

On the conditioning of multipoint and integral boundary value problems

Citation for published version (APA):

Hoog, de, F. R., & Mattheij, R. M. M. (1989). On the conditioning of multipoint and integral boundary value problems. *SIAM Journal on Mathematical Analysis*, 20(1), 200-214. <https://doi.org/10.1137/0520016>

DOI:

[10.1137/0520016](https://doi.org/10.1137/0520016)

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](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

ON THE CONDITIONING OF MULTIPOINT AND INTEGRAL BOUNDARY VALUE PROBLEMS*

F. R. DE HOOG† AND R. M. M. MATTHEIJ‡

Abstract. Linear multipoint boundary value problems are investigated from the point of view of the condition number and properties of the fundamental solution. It is found that when the condition number is not large, the solution space is polychotomic. On the other hand, if the solution space is polychotomic then there exist boundary conditions such that the associated boundary value problem is well conditioned.

Key words. boundary value problem, conditioning, Green function, integral conditions

AMS(MOS) subject classifications. 34B10, 65L10

1. Introduction. Consider a system of first-order ordinary differential equations

$$(1.1) \quad \mathcal{L}y := y' - Ay = f, \quad 0 < t < 1$$

where $A \in L_1^{n \times n}(0, 1)$ and $f \in L_1^n(0, 1)$. We are interested in the solution of (1.1) that satisfies the multipoint boundary condition (BC)

$$(1.2) \quad \mathcal{B}y := \sum_{i=1}^N B_i y(t_i) = b.$$

Here, $0 = t_1 < \dots < t_N = 1$ and the matrices $B_i \in \mathbb{R}^{n \times n}$, $k = 1, \dots, N$, have been scaled so that, for instance,

$$(1.3) \quad \sum_{i=1}^N B_i B_i^T = I.$$

The restriction $t_1 = 0$, $t_N = 1$ has been introduced for notational convenience and is not restrictive provided we allow for the possibility that $B_0 = 0$ and $B_N = 0$.

One of the simplest examples of a multipoint boundary value problem is that of a dynamical system with n states which are observed at different times. Further examples and a description of numerical schemes for the solution of such equations may be found in [12], [1], and [11].

From the theory of boundary value problems, (1.1), (1.2) has a unique solution if $\mathcal{B}Y$ is nonsingular for any fundamental solution Y of \mathcal{L} (see, for example, Keller [8]). In the sequel we assume this is the case. Then, given any fundamental solution Y of (1.1), we may write the solution of (1.1), (1.2) as

$$(1.4) \quad y(t) = \Phi(t)b + \int_0^1 G(t, s)f(s) ds, \quad 0 \leq t \leq 1$$

where

$$(1.5a) \quad \Phi(t) := Y(t)(\mathcal{B}Y)^{-1}$$

* Received by the editors February 27, 1987; accepted for publication (in revised form) May 3, 1988.

† CSIRO, Division of Mathematics and Statistics, P.O. Box 1965, Canberra ACT 2601, Australia.

‡ Department of Mathematics and Computing Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, the Netherlands.

and

$$(1.5b) \quad G(t, s) = \begin{cases} \Phi(t) \sum_{i=1}^k B_i \Phi(t_j) \Phi^{-1}(s), & t_k < s < t_{k+1}, \quad t > s, \\ -\Phi(t) \sum_{i=k+1}^N B_i \Phi(t_j) \Phi^{-1}(s), & t_k < s < t_{k+1}, \quad t < s. \end{cases}$$

The function G is the *Green function* associated with (1.1), (1.2).

We can now use (1.4) to examine the conditioning of (1.1), (1.2). Let $|\cdot|$ denote the usual Euclidean norm in \mathbb{R}^n and define

$$\|u\|_\infty := \sup_t |u(t)|, \quad u \in [L_\infty(0, 1)]^n,$$

$$\|u\|_1 = \int_0^1 |u(t)| dt, \quad u \in [L_1(0, 1)]^n.$$

Then it follows from (1.3) that

$$(1.6) \quad \|y\|_\infty \leq \beta |\mathcal{B}y| + \alpha \|\mathcal{L}y\|_1$$

where

$$(1.7a) \quad \alpha := \sup_{t,s} |G(t, s)|$$

and

$$(1.7b) \quad \beta := \sup_t |\Phi(t)|.$$

The quantities α, β defined by (1.7) serve quite well as condition numbers for the boundary value problem in the sense that they give a measure for the sensitivity of (1.1), (1.2) to changes in the data. Consequently, if α or β is large, we may expect to have difficulties in obtaining an accurate numerical approximation to the solution of (1.1), (1.2).

If α is of moderate size, the solution space of (1.1) has properties that can (and should) be used in the construction of algorithms for calculating an approximate solution of (1.1), (1.2). For the two-point case (i.e., $N = 2$), de Hoog and Mattheij [5], [6] have shown that the solution space is dichotomic when α is not too large. A dichotomic solution space (see § 4 for a more detailed discussion of dichotomy) essentially means that nonincreasing modes of the solution space can be controlled by boundary conditions imposed on the left while nondecreasing modes can be controlled by boundary conditions imposed on the right. This concept is the basis for algorithms using decoupling ideas (see, for example, [10], [11]). The aim of this paper is to generalize the results of [5], [6] to (1.1), (1.2) with $N \geq 2$. In this case the notion of dichotomy has to be generalized, and it turns out that, for well-conditioned problems, the solution space consists of modes that can be controlled at one of the points t_1, \dots, t_N (see § 4). This has allowed us to generalize the ideas of decoupling to multipoint problems, but that is discussed elsewhere [7].

In general we may say that if $N > n$ there is a redundancy in the number of conditions involved. It is therefore crucial to pick precisely n appropriate points from which modes are actually controlled by suitable conditions. It is quite natural to consider then a limit case of multipoint BC, viz., an integral condition (which incidentally generalizes two and multipoint conditions in an obvious way), so

$$(1.8) \quad \mathcal{B}y := \int_0^1 B(\tau)y(\tau) d\tau = b.$$

Such BC arise directly when L_p norms are used to scale the solution (possibly after linearization) as in eigenvalue problems.

