Lecture 1

Constructing variational principles:

Euler equations, Lagrangian and a general methodology

Fulvio Zonca

http://www.afs.enea.it/zonca

ENEA C.R. Frascati, C.P. 65 - 00044 - Frascati, Italy.

Institute for Fusion Theory and Simulation, Zhejiang University, Hangzhou 310027, P.R.C.

May 7.th, 2018

Physics of Alfvén waves and energetic particles in fusion plasmas
– Mode structures and global dispersion of Alfvén waves in tokamaks –
May 7 – 21, 2018, IFTS – ZJU, Hangzhou

Lecture Plan and Teaching Material

Abstract: This *Introduction to Physics of Alfvén waves and energetic particles in fusion plasmas* is devoted to *Mode structures and global dispersion of Alfvén waves in tokamaks* and will be lectured by Prof. Fulvio Zonca in Spring 2018. These Lectures will cover with pedagogical style the first part of *Reviews of Modern Physics* **88**, *015008 (2016)* and are a continuation of the Intensive Course on *Linear physics and stability in tokamaks* taught in Spring 2017. Prerequisite of the course is a preliminary knowledge of general plasma physics and classical mechanics, and the Lecture Notes of the 2017 Intensive Course. General theoretical aspects are treated in Lectures ($1 \div 7$), which are complemented by corresponding Q&A Sessions.

Main areas that will be explored are:

(I) Variational methods in plasma physics

i) Constructing variational principles: Euler equations and Lagrangian. A general methodology. (Lecture 1)

ii) Examples and applications of variational principles:
Power dissipation in a passive resistive network.
Energy eigenvalues in quantum mechanics. (Lecture 2)

iii) Shear Alfvén waves in tokamaks.
Variational formulation and variational principle.
Integral form of the dispersion relation. (Lecture 3)

(II) Wave propagation in slowly varying, weakly non-uniform media

iv) Eikonal formulation of wave equations.
Wave-packet amplitude evolution equation. (Lecture 4)

v) The wave kinetic equation.
Bound states and Schrödinger equation. (Lecture 5)

(III) Description of Alfvén waves in tokamaks

vi) Global eigenvalue problem in toroidal plasmas.
The formation of meso-scales. (Lecture 6)

vii) Time asymptotic solution and WKB dispersion relation (Lecture 7)

Lecture Notes: Available in electronic form. At the end of each lecture and during the Q&A time, a list of specific reading material is given explicitly. Should you have difficulty in finding literatures and papers, please contact me at fulvio.zonca@enea.it

Q&A Sessions: Devoted to detailed discussions and analysis of technical aspects. They are meant to be complementary to general lectures and are a necessary addendum for full appreciation of the material presented in the course.

Time and Location: Lectures (9am-11am) will be held at IFTS on May 7-9-11-14-16-18-21, with corresponding Q&A Sessions (3pm-5pm) of the same days.

Review of basic concepts

- Variable quantities called functionals play an important role in many problems arising in analysis, mechanics, geometry, etc. [Gelfand & Fomin 1963].
- By a functional, one means a correspondence which assigns a definite (generally complex) number to each function belonging to some class.
- Thus, one might say that a functional is a kind of function, where the independent variable is itself a function.
- Typically, a functional involves integration over the domain of the considered class of functions.
- The most developed branch of the calculus of functionals is concerned with finding the maxima and minima of functionals, and is called the calculus of variations.
- Typical examples of variational problems, that is of problems involving the determination of maxima and minima of functionals, are given below.

Some examples of historic significance

- Find the shortest plane curve joining two points A and B, i.e.,find the curve $y = y(x)$ for which the functional

$$J[y] = \int_a^b \sqrt{1 + y'^2} dx$$

achieves its minimum. The curve in question turns out to be the straight line segment joining A and B.

- This is a first example of a broad class of variational problems, which involve functionals that can be written in the general form

$$J[y] = \int_a^b F(x, y, y') dx \ .$$

- Another important problem is that of the brachistochrone, which is the curve along which a particle subject to gravity only takes the least time to go from A to B, with A and B two fixed points.
  - The brachistochrone problem was posed by John Bernoulli in 1696, and played an important part in the development of the calculus of variations.
  - The problem was solved by John Bernoulli, James Bernoulli, Newton, and L'Hospital.
  - The brachistochrone turns out to be a cycloid, lying in the vertical plane and passing through A and B.

