12 Introduction to the Calculus of Variations, optimization and an overview of variational techniques

Tact is the knack of making a point without making an enemy.
—Isaac Newton

12.1 Introductory concepts

The purpose of this half of this text is to combine the concepts from the first half, with the fundamental ideas and concepts that you previously covered (TODO: reference to Vol II) related to optimization. In earlier Chapters we introduced several dynamical models and discussed how to analyze those models and what their evolution could mean physically. In this half of the text we will demonstrate that differential equations (both ordinary and partial) are often derived as representative 'solutions' of some type of optimization principle. That is to say that the solutions of some differential equations yield states that satisfy some optimization principle. For example the harmonic oscillator can be derived as the state which minimizes the difference between the kinetic and potential energies of a spring, i.e. the equations of motion for many dynamical systems provide solutions that satisfy optimal principles and need not only be derived from Newtonian style formulations.

Optimization plays a role in many applications. One such example is driving between two locations and the question is how you choose the route to optimize your arrival time, or perhaps you are more concerned with how much fuel you expend, i.e. how can you get from one position to another expending the least amount of fuel possible? Other examples include optimizing stoplight patterns, placement of stop signs, and/or usage of roundabouts etc. in order to minimize the pollution from cars at idle (this is a very carefully analyzed problem for urban planning), and geese optimizing their flow configuration to reduce drag while migrating (it is a good thing that the geese paid attention when they took their dynamical modeling course as a young gosling).

One might also be concerned with minimizing the amount of gas used or time spent with the added constraint that there are a sufficient number of rest stops along the way so that the youngest passengers in the car can have a toilet stop (if you don't understand what this means yet, you will one day). Or perhaps, the geese are trying to optimize their flight pattern when the yearling goslings are unable to lead the formation. These two examples illustrate an optimization problem in which constraints are present, something that occurs often in real life. These different considerations define what we will refer to as a cost function. For instance if the goal is to minimize time the cost function may simply be the total time spent in transit, while on the other hand if it instead is desirable to minimize the gas, then the cost function may simply be the total amount of fuel required for the trip.

Another very different example is when George Lucas was producing the Phantom Menace. Evidently the cost function that was used was geared highly toward comedic interruptions to the plot, and relatively meaningless characters that would make high-selling action figures, hence Jar-Jar Binks was created. This cost function has been highly modified since Disney obtained LucasFilm, but the target has continued to be profitability both in the theatre and outside of it. We won't debate the merits of such an approach here nor the impact that has had on thousands of recent fans of the franchise, but it is evident what the optimization goal is from the end product.

In Calculus we learned how to optimize over a discrete set of points. For example, you may have found the maxima or minima of $f(x, y)$ on a domain $\Omega \subset \mathbb{R}^2$. In the Calculus of Variations we will find extrema (optimizers) from a class of functions (a bit bigger than Euclidean Space) for some functional $J : X \to \mathbb{R}$ that gives a real number for each function in a given class. In other words we are going to be looking for objects in a Banach space that are optimal (in some sense which we worry about later) with respect to some cost J (which we refer to as the cost functional as already mentioned above). This means that we are going to be optimizing over an infinite dimensional Banach space, rather than a finite dimensional space such as $\mathbb{R}^n$.

When we study Optimal Control later on we will consider optimization of a similar cost functional, but we will include the effects of a set of control variables that we can change to optimize our desired cost i.e. we are allowed to adjust the outcome by tweaking with the control variable to yield the best solution. An example of an optimal control problem would be to optimize fuel consumption for a trip across town where the control variable is the pressure you put on the gas pedal of the vehicle.

The primary difference between purely variational opimization problems, and control is the presence of a control variable. Variational problems suppose that the entire evolution of the system is predetermined if the system is specified precisely (this is a classic approach to deriving the equations that govern physical systems). Control problems instead suppose that although the evolution of the system is predetermined, there is a control variable that can be adjusted to reach the desired optimal goal. Typically control is used as a way of manipulating the physical environment with a specific choice of the control variable and how it adjusts to the other physical variables.

We will first focus on the Calculus of Variations, and later on introduce control variables. This is a choice for the presentation of this material, and should not be taken as a necessary approach. We emphasize that this natural progression is beneficial for the current presentation only.

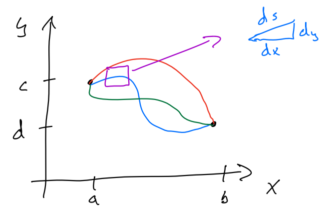

Figure 12.1: An illustration of finding the shortest path between two points in the plane as motivated in Example 12.1.1 for a rabbit trying to reach a patch of clover. Note that each of the curves depicted here will have a different total length. Of course the rabbit is interested in the curve of shortest length which leads to the variational statement given in Example 12.1.1.

