Lecture 7

Let's consider some examples.

Examples: conservation laws (revisited)

1. "no y" case from before: $L = L(x, y')$.

In new notation: $L = L(t, \dot{q})$, no q, i.e., invariance under parallel displacements. $L_q = 0$ and this is the force (why? because $L = T - U$, T is independent of q, so $L_q = -U_q =$ force). So, this case is when the system is *closed*, i.e., no forces are acting on it.

We saw earlier that *momentum* ($L_{\dot{q}}$) *is conserved*.

For multiple degrees of freedom, the general statement is that for every coordinate q_i that L does not depend on, the corresponding momentum $L_{\dot{q}_i}$ is conserved.

2. "no x" case from before: $L = L(y, y')$.

In new notation: $L = L(q, \dot{q})$, no t, i.e., invariant under time shift.

Such a system is called *conservative* [Arnol'd, Feynman], and we saw earlier that Hamiltonian = *total energy is conserved*.

In mechanics, $L = \frac{m\dot{q}^2}{2} - U$, so basically this case is always true (as long as U exists).

3. Conservation of angular momentum.

Consider planar motion in central field: (r, φ) are polar coordinates, $U = U(r)$ is independent of angle. This means that *no torques* are acting on the system, i.e., invariance under rotations.

(Proved below) "no φ" case $\Rightarrow$ as before, can show that the corresponding component of the momentum, $L_{\dot{\varphi}}$, is conserved.

(Use Euler-Lagrange for multiple degrees of freedom, recall its coordinate invariance: the second equation is

$$0 = L_\varphi = \frac{d}{dt} L_{\dot{\varphi}}$$

There is another equation for r, we don't need it.)

$L = T - U$.

$$T = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2} m(\dot{r}^2 + r^2 \dot{\varphi}^2)$$

in polar coordinates.

$$L_{\dot{\varphi}} = T_{\dot{\varphi}} = mr^2 \dot{\varphi} = m(x\dot{y} - y\dot{x})$$

back in Cartesian coordinates. This is *angular momentum*, and it is conserved.

Note: all the above examples are instances of *Noether's Theorem*, which says that invariance of the action integral under some transformation (e.g., time shift, translation, rotation) implies the existence of a first integral (conserved quantity). Read G-F, Sect. 20–; very nice and clear application of variation of a functional.

2.5 Variational problems with constraints

Recall: Lagrange multipliers (in finite dimensions).

References: G-F, p. 42–; MacCluer; Young, p. 215–.

2.5.1 Integral constraints

Problem: (2.2) for $\mathcal{C}^1$ curves $y(x)$ satisfying (2.3) (some other versions are possible here) and the constraint

$$C(y) = \int_a^b M(x, y, y')dx = c_0 \quad \text{(given constant)}$$

Here C stands for "constraint," and M is another ($\mathcal{C}^2$) function. For simplicity, we're considering the case of only one constraint.

Examples—already saw: catenary, Dido.

Assume that $y(x)$ is an extremum.

What follows is a *heuristic argument* based on the earlier finite-dimensional proof of Lagrange multipliers.

Let's consider a perturbation

$$y + \alpha\eta$$

If it preserves the constraint (i.e., if this perturbation is admissible), then we have

$$\delta C|_y(\eta) = 0$$

and we know that this means

$$\int_a^b \left(M_y - \frac{d}{dx}M_{y'}\right)\eta(x)dx = 0$$

For every such η, we must have

$$\delta J|_y(\eta) = \int_a^b \left(L_y - \frac{d}{dx}L_{y'}\right)\eta(x)dx = 0$$

In other words, we must have

$$\int_a^b \left(L_y - \frac{d}{dx}L_{y'}\right)\eta(x)dx = 0 \qquad \forall\,\eta \;\text{ such that }\; \int_a^b \left(M_y - \frac{d}{dx}M_{y'}\right)\eta(x)dx = 0 \tag{2.8}$$

This is similar to what we had in the finite-dimensional case—recall (1.5).

What does this suggest?

There exists a constant λ (a *Lagrange multiplier*) such that

$$\left(L_y - \frac{d}{dx}L_{y'}\right) + \lambda\left(M_y - \frac{d}{dx}M_{y'}\right) = 0$$

In other words, the two variations are multiples of each other. This is equivalent to saying that we have Euler-Lagrange for an augmented Lagrangian $L + \lambda M$, i.e., $y(x)$ is an extremal of the functional

$$\int_a^b (L + \lambda M) dx$$

A unique extremal is determined by two boundary conditions plus one constraint.

But wait! In the finite-dimensional case, we had a technical condition (x^* had to be a regular point, see Exercise 2). It turns out that here we need a technical condition too. Namely, we need to assume that the test curve $y(x)$ is not an extremal of the constraint functional C (i.e., Euler-Lagrange for C doesn't hold along $y(x)$). This means there should exist nearby curves such that $C(y)$ takes values both $> c_0$ and $< c_0$. (For instance, consider the length constraint: small perturbations can both increase and decrease the length.) To see why we need it, consider this example: $C = \int_0^1 \sqrt{1 + (y')^2} dx = 1$, $y(0) = y(1) = 0$, only one admissible curve, hence no necessary condition. If this assumption fails, then the second expression in (2.8) is 0 for *every* η. If (2.8) were true, then it would imply that $y(x)$ must be an extremal for L, but the above example clearly shows that this is not necessary.

We can now *conjecture* the following *necessary condition for constrained optimality:* $y(x)$ is either an extremal for M (i.e., satisfies Euler-Lagrange for M) or an extremal for $L + \lambda M$, for some $\lambda \in \mathbb{R}$.

Can actually combine these two statements into one: $y(x)$ must satisfy Euler-Lagrange for

$$\lambda_0 L + \lambda M$$

where λ_0, λ are constants, not both zero.

λ_0 is an *abnormal multiplier*. We'll see this again in optimal control.

It turns out that this conjecture is correct! However, the above argument is faulty:

- One gap was that we didn't make formal the step of passing to λ;
- Another gap was that $\delta C|_y (\eta) = 0$ was *necessary* for the perturbation to be admissible, but we don't know if it's sufficient! We omitted this step in the finite-dimensional case also (see Luenberger).

Here, we would need to show that, given η such that $\delta C|_y (\eta) = 0$, we can construct a perturbation family, perhaps of the form

$$y + \alpha\eta + \text{h.o.t.}$$

along which $C = c_0$.

There is, however, an easier alternative fix:

Exercise 7 (due Feb 14)

Write down a correct proof by considering a two-parameter family of perturbations

$$y + \alpha_1 \eta_1 + \alpha_2 \eta_2$$

and using the Inverse Function Theorem.

