Direct Method in the Calculus of Variations *(basic definitions)*

Tadeusz Iwaniec (Syracuse University)

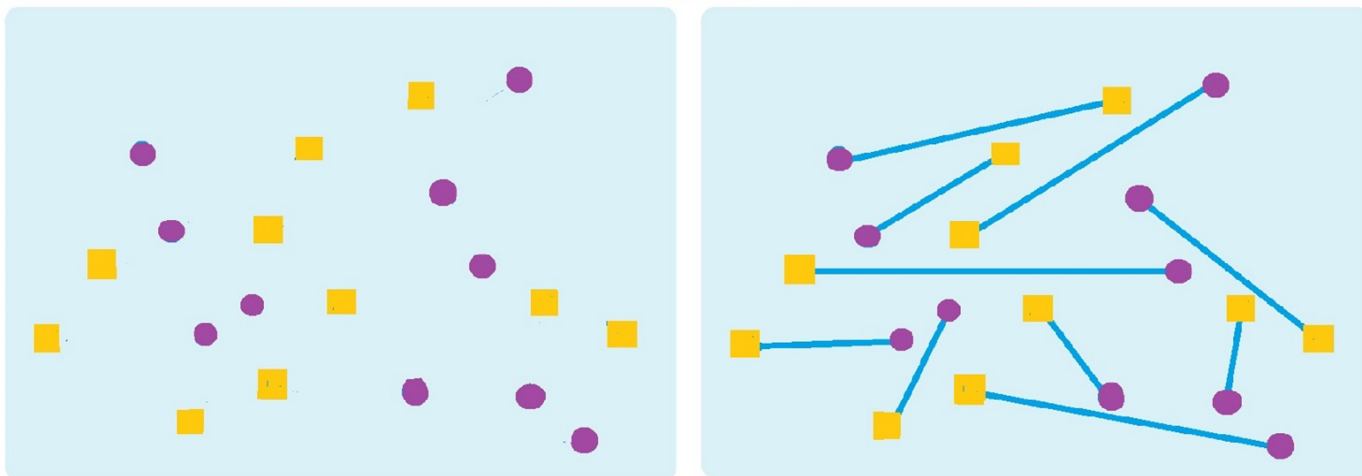
UAB Lecture 1

(Monday, July 16, 11:30am –12:20pm)

Why the Extremal Configurations are Essential?

Let me illustrate how PROBLEMS of EXISTENCE in elementary geometry can be solved by looking at the extremal configurations.

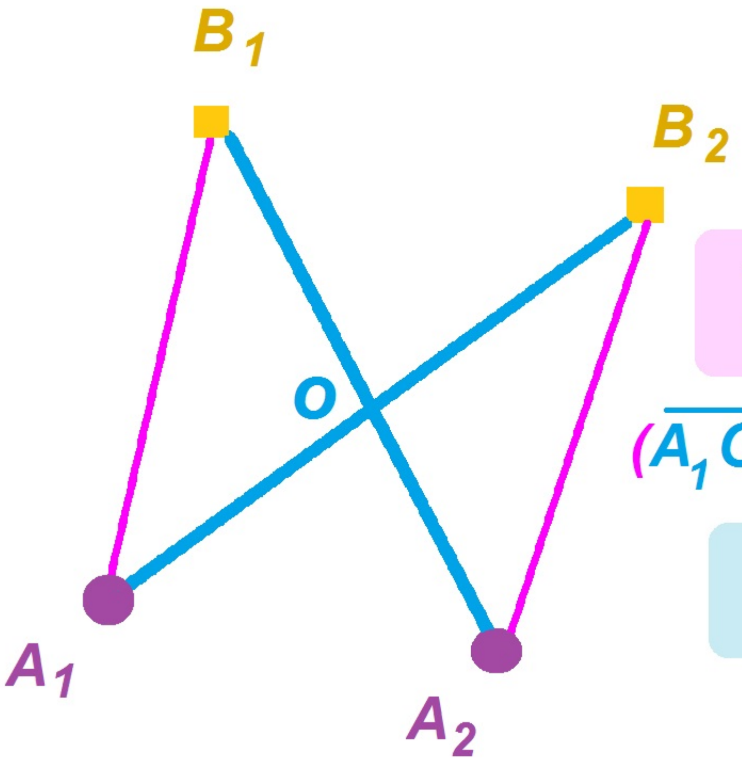
The shortest connections



A set of $2n$ points in which n are marked in purple and n in orange.

(None of three points of the set lay in the same line)

Connect the purple and yellow points pairwise by <u>disjoint</u> straight segments.



