \end{bmatrix}.$$

We have the following lemmata

Lemma 3.9.

(i) $\qquad \bar{M}_1 \, \tilde{\Phi}_1 + \bar{M}_N \, \tilde{\Phi}_N = I$

(ii) $\qquad \|\bar{M}_1\|_2 \le 1, \quad \|\bar{M}_N\|_2 \le 1$

(iii) $\qquad \|\tilde{\Phi}_1\|_2 \le 1, \quad \|\tilde{\Phi}_N\|_2 \le 1$

Proof. We have

$$\tilde{\Phi}_1 = \Phi_1 \bar{\Phi}_1^{-1} X K^{-1} = U_1 \begin{bmatrix} I_m & \varnothing & \varnothing \\ \varnothing & \bar{V} & \varnothing \\ \varnothing & \varnothing & \varnothing \end{bmatrix} \begin{bmatrix} I_m & \varnothing & \varnothing \\ \varnothing & \bar{D}^{-1} & \varnothing \\ \varnothing & \bar{E} & C^{-1} \end{bmatrix},$$

where $\bar{E} = -C^{-1} B \bar{V} \bar{D}^{-1}$.

Because of the construction of $\bar{D}^{-1}$ it directly follows that $\|\tilde{\Phi}_1\|_2 \le 1$. Furthermore we have

$$\tilde{\Phi}_N = \Phi_N \bar{\Phi}_1^{-1} X K^{-1} = U_N^T Y \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ \varnothing & \bar{\Sigma}\bar{D}^{-1} & \varnothing \\ \varnothing & \varnothing & I_l \end{bmatrix}.$$

Hence $\|\tilde{\Phi}_N\|_2 \le 1$. From these expressions of $\tilde{\Phi}_1$ and $\tilde{\Phi}_N$ the other assertions follow immediately.

$\square$

Lemma 3.10. $\|\tilde{\Phi}_i\| \le 2\kappa$.

Proof.

$$\tilde{\Phi}_i = \Phi_1 [M_1 \Phi_1 + M_N \Phi_N] [\bar{\Phi}_1^{-1} X K^{-1}] = \Phi_i M_1 \tilde{\Phi}_1 + \Phi_i M_N \tilde{\Phi}_N.$$

So the estimate follows from Lemma 3.9.

$\square$

Theorem 3.11. $\|\tilde{\Phi}_i P \tilde{\Phi}_j^\dagger\| \le (\kappa + 4\kappa^2)\gamma_j, \quad i \le j,$

$\qquad \|\tilde{\Phi}_i(1-P) \tilde{\Phi}_{j+1}^\dagger\| \le (\kappa + 4\kappa^2)\gamma_j, \quad i \ge j+1,$

$\qquad$ where $\gamma_j = \max(\|A_j\|, \|B_j\|)$.

Proof. We only show the result for $i \ge j + 1$. Let $S^{-1} := \bar{\Phi}_1^{-1} X K^{-1}$ (cf. (3.7a)) then $\tilde{\Phi}_{j+1}^\dagger = S M_1 \Phi_1 \Phi_{j+1}^\dagger + S M_N \Phi_N \Phi_{j+1}^\dagger$. So

$$\bar{\Phi}_i \bar{M}_1 \bar{\Phi}_1 \bar{\Phi}_{j+1}^+ = \bar{\Phi}_i (1 - \bar{M}_N \bar{\Phi}_N) S M_1 \Phi_1 \Phi_{j+1}^+ + \bar{\Phi}_i (\bar{M}_1 \bar{\Phi}_1) S M_N \Phi_N \bar{\Phi}_{j+1}^+$$

$$= [\Phi_i M_1 \Phi_1 \Phi_{j+1}^+] - \bar{\Phi}_i \bar{M}_N [\Phi_N M_1 \Phi_1 \Phi_{j+1}^+] + \bar{\Phi}_i \bar{M}_1 [\Phi_1 M_N \Phi_N \Phi_{j+1}^+].$$

The expressions in square brackets are each bounded in norm by $\kappa \gamma_j$. Since $\bar{M}_1 \bar{\Phi}_1 = P$, the result follows from using Lemma 3.9, Lemma 3.10 and Theorem 2.15.   ▯

We conclude from this theorem that there exists a "dichotomic" fundamental solution $\{\bar{\Phi}_i\}$ in the sense that basis modes of

(3.12a) $S_1 := \{\bar{\Phi}_i P c \mid c \in I\!\!R\}$

do not increase significantly, as long as they exist for increasing index, and similarly of

(3.12b) $S_2 := \{\bar{\Phi}_i (I - P) c \mid c \in I\!\!R\}$,

for decreasing index.
As far as parasitic modes (i.e. which do not exist till some point) is concerned, we deduce from Property 2.17

Theorem 3.13. For any index j for which $P_j \neq I$ (see (2.6)) we have

$$\|F_{ij}\|_2 \leq \kappa, \quad i > j$$

$$\|F_{ij}\|_2 \leq \kappa, \quad i \leq j.$$

So these parasitic modes do not grow, away from the point j from which they originate.

