

Precise and Approximate Representation of Numbers

- The Cartesian-Lagrangian representation of numbers.
- The homotopic representation of numbers
- Loops and deck transformations
- The maximal ideal representation
- The place cell representation
- The radix r representation

Outline of the Day

- 9:30 - 10:45 Part 1. Examples and Mathematical Background
10:45 - 11:15 Coffee break
- 11:15-12:30 Part 2. Principal components, Neural Nets, and Automata
- 12:30 - 14:30 Lunch
- 2:30- 3:15 Part 3. Precise and Approximate Representation of Numbers
- 15:45 - 16:15 Coffee break
- 16:15 - 17:30 Part 4. Quantum Computation

The Usual Analog Computing Paradigm

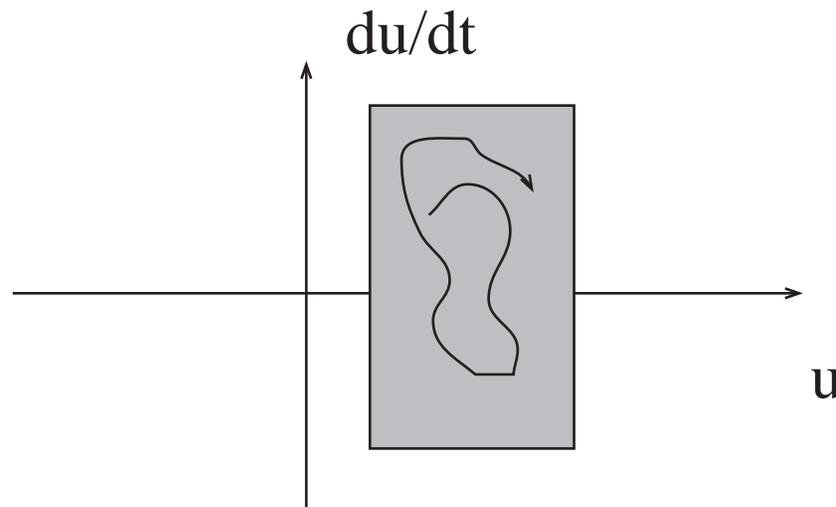
Problems arise because the dynamic range is limited by noise, and because the flows often contain hyperbolic points that expand the dynamic range. These considerations may make it more desirable to compute the geometric mean than the product, the average rather than the sum.

Some experience indicates that this representation of data is limited to about one part in 256, or eight bits.

Moreover this scheme is conceptually suspect in that it suggests that it is possible to transmit information perfectly from one location to another at arbitrary transmission rates. It fails to acknowledge either limited dynamic range or limited bandwidth.

The Use of Time

Consider the representation of numbers via pulse frequency modulation. If we can vary the pulse rate from 1 Hz. to 10^6 Hz. then we can represent one part in 10^6 . However, if we can not reliably count pulses at a rate faster than one every T seconds it will take $T \times 10^6$ seconds to transmit one number. These and other considerations suggest that. In this setting it is worthwhile to think of the communication constraints as limiting the signal space to a subset of the phase plane.

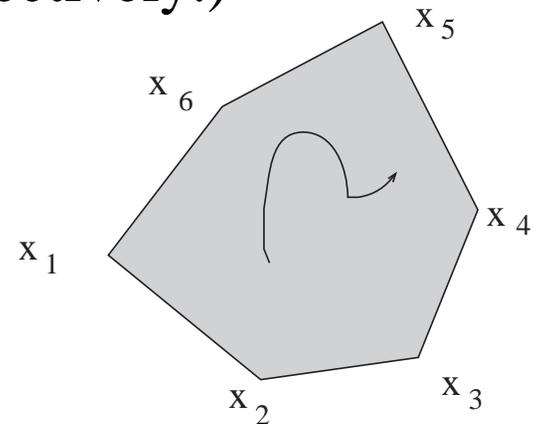


Channel Capacity

In this case the bit rate is $\ln 10^6 / T_x 10^6$. This model brings clearly into evidence the limitations on communication speed. It is impossible to know both the value and the rate of change because it takes time to transmit this information. This suggests a strong limitation on differential equations as computational models.

Place cell representations must be specified by the lowest possible pulse rate and the highest possible pulse rate. There are systems used just to give some numbers, there is a part of the auditory cortex for which these numbers are 50 and 250 respectively.)

The representation of a curve as a time varying convex combination of points.



