Reducing a singular linear two point boundary value problem to a regular one by means of Riccati transformations

Citation for published version (APA):

Document status and date:
Published: 01/01/1983

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.

Download date: 03. Aug. 2019
Reducing a singular linear two point boundary value problem to a regular one by means of Riccati transformations

by

Paul van Loon

EUT-Report 83-WSK-03

May 1983
Abstract

In this note a method for solving a two point boundary value problem in the interval $(0,1)$ for a linear first order system of ordinary differential equations with a singularity of the first kind in $t = 0$ is examined. The idea is to replace it by a regular problem on some subinterval $(0,\delta)$. To this end a singular initial value problem of Riccati type on $(0,\delta]$ has to be solved. This initial value problem is such that the spectrum of the corresponding Jacobian at $t = 0$ is in the closed left halfplane $\mathbb{C}^-$. Moreover, if the coefficients of the differential equation are analytic, so is the solution.
§1. Introduction

Let \( x : [0,1] \to \mathbb{R}^n \) satisfy the differential equation (DE)

\[
(1.1) \quad tx'(t) = A(t)x(t) + f(t), \quad t \in (0,1),
\]

with \( A : [0,1] \to \mathbb{R}^{n \times n} \) and \( f : [0,1] \to \mathbb{R}^n \) continuous in \([0,1]\) and analytic at \( t = 0 \), and the boundary conditions

\[
(1.2a) \quad x(0) = \lim_{t \to 0^+} x(t) \text{ exists}, \\
(1.2b) \quad B_0^0 x(0) + B_1^1 x(1) = b,
\]

where \( B_0^0, B_1^1 \in \mathbb{R}^{s \times n} \) such that \( \text{rank}(B_0^0, B_1^1) = s \) and \( b \in \mathbb{R}^s \) (s will be specified later).

The singular boundary value problem (BVP) \((1.1)\) and \((1.2)\) will be replaced by a regular BVP on a subinterval \((0,\delta)\) \( (\delta > 0)\). To this end some auxiliary singular initial value problems (IVP's) on \((0,\delta]\) have to be solved. It will be seen that these IVP's have two nice properties:

(i) at \( t = 0 \) the spectra of the corresponding Jacobians lie in the closed left halfplane \( \mathbb{C}^- \)

(ii) they have solutions which are analytic at \( t = 0 \).

Ad (i):

Observe that for DE's of the form

\[
(1.3) \quad ty'(t) = f(t,y(t)), \quad (t \in (0,1)),
\]

By analytic functions at \( t = 0 \) we will mean: real functions that, in \( \mathbb{C} \), can be expanded around \( t = 0 \) in power series, having a positive radius of convergence. Sometimes this radius of convergence will be given explicitly.
numerical methods with reasonable stepsize $h > 0$ that are numerically stable can be found if

$$
\delta \left( \frac{\partial f}{\partial y}(t,y(t)) \right) \leq \mathcal{C}^{-} \quad \text{for all } t \in (0,1)
$$

(Lambert, [7]), even if the equation (1.3) is (partly) stiff.

If $\frac{\partial f}{\partial y}(t,y(t))$ has also eigenvalues on the imaginary axis, then stiffly stable predictor corrector methods can be found (Gear, [1]).

In the latter case, absolute stability is required only for the fastly decaying solutions, while accuracy and relative stability are required for the other solutions. Using these methods, the real parts of the eigenvalues of the Jacobian may even be small positive. Since the interval of integration is not a priori determined, one may choose $\delta$ such that this condition is indeed satisfied in the whole interval $[0,\delta]$.

In practical applications one often finds that

$$\delta \left( \frac{\partial f}{\partial y}(0,y_0) \right) \leq \mathcal{C}^{-},
$$

which by continuity implies that there exists a large scope of numerical methods and a $\delta > 0$ such that the solution of (1.3) can be computed numerically stable on the whole interval $(0,\delta]$.

Ad (ii):

The analyticity of a solution at $t = 0$ is an important property for starting a solution method, since it implies that all derivatives at $t = 0$ exist. Moreover, most numerical integration methods are based on the assumption that the solution behaves like a polynomial, and consequently perform well if so. Hence, for small $t$, the solution can simply be computed by formal power series expansion, and for other values of $t$, stable numerical methods are available.
The fact that it gives solutions of the IVP's which are analytic at \( t = 0 \) is not obvious, since for instance the fundamental matrices belonging to a linear IVP may contain logarithmic singularities at \( t = 0 \).

§2. Preliminaries

For small \( t \) we have the power series

\[
(2.1) \quad A(t) = \sum_{k=0}^{\infty} A^k t^k \quad \text{and} \quad f(t) = \sum_{k=0}^{\infty} f^k t^k .
\]

Without restriction we may assume that \( A^0 \) takes the form

\[
(2.2) \quad A^0 = \begin{bmatrix}
A^+ & \cdots & \cdots \\
\vdots & \ddots & \vdots \\
A^0_{pq} & \cdots & A^0_{qq} \\
\vdots & \cdots & \vdots \\
A^- & \cdots & \cdots
\end{bmatrix}
\]

with \( \sigma(A^+) \subseteq \mathbb{C}^+ \), \( \sigma(A^-) \subseteq \mathbb{C}^- \), \( \sigma(A^0_{pq}) \subseteq i\mathbb{R} \) and \( p = \dim(\ker(A^0_{pq})) \), which implies that

\[
\text{rank} \begin{bmatrix} A^0_{pq} \\ A^0_{qq} \end{bmatrix} = q .
\]

Notation:

Let \( x \in \mathbb{R}^n \), then we will use two kind of partitions:

\[
x = (x_k^T, x_p^T, x_q^T, x_m^T)^T , \quad \text{where} \quad x_k \in \mathbb{R}^k \quad \text{etc.}
\]
and
\[
x = (x_1^T, x_2^T)^T,
\]
where \( x_1 \in \mathbb{R}^{k \times p} \) and \( x_2 \in \mathbb{R}^{q \times m} \).