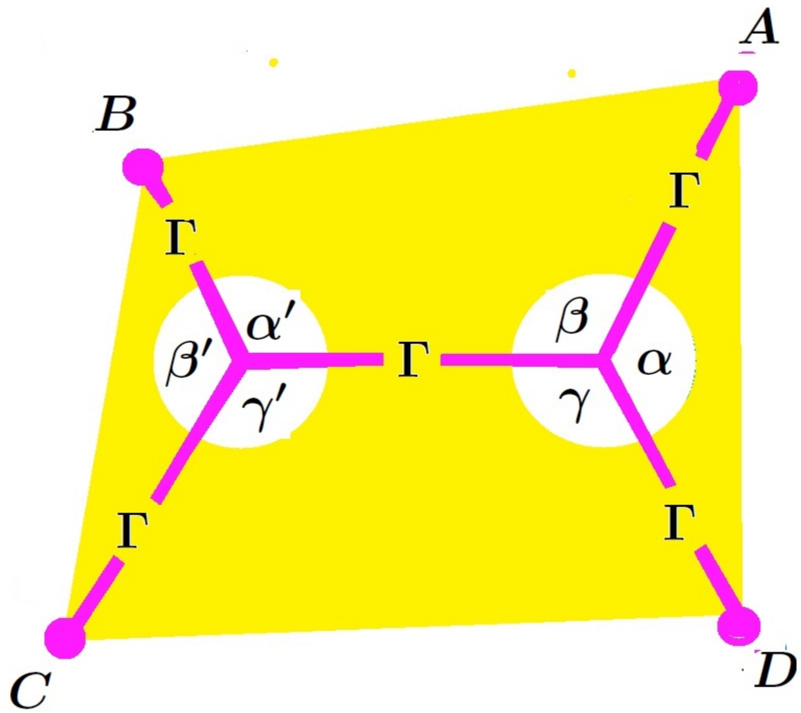
$$\overline{A_1B_1} + \overline{A_2B_2} <$$

$$(\overline{A_1O} + \overline{OB_1}) + (\overline{A_2O} + \overline{OB_2}) =$$

$$\overline{A_1B_2} + \overline{A_2B_1}$$

Among all purple-yellow connections by straight segments the one of smallest total length admits no intersection.

Homework (Minimal Continua)



Among all continua Γ , containing vertices of a quadrilateral $ABCD$, the one drawn here turns out to have smallest Hausdorf 1-dimensional measure.

Using a variational trick, show that:

$$\alpha = \beta = \gamma = \frac{2\,\pi}{3}$$

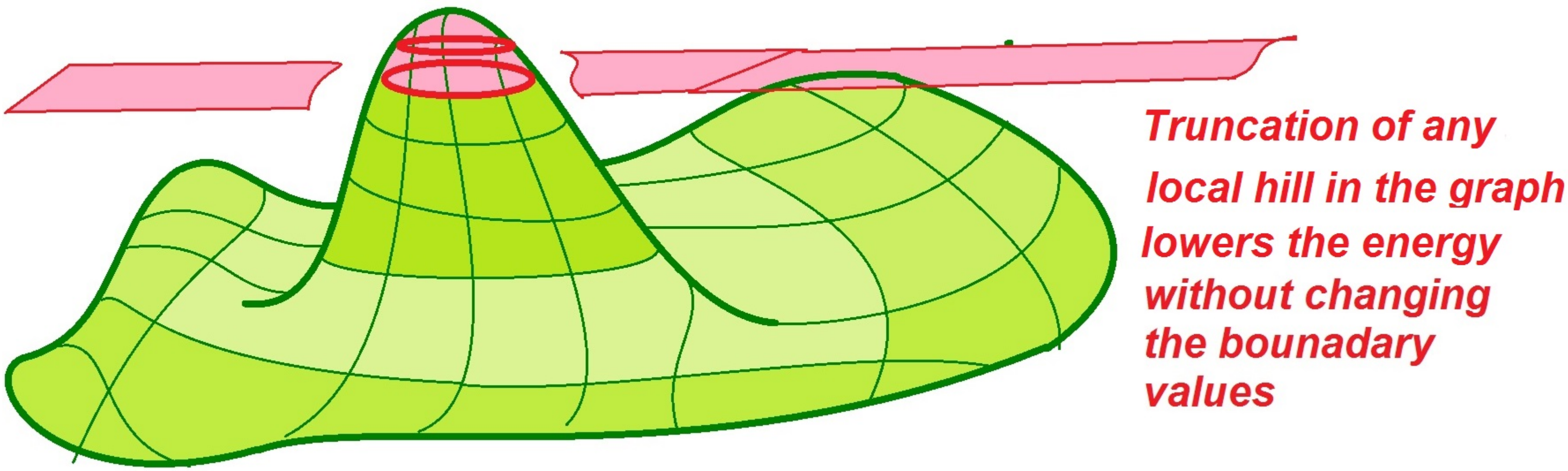
$$\alpha' = \beta' = \gamma' = \frac{2\,\pi}{3}$$

Why Limits of Sobolev Homeomorphisms ?

As we seek greater knowledge about the energy-minimal deformations the questions about Sobolev homeomorphisms and their limits (Monotone Sobolev Mappings) become ever more quintessential. In fact the weak and strong limits of planar Sobolev homeomorphisms are the same; and we shall approximate them with diffeomorphisms.

Why Generalized Solutions?

Generalized solutions (proposed by S. Zaremba at ICM in Rome, 1908) help us not only to solve the existence problems in PDEs but also to understand the regularity properties of the solutions (maximum principle for harmonic functions).



Nonlinear Hyperelasticity
(brief description)

Deformations $f : \mathbb{X} \xrightarrow{\text{onto}} \mathbb{Y}$ of smallest stored energy

$$\mathscr{E}[f] = \int_{\mathbb{X}} \mathbf{E}(x, f, Df)\, \mathrm{d}x$$

The given stored energy function

$$\mathbf{E} : \mathbb{X} \times \mathbb{Y} \times \mathbb{R}^{m \times n} \to \mathbb{R}$$

characterizes mechanical properties of the elastic material in the domains.