Remark 3.14. If the BC are separated we do not have to construct the fundamental solution $\bar{\Phi}$ as above. As in [6] the projection matrix P can then be identified with $M_1 \Phi_1$.

4. Application to special block tridiagonal matrices

Consider a discrete two point BVP with separated BC, i.e. in (2.1b) we have

$$(4.1) \qquad M_1 = \begin{bmatrix} M_1^2 \\ \varnothing \end{bmatrix} \updownarrow k \quad , \quad M_N = \begin{bmatrix} \varnothing \\ M_N^1 \end{bmatrix} \updownarrow n-k$$

A natural way of employing this zero structure (which apparently induces a decoupling) is to write the resulting system for $x_1, \dots, x_N$ as

$$(4.2a) \qquad \mathbf{A}\,\mathbf{x} = \mathbf{b} \,,$$

where

$$(4.2b) \qquad \mathbf{A} = \begin{bmatrix} M_1^2 & & & & \\ A_1 & B_1 & & & \varnothing \\ & A_2 & B_2 & & \\ & & \ddots & \ddots & \\ \varnothing & & & A_{N-1} & B_{N-1} \\ & & & & M_N^1 \end{bmatrix} \,,$$

$$(4.2b) \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix} \,, \quad \mathbf{b} = \begin{bmatrix} b^2 \\ f_1 \\ \vdots \\ f_{N-1} \\ f_{N-1} \\ b^1 \end{bmatrix} = \begin{bmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ b_N \end{bmatrix}$$

(here b^2, b^1 denote the first k and last $(n-k)$ coordinates of b). We can repartition the matrix A into a block tridiagonal one (with square blocks of size n). This results in a two step recursion

$$(4.3) \qquad S_i\, x_{i-1} + Q_i\, x_i + R_i\, x_{i+1} = b_i \,, \; i = 2, \dots, N-1$$

where (with obvious row partioning for A_i, B_i, C_i)

$$(4.4) \qquad S_i = \begin{bmatrix} A_{i-1}^2 \\ \varnothing \end{bmatrix} \updownarrow k \quad , \quad Q_i = \begin{bmatrix} B_{i-1}^2 \\ A_i^1 \end{bmatrix} \updownarrow k \quad , \quad R_i = \begin{bmatrix} \varnothing \\ B_i^1 \end{bmatrix} \updownarrow k \quad .$$

Moreover we obtain a (separated) BC

$$(4.5a) \qquad \begin{bmatrix} M_1^2 \\ A_1^1 \end{bmatrix} x_1 + \begin{bmatrix} \varnothing \\ B_1^1 \end{bmatrix} x_2 = \begin{bmatrix} b^2 \\ f^1 \end{bmatrix} \updownarrow k$$

(4.5b)
$$\begin{bmatrix} A_{N-1}^2 \\ \varnothing \end{bmatrix} x_{N-1} + \begin{bmatrix} B_{N-1}^2 \\ M_N^1 \end{bmatrix} x_N = \begin{bmatrix} f_{N-1}^2 \\ b^1 \end{bmatrix} \updownarrow k .$$

The recursion (4.3) was already discussed in [8] with respect to stability and dichotomy by rewriting it into one step form:

(4.6)
$$\begin{bmatrix} I_n & \varnothing \\ \varnothing & R_i \end{bmatrix} \begin{bmatrix} x_i \\ x_{i+1} \end{bmatrix} + \begin{bmatrix} \varnothing & -I_n \\ S_i & Q_i \end{bmatrix} \begin{bmatrix} x_{i-1} \\ x_i \end{bmatrix} + \begin{bmatrix} 0 \\ b_i \end{bmatrix} .$$

Clearly both matrices in this implicit recursion are singular; this caused some problems for application of the general theory, which were circumvented by an indirect analysis in [8].

