Dynamic Programming & Optimal Control

Adi Ben-Israel

Adi Ben-Israel, RUTCOR–Rutgers Center for Operations Research, Rutgers University, 640 Bartholomew Rd., Piscataway, NJ 08854-8003, USA

E-mail address: `benisrael@rbsmail.rutgers.edu`

LECTURE 1

Dynamic Programming

1.1. Recursion

A *recursion* is a rule for computing a value using previously computed values, for example, the rule

$$f_{k+1} = f_k + f_{k-1} \ , \tag{1.1}$$

computes the *Fibonnaci sequence*, given the initial values $f_0 = f_1 = 1$.

EXAMPLE 1.1. A projectile with mass 1 shoots up against earth gravity g. The initial velocity of the projectile is v_0. What maximal altitude will it reach?
Solution. Let $y(v)$ be the maximal altitude reachable with initial velocity v. After time Δt, the projectile advanced approximately $v\,\Delta t$, and its velocity has decreased to approximately $v - g\,\Delta t$. Therefore the recursion

$$y(v) \approx v\,\delta t + y(v - g\,\Delta t) \ , \tag{1.2}$$

gives

$$\begin{aligned} \frac{y(v) - y(v - g\,\Delta t)}{\Delta t} &\approx v \ , \\ \therefore \ y'(v) &= \frac{v}{g} \\ \therefore \ y(v) &= \frac{v^2}{2\,g} \ , \end{aligned} \tag{1.3}$$

and the maximal altitude reached by the projectile is $v_0^2/2\,g$.

EXAMPLE 1.2. Consider the partitioned matrix

$$A = \begin{pmatrix} B & c \\ r & \alpha \end{pmatrix} \tag{1.4}$$

where B is nonsingular, c is a column, r is a row, and α is a scalar. Then A is nonsingular iff

$$\alpha - rB^{-1}c \neq 0 \ , \tag{1.5}$$

(verify!) in which case

$$A^{-1} = \begin{pmatrix} B^{-1} + \beta B^{-1} c r B^{-1} & -\beta B^{-1} c \\ -\beta r B^{-1} & \beta \end{pmatrix} \ , \tag{1.6a}$$

$$\text{where } \beta = \frac{1}{\alpha - rB^{-1}c} \ . \tag{1.6b}$$

Can this result be used in a recursive computation of A^{-1}? Try for example

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 1 & 3 & 4 \end{pmatrix} \text{ , with inverse } A^{-1} = \begin{pmatrix} 4 & -1 & -2 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix} .$$

EXAMPLE 1.3. Consider the LP

$$\begin{aligned} \max \mathbf{c}^T\mathbf{x} & \\ \text{s.t. } A\,\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

and denote its optimal value by $V(A, \mathbf{b})$. Let the matrix A be partitioned as $A = (A_1, \mathbf{a}_n)$, where $\mathbf{a}_n$ is the last column, and similarly partition the vector $\mathbf{c}$ as $\mathbf{c}^T = (\mathbf{c}_1^T, c_n)$. Is the recursion

$$V(A, \mathbf{b}) = \max_{x_n \geq 0} \{c_n x_n + V(A_1, \mathbf{b} - \mathbf{a}_n\, x_n)\} \tag{1.7}$$

valid? Is it useful for solving the problem?

1.2. Multi stage optimization problems and the optimality principle

Consider a process consisting of N sequential *stages*, and requiring some decision at each stage. The information needed at the beginning of stage i is called its *state*, and denoted by x_i. The *initial state* x_1 is assumed given. A *decision* u_i is selected in stage i from the *feasible decision set* $U_i(x_i)$, that in general depends on the state x_i. This results in a *reward* $r_i(x_i, u_i)$ and the next state becomes

$$x_{i+1} = T_i(x_i, u_i) \; , \tag{1.8}$$

where $T_i(\cdot, \cdot)$ is called the *state transformation* (or *dynamics*) in stage i. The *terminal state* x_{N+1} results in a reward $S(x_{N+1})$, called the *salvage value* of x_{N+1}. If the rewards accumulate, the problem becomes

$$\max \left\{ \sum_{1=1}^{N} r_i(x_i, u_i) + S(x_{N+1}) : \begin{array}{l} x_{i+1} = T_i(x_i, u_i) \; , \; i = \overline{1, N} \; , \\ u_i \in U_i(x_i) \; , \; i = \overline{1, N} \; , \\ x_1 \text{given} \; . \end{array} \right\} . \tag{1.9}$$

This problem can be solved, in principle, as an optimization problem in the variables $u_1, \ldots, u_N$. However, this ignores the special, sequential, structure of the problem.

The *Dynamic Programming* (DP) solution is based on the following concept.

DEFINITION 1.1. The *optimal value function* (OV function) at stage k, denoted $V_k(\cdot)$, is

$$V_k(x) := \max \left\{ \sum_{1=k}^{N} r_i(x_i, u_i) + S(x_{N+1}) : \begin{array}{l} x_{i+1} = T_i(x_i, u_i) \; , \; i \in \overline{k, N} \; , \\ u_i \in U_i(x_i) \; , \; i \in \overline{k, N} \; , \\ x_k = x \; . \end{array} \right\} . \tag{1.10}$$

The *optimal value* of the original problem is $V_1(x_1)$. An *optimal policy* is a sequence of decisions $\{u_1, \ldots, u_N\}$ resulting in the value $V_1(x_1)$.

Bellman's *Principle of Optimality* (abbreviated PO) is often stated as follows:

> An optimal policy has the property that whatever the initial state and the initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision, [**5**, p. 15].

The PO can be used to recursively compute the OV functions

$$
\begin{aligned}
V_k(x) &= \max_{u \in U_k(x)} \{r_k(x,u) + V_{k+1}(T_k(x,u),u)\}\ ,\ k \in \overline{1,N}\ , && (1.11)\\
V_{N+1}(x) &= S(x)\ . && (1.12)
\end{aligned}
$$

The last equation is called the *boundary condition* (BC).

Example 1.4. It is required to partition a positive number x in N parts, $x = \sum_{i=1}^{N} u_i$, such that the sum of squares $\sum_{i=1}^{N} u_i^2$ is minimized.
For each $k \in \overline{1,N}$ define the OV function

$$V_k(x) := \min \left\{ \sum_{i=1}^{k} u_i^2 \ :\ \sum_{i=1}^{k} u_i = x \right\}\ .$$

The recursion (1.11) here becomes

$$V_k(x) := \min_{u \in [0,x]} \left\{ u^2 + V_{k-1}(x-u) \right\}\ ,$$

with $V_1(x) = x^2$ as BC.

For $k = 2$ we get

$$V_2(x) := \min_{u \in [0,x]} \left\{ u^2 + (x-u)^2 \right\} = \frac{x^2}{2}\ ,\ \text{with optimal } u = \frac{x}{2}\ .$$

Claim: For genral N, the optimal value is $V_N(x) = \dfrac{x^2}{N}$, with optimal policy of equal $u_i = \dfrac{x}{N}$.

Proof. (by induction)

$$
\begin{aligned}
V_N(x) &= \min_{u \in [0,x]} \left\{ u^2 + V_{N-1}(x-u) \right\}\\
&= \min_{u \in [0,x]} \left\{ u^2 + \frac{(x-u)^2}{N-1} \right\}\ ,\ \text{by the induction hypothesis}\ ,\\
&= \frac{x^2}{N}\ ,\ \text{for } u = \frac{x}{N}\ .
\end{aligned}
$$

□

The following example shows that the PO, as stated above, is incorrect. For details see the book [**29**] by M. Sniedovich where the PO is given a rigorous treatment.

Example 1.5. *([**29**, pp. 294–295])* Consider the graph in Figure 1.1. It is required to find the shortest path from node 1 to node 7, the length of a path is defined as the length of its longest arc. For example,

$$\text{length}\{1,2,5,7\} = 4 \ .$$

If the nodes are viewed as states, then the path $\{1,2,4,6,7\}$ is optimal w.r.t. $s = 1$. However, the path $\{2,4,6,7\}$ is not optimal w.r.t. the node $s = 2$, as its length is greater than the length of $\{2,5,7\}$.

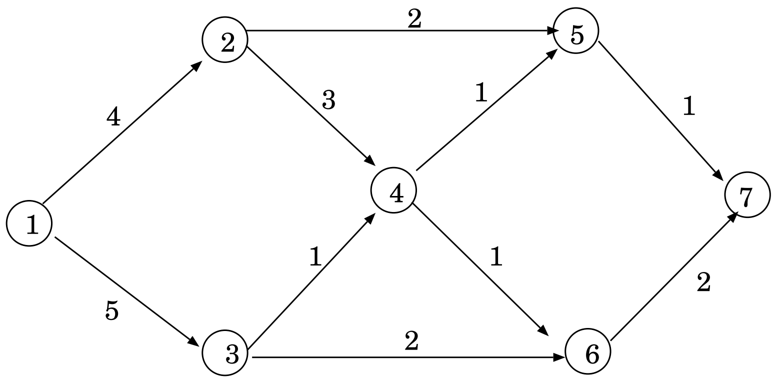

Figure 1.1. An illustration why the PO should be used carefully

1.3. Inverse Dynamic Programming

Consider a multi–stage decision process of § 1.2, where the state x is the project budget. A reasonable question is to determine the *minimal budget* that will enable achieving a given *target* v, i.e.

$$\min\{x : V_1(x) \geq v\} \ ,$$

denoted $I_1(v)$ and called the *optimal input* for v. Similarly, at any stage we define the *optimal input* as

$$I_k(v) := \begin{cases} \min\{x : V_k(x) \geq v\} \ , \\ \infty \text{ if the target } v \text{ is unattainable} \ , \end{cases} \quad k = 1, 2, \cdots, N \ . \tag{1.13}$$

The terminal optimal input is defined as

$$I_{N+1}(v) := S^{-1}(v) \tag{1.14}$$

assuming the salvage value function $S(\cdot)$ is monotonic.

A natural recursion for the optimal inputs is:

$$I_k(v) := \min\{x : \exists u \in U_k(x) \ni T_k(x,u) \geq I_{k+1}(v - r_k(x,u))\} \ , \ k = N, N-1, \cdots, 1 \ , \tag{1.15}$$

with (1.14) as BC.

Exercises.

