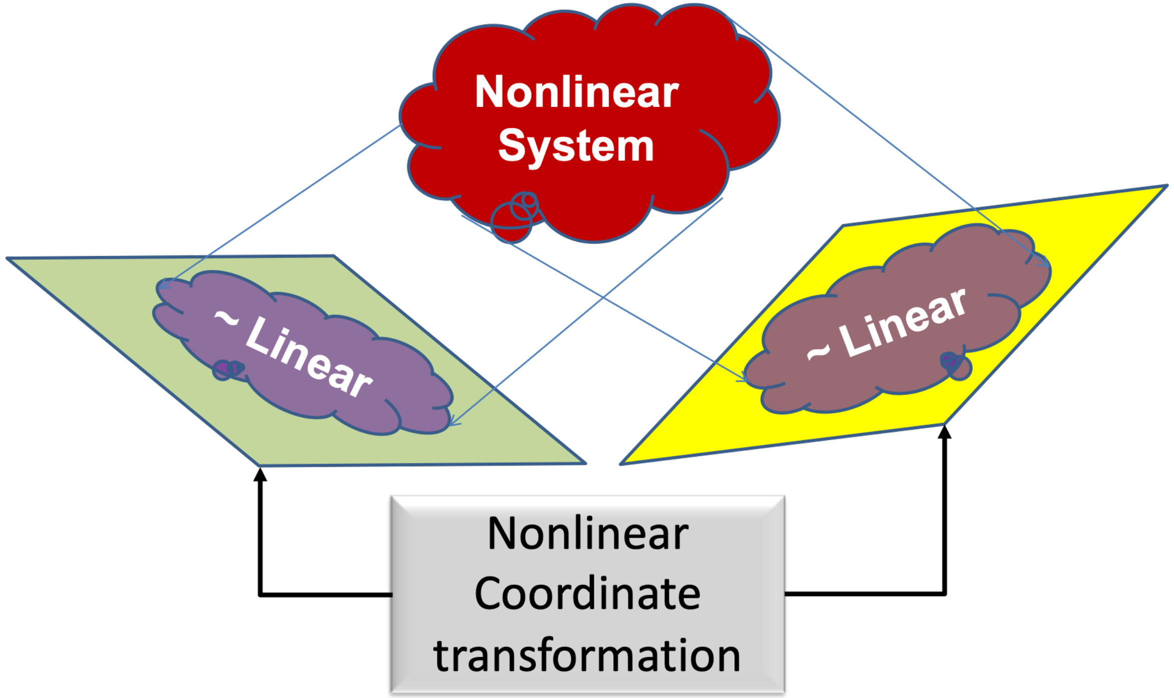
2.160 Identification, Estimation, and Learning

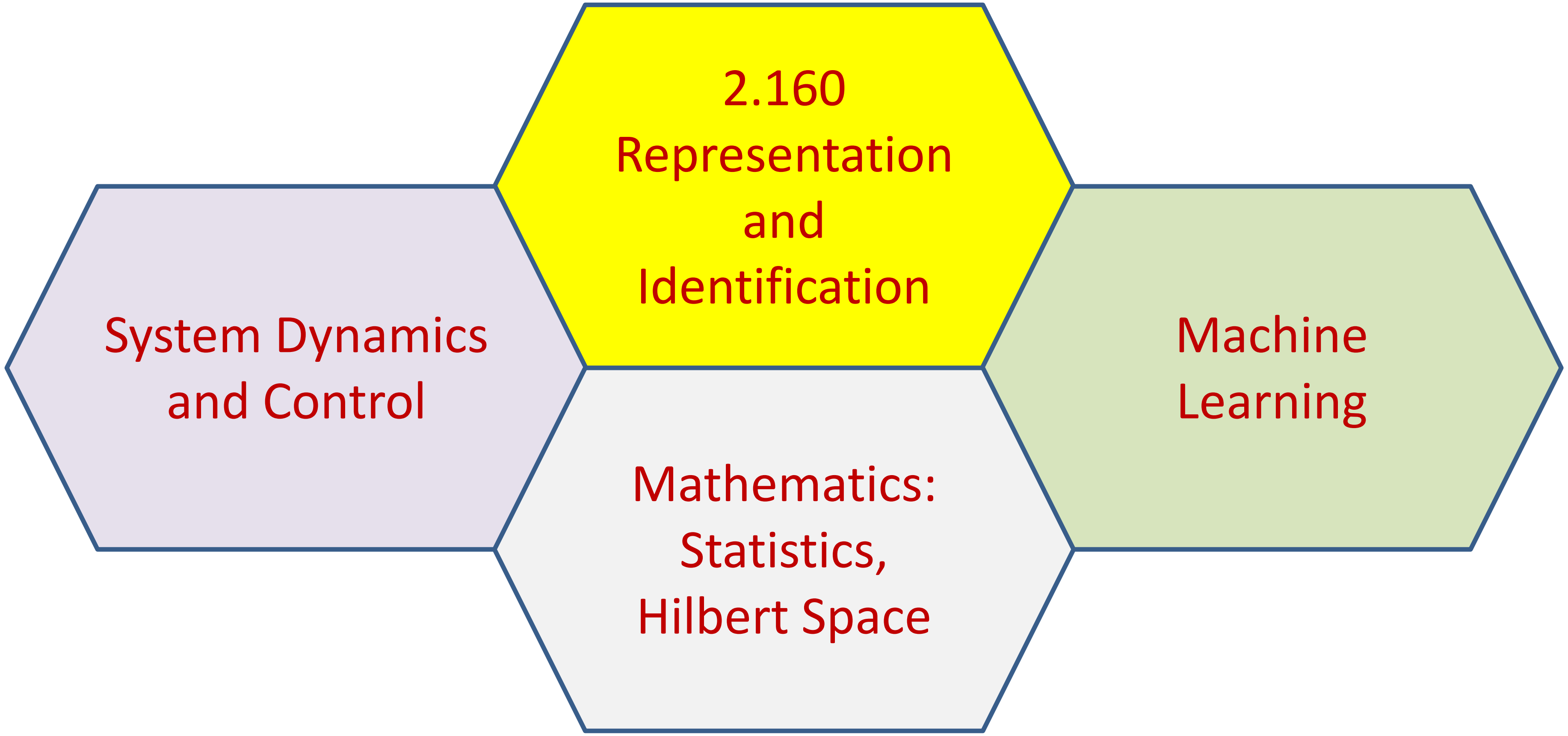
Part 4 Machine Learning and Nonlinear System Modeling

Lecture 25

Dual-Faceted Linearization with Application to Nonlinear Model Predictive Control - 2.160 Final Lecture -

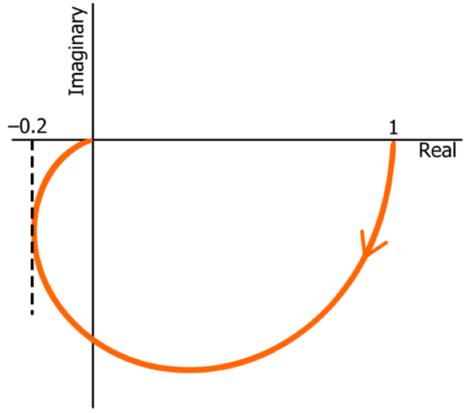
H. Harry Asada
Department of Mechanical Engineering
MIT




2.160
Representation
and
Identification
System Dynamics
and Control
Machine
Learning
Mathematics:
Statistics,
Hilbert Space

Reflection

- Modeling and representation are the foundation of engineering;
- Control engineering made groundbreaking progress, whenever new model representation methods were introduced;
  - Frequency-domain representation
  - State-space representation, etc.
- Linearization of nonlinear dynamics through “***lifting***” could be one of the new representations leading to groundbreaking progress.



$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t)$$

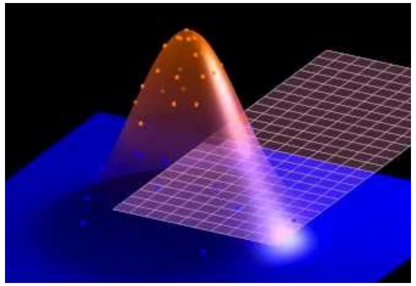
$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u(t)$$

Cover’s Theorem

- The use of non-linear **kernel methods** in machine learning: Given a set of training data that is not linearly separable, one can with high probability transform it into a training set that is linearly separable by projecting it into a higher-dimensional space via some non-linear transformation, 1965.

Thomas Cover

$$\varphi(x_1, x_2) = \begin{pmatrix} x_1 \\ x_2 \\ x_1 x_2 \end{pmatrix}$$



Koopman Operator

- Consider a discrete-time, autonomous dynamical system:

$$x(t+1) = F(x(t)) \qquad x \in \Re^n, \quad F : \Re^n \to \Re^n \text{ (Non-singular)}$$

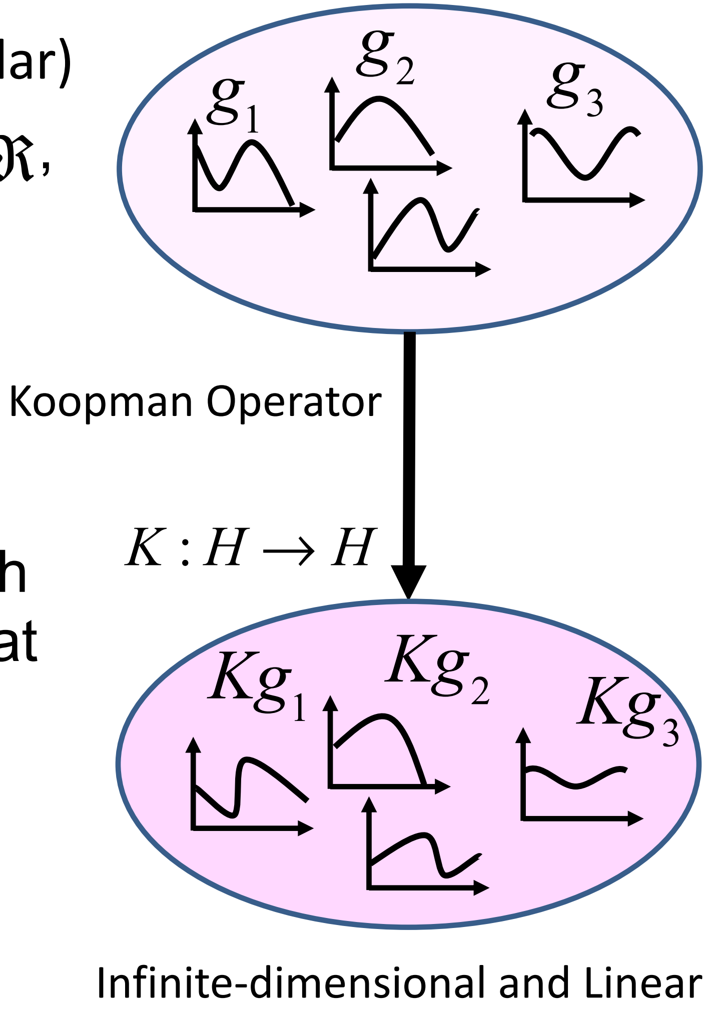
- Also, consider output functions, called observables: $g(x) : \Re^n \to \Re$, and their trajectories:

$$g_i(x(0)),\ g_i(x(1)),\ g_i(x(2)), \cdots$$

- **Definition (Discrete-time Koopman Operator)**
  Let observable $g(x) : \Re^n \to \Re$ be in an element in a Hilbert space $\boldsymbol{H}$. The Koopman Operator $\boldsymbol{K}_F : H \to H$ associated with the map $F : \Re^n \to \Re^n$ is defined to be a linear transform that meets the following compositional relation:

$$\boldsymbol{K} g(x) = g \circ F(x) \qquad \forall g \in H$$

- Note that this transform acts on functions, i.e. trajectories.
- Although the original system represented as a point-wise transformation is nonlinear, the above Koopman operator is linear.



Comparison between Evolution Operator and Koopman Operator

- Evolution Operator

$$\begin{pmatrix} g_1(t+1) \\ \vdots \\ g_m(t+1) \end{pmatrix} = \begin{pmatrix} g_1(F(x(t)) \\ \vdots \\ g_m(F(x(t)) \end{pmatrix}$$

- Koopman Operator

$$g \circ F = K\, g$$

Trajectory of i^{th} observable

$$\begin{pmatrix} g_i(1) \\ g_i(2) \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} g_i(F(x(0)) \\ g_i(F(x(1)) \\ \vdots \end{pmatrix}}_{g_i(F(x))} = \begin{pmatrix} k_{11} & k_{12} & \cdots \\ k_{21} & k_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \underbrace{\begin{pmatrix} g_i(0) \\ g_i(1) \\ \vdots \end{pmatrix}}_{g_i(x)}$$

Computation of Koopman Operator Using a Companion Matrix

- Suppose that we use m observables, $g_1(t),\cdots,g_m(t)$.
- Collecting data for time 0 through m,

$$Z_t = \begin{pmatrix} g_1(t) \\ \vdots \\ g_m(t) \end{pmatrix} \qquad \begin{aligned} Z_{0|m-1} &= \left(Z_0, Z_1, \cdots, Z_{m-1}\right) \in \Re^{m\times m} \\ Z_{1|m} &= \left(Z_1, Z_2, \cdots, Z_m\right) \in \Re^{m\times m} \end{aligned} \qquad {Z_{1|m}}^T = K_m {Z_{0|m-1}}^T$$

- The relationship between $Z_{1|m}$ and $Z_{0|m-1}$ is described with a **Companion matrix C_m**. and a residual r.

$$\begin{pmatrix} g_1(1) & g_2(1) & \cdots & g_m(1) \\ g_1(2) & g_2(2) & \cdots & g_m(2) \\ \vdots & \vdots & \vdots & \vdots \\ g_1(m) & g_2(m) & \cdots & g_m(m) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & \vdots & & 0 & 1 \\ c_0 & c_1 & c_2 & \cdots & c_{m-1} \end{pmatrix}}_{C_m} \begin{pmatrix} g_1(0) & g_2(0) & \cdots & g_m(0) \\ g_1(1) & g_2(1) & \cdots & g_m(1) \\ \vdots & \vdots & \vdots & \vdots \\ g_1(m-1) & g_2(m-1) & \cdots & g_m(m-1) \end{pmatrix}$$

$$K_m \leftrightarrow C_m$$

- The last row of the above equation:

$$\begin{pmatrix} g_1(m) & g_2(m) & \cdots & g_m(m) \end{pmatrix} = \begin{pmatrix} c_0 & c_1 & \cdots & c_{m-1} \end{pmatrix} \begin{pmatrix} g_1(0) & g_2(0) & \cdots & g_m(0) \\ g_1(1) & g_2(1) & \cdots & g_m(1) \\ \vdots & \vdots & \vdots & \vdots \\ g_1(m-1) & g_2(m-1) & \cdots & g_m(m-1) \end{pmatrix}$$

