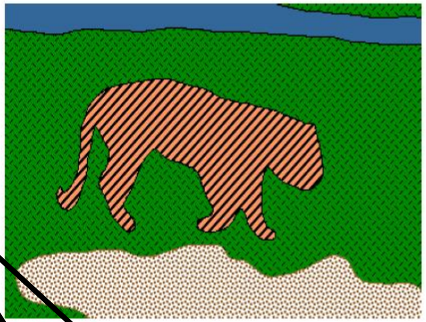
Medical Image Analysis
Lecture 03

Variational Methods & Denoising

Martin Urschler

We are not afraid of variational methods!



$$\max\left\{P(u\,|\,u_0)=\prod_{\Omega}P(u)p(u_0\,|\,u)\right\}$$

$$\min\left\{E=\frac{1}{2}\int_{\Omega}|\nabla u|^2\,dx+\frac{\lambda}{2}\int_{\Omega}(u-u_0)^2\,dx\right\}$$

$$\min\left\{\int_{\Omega}g(x)|\nabla u|\,dx+\lambda\int_{\Omega}\sum_{i\in I(x)}w_i(x)|u-f_i|\,dx\right\}$$

$$-\Delta u+\lambda(u-u_0)=0$$

Much of the material taken from the PhD thesis of Dr. Thomas Pock!

Basic Problem of Computer Vision

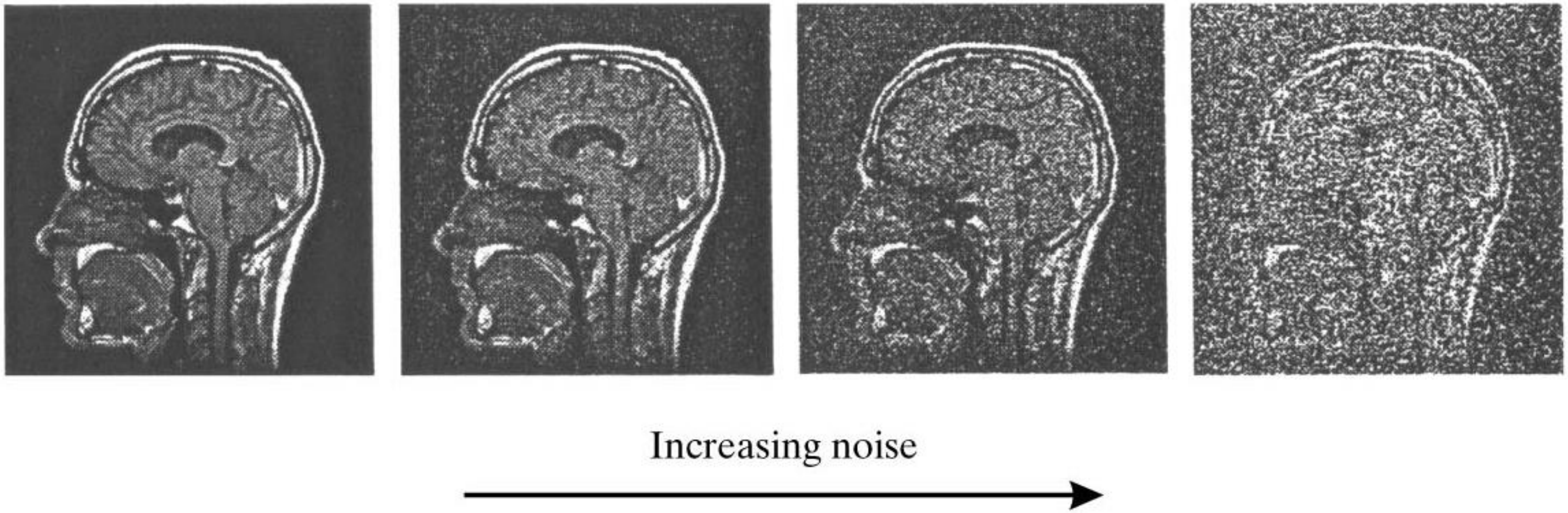
- CV deals with inverse (often ill-posed) problems:
  - Given observed data: estimate unknown quantities!
- Image Denoising / Restoration

Observed data: Noisy Image

Unknown Quantity: Clean Image

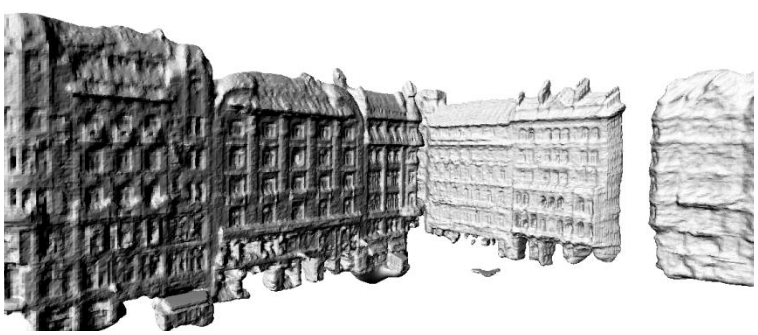
Denoising

- Inherent problem in (medical) image acquisition
- Physical processes involved often lead to compromises w.r.t signal to noise ratio



Basic Problem of Computer Vision

- CV deals with inverse (often ill-posed) problems:
  - Given observed data: estimate unknown quantities!
- 3D Reconstruction



Observed data: Stereo Images & Depth Maps

Unknown Quantity: 3D Model

Basic Problem of Computer Vision

- CV deals with inverse (often ill-posed) problems:
  - Given observed data: estimate unknown quantities!
- Image Segmentation