- Last example given here is the isoperimetric problem, a variational problem solved by Euler.

  $\Rightarrow$ Among all closed curves of a given length L, find the curve enclosing the greatest area.

  $\Rightarrow$ The required curve turns out to be a circle.

The method of finite differences

- The method of finite differences is widely used in numerical analysis and is typically connected with the weak formulation of the problem of interest, which can be seen as the construction of the integral formulation of the same problem and, thus, as the formulation of a proper variational principle.
- The method of finite differences was introduced by Euler, who wanted to illuminate the fundamental connection of the problems of the calculus of variations with those of classical analysis, that is of the calculus of functions of n variables.
- Consider again the variational problem generated by a functional in the form

$$J[y] = \int_a^b F(x, y, y')dx$$

Here, assume that $y(a) = A$ and $y(b) = B$ are fixed.

- Consider the points

$$x_0 = a, \quad x_1, \ldots, \quad x_n, \quad x_{n+1} = b \quad ,$$

and divide the interval $[a, b]$ in $n + 1$ equal parts.

- Now replace the curve $y = y(x)$ with the polygonal line with vertices

$$(x_0, A), \quad (x_1, y_1), \ldots, \quad (x_n, y_n), \quad (x_{n+1}, B) \quad ,$$

with $y_i = y(x_i)$.

- It is then possible to approximate the functional $J[y]$ by the sum

$$J(y_1, \ldots, y_n) = \sum_{i=1}^{n+1} F\left(x_i, y_i, \frac{y_i - y_{i-1}}{h}\right) h \quad ,$$

with $h = x_i - x_{i-1}$.

- Since $y(a) = A$ and $y(b) = B$ are fixed, each polygonal line is uniquely identified by the values $y_1, ..., y_n$. Thus, as an approximation, we can regard the variational problem as the problem of finding the extrema of the function $J(y_1, ..., y_n)$.
- By replacing smooth curves by polygonal lines, Euler reduced the problem of finding extrema of a functional to the problem of finding extrema of a function of n variables: Then, in the limit of $n \to \infty$, he was able to obtain exact solutions.
- In this sense, functionals can b.e regarded as functions of infinitely many variables [i.e., the values of the function $y(x)$ at separate points], and the calculus of variations can be regarded as the corresponding analog of differential calculus.

Brief introduction to function spaces

- In the study of functions of n variables, it is convenient to use geometric language, by regarding a set of n numbers $(y_1, ..., y_n)$ as a point in an n-dimensional space.
- In just the same way, geometric language is useful when studying functionals. Thus, we can regard each function $y(x)$ belonging to some class as a point in some space, and spaces whose elements are functions may be called function spaces.
- The concept of continuity plays an important role for functionals, just as it does for the ordinary functions considered in classical analysis.
- In order to formulate this concept for functionals, we must somehow introduce a concept of closeness for elements in a function space. The concept of the norm of a function, analogous to the concept of the distance between a point in Euclidean space and the origin of coordinates, is introduced in the standard and abstract form, using the concept of a normed linear space.

- The concept of a linear space $\Re$ as set of elements $x, y, z, \ldots$ obeying the following axioms, with α, β real numbers

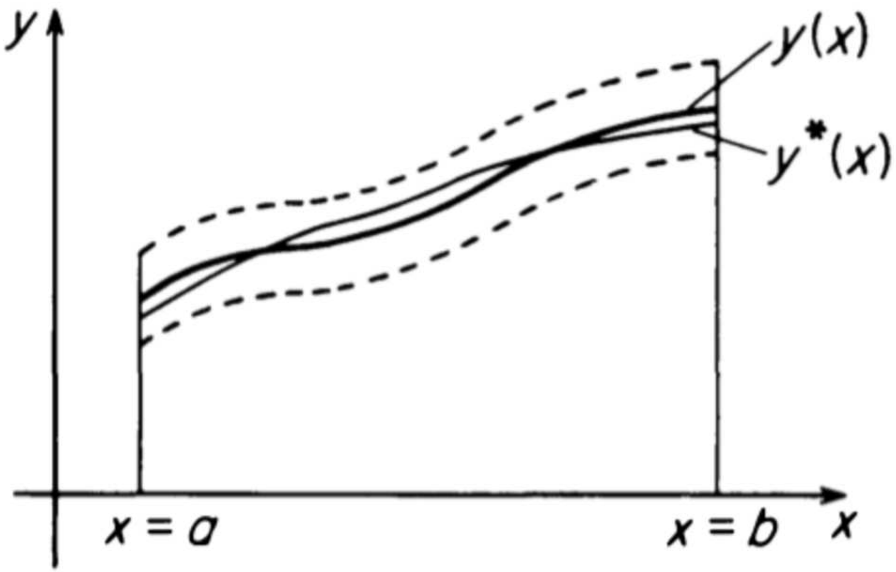
1. $x + y = y + x$;
2. $(x + y) + z = x + (y + z)$;
3. There exists an element 0 (the *zero element*) such that $x + 0 = x$ for any $x \in \mathscr{R}$;[4]
4. For each $x \in \mathscr{R}$, there exists an element $-x$ such that $x + (-x) = 0$;
5. $1 \cdot x = x$;
6. $\alpha(\beta x) = (\alpha\beta)x$;
7. $(\alpha + \beta)x = \alpha x + \beta x$;
8. $\alpha(x + y) = \alpha x + \alpha y$.

