MODELING, MODULARITY

and

MODULES

Jan C. Willems
University of Groningen, NL

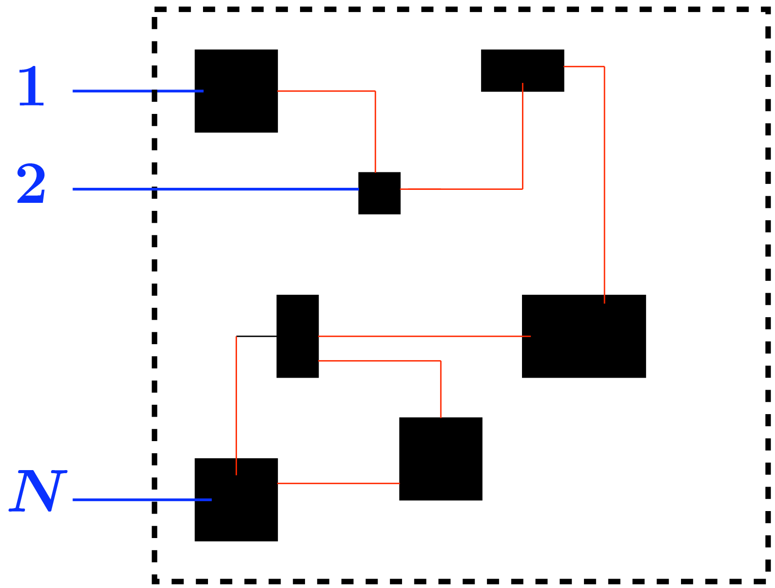
BOYDFEST, Rice University, Houston, TX

March 9, 2001

RuG

Dedicated to Boyd Pearson

How do we model an interconnected system?

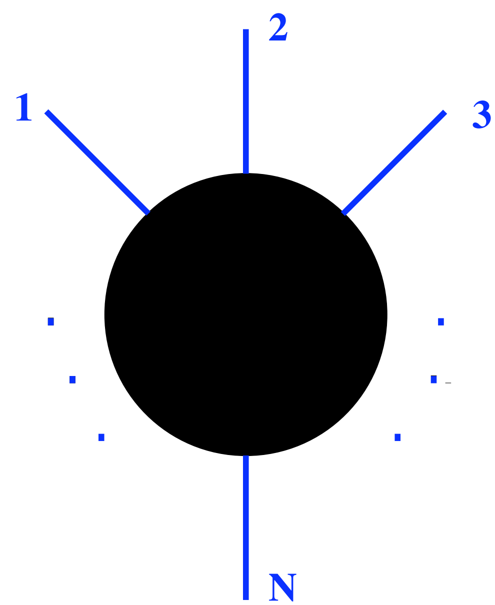


Interconnected system

Exs.: circuits, robots, chemical plants, etc. $\rightsquigarrow$ **Modularity**

$\rightsquigarrow$ **Object-oriented modeling**

How do we model a building block?



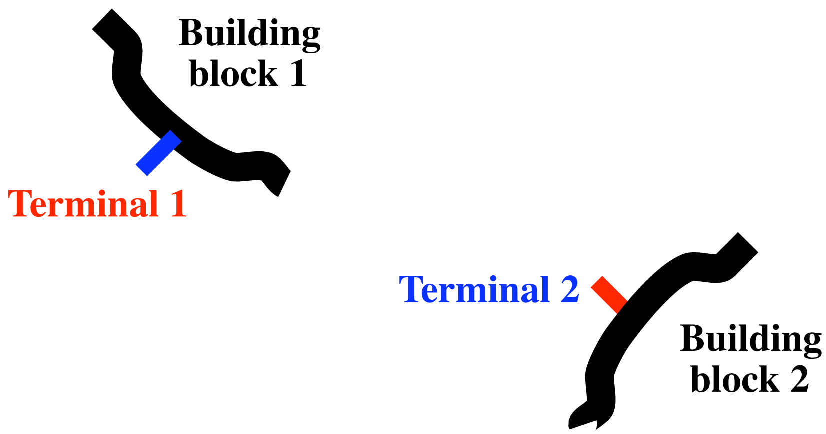
Building block

Exs.: resistor, capacitor, mass, spring, damper, tank, heat bath, etc.