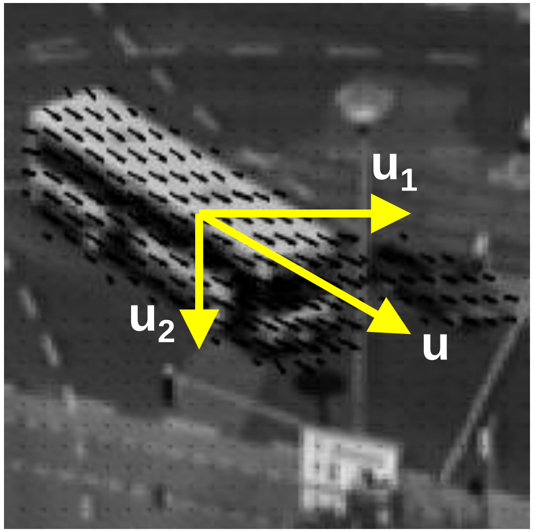
Observed data: Lena

Unknown Quantity: Fore-/Background

Basic Problem of Computer Vision

- CV deals with inverse (often ill-posed) problems:
  - Given observed data: estimate unknown quantities!
- Motion Fields & Image Registration

Observed data: Bus in Motion



Unknown Quantity: Optical Flow

Ill-Posed Problems

- Many solutions possible
- Regularization of the solution is needed.
  - Restrict space of possible solutions!

(a) True Image $u(x,y)$ $\leftarrow$ $\boxed{f(x,y)}$ (b) Degraded Image

Figure 1.1: A degraded image using motion blur and 2% additive Gaussian noise.

- How to choose regularization of a given problem?
  - A priori knowledge has to be incorporated to restrict the solution space.

Bayesian Inference

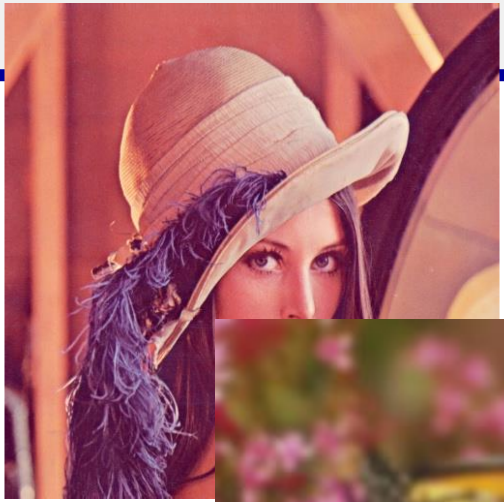
- No direct solutions -> constrict space of possible solutions to physically meaningful ones
- Statistical interpretation
  - Image u: random variable (drawn from probability distribution)
  - Assume it‘s possible to compute belief of hypothesis u being true p(u | f)
  - *We want to find u‘, the maximally probable hypothesis u solving our inverse problem given the observed image f.*

$$u' = \max_u \{ p(u \mid f) \}$$

  - Maximum a posteriori estimation (MAP)

Bayesian Inference

Observed data f



A: Reconstructed data u1
p(u1|f) very low

B: u2
p(u2|f)

C: u3
p(u3|f) maximum
among A,B,C

Bayes Rule

$$u' = \max_{u}\{p(u \mid f) = \frac{p(f \mid u)\,p(u)}{p(f)}\}$$

Conditional Probability,
Likelihood, Data model

Prior probability
Prior model

Normalization Factor

<u>Bayes Rule:</u>
Tells us how to update probability of hypothesis u given new observations f

Example: Tikhonov Denoising

Independence assumption: Pixelwise product of distributions

$$\max\left\{ p(u \mid f) = \prod_{\Omega} p(u)p(f \mid u) \right\}$$

Probabilistic Model

Model probabilities $\Updownarrow$ as normal distributions

$$p(u) = \exp\left(-\frac{|\nabla u|^2}{2} \right), \quad p(f \mid u) = \exp\left(-\frac{\lambda}{2}(u-f)^2 \right)$$

Mathematical Realization

Maximizing exp is $\Updownarrow$ minimizing its argument

$$\min\left\{ E = \frac{1}{2}\int_{\Omega} |\nabla u|^2 dx + \frac{\lambda}{2}\int_{\Omega} (u-f)^2 dx \right\}$$