- The concept of a normed linear space $\Re$ (note: often an inner product is defined, from which the norm follows; with the addition of completeness this becomes a Hilbert space)

A linear space $\mathscr{R}$ is said to be *normed*, if each element $x \in \mathscr{R}$ is assigned a nonnegative number $\|x\|$, called the *norm* of x, such that

1. $\|x\| = 0$ **if and only if** $x = 0$;
2. $\|\alpha x\| = |\alpha| \, \|x\|$;
3. $\|x + y\| \leqslant \|x\| + \|y\|$.

- Define $\mathcal{C}(a, b)$ as the space of all continuous functions $y(x)$ defined on a (closed) interval $[a, b]$.



- Two functions are close if

$$||y - y^*|| < \epsilon \ ;$$

i.e., $|y(x) - y^*(x)| < \epsilon \ , \forall x \in [a, b]$.

- The same approach is extended by the definition of $\mathcal{D}_1(a,b)$ as the space of all continuous functions $y(x)$ with continuous first derivative defined on a (closed) interval $[a,b]$. The norm, in this case, is

$$||y-z|| = \max_{a\leq x\leq b} |y(x)-z(x)| + \max_{a\leq x\leq b} |y'(x)-z'(x)| \; .$$

  Thus,

$$||y-z|| < \epsilon \quad \Rightarrow |y(x)-z(x)| < \epsilon \;\; , |y'(x)-z'(x)| < \epsilon \;\; \forall x \in [a,b] \; .$$

  This means that two functions in $\mathcal{D}_1(a,b)$ are close if both the function and its first derivative are close over the whole domain.

- In general one can introduce the definition $\mathcal{D}_n(a,b)$ as the space of all continuous functions $y(x)$ with continuous derivatives up to n-th order defined on a (closed) interval $[a,b]$. The norm, in this case, is

$$||y-z|| = \sum_{i=0}^{n} \max_{a\leq x\leq b} |y^{(i)}(x)-z^{(i)}(x)| \; .$$

The variation of a functional and necessary condition for an extremum

- Continuity of a functional: a functional $J[y]$ is said to be continuous at the point $\hat{y} \in \Re$ if, for any $\epsilon > 0$, there is a $\delta > 0$ such that

$$|J[y] - J[\hat{y}]| < \epsilon \ ,$$

  provided $||y - \hat{y}|| < \delta$.

- With this definition, it follows that $J[y] = \int_a^b F(x, y, y')dx$ is continuous when the norm is that of $\mathcal{D}_1(a, b)$, but this is not the case should the norm be that of $\mathcal{C}(a, b)$.

E: Discuss the above remark in the light of the definitions given so far. How would the continuity of $J[y]$ be affected should the functional also depend on y''?

- Lemma (fundamental Lemma of variational calculus): If $\alpha(x)$ is continuous in $[a, b]$, and if

$$\int_a^b \alpha(x)h(x)dx = 0$$

for every $h(x) \in \mathcal{C}(a, b)$ such that $h(a) = h(b) = 0$, then $\alpha(x) = 0 \; \forall x \in [a, b]$. (Lemma given without Proof)

- The variation or differential of a functional $J[y]$ is defined as

$$\Delta J[h] \equiv J[y + h] - J[y] \; .$$

For fixed y, $\Delta J[h]$ is generally a nonlinear functional of h.