With the more general results of the preceding sections we can tackle this problem more naturally now. Rather than (4.6) we consider a recursion of which we shall only analyse the essential, i.e. the homogeneous part:

(4.7a)
$$\begin{bmatrix} \varnothing & -I_n \\ A_{i-1}^2 & B_{i-1}^2 \\ \varnothing & \varnothing \end{bmatrix} \begin{bmatrix} x_{i-1} \\ x_i \end{bmatrix} + \begin{bmatrix} I_n & \varnothing \\ \varnothing & \varnothing \\ A_i^1 & B_i^1 \end{bmatrix} \begin{bmatrix} x_i \\ x_{i+1} \end{bmatrix} = 0, \quad 2 \le i \le N-1$$

(note that (4.7a) is directly found from (4.6) through left multiplication by a suitable nonsingular matrix and setting $b_i = 0$). The form (4.7a) is unbiassed with respect to the increase or decrease of the index sequence

As BC we have

(4.7b)
$$\begin{bmatrix} M_1^2 & \varnothing \\ A_1 & B_1^1 \\ \varnothing & \varnothing \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \varnothing & \varnothing \\ A_{N-1}^2 & B_{N-1}^2 \\ \varnothing & M_N^1 \end{bmatrix} \begin{bmatrix} x_{N-1} \\ x_N \end{bmatrix} = \begin{bmatrix} b^2 \\ f_1^1 \\ f_{N-1}^2 \\ b^1 \end{bmatrix} ,$$

formally written as

$$\hat{M}_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \hat{M}_{N-1} \begin{bmatrix} x_{N-1} \\ x_N \end{bmatrix} = \hat{b}.$$

In order to avoid unnecessary complication let us assume that the original recursion (2.1a) involves nonsingular matrices only. Let $\{\hat{\Phi}_i\}_{i=1}^{N-1}$ denote a fundamental solution of (4.7) with $\hat{M}_1 \hat{\Phi}_1 + \hat{M}_{N-1} \hat{\Phi}_{N-1} = I$, so $\hat{\Phi}_i, \hat{M}_1, \hat{M}_{N-1} = \mathbb{R}^{2n \times 2n}$.

From (2.3) we see that it is meaningful to define the projections

- 14 -

$$(4.8) \qquad P := M_1 \Phi_1 = \begin{bmatrix} I_k & \varnothing \\ \varnothing & \varnothing \end{bmatrix} , \quad 1 - P = M_N \Phi_N.$$

Since we here assume that (2.1) is nondefective we can write for the Green's function of (2.1), see (2.4)

$$(4.9) \qquad \begin{aligned} G_{ij} &= \Phi_i P (B_j \Phi_{j+1})^{-1}, \quad i \geq j+1 \\ G_{ij} &= \Phi_i (I-P)(A_j \Phi_j)^{-1} \quad i \leq j. \end{aligned}$$

We now have

Property 4.10. Let $P_1, P_N \in I\!R^{n^2}$ be such that $P_1 = \begin{bmatrix} I_k \\ \varnothing \end{bmatrix}$, $P_N = \begin{bmatrix} 0 \\ I_{n-k} \end{bmatrix}$.

Define $\hat\Phi_i$ as $\begin{bmatrix} \hat\Phi_i^{11} & \hat\Phi_i^{12} & \hat\Phi_i^{13} & \hat\Phi_i^{14} \\ \hat\Phi_i^{21} & \hat\Phi_i^{22} & \hat\Phi_i^{23} & \hat\Phi_i^{24} \end{bmatrix}$, $\hat\Phi_i^{11}, \hat\Phi_i^{21}, \hat\Phi_i^{13}, \hat\Phi_i^{23} \in I\!R^{n \times (n-k)}$.

Then for $2 \leq i \leq N-1$ we have

$$\hat\Phi_i^{2j} = \hat\Phi_i^{1j}, \quad j = 1, \dots, 4$$

$$\hat\Phi_i^{11} = \Phi_i P_1, \quad \hat\Phi_i^{12} = G_{i1} P_N, \quad \hat\Phi_i^{13} = G_{i,N-1} P_1, \quad \hat\Phi_i^{14} = \Phi_i P_N.$$

Proof. Write $\hat\Phi_i = \begin{bmatrix} K_i^{11} & K_i^{12} \\ K_i^{21} & K_i^{22} \end{bmatrix}$, with square blocks. For $2 \leq i \leq N-2$ we then see