In accordance with this last partition let
\[
A(t) = \begin{bmatrix}
A_{11}(t) & A_{12}(t) \\
A_{21}(t) & A_{22}(t)
\end{bmatrix}
\]
\[
\uparrow \quad k+p
\]
\[
\downarrow q+m
\]

First we remark that a part of \( x(0) \) can be computed directly.

**Theorem 2.1.** Let \( x(t) \) satisfy (1.1) and (1.2a). Then

(2.3)
\[
A^0 x(0) + f^0 = 0,
\]
where \( A^0 \) and \( f^0 \) are defined in (2.1).

**Proof.** Suppose \( A^0 x(0) + f^0 \neq 0 \).

From the differential equation (1.1) and the continuity of the solution we obtain

(2.4)
\[
t x'(t) = A^0 x(0) + f^0 + \varepsilon(t), \quad t \to 0,
\]
where \( \varepsilon(t) = o(1), \quad t \to 0. \)

Integration of (2.4) leads to

(2.5)
\[
x(t) - x(t_0) = \log \left( \frac{t}{t_0} \right) [A^0 x(0) + f^0] + \int_{t_0}^{t} \frac{\varepsilon(t)}{t} \, dt, \quad t_0 > 0.
\]

For \( t \) and \( t_0 \) sufficiently small we have, since \( \varepsilon(t) = o(1) \),

(2.6)
\[
\| x(t) - x(t_0) \| \geq \frac{1}{2} \| A^0 x(0) + f^0 \| \cdot | \log \left( \frac{t}{t_0} \right) |.
\]

Hence,
\[
\lim_{t \to 0} \| x(t) - x(t_0) \| \text{ does not exist},
\]
which is in contradiction with (1.2a). \( \square \)
Corollary 2.2. A necessary condition for the existence of solutions of (1.1) subject to (1.2a) is

\[
\begin{bmatrix}
f_0^0 \\ p \\ f_0^0 \\ q \\
\end{bmatrix} \in \mathbb{R} \begin{pmatrix}
0 \\ A_{pq} \\ 0 \\ A_{qq}
\end{pmatrix}.
\]

Observing that (2.3) can be written as

\[
\begin{pmatrix}
A^+ \\
0 \\
A_{pq} \\
0 \\
A_{qq} \\
A^-
\end{pmatrix}
\begin{pmatrix}
x_k(0) \\ x_q(0) \\ x_m(0)
\end{pmatrix} + \begin{pmatrix}
f_k^0 \\ f_p^0 \\ f_q^0 \\ f_m^0
\end{pmatrix} = 0,
\]

we see that \((x_k(0)^T, x_q(0)^T, x_m(0)^T)^T\) is uniquely defined by the boundary condition (1.2a), since

\[
\text{rank} \begin{pmatrix}
A^+ \\
0 \\
A_{pq} \\
0 \\
A_{qq} \\
A^-
\end{pmatrix} = k + q + m.
\]

Hence, if the ODE (1.1) is homogeneous then only \(x_p(0)\) may be nonzero.

Theorem 2.3. Let \(V\) be the subspace of \(C^1(0,1]\) formed by all solutions of

(2.7) \(tx'(t) = A(t)x(t), \quad t \in (0,1)\),

for which \(\lim_{t \to 0} x(t)\) exists.

Then \(\dim V = k + p\).
Proof. By Theorem 2.1 and the variation of constants formula (Henrici [3], p. 199) we find that any element \( \phi \in V \) must satisfy

\[
\phi_1(t) = \left( \frac{t}{t_0} \right)^A \xi + \int_{t_0}^{t} \left( \frac{t}{\tau} \right)^A \frac{\phi_2(\tau)}{\tau} + \left[ A_{11}^* (\tau) A_{12}^* (\tau) \right] \phi(\tau) d\tau
\]

(2.8)

\[
\phi_2(t) = \int_{0}^{t} \left( \frac{t}{\tau} \right)^{A_{22}} \frac{[A_{21}^* (\tau) A_{22}^* (\tau)] \phi(\tau)}{\tau} d\tau
\]

where \( \xi \in \mathbb{R}^{k+p} \), \( t_0 \in (0,1] \) fixed and \( A^* (t) = A(t) - A^0 \). Existence of a solution \( \phi \) of (2.8) for any \( \xi \in \mathbb{R}^{k+p} \) can be proven by the method of successive approximations, starting with \( \phi^0 = 0 \). At the same time we obtain that there exist constants \( C_1, C_2 > 0 \) such that, for \( t \) sufficiently small, \( \| \phi_1(t) \| \leq C_1 \) and \( \| \phi_2(t) \| \leq C_2 \cdot |t \log t| \). For other values of \( t \) the boundedness of \( \phi \) is obvious which implies that \( \phi \in V \).

Moreover, if \( \xi = 0 \) then \( \phi = 0 \), which implies that any solution \( \phi \) of (2.8) is uniquely determined by \( \xi \). Thus, by the linearity of (2.7) and (2.8), we have obtained that \( \dim V = k + p \).

By Theorem 2.3 we see that the boundary condition (1.2a) imposes \( (q+m) \) restrictions on the solution. In order to have a unique solution of (1.1) and (1.2) it is therefore necessary that \( s = k + p \) (cf. de Hoog, Weiss [4]).

To conclude this section, we investigate the behaviour of a solution of a special type of singular initial value problems in the complex plane. This complexification is necessary since the space of analytic functions at \( t = 0 \), defined on \([0,1)\), is not complete.
Theorem 2.4. Define the set $A$ by

$$A := \{(z,x) \in \mathbb{C} \times \mathbb{C}^n | |z| \leq \delta_1 \land \|x\| \leq \delta_2\},$$

where $\delta_1, \delta_2 > 0$ are fixed.