Types of Quantizers

General Definition of “Digital” System

Partition the input space

$u \mapsto [u]$ = equivalence class

$$\dot{x}(t) = f(x(t), u(t)); y(t) = h(x(t))$$

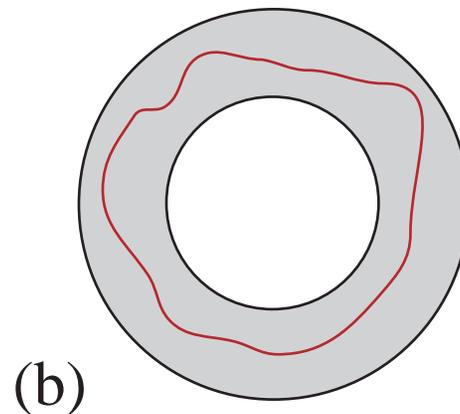
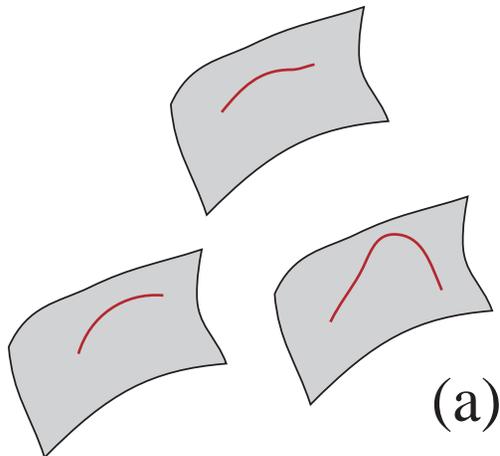
such that

$$\lim_{t \rightarrow \infty} x(t) = \phi([u])$$

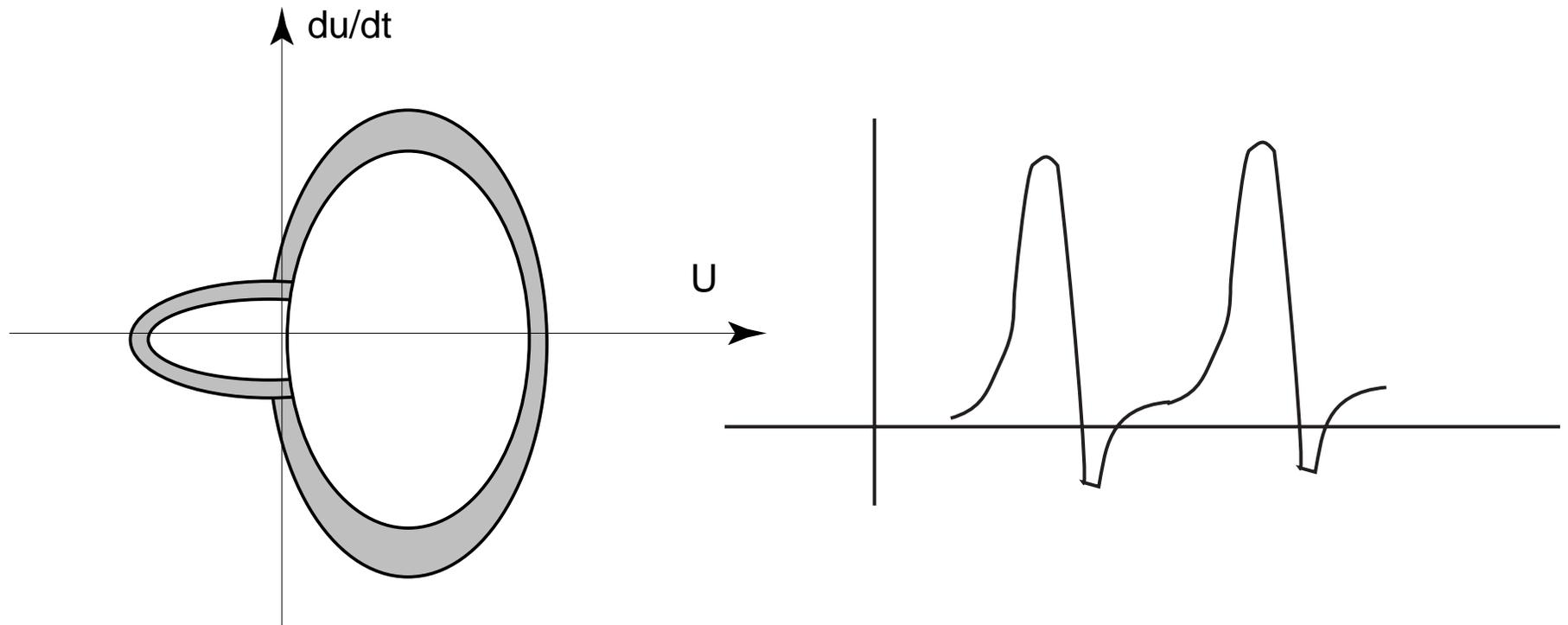
Types of Quantizers

In topology there is a distinction between π_0 —the number of connected components of a topological space and π_1 , the set of equivalence classes of closed curves.