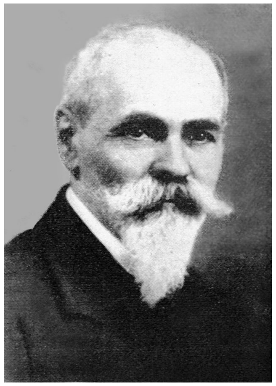
Are you hyperelastic enough to perform such gymnastic feats?

Direct Method in the Calculus of Variations

(Let us take a quick look at the Direct Method in the Calculus of Variation)

This is a general scheme for constructing a minimizer for a given functional. It was introduced by David Hilbert and Stanisław Zaremba around 1900 (*at the International Congress of Mathematicians in Rome in 1908*). The technique relies on methods of functional analysis and topology.


Hilbert


Zaremba


Tonelli

Stanisław Zaremba's early publications

[1] **S. Zaremba**, *Sur le principe de Dirichlet*, Atti del IV Congresso Internazionale dei Matematici (Roma, 6-11 Aprile 1908), vol.II, Communicazioni delle sezioni I e II, Roa 1909, 194-199.

[2] **S. Zaremba**, *Sur le principe du minimum*, Bulletin Internationale de l Academie des Sciences de Cracovie, Classe des Sciences Mathematiques et Naturelles, 1909,(7),197-264.[1]

[3] **S. Zaremba**, *Sur le principe de Dirichlet*, Acta Mathematica,34 (1911), 293-316.

[1] Zaremba introduced generalized solutions into the direct method of variational calculus (built up by Hilbert's ideas).

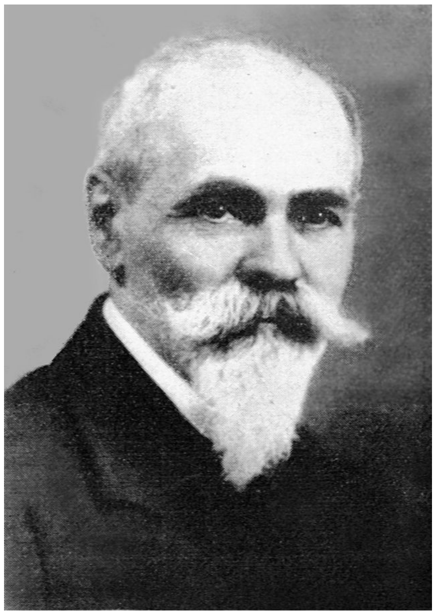
A Digression (following Andrzej Pelczar's comments)

Stanisław Zaremba was the first President of the *Polish Mathematical Society*; among names of founders we can find Stefan Banach. Mathematical activities of great significance took place in Kraków during the war, when the Jagiellonian University was acting as an ***underground university*** (being formally closed by German occupants since November 1939).

For instance, important results in PDEs were obtained during the war by Tadeusz Ważewski after his return from the Sachsenhausen concentration camp in 1940.

Tadeusz Ważewski (1896-1972), one of the outstanding pupils of Stanisław Zaremba (and his successor), created the school of differential equations called often: ***the Kraków School of Differential Equations***.

One can say that the academic society in Kraków was in that way fighting, without arms - against the occupants.



Grey-headed Stanisław Zaremba, however, did not live to the end of the Second World War. Luckily for him, he did not go through Soviet's dictatorship in Poland after the war.

General Scheme of the Direct Method

Given a nonlinear functional $\mathscr{E} : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ defined on a subset $\mathcal{H} \subset \mathfrak{B}$ of a separable reflexive Banach space $\mathfrak{B}$
Find $h_\circ \in \mathcal{H}$ such that;

$$\mathscr{E}[h_\circ] = \mathscr{E}_{\mathsf{inf}}(\mathcal{H}) := \inf_{h \in \mathcal{H}} \mathscr{E}[h] > -\infty \quad \big(\text{ this is an assumption on } \mathcal{H} \big)$$

Step 1. Show that the minimizing sequence $\{h_n\} \subset \mathcal{H}$, $\mathscr{E}[h_n] \to \mathscr{E}_{\mathsf{inf}}(\mathcal{H})$, admits a subsequence $h_{n_k} \rightsquigarrow h_\circ \in \mathcal{H}$ (usually converging in weak topology of $\mathfrak{B}$)

Step 2. Show that $\mathscr{E}$ is sequentially lower semi-continuous ; that is,

$$\mathscr{E}[h_*] \leqslant \liminf \mathscr{E}[h_n] \quad , \qquad \text{whenever } \; h_n \rightsquigarrow h_* \in \mathcal{H}.$$

Direct method in the Calculus of Variations

We shall discuss variational functionals whose prototype is the Dirichlet energy integral

$$\mathscr{E}[u] \; = \; \int_\Omega \Big| \, \nabla u(x) \, \Big|^2 \, \mathrm{d}x \;\; , \quad \Omega \subset \mathbb{R}^n$$

subject to functions with prescribed boundary values. The smallest complete linear metric space containing smooth functions of finite energy is none other than the familiar Sobolev space $\mathscr{W}^{1,2}(\Omega)$. Let $\mathscr{W}_\circ^{1,2}(\Omega)$ denote the subspace of $\mathscr{W}^{1,2}(\Omega)$ that consists of Sobolev functions vanishing on $\partial\Omega$. By definition, this is the closure of $\mathscr{C}_\circ^\infty(\Omega)$.

