Notions of Energy and Entropy

Jean-Charles Delvenne, Henrik Sandberg,

and John Doyle

14 August 2006

Overview:

- **Tutorial on some notions of thermodynamics, information theory, and dynamical systems, in particular energy, entropy, and temperature**

Summary

Energy

Entropy in information theory

Large isolated systems

Physical entropy

Large interconnected systems

Temperature, heat, work

The second law, Carnot cycle

Entropy in linear systems

Energy in Linear Systems

Linear time-invariant system:

$$\dot{x} = Ax + Bu \quad (1)$$
$$y = Cx \quad (2)$$

Through input and output, we can supply energy to the system.

We often take $\text{Supply} = u^T y$

(eg, voltage – current, force – velocity)

A physical system preserves stored energy *U(x)* (is lossless) if

$$\dot{U}(x(t)) = u(t)^T y(t)$$

<u>Dissipativity theory</u> (Willems)

Energy

The system (1)-(2) is lossless iff there is a sym. pos. def. matrix K s.t.

$$\begin{aligned} 0 &= A^T K + KA \\ KB &= C^T \end{aligned}$$

(Kalman-Yakubovich-Popov Lemma)

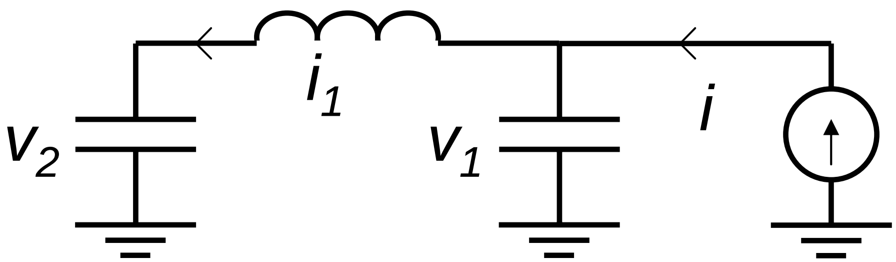
The energy can be chosen as $U(x) = \frac{1}{2} x^T K x$

Through change of variables, we can always suppose

$$\begin{aligned} K &= I \\ A^T + A &= 0 \\ B &= C^T \end{aligned}$$

Example: LC Circuit

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



Input: current i
Output: voltage v_1

Dynamics:

$$\dot{x} = \begin{pmatrix} 0 & -1/\sqrt{C_1 L_1} & 0 \\ 1/\sqrt{C_1 L_1} & 0 & -1/\sqrt{L_1 C_2} \\ 0 & 1/\sqrt{L_1 C_2} & 0 \end{pmatrix} x + \begin{pmatrix} 1/\sqrt{C}_1 \\ 0 \\ 0 \end{pmatrix} u$$

Output, state:

$$y = \begin{pmatrix} 1/\sqrt{C}_1 & 0 & 0 \end{pmatrix} x, \quad x^T = \begin{pmatrix} \sqrt{C}_1 v_1 & \sqrt{L}_1 i_1 & \sqrt{C}_2 v_2 \end{pmatrix}$$

Energy, work rate:

$$U = \frac{1}{2} x^T x = \frac{1}{2}(C_1 v_1^2 + L_1 i_1^2 + C_2 v_2^2), \quad w = yu = v_1 i.$$

Entropy in Information Theory

Random variable with discrete values *1,2,…,N* with probability *p(i).*

'Surprise' to find value *i* is $-\log p(i)$
(p(i) small $\Rightarrow$ big surprise if *i* occurs $\Rightarrow$ a lot of information)

Average 'surprise' =average 'uncertainty' = average 'information' =

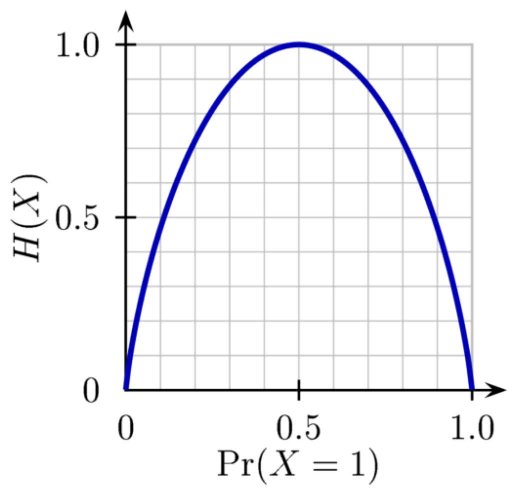
$$H(p) = -\sum_i p(i) \log p(i)$$

(with 0 log 0 = 0)