Several constraints $\Rightarrow \lambda_1, \dots, \lambda_m$ [G-F].

Lagrange's original thinking was: replace minimization of J with respect to x by minimization of

$$\int L + \lambda \left(\int M - c_0 \right)$$

with respect to $y(x)$ and λ (cf. the discussion in the finite-dimensional case).

2.5.2 Non-integral constraints

Suppose the constraint is given by just

$$M(x, y, y') = 0 \tag{2.9}$$

where M is a (possibly vector-valued) function.

$\longrightarrow$ For now, take M scalar-valued for simplicity.

Let $y(x)$ be a test curve.

So, can we guess the necessary condition for constrained optimality in this case?

Answer: same as before, but the Lagrange multiplier is now a function of x: Euler-Lagrange holds for the augmented Lagrangian

$$L + \lambda(x)M$$

As before, we can introduce the abnormal multiplier: work with

$$\lambda_0 L + \lambda(x)M$$

Again, we need a technical assumption to rule out abnormal cases. Following G-F and MacCluer, can assume that $M_y \neq 0$ everywhere along the curve. This guarantees that we can perturb the curve anywhere and increase or decrease M.

We won't give a proof (a bit hard), but let's discuss the reason why $\lambda = \lambda(x)$.

The integral constraint $\int_a^b M dx = c_0$ is "global"—it applies to the whole curve. In contrast, the non-integral constraint $M = 0$ is "local"—it applies to each point on the curve. Locally,

$$\int_x^{x+\Delta x} M dx \approx M \Delta x$$

so locally there is no difference between the two. This is one way to see the correspondence.

Another way: if $M(x, y, y') = 0$, then

$$J = \int L dx = \int (L + \lambda(x)M) dx$$

for every function $\lambda(\cdot)$. Then we can consider minimization of this new functional with respect to $y(x)$ and $\lambda(x)$. This was Lagrange's original idea. Note that M is still integrated along the curve (we need to get a cost for each curve). In the integral case, λ was outside and constant, but now it is inside the integral and depends on x.

Suppose we want to solve the constraint (2.9) for y' as a function of x and y. In general, M gives fewer constraints than the dimension of y', i.e., the system is under-determined. Then we have some free parameters—call them u, and we get

$$y' = f(x, y, u)$$

which is a control system!

Example 3 $n = 2$, $M = y_1' - y_2$, then we get

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= u \quad \text{(free)} \end{aligned}$$

□

And the case of no constraints (the case we started with) then corresponds to $y' = u$.

Much more on this later, the above was just a brief introductory discussion. We will revisit this idea in optimal control [A-F]. There, $\lambda(x) \leftrightarrow p$ (costate), hard to see now.

Holonomic constraints

Suppose that the constraint function M doesn't depend on y':

$$M(x, y) = 0$$

or that we can integrate the constraints to get them to this form. Then we have a constraint surface, and we have two options:

1) Use the Lagrange multiplier condition

2) Use the constraints to find fewer independent variables and reformulate the problem in terms of these variables.

The second option might be easier. The next example illustrates this.

Example 4 Pendulum [MacCluer, pp. 37 and 83]

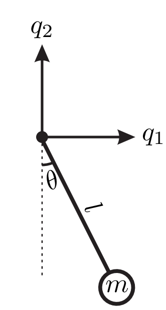

Figure 2.11: Pendulum

We want to derive trajectories of motion of this system, using the principle of least action.

Take $l = 1$, $m = 1$.

Since this is a mechanical example, let us use the notation $t, q, \dot{q}$.

Attach coordinates q_1, q_2 to the pivot point.

Kinetic energy: $T = \frac{1}{2}(\dot{q}_1^2 + \dot{q}_2^2)$.

Potential energy: $U = q_2$ (normalizing g to 1). (Or could let $U = 1 + q_2$ to get $U = 0$ in the downward equilibrium.)

Recall the principle of least action: $L = T - U = \frac{1}{2}(\dot{q}_1^2 + \dot{q}_2^2) - q_2$.

Constraint: $M(q) = q_1^2 + q_2^2 - 1 = 0$ (holonomic).

Lagrange multiplier condition:

$$\widetilde{L} := L + \lambda(t)M = \frac{1}{2}(\dot{q}_1^2 + \dot{q}_2^2) - q_2 + \lambda(t)(q_1^2 + q_2^2 - 1)$$

Need to write Euler-Lagrange for this:

$$\frac{d}{dt}\widetilde{L}_{\dot{q}_i} = \widetilde{L}_{q_i}, \qquad i = 1, 2$$

We get:

$$\frac{d}{dt}(\dot{q}_1) = \ddot{q}_1 = 2\lambda(t)q_1$$
$$\frac{d}{dt}(\dot{q}_2) = \ddot{q}_2 = 2\lambda(t)q_2 - 1$$

These are equations of motion for the pendulum, but we don't know $\lambda(t)$ so this is not very helpful.

Let's instead take a different approach and work with the angle θ as a single degree of freedom. We get the problem in terms of θ, this is an unconstrained problem.

$T = \frac{1}{2}\dot{\theta}^2$ (cf. earlier angular momentum example). To check, just substitute $q_1 = r\cos\theta$ and $q_2 = r\sin\theta$ and use the fact that $r \equiv 1$.

$U = -\cos\theta$ (or $1 - \cos\theta$ if we want 0 when $\theta = 0$).

$L = \frac{1}{2}\dot{\theta}^2 + \cos\theta$.

Euler-Lagrange :

$$\frac{d}{dt}(\dot{\theta}) = \ddot{\theta} = -\sin\theta$$

which is the familiar pendulum equation!

□

$\longrightarrow$ In control theory, we will be interested in the opposite case (nonholonomic) where constraints $M(x, y, y')$ cannot be integrated, i.e., the system can be driven to any configuration.

The above Lagrange multiplier condition (for the case of integral constraints) can be used, for example, to solve Dido's isoperimetric problem and also to solve the catenary problem.

Exercise 8 (due Feb 21) Derive solutions of Dido's isoperimetric problem and of the catenary problem.

Hint: for catenary, use the "no x" form of Euler-Lagrange equation.

2.6 Second-order conditions

(First necessary, then sufficient; first weak minima, eventually move to strong.)

Reference: G-F, Chap. 5.

So far, we only worked with the first variation. Recall the second-order expansion (same as in Section 1.3 but with slightly different notation):

$$J(y+\alpha\eta) = J(y) + \delta J\big|_y(\eta)\alpha + \delta^2 J\big|_y(\eta)\alpha^2 + o(\alpha^2)$$

Here $\delta^2 J\big|_y(\eta)$ is the second variation; $\delta^2 J\big|_y$ is a quadratic form (see Section 1.3.3).

