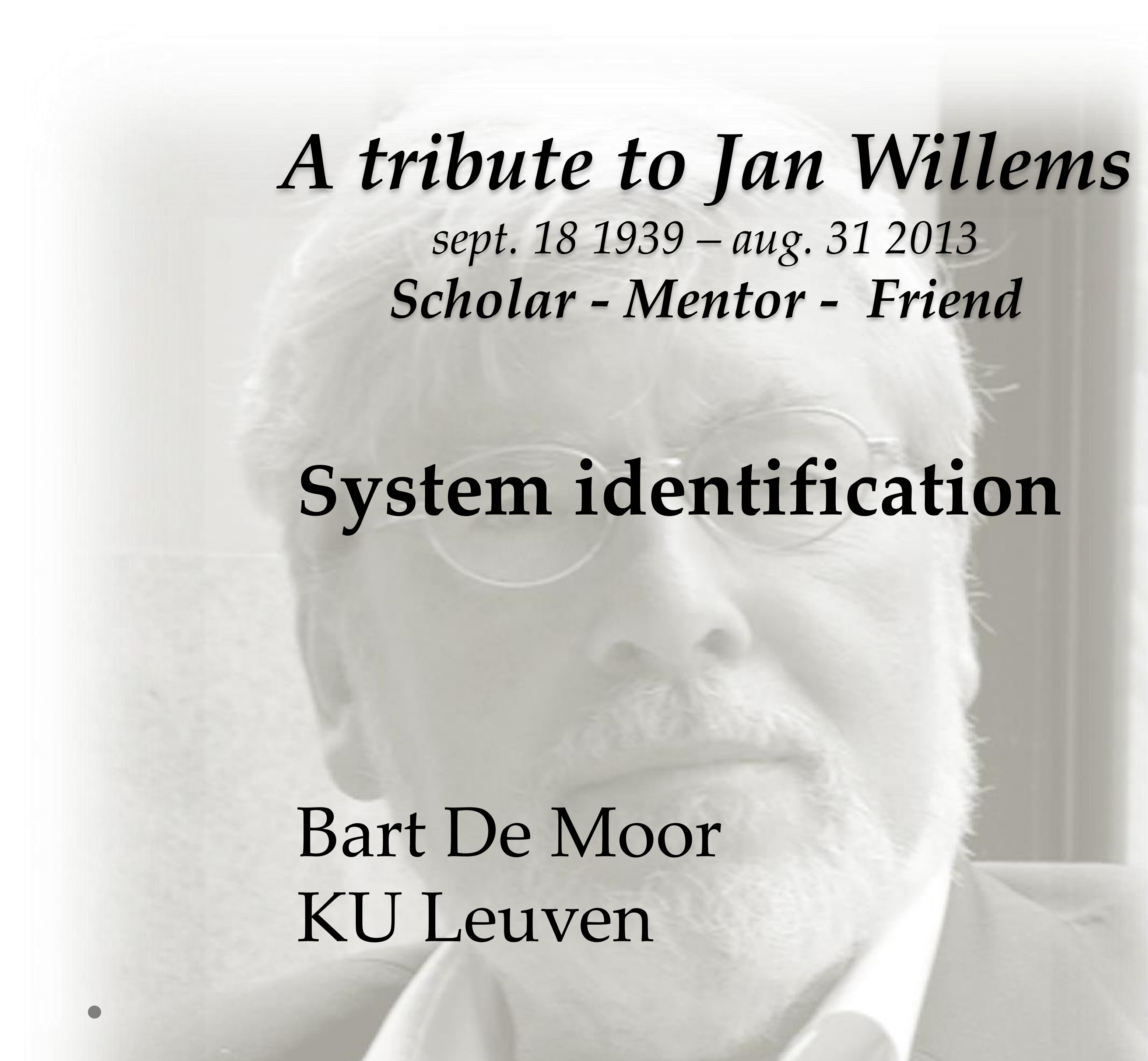


A tribute to Jan Willems

sept. 18 1939 – aug. 31 2013

Scholar - Mentor - Friend

System identification

Bart De Moor
KU Leuven

Outline

- 10 years ago

- My first encounter with Jan

- The three famous Automatica papers and their impact
  - Behaviors
  - Exact modelling
  - Approximate modeling

- Remembering Jan

JOHANN BERNOULLI LECTURE

Prof.dr.ir. Bart De Moor

Systems biology: A new mathematical frontier

Tuesday 23 March 2004
7.30 - 10 p.m.
Aula, Academiegebouw
Broerstraat 5, Groningen
Admission free

Information:
Department of Mathematics and Computing Science
University of Groningen
(050) 363 3939
www.rugnl/wiskunde/nieuws/agenda

The Johann Bernoulli Lecture is organised by the Johann Bernoulli Stichting in cooperation with Studium Generale Groningen.

Come forth into the light of things
Let nature be your teacher
Wordsworth

At first blush, mathematics on the one hand and biology on the other, are an unlikely pairing: abstract, symbolic-numeric computation, versus 'wet', evolving and living organisms. However, we find that there is a great abundance of mathematical structure in biological objects, from fractals found in the branches of an oak tree to the symmetries of DNA's double helix. Throughout history, mathematicians have been fascinated by biology. the classic studies of inheritance by Gregor Mendel, were an exercise not in biology, but in statistical inference. Alan Turing studied the morphogenesis of embryos invoking reaction-diffusion equations. And Erwin Schrödinger in *What's life?* envisioned life as an aperiodic crystal.
Likewise, biology has inspired mathematicians and engineers to work for instance on massive analog synaptic computing, neural networks and machine learning. Or on genetic algorithms and evolutionary programming techniques that mimic Darwin's 'survival-of-the-fittest'.

In recent years, biology and mathematics have undergone an increasingly intense interdisciplinary merger, leading to new scientific fields such as *bioinformatics* and *systems biology*. Several major breakthroughs have catalyzed this development. In biology for instance, the
description of the double helix 50 years ago, and the unraveling of the human genome has provided us with some deep insights into the mathematical codes of life itself. In engineering, new devices have been developed, such as microarrays and DNA chips, that allow for the simultaneous measurement of the expression levels of thousands of genes. Moreover, global connectivity over the World Wide Web ensures that we can access biological databases, the floating point operations. This has resulted in new emerging scientific fields such as functional and structural genomics, computational biomedicine, metabolic pathway analysis, combinatorial drug design, systems biology, evolutionary modeling and phylogenetic footprinting, and a whole collection of neologisms ending in 'omics' such as transcriptomics, proteomics, metabolomics and '-omes' such as the transcriptome, the proteome, the metabolome, the interactome etc…

In this lecture, we will guide the audience through the historical origins of bioinformatics and systems biology, by elaborating on those major breakthroughs. We will illustrate the presentation with many examples of current mathematical and biological challenges, including disease management, computational biomedicine and diagnostics in oncology, unraveling the mitotic cycle of yeast and motif detection in DNA sequences of plants.

Bart De Moor (1960) obtained in 1983 his Master Degree in Electrical Engineering at the Katholieke Universiteit Leuven, Belgium, and a PhD in Engineering at the same university in 1988. He spent two years as a Visiting Research Associate at Stanford University (1988-1990) at the departments of Electrical Engineering (ISL, Prof. Kailath) and Computer Science (Prof. Golub). Currently, he is a full professor at the Department of Electrical Engineering in Leuven. Currently, he is leading a research group of 39 PhD students and 8 postdocs and in the recent past, 16 PhDs were obtained under his guidance. He has been teaching at and been a member of PhD jury's in several universities in Europe and the US.
His work has won him several scientific awards (Leybold-Heraeus Prize (1986), Leslie Fox Prize (1989), Guillemin-Cauer best paper Award of the IEEE Transaction on Circuits and Systems (1990), Laureate of the Belgian Royal Academy of Sciences (1992), bi-annual Siemens Award (1994), best paper award of Automatica (IFAC, 1996), IEEE Signal Processing Society Best Paper Award (1999). He is an associate editor of several scientific journals.
From 1991-1999 he was the chief advisor on Science and Technology of several ministers of the Belgian Federal Government and the Flanders Regional Governments.
He was and/or is in the board of 3 spin-off companies (www.ipcos.be, www.data4s.com, www.tml.be), of the Flemish Interuniversity Institute for Biotechnology, the Study Center for Nuclear Energy, and several other scientific and cultural organizations. He was a member of the Academic Council of the Katholieke Universiteit Leuven, and still is a member of its Research Policy Council. Since 2002 he also makes regular television appearances in the Science Show 'Hoe?Zo!' on national television in Belgium.
Full details on his CV can be found at www.esat.kuleuven.ac.be/~demoor

1983: First encounter

- What are data/measurements/observations ?
- What are models ? Deduction, induction, inspiration
- What is noise ?
- What are statistical assumptions worth ?
- What are stochastic systems ?

