MA209
Variational Principles

Revision Guide

Written by Joy Tolia and Matt Rigby

Contents

Introduction

This revision guide for MA209 Variational Principles has been designed as an aid to revision, not a substitute for it. This module consists a few derivations which are important and there is a lot of solving linear coefficient ODEs. The exams are very similar each year so practice all derivations and do examples from the past papers.

Disclaimer: Use at your own risk. No guarantee is made that this revision guide is accurate or complete, or that it will improve your exam performance. Use of this guide *will* increase entropy, contributing to the heat death of the universe. Contains no GM ingredients. Your mileage may vary. All your base are belong to us.

Authors

This revision guide for MA209 has been designed as an aid to revision, not a substitute for it. Use at your own risk. Written by Joy Tolia and Matt Rigby.
Based upon lectures given by Prof. John Rawnsley at the University of Warwick, 2012-2013.
Any corrections or improvements should be entered into our feedback form at http://tinyurl.com/WMSGuides (alternatively email revision.guides@warwickmaths.org).

1 Fundamental Theorem of Calculus of Variations

Theorem 1.1 (Fundamental Theorem of Calculus of Variations)**.** If $v(x)$ is a continuous function on $[x_1, x_2]$ such that

$$\int_{x_1}^{x_2} v(x)u(x)dx = 0$$

for all $u \in C^2$ with $u(x_1) = u(x_2) = 0$ then

$$v(x) = 0 \qquad \forall x \in [x_1, x_2]\,.$$

Proof. Assume that $\int_{x_1}^{x_2} v(x)u(x)dx = 0$ for all $u \in C^2$ with $u(x_1) = u(x_2) = 0$. Suppose $\exists x_0 \in (x_1, x_2)$ where $v(x_0) > 0$. Then since v is continuous, $\exists \delta > 0$ such that $v(x) > 0$ $\forall x \in (x_0 - \delta, x_0 + \delta)$. Suppose we can find $u \in C^2$ such that

$$u(x) = \begin{cases} u(x) > 0 & \forall x \in (x_0 - \delta, x_0 + \delta) \\ 0 & \text{otherwise} \end{cases}$$

Then $u(x)v(x)$ is strictly positive for all $x \in (x_0 - \delta, x_0 + \delta)$ and zero otherwise. Hence we have:

$$\int_{x_1}^{x_2} v(x)u(x)dx = \int_{x_0-\delta}^{x_0+\delta} v(x)u(x)dx > 0$$

This is a contradiction, hence $v(x) = 0$ for all $x \in [x_1, x_2]$. □

Remark 1.2. We can always find a u that we need in the above proof, for example:

$$u(x) = \begin{cases} (x_0 + \delta - x)^3(x - x_0 + \delta)^3 & \forall x \in (x_0 - \delta, x_0 + \delta) \\ 0 & \text{otherwise} \end{cases}$$

2 Euler-Lagrange Equation

Definition 2.1. Let y be a function of x, f be a function of x, y, y' and $I(y) = \int_{x_1}^{x_2} f(x, y, y')dx$ be a functional then its *Euler-Lagrange equation* is given by

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

Theorem 2.2. Suppose $y \in C^2$ is a function of x and $f \in C^2$ is a function of x, y, y'. Then, any critical point $y(x)$ of the functional

$$I(y) = \int_{x_1}^{x_2} f(x, y, y')dx$$

satisfies its Euler-Lagrange equation:

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

Proof. Suppose $y \in C^2$ is a critical point of I. Let $u \in C^2$ satisfy $u(x_1) = u(x_2) = 0$, and consider

$g_u(t) = I(y+tu)$. $g_u(t)$ has a critical point at $t=0$, hence:

$$\begin{aligned}
\frac{d}{dt}(g_u(t))\bigg|_{t=0} &= \frac{d}{dt}(I(y+tu))\bigg|_{t=0} \\
&= \frac{d}{dt}\int_{x_1}^{x_2} f(x, y+tu, y'+tu')dx\bigg|_{t=0} \\
&= \int_{x_1}^{x_2} \left(\frac{d}{dt}f(x, y+tu, y'+tu')\bigg|_{t=0}\right)dx \\
&= \int_{x_1}^{x_2} \left(u\frac{\partial f}{\partial y} + u'\frac{\partial f}{\partial y'}\right)dx \\
&\ \text{Integrating by parts we get:} \\
&= \int_{x_1}^{x_2} u\frac{\partial f}{\partial y}dx + \left[u\frac{\partial f}{\partial y'}\right]_{x_1}^{x_2} - \int_{x_1}^{x_2} u\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)dx \\
&\ \text{As } u(x_1) = 0 = u(x_2): \\
&= \int_{x_1}^{x_2} u\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)dx = 0
\end{aligned}$$