$[K_{i+1}^{11} \ K_{i+1}^{12}] = [K_i^{21} \ K_i^{22}]$. It follows from the BC that the first k columns of $\{K_i^{11}\}$ and the last $(n-k)$ columns of $\{K_i^{12}\}$ form the (normalized) fundamental solution $\{\Phi_i\}$ of (2.1a), (4.1). Furthermore, since the last $(n-k)$ columns of $\{K_i^{11}\}$ satisfy homogeneous BC and the recursion, except for

$$[A_1^1 K_1^{11} + B_1^1 K_2^{11}] \begin{bmatrix} \varnothing \\ I_{n-k} \end{bmatrix} = I_{n-k} ,$$

it is clear that they are identical to the last $n-k$ columns of the Green's function $\{G_{i1}\}$. The argument for the first k columns of $\{K_i^{12}\}$ is similar. □

Note that for all j the projections like P_j in (2.6) are identical to $\hat P$, where

$$(4.11) \qquad \hat P = \begin{bmatrix} I_n & \varnothing \\ \varnothing & \varnothing \end{bmatrix}.$$

Property 4.12. Let $\{\hat G_{ij}\}$ denote the Green's function of (4.7), Then, using the notation of Property 2.13, i.e. $\bar G_{ij} = \hat G_{ij} \hat P$ (where $\hat G_{ij} \in I\!R^{2n \times 2n}$ now)

$$\bar{G}_{ij} = \begin{bmatrix} \Phi_i \ P \ \Phi_{j+1}^{-1} & \varnothing \\ \Phi_{i+1} \ P \ \Phi_{j+1}^{-1} & \varnothing \end{bmatrix} , \quad i \geq j+1$$

$$\bar{G}_{ij} = \begin{bmatrix} \Phi_i \ (I-P) \ \Phi_{j+1}^{-1} & \varnothing \\ \Phi_{i+1} \ (I-P) \ \Phi_{j+1}^{-1} & \varnothing \end{bmatrix} , \quad i \leq j.$$

Proof. Apparently, for $i > j$, we have $\bar{G}_{ij} = \hat{\Phi}_i \hat{P} (\hat{B}_j \hat{\Phi}_{j+1})^+$ (cf. (2.13b)), where

$$\hat{B}_j = \begin{bmatrix} I_n & \varnothing \\ \varnothing & \varnothing \\ A_{j+1}^1 & B_{j+1}^1 \end{bmatrix}.$$

Now $\hat{B}_j \hat{\Phi}_{j+1} = \hat{P} \hat{\Phi}_{j+1} = \Phi_{j+1} \begin{bmatrix} P_1 \ | K \, P_N \, | \, L \, P_1 \ | \, P_N \} \\ \varnothing \end{bmatrix}$, if $1 \leq j \leq N-2$, where

$K := P(A_1 \Phi_1)^{-1}$, $L := (I-P) (B_{N-1} \Phi_N)^{-1}$.

Hence $\hat{B}_j \hat{\Phi}_{j+1}$ is of the form $\Phi_{j+1} \begin{bmatrix} I_k \ L_1 & \varnothing & \varnothing \\ \varnothing & \varnothing & L_2 \ I_{n-k} \\ & \varnothing & \end{bmatrix} \updownarrow n$, for some $L_1 \in \mathbb{R}^{k \times (n-k)}$,

$L_2 \in \mathbb{R}^{(n-k) \times k}$.

So

$$(\hat{B}_j \hat{\Phi}_{j+1})^+ = \left[\begin{bmatrix} (I_k + L_1 L_1^T)^{-1} & \varnothing \\ L_1^T (I_k + L_1 L_1^T)^{-1} & \varnothing \\ \varnothing & L_2^T (I_{n-k} + L_2 L_2^T)^{-1} \\ \varnothing & (I_{n-k} + L_2 L_2^T)^{-1} \end{bmatrix} \Phi_{j+1}^{-1} \; \middle| \; \varnothing \right]$$
$$\underset{n}{\longleftrightarrow}$$

Substitution in the expression for $\bar{G}_{ij}$ proves the assertion for $i > j$, $2 \leq i \leq N-1$. A similar proof can be carried out for $i \leq j$, $2 \leq i \leq N-1$. Further inspection finally shows that for the boundary cases $\bar{G}_{Nj}$ and $\bar{G}_{1j}$ the assertion is true as well. □

It is convenient to associate to the $2n$-th order system (4.7) quantities as defined in section 2, but now provided with a cap to avoid confusion. Hence the Green's function $\hat{G}_{ij}$ (see Property 4.12) actually consists of two parts (cf. (2.7)ff), viz. $\hat{G}_{ij} \hat{P}_1 = \bar{G}_{ij}$ and $\hat{G}_{ij}(I-\hat{P}_1) =: \hat{F}_{ij}$. To the latter "parasitic part" we can associate projection matrices $\hat{Q}_1$ and $I-\hat{Q}_1$, defined for the larger system just like in Property 2.15(iii), realizing that they are index independent. The proof of the following Property about "parasitic solutions" is immediate.