- Kalman: the Frisch scheme
- Least squares
- Mathematical rigor; accuracy and precision

KATHOLIEKE UNIVERSITEIT LEUVEN

Fakulteit der Toegepaste Wetenschappen

Departement Elektrotechniek
Kard. Mercierlaan 94, 3030 Heverlee

Mathematical Concepts and Techniques for Modelling of Static and Dynamic Systems

Jury:
Prof.Dr.Ir. J. Delrue, vice-dekaan, voorzitter
Prof.Dr.Ir. J. Vandewalle, promotor
Prof.Dr. A. Bultheel
Dr.Ir. A. Barbé
Prof.Dr.Ir. J. Peperstraete
Prof.Dr.Ir. H. Van Brussel
Prof.Dr.Ir. M. Gevers

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de Toegepaste Wetenschappen
door
Bart De Moor

UDC 519.71

Juni 1988

Chapter 5

Identification of Linear Relations in Noisy Data

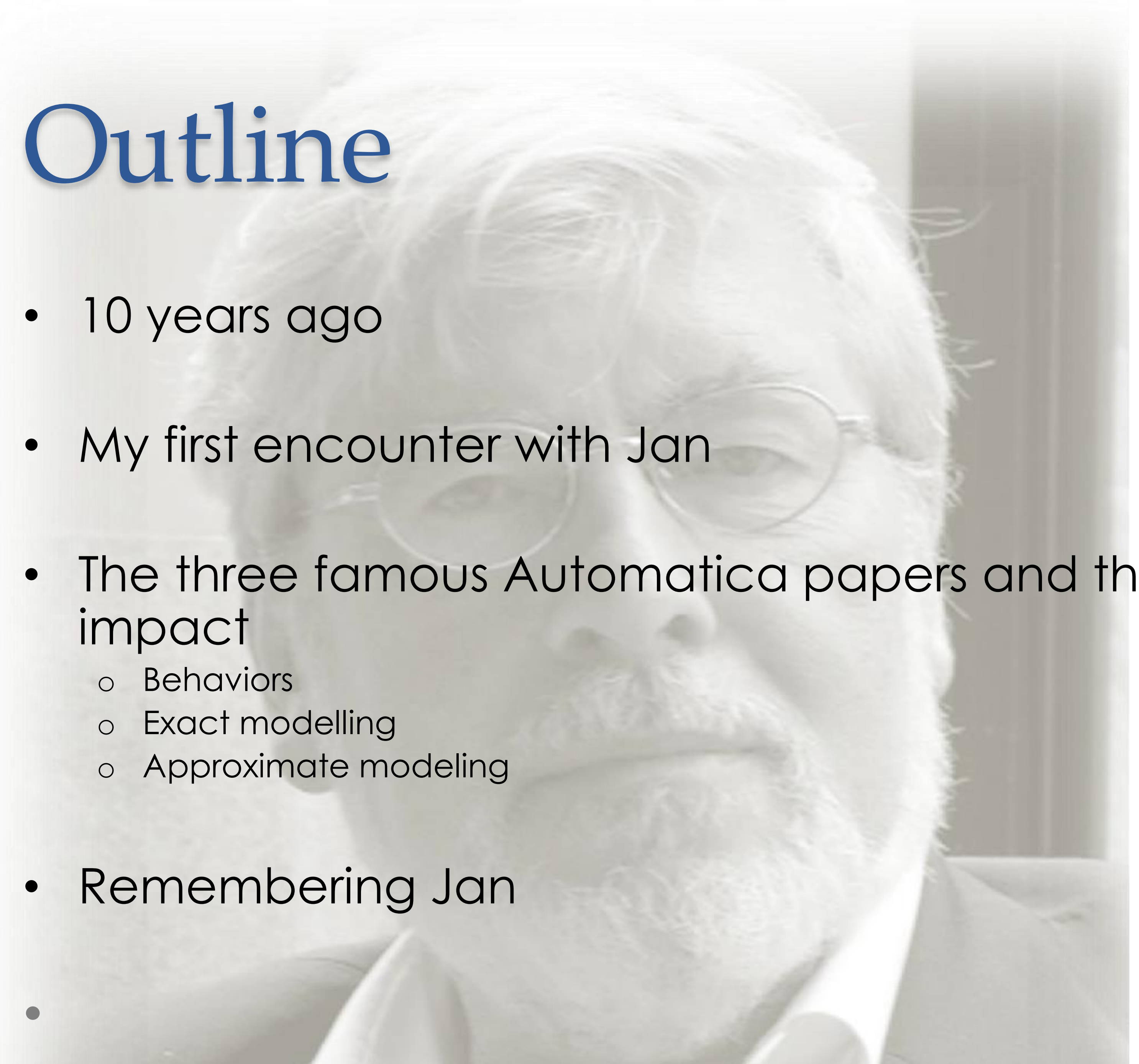
Models are a matter of inspiration,
not of deduction.
Jan Willems.

The reader is referred to some recent papers by Kalman (1982, 1983) in which he discusses (mainly in an econometrics context) the problem of modelling on the basis of data and in which he argues the limited value of the statistician's stochastic approach.

The Frisch scheme:

Given a positive definite $n \times n$ matrix Σ. Find all nonnegative diagonal matrices $\tilde{\Sigma}$ and all n-vectors x such that:

1. $\hat{\Sigma} = \Sigma - \tilde{\Sigma}$ is nonnegative definite
2. $\text{corank}(\hat{\Sigma})$ is maximal
3. $\hat{\Sigma} x = 0$



Outline

- 10 years ago
- My first encounter with Jan
- The three famous Automatica papers and their impact
  - Behaviors
  - Exact modelling
  - Approximate modeling
- Remembering Jan

Automatica, Vol. 22, No. 5, pp. 561–580, 1986
Printed in Great Britain.

0005–1098/86 $3.00 + 0.00
Pergamon Journals Ltd.

From Time Series to Linear System—Part I. Finite Dimensional Linear Time Invariant Systems*

JAN C. WILLEMS†

Automatica, Vol. 22. No. 6, pp. 675–694, 1986
Printed in Great Britain.

0005–1098/86 $3.00 + 0.00
Pergamon Journals Ltd.

From Time Series to Linear System—Part II. Exact Modelling*

JAN C. WILLEMS†

Automatica, Vol. 23, No. 1, pp. 87–115, 1987
Printed in Great Britain.

0005–1098/87 $3.00+0.00
Pergamon Journals Ltd.

From Time Series to Linear System—Part III. Approximate Modelling*

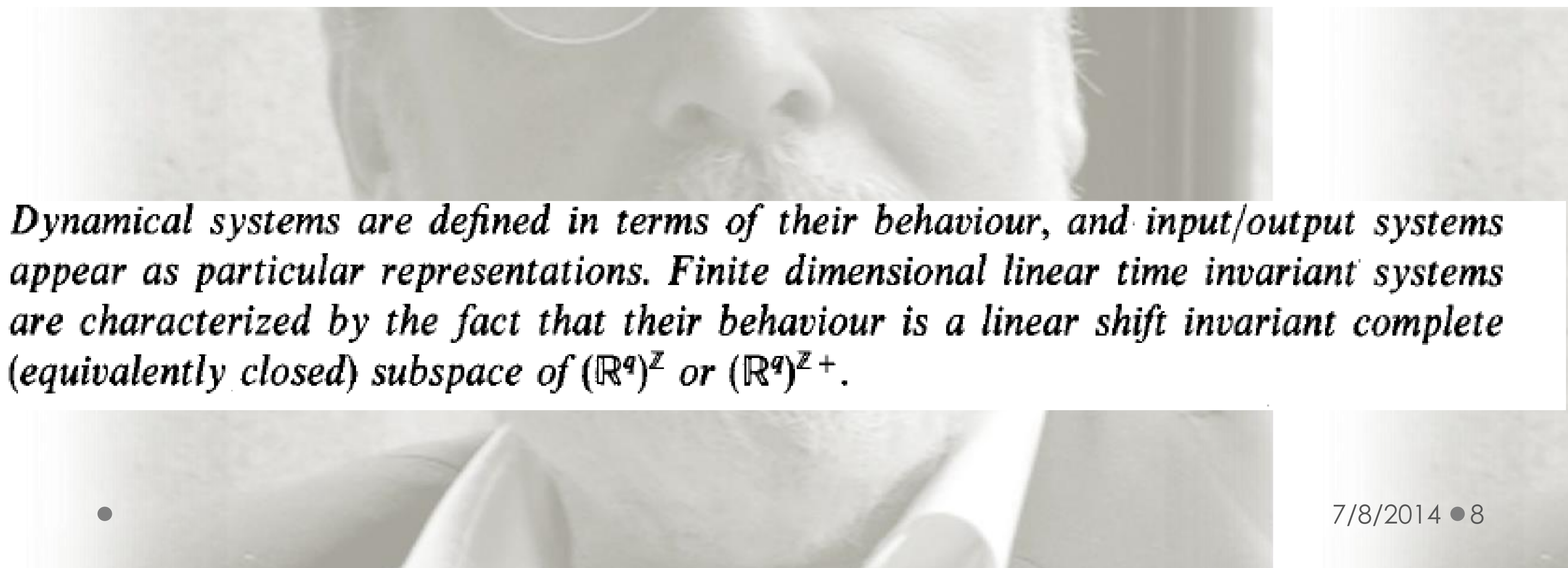
JAN C. WILLEMS†

Automatica, Vol. 22, No. 5, pp. 561–580, 1986
Printed in Great Britain.

0005–1098/86 $3.00 + 0.00
Pergamon Journals Ltd.

From Time Series to Linear System—Part I. Finite Dimensional Linear Time Invariant Systems*

JAN C. WILLEMS†



Dynamical systems are defined in terms of their behaviour, and input/output systems appear as particular representations. Finite dimensional linear time invariant systems are characterized by the fact that their behaviour is a linear shift invariant complete (equivalently closed) subspace of $(\mathbb{R}^q)^{\mathbb{Z}}$ *or* $(\mathbb{R}^q)^{\mathbb{Z}_+}$.

Abstract—In the first part of this paper the definition of a dynamical system as simply consisting of a family of time series will be developed. In this context the notions of linearity, time invariance, and finite dimensionality will be introduced. It will be shown that a given family of time series may be represented by a system of (AR) equations: $R_l\mathbf{w}(t+l) + R_{l-1}\mathbf{w}(t+l-1) + \cdots + R_0\mathbf{w}(t) = 0$, or, equivalently, by a finite dimensional linear time invariant system: $\mathbf{x}(t+1) = A\mathbf{x}(t) + B\mathbf{u}(t)$; $\mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t)$; $\mathbf{w} = (\mathbf{u}, \mathbf{y})$, if and only if this family is *linear*, *shift invariant* and *complete* (or, as is equivalent, closed in the topology of pointwise convergence). This yields a very high level and elegant set of axioms which characterize these familiar objects. It is emphasized, however, that no *a priori* choice is made as to which components of $\mathbf{w}$ are inputs and which are outputs. Such a separation always exists in any specific linear time invariant model. Starting from these definitions, the structural indices of such systems are introduced and it is shown how an (AR) representation of a system having a given behaviour can be constructed. These results will be used in a modelling context in Part II of the paper.

The totality of possible events (*before* we have modeled the phenomenon) forms a set $\mathbb{U}$, called the *universum*. A *mathematical model* of the phenomenon restricts the outcomes that are declared possible to a subset $\mathscr{B}$ of $\mathbb{U}$; $\mathscr{B}$ is called the *behavior* of the model. I refer to $(\mathbb{U}, \mathscr{B})$ (or to $\mathscr{B}$ by itself, since $\mathbb{U}$ usually follows from the context) as a mathematical model.