- The functional $J[y]$ is said to be differentiable if

$$\Delta J[h] = \varphi[h] + \epsilon ||h|| \; ,$$

where $\varphi[h]$ is a linear functional and $\epsilon \to 0$ as $||h|| \to 0$.

- Theorem 1: The differential of a differentiable functional is unique (Theorem given without Proof) and is denoted by

$$\delta J[h] = \varphi[h]$$

- Theorem 2: A necessary condition for the differentiable functional $J[y]$ to have an extremum for $y = \hat{y}$ is that its variation vanishes for $y = \hat{y}$, i.e. that

$$\delta J[h] = 0$$

for $y = \hat{y}$ and all admissible h. (Theorem given without Proof)

Euler's equation and variational derivative

- The simplest variational problem can be formulated as follows: Let $F(x, y, z)$ be a function with continuous first and second (partial) derivatives with respect to all its arguments. Then, among all continuous differentiable functions $y(x)$ in the interval $a \leq x \leq b$, which satisfy the boundary conditions $y(a) = A$ and $y(b) = B$, find the function for which the functional

$$J[y] = \int_a^b F(x, y, y')dx$$

has a weak extremum.

- To solve the problem, assume to give $y(x)$ an increment $h(x)$, that is $y(x) \to y(x) + h(x)$. Then

$$\Delta J[h] = J[y + h] - J[y] = \int_a^b [F(x, y + h, y' + h') - F(x, y, y')] \, dx$$

- Taylor expanding and integrating by parts

$$\delta J[h] = \int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \right] h(x) dx \ .$$

- Using $\delta J[h] = 0$ as necessary condition for a weak extremum and Lemma (p. 16), one derives the Euler's equation

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0 \ .$$

- The Euler's equation is a second order differential equation, whose solutions are called extremals. Together with the boundary conditions $y(a) = A$ and $y(b) = B$, the Euler's equation provides a well posed Cauchy's problem.

E: Generalize the Euler's equation to the case of functionals involving higher order derivatives. Discuss the properties of the obtained equation and corresponding boundary conditions.

E: Generalize the Euler's equation to the case of functionals involving more dimensions/degrees of freedom.

- The variational derivative is defined as limit of the incremental ratio

$$\frac{J[y+h]-J[y]}{\Delta\sigma}$$

of $\Delta J[h]$ and the area $\Delta\sigma$ delimited by the function $h(x)$.

- Letting $\Delta\sigma$ going to zero such that $||h||$ and the interval over which $h(x)$ is non-vanishing go to zero, the variational derivative of $J[y]$ at x_0 (the limiting point where $h(x)$ is non-vanishing) is the limit of the incremental ratio above, if it exists

$$\left.\frac{\delta J}{\delta y}\right|_{x=x_0} = \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]_{x=x_0} .$$

- The Euler's equation requires that the variational derivative of the consider functional vanishes as necessary condition for a weak extremum to exist.

Examples solved: (I) The shortest plane curve

- Reconsider the problem of p.6. Find the shortest plane curve joining two points A and B, i.e.,find the curve $y = y(x)$ for which the functional

$$J[y] = \int_a^b \sqrt{1 + y'^2} dx$$

achieves its minimum.

- In this case the functional is independent of $y(x)$ and the Euler's equation becomes

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = \frac{y''}{\sqrt{1 + y'^2}} = 0 \ .$$

- The curve in question turns out to be the straight line segment joining A and B.

A simple variable end problem

- It may occur that considering the variational problem generated by the functional $J[y] = \int_a^b F(x, y, y')dx$, the boundary conditions are not $y(a) = A$ and $y(b) = B$ but either one or both values of $y(x)$ are left free to vary at the end points.

- Consider again $y(x) \to y(x)+h(x)$ (as on p. 18) and compute the increment

$$\delta J[h] = \int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right] h(x)dx + \left(\frac{\partial F}{\partial y'}h(x)\right)_{x=b} - \left(\frac{\partial F}{\partial y'}h(x)\right)_{x=a}$$

- Thus, if $h(a)$ and/or $h(b)$ are non-vanishing, the boundary conditions to the Euler's equation must be replaced by $\partial_{y'}F|_{x=a} = 0$ and/or $\partial_{y'}F|_{x=b} = 0$.

- A simple application of this type of problem is given by the brachistochrone.

