Finite-Time Thermodynamics of Port-Hamiltonian Systems

ROYAL INSTITUTE OF TECHNOLOGY

Henrik Sandberg

Automatic Control Lab, ACCESS Linnaeus Centre, KTH

(Currently on sabbatical leave at LIDS, MIT)

Jean-Charles Delvenne

CORE, UC Louvain

ACCESS the future

Frontier of Statistical Physics and Information Processing

Kyoto, Japan, July 11-14, 2013

Vetenskapsrådet

ROYAL INSTITUTE OF TECHNOLOGY

Outline

- Background and motivation
- Part 1: Synthesis
  - Lossless port-Hamiltonian systems
  - Dissipative port-Hamiltonian systems
  - Synthesis of simple components
- Part 2: Optimal control theory
  - Finite-time work extraction devices
  - Periodic work extraction devices
- Summary

ROYAL INSTITUTE OF TECHNOLOGY

Background and Motivation

- What are the **performance limits** of control devices when **resources,** such as DOFs (size), energy, temperature, and time, are **finite** and **small** (nonzero)?

Related work:

- How well — and what — can we actually implement from a small collection of simple physical building blocks? Related to *network synthesis*

[B.D.O. Anderson, S. Vongpanitlerd, 1973]
[M. Smith, 2002]

- Similar questions currently being asked in *synthetic biology*

[P. Varadarajan, D. Del Vecchio, 2009]
[J. Lavei, S. Sojoudi, R. Murray, 2010]

ROYAL INSTITUTE OF TECHNOLOGY

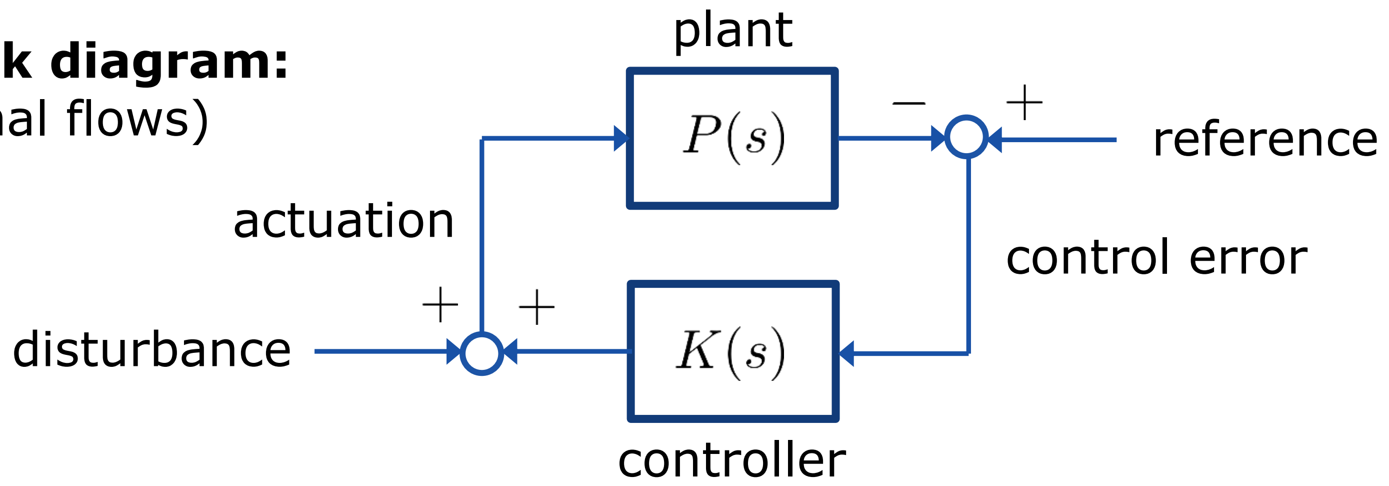
Two Problems from Control Theory

1. What is the optimal protocol for controlling a system's behavior?
   - Often feedback policies (**bonus:** robust to noise and to model uncertainty)
   - Hamilton-Jacobi-Bellman equation, dynamic programming

2. When is it possible to "physically" implement the optimal *continuous-time* controller exactly?
   - Approximate implementation is "easy" using computers (at least on the macro scale)
   - Need to carefully select model class, and how to model interconnection of systems

Feedback Loop and Transfer Functions

Block diagram:
(signal flows)



- **Transfer function** $K(s)$ is Laplace transform of impulse response of controller, i.e., $\tilde{\chi}(\omega) = K(i\omega)$ is the Fourier transform of its **linear response function**

- Typical optimal control design problem:

Given $P(s)$ with possible uncertainty, solve

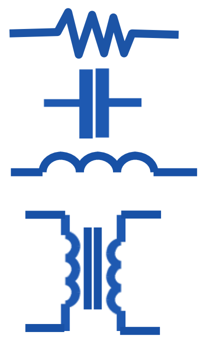
$$K^\star := \arg \min_{\text{stabilizing } K(s)} \|\text{disturbance} \mapsto \text{control error}\|_*$$

Physical Implementations of Controllers

- **Given** a transfer function $K^\star(s)$ of a controller. How do we **implement** it using a fixed set of physical components?

$$\text{input} \longrightarrow \boxed{K^\star(s)} \longrightarrow \text{output}$$

- Any rational $K^\star(s)$ can be implemented (impedance/admittance= $K^\star(s)$) using
  - Resistors
  - Capacitors
  - Inductors
  - Transformers
  - (Gyrators)



iff controller is **passive**

- *Network synthesis* in circuit theory (~1930-)
  [B.D.O. Anderson, S. Vongpanitlerd, 1973]

(Passive $\Leftrightarrow \int_{-\infty}^{t} \mathrm{input}(t') \cdot \mathrm{output}(t')\, dt' \geq 0 \Leftrightarrow$ no internal energy source)

Why Lossless Components?