In the study of dynamical systems we are, more specifically, interested in situations where the events are signals, trajectories, i.e. maps from a set of *independent variables* (time, or space, or time and space) to a set of *dependent variables* (the values taken on by the signals). In this case the universum is the collection of all maps from the set of independent variables to the set of dependent variables. It is convenient to distinguish these sets explicitly in the notation: $\mathbb{T}$ for the set of independent variables, and $\mathbb{W}$ for the set of dependent variables. $\mathbb{T}$ suggests 'time', but in distributed parameter systems $\mathbb{T}$ is often time and space. I have incorporated distributed systems because of their importance in engineering models. Whence a (dynamical) *system* is defined as a triple

$$\boxed{\Sigma = (\mathbb{W}, \mathbb{T}, \mathscr{B})}$$

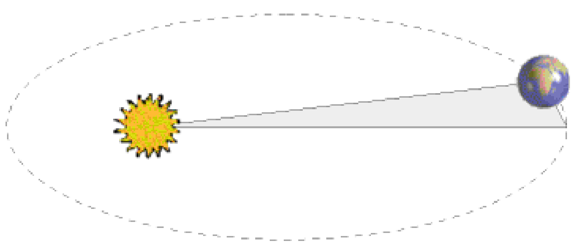
with $\mathscr{B}$, *the behavior,* a subset of $\mathbb{W}^{\mathbb{T}}$ ($\mathbb{W}^{\mathbb{T}}$ is the standard mathematical notation for the set of all maps from $\mathbb{T}$ to $\mathbb{W}$). The behavior is the central object in this definition. It formalizes which signals $w : \mathbb{T} \to \mathbb{W}$ are possible, according to the model: those in $\mathscr{B}$, and which are not: those not in $\mathscr{B}$.

4. *Kepler's laws* describe the possible motions of the planets in the solar system. This defines a dynamical system with $\mathbb{T} = \mathbb{R}$, $\mathbb{W} = \mathbb{R}^3$, and $\mathscr{B}$ the set of maps $w : \mathbb{R} \to \mathbb{R}^3$ that satisfy Kepler's laws: 1. the orbits must be ellipses in $\mathbb{R}^3$ with the sun (assumed in fixed position, say the origin of $\mathbb{R}^3$) in one of the foci; 2. the radius vector from the sun to the planet must sweep out equal areas in equal time, and 3. the ratio of the period of revolution around the ellipse to the major axis must be the same for all w's in $\mathscr{B}$.

Law 1: Orbit is ellips with Sun in focus

Law 2: Radius vector sweeps out equal areas in equal time

Law 3: $\frac{T_1^2}{T_2^2} = \frac{a_1^3}{a_2^3}$



From conic sections to centripetal forces and states

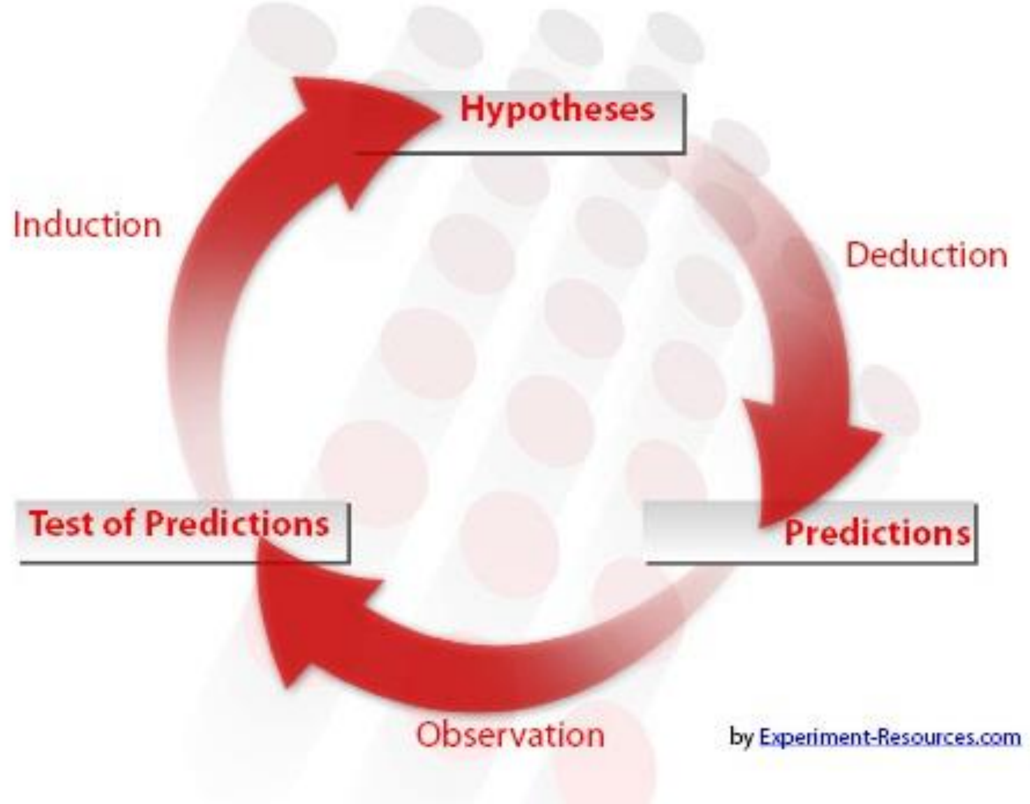
$$\mathbf{F} = m.\mathbf{a}$$

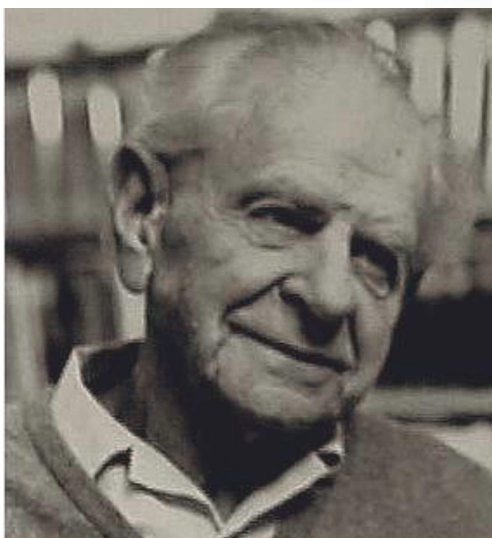
$$F = G\frac{m.M}{r^2}$$

1. *Newton's second law* imposes a restriction that relates the position $\vec{q}$ of a point mass with mass m to the force $\vec{F}$ acting on it: $\vec{F} = m\frac{d^2}{dt^2}\vec{q}$. This is a dynamical system with $\mathbb{T} = \mathbb{R}, \mathbb{W} = \mathbb{R}^3 \times \mathbb{R}^3$ (typical elements of $\mathscr{B}$ are maps $(\vec{q}, \vec{F}) : \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}^3$), and behavior $\mathscr{B}$ consisting of all maps $t \in \mathbb{R} \mapsto (\vec{q}, \vec{F})(t) \in \mathbb{R}^3 \times \mathbb{R}^3$ that satisfy $\vec{F} = m\frac{d^2}{dt^2}\vec{q}$. I do not specify the precise sense of what it means that a function satisfies a differential equation (I will pay no attention to such secondary issues).

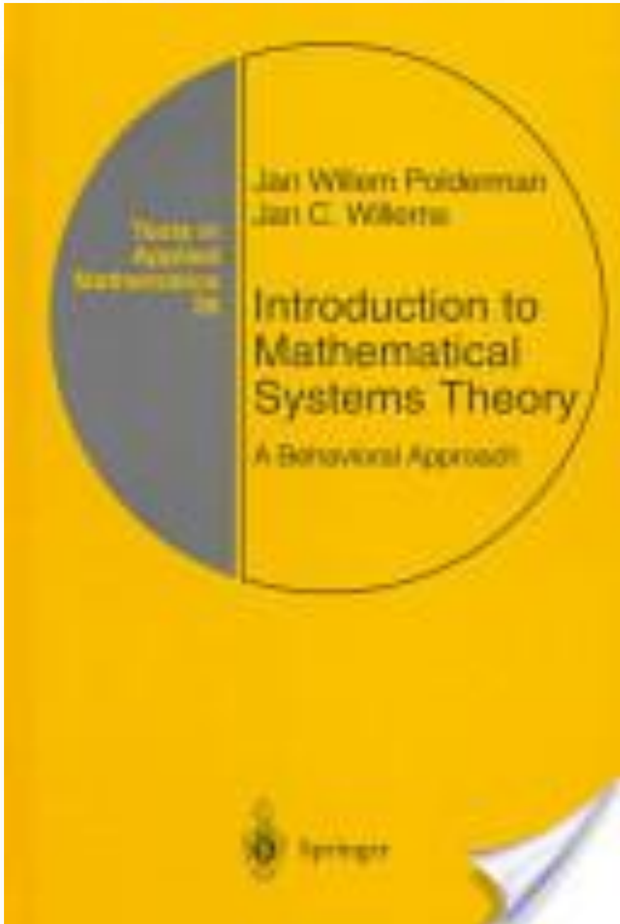
THE CLASSICAL view of modelling is the descriptive one of the physicist: nature functions consistently according to some universal laws and the task is to discover them.

In some cases, it may, in principle, be possible to obtain such laws by deduction or extrapolation from observed data. In Willems (1979) it has been shown how one can for example view Newton's laws as a logical extrapolation of Kepler's laws. However, the practice in the descriptive sciences is really not this; it is much more the concept of *falsification* than that of *deduction* which is the central idea. This observation, in fact, has formed the cornerstone of the philosophy of science since Popper. Models and laws are postulated, often on the basis of an Aristotelian philosophical view and aesthetic appeal, and it is only later that one discovers that, to some extent, they could also have been deduced from already existing knowledge and observed facts. In this sense, models are obtained neither by deduction, nor by induction, but by inspiration.




Sir Karl Popper (1902-1994)

† Cf. Popper (1963, p. 36). "Every good scientific theory is prohibition, it forbids certain things to happen. The more it forbids, the better it is. A theory which is not refutable by any conceivable event is non-scientific. Irrefutability is not a virtue of a theory (as people often think) but a vice."