Example 12.1.1. One of the most classic examples of a variational problem is to find the shortest curve connecting two points (a, c) and (b, d) in the plane. For instance it is important for a rabbit that spots some clover to reach that patch of clover as quickly as possible so that their siblings (which there are a lot of) don't get to it first. For any $y = y(x)$ the length of the curve will be $\int ds$ where ds is the infinitesimal length of the curve over any given infinitesimal interval. To make this tractable, we note that $ds^2 = dx^2 + dy^2$ in the plane so long as all of these terms are infinitesimally small (and hence the curve can be considered linear). This means that $ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}dx$.

Inserting this back into the integral, we see that the distance along the curve is actually given by

$$J[y] = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx.$$

In this case, we would like to minimize $J[y]$ where $y(x)$ is restricted to the two end-points, that is $y(a) = c$ and $y(b) = d$. At this point we have not really completed the specification of the problem however. We must specify what type of curves $y(x)$ are admissible, i.e. we would typically require $y(x)$ to be continuously differentiable, but maybe we are interested in an infinitely continuously differentiable curve instead. Once the class of admissible functions $y(x)$ is specified than the problem is completely stated. We need to keep in mind throughout this process what the problem is originally asking. For instance, if we are talking about the rabbit, does it make sense for the rabbit to travel along a discontinuous route, and if so what would this mean? In other words, can the rabbit teleport, or how are we modeling the fact that the rabbit is hopping in three-dimensional space but our model is on the 2D plane? What about a continuous route that is not differentiable, i.e. the rabbit has the ability to change directions instantaneously. If you have ever tried to chase a rabbit you may find yourself agreeing that continuous but not differentiable paths are perfectly reasonable. The simplest assumption is to have $y(x)$ be infinitely continuously differentiable but we should always re-examine this assumption for the particular setting we are investigating.

Remark 12.1.2. This is a classic example primarily because the optimization (maximization or minimization) of a functional of the form

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

appears very often in classical physics and applied mathematics. This forms what is often referred to as the simplest problem in the Calculus of Variations.

Example 12.1.3. Now suppose that the rabbit's path is chosen, and now the rabbit needs to decide how to expend its energy as it is running. If we let $u(t)$ be the energy expended in accelerating or decelerating, then we have a new optimization problem. In this case we are seeking to minimize the amount of time to reach the patch of clover where the rabbit's velocity (more likely acceleration) is some function of the energy expended. In this case it doesn't make sense for the rabbit to have an infinite amount of energy, so we expect that there is also some cost associated with $u(t)$, i.e. the rabbit doesn't want to use all its energy just to reach the clover patch, neither can it physically do so.

This problem represents an optimal control problem with control variable $u(t)$. The cost in this case is a bit more complicated than when we were just selecting the path we wanted to travel on. An example cost function in this case is that we seek to minimize

$$J[u] = t_f + \int_0^{t_f} u^2(s)ds,$$

where t_f is the final time that it takes the rabbit to run to the desired patch of clover.

Remark 12.1.4. Clearly the previous example is not complete. We have stated that we wish to minimize time t_f without stating how this time is related to our control variable $u(t)$ which we were using to denote the energy the rabbit can expend. To complete the statement of the problem we need to specify how the energy is related to the velocity, and hence the time it takes for the rabbit to reach its destination. This is considered in the next example where we make some explicit assumptions on the relationship between the energy expended $u(t)$ and the time it takes to reach the destination.

Example 12.1.5 (Continuation of the rabbit's energy). To complete this example of the running, hungry rabbit, we need to demonstrate how $u(t)$ and the final time t_f are related. Suppose that the rabbit must travel across a total distance D. Ignoring units for now (very bad form), if the rabbit is traveling at a velocity $x'(t)$ then we will make the assumption that the energy expended is directly proportional to the rabbit's acceleration, i.e. $x''(t) = u(t)$. Hence we can succinctly state our optimization problem as minimization of the following cost functional

$$J[u] = \int_0^{t_f} \left[1 + u(s)^2\right] ds,$$

under the constraints that

$$x''(t) = u(t), \quad x(0) = 0, \quad x(t_f) = D,$$

where D is the distance the rabbit must travel to reach the patch of clover. Note that we are assuming that the optimal path has already been chosen so that D is a minimal value already.

