Control and Statistical Mechanics

Henrik Sandberg, Jean-Charles Delvenne, and John Doyle

15 August 2006

Topics:
• Statistical mechanics formalism using linear system theory
• How come real systems look dissipative when energy is conserved? Does this give limitations?
• Back action of measurements (without quantum effects)
• How do we convert heat into work using control theory?
Motivation and Background

• **Hard limits**: With ideal devices causality alone gives limitations (John). But what other limitations are there?

• **Uncertainty**: Relation information theory, control theory, statistical mechanics? Channel capacity, Bode-Shannon, Carnot cycles,… Entropy is everywhere!

• **Energy conservation**: Measurements are not unproblematic, even using classical physics

• How much of this can we understand with simple linear systems and some approximation theory?
Outline

Focus here: Control theory and statistical mechanics. Connection to information theory not clear yet.

1. Energy and lossless/dissipative models (physical models)
2. Lossless approximations of dissipative models, emergence of noise and limitations
3. Back action of measurements
4. Examples: Conversion of heat into useful work (compare with Carnot heat engine)
Physical Linear System: Lossless and Strictly Causal

We assume linear physical models have the structure:

\[ \dot{x}(t) = Jx(t) + Bu(t), \quad x(t) \in \mathbb{R}^n, \]
\[ y(t) = B^T x(t) \]

- **Internal energy**: \( U(x(t)) = \frac{1}{2} x(t)^T x(t) \)
- **Energy preserving**: \( J = -J^T \) (anti-symmetric)
- **Work rate**: \( w(t) = y(t)^T u(t) \)
- **Lossless**: 
  \[ \frac{dU(x(t))}{dt} = x(t)^T \dot{x}(t) = x(t)^T (Jx(t) + Bu(t)) = y(t)^T u(t) \]
- **Strictly causal and \((J,B)\) controllable**
Example: LC Circuit

Choice of inputs/outputs not always simple. Assume choice is already made.

\[
\begin{align*}
\dot{x} &= \begin{pmatrix}
0 & -1/\sqrt{C_1 L_1} \\
1/\sqrt{C_1 L_1} & 0 \\
0 & 1/\sqrt{L_1 C_2} \\
-1/\sqrt{L_1 C_2} & 0
\end{pmatrix} x + \begin{pmatrix}
1/\sqrt{C_1} \\
0 \\
0 \\
0
\end{pmatrix} u \\
y &= \begin{pmatrix}
1/\sqrt{C_1} & 0 & 0
\end{pmatrix} x, \quad x^T = \begin{pmatrix}
\sqrt{C_1} v_1 \\
\sqrt{L_1} i_1 \\
\sqrt{C_2} v_2
\end{pmatrix} \\
U &= \frac{1}{2} x^T x = \frac{1}{2} (C_1 v_1^2 + L_1 i_1^2 + C_2 v_2^2), \quad w = yu = v_1 i.
\end{align*}
\]
Macroscopic Dissipative Systems
(Example: Resistor)

• Consider a linear static input-output device
  \[ y(t) = ku(t) \]

• If \( k > 0 \), the device is dissipative
  \[ w(t) = y(t)u(t) = ku(t)^2 \geq 0 \]

• Can also be modeled by the convolution
  \[ y(t) = \int_{0}^{\infty} k\delta(t-s)u(s)ds \]

  for smooth \( u(t) \)

• Choose a time interval \([0, \tau]\) of interest
Model with Recurrence Time $\tau$

\[
g(t) = \sum_{l=-\infty}^{\infty} k \delta(t - l\tau)
\]

\[y(t) = ku(t), \ t \in [0, \tau]\]
Fourier Expansion and Physical Realization

- Expansion of impulse response (in distribution sense)

\[ g(t) \sim \frac{k}{2\tau} + \sum_{l=1}^{\infty} \frac{k}{\tau} \cos l\omega_0 t \overset{\Delta}{=} g_+(t) + g_-(t), \quad \omega_0 = \frac{2\pi}{\tau} \]

- \( g_+(t) \) causal and \( g_-(t) \) anti-causal

- Truncate and keep \( N \) terms and realize \( 2g_+(t) \) with lossless and strictly causal linear system \( G_N \):

\[
J = \begin{bmatrix}
0 & \Omega & 0 \\
-\Omega^T & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}, \quad \Omega = \text{diag}\{\omega_0, 2\omega_0, \ldots, N\omega_0\}
\]

\[
C = 2\sqrt{\frac{k}{\tau}} \begin{pmatrix}
1 & \ldots & 1 & 0 & \ldots & 0 & \frac{1}{\sqrt{2}}
\end{pmatrix}
\]

\[
B = C^T.
\]
Simulations of \( y(t) = ku(t) \) and \( y_N(t) = G_N u(t) \) (\( \tau = 1 \))

\[
g_N(t)
\]

\[
u(t) = 1 - \cos 40\pi t
\]