Given Dirichlet data $u_\circ \in \mathscr{W}^{1,2}(\Omega)$, we are looking for $u \in u_\circ + \mathscr{W}_\circ^{1,2}(\Omega)$ with minimal energy. The primary non-quadratic analogue of the Dirichlet integral, from which we learn a little about other convex variational problems, is the p-harmonic integral

$$\mathscr{E}_p[u] \;=\; \int_\Omega \Big|\, \nabla u(x) \,\Big|^p \,\mathrm{d}x\,, \qquad \text{for} \qquad u \in u_\circ + \mathscr{W}_\circ^{1,p}(\Omega),$$

It is worthwhile making a rather lengthy digression on the techniques involved. To prepare for somewhat greater generality, we formulate the minimization problem for Sobolev mappings $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, with prescribed boundary data $f_\circ \in \mathscr{W}^{1,p}(\Omega, \mathbb{R}^m)$.

$$\mathscr{E}[f] \;= \mathscr{E}_{\mathbf{E}}[f] \;=\; \int_\Omega \mathbf{E}(x, Df)\,\mathrm{d}x\,,\;\; f \in f_\circ + \mathscr{W}_\circ^{1,p}(\Omega, \mathbb{R}^m),$$

The variational problem about to be discussed can only be rigorously justified for integrands $\mathbf{E}(x,\xi)$ satisfying certain convexity and growth conditions. Later, these conditions will be subjected to an in-dept analysis of polyconvex functionals. Thus these are at present only formal considerations. The concerned reader may wish to think of $\mathbf{E} : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ as the p-harmonic integrand, $\mathbf{E}(x,\xi) = |\xi|^p$, with $\xi \in \mathbb{R}^{m\times n}$. The first requisite for the Direct Method is the *coercivity estimate*, which allows us to control the $\mathscr{W}^{1,p}$-norm of f by means of its energy and boundary data.

$$\|f\|_{\mathscr{W}^{1,p}(\Omega,\mathbb{R}^m)} \preccurlyeq \|f_\circ\|_{\mathscr{W}^{1,p}(\Omega,\mathbb{R}^m)} + \left[\,\mathscr{E}[f]\,\right]^{\frac{1}{p}}$$

Suppose next that $f_\kappa \in f_\circ + \mathscr{W}_\circ^{1,p}(\Omega, \mathbb{R}^m)$ is a minimizing sequence, meaning that

$$\inf \left\{ \int_\Omega |\, Dh \,|^p \,;\; h \in f_\circ + \mathscr{W}_\circ^{1,p}(\Omega, \mathbb{R}^m) \right\} = \lim_{\kappa \to \infty} \int_\Omega |\, Df_\kappa \,|^p$$

As $\mathscr{W}^{1,p}(\Omega, \mathbb{R}^m)$ is a reflexive Banach space we may extract from $\{f_\kappa\}$ a subsequence, again denoted by $\{f_\kappa\}$, converging weakly to a mapping $f \in f_\circ + \mathscr{W}_\circ^{1,p}(\Omega, \mathbb{R}^m)$. Now, everything hinges on showing that

$$\mathscr{E}[f] \leqslant \liminf_{\kappa \to \infty} \mathscr{E}[f_\kappa] \,, \tag{1}$$

whenever $f_\kappa \rightharpoonup f$, weakly in $\mathscr{W}^{1,p}(\Omega, \mathbb{R}^m)$. A customary terminology refers to such energy functional $\mathscr{E} : \mathscr{W}^{1,p}(\Omega, \mathbb{R}^m) \to \mathbb{R}$ as being

sequentially weakly lower semicontinuous.

For convenience we shall ignore the word sequentially.

That this is the case for the p-harmonic integral, $1 < p < \infty$, is quite straightforward.

$$
\begin{aligned}
\int_\Omega |Df|^p \quad &= \quad \lim_{\kappa\to\infty} \int_\Omega |Df|^{p-2}\langle\, Df \mid Df_\kappa \,\rangle \\
&\leqslant \liminf_{\kappa\to\infty} \int_\Omega |Df|^{p-1}|Df_\kappa| \\
&\leqslant \liminf_{\kappa\to\infty} \left(\int_\Omega |Df_\kappa|^p\right)^{\frac{1}{p}} \cdot \left(\int_\Omega |Df|^p\right)^{\frac{p-1}{p}}
\end{aligned}
$$

Hence

$$
\int_\Omega |Df|^p \leqslant \liminf_{\kappa\to\infty} \int_\Omega |Df_\kappa|^p
$$

There are several general methods for proving lower semicontinuity of a convex functional. When dealing with weak convergence in Banach spaces it is impossible to overlook the classical lemma of S. Mazur.

