Perception Viewed as an Inverse Problem

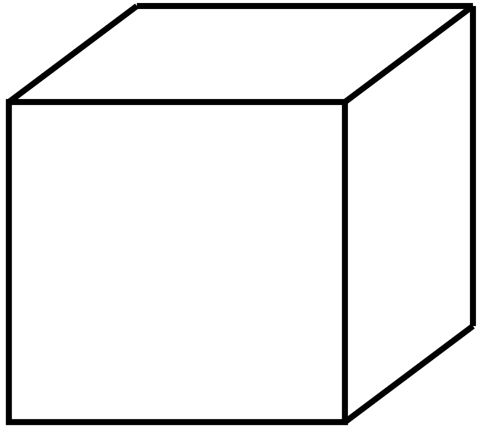
Fechnerian Causal Chain of Events - an evaluation

- Fechner’s study of outer psychophysics assumes that the percept is a result of a causal chain of events:

Distal stimulus $\longrightarrow$ proximal stimulus $\longrightarrow$ percept

- In this framework, the percept is a mental measurement of the physical stimulus.
- It seems reasonable to think that some properties of the *proximal* stimulus can be measured by the perceptual system (intensity, distance etc.)

It is less obvious how properties of the *distal* stimulus (e.g., 3D shape) can be measured by the perceptual system



- Consider the percept of a 3D object from a single image, like that shown on the left.
- There exists an infinite number of different 3D interpretations, corresponding to 3D objects that could produce this 2D image.
- Yet, we usually perceive only one 3D object. Here, the percept corresponds to a cube.

- Even if we assume that everything about the 2D retinal image can be perfectly “measured,” there is still not enough information to account for a unique (and usually, “veridical”) percept of a 3D object.
- To avoid the theoretical and experimental problems related to the inadequacy of the ‘*causal chain of events*’ framework, researchers often limited psychophysical studies to the case of simple stimuli (e.g. light and sound intensity), where the assumption about the ‘mental measurement’ seemed reasonable.
- To study complex stimuli, however, one needs to reformulate Fechnerian framework.

A New Approach

Distal stimulus $\longrightarrow$ Proximal stimulus

This mapping is called a forward (direct) problem

A forward problem is expressed in the rules of physics.

***The task (goal) for the perceptual system is to infer (reconstruct) the properties of the distal stimulus given the proximal stimulus*:**

Proximal stimulus $\longrightarrow$ Percept

This mapping is called an inverse problem.

New Framework

Let X be a distal stimulus and Y a corresponding proximal stimulus. Let A be a (linear) transformation. Then, the forward problem is formally expressed as follows:

$$\mathbf{Y = AX}$$

Most forward problems are well-posed and well-conditioned.

Forward Problems are well-posed and well-conditioned

- A problem is well-posed when
  - There is a solution
  - The solution is unique
  - The solution depends continuously on the data
- A problem is well-conditioned when
  - The solution is computationally stable in the presence of noise in the data

Inverse Problems

- The inverse problem of perceptual reconstruction of the distal stimulus from the proximal stimulus is formally expressed as follows:

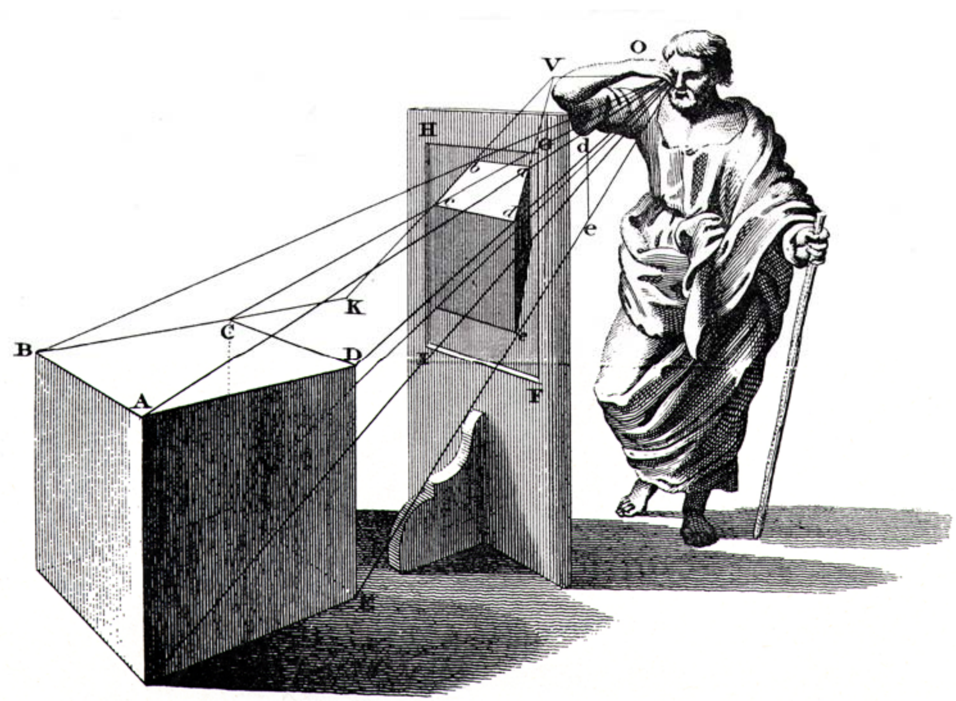
  $$X' = A^{-1}Y$$

  A^{-1} may not exist, may not be unique, or it may be unstable. All interesting inverse problems in science (and engineering) are ill-posed and/or ill-conditioned.