Recall the second-order optimality conditions.

Necessary: y is a weak min $\Rightarrow \delta J\big|_y = 0,\ \delta^2 J\big|_y(\eta) \geq 0\ \forall\,\eta$.

Sufficient: $\delta J\big|_y = 0,\ \delta^2 J\big|_y(\eta) \geq c\|\eta\|^2$—or something of this sort, to dominate h.o.t.

Of course, conditions for maxima are the same but with ≤ 0 instead of ≥ 0.

2.6.1 Legendre's necessary condition for a weak extremum

We'll study necessary conditions first (weak minima). Go back to the basic calculus of variations problem (2.2), (2.3).

Single degree of freedom for now.

Smoothness: assume $L \in \mathcal{C}^3$ (third derivatives will appear).

Perturbation: since we are in the (standard) case of fixed endpoints, $y+\alpha\eta$, $\eta(a)=\eta(b)=0$.

Let's compute $\delta^2 J$. The second-order Taylor expansion of $L(x, y+\alpha\eta, y'+\alpha\eta')$ gives

$$\begin{aligned} J(y+\alpha\eta) = \int_a^b L(x, y+\alpha\eta, y'+\alpha\eta')dx = J(y) + \int_a^b \left(L_y\eta + L_{y'}\eta'\right)dx\cdot\alpha \quad \text{(first variation)} \\ + \frac{1}{2}\int_a^b \left(L_{yy}\eta^2 + 2L_{yy'}\eta\eta' + L_{y'y'}(\eta')^2\right)dx\cdot\alpha^2 + o(\alpha^2) \quad \text{(second variation)} \end{aligned}$$

This is the second variation—it's a quadratic form which involves η and η'.

If the above step was not clear: differentiating with respect to α,

$$L(y+\alpha\eta) \longrightarrow L_y(y+\alpha\eta)\eta \longrightarrow L_{yy}(y+\alpha\eta)\eta^2$$

then set $\alpha = 0$. Same idea for y'.

As we did with Euler-Lagrange and the first variation, let us integrate by parts (just one of the terms):

$$\int_a^b 2L_{yy'}\eta\eta' dx = \int_a^b L_{yy'}\frac{d}{dx}(\eta^2)dx = L_{yy'}\eta^2\big|_{x=a}^b - \int_a^b \frac{d}{dx}(L_{yy'})\eta^2 dx$$

The first term on the right-hand side is 0 by the boundary conditions. So, the second variation can be written as

$$\delta^2 J\big|_y(\eta) = \int_a^b \frac{1}{2}\left(L_{y'y'}(\eta')^2 + \left(L_{yy} - \frac{d}{dx}L_{yy'}\right)\eta^2\right) dx =: \int_a^b \left(P(x)(\eta')^2 + Q(x)\eta^2\right) dx$$

where

$$P := \frac{1}{2}L_{y'y'}, \qquad Q := \frac{1}{2}\left(L_{yy} - \frac{d}{dx}L_{yy'}\right).$$

Remark 2 By Taylor's theorem with remainder (see, e.g., Rudin), the h.o.t. $o(\alpha^2)$ take the form (go back to the derivation)

$$\frac{1}{2}\left(L_{yy}(x, y + \bar{\alpha}\eta, y' + \bar{\alpha}\eta') - L_{yy}(x, y, y')\right)\eta^2 + \dots, \qquad \bar{\alpha} \in [0, \alpha]$$

and we can integrate by parts to make it into the form similar to the second variation:

$$o(\alpha^2) = \int_a^b \left(\bar{P}(x,\alpha)(\eta')^2 + \bar{Q}(x,\alpha)\eta^2\right) dx \cdot \alpha^2$$

where $\bar{P}(x,\alpha), \bar{Q}(x,\alpha) \to 0$ as $\alpha \to 0$. This will be needed later (in the proof of the sufficient condition for a weak minimum). □

So, if $y(x)$ is a weak minimum then we must have

$$\int_a^b \left(P(x)(\eta')^2 + Q(x)\eta^2\right) dx \geq 0 \qquad \forall \eta$$

We would like to restate it in terms of just P and Q, not to have to check it for all η (cf. Euler-Lagrange derivation).

What does the above inequality imply? Does it mean that we must have $P(x) \geq 0$? $Q(x) \geq 0$? Both? Note that η and η' are related.

Let's see if maybe one of them can be negative, but the other must be non-negative. Which of the terms is more important? Can it happen that η is large (in magnitude) while η' is small? Can it be the other way around?

Pick $\varepsilon > 0$ and consider the perturbation $\eta_\varepsilon(x)$:

$\longrightarrow$ Each pulse has width $\sim \varepsilon$ and height $\sim 1/\varepsilon$. Can find formulas for these functions in [G-F, p. 103].

Then

$$\int_a^b Q(x)\eta^2(x)dx \approx \int_{x_0}^{x_0+\Delta x} Q(x)dx \approx Q(x_0)\Delta x$$

is not really affected by ε.

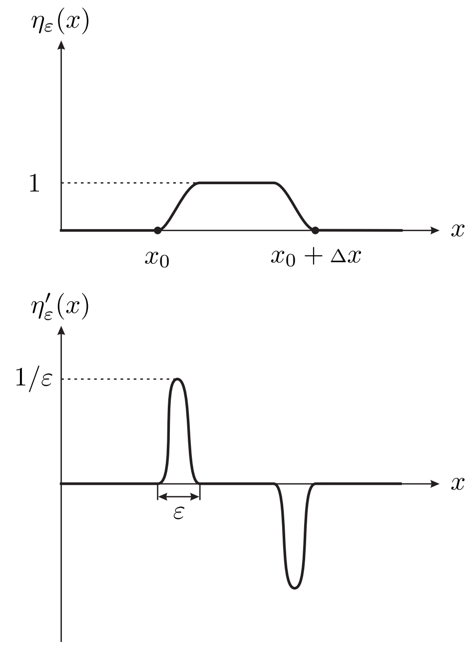

Figure 2.12: The graphs of η_ε and its derivative

On the other hand,

$$\int_a^b P(x)(\eta'(x))^2 dx \approx 2P(x_0)\varepsilon\frac{1}{\varepsilon^2}$$

blows up for small ε. So, if $P(x_0) < 0$ for some x_0, then the second variation cannot be ≥ 0.

