ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

FOUNDING EDITOR G. C. ROTA

All the titles listed below can be obtained from good booksellers or from Cambridge University Press. For a complete series listing visit http://publishing.cambridge.org/stm/mathematics/eom/.

88. Teo Mora *Solving Polynomial Equation Systems, I*
89. Klaus Bichteler *Stochastic Integration with Jumps*
90. M. Lothaire *Algebraic Combinatorics on Words*
91. A. A. Ivanov & S. V. Shpectorov *Geometry of Sporadic Groups, 2*
92. Peter McMullen & Egon Schulte *Abstract Regular Polytopes*
93. G. Gierz *et al. Continuous Lattices and Domains*
94. Steven R. Finch *Mathematical Constants*
95. Youssef Jabri *The Mountain Pass Theorem*
96. George Gasper & Mizan Rahman *Basic Hypergeometric Series, 2nd ed.*
97. Maria Cristina Pedicchio & Walter Tholen *Categorical Foundations*
100. Enzo Olivieri & Maria Eulalia Vares *Large Deviations and Metastability*
102. R. J. Wilson & L. Beineke *Topics in Algebraic Graph Theory*

Well-Posed Linear Systems

OLOF STAFFANS

Department of Mathematics
Åbo Akademi University, Finland

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

www.cambridge.org
Information on this title: www.cambridge.org/9780521825849

First published 2005

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this book is available from the British Library

Library of Congress Cataloging in Publication data

Staffans, Olof J., 1947–
Well-posed linear systems / Olof Staffans.
p. cm. – (Encyclopedia of mathematics and its applications)
Includes bibliographical references and index.
ISBN 0-521-82584-9
1. Linear systems. 2. System theory. I. Staffans, Olof J., 1947–
II. Title. III. Series.
QA372.S784 2004
003′.74–dc22 2004051868

ISBN-13 978-0-521-82584-9 hardback
ISBN-10 0-521-82584-9 hardback

Contents

Figures

Preface

This main purpose of this book is to present the basic theory of well-posed linear systems in a form which makes it available to a larger audience, thereby opening up the possibility of applying it to a wider range of problems. Up to now the theory has existed in a distributed form, scattered between different papers with different (and often noncompatible) notation. For many years this has forced authors in the field (myself included) to start each paper with a long background section to first bring the reader up to date with the existing theory. Hopefully, the existence of this monograph will make it possible to dispense with this in future.

My personal history in the field of abstract systems theory is rather short but intensive. It started in about 1995 when I wanted to understand the true nature of the solution of the quadratic cost minimization problem for a linear Volterra integral equation. It soon became apparent that the most appropriate setting was not the one familiar to me which has classically been used in the field of Volterra integral equations (as presented in, e.g., Gripenberg *et al.* [1990]). It also became clear that the solution was not tied to the class of Volterra integral equations, but that it could be formulated in a much more general framework. From this simple observation I gradually plunged deeper and deeper into the theory of well-posed (and even non-well-posed) linear systems.

One of the first major decisions that I had to make when I began to write this monograph was how much of the existing theory to include. Because of the nonhomogeneous background of the existing theory (several strains have been developing in parallel independently of each other), it is clear that it is impossible to write a monograph which will be fully accepted by every worker in the field. I have therefore largely allowed my personal taste to influence the final result, meaning that results which lie closer to my own research interests are included to a greater extent than others. It is also true that results which blend more easily into the general theory have had a greater chance of being included than those which are of a more specialist nature. Generally speaking, instead of borrowing results directly from various sources I have reinterpreted and reformulated many existing results into a coherent setting and, above all, using a coherent notation.

The original motivation for writing this book was to develop the background which is needed for an appropriate understanding of the quadratic cost minimization problem (and its indefinite minimax version). However, due to page and time limitations, I have not yet been able to include any optimal control in this volume (only the background needed to attack optimal control problems). The book on optimal control still remains to be written.