The Role of Constraints

- A standard way to solve ill-posed inverse problems is to impose *a priori* constraints on the family of possible solutions (interpretations).
- For the percept to be unique, stable and accurate (veridical), the constraints should reflect the properties of (regularities in) the natural environment.
- Examples of regularities present in objects from our environment: continuity and piece-wise smoothness of surfaces, symmetry, familiarity.

3D Percept from a 2D Image



An image of a regular polyhedron in the observer's eye can be produced by an object whose edges are curved, and faces non-planar. But we perceive a regular (*simple*) polyhedron.

Solving Inverse Problems

- Inverse problems are ill-posed and/or ill-conditioned
- In order to produce a unique, stable and accurate interpretation, the visual system has to impose *constraints* on the family of possible interpretations (regularization, MAP methods).

Contributions of Gestalt Psychology

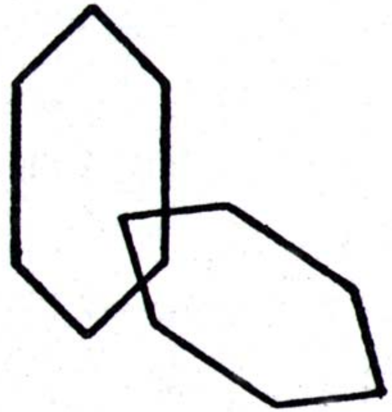


Fig. 26

Koffka (1935)

Two concave octagons or two convex hexagons?

The visual system tests two hypotheses and "chooses" the "*simpler*".

Spatially Global Relations vs. Spatially Local Interpretations

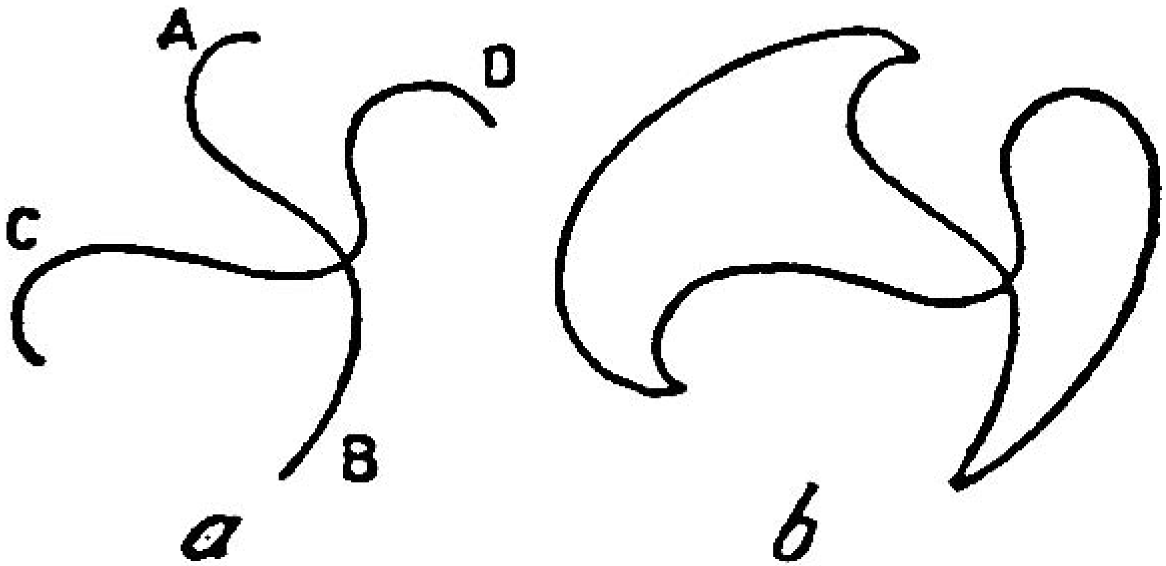


Fig. 46

Koffka (1935)