Energy Functional

Minimizing functional by $\Updownarrow$ solving Euler-Lagrange equations

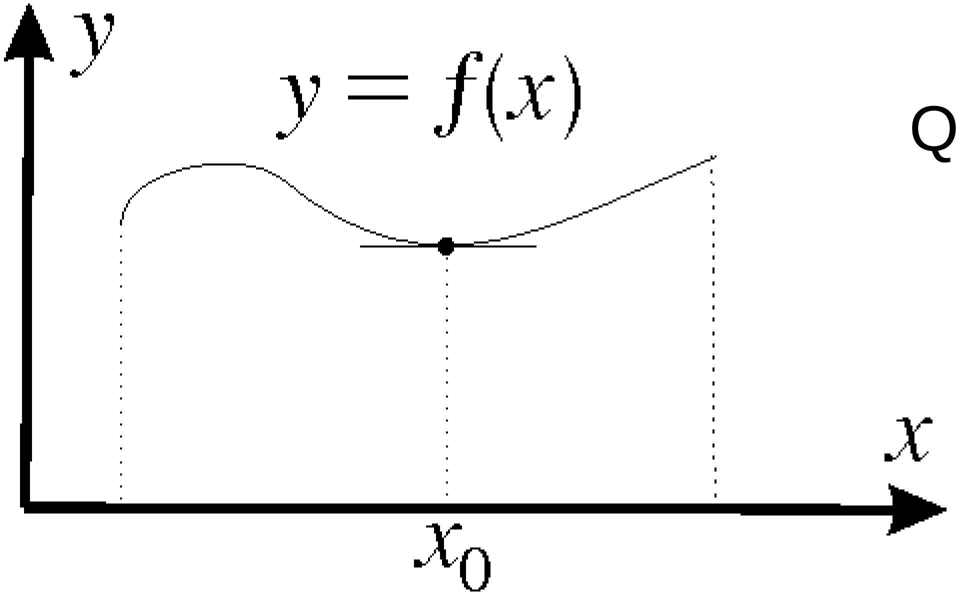
Calculus of Variations

$$-\Delta u + \lambda(u-f) = 0$$

PDE

A. Tikhonov, On the Stability of Inverse Problems, 1943

Calculus of Variations



Quiz: What makes x_0 special?

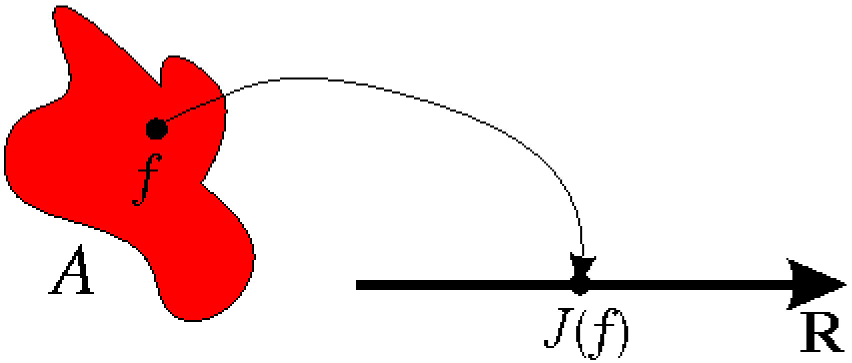
$$f'(x_0) = 0.$$

$$f : \mathbf{R} \rightarrow \mathbf{R}$$

Calculus of Variations

Functional

$J : A \to \mathbf{R},$



A is a set of admissible functions -> function space

Fundamental Problem of Calculus of Variations:

Given a functional J and a set A of admissible functions, find the <u>function(s) in A</u> that give(s) an extreme value to J.

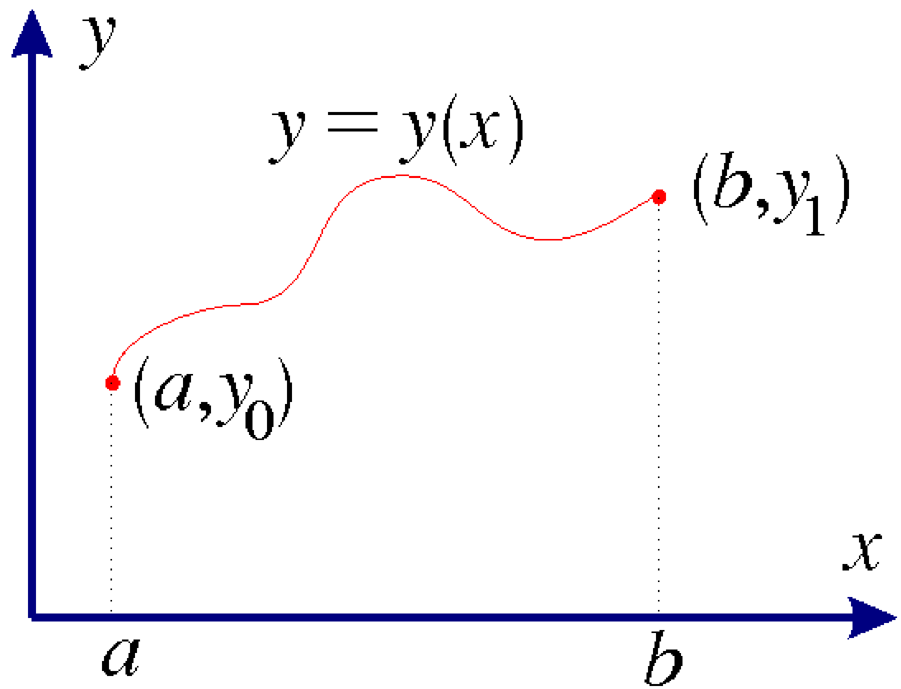
Calc. of Variations: Example 1

$$A = \{f \in C^1[a,b],\ y(a) = y_0,\ y(b) = y_1\}.$$

$$J(y) = \int_a^b \sqrt{1 + \left(y'(x)\right)^2}\, dx,\ \ y \in A.$$

Quiz: What is this functional about?

Quiz: What is the obvious solution when minimizing J(y)?



Calculus of Variations

- Generic Variational Formulation

$$J(y) = \int_a^b L(x, y, y')dx, \qquad y(a) = y_0, \; y(b) = y_1,$$

- Euler Lagrange Equation (P.D.E.)