Apply the fundamental theorem of calculus of variations with $v = \frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)$ then we have $\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$. ☐

Proposition 2.3. Suppose f has no explicit x dependence and satisfies the Euler-Lagrange equation. Then, $f - y'\frac{\partial f}{\partial y'}$ is constant.

Proof. As f has no explicit x dependence, $\frac{\partial f}{\partial x} = 0$.

$$\begin{aligned}
\frac{d}{dx}\left(f - y'\frac{\partial f}{\partial y'}\right) &= \frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'} - y''\frac{\partial f}{\partial y'} - y'\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) \\
&= y'\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right) \\
&= 0 \qquad \text{(By E-L equation)}
\end{aligned}$$

Therefore $f - y'\frac{\partial f}{\partial y'}$ is constant. ☐

Definition 2.4. Let $I(y) = \int_{x_1}^{x_2} f(x, y, y')dx$ be a functional then if it has no explicit x dependence. Then ***its*** 1^{st} *Integral* is given by

$$f - y'\frac{\partial f}{\partial y'}.$$

When y is sufficiently smooth, this theory above extends to when f is a function of multiple derivatives of y.

Proposition 2.5. Suppose $y \in C^{n+1}$ is a function of x and $f \in C^{n+1}$ is a function of $x, y, y', \ldots, y^{(n)}$, then any critical point $y(x)$ of the functional

$$I(y) = \int_{x_1}^{x_2} f(x, y, y', \ldots, y^{(n)})dx$$

satisfies the following equation:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''}\right) + \ldots + (-1)^n\frac{d^n}{dx^n}\left(\frac{\partial f}{\partial y^{(n)}}\right) = 0$$

Proof. Suppose $y \in C^{n+1}$ is critical point of I. Let $u \in C^n$ such that $u(x_1) = u(x_2) = \ldots = u^{(n-1)}(x_1) = u^{(n-1)}(x_2) = 0$, and consider $g_u(t) = I(y + tu)$. $g_u(t)$ has a critical point at $t = 0$, hence:

$$\begin{aligned}
\frac{d}{dt} g_u(t)\bigg|_{t=0} &= \int_{x_1}^{x_2} u\frac{\partial f}{\partial y} + u'\frac{\partial f}{\partial y'} + \ldots + u^{(n)}\frac{\partial f}{\partial y^{(n)}} dx \\
&\quad \text{Integrating by parts multple times and} \\
&\quad \text{as } u(x_1) = u(x_2) = \ldots = u^{(n-1)}(x_1) = u^{(n-1)}(x_2) = 0 \text{ we have:} \\
&= \int_{x_1}^{x_2} u\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''}\right) + \ldots + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial f}{\partial y^{(n)}}\right)\right) dx \\
&= 0
\end{aligned}$$

Applying the Fundamental Theorem of Calculus of Variations we get the result. □

This theory also extends to the case where f is a function of more than one variable.

Proposition 2.6. Suppose $x, y \in C^2$ are functions of t and $f \in C^2$ is function of $t, x, y, \dot{x}, \dot{y}$, then any critical point $(x(t), y(t))$ of the functional $I(x, y) = \int_{t_1}^{t_2} f(t, x, y, \dot{x}, \dot{y})dt$ satisfies the following equations:

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) = 0$$

$$\frac{\partial f}{\partial y} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{y}}\right) = 0$$

Proof. Let $u_1, u_2 \in C^2$ such that $u_1(x_1) = u_1(x_2) = u_2(x_1) = u_2(x_2) = 0$ and consider

$$g_{u_1,u_2}(h_1, h_2) = I(x + h_1u_1, y + h_2u_2).$$

If $(x(t), y(t))$ is a critical point of $I(x, y)$ then $g_{u_1,u_2}(h_1, h_2)$ has a critical point at $(h_1, h_2) = (0, 0)$ so we have:

$$\frac{\partial g_{u_1,u_2}}{\partial h_1}(0, 0) = 0, \qquad \frac{\partial g_{u_1,u_2}}{\partial h_2}(0, 0) = 0$$