Let $F: \mathbb{C} \times \mathbb{C}^n \rightarrow \mathbb{C}^n$ satisfy

(i) $F(0,0) = 0$

(ii) $F$ analytic in both arguments on $A$, i.e.,

$$F(z,x) = \sum_{i,j,\ldots,j_n=0}^{\infty} f_{ij_1\ldots j_n} z^i x_1^{j_1} x_2^{j_2} \ldots x_n^{j_n} \quad ((z,x) \in A),$$

where $f_{ij_1\ldots j_n} \in \mathbb{C}^n$.

Define $F_0(z) := F(z,0)$, $F_1 := \frac{\partial F}{\partial x}(0,0) \in \mathbb{C}^{n \times n}$ and $F^*(z,x) := F(z,x) - F_0(z) - F_1 x$.

Suppose $\sigma(F_1) \subset \mathbb{C}^-$.

Then there exists a $\delta_0 > 0$ such that the IVP

(2.9a) $z \frac{dx}{dz} = F(z,x)$

subject to

(2.9b) $\lim_{t \rightarrow 0^+} x(t) = 0$,

has exactly one analytic solution for $|z| \leq \delta_0$.

Proof. By the variation of constants formula, we see that any solution of

(2.9a) is also a solution of

$$x(z) = \int_0^1 \left[ -F_1 \frac{F_0(\tau z) + F^*(\tau z, x(\tau z))}{\tau} \right] d\tau.$$
Using successive substitution, we shall construct a sequence of functions $x^i$ ($i = 0, 1, 2, \ldots$) that converges to a solution of (2.10).

**Lemma 2.4a.** Define the sequence $\{x^i\}$ by

\[
\begin{cases}
  x^0 \equiv 0 \\
  x^{i+1}(z) := \int_0^1 \frac{-F_1 F_0(tz) + F^*(tz,x^i(tz))}{\tau} \, dt.
\end{cases}
\]

Then there exists a $\delta_3$, $0 < \delta_3 \leq \delta_1$ such that for $|z| \leq \delta_3$ and $i = 0, 1, 2, \ldots$

(i) $x^i(z)$ is well defined,

(ii) $\exists \epsilon \in \mathbb{R}^+ \text{ s.t. } \|x^i(z)\| \leq C|z|$, 

(iii) there exists a $\delta_0$, $0 < \delta_0 \leq \delta_3$, such that at $\{z \in \mathbb{C} | |z| \leq \delta_0\}$ the sequence $\{x^i\}$ converges to a solution $x$ of (2.10), which is analytic at $t = 0$.

**Proof.** Since (ii) implies (i) it suffices to show (ii) and (iii).

(ii) (by induction): Suppose (ii) is valid for $i = j$ ($j \in \mathbb{N}_0$). Then

\[
\|x^{j+1}(z)\| \leq \int_0^1 \|F_1\| \|x^j(tz) + F^*(tz,x^j(tz))\| \, dt.
\]

Note that $\|F_0(z)\| \leq C_1|z|$, $0 \leq |z| \leq \delta_1$, and since $\frac{\partial F^*}{\partial x}(0,0) = 0$, 

$\|\frac{\partial F^*}{\partial x}\| \leq C_2(|z| + \|x\|)$, $(z,x) \in A$.

Moreover, $F^*(z,0) = 0$, so

$\|F^*(z,x)\| \leq C_3(|z|\|x\| + \|x\|^2)$, $(z,x) \in A$.

Together with the inequality (Wasow [11], p. 93)

$$\|F_1\| \leq C_4 |\ln^{-1}(\tau)| + C_5, \quad 0 < \tau \leq 1,$$

where $\nu$ is the size of the largest Jordan-block belonging to an eigenvalue
\[ \lambda \text{ of } F_1 \text{ with } \Re \lambda = 0, \text{ we obtain} \]

\[ \| x^{i+1}(z) \| \leq |z| \int_0^1 (C_4 |1 + \nu - 1(\tau)| + C_5(C_1 + C_3(C + C^2)\tau |z|) \, d\tau \]

\[ = E_1 |z| + E_2 |z|^2, \quad 0 \leq |z| \leq \min(\delta_1, \delta_2/C), \]

where \( E_1 = C_1(C_4(\nu - 1)! + C_5) \) and \( E_2 = C_3(C + C^2)(\frac{1}{2}C_5 + C_4(\nu - 1)!/2^\nu) \).

Since \( E_1 \) and \( E_2 \) are independent from \( \tau \) and only \( E_2 \) depends on \( C \), we can choose \( C \in \mathbb{R}^+ \) and \( 0 < \delta_3 \leq \min(\delta_1, \delta_2/C) \) independent from \( \tau \), such that

\[ \| x^{i+1}(z) \| \leq C|z|, \quad 0 \leq |z| \leq \delta_3. \]

This inequality is certainly satisfied for \( \tau = 0 \), which implies (ii).

(iii) is proven in a similar way. Namely,

\[ \| F^*(z, x^i(z)) - F^*(z, x^{i-1}(z)) \| = \]

\[ = \| \int_0^1 \frac{\partial F^*}{\partial x}(z, x^{i-1}(z) + \theta(x^i(z) - x^{i-1}(z))) \cdot (x^i(z) - x^{i-1}(z)) \, d\theta \|
\]

\[ \leq \sup_{0 \leq \theta \leq 1} \| \frac{\partial F^*}{\partial x}(z, x^{i-1}(z) + \theta(x^i(z) - x^{i-1}(z))) \| \cdot \| x^i(z) - x^{i-1}(z) \|
\]

\[ \leq C_2(1 + C)|z| \| x^i(z) - x^{i-1}(z) \|, \quad 0 \leq |z| \leq \delta_3. \]

From this it follows that

\[ \| x^{i+1}(z) - x^i(z) \| = \]

\[ = \| \int_0^1 \frac{1}{\tau} F^*(\tau z, x^i(\tau z)) - F^*(\tau z, x^{i-1}(\tau z)) \, d\tau \|
\]

\[ \leq C_6 |z| \sup_{0 \leq \tau \leq 1} \| x^i(\tau z) - x^{i-1}(\tau z) \|. \]
Hence, taking the sup-norm and choosing $0 < \delta_0 \leq \delta_3$ such that $C_6 \cdot \delta_0 \leq L_0 < 1$, the sequence $\{x^i\}$ converges uniformly in $\{z \in \mathbb{C} \mid |z| \leq \delta_0\}$ and its limit, say $x$, satisfies (2.10).