Property 4.13.

(i)　The projection matrix $\hat{Q}_j$ is given by $\begin{bmatrix} \varnothing & \varnothing & \varnothing \\ \varnothing & I_k & \varnothing \\ \varnothing & \varnothing & \varnothing \end{bmatrix} \begin{matrix} \updownarrow n \\ \\ \updownarrow n \end{matrix}$.

(ii)　$F_i^j = \begin{bmatrix} \varnothing & G_{ij} \\ \varnothing & G_{i+1,j} \end{bmatrix}$.

<div align="right">☐</div>

Corollary 4.14. The Green's function for (4.7a,b) is given by

$$\begin{bmatrix} \Phi_i P \Phi_{j+1}^{-1} & G_{ij} \\ \Phi_{i+1} P \Phi_{j+1}^{-1} & G_{i+1j} \end{bmatrix} , \quad i \geq j+1$$

$$\begin{bmatrix} -\Phi_i(I-P)\Phi_{j+1}^{-1} & G_{ij} \\ -\Phi_i(I-P)\Phi_{j+1}^{-1} & G_{i+1j} \end{bmatrix} , \quad i \leq j.$$

Of course, the results of this section might also be obtained by manipulating directly on the matrix resulting from (4.7a,b) in relation to (4.2). However, the construction in this section demonstrates the validity of our analysis given in Section 3.

From Theorem 3.11 we can now see that the (inflated) one-step system (4.7a,b) has a dichotomic fundamental solution.

5. Decoupling and singular matrices

In this section we consider the BVP (2.1), (2.2), assuming that it is well-conditioned, but allowing any of the matrices involved to be singular.

A general decoupling method which employs orthogonal matrices was given in [8]. We first describe it here briefly: Let Q_0 be an orthogonal matrix. Then perform a QU-decomposition

$$(5.1a) \qquad A_1 Q_1 = R_1 U_1$$

(i.e. R_1 orthogonal and U_1 (block) upper triangular). Next, perform a "(block) UQ-decomposition" on $R_1^{-1} B_1$, i.e.

$$(5.1b) \qquad R_1^{-1} B_1 = V_1 Q_2^{-1}$$

where V_1 is (block) upper triangular. Initially one can take both U_1 and V_1 to be upper triangular. The block form follows from the construction in Theorem 5.6. We can continue this process giving at the i-th step

$$(5.2a) \qquad A_i Q_i =: R_i U_i$$

$$(5.2b) \qquad R_i^{-1} B_i =: V_i Q_{i+1}^{-1}$$

By defining

$$(5.3a) \qquad \overline{M}_1 := M_1 Q_1$$

$$(5.3b) \qquad \overline{M}_N := M_N Q_N$$

$$(5.4a) \qquad \overline{x}_i := Q_i^{-1} x_i$$

$$(5.4b) \qquad \overline{f}_i := R_i^{-1} f_i ,$$

we obtain an upper triangular recursion

$$(5.5) \qquad U_i \overline{x}_i + V_i \overline{x}_{i+1} = \overline{f}_i.$$

In order to employ the decoupling for separate computation of non-increasing and non-decreasing modes, we need appropriate blocks of U_i and V_i to be nonsingular. In the following theorem we shall show that a suitable permutation gives *block* upper triangular U_i and V_i, which allow for a globally meaningful decoupling.

Theorem 5.6. Let rank(A_i) = k_i and rank(B_i) = $n - l_i$. Then there exists an orthogonal matrix Q_1 such that U_i has zero lower right $(n-k_i) \times (n-k_i)$ blocks and V_i has zero upper left $l_i \times l_i$ blocks. Moreover $l_i \le k_i$ for all i and there exists a k, with $l_i \le k \le k_i$ for all such that the upper left $k \times k$ blocks of the U_i and the lower $(n-k) \times (n-k)$ blocks of the V_i are nonsingular.

