Variational principles: summary and problems

David Stuart
dmas2@cam.ac.uk

1 Introduction

Below is an expanded version of parts of the syllabus, intended to fix notation and terminology for doing the problems. It is not a complete summary. For learning all the material some combination of the lectures and the books

- Perfect Form, by Lemons (PUP), general
- Calculus of Variations, by Gelfand and Fomin (Dover) for calculus of variations
- Variational principles in dynamics and quantum theory, by Yourgrau and Mandelstam (Dover) for applications
- Convex optimization, Chapter 3, Boyd S., Vandneberghe L.(CUP) for convexity

should be used. (The last three books give much more detailed treatments than possible/necessary for this course.) The problems are at the end, starred problems being more difficult and not intended for supervision. Please send errors and corrections to the email address above.

2 Variational problems for functions on $\mathbb{R}^n$

$\mathbb{R}^n$ is the the vector space with typical element $\{\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i\}$ where $\mathbf{e}_1 = (1, 0, \ldots, 0)$ etc.

2.1 Differentiability and first order conditions

If a function $f : \mathbb{R}^n \to \mathbb{R}$ has partial derivatives $\partial_i f(\mathbf{x}) = \lim_{t\to 0} t^{-1}(f(\mathbf{x} + t\mathbf{e}_i) - f(\mathbf{x}))$ which exist and are *continuous* on $\mathbb{R}^n$, it is a $C^1(\mathbb{R}^n)$ function, and is differentiable at every $\mathbf{x}$ in the sense that $f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - \nabla f(\mathbf{x}) \cdot \mathbf{h} = o(\|\mathbf{h}\|)$ as $\mathbf{h} \to 0$. This means it can be approximated linearly, and the derivative is the linear map on $\mathbb{R}^n$ given by $Df(\mathbf{x})(\mathbf{h}) = \nabla f(\mathbf{x}) \cdot \mathbf{h}$, which is linear in $\mathbf{h}$.

Lemma 2.1.1 (First order necessary condition) *A local minimum (or maximum) of a C^1 function is a stationary point, i.e. the derivative vanishes there.*

2.2 Second order conditions

If the partial derivatives up to order $r \in \mathbb{N}$ exist and are continuous the function lies in $C^r(\mathbb{R}^n)$. Write the second order partial derivatives $\partial^2_{ij} f = \frac{\partial^2 f}{\partial x_i \partial x_j}$. For a C^2 function $f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - \nabla f(\mathbf{x}) \cdot \mathbf{h} - \frac{1}{2}\sum_{ij} \partial^2_{ij} f(\mathbf{x}) h_i h_j = o(\|\mathbf{h}\|^2)$ as $\mathbf{h} \to 0$.

A real symmetric matrix is positive (resp. non-negative) if $\sum_{ij} A_{ij} v_i v_j > 0$ (resp. ≥ 0) for all non-zero vectors $\mathbf{v}$, or equivalently if all its eigenvalues are positive (resp. non-negative).

Lemma 2.2.1 (Second order necessary conditions) *If a stationary point $\mathbf{x}$ of a $f \in C^2(\mathbb{R}^n)$ is a local maximum (resp. minimum) then $\partial^2_{ij} f(\mathbf{x})$ is a non-positive (resp. non-negative) symmetric matrix.*

Lemma 2.2.2 (Second order sufficient conditions) *If* $f \in C^2(\mathbb{R}^n)$ *and* $Df(\mathbf{x}) = 0$ *and* $\partial^2_{ij} f(\mathbf{x})$ *is a positive (resp. negative) symmetric matrix then* $\mathbf{x}$ *is a strict local minimum (resp. maximum).*

2.3 Convexity

A subset $S \subset \mathbb{R}^n$ is *convex* if for any $\mathbf{x}, \mathbf{y}$ in S and any $t \in [0,1]$ the point $(1-t)\mathbf{x} + t\mathbf{y} \in S$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if $f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y})$ for any $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$ and any $t \in [0,1]$ (or more generally it is convex on a convex subset S if this inequality holds for any $\mathbf{x}, \mathbf{y}$ in S and any $t \in [0,1]$.) Further f is called *strictly convex* if the above inequality is strict whenever it can be i.e. for $0 < t < 1$ and $\mathbf{x} \neq \mathbf{y}$. *Affine* functions, i.e. functions of the form $f(\mathbf{x}) = a + \mathbf{b} \cdot \mathbf{x}$, are examples of functions which are convex but not strictly convex.

Lemma 2.3.1 (Convexity: first order conditions) $f \in C^1(\mathbb{R}^n)$ *convex* $\iff$ $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x})$ *for all* $\mathbf{x}, \mathbf{y}$ $\iff$ $(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})) \cdot (\mathbf{x} - \mathbf{y}) \geq 0$, *for all* $\mathbf{x}, \mathbf{y}$.

As a corollary, this implies that if $\mathbf{x}$ is a stationary point of a convex C^1 function then it is a global minimum.

Also this shows that C^1 convex functions lie above their tangent planes.

Lemma 2.3.2 (Strict convexity: first order conditions) $f \in C^1(\mathbb{R}^n)$ *strictly convex* $\iff$ $f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x})$ *for all* $\mathbf{x} \neq \mathbf{y}$, $\iff$ $(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})) \cdot (\mathbf{x} - \mathbf{y}) > 0$ *for all* $\mathbf{x} \neq \mathbf{y}$.

As a corollary, this implies that if $f \in C^1(\mathbb{R}^n)$ is strictly convex, the equation $\nabla f(\mathbf{x}) = \mathbf{b}$ can have no more than one solution. In particular, stationary points for strictly convex functions are unique.