Also defined for countably many events

Called <u>Entropy</u> (Shannon, 1948)

Example for $N = 2$



Maximum entropy

Highest possible entropy for a *N*-valued probability distribution?

Distribution maximizing entropy= uniform distribution

Maximal entropy = $\log N$

Proof: convexity of logarithm

One interpretation:

If no a priori knowledge on *N* events, then assign uniform a priori distribution

(Principle of Indifference)

Noiseless coding

Distribution on *N* events

Code every event by a binary word

No word is a prefix of another

Then <u>average length ≥ entropy of distribution</u>

(Shannon noiseless coding theorem)

Gives interpretation of entropy

Relative (differential) entropy

What if uncountably many outcomes?

Example: distribution on the real numbers with density $\phi(x)$

$$\text{Relative entropy} = h(\phi) = -\int \phi(x) \log \phi(x) dx$$

Taken to $-\infty$ **if no density**

Relative to a measure (e.g., usual measure on the real numbers)

If we quantize the real line in intervals of size $\varepsilon << 1$**, then discrete entropy makes sense and**

$$\text{Relative entropy} \simeq \text{Discrete entropy} + \log \epsilon$$

Boltzmann distribution

Let *U* = given function on the real numbers (energy for example)

How to maximize entropy for given expected value of *U* ?

$$\max h(\phi) \text{ s.t. } \int U(x)\phi(x)dx = U_0$$

Solution: $\phi(x) = \alpha e^{-\beta U(x)}$

(Boltzmann distribution)

Examples:

- If support on a set of measure *V*, then uniform distribution optimal with <u>entropy log *V*</u>
- If $U(x) = x^T K x$ then normal distribution with covariance matrix K^{-1} is optimal, with entropy $\frac{1}{2}\log\det K^{-1} + \frac{1}{2}n\log 2\pi e$

Large isolated systems

Physical systems with many degrees of freedom

We cannot measure all variables

We know only a few quantities (e.g., energy)

We put a probability distribution on the state

What probability distribution?

Axiom of statistical physics:

Isolated system at equilibrium with given energy is endowed with uniform distribution

Why? Difficult to justify

Ergodicity

The uniform distribution is invariant:

- Linear systems: rotations
- In general: Liouville's theorem

<u>Ergodic hypothesis</u>: The uniform distribution is the only invariant distribution

Not true for all systems

Never true for linear systems (decouple all degrees of freedom)

MaxEnt principle

Toss a coin: uniform probability on the result

If the coin is known unfair: uniform probability on the result

In general: assign the distribution compatible with what you know, that maximizes the entropy

=<u>MaxEnt principle</u> (Jaynes 1957)

An isolated system with given energy: uniform distribution on all states of same energy satisfies MaxEnt principle

Bayesian vs. Frequentist view of probabilities

Physical entropy

The entropy of a physical system:

- Partition the phase space in cells of size ε
- Measure some macroscopic variables of interest (energy)
- Discrete entropy is the physical entropy

If isolated system at equilibrium, then entropy is logarithm of number of cells (Boltzmann)

$$S = k_B \log \Omega$$

Large interconnected systems

A system connected to a much larger system (called <u>bath</u>)

Energy of small system fluctuates

Only average energy is known

From MaxEnt principle: distribution is $\phi(x) = \alpha e^{-\beta U(x)}$
(Boltzmann)

Can also be proved from Axiom. Bath= many interconnected copies of the same system

For linear systems with energy $U(x) = \frac{1}{2} x^T x$
the probability distribution is normal with covariance $\beta^{-1} I$

Temperature

Boltzmann distribution $\phi(x) = \alpha e^{-\beta U(x)}$

Lagrange coefficient $\beta = \frac{1}{k_B T}$

Defines temperature = how much entropy increases when energy increases.