We may treat the (discrete) multipoint case separately from (1.8). However, as it turns out, it is possible to construct a general mechanism that handles the integral BC as well. The price to be paid for this is that our proofs will be based on functional analytic arguments and thus are less constructive than could be given for the discrete case. The reward though is that we have been able to get sharp bounds in our estimates, sharpening even the bounds given for the two-point case in [6].

2. Notation and assumptions. In this section we review some basic results that we need later in our analysis. For some general references regarding Green functions we may consult, e.g., [2] and [9].

2.1. Boundary conditions and their normalization. Consider the general *boundary condition* (BC):

$$(2.1) \quad \mathcal{B}y = b$$

where \mathcal{B} is a bounded linear operator from $L_{1,1}^n(0, 1)$ to \mathbb{R}^n . Note that this includes the BC of type (1.2) and (1.8) as well. By $L_{1,1}^n(0, 1)$ we mean those functions the first derivative of which is in $L_1^n(0, 1)$. We introduce the norm

$$\|u\|_\infty = \max_{0 \leq t \leq 1} |u(t)|, \quad u \in L_{1,1}^n(0, 1)$$

where

$$|a| = \left(\sum_{i=1}^n a_i^2 \right)^{1/2}, \quad a \in \mathbb{R}^n.$$

For any $a \in \mathbb{R}^n$, $a^T \mathcal{B}$ is a linear functional from $L_{1,1}^n[0, 1]$ to \mathbb{R} . We define

$$\begin{aligned} \|a^T \mathcal{B}\|_\infty &:= \sup_{u \in L_{1,1}^n(0,1)} \frac{|a^T \mathcal{B}u|}{\|u\|_\infty}, \\ \rho_1(\mathcal{B}) &:= \max_{a \in \mathbb{R}^n} \frac{\|a^T \mathcal{B}\|_\infty}{|a|} = \|\mathcal{B}\|_\infty, \\ \rho_n(\mathcal{B}) &:= \min_{a \in \mathbb{R}^n} \frac{\|a^T \mathcal{B}\|_\infty}{|a|}. \end{aligned}$$

LEMMA 2.1. *Let $0 < \rho_1(\mathcal{B}) < \infty$. Then, there exists a matrix $C \in \mathbb{R}^{n \times n}$ such that*

$$\|C\mathcal{B}\|_\infty = \rho_1(C\mathcal{B}) = 1$$

and

$$\rho_n(C\mathcal{B}) \geq \frac{\rho_n(E\mathcal{B})}{\rho_1(E\mathcal{B})} \quad \forall E \in \mathbb{R}^{n \times n}.$$

Proof. If $\rho_n(\mathcal{B}) = 0$, then the result is trivial. We therefore assume $\rho_n(\mathcal{B}) > 0$ and let

$$\mathcal{D} = \{E \in \mathbb{R}^{n \times n} \mid \rho_1(E\mathcal{B}) = 1\}.$$

Since $\rho_n(E\mathcal{B})$ is continuous in E and \mathcal{D} is closed and bounded, it follows that there is a matrix $C \in \mathcal{D}$ such that

$$\rho_n(C\mathcal{B}) \geq \rho_n(E\mathcal{B}) \quad \forall E \in \mathcal{D}.$$

This is equivalent to the statement of the lemma. \square

This now gives us the possibility of scaling the BC, cf. (1.3), in a meaningful way.

Assumption 2.1. In the sequel, we shall assume that the BC (2.1) has been scaled so that

$$(2.2a) \quad \rho_1(\mathcal{B}) = \|\mathcal{B}\|_\infty = 1$$

and

$$(2.2b) \quad \rho_n(\mathcal{B}) \geq \rho_n(E\mathcal{B})/\rho_1(E\mathcal{B}) \quad \forall E \in \mathbb{R}^{n \times n}.$$

In addition to Assumption 2.1 we have the following assumption.

Assumption 2.2. Let (1.1), (2.1) have a solution for every $f \in L_1^n(0, 1)$ and $b \in \mathbb{R}^n$. Then, $\mathcal{B}Y \in \mathbb{R}^{n \times n}$ is nonsingular, where $Y \in L_{1,1}^{n \times n}(0, 1)$ is the solution of

$$(2.3a) \quad \mathcal{L}Y = 0, \quad Y(0) = F$$

and $F \in \mathbb{R}^{n \times n}$ is nonsingular.

On defining

$$(2.3b) \quad \Phi(t) := Y(t)(\mathcal{B}Y)^{-1},$$

we can write any function $y \in L_{1,1}^n(0, 1)$ as

$$(2.4) \quad y = \mathcal{P}y + (I - \mathcal{P})y = \mathcal{P}y + \mathcal{G}(\mathcal{L}y),$$

where

$$(2.5a) \quad \mathcal{P}y := \Phi(\mathcal{B}y),$$

$$(2.5b) \quad \mathcal{G}f := \int_0^1 \mathcal{G}(t, s)f(s) ds, \quad f \in L_1^n(0, 1)$$

and \mathcal{G} is the *Green function* defined by

$$(2.6a) \quad \mathcal{G}(t, s) = \Phi(t)\{H(t, s) - \mathcal{B}(\Phi H(\cdot, s))\}\Phi^{-1}(s)$$

with

$$(2.6b) \quad H(t, s) = \begin{cases} I, & t > s, \\ 0, & t < s \end{cases}$$

(cf. the special case (1.4), where \mathcal{B} is given by (1.2)).

Remark 2.1. The operator \mathcal{B} in the term $\mathcal{B}(\Phi H(\cdot, s))$ above should be interpreted as an extension of \mathcal{B} to an operator from $L_\infty^n(0, 1)$ to \mathbb{R}^n . Note however that a sensible extension of \mathcal{B} to $L_\infty^n(0, 1)$ is assured by the Hahn-Banach theorem.