Lemma 2.3.3 (Convexity: necessary and sufficient second order condition) $f \in C^2(\mathbb{R}^n)$ *is convex* $\iff$ $\partial^2 f_{ij}(\mathbf{x}) \geq 0\ \forall \mathbf{x}$.

Lemma 2.3.4 (Strict convexity: sufficient second order condition) $f \in C^2(\mathbb{R}^n)$ *is strictly convex if* $\partial^2 f_{ij}(\mathbf{x}) > 0\ \forall \mathbf{x}$.

2.4 Lagrange multipliers

Consider a hypersurface $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0\}$ where $g \in C^2(\mathbb{R}^n)$ satisfies $\nabla g(\mathbf{x}) \neq 0$ for all $\mathbf{x}$. The vector $\mathbf{n}(\mathbf{x}) = \nabla g(\mathbf{x}) / \|\nabla g(\mathbf{x})\|$ is everywhere normal to $\mathcal{C}$.

Lemma 2.4.1 *Let* $f \in C^2(\mathbb{R}^n)$. *Then if* $f\big|_{\mathcal{C}}$ *has a maximum (resp. minimum) at* $\mathbf{x} \in \mathcal{C}$ *then there exists* $\lambda \in \mathbb{R}$ *such that* $\nabla h(\mathbf{x}, \lambda) = 0$ *where* $h(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x})$, *and furthermore* $\sum_{ij} \partial^2 h_{ij}(\mathbf{x}, \lambda) v_i v_j$ *is* ≤ 0 *(resp.* ≥ 0*) for all vectors* $\mathbf{v}$ *such that* $\mathbf{v} \cdot \mathbf{n} = 0$.

The function h is the Lagrange augmented function. The number λ is called the Lagrange multiplier.

For problems with several constraints $\{g_\alpha\}_{\alpha=1}^{l}$, assume they are independent (in the sense that the matrix $\partial_i g_\alpha(\mathbf{x})$ has rank l) and consider $h(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum \lambda_\alpha g_\alpha(\mathbf{x})$, and the corresponding result holds.

2.5 Legendre Transform

Given $f : \mathbb{R}^n \to \mathbb{R}$ its Legendre transform $g = f^*$ is given by $g(\mathbf{p}) = \sup(\mathbf{p} \cdot \mathbf{x} - f(\mathbf{x}))$, defined only for $\mathbf{p}$ such that this supremum is finite. The Legendre transform is automatically convex, and the generalized Young inequality

$$f(\mathbf{x}) + g(\mathbf{p}) \geq \mathbf{p} \cdot \mathbf{x}$$

follows immediately from the definition of $g = f^*$. The inequality $xy \leq a^{-1}x^a + b^{-1}y^b$ for $a^{-1} + b^{-1} = 1$ and $a > 1$ is a well-known special case (see exercises).

Theorem 2.5.1 *If f is convex $f^{**} = f$.*

This implies that a convex functions can always be expressed as a supremum of a family of affine functions. This fact also follows from lemma 2.3.1 - just take the family of affine functions to be those lying below the graph of f, and show that this family is non-empty (since it contains the tangent planes) and the supremum gives back f.

3 Variational problems for functionals

3.1 Generalities on functionals

Terminology: $C_0^\infty(a,b)$ is the space of smooth functions whose support is a closed bounded subset of the interval (a,b). The support of a function is the closure of the set where it is non-zero. A bump function in an interval $(x_0-\epsilon, x_0+\epsilon)$ is a function $b \in C_0^\infty(\mathbb{R})$ which is positive in $(x_0-\epsilon, x_0+\epsilon)$ and vanishes for $|x-x_0| \geq 0$. These can be constructed by translating and scaling the bump function on the interval $(-1,1)$ given by $e^{\frac{-1}{(1-x^2)^2}}$ for $x^2 < 1$ and extended with value zero outside the interval (exercise).

A functional is just a function on a set of functions. Since spaces of funtions can be topologized in many inequivalent ways, the continuity and differentiability of functionals is more subtle. For example the Dirac functional $\delta_0(\phi) = \phi(0)$ is continuous on $C(\mathbb{R})$ with the topology determined by the supremum (L^∞) norm $\|\phi\|_{L^\infty} = \max|\phi(x)|$, but not with respect to that determined by the L^2 norm (defined by $\|\phi\|_{L^2}^2 = \int |\phi(x)|^2 dx$). In contrast all norms on finite dimensional vector spaces define equivalent topologies. For this reason we will study differentiability of functionals only one direction at a time, i.e. will consider directional derivatives. The following lemma is useful:

Lemma 3.1.1 *Let $g \in C([a,b])$ have the property that $\int_a^b g(x)\phi(x)dx = 0$ for all $\phi \in C_0^\infty(a,b)$. Then g vanishes identically throughout the interval.*

Proof This follows using continuity and bump functions (exercise).

A slight variation on this lemma states that if $\int_a^b g(x)\phi'(x)dx = 0$ for all $\phi \in C_0^\infty(a,b)$ (notice the prime on ϕ) then g is a constant.