- Here we strive to use *lossless* (energy-conserving / Hamiltonian) components such as:

capacitors and inductors

springs and masses

- Postulate from physics: “Everything” is lossless on the microscopic scale
- How well can “arbitrary” controllers be built from *finite number* of *lossless components*? $\Rightarrow$ Hopefully **control-relevant design trade-offs**
- What active/passive devices can we hope to build easily on a microscopic level?

ROYAL INSTITUTE OF TECHNOLOGY

Outline

- Background and motivation
- **Part 1: Synthesis**
  - **Lossless port-Hamiltonian systems**
  - Dissipative port-Hamiltonian systems
  - Synthesis of simple components
- Part 2: Optimal control theory
  - Finite-time work extraction devices
  - Periodic work extraction devices
- Summary

Lossless Port-Hamiltonian System

$$\dot{x} = J(x,v)\frac{\partial H(x,v)}{\partial x} + g(x,v)u$$

- the *symplectic form* $J(x,v)$ is skew-symmetric (i.e., $J + J^T = 0$), invertible and closed in x (i.e. satisfying $\partial_i K_{jk} + \partial_k K_{ij} + \partial_j K_{ki} = 0$ for $K = J^{-1}$);

- $g(x,v)$ derives locally from a gradient in x, i.e. can be expressed as $g(x,v) = J(x,v)\frac{\partial G(x,v)}{\partial x}$ for some scalar 'potential' $G(x,v)$;

- $u(t)$ (called the *linear input*) and $v(t)$ (called the *nonlinear input*) are time-varying parameters representing the influence of the environment on the dynamics of the system;

- the *linear output* y is defined as $y = g^T(x,v)\frac{\partial H(x,v)}{\partial x}$; and

- the *nonlinear output* z is defined as $\frac{\partial H(x,v)}{\partial v}$.

ROYAL INSTITUTE OF TECHNOLOGY

Properties

- Energy is conserved: $\dot{H} = \underbrace{uy}_{\text{linear supply}} + \underbrace{\dot{v}z}_{\text{nonlinear supply}}$

- Conditions on symplectic form ensures canonical coordinates exists (by Darboux's theorem):

$$\dot{q} = \partial H/\partial p, \quad \dot{p} = -\partial H/\partial q$$

- Under linear lossless interconnection (Kirchoff's laws)

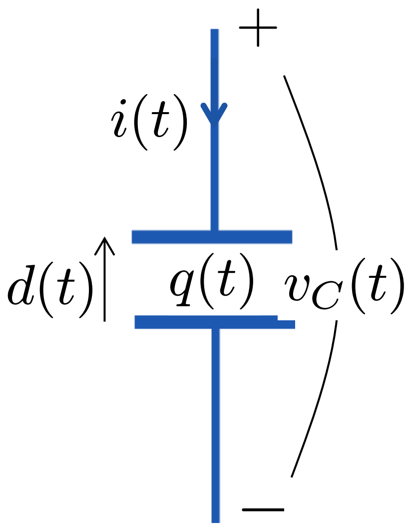
$$u_1 + y_2 = 0, \quad u_2 = y_1$$

the symplectic form of interconnected system is *closed:*

$$\begin{pmatrix} J_1 & -g_1 g_2^T \\ g_2 g_1^T & J_2 \end{pmatrix}$$

Example: Time-Varying Capacitor

$$\dot{x} = J\frac{\partial H}{\partial x} + gu \Leftrightarrow \frac{d}{dt}\underbrace{(C(t)v_C(t))}_{x_2=q} = \underbrace{i(t)}_{u} \Leftrightarrow \dot{v}_C(t) = -\frac{\dot{C}(t)}{C(t)}v_C(t) + \frac{i(t)}{C(t)}$$



- Capacitance $\quad C(d) = \epsilon A/d$
- Stored electrical energy $\quad H(v_C, d) = \frac{1}{2}C(d)v_C^2$
- Force between plates $\quad F_e = \partial H/\partial d = q^2/2\epsilon A$
- Energy balance

$$\underbrace{\dot{H}(t)}_{\text{Change of stored energy}} = \underbrace{F_e(t)\dot{d}(t)}_{\text{nonlinear supply}} + \underbrace{v_C(t)i(t)}_{\text{linear supply}}$$

[R. Brockett, J.C. Willems, Proc. IEEE CDC, 1978]

Linear vs. Nonlinear Inputs/Outputs

- Nonlinear i/o are more independent of precise physical implementation of interaction with environment

- **Example:** Instead of

$$v(t) = d(t), \quad z(t) = \frac{q(t)^2}{2\epsilon A}(= F_e)$$

choose

$$v(t) = C(t), \quad z(t) = -\frac{q(t)^2}{2C(t)^2}$$

- Nonlinear i/o can be made linear i/o with special implementation
- Linear i/o can be made nonlinear i/o with special choice of Hamiltonian

Dissipative Port-Hamiltonian System

$$\dot{x} = \left(J(x,v) - \sum_i R_i(x,v) \right) \frac{\partial H(x,v)}{\partial x} + g(x,v)u + \sum_i \sqrt{2R_i(x,v)T_i}\, n_i(t)$$

- v is a nonlinear input vector, u is a linear input vector;
- J and g satisfy the conditions of a lossless port-Hamiltonian system;
- $R_i(x,v)$ is a symmetric nonnegative definite matrix, a square root of which is denoted $\sqrt{R_i}$; and
- T_i is the temperature of R_i, and n_i are independent unit intensity Gaussian white noises.