Remark 12.1.6. We can try to answer the rabbit's question of reaching the clover patch in the context of several different optimal control problems, the previous example being just one potential choice. For instance, there is no reason to suppose that the quadratic cost on $u(t)$ over the entire time interval is the most logical especially if we suppose that $u(t)$ is equivalent to the acceleration. Rabbits have several different gaits when they run, and some of these require more energy to accelerate in than others. In addition, it may make a difference what type of ground the rabbit is running on i.e. grass, rocks etc..

At a more mathematical level, why don't we consider a cost functional like $J[u] = \int_0^{t_f} [1 + |u(s)|]\, ds$ where $|\cdot|$ is the absolute value of the acceleration, or some other way of representing the cost of expending energy? Also, is it reasonable to just state that there is a maximum acceleration/energy level that can be expended, i.e. $U_l \leq u(t) \leq U_u$? There are a lot of options when setting this problem up, and each of these options will lead to a different type of solution although each one is a reasonable choice. We need to know something of the nature of the solution which is dependent on the type of cost functional assigned to the problem, before we can decide which formulation is the right one.

The skeptic (or the fox watching from the top of the hill) may also point out that our problem formulation has only forced the rabbit to reach the patch of clover at time t_f but has said nothing of the rabbit's velocity. If we don't consider the rabbit's velocity then it may very well be going so fast that it reaches the clover in record time but blows right through it without getting to eat any of it. Hence, we really should also require that the rabbit has zero velocity at the final time, i.e. $x'(t_f) = 0$, or at the very least we could require that $x'(t_f)$ be sufficiently small that the rabbit can consume some of the desired clover.

We now consider some additional examples that hopefully illustrate how we might set up these various optimization problems.

Example 12.1.7. Let $\mathbf{y}(\mathbf{x})$ be the velocity of a given fluid with constant density ρ, in a pipe. Then $K[\mathbf{y}] = \frac{1}{2\rho}\int_\Omega |\mathbf{y}(\mathbf{x})|^2 d\mathbf{x}$ is the kinetic energy of the fluid in the region Ω. We can try to maximize $K[\mathbf{y}]$ under certain constraints to find out how much energy may be in the flow and converted into a force acting on the pipe. This is an example of how these methods of optimization can be used to understand some fundamental processes in classical physics. The mentality in this approach is that nature tends to optimize things in some sense (we will return to this more frequently in this Part of the book). Note that as in the previous examples, we would need to specify what type of functions $\mathbf{y}(\mathbf{x})$ are permissible before the problem is properly stated, i.e. is the velocity continuous, differentiable etc.?

To make this even more specific, we may have an evolution equation for $\mathbf{y}(\mathbf{x})$ so that $\mathbf{y}(\mathbf{x}, t)$ is actually a function of time t as well. In this case maximizing the kinetic energy in some region may be dictated by a certain type of forcing given to the fluid. For instance, unstable buoyant effects can generate motion in the fluid. What types of buoyancy instability are most likely to maximize the kinetic energy?

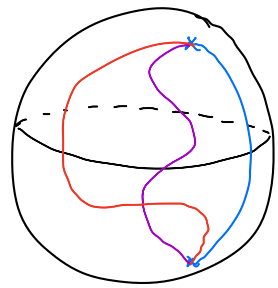

Figure 12.2: An illustration of finding the shortest path between two points on a sphere otherwise known as identifying the geodesics on the sphere as described in Example 12.1.8.

Example 12.1.8. The initial example we gave for finding the shortest path between two points on a plane is of significant interest, but what if we considered the shortest path constrained to a surface different from the plane? For instance, the Arctic Tern has been known to migrate from Greenland to Antarctica. The brightest of these birds will realize that they reside on an elongated sphere, and hence they are interested in finding the shortest path on that elongated sphere. The tern then looks for the shortest path between two points on a sphere (close enough) of radius R.

For such a sphere, $ds^2 = R^2 d\phi^2 + R^2 \sin^2(\phi) d\theta^2$ is the length of an infinitesimal distance on the surface of the sphere, and hence $ds = R\sqrt{(\frac{d\phi}{d\theta})^2 + \sin^2\phi}d\theta$. Thus the length of the path between (θ_1,ϕ_1) and (θ_2,ϕ_2) is given by

$$J[\phi(\theta)] = R\int_{\theta_1}^{\theta_2} \sqrt{\left(\frac{d\phi}{d\theta}\right)^2 + \sin^2\phi d\theta},$$

where once again we must determine how smooth we want ϕ to be as a function of θ to classify the functions we are optimizing over. For most terns, infinitely continuously differentiable paths are likely the best options, but there are of course exceptions to every species of bird and there may be some tern that is capable of flying a non differentiable or perhaps even non-continuous path (advanced birds maybe will achieve tern teleporting?).