Temperature adjusted to match the average energy of the system

Linear systems with energy $U(x) = \frac{1}{2} x^T x = \frac{1}{2} \sum_i x_i^2$ **:**

$$\mathbb{E}(\frac{1}{2} x_i^2) = \frac{1}{2} k_B T$$

Hence, temperature also average energy of any degree of freedom (<u>Equipartion of energy</u>)

Heat, work, and first law of thermodynamics

<u>First law of thermodynamics</u>: Energy is decomposed in heat and work

$$\dot{U} = w + q$$

Classical thermodynamics: heat, understood intuitively, is a form of energy

For statistical mechanics:

Heat = unknown mechanical energy (e.g., because distributed many degrees of freedom) ~ associated with variance

Work= known mechanical energy ~ associated with mean motion

Example: In linear systems with internal energy $U(x) = \frac{1}{2}x^T x$**:**

$$\mathbb{E}(U) = \frac{1}{2}(\text{Mean}^2 + \text{Variance})$$

Second law of thermodynamics

(Physical) Entropy increases in a system when heat *q* enters at a point of temperature *T* according to:

$$\frac{dS}{dt} \geq q/T$$

Equality holds for reversible (idealized infinitely slow changes)

Hence a system in contact with one temperature cannot convert heat into work (*Δ W* negative) in a cycle (entropy and internal energy are state functions):

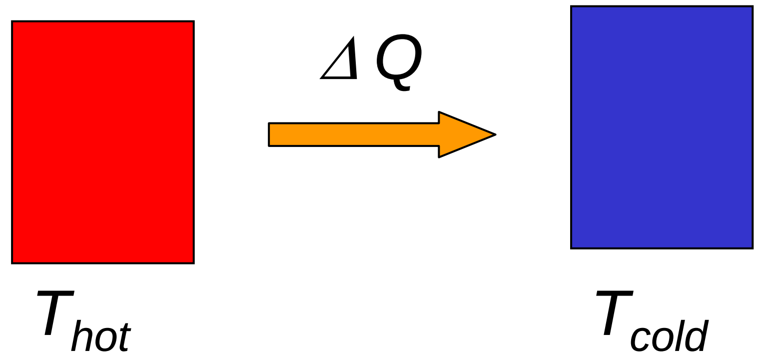
$$0 = S(T) - S(0) \geq \int_0^T \frac{q(t)}{T} dt = \frac{\Delta Q}{T}$$

$$0 = U(T) - U(0) = \int_0^T w(t) + q(t) dt = \Delta W + \Delta Q$$

$$\Rightarrow \Delta W \geq 0$$

Illustration of the second law

An intuitive example of natural increase of entropy:



$$\Delta S_{hot} \geq \frac{-\Delta Q}{T_{hot}}, \qquad \Delta S_{cold} \geq \frac{\Delta Q}{T_{cold}}$$

$$\Rightarrow \Delta S_{total} = \Delta S_{hot} + \Delta S_{cold} \geq 0.$$

Second law: The <u>total</u> entropy of a system naturally increases

Two temperatures and Carnot heat engines

If a system has access to two temperature baths:

During a cycle,

$$0 = \int \dot{S} dt \geq \int \frac{q_{hot}(t)}{T_{hot}} + \frac{q_{cold}(t)}{T_{cold}} dt$$

Work extracted is

$$\int q_{hot}(t) + q_{cold}(t) dt$$

Efficiency $\frac{\text{Work extracted}}{\text{Heat received}} = \frac{-\int w(t)dt}{\int q_{hot}(t)dt} \leq 1 - \frac{T_{cold}}{T_{hot}}$

How to reach maximal efficiency: contact with one bath at a time, infinitely slow process = <u>Carnot cycle</u>

Entropy in linear systems

A stochastic (white) signal *w* heating a linear system

$$\dot{x} = Ax + Gw,$$
$$\mathbb{E}x(0) = \bar{x}_0, \quad \mathbb{E}x(0)x(0)^T = R$$
$$\mathbb{E}w = 0, \quad \mathbb{E}w(t)w(\tau) = W\delta(t-\tau)$$