Experience in Perception

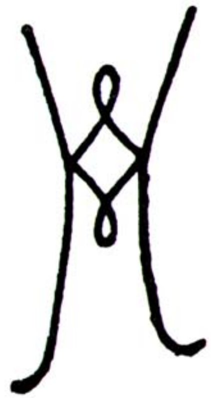
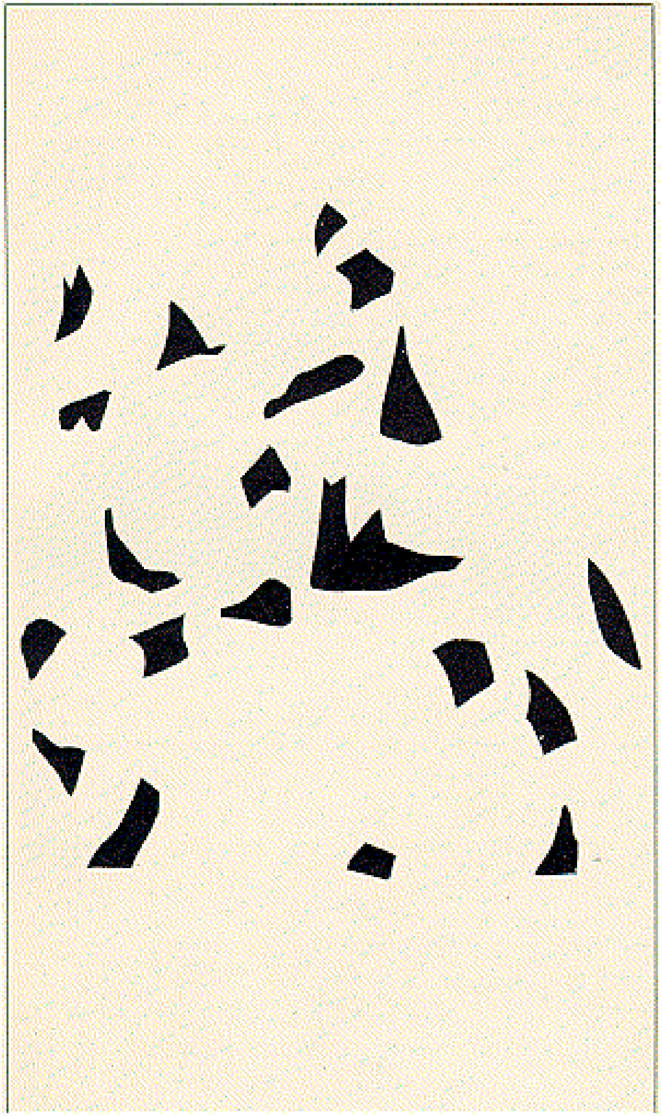


Fig. 29

Koffka (1935)

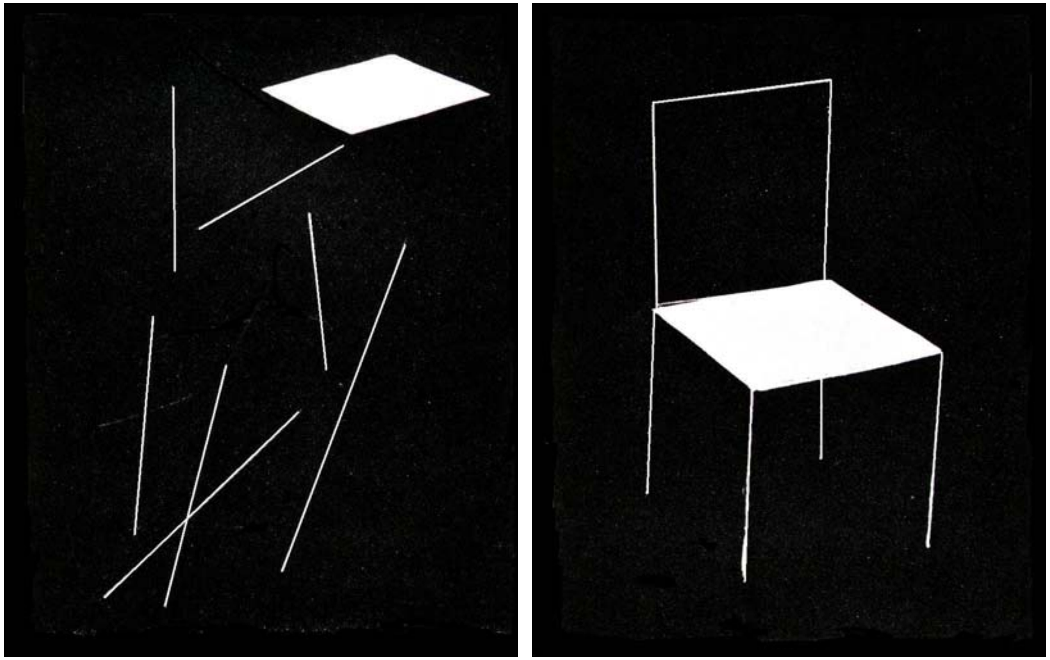
When presented with two perceptual hypotheses, the *simpler* wins over the more familiar.