Remark 2.2. \mathcal{P} is a projection from $L_{1,1}^n(0, 1)$ onto the solution space $\{Ya \mid a \in \mathbb{R}^n\}$. Given such a projection \mathcal{P} , we can define a linear operator

$$\mathcal{B} = CY^{-1}\mathcal{P}$$

where $C \in \mathbb{R}^{n \times n}$ is a scaling matrix chosen so that (1.1), (2.2a), and (2.2b) hold. Lemma 2.1 ensures the existence of such a matrix.

Remark 2.3. It is easy to verify that the Green function has the form

$$(2.7) \quad \mathcal{G}(t, s) = \begin{cases} Y(t)(I - E(s))Y^{-1}(s), & t > s, \\ -Y(t)(E(s))Y^{-1}(s), & t < s \end{cases}$$

where $E \in L_\infty^{n \times n}(0, 1)$. Conversely, given a function of the form (2.7), we have

$$\mathcal{L}\left\{\int_0^1 \mathcal{G}(\cdot, s)f(s) ds\right\} = f, \quad f \in L_1^n(0, 1).$$

In addition, if we define

$$(\mathcal{P}y)(t) := y(t) - \int_0^1 \mathcal{G}(t, s)(\mathcal{L}y)(s) ds,$$

then

$$\begin{aligned} (\mathcal{P}y)(t) &= y(t) - \int_0^t Y(t)Y^{-1}(s)(\mathcal{L}y)(s) ds + Y(t) \int_0^1 E(s)Y^{-1}(s)(\mathcal{L}y)(s) ds \\ &= Y(t) \left\{ Y^{-1}(0)y(0) + \int_0^1 E(s)Y^{-1}(s)(\mathcal{L}y)(s) ds \right\}. \end{aligned}$$

We can easily verify that \mathcal{P} is a projection. Thus, \mathcal{B} defined by

$$\mathcal{B}y := C \left\{ Y^{-1}(0)y(0) + \int_0^1 E(s)Y^{-1}(s)(\mathcal{L}y)(s) ds \right\},$$

where $C \in \mathbb{R}^{n \times n}$ is a scaling matrix chosen so that (2.2a), (2.2b) holds, gives a bounded linear operator for which \mathcal{G} is the associated Green function.

2.2. Auerbach’s lemma. Let \mathcal{V} be a normed linear space of dimension k with norm denoted by $\|\cdot\|$ and let \mathcal{V}^* be the space of all linear functionals from $\mathcal{V} \rightarrow \mathbb{R}$.

Define a norm on \mathcal{V}^* by

$$(2.8) \quad \|y^*\|^* = \sup_{x \in \mathcal{V}} \frac{y^*(x)}{\|x\|}, \quad y^* \in \mathcal{V}^*.$$

DEFINITION 2.1. A *boundary* of \mathcal{V} is any set

$$\mathcal{D} \subseteq \{y^* \in \mathcal{V}^* \mid \|y^*\|^* \leq 1\}$$

such that

$$\|x\| = \sup_{y^* \in \mathcal{D}} y^*(x) \quad \forall x \in \mathcal{V}.$$

LEMMA 2.2 (for Auerbach’s lemma see [4, Lemma 4]). If \mathcal{D} is a closed boundary of \mathcal{V} then there exist $y_i^* \in \mathcal{D}$, $y_j \in \mathcal{V}$; $i, j = 1, \dots, k$ such that

$$y_i^*(y_j) = \delta_{ij}, \quad \|y_i^*\|^* = 1, \quad \|y_j\| = 1, \quad i, j = 1, \dots, k.$$

Since $\{y^* \in \mathcal{V}^* \mid \|y^*\|^* \leq 1\}$ is a closed boundary, Corollary 2.1 follows immediately.

COROLLARY 2.1. There exist $y_i^* \in \mathcal{V}^*$, $y_j \in \mathcal{V}$; $i, j = 1, \dots, k$ such that

$$y_i^*(y_j) = \delta_{ij}, \quad \|y_i^*\|^* = 1, \quad \|y_j\| = 1, \quad i, j = 1, \dots, k.$$

3. Conditioning of differential equations. In this section we consider the relation between α and β and the effect of the normalization of the BC as in Assumption 2.1. Recall that for $y \in L_{1,1}^n(0, 1)$ (cf. (2.4))

$$y(t) = \Phi(t)\mathcal{B}y + \int_0^1 \mathcal{G}(t, s)(\mathcal{L}y)(s) ds.$$

Hence, on taking norms

$$\|y\|_\infty \leq \beta \|\mathcal{B}y\| + \alpha \|\mathcal{L}y\|_1$$

where

$$\beta = \|\Phi\|_\infty = \max_{a \in \mathbb{R}^n} \frac{\|\Phi a\|_\infty}{|a|}, \quad \alpha = \sup_{t,s} |G(t, s)|.$$

In addition to α and β , it is also useful to consider

$$\mathcal{P} := Y(\mathcal{B}Y)^{-1}\mathcal{B}.$$

LEMMA 3.1. $\rho_n(\mathcal{B})\beta \leq \|\mathcal{P}\|_\infty \leq \rho_1(\mathcal{B})\beta$.

Proof. The result follows immediately from the definition of $\rho_1(\mathcal{B})$ and $\rho_n(\mathcal{B})$. \square

LEMMA 3.2. Let $\hat{\mathcal{B}}$ be a linear operator from $L_{1,1}^n(0,1)$ to \mathbb{R}^n , and let $\hat{\alpha}$ be the constant associated with $\hat{\mathcal{B}}$ and the differential equation (1.1). Then,

$$\hat{\alpha} \leq (1 + \|\hat{\mathcal{P}}\|_\infty)\alpha, \quad \text{where } \hat{\mathcal{P}}Y = Y(\hat{\mathcal{B}}Y)^{-1}\hat{\mathcal{B}}Y.$$