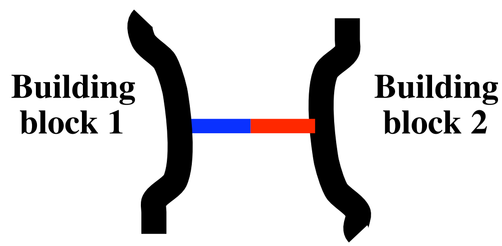
$\rightsquigarrow$ **Behavior of the terminal variables**

How do we model an interconnection?

Before:



After:



Interconnection

$\rightsquigarrow$ **Identification of terminal variables**

Examples of terminal variables:

Type of terminal	Variables	Signal space
electrical	(voltage, current)	$\mathbb{R}^2$
mechanical (1-D)	(force, position)	$\mathbb{R}^2$
mechanical (2-D)	((position, attitude), (force, torque))	$(\mathbb{R}^2 \times S^1) \times (\mathbb{R}^2 \times T^*S^1)$
mechanical (3-D)	((position, attitude), (force, torque))	$(\mathbb{R}^2 \times S^2) \times (\mathbb{R}^2 \times T^*S^2)$
thermal	(temp., heat flow)	$\mathbb{R}^2$
fluidic	(pressure, flow)	$\mathbb{R}^2$
fluidic - thermal	(pressure, flow, temp., heat flow)	$\mathbb{R}^4$

Examples of interconnection constraints:

Pair of terminals	Terminal 1	Terminal 2	Law
electrical	(V_1, I_1)	(V_2, I_2)	$V_1 = V_2, I_1 + I_2 = 0$
1-D mech.	(F_1, q_1)	(F_2, q_2)	$F_1 + F_2 = 0, q_1 = q_2$
2-D mech.			
thermal	(T_1, Q_1)	(T_2, Q_2)	$T_1 = T_2, Q_1 + Q_2 = 0$
fluidic	(p_1, f_1)	(p_2, f_2)	$p_1 = p_2, f_1 + f_2 = 0$
fluidic - thermal	(p_1, f_1, T_1, Q_1)	(p_2, f_2, T_2, Q_2)	$p_1 = p_2, f_1 + f_2 = 0, T_1 = T_2, Q_1 + Q_2 = 0$

Classical approach

Building blocks:

- **input/output:**

  Recognize input and output variables

  Model the input-to-output map or relation

- **input/state/output:**

  Recognize input, output, and state variables

  Model the input-to-state and the state-to-output maps

$$\rightsquigarrow \quad \frac{d}{dt}x = f(x,u) \quad y = h(x)$$

Interconnections:

Identify inputs with outputs

Combine series, parallel, feedback connection.

Beautiful concepts, very effective algorithms, but i/o is simply

not suitable as a 'first principles' starting point.

For building blocks:

Terminal variables are localized $\neq$ $\Rightarrow$ **System** $\Rightarrow$

A physical system is not a signal processor.

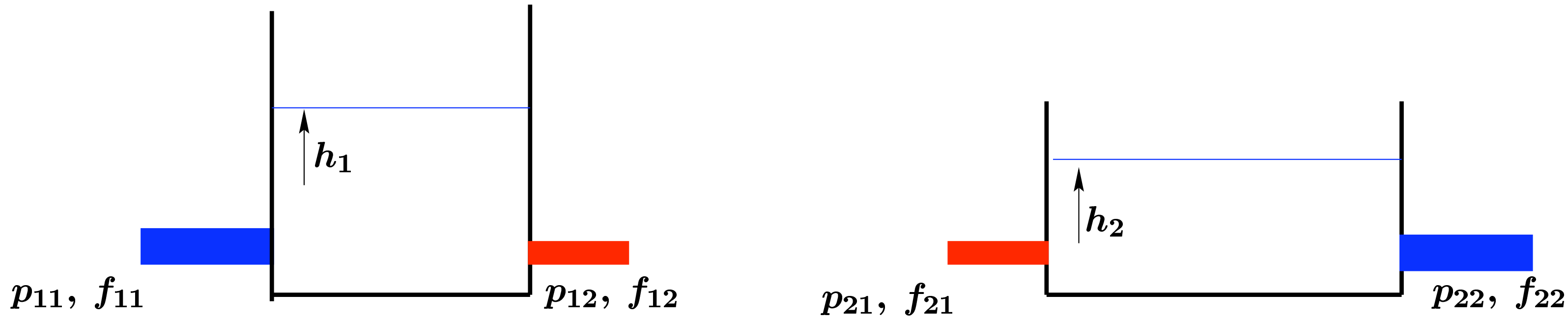
But: even CS and DES do not use the i/o approach!