Let $\{f_\kappa\}$ *weakly converge to* f *in a Banach space* $\mathscr{B}$. *Then there exist* $0 \leqslant \lambda_1^\kappa, ..., \lambda_\kappa^\kappa \leqslant 1$, $\lambda_1^\kappa + ... + \lambda_\kappa^\kappa = 1$, *such that the convex combinations*

$$F_\kappa = \lambda_1^\kappa f_1 + \lambda_2^\kappa f_2 + \cdots + \lambda_\kappa^\kappa f_\kappa \, , \quad \textit{converge to } f \textit{ strongly in } \mathscr{B}.$$

From here it is very simple to deduce that the energy functional $\mathscr{E} = \mathscr{E}_{\mathbf{E}}$ is weakly lower semicontinuous on $\mathscr{W}^{1,p}(\Omega, \mathbb{R}^m)$ if the integrand satisfies the following conditions:

1) For almost every $x \in \Omega$ the function $\mathbf{E}(x, \;\;) : \mathbb{R}^{m \times n} \to \mathbb{R}$ is convex
2) Its gradient satisfies $|\mathbf{E}'(x, \xi)| \preccurlyeq 1 + |\xi|^{p-1}$, for all $\xi \in \mathbb{R}^{m \times n}$

In the proof that we do not write here, these conditions only serve to ensure that the $\mathscr{W}^{1,p}$ -norms make sense. Unfortunately, Mazur's lemma does not work for non convex functionals. The route to essential innovations goes through the so-called *subgradient inequality* :

$$\mathbf{E}(x,\xi) \,-\, \mathbf{E}(x,\xi_\circ) \;\;\geqslant\;\; \langle\, \mathbf{E}'(x,\xi_\circ) \mid \xi - \xi_\circ \,\rangle \tag{2}$$

It says that the graph of a convex function $\mathbf{E}(x, \quad) : \mathbb{R}^{m\times n} \to \mathbb{R}$ lies above its tangent hyperplane. Should $\mathbf{E}(x, \quad)$ fail to be differentiable at $\xi_\circ$, we may wish to put any subgradient at $\xi_\circ$ in place of $\mathbf{E}'(x,\xi_\circ)$,.

The lower semicontinuity becomes almost entirely trivial by applying this inequality to the differential matrices

$$\mathscr{E}[f_\kappa] \,-\, \mathscr{E}[f] \;\;\geqslant\;\; \int_\Omega \langle\, \mathbf{E}'(x,Df) \mid Df_\kappa - Df \,\rangle \, \mathrm{d}x \;\longrightarrow\; 0 \tag{3}$$

The Concept of Polyconvexity

(Morrey's quasiconvexity is practically impossible to verify)

In his mathematical models of Nonlinear Elasticity John Ball has made the crucial observation that, if the convexity of the stored energy integrand $\mathbf{E}(x, h, Dh)$ must be ruled out, it can be replaced by the weaker requirement; by expressing the integrand as a convex function of subdeterminants of the deformation gradient Dh.

$$\mathbf{E}(x, h, Dh) = \mathbf{E}^*(x, h, \text{subdeterminants of } Dh) \qquad \begin{pmatrix} \text{convex with respect} \\ \text{to subdeterminants} \end{pmatrix}$$

The linear combinations of subdeterminants of Dh turn out to be none other than ***First Order Null-Lagrangians.*** Above all, this is the only practical idea that offers substantially more than the convex calculus.

Null-Lagrangians *J.M. Ball, J.C. Currie, P.J. Olver*

The term null-Lagrangian refers to a differential n- form $\mathbf{N}(x, h, Dh)\mathrm{d}x$, defined for Sobolev mappings $h : \mathbb{X} \subset \mathbb{R}^n \to \mathbb{R}^m$, such that:

$$\int_{\mathbb{X}} \mathbf{N}(x, h_1, Dh_1)\mathrm{d}x = \int_{\mathbb{X}} \mathbf{N}(x, h_2, Dh_2)\mathrm{d}x \ , \quad \text{whenever } h_1 = h_2 \text{ on } \partial\mathbb{X} \ ,$$

The latter equation is understood in the sense of the inclusion; $h_1 - h_2 \in \mathscr{W}_0^{1,p}(\mathbb{X}, \mathbb{R}^m)$. For the Jacobian determinant, we have the identity

$$\int_{\mathbb{X}} J(x, h_1)\,\mathrm{d}x \ = \ \int_{\mathbb{X}} J(x, h_2)\,\mathrm{d}x \ , \text{ whenever } h_1 - h_2 \in \mathscr{W}_0^{1,n}(\mathbb{X}, \mathbb{R}^n)$$

Note that the Lagrange-Euler equation is identically satisfied, hence the name null-Lagrangian. Adding null-Lagrangian to an energy integral does not change the Lagrange-Euler equation.

The Differential Matrix and its Subdeterminants form a basis for null-Lagrangians

We shall study mappings $f = (f^1, f^2, ..., f^m) : \mathbb{X} \longrightarrow \mathbb{R}^m$ defined on an open region $\mathbb{X} \subset \mathbb{R}^n$ and valued in an open region $\mathbb{Y} \subset \mathbb{R}^m$, which belong to the Sobolev class $\mathscr{W}^{1,p}(\mathbb{X}, \mathbb{R}^m)$. Their differential, also called the *gradient matrix* (in the theory of elasticity we call it deformation gradient), consists of the first order partial derivatives of the coordinate functions

$$Df(x) = \begin{bmatrix} \frac{\partial f^1}{\partial x_1} & \frac{\partial f^1}{\partial x_2} & \cdots & \frac{\partial f^1}{\partial x_n} \\ \\ \frac{\partial f^2}{\partial x_1} & \frac{\partial f^2}{\partial x_2} & \cdots & \frac{\partial f^2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f^m}{\partial x_1} & \frac{\partial f^m}{\partial x_2} & \cdots & \frac{\partial f^m}{\partial x_n} \end{bmatrix} = \left[\frac{\partial f^i}{\partial x_j}\right]_{j=1,\dots,n}^{i=1,\dots,m} \tag{4}$$