- In general, the last row is an approximation with some residual r_i.

$$g_i(m) = \sum_{j=0}^{m-1} c_j g_i(j) + r_i, \quad i = 1, \cdots, m$$

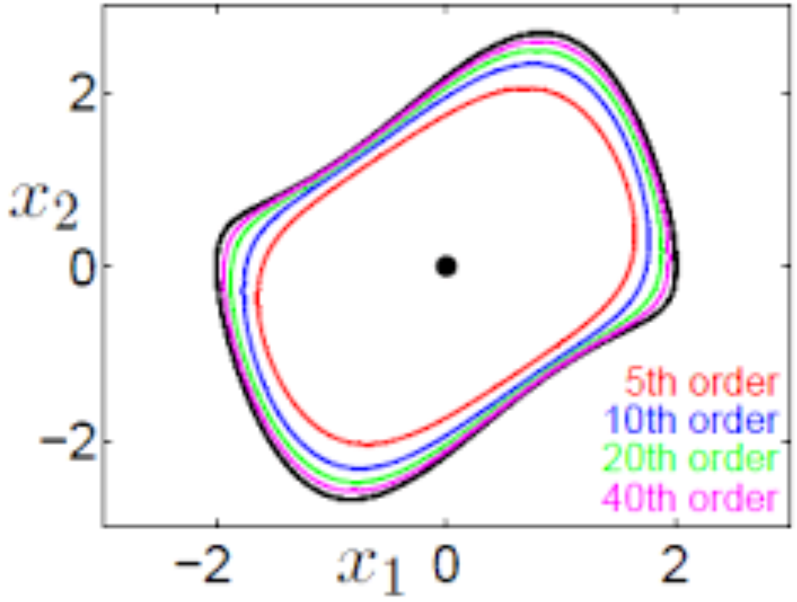
- The squared residual $R^2 = \sum r_i^2$ can be minimized by optimizing the coefficients c_i.

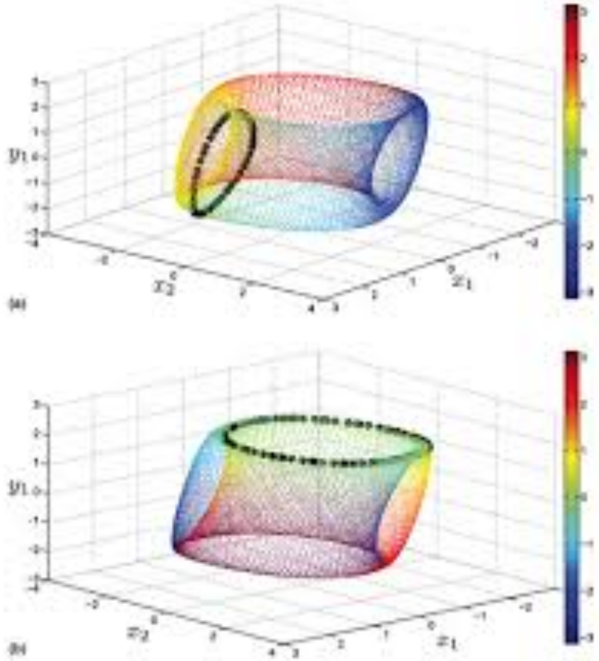
$$(c_1, \cdots, c_m) = \arg\min_{c_1, \cdots, c_m} \sum_{i=1}^{m} \left(g_i(m) - \sum_{j=0}^{m-1} c_j g_i(j) \right)^2$$

- The magnitude of R^2 represents the accuracy of the approximation.
- Now an interesting question is whether R^2 converges to 0 as m tends to infinity for all initial conditions. If all observables g_i are in a Hilbert space, where all Cauchy sequences converge within the Hilbert space, this residual sequence, too, converges.

So far, exact linear models have been found for

- Hamiltonian systems discussed in the original paper by Koopman;
- Nonlinear autonomous systems with a single equilibrium, a single attractor / limit cycle, and a single torus.

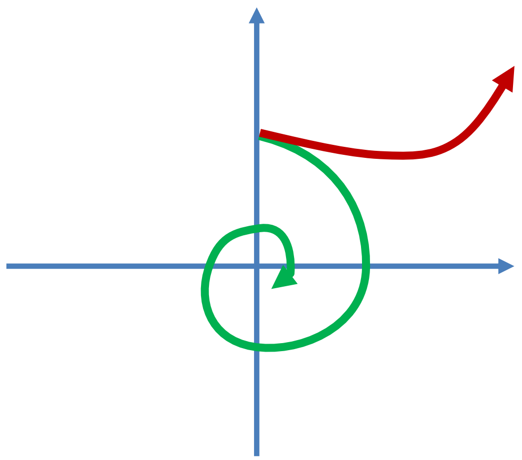




Koopman Eigenvalues and Eigenfunctions

- From the companion matrix we can obtain eigenvalues and eigen vectors of Koopman operator – empirical Ritz values and Ritz vectors.

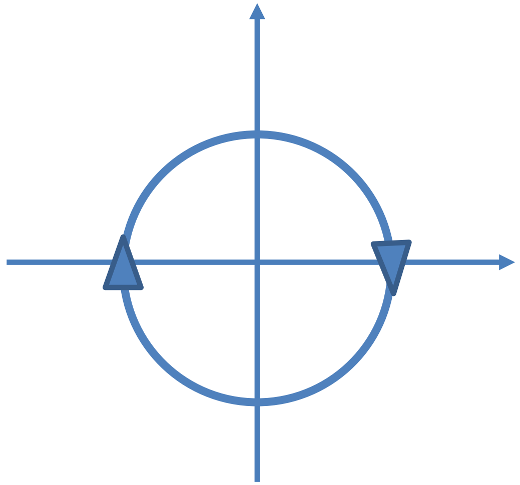
- The temporal behaviors of observables can be represented with the Koopman eigenvalues, eigen-functions, and modes.



$$\boldsymbol{g}(x_k) = \sum_{j=1}^{\infty} \lambda_j^{\,k} \varphi_j(x_0) \boldsymbol{v}_j$$

Mode:
Representing the observable w.r.t. eigen functions

Eigen-function:
Bases spanning the function space



- If one of the eigenvalues is greater than 1, that mode diverges;
- Those modes of $|\lambda_j|<1$ converge; and
- The one on the unit circle evolves on an attractor (limit cycle).

- Truncation based on eigenvalues. Fast decaying modes can be eliminated.

Drawbacks and Open Questions of the Koopman Operator Theory

- Koopman Operator for exact linearization is applicable only to autonomous systems with no control input.
- For exact linearization, the system must be lifted to an infinite dimensional space.

- There is no systematic method for finding an effective set of observables. Typically it requires a trial-and-error effort to find a good collection of observable functions.

Goals and Needs for Research

1. Establish a methodology of lifting linearization for non-autonomous systems;
2. Keep the order of a lifted system low, yet accurate enough to predict the true dynamic behavior over a required period of time; and
3. Make the lifted dynamic system causal and physically meaningful.

Goal: Linearized State Equations

The original nonlinear state equation: $$\frac{dx}{dt} = f(x,u)$$

Lifting

$$\begin{pmatrix} x \\ \eta \end{pmatrix}$$

x Independent State Variables

η Auxiliary Variables (observables)

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}_x\mathbf{x} + \mathbf{A}_\eta \eta + \mathbf{B}_x\mathbf{u} \qquad x \in \Re^n$$

$$\frac{d\eta}{dt} = \mathbf{H}_x\mathbf{x} + \mathbf{H}_\eta \eta + \mathbf{H}_u\mathbf{u} \qquad \eta \in \Re^{n_a}$$

Ad hoc methods do not work for non-autonomous systems

- Consider the following non-autonomous system with input u.

$$\frac{dx}{dt} = f(x,u) \longrightarrow \dot{x} = ax^3 + bux \quad \text{Example}$$

- Pick the following auxiliary variables (observables) for linearizing the state equation.

$$\eta_1 = x^3, \quad \eta_2 = ux$$

$$\dot{x} = ax^3 + bux \quad \rightarrow \quad \dot{x} = a\eta_1 + b\eta_2$$

- In lifting η_2, the time-derivative includes the derivative of input u, which is not causal.

$$\frac{d}{dt}\eta_2 = \frac{\partial \eta_2}{\partial x}\frac{dx}{dt} + \frac{\partial \eta_2}{\partial u}\frac{du}{dt} = u\dot{x} + x\dot{u}$$

This cannot be used as part of state equations.

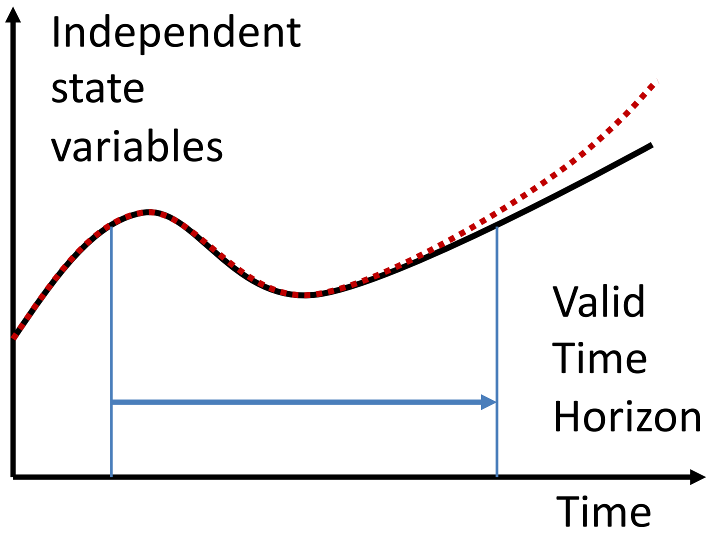
Dual-Faceted Linearization

- A systematic method for finding causal observables / auxiliary variables for lifting linearization of non-autonomous systems.

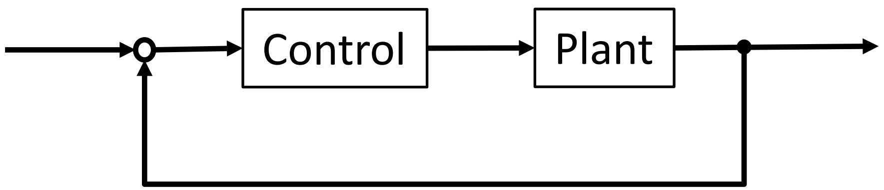
$$\frac{d\mathbf{x}}{dt} = \mathbf{A}_x \mathbf{x} + \mathbf{A}_\eta \eta + \mathbf{B}_x \mathbf{u} \quad ; \text{ Exact, No Approximation}$$

$$\frac{d\eta}{dt} = \mathbf{H}_x \mathbf{x} + \mathbf{H}_\eta \eta + \mathbf{H}_u \mathbf{u} \quad ; \text{ Lowest dimension, minimum order}$$

- Accurate; and
- Low dimension



Control Design
Real-time computation, implementation



Application: Model Predictive Control

Outline

- Bringing a physical insight into the selection of observables / auxiliary variables;
- Physical system modeling;
- Definition of auxiliary variables;
- Linear regression of the auxiliary state equations;
- Numerical examples;
- Application to Model Predictive Control; and
- Comparison to Koopman Operator

Towards a New Modeling Methodology for Nonlinear Dynamical Systems - Lifted Dynamics: Data and Physics

From the Semi-Plenary Lecture by Harry Asada

2017 American Control Conference

Seattle, May 26, 2017

d'Arbeloff Laboratory

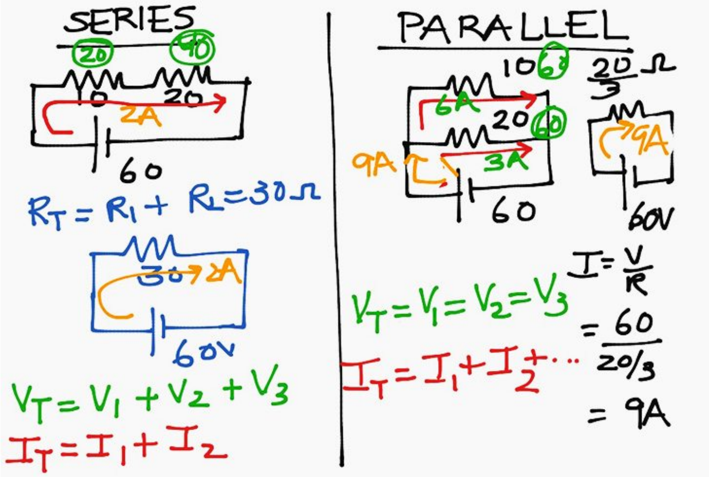
Massachusetts Institute of Technology

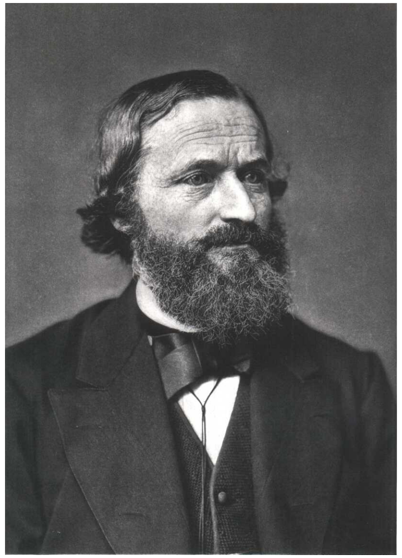
Physical systems are linear in structure;

Basic elements, like mass, spring, and dampers, are linearly connected in all dynamical systems.