For interconnected systems:

It is *not feasible to recognize the signal flow graph* before we have a model. The signal flow graph should be deduced from a model!

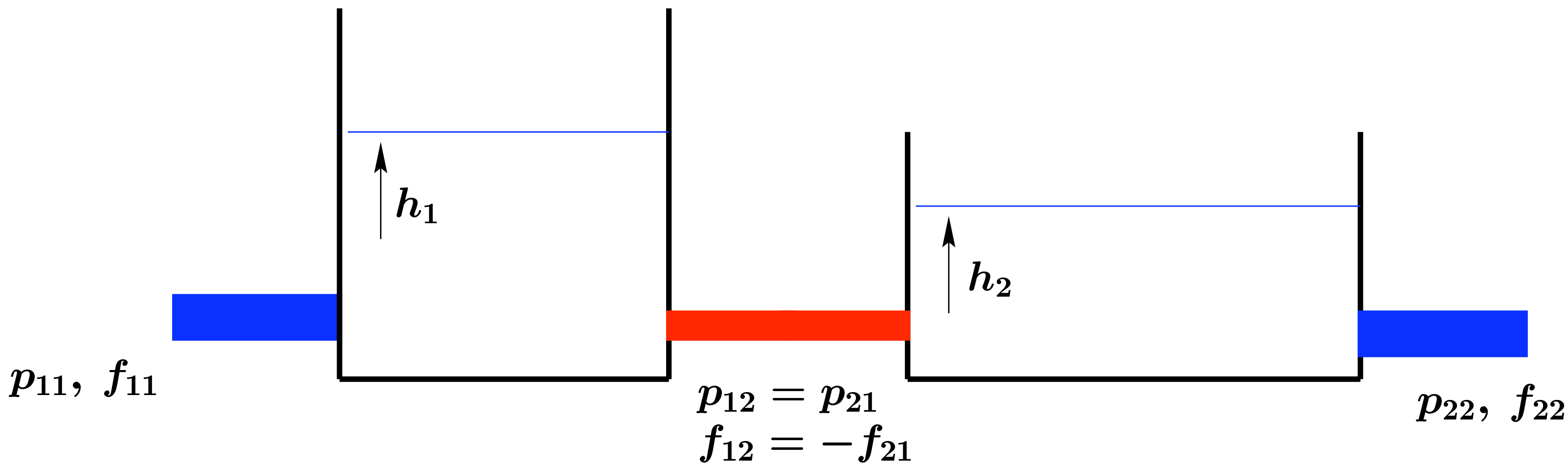
More suitable approach for dealing with interconnections $\leadsto$ **Bondgraphs.**

The inappropriateness of input - to - output connections is illustrated well by the following simple physical example:



Logical choice of inputs: the pressures $p_{11}, p_{12}, p_{21}, p_{22}$,
and of outputs: the flows $f_{11}, f_{12}, f_{21}, f_{22}$
(h_1, h_2: state variables)

In any case, the input/output choice should be 'symmetric'.



Interconnection constraints:

$$p_{12} = p_{21}, \quad f_{12} = -f_{21}.$$

Equates two inputs and two outputs.

$\neq$ **equating inputs with outputs.**

BEHAVIORAL SYSTEMS

A system **:=** $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$

$\mathbb{T}$ = **the set of *independent* variables**

time, space, time and space

$\mathbb{W}$ = **the set of *dependent* variables**

(= where the variables take on their values),

signal space, space of field variables, . . .

$\mathfrak{B} \subseteq \mathbb{W}^{\mathbb{T}}$ **:** ***the behavior*** **= the admissible trajectories**

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

for a trajectory $w : \mathbb{T} \to \mathbb{W}$, **we thus have:**

$w \in \mathfrak{B}$ **: the model allows the trajectory** w,

$w \notin \mathfrak{B}$ **: the model forbids the trajectory** w.