Proof. We shall use induction. We can always choose the columns in Q_i such that the first k_1

columns are orthogonal to $\mathrm{Null}(A_1)$, so the last $(n-k_1)$ rows of U_1 are zero. Since the BVP is well-posed, i.e. the "multiple shooting" matrix is nonsingular it follows that the last $(n-k_1)$ rows of $R_1^{-1} B_1^{-1}$ must have full rank, so the lower right $(n-k_1) \times (n-k_1)$ block of V_1 is nonsingular. By a suitable choice of R, we can guarantee that V_1 has zeros on the first l_1 diagonal positions, where l_1 is such that $\mathrm{rank}(B_1) = n - l_1$.

Now let the statement be true for $i-1$. Due to the zero $l_{i-1} \times l_{i-1}$ left upper block of V_{i-1}, the corresponding block of U_i must be nonsingular, i.e. $k_i \ge l_{i-1}$. It is always possible to permute the last $n - l_{i-1}$ columns of B_{i-1} and A_i such that U_i, when singular, i.e. of $k_i \le n$, is such that its right lower $(n-k_i) \times (n-k_i)$ is zero; this affects only the orthogonal matrix Q_i: if the permutation P_i, say, then one should replace Q_i by $P_i Q_i$.
The zero block in U_i implies a nonsingular block at the same position in V_i hence $l_i \le k_i$. By permuting rows of A_i and B_i it can be assured that the left upper $(l_i \times l_i)$ block of V_i is zero. The existence of the integer k follows from the construction above. $\qquad \Box$

The proof of the preceding theorem actually gives a constructive method for finding a partioning integer k and $\{Q_i, R_i\}$. We can go a step further and construct Q_i such that it produces an appropriately ordered diagonal of the U_i and V_i as far as the nondisappearing parts of the fundamental solution of (5.5), $\{\bar{\Phi}_i\}_{i=1}^N$ say, are concerned.
Now let $\bar{k}$ be the largest possible integer such that for all i, the $\bar{k} \times \bar{k}$ left upper blocks of U_i are nonsingular (see Theorem 5.6).

Let us denote by $\bar{U}_i$ and $\bar{V}_i$ the $\bar{k} \times \bar{k}$ left upper blocks of U_i and V_i respectively. This induces the recursion

$$(5.7) \qquad \bar{U}_i \tilde{x}_i + \bar{V}_i \tilde{x}_{i+1} = 0.$$

Property 5.8. Let $\bar{l} := \max l_i$. Then there exists a matrix solution $\{\bar{\Psi}_i\}$ of (5.7), so $\bar{\Psi}_i \in \mathbb{R}^{\bar{k} \times \bar{k}}$ for all i and a projection $\bar{P}$, with $\mathrm{rank}(\bar{P}) \le \bar{k} - \bar{l}$ such that

$$\| \bar{\Psi}_i \bar{P} \bar{\Psi}_j^\dagger \| \le (\kappa + 4\kappa^2)\gamma_j, \quad i \le j$$

$$\| \bar{\Psi}_i (I - \bar{P}) \bar{\Psi}_{j+1}^\dagger \| \le (\kappa + 4\kappa^2)\gamma_j, \quad i \ge j+1,$$

where $\gamma_j = \max (\|A_j\|, \|B_j\|)$.

Proof. We use the orthogonal invariance of the 2-norm to see that the problem (5.3), (5.5) induces a matrix with the same bound for its inverse as (2.1); hence because of the orthogonality of the transformation $\{Q_i\}$ and $\{R_i\}$ the estimates essentially follow from Theorem 3.11, constructing $\bar{P}$ like P in (3.5). $\qquad \Box$

From Property 5.8 we deduce that we can choose the matrix Q_1 such that the nondecreasing modes, cf. $\{\bar{\Psi}_i \bar{P}\}$ in Property 5.8, appear before the nonincreasing ones in $\{\Psi_i\}$, i.e. $\bar{P}$ is of the

form diag $(0, 0, \ldots, 0, 1, \ldots, 1)$. This then finally gives us the result that there exists a stable decoupling, also in the rank deficient case.

6. Examples

In this section we give two-examples to demonstrate the preceding analysis

Example 6.1. Consider the ODE

$$(6.1) \qquad \frac{dx}{dt} = \begin{bmatrix} c^2\lambda + s^2\mu & cs(\mu-\lambda) \\ cs(\mu-\lambda) & s^2\lambda + c^2\mu \end{bmatrix} x + f(t),$$

where $c = \cos t$, $s = \sin t$ and f is chosen such that $x(t) = (\cos t, \sin t)$ is a solution. As BC we take

$$(6.2) \qquad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x(0) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x(\pi) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

For λ large positive and μ large negative this constitutes a stiff BVP, having layers at both ends of the interval $[0, \pi]$. Therefore we use the midpoint schema on a nonuniform interval: Writing (6.1) as

$$\frac{dx}{dt} = A(t)x + f,$$

we have at the grid point t_i

$$(6.3) \qquad [I - \tfrac{1}{2} h A_{i+\frac{1}{2}}] x_{i+1} = [I + \tfrac{1}{2} h A_{i+\frac{1}{2}}] x_i + h f_{i+\frac{1}{2}},$$

where $A_{i+\frac{1}{2}} := A(\tfrac{1}{2}(t_{i+1} + t_i))$.

