Proceedings of the 35th
Conference on Decision and Control
Kobe, Japan • December 1996

MODELLING USING MANIFEST AND LATENT VARIABLES

Jan C. Willems
Department of Mathematics
University of Groningen
P.O. Box 800
9700 AV Groningen, The Netherlands
email: J.C.Willems@math.rug.nl

Control questions often occur in the context of interconnected systems. In choosing a framework for a mathematical theory of control design, it is natural to take the model class which will be the end point of the modelling process as the starting point of the design process. Let us reflect on how such models are likely to look like. The main point that we are trying to make in this talk is the emergence of two types of variables: *manifest* and *latent* variables.

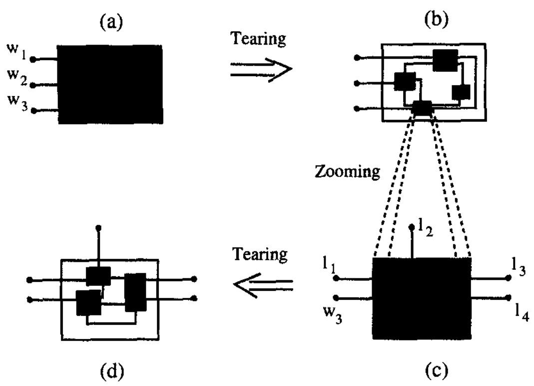

Fig. 1: An interconnected system

Consider an interconnected system, see figure 1. Assume that the purpose is to model the relation among certain variables, visualized as variables on the external terminals 1,2,3 of figure 1(a). Assume, for simplicity, that associated with each of these terminals there is one real number. Thus we are looking for the dynamic relation among the components of the vector $w = col(w_1, w_2, w_3)$. In order to come up with the required relation, look inside the black box. Typically this will lead to the situation of figure 1(b), obtained by *tearing* the black box of figure 1(a) into, say, four interconnected black boxes. Model the 4 new black boxes. *Zoom* in on, say, the one in the lower left corner, enlarged in figure 1(c). View this black box in the same way as the original one in figure 1(a), call the variables on its terminals $(w_3, \ell_1, \ell_2, \ell_3, \ell_4)$, and proceed by tearing 1(c) into new black boxes as shown in figure 1(d), zoom in on the subsystem thus obtained, etc. The modelling process will be complete when we reach subsystems which we call *modules*, when by tearing in this hierarchical way, we reach a point where the modules of all the sub-...-sub-black-boxes are known: as standard components, or already available as modules in a database, or identified through experiments.

This is the usual way modelling proceeds. *How will the final model look like?* It will consist of relations involving the external variables w_1, w_2, w_3 and many internal variables ℓ_1, ℓ_2, ... introduced as auxiliary variables in the modelling process. Two types of relations will occur: on the one hand, the laws of the modules relating external and internal variables, and on the other the interconnection laws, relating the variables of one module to those of another. Some of the resulting relations will be dynamical; some will be first order differential relations, some will be second order, some may be partial differential equations. However, there will also be (in fact, usually many) static equations (interconnection laws, resistors, springs). We will thus typically arrive at the following type of model

$$f_1(w, \frac{dw}{dt}, \ldots, \frac{d^n w}{dt^n}, \ell, \frac{d\ell}{dt}, \ldots, \frac{d^n \ell}{dt^n}) = f_2(w, \frac{dw}{dt}, \ldots, \frac{d^n w}{dt^n}, \ell, \frac{d\ell}{dt}, \ldots, \frac{d^n \ell}{dt^n}) \qquad (1)$$

with $w = col(w_1, w_2, \ldots) \in \mathbb{W}$ the vector of to-the-modelled variables (we will call them *manifest* variables), $\ell = col(\ell_1, \ell_2, \ldots) \in \mathbb{L}$ the vector of the other variables introduced in the modelling process, for example in the tearing and zooming stage (we will call them *latent* variables), and f_1, f_2 maps into an appropriate space. For exapmle, when modelling an electrical circuit, (1) will consist of the constitutive equations of the resistors, inductors, capacitors, transformers, etc., Kirchhoff's voltage laws (one for each loop), and Kirchhoff's current laws (one for each node).

The differential equations (1) can be considered as a complete model of the dynamical system. They specify the desired relation among the manifest variables, i.e., they specify the family of time functions which, according to the model (for example the black box of figure 1(a)), are in principle possible:

$$\mathfrak{B} = \{w : \mathbb{R} \to \mathbb{W} \mid \exists \ell : \mathbb{R} \to \mathbb{L} \text{ such that (1) holds}\} \tag{2}$$

We will call $\mathfrak{B}$ the (*manifest*) *behavior* of the system and denote the system itself by

$$\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B}) \tag{3}$$

with $\mathbb{T}$ the time set (in our case, $\mathbb{T} = \mathbb{R}$), $\mathbb{W}$ the space of manifest variables, and $\mathfrak{B}$ given by (??).

Of course, an even more basic model is the one involving the latent variables explicitly, rather than implicitly. Thus we will call the system

$$\Sigma_L = (\mathbb{T}, \mathbb{W}, \mathbb{L}, \mathfrak{B}_f) \tag{4}$$

with the *full behavior* defined by

$$\mathfrak{B}_f = \{(w, \ell) : \mathbb{T} \to \mathbb{W} \times \mathbb{L} \mid \text{(??) is satisfied}\} \tag{5}$$

the *latent variable system* defined by (1), and (??) the *manifest system* induced by (??).

The model (1) is the appropriate starting point for a theory of differential systems and of control design. Note that usually there will be a far distance from (1) to

$$\begin{aligned} \frac{dx}{dt} &= f(x,u) \\ y &= g(x,u) \\ w &= \begin{pmatrix} u \\ y \end{pmatrix} \end{aligned} \tag{6}$$

or

$$\hat{y}(s) = G(s)\hat{u}(s), \tag{7}$$

the common starting points of the theory of control!

References

[1] J. C. Willems: From time series to linear system. Part I: Finite dimensional linear time invariant systems; Part II: Exact modelling, Part III: Approximate modelling, *Automatica*, **22**, pp. 561–580, pp. 675–694 (1986) **23**, pages 87–115 (1987).

[2] J.C. Willems: Models for Dynamics, *Dynamics Reported*, **2**, pages 171–269 (1989).

[3] J.C. Willems: Paradigms and puzzles in the theory of dynamical systems, *IEEE Transactions on Automatic Control*, **36**, pages 259–294 (1991).

[4] J.C. Willems: On interconnections, control, and feedback, *IEEE Transactions on Automatic Control*, to appear.