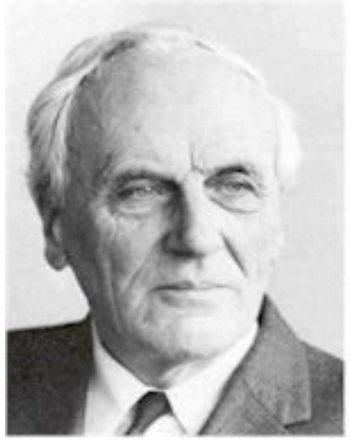
Electric Circuit Theory
(Physical Network Theory)

Kirchhoff's Voltage Law
Kirchhoff's Current Law





Gustav Kirchhoff
1882-1887



Bernard Tellegen
1900-1990

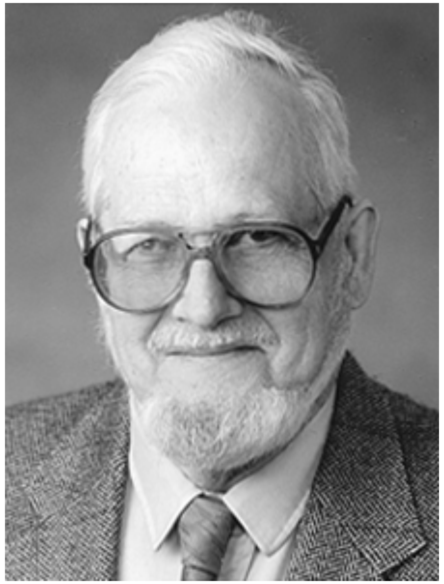
Tellegen's Theorem

$$\sum_{k=1}^{b} W_k F_k = 0.$$

Bond Graph

Professor Henry M. Paynter, MIT
(1923 – 2002)

Principle of
Physical System Modeling



Formal (NAE photo)

Casual/Natural

Bi-directional energetic interactions

Force; Voltage; Pressure

effort : $\boldsymbol{e}$

Subsystem: A ———————— Subsystem: B

flow : $\boldsymbol{f}$

Velocity; Current; Flow Rate

Causality

Connectivity: Junctions in Bond Graph

Kirchhoff's Voltage Law

||

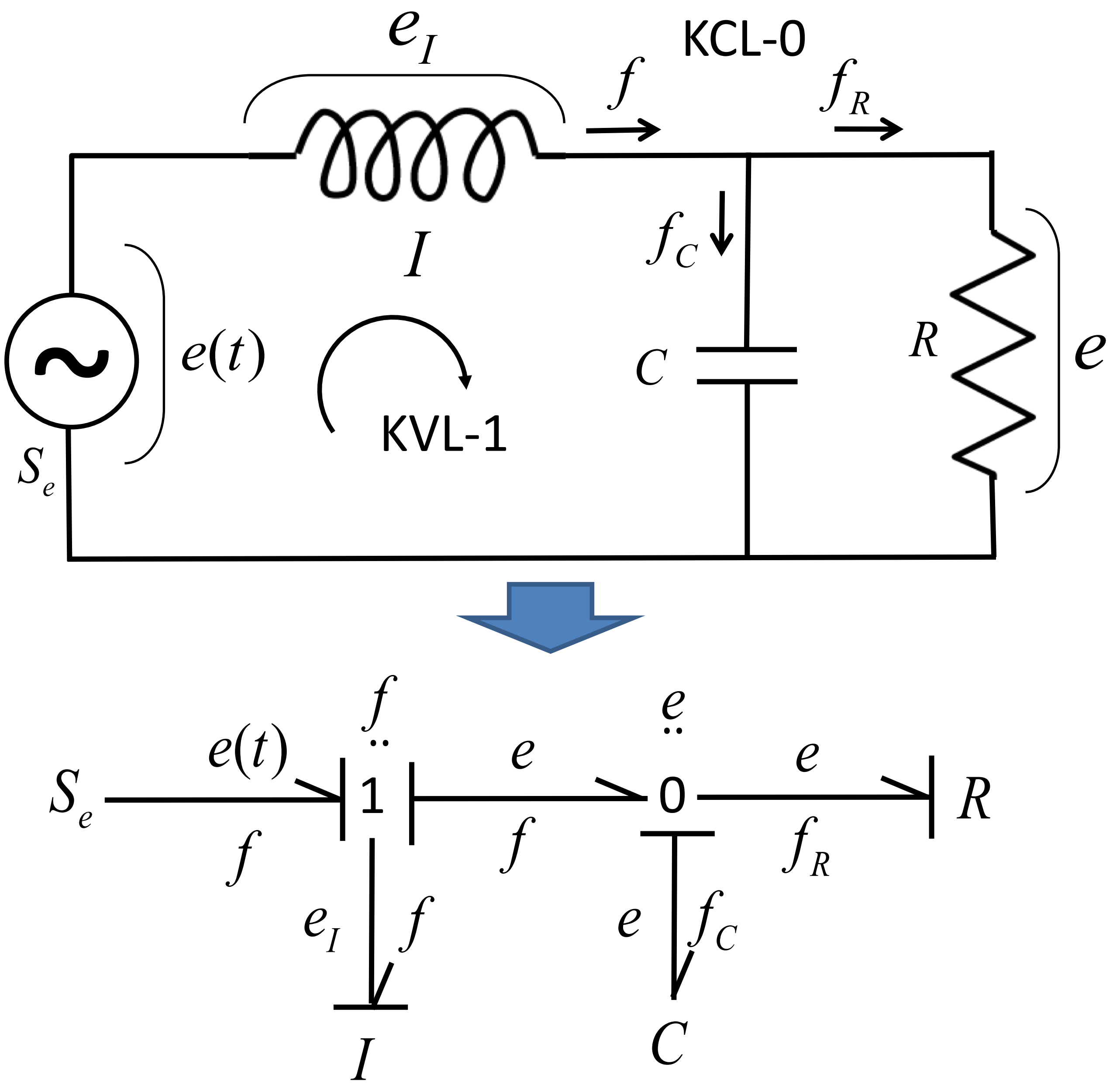
Bond Graph "1" Junction

Kirchhoff's Current Law

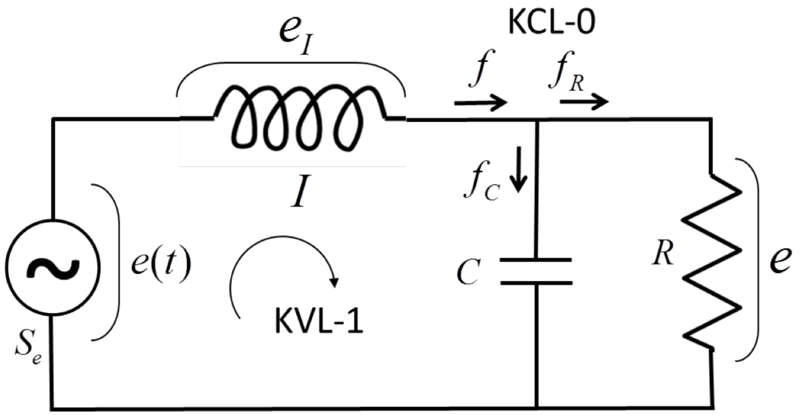
||

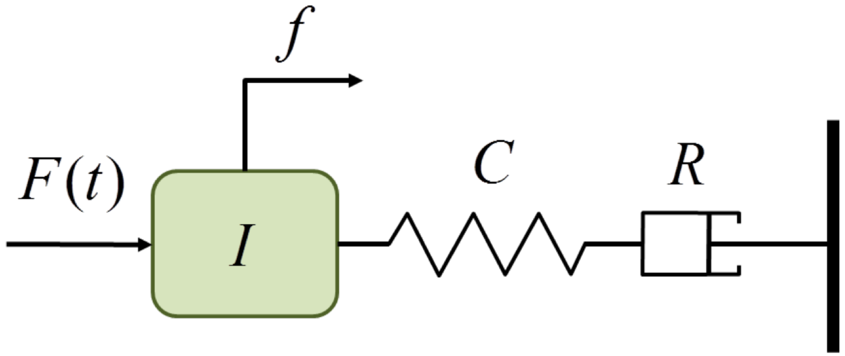
Bond Graph "0" Junction

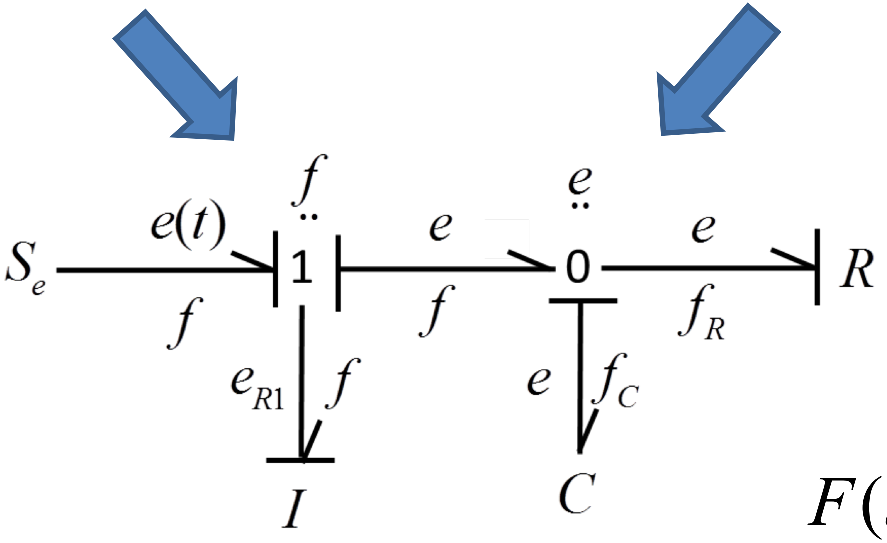
Equivalent Bond Graph



Electrical-Mechanical Equivalence







$$e(t) - e_I - e = 0$$

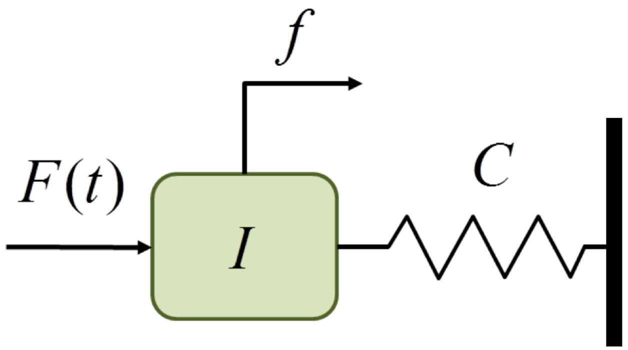
$$F(t) - m\frac{df}{dt} - e = 0$$

Kirchhoff's Voltage Law

Newton's equation of Motion

All the junction conditions are linear.

Connection v.s. Element's Property (Constitutive Law)



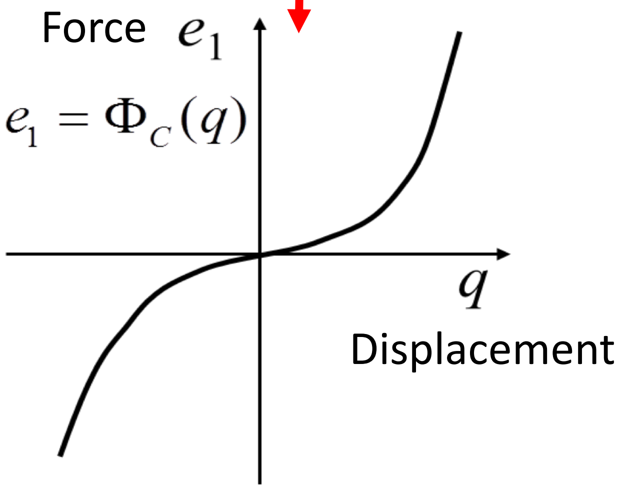
Newton's equation of Motion

$$F(t) - m\frac{df}{dt} - e_1 = 0$$

Nonlinear spring:

Force is a nonlinear function of Displacement.

Note that, if you do not replace force e by the nonlinear constitutive law with displacement q, <u>the equation is linear</u>.

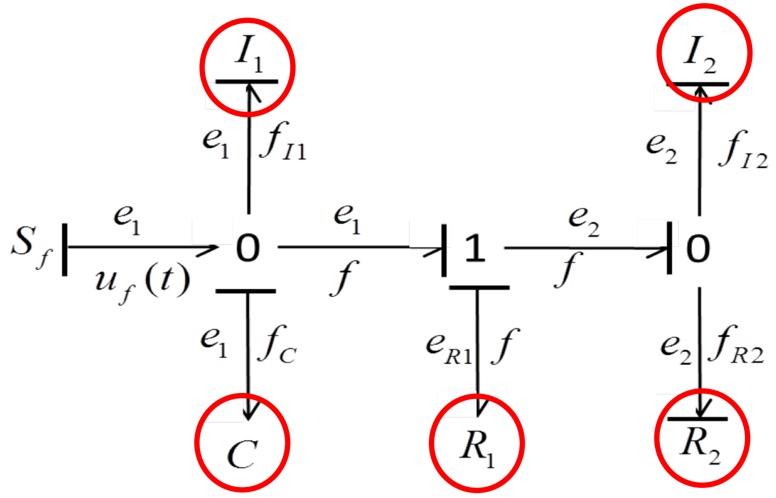


Element's Constitutive Law

Introducing Auxiliary Variables

Definition (a complete set of auxiliary variables):

Output variables of all the nonlinear elements

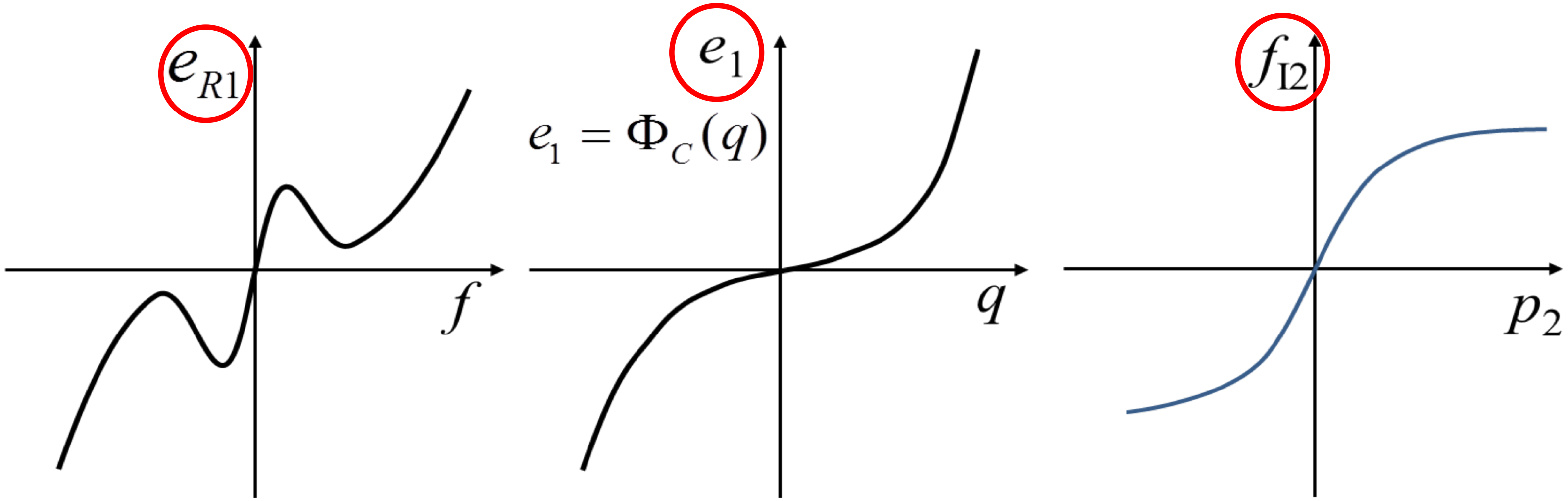


$$\eta = \begin{pmatrix} e_1 \\ f_{I1} \\ f_{I2} \\ e_{R1} \\ f_{R2} \end{pmatrix} \begin{array}{l} \cdots C \\ \cdots I_1 \\ \cdots I_2 \\ \cdots R_1 \\ \cdots R_2 \end{array}$$