Not only was it difficult to decide exactly what parts of the existing theory to include, but also in which form it should be included. One such decision was whether to work in a *Hilbert space* or in a *Banach space* setting. Optimal control is typically done in Hilbert spaces. On the other hand, in the basic theory it does not matter if we are working in a Hilbert space or a Banach space (the technical differences are minimal, compared to the general level of difficulty of the theory). Moreover, there are several interesting applications which require the use of Banach spaces. For example, the natural norm in population dynamics is often the L^1-norm (representing the total mass), parabolic equations have a well-developed L^p-theory with $p \neq 2$, and in nonlinear equations it is often more convenient to use L^∞-norms than L^2-norms. The natural decision was to present the basic theory in an arbitrary Banach space, but to specialize to Hilbert spaces whenever this additional structure was important. As a consequence of this decision, the present monograph contains the first comprehensive treatment of a well-posed linear system in a setting where the input and output signals are continuous (as opposed to belonging to some L^p-space) but do not have any further differentiability properties (such as belonging to some Sobolev spaces). (More precisely, they are continuous apart from possible jump discontinuities.)

The first version of the manuscript was devoted exclusively to *well-posed* problems, and the main part of the book still deals with problems that are well posed. However, especially in H^∞-optimal control, one naturally runs into non-well-posed problems, and this is also true in circuit theory in the impedance and transmission settings. The final incident that convinced me that I also had to include some classes of non-well-posed systems in this monograph was my discovery in 2002 that every passive impedance system which satisfies a certain algebraic condition can be represented by a (possibly non-well-posed) *system node*. System nodes are a central part of the theory of well-posed systems, and the well-posedness property is not always essential. My decision not to stay strictly within the class of well-posed systems had the consequence that this monograph is also the the first comprehensive treatment of (possibly non-well-posed) systems generated by arbitrary system nodes.

The last three chapters of this book have a slightly different flavor from the earlier chapters. There the general Banach space setting is replaced by a standard Hilbert space setting, and connections are explored between well-posed linear systems, Fourier analysis, and operator theory. In particular, the admissibility of scalar control and observation operators for contraction semigroups is characterized by means of the Carleson measure theorem, and systems theory interpretations are given of the basic dilation and model theory for contractions and continuous-time contraction semigroups in Hilbert spaces.

It took me approximately six years to write this monograph. The work has primarily been carried out at the Mathematics Institute of Åbo Akademi, which has offered me excellent working conditions and facilities. The Academy of Finland has supported me by relieving me of teaching duties for a total of two years, and without this support I would not have been able to complete the manuscript in this amount of time.

I am grateful to several students and colleagues for helping me find errors and misprints in the manuscript, most particularly Mikael Kurula, Jarmo Malinen and Kalle Mikkola.

Above all I am grateful to my wife Marjatta for her understanding and patience while I wrote this book.