EXERCISE 1.1. Use DP to maximize the entropy

$$\max \left\{ -\sum_{i=1}^{N} p_i \log p_i : \sum_{i=1}^{N} p_i = 1 \right\} .$$

EXERCISE 1.2. Let $\mathbb{Z}_+$ denote the nonnegative integers. Use DP to write a recursion for the knapsack problem

$$\max \left\{ \sum_{i=1}^{N} f_i(x_i) : \sum_{i=1}^{N} w_i(x_i) \leq W, \ , x_i \in \mathbb{Z}_+ \right\} ,$$

where $c_i(\cdot), w_i(\cdot)$ are given functions: $\mathbb{Z}_+ \to \mathbb{R}$ and $W > 0$ is given. State any additional properties of $f_i(\cdot)$ and $w_i(\cdot)$ that are needed in your analysis.

EXERCISE 1.3. ([**7**]) A cash amount of x cents can be represented by

$$\begin{aligned} &x_1 \text{ coins of 50 cents} , \\ &x_2 \text{ coins of 25 cents} , \\ &x_3 \text{ coins of 10 cents} , \\ &x_4 \text{ coins of 5 cents} , \text{ and} \\ &x_5 \text{ coins of 1 cent} . \end{aligned}$$

The representation is:

$$x = 50x_1 + 25x_2 + 10x_3 + 5x_4 + x_5 .$$

(a) Use DP to find the representation with the minimal number of coins.
(b) Show that your solution agrees with the "greedy" solution:

$$x_1 = \left\lfloor \frac{x}{50} \right\rfloor , \ x_2 = \left\lfloor \frac{x - 50x_1}{25} \right\rfloor , \text{ etc.}$$

where $\lfloor \alpha \rfloor$ is the greatest integer $\leq \alpha$.
(c) Suppose a new coin of 20 cents is introduced. Will the DP solution still agree with the greedy solution?

EXERCISE 1.4. The *energy* required to compress a gas from *pressure* p_1 to pressure p_{N+1} in N stages is proportional to

$$\left(\frac{p_2}{p_1}\right)^\alpha + \left(\frac{p_3}{p_2}\right)^\alpha + \cdots + \left(\frac{p_{N+1}}{p_N}\right)^\alpha$$

with α a positive constant. Show how to choose the intermediate pressures $p_2 , \cdots , p_N$ so as to minimize the energy requirement.

EXERCISE 1.5. Consider the following variation of the game NIM, defined in terms of N piles of matches containing $x_1, x_2, \cdots , x_N$ matches. The rules are:
(i) two players make moves in alternating turns,
(ii) if matches remain, a move consists of removing any number of matches all from the same pile,

(iii) the last player to move loses.
For example, consider a game with initial piles $\{x_1, x_2, x_3\} = \{1, 4, 7\}$ where moves by players I, II are denoted by $\stackrel{\text{I}}{\to}$ and $\stackrel{\text{II}}{\to}$ resp.,

$$\begin{aligned} &\{1,4,7\} \stackrel{\text{I}}{\to} \{1,4,5\} \stackrel{\text{II}}{\to} \{0,4,5\} \stackrel{\text{I}}{\to} \{0,4,4\} \stackrel{\text{II}}{\to} \{0,3,4\} \stackrel{\text{I}}{\to} \{0,3,3\} \\ &\stackrel{\text{II}}{\to} \{0,2,3\} \stackrel{\text{I}}{\to} \{0,2,2\} \stackrel{\text{II}}{\to} \{0,2,1\} \stackrel{\text{I}}{\to} \{0,0,1\} \stackrel{\text{II}}{\to} \{0,0,0\}\ , \text{ and II loses .} \end{aligned}$$

(a) Use DP to find an optimal move for an initial state $\{x_1, x_2, \cdots, x_N\}$.
(b) Find a simple rule to determine if an initial state is a winning position.
(c) Is $\{x_1, x_2, x_3, x_4\} = \{1, 3, 5, 7\}$ a winning position?
Hint: Let $V(x_1, x_2, \cdots, x_N)$ be the optimal value of having the piles $\{x_1, x_2, \cdots, x_N\}$ when it is your turn to play, with

$$V(x_1, x_2, \cdots, x_N) = \begin{cases} 0, & \text{if } \{x_1, x_2, \cdots, x_N\} \text{ is a losing position} \\ 1, & \text{otherwise.} \end{cases}$$

Let y_i be the number of matches removed from pile i. Then

$$V(x_1, x_2, \cdots, x_N) = 1 - \max_{i=1,\ldots,N} \left\{ \max_{1 \le y_i \le x_i} V(x_1, x_2, \cdots, x_i - u_i, \cdots, x_N) \right\}$$

where

$$V(x_1, x_2, \cdots, x_N) = 0 \text{ if all } x_i = 0 \text{ except, say } x_j = 1$$

Exercise 1.6. We have a number of coins, all of the same weight except for one which is of different weight, and a balance.
(a) Determine the weighing procedures which minimize the maximum time required to locate the distinctive coin in the following cases:

- the coin is known to be heavier,
- it is not known whether the coin is heavier or lighter.

(b) Determine the weighing procedures which minimize the expected time required to locate the coin.
(c) Consider the more general problem where there are two or more distinctive coins, under various assumptions concerning the distinctive coins.

Exercise 1.7. A rocket consists of k stages carrying fuel and a nose cone carrying the pay load. After the fuel carried in stage k is consumed, this stage drops off, leaving a $k-1$ stage rocket. Let

$$\begin{aligned} W_0 &= \text{weight of nose cone ,} \\ w_k &= \text{initial gross weight of stage } k \text{ ,} \\ W_k &= W_{k-1} + w_k \text{ , initial gross weight of sub-rocket } k \text{ ,} \\ p_k &= \text{initial propellant weight of stage } k \text{ ,} \\ v_k &= \text{change in rocket velocity during burning of stage } k \text{ .} \end{aligned}$$

Assume that the change in velocity v_k is a known function of W_k and p_k, so that .

$$v_k = v(W_k, p_k)$$

from which

$$p_k = p(W_k, v_k)$$

Since $W_k = W_{k-1} + w_k$, and the weight of the kth stage is a known function, $g(p_k)$, of the propellant carried in the stage, we have

$$w_k = w(p(W_{k-1} + w_k, v_k))$$

whence, solving for w_k, we have

$$w_k = w(W_{k-1}, v_k)$$

(a) Use DP to design a k-stage rocket of minimum weight which will attain a final velocity v.
(b) Describe an algorithm for finding the optimal number of stages k^*.
(c) Discuss the factors resulting in an increase of k^* i.e. in more stages of smaller size.

EXERCISE 1.8. Suppose that we are given the information that a ball is in one of N boxes, and the a priori probability, p_k, that it is in the kth box.
(a) Show that the procedure which minimizes the expected time required to find the ball consists of looking in the most likely box first.
(b) Consider the more general problem where the time consumed in examining the kth box is t_k, and where there is a probability q_k that the examination of the kth box will yield no information about its contents. When this happens, we continue the search with the information already available. Let $F(p_1, p_2, \cdots, p_N)$ be the expected time required to find the ball under an optimal policy. Find the functional equation that F satisfies.
(c) Prove that if we wish to "obtain" the ball, the optimal policy consists of examining first the box for which

$$\frac{p_k(1 - q_k)}{t_k}$$

is a maximum. On the other hand, if we merely wish to "locate" the box containing the ball in the minimum expected time, the box for which this quantity is maximum is examined first, or not at all.

EXERCISE 1.9. A company has m *jobs* that are numbered 1 through m. Job i requires k_i *employees*. The "natural" *monthly wage* of job i is w_i dollars, with $w_i \leq w_{i+1}$ for all i. The jobs are to be grouped into n *labor grades*, each grade consisting of several consecutive jobs. All employees in a given labor grade receive the highest of the natural wages of the jobs in that grade. A fraction r_j of the employees in each jobs quit in each month. Vacancies must be filled by promoting from the next lower grade. For instance, a vacancy in the highest of n labor grades causes $n-1$ promotions and a hire into the lowest labor grade. It costs t dollars to train an employee to do any job. Write a functional equation whose solution determines the number n of labor grades and the set of jobs in each labor grade that minimizes the sum of the payroll and training costs.

EXERCISE 1.10. *(The Jeep problem)* Quoting from D. Gale, [**18**, p. 493],

> $\cdots$ the problem concerns a jeep which is able to carry enough fuel to travel a distance d, but is required to cross a desert whose distance is greater then d (for example $2d$). It is to do this by carrying fuel from its home base

> and establishing fuel depots at various points along its route so that it can refuel as it moves further out. It is then required to cross the desert on the minimum possible amount of fuel.

Without loss of generality, we can assume $d = 1$.

A desert of length 4/3 can be crossed using 2 units of fuel, as follows:

Trip 1: The jeep travels a distance of 1/3, deposits 1/3 unit of fuel and returns to base.

Trip 2: The jeep travels a distance of 1/3, refuels, and travels 1 unit of distance.

The inverse problem is to determine the maximal desert that can be crossed, given the quantity of fuel is x. The answer is:

$$1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2\lceil x \rceil - 3} + \frac{x - \lceil x \rceil + 1}{2\lceil x \rceil - 1}.$$

References: [**1**], [**9**], [**11**], [**15**], [**17**], [**18**], [**21**], [**25**].

LECTURE 2

Dynamic Programming: Applications

2.1. Inventory Control

An *inventory* is a dynamic system described by three variables: *state* x_t, *decision* u and *demand* w, an exogeneous variable that may be deterministic or random (the interesting case).

In the simplest case the state transformation is

$$x_{t+1} := x_t + u_t - w_t \ , \ t = 0, 1, \cdots \tag{2.1}$$

where x_t is the stock level at the beginning of day t, u_t is the amount ordered from a supplier that arrives at the beginning of day t, and w_t is the demand during day t.

The *initial stock* x_0 is given. Note that (2.1) allows for negative stocks.

The economic functions are:

- a *holding cost* $A(x)$ for holding x units overnight (*shortage cost* if $x < 0$),
- an *order cost* $B(u)$ of placing an order u, and
- *revenue* the $R(w)$ from selling w units.

Let $V_t(x)$ be the *minimal cost* of operations if day t begins with x units in stock. Then

$$V_t(x) := \min_u \{A(x) + B(u) - \mathbf{E}R(w) + \mathbf{E}V_{t+1}(x + u - w)\} \tag{2.2}$$

where $\mathbf{E}$ is expectation w.r.t. w. We assume the boundary condition

$$V_{N+1}(x) = S(x) \tag{2.3}$$

where $S(\cdot)$ is the cost of terminal stock.