$e_{R1} = \Phi_{R1}(f)$

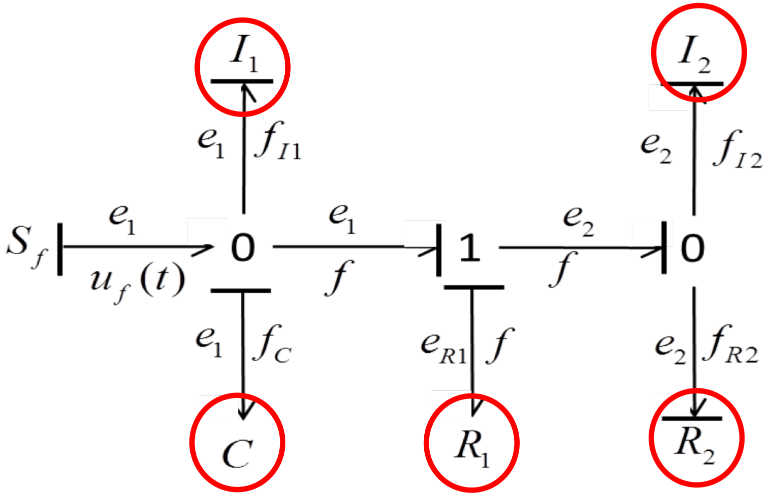
$e_1 = \Phi_C(q)$

$f_{I2} = \Phi_I(p_2)$



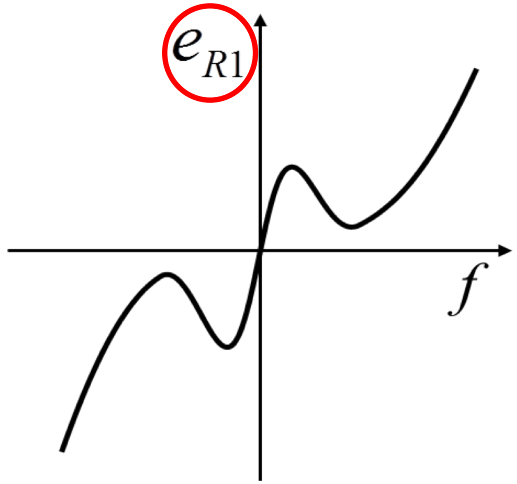
These Auxiliary Variables are Minimum.

Output variables of all the nonlinear elements:



$$
\eta = \begin{array}{c}\left(\begin{array}{c} e_1 \\ f_{I1} \\ f_{I2} \\ e_{R1} \\ f_{R2} \end{array}\right)\end{array}
\begin{array}{l} \cdots C \\ \cdots I_1 \\ \cdots I_2 \\ \cdots R_1 \\ \cdots R_2 \end{array}
$$

$$e_{R1} = \Phi_{R1}(f)$$



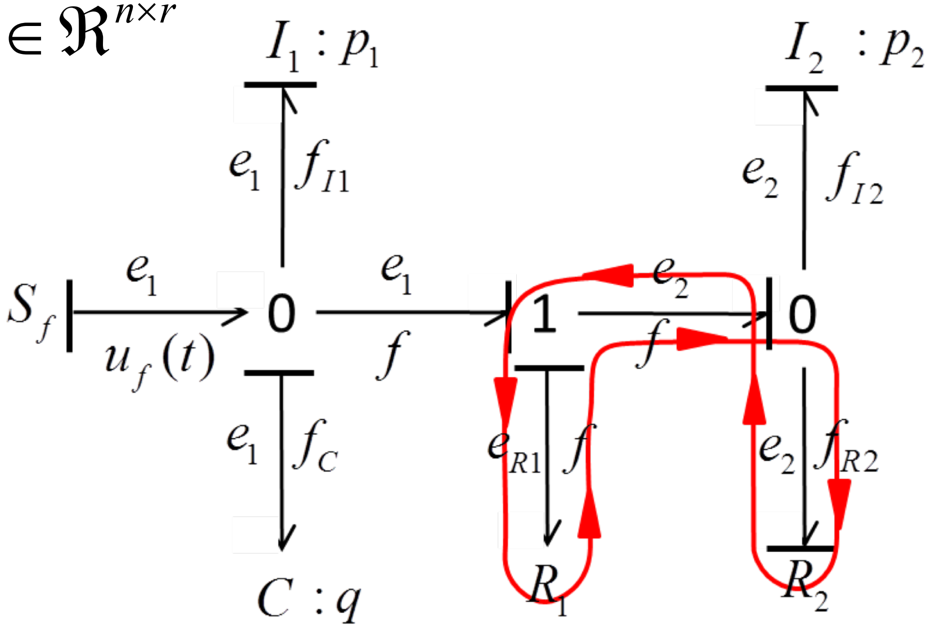
- With these auxiliary variables, all the junction conditions (KVL, KCL) can be written as linear equations.
- If any auxiliary variable is missing, then some junction conditions cannot be written as linear equations.
- Therefore, this set of auxiliary variables is the minimum set for obtaining an exact linear dynamic equation.

Theorem 1

State equations of a nonlinear lumped-parameter system that possesses a Bond Graph of integral causality and that contains n energy storage elements, n_a nonlinear elements, and r sources or exogenous inputs can be expressed as a linear equation in terms of n state variables, $\mathbf{x} \in \Re^n$, n_a auxiliary variables, $\eta \in \Re^{n_a}$, and inputs, $\mathbf{u} \in \Re^r$, as

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}_x \mathbf{x} + \mathbf{A}_\eta \eta + \mathbf{B}_x \mathbf{u}$$

where. $\mathbf{A}_x \in \Re^{n \times n}, \mathbf{A}_\eta \in \Re^{n \times n_a}, \mathbf{B}_x \in \Re^{n \times r}$



No algebraic problem occurs.

How will auxiliary variables evolve with time?

What should we do for the dynamics of auxiliary variables?

$$\eta = \eta(\mathbf{x}) \quad ^* \qquad \frac{d\mathbf{x}}{dt} = \mathbf{A}_x \mathbf{x} + \mathbf{A}_\eta \eta + \mathbf{B}_x \mathbf{u}$$

A naïve method; Take Taylor expansion and evaluate the Jacobian at a point.

$$\dot{\eta} = \mathbf{J}(\mathbf{x}) \cdot \dot{\mathbf{x}} \cong \bar{\mathbf{J}} \cdot \dot{\mathbf{x}} \quad \text{where} \quad \mathbf{J}(\mathbf{x}) = \frac{\partial \eta}{\partial \mathbf{x}}, \quad \bar{\mathbf{J}} = \left. \frac{\partial \eta}{\partial \mathbf{x}} \right|_{\bar{\mathbf{x}}}$$

then $\dot{\eta}$ and $\dot{\mathbf{x}}$ are collinear. The second state equation of $\dot{\eta}$ does not provide any new information.

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}_x \mathbf{x} + \mathbf{A}_\eta \eta + \mathbf{B}_x \mathbf{u}$$

$$\frac{d\eta}{dt} \cong \bar{\mathbf{J}} \cdot \dot{\mathbf{x}} = \bar{\mathbf{J}} \mathbf{A}_x \mathbf{x} + \bar{\mathbf{J}} \mathbf{A}_\eta \eta + \bar{\mathbf{J}} \mathbf{B}_x \mathbf{u}$$

* If an auxiliary variable is explicitly dependent on input u, we have to treat it differently.

Lifting Linearization

Let us first assume that all the auxiliary variables are causal, independent of input u, that is,

$$\eta = \eta(\mathbf{x}) \qquad \frac{d\mathbf{x}}{dt} = \mathbf{A}_x \mathbf{x} + \mathbf{A}_\eta \eta + \mathbf{B}_x \mathbf{u}$$

We apply statistic linearization to the time derivative of the auxiliary variables: Linear Regression.

$$\hat{\dot{\eta}} = \mathbf{H}_x \mathbf{x} + \mathbf{H}_\eta \eta + \mathbf{H}_u \mathbf{u}$$

$$\mathbf{H} \triangleq (\mathbf{H}_x, \mathbf{H}_\eta, \mathbf{H}_u) = \arg\min_{\mathbf{H}} E\left[|\hat{\dot{\eta}} - \dot{\eta}|^2 \right]$$

Lifting Linearization - 2

$$\mathbf{H}^0 = \arg\min_{\mathbf{H}} E\left[|\hat{\dot{\eta}} - \dot{\eta}|^2 \right] \qquad \hat{\dot{\eta}} = \mathbf{H}_x \mathbf{x} + \mathbf{H}_\eta \eta + \mathbf{H}_u \mathbf{u} = \mathbf{H}\xi$$

where $\mathbf{H} \triangleq (\mathbf{H}_x, \mathbf{H}_\eta, \mathbf{H}_u) \in \Re^{n_a \times \ell}$;

$$\xi \triangleq \begin{pmatrix} \mathbf{x} \\ \eta \\ \mathbf{u} \end{pmatrix} \in \Re^{\ell \times 1}, \qquad \ell = n + n_a + r$$

Using the standard least squares estimate:

$$\mathbf{H}^0 = E\left[\dot{\eta}\xi^T\right](E\left[\xi\xi^T\right])^{-1}$$

Data matrix $\qquad \Xi = \left\{\xi^1, \xi^2, \cdots, \xi^N\right\}$

Lifting Linearization - 3

$$\mathbf{H}^0 = E\left[\dot{\eta}\xi^T\right]\left(E\left[\xi\xi^T\right]\right)^{-1} \qquad \xi \triangleq \begin{pmatrix} \mathbf{x} \\ \eta \\ \mathbf{u} \end{pmatrix} \in \Re^{\ell \times 1}$$

The data matrix $\Xi = \left\{\xi^1, \xi^2, \cdots, \xi^N\right\}$ must satisfy:

- Persistent excitation;
- No feedback: **x** and **u** are not correlated; and
- All the elements are nonlinear.

Note: If some elements would have linear constitutive laws, the auxiliary variables would be collinear with the state variables, making matrix $E\left[\xi\xi^T\right]$ singular.

$$\hat{\dot{\eta}} = \mathbf{H}_x\mathbf{x} + \mathbf{H}_\eta\eta + \mathbf{H}_u\mathbf{u} = \mathbf{H}\xi \qquad \mathbf{H} \triangleq (\mathbf{H}_x, \mathbf{H}_\eta, \mathbf{H}_u) \in \Re^{n_a \times \ell}$$

Lifting Linearization - 5

$$\dot{\eta} = \mathbf{J}(\mathbf{x}) \cdot \dot{\mathbf{x}} \cong \bar{\mathbf{J}} \cdot \dot{\mathbf{x}} \iff \hat{\dot{\eta}} = \mathbf{H}_x \mathbf{x} + \mathbf{H}_\eta \eta + \mathbf{H}_u \mathbf{u}$$

$$= (\mathbf{H}_x, \mathbf{H}_\eta, \mathbf{H}_u) \begin{pmatrix} \mathbf{x} \\ \eta \\ \mathbf{u} \end{pmatrix} = \mathbf{H}\xi$$

Theorem 2

The linear regression based on statistical linearization using data matrix Ξ provides more accurate estimate of $\dot{\eta}$ than that of a fixed Jacobian $\bar{\mathbf{J}}$.

$$\min_{\mathbf{H} \in \Re^{n_a \times \ell}} E\left[|\dot{\eta} - \mathbf{H}\xi|^2 \right] \le \min_{\mathbf{J} \in \Re^{n_a \times n}} E\left[|\dot{\eta} - \bar{\mathbf{J}}\dot{\mathbf{x}}|^2 \right]$$

Proof : $\bar{\mathbf{J}}\dot{\mathbf{x}} = \bar{\mathbf{H}}\xi$ is a special case of $\mathbf{H}$; $\bar{\mathbf{H}} = \left[\bar{\mathbf{J}}\mathbf{A}_x \quad \bar{\mathbf{J}}\mathbf{A}_\eta \quad \bar{\mathbf{J}}\mathbf{B}_x \right]$

Reconsider State Variables

If a system is linear, transformation of state variables does not change the structure of state equations.
For example,

$$F = -\,k\,x \quad \text{--------} \quad \text{Linear spring}$$

The state equation using x and the one with F are basically the same.

$$F = m\frac{d^2x}{dt} \rightarrow m\frac{d^2x}{dt} + kx = 0$$

$$F = m\frac{d^2x}{dt} \rightarrow \frac{m}{k}\frac{d^2F}{dt^2} + F = 0$$

Reconsider State Variables

$F = -\,k\,x$ -------- Linear spring

The two state equations are basically the same.

It is not the case for a nonlinear system with a nonlinear element constitutive law.