Notation

Basic sets and symbols

$\mathbb{C}$	The complex plane.
$\mathbb{C}^+_\omega$, $\overline{\mathbb{C}}^+_\omega$	$\mathbb{C}^+_\omega := \{z \in \mathbb{C} \mid \Re z > \omega\}$ and $\overline{\mathbb{C}}^+_\omega := \{z \in \mathbb{C} \mid \Re z \geq \omega\}$.
$\mathbb{C}^-_\omega$, $\overline{\mathbb{C}}^-_\omega$	$\mathbb{C}^-_\omega := \{z \in \mathbb{C} \mid \Re z < \omega\}$ and $\overline{\mathbb{C}}^-_\omega := \{z \in \mathbb{C} \mid \Re z \leq \omega\}$.
$\mathbb{C}^+$, $\overline{\mathbb{C}}^+$	$\mathbb{C}^+ := \mathbb{C}^+_0$ and $\overline{\mathbb{C}}^+ := \overline{\mathbb{C}}^+_0$.
$\mathbb{C}^-$, $\overline{\mathbb{C}}^-$	$\mathbb{C}^- := \mathbb{C}^-_0$ and $\overline{\mathbb{C}}^- := \overline{\mathbb{C}}^-_0$.
$\mathbb{D}^+_r$, $\overline{\mathbb{D}}^+_r$	$\mathbb{D}^+_r := \{z \in \mathbb{C} \mid \Re z > r\}$ and $\overline{\mathbb{D}}^+_r := \{z \in \mathbb{C} \mid \lvert z\rvert \geq r\}$.
$\mathbb{D}^-_r$, $\overline{\mathbb{D}}^-_r$	$\mathbb{D}^-_r := \{z \in \mathbb{C} \mid \Re z < r\}$ and $\overline{\mathbb{D}}^-_r := \{z \in \mathbb{C} \mid \lvert z\rvert \leq r\}$.
$\mathbb{D}^+$, $\overline{\mathbb{D}}^+$	$\mathbb{D}^+ := \mathbb{D}^+_1$ and $\overline{\mathbb{D}}^+ := \overline{\mathbb{D}}^+_1$.
$\mathbb{D}^-$, $\overline{\mathbb{D}}^-$	$\mathbb{D}^- := \mathbb{D}^-_1$ and $\overline{\mathbb{D}}^- := \overline{\mathbb{D}}^-_1$.
$\mathbb{R}$	$\mathbb{R} := (-\infty, \infty)$.
$\mathbb{R}^+$, $\overline{\mathbb{R}}^+$	$\mathbb{R}^+ := (0, \infty)$ and $\overline{\mathbb{R}}^+ := [0, \infty)$.
$\mathbb{R}^-$, $\overline{\mathbb{R}}^-$	$\mathbb{R}^- := (-\infty, 0)$ and $\overline{\mathbb{R}}^- := (-\infty, 0]$.
$\mathbb{T}$	The unit circle in the complex plane.
$\mathbb{T}_T$	The real line $\mathbb{R}$ where the points $t + mT$, $m = 0, \pm 1, \pm 2, \ldots$ are identified.
$\mathbb{Z}$	The set of all integers.
$\mathbb{Z}^+$, $\mathbb{Z}^-$	$\mathbb{Z}^+ := \{0, 1, 2, \ldots\}$ and $\mathbb{Z}^- := \{-1, -2, -3, \ldots\}$.
j	$j := \sqrt{-1}$.
0	The number zero, or the zero vector in a vector space, or the zero operator, or the zero-dimensional vector space $\{0\}$.
1	The number one and also the identity operator on any set.

Operators and related symbols

A, B, C, D	In connection with an $L^p\vert Reg$-well-posed linear system or an operator node, A is usually the main operator, B the control