Example 12.1.9. Returning to one of the more ambiguous examples considered above, one might classify the functional $J[y]$ as the amount of fuel required in driving from Boston to New York City. Minimizing $J[y]$ under certain constraints (the beginning point is in Boston, and the end is in New York, along with the speed limit and gas mileage of the vehicle) is the goal of the problem. Changes in this problem could be allowing for a choice of different transportation methods. For example, perhaps a climate conscious traveler is primarily concerned with their carbon footprint, and is not concerned with the required time to get from one place to another. In such a case, the functional $J[y]$ would consider the carbon footprint of all the possible types of transportation under the constraints that are inherent to each. In reality, this may be reasonable but likely with the additional constraint that the traveler must arrive in New York on a certain day or at least in a certain week. What type of functions y would be ideal in such a circumstance? Does continuity and differentiability even make sense then?

Remark 12.1.10. Now a few remarks are in order (of course). First, we note that in all of the examples above there was quite a bit of wiggle room in the choice of cost functional. Indeed, other than dealing with kinetic energy (something that has a firm definition) it seemed that we were arbitrarily choosing the form of the cost function, which is true to some extent. This is part of the 'modeling' in optimization problems. Something that you have already seen, even if you didn't realize it is that optimization of some functionals/functions is much easier than others.

The general concept is that you want to have a globally convex functional (a convex function in multivariable calculus) because then there is a unique minimizer. Sometimes you don't have a convex cost functional and there is nothing you can do about it, but there are also many times when you can adjust the cost function so it is convex. The easiest way to do this is to make the cost quadratic in all of the relevant variables. For example when we are considering the rabbit's expenditure of energy we anticipate that a quadratic cost in the energy will be easier to work with. There is no physical law that dictates to us that we can't assume a quadratic cost in the energy, so we may as well go with that.

For optimal control problems there is a specific class of problems when the cost is quadratic in all the variables, and the evolution equation is linear. These problems are referred to as the Linear Quadratic Regulator (LQR) and have a very efficient and satsifying solution that is not only efficient to compute, but is actually quite easy to derive ('easy' is a relative term here of course). The following example is just one example of the setup for an LQR problem.

Example 12.1.11. Consider poor Bob stuck on the damped oscillator. Suppose that Bob is interested in reaching a final point at a specific time, i.e. he would like to hop off of the spring at position $y(t) = 1$ at precisely time $t = t_f$. This means that he has the endpoint condition $y(t_f) = 1$. Now suppose that Bob has control of some forcing function for the spring (maybe he has a braking mechanism he can employ or some type of acceleration he can apply) so we will refer to this forcing function as the control variable u. The dynamics governing the spring were already discussed in this text, and are given by the linear system:

$$y''(t) + \omega^2 y(t) = u(t), \tag{12.1}$$

which we rewrite as the linear system

$$\mathbf{x}' = A\mathbf{x} + Bu, \tag{12.2}$$

where $\mathbf{x} = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}$, and the matrices A and B (B is actually just a vector in this setting) are given by:

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{12.3}$$

To augment this system with a cost, we recognize that the acceleration/braking mechanism is costly (in whatever sense we make of things) so that we don't want to allow for an infinite acceleration/braking, and we also want to ensure that Bob doesn't oscillate to wildly (he is sensitive to motion sickness). This means that we will seek to minimize the cost functional

$$J[u] = \int_0^{t_f} \left[Ru^2(t) + \alpha y(t)^2 + \beta y'(t)^2\right] dt, \tag{12.4}$$

with $y(t_f) = 1$. This cost can be rewritten as:

$$J[u] = \int_0^{t_f} [\mathbf{x}(t)^T Q\mathbf{x}(t) + u(t)Ru(t)]dt, \tag{12.5}$$

where $Q = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$.

This is the standard LQR setup, a linear evolution equation for the state and a quadratic cost in both the state $\mathbf{x}(t)$ and the control $u(t)$. In general the control may be vector valued so that R is actually a matrix, but in this simple case $u(t)$ is a scalar and thus so is R.

This approach may seem restrictive because we are requiring a linear evolution equation, but recall that most of the theory we developed for dynamically evolving systems relies on linearization. In reality many control problems can be linearized periodically and the linear state is then used to set up an LQR problem that can be solved very efficiently.