3.2 Directional derivatives of functionals

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be smooth and consider the functional $I[y] = \int_a^b f(x,y,y')dx$ as a function on the space V of C^1 functions with $y(a) = \alpha$ and $y(b) = \beta$. Assume $I[y] = \min_{w \in V} I[w]$ then the function $i(\epsilon) = I[y + \epsilon\phi]$ has a minimum at $\epsilon = 0$ for all $\phi \in C_0^\infty(a,b)$, so that $i'(0) = DI[y](\phi) = \int_a^b (f_y \phi + f_{y'}\phi')dx$ vanishes for each such ϕ. The quantity $DI[y](\phi)$ is called the directional derivative of the functional I along ϕ. Assume further that $y \in C^2(a,b)$, then integration by parts gives, for $\phi \in C_0^\infty(a,b)$:

$$DI[y](\phi) = \int_a^b \big(f_y - \frac{d}{dx}(f_{y'})\big)\phi dx$$

and by lemma 3.1.1, we deduce that

$$\frac{\delta I}{\delta y} = \big(f_y - \frac{d}{dx}(f_{y'})\big) = 0$$

for y a C^2 minimizer. The quantity $\frac{\delta I}{\delta y}$ is sometimes known as the functional derivative, and the mapping $DI[y] : \phi \mapsto DI[y](\phi)$ is called the first variation, and sometimes written δI. The equation

$$\frac{d}{dx}(f_{y'}) - f_y = 0$$

is the Euler-Lagrange equation associated to I. In fact it holds in integrated form $f_{y'} - \int_a^x f_y = constant$ even for C^1 minimizers - this can be deduced using the variation on lemma 3.1.1 mentioned above and an integration by parts trick.

4 Applications

4.1 Fermat principle

Light rays follows paths γ which minimize (or make stationary) the time $T = \int_\gamma \frac{1}{c} ds$, where $ds = \|\dot{\gamma}(t)\| dt$ is the element of arclength along γ and c is the speed of light, which may depend on position.

4.2 Geodesics

A (smooth) Riemannian metric on an open subset $U \subset \mathbb{R}^n$ is a (smooth) function $\mathbf{x} \mapsto g_{ij}(\mathbf{x})$ from U into the space of real positive symmetric $n \times n$ matrices. The geodesics are C^2 curves which are stationary points for the length functional $l[\mathbf{x}] = \int (g_{ij}\dot{x}_i\dot{x}_j)^{\frac{1}{2}} dt$, (where summation convention is understood.) They solve the equation

$$\frac{d}{dt}\Big(\frac{g_{ij}\dot{x}_j}{\sqrt{g_{lm}\dot{x}_l\dot{x}_m}}\Big) - \frac{1}{2}\frac{\partial g_{jk}}{\partial x_i}\frac{\dot{x}_j\dot{x}_k}{\sqrt{g_{lm}\dot{x}_l\dot{x}_m}} = 0.$$

Since the length functional is parametrization invariant, it is possible to choose the parameter t to be the arclength so that $g_{ij}\dot{x}_i\dot{x}_j = 1$, in which case the equation simplifies to

$$\frac{d}{dt}\Big(g_{ij}\dot{x}_j\Big) - \frac{1}{2}\frac{\partial g_{jk}}{\partial x_i}\dot{x}_j\dot{x}_k = 0.$$

This equation is the Euler-Lagrange equation associated to the "kinetic energy integral" $I[\mathbf{x}] = \int g_{ij}\dot{x}_i\dot{x}_j \, dt$, so that an alternative definition of geodesic is a C^2 curve for which I is stationary- this definition automatically gives geodesics with a parametrization for which $g_{ij}\dot{x}_i\dot{x}_j = constant$, by the second conservation law (Noether theorem).

4.3 Lagrangian and Hamiltonian mechanics

The equation

$$m\ddot{\mathbf{x}} + \nabla V = 0 \tag{4.1}$$

for a particle of mass $m > 0$ moving in a potential $V(\mathbf{x})$ can be derived as the Euler-Lagrangian associated to the action functional $S[\mathbf{x}] = \int L(\mathbf{x}, \dot{\mathbf{x}})dt$, where $L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}m\|\dot{\mathbf{x}}\|^2 - V(\mathbf{x})$ is called the Lagrangian. This is the *Lagrangian formulation* of Newtonian mechanics. Since L is convex in $\dot{\mathbf{x}}$ the Legendre transformation in the velocity variables gives a function $H(\mathbf{x}, \mathbf{p}) = \sup_{\dot{\mathbf{x}}} (\mathbf{p} \cdot \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}}))$ from which L can be recovered just by applying the Legendre transform again. The function H is the Hamiltonian, and gives an equivalent formulation of (4.1) in *Hamiltonian form* :

$$\dot{x}_j = \frac{\partial H}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial H}{\partial x_j}$$

Convexity of the Lagrangian in the velocity variables ensures the possibility of going back and forth between the two formulations. Notice that the supremum in the definition of H is attained at the unique $\dot{\mathbf{x}}$ given by $\mathbf{p} = m\dot{\mathbf{x}}$: this defines the *conjugate momentum.*

5 The second variation

Consider the functional $I[y] = \int_a^b f(x, y, y')dx$ on the space V of C^1 functions with $y(a) = \alpha$ and $y(b) = \beta$. Let V_0 be the vector space of C^1 functions with $y(a) = 0$ and $y(b) = 0$.

Definition 5.0.1 *A function $y \in V$ is a* weak *local minimizer for I if $I[y+\phi] \geq I[y]$ for all $\phi \in V_0$ with $\|\phi\|_{C^1} = \max_{[a,b]} |\phi(x)| + \max_{[a,b]} |\phi'(x)|$ sufficiently small. If the inequality is strict for such ϕ not identically zero, the minimum is strict. There is a corresponding definition for weak maximum.*

(There is also a corresponding notion of *strong* minimizer for I with the norm $\|\phi\|_{C^0} = \max_{[a,b]} |\phi(x)|$ used instead of $\|\phi\|_{C^1}$, see Chapter 6 in Gelfand and Fomin.)