We have just proved a second-order necessary condition for a weak minimum—*Legendre's condition* (1786, according to Sussmann). Recalling the definition of P in terms of L:

$$\boxed{L_{y'y'}(x, y(x), y'(x)) \geq 0 \qquad \forall\, x \in [a, b]} \tag{2.10}$$

It says nothing about Q, which is ugly anyway.

$\longrightarrow$ We will revisit this in the variational approach to optimal control [A-F, p. 264].

Digression: recall the "control" Hamiltonian

$$H(x, y, y', p) = py' - L(x, y, y')$$

where $p := L_{y'}$. We noted earlier that $H_{y'} = p - L_{y'} = 0$. Legendre's condition gives $H_{y'y'} = -L_{y'y'} \leq 0$. This confirms the idea that H attains a maximum as a function of y' along the optimal trajectory! Much more on this later.

$\longrightarrow$ Max vs. min is just a matter of notation: if we define $H := py' + L$, we get a min (this is common convention in optimal control).

What about *multiple degrees of freedom?*

$$\delta^2 J\big|_y(\eta) = \int_a^b \left((\eta')^T P(x)\eta' + \eta^T Q(x)\eta\right) dx$$

where $P(x) = \frac{1}{2}L_{y'y'}$ is the *Hessian matrix.* Legendre's condition still says $P(x) \geq 0$, but in matrix sense now. The derivation is the same—read [G-F, Sect. 29].

OK, so

$$\text{weak min} \ \Rightarrow \ \delta^2 J\Big|_y = \int (P(\eta')^2 + Q\eta^2)dx \geq 0 \ \Rightarrow \ L_{y'y'} \geq 0$$

2.6.2 A sufficient condition for a weak extremum

What about a *sufficient condition* for a weak minimum? We need at least $\int (P(\eta')^2 + Q\eta^2)dx > 0 \ \forall \eta$ (plus need some uniformity to dominate $o(\alpha^2)$).

When does the above strict inequality hold?

A natural guess is $P(x) > 0 \ \forall x$. Let us assume this. We will see that it is not enough. Legendre tried to prove that $P > 0$ implies $\delta^2 J > 0$ using the following idea: for some function $w(x)$, use the identity

$$0 = \int_a^b \frac{d}{dx}(w\eta^2) = \int_a^b (w'\eta^2 + 2w\eta\eta')dx$$

(recall that $\eta(a) = \eta(b) = 0$) to get

$$\int_a^b \left(P(\eta')^2 + Q\eta^2\right) dx = \int_a^b \left(P(\eta')^2 + 2w\eta\eta' + (Q + w')\eta^2\right) dx$$

Now, try to find w such that this is a *perfect square*. For this we need $\sqrt{P}\sqrt{Q + w'} = w$, or

$$P(x)\left(Q(x) + w'(x)\right) = w^2(x) \tag{2.11}$$

which is an ODE for w. This is a quadratic ODE, of *Riccati* type.

Lecture 9

Let us assume that we solved (2.11) and have w. Then our second variation can be written as

$$\int_a^b \left(\sqrt{P}\eta' + \frac{w}{\sqrt{P}}\eta\right)^2 dx = \int_a^b P\left(\eta' + \frac{w}{P}\eta\right)^2 dx$$

(recall the assumption that $P(x) > 0$, so dividing by P is OK). This is clearly ≥ 0.

Claim: This is $> 0 \; \forall \, \eta \not\equiv 0$.

Proof. If it is 0, then $\eta'(x) + \frac{w(x)}{P(x)}\eta(x) \equiv 0$. Also, we have $\eta(a) = 0$. But there is only one solution of this first-order ODE with this initial condition, and this solution is $\eta \equiv 0$. □

So, it seems that we have $\delta^2 J\big|_y(\eta) > 0 \; \forall \, \eta \not\equiv 0$. This is what Legendre originally thought. Very clever, but wrong. Why?

The Riccati differential equation may have a finite escape time, i.e., the solution $w(x)$ may not exist on $[a, b]$!

For example, $w' = -1 - w^2$ has the solution $w(x) = \tan(c - x)$ where c depends on the initial condition (this is for $P = -1$, $Q = 1$). This blows up for $c - x = \pm k\pi/2$ so we are in trouble if $b - a > \pi$. Similarly, $w' = w^2 + 1$ has solution $w = \tan(x - c)$ (this is for $P = 1$, $Q = -1$). Or for $P = 1$, $Q = 0$ we get $w' = w^2$, also blows up in finite time.

In fact, it is clear that a sufficient condition should involve some "global" considerations, because a concatenation of optimal curves might not be optimal. (Here, "global" means a condition on the whole curve as opposed to "pointwise" in x; the minimum is still local among nearby curves.)

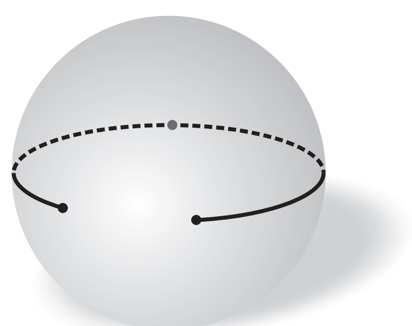

Figure 2.13: Concatenation of shortest-distance curves on a sphere is not shortest-distance

The concatenation is no longer optimal—in fact, there exist nearby curves that are shorter—but it would still satisfy any local condition that is satisfied by the two pieces.

So, we need to know when the Riccati DE (2.11) has a solution on the whole interval $[a, b]$.

Standard trick: reduce this quadratic ODE to another ODE, linear but of order twice the given one (in this case, 2) by making the substitution

$$w(x) = -\frac{Pu'(x)}{u(x)} \tag{2.12}$$

where $u(x)$ is a new (unknown) function, not equal to 0 anywhere.

$$w' = -\frac{\frac{d}{dx}(Pu')u - P(u')^2}{u^2}$$

so (2.11) gives

$$P\left(Q - \frac{\frac{d}{dx}(Pu')u - P(u')^2}{u^2}\right) = \frac{P^2(u')^2}{u^2}$$

Multiplying by $u(\neq 0)$:

$$P\left(Qu - \frac{d}{dx}(Pu')\right) = 0$$

Divide by $P(> 0)$:

$$\boxed{\tfrac{d}{dx}(Pu') = Qu}$$

This is the *accessory*, or *Jacobi*, equation. We will be done if we can find a solution $u(x)$ of the accessory equation which does not vanish anywhere on $[a, b]$, because then we can obtain a solution w to (2.11) via (2.12).

Let's take the point $x = a$ and consider initial data $u(a) = 0$, $u'(a) = 1$ (second-order $\Rightarrow$ need to specify u and u', $u' = 1$ is no loss of generality because if u is a solution then cu, $c \in \mathbb{R}$ is also a solution). The point $x = c > a$ at which u hits 0 again is called the *conjugate point* (with respect to a). This point depends only on P and Q, which is the data of the original variational problem.