LQR is such a convenient setup that it is best practice in any control problem to avoid any other formulation of the problem if at all possible. Linearization is often a small cost compared to the gain that can be obtained by using LQR instead of a more expensive and often less accurate method.

12.2 Formalizing the Optimization Process

Now with some examples in place, we need to formalize exactly what we mean by an optimization problem, and we need to specify a class of functions to optimize over. In other words we need to create a mathematical setting under which the optimization problems can be performed. Recall the following from analysis (TODO: reference to Volume 1).

Definition 12.2.1. *$C^p[a,b]$ is the space of p-times continuously differentiable functions. This space is coupled with the norm given by $||f||_{C^p} = \sum_{i=0}^{p} ||f^{(i)}(x)||_\infty$, and this forms a normed linear space.*

Recall that in this context $f^{(i)}(x)$ refers to the ith derivative of $f(x)$. In other words this space is endowed with a norm that equally weights the maximum value of all derivatives up to order p. In the following chapters we will often refer to Banach spaces $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ over $\mathbb{R}$ but most frequently $X = Y = C^p[a,b]$ for some finite interval $[a,b]$.

Definition 12.2.2. *Let $J : X \to Y$ be a function and let $x, h \in X$. If the limit exists then we call*

$$\delta_h J(x) := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[J(x + \varepsilon h) - J(x) \right] = \frac{d}{d\varepsilon} \left. J(x + \varepsilon h) \right|_{\varepsilon=0} \tag{12.6}$$

the directional derivative of J at x in the direction h.

If the directional derivative of J at x exists for all $h \in X$ and changes continuously in h then we say that J is Gateaux differentiable at x.

Definition 12.2.3. *Let $J : X \to Y$ be a functional and let $x \in X$. If the directional derivative of J exists for all $h \in X$ and if there exists a bounded linear operator T_x such that $\delta_h J(x) = T_x h$ for all $h \in X$ then J is said to be Gateaux differentiable with Gateaux derivative T_x and Gateaux differential*

$$T_x h = \delta J(x; h) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[J(x + \varepsilon h) - J(x) \right] = \frac{d}{d\varepsilon} \left. J(x + \varepsilon h) \right|_{\varepsilon=0}. \tag{12.7}$$

In general we define the nth Gateaux differential of J at x as

$$\delta^n J(x; h) = \frac{d^n}{d\varepsilon^n} \left. J(x + \varepsilon h) \right|_{\varepsilon=0}, \tag{12.8}$$

whenever this differential exists for all $h \in X$ and can be computed as a bounded linear operator.

Definition 12.2.4. *A function $J : X \to Y$ is Fréchet differentiable at $x \in X$, if there is a bounded linear transformation $DJ(x) : X \to Y$ so that*

$$\lim_{\|h\|_X \to 0} \frac{\|J(x+h) - J(x) - DJ(x)h\|_Y}{\|h\|_X} = 0,$$

and $DJ(x)$ is called the Fréchet derivative of J at x. If J is differentiable at every point $x \in X$ then we say that J is differentiable on X. We call $DJ(x)h$ the Fréchet differential of J at x in the direction of h.

Remark 12.2.5. We emphasize at this point that the Fréchet derivative extends the concept of a derivative on functions of a single independent variable to derivatives on functionals where the independent variables are functions themselves.

To review some of the properties of the Fréchet derivative, recall that:

- If the Fréchet derivative exists then it is unique
- The definition inherently relies on the choice of the norm. However equivalent norms will give the same Fréchet derivative
- If J is Fréchet differentiable at x, then J is also Lipschitz and therefore continuous at x.
- When $Y = \mathbb{R}$ then the need for a bounded linear transformation can be replaced with a 'bounded linear functional'. This is the standard context that we will use in most of the rest of this book.

Example 12.2.6. Let $J : \mathbb{R}^2 \to \mathbb{R}^2$ be given by $J(x, y) = x^2 + 2yx$. We claim that

$$DJ(x,y) = \begin{pmatrix} 2x + 2y \\ 2x \end{pmatrix}$$
$$\delta J((x,y);(h_1,0)) = 2(x+y)h_1,$$
$$\delta J((x,y);(0,h_2)) = 2xh_2.$$

To see this, note that if $x, y, h_1, h_2 \in \mathbb{R}$ then

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left[J(x+\varepsilon h_1, y) - J(x,y)\right] = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left[(x+\varepsilon h_1)^2 + 2y(x+\varepsilon h_1) - (x^2 + 2yx)\right]$$
$$= 2(x+y)h_1.$$

A similar calculation holds for the differential with respect to the second coordinate.