The Behavioral Approach to Open and Interconnected Systems

JAN C. WILLEMS

MODELING BY TEARING, ZOOMING, AND LINKING

During the opening lecture of the 16th IFAC World Congress in Prague on July 4, 2005, Rudy Kalman articulated a principle that resonated very well with me. He put forward the following paradigm for research domains that combine models and mathematics:

1) Get the physics right.
2) The rest is mathematics.

Did we, system theorists, get the physics right? Do our basic model structures adequately translate physical reality? Does the way in which we view interconnections respect the physics? These questions, in a nutshell, are the theme of this article.

The motivation for the behavioral approach stems from the observation that classical system-theoretic thinking is unsuitable for dealing on an appropriately general level with the basic tenets at which system theory aims, namely, open and interconnected systems. By an *open* system, we mean a system that interacts with its environment, for example, by exchanging matter, energy, or information. By an *interconnected* system, we mean a system that consists of interacting subsystems. Classical system theory introduces inputs, outputs, and signal-flow graphs *ab initio.* Inputs serve to capture the influence of the environment on the system, outputs serve to capture the influence of the system on the environment, while output-to-input assignments, such as series and feedback connection, serve to capture interconnections. A system is thus viewed as transmitting and transforming signals from the input channel to the output channel, and interconnections are viewed as pathways through which outputs of one system are imposed as inputs to another system.

Laws that govern physical phenomena, however, merely impose relations on the system variables, while interconnection means that variables are shared among subsystems. For

TEAR

Digital Object Identifier 10.1109/MCS.2007.906923

1066-033X/07/$25.00©2007IEEE

Automatica, Vol. 22. No. 6, pp. 675–694, 1986
Printed in Great Britain.

0005–1098/86 \$3.00 + 0.00
Pergamon Journals Ltd.

From Time Series to Linear System—Part II. Exact Modelling*

JAN C. WILLEMS†

The most powerful unfalsified model is defined as that element in a model class which explains a given set of observations and as little else as possible. Algorithms are developed which compute the most powerful unfalsified linear time invariant model starting from an observed time series.

ONE OF THE central issues in the modelling of dynamical phenomena may succinctly be formulated as follows:

Given an observed q-dimensional vector time series

$$\tilde{\mathbf{w}}(t_0), \tilde{\mathbf{w}}(t_0 + 1), \ldots, \tilde{\mathbf{w}}(t_1) \qquad (-\infty \leqslant t_0 \leqslant t \leqslant t_1 \leqslant \infty)$$

with $\tilde{\mathbf{w}}(t) \in \mathbb{R}^q$, find a dynamical model which explains these observations.

as well. Before spelling out the algorithm the notion of the relative row rank $r(M_1; M_2)$ of a partitioned (infinite) matrix $M = \begin{bmatrix} M_1 \\ \hdashline M_2 \end{bmatrix}$ is introduced. Assume first that $M = \text{col}(M_1;\ M_2)$ is finite: $M_1 \in \mathbb{R}^{k_1 \times k}$, $M_2 \in \mathbb{R}^{k_2 \times k}$. Then $r(M_1; M_2) := \text{rank}\ M_1 + \text{rank}\ M_2 - \text{rank}\ M$. Next, assume that M has an

Now consider $\tilde{\mathbf{w}}: \mathbb{Z} \to \mathbb{R}^q$ and define the following partitioned (4 way infinite) Hankel matrix

$$\begin{bmatrix} \mathscr{H}_-(\tilde{\mathbf{w}}) \\ \hdashline \mathscr{H}_+(\tilde{\mathbf{w}}) \end{bmatrix} :=$$

$$\begin{bmatrix}
 & \vdots & \vdots & & \vdots & \\
\cdots & \tilde{\mathbf{w}}(-t-1) & \tilde{\mathbf{w}}(-t) & \cdots & \tilde{\mathbf{w}}(0) & \cdots \\
 & \vdots & \vdots & & \vdots & \\
\cdots & \tilde{\mathbf{w}}(-2) & \tilde{\mathbf{w}}(-1) & \cdots & \tilde{\mathbf{w}}(t'-1) & \cdots \\
\cdots & \tilde{\mathbf{w}}(-1) & \tilde{\mathbf{w}}(0) & \cdots & \tilde{\mathbf{w}}(t') & \cdots \\
\hdashline
\cdots & \tilde{\mathbf{w}}(0) & \tilde{\mathbf{w}}(1) & \cdots & \tilde{\mathbf{w}}(t'+1) & \cdots \\
 & \vdots & \vdots & & \vdots & \\
\cdots & \tilde{\mathbf{w}}(t-1) & \tilde{\mathbf{w}}(t) & \cdots & \tilde{\mathbf{w}}(t+t') & \cdots \\
 & \vdots & \vdots & & \vdots &
\end{bmatrix}.$$

The relevance of the motion of relative row rank follows from the following result.

Proposition 20. $r(\mathscr{H}_-(\tilde{\mathbf{w}});\ \mathscr{H}_+(\tilde{\mathbf{w}})) < \infty$. In fact, it equals the dimension of a minimal state space representation $\Sigma_s(A', B', C', D')$ of $\mathscr{B}(R^*_{\tilde{\mathbf{w}}})$, the most powerful unfalsified (AR) model for $\tilde{\mathbf{w}}$.

Exact subspace identification

$$
\begin{aligned}
\dim[R(A) \cap R(B)] &= \dim R(A) + \dim R(B) - \dim R\begin{pmatrix} A & B \end{pmatrix} \\
&= r_A + r_B - r_{AB}
\end{aligned}
$$

$$
R[\, R(W_{0|i-1}^T) \cap R(W_{i|2i-1}^T)\,] = R((X_p^d)^T)
$$

$$
\underbrace{\begin{pmatrix} X_{i+1}^d \\ Y_{i|i} \end{pmatrix}}_{\text{known}} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \underbrace{\begin{pmatrix} X_i^d \\ U_{i|i} \end{pmatrix}}_{\text{known}}
$$

Automatica, Vol. 23, No. 1, pp. 87–115, 1987
Printed in Great Britain.

0005–1098/87 \$3.00+0.00
Pergamon Journals Ltd.

From Time Series to Linear System—Part III. Approximate Modelling*

JAN C. WILLEMS†

An optimal approximate model, defined in terms of the complexity of a model and the misfit between a model and the observed data, yields algorithms for computing an optimally fitting model with a maximal admissible complexity or, alternatively, a minimally complex model which explains an observed time series up to a maximally tolerated misfit.

The complexity can be viewed as the inverse of the power of a model and is hence a quantitative measure for expressing how powerful a model actually is. The misfit $\varepsilon(Z, M)$ indicates how far the model M fails to explain the measurements Z. Large complexity and large misfit are both undesirable properties of a model. Models with large complexity explain too much, while models for which the misfit is large explain the observations poorly and therefore do not inspire much confidence.

Fix the *maximal admissible complexity*, c^{adm}.

In this methodology the optimal approximate model has an allowed complexity level (i) and, within this class, a minimal misfit (ii). However, if there are many models achieving this minimum, then it is logical to choose the one which has smallest complexity (iii). This often induces uniqueness, while (i) and (ii) alone may not.

Fix the *maximal tolerated misfit*, ε^{tol}.

In this methodology the optimal approximate model has a tolerated error level (i) and, within this class, a minimal complexity (ii). However, if there are many models achieving this minimum, then it is logical to choose the one which has smallest misfit (iii). This often induces uniqueness, while (i) and (ii) alone may not.

William van Occam (1290-1349):
“Entia non sunt multiplicanda praeter necessitatem”
(Wezensbegrippen moeten niet onnodig vermeerderd worden)

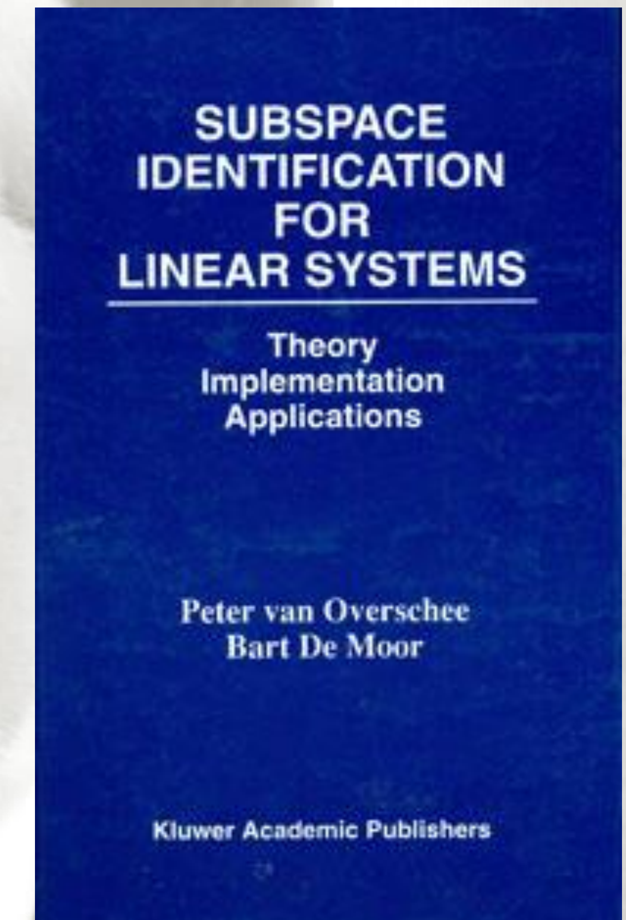
Simpler explanations of the same phenomenon are preferable over complicated ones