	operator, C the observation operator and D a feedthrough operator. See Chapters 3 and 4.
$C\&D$	The observation/feedthrough operator of an $L^p\|Reg$-well-posed linear system or an operator node. See Definition 4.7.2.
$\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{D}$	The semigroup, input map, output map, and input/output map of an $L^p\|Reg$-well-posed linear system, respectively. See Definitions 2.2.1 and 2.2.3.
$\widehat{\mathfrak{D}}$	The transfer function of an $L^p\|Reg$-well-posed linear system or an operator node. See Definitions 4.6.1 and 4.7.4.
$\mathcal{B}(U;Y)$, $\mathcal{B}(U)$	The set of bounded linear operators from U into Y or from U into itself, respectively.
$\mathcal{C}$, $\mathcal{L}$	The Cayley and Laguerre transforms. See Definition 12.3.2.
τ^t	The bilateral time shift operator $\tau^t u(s) := u(t+s)$ (this is a left-shift when $t>0$ and a right-shift when $t<0$). See Example 2.5.3 for some additional shift operators.
γ_λ	The time compression or dilation operator $(\gamma_\lambda u)(s) := u(\lambda s)$. Here $\lambda > 0$.
π_J	$(\pi_J u)(s) := u(s)$ if $s \in J$ and $(\pi_J u)(s) := 0$ if $s \notin J$. Here $J \subset \mathbb{R}$.
π_+, π_-	$\pi_+ := \pi_{[0,\infty)}$ and $\pi_- := \pi_{(-\infty,0)}$.
$\boldsymbol{Я}$	The time reflection operator about zero: $(\boldsymbol{Я}u)(s) := u(-s)$ (in the L^p-case) or $(\boldsymbol{Я}u)(s) := \lim_{t\downarrow -s} u(t)$ (in the Reg-case). See Definition 3.5.12.
$\boldsymbol{Я}_h$	The time reflection operator about the point h. See Lemma 6.1.8.
$\boldsymbol{\sigma}$	The discrete-time bilateral left-shift operator $(\boldsymbol{\sigma}\mathbf{u})_k := u_{k+1}$, where $\mathbf{u} = \{u_k\}_{k\in\mathbb{Z}}$. See Section 12.1 for the definitions of $\boldsymbol{\sigma}_+$ and $\boldsymbol{\sigma}_-$.
$\boldsymbol{\pi}_J$	$(\boldsymbol{\pi}_J \mathbf{u})_k := u_k$ if $k \in J$ and $(\boldsymbol{\pi}_J \mathbf{u})_k := 0$ if $k \notin J$. Here $J \subset \mathbb{Z}$ and $\mathbf{u} = \{u_k\}_{k\in\mathbb{Z}}$.
$\boldsymbol{\pi}_+$, $\boldsymbol{\pi}_-$	$\boldsymbol{\pi}_+ := \boldsymbol{\pi}_{\mathbb{Z}^+}$ and $\boldsymbol{\pi}_- := \boldsymbol{\pi}_{\mathbb{Z}^-}$.
w-lim	The weak limit in a Banach space. Thus $\text{w-lim}_{n\to\infty} x_n = x$ in X iff $\lim_{n\to\infty} x^* x_n = x^* x$ for all $x^* \in X^*$. See Section 3.5.
$\langle x, x^* \rangle$	In a Banach space setting $x^* x := \langle x, x^* \rangle$ is the continuous linear functional x^* evaluated at x. In a Hilbert space setting this is the inner product of x and x^*. See Section 3.5.
$E^\perp$	$E^\perp := \{x^* \in X^* \mid \langle x, x^* \rangle = 0 \text{ for all } x \in E\}$. This is the annihilator of $E \subset X$. See Lemma 9.6.4.
${}^\perp F$	${}^\perp F := \{x \in X \mid \langle x, x^* \rangle = 0 \text{ for all } x^* \in F\}$. This is the preannihilator of $F \subset X^*$. See Lemma 9.6.4. In the reflexive case ${}^\perp F = F^\perp$, and in the nonreflexive case ${}^\perp F = F^\perp \cap X$.

A^*	The (anti-linear) dual of the operator A. See Section 3.5.
$A \geq 0$	A is (self-adjoint and) positive definite.
$A \gg 0$	$A \geq \epsilon$ for some $\epsilon > 0$, hence A is invertible.
$\mathcal{D}(A)$	The domain of the (unbounded) operator A.
$\mathcal{R}(A)$	The range of the operator A.
$\mathcal{N}(A)$	The null space (kernel) of the operator A.
$\mathrm{rank}(A)$	The rank of the operator A.
$\dim(X)$	The dimension of the space X.
$\rho(A)$	The resolvent set of A (see Definition 3.2.7). The resolvent set is always open.
$\sigma(A)$	The spectrum of A (see Definition 3.2.7). The spectrum is always closed.
$\sigma_p(A)$	The point spectrum of A, or equivalently, the set of eigenvalues of A (see Definition 3.2.7).
$\sigma_r(A)$	The residual spectrum of A (see Definition 3.2.7).
$\sigma_c(A)$	The continuous spectrum of A (see Definition 3.2.7).
$\omega_{\mathfrak{A}}$	The growth bound of the semigroup $\mathfrak{A}$. See Definition 2.5.6.
TI, TIC	TI stands for the set of all time-invariant, and TIC stands for the set of all time-invariant and causal operators. See Definition 2.6.2 for details.
$A\&B$, $C\&D$	$A\&B$ stands for an operator (typically unbounded) whose domain $\mathcal{D}(A\&B)$ is a subspace of the cross-product $\left[\begin{smallmatrix} X \\ U \end{smallmatrix}\right]$ of two Banach spaces X and U, and whose values lie in a third Banach space Z. If $\mathcal{D}(A\&B)$ splits into $\mathcal{D}(A\&B) = X_1 \dotplus U_1$ where $X_1 \subset X$ and $U_1 \subset U$, then $A\&B$ can be written in block matrix form as $A\&B = [A \;\; B]$, where $A = A\&B_{\vert X_1}$ and $B = A\&B_{\vert U_1}$. We alternatively write these identities in the form $Ax = A\&B\left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right]$ and $Bu = A\&B\left[\begin{smallmatrix} 0 \\ u \end{smallmatrix}\right]$, interpreting $\mathcal{D}(A\&B)$ as the cross-product of X_1 and U_1.