Now, suppose that $[a, b]$ has no points conjugate to a. Then we can take $x = a - \varepsilon$, $\varepsilon > 0$ arbitrarily small, and consider initial data $u(a - \varepsilon) = 0$, $u'(a - \varepsilon) = 1$. By continuity with respect to initial data for solutions of ODEs, the corresponding solution is not zero anywhere in the interval $[a, b]$.

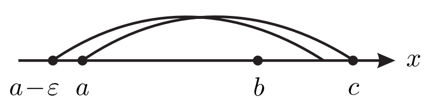

Figure 2.14: Conjugate point

$\longrightarrow$ So, in addition to the Legendre condition (2.10), let us assume that $[a, b]$ has no points conjugate to a.

We proved that if $P > 0$ and there are no points conjugate to a on $[a, b]$, then $\delta^2 J\big|_y$ is positive definite (this is Theorem 1 in G-F, p. 106).

Begin optional:

It can also be shown that absence of conjugate points on $[a, b]$ is not only sufficient but also necessary for $\delta^2 J\big|_y > 0$ (this is Theorem 2 in G-F, p. 109). And if $\delta^2 J\big|_y \geq 0$ (nonnegative definite), then there are no conjugate points on (a, b).
End optional

Conjugate points have many interesting interpretations, and their theory is outside the scope of the course (read G-F). Just mention this briefly [G-F, p. 115]:

Conjugate point is where different (neighboring) extremals starting at the same point meet again (roughly). Also related to Sylvester criterion for quadratic forms [G-F, Sect. 30].

We are now ready to state the sufficient condition. (It is not really as constructive and practical as the necessary conditions, because of the need to study conjugate points. In practice, we would try to exploit the simpler necessary conditions first to see if we can narrow down the candidates for an optimum.)

Theorem 2 (sufficient condition for a weak minimum) *Consider the problem* (2.2), (2.3) *under the usual smoothness assumptions (in fact, here we take* $L \in \mathcal{C}^3$, $y \in \mathcal{C}^1$*). Then* y *is a weak minimum if it satisfies the following three conditions:*

1) It is an extremal, i.e., it satisfies Euler-Lagrange

$$\frac{d}{dx}L_{y'} = L_y$$

2) $L_{y'y'} > 0 \;\forall\, x \in [a,b]$

3) $[a,b]$ *contains no points conjugate to* a.

Proof. The only thing that remains to be shown is that $\delta^2 J\big|_y$ dominates the h.o.t. $o(\alpha^2)$ in the second-order expansion ($\delta^2 J\big|_y > 0$ by itself is not enough in general, but here we can show that all is well).

Recall our earlier derivation:

$$J(x, y+\alpha\eta, y'+\alpha\eta') = J(x,y,y') + \underbrace{\int_a^b \left(L_y - \frac{d}{dx}L_{y'}\right)\eta dx}_{=\delta J|_y(\eta)=0} \cdot\alpha$$
$$+ \underbrace{\int_a^b \left(P(x)(\eta')^2 + Q(x)\eta^2\right)dx}_{=\delta^2 J|_y(\eta)>0} \cdot\alpha^2 + \int_a^b \left(\bar{P}(x,\alpha)(\eta')^2 + \bar{Q}(x,\alpha)\eta^2\right)dx\cdot\alpha^2$$

where $\bar{P}(x,\alpha), \bar{Q}(x,\alpha) \to 0$ as $\alpha \to 0$.

Since $P = \frac{1}{2}L_{y'y'} > 0$, we can pick a small enough $\delta > 0$ such that $P(x) > \delta \;\forall\, x \in [a,b]$. Consider

$$\int_a^b \left((P-\delta)(\eta')^2 + Q\eta^2\right)dx \tag{2.13}$$

If δ is small enough, no conjugate points are introduced as we pass from P to $P-\delta$ (by continuity of solutions of the accessory equation with respect to parameter variations). So, the functional (2.13) is still positive definite, hence

$$\int_a^b \left(P(\eta')^2 + Q\eta^2\right)dx > \delta\int_a^b (\eta')^2 dx \qquad \forall\,\eta$$

Idea: We are able to handle $\int(\eta')^2$ in $o(\alpha^2)$, and we know that this is the dominant term (cf. Legendre's necessary condition), so we should be OK.

In fact, by Cauchy-Schwarz with respect to L_2 norm ($(\int fg)^2 \le \int f^2 \cdot \int g^2$):

$$\eta^2(x) = \left(\int_a^x 1\cdot\eta'(t)dt\right)^2 \le (x-a)\int_a^x (\eta'(t))^2 dt \le (x-a)\int_a^b (\eta'(t))^2 dt$$

Hence

$$\int_a^b \eta^2(x)dx \le \int_a^b (x-a)dx \int_a^b (\eta'(t))^2 dt \le \frac{(b-a)^2}{2}\int_a^b (\eta'(x))^2 dx$$

Now it is easy to see that *for α small enough*, we can get both $\bar{P}$ and $\bar{Q}\frac{(b-a)^2}{2}$ to be $< \delta/2$ and then $J(x, y+\alpha\eta, y'+\alpha\eta') > J(x, y, y')$. (Actually, the weak minimum is strict.) □

Multiple degrees of freedom: similar. P, Q, W —matrices, square completion requires

$$Q + W' = WP^{-1}W$$

which is the *matrix Riccati DE*! Can be reduced to a second-order linear ODE by the substitution $W = -PU'U^{-1}$, where U is a matrix. Relates to canonical equations, more on this later. Read: [ECE 515 Class Notes, Sect. 10.4, 11.4]; also [Brockett, FDLS].

Chapter 3

From calculus of variations to optimal control

3.1 Necessary conditions for strong extrema

Variational problems and methods discussed so far apply primarily to weak minima over $\mathcal{C}^1$ curves. We saw that we really want to study strong minima over a.e. $\mathcal{C}^1$ curves. This will be the primary goal of the next part—*optimal control*. However, we first briefly discuss two results from calculus of variations on strong minima over a.e. $\mathcal{C}^1$ curves. Reasons for doing this:

- Technical "warm-up" before hitting the maximum principle;
- Tracing the historical development of the subject (this is why we began with calculus of variations in the first place).

$\longrightarrow$ Note: if a strong minimum is $\mathcal{C}^1$ then it is also a weak minimum, so the previous necessary conditions are still useful.

From the Maximum Principle we'll be able to recover these conditions, and more.