Typical constraints on u are: $u \geq 0$ and $u \leq C - x$ where C is a *storage capacity.*

The order cost is typically of the form

$$B(u) = \begin{cases} 0 & \text{if } u = 0 \\ K + cu & \text{if } u > 0 \end{cases} \tag{2.4}$$

where K is a *fixed cost.*

It is convenient to use the normalized cost functions

$$V^*(x) := V(x) + cx \tag{2.5}$$

in which case (2.2) becomes

$$V_t^*(x) := \min_u \{A(x) + KH(u) - (\mathbf{E}R(w) - c\mathbf{E}w) + \mathbf{E}V_{t+1}^*(x + u - w)\} \tag{2.6}$$

where $H(\cdot)$ is the Heaviside function, $H(u) = 1$ if $u > 0$ and 0 otherwise.

THEOREM 2.1. *Let $A(x)$ and $S(x)$ be convex with limit $+\infty$ at $x = \pm\infty$, and let $K = 0$. Then the optimal policy is*

$$u_t = \begin{cases} 0 & \text{if } x_t > S_t \ , \\ S_t - x_t & \text{otherwise} \ , \end{cases} \tag{2.7}$$

for some $S_t > 0$ (the order to level on day t).

PROOF. Let $\mathcal{C}$ be the class of convex functions with limit $+\infty$ as the argument approaches $\pm\infty$. The function minimized in (2.6) is

$$\theta(x+u) = \mathbf{E} V^*_{t+1}(x + u - w) \tag{2.8}$$

If $\theta(x) \in \mathcal{C}$ and has minimum at S_t then the rule (2.7) is optimal. The minimized function is

$$\min_u \theta(x+u) = \begin{cases} \theta(S_t) & \text{if } x \le S_t \ , \\ \theta(x) & \text{if } x > S_t \ , \end{cases} \tag{2.9}$$

that is constant for $x \le S_t$.

Writing (2.6) as $V_t := \mathbf{L} V_{t+1}$, it follows that $V_t \in \mathcal{C}$ if $V_{t+1} \in \mathcal{C}$. Since $V_{N+1} \in \mathcal{C}$ by (2.3), it follows from $V_t = \mathbf{L}^{N-t} V_{N+1}$ is in $\mathcal{C}$ □

The determination of the optimal S_t is from

$$V^*_t(x) = A(x) + \mathbf{E} V^*_{t+1}(S - w) = A(x) + \theta(S) \ , \ x \le S \ ,$$

and the optimal S minimizes $\mathbf{E} A(S - w)$.

THEOREM 2.2. *If $A(x)$ and $S(x)$ are convex with the limit $+\infty$ at $x = \pm\infty$ and if $K \neq 0$ then the optimal order policy is*

$$u = \begin{cases} 0 & \text{if } x > s \ , \\ S - x & \text{otherwise} \ , \end{cases} \tag{2.10}$$

where $S > s$ if $K > 0$.

PROOF. The relevant optimality condition is (2.6). The expression to be minimized is

$$K H(u) + \theta(x+u) = \begin{cases} \theta(x) & \text{if } u = 0 \ , \\ K + \theta(x+u) & \text{if } u > 0 \ , \end{cases} \tag{2.11}$$

and θ is (2.8). If $\theta \in \mathcal{C}$ with minimum at S then the rule (2.10) is optimal for s the smaller root of

$$\theta(s) = K + \theta(S) \ .$$

In general the functions V^* and θ are not convex. The proof resumes after Lemma 2.4 below. □

DEFINITION 2.1. A scalar function $\phi(x)$ is *K-convex* if

$$\phi(x+u) \ge \phi(x) + u\phi'(x) - K \ , \ \forall u \ , \tag{2.12}$$

where ϕ' is the derivative from the right,

$$\phi'(x) = \lim_{u \downarrow 0} \frac{\phi(x+u) - \phi(x)}{u} \ .$$

The class of K-convex functions with limit $+\infty$ at $x = \pm\infty$ is denoted by $\mathcal{C}_K$. $\square$

Immediate consequences of the definition:

(a) K-convexity implies L-convexity for $L \geq K$
(b) 0-convexity is ordinary convexity
(c) if $\phi(x)$ is K-convex, so is $\phi(x-w)$
(d) if ϕ_1 and ϕ_2 are K_1 and K_2 convex respectively then $\alpha_1\phi_1 + \alpha_2\phi_2$ is $\alpha_1 K_1 + \alpha_2 K_2$ convex for $\alpha_1, \alpha_2 > 0$
(e) if ϕ is K-convex so is $\mathbf{E}\phi(x-w)$

The corresponding statements for $\mathcal{C}_K$ are:

(a) $\mathcal{C}_K$ increases with K
(b) $\mathcal{C}_0 = \mathcal{C}$
(c) $\mathcal{C}_K$ is closed under translations
(d) $\phi_i \in \mathcal{C}_{K_i}$ and $\alpha_i \geq 0$ imply $\sum \alpha_i\phi_i \in \mathcal{C}_K$ where $K = \sum \alpha_i K_i$
(e) $\mathcal{C}_K$ is closed under averaging.

Lemma 2.1. *Let $K > 0$ and suppose θ is left-continuous. If one orders for all x in some interval, then one orders up to a common level.*

Proof. Let $x' < x''$ be two points in the interval from where one orders up to S' and S'' respectively. Then $\theta(S') \leq \theta(S'')$, and $S' < x''$ (otherwise at x'' one should order up to S'). Therefore S' belongs to the interval, and in (x', S') one should order up to S' if at all. For x near S' the inequality $K + \theta(S') < \theta(x)$ is violated, and one does not order, a contradiction. $\square$

Lemma 2.2. *Let $\theta(x)$ is left-continuous and K-convex, and let $x_1 < x_2 < x_3 < x_4$, where it is optimal not to order in (x_1, x_2) and (x_3, x_4). Then it is not optimal to order in (x_2, x_3).*

Proof. Follows from Lemma 2.1. $\square$

Lemma 2.3. *If $\theta(x) \in \mathcal{C}_K$ then an (s, S) policy is optimal.*

Proof. By Lemma 2.2 the optimal policy is either

$$\begin{cases} \text{order} & \text{if } x \leq s \ , \\ \text{do not order} & \text{otherwise} \ , \end{cases}$$

or

$$\begin{cases} \text{do not order} & \text{if } x \leq s \ , \\ \text{order} & \text{otherwise} \ , \end{cases}$$

for some s. By Lemma 2.1 the first case orders up to some $S > s$, and in the second case one orders up to ∞. $\square$

Lemma 2.4. *If $\theta(x)$ is K-convex then so is*

$$\psi(x) = \min_{u \geq 0}\{KH(u) + \theta(x+u)\} \ .$$

Lemma 2.4 says that $\mathbf{L}\phi$ is K-convex if ϕ is. In fact, $\mathbf{L}\phi \in \mathcal{C}_K$ if $\phi \in \mathcal{C}_K$. Since $V_t^* = L^{N-t}\phi$ it follows that V_t^* is K-convex. Optimality of the (s, S)-policy follows from Lemma 2.3.

LECTURE 3

Calculus of Variations

3.1. Necessary Conditions

Consider the problem

$$\min J(y) := \int_a^b F(y(x), y'(x), x)dx \tag{3.1}$$

subject to some conditions on $y(x)$ at $x = a$ and $x = b$. If $y(x)$ is a local minimizer of (3.1) then for any "nearby" function $z(x)$

$$J(z) \geq J(y) \ . \tag{3.2}$$

In particular, consider

$$\begin{aligned} z(x) &:= y(x) + \delta y(x) \ , \ (\text{a } \textbf{test-function}) \ , & (3.3\text{a}) \\ \text{with } \delta y(x) &= \epsilon\eta(x) \ , \ (\text{a } \textbf{variation}) \ , & (3.3\text{b}) \end{aligned}$$

where ϵ is a parameter with small modulus. The test-function z satisfies the same boundary conditions as y, imposing certain conditions on η, see e.g. (3.11). The variations of the derivatives of y are

$$\delta y'(x) = \epsilon\eta'(x) \ , \ \delta y''(x) = \epsilon\eta''(x) \ , \ldots \tag{3.4}$$

provided η is differentiable.

Notation and terminology: Let $\mathcal{Y}$ denote the class of functions:$[a, b] \to \mathbb{R}$ admissible for the above problem. The variation δy is **strong** if $\delta y = \epsilon\eta \to 0$ as $\epsilon \to 0$, and **weak** if also $\delta y' = \epsilon\eta' \to 0$.

We use the following norms,

$$\begin{aligned} \|y\|_0 &:= \sup\{|y(x)| : a \leq x \leq b\} \ , & (3.5\text{a}) \\ \|y\|_1 &:= \sup\{|y(x)| + |y'(x)| : a \leq x \leq b\} \ , & (3.5\text{b}) \end{aligned}$$

and associated neighborhoods of $y \in \mathcal{Y}$,

$$\begin{aligned} U_0(y, \delta) &:= \{z : \|z - y\|_0 \leq \delta\} \ , \ (\text{a } \textbf{strong neighborhood}) \ , & (3.6\text{a}) \\ U_1(y, \delta) &:= \{z : \|z - y\|_1 \leq \delta\} \ , \ (\text{a } \textbf{weak neighborhood}) \ . & (3.6\text{b}) \end{aligned}$$

Note that $U_1(y, \delta) \subset U_0(y, \delta)$. The value $J(y)$ is:

a **strong local minimum** if $J(z) \geq J(y)$ for all $z \in \mathcal{Y} \bigcap U_0(y, \delta)$,

a **weak local minimum** if $J(z) \geq J(y)$ for all $z \in \mathcal{Y} \bigcap U_1(y, \delta)$,

for some $\delta > 0$.