Examples solved: (II) The brachistochrone

- Starting from the point $P = (a, A)$, a heavy particle slides down a curve in the vertical plane. Find the curve such that the particle reaches the vertical line $x = b\ (\neq a)$ in the shortest time.
- This is a variant of the classical brachistochrone problem, for which the arrival point is also given at (b, B).
- By definition

$$v = \sqrt{1+y'^2}\frac{dx}{dt} \quad \Rightarrow \quad dt = \frac{\sqrt{1+y'^2}}{v}dx = \frac{\sqrt{1+y'^2}}{\sqrt{2gy}}dx$$

- Thus, the functional is the time $J[y] = T = \int_a^b dt$

$$J[y] = \int_a^b \frac{\sqrt{1+y'^2}}{\sqrt{2gy}}dx \ .$$

- Since the function $F(x,y,y') = \sqrt{1+y'^2}/\sqrt{2gy}$ does not depend explicitly on x, the Euler's equation can be integrated once and, with C an integration constant, yields

$$F - y'\frac{\partial F}{\partial y'} = C \; ;$$

that is

$$\frac{\sqrt{1+y'^2}}{\sqrt{2gy}} - \frac{y'^2}{\sqrt{2gy}\sqrt{1+y'^2}} = \frac{1}{\sqrt{2gy}\sqrt{1+y'^2}} = C \; .$$

- We can rewrite this equation as

$$y\left(1+y'^2\right) = \frac{1}{2gC^2} = 2r \; ,$$

introducing the integration constant $r = (4gC^2)^{-1}$.

- Using the change of variables

$$y = r(1-\cos\theta) \quad \Rightarrow \quad y'^2 = \frac{1+\cos\theta}{1-\cos\theta} = \frac{\cos^2(\theta/2)}{\sin^2(\theta/2)} \; .$$

Thus, introducing the integration constant c, the solution for x becomes

$$dx = \frac{\sin(\theta/2)}{\cos(\theta/2)} dy = 2r\sin^2(\theta/2)d\theta \quad \Rightarrow \quad x = r(\theta - \sin\theta) + c \; .$$

E: Derive the parametric form of the solution of the brachistochrone problem and show it is a cycloid.

- Using (arbitrarily) that the starting point is the origin, then $c = 0$. Meanwhile, the condition at $x = b$ must be $\partial_{y'} F|_{x=b} = 0 \Rightarrow y'(b) = 0$. Thus $r = b/\pi$, and the solution of the problem is

$$x = \frac{b}{\pi}(\theta - \sin\theta) \; , \qquad y = \frac{b}{\pi}(1-\cos\theta) \; .$$

Examples: (III) The isoperimetric problem

- The isoperimetric problem can be formulated as follows: Among all curves of length L in the upper half-plane passing through the points $(-a, 0)$ and $(a, 0)$, find the one which together with the interval $[-a, a]$ encloses the largest area.
- In this case, the functional of interest is

$$J[y] = \int_{-a}^{a} y dx \ ,$$

subject to the boundary conditions $y(-a) = y(a) = 0$ and to the additional constraint

$$B[y] = \int_{-a}^{a} \sqrt{1 + y'^2} dx - L = 0 \ .$$

- Similar to the standard procedure in the classical case of calculus of functions of n variables, the effect of constraints are introduced by proper Lagrange multipliers. Thus, we will find a weak extremum of the functional

$$\bar{J}[y] = J[y] + \lambda B[y]$$

subject to the boundary conditions $y(-a) = y(a) = 0$.

- Note that the requirement that variation with respect to the Lagrange multiplier λ vanishes, yields $B[y] = 0$, that is the constraint itself. Meanwhile, The Euler's equation for the functional $\bar{J}[y]$ becomes

$$1 - \lambda \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) = 0 \; .$$

- Integrating once in x, with C_1 integration constant

$$x - \lambda \frac{y'}{\sqrt{1+y'^2}} = C_1 \; .$$

- Solving for $y'(x)$, and introducing another integration constant, C_2:

$$\frac{dy}{dx} = -\frac{(x-C_1)/\lambda}{\sqrt{1-(x-C_1)^2/\lambda^2}} \,, \qquad \Rightarrow \qquad \frac{y-C_2}{\lambda} = \sqrt{1-(x-C_1)^2/\lambda^2} \,.$$

- Thus, the solution is a circle:

$$(x-C_1)^2 + (y-C_2)^2 = \lambda^2 \,.$$

- That $y(-a) = y(a) = 0$ imposes symmetry, i.e. $y'(0) = 0$ and, thus $C_1 = 0$ as well as $C_2 = -\sqrt{\lambda^2 - a^2}$. Meanwhile, the constraint on the arc length L yields $(\lambda/a)\sin^{-1}(a/\lambda) = L/(2a)$.