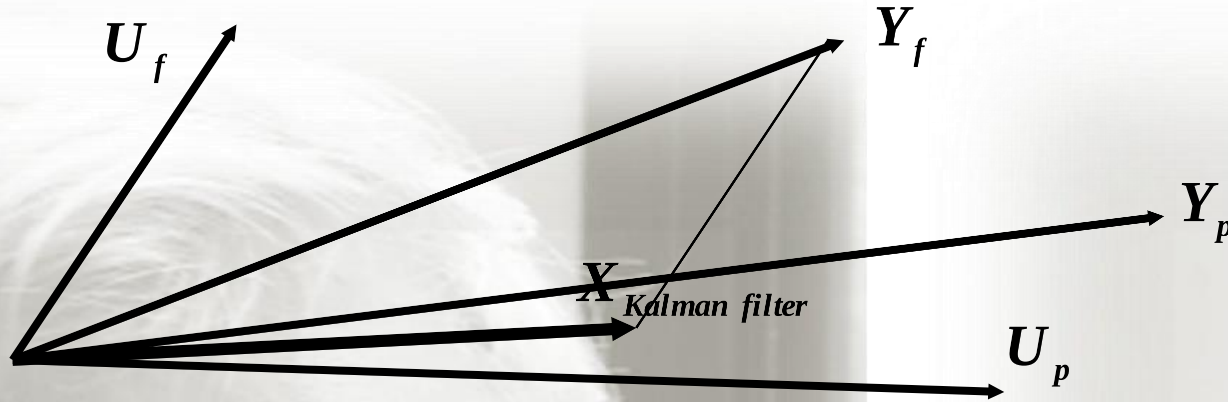
Reality is just another model

Subspace identification

- Subspace identification = Jan Willems + numerical linear algebra (SVD, QR, EVD)

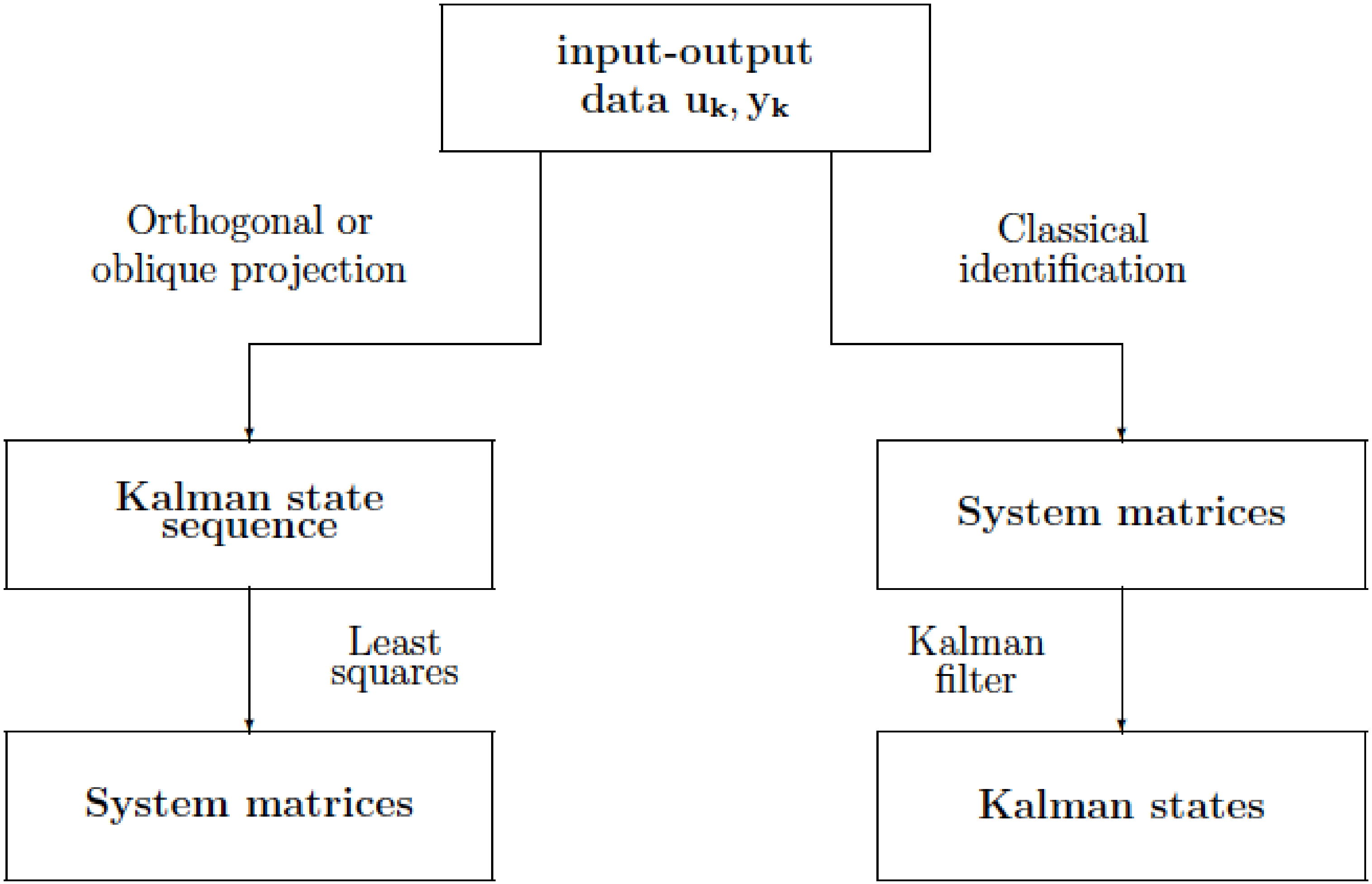
- "The development of subspace identification methods is the most exciting thing that has happened to system identification the last 5 years or so."
Lennart Ljung, ERNSI, LLN, 1993