Proof. Let

$$\hat{\Phi} := Y(\hat{\mathcal{B}}Y)^{-1} \quad \text{and} \quad \hat{\mathcal{G}}f := \int_0^1 \hat{\mathcal{G}}(\cdot, s)f(s) ds,$$

where $\hat{\mathcal{G}}$ is defined similarly to \mathcal{G} in (2.6a), i.e., \mathcal{B} replaced by $\hat{\mathcal{B}}$. Clearly, $\hat{\Phi} = Y(\hat{\mathcal{B}}Y)^{-1}$ and consequently $\hat{\mathcal{P}} = \hat{\Phi}\hat{\mathcal{P}}$. That is, $\hat{\mathcal{G}}f = (I - \hat{\mathcal{P}})\mathcal{G}f$, and hence

$$\|\hat{\mathcal{G}}f\|_\infty \leq (1 + \|\hat{\mathcal{P}}\|_\infty)\|\mathcal{G}f\|_\infty.$$

Thus, $\hat{\alpha} \leq (1 + \|\hat{\mathcal{P}}\|_\infty)\alpha$. \square

It is clear that the result of Lemmas 3.1 and 3.2 can be combined to give

$$\hat{\alpha} \leq (1 + \rho_1(\hat{\mathcal{B}})\hat{\beta})\alpha.$$

Since it has been assumed that (2.2a), (2.2b) hold, we obtain the estimate

$$(3.1) \quad \hat{\alpha} \leq (1 + \hat{\beta})\alpha.$$

Note, however, that α and $\|\mathcal{P}\|_\infty$ are independent of the scaling (2.2a), (2.2b) but that $\rho_1(\mathcal{B})$, $\rho_n(\mathcal{B})$, and β are not. Therefore we examine some of the ramifications of Assumption 2.1.

LEMMA 3.3. $\rho_n(\mathcal{B}) \geq n^{-1}$.

Proof. Let

$$\mathcal{V} = \{a^T\mathcal{B} \mid a \in \mathbb{R}^n\}.$$

That is, \mathcal{V} are the linear functionals of the form $a^T\mathcal{B}$. Since $\mathcal{B}\Phi = I$, $\dim(\mathcal{V}) = n$. For $\ell \in \mathcal{V}$, define

$$\|\ell\| = \sup_{y \in L_{1,1}^n(0,1)} \frac{(\ell y)}{\|y\|_\infty} = \|\ell\|_\infty.$$

\mathcal{V} equipped with the norm $\|\cdot\|$ is an n -dimensional normed space. From Auerbach's theorem (Corollary 2.1), there exist $\ell_j^* \in \mathcal{V}^*$, $\ell_i \in \mathcal{V}$; $i, j = 1, \dots, n$ such that

$$\ell_j^*(\ell_i) = \delta_{ij}, \quad \|\ell_j^*\|^* = \|\ell_i\| = 1, \quad i, j = 1, \dots, n.$$

Clearly, for some $E \in \mathbb{R}^{n \times n}$,

$$a^TE\mathcal{B} = \sum_{i=1}^n a_i \ell_i \quad \forall a = (a_1, \dots, a_n)^T \in \mathbb{R}^n.$$

Furthermore,

$$\begin{aligned} \|a^TE\mathcal{B}\|_\infty &= \left\| \sum_{i=1}^n a_i \ell_i \right\|_\infty \\ &\leq \frac{|\sum_{j=1}^n a_j \ell_j^*(\sum_{i=1}^n a_i \ell_i)|}{\|\sum_{j=1}^n a_j \ell_j^*\|^*} \\ &\leq \frac{|a|}{\sqrt{n}}. \end{aligned}$$

Thus, $\rho_n(E\mathcal{B}) \geq 1/\sqrt{n}$. In addition,

$$\begin{aligned} \|a^T E\mathcal{B}\|_\infty &= \left\| \sum_{i=1}^n a_i \ell_i \right\|_\infty \\ &\leq \sum_{i=1}^n |a_i| \|\ell_i\|_\infty \leq n^{1/2} |a|. \end{aligned}$$

Thus, $\rho_1(E\mathcal{B}) \leq n^{1/2}$, and hence from (2.2b)

$$\rho_n(\mathcal{B}) \geq \frac{\rho_n(E\mathcal{B})}{\rho_1(E\mathcal{B})} \geq n^{-1}.$$

For boundary conditions of the form (1.2) we can obtain somewhat sharper estimates.

LEMMA 3.4. *For \mathcal{B} given by (1.2) and satisfying (1.1), (2.1), we have*

$$\rho_n(\mathcal{B}) \geq N_1^{-1/2}$$

where N_1 is the number of nontrivial matrices B_i in (1.2).

Proof. Without loss of generality, we take $N_1 = N$

$$\begin{aligned} \|a^T E\mathcal{B}\|_\infty &= \sum_{i=1}^N |B_i^T E^T a| \\ &\leq N^{1/2} \left(a^T E \sum_{i=1}^N B_i B_i^T E^T a \right)^{1/2} \\ &\leq N^{1/2} \left| E \sum_{i=1}^N B_i B_i^T E^T \right|^{1/2} |a|. \end{aligned}$$

Thus, $\rho_1(E\mathcal{B}) \leq N^{1/2} |E \sum_{i=1}^N B_i B_i^T E^T|^{1/2}$. On the other hand,

$$\begin{aligned} \|a^T E\mathcal{B}\|_\infty &= \sum_{i=1}^N |B_i^T E^T a| \\ &\geq \left(a^T E \sum_{i=1}^N B_i B_i^T E^T a \right)^{1/2} \geq |a| / \left| \left(E \sum_{i=1}^N B_i B_i^T E^T \right)^{-1} \right|^{1/2}. \end{aligned}$$

Thus, $\rho_n(E\mathcal{B}) \geq 1 / \left(E \sum_{i=1}^N B_i B_i^T E^T \right)^{-1/2}$. Now if we take $E = \left(\sum_{i=1}^N B_i B_i^T \right)^{-1/2}$, then, from (2.2b), $\rho_n(\mathcal{B}) \geq \rho_n(E\mathcal{B}) / \rho_1(E\mathcal{B}) \geq N^{-1/2}$. \square

For an important class of boundary conditions, the bound in Lemma 3.4 is attained.