In the remainder of this lecture, $\mathbb{T} = \mathbb{R}^n$, $\mathbb{W} = \mathbb{R}^{\mathtt{w}}$,

$w : \mathbb{R}^n \to \mathbb{R}^{\mathtt{w}}, (w_1(x_1, \cdots, x_{\mathtt{n}}), \cdots, w_{\mathtt{w}}(x_1, \cdots, x_{\mathtt{n}}))$,

often, $\mathtt{n} = 1$, **independent variable time,**

or $\mathtt{n} = 4$, **independent variables** (t, x, y, z),

$\mathfrak{B}$ = **solutions of a system of constant coefficient linear ODE's or PDE's.**

Linear constant coefficient ODE's.

Variables: $w_1, w_2, \ldots w_{\mathtt{w}}$, **their derivatives, combined in any number of linear equations. In vector/matrix notation:**

$$w = \begin{bmatrix} w_1 \\ w_2, \\ \vdots \\ w_{\mathtt{w}} \end{bmatrix}, \quad R_{\mathtt{k}} = \begin{bmatrix} R^{\mathtt{k}}_{1,1} & R^{\mathtt{k}}_{1,2} & \cdots & R^{\mathtt{k}}_{1,\mathtt{w}} \\ R^{\mathtt{k}}_{2,1} & R^{\mathtt{k}}_{2,2} & \cdots & R^{\mathtt{k}}_{2,\mathtt{w}} \\ \vdots & \vdots & \cdots & \vdots \\ R^{\mathtt{k}}_{\mathtt{g},1} & R^{\mathtt{k}}_{\mathtt{g},2} & \cdots & R^{\mathtt{k}}_{\mathtt{g},\mathtt{w}} \end{bmatrix}.$$

Yields

$$\boxed{R_0 w + R_1 \frac{d}{dt} w + \cdots + R_{\mathtt{n}} \frac{d^{\mathtt{n}}}{dt^{\mathtt{n}}} w = 0,}$$

with $R_0, R_1, \cdots, R_{\mathtt{n}} \in \mathbb{R}^{\mathtt{g} \times \mathtt{w}}$.

Combined with the polynomial matrix

$$R(\xi) = R_0 + R_1\xi + \cdots + R_n\xi^n,$$

we obtain

$$\boxed{R(\frac{d}{dt})w = 0.}$$

Examples:

- The ubiquitous

$$P(\tfrac{d}{dt})y = Q(\tfrac{d}{dt})u, \quad w = (u, y)$$

with $P, Q \in \mathbb{R}^{\bullet\times\bullet}[\xi]$, $\det(P) \neq 0$ and, perhaps, $P^{-1}Q$ proper.

- The ubiquitous

$$\tfrac{d}{dt}x = Ax + Bu; \quad y = Cx + Du, \quad w = (u, y).$$

- The descriptor systems

$$\frac{d}{dt}Ex + Fx + Gw = 0.$$

- etc., etc.

Notation:

Ring of real polynomials in $\mathtt{n}$ variables $\rightsquigarrow \mathbb{R}[\xi_1, \cdots, \xi_{\mathtt{n}}]$.

$\mathbb{R}^{\mathtt{n}}[\xi_1, \cdots, \xi_{\mathtt{n}}], \mathbb{R}^{\bullet}[\xi_1, \cdots, \xi_{\mathtt{n}}], \mathbb{R}^{\mathtt{n}_1 \times \mathtt{n}_2}[\xi_1, \cdots, \xi_{\mathtt{n}}],$
$\mathbb{R}^{\bullet \times \mathtt{n}}[\xi_1, \cdots, \xi_{\mathtt{n}}], \mathbb{R}^{\mathtt{n} \times \bullet}[\xi_1, \cdots, \xi_{\mathtt{n}}],$
$\mathbb{R}^{\bullet \times \bullet}[\xi_1, \cdots, \xi_{\mathtt{n}}]$.

$\mathbb{R}[\xi_1, \cdots, \xi_{\mathtt{n}}]$ **has much less convenient properties than** $\mathbb{R}[\xi]$**:**

not Euclidean domain, hence not p.i.d., no Smith form, etc.