The determinant of the $n \times n$ matrix is called *Jacobian* of f, after Carl Gustav Jacob Jacobi (1804-1851), a prominent mathematical figure who studied various nonlinear differential expressions. We reserve several different symbols to denote the Jacobian, most common are:

$$\mathbf{J}(x,f) = \det\,[\,Df(x)\,] \;=\; \frac{\partial(f^1,\, f^2, \dots,\, f^n)}{\partial(x_1,\, x_2,\, \dots, x_n)}$$

$$= \; \mathbf{J}_{_f}(x) \; = \mathbf{J}(f^1,\, f^2, \dots, f^n)$$

In any instance we shall take whichever notation seems temporarily most suitable. We shall make use of continuity of subdeterminants in the weak topology of $\mathscr{W}^{1,\mathfrak{s}}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$. To appreciate the best out of it, let us mention two well known results.

THEOREM. **[Weak continuity of the Jacobians]**

Suppose that the mappings $f_\nu : \Omega \to \mathbb{R}^n$, $\nu = 1, 2, \ldots,$ converge to f weakly in $\mathscr{W}^{1,n}(\Omega, \mathbb{R}^n)$. Then for every test function $\varphi \in \mathscr{C}_\circ^\infty(\Omega)$,

$$\lim_{\nu \to \infty} \int_\Omega \varphi(x)\, \mathbf{J}(x, f_\nu)\, \mathrm{d}x = \int_\Omega \varphi(x)\, \mathbf{J}(x, f)\, \mathrm{d}x$$

$$\left| \int_\Omega \varphi(x) \left[\, \mathbf{J}(x, f_\nu) \; - \mathbf{J}(x, f)\, \right] \mathrm{d}x \right| \leqslant$$

$$\|\nabla \varphi\|_{\mathscr{L}^\infty(\Omega)} \, \|f_\nu - f\|_{\mathscr{L}^n(\Omega)} \left(\|Df_\nu\|_{\mathscr{L}^n(\Omega)} \; + \; \|Df\|_{\mathscr{L}^n(\Omega)} \right)^{n-1}$$

More generally, for every pair of ordered ℓ-tuples $I;\ 1 \leqslant i_1 < i_2 < \ldots < i_\ell \leqslant m$ and $J;\ 1 \leqslant j_1 < j_2 < \ldots < j_\ell \leqslant n$, weak continuity of Jacobian subdeterminants reads as:

$$\lim_{\nu\to\infty} \int_\Omega \frac{\partial f_\nu^I}{\partial x_J}\ \varphi(x)\ \mathrm{d}x = \int_\Omega \frac{\partial f^I}{\partial x_J}\ \varphi(x)\ \mathrm{d}x$$

provided the mappings $f_\nu : \Omega \longrightarrow \mathbb{R}^m$ converge to f weakly in $\mathscr{W}^{1,\ell}(\Omega, \mathbb{R}^m)$, where $1 \leq \ell \leqslant \min\{m, n\}$. Here we have used the notation for subdeterminants

$$\frac{\partial f_\nu^I}{\partial x_J} = \frac{\partial(f_\nu^{i_1}, \ldots, f_\nu^{i_\ell})}{\partial(x_{j_1}, \ldots, x_{j_\ell})} \qquad \text{and} \qquad \frac{\partial f^I}{\partial x_J} = \frac{\partial(f^{i_1}, \ldots, f^{i_\ell})}{\partial(x_{j_1}, \ldots, x_{j_\ell})}$$

It is instructive to compare this situation with the familiar case of $\ell = 1$. For $i = 1, ..., m$ and $j = 1, ..., n$ fixed, we have

$$\lim_{\nu \to \infty} \int_\Omega \frac{\partial f_\nu^i}{\partial x_j} \varphi = - \lim_{\nu \to \infty} \int_\Omega \frac{\partial \varphi}{\partial x_j} f_\nu^i = - \int_\Omega \frac{\partial \varphi}{\partial x_j} f^i = \int_\Omega \frac{\partial f^i}{\partial x_j} \varphi$$

Given any matrix $M = [\, M^i_j \,] \in \mathbb{R}^{m \times n}$, we denote by $M_\boxplus$ the ordered list of all subdeterminants of M. This includes the number 1 as 0×0-minor and the entries of M as 1×1-minors. The highest order subdeterminants considered are the $\ell \times \ell$-minors, $\ell =$ min$\{m, n\}$. $M_\boxplus$ will be viewed as a point in the Euclidean space , in which one can speak of convex sets, $M_\boxplus \in \mathbb{R}^{\binom{m+n}{n}}$.