Example 12.2.7. Let $J : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$J(x,y) = \begin{cases} 1 & \text{if } y = x^2, (x,y) \neq (0,0), \\ 0 & \text{otherwise.} \end{cases}$$

The functional J is Gateux differentiable at $(0, 0)$ but it is not Fréchet differentiable at $(0, 0)$.

This follows because along any line the Gateaux derivative is 0, but the limit in the definition of the Fréchet derivative does not exist along the path $(h_1, h_2)_n = \left(\frac{1}{n}, \frac{1}{2n^2}\left[1 + (-1)^n\right]\right)$ for $n \in \mathbb{N}$. To verify this, note that for even n, $h_{2,n} = h_{1,n}^2$ hence $J(h_{1,n}, h_{2,n}) = 1$ for even n but $J(h_{1,n}, h_{2,n}) = 0$ for odd values of n. If we try to calculate the Fréchet derivative at this $(0, 0)$ along this limit (which would necessarily be equal to the Gateaux derivative which is $\delta J = 0$ then we would be trying to verify that

$$\lim_{n\to\infty} \frac{\|J(h_{1,n}, h_{2,n}) - J(0,0)\|}{1/n} = 0,$$

but such a limit does not exist because $\lim_{n\to\infty} J(h_{1,n}, h_{2,n})$ does not exist.

In a more traditional context that typically appears in classical physics and most texts on the Calculus of Variations, the Fréchet differential with increment h is written as $\delta J(y; h) = DJ(y)h$. At this point it is important to stop and discuss what this actually means. $DJ(y)$ is an operator (linear and bounded at that) that acts on the functions $h(x)$ (note the shift from x to y). Thus when we write $DJ(y)h$ we are referring to the action of the linear operator $DJ(y)$ on the function $h(x)$. The result of $DJ(y)h$ is a real number for every specific function $h(x)$ meaning that $DJ(y)$ is a functional acting on $h(x)$.

For a functional $J : X \to \mathbb{R}$, the differential δJ is often referred to as the first variation of J. This differential in this context is written out as an integral functional that depends on both y and h, i.e. it takes on potentially different values in $\mathbb{R}$ depending on y and h.

Example 12.2.8. Consider the functional

$$J[y] = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

derived in the previous section to find the shortest route between two points on the plane. The first variation of J is

$$\delta J[y; h] = \int_a^b \frac{y'h'}{\sqrt{1 + (y')^2}} dx,$$

where $y' = \frac{dy}{dx}$ and $h' = \frac{dh}{dx}$.

To prove this, note that

$$J[y+h] - J[y] - \delta J[y; h] = \int_a^b \left[\sqrt{1 + (y' + h')^2} - \sqrt{1 + (y')^2} - \frac{y'h'}{\sqrt{1 + (y')^2}} \right] dx$$

$$= \int_a^b \left[\sqrt{1 + (y')^2 + 2(y'h') + (h')^2} - \sqrt{1 + (y')^2} - \frac{y'h'}{\sqrt{1 + (y')^2}} \right] dx$$

$$= \int_a^b \left[\sqrt{1 + (y')^2} + \frac{y'h'}{\sqrt{1 + (y')^2}} + O(\|h'\|^2) - \sqrt{1 + (y')^2} - \frac{y'h'}{\sqrt{1 + (y')^2}} \right] dx,$$

as h and h' get smaller and smaller, i.e. as $h \to 0$ in $C^1[a, b]$ (which is the definition of the Frechét derivative). In this setting, $\int_a^b O(\|h'\|^2)dx \to 0$ as $h \to 0$ in $X = C^1[a, b]$.

Remark 12.2.9. The first variation (Fréchet differential) is also often written as $\delta J(y; \delta y) = DJ\delta y$ where DJ is the Fréchet derivative. In other words, sometimes rather than using the increment function $h(x)$ we denote increments as $\delta y(x)$. Just keep in mind that these are still functions acting on x, and are elements of the infinite dimensional (typically Banach) space X.

Example 12.2.10. For a scalar valued function $f(\mathbf{x})$, and variation/increment $\mathbf{h}$,

$$\delta f(\mathbf{x}; \mathbf{h}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i,$$

where h_i refers to the ith component of the vector valued function $\mathbf{h}(\mathbf{x})$.