Special Banach spaces

U	Frequently the input space of the system.
X	Frequently the state space of the system.
Y	Frequently the output space of the system.
X_n	Spaces constructed from the state space X with the help of the generator of a semigroup $\mathfrak{A}$. In particular, X_1 is the domain of the semigroup generator. See Section 3.6.
X_n^*	$X_n^* := (X^*)_n = (X_{-n})^*$. See Remark 3.6.1.
$\dotplus$	$X = X_1 \dotplus X_2$ means that the Banach space X is the direct sum of X_1 and X_2, i.e., both X_1 and X_2 are closed subspaces

	of X, and every $x \in X$ has a unique representation of the form $x = x_1 + x_2$ where $x_1 \in X_1$ and $x_2 \in X_2$.
$\oplus$	$X = X_1 \oplus X_2$ means that the Hilbert space X is the orthogonal direct sum of the Hilbert spaces X_1 and X_2, i.e., $X = X_1 \dotplus X_2$ and $X_1 \perp X_2$.
$\left[\begin{smallmatrix} X \\ Y \end{smallmatrix}\right]$	The cross-product of the two Banach spaces X and Y. Thus, $\left[\begin{smallmatrix} X \\ Y \end{smallmatrix}\right] = \left[\begin{smallmatrix} X \\ 0 \end{smallmatrix}\right] \dotplus \left[\begin{smallmatrix} 0 \\ Y \end{smallmatrix}\right]$.

Special functions

χ_I	The characteristic function of the set I.
1_+	The Heaviside function: $1_+ = \chi_{\mathbb{R}^+}$. Thus $(1_+)(t) = 1$ for $t \geq 0$ and $(1_+)(t) = 0$ for $t < 0$.
B	The Beta function (see (5.3.1)).
Γ	The Gamma function (see (3.9.7)).
e_ω	$e_\omega(t) = e^{\omega t}$ for $\omega, t \in \mathbb{R}$.
log	The natural logarithm.

Function spaces

$V(J;U)$	Functions of type V (= L^p, BC, etc.) on the interval $J \subset \mathbb{R}$ with range in U.
$V_{\mathrm{loc}}(J;U)$	Functions which are locally of type V, i.e., they are defined on $J \subset \mathbb{R}$ with range in U and they belong to $V(K;U)$ for every bounded subinterval $K \subset J$.
$V_c(J;U)$	Functions in $V(J;U)$ with bounded support.
$V_{c,\mathrm{loc}}(J;U)$	Functions in $V_{\mathrm{loc}}(J;U)$ whose support is bounded to the left.
$V_{\mathrm{loc},c}(J;U)$	Functions in $V_{\mathrm{loc}}(J;U)$ whose support is bounded to the right.
$V_0(J;U)$	Functions in $V(J;U)$ vanishing at $\pm\infty$. See also the special cases listed below.
$V_\omega(J;U)$	The set of functions u for which $(t \mapsto e^{-\omega t}u(t)) \in V(J;U)$. See also the special cases listed below.
$V_{\omega,\mathrm{loc}}(\mathbb{R};U)$	The set of functions $u \in V_{\mathrm{loc}}(\mathbb{R};U)$ which satisfy $\pi_- u \in V_\omega(\mathbb{R}^-;U)$.
$V(\mathbb{T}_T;U)$	The set of T-periodic functions of type V on $\mathbb{R}$. The norm in this space is the V-norm over one arbitrary interval of length T.
BC	Bounded continuous functions; sup-norm.
BC_0	Functions in BC that tend to zero at $\pm\infty$.
BC_ω	Functions u for which $(t \mapsto e^{-\omega t}u(t)) \in BC$.