Linear differential systems (PDE's)

$\mathbb{T} = \mathbb{R}^{\mathtt{n}}$, $\mathtt{n}$ **independent variables,**

$\mathbb{W} = \mathbb{R}^{\mathtt{w}}$, $\mathtt{w}$ **dependent variables,**

$\mathfrak{B} =$ **the solutions of a linear constant coefficient system of PDE's.**

Let $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi_1, \cdots, \xi_{\mathtt{n}}]$, **and consider**

$$R\left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathtt{n}}}\right) w = 0 \quad (*)$$

Define its behavior

$$\mathfrak{B} = \{ w \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}) \mid (*) \text{ holds } \}$$

$\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$ **mainly for convenience, but important for some results.**

An example of a DPS: *Maxwell's equations*

$$\begin{aligned}
\nabla \cdot \vec{E} &= \frac{1}{\varepsilon_0}\rho\,, \\
\nabla \times \vec{E} &= -\frac{\partial}{\partial t}\vec{B}, \\
\nabla \cdot \vec{B} &= 0\,, \\
c^2 \nabla \times \vec{B} &= \frac{1}{\varepsilon_0}\vec{j} + \frac{\partial}{\partial t}\vec{E}.
\end{aligned}$$

$\mathbb{T} = \mathbb{R} \times \mathbb{R}^3$ (time and space),

$w = (\vec{E}, \vec{B}, \vec{j}, \rho)$

(electric field, magnetic field, current density, charge density),

$\mathbb{W} = \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$,

$\mathfrak{B}$ = set of solutions to these PDE's.

Note: 10 variables, 8 equations! $\Rightarrow \exists$ free variables.

Notation:

$$(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}, \mathfrak{B}) \in \mathfrak{L}^{\mathtt{w}}_{\mathtt{n}}, \quad \text{or } \mathfrak{B} \in \mathfrak{L}^{\mathtt{w}}_{\mathtt{n}},$$

$$\mathfrak{B} = \ker(R(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathtt{n}}})).$$

'kernel representation'.

R **defines** $\mathfrak{B} = \ker(R(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}))$**, but not vice-versa!**

¿¿ $\exists$ 'intrinsic' characterization of $\mathfrak{B} \in \mathfrak{L}^{w}_{n}$??

Define the *annihilators* of $\mathfrak{B}$ by

$$\mathfrak{N}_{\mathfrak{B}} := \{n \in \mathbb{R}^{w}[\xi_1, \cdots, \xi_n] \mid n^{\top}(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n})\mathfrak{B} = 0\}.$$

$\mathfrak{N}_{\mathfrak{B}}$ **is clearly a sub-module of** $\mathbb{R}^{w}[\xi_1, \cdots, \xi_n]$.

Let $< R >$ **denote the sub-module of** $\mathbb{R}^{w}[\xi_1, \cdots, \xi_n]$ **spanned by the transposes of the rows of** R**. Obviously** $< R > \subseteq \mathfrak{N}_{\mathfrak{B}}$**. But, in fact:**

$$\mathfrak{N}_{\mathfrak{B}} = < R >$$

<u>Therefore</u>

$$\mathfrak{L}^{w}_{n} \stackrel{1:1}{\longleftrightarrow} \text{sub-modules of } \mathbb{R}^{w}[\xi_1, \cdots, \xi_n]$$

Controllability

Definition: $\mathfrak{B} \in \mathfrak{L}^{\mathtt{w}}_{\mathtt{n}}$ is said to be