Let $J(y)$ be a weak local minimum. Then for all $y + \epsilon\eta$ in a weak neighborhood of y,

$$\begin{aligned} J(y + \epsilon\eta) &\geq J(y) & (3.7\text{a}) \\ \text{i.e.} \int_a^b F(y + \epsilon\eta, y' + \epsilon\eta', x)dx &\geq \int_a^b F(y, y', x)dx & (3.7\text{b}) \end{aligned}$$

Expanding the LHS of (3.7b) in ϵ we get

$$J(y)+\epsilon\left[\int_a^b (F_y\eta + F_{y'}\eta')\,dx\right] + O(\epsilon^2) \geq J(y) \tag{3.8}$$

and as $\epsilon \to 0$ (since $\epsilon\eta$ is in a weak neighborhood of 0, $\epsilon\eta \to 0$ and $\epsilon\eta' \to 0$),

$$\int_a^b (F_y\eta + F_{y'}\eta')\,dx = 0\ . \tag{3.9}$$

Integrating the second term by parts we get

$$\int_a^b \left[\eta(x)F_y - \eta(x)\frac{d}{dx}(F_{y'})\right] dx + [\eta(x)F_{y'}]_a^b = 0\ . \tag{3.10}$$

The requirement that the test-function z of (3.3) satisfies the same boundary conditions as y implies that

$$\eta(a) = \eta(b) = 0\ , \tag{3.11}$$

killing the last term in (3.10), leaving

$$\int_a^b \eta(x)\left[F_y - \frac{d}{dx}(F_{y'})\right] dx = 0 \tag{3.12}$$

which holds for all admissible $\eta(x)$. It follows from Lemma 3.1 below that

$$F_y - \frac{d}{dx}F_{y'} \equiv 0\ ,\ a \leq x \leq b\ , \tag{3.13}$$

the differential form of the **Euler–Lagrange equation**. This is a 2nd order differential equation in y:

$$F_y - \frac{\partial^2 F}{\partial x \partial y'} - \frac{\partial^2 F}{\partial y \partial y'}\,y' - \frac{\partial^2 F}{\partial y'^2}\,y'' = 0\ . \tag{3.14}$$

LEMMA 3.1. If Φ is continuous in $[a, b]$, and

$$\int_a^b \Phi(x)\,\eta(x)\,dx\ =\ 0\ ,$$

for every continuous function $\eta(x)$ that vanishes at a and b, then $\Phi(x) \equiv 0$ in $[a, b]$.

PROOF. Suppose $\Phi(\xi) > 0$ for some $\xi \in (a, b)$. Then $\Phi(x) > 0$ in some neighborhood of ξ, say $\alpha \leq x \leq \beta$. Let

$$\eta(x)\ :=\ \begin{cases} (x-\alpha)^2(x-\beta)^2 & \text{for } \alpha \leq x \leq \beta\ , \\ 0 & \text{otherwise}\ . \end{cases}$$

Then $\int_a^b \Phi(x)\,\eta(x)\,dx\ >\ 0$. Therefore Φ cannot be positive in $[a, b]$. It similarly cannot be negative. □

Exercises.

EXERCISE 3.1. Derive (3.9) by the condition that $\partial J/\partial\epsilon$ must be zero at $\epsilon = 0$.

EXERCISE 3.2. Does (3.13) follow from (3.12) by using $\eta(x) := F_y - \frac{d}{dx}F_{y'}$ in (3.12)?

3.2. The second variation and sufficient conditions

Maximizing the value of the functional (3.1), instead of minimizing it as in § 3.1, would give exactly the same necessary condition, the Euler–Lagrange equation (3.13). To distinguish between minima and maxima we need a second variation,

Let

$$\phi(\epsilon) := \int_a^b F(y + \epsilon\eta, y' + \epsilon\eta', x)dx \tag{3.15}$$

where y is an extremal solution, and η is fixed. Then

$$\phi'(0) = \int_a^b (F_y\eta(x) + F_{y'}\eta'(x))\,dx \ , \text{ compare with (3.9) }, \tag{3.16a}$$

$$\phi''(0) = \int_a^b \Big(F_{yy}\eta^2(x) + 2F_{yy'}\eta(x)\eta'(x) + F_{y'y'}\eta'^2(x)\Big)\,dx \tag{3.16b}$$

$$= \int_a^b (\eta(x), \eta'(x)) \begin{pmatrix} F_{yy} & F_{yy'} \\ F_{yy'} & F_{y'y'} \end{pmatrix} \begin{pmatrix} \eta(x) \\ \eta'(x) \end{pmatrix} dx \tag{3.16c}$$

We call (3.16a) and (3.16b) the **first** and **second variations** of $J(y)$, respectively.

Assuming the matrix in (3.16c) is positive definite, along the extremal $y(x)$, for all $a \le x \le b$, a sufficient condition for minimum is the **Legendre condition**

$$F_{y'y'} \ > \ 0 \ , \tag{3.17}$$

the reverse inequality sufficient for maximum.

The Legendre condition (3.17) means that F is convex as a function of y' along the extremal trajectory. Another aspect of convexity appears in the **Weierstrass condition**

$$F(y, Y', x) - F(y, y', x) + (Y' - y')F_{y'} \ \ge \ 0 \ , \tag{3.18}$$

for all admissible functions $Y' = Y'(x, y)$. We recognize that (3.18) is the gradient inequality for F, i.e. there is an implicit assumption that F is convex as a function of y'.

Both sufficient conditions (3.17) and (3.18) are strong, and difficult to check even if they hold.

3.3. Calculus of Variations and DP

Consider the problem of minimizing the functional

$$J(y) := \int_a^b F(x, y, y')dx \tag{3.19}$$

subject to the initial condition $y(a) = c$. Introduce the function

$$f(a, c) := \min J(y) \tag{3.20}$$

over all feasible y. Here (a, c) are parameters with $-\infty < a < b$ and $-\infty < c < \infty$. Breaking the integral

$$\int_a^b = \int_a^{a+\Delta} + \int_{a+\Delta}^b \tag{3.21}$$

the *Principle of Optimality* gives

$$f(a,c) = \min_y \left[\int_a^{a+\Delta} F(x,y,y')dx + f(a+\Delta, c(y)\right] \tag{3.22}$$

the minimization is over all y defined on $[a, a+\Delta]$ such that $y(a) = c$, and $c(y) = y(a+\Delta)$.

If Δ is small, the function $y(x)$ can be approximated in $[a, a+\Delta]$ by $y(x) \approx y(a)+y'(a)x = c + y'(a)x$. The choice of y in $[a, a+\Delta]$ is therefore equivalent to the choice of

$$v := y'(a) \ . \tag{3.23}$$

Therefore

$$\int_a^{a+\Delta} F(x,y,y')dx = F(a,c,v)\Delta + o(\Delta) \ , \ c(y) = c + v\Delta + o(\Delta) \ . \tag{3.24}$$

The optimality condition (3.22) then becomes

$$f(a,c) = \min_v \left[F(a,,c,v)\Delta + f(a+\Delta, c+v\Delta\right] + o(\Delta) \ , \tag{3.25}$$

with limit, as $\Delta \to 0$,

$$-\frac{\partial f}{\partial a} = \min_v \left[F(a,c,v) + v\frac{\partial f}{\partial c}\right] \ , \tag{3.26}$$

holding for all $a < b$. The initial condition here is $f(b,c) = 0$ for all c.

To derive the Euler–Lagrange equation, we rewrite (3.25), denoting the endpoint a by x and the initial velocity v by y',

$$f(x,y) = \min_{y'} \left[F(x,y,y')\Delta + f(x,y) + \frac{\partial f}{\partial x}\Delta + \frac{\partial f}{\partial y}y'\Delta + \cdots\right] \tag{3.27}$$

with limit, as $\Delta \to 0$,

$$0 = \min_{y'} \left[F(x,y,y') + \frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y}\right] \tag{3.28}$$

a rewrite of (3.26). This equation is equivalent to the two equations

$$F_{y'} + \frac{\partial f}{\partial y} = 0 \tag{3.29}$$

obtained by partial differentiation w.r.t. y', and

$$F + \frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} = 0 \tag{3.30}$$

which holds form all x, y, y' satisfying (3.29). Differentiating (3.29) w.r.t. x, and (3.30) w.r.t. y we get two equations

$$\frac{d}{dx}F_{y'} + \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2}y' \quad = \quad 0 \ , \tag{3.31}$$

$$F_y + F_{y'}\frac{\partial y'}{\partial y} + \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2}y' + \frac{\partial f}{\partial y}\frac{\partial y'}{\partial y} \quad = \quad 0 \ , \tag{3.32}$$

which, together with (3.29) give the Euler-Lagrange equation

$$\tfrac{d}{dx}F_{y'} - F_y = 0 \ . \tag{3.13}$$

3.4. Special cases

The Euler–Lagrance equation is simplified in the following cases:
(a) The function $F(x, y, y')$ is independent of y'. Then (3.13) reduces to

$$\frac{\partial F(x,y)}{\partial y} = 0$$

and y is implicitly defined (with no guarantee that the boundary conditions are satisfied; these conditions cannot be arbitrary).
(b) $F(x, y, y')$ is independent of y. Then (3.13) gives

$$\frac{d}{dx} F_{y'} = 0 \ , \ \text{ or } \ \frac{\partial F(x,y')}{\partial y'} = C \text{ (constant)}$$

which may be solved for y',

$$y'(x) = f(x, C) \ , \ \therefore \ y(x) = \int f(x, C)\, dx$$

(c) The **autonomous case** where $\frac{\partial F}{\partial x} = 0$. Multiply (3.13) by $y' = \frac{dy}{dx}$

$$\begin{aligned} y' \frac{d}{dx} F_{y'} - y' F_y &= 0 \ . \\ \frac{d}{dx}\left(y' \frac{\partial F}{\partial y'}\right) - y'' \frac{\partial F}{\partial y'} - y' \frac{\partial F}{\partial y} &= 0 \ . \\ \frac{d}{dx}\left(y' \frac{\partial F}{\partial y'} - F(y, y')\right) &= 0 \ . \\ \therefore \ y' \frac{\partial F}{\partial y'} - F(y, y') & \quad \text{is constant} \end{aligned} \tag{3.33}$$

3.5. Extensions

The Euler–Lagrange equation is extended in three ways:
(a) Higher order derivatives. It is required a stationary value of

$$J(y) := \int_a^b F\left(y, \frac{dy}{dx}, \ldots, \frac{d^n y}{dx^n}, x\right) dx \tag{3.34}$$

Using (3.3), the first variation of J is

$$\delta J = \int_a^b \left(\eta F_y + \frac{d\eta}{dx} F_{y'} + \cdots + \frac{d^n \eta}{dx^n} F_{y^{(n)}}\right) dx$$