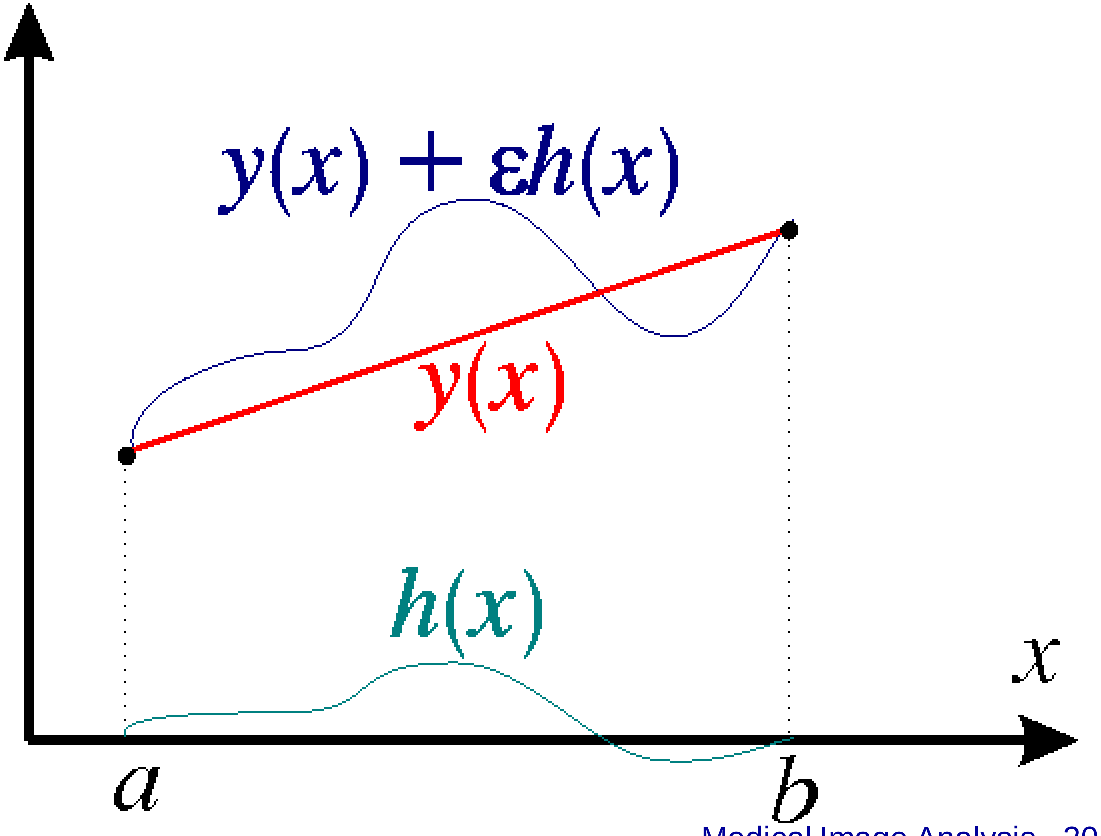
min {J(y)} $\Rightarrow$ $L'_y(x, y, y') - \frac{d}{dx} L'_{y'}(x, y, y') \overset{!}{=} 0, \; x \in [a, b].$

- Setting the Functional Derivative to Zero!

Functional (Gateaux) Derivative

$$\delta J(u;v) = \lim_{\varepsilon \to 0} \frac{J(u+\varepsilon v) - J(u)}{\varepsilon},$$

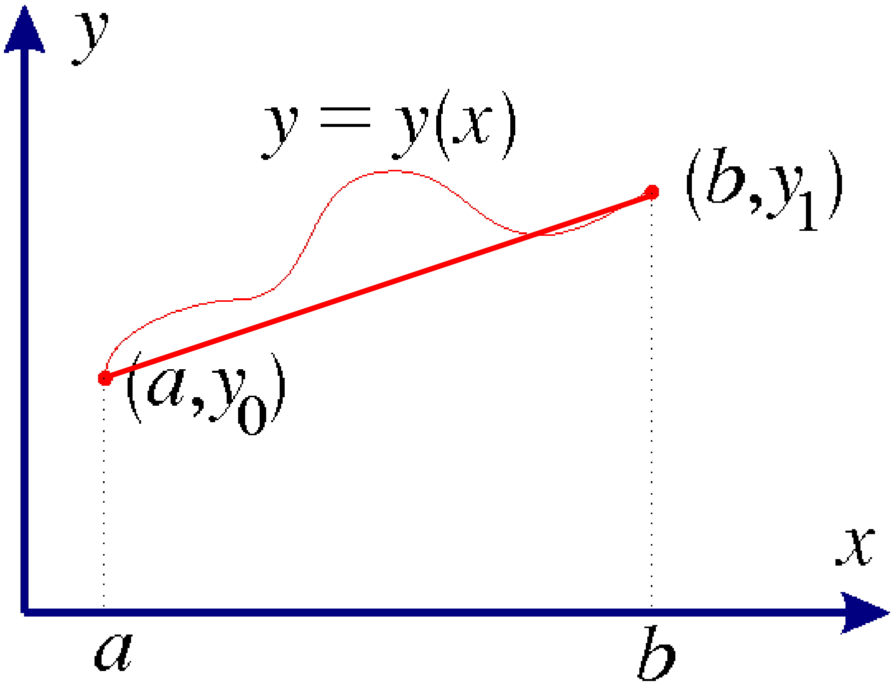
“Competing Curves”