Looking at the partial derivative of g_{u_1,u_2} with respect to h_1 we have:

$$\frac{\partial}{\partial h_1} g_{u_1,u_2}(0, 0) = \frac{d}{dh_1} I(x + h_1u_1, y)\bigg|_{h_1=0}$$

Therefore x is a critical point of the functional $x \mapsto I(x, y)$ with y fixed. Hence x satisfies its Euler-Lagrange equation:

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) = 0$$

Similarly considering the partial derivative of g_{u_1,u_2} with respect to h_2 we get that y satisfies its Euler-Lagrange equation. □

3 Fermat's Principle for Optics

Light in a transparent medium travels along trajectories whose shape is determined by the speed of light c. In a 2-D medium the speed at (x, y) is given by a function $c(x, y)$.

Fermat's Principle: Light travels along a path in a transparent medium between two points chosen to minimise the time taken amongst all possible paths joining those two points.

Proposition 3.1. Let the points (x_1, y_1) and (x_2, y_2) be in a medium where the speed of light is $c(x, y)$. Then the path of light, $y(x)$ between the two points is given by the critical point of the following functional:

$$T(y) = \int_{x_1}^{x_2} \frac{\sqrt{1 + (y')^2}}{c(x, y)} dx$$

Proof. The time taken by light to travel from (x_1, y_1) to (x_2, y_2) in a medium where the speed of light is $c(x, y)$ is given by:

$$\begin{aligned} T(y) &= t_2 - t_1 \\ &= \int_{t_1}^{t_2} dt \\ &= \int_{s_1}^{s_2} \frac{ds}{\frac{ds}{dt}} \\ &= \int_{s_1}^{s_2} \frac{ds}{c(x,y)} \\ &= \int_{x_1}^{x_2} \frac{\sqrt{1+(y')^2}}{c(x,y)} dx \end{aligned}$$

Then by Fermat's Principle we have that light travels on the path $y(x)$ which minimises the above functional. □

Proposition 3.2. Let the points (x_1, y_1) and (x_2, y_2) be in a medium where the speed of light only depends on y, hence $c = c(y)$. Then $\exists D \in \mathbb{R}$ such that

$$\frac{1}{c(y)\sqrt{1+(y')^2}} = D.$$

Proof. Observing that $f(x, y, y') = \frac{\sqrt{1+(y')^2}}{c(y)}$ as f has no explicit x dependence, and that any path light takes is an extremal of the functional $T(y)$ defined above, the first integral of T, $f - y'\frac{\partial f}{\partial y'}$ is constant. Also we see:

$$\begin{aligned} f - y'\frac{\partial f}{\partial y'} &= \frac{\sqrt{1+(y')^2}}{c(y)} - y'\frac{y'}{c(y)\sqrt{1+(y')^2}} \\ &= \frac{1+(y')^2}{c(y)\sqrt{1+(y')^2}} - \frac{(y')^2}{c(y)\sqrt{1+(y')^2}} \\ &= \frac{1}{c(y)\sqrt{1+(y')^2}} \end{aligned}$$

□

4 Hamilton's Principle for Conservative Mechanics

A path of a (system of) particle(s) is a path in a Euclidean space of some dimension $(\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3, \mathbb{R}^{3n}, \ldots)$ depending on the number of particles and degrees of freedom.

Let $\mathbf{x}(t)$ be the path of a particle (or system). In 3-D, a particle has a mass m and determines a kinetic energy $\frac{1}{2}m|\dot{\mathbf{x}}|^2$. For a system of masses m_i, positions $\mathbf{x}_i(t)$ then add together all the kinetic energies for total kinetic energy:

$$T = \sum_i \frac{1}{2}m_i|\dot{\mathbf{x}}_i|^2$$

If force $\mathbf{F}$ acting on a system of particles is conservative then:

$$\mathbf{F} = -\nabla V$$

for some V which is a function of $\mathbf{x}_i$, called the potential energy of the system.

We can change variables to some conveniently chosen system of generalised *unconstrained* coordinates.

Denote the generalised unconstrained coordinates by $q_1, q_2, \ldots$. As the system moves the q_i will be a function of t. Using chain rule, T and V become functions of $q_i, \dot{q}_i$:

$$T(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n)$$

$$V(q_1, \ldots, q_n)$$

$L = T - V$ is a function of Lagrangian $q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n$.