Assuming, as always, that f is smooth, Taylor's theorm implies that $\forall \epsilon > 0\, \exists \delta(\epsilon) > 0$ such that for all $x \in [a, b]$ and $\|\phi\|_{C^1} < \delta$:

$$|f(x, y+\phi, y'+\phi') - f(x, y, y') - \phi f_y(x, y, y') - \phi' f_{y'}(x, y, y') - Q| < \epsilon(|\phi|^2 + |\phi'|^2)$$

where Q is the quadratic part of the Taylor expansion

$$Q = \frac{1}{2}\big(\phi^2 f_{yy}(x, y, y') + 2\phi\phi' f_{yy'}(x, y, y') + \phi'^2 f_{y'y'}(x, y, y')\big).$$

Here ϕ, ϕ' are evaluated with argument x. From this follows a corresponding Taylor expansion for the functional I:

$$I[y+\phi] = I[y] + DI[y](\phi) + \frac{1}{2}D^2I[y](\phi) + \mathcal{R}$$

where $|\mathcal{R}| < \epsilon \int_a^b (|\phi|^2 + |\phi'|^2)dx$ for $\|\phi\|_{C^1} < \delta(\epsilon)$. The quadratic part

$$D^2I[y](\phi) = \int \big(\phi^2 f_{yy}(x, y, y') + 2\phi\phi' f_{yy'}(x, y, y') + \phi'^2 f_{y'y'}(x, y, y')\big)dx$$

is sometimes called the second variation, and denoted $\delta^2 I$. From this we can read off:

Lemma 5.0.2 (Necessary conditions) *If $y \in V$ is a weak minimum then $DI[y](\phi) = 0 \forall \phi \in V_0$ and the second variation $D^2I[y](\phi) \geq 0 \forall \phi \in V_0$.*

Lemma 5.0.3 (Sufficient conditions) *Assume $y \in V$ is such that $DI[y](\phi) = 0 \forall \phi \in V_0$ and the second variation satisfies, for some $c > 0$,*

$$D^2I[y](\phi) \geq c \int_a^b (|\phi|^2 + |\phi'|^2)dx \ \forall \phi \in V_0. \tag{5.2}$$

Then y is a weak local minimum.

Recall that if y is C^2 it solves the Euler-Lagrange equation if $DI[y](\phi) = 0 \forall \phi \in V_0$. The fact that $\phi(a) = 0 = \phi(b)$ means that in this case the formula for the second variation can be put into Sturm-Liouville form:

$$D^2I[y](\phi) = \int_a^b \big(p(x)\phi'^2 + q(x)\phi^2\big)dx$$

where $p(x) = f_{y'y'}(x, y(x), y'(x))$ and $q(x) = f_{yy}(x, y(x), y'(x)) - \frac{d}{dx}(f_{yy'}(x, y(x), y'(x))$. One explicit approach to determining whether (5.2) holds for some $c > 0$ is to calculate the eigenvalues of the Sturm-Liouville operator $L = -(p\phi')' + q\phi$. There are also general conditions which ensure (5.2): it is sufficient that $p(x) > 0$ on $[a, b]$ and that there are no conjugate points, i.e. there are no points $\tilde{a} \in (a, b]$ such that there is a non-trivial function h such that $Lh = 0$ and $h(a) = 0 = h(\tilde{a})$. This is proved in theorem 1 in section 26 of Gelfand and Fomin.

6 Example sheet 1

1. Prove that if $f \in C^1(\mathbb{R})$ has only one stationary point which is a local minimum, then it must be a global minimum. Give a counter-example to show this is false in $\mathbb{R}^2$.
2. * Prove that a real symmetric matrix A_{ij} is > 0, in the sense defined in §2.2, iff all its eigenvalues are positive.
3. * Prove, using the Bolzano-Weierstrass property, but without using diagonalizability, that if a real symmetric matrix $A_{ij} > 0$ then $\sum_{ij} A_{ij} v_i v_j \geq c\|\mathbf{v}\|^2$ for some $c > 0$. (After analysis II).
4. * Let $f \in C^2(\mathbb{R}^2)$ have a stationary point $\mathbf{x} = (x^1, x^2)$ and let $A_{ij} = \partial^2_{ij} f(\mathbf{x})$. Show that $A_{11} + A_{22} > 0$ and $A_{11}A_{22} - A_{12}^2 > 0$ implies $A_{ij} > 0$ so that $\mathbf{x}$ is a strict local minimum.
5. Given $f : \mathbb{R}^n \to \mathbb{R}$ define its epigraph to be $E_f = \{(\mathbf{x}, z) : z \geq f(\mathbf{x})\} \subset \mathbb{R}^{n+1}$. Show that f is a convex function iff E_f is convex subset.
6. Give an example of a function which is strictly convex but whose second derivative is not everywhere > 0.
7. Show that x^2/y is convex on the upper half plane $(x, y) : y > 0$. * Show that if $f \in C^2(\mathbb{R})$ is convex then the function $yf(y^{-1}x)$ is convex on $(x, y) : y > 0$.
8. Given a family $L^\alpha(\mathbf{x})$ of affine functions indexed by $\alpha \in \mathbb{N}$, (or in fact an arbitrary index set) show that $f(\mathbf{x}) = \sup_\alpha L^\alpha(\mathbf{x})$ is convex. * Show that all C^1 convex functions arise in this way.
9. * With L^α as in the previous question, show that the function $f(\mathbf{x}) = \inf_\alpha L^\alpha(\mathbf{x})$ is concave.
10. For A any real symmetric $n \times n$ matrix consider $\lambda(A) = \sup_{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\| = 1} \mathbf{v} \cdot (A\mathbf{v})$. Use Lagrange multipliers to show that $\lambda(A)$ is the largest eigenvalue of A. * Also prove that λ is a convex function of A. (Assume the fact from analysis II that a continuous function on the sphere $\{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\| = 1\}$ attains its supremum.)
11. The area A of a triangle with sides a, b, c is given by

$$A = \sqrt{[s(s-a)(s-b)(s-c)]}, \quad \text{where } s = \tfrac{1}{2}(a+b+c).$$

(i) Show that of all triangles of given perimeter $2s$, the triangle of largest area is equilateral.
(ii) Find (in terms of the perimeter) the largest possible area of a right-angled triangle of given perimeter.