Example 1: Derivation

$$A = \{f \in C^1[a,b],\ y(a) = y_0,\ y(b) = y_1\}.$$

$$J(y) = \int_a^b \sqrt{1 + \left(y'(x)\right)^2}\, dx,\ \ y \in A.$$

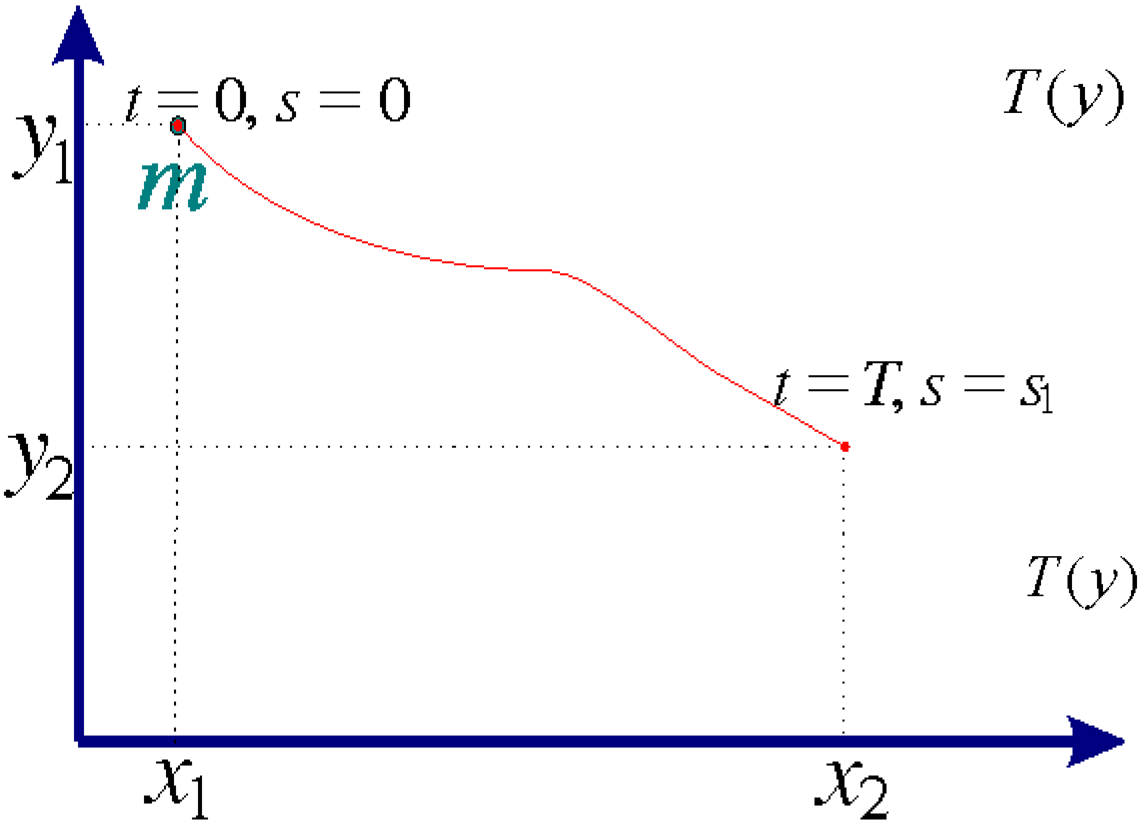


min{J(y)} -> Blackboard!

Solution:

$$y = \frac{y_1 - y_0}{b - a} x + \frac{b y_0 - a y_1}{b - a}.$$

Example 2: Brachistochrone Problem



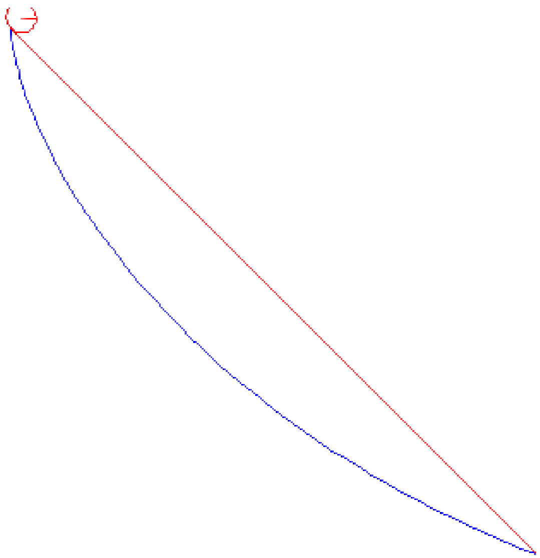
$$T(y) = \int_{x_1}^{x_2} \frac{ds}{v} = \int_{x_1}^{x_2} \frac{\sqrt{1+\left(y'(x)\right)^2}}{v} dx.$$

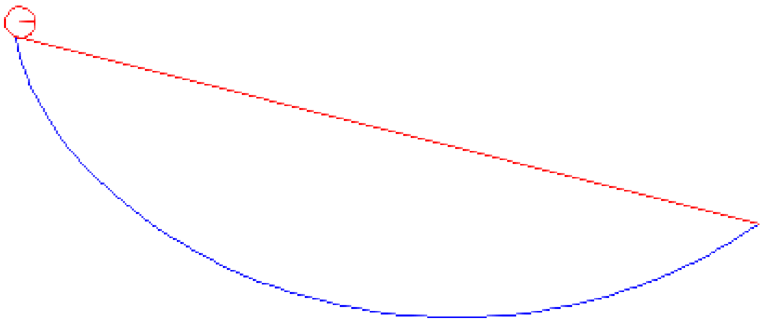
$$\frac{1}{2} m v(x)^2 = mg\left(y_1 - y(x)\right),$$

$$T(y) = \int_{x_1}^{x_2} \frac{\sqrt{1+\left(y'(x)\right)^2}}{\sqrt{2g\left(y_1 - y(x)\right)}} dx, \; y \in A,$$

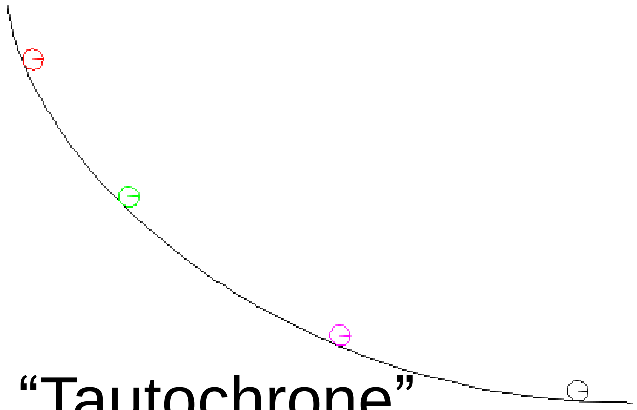
min{T(y)}!