Familiarity (Experience) in Perception

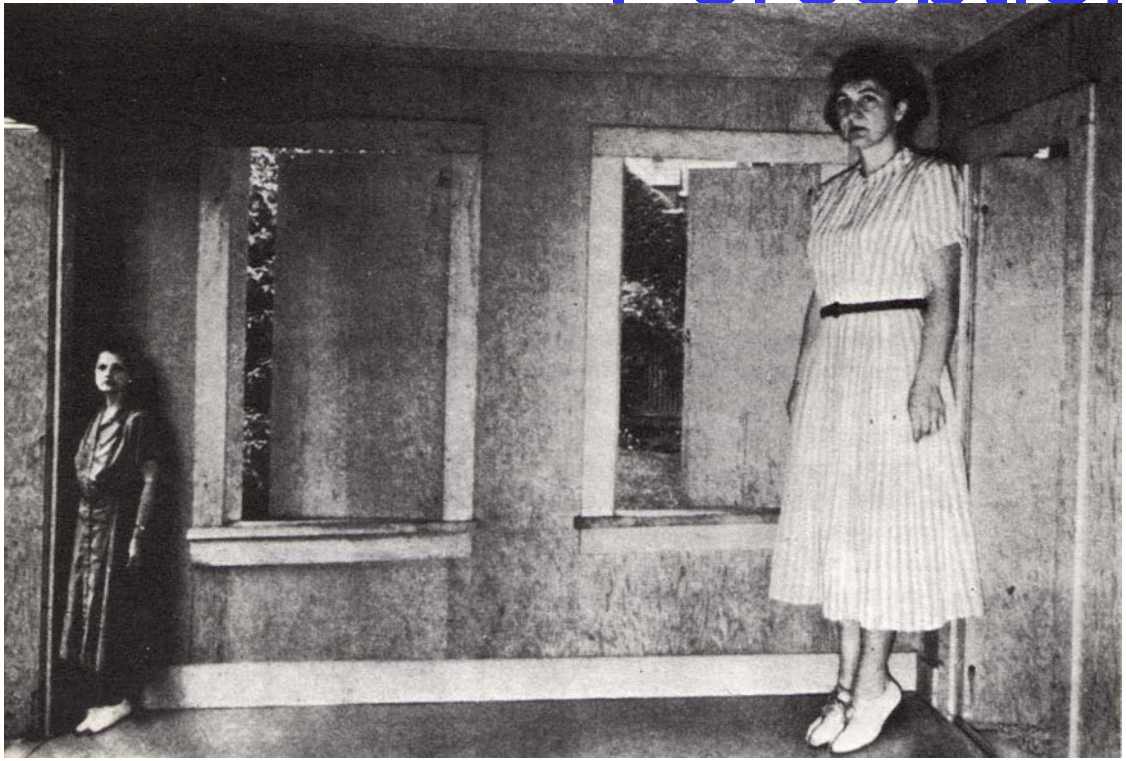


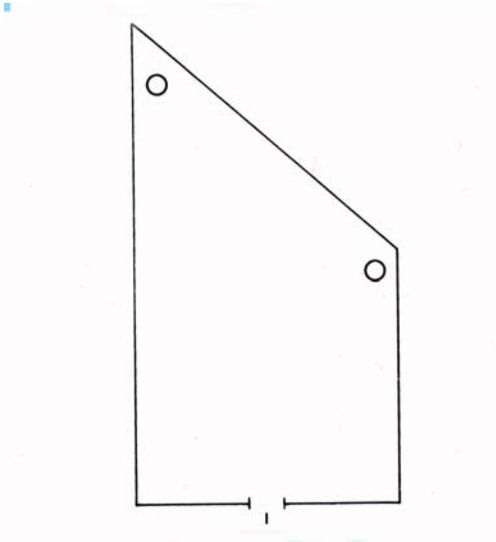
Street (1931)