$F = -a\,x - b\,x^3$ ---- Nonlinear spring

The state equation in terms of *F* and the one with *x* differ in representation, exhibiting different facets of the nonlinear system.

$$F = m\frac{d^2x}{dt}$$

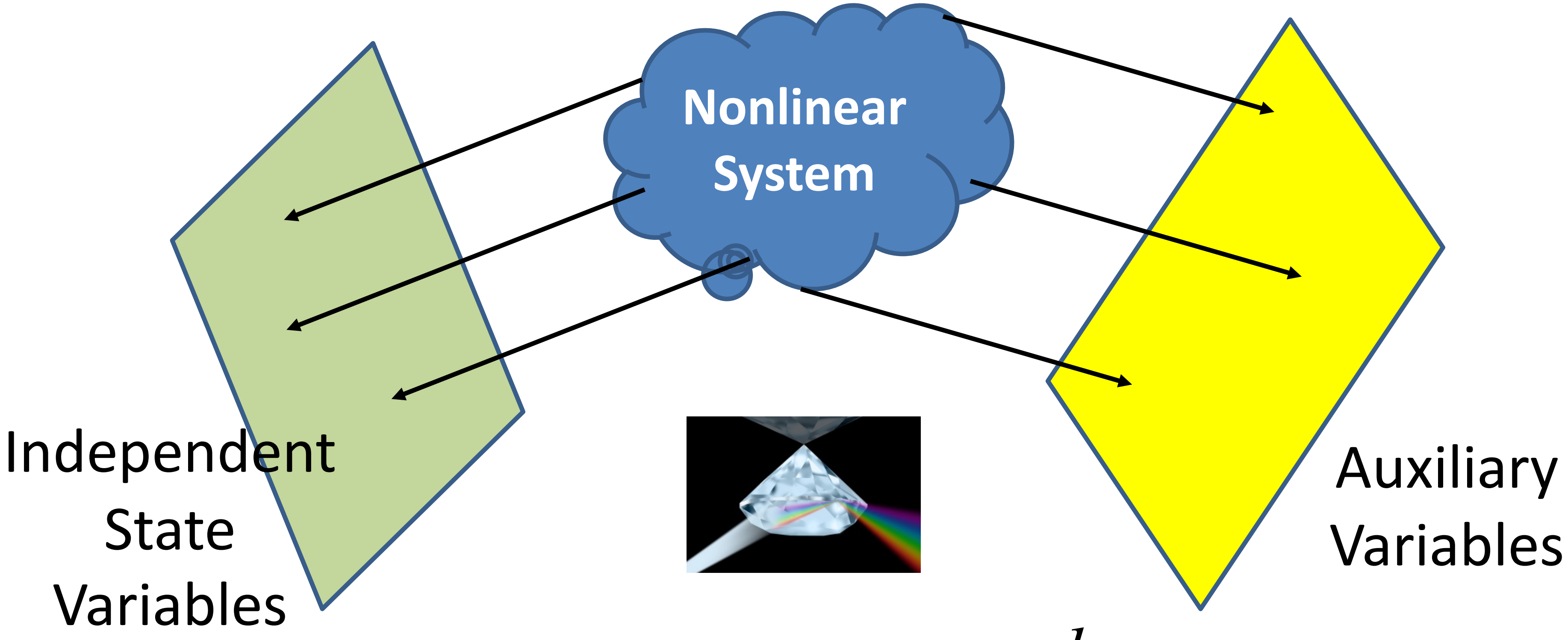
$$m\ddot{x} + ax + bx^3 = 0$$

$$mg'(F)\ddot{F} + mg''(F)\dot{F}^2 + F = 0$$

$$x = g(F),\; g^{(n)}(F) = \frac{d^nF}{dF^n}$$

Dual Faceted Representation

Viewing a nonlinear system from two sets of variables, we can obtain a richer representation of the nonlinear system.

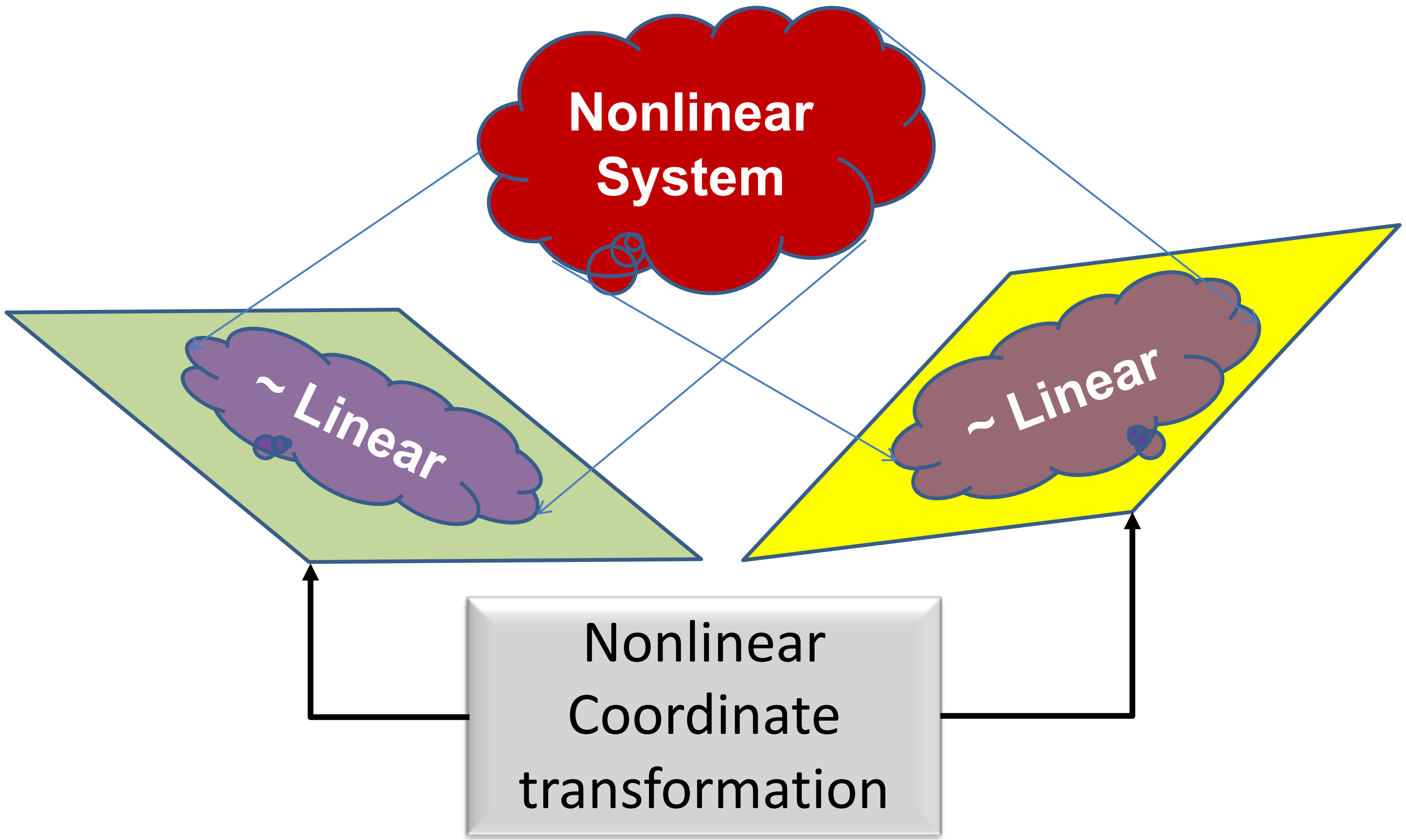


$$\frac{d\mathbf{x}}{dt} = \mathbf{A}_x\mathbf{x} + \mathbf{A}_\eta \eta + \mathbf{B}_x\mathbf{u}$$

$$\frac{d\eta}{dt} = g(\eta, x, u)$$

Linearize this

Dual Faceted Linearization



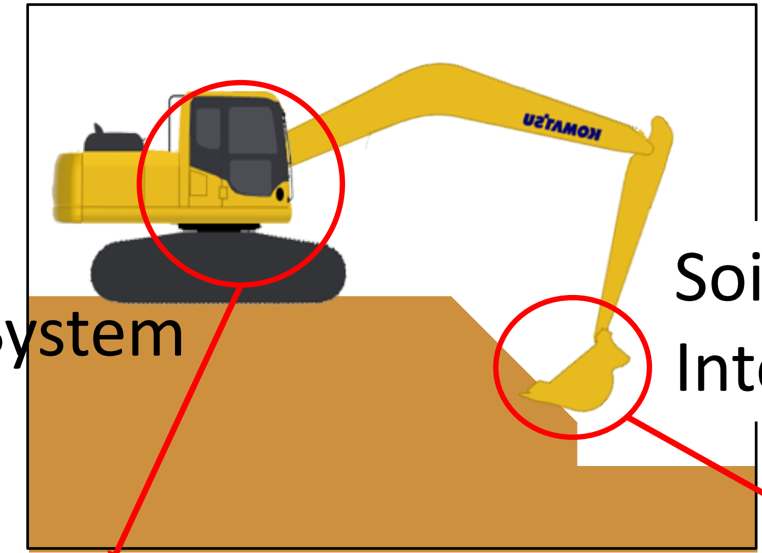
$$\frac{d\mathbf{x}}{dt} = \mathbf{A}_x\mathbf{x} + \mathbf{A}_\eta\eta + \mathbf{B}_x\mathbf{u}$$

$$\eta = \eta(\mathbf{x})$$

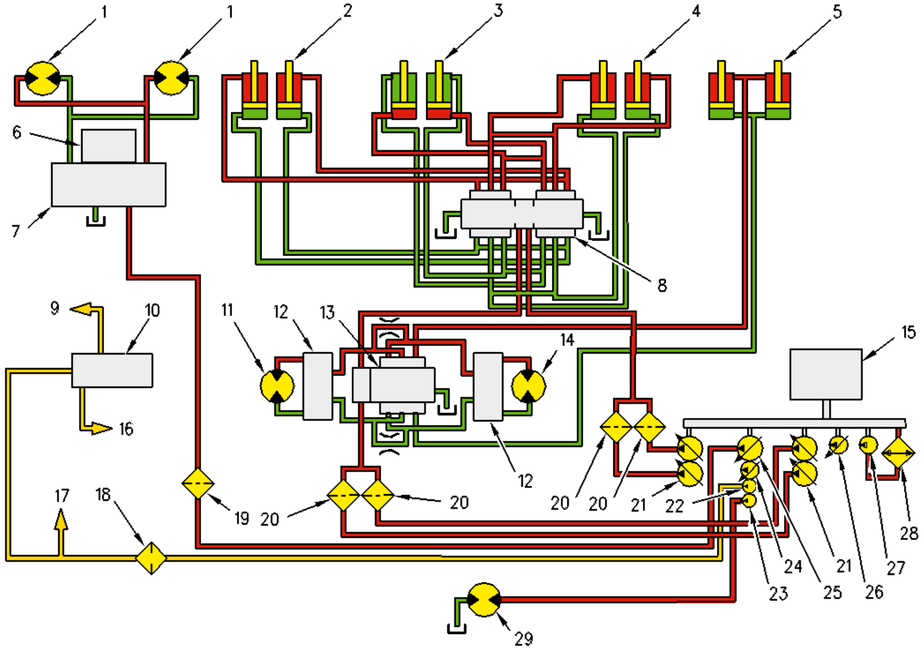
$$\frac{d\eta}{dt} = \mathbf{H}_x\mathbf{x} + \mathbf{H}_\eta\eta + \mathbf{H}_u\mathbf{u}$$