- Initial conditions zero
- \[ |y(t) - y_N(t)| \leq \frac{\text{const.}}{N} \left( |\dot{u}(t)| + \|\ddot{u}\|_{L_1[0,t]} \right), \quad t \in [0, \tau] \]
Input-Output Relation of Lossless Approximation

\[ y_N(t) = B^T e^{Jt} x_N(0) + \int_0^t B^T e^{J(t-s)} B u(s) ds \]

- Assume \( u(t) = 0 \), then the covariance function of \( y_N(t) \) is

\[ R_{y_N}(t) = \mathbf{E} y_N(t) y_N(0)^T = B^T e^{Jt} (\mathbf{E} x_N(0) x_N(0)^T) B \]

- Stat. mech.: Assume thermal equilibrium, temperature \( T \)

\[ \mathbf{E} x_N(0) x_N(0)^T = T \cdot I \]

- As \( N \) increases, we have (with \( v(t) \) white noise)

\[ R_{y_N}(t) \to 2T k \delta(t), \quad t \in [0, \tau] \]

\[ y_\infty(t) = k u(t) + \sqrt{2T} k v(t) \]
Implications and Discussion: Dissipation-Fluctuation

• We wanted to realize the dissipative model

\[ y(t) = k u(t) \]

with lossless linear physical model

• Assumptions: \( u(t) \) is smooth and finite-time interval \([0, \tau]\)

• We needed “many” states \( N \), and uncertainty in the initial state gave stochastic white noise:

\[ y_\infty(t) = k u(t) + \sqrt{2T} k v(t) \]

• Generalization of Johnson-Nyquist noise in resistors

• \( N \sim \) number of atoms in a resistor. The recurrence time \( \tau \) is very large
Interconnections of Lossless Systems

• Preserve the total energy
• Use feedback $u_1(t) = -y_2(t)$ and $u_2(t) = y_1(t)$
• Dynamics

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} J_1 & -B_1B_2^T \\ B_2B_1^T & J_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

• Overall system still lossless
Interconnection to Heat Bath

- Assume the lossless system $S_2$ is “large” as before and can be modeled by $u_1(t) = ky_1(t) + \sqrt{2Tk}v(t)$
- Call $S_2$ a heat bath of temperature $T$
- Dynamics $\dot{x}_1 = (J_1 - kB_1B_1^T)x_1 - B_1\sqrt{2Tk}v$
- $S_1$ looks dissipative while connected to heat bath $S_2$
- Stochastic model because of unknown initial state in $S_2$
Heat Dynamics, Equilibrium, and Energy Equipartition

Mean evolution in $S_1$:
\[
\frac{d}{dt} \mathbf{E}x_1 = (J_1 - kB_1B_1^T)\mathbf{E}x_1
\]

Variance evolution $S_1$:
\[
\dot{X}_1 = (J_1 - kB_1B_1^T)X_1 + X_1(J_1 - kB_1B_1^T)^T + \underbrace{2TkB_1B_1^T}_{"heat flow out" from $S_1"} + \underbrace{2TkB_1B_1^T}_{"heat flow in"
\]

Expected internal energy $S_1$:
\[
\mathbf{E}U_1 = \frac{1}{2}(\mathbf{E}x_1)^T \mathbf{E}x_1 + \frac{1}{2} \text{Tr}X_1
\]

If $(J_1, B_1)$ controllable:
\[
\begin{align*}
X_1(t) &\to T \cdot I, \\
\mathbf{E}x_1(t) &\to 0, \\
\mathbf{E}U_1(t) &\to \frac{1}{2}nT, \quad t \to \infty
\end{align*}
\]
Linear Measurements

- Measure the state $x_1$ (scalar) of lossless system $S_1$
- Choose a linear measurement device: $y(t) = kx_1(t)$
- A lossless implementation of $S_2$ gives

$$y_1(t) = kx_1(t) + \sqrt{2T_m}kv(t)$$

if measurement device has temperature $T_m$

- $S_1$ connected to measurement device $S_2$:

\[
\begin{align*}
\dot{x}_1(t) &= (J_1 - kB_1B_1^T)x_1(t) - B_1\sqrt{2kT_m}v(t) \\
\hat{x}_1(t) &= x_1(t) + \sqrt{\frac{2T_m}{k}}v(t)
\end{align*}
\]