More specifically, suppose that $f(\mathbf{x}) = x_1^2 + x_2^2$ then for $\mathbf{h} = (h_1, h_2)^T$ then

$$\delta f(\mathbf{x}; \mathbf{h}) = 2x_1 h_1 + 2x_2 h_2.$$

Example 12.2.11. For the function space $X = C[0, 1]$ and functional

$$J(x) = \int_0^1 g(x(t))dt,$$

(where g_x is continuous) then

$$\delta J(x;h) = \int_0^1 g_x(x(t))h(t)dt.$$

More particularly, if we are looking at the specific cost functional

$$J[x] = \int_0^1 (x(t)^2)dt,$$

then

$$\delta J(x;h) = 2\int_0^1 x(t)h(t)dt.$$

Remark 12.2.12. It is clear that if the Fréchet differential exists then so does the Gateaux differential and they are equal. This does not go the other way however as pointed out in Example 12.2.7. There are special cases where the Gateaux derivative exists for certain directions h but the Fréchet derivative does not i.e. the Gateaux differential is valid in a certain direction, but may not be defined for all increments $h(x)$. For our purposes we will typically refer to the Gauteaux differential because it is more practical to compute. For a particular h we can consider the Gateaux differential $\delta J(y;h)$ as the directional derivative in the direction of h, that is $\delta J(y;h)$ is a measure of the rate of growth in the direction dictated by the increment h which typically is a function itself.

Example 12.2.13. Compute the Gateaux differential of

$$J[y] = \int_0^1 (y^2 + x)dx,$$

where $y(x) = x^2$ and $h(x) = x$.

This can be written out as

$$\begin{aligned}
\delta J[y;h] &= \frac{d}{d\varepsilon} J[y+\varepsilon h]\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\int_0^1 \left[(y+\varepsilon h)^2 + x\right] dx\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\int_0^1 \left[(x^2+\varepsilon x)^2 + x\right] dx\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\int_0^1 \left[x^4 + 2\varepsilon x^3 + \varepsilon^2 x^2 + x\right] dx\Big|_{\varepsilon=0} \\
&= \int_0^1 2x^3 dx \\
&= \frac{x^4}{2}\Big|_0^1 \\
&= \frac{1}{2}.
\end{aligned}$$

This example demonstrates what the Gateaux differential in a specific direction h and at a specific value y actually is, it is nothing more than a harmless number. But just as we normally don't want to compute the derivative of a function one number at a time, we would hope to be able to compute a Gateux differential more generically. The following example does exactly this for the same functional:

Example 12.2.14. Returning to the previous example, we now consider the generic formula for $\delta J[y; h]$ as follows:

$$\begin{aligned} \delta J[y;h] = \frac{d}{d\varepsilon}\int_0^1 \left\{(y+\varepsilon h)^2 + x\right\} dx\Big|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}\int_0^1 \left\{y^2 + 2\varepsilon hy + \varepsilon^2 h^2 + x\right\} dx\Big|_{\varepsilon=0} \\ &= \int_0^1 2yhdx. \end{aligned}$$

From the previous two examples we see that for each different $y(x)$ and direction $h(x)$, the Gateaux differential yields a different numerical value, which is consistent with the earlier observation that the Gateaux differential is akin to the directional derivative. This analogy is useful when we are trying to determine the optimizers of some cost functional. In the finite dimensional setting we found that the presence of a local minima/maxima would imply that the gradient vanishes at that point, or equivalently that the directional derivative vanishes for all possible directions. Before examining this idea for the infinite dimensional setting we first define what we mean by minima/maxima for a cost functional.

Definition 12.2.15. *Let $J : \Omega \to \mathbb{R}$ where $\Omega \subset X$ and X is endowed with norm $\| * \|_X$. The point $x_0 \in \Omega$ is a 'local minimum' of J on Ω if there is an $\varepsilon > 0$ such that $J(x_0) \leq J(x)$ for all $x \in \Omega \cap B(x_0, \varepsilon)$. The point x_0 is a 'strict local minimum' of J on Ω if $J(x_0) < J(x)$ for all $x \in \Omega \cap B(x_0, \varepsilon) \setminus \{x_0\}$.*

Relative Maxima are defined similarily. The set $\Omega \subset X$ is typically referred to as the admissible set. Local maxima and minima are referred to as local extrema. Now we are ready to see how local maxima/minima are related to a vanishing Gateaux differential.