Models interconnection to *macroscopic* heat baths R_i
(Interpret equation in Itō sense, and $T = k_B T_K$)

ROYAL INSTITUTE OF TECHNOLOGY

Properties: Laws of Thermodynamics

- Internal energy $U = \mathbf{E}_x H(x, v)$
- Work rate $w = -\mathbf{E}_x u^T y - \mathbf{E}_x \dot{v}^T \frac{\partial H}{\partial v}$
- Heat rate $q = -\mathbf{E}_x \frac{\partial H}{\partial x}^T R \frac{\partial H}{\partial x} + T \mathbf{E}_x \mathrm{Tr}\; R \frac{\partial^2 H}{\partial x^2}$
- Entropy $S = \mathbf{E}_x(-\ln \rho)$ (Shannon differential entropy)

Dissipative port-Hamiltonian systems satisfy

- **First law:** $\dot{U} = q - w$
- **Second law:** $\dot{S} \geq \sum_i \frac{q_i}{T_i}$ (q_i heat rate exchange with heat bath of temperature T_i)

ROYAL INSTITUTE OF TECHNOLOGY

Outline

- Background and motivation

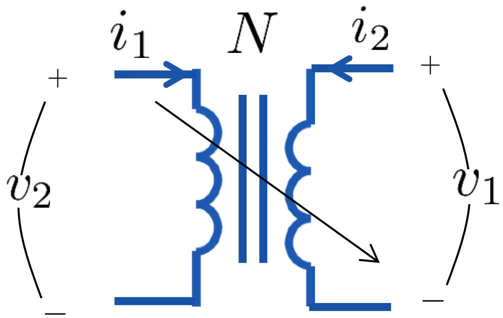
- **Part 1: Synthesis**
  - Lossless port-Hamiltonian systems
  - Dissipative port-Hamiltonian systems
  - **Synthesis of simple components**

- Part 2: Optimal control theory
  - Finite-time work extraction devices
  - Periodic work extraction devices

- Summary

Time-Varying Ideal Lossless Transformer

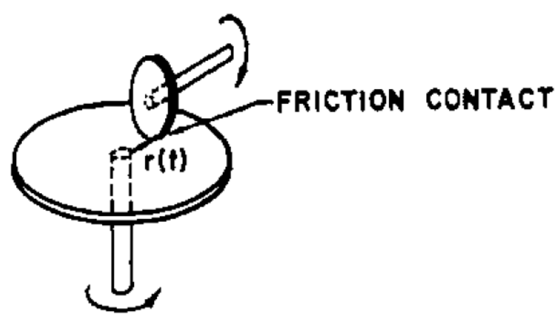
- Time-varying transformer ratio $N(t)$



$$v_2(t) = N(t)v_1(t)$$

$$i_2(t) = \frac{-1}{N(t)} i_1(t)$$

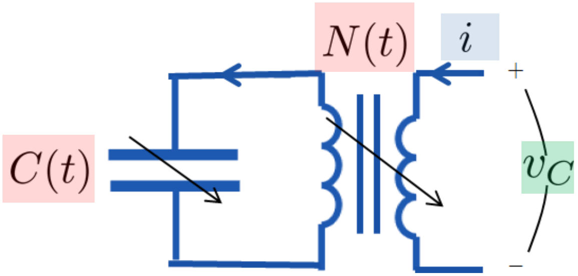
- Mechanical analogue:



Variable gear ratio transformer.

[B.D.O. Anderson *et al.*, "*The time-varying transformer*", Proc. IEEE, 1965]

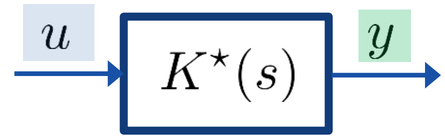
Synthesis Using Time-Varying Lossless Circuit



$$\dot{\tilde{x}} = -\frac{\dot{C}(t)}{2C(t)}\tilde{x} + \frac{N(t)}{\sqrt{C(t)}}i$$

$$v_C = \frac{N(t)}{\sqrt{C(t)}}\tilde{x},$$

$$\overset{?}{\Leftrightarrow}$$



(a, b, c) constant

$\Rightarrow K^\star(s) = \frac{bc}{s-a}$

Theorem: The linear system

$$\dot{p} = a(t)p + b(t)u, \quad y = c(t)p, \quad p(0) = 0,$$

where $a \in \mathcal{C}^0$, $b, c \in \mathcal{C}^1$, can be exactly implemented using the lossless circuit with $\tilde{x}(0) = 0$ if and only if $b(t)c(t) > 0$ for all t. Choose

linear input $\quad i(t) = u(t), \quad v_C(t) = y(t), \quad$ linear output

nonlinear inputs

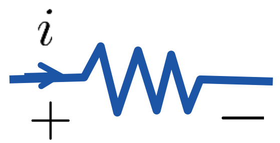
$$C(t) = \exp\left[-2\int_0^t a(s)ds\right]\frac{b(t)c(0)}{c(t)b(0)}C(0), \quad N(t) = \sqrt{b(t)c(t)C(t)},$$

with any $C(0) > 0$.