HOMEWORK. Using elementary combinatorics show that the number of all subdeterminants is in fact equal to:

$$\sum_{\ell \geqslant 0} \binom{m}{\ell}\binom{n}{\ell} = \binom{m+n}{n} = \frac{(m+n)!}{m!\ n!}. \tag{5}$$

Accordingly, to every pair of ordered ℓ-tuples $I:\ 1 \leqslant i_1 <, ..., < i_\ell \leqslant m$ and $J:\ 1 \leqslant j_1 <, ..., < j_\ell \leqslant n$, where $0 \leqslant \ell \leqslant \min\{m, n\}$, there corresponds a function, also called $[I, J]$-subdeterminant

$$\mathbf{M}^I_J :\ \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}, \qquad \mathbf{M}^I_J(X) \ = \ \det X^I_J \ = \begin{vmatrix} X^{i_1}_{j_1} & \cdots & X^{i_1}_{j_\ell} \\ \cdots & \cdots & \cdots \\ X^{i_\ell}_{j_1} & \cdots & X^{i_\ell}_{j_\ell} \end{vmatrix}$$

We adhere to the convention that the function $\mathbf{M}^I_J(X) \equiv 1$, if $\ell = 0$. The following definition provides us with a complete algebraic description of the first order null Lagrangians, of which we shall make repeated use throughout these lectures.

First Order Null Lagrangians

THEOREM. Null Lagrangians $\mathbf{N} : \mathbb{R}^{m\times n} \to \mathbb{R}$ form a linear space of dimension $\binom{m+n}{n}$. Its basis is furnished by the $\mathbf{M}^I_J$ -functions.

$$\mathbf{N}(Df) \;=\; \sum_{\ell\,=\,0}^{\min\{m,n\}} \;\; \sum_{\substack{1\leqslant i_1<\ldots<i_\ell\leqslant m \\ 1\leqslant j_1<\ldots<j_\ell\leqslant n}} \lambda^{j_1\ldots j_\ell}_{i_1\ldots i_\ell} \; \frac{\partial(f^{i_1},\ldots,f^{i_\ell})}{\partial(x_{j_1},\ldots,x_{j_\ell})}$$

$$\int_{\mathbb{X}} \mathbf{N}(Df)\,\mathrm{d}x = \int_{\mathbb{X}} \mathbf{N}(Dh)\,\mathrm{d}x \;, \quad \text{whenever } f_{|\partial\mathbb{X}} = h_{|\partial\mathbb{X}}$$

Definition of Polyconvexity

The energy integral

$$\mathscr{E}[f] = \int_{\mathbb{X}} \mathbf{E}(x, f, Df)\, \mathrm{d}x \,, \qquad f : \mathbb{X} \to \mathbb{Y} \subset \mathbb{R}^m \quad , \quad \mathbb{X} \subset \mathbb{R}^n$$

is polyconvex if $\mathbf{E} : \mathbb{X} \times \mathbb{Y} \times \mathbb{R}^{m \times n} \to \mathbb{R}$ **can be written as**

$$\mathbf{E}(x, y, M) \;=\; \Xi(x, y, M_{\boxplus})$$

where, for every $x \in \mathbb{X}$ and $y \in \mathbb{Y}$ **, the function** $\Xi(x, y, \quad) : \mathbb{R}^{\binom{m+n}{n}} \to \mathbb{R}$ **is convex.**

Subgradient Inequality

Let $\mathbf{E}(x,y,M) \;=\; \Xi(x,y,M_{\boxplus})$ be polyconvex and $A \in \mathbb{R}^{m\times n}$. Then for every matrix $M \in \mathbb{R}^{m\times n}$, we have

$$\mathbf{E}(x,y,M) \;-\; \mathbf{E}(x,y,A) \;\; \geqslant \sum_{I,J} \mathrm{E}^I_J(x,y,A)\left[M^I_J \;-\; \mathbb{A}^I_J\right]$$

where the E^I_J -coefficients are functions in $(x,y,A) \in \mathbb{X}\times\mathbb{Y}\times\mathbb{R}^{m\times n}$ only; they do not depend on the matrix $M \in \mathbb{R}^{m\times n}$. Precisely, we have

$$\mathrm{E}^I_J(x,y,A) = \frac{\partial\, \Xi(x,y,A_{\boxplus})}{\partial\, A^I_J}$$

Lower Semicontinuity

Consider a polyconvex energy integral

$$\mathscr{E}[f] = \int_{\mathbb{X}} \mathbf{E}(x, f, Df)\, \mathrm{d}x \;, \qquad f : \mathbb{X} \to \mathbb{Y} \subset \mathbb{R}^m \;\;, \quad \mathbb{X} \subset \mathbb{R}^n$$

defined in a class of Sobolev mappings $f = (f^1, \dots f^m) \, : \mathbb{X} \to \mathbb{Y} \subset \mathbb{R}^m$ of finite energy.