Moreover, $\lim_{t \to 0+} x(t) = 0$, which implies that $x$ is also a solution of (2.9).

To prove that $x$ is analytic for $|z| \leq \delta_0$ we show by induction that $x^i$ is analytic for $|z| \leq \delta_3$ ($i = 0, 1, \ldots$).

(a) $x^0$ is analytic for $|z| \leq \delta_3$.

(b) suppose $x^i$ analytic for $|z| \leq \delta_3$.

Then $G_j(z) = F_0(z) + F^*(z, x^i(z))$ is analytic for $|z| \leq \delta_3$.

By definition we have that $G_j(0) = 0$, so we may write

$$G_j(\tau z) = \sum_{k=1}^{\infty} (\tau z)^k g_j^k \quad (g_j^k \in \mathbb{C}^n).$$

Then

$$x^{i+1}(z) = \int_0^1 (I + F_1)^{-1} \sum_{k=1}^{\infty} (\tau z)^k g_j^k d\tau$$

$$= \sum_{k=1}^{\infty} \frac{z^k}{k!} [kI - F_1]^{-1} g_j^k,$$

where $[kI - F_1]^{-1}$ exists, because $\sigma(F_1) \subset \mathbb{C}^-$.

Moreover, $\| [kI - F_1]^{-1} \|$ is bounded uniformly in $k$, so $x^{i+1}$ is analytic for $|z| \leq \delta_3$. Hence,

$$x = \lim_{i \to \infty} x^i(t)$$

is analytic for $|z| \leq \delta_0$.

This proves the lemma.

By Lemma 2.4 it is shown that an analytic solution of (2.9) exists, so we
only have to prove the unicity of this solution.

Let \( \bar{x} \) be another analytic solution of (2.9) and define \( e(z) := x(z) - \bar{x}(z) \).

Then, for \( |z| \) sufficiently small, \( 0 \leq e(z) \leq \tilde{C}|z| \). So,

\[
0 \leq e(z) \leq C_2(1 + C)|z| \int_0^1 \left( C_4|1n^{\nu-1}(\tau)| + C_5 \right) e(\tau z) d\tau
\]

\[
\leq C_7|z| \max_{0 \leq \tau \leq 1} e(\tau z) .
\]

Hence, for \( 0 \leq \varepsilon \leq \delta_0 \) sufficiently small,

\[
0 \leq \sup_{|z| \leq \varepsilon} e(z) \leq C_7 \cdot \varepsilon = \sup_{|z| \leq \varepsilon} e(z) ,
\]

which implies that \( e(z) = 0 \), for \( z \leq \varepsilon \).

By unicity of the power series expansion of an analytic function we obtain that \( x(z) = \bar{x}(z) \), for \( |z| \leq \delta_0 \).

Remarks 2.5.

(i) Under the given conditions we cannot maintain that (2.9) has a unique solution. For example let \( F(z, x) = x^2 \). Then (2.9) has infinitely many solutions, namely \( x(z) = 0 \lor x(z) = (C - \log z)^{-1} \).

Of course, only the first one of these solutions is analytic.

However, easy manipulations show that each of the following restrictions makes the solution unique:

- restrictions to the differential equation:
  - (a) \( \sigma(\mathcal{F}_1) \subset \mathbb{C}^- \) or
  - (b) \( F \) linear in \( x \),

- restrictions to the solution:
  - (c) \( \|x(z)\| = O(|z|^\varepsilon) \) (\( \varepsilon > 0 \)).
(ii) The same result could have been derived by construction of formal power series, satisfying (2.9). In Hautus [2] it is proven that this (unique) power series has a positive radius of convergence.

\begin{corollary}
If all coefficients of $F$ in Theorem 2.4 are real, i.e., $f_{ij_1 \ldots j_n} \in \mathbb{R}^n$, then the analytic solution of (2.9), restricted to the domain $[0, \delta_0)$, is real.

\begin{proof}
One directly verifies that $x^i(t) \in \mathbb{R}^n$, $i \in \mathbb{N}_0$ and $t \in [0, \delta_0)$, so $x(t) \in \mathbb{R}^n$, for $t \in [0, \delta_0)$.
\end{proof}
\end{corollary}

\begin{example}
Let $x : [0,1] \rightarrow \mathbb{R}^n$ satisfy the initial value problem

\begin{equation}
(2.14) \begin{cases}
tx'(t) = A(t)x(t) + f(t), & t \in (0,1] \\
x(0) = x_0, & 
\end{cases}
\end{equation}

where $A,f$ analytic at $t = 0$ and $\sigma(A(0)) \subseteq \mathbb{T}^-$. By Theorem 2.1 we know that $A(0)x_0 + f(0) = 0$.

Define $z(t) := x(t) - x_0, t \in [0,1]$. Then $z$ is a solution of

\begin{equation}
\begin{cases}
tz'(t) = A(t)z(t) + g(t), & t \in [0,1] \\
z(0) = 0, & 
\end{cases}
\end{equation}

where $g(t) = f(t) - (A(t) - tI)x_0$.

By remark 2.5.(i) this solution is unique and analytic at $t = 0$, so (2.14) has a unique solution, which is analytic at $t = 0$.
§3. Riccati transformations

To transform (1.1) and (1.2) into a regular system on \((\delta, 1) \, (\delta > 0)\) we use a so-called Riccati transformation

\[
\begin{pmatrix}
    y_1(t) \\
    y_2(t)
\end{pmatrix}
= \begin{bmatrix}
    I_{k+p} & 0 \\
    -R_{21}(t) & I_{q+m}
\end{bmatrix}
\begin{pmatrix}
    x_1(t) \\
    x_2(t)
\end{pmatrix}, \quad t \in [0, \delta],
\]

where \(R_{21}\) is chosen such that the transformed system is block upper-triangular and can be solved in a numerically stable way.