Theorem 12.2.16. *Let $J : \Omega \to \mathbb{R}$ be Gateaux diffentiable at $x_0 \in \Omega$. Then a necessary condition for x_0 to be a local extremum is for $\delta J(x_0; h) = 0$ for all admissible h (meaning that $x + \varepsilon h \in \Omega$ for sufficiently small ε).*

Proof. For every admissible h, if x_0 is an extremum of $J(x)$ then $J(x_0 + \alpha h)$ (which is a function of a single variable) must also have an extremum at $\alpha = 0$. Thus from calculus we see that

$$\frac{d}{d\alpha} J(x_0 + \alpha h)|_{\alpha=0} = 0.$$

□

Recall in multivariable calculus there were conditions on the second derivatives that gave ‘sufficient’ conditions for a max/min. There are similar circumstances here, but they are quite complicated and not very practical for computational purposes, so we will leave that discussion to later. Most often, physical considerations indicate a type of sufficient condition, or careful construction of the cost functional (guaranteeing convexity for example) can also provide a sufficient condition.

Exercises

Note to the student: Each section of this chapter has several corresponding exercises, all collected here at the end of the chapter. The exercises between the first and second line are for Section 1, the exercises between the second and third lines are for Section 2, and so forth.

You should **work every exercise** (your instructor may choose to let you skip some of the advanced exercises marked with *). We have carefully selected them, and each is important for your ability to understand subsequent material. Many of the examples and results proved in the exercises are used again later in the text. Exercises marked with ⚠ are especially important and are likely to be used later in this book and beyond. Those marked with † are harder than average, but should still be done.

Although they are gathered together at the end of the chapter, we strongly recommend you do the exercises for each section as soon as you have completed the section, rather than saving them until you have finished the entire chapter.

12.1. Consider the functional $J(y) = \int_0^1 (1+x)(y')^2 dx$, where $y \in C^2[0,1]$ and $y(0) = 0,\ y(1) = 1$. Of all functions of the form $y(x) = x + c_1 x(1-x) + c_2 x^2(1-x)$, where c_1 and c_2 are constants, find the one that minimizes J.

12.2. Suppose that the rabbit isn’t trying to reach a patch of clover, but is instead trying to grab a carrot off the ground at a distance P away. This means that the rabbit doesn’t need to come to a complete stop when it reaches the carrot, but shouldn’t be going to fast. What type of conditions would this introduce to the optimal control problem?

12.3. Rather than minimizing the time to reach the patch of clover, the rabbit is really concerned with minimizing the amount of time it is exposed to hawks overhead. If this only happens for certain values of x e.g. $5 \le x \le 6$ then how could you modify the optimal control problem to account for that?

12.4. Following Example 12.1.8, suppose that trade winds tend to blow the Arctic tern off course. Modify the cost functional for the shortest path between two points on the sphere, if it is harder for the tern to change ϕ than it is to change θ.

12.5. Specify a cost function for Bob to exit the spring if he is extremely sensitive to rapid motions, i.e. he wants to keep $|y'(t)|$ as small as possible. What if he really doesn’t care about being jerked around, i.e. $|y'(t)|$ can be quite large?

12.6. For $x(t) \in C[0,1]$ define

$$J(x) = \max_{0 \le t \le 1} x(t).$$

Determine for which class of functions $x(t)$ the Gateaux differential $\delta J(x; h)$ exists and is linear in h. *HINT: The Gateaux differential will depend explicitly on $h(t)$ but not on $x(t)$.*

12.7. Prove that a *linear* functional $J : X \to \mathbb{R}$ cannot have an extremum unless $J(x)$ is constant for all $x(t)$.

12.8. Consider the functional $J(y) = \int_0^1 (1+x)(y')^2 dx$ where $y \in C^2[0,1]$ and $y(0) = 0$ and $y(1) = 1$. Show that the Gateaux differential $\delta J(y_0; h) = 0$ for all $h \in C^2[0,1]$ with $h(0) = h(1) = 0$, where $y_0(x) = \log(1+x)/log(2)$.

12.9. Compute the Gateaux differential $\delta J(y; h)$ of $J(y) = \int_0^1 (3y^2 + x)dx + \frac{1}{2}y(0)^2$, where $y = x$ and $h = x + 1$.

12.10. Let $\phi(t) \in C^1[0,1]$ and define the functional $J[y] = \int_0^1 \phi(y(x))dx$. Show that the Gateaux differential exists and is given by $\delta J(y; h) = \int_0^1 \phi'(y(x))h(x)dx$.

12.11. Show that the first variation (Gateaux differential) satisfies the homogeneity condition $\delta J[y; \alpha h] = \alpha \delta J[y; h]$ for $\alpha \in \mathbb{R}$.

Notes