Back action: $E(\text{process noise x measurement noise}) = 2T_m$
Control Loops of Increasing Complexity and Realism

**Ideal devices**: Causality alone gives Bode integral limitations

```plaintext
Plant
       Actuator
            -1
              |
            +
            |
 Sensor
```

**Lossless devices**: Energy conservation gives back action from sensor and actuator and additional limitations!

```plaintext
Plant
       Actuator
            -1
              |
            +
            |
 Sensor
```
Example: Conversion of Heat into Work I

- Classic thermodynamics problem. Here a controls version using increasingly realistic models with back action.
- Assume $S_1$ is lossless and $E_{x_1}(0)=0$
- Extracted work: $W = - \int y_1(t)u_1(t)dt$
- With no extra information about the initial state of $S_1$:
  \[
  E_{y_1}(t) = B_1^T e^{J_1 t} E_{x_1}(0) + \int_0^t B_1^T e^{J_1(t-s)} B_1 u_1(s) ds
  \]
- Then $E_W = 0$, since on average we only get out what we put in.
- With information about the initial state, we can do better!
Heat into Work I

Idea (assuming we have correct model of $S_1$):
1. First heat (lossless) capacitor ($S_1$) to temperature $T_1$
2. Disconnect it, and somehow estimate its state: $\bar{x}_1(0)$
3. Use estimate to pull out energy to ideal current source, using open loop signal $u_1^0(t)$

How efficiently can we do this?
Heating of Capacitor – Step 1

- The charge in the capacitor fluctuates during heating
- The variance $E x_1(t)^2$ (temperature) increases with time constant $1/RC$ towards equilibrium $T_1$
- $Ex_1(t)=0$, but it is unlikely that the charge is close to zero after long time of heating
Heat into Work I: Ideal Sensor and Actuator

• Assume we have an optimal linear estimate $\bar{x}_1(0)$ of $x_1(0)$:

$$x_1(0) = \bar{x}_1(0) + \delta x_1(0), \quad \mathbb{E}x_1(0) = \mathbb{E}\bar{x}_1(0),$$

$$\mathbb{E}\bar{x}_1(0)^T\delta x_1(0) = 0,$$

$$\mathbb{E}x_1(0)^T x_1(0) = \mathbb{E}Q = nT_1,$$

"available energy"

$$\mathbb{E}\delta x_1(0)^T \delta x_1(0) = nT_2, \quad T_2 \leq T_1$$

"remaining uncertainty"

• Based on $\bar{x}_1(0)$, we can compute an input $u_1^0(t)$ so that $\bar{x}_1(t) \to 0$ in the model

$$\dot{\bar{x}}_1(t) = J_1 \bar{x}_1(t) + B_1 u_1(t)$$

$$\bar{y}_1(t) = B_1^T \bar{x}_1(t)$$

• Expected extracted work to input heat ratio (compare Carnot):

$$\frac{\mathbb{E}W}{\mathbb{E}Q} = -\mathbb{E} \int \frac{y_1(t) u_1^0(t)}{nT_1} dt = \frac{T_1 - T_2}{T_1} = 1 - \frac{T_2}{T_1}$$
Heat to Work I, Sensor with Back Action (Ideal Actuator)

• **More realism**: Use a linear measurement device of temperature $T_m$ and a Kalman filter to obtain $\bar{x}_1(0)$

Tradeoffs:

• **Measure long time**: Dissipation in measurement device
• **Measure short time**: Bad estimate
Example: Expected Work from Heated Capacitor Using Kalman Filter ($T_m=1$, $T_1(0)=5$, $R=3$, $C=1$)

- $T_1(t)$ – Temperature of capacitor
- $T_2(t)$ – Uncertainty (or temperature) of estimate
- “Efficiency” – $\frac{\text{expected work}}{\text{initial heat}} = \frac{T_1(t) - T_2(t)}{T_1(0)}$

- There is an optimal amount of time to measure before pulling out work ($t_{opt} \approx 0.4$)!
Heat to Work I: Summary

• How to extract work from heat?