E: Considering $\lambda \geq a$ (discuss why it is so), $C_2 \leq 0$. What is the physics meaning of C_2? Discuss what happens when λ varies and give a graphic interpretation for it. Can you explain what happens when $\lambda = a$? And when $\lambda \to \infty$?

Lagrangian field theory: a brief introduction

- Lagrangian field theory belongs to classical field theory. Lagrangian field theory applies to continua and fields, which have an infinite number of degrees of freedom.
- Lagrangian field theory is introduced by extension of Lagrangian mechanics, which is used for discrete particles each with a finite number of degrees of freedom.
- Thus, Lagrangian field theory is the field-theoretic analogue of Lagrangian mechanics.
- In general, the time integral of the Lagrangian is said to be the action, $\mathcal{S}$,
$$\mathcal{S} = \int L dt \; .$$
- The Lagrangian itself can be seen a volume integral of the Lagrangian density, $\mathcal{L}$,
$$L = \int \mathcal{L} d\mathbf{r} \; .$$

- For a scalar field $\varphi(\boldsymbol{r}, t)$, the Lagrangian density (often simply Lagrangian) is generally expressed as

$$\mathcal{L} = \mathcal{L}\left(\varphi, \boldsymbol{\nabla}\varphi, \partial_t \varphi, \boldsymbol{r}, t\right) \ .$$

- Thus, the action is a functional $\mathcal{S}[\varphi]$, and the condition for it to be stationary is given by the Euler's equation

$$\frac{\delta \mathcal{S}}{\delta \varphi} = \frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right) = 0 \ .$$

- The left hand side is the functional derivative of the action with respect to φ. Thus, the functional derivative is a generalization of the variational derivative introduced on p. 20 (the terminologies are synonyms).

- A further generalization of the theory typically refers to m real-valued scalar fields, $\varphi_1, ..., \varphi_m$, which are defined over a n-dimensional manifold M (the space-time in the case above).

A general methodology for constructing variational principles.

- So far we have introduced the concepts of functionals, of function spaces and of their calculus, that is of functional calculus.
- We have provided examples of historic significance, which motivated the development of functional calculus and of calculus of variations, but nothing has been said on the general methodology for constructing functionals and the corresponding variational principles.
- An important class of variational principles is constructed for systems, whose governing equations can be derived from a Lagrangian. That is, this class of variational principles is analogous to Hamilton's principle in classical mechanics and the corresponding field equations can be formulated as Euler's equations.

- In plasma physics, this is the case, for example, of the well-known energy principle in ideal Magneto-Hydro-Dynamics (MHD) [Bernstein 58, Kruskal 58, Rosenbluth 59], and it is the consequence of the MHD force operator being self-adjoint (cf., *e.g.*, [Freidberg 87, White 89]).
- However, there exists a general method, which allows to systematically construct a variational principle for just about any quantity (integral form) defined as a functional of fields that are uniquely defined by a set of equations [Gerjuoy 1983].
- The field equations may involve difference or differential or integral operators, they may be homogeneous or inhomogeneous, linear or nonlinear, self-adjoint or not, and they may or may not represent time-reversible systems.
- The general method and two of its simple yet paradigmatic applications are subject of Lecture 2.

E: Read about the MHD energy principle in your favorite text-book. Discuss it and its implications in the light of the theoretical framework of Lecture 1.

References and reading material

I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Prentice-Hall, Inc. (Englewood Cliffs, NJ, 1963).

E. Gerjuoy, A. Rau, and L. Spruch, 1983, Rev. Mod. Phys. **55**, 725 (1983).

I. B. Bernstein, E. A. Frieman, M. D. Kruskal and R. M. Kulsrud", Proc. Roy. Soc. Ser. A **244**, 17 (1958).

M. D. Kruskal and C. R. Oberman, Phys. Fluids **1**, 275 (1958).

M. N. Rosenbluth and N. Rostoker, Phys. Fluids **2**, 23 (1959).

J. P. Freidberg, *Ideal Magnetohydrodynamics*, Plenum Press (New York and London, 1987).

R. B. White, *Theory of Tokamak Plasmas*, North Holland (Amsterdam, 1989).