$BC_{\omega,\mathrm{loc}}(\mathbb{R};U)$	Functions $u \in C(\mathbb{R};U)$ which satisfy $\pi_- u \in BC_\omega(\mathbb{R}^-;U)$.
$BC_{0,\omega}$	Functions u for which $(t \mapsto \mathrm{e}^{-\omega t}u(t)) \in BC_0$.
$BC_{0,\omega,\mathrm{loc}}(\mathbb{R};U)$	Functions $u \in C(\mathbb{R};U)$ which satisfy $\pi_- u \in BC_{0,\omega}(\mathbb{R}^-;U)$.
BUC	Bounded uniformly continuous functions; sup-norm.
BUC^n	Functions which together with their n first derivatives belong to BUC. See Definition 3.2.2.
C	Continuous functions. The same space as BC_{loc}.
C^n	n times continuously differentiable functions. The same space as BC^n_{loc}.
C^∞	Infinitely many times differentiable functions. The same space as BC^∞_{loc}.
L^p, $1 \le p < \infty$	Strongly measurable functions with norm $\left\{\int |u(t)|^p\,dt\right\}^{1/p}$.
L^∞	Strongly measurable functions with norm $\operatorname{ess\,sup}|u(t)|$.
L^p_0	$L^p_0 = L^p$ if $1 \le p < \infty$, and L^∞_0 consists of those $u \in L^\infty$ which vanish at $\pm\infty$, i.e., $\lim_{t\to\infty} \operatorname{ess\,sup}_{|s|\ge t}|u(s)| = 0$.
L^p_ω	Functions u for which $(t \mapsto \mathrm{e}^{-\omega t}u(t)) \in L^p$.
$L^p_{\omega,\mathrm{loc}}(\mathbb{R};U)$	Functions $u \in L^p_{\mathrm{loc}}(\mathbb{R};U)$ which satisfy $\pi_- u \in L^p_\omega(\mathbb{R}^-;U)$.
$L^p_{0,\omega}$	Functions u for which $(t \mapsto \mathrm{e}^{-\omega t}u(t)) \in L^p_0$.
$L^p_{0,\omega,\mathrm{loc}}(\mathbb{R};U)$	Functions $u \in L^p_{\mathrm{loc}}(\mathbb{R};U)$ which satisfy $\pi_- u \in L^p_{0,\omega}(\mathbb{R}^-;U)$.
$W^{n,p}$	Functions which together with their n first (distribution) derivatives belong to L^p. See Definition 3.2.2.
Reg	Bounded right-continuous functions which have a left hand limit at each finite point.
Reg_0	Functions in Reg which tend to zero at $\pm\infty$.
Reg_ω	The set of functions u for which $(t \mapsto \mathrm{e}^{-\omega t}u(t)) \in Reg$.
$Reg_{\omega,\mathrm{loc}}(\mathbb{R};U)$	The set of functions $u \in Reg_{\mathrm{loc}}(\mathbb{R};U)$ which satisfy $\pi_- u \in Reg_\omega(\mathbb{R}^-;U)$.
$Reg_{0,\omega}$	The set of functions u for which $(t \mapsto \mathrm{e}^{-\omega t}u(t)) \in Reg_0$.
$Reg_{0,\omega,\mathrm{loc}}(\mathbb{R};U)$	Functions $u \in Reg_{\mathrm{loc}}(\mathbb{R};U)$ which satisfy $\pi_- u \in Reg_{0,\omega}(\mathbb{R}^-;U)$.
Reg^n	Functions which together with their n first derivatives belong to Reg. See Definition 3.2.2.
$L^p\vert Reg$	This stands for either L^p or Reg, whichever is appropriate.