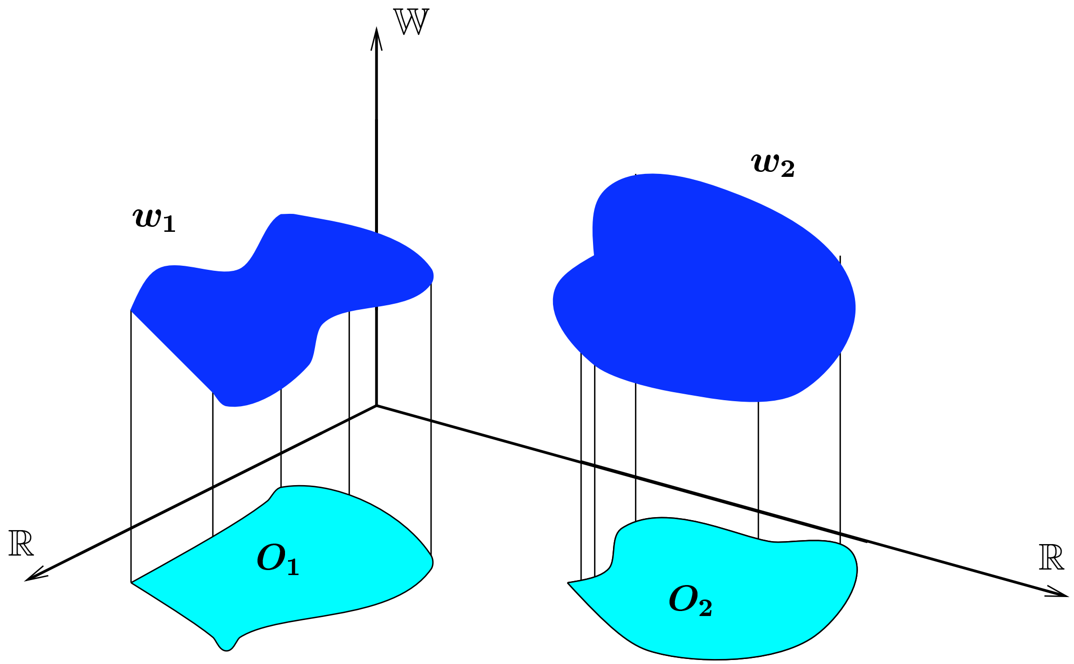
controllable

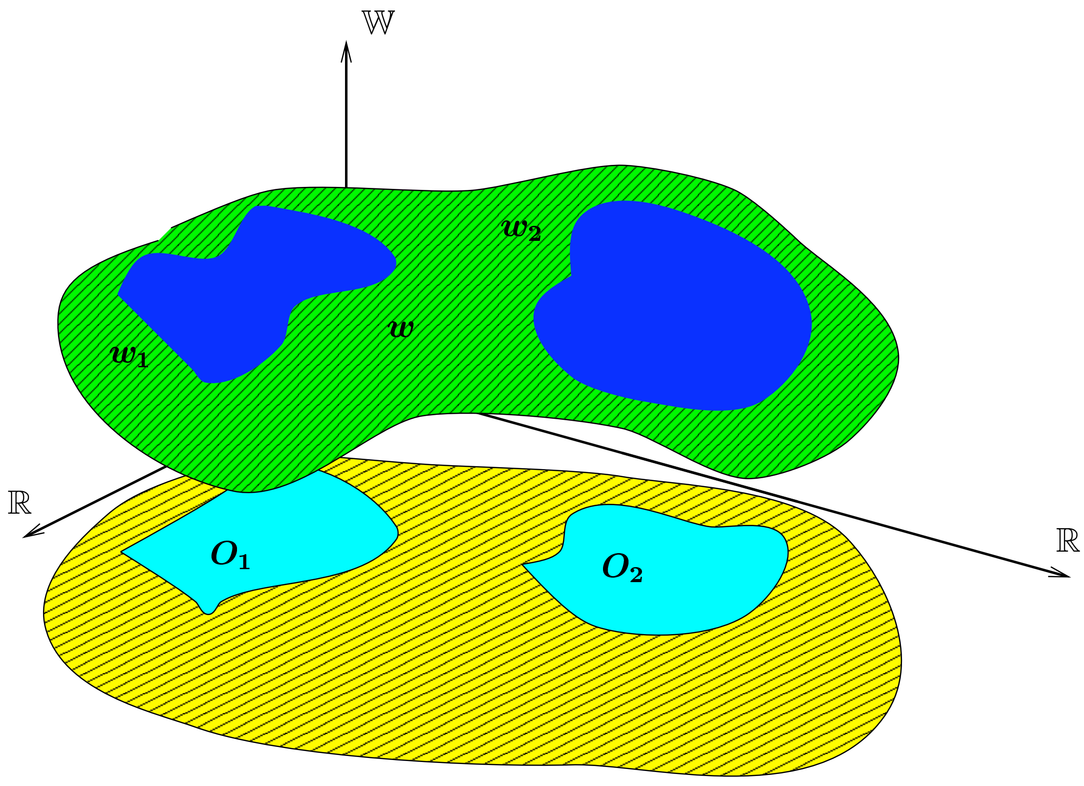
if for all $w_1, w_2 \in \mathfrak{B}$ and
for all $O_1, O_2 \subset \mathbb{R}^n$, non-overlapping closure,
there exists $w \in \mathfrak{B}$ such that $w|_{O_1} = w_1|_{O_1}$ and $w|_{O_2} = w_2|_{O_2}$.