Brachistochrone Solution





Fastest Solution Path is a Cycloid Curve!



Cycloid is a “Tautochrone”

Euler-Lagrange for Two Dimensions

$$J(u) = \iint_{\Omega} L\Big(x, y, u(x,y), u'_x(x,y), u'_y(x,y)\Big)\,dxdy,$$

$$\min\{J(u)\} \Rightarrow L'_u - \frac{\partial}{\partial x} L'_{u'_x} - \frac{\partial}{\partial y} L'_{u'_y} = 0.$$

Now we are back at our Tikhonov Denoising functional:

$$\min\left\{E = \frac{1}{2}\int_{\Omega} |\nabla u|^2 dx + \frac{\lambda}{2}\int_{\Omega} (u - f)^2 dx\right\}$$

Tikhonov Functional Derivation

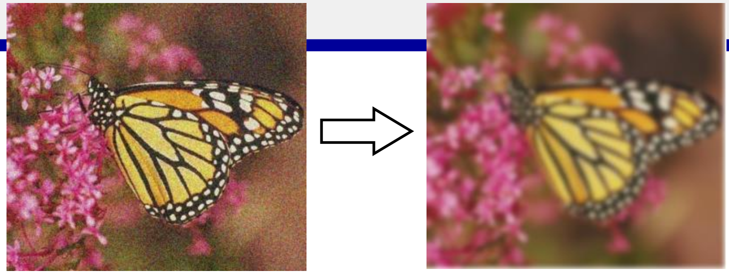
- Blackboard

- Euler-Lagrange Equation

$$-\Delta u + \lambda \left(u - f \right) = 0$$

Summary Tikhonov

$$\min\left\{E = \frac{1}{2}\int_{\Omega}|\nabla u|^{2}dx + \frac{\lambda}{2}\int_{\Omega}(u-f)^{2}dx\right\}$$



Example: *Tikhonov Denoising*

- Energy Functional
  - dependence on unknown function u (continuous domain)
- Calculus of Variations gives theorem to describe a functional at stationary points
  - Setting Functional (Gateaux) derivative to zero leads to the Euler-Lagrange PDE
  - The functional has to fulfill the Euler-Lagrange equation!

$$-\Delta u + \lambda(u-f) = 0$$

0	1	0
1	-4	1
0	1	0

Laplace Operator

Numerical Implementation Tikhonov

$$-\Delta u + \lambda\left(u - f\right) = 0$$

- Discretization necessary:
  Rather easy for quadratic Tikhonov model -> only Laplace operator
- Numerical Solver:
  - Gradient Descent Optimization: u=u(t)

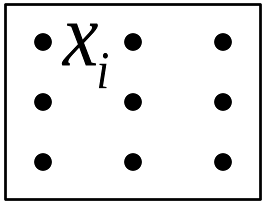
timestep

$$u^{t+1} = u^t - \tau\left(-\Delta u^t + \lambda(u^t - f)\right)$$

  - Direct (semi-implicit) Method: Huge equation system over all pixels to solve for solution u

$$Au(x_i) = \lambda f(x_i)$$

huge band matrix



$$i = 1 \ldots N$$

Example: Tikhonov Denoising

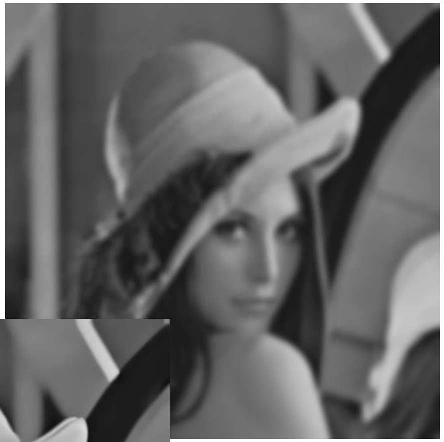
- What the heck is this all about?

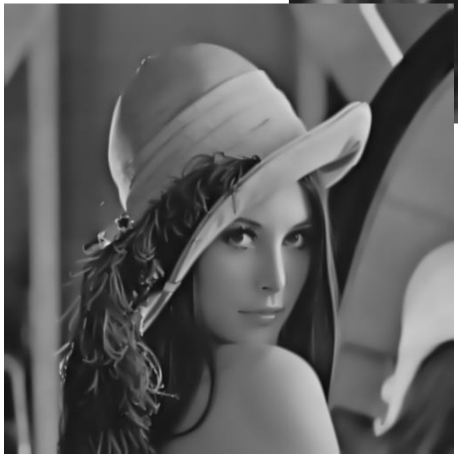
Quiz: Scale of Gauss? -> Matlab

- By solving the partial differential equation $-\Delta u + \lambda(u - f) = 0$ numerically, we implemented a Gauss blurring filter!
  - Excellent, but why is this interesting, people do this for decades?
  - *We now work in a mathematical framework for the analysis of inverse problems including their modeling, regularization & numerical implementation! (Variational Framework)*

Extensions of Tikhonov

- Literature proposed edge-preserving denoising methods using the quadratic L2-norm (Tikhonov) regularization
  - Bilateral Filtering (Tomasi-Manduchi)
  - Mean Shift Filtering (Comaniciu-Meer)
  - Anisotropic Diffusion (Perona-Malik, Weickert)
- These methods model smoothing dependent on image gradient, while Tikhonov (i.e. Gauss) ignores image gradient ∇f !