Exercise: To Synthesize a Resistor

- Usually a resistor is synthesized using many DOFs (a macro system)

- With many DOFs and thermal uncertainty in initial conditions comes **undesired fluctuation** (Nyquist noise, FD-theorem):



$$v_C(t) = Ri(t)$$

$\approx$

$$v_C(t) = Ri(t) + \sqrt{2TR}n(t)$$

(n white noise)

- **Can we do better??**

Synthesizing a Resistor without Noise

- **Goal:** To build a resistor without Nyquist noise

- If capacitance C is small, it holds

$$C\dot{v} = -\frac{1}{R}v + i$$

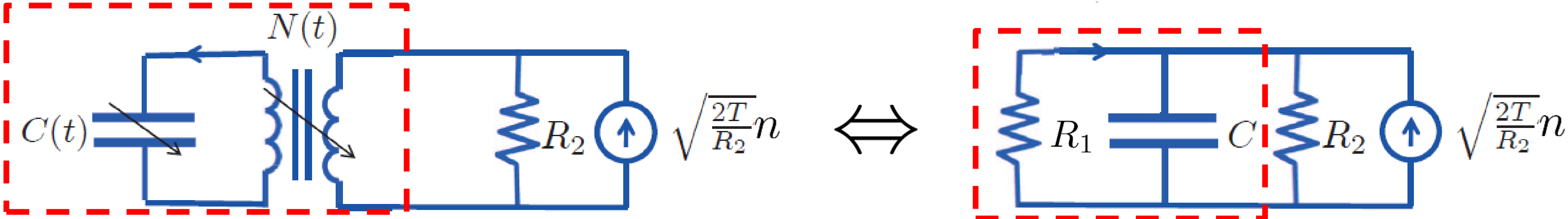
- $G(s) = R/(s\tau + 1) \approx R$ for frequencies up to bandwidth $1/\tau = 1/(RC)$

- Choose time-varying lossless circuit with

$$a = -1/(RC) = -1/\tau, \quad b = 1/C, \quad c = 1 \quad \Longrightarrow$$

$$C(t) = e^{2t/\tau}C(0) \quad (\Leftrightarrow d(t) = e^{-2t/\tau}x(0)), \quad N(t) = \sqrt{C(t)/C}$$

A Lossless Cooling Device



$$C(t) = e^{2t/\tau_1} C(0), \quad N(t) = \sqrt{C(t)/C}$$

$$\tau_1 = R_1 C$$

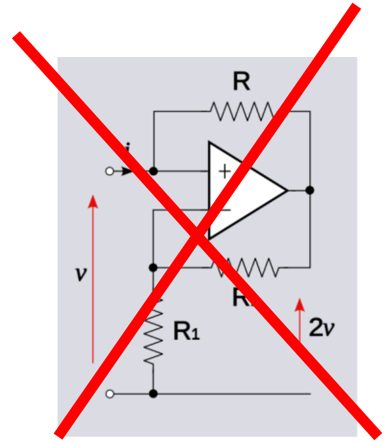
$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{R_2 C}$$

- Internal energy $U(t) = \frac{T'}{2} + \left(U(0) - \frac{T'}{2}\right) e^{-2t/\tau} \quad \left(U = \frac{1}{2} C \mathbf{E}\, v_C^2\right)$
- Effective temperature $T' = \dfrac{T}{1 + R_2/R_1}$
- Mechanical work extracted ($\Leftrightarrow$ energy dissipated in R_1)

$$W = \int_0^{t_f} w\, dt = -\int_0^{t_f} F_e \dot{d}\, dt = \frac{t_f}{\tau_1} T' - (U(t_f) - U(0))\,(1 + R_1/R_2)$$

$$\sim \frac{\text{time} \times \text{effective temp.}}{\text{compresssion time const.}}$$

Exercise: Synthesize a *Negative* Resistor

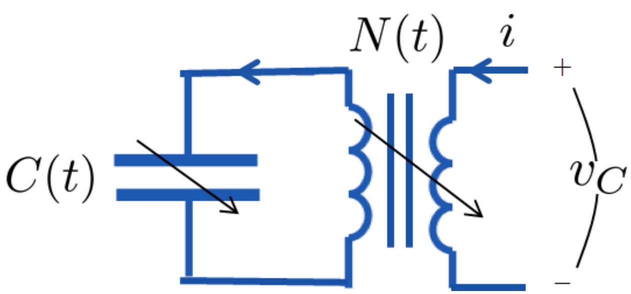


- **Goal:** To build simple **active** device

  $v = -Ri, \quad R > 0$ *without using op-amps*

- **Idea:** Simply time-reverse the plate trajectory!