Numerical Examples and Applications

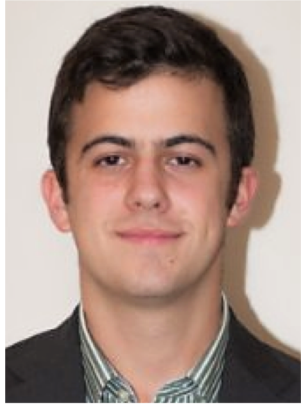
Autonomous Excavators and Dump Trucks



Hydraulic System



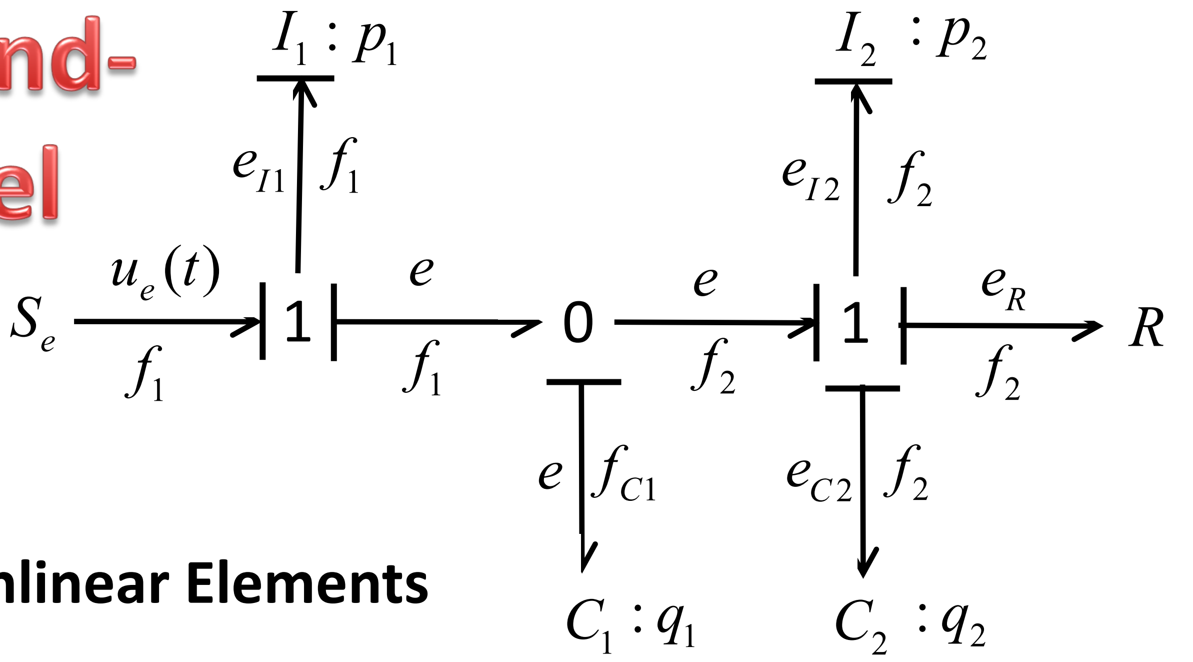
Soil-Bucket Interactions



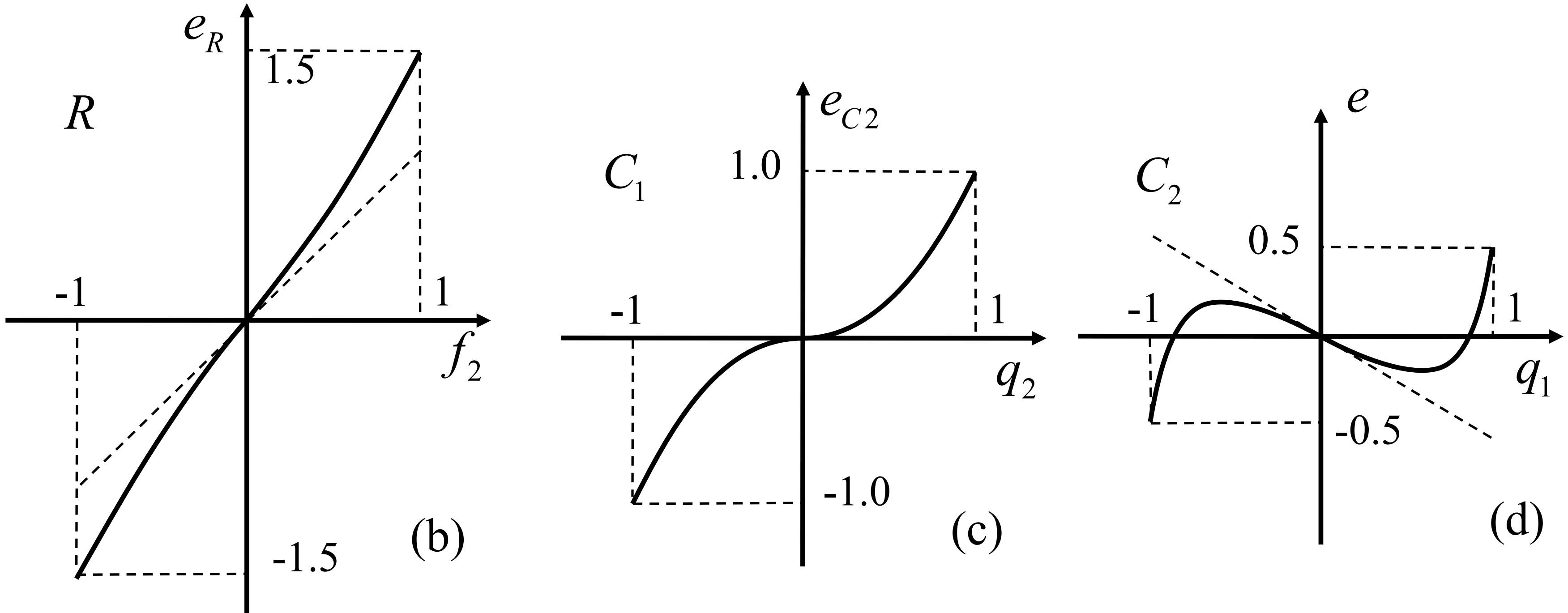
Filippos Sotiropoulos

Nonlinearity is everywhere.

Simplified Bond-Graph Model

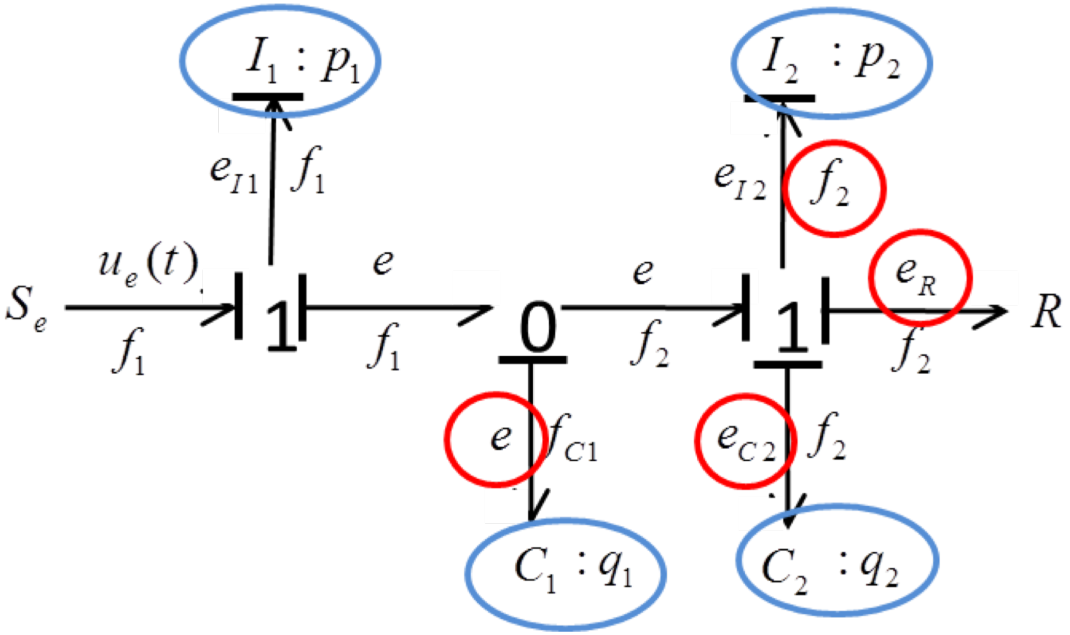


Constitutive Laws of Nonlinear Elements



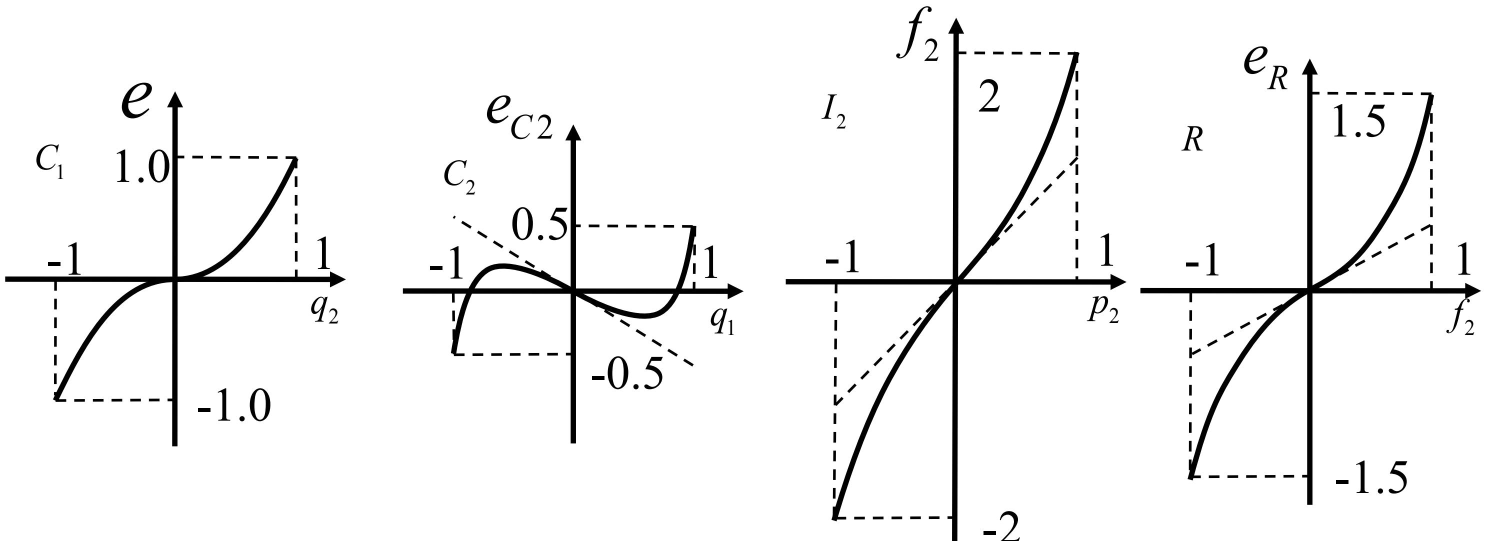
Independent State Variables

$$x = \begin{pmatrix} q_1 \\ q_2 \\ p_1 \\ p_2 \end{pmatrix}$$



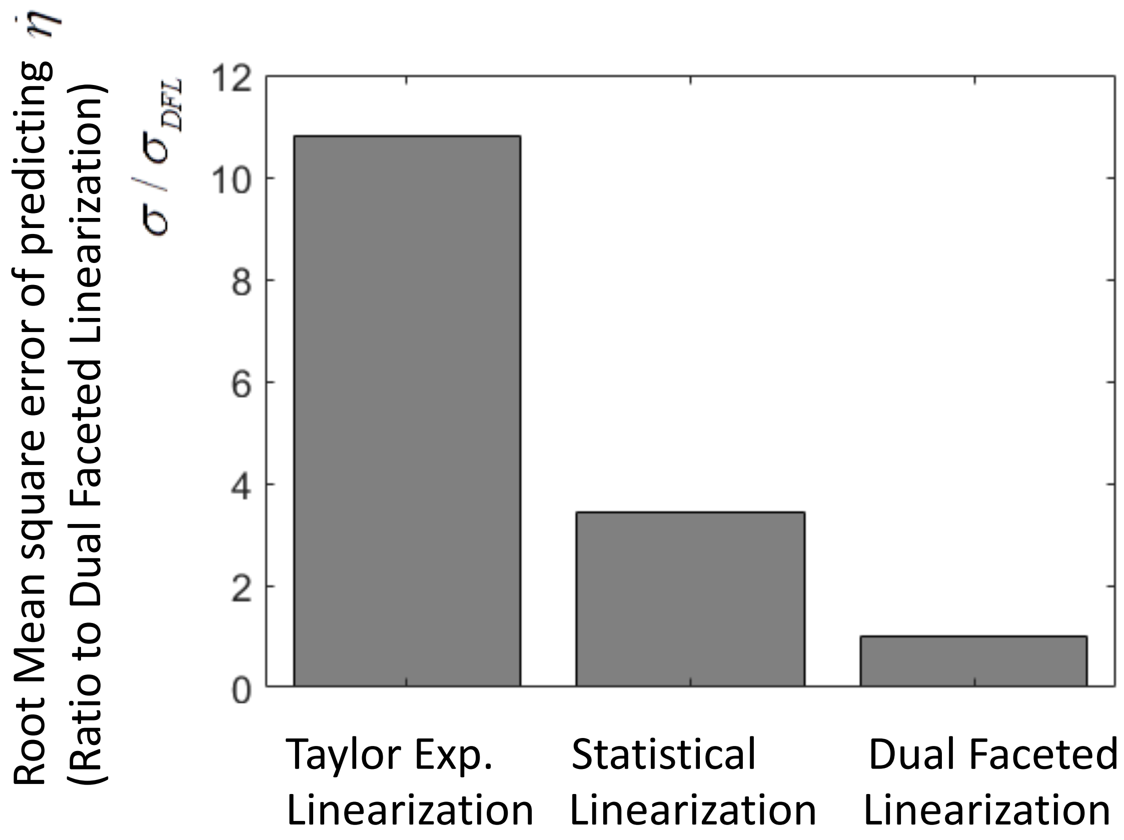
Auxiliary Variables

$$\eta = \begin{pmatrix} e \\ e_{C2} \\ f_2 \\ e_R \end{pmatrix}$$

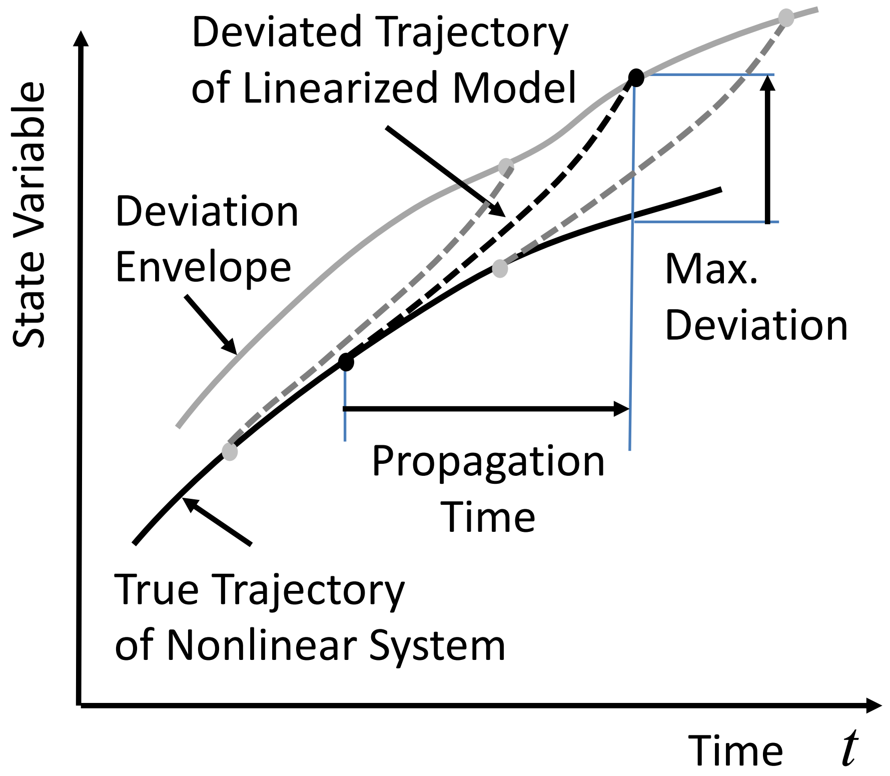


Comparison of Linearization Methods

$$\sigma = \sqrt{E\left[\,|\hat{\dot{\eta}} - \dot{\eta}|^2\,\right]} \qquad \eta = \begin{pmatrix} e & e_{C2} & f_2 & e_R \end{pmatrix}^T$$



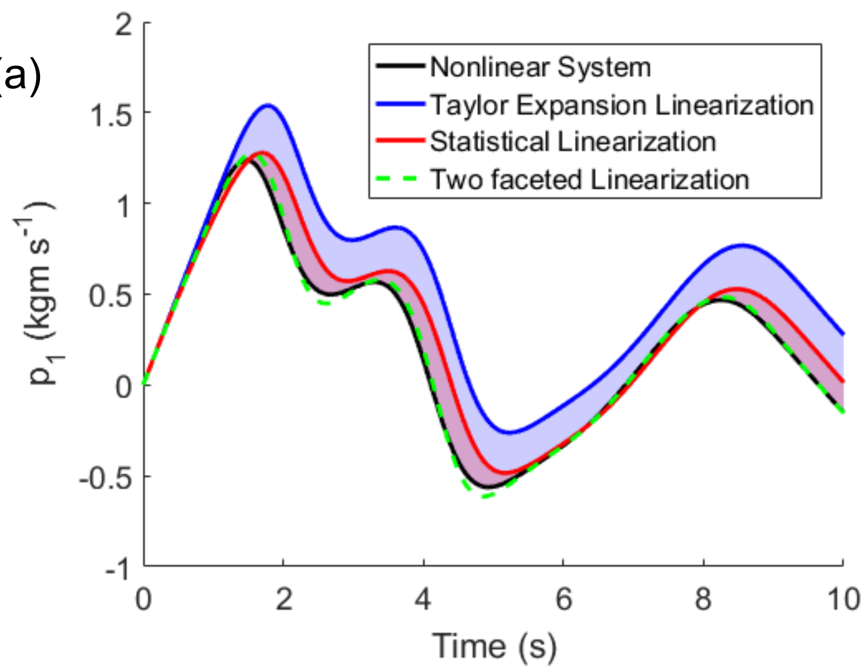
Linearization Accuracy v.s. Time Horizon