Extensions of Tikhonov

- Anisotropic Diffusion (Taxonomy of Weickert)
  - Generalize Quadratic Regularization

    $$\int_\Omega |\nabla u|^2 dx = \int_\Omega \nabla u^T \nabla u dx = \int_\Omega \nabla u^T D \nabla u dx$$

    $$D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$ D … Diffusion Tensor

  - Perona-Malik

    $$D = \begin{pmatrix} g(|\nabla f|^2) & 0 \\ 0 & g(|\nabla f|^2) \end{pmatrix}$$

  - Also Incorporate Gradient Orientation!

    $$D = \begin{pmatrix} d_{11}(\nabla f) & d_{12}(\nabla f) \\ d_{12}(\nabla f) & d_{22}(\nabla f) \end{pmatrix}$$

Alternative Extension: Total Variation Denoising

- Remember: Tikhonov used quadratic prior

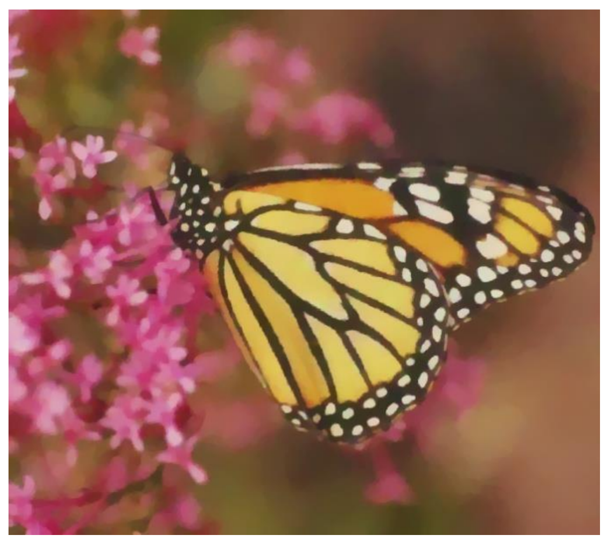
$$\min\left\{ E = \frac{1}{2}\int_{\Omega} |\nabla u|^2 dx + \frac{\lambda}{2}\int_{\Omega} (u-f)^2 dx \right\}$$

$$\Updownarrow$$

- Different, robust norm for prior?

$$\int_{\Omega} |\nabla u| dx = \int_{\Omega} \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2}\, dx$$

Total Variation Norm



Suddenly edges are preserved!

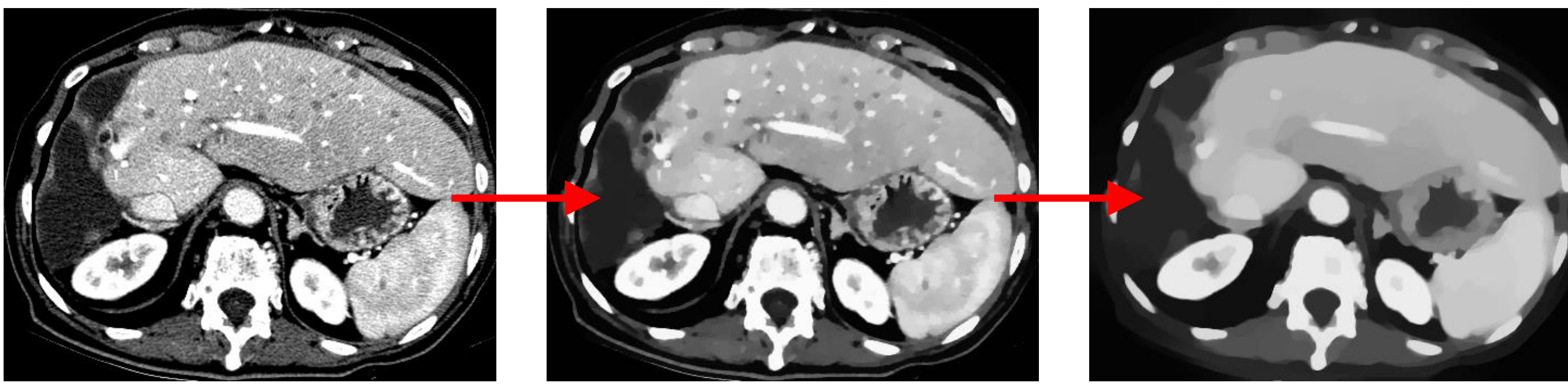
Quadratic prior does not allow sharp edges!