The idea to use such a transformation is not new. Russell [10] used this transformation in 1970 to identify the \(k\)-dimensional affine manifold \(V\) in \(\mathbb{R}^n\) in which a solution of (1.1) and (1.2a) is found, where he assumed that \(\ell = 0\). Nelson, Sagong and Elder [9] tried to find the solution of a homogeneous singular linear BVP by a Riccati transformation without first reducing it to a regular problem on some subinterval. However, difficult problems then arise in the course of solving the differential equations for \(y_1\), because its solution is in general not analytic at \(t = 0\).

After transformation (3.1) the system (1.1) is changed into

\[
\begin{pmatrix}
    \tilde{A}_{11}(t) & \tilde{A}_{12}(t) \\
    \tilde{A}_{21}(t) & \tilde{A}_{22}(t)
\end{pmatrix}
\begin{pmatrix}
    y(t) \\
    \tilde{y}(t)
\end{pmatrix}
= \begin{pmatrix}
    f_1(t) \\
    \tilde{f}_2(t)
\end{pmatrix}, \quad t \in (0, \delta],
\]

where the same partitioning is used as in Section 2 and
Theorem 3.1. There exists an $\varepsilon > 0$ such that \( (3.3) \) has exactly one solution on \([0, \varepsilon)\), say $R_{21}$, which is analytic at $t = 0$.

Proof. Consider the entries of $U(t) \in \mathbb{R}^{(q+m)(k+p)}$ as entries of a vector $u(t) \in \mathbb{R}^{(q+m)(k+p)}$. Let $F(t,U) := A_{21}(t) + A_{22}(t)U - UA_{11}(t) - UA_{12}(t)U$ and $\tilde{F}(t,u)$ the associated vectorfunction. Now it is evident that $\tilde{F}(t,u)$ satisfies the conditions (i) and (ii) of Theorem 2.4 and the condition of Corollary 2.6.

Moreover,

$$F_1 u := \frac{\partial \tilde{F}}{\partial u}(0,0)u = (A_{22} \oplus (-A_{11}))u,$$

where $\oplus$ denotes the Kronecker sum.

So

$$\sigma(F_1) = \sigma(A_{22}) \oplus \sigma(-A_{11}) := \{ \gamma \in \mathbb{C} \mid \gamma = \alpha - \beta, \alpha \in \sigma(A_{22}), \beta \in \sigma(A_{11}) \}.$$
The partitioning (3.2) is chosen such that $\sigma(F_1) \subset \overline{\mathcal{C}^-}$, so all conditions to guarantee the existence of exactly one solution of (3.3) on $[0, \varepsilon)$ ($\varepsilon > 0$), which is analytic at $t = 0$, are satisfied.

As direct consequence of Theorem 3.1 we obtain that the righthand side of the DE for $y_2$ is analytic at $t = 0$. Since $\tilde{A}_{22}(0) = A_{22}(0)$, which has its spectrum in $\mathcal{C}^-$, $y_2(0) = x_2(0)$ and by Corollary 2.6, also $y_2$ is analytic at $t = 0$.

Remark 3.2. Writing $R_{21}(t) = \sum_{k=1}^{\infty} C_k t^k$, the following relation for $C_k$ ($k=1,2,\ldots$) is found:

$$(kI - A_{22}^0)C_k + C_k A_{11}^0 =$$

$$= A_{21}^k + \sum_{m=1}^{k-1} \left( A_{22}^m C_{k-m} - C_{k-m} A_{11}^m - C_m \sum_{n=0}^{k-m-1} A_{12}^n C_{k-m-n} \right),$$

which has a unique solution for all $k \geq 1$ (Wasow [11], ch.2). By remark 2.5.iii this also implies the local existence of $R_{21}$.

A result that will be used later is the following:

**Theorem 3.3.** There exists a fundamental matrix $X : [0,1] \rightarrow \mathbb{R}^{n \times n}$ corresponding to the system (1.1) such that $R_{21}(t) = X_{21}(t)X_{11}^{-1}(t)$, for all $t \in [0, \varepsilon]$, where $\varepsilon$ is defined by Theorem 3.1.

**Proof.** Let $X$ be a fundamental matrix for which $X_{11}(t)$ ($t \in (0, \varepsilon]$) is regular. Then it follows from easy manipulations that $X_{21}(t)X_{11}^{-1}(t)$ satisfies the same differential equation as $R_{21}$, so it suffices to show that
If \( A^0 \) has no eigenvalues that differ by positive integers, then there exists a fundamental matrix \( X \) corresponding to (1.1) of the form

\[
X(t) = P(t)t^{A^0}, \quad t \in [0,1],
\]

where \( P \) is analytic at \( t = 0 \) and \( P(0) = I \) (Henrici [3], Th.9.5.c). This matrix \( X \) already has the desired property.

If \( A^0 \) has eigenvalues that do differ by a positive integer then the proof is rather technical.

Elaborating Theorem 9.5.d in Henrici [3] one finds that there exists a fundamental matrix \( X \) such that

\[
\begin{bmatrix}
X_{11}(t) \\
X_{21}(t)
\end{bmatrix}
= V(t)\bar{P}(t)
\begin{bmatrix}
U_{11} \\
0
\end{bmatrix}, \quad t \in [0,1],
\]

where

\[
V(t) = \begin{bmatrix}
V_{kk}(t) & 0 \\
0 & I_p \\
0 & I_{q+m}
\end{bmatrix}
\text{ is regular for } t > 0,
\]

\[
U = \begin{bmatrix}
U_{kk} & * & * & * \\
0 & A_{pq}^0 \\
0 & A_{qq}^0 \\
\bar{A}
\end{bmatrix}
\text{ with } U_{kk} \text{ upper-triangular}
\]

and \( \bar{P} \) analytic at \( t = 0 \) with \( \bar{P}(0) = I_n \).
Now

\[(3.4) \quad x_{21}(t)x^{-1}_{11}(t) = \tilde{P}_{21}(t)\tilde{P}^{-1}_{11}(t) \begin{bmatrix} v^{-1}_{kk}(t) & 0 \\ 0 & I_p \end{bmatrix}, \]

which already implies that the last p columns of \(x_{21}x^{-1}_{11}\) are of order \(t\).