LEMMA 3.5. *Let \mathcal{B} be given by (1.2),*

$$\sum_{i=1}^N \text{rank}(B_i) = n$$

and N_1 be the number of nontrivial matrices B_i in (1.2). Then,

$$\frac{\rho_n(\mathcal{B})}{\rho_1(\mathcal{B})} \leq N_1^{-1/2}.$$

In addition, (2.2a), (2.2b) hold if and only if

$$\sum_{i=1}^N B_i B_i^T = N_1^{-1} I.$$

Proof. Let us assume without loss of generality that $N_1 = N$,

$$B_i^T B_i \eta_i = \sigma_i^2 \eta_i, \quad |\eta_i| = 1, \quad i = 1, \dots, N$$

and

$$w_1 = 1, \quad w_k = \text{sign} \left\{ \eta_k^T B_k^T \sum_{i=1}^{k-1} w_i B_i \eta_i \right\}, \quad k = 2, \dots, N.$$

Now,

$$\begin{aligned} \rho_1(\mathcal{B}) &= \max_a \left\{ \frac{\sum_{i=1}^N |a^T B_i|}{|a|} \right\} \cong \max_a \left\{ \frac{|\sum_{i=1}^N w_i a^T B_i \eta_i|}{|a|} \right\} \\ &= \left| \sum_{i=1}^N w_i B_i \eta_i \right| \cong \left(\sum_{i=1}^N \eta_i^T B_i^T B_i \eta_i \right)^{1/2} = \left(\sum_{i=1}^N \sigma_i^2 \right)^{1/2}. \end{aligned}$$

This result holds for all singular values σ_i , and we may therefore take $\sigma_i = |B_i|$. Then $\rho_1(\mathcal{B}) \cong (\sum_{i=1}^N |B_i|^2)^{1/2}$.

In addition, for $\sigma_k \neq 0$,

$$\begin{aligned} \rho_n(\mathcal{B}) &= \min_a \left\{ \frac{\sum_{i=1}^N |a^T B_i|}{|a|} \right\} = \sigma_k \min_a \left\{ \frac{\sum_{i=1}^N |a^T B_i|}{|a| |B_k \eta_k|} \right\} \\ &\leq \sigma_k \min_a \left\{ \frac{\sum_{i=1}^N |a^T B_i|}{|a^T B_k \eta_k|} \right\} = \sigma_k. \end{aligned}$$

Note that the last equality is not valid if $\sum \text{rank}(B_i) > n$. Nor is it valid for an arbitrary vector η_k . Thus,

$$\frac{\rho_n(\mathcal{B})}{\rho_1(\mathcal{B})} \leq \min_k \frac{\sigma_k}{(\sum_{i=1}^N |B_i|^2)^{1/2}} \leq N^{-1/2},$$

which proves the first part of the lemma.

Now let (2.2a), (2.2b) hold. From Lemma 3.4 and the result above

$$\sigma_k = N^{-1/2} \left(\sum_{i=1}^N |B_i|^2 \right)^{1/2}.$$

Since, σ_k is an arbitrary singular value, all the singular values are equal, and using (2.2a) we obtain that $\sum_{i=1}^N B_i B_i^T = N^{-1} I$.

Finally, let $\sum_{i=1}^N B_i B_i^T = N^{-1} I$. Then, as previously,

$$\rho_1(\mathcal{B}) \cong \left(\sum_{i=1}^N |B_i|^2 \right)^{1/2} = 1 \quad \text{and} \quad \rho_n(\mathcal{B}) \leq N^{1/2} \left| \sum_{i=1}^N B_i B_i^T \right|^{1/2} = 1.$$

Thus, $\rho_1(\mathcal{B}) = 1$. In addition, as in Lemma 3.4,

$$\rho_n(\mathcal{B}) \geq 1 / \left| \left(\sum_{i=1}^N B_i B_i^T \right)^{-1} \right|^{1/2} = N^{-1/2}$$

and since this is the best possible, (2.2b) holds. \square

We now have the tools to assess the condition numbers α, β . Let us consider in particular (1.1) and the multipoint BC (1.2),

$$\mathcal{B}y = \sum_{i=1}^N B_i y(t_i),$$

for which we have the following useful properties:

$$(3.2) \quad \Phi(t) B_i = G^+(t, t_i) - G^-(t, t_i), \quad i = 1, \dots, N,$$

where

$$(3.3a) \quad G^+(t, t_i) = \lim_{s \rightarrow t_i^+} G(t, s), \quad i = 1, \dots, N-1,$$

$$(3.3b) \quad G^-(t, t_i) = \lim_{s \rightarrow t_i^-} G(t, s), \quad i = 2, \dots, N,$$

$$(3.3c) \quad G^+(t, 1) = G^-(t, 0) = 0.$$

THEOREM 3.1. *For \mathcal{B} given by (2.1) and satisfying (2.2a), (2.2b), we have*

$$\beta \leq \frac{2N_1\alpha}{\rho_n(\mathcal{B})} \leq 2N_1\alpha \min(n, N^{1/2}),$$

where N_1 is the number of nontrivial matrices B_i in (3.2). If, in addition $\sum_{i=1}^N \text{rank}(B_i) = n$, then $\beta \leq 2N_1\alpha$.

Proof. Without loss of generality, we take $N_1 = N$. From (3.2), (3.3)

$$|\Phi(t)B_i| \leq 2\alpha,$$

and hence

$$\begin{aligned} |\Phi(t)| &\leq \left(\sum_{i=1}^N |\Phi(t)B_i|^2 \right)^{1/2} \left| \left(\sum_{i=1}^N B_i B_i^T \right)^{-1} \right|^{1/2} \\ &\leq 2\alpha N^{1/2} \left| \left(\sum_{i=1}^N B_i B_i^T \right)^{-1} \right|^{1/2}. \end{aligned}$$

The first result now follows from the inequality

$$\rho_n(\mathcal{B}) \leq N^{1/2} \left| \left(\sum_{i=1}^N B_i B_i^T \right)^{-1} \right|^{1/2}$$

and Lemmas 3.3 and 3.4.