Total Variation Denoising

- Image denoising model introduced by Rudin, Osher and Fatemi in 1992 (a.k.a. ROF, TV-L2 model)

TV

$$\min_u \left\{ \int_\Omega |\nabla u|\, dx + \frac{\lambda}{2} \int_\Omega (u - f)^2\, dx \right\}$$



$\lambda = 10$ $\lambda = 1$

Quadratic vs. Total Variation

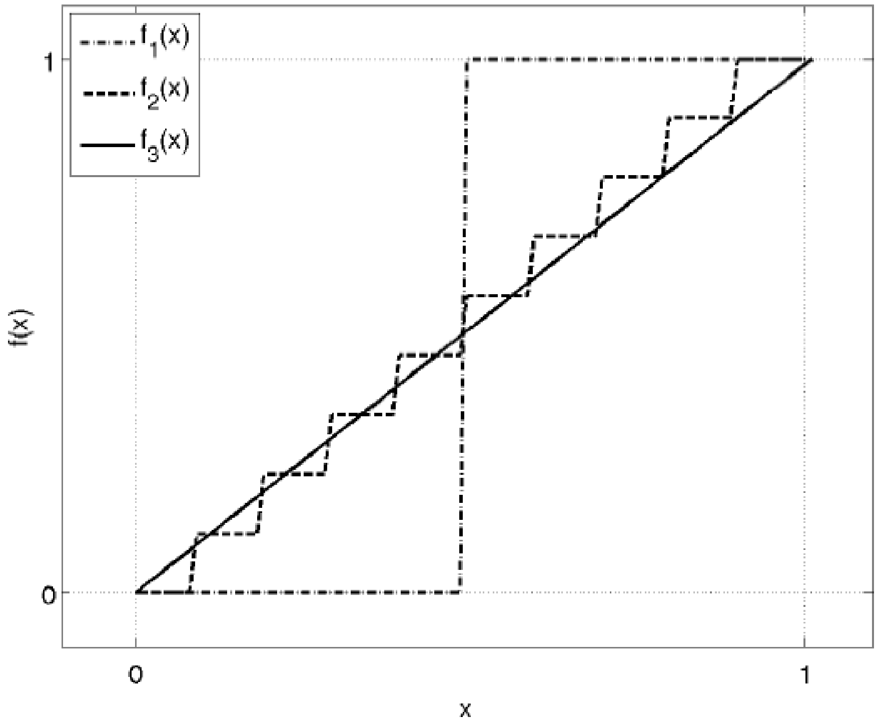


Figure 2.3: Total Variation does not see any difference between these three functions

Let‘s have a closer look at edges in f!

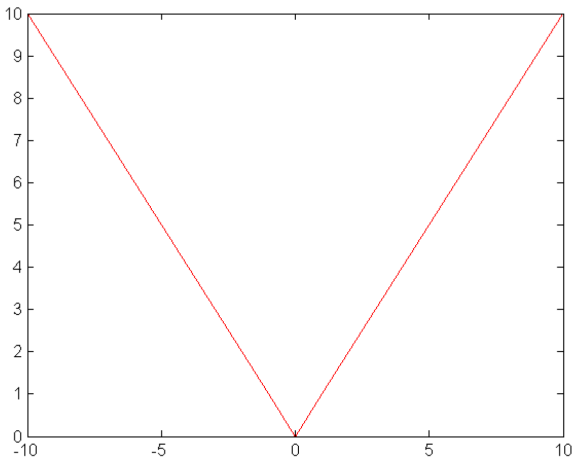
Sample f at 100 locations

Functions	Total Variation	Quadratic
$f_1(x)$	1.0	1.0
$f_2(x)$	1.0	0.11
$f_3(x)$	1.0	0.01

Minimizing quadratic norm favors f3!
Minimizing TV norm makes no distinction

Numerical Implementation - TV

- However, Total Variation (TV) model unfortunateley harder to minimize compared to quadratic!
  - Why? : Derivative undefined at zero!
  - Remember: Euler Lagrange eq. leads to derivative!
- Approaches
  - Slow Gradient Descent Methods
  - Sophisticated Primal-Dual Methods



Numerical Implementation - ROF

Minimize the following energy:

$$\min_u \left\{ \int_\Omega |\nabla u| dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 dx \right\}$$

Euler-Lagrange for $J(u) = \frac{1}{p} \iint_\Omega |\nabla u|^p dxdy, \; 1 \leq p < \infty,$

is

$$-\mathrm{div}\left(|\nabla u|^{p-2} \nabla u \right) = 0.$$

So: Associated Euler-Lagrange equation of our energy is:

$$-\nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda (u - f) = 0$$

Numerical Implementation - ROF

Explicit (Gradient Descent) Optimization:

$$u^{t+1} = u^t - \tau \left[-\nabla \cdot \left(\frac{\nabla u^t}{\sqrt{\left|\nabla u^t\right|^2 + \varepsilon}} \right) + \lambda (u^t - f) \right]$$

Choice of ε difficult & critical!

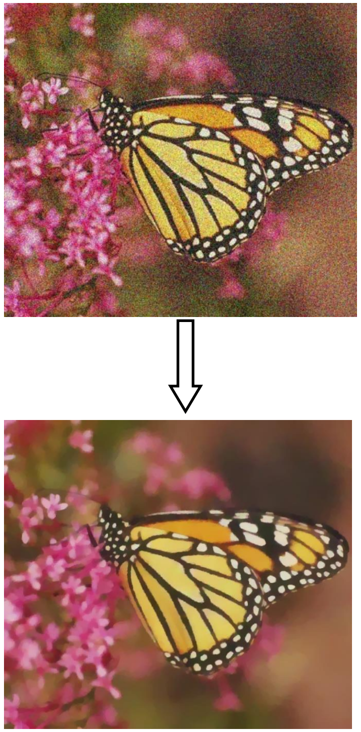
large: slow convergence, smooth over edges

small: divide by nearly zero (numerically unstable)

We call this solution: **ROF-primal**

ROF for Color Images

- Sophisticated models available combining RGB channels (e.g. vector TV)
- Simple:
  - Treat R,G,B planes separately
  - Three, uncoupled ROF steps & combine denoised RGB again



END

One minute paper:

a) What did I learn today?

b) Which topics remained open?

See you next week!