As grid points we use $\frac{j}{10}, j = 0, \dots, 30$ and moreover in the left layer: $t = \frac{1}{1200}, \frac{1}{600}, \frac{1}{400}, \frac{1}{300}, \frac{1}{250}, \frac{1}{200}, \frac{1}{150}, \frac{1}{100}, \frac{1}{50}, \frac{1}{25}$ and in the right layer the points $t = \pi - \frac{j}{10}$, for $j = 0; \cdots, 30$.

In Table 6.1 we display some typical values for U_i and V_i

t	U_i		V_i	
0.04	1.4999	-0.0125	0.5000	$.6\ 10^{-18}$
	0	$.7\ 10^{-10}$	0	2.0000
0.10	1.9998	0.0225	$.9\ 10^{-16}$	-0.0225
	0	1.0001	0	2.9999
1.6	6.1809	-1.5710	4.1933	-3.0522
	0	8.7366	0	10.4927
2.9	6.1809	-1.5710	4.1933	-3.0522
	0	8.7366	0	10.4927
3.1316	1.9960	0.1097	$.9\ 10^{-15}$	-0.1100
	0	1.0020	0	2.9980
3.1333	1.4998	0.0225	0.5000	$.-4\ 10^{-15}$
	0	$.23\ 10^{-13}$	0	2.0000

Table 6.1.

For $t = 0.04$, 3.1333 we see that U_i is almost singular and similarly V_i for $t = 0.10$, 3.1316; latter points mark the end of the layers, and if the decoupling is done adequately this should not have a noticeable global effect (cf. [1]). Though, clearly, A is not constant, the matrices U_i and V_i appear to be almost constant (due to the fact that $h\lambda$, $|h\mu|$ are "large", a well-known by-product of di-stable methods, see [3]). It is worth noting that the dichotomy can be deduced from the quotient of diagonal elements of U_i and V_i: $|(U_i)_{22} / (V_i)_{22}| < 1$ and $|(V_i)_{11} / (U_i)_{11}| < 1$.

In tabel 6.2 we have displayed the discrete fundamental solution $\overline{\Phi}_i$ of the transformed recursion 5.5 for some values of t_i, and in the last column the absolute error vector (after transforming back).

t	$\overline{\Phi}_i$		error	
0	0.365 10^{-39}	0.344 10^{-2}	0	
	0	1.00	0.274 10^{-2}	
0.04	0.299 10^{-38}	-0.742 10^{-19}	0.639 10^{-4}	
	0	.495 10^{-17}	0.181 10^{-5}	
0.10	0.639 10^{-22}	-0.188 10^{-18}	0.121 10^{-2}	
	0	0.165 10^{-17}	0.195 10^{-4}	
1.6	0.447 10^{-19}	-0.263 10^{-19}	0.265 10^{-4}	
	0	0.728 10^{-19}	0.125 10^{-2}	
2.9	0.102 10^{-16}	-0.118 10^{-20}	0.195 10^{-2}	
	0	0.560 10^{-20}	0.205 10^{-4}	
3.1316	0.122	-0.126 10^{-22}	0.222 10^{-6}	
	0	0.837 10^{-21}	0.351 10^{-6}	
3.1333	0.367	0.776 10^{-37}	0.351 10^{-6}	
	0	0.965 10^{-35}	0.198 10^{-6}	
3.1416	1.000	0	0	
	0	0.128 10^{-35}	0.481 10^{-7}	

Table 6.2.

The table shows that there is a nice layer resolution indeed; moreover the global error is of a proper order. These computations are essentially based on employing the stable decoupling.