Successive integration by parts (assuming that $y, y', \cdots, y^{(n)}$ are given at $x = a$ and $x = b$) gives

$$\int_a^b \left(F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n} F_{y^{(n)}}\right) \eta(x)\, dx = 0$$

and by Lemma 3.1,

$$F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n} F_{y^{(n)}} \equiv 0 \ , \ a \le x \le b \ . \tag{3.35}$$

(b) Several dependent variables. Here

$$J(y) := \int_a^b F\left(y_1, y_2, \cdots, y_n, y_1', y_2', \cdots, y_n', x\right) dx$$

and a similar analysis gives the necessary conditions

$$F_{y_i} - \frac{d}{dx} F_{y_i'} \equiv 0\ ,\ i \in \overline{1,n} \tag{3.36}$$

(c) Several independent variables. Let $y = y(x_1, \cdots, x_n)$ and

$$J(y) = \int\int_D \cdots \int F\left(y, \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \cdots, \frac{\partial y}{\partial x_n}, x_1, x_2, \cdots, x_n\right) dx_1\, dx_2 \cdots dx_n$$

The first variation is

$$\epsilon \int\int_D \cdots \int \left(\eta \frac{\partial F}{\partial y} + \sum_{k=1}^{n} \frac{\partial \eta}{\partial x_k} F_k\right) dx_1\, dx_2 \cdots dx_n$$

where F_k is the derivative of F w.r.t. $\dfrac{\partial y}{\partial x_k}$. The analog of integration by parts is Green's Theorem. Let ∂D denote the boundary of D. Then

$$\begin{aligned} \int\int_D \cdots \int \left(\sum_{k=1}^{n} \frac{\partial \eta}{\partial x_k} F_k\right) dx_1\, dx_2 \cdots dx_n \ &= \ \int\int_{\partial D} \cdots \int \sum_{k=1}^{n} (\eta F_k)\, \mathbf{n}_k \cdot d\mathbf{s} - \\ &\quad - \int\int_D \cdots \int \sum_{k=1}^{n} \eta \left(\frac{\partial F_k}{\partial x_k}\right) dx_1\, dx_2 \cdots dx_n \end{aligned}$$

where $\mathbf{n}_k$ is unit normal to the surface ∂D. If y is given on ∂D then $\eta = 0$ there, and the Euler–Lagrange equation becomes

$$F_y - \sum_{k=1}^{n} \frac{\partial F_k}{\partial x_k} = 0\ . \tag{3.37}$$

3.6. Hidden variational principles

Given a differential equation, is it the Euler–Lagrange equation of a variational problem? We illustrate this for the 2nd order linear ODE

$$\frac{d}{dx}\left(a(x)\frac{dy}{dx}\right) + b(x)y + \lambda c(x)y \ = \ 0 \tag{3.38a}$$

$$y(x_0) = y(x_1) \ = \ 0 \tag{3.38b}$$

The values of λ for which a solution exists are the **eigenvalues** of (3.38). Multiply (3.38a) by the variation $\delta y(x)$ and integrate

$$\int_{x_0}^{x_1} dx \left(\delta y \frac{d}{dx}\left(a(x)\frac{dy}{dx}\right) + b(x)y\delta y + \lambda c(x) y \delta y\right) = 0$$

Using $y\,\delta y = \frac{1}{2}\,\delta(y^2)$ and integrating the first term by parts we get

$$\int_{x_0}^{x_1} dx\,\left(-a(x)\frac{dy}{dx}\,(\delta y)' + \frac{b(x)}{2}\delta y^2 + \frac{\lambda c(x)}{2}\delta y^2\right) = 0$$

Using the same notation for the variation of $(y')^2$,

$$\frac{dy}{dx}(\delta y)' = y'(\delta y)' = y'\delta y' = \frac{1}{2}\delta y'^2$$

and therefore

$$\begin{aligned}
\int_{x_0}^{x_1} \delta\left(-a(x)y'^2(x) + [b(x)+\lambda c(x)]\,y^2(x)\right)\,dx &= 0\\
\therefore\ \lambda\,\delta\int_{x_0}^{x_1} c(x)y^2(x)\,dx - \delta\int_{x_0}^{x_1}\left(a(x)y'^2(x) - b(x)y^2(x)\right)\,dx &= 0\\
\text{or } \delta(J_1 - \lambda J_2) &= 0
\end{aligned}$$

$$\begin{aligned}
\text{where } J_1(y) &= \int_{x_0}^{x_1}\left(a(x)y'^2(x) - b(x)y^2(x)\right)\,dx\\
J_2(y) &= \int_{x_0}^{x_1} c(x)y^2(x)\,dx
\end{aligned}$$

Therefore any solution of (3.38) is an extremal of the variational problem of finding the stationary values of

$$\lambda = \frac{J_1}{J_2} = \frac{\int_{x_0}^{x_1}\left(a(x)y'^2(x) - b(x)y^2(x)\right)\,dx}{\int_{x_0}^{x_1} c(x)y^2(x)\,dx} \tag{3.39}$$

EXAMPLE 3.1. The eigenvalues of

$$y'' + \lambda y = 0 \tag{3.40a}$$

$$y(0) = y(1) = 0 \tag{3.40b}$$

are π^2, $4\pi^2$, $9\pi^2, \cdots$. Using (3.39) these eigenvalues are the stationary values of the ratio

$$\lambda = \frac{\int_0^1 y'^2\,dx}{\int_0^1 y^2\,dx} \tag{3.41}$$

Exercises.

EXERCISE 3.3. Reverse the above argument, starting from the variational problem (3.39) and deriving the Euler–Lagrange equation.

EXERCISE 3.4. Use a trial solution of (3.40)

$$y_0(x) \;=\; x(1-x)(1+\alpha x)$$

where α is a changing parameter. Compute RHS(3.41), and minimize w.r.t α to obtain an extimate of the smallest eigenvalue $\lambda = \pi^2$.

3.7. Integral constraints

Consider the problem of minimizing (3.19) subject to the additional constraint

$$\int_a^b G(x,y,y')dx = z \;, \tag{3.42}$$

where z is given. We denote

$$f(x,y,z) := \min_y \int_a^b F(x,y,y')dx \tag{3.43}$$

subject to (3.42). Such a problem is called **isoperimetric**, for historic reasons (see Example 3.2). The analog of (3.25) is

$$f(x,y,z) = \min_{y'} \left[F(x,y,y')\Delta + f(x+\Delta, y+y'\Delta, z - G(x,y,y')\Delta\right] \tag{3.44}$$

with limit

$$0 = \min_{y'} \left[F(x,y,y') + \frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} - G(x,y,y')\frac{\partial f}{\partial z}\right] \;. \tag{3.45}$$

Therefore,

$$0 = F_{y'} + \frac{\partial f}{\partial y} - G_y \frac{\partial f}{\partial z} \;, \tag{3.46}$$

and

$$0 = F + \frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} - G\frac{\partial f}{\partial z} \;, \tag{3.47}$$

holding for all x, y, z, y' satisfying (3.46). Differentiating (3.46) w.r.t. x and (3.47) w.r.t. y, and combining, we get

$$\frac{\partial}{\partial y'}\left(F - \frac{\partial f}{\partial z}G\right) - \frac{\partial}{\partial y}\left(F - \frac{\partial f}{\partial z}G\right) = 0 \tag{3.48}$$

Partial differentiation of (3.47) w.r.t. z yields

$$\frac{\partial^2 f}{\partial x \partial z} + y'\frac{\partial^2 f}{\partial y \partial z} - G\frac{\partial^2 f}{\partial z^2} \;=\; 0 \;. \tag{3.49}$$

$$\therefore \;\; \frac{d}{dx}\left(\frac{\partial f}{\partial z}\right) \;=\; 0 \;, \tag{3.50}$$

$$\tag{3.51}$$

i.e. $\frac{\partial f}{\partial z}$ is a constant, say λ, and (3.48) is the Euler-Lagrange equation for $F - \lambda G$. The parameter λ is the Lagrange multiplier of the constraint (3.42).

EXAMPLE 3.2. ([**10**, p. 22]) The **Dido problem** is to find the smooth plane curve of perimeter L which encloses the maximum area. We show that this curve is a circle. Fix any point P of the curve C, and use polar coordinates with origin at P, and $\theta = 0$ along the (half) tangent of C at P. The area enclosed by C is $A = \int_0^\pi \frac{1}{2} r^2 \, d\theta$, and the perimeter is $L = \int_0^\pi \left(\left(\frac{dr}{d\theta}\right)^2 + r^2 \right)^{1/2} d\theta$. Define $F := A + \lambda L$, and maximize

$$F = \int_0^\pi \left[\tfrac{1}{2} r^2 + \lambda \left(\left(\frac{dr}{d\theta} \right)^2 + r^2 \right)^{1/2} \right] d\theta = \int_0^\pi \left[\tfrac{1}{2} r^2 + \lambda T \right] d\theta$$

subject to $r(0) = r(\pi) = 0$. The Euler-Lagrange equation gives, using (3.33),

$$\begin{aligned}
\frac{dr}{d\theta} \frac{\lambda \, dr/d\theta}{T} - \lambda T - \tfrac{1}{2} r^2 &= \text{constant} \\
\therefore \ \frac{\lambda (dr/d\theta)^2 - \lambda T^2}{T} - \tfrac{1}{2} r^2 &= \text{constant} \\
\therefore \ \frac{\lambda r^2}{T} + \tfrac{1}{2} r^2 &= \text{constant} = 0 \ , \ \text{since } r = 0 \text{ is on } C \\
\therefore \ 2\lambda + \left(\left(\frac{dr}{d\theta} \right)^2 + r^2 \right)^{1/2} &= 0 \ . \quad \therefore \ \left(\frac{dr}{d\theta} \right)^2 + r^2 = 4\lambda^2 \\
\therefore \ \frac{d\theta}{dr} &= \frac{1}{(4\lambda^2 - r^2)^{1/2}} \ . \quad \therefore \ \theta = \sin^{-1}\left(-\frac{r}{2\lambda} \right) \\
\therefore \ r &= -2\lambda \sin\theta
\end{aligned}$$

a circle of radius $-\lambda$. The radius is determined from

$$L = \int_0^\pi \left(\left(\frac{dr}{d\theta} \right)^2 + r^2 \right)^{1/2} d\theta = -\int_0^\pi 2\lambda d\theta = -2\pi\lambda \ ,$$

therefore the radius is $L/2\pi$.