Suppose that, within this class, we are given a sequence of Sobolev mappings $f_\nu = (f_\nu^1, \dots, f_\nu^m) \, : \mathbb{X} \to \mathbb{R}^m \, , \nu = 1, 2, \dots$ converging weakly to $f = (f^1, \dots, f^m) \, : \mathbb{X} \to \overline{\mathbb{Y}} \subset \mathbb{R}^m$. Then

$$\mathscr{E}[f_\nu] \; - \; \mathscr{E}[f] \; \geqslant$$

$$\sum_{\ell=0}^{\min\{m,n\}} \sum_{\substack{1\leqslant i_1 < \ldots < i_\ell \leqslant m \\ 1 \leqslant j_1 < \ldots < j_\ell \leqslant n}} \int_{\mathbb{X}} \lambda^{j_1 \ldots j_\ell}_{i_1 \ldots i_\ell} \left[\frac{\partial(f_\nu^{i_1}, \ldots, f_\nu^{i_\ell})}{\partial(x_{j_1}, \ldots, x_{j_\ell})} - \frac{\partial(f^{i_1}, \ldots, f^{i_\ell})}{\partial(x_{j_1}, \ldots, x_{j_\ell})} \right]$$

$$\to 0$$

because the coefficients $\lambda^{j_1 \ldots j_\ell}_{i_1 \ldots i_\ell}(x)$ are independent of the sequence f_ν. Thus

$$\mathscr{E}[f] \; \leqslant \; \liminf \mathscr{E}[f_\nu]$$

Hessian in Three Dimensions

Guessing a differential identity that might lead to the definition of a weak Hessian in three and higher dimensions can be very sophisticated. The readers patient with lengthy though elementary computation may wish to verify the following formula

$$3\,\mathcal{H}u \;=\; 3\begin{vmatrix} u_{xx} & u_{xy} & u_{xz} \\ u_{yx} & u_{yy} & u_{yz} \\ u_{zx} & u_{zy} & u_{zz} \end{vmatrix} \;=$$

$$\left(\begin{vmatrix} u_{yy} & u_{yz} \\ u_{zy} & u_{zz} \end{vmatrix} u\right)_{xx} + \left(\begin{vmatrix} u_{xx} & u_{xz} \\ u_{zx} & u_{zz} \end{vmatrix} u\right)_{yy} + \left(\begin{vmatrix} u_{xx} & u_{xy} \\ u_{yx} & u_{yy} \end{vmatrix} u\right)_{zz} -$$

$$2\left(\begin{vmatrix} u_{xy} & u_{xz} \\ u_{zy} & u_{zz} \end{vmatrix} u\right)_{xy} - 2\left(\begin{vmatrix} u_{yz} & u_{yx} \\ u_{xz} & u_{xx} \end{vmatrix} u\right)_{yz} - 2\left(\begin{vmatrix} u_{zx} & u_{zy} \\ u_{yx} & u_{yy} \end{vmatrix} u\right)_{zx}$$

After integration by parts (twice), we arrive at an integrand which depends linearly on u and quadratically on the second order derivatives of u. By contrast, the point-wise Hessian is a cubic polynomial in $\nabla^2 u$. Take notice that the integral expression of this formula will define a distribution in $\mathscr{D}_2'(\Omega)$, for $u \in \mathscr{W}_{\mathrm{loc}}^{2,\,2}(\Omega) \subset \mathscr{C}_{\mathrm{loc}}^{1/2}(\Omega)$. Call it *second order Hessian*. Somewhat more sophisticated computation leads to the integrands depending linearly on $\nabla^2 u$. But this can be done only at the expense of producing quadratic terms with respect to ∇u. We yield to curiosity and mention the following formula for the *very weak Hessian* in three dimensions.

$$-6\,\mathcal{H}u \;=\; -6\begin{vmatrix} u_{xx} & u_{xy} & u_{xz} \\ u_{yx} & u_{yy} & u_{yz} \\ u_{zx} & u_{zy} & u_{zz} \end{vmatrix} \;= \tag{6}$$

$$\Big(\,u_y^2\,u_{zz} \;+\; u_z^2\,u_{yy} \;-\; 2\,u_y u_z\,u_{yz}\,\Big)_{xx} \;+$$

$$\Big(\,u_z^2\,u_{xx} \;+\; u_x^2\,u_{zz} \;-\; 2\,u_z u_x\,u_{zx}\,\Big)_{yy} \;+$$

$$\Big(u_x^2\, u_{yy} \;+\; u_y^2\, u_{xx} \;-\; 2\, u_x u_y\, u_{xy} \Big)_{zz} \;+$$

$$2\Big(u_z u_x\, u_{yz} \;+\; u_z u_y\, u_{zx} \;-\; u_x u_y\, u_{zz} \;-\; u_z^2\, u_{xy} \Big)_{xy} \;+$$

$$2\Big(u_x u_z\, u_{xy} \;+\; u_x u_y\, u_{zx} \;-\; u_y u_z\, u_{xx} \;-\; u_x^2\, u_{yz} \Big)_{yz} \;+$$

$$2\Big(u_y u_x\, u_{yz} \;+\; u_y u_z\, u_{xy} \;-\; u_z u_x\, u_{yy} \;-\; u_y^2\, u_{zx} \Big)_{zx}$$

It follows from the imbedding theorem that the very weak Hessian is well defined for $u \in \mathscr{W}^{2,\,9/5}_{\mathrm{loc}}(\Omega) \subset \mathscr{W}^{1,\,9/2}_{\mathrm{loc}}(\Omega)$, where we have Hölder's relation $\frac{5}{9} + \frac{2}{9} + \frac{2}{9} = 1$. One may ask at this stage if some amount of juggling with the integration by parts would result in a complete absence of the second gradient of u, as for the Hessian in dimension two.

Finding differential identities (like the above for Hessian) seems to be a small job. However,

If you are too big for small jobs,
you are too small for big ones.

See you tomorrow at 10:00 am