Familiarity as a Constraint:
Transactional Psychology



The Role of Experience in Perception



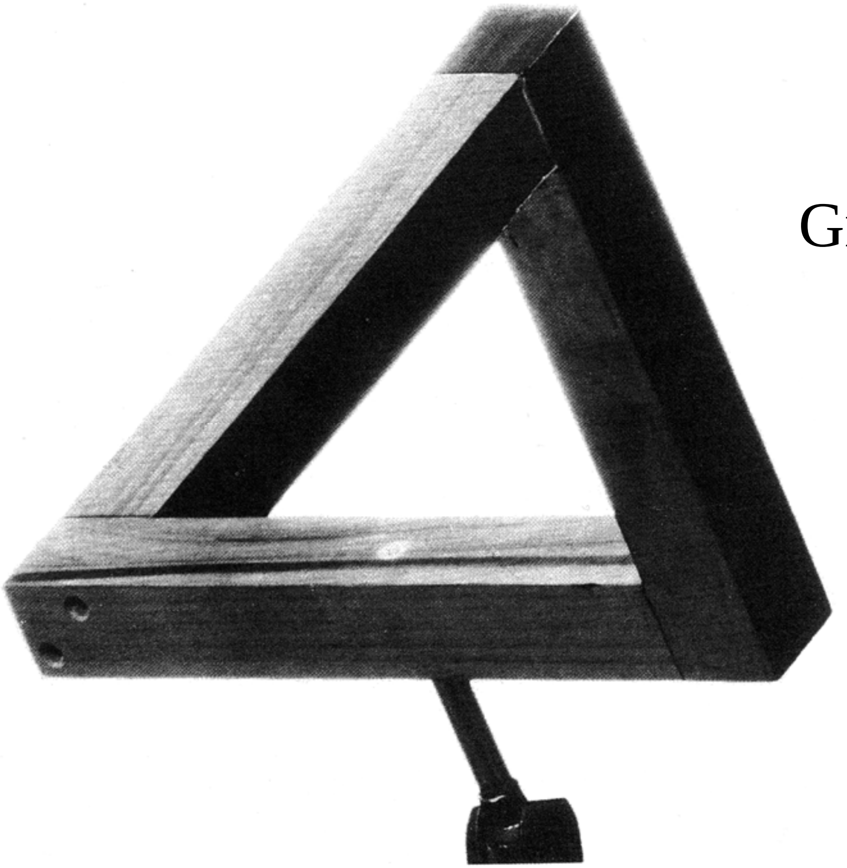


Ittelson & Kilpatrick, 1961

- If perceptual learning were the source of constraints, it would be easy to demonstrate the effect of experience on perception.

It is not!

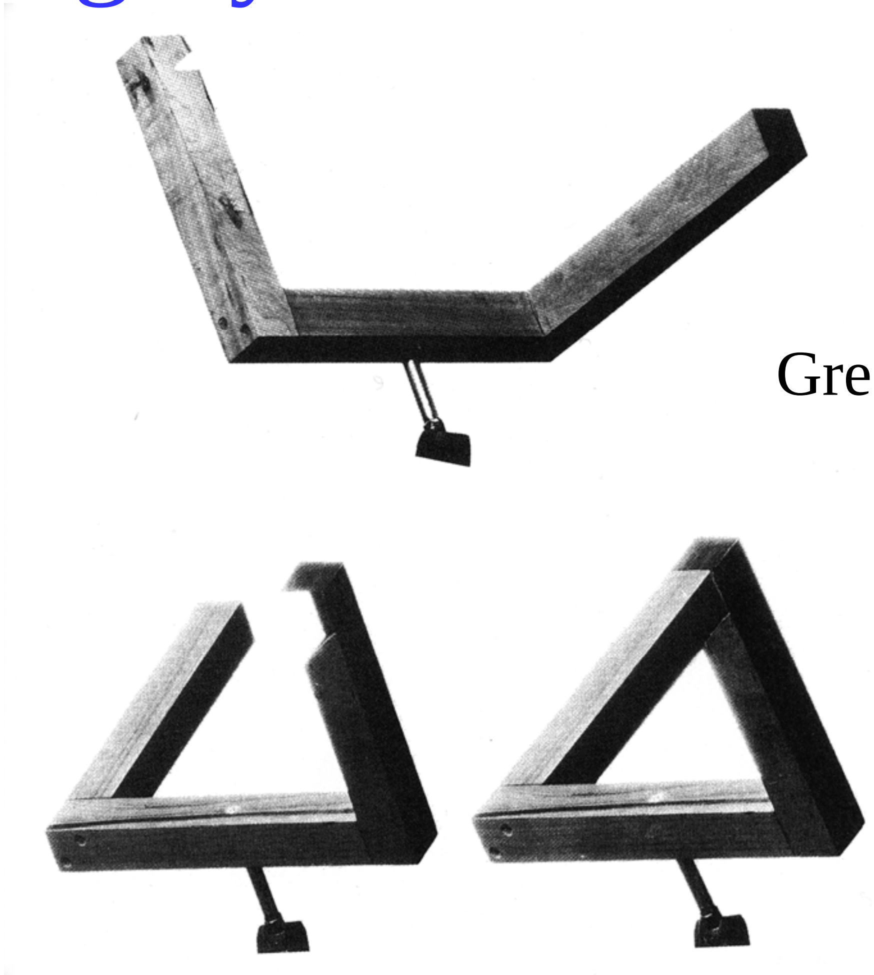
Penroses’ (1958) triangle



Gregory (1970)

It looks like an impossible object!

Gregory's Construction



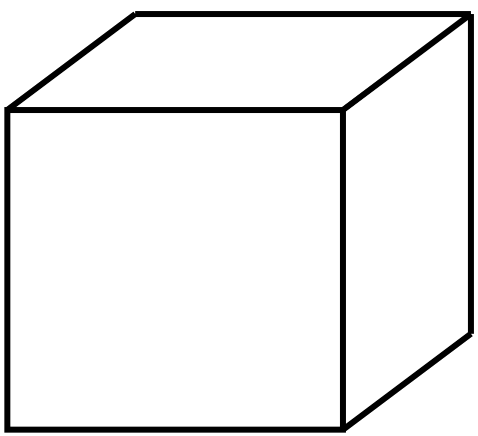
Gregory (1968)

Impossible interpretation cannot be the simplest or the most likely, when there is an object which is a possible interpretation…

Example

- Consider the problem of determining a 3D object from the image of a cube. This ill-posed problem can be solved by finding an object, which: is consistent with the given image, has planar faces, and has minimum variance of interior angles:

$$E(X)=||AX-Y||+\lambda_1||D_{pl}(X)||+ \lambda_2||Var_{\alpha}(X)||$$



The global minimum of E(X) corresponds to a cube. This solution agrees with the percept.