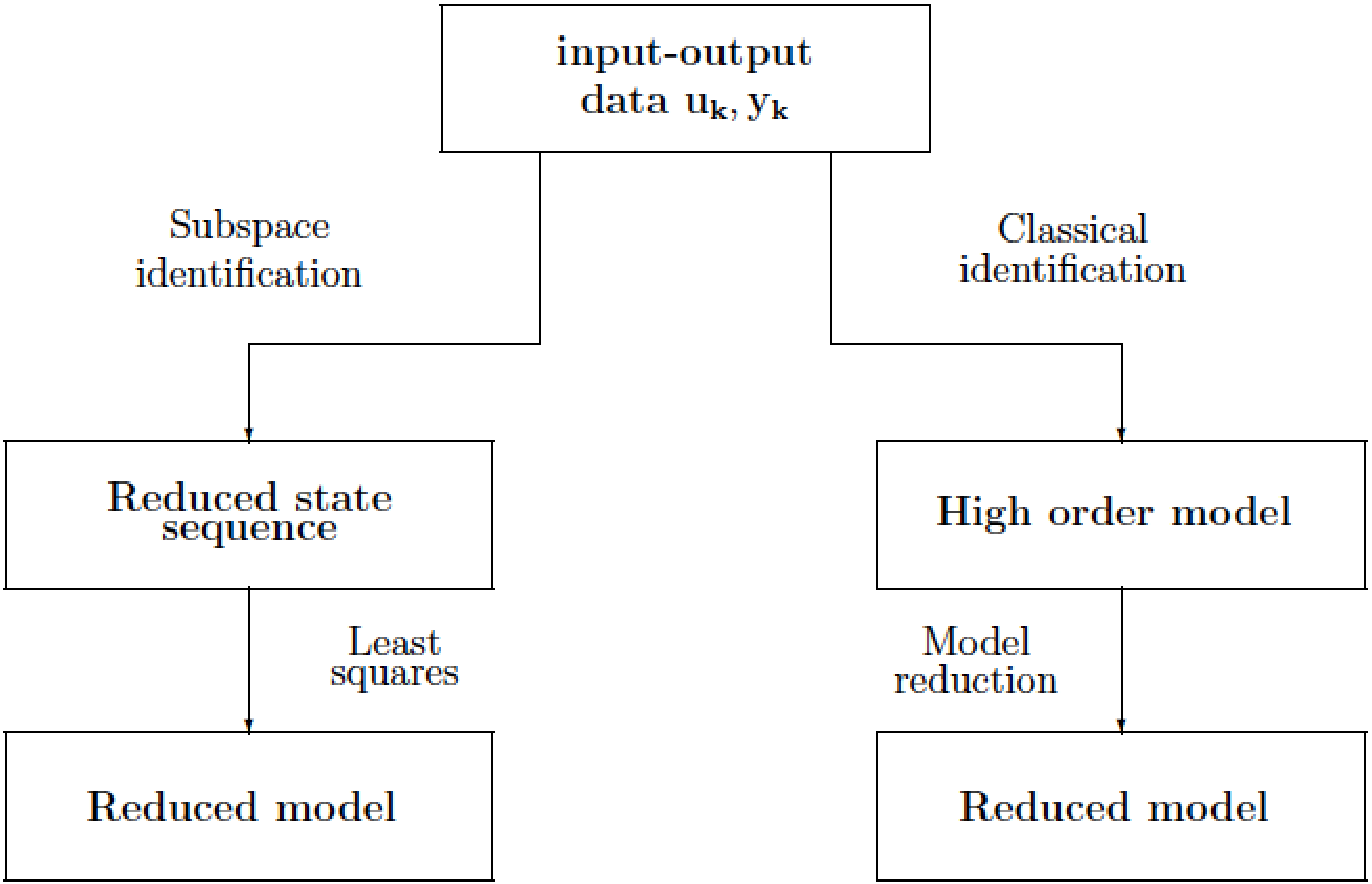




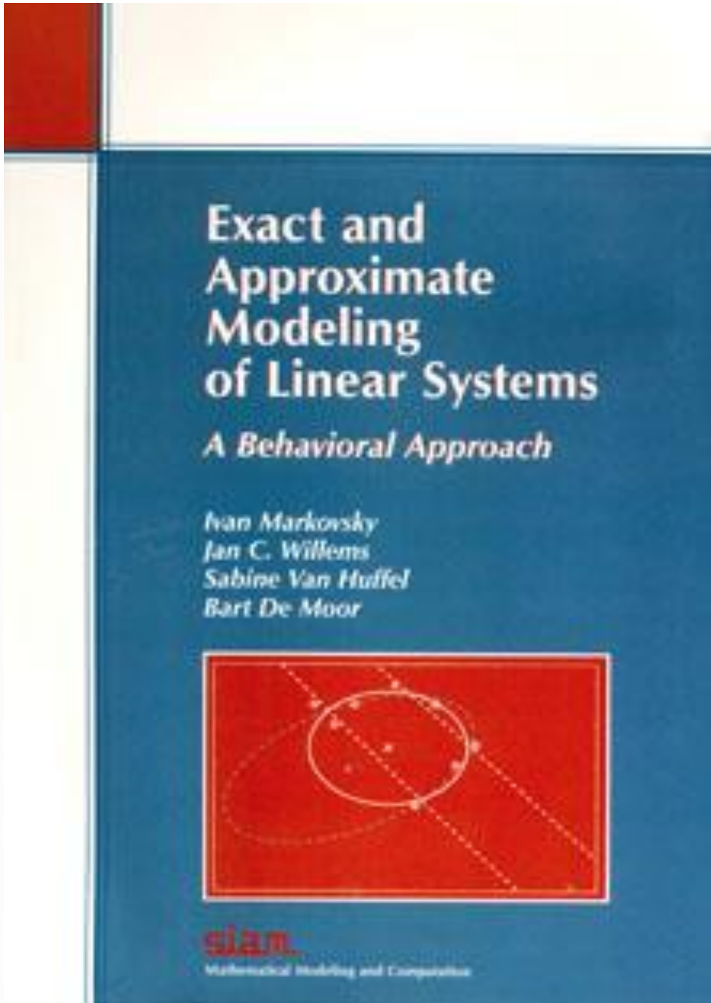
As $j \to \infty$

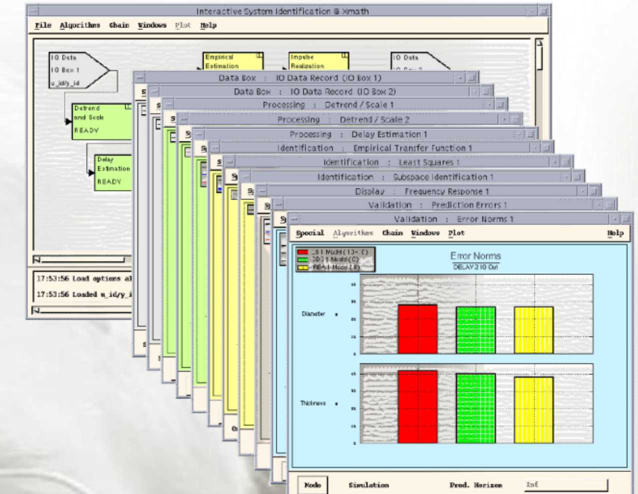
$$\begin{aligned}
\textbf{rank}\,(Y_f/_{U_f}\boldsymbol{W_p}) &= n \\
\textbf{row space}\,(Y_f/_{U_f}\boldsymbol{W_p}) &= \textbf{row space}\,(\widetilde{X}_i) \\
\textbf{column space}\,(Y_f/_{U_f}\boldsymbol{W_p}) &= \textbf{column space}\,(\Gamma_i)
\end{aligned}$$


input-output
data u_k, y_k
Orthogonal or
oblique projection
Classical
identification
Kalman state
sequence
System matrices
Least
squares
Kalman
filter
System matrices
Kalman states



input-output
data u_k, y_k
Subspace
identification
Classical
identification
Reduced state
sequence
High order model
Least
squares
Model
reduction
Reduced model
Reduced model





MathWorks

The toolbox provides identification techniques such as maximum likelihood, prediction-error minimization and subspace system identification.

Process industry

Oil exploration

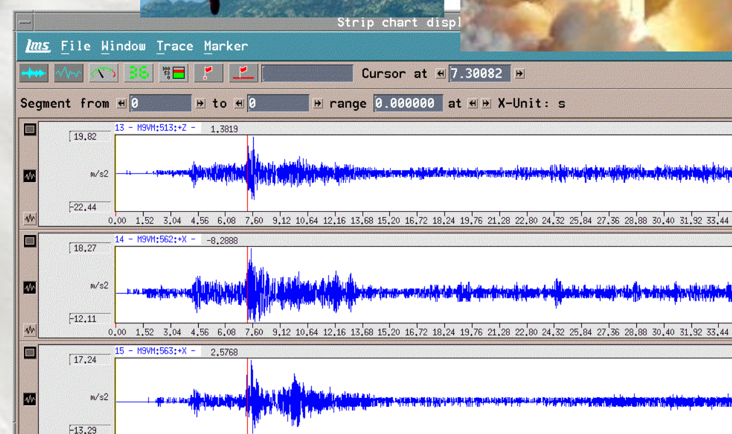
Mechatronics

Biomedical SP

Power

Civil Engineering

Telecommunications

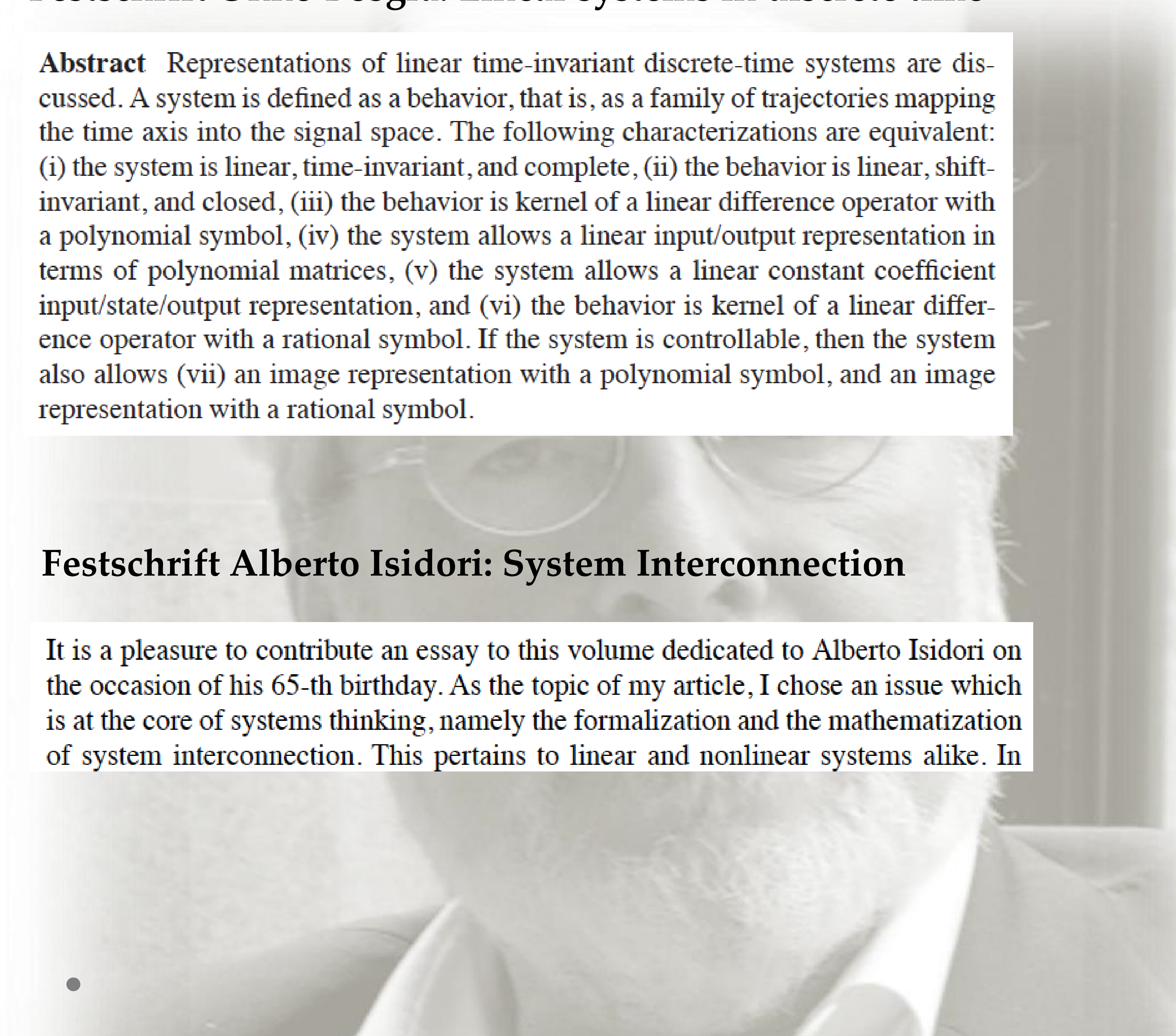


LMS A Siemens Business

IPCOS

Festschrift Okko Bosgra: Linear systems in discrete time

Abstract Representations of linear time-invariant discrete-time systems are discussed. A system is defined as a behavior, that is, as a family of trajectories mapping the time axis into the signal space. The following characterizations are equivalent: (i) the system is linear, time-invariant, and complete, (ii) the behavior is linear, shift-invariant, and closed, (iii) the behavior is kernel of a linear difference operator with a polynomial symbol, (iv) the system allows a linear input/output representation in terms of polynomial matrices, (v) the system allows a linear constant coefficient input/state/output representation, and (vi) the behavior is kernel of a linear difference operator with a rational symbol. If the system is controllable, then the system also allows (vii) an image representation with a polynomial symbol, and an image representation with a rational symbol.



Festschrift Alberto Isidori: System Interconnection

It is a pleasure to contribute an essay to this volume dedicated to Alberto Isidori on the occasion of his 65-th birthday. As the topic of my article, I chose an issue which is at the core of systems thinking, namely the formalization and the mathematization of system interconnection. This pertains to linear and nonlinear systems alike. In

Festschrift Keith Glover

A Festschrift is a welcome occasion to write an article with a personal and historical flavor. Because of the occasion, I chose the subject of Keith Glover's Ph.D. dissertation, SYSID. My interest in this area remained originally limited to the implications of the structure of linear systems. This situation changed with the *Automatica* papers [35]. These contain, in addition to the first comprehensive exposition of the behavioral approach to systems theory, a number of new ideas and subspace type-algorithms for SYSID. The aim of the present article is to explain in a somewhat informal style my own personal point of view on SYSID. Among other things, I will describe in some detail the many representations of linear time-invariant systems, leading up to some exact deterministic SYSID algorithms based on the notion of the most powerful unfalsified model. I will then explain the idea behind subspace algorithms, where the state trajectory is constructed directly from the observations and a system model in state form is deduced from there. Subsequently, I will discuss the role of latent variables in SYSID. This leads in a natural way to stochastic models. I will finish with some remarks on the rationale, or lack of it, of viewing SYSID in a stochastic framework.