12. Prove that the Legendre transform of a function is always convex.
13. Find the Legendre transform of $f(x) = e^x$, (giving its domain also). Find the Legendre transform of $f(x) = a^{-1}x^a, a > 1$ defined on $x > 0$, and deduce $xy \leq a^{-1}x^a + b^{-1}y^b$ for $a^{-1} + b^{-1} = 1$ (Young).
14. * Find the Legendre transform of $f(\mathbf{x}) = \frac{1}{2}\sum_{ij} A_{ij} x_i x_j$ where A_{ij} is a positive symmetric matrix.
15. For an ideal gas, the internal energy $U = U(S, V)$ as a function of entropy and volume is

$$U = U_0 + \alpha n R T_0 \Big[\big(\frac{V_0}{V}\big)^{\frac{1}{\alpha}} e^{\frac{S - S_0}{\alpha n R}} - 1 \Big]$$

for some constants $U_0, T_0, V_0, S_0, \alpha, n, R$. Calculate the pressure and temperature (defined by $dU = TdS - pdV$), and verify that $pV = nRT$ (ideal gas equation of state). Calculate also the constant volume heat capacity $C_V = T\frac{\partial S}{\partial T}|_V$, and comment on the convexity of U as a function of S. Calculate the Helmholtz free energy $F = F(T, V)$ defined by $F(T, V) = \min_S(U(S, V) - TS)$. [*In this formula T is a fixed number - do not substitute for T from the formula you derived in the first part of the question!*]

16. * For black body radiation the internal energy $U = U(S, V)$ as a function of entropy and volume is

$$U(S, V) = \Big(\frac{3S}{4}\Big)^{\frac{4}{3}} \Big(\frac{1}{CV}\Big)^{\frac{1}{3}}$$

where C is a constant. Calculate P, T as in the previous question and verify that the energy density (i.e. the internal energy per unit volume) is CT^4 and that the value of the pressure is $\frac{1}{3}$ of the energy density. Calculate the Helmholtz free energy $F = F(T, V)$ defined by $F(T, V) = \min_S(U(S, V) - TS)$, and show that its value is $-\frac{1}{3}U$.

17. Show that the Euler-Lagrange equation of the functional

$$I[y] = \int_{x_1}^{x_2} f(y, y')dx = 0, \quad y(x_1) = y_1 \text{ and } y(x_2) = y_2 \text{ fixed}$$

has the first integral $f(y, y') - y' \frac{\partial}{\partial y'} f(y, y') =$ constant. The curve assumed by a uniform cable which is suspended between two points $(-a, b)$ and (a, b) minimises the potential energy

$$\int_{-a}^{a} y(1 + y'^2)^{1/2} dx$$

subject to the constraint that its length remains fixed,

$$\int_{-a}^{a} (1 + y'^2)^{1/2} dx = 2L,$$

where $L > a$. Using the Lagrange multiplier method, show that the curve is a catenary

$$y - y_0 = c \cosh\left(\frac{x - x_0}{c}\right),$$

where c, x_0 and y_0 are constants. * Find an equation for c, and show that it has a unique positive solution.

18. Write down the Euler-Lagrange equation for the functional

$$I[u] = \int_{-\infty}^{+\infty} \frac{1}{2} u'^2 + (1 - \cos u) dx$$

and find all solutions which satisfy $\lim_{x\to-\infty} u(x) = 0$ and $\lim_{x\to+\infty} u(x) = 2\pi$.
Show that if $u \in C^1(\mathbb{R})$ satisfies $\lim_{x\to-\infty} u(x) = 0$ and $\lim_{x\to+\infty} u(x) = 2\pi$

$$I[u] = \frac{1}{2} \int_{-\infty}^{+\infty} (u' - 2\sin\frac{u}{2})^2 \, dx + 8.$$

Deduce that a lower bound for $I[u]$ amongst such functions is 8, and give a *first order* differential equation which u must satisfy in order to realize this lower bound. Show that any solution of this first order equation solves the Euler-Lagrange equation you derived in the first part of the question. Give all the functions satisfying $I[u] = 8$.

19. * The brachistochrone problem leads to the study of the functional $I[y] = \int_0^X \frac{\sqrt{(1+y'^2)}}{\sqrt{y}} dx$ for C^1 curves $y = y(x) > 0$ such that $y(0) = 0$ and $y(X) = Y > 0$. Make the change of variables $y = \phi^2$, and show that $J[\phi] = I[\phi^2] = \int_0^X (\phi^{-2} + 4\phi'^2)^{\frac{1}{2}} dx$. Show that the function $l(u, v) = (u^{-2} + 4v^2)^{\frac{1}{2}}$ is strictly convex on $\{(u, v) : u > 0\} \in \mathbb{R}^2$. (This can be used to prove the cycloid solution which we obtained as a solution of the Euler-Lagrange equation, which is only a necessary condition for a minimizer, actually does minimze I.) Write down the Euler-Lagrange equation for $J[\phi]$, solve for ϕ and show that the solutions are cycloids, as for the Euler-Lgrange equation for I.

20. Obtain the Euler-Lagrange equation for the function $x(t)$ that makes stationary the integral

$$\int_{t_1}^{t_2} f(t, x(t), \dot{x}(t), \ddot{x}(t)) dt$$

for fixed values of both $x(t)$ and $\dot{x}(t)$ at both $t = t_1$ and $t = t_2$.
Find the function $x(t)$ with $x(1) = 1, \dot{x}(1) = -2, x(2) = \frac{1}{4}$ and $\dot{x}(2) = -\frac{1}{4}$, that minimises $\int_1^2 t^4 [\ddot{x}(t)]^2 dt$, including a demonstration that it is a minimizer (not just a stationary point) for the integral.