EXAMPLE 3.3. *(*[**10**, p. 23]*)*

$$\text{minimize} \int_0^1 \dot{x}^2 \, dt \quad \text{s.t.} \quad \int_0^1 x \, dt = 0 \ , \ \int_0^1 xt \, dt = 1 \text{ and } x(0) = x(1) = 0 \ .$$

Let $F(x, \dot{x}, t) := \dot{x}^2 + \lambda x + \mu x t$. The Euler–Lagrange equation is $2\ddot{x} - \lambda - \mu t = 0$

$$\begin{aligned}
\therefore \ x &= \tfrac{1}{4}\lambda t^2 + \tfrac{1}{6}\mu t^3 + At \ , \ \text{since } x(0) = 0 \\
\therefore \ \tfrac{1}{4}\lambda + \tfrac{1}{6}\mu + A &= 0 \ , \ \text{since } x(1) = 0 \\
\therefore \ \tfrac{1}{12}\lambda + \tfrac{1}{24}\mu + \tfrac{1}{2}A &= 0 \ , \ \text{since} \int_0^1 x \, dt = 0 \\
\therefore \ \tfrac{1}{16}\lambda + \tfrac{1}{30}\mu + \tfrac{1}{3}A &= 1 \ , \ \text{since} \int_0^1 xt \, dt = 1
\end{aligned}$$

three linear equations with solution $\lambda = -\mu = 360$, $A = 30$. Finally $x = 90t^2 - 60t^3 + 30$.

Exercises.

EXERCISE 3.5. A flexible fence of length L is used to enclose an area bounded on one side by a straight wall. Find the maximum area that can be enclosed.

EXERCISE 3.6. (The **catenary equation**) A flexible uniform cable of length $2a$ hangs between the fixed points $(0,0)$ and $(2b,0)$, where $b < a$. Find the curve $y = y(x)$ minimizing

$$\int_0^{2b} (1 + (y')^2)^{1/2}\, y\, dx \quad \text{(potential energy)}$$

$$s.t. \quad \int_0^{2b} (1 + (y')^2)^{1/2}\, dx \;=\; 2a \quad \text{(given length)}$$

Answer: $y/k = \cosh(b/k) - \cosh\{(x-b)/k\}$, where k is from $a = k \sinh b/k$.

EXERCISE 3.7.

$$\text{Minimize} \int\int_A \left(\frac{\partial f}{\partial x}^2 + \frac{\partial f}{\partial y}^2 \right) dxdy \ , \ \text{where } A := \{(x,y) : x^2 + y^2 \leq 1\}$$

s.t. $\int\int_A f\, dxdy \;=\; B$ with f given on ∂A. Show that f satisfies the partial differential equation

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + kf = 0$$

where k is a constant to be determined.

3.8. Natural boundary conditions

If $y(a)$ is not specified then

$$\frac{\partial f}{\partial y} \mid_{x=a} = 0 \tag{3.52}$$

for otherwise there is a better starting point. Then (3.29) gives

$$F_{y'} \mid_{x=a} = 0 \ , \tag{3.53}$$

a *natural boundary condition* at the free endpoint $x = a$.

Another case is where even $x = a$ is unspecified, and the y curve is only required to start somewhere on a given curve $y = g(x)$. Then the change in f as the initial point varies along $y = g(x)$ must be zero,

$$\frac{\partial f}{\partial x} + g' \frac{\partial f}{\partial y} = 0 \tag{3.54}$$

and by (3.30),

$$F + y' \frac{\partial f}{\partial y} - g' \frac{\partial f}{\partial y} = 0 \tag{3.55}$$

and finally, by (3.29),

$$F + (g' - y')F_{y'} = 0 \ , \tag{3.56}$$

the *transversality condition* at the free initial point.

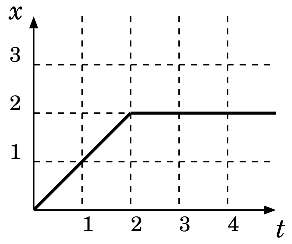

FIGURE 3.1. At the corner, $t = 2$, the solution switches from $x' = 1$ to $x = 2$

3.9. Corners

Consider the Calculus of Variations problem

$$\begin{array}{ll} \text{opt} & \displaystyle\int_o^T F(t, x, x')\, dt \\ \text{s.t.} & x(0), x(T) \text{ given} \\ & x \text{ piecewise smooth} \end{array}$$

An optimal x^* satisfies the Euler-Lagrange equation

$$F_x = \frac{d}{dt} F_{x'}$$

in any subinterval of $[0, T]$ where x^* is continuously differentiable. The discontinuity points of $\frac{d}{dt}x^*$ are called its **corners**.
The **Weierstrass-Erdmann corner conditions**: The functions $F_{x'}$, $F - x'F_{x'}$ are continuous in corners.

EXAMPLE 3.4. Consider

$$\begin{array}{ll} \min & \displaystyle\int_0^3 (x-2)^2(x'-1)^2 dt \\ \text{s.t.} & x(0) = 0, x(3) = 2 \end{array}$$

The lower bound, 0, is attained if

$$x = 2 \text{ or } x' = 1\ ,\ \ 0 \le t \le 3\ .$$

This function has a corner at $t = 2$, but

$$F_{x'} = 2(x-2)^2(x'-1) \text{ and } F - x'F_{x'} = (x-2)^2(x'-1)(x'+1)$$

are continuous at $t = 2$.

LECTURE 4

The Variational Principles of Mechanics

This lecture is based on [**23**] and [**24**]. Other useful refernces are [**24**], [**20**] and [**32**].

4.1. Mechanical Systems

A **particle** is a point mass, i.e. mass concentrated in a point. The **position r**, **velocity v** and **acceleration a** of a particle in $\mathbb{R}^3$ are

$$\mathbf{r} = \mathbf{r}(t) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \ , \ \mathbf{v} = \dot{\mathbf{r}} = \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} \ , \ \mathbf{a} = \dot{\mathbf{v}} = \ddot{\mathbf{r}} \ .$$

The norm of the velocity vector $\mathbf{v}$ is called **speed**, and denoted by $v = \|\mathbf{v}\|$.

Consider a system of N particles in $\mathbb{R}^3$. Describing it requires $3N$ coordinates, less if the N particles are not independent. The number of degrees of freedom of the system is the minimal number of coordinates describing it. If the system has s degrees of freedom, let

$\mathbf{q} := (q_1, q_2, \ldots, q_s)$ denote its **generalized coordinates**,

$\dot{\mathbf{q}} := (\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_s)$ its **generalized velocities**.

The position & motion of the system are determined by the $2s$ numbers $(\mathbf{q}, \dot{\mathbf{q}})$.

Example 4.1. A double pendulum in planar motion, see Fig. 4.1, has two degrees of freedom. It is described in terms of the angles ϕ_1 and $\phi_2\}$. The generalized coordinates and velocities can be taken as $\{\phi_1, \phi_2\}$, and $\{\dot{\phi}_1, \dot{\phi}_2\}$ respectively.

4.2. Hamilton's Principle of Least Action

A mechanical system is characterized by a function

$$L = L(\mathbf{q}, \dot{\mathbf{q}}, t) = L(q_1, \ldots, q_s, \dot{q}_1, \ldots, \dot{q}_s, t) \tag{4.1}$$

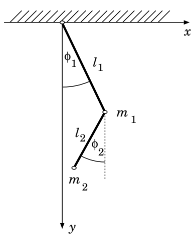

Figure 4.1. A double pendulum.

Hamilton's Principle, [**23**, pp. 111-114]: The motion of a mechanical system between any two points $(t_1, \mathbf{q}(t_1))$ and $(t_2, \mathbf{q}(t_2))$ occurs in such a way that the definite integral

$$S := \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t)\, dt \tag{4.2}$$

becomes stationary for arbitrary feasible variations.
L is called the **Lagrangian** of the system, and $S = \int L\, dt$ its **action**.

4.3. The Euler-Lagrange Equation

Let $s = 1$, $q = q(t) \in \operatorname{argmin} S$ and consider a perturbed path $q(t) + \delta q(t)$ where

$$\delta q(t_1) = \delta q(t_2) = 0\ .$$

$$\begin{aligned}
\delta S &:= \int_{t_1}^{t_2} L(q + \delta q, \dot{q} + \delta \dot{q}, t)\, dt -_{t_1}^{t_2} L(q, \dot{q}, t)\, dt \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}}\, \delta \dot{q} \right) dt \\
&= \left[\frac{\partial L}{\partial \dot{q}}\, \delta q \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q\, dt\ ,\ \text{since } \delta \dot{q} = \frac{d}{dt} \delta q \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q\, dt\ ,\ \text{since } \delta q(t_1) = \delta q(t_2) = 0\ .
\end{aligned}$$

Since δq is arbitrary, we conclude that $\delta S = 0$ iff

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0\ , \tag{4.3}$$

the **Euler-Lagrange equation** (3.13), a necessary condition for minimal action.

The Euler–Lagrange equations for a system with s degrees of freedom are

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0\ ,\ i \in \overline{1, s}\ . \tag{4.4}$$

4.4. Nonuniqueness of the Lagrangian

Consider

$$\widehat{L}(\mathbf{q}, \dot{\mathbf{q}}, t) := L(\mathbf{q}, \dot{\mathbf{q}}, t) + \frac{d}{dt} f(\mathbf{q}, t)\ , \tag{4.5}$$

the last term does not depend on $\dot{\mathbf{q}}$. Then

$$\begin{aligned}
\widehat{S} &= \int_{t_1}^{t_2} \widehat{L}(\mathbf{q}, \dot{\mathbf{q}}, t)\, dt = S + f(\mathbf{q}(t_2), t_2) - f(\mathbf{q}(t_1), t_1)\ . \\
\therefore\ \delta \widehat{S} &= 0 \quad \Longleftrightarrow \quad \delta S = 0\ .
\end{aligned}$$

The Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ is therefore defined up to derivative $\frac{d}{dt} f(\mathbf{q}, t)$.

4.5. The Galileo Relativity Principle

An **inertial frame** (of reference) is one in which **time** is homogeneous and **space** is both **homogeneous** and **isotropic**. In such a system, a free particle at rest remains at rest.