3.1.1 Weierstrass-Erdmann corner conditions

We saw: in optimal control we want $\mathcal{C}^1$ almost everywhere, not everywhere (but of course $\mathcal{C}^0$ everywhere). Another example: catenary. (It has two corners, though.)

See also example in [G-F, pp. 61–62].

So, let us derive a necessary condition for a *strong* extremum for the problem (2.2), (2.3) (G-F say *weak* but they seem to be wrong), where $y \in \mathcal{C}^1$ except possibly at some unspecified point $c \in (a, b)$.

Begin optional:

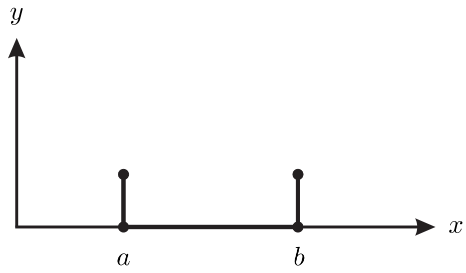

Figure 3.1: Extremal with two corners for catenary problem

Let us consider two separate perturbations η_1 and η_2 on the two pieces—will glue them together later. Clearly, $\eta_1(a) = \eta_2(b) = 0$. Also note that we should perturb the corner point c since it is not fixed.

$$J_1(y + \alpha\eta_1) = \int_a^{c+\alpha\Delta x} L(x, y + \alpha\eta_1, y' + \alpha\eta_1')dx$$

so the first variation is

$$\delta J_1\big|_y = \int_a^c \left(L_y\eta_1 + L_{y'}\eta_1'\right) dx + \Delta x \cdot L(x, y, y')\big|_{x=c^-}$$

If $\Delta x < 0$, define η_1 up to c in some way (e.g., linear continuation). Integrating by parts:

$$= \int_a^c \left(L_y - \frac{d}{dx}L_{y'}\right)\eta_1 dx + L_{y'}\ \eta_1\big|_{x=c^-} + \Delta x \cdot L(c^-)$$

(recall that $\eta_1(a) = 0$). Similarly,

$$J_2(y + \alpha\eta_2) = \int_{c+\alpha\Delta x}^b L(x, y + \alpha\eta_2, y' + \alpha\eta_2')dx$$

hence

$$\delta J_2\big|_y = \int_b^c \left(L_y - \frac{d}{dx}L_{y'}\right)\eta_2 dx - L_{y'}\eta_2(c^+) - \Delta x \cdot L(c^+)$$

$\longrightarrow$ (note: need to distinguish $c^- := \lim_{h\nearrow 0}(c + h)$ and $c^+ := \lim_{h\searrow 0}(c + h)$ since y' is not continuous at $x = c$). The total first variation is

$$\delta J = \delta J_1 + \delta J_2 = \int_a^b \left(L_y - \frac{d}{dx}L_{y'}\right)\eta dx \quad \text{(must be 0 because the individual pieces must be extremals)}$$
$$+ L_{y'}\eta_1(c^-) + \Delta x \cdot L(c^-) - L_{y'}\eta_2(c^+) - \Delta x \cdot L(c^+)$$

and this must be 0.

Now we need to make sure that the two perturbations match, in the sense that they preserve continuity of the curve.

We used the same Δx in J_1 and J_2, so that's OK. But vertical displacements must also be the same. How does this relate Δx, $\eta_1(c^-)$, and $\eta_2(c^+)$?

$\longrightarrow$ (Note: the quantities $\eta_1(c^-)$ and $\eta_2(c^+)$ are by themselves not very revealing, need to use Δx as well.)

$$y(c^- + \alpha\Delta x) + \alpha\eta_1(c^- + \alpha\Delta x) = y(c) + \delta y \cdot \alpha = y(c^+ + \alpha\Delta x) + \alpha\eta_2(c^+ + \alpha\Delta x)$$

Matching the first-order terms in α:

$$y'(c^-)\Delta x + \eta_1(c^-) = \delta y = y'(c^+)\Delta x + \eta_2(c^+)$$

So

$$\eta_1(c^-) = \delta y - y'(c^-)\Delta x, \qquad \eta_2(c^+) = \delta y - y'(c^+)\Delta x$$

Substitute these into the first variation:

$$0 = \delta J = \left(L_{y'}\big|_{x=c^-} - L_{y'}\big|_{x=c^+} \right) \delta y - \left((y'L_{y'} - L)\big|_{x=c^-} - (y'L_{y'} - L)\big|_{x=c^+} \right) \Delta x$$

Since Δx and δy are arbitrary, we conclude that

End optional

$L_{y'}$ and $y'L_{y'} - L$ *must be continuous at* c. More precisely, their discontinuities (due to the fact that y' does not exist at $x = c$) must be *removable.*

These are *Weierstrass-Erdmann corner conditions*. Recall that $y'L_{y'} - L$ is the Hamiltonian and $L_{y'}$ is the momentum, already familiar (Section 2.4). The condition says that these must be continuous.

- Euler-Lagrange equation should still hold for each piece of the extremal. Note that W-E conditions provide 4 (together with boundary conditions) relations—the correct number to specify 2 pieces of the extremal (each given by the second-order Euler-Lagrange equation);

3.1.2 Weierstrass' excess function

References: G-F, MacCluer, Sussmann.

Given the Lagrangian $L(x, y, z)$, *Weierstrass' excess function* is defined as

$$E(x, y, z, w) := L(x, y, w) - L(x, y, z) - (w - z) \cdot L_z(x, y, z)$$

(Weierstrass, ∼1880).

Note that $L(x, y, z) + (w - z)L_z(x, y, z)$ is the first-order Taylor approximation of $L(x, y, w)$, viewed as a function of w, around $w = z$. This gives a geometric interpretation of the E-function:

Theorem 3 (Weierstrass' necessary condition for a strong minimum) *Consider the problem* (2.2), (2.3)*. If* $y(\cdot)$ *is a strong minimum, then* $E(x, y, y', w) \geq 0$ *along* $y(x)$ *for every* w*.*

Geometrically: the graph lies above the tangent at y'.

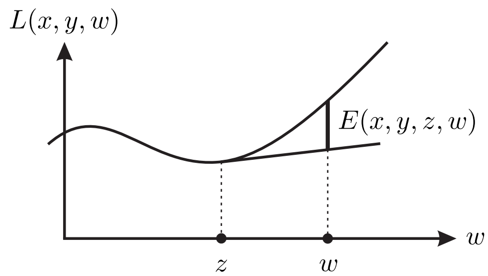

Figure 3.2: Weierstrass' excess function

Proof (sketch) [MacCluer, p. 214]; [G-F, p. 149, 151].