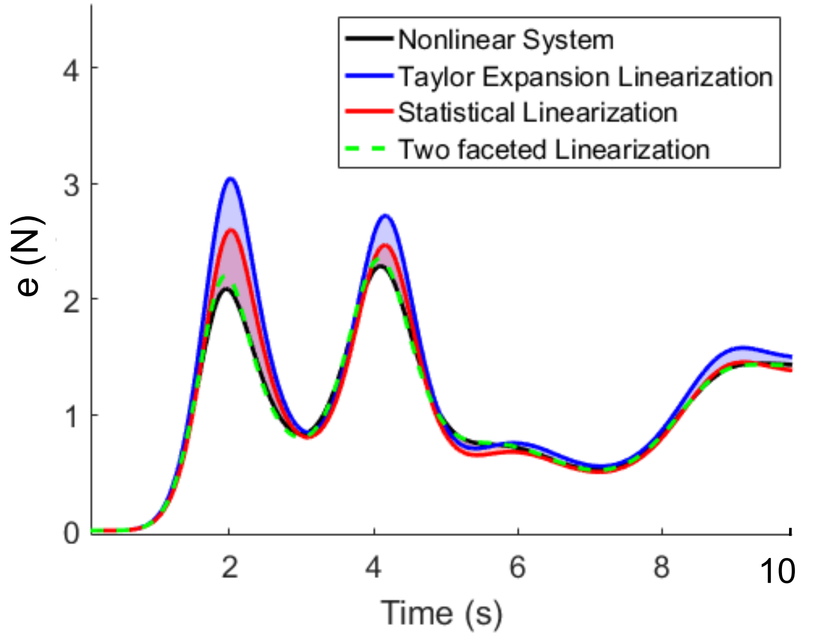
Application: Model Predictive Control

Oyama, Yamakita, Asada, “Approximated Stochastic Model Predictive Control using Statistical Linearization of Nonlinear Dynamical System in Latent Space”, CDC 2016

(a)

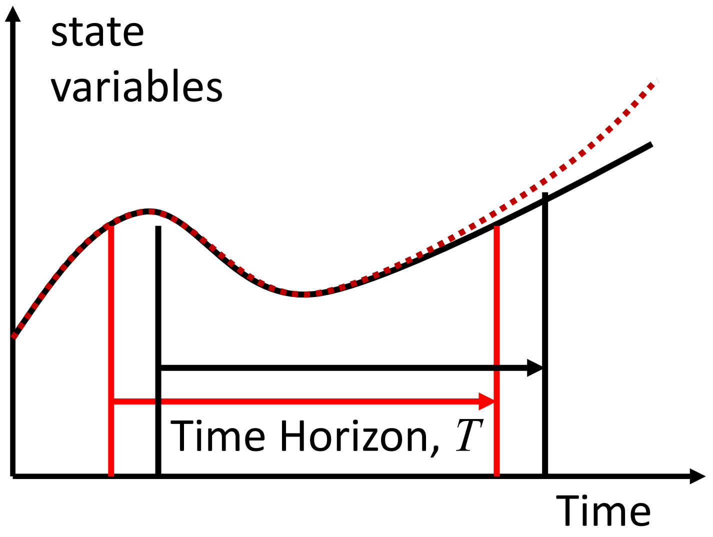


(b)



Application to Model Predictive Control

- Optimal control over a finite time horizon;
- Execute the optimal control for the current timeslot only;
- Repeat the computation of optimal control over the succeeding time horizon.

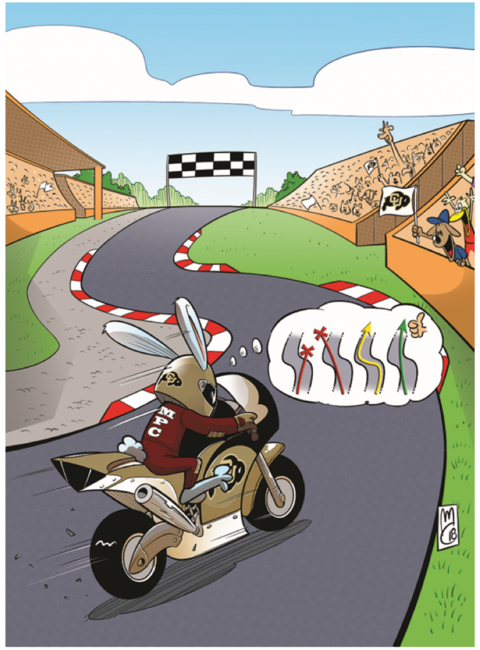


Minimize

$$V = \int_t^{t+T} \left(x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau) \right) d\tau + (Terminal\ Costs)$$

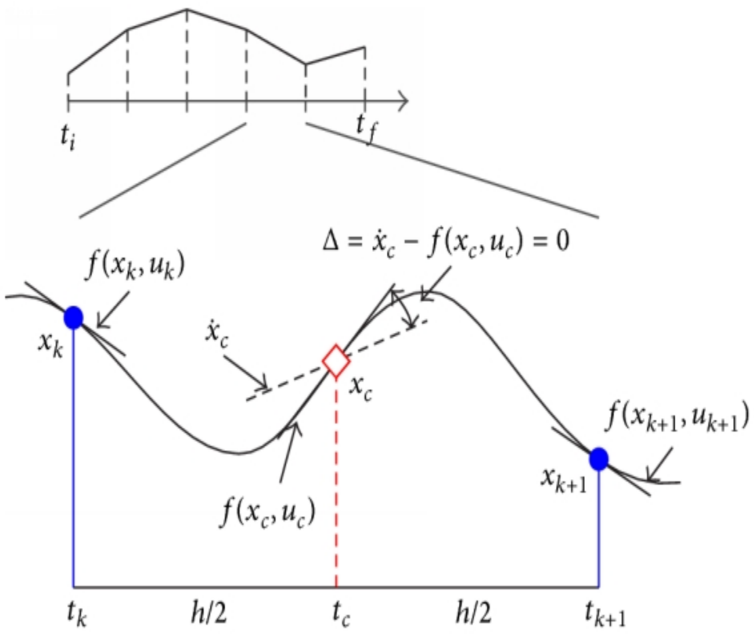
Subject to

$$\frac{dx}{dt} = f(x,u)$$



Nonlinear Model Predictive Control

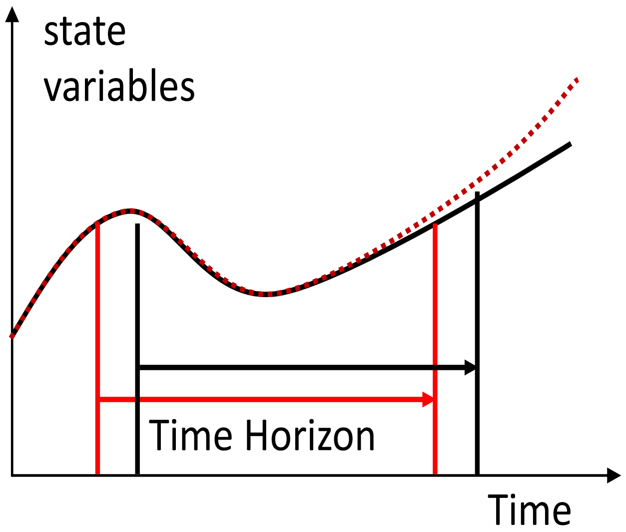
- It is hard, very hard for real-time applications.
- Direct Collocation Method, etc. Effective solution algorithms for solving nonlinear optimal control problems
- Computationally expensive Curse of Dimensionality
- Local minima problem



MPC using Dual-Faceted Linearization

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}_x\mathbf{x} + \mathbf{A}_\eta\eta + \mathbf{B}_x\mathbf{u} \qquad \frac{d\eta}{dt} = \mathbf{H}_x\mathbf{x} + \mathbf{H}_\eta\eta + \mathbf{H}_u\mathbf{u}$$

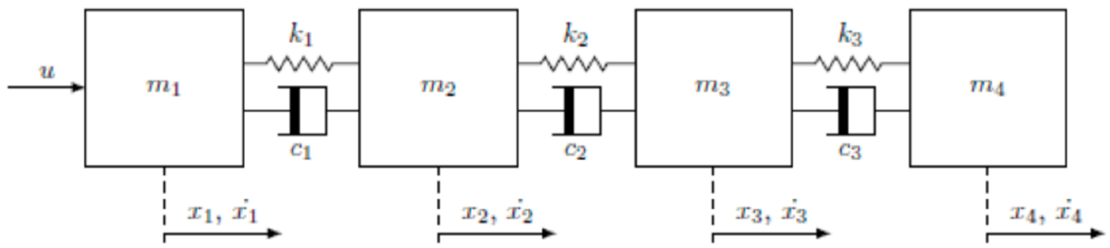
- Linear MPC is a convex optimization problem:
  - No local minima problem
  - More than 100 times faster than nonlinear MPC (Direct Collocation Method)
- High accuracy approximation over a finite time horizon



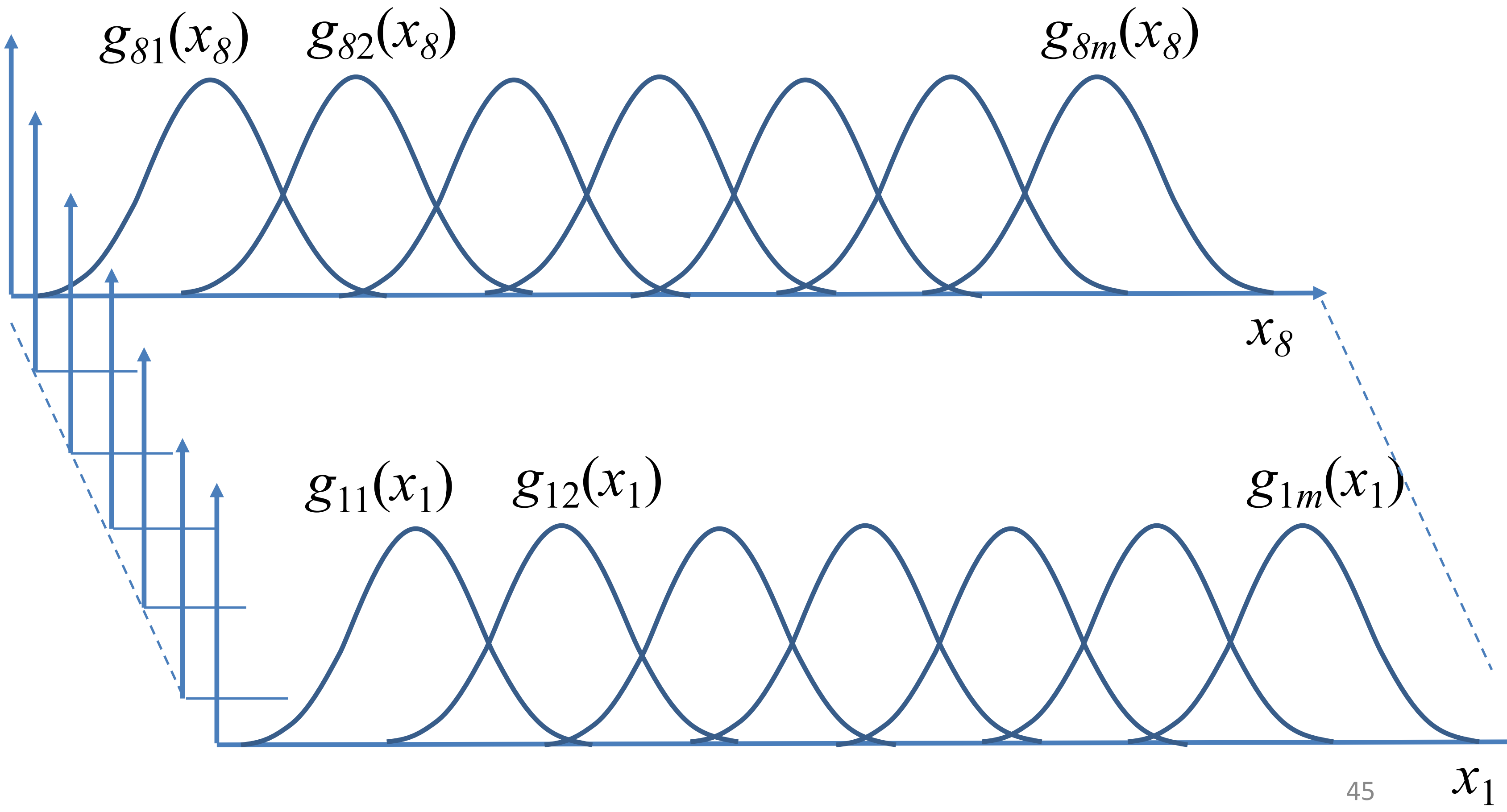
Simulation Comparison

	MSE	Calculation time[sec]
Koopman	0.0389	0.00215
DFL	0.0364	0.00135
Taylor	0.121	0.00131
NP	0.0371	0.0845
NP(Direct Collocation)	0.0376	0.212

Koopman Observables

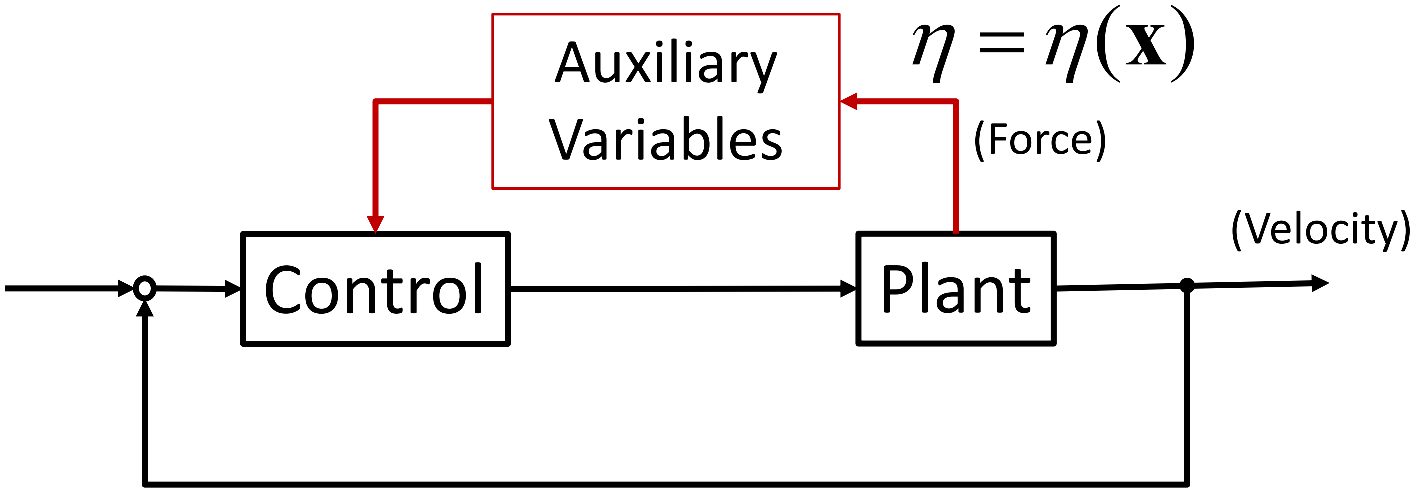


- ❑ Used Radial Basis Functions for observables.
- ❑ For each state variable, we defined 25 RBFs.
- ❑ The order of the system is 8.
- ❑ Total 200 RBFs. The system order 208.