Examples

- Cuboid illusion:

http://bigbird.psych.purdue.edu/~pizlo/cuboid_illusion/

- Magniphi:

http://psych.purdue.edu/Magniphi/

Computational Methods – Regularization

- This method of solving inverse problems is called regularization.
  - In standard regularization (Tikhonov, 1963), the constraints are represented by a linear combination of the first *p* derivatives of the distal stimulus X. This corresponds to ‘smoothness’ of X.
  - λ is a parameter, whose value depends on how reliable is the proximal stimulus relative to constraints.

Computational Methods – MAP

- The stochastic version:

  $p(X|Y) = p(Y|X) \cdot p(X)/p(Y)$

  – $p(X|Y)$ is the posterior, $p(Y|X)$ is the likelihood, and $p(X)$ is the prior. The prior represents knowledge about distal stimuli. $p(Y)$ is the probability density of obtaining the proximal stimulus. It is a normalizing constant and, thus can be ignored.

- The inverse problem is solved by finding an X´ which maximizes the posterior (MAP estimate).

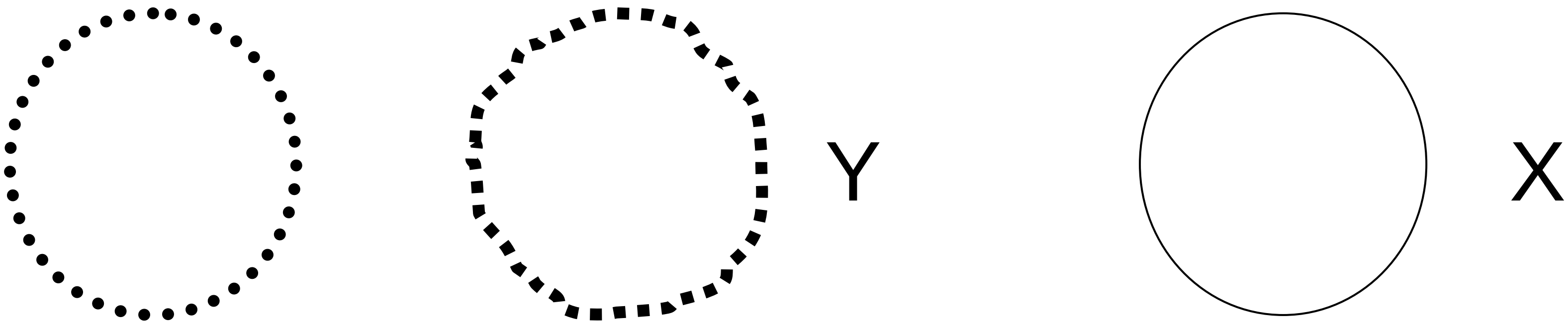
Regularization and MAP

- Regularization: $E = ||AX-Y|| + \lambda||P_X||$ (1)
- MAP: $p(X|Y) = p(Y|X) \cdot p(X)/p(Y)$

  or equivalently,

  - $\log p(X|Y) = -\log p(Y|X) - \log p(X)$ (2)
- Under some assumptions (quadratic norms, Gaussian pdf's), equations (1) and (2) are mathematically equivalent.
- The regularizing parameter λ in (1) is implicitly represented by the ratio of the variances of $p(Y|X)$ and $p(X)$ in (2).

Contemporary View of Perception



Retinal data is almost never sufficient for a unique and veridical percept (inverse problem is difficult)

$$\text{Percept}(Y) = \underset{X}{\arg\min}\left(||Y-X||^2 + \lambda||d\kappa_X/ds||^2\right)$$

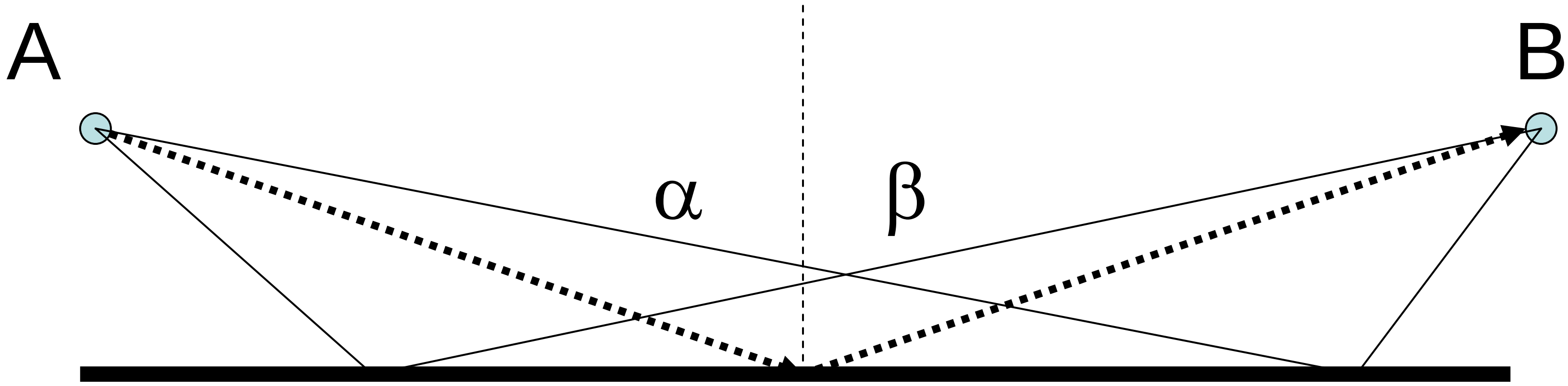
This is a quadratic variational problem

Minimum Principle in Physics

- Minimum (Simplicity) Principle in perception corresponds to the Minimum Principle in physics:
  - Hamilton’s least action principle in dynamics
  - Fermat’s least time principle in optics
  - Maxwell’s minimum heat theorem in electricity