Ordinary quantizers can be thought of as π_0 quantizers. Pulse counters can be thought of as π_1 quantizers



The Pulse Annulus and its Winding Number



The annular characterization allows arbitrary spacing between pulses, a characterization of a set of functions that is very different from, say, the characterization of band-limited functions.

Computing with Pulse Representations

A matched filter for pulses might take the form

$$\dot{x} = -\sin(2\pi x) + u$$

If the area under the pulse is near 1 and if the refractory period is long enough, this system will count pulses with no error.

In symbols

$\lim_{t \rightarrow \infty} x(t) = \nu[u]$ where ν denotes the winding number in the annulus sense.

$$\dot{x}(t) = -\sin x(t) + u(t)$$

u is “pulse-like” with area $\approx 2\pi$

Suppose that $x(0) \approx 0$

Is $x(t) \approx 2n\pi$ most of the time?

Yes, if the pulses are sharp enough
 x will advance in units of 2π

Applicable to

$$\dot{x}(t) = x(t) - x^3(t) + u(t)$$

$$x(t) - x(0) - \int_0^t u(\sigma) d\sigma = \\ - \int_0^t \sin x(\sigma) d\sigma$$

$$y(t) = x(t) - \int_0^t u(\sigma) d\sigma$$

$$y(t) - y(0) = \\ - \int_0^t \sin(y(\sigma) + \int_0^\sigma u(\eta) d\eta) d\sigma$$

$$\sin(y(\sigma) + \int_0^\sigma u(\eta) d\eta) = \\ \sin y(\sigma) \cos \int_0^\sigma u(\eta) d\eta + \\ \cos y(\sigma) \sin \int_0^\sigma u(\eta) d\eta$$

But if u is “pulse-like”

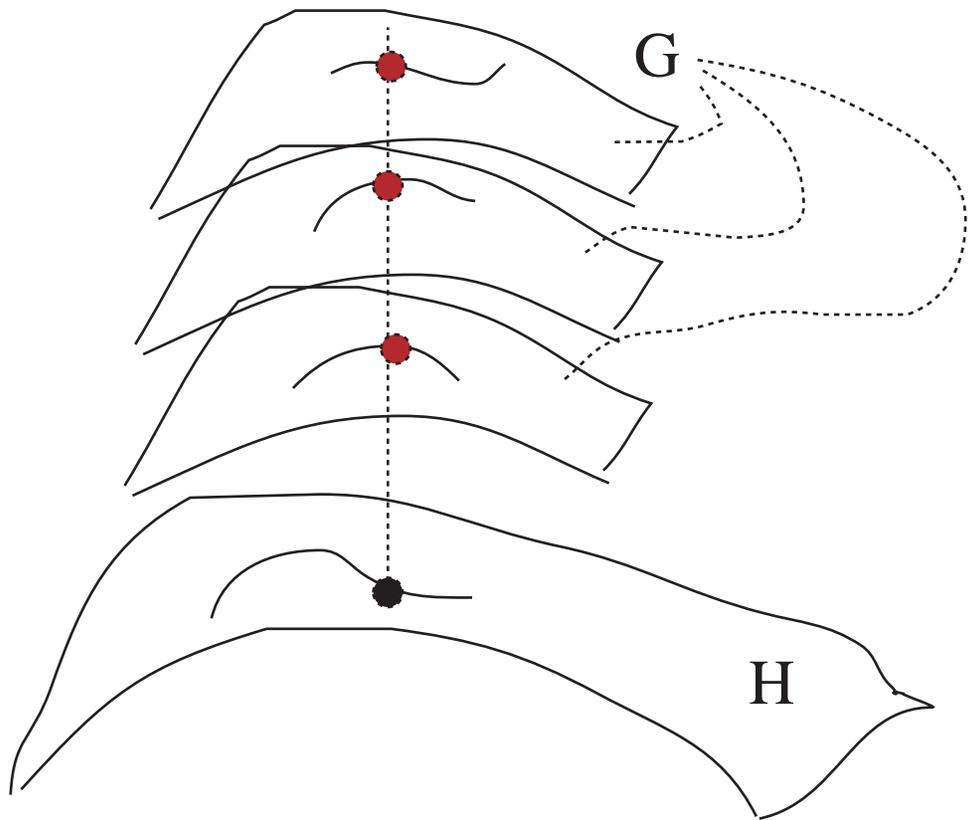
$$\cos\left(\int_0^\sigma u(\eta) d\eta\right) \approx 1$$