$$d(t) = e^{2t/\tau} d(0)$$



$$\dot{\tilde{x}} = \frac{1}{\tau}\tilde{x} + \sqrt{\frac{R}{\tau}}\, i$$

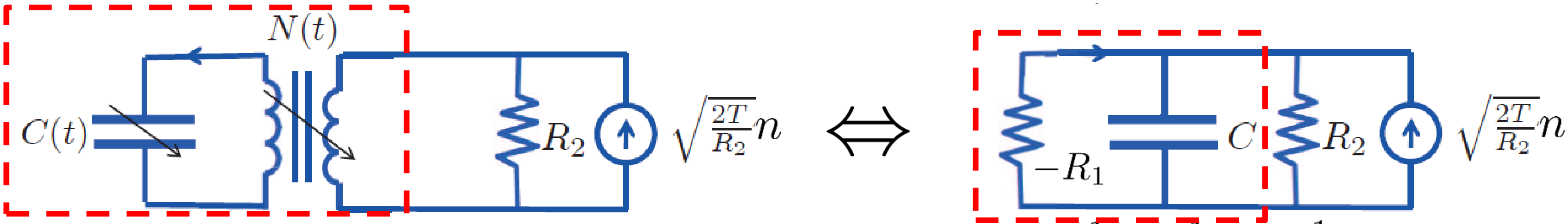
$$v = \sqrt{\frac{R}{\tau}}\,\tilde{x},$$

$$H = \frac{1}{2}\tilde{x}^2$$

- Circuit is **unstable**! Externally supplied mechanical energy continuously converted into stored electrical energy

KTH VETENSKAP OCH KONST
ROYAL INSTITUTE OF TECHNOLOGY

A Lossless Heating Device



$C(t) = e^{-2t/\tau_1} C(0), \quad N(t) = \sqrt{C(t)/C}$

$\tau_1 = R_1 C$

$$\frac{1}{\tau} = \frac{1}{-\tau_1} + \frac{1}{R_2 C}$$

- Internal energy $\quad U(t) = \frac{T'}{2} + \left(U(0) - \frac{T'}{2}\right) e^{-2t/\tau}$
- Effective temperature $\quad T' = \frac{T}{1 - R_2/R_1}$
- Stability condition $\quad R_1 > R_2$
- Mechanical work injected ($\Leftrightarrow$ energy extracted from $-R_1$)

$$W = \int_0^{t_f} w\, dt = -\frac{t_f}{\tau_1} T' - (U(t_f) - U(0))\,(1 + R_1/R_2)$$

ROYAL INSTITUTE OF TECHNOLOGY

Outline

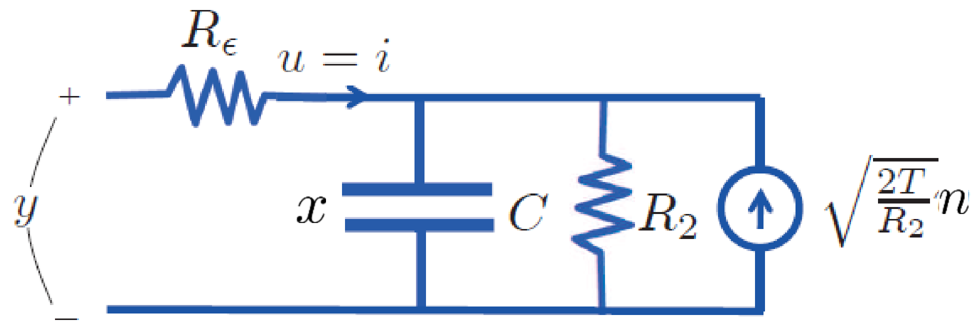
- Background and motivation

- Part 1: Synthesis
  - Lossless port-Hamiltonian systems
  - Dissipative port-Hamiltonian systems
  - Synthesis of simple components

- **Part 2: Optimal control theory**
  - **Finite-time work extraction devices**
  - Periodic work extraction devices

- Summary

A More Interesting Control Problem



- Assume we can control the injected current $u = i$ and measure the voltage y continuously in time

- What is the maximum power $w = -\mathbf{E}_x uy$ we can extract?

  ("optimal" Maxwell's demon in continuous time?)

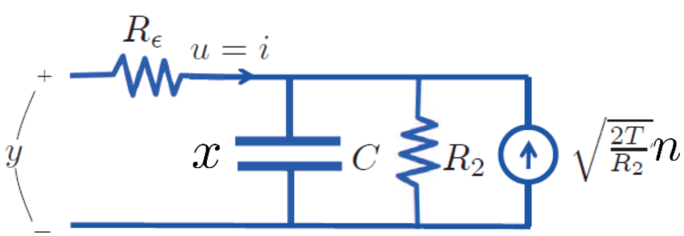
- Back-of-the-envelope calculation:
  - Nyquist: frequency band Δf carries power $T\Delta f$
  - Bandwidth $\Delta f \sim 1/\tau_2 = 1/(R_2 C)$
  - Total available power $\sim T/\tau_2$ (?)
- Is this the optimum, and can it be realized?