7 Example sheet 2

1. Consider the problem of maximizing the area $\frac{1}{2}\int_0^{2\pi}(x\dot{y} - y\dot{x})dt$ enclosed by a *closed* curve of fixed length $l = \int_0^{2\pi}(\dot{x}^2 + \dot{y}^2)^{\frac{1}{2}}dt$. Write down and solve the Euler-Lagrange equations for this constrained problem in parametric form.
2. Consider the problem of minimizing $I[\psi] = \int_{-\infty}^{+\infty}\left(\psi'^2 + x^2\psi^2\right)dx$ amongst functions with $\int \psi^2 dx = 1$.
(i) Write down the corresponding Euler-Lagrange equation for this constrained problem.
(ii) Show that under the assumption $x\psi(x)^2 \to 0$ as $x \to +\infty$ it is possible to write $I[\psi] = 1 + \int_{-\infty}^{+\infty}(\psi' + x\psi)^2 dx$, and hence show that amongst such functions the minimum value of I is 1 and is attained on a function which should be given explicitly. Verify that this function satisfies the Euler-Lagrange equation you wrote down in (i), for an appropriate value of the Lagrange multiplier.
(iii) * Use the method of power series solutions to solve the Euler-Lagrange equation in (i), and comment on the relation with the minimizing function you obtained in (ii). (Here you may find it useful to rewrite the Euler-Lagrange equation as an equation for $f = e^{\frac{x^2}{2}}\psi(x)$).
3. Obtain the Euler-Lagrange equations associated to the functionals
(i) $I[u] = \int(\frac{1}{2}u_t^2 - F(u_x))dxdt$,
(ii) * $I[u] = \frac{1}{2}\int(u_t^2 - c(u)^2u_x^2)dxdt$,
where $u = u(t,x)$ is a function on $\mathbb{R}^2$, where F and c are given smooth functions.
4. Obtain the Euler-Lagrange equations associated to the functionals
(i) $I[u] = \int(|\nabla u|^2 + e^{2u})dxdy$,
where $u = u(x,y)$ is a function on $\mathbb{R}^2$, and
(ii) * $I[u] = \int(\det Du)dxdy$,
where $u : \mathbb{R}^2 \to \mathbb{R}^2$, and $\det Du$ means the Jacobian determinant. What is unusual about the second example?
5. Consider $I[y] = \int_{-1}^{+1}(xy')^2dx$ for $y(x)$ in the set S of C^1 functions such that $y(1) = 1$ and $y(-1) = -1$. By considering $y_\epsilon(x) = \frac{\arctan x/\epsilon}{\arctan 1/\epsilon}$ show that $\inf_{y\in S} I[y] = 0$. Show that this infimum is not attained in S.
6. Consider $I[y] = \int_{-1}^{1}(1 - y_x^2)^2dx$ with $y = y(x)$ lying in the set S' of *piecewise* C^1 functions such that $y(\pm 1) = 1$. By considering $y(x) = |x|$ show that the $\min_{y\in S'} I[y] = 0$. Does there exist a C^1 (not just piecewise C^1) function for which this value is attained?
7. The smooth functions $p(x), q(x)$ and $w(x) \geq 0$ are prescribed on $[a,b]$, with w not identically zero. Show that the following three conditions are equivalent for C^2 functions $y(x)$ satisfying $y(a) = 0 = y(b)$:
(i) y satisfies: $(py')' - qy = -\lambda wy$;
(ii) $I[u] = \int_a^b(pu'^2 + qu^2)dx$ is stationary at $u = y$ amongst C^1 functions satisfying the boundary conditions and subject to the constraint $\int_a^b wu^2dx =$ constant;
(iii) $Q[u] = \int_a^b(pu'^2 + qu^2)dx / \int_a^b wu^2dx$, is stationary amongst C^1 functions satisfying the boundary conditions at $u = y$. What is the value of $Q[y]$?
(Assume that y is not identically zero, and that $w > 0$ in (a,b) so that so that the denominator $\int_a^b wy^2dx$ in (iii) is non-zero.)
8. Let $\mathbf{x}(t) \in \mathbb{R}^3$ be a curve which is constrained to lie on the sphere $S^2 = \{\mathbf{x} : \|\mathbf{x}\| = 1\}$. Use the Lagrange multiplier function formalism to obtain the following Euler-Lagrange equation
$$\ddot{\mathbf{x}} + \|\dot{\mathbf{x}}\|^2\mathbf{x} = 0 \tag{7.3}$$
for the problem of minimizing $I[\mathbf{x}] = \int \|\dot{\mathbf{x}}\|^2 dt$ amongst curves satisfying the constraint $\mathbf{x}(t) \in S^2$. Show that the solutions of the Euler-Lagrange equation lie on a plane through the origin (they are great circles.)
9. * As an alternative approach to (7.3), let θ, ϕ be standard angles given by spherical coordinates, and assume the curve on S^2 is given as $\phi = \phi(\theta)$. Show that the length integral

is $l[\phi] = \int (1 + \sin^2\theta\phi'^2)^{\frac{1}{2}} d\theta$. Obtain the Euler-Lagrange equation associated to this functional, integrate it and show that the resulting solutions are great circles.

10. * Obtain (7.3) by considering variations of the curve $\mathbf{x}(t)$ of the form

$$\mathbf{x}^\epsilon(t) \equiv \frac{\mathbf{x}(t) + \epsilon\mathbf{z}(t)}{\|\mathbf{x}(t) + \epsilon\mathbf{z}(t)\|}$$

which lie on S^2 and requiring $\frac{d}{d\epsilon} I[\mathbf{x}^\epsilon] = 0$ at $\epsilon = 0$ for every smooth $\mathbf{z}(t)$.