Example 6.2. Consider the second order DAE

(6.4) $E\ddot{x} = Ax + q$,

where E is the (constant, singular) matrix $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Introducing an additional variable y, we can write

(6.5) $\begin{cases} E\dot{x} = y \\ Ey = Ax + q. \end{cases}$

For this simple index 1 probleem, we can use the trapezoidal method with grid spacing h resulting in (cf. [5])

(6.6) $\quad A_i\,z_i + B_i\,z_{i+1} = b_i\,, \quad i = 0,\dots,N-1$

where

(6.7) $\quad A_i := \begin{bmatrix} -\dfrac{1}{h}E & -\dfrac{1}{2}E \\[2mm] \varnothing & \varnothing \end{bmatrix}, \quad B_i := \begin{bmatrix} -\dfrac{1}{h}E & -\dfrac{1}{2}E-\Gamma \\[2mm] -A & E \end{bmatrix},$

and $z_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$, with $x_i \doteq x(ih)$, $y_i \doteq y(ih)$.

Assume we have a BC for (6.4)

(6.8) $\quad M_a\,x(a) + M_b\,x(b) = c\,,$

where $M_a, M_b \in \mathbb{R}^{1\times2}$.

Then a complete set of BC for (6.6) is given by (cf. [5])

(6.9) $\quad M_0\,z_0 + M_N\,z_n = d\,,$

where

(6.10) $\quad M_0 := \begin{bmatrix} M_a & I-E \\ -A & E \end{bmatrix}, \quad M_N := \begin{bmatrix} M_b & \varnothing \\ \varnothing & \varnothing \end{bmatrix}, \quad d := \begin{bmatrix} c \\ 0 \\ q(a) \end{bmatrix}.$

The BVP (6.6), (6.9) has a unique solution. A typical block row in the matrix resulting from the set of equations for $\{z_i\}$ is given by

(6.11) $\quad \begin{bmatrix} \varnothing & \begin{array}{|cccc|cccc|} \hline -\dfrac{1}{h} & 0 & -\dfrac{1}{2} & 0 & -\dfrac{1}{h} & 0 & \dfrac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & * & * & 1 & 0 \\ 0 & 0 & 0 & 0 & * & * & 0 & 0 \\ \hline \end{array} & \varnothing \end{bmatrix}$

(where $*$ denotes an element of A). In view of the algorithm outlined in section 5, we note that whatever choice of Q_i is made, $R_i\,U_i$, so U_i, will have zeros in the last three rows. Hence we can take $R_i = I$. As for the matrix B_i we note that it is nonsingular if A is (unless h is such that $a_{11} + 2/h \doteq a_{12}\,a_{21}/a_{22}$). Hence we can take Q_{i+1} simply such that $B_i\,Q_{i+1}$ is upper triangular (by Householder's method e.g.). The resulting matrices U_i and V_i induce a simple decoupling: The dimension k of the nonsingular diagonal block is 1. From this it follows that the projection matrix $\bar{P}$ as in Property (5.8) must equal

$$\bar{P} = \begin{bmatrix} 1 & \varnothing \\ \varnothing & \varnothing \end{bmatrix}.$$

Hence it follows that the parasitic solutions generated by the Green's function are zero, except at the point where they arise. This then implies numerical stability of the decoupling if only the decoupled scalar recursion employing the first diagonal elements of U_i and V_i is used in a stable

direction. The existence of a stable direction follows directly from the well conditioning of the DAE (6.4), (6.8) (cf. [4]).

References

[1] U.M. Ascher, R.M.M. Mattheij, General Framework, Stability and Error Analysis for Numerical Stiff Boundary Value Methods, Num. Math 54 (1988), 355-372.

[2] U.M. Ascher, R.M.M. Mattheij, R.D. Russell, Numerical Solution of Boundary Value Problems for Ordinary Differential Equations, Prentice-Hall, 1988, Englewood-Cliffs.

[3] R. England, R.M.M. Mattheij, Boundary Value Problems and Dichotomic Stability, SIAM J. Numer. Anal. 25 (1988), 1037-1054.

[4] Griepentrog, R. März, Differential Algebraic Equations and their Numerical Treatment, Tubner, Reipzig (1986).

[5] G.H. Golub, C.F. van Loan, Matrix Computations, John Hopkins Univ. Press, Baltimorre, 1983.

[6] F.R. de Hoog, R.M.M. Mattheij, On dichotomy and well-conditioning in BVP, SIAM J. Numer. Anal. 24 (1987), 89-105.

[7] R.M.M. Mattheij, Estimates for the errors in the solutions of linear boundary value problems, due to perturbations, Computing 27 (1981), 299-318.

[8] R.M.M. Mattheij, Decoupling and stability of algorithms for boundary value problems, SIAM Review 27 (1984), 314-331.