Apply Optimal Control Theory

- Solve the stochastic optimal control problem

$$W^\star := \max_u \int_0^{t_f} w(t)\,dt = -\min_u \mathbf{E}_x \int_0^{t_f} y(t)u(t)\,dt$$

subject to RC-circuit dynamics



$$\begin{cases} \dot{x} = -\dfrac{1}{\tau_2}x + \dfrac{1}{\sqrt{C}}u + \sqrt{\dfrac{2T}{\tau_2}}n, & \mathbf{E}_x x(0) = 0 \\ y = \dfrac{1}{\sqrt{C}}x + R_\epsilon u, & \mathbf{E}_x x(0)^2 = T \end{cases}$$

- **A standard linear-quadratic optimal control problem**

Solution

- Optimal solution is the time-varying feedback control

$$\boxed{u(t) = -\frac{S(t)+1/2}{R_\epsilon\sqrt{C}}x(t) =: -\frac{y(t)}{R_1(t)}}$$

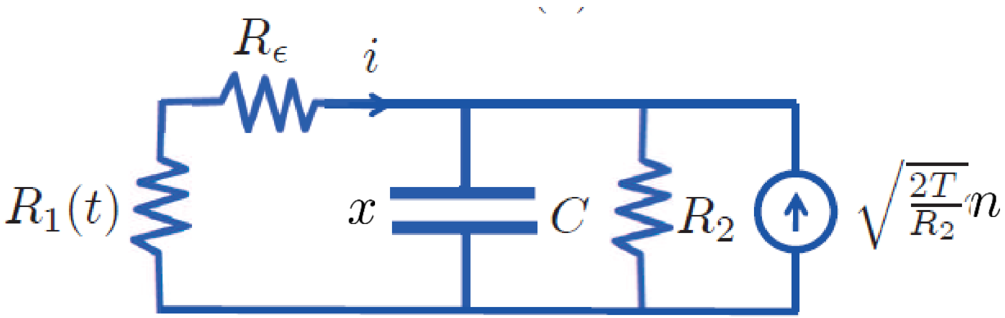
where S solves the Riccati equation

$$\dot{S} = \frac{2}{\tau_2}S + \frac{1}{R_\epsilon C}\left(S+\frac{1}{2}\right)^2, \quad S(t_f) = 0$$

- Optimal amount of extracted work

$$W^\star = \left(\frac{1}{2}+\frac{t_f}{\tau_2}\right)T - \sqrt{\kappa}\left(1+\frac{2t_f}{\tau_2}\right)T + O(\kappa), \quad \kappa = \frac{R_\epsilon}{R_2} \to 0$$

- Equivalent physical realization:

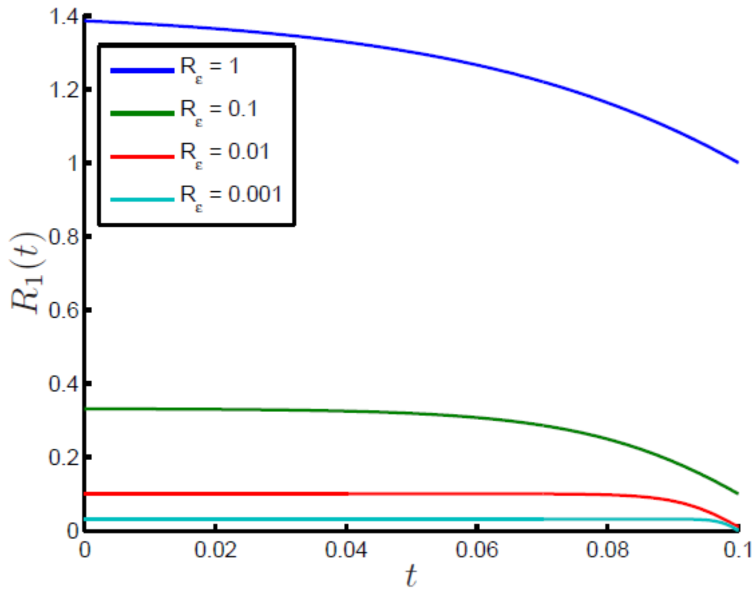


ROYAL INSTITUTE OF TECHNOLOGY

Interpretation

- Feedback converges exponentially fast to the constant

$$R_1 = \sqrt{R_2 R_\epsilon + R_\epsilon^2}$$



- As $R_\epsilon \to 0$ we indeed come arbitrarily close to

$$W^\star = \left(\frac{1}{2} + \frac{t_f}{\tau_2}\right) T \quad \Rightarrow \quad w_{\max} = \frac{T}{\tau_2}$$

- **The synthesized positive resistor from Part 1 is an optimal work extraction device (with proper choice of R_1)**

- How does this relate to Carnot's theorem?

ROYAL INSTITUTE OF TECHNOLOGY

Outline

- Background and motivation

- Part 1: Synthesis
  - Lossless port-Hamiltonian systems
  - Dissipative port-Hamiltonian systems
  - Synthesis of simple components

- **Part 2: Optimal control theory**
  - Finite-time work extraction
  - **Periodic work extraction**

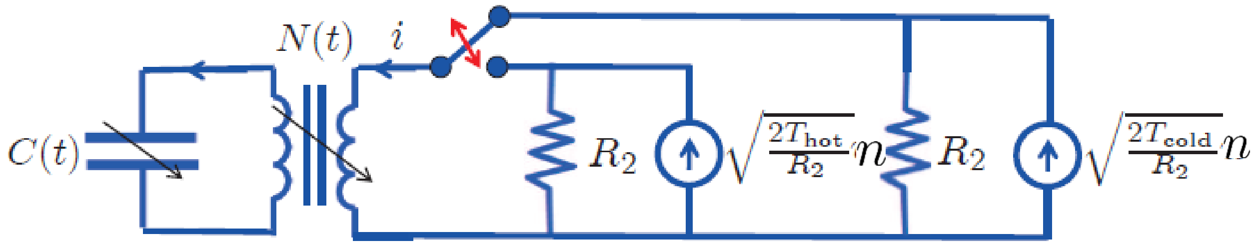
- Summary

Periodic Work Extraction

- Compression of capacitor (work extraction) cannot go on forever (model breakdown etc.)