Misfit, complexity, latency

IEEE TRANSACTIONS ON AUTOMATIC CONTROL, VOL. 50, NO. 10, OCTOBER 2005

Application of Structured Total Least Squares for System Identification and Model Reduction

Ivan Markovsky, Jan C. Willems, Sabine Van Huffel, Bart De Moor, and Rik Pintelon, *Fellow, IEEE*

***Abstract*—The following identification problem is considered: Minimize the ℓ_2 norm of the difference between a given time series and an approximating one under the constraint that the approximating time series is a trajectory of a linear time invariant system of a fixed complexity. The complexity is measured by the input dimension and the maximum lag. The question leads to a problem that is known as the global total least squares problem and alternatively can be viewed as maximum likelihood identification in the errors-in-variables setup. Multiple time series and latent variables can be considered in the same setting. Special cases of the problem are autonomous system identification, approximate realization, and finite time optimal ℓ_2 model reduction. The identification problem is related to the structured total least squares problem. This paper presents an efficient software package that implements the theory. The proposed method and software are tested on data sets from the database for the identification of systems DAISY.**

Stochastic systems

- Akaike
- Caines, Lindquist, Picci, etc….

$$
\begin{aligned}
\textbf{rank}\,(\,Y_f/\mathbf{Y_p}\,) &= n \\
\textbf{row space}\,(\,Y_f/\mathbf{Y_p}\,) &= \textbf{row space}\,(\,\widehat{X}_i\,) \\
\textbf{column space}\,(\,Y_f/\mathbf{Y_p}\,) &= \textbf{column space}\,(\,\Gamma_i\,)
\end{aligned}
$$

Open Stochastic Systems

Jan C. Willems *Life Fellow, IEEE*

IEEE Transactions on Automatic Control
Volume 58, pages 406-421, 2013

Abstract—The problem of providing an adequate definition of a stochastic system is addressed and motivated using examples. A stochastic system is defined as a probability triple. The specification of the set of events is an essential part of a stochastic model and it is argued that for phenomena with as outcome space a finite dimensional vector space, the framework of classical random vectors with the Borel sigma-algebra as events is inadequate even for elementary applications. Models very often require a coarse event sigma-algebra. A stochastic system is linear if the events are cylinders with fibers parallel to a linear subspace of a vector space. We address interconnection of stochastic systems. Two stochastic systems can be interconnected if they are complementary. We discuss aspects of the identification problem from this vantage point. A notion that emerges is constrained probability, a concept that is reminiscent but distinct from conditional probability. We end up with a comparison of open stochastic systems with probability kernels.

Multidim Syst Sign Process (2011) 22:45–53
DOI 10.1007/s11045-010-0126-z

Markovian properties for 2D behavioral systems described by PDE's: the scalar case

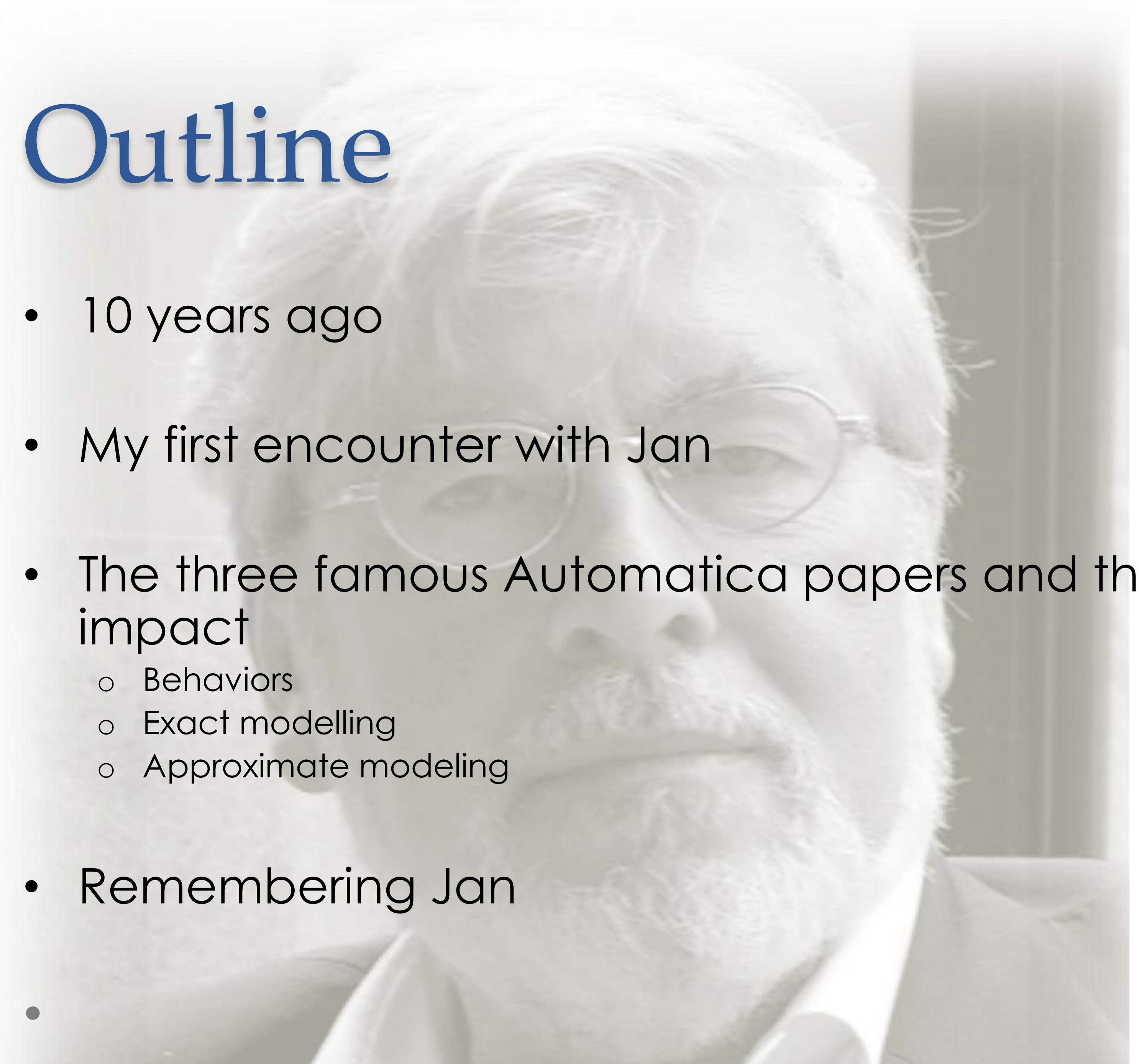
Paula Rocha · Jan C. Willems

Received: 1 April 2010 / Revised: 9 June 2010 / Accepted: 15 June 2010 /
Published online: 1 July 2010

Abstract In this paper we study the characterization of deterministic Markovian properties for 2D behavioral systems in terms of their descriptions by PDE's. In particular, we consider scalar systems and show that in this case strong-Markovianity is equivalent to the existence of a first order PDE description.

Keywords 2D systems · Behavioral approach · Markovian properties

Subspace identification for nD systems ?