11. * For the brachistochrone problem, show that the minimum travel time between two points at the same level and a distance l apart is $(2\pi l/g)^{1/2}$ (for a bead moving on a wire under the action of gravity without friction. The acceleration due to gravity is g.)
12. * For the brachistochrone problem, show that there is a unique arc of a cycloid (without a cusp) from the starting point $(0, 0)$ to a point (X, Y) below the starting point.
13. In an optical medium filling the region $0 < y < h$, the speed of light is

$$c(y) = \frac{c_0}{(1 - ky)^{1/2}} \quad (0 < k < 1/h).$$

Show that the paths of light rays in the medium are parabolic. Show also that, if a ray enters the medium at $(-x_0, 0)$ and leaves it at $(x_0, 0)$, then

$$(kx_0)^2 = 4ky_0(1 - ky_0),$$

where y_0 $(< h)$ is the greatest value of y attained on the ray path.

14. * Hamilton's Principle is applicable also to the *relativistic* dynamics of a charged particle in an electromagnetic field. The appropriate choice of Lagrangian $L[t, \mathbf{x}(t), \dot{\mathbf{x}}(t)]$ is

$$L = -m_0c^2\gamma^{-1} + qA_0 + q\mathbf{v} \cdot \mathbf{A},$$

with the Lorentz factor $\gamma = (1 - v^2/c^2)^{-1/2}$, and where $\mathbf{x}$ is the position and $\mathbf{v} = \dot{\mathbf{x}}(t)$ is the velocity of a particle of rest-mass m_0 and charge q in fields determined by a given scalar potential $A_0(\mathbf{x}, t)$ and a given vector potential $\mathbf{A}(\mathbf{x}, t)$. Verify that the Euler-Lagrange equations, with this choice of L, yield the equation of motion

$$\frac{d}{dt}(m_0\gamma\mathbf{v}) = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

where the electric field $\mathbf{E} = \nabla A_0 - \frac{\partial \mathbf{A}}{\partial t}$ and the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$.

15. * With $\mathbf{E}$ and $\mathbf{B}$ as in the previous question, obtain the Euler-Lagrange equations associated to the functional $I[A] = \int (\mathbf{E}^2 - \mathbf{B}^2) dx dt$. (This gives two of Maxwell's equations).
16. For the length functional for curves in the plane $I[y] = \int_a^b (1 + y'^2)^{\frac{1}{2}} dx$, with $y(a) = \alpha$ and $y(b) = \beta$ show that the straight line $y = y_0(x)$ joining (a, α) to (b, β) solves the Euler-Lagrange equation. Compute the second variation of I at y_0 and show that it is positive.
17. For $I[y] = \int_a^b (y'^2 + y^4) dx$ with $y(a) = \alpha$, $y(b) = \beta$ find the Euler-Lagrange equation and the second variation. For the case $\alpha = 0 = \beta$ write down the solution of the Euler-Lagrange equation and the second variation explicitly, and show that the second variation is strictly positive.
18. For $I[y] = \int_0^1 \left(\frac{1}{2}y'^2 + F(y)\right) dx$ with $y(0) = 0 = y(1)$. Assume that $F \in C^2(\mathbb{R})$ satisfies $F'(0) = 0$. Write down the associated Euler-Lagrange equation, and show that $y_0(x) = 0$ is a solution. Find the second variation. Give (i) a condition on $F''(0)$ which ensures that the second variation is positive, and (ii) a condition which ensures the second variation has at least one negative eigenvalue.

8 Additional questions

1. The following questions from recent methods exams are good for practice with Lagrange multipliers, Euler-Lagrange equations etc: 2008 1/II/14D and 2/I/5D, 2007: 3/I/6E and 4/II/16E, 2006: 2/I/5A and 4/II/16B.
2. At how many points in R^3 does the function

$$\phi(x_1, x_2, x_3) = \frac{1}{4}(x_1^4 + x_2^4 + x_3^4) - x_2x_3 - x_3x_1 - x_1x_2$$