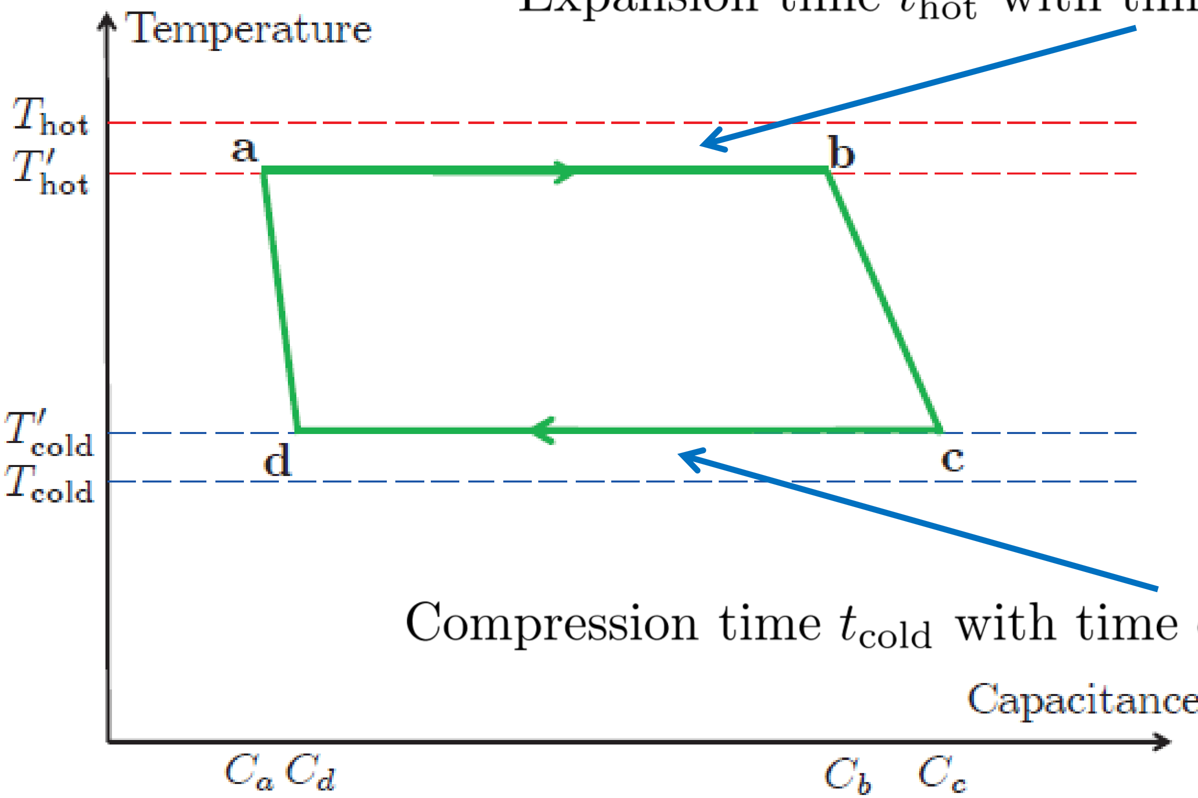
- How do we reset it to enable periodic operation and its use as heat engine?

- **Idea:** Operate it in a **finite-time** Carnot cycle:
  1. Extract work/compress cap. at temp. $T'_{\mathrm{hot}} < T_{\mathrm{hot}}$
  2. Lower temperature to T'_{cold} adiabatically
  3. Expand capacitor at temperature $T'_{\mathrm{cold}} > T_{\mathrm{cold}}$
  4. Increase temp. to T'_{hot} adiabatically, and repeat



Finite-Time Carnot Cycle

Expansion time $t_{\rm hot}$ with time constant $\tau_1^{\rm hot} = R_1^{\rm hot} C$



$$T'_{\rm hot} = \frac{T_{\rm hot}}{1 + R_2/R_1^{\rm hot}}$$

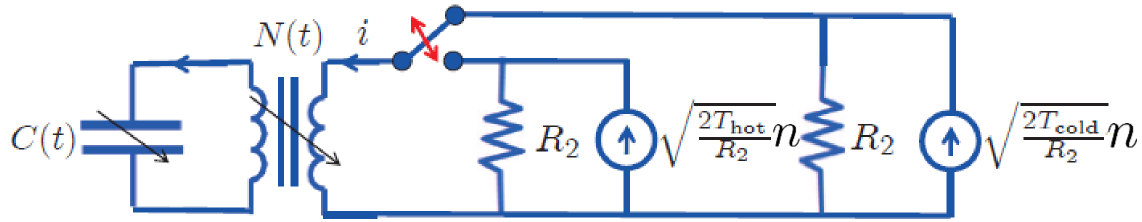
$$T'_{\rm cold} = \frac{T_{\rm cold}}{1 - R_2/R_1^{\rm cold}}$$