Hamilton's Principle: If a mechanical system evolves from position p_1 at time t_1 to position p_2 at time t_2 then amongst all paths joining p_1 to p_2 at times t_1 and t_2, the actual path is a critical point of $I(q_1, \ldots, q_n) = \int_{t_1}^{t_2} L dt$.

5 Constraints and Lagrange Multipliers

If we want to find an extremum on a constrained set $X = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$ then the following theorem is very important.

Theorem 5.1. Suppose $f, g \in C^1$ are functions of two variables x, y and g is regular (i.e. $\nabla g \neq 0$). If (x_0, y_0) is an extremum of f on $X = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$ then there exists $\lambda \in \mathbb{R}$ such that $f - \lambda g$ has an unconstrained critical point at (x_0, y_0).

Proof. Let $f, g \in C^1$ and $\nabla g \neq 0$ so without loss of generality assume $\frac{\partial g}{\partial y} \neq 0$. Let (x_0, y_0) be an extremum of f on X. By implicit function theorem there exists a function $\eta(x) \in C^1$ defined near x_0 with $\eta(x_0) = y_0$ such that $y = \eta(x)$ for all (x, y) near (x_0, y_0), so for all (x, y) near (x_0, y_0) we have:

$$g(x, \eta(x)) = 0$$

We also have that $f(x, y) = f(x, \eta(x))$ near (x_0, y_0) so $f(x, \eta(x))$ has an extremum at x_0 so:

$$\frac{d}{dx}(f(x, \eta(x)))\bigg|_{x=x_0} = 0$$

Which is the same as:

$$\frac{\partial f}{dx}(x_0, y_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{d\eta}{dx}(x_0) = 0 \tag{1}$$

As $g(x, \eta(x)) = 0$ for all x near x_0 we have

$$\frac{d}{dx}(g(x, \eta(x)))\bigg|_{x=x_0} = 0$$

Which is the same as:

$$\frac{\partial g}{dx}(x_0, y_0) + \frac{\partial g}{\partial y}(x_0, y_0) \cdot \frac{d\eta}{dx}(x_0) = 0 \tag{2}$$

As $\frac{\partial g}{\partial y}(x_0, y_0) \neq 0$ by assumption, set:

$$\lambda := \frac{\frac{\partial f}{\partial y}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)} \tag{3}$$

From equations (1), (2) and (3) we have:

$$\begin{aligned}
\frac{\partial f}{dx}(x_0, y_0) &= -\frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{d\eta}{dx}(x_0) \\
&= -\lambda \frac{\partial g}{\partial y}(x_0, y_0) \cdot \frac{d\eta}{dx}(x_0) \\
&= \lambda \frac{\partial g}{dx}(x_0, y_0)
\end{aligned}$$

Finally we have:

$$\begin{aligned}\nabla f(x_0,y_0) &= \left(\frac{\partial f}{dx}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0)\right)\\ &= \left(\lambda\frac{\partial g}{dx}(x_0,y_0), \lambda\frac{\partial g}{\partial y}(x_0,y_0)\right)\\ &= \lambda\nabla g(x_0,y_0)\end{aligned}$$

Therefore we have $\nabla(f-\lambda g)(x_0,y_0)=0$ and hence $f-\lambda g$ has a critical point at (x_0,y_0). □

Remark 5.2. The three important equations you will need, to find an extremum (x_0,y_0) of f on $X=\{(x,y):g(x,y)=0\}$ are:

$$\begin{aligned}\frac{\partial(f-\lambda g)}{\partial x}(x_0,y_0) &= 0\\ \frac{\partial(f-\lambda g)}{\partial y}(x_0,y_0) &= 0\\ g(x_0,y_0) &= 0\end{aligned}$$

The above theorem can we extended to functions f,g of several variables.

Theorem 5.3. Suppose $f,g\in C^1$ are functions of n variables $x_1,\dots,x_n$ and g is regular (i.e. $\nabla g\neq 0$). Using the notation $\mathbf{x}=(x_1,\dots,x_n)$. If $\mathbf{x_0}$ is an extremum of f on $X=\{\mathbf{x}\in\mathbb{R}^n : g(\mathbf{x})=0\}$ then there exists $\lambda\in\mathbb{R}$ such that $f-\lambda g$ has an unconstrained critical point at $\mathbf{x_0}$.