Evolution of mean and covariance:

$$\dot{\bar{x}} = A\bar{x}, \quad \bar{x}(0) = \bar{x}_0$$
$$\dot{X} = AX + XA^T + GWG^T, \quad X(0) = R$$

Internal energy and entropy (definitions):

$$U = \mathbb{E}\frac{1}{2}x^T x = \frac{1}{2}\bar{x}^T\bar{x} + \frac{1}{2}\mathrm{Tr}X$$
$$S = \frac{1}{2}\log\det X$$

Entropy in linear systems

Heat flow into system: $q = \frac{1}{2}\mathrm{Tr}\dot{X}$

Entropy change (use Jacobi's formula):

$$\frac{dS}{dt} = \frac{1}{2}\mathrm{Tr}(X^{-1}\dot{X})$$

In an equilibrium with equipartition X = *T I*:

$$\frac{dS}{dt} = \frac{1}{2}\mathrm{Tr}(X^{-1}\dot{X}) = 0$$

A linear lossless system connected to heat bath of temperature *T* (details on Tuesday!)*:*

$$\frac{dS}{dt} = \frac{1}{2}\mathrm{Tr}(X^{-1}\dot{X}) \geq q/T$$

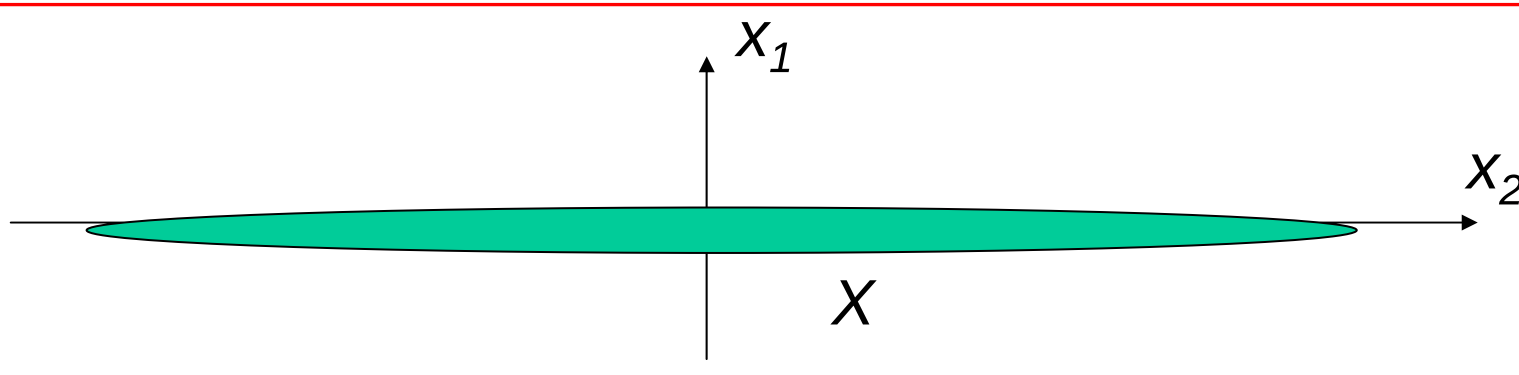
Entropy vs. Energy

Assume covariance has eigenvalues $\lambda_1,\ldots,\lambda_n$

Energy: $$U = \frac{1}{2}\mathrm{Tr}X = \frac{1}{2}\sum_i \lambda_i$$

Entropy: $$S = \frac{1}{2}\log \underbrace{\det X}_{\text{"volume"}} = \frac{1}{2}\sum_i \log \lambda_i$$

Energy can be very large while the entropy is very small!



Free Energy

First and second law gives:

$$\dot{U} = w + q \leq w + T\dot{S}$$

Define (Helmholtz) *free energy* as:

$$A = U - TS$$

Gives bound on possible amount of extracted work:

$$-\int_0^T w(t)dt \leq A(0) - A(T)$$

A large entropy *S* means a lot of the internal energy *U* is not available for work!