Compression time $t_{\rm cold}$ with time constant $\tau_1^{\rm cold} = R_1^{\rm cold} C$

Work extracted in isotherms:

$$W_{\rm hot} = Q_{\rm hot} = T'_{\rm hot} \frac{t_{\rm hot}}{\tau_1^{\rm hot}}, \quad W_{\rm cold} = Q_{\rm cold} = -T'_{\rm cold} \frac{t_{\rm cold}}{\tau_1^{\rm cold}}$$

Efficiency



$$\tau_1^{\text{hot}} = R_1^{\text{hot}} C, \quad \tau_1^{\text{cold}} = R_1^{\text{cold}} C$$

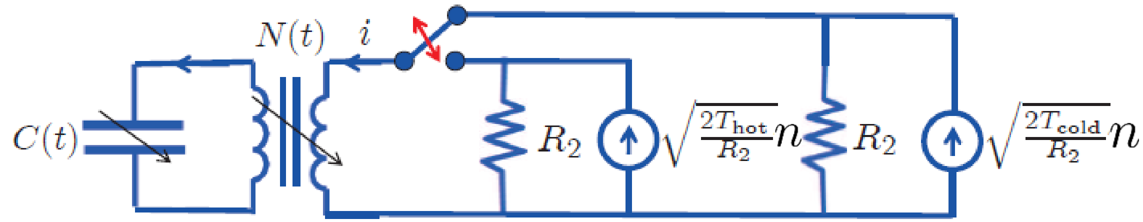
- Efficiency

$$\eta = \frac{W_{\text{hot}} + W_{\text{ad}} + W_{\text{cold}} - W_{\text{ad}}}{Q_{\text{hot}}} = \frac{Q_{\text{hot}} + Q_{\text{cold}}}{Q_{\text{hot}}}$$

$$= 1 - \frac{T'_{\text{cold}}}{T'_{\text{hot}}} = 1 - \frac{T_{\text{cold}}}{T_{\text{hot}}} \frac{1 + R_2/R_1^{\text{hot}}}{1 - R_2/R_1^{\text{cold}}}$$

- Similar to Carnot heat engine efficiency, but…
  - Heat bath temp. replaced by effective temp.
  - Efficiency independent on time period
  - **Efficiency arbitrarily close to theoretical max** by choosing $R_1^{\text{hot}}, R_2^{\text{cold}} \gg R_2$

Power



$$\tau_1^{\rm hot} = R_1^{\rm hot} C, \quad \tau_1^{\rm cold} = R_1^{\rm cold} C$$

- Power

$$\bar{w} := \frac{W_{\rm hot} + W_{\rm ad} + W_{\rm cold} - W_{\rm ad}}{t_{\rm hot} + t_{\rm cold}}$$

$$= \frac{1}{C(R_1^{\rm hot} + R_1^{\rm cold})}\left(\frac{T_{\rm hot}}{1 + R_2/R_1^{\rm hot}} - \frac{T_{\rm cold}}{1 - R_2/R_1^{\rm cold}}\right)$$

- **Power can be made arbitrarily large** by choosing a small C. Assuming a lower bound on $\tau_2 = R_2 C$ there is **trade-off in efficiency vs. power**
  - Optimal resistors $\frac{R_1^{\rm cold}}{R_2} = \frac{R_1^{\rm hot}}{R_2} = \frac{\sqrt{T_{\rm hot}} + \sqrt{T_{\rm cold}}}{\sqrt{T_{\rm hot}} - \sqrt{T_{\rm cold}}}$
  - Max power $(\sqrt{T_{\rm hot}} - \sqrt{T_{\rm cold}})^2/4\tau_2$
  - Recovers the Chambadal-Novikov-Curzon-Ahlborn efficiency $1 - \sqrt{T_{\rm cold}/T_{\rm hot}}$

ROYAL INSTITUTE OF TECHNOLOGY

Summary

- Port-Hamiltonian systems a good class of models
  - Satisfy laws of thermodynamics and Hamiltonian mechanics
  - Well-behaved under interconnection

- Simple electro-mechanical device (time-varying cap.) introduced and used to synthesize simple active and passive systems
  - Iff characterization of possible implementations given

- An optimal control problem used to synthesize a (sub)-optimal finite-time heat engine
  - Arbitrarily close to Carnot efficiency
  - Recovers standard finite-time thermodynamics

References

[1] J.-C. Delvenne and H. Sandberg, *Finite-time thermodynamics of port-Hamiltonian systems*. Physica D, 2013. To appear, available on ScienceDirect. (Link)

[2] H. Sandberg, J.-C. Delvenne, and J. C. Doyle, *On Lossless Approximations, the Fluctuation-Dissipation Theorem, and Limitations of Measurements*. IEEE Transactions on Automatic Control, 56:2, pp. 293–308, February, 2011. (Link)

[3] H. Sandberg and J.-C. Delvenne, *The Observer Effect in Estimation with Physical Communication Constraints*. In Preprints of the 18th IFAC World Congress, Milano, Italy, August/September 2011. (Link)

Contact: hsan@kth.se

Web: http://www.ee.kth.se/~hsan/

Recommended reading (by Hugo Touchette):

J. Bechhoefer, *Feedback for physicists: A tutorial essay on control*, Rev. Mod. Phys. 77, 783–836, 2005