More detailed we have for the first \(k\) columns

\[(3.5) \quad v_{kk}(t) = \begin{bmatrix} v_1(t) & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & v_s(t) \end{bmatrix}, \]

where \(v_i(t) = t^{a_i}w_i\) (\(i = 1, \ldots, s\)) with \(a_1, \ldots, a_s\) integers satisfying

\[a_1 \geq a_2 \geq \ldots \geq a_s \geq 0\] and regular \(w_i\).

With \((\tilde{P}_{21}(t))_i\) we will mean those columns of \(\tilde{P}_{21}(t)\) that correspond to \(v_i(t)\) (\(i = 1, \ldots, s\)). Then by induction one can prove that

\[\| (\tilde{P}_{21}(t))_i \| = o(t^{a_i+1}), \quad t \to 0.\]

Since (with similar notation)

\[(x_{21}(t)x^{-1}_{11}(t))_i = (\tilde{P}_{21}(t)\tilde{P}^{-1}_{11}(t))_i w_i^{-1} t^{-a_i}\]

and

\[\| \tilde{P}_{11}(t) - I_{k+p} \| = o(t), \quad t \to 0,\]

the theorem is proven.

(Observe that by (3.4) and (3.5) we have a third proof of Theorem 3.1.)
Note that the function $R_{21}$ defined by (3.3) can be seen as a generalization of the Riccati matrix defined by Meyer ([8], p. 68). This follows from the fact that the boundary conditions (1.2) may also be written as

\[
\begin{pmatrix}
B^0 \\
0
\end{pmatrix}
\begin{pmatrix}
x(0) \\
x(1)
\end{pmatrix}
= \begin{pmatrix}
B^1 \\
s(m+n)
\end{pmatrix}
\begin{pmatrix}
b \\
x_2(0)
\end{pmatrix},
\]

where $x_2(0)$ is found by Theorem 2.1.

However, in our case numerical stability of the integration method can be guaranteed.

§4. Boundary conditions

Choose $\delta > 0$ such that $R_{21}$ exists on $[0, \delta]$. Now the boundary conditions at $t = 0$ will be translated in conditions at $t = \delta$.

If we return to system (3.2) we see that the differential equation for $R_{21}$ has at $t = 0$ a Jacobian with spectrum in $\mathbb{C}^-$, but also that the sets of bounded and unbounded solutions of the differential equation (1.1) are decoupled for $t$ sufficiently small.

From the resulting system, $y_2$ can be computed directly, since $y_2(0) = x_2(0)$ is known, by Theorem 2.1 and, by Theorem 2.4, $y_2$ is analytic at $t = 0$. Moreover, this computation can be done numerically stable because

\[
\sigma(\tilde{A}_{22}^0) = \sigma(A^0_{22}) \subset \mathbb{C}^-.
\]

When $y_2(\delta)$ has been computed, the boundary condition (1.2a) can be replaced by

\[
(4.1) \quad -R_{21}(\delta)x_1(\delta) + x_2(\delta) = y_2(\delta),
\]
which indeed imposes \((q+m)\) restrictions to the solution of \((1.1)\).

Since \(y_2(t)\) is known for \(t \in [0,\delta]\) we obtain for \(x_1 = y_1\) the differential equation

\[
\frac{t}{x_1(t)} = \widetilde{A}_1(t)x_1(t) + A_{12}(t)y_2(t) + f_1(t), \quad t \in (0,\delta].
\]

We want to express \(x_p(0)\) in terms of \(x_1(\delta)\). To this end let us assume there exist functions \(R_{p1}(t)\) and \(z_p(t)\) such that

\[
x_p(0) = R_{p1}(t)x_1(t) + z_p(t),
\]

with

\[
R_{p1}(0) = [0 \ I_p] \quad \text{and} \quad z_p(0) = 0.
\]

Differentiating (4.3) leads to

\[
0 = (tR'_{p1}(t) + R_{p1}(t))\widetilde{A}_1(t)x_1(t) +
\]

\[
+ tz'_p(t) + R_{p1}(t)(A_{12}(t)y_2(t) + f_1(t)).
\]

Hence, it suffices to take \(R_{p1}\) and \(z_p\) such that

\[
\begin{align*}
\begin{cases}
\frac{tR'_{p1}(t)}{} = -R_{p1}(t)A_{11}(t), & t \in (0,\delta] \\
R_{p1}(0) = [0 \ I_p]
\end{cases}
\end{align*}
\]

and

\[
\begin{align*}
\begin{cases}
\frac{tz'_p(t)}{} = -R_{p1}(t)(A_{12}(t)y_2(t) + f_1(t)), & t \in (0,\delta] \\
z_p(0) = 0
\end{cases}
\end{align*}
\]
Both systems of differential equations have a unique solution, which is analytic at $t = 0$ and that can be computed in a numerically stable way.

Remark 4.1. $R_p$ is the matrix consistency of the last $p$ rows of $X^{-1}_{11}$, where $X_{11}$ is defined by Theorem 3.3.