Suppose $E(x_0, y(x_0), y'(x_0), w) < 0$ for some x_0 and some w. Consider another curve $\bar{y}(x)$ which coincides with $y(x)$ everywhere except on a small interval $[x_0, x_0 + \Delta x]$ where $\bar{y}'(x_0) = w$ (except near $x_0 + \Delta x$, to ensure continuity).

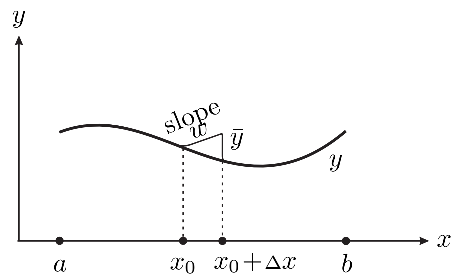

Figure 3.3: The graphs of y and $\bar{y}$

On $[x_0, x_0 + \Delta x]$ we have $\bar{y}(x) \approx y(x_0) + w(x - x_0)$, $y(x) \approx y(x_0) + y'(x_0)(x - x_0)$, hence $|\bar{y} - y| \leq |w - y'(x_0)|\Delta x$. Here and below, $\approx$ means equality up to $o(\Delta x)$. We have

$$\begin{aligned}
0 \leq J(\bar{y}) - J(y) &= \int_{x_0}^{x_0+\Delta x} (L(x, \bar{y}, \bar{y}') - L(x, y, y'))dx \\
&= \int_{x_0}^{x_0+\Delta x} (L(x, \bar{y}, \bar{y}') - L(x, y, \bar{y}') + L(x, y, \bar{y}') - L(x, y, y'))dx \\
&\approx \int_{x_0}^{x_0+\Delta x} ((\bar{y} - y)L_y(x, y, \bar{y}') + L(x, y, \bar{y}') - L(x, y, y'))dx \\
&= \int_{x_0}^{x_0+\Delta x} (\bar{y} - y)L_y(x, y, \bar{y}')dx + \int_{x_0}^{x_0+\Delta x} E(x, y, y', \bar{y}')dx + \int_{x_0}^{x_0+\Delta x} (\bar{y}' - y')L_z(x, y, y')dx
\end{aligned}$$

Note that the first term is of order $o(\Delta x)$ (since the integrand is $O(\Delta x)$ as we saw earlier, and the length of the interval of integration is also $O(\Delta x)$). Integrating the last term by parts, we obtain

$$\underbrace{(\bar{y} - y)L_z|_{x_0}^{x_0+\Delta x}}_{=0} - \int_{x_0}^{x_0+\Delta x} (\bar{y} - y)\frac{d}{dx}L_z(x, y, y')dx$$

and this integral is again of order $o(\Delta x)$. This implies that the dominant first-order term

$$\int_{x_0}^{x_0+\Delta x} E(x, y, y', \bar{y}')dx$$

must be ≥ 0. But by construction of $\bar{y}$ and since $E(x_0, y(x_0), y'(x_0), w) < 0$, the integral is negative for small Δx, a contradiction. □

$\longrightarrow$ Combining nonnegativity of the E-function with sufficient conditions for a weak extremum (Legendre's condition $L_{y'y'} > 0$ plus the absence of conjugate points), one can obtain a *sufficient* condition for a strong extremum. Precise formulation requires the concept of a field [G-F, Chap. 6].

Reference info:

The above proof of Weierstrass' condition seems to be due to McShane (1939). This is why needle variations in the proof of the Maximum Principle are also called *Pontryagin-McShane variations*. McShane's paper is based on a monograph by Bliss (1930). Bliss also has nice historical remarks at the end. Both papers appeared in American Journal of Math and are available on-line.

Digression: note that we can write

$$\begin{aligned} E(x, y, y', w) &= \big(L(x, y, w) - wL_z(x, y, y')\big) - \big(L(x, y, y') - y'L_z(x, y, y')\big) \\ &= H(x, y, y', p) - H(x, y, w, p) \geq 0 \end{aligned}$$

by Weierstrass' condition (be mindful of the sign in the above formula), where $p = L_z(x, y, y')$ —same momentum in both, but the third argument is y' in one and w in the other. (Recall: $H(x, y, y', p) := py' - L(x, y, y')$, $p = L_z$, treating y' and p as independent arguments.)

So Weierstrass' condition means: if $y(x)$ is an optimal trajectory and $p(x)$ is the momentum computed along the optimal trajectory, then H as a function of y' is *maximized* along the optimal trajectory.

Weierstrass didn't know this! Neither did Carathéodory [Sussmann].

This is again the maximum principle, and Weierstrass' condition can be deduced from it. *But:* the Maximum Principle is more general. It applies to constrained $\mathcal{U}$ spaces, but Weierstrass' condition may not hold at boundary points.

The perturbation used in the above proof is already quite close to perturbations we'll use in proving the Maximum Principle .

Exercise 9 (due Feb 28) Deduce (rigorously!) the two Weierstrass-Erdmann corner conditions given earlier from the integral form of Euler-Lagrange :

$$L_{y'} = \int_a^x L_y dx + C$$

and the Weierstrass necessary condition ($E \geq 0$).

3.2 Calculus of variations vs. optimal control

In calculus of variations , we minimize the cost $\int L(x, y, y')dx$ over a family of curves $y(x)$.

Optimal control takes a more *dynamic* view: $y' = u$—view this as a control system.

At each point along the curve, we choose the direction and speed (u) such that the cost is minimized (the control is infinitesimal while the cost is aggregate).

We know that not all directions may be available (constraints). In calculus of variations , we model this as $M(x, y, y') = 0$. We already discussed that this is equivalent to a control system:

$$y' = f(x, y, u)$$

Here $u \in \mathbb{R}^m$ gives available controls, and they affect the system and induce available direction of motion (again, we think dynamically as moving along the curve and incurring cost along the way).

In applications, $u \in \mathbb{R}^m$ is not always reasonable, might have $u \in \mathcal{U}$—some (closed) subset of $\mathbb{R}^m$ (think force, velocity, etc.) In optimal control, as we will see, such constraints are incorporated very naturally, but calculus of variations has trouble. (Though there exists some work on this; see reference [D-8] in [A-F], which is 1962—after the Maximum Principle ?)

Also may want to consider costs like $\int |u| dx$, where L is not differentiable with respect to u.