Theorem 5.4. Suppose $y(x)\in C^2$, $f,g\in C^2$ are functions of x,y,y'. If $y(x)$ extremises $I(y)=\int_{x_1}^{x_2} f(x,y,y')dx$ while $J(y)=\int_{x_1}^{x_2} f(x,y,y')dx=j_0$ where $j_0\in\mathbb{R}$, then there exists $\lambda\in\mathbb{R}$ such that $y(x)$ is a critical point of $I-\lambda J$. Or in other words $f-\lambda g$ satisfies its Euler-Lagrange equation.

Proof. Suppose y is an extremum of I on the set of functions $J(y)=j_0$. Let $u,v\in C^2$ be functions of x such that $u(x_1)=u(x_2)=v(x_1)=v(x_2)=0$ and define:

$$\begin{aligned}F_{u,v}(h,k) &:= I(y+hu+kv)\\ G_{u,v}(h,k) &:= J(y+hu+kv)-j_0\end{aligned}$$

As y is an extremum of I on the set of functions $J(y)=j_0$, $(0,0)$ is an extremum of $F_{u,v}$ on the set of h,k where $G_{u,v}(h,k)=0$. Then there exists $\lambda_{u,v}$ such that $F_{u,v}-\lambda G_{u,v}$ has a critical point at $(h,k)=(0,0)$. Therefore we get the following equations:

$$\begin{aligned}\frac{\partial}{\partial h}(F_{u,v}-\lambda_{u,v}G_{u,v})(h,k)\Big|_{(h,k)=(0,0)} &= 0\\ \frac{\partial}{\partial k}(F_{u,v}-\lambda_{u,v}G_{u,v})(h,k)\Big|_{(h,k)=(0,0)} &= 0\end{aligned}$$

This means we get the following equations:

$$\begin{aligned}\frac{d}{dh}\left(I(y+hu)-\lambda_{u,v}J(y+hu)\right)\Big|_{h=0} &= 0\\ \frac{d}{dk}\left(I(y+kv)-\lambda_{u,v}J(y+kv)\right)\Big|_{k=0} &= 0\end{aligned}$$

Therefore y satisfies the following equations:

$$\int_{x_1}^{x_2} u\left(\frac{\partial(f-\lambda_{u,v}g)}{\partial y}-\frac{d}{dx}\left(\frac{\partial(f-\lambda_{u,v}g)}{\partial y'}\right)\right)dx = 0 \tag{4}$$

$$\int_{x_1}^{x_2} v\left(\frac{\partial(f-\lambda_{u,v}g)}{\partial y}-\frac{d}{dx}\left(\frac{\partial(f-\lambda_{u,v}g)}{\partial y'}\right)\right)dx = 0 \tag{5}$$

From equation (4) we get:

$$\int_{x_1}^{x_2} u\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)dx = \lambda_{u,v}\int_{x_1}^{x_2} u\left(\frac{\partial g}{\partial y} - \frac{d}{dx}\left(\frac{\partial g}{\partial y'}\right)\right)dx$$

Pick u_0 such that $\int_{x_1}^{x_2} u_0\left(\frac{\partial g}{\partial y} - \frac{d}{dx}\left(\frac{\partial g}{\partial y'}\right)\right)dx \neq 0$. Then:

$$\lambda_{u_0,v} = \frac{\int_{x_1}^{x_2} u_0\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)dx}{\int_{x_1}^{x_2} u_0\left(\frac{\partial g}{\partial y} - \frac{d}{dx}\left(\frac{\partial g}{\partial y'}\right)\right)dx} \tag{6}$$

As the right hand side of equation (6) is independent of v, we can write $\lambda_{u_0,v} = \lambda$. Then for all $v \in C^2$ such that $v(x_1) = v(x_2) = 0$ we have:

$$\int_{x_1}^{x_2} v\left(\frac{\partial(f-\lambda g)}{\partial y} - \frac{d}{dx}\left(\frac{\partial(f-\lambda g)}{\partial y'}\right)\right)dx$$

Then by the Fundamental Theorem of Calculus of Variations we get $f - \lambda g$ satisfies its Euler-Lagrange equation and hence y is a critical point of $I - \lambda J$. $\square$