$$\sin\left(\int_0^\sigma u(\eta) d\eta\right) \approx 0$$

in the sense that the integral of the deviation is small

A Few Facts from the Topology of Adjoint Orbits

The n -dimensional orthogonal group $So(n)$ is an 2^{n-1} -fold cover of $\text{Sym}(\Lambda)$. Interpretation: First of all $\text{Sym}(\Lambda)$ is a manifold of the same dimension as $So(n)$. Secondly, for each $H \in \text{Sym}(\Lambda)$ there are 2^{n-1} different values of Θ such that $\Theta_i^T H \Theta_i = H$. If $\Theta(\alpha) \in So(n)$ for $\alpha \in [0,1]$ defines a curve joining the identity to Θ_i then there is a closed curve $H(\alpha) = \Theta^T(\alpha) H_0 \Theta(\alpha)$ in $\text{Sym}(\Lambda)$. This shows that there are 2^{n-1} distinct elements of $\pi_1(\text{Sym}(\Lambda))$.



Moving Inverse Images Around

Consider the system

$$\dot{\Theta} = [H, N]\Theta$$

with H considered to be an input, subject to the constraint that $H(0) = H(T)$. Then the value of $\Theta(T)\Theta^{-1}(0)$ is independent of the details of the path and only depends its homotopy class in $\text{Sym}(\Lambda)$. There are 2^{n-1} distinct values for $\Theta(T)\Theta^{-1}(0)$. These are separated by at least π units of length.

Using Systems with Many Stable Equilibria.

The system

$$\dot{H} = [H[H, \text{diag}(H)]]$$

has $n!$ stable equilibria. If we add a control in the form of a diagonal matrix U we can steer the system between these equilibria, at will. By coupling it with a slower time scale copy of the system we can realize automata as

$$\dot{H} = [H[H, \phi(U, J)]]$$

$$\dot{J} = [J[J, H]]$$

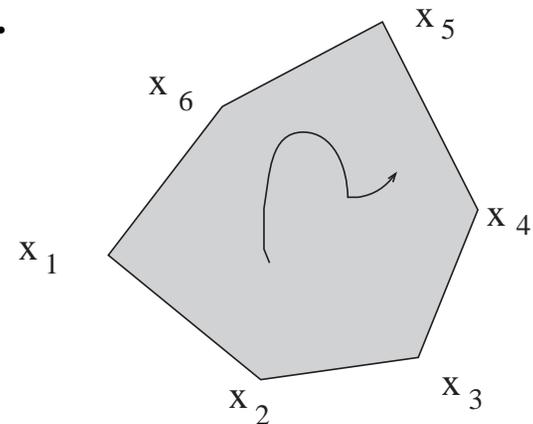
This is much more like standard digital implementations.

Robustness via Redundant Convex Combinations.

A second, completely different, way to achieve robustness of the representation of a real number is to represent it as the redundant weighted combination of fixed numbers. The number is represented by weights and a function of time is represented by an evolution in “weight space”. In the place cell picture the weights are organized topographically. It is also to think of them as being the coefficients in a radix r representation,

$$x(t) = \sum a_i(t)r^i$$

In this case there is no redundancy however. More typically the situation is as suggested On the right.



Probabilities on Vectors Spaces Lead to Averages

This representation of points is conceptually different from averages
In a probabilistic setting but when implemented it looks almost the
same.

Conditional Density Flow Takes away More or Less

$$\frac{d}{dt} \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix} +$$
$$\frac{-1}{2} \begin{bmatrix} (d_{11} - y)^2 & 0 & \dots & 0 \\ 0 & (d_{22} - y)^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & (d_{nn} - y)^2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix}$$

Computation with an Argmax (Place Cell) Representation

Fact: The conditional density for the usual gauss markov process observed with additive white noise evolves as a Gaussian.

The argmax of a Gaussian can be computed via the mean. Thus it is possible to convert ordinary differential algorithms into density evolution equations which will do the same calculations.