Example 5 Brachistochrone [Sussmann-Willems]

Our old formulation was: among curves $y(x)$ connecting two given points, minimize

$$\text{Time} = \int \frac{\text{Arclength}}{\text{Velocity}} = \int_a^b \frac{\sqrt{1+(y')^2}}{v} dx$$

where v was found from conservation of energy:

Take $A = 0$ so that $E = 0$.

$$\frac{mv^2}{2} - mgy = 0$$

Normalizing $m = 1$, $2g = 1$: $v^2 = y$, or $v = \sqrt{y}$. So we get

$$J = \int_a^b \frac{\sqrt{1+(y')^2}}{\sqrt{y}} dx$$

But: this is just time in disguise, so let us switch notation to make time explicit. Let us reparameterize everything in terms of t: $x(t), y(t)$. Constraint (from conservation of energy):

$$v^2 = y \;\Leftrightarrow\; \dot{x}^2 + \dot{y}^2 = y$$

Let us rewrite the constraint as a control system with constraint on u:

$$\dot{x} = u\sqrt{y}$$
$$\dot{y} = v\sqrt{y}$$

where

$$\left\| \begin{pmatrix} u \\ v \end{pmatrix} \right\| = 1$$

or $u^2 + v^2 = 1$ (unit circle).

$J =$ time when we hit the target: $x = b$, $y = B$. If we want to write this in terms of the Lagrangian, and if the curve is parameterized by $t \in [t_0, t_1]$:

$$J = t_1 - t_0 = \int_{t_0}^{t_1} 1 \cdot dt$$

so $L \equiv 1$, very simple. Here $(x(t_0), y(t_0)) = (a, A)$ is the initial point and $(x(t_1), y(t_1)) = (b, B)$ is the terminal point.

The obtained formulation is equivalent to the original one.

$\longrightarrow$ Note that the emphasis (in terms of complexity) has shifted from the cost to the right-hand side of the control system. □

Read [Sussmann-Willems], where the solution of the brachistochrone problem is rederived using the new control formulation and the maximum principle.

- We emphasize that in the above formulation, the control set is constrained (unit circle).
- In what follows, we will start using t as the independent variable. Dependent variables will be $x = (x_1, \dots, x_n)$, and controls will be $u = (u_1, \dots, u_m)$.
- The fact that $L = 1$ is because this was a special case—*time-optimal control*. In general, L can be any function.

3.3 The optimal control problem formulation and assumptions

From now on, primary reference: [A-F] (however, our sign convention—minimum vs. maximum principle—will be different, to make things more consistent with the calculus of variations part). Also [Pontryagin et al.]. More technical details in [Sussmann].

3.3.1 System

$$\dot{x} = f(t, x, u), \qquad x(t_0) = x_0 \tag{3.1}$$

where $x \in \mathbb{R}^n$, $u \in \mathcal{U} \subset \mathbb{R}^m$ (or, more generally, $\mathcal{U}(t)$) , $t \in [t_0, \infty)$, or $(-\infty, \infty)$ (to flow back), can just write $t \in \mathbb{R}$.

Need to know: $\forall\, t_0, x_0$ there exists a unique solution on some time interval $[t_0, t_1)$.

What conditions does this impose on f and u?

Let us first discuss the case of no controls:

$$\dot{x} = f(t, x)$$

Conditions for existence and uniqueness – ?

f needs to be "sufficiently regular"—what does that mean?

[A-F, Sect. 3-18]; [Khalil] (no u); [Sontag].

With respect to t: $f(\cdot, x)$ is *piecewise continuous*, for each fixed x. (Actually, *measurable* is enough.)

Piecewise continuous means:

- Continuous except possibly a finite number of discontinuities on each finite interval (countable number overall);
- There exist left and right limits everywhere;
- By convention, the value at each discontinuity equals either left or right limit (continuous from left or from right).

With respect to x: $f(t, \cdot)$ must be locally Lipschitz, uniformly in t:

$$\|f(t, x_1) - f(t, x_2)\| \le L\|x_1 - x_2\|$$

with respect to standard Euclidean norm on $\mathbb{R}^n$. The Lipschitz constant L is valid in some neighborhood of (t_0, x_0) in $\mathbb{R} \times \mathbb{R}^n$.

Let's actually be more generous. By MVT, the above condition is guaranteed if:

- $f(t,\cdot)$ is $\mathcal{C}^1$ for each fixed t;
- f_x is (piecewise) continuous in t (actually, all we need is $|f_x| \le \alpha(t)$ for α locally integrable [Sontag, p. 480]).

Then on some interval $[t_0, t_1)$ there exists a unique solution $x(\cdot)$. Note that $f(\cdot, x)$ is not necessarily $\mathcal{C}^0$ everywhere, so we need to be careful: $x(\cdot) \in \mathcal{C}^0$ everywhere and $\mathcal{C}^1$ almost everywhere, and satisfies

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds$$

Such $x(\cdot)$ are called *absolutely continuous.*

Now back to the control system (3.1). To get the above condition ($\mathcal{C}^1$) for

$$\tilde{f}(t,x) := f(t,x,u(t))$$

we can assume (overkill, but reasonable):

- $u(\cdot)$ is piecewise continuous (this allows quite general dependance on t). Or just measurable and locally (essentially) bounded. This will come up (Fuller's problem)!
- f is $\mathcal{C}^1$ in x;
- f is $\mathcal{C}^0$ or even $\mathcal{C}^1$ in t ($\mathcal{C}^1$—to be able to absorb t into x via $\dot{t} = 1$). Can relax to piecewise $\mathcal{C}^0$, but this is easier.
- f_x is (piecewise) $\mathcal{C}^0$ in t;
- f, f_x are $\mathcal{C}^0$ with respect to u.

In the above, the conditions imposed on f_x (namely, its existence and continuity properties with respect to t and u) can be relaxed to Lipschitz property. (See below for details.)

$\longrightarrow$ Note: existence of f_u is *not* assumed.

Can check that the above works—do it yourselves. (Need to know that (continuous)$\circ$(piecewise continuous) is piecewise continuous.)

Begin optional: ______

As far as Lipschitz relaxation, need

$$\|f(t,x_1,u) - f(t,x_2,u)\| \le L\|x_1 - x_2\|$$

for bounded ranges of t, x, u. This is ensured if f is locally Lipschitz in $\binom{x}{u}$, locally uniformly over t. The last condition is especially convenient when f doesn't depend on t.

End optional ______

$\longrightarrow$ Again, by a solution we mean an absolutely continuous $x(\cdot)$ satisfying

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s), u(s))ds$$

$\longrightarrow$ In the future, the control system is always assumed to satisfy local existence and uniqueness. Actually, when we consider $[t_0, t_1]$, we assume that at least the candidate optimal trajectory (and hence nearby ones) exists over $[t_0, t_1]$, so that the cost is finite for it.