However, if $\sum_{i=1}^N \text{rank}(B_i) = n$, it follows from Lemma 3.5 that $\left| \left(\sum_{i=1}^N B_i B_i^T \right)^{-1} \right|^{1/2} = GN^{1/2}$ and this establishes the second part of the theorem. \square

Thus, when \mathcal{B} is given by (2.1) and N is not too large, the single parameter α is a suitable measure of the conditioning of the problem. However, as $N \rightarrow \infty$ we cannot bound β in terms of α using the results of Theorem 3.1, which suggests that in general it is not possible to obtain such bounds. This is confirmed by the following example.

Example 3.1. Consider the problem

$$\mathcal{L}y = y' + ay, \quad a > 0.$$

$$\mathcal{B}y = \int_0^1 y(s) ds,$$

for which $\alpha = 1$, $\beta = a(1 - e^{-a})$ and $\rho_1(\mathcal{B}) = 1$. Clearly, β becomes unbounded as $a \rightarrow \infty$.

Thus, in general both α and β need to be addressed in a discussion of stability.

4. Polychotomy. For two-point boundary value problems (i.e., $N = 2$) it has become almost traditional to assume that the solution space

$$\mathcal{S}(t) = \{ \Phi(t)c \mid c \in \mathbb{R}^n \}$$

can be separated into a space

$$\mathcal{I}(t) = \{ \Phi(t)Pc \mid c \in \mathbb{R}^n \}, \quad P^2 = P$$

of nondecreasing solutions and a space

$$\mathcal{D}(t) = \{ \Phi(t)(I - P)c \mid c \in \mathbb{R}^n \}$$

of nonincreasing solutions. In addition, if neither $\mathcal{F}(t)$ nor $\mathcal{D}(t)$ is trivial (i.e., $P \neq 0, I$), it is usually assumed that the angle $0 < \eta(t) < \pi/2$ between $\mathcal{F}(t)$ and $\mathcal{D}(t)$, defined by

$$\cos \eta(t) = \max_{y_1 \in \mathcal{F}(t), y_2 \in \mathcal{D}(t)} \frac{|y_1^T y_2|}{|y_1| |y_2|}$$

is not too small. This has led to the following definition.

DEFINITION 4.1. The solution space is *dichotomic* if there exists a projector P and a constant κ such that

$$(4.1a) \quad |\Phi(t)P\Phi^{-1}(s)| < \kappa, \quad t > s,$$

$$(4.1b) \quad |\Phi(t)(I - P)\Phi^{-1}(s)| < \kappa, \quad t < s;$$

κ is called the *dichotomy constant*.

Although a projector always exists such that (4.1) is valid for some constant κ , we are primarily interested in the case when κ is of moderate size. In fact a more precise definition would involve the size of κ as well; we do not dwell on this, however. It turns out that dichotomy is intimately connected with the conditioning of two-point boundary value problems. Specifically, de Hoog and Mattheij [5], [6] have shown the following.

THEOREM 4.1. When $N = 2$, there exists a projector P such that (4.1) holds with $\kappa = \alpha + 4\alpha^2$. Alternatively, if (4.1) holds, then there exist matrices $B_1, B_2 \in \mathbb{R}^{n \times n}$ such that $\alpha \leq \kappa$.

Thus, if $N = 2$ and α is of moderate size, the solution space is dichotomic (i.e., κ is also of moderate size). Conversely, if the solution space is dichotomic, there is a two-point boundary value problem for which the condition number is not too large.

However, a well-conditioned multipoint problem does not necessarily have a dichotomic solution space as can be seen from Example 4.1.

Example 4.1. Consider the problem

$$y' + 2\lambda(t - \frac{1}{2})y = f, \quad \lambda > 0,$$

$$y(\frac{1}{2}) = 1.$$

For this example,

$$\Phi(t) = \exp(-\lambda(t - \frac{1}{2})^2),$$

$$y(t) = \Phi(t) + \int_{1/2}^t \Phi(t)\Phi^{-1}(s)f(s) ds,$$

and hence

$$\alpha = 1 \quad (\text{for all } \lambda).$$

Thus the problem is well conditioned but the fundamental solution now increases on the interval $0 < t < \frac{1}{2}$ and decreases on $\frac{1}{2} < t < 1$. Such behavior is quite common in multipoint problems. Indeed, the results of de Hoog and Mattheij [5], [6] can be used to show that there exist projectors $\hat{P}_i, i = 1, \dots, N - 1$ such that

$$|\Phi(t)\hat{P}_i\Phi^{-1}(s)| < \kappa, \quad t_i < s < t < t_{i+1},$$

$$|\Phi(t)(I - \hat{P}_i)\Phi^{-1}(s)| < \kappa, \quad t_i < t < s < t_{i+1},$$

where κ is of moderate size if α is not large. Thus, on each interval $t_i < t < t_{i+1}, i = 1, \dots, N - 1$ the solution space is dichotomic.

However, the examination of a number of well-conditioned multipoint problems has suggested that additional structure is present in the solution space. This leads to the following generalization of dichotomy.

DEFINITION 4.2. The solution space $\mathcal{S}(t)$ is *polychotomic* if, for some $M \in \mathbb{N}$, and $0 = x_1 \leq x_2 \leq \dots \leq x_M = 1$, there exist projectors P_k , $k = 1, \dots, M$ and a constant κ such that

$$\sum_{k=1}^M P_k = I, \quad P_i P_j = P_j P_i = \delta_{ij} P_j,$$

$$(4.2a) \quad \left| \Phi(t) \sum_{j=1}^k P_j \Phi^{-1}(s) \right| < \kappa, \quad x_k < s < x_{k+1}, \quad t > s,$$

$$(4.2b) \quad \left| \Phi(t) \sum_{j=k+1}^M P_j \Phi^{-1}(s) \right| < \kappa, \quad x_k < s < x_{k+1}, \quad t < s.$$

In § 5 we show that the concept of polychotomy is closely related to the conditioning of multipoint boundary value problems in the sense that κ will be of moderate size when α is not too large. It turns out that this relationship can be exploited in the construction of efficient numerical schemes for the solution of (1.1), (1.2); this is discussed in detail in [7].

5. Bounds for polychotomy. In this section we show how the condition number α can be used to obtain bounds for κ . Initially we consider separable boundary conditions.