Consider a particle moving freely in an inertial frame. Its Lagrangian must be independent of t, $\mathbf{r}$ and the direction of $\mathbf{v}$.

$$\begin{aligned} \therefore\ L &= L(\|\mathbf{v}\|) = L(v^2)\ . \\ \therefore\ \frac{d}{dt}\left(\frac{\partial L}{\partial \mathbf{v}}\right) &= 0\ ,\ \text{by the Euler–Lagrange equation}\ . \\ \therefore\ \frac{\partial L}{\partial \mathbf{v}} &= 0 \quad \therefore\ \mathbf{v}\ =\ \text{constant}\ . \end{aligned}$$

Law of Inertia. In an inertial frame a free motion has a constant velocity.

Consider two frames, K and $\widehat{K}$, where K is inertial, and $\widehat{K}$ moves relative to K with velocity $\mathbf{v}$. The coordinates $\mathbf{r}$ and $\widehat{\mathbf{r}}$ of a given point are related by

$$\mathbf{r} = \widehat{\mathbf{r}} + \mathbf{v}\,t$$

(time is absolute, i.e. $t = \widehat{t}$).

Galileo's Principle. The laws of mechanics are the same in K and $\widehat{K}$.

4.6. The Lagrangian of a Free Particle

Let an inertial frame K move with an infitesimal velocity $\boldsymbol{\varepsilon}$ relative to another inertial frame $\widehat{K}$. A free particle has velocities $\{\mathbf{v}, \widehat{\mathbf{v}}\}$ and Lagrangians $\{L, \widehat{L}\}$ in these two frames.

$$\begin{aligned} \therefore\ \widehat{\mathbf{v}} &= \mathbf{v} + \boldsymbol{\varepsilon}\ . \\ \therefore\ \widehat{L} &= L(\widehat{v}^2) = L\left(v^2 + 2\mathbf{v}\cdot\boldsymbol{\varepsilon} + \varepsilon^2\right) \\ &\approx L(v^2) + \frac{\partial L}{\partial(v^2)} 2\mathbf{v}\cdot\boldsymbol{\varepsilon} \\ &= L + \frac{d}{dt} f(\mathbf{r}, t)\ ,\ \text{see (4.5)}\ . \\ &\therefore\ \frac{\partial L}{\partial(v^2)} 2\mathbf{v}\cdot\boldsymbol{\varepsilon}\ \text{is linear in}\ \mathbf{v}\ . \\ &\therefore\ \frac{\partial L}{\partial(v^2)}\ \text{is independent of}\ \mathbf{v}\ . \\ \therefore\ L &= \tfrac{1}{2} m\, v^2\ , \end{aligned} \tag{4.6}$$

where m is the **mass** of the particle.

4.7. The Lagrangian of a System

Consider a system with several particles. If they do not interact, the Lagrangian of the system is the sum of (4.6) terms,

$$L = \sum \tfrac{1}{2} m_k\, v_k^2\ . \tag{4.7}$$

If the particles interact with each other, but not with anything outside the system (in which case the system is called **closed**), the Lagrangian is

$$L = \sum \tfrac{1}{2}\, m_k\, v_k^2 - U(\mathbf{r}_1, \mathbf{r}_2, \ldots) \tag{4.8}$$

where $\mathbf{r}_k$ is the position of the kth particle. Here
$T := \sum \frac{1}{2}\, m_k\, v_k^2$ the **kinetic energy**,
$U := U(\mathbf{r}_1, \mathbf{r}_2, \ldots)$ the **potential energy**.
If t is replaced by $-t$ the Lagrangian does not change (time is isotropic).

4.8. Newton's Equation

The equations of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) = \frac{\partial L}{\partial q_k}\ ,\ \ k = 1, 2, \ldots$$

can be written, using the **generalized momentums**,

$$\begin{aligned} p_k \ &:= \ \frac{\partial L}{\partial \dot{q}_k} \qquad (4.9)\\ \text{as}\ \ \frac{dp_k}{dt} \ &= \ -\frac{\partial U}{\partial q_k}\ ,\ \ k = 1, 2, \cdots \end{aligned}$$

In particular, the Lagrangian (4.8) gives

$$\frac{d}{dt}\,(m_k \mathbf{v}_k) = -\frac{\partial U}{\partial \mathbf{r}_k}\ ,\ \ k = 1, 2, \cdots \tag{4.10}$$

If m_k is constant, (4.10) becomes

$$m_k \frac{d\mathbf{v}_k}{dt} = -\frac{\partial U}{\partial \mathbf{r}_k} \tag{4.11}$$

4.9. Interacting Systems

Let $\mathcal{A}$ be a system interacting with another system $\mathcal{B}$ ($\mathcal{A}$ moves in a given external field due to $\mathcal{B}$), and let $\mathcal{A} + \mathcal{B}$ be closed.

Using generalized coordinates $\mathbf{q}_{\mathcal{A}}, \mathbf{q}_{\mathcal{B}}$,

$$\begin{aligned} L_{\mathcal{A}+\mathcal{B}} \ &= \ T_{\mathcal{A}}(\mathbf{q}_{\mathcal{A}}, \dot{\mathbf{q}}_{\mathcal{A}}) + T_{\mathcal{B}}(\mathbf{q}_{\mathcal{B}}, \dot{\mathbf{q}}_{\mathcal{B}}) - U(\mathbf{q}_{\mathcal{A}}, \mathbf{q}_{\mathcal{B}})\ .\\ \therefore\ L_{\mathcal{A}} \ &= \ T_{\mathcal{A}}(\mathbf{q}_{\mathcal{A}}, \dot{\mathbf{q}}_{\mathcal{A}}) - U(q_{\mathcal{A}}, q_{\mathcal{B}}(t))\ . \end{aligned}$$

The potential energy may depend on t.

Example 4.2. A single particle in an external field,

$$\begin{aligned} L \ &= \ \tfrac{1}{2}\, m\, v^2 - U(\mathbf{r}, t)\ .\\ \therefore\ m\dot{\mathbf{v}} \ &= \ -\frac{\partial U}{\partial \mathbf{r}}\ . \end{aligned}$$

A **uniform field** is where $U = -\mathbf{F}\cdot\mathbf{r}$ where $\mathbf{F}$ is a constant force.

Example 4.3. Consider the double pendulum of Fig. 4.1. For the first particle,

$$\begin{aligned} T_1 &= \tfrac{1}{2} m_1 l_1^2 \dot{\phi}_1^{\,2} \\ U_1 &= -m_1 g\, l_1 \cos\phi_1 \end{aligned}$$

For the second particle,

$$\begin{aligned} x_2 &= l_1 \sin\phi_1 + l_2 \sin\phi_2 \\ y_2 &= l_1 \cos\phi_1 + l_2 \cos\phi_2 \\ \therefore\ T_2 &= \tfrac{1}{2} m_2 \left(\dot{x}_2^2 + \dot{y}_2^2\right) \\ &= \tfrac{1}{2} m_2 \left(l_1^2 \dot{\phi}_1^{\,2} + l_2^2 \dot{\phi}_2^{\,2} + 2 l_1 l_2 \cos(\phi_1 - \phi_2)\dot{\phi}_1\dot{\phi}_2\right) \\ \therefore\ L &= \tfrac{1}{2}(m_1 + m_2) l_1^2 \dot{\phi}_1^{\,2} + \tfrac{1}{2} m_2 l_2^2 \dot{\phi}_2^{\,2} + m_2 l_1 l_2 \cos(\phi_1 - \phi_2)\dot{\phi}_1\dot{\phi}_2 + \\ &\qquad + (m_1 + m_2) g l_1 \cos\phi_1 + m_2 g l_2 \cos\phi_2\ . \end{aligned}$$

4.10. Conservation of Energy

The homogeneity of time means that the Lagrangian of a closed system does not depend explicitly on time,

$$\frac{dL}{dt} = \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i$$

with no term $\frac{\partial L}{\partial t}$. Using the Euler-Lagrange equations $\frac{\partial L}{\partial q_i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}$ we write

$$\begin{aligned} \frac{dL}{dt} = \sum_i \dot{q}_i \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i &= \sum_i \frac{d}{dt}\left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) \\ \therefore\ \frac{d}{dt}\left(\sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L\right) &= 0\ . \end{aligned}$$

Therefore

$$E := \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L \tag{4.12}$$

remains constant during the motion of a closed system, see also (3.33).

If $T(\mathbf{q}, \dot{\mathbf{q}})$ is a quadratic function of $\dot{\mathbf{q}}$, and $L = T(\mathbf{q}, \dot{\mathbf{q}}) - U(\mathbf{q})$, then (4.12) becomes

$$\begin{aligned} E &= \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = 2T - L \\ &= T + U \end{aligned} \tag{4.13}$$

the total energy.

4.11. Conservation of Momentum

The homogeneity of space implies that the Lagrangian is unchanged under a translation of the space,

$$\begin{aligned} \widehat{\mathbf{r}} &= \mathbf{r} + \boldsymbol{\varepsilon} \\ \therefore\ \delta L &= \sum_i \frac{\partial L}{\partial \mathbf{r}_i} \cdot \delta \mathbf{r}_i = \boldsymbol{\varepsilon} \cdot \sum_i \frac{\partial L}{\partial \mathbf{r}_i} = 0 \end{aligned}$$

Since $\boldsymbol{\varepsilon}$ is arbitrary,

$$\delta L = 0 \iff \sum_i \frac{\partial L}{\partial \mathbf{r}_i} = 0$$

$$\begin{aligned} \therefore\ \sum_i \frac{d}{dt}\frac{\partial L}{\partial \mathbf{v}_i} &= 0\ ,\ \text{by Euler–Lagrange}\ . \\ \therefore\ \frac{d}{dt}\sum_i \frac{\partial L}{\partial \mathbf{v}_i} &= 0\ . \end{aligned}$$

Therefore

$$\mathbf{P} := \sum_i \frac{\partial L}{\partial \mathbf{v}_i} \tag{4.14}$$

remains constant during motion. For the Lagrangian (4.8),

$$\mathbf{P} = \sum_i m_i \mathbf{v}_i \tag{4.15}$$

the **momentum** of the system.

Newton's 3rd Law. From $\sum_i \frac{\partial L}{\partial \mathbf{r}_i} = 0$ we conclude $\sum_i \mathbf{F}_i = 0$ where $\mathbf{F}_i = -\frac{\partial U}{\partial \mathbf{r}_i}$, i.e. the sum of forces on all particles in a closed system is zero.