Proof:

$$tX'_{11}(t) = A_{11}(t)X_{11}(t) + A_{12}(t)X_{21}(t)$$

$$= (A_{11}(t) + A_{12}(t)R_{21}(t))X_{11}(t)$$

$$= \tilde{A}_{11}(t)X_{11}(t)$$

Hence

$$t(X^{-1}_{11})'(t) = -X^{-1}_{11}(t)\tilde{A}_{11}(t).$$

Moreover, since $P_{11}(0)$ is invertible,

$$\lim_{t \to 0} [0\ L_p]X^{-1}_{11}(t) = \lim_{t \to 0} [0\ L_p]t^{-U}_{11}(V_{11}(t)\tilde{P}_{11}(t))^{-1}$$

$$= \lim_{t \to 0} [0\ L_p] [\begin{array}{cc} -U_{kk} & 0 \\ 0 & I_p \end{array}] [\begin{array}{cc} \tilde{V}_{kk}(t) & 0 \\ 0 & I_p \end{array}]$$

$$= \lim_{t \to 0} [0\ L_p] \tilde{P}_{11}(t) = [0\ L_p].$$

By the unicity of the solution of (4.4) the proof is now complete.
Computing \( R_{21}(δ), y_2(δ), R_{p1}(δ) \) and \( z_p(δ) \) and writing

\[
B_i = \begin{bmatrix} B_k^i & B_p^i & B_2^i \end{bmatrix} k^p q^m (i = 0, 1),
\]

the boundary conditions (1.2) can be replaced by

\[
\begin{bmatrix} B_0^0 R_{p1}(δ) & 0 \\ -R_{21}(δ) & I_q^m \end{bmatrix} x(δ) + \begin{bmatrix} B_k^1 & B_p^1 & B_2^1 \\ 0 & 0 & 0 \end{bmatrix} x(1) = \begin{bmatrix} b - B_k^0 x_k(0) - B_2^0 x_2(0) - B_p^0 z_p(δ) \\ y_2(δ) \end{bmatrix}.
\]

(4.6)

Now our mission is fulfilled since the singular boundary value problem (1.1) and (1.2) is replaced by a regular one, namely (1.1) and (4.6). By construction we have that the solution of (1.1) and (1.2) is also a solution of (1.1) and (4.6). To prove that (1.1) and (4.6) have a unique solution let

\[
B^δ := \begin{bmatrix} B_0^0 R_{p1}(δ) & 0 \\ -R_{21}(δ) & I_q^m \end{bmatrix} = \begin{bmatrix} 0 & B_0^0 x_1^{-1}(δ) \\ -x_2(δ)x_1^{-1}(δ) & I_q^m \end{bmatrix}.
\]
where $X$ is defined by Theorem 3.3, and

$$B' = \begin{bmatrix} B_1^k & B_1^p & B_2^l \\ \varnothing & \varnothing & \varnothing \end{bmatrix}.$$ 

Now (1.1) and (4.6) have a unique solution if and only if $[B^gX(8) + B'X(1)]$ is invertible (Keller [6]).

**Theorem 4.2.** (1.1) and (4.6) have a unique solution if and only if (1.1) and (1.2) have a unique solution.

**Proof.** Suppose (1.1) and (1.2) have a unique solution $x$. Let $\bar{x}$ be that particular solution of (1.1) for which

$$\bar{x}(0) = \begin{bmatrix} x_k(0) \\ x_2(0) \end{bmatrix}.$$ 

Then $x(t) = X(t)\begin{bmatrix} c_1 \\ 0 \end{bmatrix}$ and

$$\begin{bmatrix} 0 & B_0^p & 0 \\ B_0^k & 0 & 0 \end{bmatrix} X(0) + B'X(1) = \begin{bmatrix} c_1 \\ 0 \end{bmatrix} = b - B_0^k x_k(0) - B_2^0 x_2(0) - B_1^l x(1)$$

has a unique solution, which implies that
(4.7) \[ 0 \begin{bmatrix} B_0^0 & B_1^1 \end{bmatrix} X_{11}(0) + \begin{bmatrix} B_0^1 & B_1^0 \end{bmatrix} X_{11}(1) + B_2^1 X_{21}(1) \]

is invertible.

Since (4.7) is equal to \( [B^0 X(\delta) + B^1 X(1)]_{11} \) and \( [B^0 X(\delta) + B^1 X(1)]_{21} = 0 \) it suffices to show that

\[ [B^0 X(\delta) + B^1 X(1)]_{22} = X_{22}(\delta) - X_{21}(\delta) X_{11}(\delta)^{-1} X_{12}(\delta) \]

is invertible. This is however a direct consequence of the fact that \( X(\delta) \) and \( X_{11}(\delta) \) are invertible. The other way around is proved by reversing the argument.

Remark 4.3. To find \( y_2', R_p^1 \) and \( z_p \) we need \( R_{21}' \). However, it is not necessary to store computed values of \( R_{21}(t) \), since all these functions can be computed by solving just one \((p+q+m) \times (k+p+1)\) Riccati differential equation. Namely, by writing

\[
W(t) = \begin{bmatrix} R_{p1}(t) & z_p(t) \\ R_{21}(t) & y_2(t) \end{bmatrix}, \quad (t \in [0, \delta]),
\]
we obtain

\[
W'(t) = A(t) \cdot W(t) + W(t) \cdot B(t) + C(t)
\]

(4.8)

By Theorem 2.4, (4.8) has exactly one solution, which is analytic at \( t = 0 \), although the differential equation may be stiff.

§5. Bounded solutions

A second class of problems for which this method is applicable is obtained by replacing the condition (1.2a) by the boundedness condition

(5.1) \[ \sup_{t \in (0,1)} \| x(t) \| < \infty. \]
In that case we assume the following partitioning:

\[ A^0 = \begin{bmatrix}
A^+ & & & \\
& & & \\
& & D & A^0_{q_1 q_2} \\
& & & \\
& & & A^-
\end{bmatrix} \]

where \( D \in \mathbb{R}^{q_1 \times q_1} \) is a diagonal matrix with purely imaginary (but non-zero) eigenvalues and \( q_1 \) is the sum of the geometric multiplicities of all purely imaginary eigenvalues of \( A^0 \).