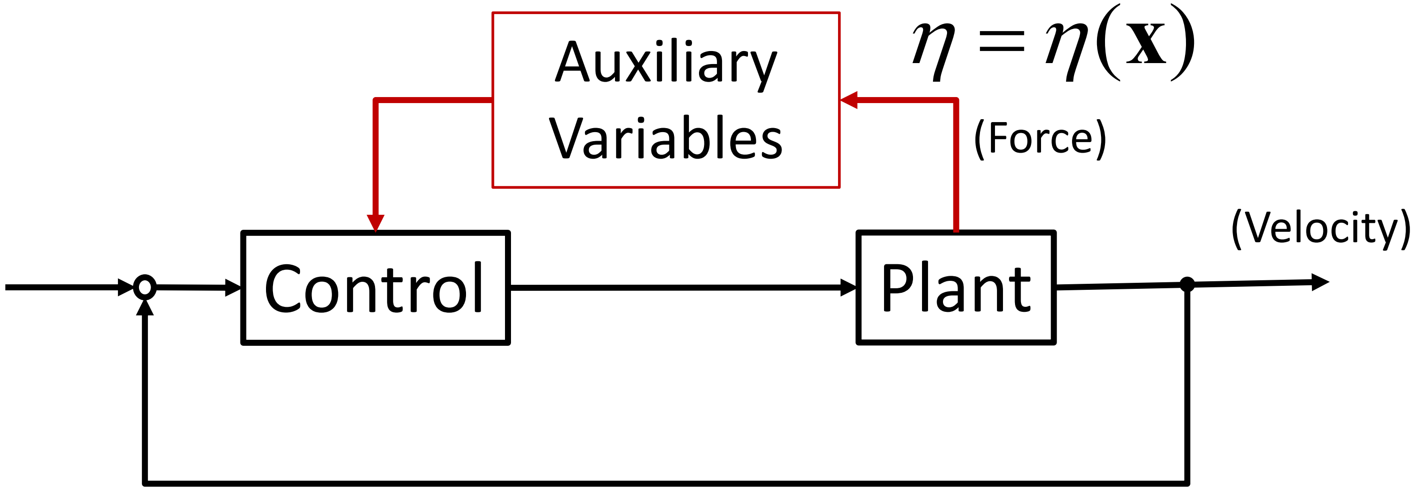


Lifted State Feedback better informs the controller

- For a linear system, feedback of additional variables does not make sense;
- However, for a nonlinear system, the use of auxiliary variables may better inform the controller;
- In fact, DFL-MPC outperformed nonlinear MPC.



Lifted State Feedback better informs the controller



$$V^* = \int_t^{t+T} \left(\begin{pmatrix} x(\tau) \\ \eta(\tau) \end{pmatrix}^T Q^* \begin{pmatrix} x(\tau) \\ \eta(\tau) \end{pmatrix} + u(\tau)^T R u(\tau) \right) d\tau$$

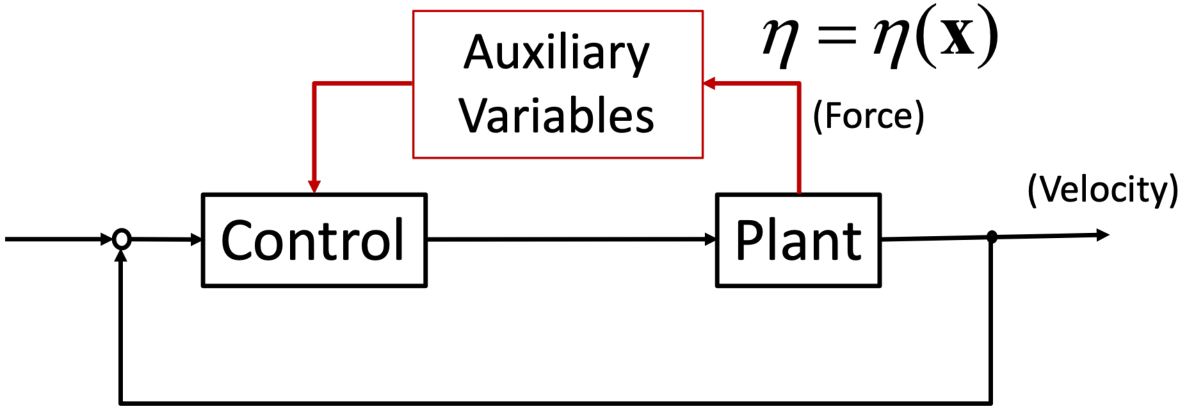
Standard $\quad Q^* = \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}, \qquad Q^* = \begin{pmatrix} Q & Q_{x\eta}^T \\ Q_{x\eta} & Q_{\eta\eta} \end{pmatrix}$

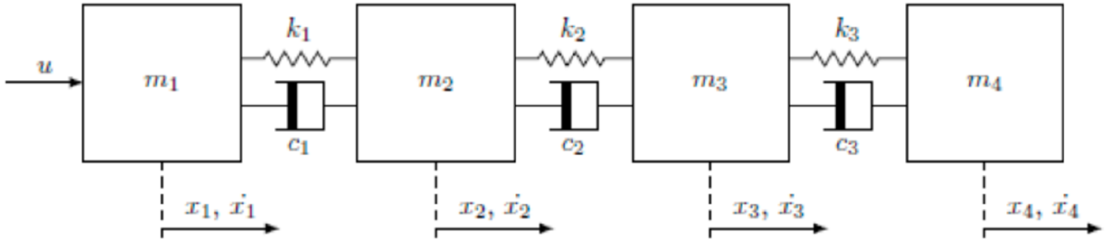
Augmented Cost Functional

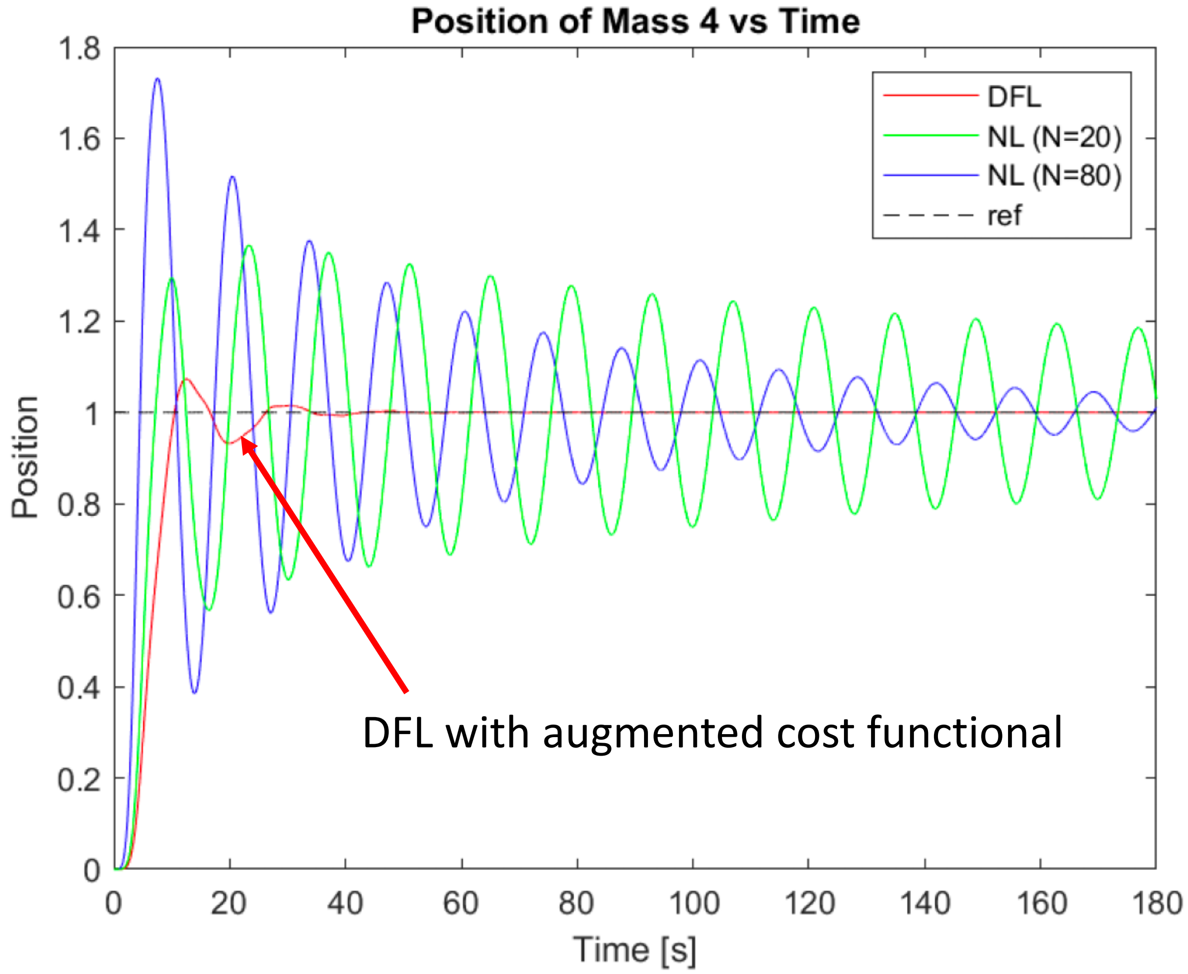
LQR solution

$$u(t) = -K_x x(t) - K_\eta \eta(t)$$

Lifted State Feedback better informs the controller

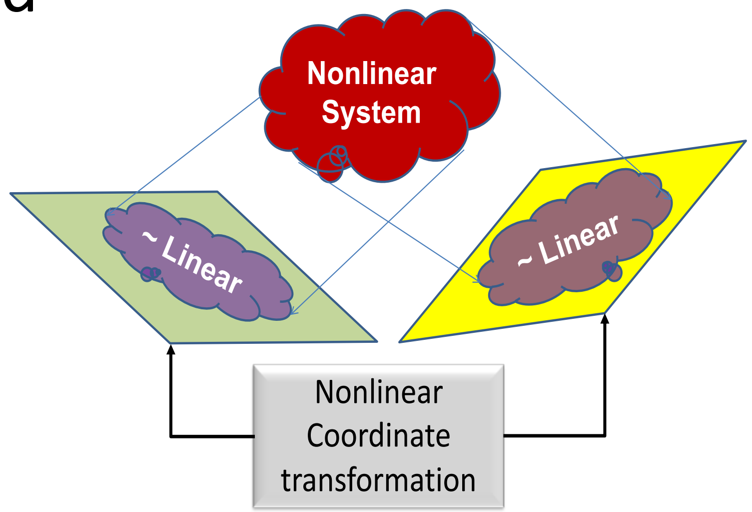


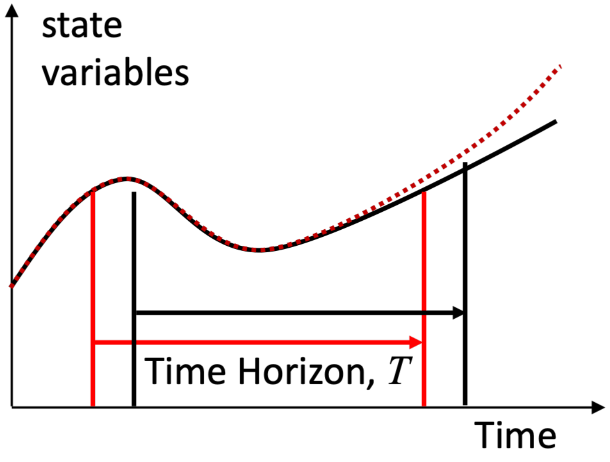




Summary of DFL-MPC

- Based on physical modeling theory, low-dimensional lifted linear state equations have been obtained for non-autonomous systems.
- Those observables, or auxiliary variables, are physically meaningful, and may be measurable.
- Augmented state feedback with auxiliary variables can better inform an optimal controller.
- DFL is particularly useful for Model Predictive Control for a finite time horizon, where DFL provides accurate prediction.
- Linear representation of nonlinear dynamics opens up new possibilities and challenges:
  - Observability and controllability of lifted dynamical systems
  - Observer design for augmented state feedback
  - Conversion of non-convex optimization to convex optimization
  - Approximation error analysis





Conclusion of 2.160: Identification, Estimation, and Learning

- The cross-disciplinary field of system dynamics, statistics, and machine learning is an exciting, emerging area.
- System identification, state-parameter-function estimation, and supervised learning can be studied in a cohesive manner.

- I hope that you will find the materials of 2.160 useful for your future work.

Thank you.

While the merry bells keep ringing