• First idea:
  • Linear time-invariant system;
  • Apply input in open loop;
  • Work extracted is \[ - \int u(t)^T y(t) dt \]

• But: cannot convert heat into work

• Because: equation of mean and equation of variance are decoupled
Heat to Work I: Summary

• Second idea
  • Linear time-invariant system
  • Observe the output
  • Estimate the state
  • Apply an input

• Apparently: all heat into work
• Maxwell’s demon
• But: Observation, estimation, actuation must be realized by (linear) physical systems
• Must be dissipative

• Kalman filter: linear time-varying
Heat to Work II: Time-Varying Capacitor

• Brockett-Willems 1978
• How to modify capacity and connection to maximize work?
• Solution:
  • Capacitor connected to $T_1$ (high), capacitance increased slowly, energy $\frac{1}{2} T_1$
  • Capacitor disconnected, capacitance increased until energy $\frac{1}{2} T_2$
  • Capacitor connected to $T_2$ (low), capacitance decreased slowly, energy $\frac{1}{2} T_2$
  • Capacitor disconnected, capacitance decreased until energy $\frac{1}{2} T_1$
• Carnot cycle, efficiency $1 - \frac{T_2}{T_1}$
General Formalism: Linear Time-Varying Systems

• Dissipative linear time-varying system:

\[ \dot{x} = (A(t) - \sum_i M_i(t))x + \sum_i \sqrt{2T_i} F_i(t)v_i(t) \]

• Energy \[ U = \frac{1}{2} x^T K(t) x \]

• We can choose \[ A(t), \ F_i(t), \ K(t) \]

• Without loss of generality, \[ A^T K + K A = 0 \]

• How to extract work in an optimal way? (to be defined)

• Generalizes time-varying capacitor

• Generalizes Plant + Time-varying controller

• Nonlinear open-loop control in disguise
Work and Heat

• Variance: \( X = \mathbb{E}xx^T \)

• Average energy: \( U = \mathbb{E} \frac{1}{2} x^T K x = \frac{1}{2} \text{Tr} K X \)

• Energy equation:

\[
\dot{U} = \frac{1}{2} \text{Tr} \dot{K} X + \frac{1}{2} \text{Tr} K \dot{X}
\]

• Work extracted is \( -\frac{1}{2} \text{Tr} \dot{K} X \)

• Heat flow =

\[
\frac{1}{2} \text{Tr} \dot{X} K = -\frac{1}{2} \sum_i \text{Tr} K (M_i X + X M_i) + \sum_i T_i \text{Tr} F_i F_i^T K
\]

• Heat in-flow = fluctuation term = \( \sum_i T_i \text{Tr} F_i F_i^T K \)
Efficiency

• During a cycle:

\[
\text{Efficiency} = \frac{\text{Work}}{\text{Heat received}} = \frac{-\frac{1}{2} \int \text{Tr} \dot{K} X}{\sum_i T_i \int \text{Tr} F_i F_i^T K}
\]

• If one temperature, efficiency \( \equiv 0 \)

• If two temperatures, efficiency \( \leq 1 - \frac{T_{\text{low}}}{T_{\text{high}}} \)

• Design the optimal cycle? In finite time?

• Given an open system, design the optimal controller?
Entropy

- To prove efficiency bounds: entropy
- Entropy is
  \[ S = \frac{1}{2} \log \det X(t) \]

- Second Law
  \[ \dot{S} \geq \sum_i \frac{Q_i}{T_i} = - \sum_i \frac{1}{T_i} \left( \frac{1}{2} \text{Tr} K (M_i X + X M_i) + T_i \text{Tr} F_i F_i^T K \right) \]

- In control and in physics,
  - entropy not interesting for itself
  - gives bound on work and heat
Related Work

• *Willems and Brockett* (1978): Carnot cycles
• *Brockett* (1999): Control of stochastic ensembles
• *Newton and Mitter*: Kalman filter, information theory, and thermodynamics.
• *Lloyd and Touchette*: Control and information theory
• *Haddad, Chellaboina, and Nersesov*: Thermodynamics using storage function concepts.
• *Willems* (2006): Connections to behaviors
• “Finite-time thermodynamics”
Summary

- **Limitation I**: Dissipative devices give noise when realized physically (from uncertainty in initial state). Example: measurement devices

- **Limitation II**: Measurements disturb the measured system (*back action*) through conservation of energy

- **Limitation III**: Work can be partly extracted from heated system using (noisy) measurements and models. (Compare with Carnot)

- **Limitation IV**: Work can be partly extracted from heated system by allowing time-varying systems (compare with Carnot)

- Possible connections to Bode-Shannon?