Minimum Principle in Optics

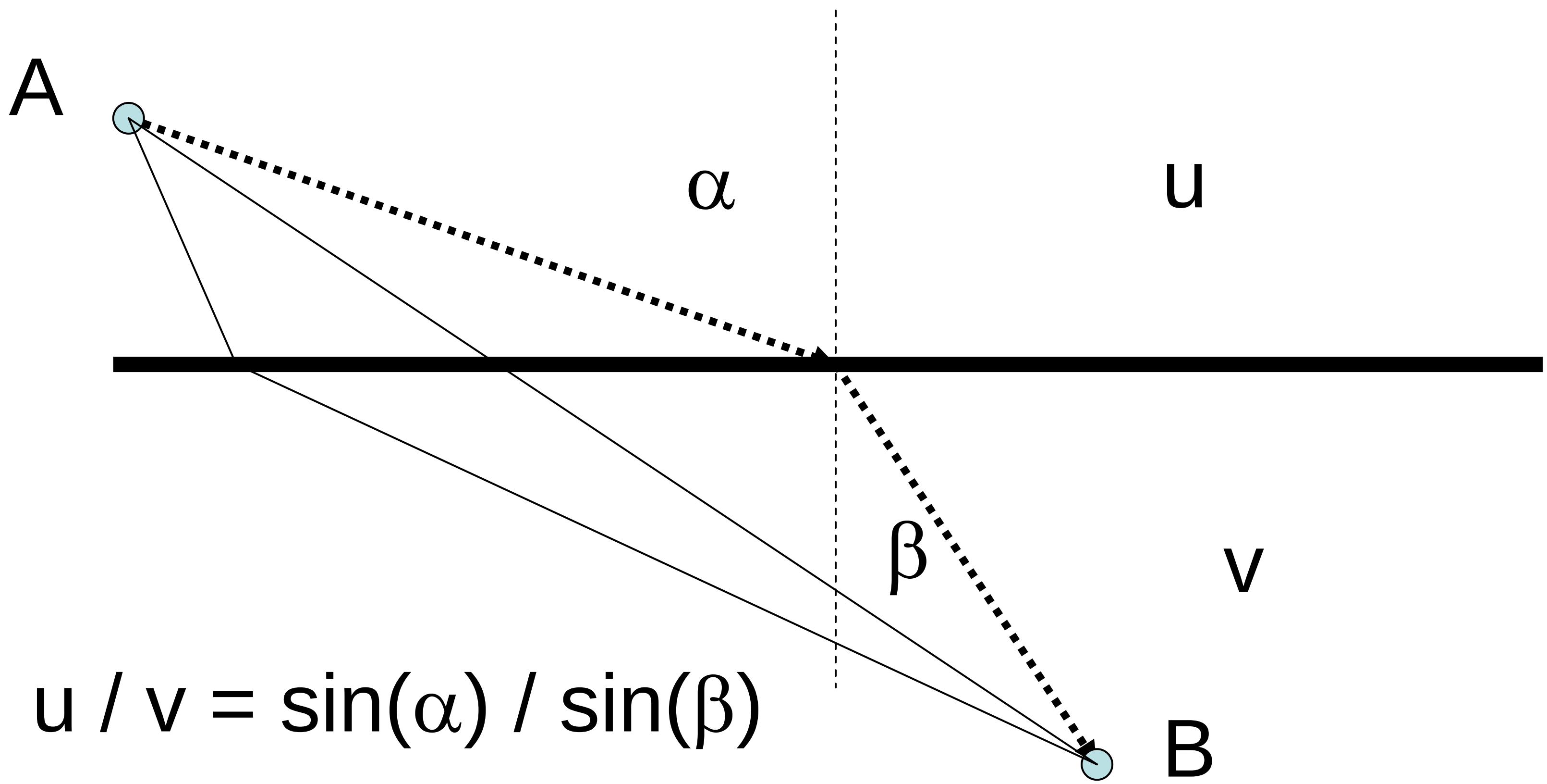
Law of reflection:



$\alpha=\beta$

Minimum Principle in Optics

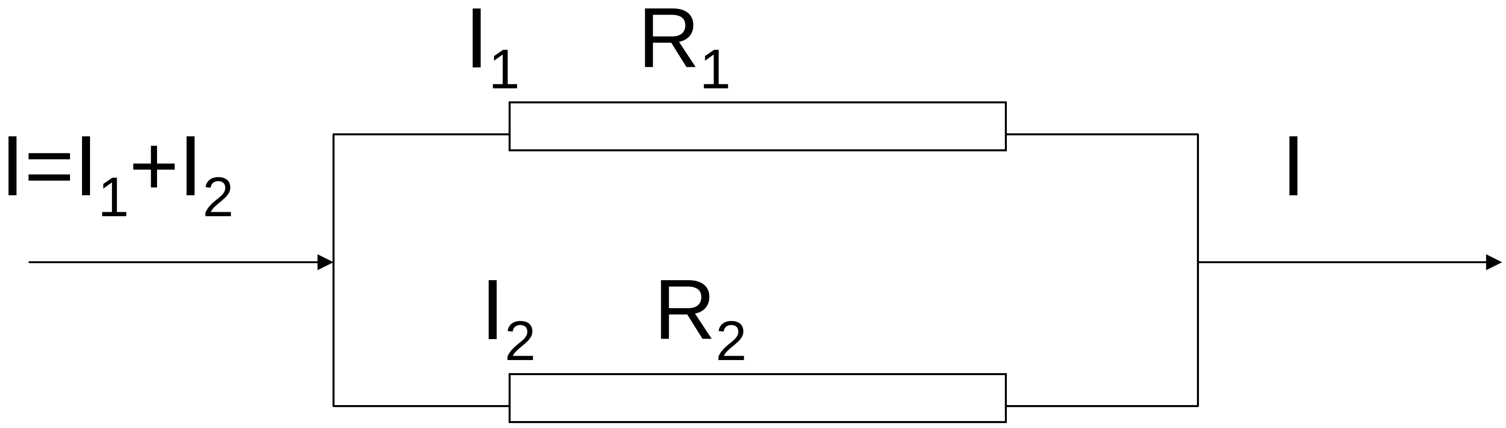
Law of refraction:



u / v = sin(α) / sin(β)

Minimum Principle in Electrical Circuits

- Kirchhoff's laws for electrical circuits:



min_heat = $\min(I_1^2R_1 + I_2^2R_2)$

$I_1/I_2 = R_2/R_1$

Köhler's (1920) Physiological Model

- The brain is a volume conductor whose minimum state (minimum heat) corresponds to perceptual simplicity
- This model was discredited in 1951 in experiments where short-circuiting the surface of the brain by metal plates, and inserting isolating plates into the brain of monkey did not affect her visual perception

Relevance of Köhler's Model

- For every quadratic variational problem with a unique solution, there exists a corresponding electrical network consisting of resistances and voltage or current sources having the same solution (Poggio & Koch, 1985)
- Electrical network (i.e., analog system) solves the minimum problem instantly, compared to a time-consuming process of solving a nonlinear optimization problem on a digital computer.

Köhler's Simplicity Principle

- Köhler was right! - except for his claim about physiological relevance of the electrical network.
- Electrical network is a physical model of perceptual processes.
- Different perceptual mechanisms require different electrical networks, representing different cost functions.

Examples of the Application of the Minimum Principle in Vision

Edge Detection

$$\int [(f - i)^2 + \lambda (f_{xx})^2 dx$$

f(x) – percept

i(x) – image

Optical Flow (edges)

$$\int [(V \cdot N - V^N)^2 + \lambda((\partial / \partial_s)V)^2]ds$$

s – arc length

V – perceived velocity

V^N – normal velocity component

N – normal unit vector to the contour

Optical Flow (area)

$$\int [(i_x u + i_y v + i_t)^2 + \lambda(u_x^2 + u_y^2 + v_x^2 + v_y^2)]dxdy$$

i(x,y,t) – image

(u,v) – perceived velocity field

Surface from Depth

$$\int [(f - d)^2 + \lambda (f_{xx}^2 + 2 f_{xy}^2 + f_{yy}^2)] dxdy$$

d(x,y) – depth data

f(x,y) – perceived surface

Perceived Depth (Binocular Disparity)

$$\int \{[\nabla^2 G * (L(x,y) - R(x + d(x,y), y))]^2 + \lambda (\nabla d)^2\} dxdy$$

d(x,y) – perceived depth

L(x,y) – left image

R(x,y) – right image

Binocular shape reconstruction

- In the absence of visual noise, binocular reconstruction is a well-posed problem (assuming known correspondence)
- In the presence of noise, however, the problem is ill-conditioned:

  http://bigbird.psych.purdue.edu/binshape/

  To stabilize the solution, several constraints are used, such as compactness and planarity.

References

- Pizlo, Z. (2001) Perception viewed as an inverse problem. *Vision Research*, 41, 3145-3161.
- Pizlo, Z. & Stevenson, A.K. (1999) Shape constancy from novel views. *Perception & Psychophysics* 61, 1299-1307.
- Poggio, T., Torre, V. & Koch, C. (1985) Computational vision and regularization theory. *Nature* 317, 314-317.
- Tarantola, A. (1987) *Inverse Problem Theory*. New York: Elsevier.