5.1. Separable boundary conditions.

DEFINITION 5.1. The boundary condition (1.2) is called *separable* if

$$\sum_{i=1}^N \text{rank}(B_i) = n.$$

Thus for separable boundary conditions, the solution space consists of a number of modes each of which is controlled by a condition at one of the points when $\text{rank}(B_i) \neq 0$.

We shall see that when the boundary condition (1.2) is separable, the solution space is polychotomic with constant $\kappa = \alpha$. Before we can show this, however, some preliminary results are required.

LEMMA 5.1. If $C_k \in \mathbb{R}^{n \times n}$, $k = 1, \dots, N$

$$\sum_{k=1}^N C_k = I \quad \text{and} \quad \sum_{k=1}^N \text{rank}(C_k) = n,$$

then C_k , $k = 1, \dots, N$ are projectors (i.e., $C_i C_j = C_j C_i = \delta_{ij} C_j$).

Proof. The result follows from the arguments used in [6, Thm. 3.2]. □

LEMMA 5.2. For $E_k \in \mathbb{R}^{n \times n}$, $k = 1, \dots, N$, let

$$\sum_{k=1}^N E_k = I, \quad \sum_{k=1}^N \text{rank}(E_k) = n,$$

and define

$$\hat{G}(t, s) = \begin{cases} Y(t) \sum_{k=1}^i E_k Y^{-1}(s), & t_i < s < t_{i+1}, \quad t > s, \\ -Y(t) \sum_{k=i+1}^N E_k Y^{-1}(s), & t_i < s < t_{i+1}, \quad t < s, \end{cases}$$

where Y is a fundamental solution of (1.1). Then there exists a boundary condition

$$(5.1) \quad \hat{\mathcal{B}}y := \sum_{i=1}^N \hat{B}_i y(t_i)$$

satisfying $\text{rank}(\hat{B}_i) = \text{rank}(E_i)$ and

$$\sum_{i=1}^N \hat{B}_i \hat{B}_i^T = N_1^{-1} I$$

such that \hat{G} is the Green function associated with (1.1), (5.1) and N_1 is the number of nontrivial matrices E_i .

Proof. Consider the LQ^T decomposition

$$[E_1 Y^{-1}(t_1) | E_2 Y^{-1}(t_2) | \cdots | E_N Y^{-1}(t_N)] = LQ^T$$

where $L \in \mathbb{R}^{n \times n}$ is lower triangular and nonsingular and $Q \in \mathbb{R}^{(N+1)n \times n}$ is orthogonal (i.e., $Q^T Q = I$). Now define $\hat{B}_k \in \mathbb{R}^{n \times n}$, $k = 1, \dots, N$ by

$$[\hat{B}_1 | \hat{B}_2 | \cdots | \hat{B}_N] := N_1^{-1} Q^T.$$

If we define

$$\hat{\Phi}(t) := Y(t)(\hat{\mathcal{B}}Y)^{-1},$$

we see that $\hat{\Phi}(t) = Y(t)L$. Then it is easy to verify that \hat{G} is the Green function associated with (1.1), (5.1), viz.,

$$\hat{G}(t, s) = \begin{cases} \hat{\Phi}(t) \sum_{i=1}^K \hat{B}_i \Phi(t_i) \Phi^{-1}(s), & t > s, \\ -\hat{\Phi}(t) \sum_{i=k+1}^N \hat{B}_i \Phi(t_i) \Phi^{-1}(s), & t < s \end{cases}$$

can be identified with $\hat{G}(t, s)$. \square

The relationship between polychotomy and the condition number for separable boundary conditions is now straightforward. Specifically we have the following theorem.

THEOREM 5.1. *If the boundary condition (1.2), is separable, then the solution space is polychotomic with $\kappa \leq \alpha$.*

Conversely, if the solution space of (1.1) is polychotomic with constant κ , then there exists a separable boundary condition (1.2), satisfying Assumption 2.1, such that $\alpha \leq \kappa$.

Proof. If the boundary condition (1.2) is separable

$$\sum_{i=1}^N \text{rank}(B_i) = n$$

and

$$\sum_{i=1}^N B_i \Phi(t_i) = I \quad (\text{cf. (2.3b)}).$$

Thus

$$\sum_{i=1}^N \text{rank}(B_i \Phi(t_i)) = n$$

and from Lemma 5.1,

$$P_i = B_i \Phi(t_i), \quad i = 1, \dots, N$$

are projectors. On substituting for P_i in the Green function (1.5) and comparing the resulting expression with the definition of polychotomy (see Definition 5.1), we find that (4.2) holds with $\kappa = \alpha$, $M = N$ and $x_j = t_j$.

If on the other hand the solution is polychotomic, then

$$|G(t, s)| \leq \kappa$$

where

$$G(t, s) = \begin{cases} Y(t) \sum_{i=1}^k P_i Y^{-1}(s), & x_k < s < x_{k+1}, \quad t > s, \\ -Y(t) \sum_{i=k+1}^M P_i Y^{-1}(s), & x_k < s < x_{k+1}, \quad t < s \end{cases}$$

and

$$\sum_{i=1}^M P_i = I, \quad P_i P_j = P_j P_i = \delta_{ij} P_j.$$

But from Lemmas 5.2 and 3.5 there exists a separable boundary condition of the form (1.2) which satisfies Assumption 2.1 and is such that G is the Green function associated with (1.1), (1.2) when $N = M$ and $t_i = x_i$. \square

5.2. General boundary condition. We again turn to the general BC (2.1) and show how we can select appropriate separable BC from them; this is based on the theory given in § 2.