It seems likely that the brute force way of doing this is not the most efficient and from our knowledge of completely integrable systems It seems likely that one can find soliton equations that will perform these calculations robustly.

Conditional Density Equation: x Real Valued

Let $\rho(t, x)$ denote the conditional density of x , given the past observations.

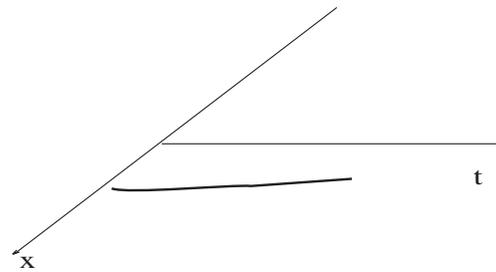
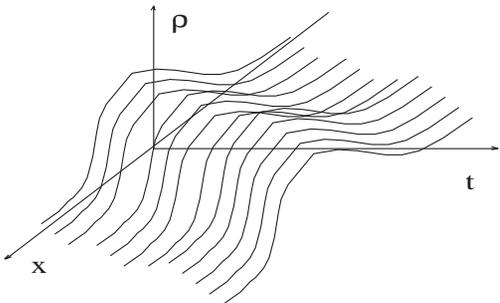
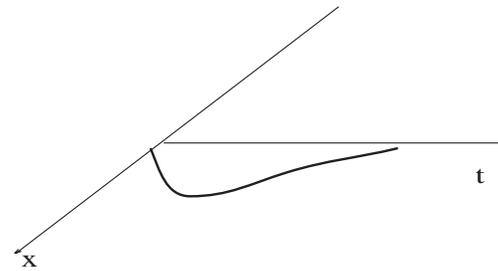
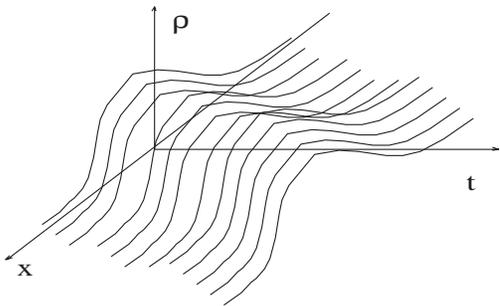
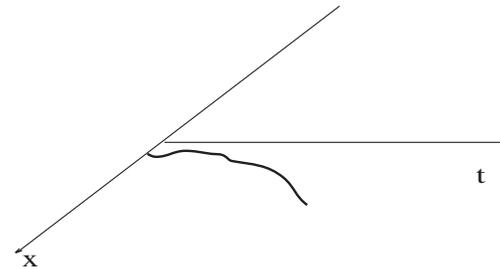
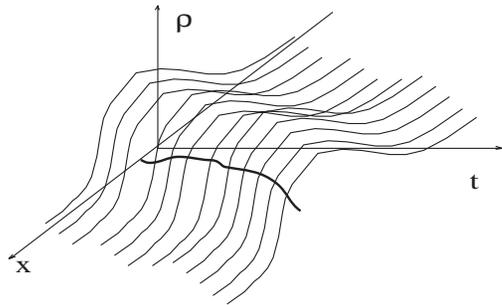
with no observations we have

$$\frac{\partial \rho}{\partial t} = L\rho$$

Observations is to change this to

$$\frac{\partial \rho}{\partial t} = L\rho - \frac{1}{2} \sum \phi_{s_k}^2(x) \rho + \sum \frac{dy_{s_k}}{dt} \phi_{s_k}(x) \rho$$

The Argmax Representation



A Distributed (Argmax) Model for Analog Computation

$$\frac{\partial \rho_1(t, x)}{\partial t} = L\rho_1(t, x) + F_1(\rho_1, \dots, \rho_n)\rho_1(t, x)$$
$$\frac{\partial \rho_2(t, x)}{\partial t} = L\rho_2(t, x) + F_2(\rho_1, \dots, \rho_n)\rho_2(t, x)$$

.....

$$\frac{\partial \rho_n(t, x)}{\partial t} = L\rho_n(t, x) + F_n(\rho_1, \dots, \rho_n)\rho_n(t, x)$$

Summary of Part 3

1. Quantization is necessary for communication, computation, reasoning and data storage. It must be adapted to the computational “hardware” available.
2. Topological invariants are attractive examples of robust quantization and can be related to computation in some cases. There are arguments that show one can compute with this type of representational scheme.
3. We have illustrated the realization of finite state machines using homotopy classes to represent states.
4. Conditional probabilities evolve and their evolution defines one of the most basic analog computers.