Controllability :$\Leftrightarrow$ the elements of $\mathfrak{B}$ are ‘patch-able’.

Special case: Kalman controllability for input/state systems.

In pictures:




𝕎
w_2
w
w_1
ℝ
ℝ
O_1
O_2

Conditions for controllability

Representations of $\mathfrak{L}_{\mathtt{n}}^{\mathtt{w}}$:

$$R(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathrm{n}}})w = 0 \qquad (*)$$

called a *‘kernel’ representation* of $\mathfrak{B} = \ker(R(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathrm{n}}}))$;

$$R(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathrm{n}}})w = M(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathrm{n}}})\ell \quad (**)$$

called a *‘latent variable’ representation* of the manifest behavior $\mathfrak{B} = (R(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathrm{n}}}))^{-1} M(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathrm{n}}})\mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^\ell)$.

Missing link:

$$w = M(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathrm{n}}})\ell \quad (***)$$

called an *‘image’ representation* of $\mathfrak{B} = \mathrm{im}(M(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathrm{n}}}))$.

Elimination theorem $\Rightarrow$

every image (of a linear constant coefficient PDO) is also a kernel.

¿¿ Which kernels are also images ??

Theorem: The following are equivalent for $\mathfrak{B} \in \mathfrak{L}^{\mathtt{w}}_{\mathtt{n}}$:

1. **$\mathfrak{B}$ is controllable,**
2. **$\mathfrak{B}$ admits an image representation,**
3. **for any $a \in \mathbb{R}^{\mathtt{w}}[\xi_1, \cdots, \xi_{\mathtt{n}}]$,**
   $$a^\top [\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_{\mathtt{n}}}]\mathfrak{B} \text{ equals } 0 \text{ or all of } \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}),$$
4. **$\mathbb{R}^{\mathtt{w}}[\xi_1, \cdots, \xi_{\mathtt{n}}]/\mathfrak{N}_{\mathfrak{B}}$ is torsion free,**

etc.

Algorithm: R + syzygies + Gröbner basis $\Rightarrow$ numerical test on coefficients of R.

ARE MAXWELL'S EQUATIONS CONTROLLABLE ?

The following equations in the *scalar potential* $\phi : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}$ and the *vector potential* $\vec{A} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$, generate exactly the solutions to Maxwell's equations:

$$
\begin{aligned}
\vec{E} &= -\frac{\partial}{\partial t}\vec{A} - \nabla\phi, \\
\vec{B} &= \nabla \times \vec{A}, \\
\vec{j} &= \varepsilon_0 \frac{\partial^2}{\partial t^2}\vec{A} - \varepsilon_0 c^2 \nabla^2 \vec{A} + \varepsilon_0 c^2 \nabla(\nabla \cdot \vec{A}) + \varepsilon_0 \frac{\partial}{\partial t}\nabla\phi, \\
\rho &= -\varepsilon_0 \frac{\partial}{\partial t}\nabla \cdot \vec{A} - \varepsilon_0 \nabla^2 \phi.
\end{aligned}
$$

Proves controllability. Illustrates the interesting connection

controllability $\Leftrightarrow$ $\exists$ potential!

SUMMARY

- **The i/s/o paradigm is inadequate for first principles *modeling.* It fails in the first examples, it is unsuited for interconnection, for *modularity*, for object-oriented modeling.**
- **Universal paradigm: *Behavioral systems*. Illustrated via PDE's.**
- **Linear shift-invariant differential systems**

  $$\overset{1:1}{\longleftrightarrow}$$ ***sub-modules* of $\mathbb{R}^{\mathtt{w}}[\xi_1, \cdots, \xi_{\mathtt{n}}]$.**
- **Controllability $\Leftrightarrow$ *sub-module* is torsion-free.**
- **$\exists$ extensive theory, adapted to modeling, covering all the classical results, unifying physical models with DES, etc.**

THANK YOU

&

BEST WISHES, BOYD !