Let

$$\mathcal{S} = \{Ya \mid a \in \mathbb{R}^n\}$$

with

$$\|y\| = \|y\|_\infty, \quad y \in \mathcal{S}.$$

Clearly, \mathcal{S} equipped with the norm $\|\cdot\|$ is a normed space of dimension n . In addition,

$$\mathcal{D} = \{y^* \in \mathcal{S}^* \mid y^*(y) = c^T y(t), \quad |c| = 1, \quad 0 \leq t \leq 1\}$$

is a closed boundary for \mathcal{S} . Hence, from Auerbach's lemma (Lemma 2.2) there exist $y_j^* \in \mathcal{D}$, $y_i \in \mathcal{S}$; $i, j = 1, \dots, n$ such that

$$y_j^*(y_i) = \delta_{ij}, \quad \|y_j^*\|^* = 1, \quad \|y_i\|_\infty = 1, \quad i, j = 1, \dots, n.$$

That is, there exist $c_j \in \mathbb{R}^n$, $|c_j| = 1$, points t_j with $0 \leq t_j \leq 1$, $j = 1, \dots, n$ and $y_i \in \mathcal{S}$, $i = 1, \dots, n$ such that

$$(5.2) \quad c_j^T y_i(t_j) = \delta_{ij}, \quad |c_j| = \|y_i\|_\infty = 1, \quad i, j = 1, \dots, n.$$

Furthermore,

$$c_j = y_j(t_j),$$

and hence

$$(5.3) \quad c_i^T c_j = 0 \quad \text{if } i \neq j \text{ and } t_i = t_j.$$

Let

$$(\hat{\mathcal{P}}y)(t) := \sum_{i=1}^n y_i(t) c_i^T y(t_i).$$

Thus,

$$\begin{aligned} \|\hat{\mathcal{P}}y\|_\infty &\leq \sum_{i=1}^n \|y_i\|_\infty \|y\|_\infty \\ &\leq n \|y\|_\infty. \end{aligned}$$

Hence

$$\|\hat{\mathcal{P}}\|_\infty \leq n$$

and, as in Lemma 3.2, we find that

$$\begin{aligned} \hat{\alpha} &\leq (1 + \|\hat{\mathcal{P}}\|_\infty)\alpha \\ &\leq (n + 1)\alpha. \end{aligned}$$

In addition, we have

$$\hat{\mathcal{B}}\hat{\Phi} = I$$

where

$$(5.4) \quad \hat{\Phi} = N_1^{1/2}[|y_1| \cdots |y_n|],$$

$$(5.5) \quad \hat{\mathcal{B}}y := \sum_{i=1}^N \hat{\mathcal{B}}y(t_i),$$

$$B_k = N_1^{-1/2} \begin{bmatrix} 0 \\ C_k \\ 0 \end{bmatrix} \leftarrow k\text{th position},$$

and N_1 is the number of *distinct points* in the set $\{t_k\}$. From (5.2), (5.3)

$$\sum_{k=1}^n \hat{B}_k \hat{B}_k^T = N_1^{-1}I,$$

and hence from Lemma 3.5, the boundary condition \tilde{B} defined by (5.5), which is clearly separable, satisfies (2.2a), (2.2b). Finally from (5.2), (5.5)

$$|\hat{\Phi}(t)| \leq N_1^{1/2}n^{1/2}.$$

Thus, we have shown the following theorem.

THEOREM 5.2. *For a general BC (2.1) we can construct a separable BC $\hat{\mathcal{B}}$ of the form $\hat{\mathcal{B}}y := \sum_{i=1}^n \hat{B}_i y(t_i)$, with $t_i \in [0, 1]$, such that $\hat{\mathcal{B}}$ satisfies (2.2a) and (2.2b) and for which (cf. (1.7))*

$$\hat{\beta} := \sup_t |\hat{\Phi}(t)| \leq n, \quad \hat{\alpha} := \sup_{s,t} |\hat{G}(s, t)| \leq (n + 1)\alpha.$$

COROLLARY 5.1. *If the BVP (1.1), (2.1) has a condition number α , then the solution space is polychotomic with*

$$\kappa \leq (n + 1)\alpha.$$

Note that the result of this corollary is somewhat different from Theorem 3.16 of [6], where bounds are derived $\sim \alpha^2$ for the two-point case. For large α we may therefore say that this more general result is sharper, though not constructive.

REFERENCES

[1] R. P. AGARWAL, *The numerical solution of multipoint boundary value problems*. J. Comp. Appl. Math., 5 (1979), pp. 17-24.
 [2] F. V. ATKINSON, *Discrete and Continuous Boundary Problems*, Academic Press, New York, 1964.
 [3] C. DE BOOR AND H.-O. KREISS, *On the condition of the linear system associated with discretized BVPs of ODEs*, SIAM J. Numer. Anal., 23 (1985), pp. 936-939.

- [4] E. W. CHENEY AND K. H. PRICE, *Minimal projections in approximation theory*, in *Approximation Theory*, A. Talbot, ed., Academic Press, New York, 1970, pp. 261–289.
- [5] F. R. DE HOOG AND R. M. M. MATTHEIJ, *The role of conditioning in shooting techniques*, in *Numerical Boundary Value ODEs*, U. Ascher and R. Russell, eds., Birkhäuser, Boston, 1985, pp. 21–54.
- [6] ———, *On dichotomy and well conditioning in BVP*, *SIAM J. Numer. Anal.*, 24 (1987), pp. 89–105.
- [7] ———, *An algorithm for solving multipoint boundary value problems*, *Computing*, 38 (1987), pp. 219–234.
- [8] H. B. KELLER, *Numerical Solution of Two-Point Boundary Value Problems*, CBMS-NSF Regional Conference Series in Applied Mathematics 24, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1976.
- [9] W. S. LOUD, *Generalized inverses and generalized Green functions*, *J. Soc. Indust. Appl. Math.*, 14 (1966), pp. 342–369.
- [10] R. M. M. MATTHEIJ, *Decoupling and stability of algorithms for boundary value problems*, *SIAM Rev.*, 27 (1985), pp. 1–44.
- [11] R. M. M. MATTHEIJ AND G. W. M. STAARINK, *An efficient algorithm for solving general linear two-point BVP*, *SIAM J. Sci. Statist. Comput.*, 5 (1984), pp. 745–763.
- [12] W. WELSH AND T. OJIKI, *Multipoint boundary value problems with discontinuities*, I. *Algorithms and Applications*, *J. Comp. Appl. Math.*, 6 (1980), pp. 133–143.