take its minimum value? Show that this least value is -3. Show also that ϕ has one saddle point, at which the surface of vanishing ϕ is tangent to a double cone of semi-angle $\tan^{-1}(\sqrt{2})$.
3. Find the maximum volume of a rectangular parallelopiped inscribed inside an ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$.
4. *Show that if $f : (a, b) \to \mathbb{R}$ is convex the one-sided difference quotients $\phi_x(h) = h^{-1}(f(x+h) - f(x))$, $h > 0$ are non-decreasing i.e. $\phi_x(h) \le \phi_x(k)$ if $0 < h \le k$. Deduce that the right derivative $D^+f(x) \equiv \lim_{h\to 0+, h>0} \phi_x(h)$ exists in $-\infty \cup \mathbb{R}$. By considering $\phi_{x-l}(l)$ for $l > 0$ show that for any $x \in \mathbb{R}$ the $\phi_x(h)$ are bounded below for $h > 0$ so that the right derivative $D^+f(x)$ just defined is finite for all x for a convex function with domain $\mathbb{R}$ like f. Show that if the domain of f is only an interval that the same is true for x an interior point of the interval. Give an example of a convex function defined only on $[0, \infty)$ for which the right derivative at $x = 0$ is $-\infty$.
5. *Consider $I[y] = \int_a^b f(x, y, y')dx$ with $y(a) = \alpha, y(b) = \beta$, where f is a smooth function $f : \mathbb{R}^3 \to \mathbb{R}$. Consider variations of the form $y^\epsilon(x) = y(x + \epsilon\phi(x))$ where $\phi \in C_0^\infty(a, b)$, and compute $\frac{d}{d\epsilon}I[y^\epsilon]|_{\epsilon=0}$; show that if y is such that this is zero for all such ϕ then the conservation law $y'f_{y'} - f = constant$ holds.
6. Consider the area of a surface obtained by rotating a curve $y = y(x)$ with $y(a) = \alpha$ and $y(b) = \beta$ about the y-axis. Write down an integral for the area, and solve the associated Euler-Lagrange equation.
7. Consider $I[y] = \int_a^b f(x, y, y')dx$ with $y(a) = \alpha$ but $y(b)$ is not fixed. As usual f is a smooth function $f : \mathbb{R}^3 \to \mathbb{R}$. Show that if $y \in C^2$ minimizes I amongst C^1 functions with $y(a) = \alpha$ then as well as the Euler-Lagrange eqyation it satisfies the additional boundary condition $f_{y'}(b, y(b), y'(b)) = 0$. Together with the initial condition this gives the correct number of boundary conditions for the second order Euler-Lagrange equation. Boundary conditions which are a consequence of a variational problem in this way are called *natural*. What is the natural boundary condition for $I[u] = \int_B \big(\frac{1}{2}|\nabla u|^2 - gu\big)dx$ where B is the unit ball in $\mathbb{R}^n$?
8. Find the Hamiltonian obtained via the Legendre transformation from the Lagrangian $L = \frac{1}{2}g_{ij}\dot{x}_i\dot{x}_j - V(\mathbf{x})$ (summation convention assumed).
9. Find the Hamiltonian for the relativistic dynamics of a charged particle by applying the Legendre transformation to the Lagrangian $L = -m_0c^2\gamma^{-1} - qA_0 - q\mathbf{v}\cdot\mathbf{A}$, which appears in sheet II.
10. Write down the Euler-Lagrange equation associated to $I[u] = \int_{-\infty}^{+\infty} \frac{1}{2}u'^2 + (1 - \cos u)dx$ and show that $u(x) = 4\arctan e^x$ is a solution with boundary conditions $\lim_{x\to-\infty} u(x) = 0$ and $\lim_{x\to+\infty} u(x) = 2\pi$. (i) Calculate the second variation, and (ii)* use the method of power series to find the eigenvalues of the associated Sturm-Liouville operator.
11. (i) Consider the functional $I[u] = \int_{-\pi}^{+\pi} \big(\frac{u_x^2}{2} - fu\big)dx$ where u and f are real 2π- periodic functions with zero mean:$\int_{-\pi}^{+\pi} u(x)dx = 0 = \int_{-\pi}^{+\pi} f(x)dx$. Write down the Euler-Lagrange equation.
(ii) Now consider the case that u, f are given by finite sums of exponentials:

$$u(x) = \sum_{0<|n|\le N} u_n e^{inx}, \quad f(x) = \sum_{0<|n|\le N} f_n e^{inx}$$

with the reality conditions $\bar{u}_n = u_{-n}, \bar{f}_n = f_{-n}$ and N any positive integer. Show that

$I[u] = 2\pi J_N[\underline{u}]$ where $\underline{u} = (u_1, u_2, \ldots u_n) \in \mathbb{C}^N$ and

$$J_N[\underline{u}] = \sum_{n=1}^{N} n^2|u_n|^2 - \bar{f}_n u_n - f_n \bar{u}_n$$

Use completion of the square to show that the minimum of J_N is attained for some unique $\underline{u}$, and show that the corresponding function u solves the Euler-Lagrange equation in (i).
(iii)* Use the direct method to prove the existence of a minimizer for J_N as follows. First show that J_N is bounded below, and let $\{\underline{u}^\alpha\}_{\alpha=1}^{\infty}$ be a sequence such that $J_N[\underline{u}^\alpha] \to \inf_{\underline{v}\in\mathbb{C}^N} J_N[\underline{v}]$ as $\alpha \to \infty$. Show that there is a subsequence which converges to a limit point $\underline{u}$ which is a minimizer , i.e. $J_N[\underline{u}] = \inf_{\underline{v}\in\mathbb{C}^N} J_N[\underline{v}]$. Finally, deduce by considering the stationary condition satisfied by minimizers for J_N, that this minimizer is the same as the one you obtained in (ii).
(iv)* [After Methods and Analysis II] Extend your argument in (iii) to the case $N = +\infty$ and show that amongst sequences such that $\sum_{n=1}^{\infty} n^2|u_n|^2 < \infty$ there is one that minimizes J_∞. Work under the assumption that f is given by an absolutely convergent Fourier series. (Hint: look up Cantor diagonalization.)

Variational principles in dynamics and quantum theory, by Yourgrau and Mandelstam(Dover) for applications. Convex optimization, Chapter 3, Boyd S., Vandneberghe L.(CUP) for convexityshould be used. (The last three books give much more detailed treatments than possible/necessaryfor this course.) The problems are at the end, starred problems being more difficult and not in-tended for supervision. Please send errors and corrections to the email address above. 2 Variational problems for functions on Rn. Rn is the the vector space with typical element {x = ni=1 xiei} where e1 = (1, 0, . . . , 0) et Physical laws may be expressed by variational principle not only by dierential formulas, in fact, dierential formulas can be derived from variational principle[1], and there are the variational principles of dierential and integral styles[2, 3]. Hilbert outlined 23 major mathematical problems to be studied in the coming century, variational problem is one of the problems.Â IV represents Noether conservation charges of all integral variational principles, the last Sect. is summary and conclusions. 2 Unied Expressions Of All Integral Variational Principles. Ref.[6] rigorously gives the expression of QCP derived from the no-loss-no-gain principle.