Consequently we have

\[
A(t) = \begin{bmatrix}
A_{11}(t) & A_{12}(t) \\
A_{21}(t) & A_{22}(t)
\end{bmatrix}
\]

\[
\begin{array}{c}
k+p+q_1 \\
k+p+q_1
\end{array}
\]

\[
\begin{array}{c}
k+p+q_1 \\
q_2+m
\end{array}
\]

and \( x(t) \) partitioned likewise or

\[
x(t) = (x_k(t)^T, x_p(t)^T, x_{q_1}(t)^T, x_{q_2}(t)^T, x_m(t)^T)^T,
\]

where \( x_k(t) \in \mathbb{R}^k \), etc.
Theorem 5.1. Let $x(t)$ satisfy (1.1) and (5.1).

Then

$$
\left[
\begin{array}{c}
x_k(0) \\
x_2(0)
\end{array}
\right] := \lim_{t \to 0} \left[
\begin{array}{c}
x_k(t) \\
x_2(t)
\end{array}
\right]
$$

exists and satisfies the relation

$$
\begin{bmatrix}
A^+ & 0 \\
0 & A_{pq2}
\end{bmatrix}
\begin{bmatrix}
x_k(0) \\
x_2(0) \\
x_2(0)
\end{bmatrix}
+ 
\begin{bmatrix}
f_u^0 \\
f_p^0 \\
f_q^0 \\
f_m^0
\end{bmatrix} = 0.
$$

Proof. The first part of the theorem may be proven by investigation of differential equations of the form

$$
tu'(t) = Ju(t) + f(t), \quad t \in (0,1),
$$

where $J$ is a Jordan-block, like is done in de Hoog, Weiss [5].

The relations

$$
A^+ x_k(0) + f_k^0 = 0
$$

and

$$
A_{22} x_2(0) + f_2^0 = 0
$$

are found similarly to Theorem 2.1.
Write
\[ a_p = A^0_{pq_2} x_{q_2}(0) + f^0_p, \]
then \( t x'(t) = a_p + \epsilon_p(t), \quad t \in (0,1), \)
where \( \epsilon_p(t) = o(1), \quad t \to 0. \)
Hence,
\[ x_p(t) - x_p(t_0) = a_p \log(t/t_0) + \int_{t_0}^{t} \frac{\epsilon_p(u)}{u} \, du, \]
where \( t_0 \) is fixed.
Thus, by (5.1), \( a_p := A^0_{pq_2} x_{q_2}(0) + f^0_p = 0. \)

In general \( x_2(0) \) is not uniquely defined by (5.3). We would like to derive a relation of the form
\[ A^0_{q_1 q_2 q_2} x_{q_2}(0) + f^0_{q_1} = 0. \]
However, observing that a solution \( x_{q_2} \) of
\[ t x_{q_1}'(t) = D x_{q_1}(t) + A^0_{q_1 q_2 q_2} x_{q_2}(0) + f^0_{q_1} + \epsilon_{q_1}(t), \quad t \in (0,1), \]
where \( \epsilon_{q_1}(t) = o(1), \quad t \to 0, \) is always oscillatory and bounded, no extra conditions for \( x_{q_2}(0) \) follow from this differential equation.
Hence, a necessary condition for \( x_2(0) \) to be determined explicitly from (1.1) and (5.1) is that
\[ \text{rank}
\begin{bmatrix}
  A^0_{p q_2} \\
  A^0_{q_1 q_2 q_2}
\end{bmatrix}
= q_2. \]
Theorem 5.2. The condition (5.4) is always satisfied.

Proof.
Suppose
\[
\begin{bmatrix}
A_0^0 & A_{pq_2}^0 \\
\emptyset & x_2 = 0 \\
A_0 & A_{q_2 q_2}^0
\end{bmatrix}
\]
Since D is regular there exists an \( x_1 \in \mathbb{R}^q \) such that \( Dx_1 + A_{q_1 q_2}^0 x_2 = 0 \). So,
\[
\begin{bmatrix}
A_0^0 & A_{pq_2}^0 \\
D & A_{q_1 q_2}^0 \\
\emptyset & A_{q_2 q_2}^0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} = 0,
\]
which implies \( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \), because \( p = \text{dim}(\ker(A^0)) \).

Similarly to Theorem 2.3 we have

Theorem 5.3. Let \( V \) be the subspace of \( C^1(0,1] \) formed by all bounded solutions of
\[
tx'(t) = A(t)x(t), \quad t \in (0,1].
\]
Then \( \text{dim} \ V = k + p + q_1 \).

Hence, for the existence of a unique bounded solution (1.1), subject to (1.2b), we need that \( s = k + p + q_1 \) and
If this condition is satisfied, the same technique to replace the boundary conditions (5.1) and (1.2b) as in Chapters 3 and 4 can be used. Finally, we have to compute the only analytic solution (cf. Theorem 2.4) of

\[
\begin{pmatrix}
\dot{x}_1^0 \\
\dot{x}_p^0 \\
\dot{x}_{q_2}^0
\end{pmatrix}
\in \mathbb{R}^n \begin{pmatrix}
A_{pq_2}^0 \\
0 \\
A_{q_2q_2}^0
\end{pmatrix}
\]

subject to

\[
-W(t) - W(t) 
\begin{bmatrix}
A_{11}(t) & f_1(t) \\
0 & 0
\end{bmatrix}
- W(t) 
\begin{bmatrix}
0 & 0
\end{bmatrix}
W(t),
\]

with

\[
JW'(t) = \begin{bmatrix}
0 & 0 \\
A_{21}(t) & f_2(t) \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
0 & 0 \\
0 & 0
\end{bmatrix} W(t)
\]

(5.5)

subject to

\[
W(0) = \begin{bmatrix}
0 & I_p & 0 \\
0 & 0 & 0 \\
0 & 0 & x_2(0)
\end{bmatrix}
W(t) \in \mathbb{R}^{(p+q_2+m) \times (k+p+q_1+1)}.
\]

As soon as \(W(\delta)\) is known, boundary conditions like (4.6) can simply be derived \((R_{q_1}^0 = 0!)\).
References


London, etc., Wiley 1977 (Pure and Applied Mathematics.)