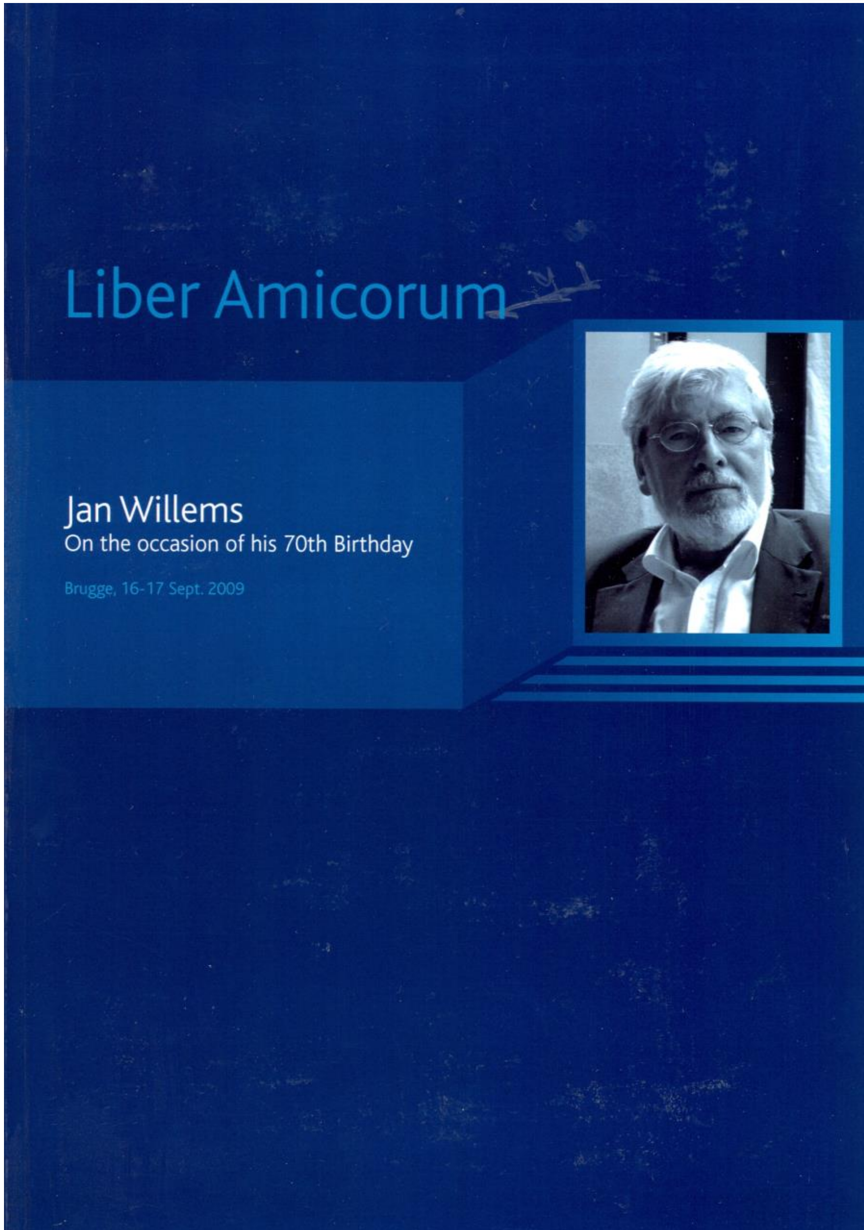
Outline

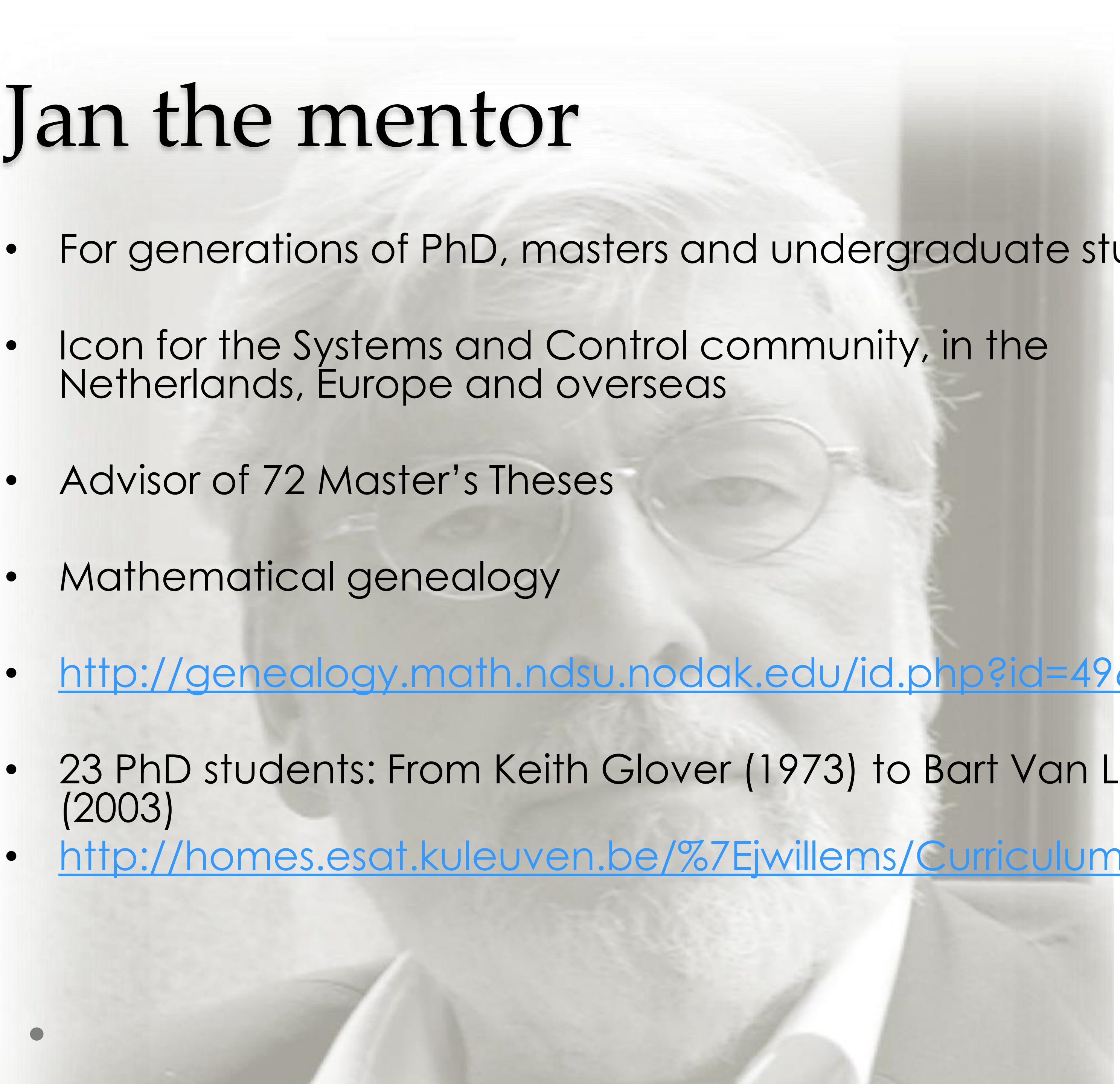
- 10 years ago

- My first encounter with Jan

- The three famous Automatica papers and their impact
    - Behaviors
    - Exact modelling
    - Approximate modeling

- Remembering Jan


Liber Amicorum
Jan Willems
On the occasion of his 70th Birthday
Brugge, 16-17 Sept. 2009



Jan the mentor

- For generations of PhD, masters and undergraduate students
- Icon for the Systems and Control community, in the Netherlands, Europe and overseas
- Advisor of 72 Master's Theses
- Mathematical genealogy
- http://genealogy.math.ndsu.nodak.edu/id.php?id=49680
- 23 PhD students: From Keith Glover (1973) to Bart Van Luyten (2003)
- http://homes.esat.kuleuven.be/%7Ejwillems/Curriculum.html#

1999

60th birthday

Keith Glover, Madhu Belur, Siep Weiland, Arjan van der Schaft, Henk Nijmeijer, Harry Trentelman, Jan Willem Polderman; second row: Tommaso Cotroneo, Paolo Rapisarda, Paula Rocha, Fabio Fagnani, Berend Roorda, Christiaan Heij, Tonny ten Dam.

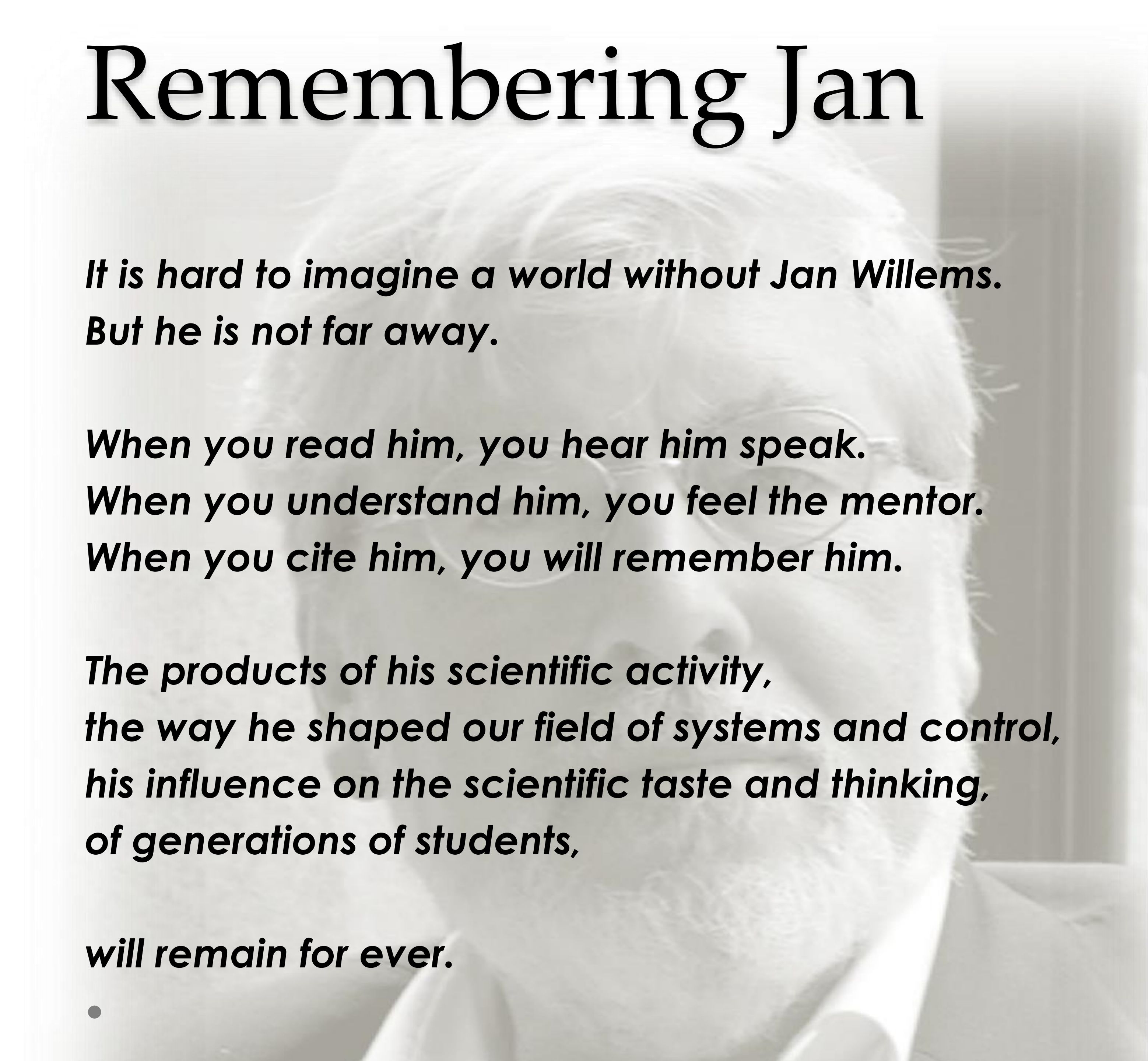
Jan the scholar

- One of the founding fathers of mathematical system theory; pursuer of system theoretic paradigms; helped shaping the field
- Huge contributions to the field of systems and control, as a scientist and an organizer
- Many people benefitted from his vision and personal perspective
- Critical but positive and constructive thinker
- Unique mix of creativity, associative power, ability for deep insights that he loved to share
- Responsibility and dedication, true scholar
- Argued with energy but also listened empathically
- Perfectionist care for rigor and details
- ‘Science should be left to scientists, not to administrators’

Jan the friend

- ‘Un grand monsieur’
- Cheerful, enthusiastic, inspiring
- Wonderful, considerate and animated
- Unquenchable amount of scientific and intellectual energy
- Charisma (‘the ‘X’-factor)
- Natural charm and skills in
  diplomacy and persuasion
- Talented story teller
- Enjoyed company with
- good glass and meal

Remembering Jan



It is hard to imagine a world without Jan Willems.
But he is not far away.

When you read him, you hear him speak.
When you understand him, you feel the mentor.
When you cite him, you will remember him.

The products of his scientific activity,
the way he shaped our field of systems and control,
his influence on the scientific taste and thinking,
of generations of students,

will remain for ever.

Bedankt, Jan

Thank you, Jan

http://janwillems-memoriam.net
http://homes.esat.kuleuven.be/~jwillems/