4.12. Conservation of Angular Momentum

The isotropy of space implies that the Lagrangian is invariant under an infinitesimal rotation $\delta\phi$, see Fig. 4.2

$$\delta \mathbf{r} = \delta\phi \times \mathbf{r}\ ,$$

or $\|\delta \mathbf{r}\| = r \sin\theta \|\delta\phi\|$. Similarly,

$$\delta \mathbf{v} = \delta\phi \times \mathbf{v}\ .$$

Substituting these in

$$\delta L = \sum_i \left(\frac{\partial L}{\partial \mathbf{r}_i} \cdot \delta \mathbf{r}_i + \frac{\partial L}{\partial \mathbf{v}_i} \cdot \delta \mathbf{v}_i \right) = 0\ ,$$

and using $\mathbf{p}_i = \frac{\partial L}{\partial \mathbf{v}_i}$, $\dot{\mathbf{p}}_i = \frac{\partial L}{\partial \mathbf{r}_i}$ we get

$$\sum_i \left(\dot{\mathbf{p}}_i \cdot \delta\phi \times \mathbf{r}_i + \mathbf{p}_i \cdot \delta\phi \times \mathbf{v}_i \right) = 0$$

or

$$\delta\phi \cdot \sum_i \left(\mathbf{r}_i \times \dot{\mathbf{p}}_i + \mathbf{v}_i \times \mathbf{p}_i \right) = \delta\phi \cdot \frac{d}{dt} \sum_i \mathbf{r}_i \times \mathbf{p}_i = 0\ .$$

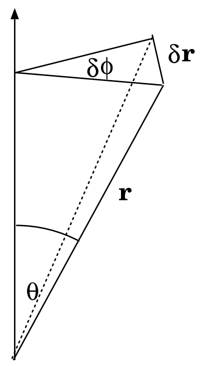

FIGURE 4.2. An infitesimal rotation.

Since $\delta\phi$ is arbitrary, we conclude

$$\frac{d}{dt}\sum_i \mathbf{r}_i \times \mathbf{p}_i = 0$$

or, the **angular momentum**

$$\mathbf{M} = \sum_i \mathbf{r}_i \times \mathbf{p}_i \tag{4.16}$$

is conserved in the motion of a closed system.

4.13. Mechanical Similarity

Let the potential energy be a homogeneous function of degree k, i.e.

$$U(\alpha\mathbf{r}_1, \alpha\mathbf{r}_2, \ldots, \alpha\mathbf{r}_n) = \alpha^k\, U(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) \tag{4.17}$$

for all α. Consider a change of units,

$$\begin{aligned} \widehat{\mathbf{r}}_i &= \alpha\mathbf{r}_i \ , \\ \widehat{t} &= \beta t \ . \\ \therefore\ \widehat{\mathbf{v}}_i &= \frac{\alpha}{\beta}\mathbf{v}_i \ . \\ \therefore\ \widehat{T} &= \frac{\alpha^2}{\beta^2} T \ . \end{aligned}$$

If

$$\frac{\alpha^2}{\beta^2} = \alpha^k \ , \text{ i.e. if } \beta = \alpha^{1-k/2} \ ,$$

then

$$\widehat{L} = \widehat{T} - \widehat{U} = \alpha^k(T - U) = \alpha^k\, L$$

and the equations of motion are unchanged. If a path of length l is travelled in time t, then the corresponding path of length $\widehat{l}$ is travelled in time $\widehat{t}$ where

$$\frac{\widehat{t}}{t} = \left(\frac{\widehat{l}}{l}\right)^{1-k/2} \ . \tag{4.18}$$

EXAMPLE 4.4. [Coulomb Force] If $k = -1$ then (4.18) gives

$$\frac{\widehat{t}}{t} = \left(\frac{\widehat{l}}{l}\right)^{3/2} \tag{4.19}$$

that is **Kepler's 3rd Law**.

EXAMPLE 4.5. For small oscilations and $k = 2$,

$$\frac{\widehat{t}}{t} = 1$$

independet of amplitude.

4.14. Hamilton's equations

Writing

$$\begin{aligned}
dL &= \sum_i \frac{\partial L}{\partial q_i} dq_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \\
&= \sum_i \dot{p}_i dq_i + \sum_i p_i d\dot{q}_i \\
&= \sum_i \dot{p}_i dq_i + d\left(\sum_i p_i \dot{q}_i\right) - \sum_i \dot{q}_i dp_i \\
\therefore\ d\left(\sum p_i \dot{q}_i - L\right) &= -sum_i \dot{p}_i dq_i + \sum_i \dot{q}_i dp_i
\end{aligned}$$

Defining the **Hamiltonian** (**Legendre transform** of L),

$$H = H(p, q, t) := \sum_i p_i \dot{q}_i - L \tag{4.20}$$

we have

$$dH = -\sum_i \dot{p}_i dq_i + \sum_i \dot{q}_i dp_i\ .$$

Comparing with

$$dH = \sum_i \frac{\partial H}{\partial q_i} dq_i + \sum_i \frac{\partial H}{\partial p_i} dp_i$$

we get the **Hamilton equations**

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{4.21}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{4.22}$$

The total time derivative of H is

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_i \frac{\partial H}{\partial q_i} \dot{q}_i + \sum_i \frac{\partial H}{\partial p_i} \dot{p}_i\ ,$$

and by (4.21)–(4.22),

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \ .$$

4.15. The Action as Function of Coordinates

The action $S = \int_{t_1}^{t_2} L dt$ has the Lagrangian as its time derivative, $\frac{dS}{dt} = L$.

$$\begin{aligned}
\delta S &= \left[\frac{\partial L}{\partial \dot q}\delta q\right]_{t_1}^{t_2} + \int_{t_1}^{t_2}\left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot q}\right)\delta q dt \\
&= \sum_i p_i \delta q_i \quad \therefore \ \frac{\partial S}{\partial q_i} = p_i \\
\therefore \ L &= \frac{dS}{dt} = \frac{\partial S}{\partial t} + \sum_i \frac{\partial S}{\partial q_i}\dot q_i = \frac{\partial S}{\partial t} + \sum_i p_i \dot q_i \\
\therefore \ \frac{\partial S}{\partial t} &= L - \sum_i p_i \dot q_i = -H \\
\therefore \ dS &= \sum_i p_i dq_i - H dt
\end{aligned}$$

Maupertuis' Principle. The motion of a mechaincal system minimizes

$$\int \left(\sum_i p_i dq_i\right) dt$$

True, if energy is conserved.

4.16. The Hamilton–Jacoby equations

$$\frac{\partial S}{\partial t} + H = 0 \tag{4.23}$$

or

$$\frac{\partial S}{\partial t} + H(q_1, \ldots, q_s, p_1, \ldots, p_s, t) = 0 \tag{4.24}$$

Solution. If $\frac{\partial H}{\partial t} = 0$ then

$$\frac{\partial S}{\partial t} + H\left(q_1, \ldots, q_s, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_s}\right) = 0$$

with solution

$$S = -ht + V(q_1, \ldots, q_s, \alpha_1, \ldots, \alpha_s, h)$$

where $h, \alpha_1, \ldots, \alpha_s$ are arbitrary constants.

$$\therefore \ H\left(q_1, \ldots, q_s, \frac{\partial V}{\partial q_1}, \ldots, \frac{\partial L}{\partial q_s}\right) = h \ .$$

Separation of variables. If

$$H = G(f_1(q_1, p_1), f_2(q_2, p_2), \ldots, f_s(q_s, p_s))$$

then

$$G\left(f_1\left(q_1, \frac{\partial V}{\partial q_1}\right), f_2\left(q_2, \frac{\partial V}{\partial q_2}\right), \ldots, f_s\left(p_s, \frac{\partial V}{\partial q_s}\right)\right) = h$$

Let

$$\begin{aligned}
\alpha_i &= f_i\left(q_i, \frac{\partial V}{\partial q_i}\right) \ , \ i \in \overline{1, s} \\
\therefore \ V &= \sum_{i=1}^{s} \int f_i(q_i, \alpha_i) dq_i \\
\therefore \ S &= -G(\alpha_1, \ldots, \alpha_s)t + \sum_{i=1}^{s} \int f_i(q_i, \alpha_i) dq_i
\end{aligned}$$

LECTURE 5

Optimal Control, Unconstrained Control

5.1. The problem

The simplest control problem

$$\begin{aligned} \max \quad & \int_0^T f(t, x(t), u(t))\, dt && (5.1) \\ \text{s.t.} \quad & x'(t) = g(t, x(t), u(t)) && (5.2) \\ & x(0) \text{ fixed}\ , && (5.3) \end{aligned}$$

where $f, g \in C^1[0, T]$, $u \in C_S[0, T]$. This is a generalization of Calculus of Variations. Indeed,

$$\max \int_0^T f(t, x(t), x'(t))\, dt\ ,\ x(0) \text{ given}\ ,$$

can be written as (5.1) with $x' = u$ in (5.2).

5.2. Fixed Endpoint Problems: Necessary Conditions

Let the endpoint $x(T)$ be fixed (variable endpoints are studied in the next section). For any x, u satisfying (5.2),(5.3) and any $\lambda(t) \in C^1[0, T]$,

$$\int_0^T f(t, x, u)\, dt = \int_0^T \left[f(t, x, u) + \lambda(g(t, x, u) - x')\right] dt\ .$$

Integrating by parts,

$$\begin{aligned} -\int_0^T \lambda(t) x'(t)\, dt &= -\lambda(T)x(T) + \lambda(0)x(0) + \int_0^T x(t)\lambda'(t)\, dt \\ \therefore\ \int_0^T f(t, x, u) dt &= \int_0^T \left[f(t, x, u) + \lambda g(t, x, u) + x\lambda'\right] dt - \lambda(T)x(T) + \lambda(0)x(0)\ . \end{aligned}$$

Let

- $u^*(t)$ be an optimal control,
- $u^*(t) + \epsilon h(t)$ a comparison control, with parameter ϵ and h fixed,
- $y(t, \epsilon)$ the resulting state, $y(t, 0) = x^*(t)$ and $y(0, \epsilon) = x(0)$, $y(T, \epsilon) = x(T)$ for all ϵ.

Define

$$\begin{aligned} J(\epsilon) &:= \int_0^T f(t, y(t,\epsilon), u^*(t) + \epsilon h(t))\, dt \\ \therefore\ J(\epsilon) &= \int_0^T [f(t, y(t,\epsilon), u^*(t) + \epsilon h(t)) + \lambda(t) g(t, y(t,\epsilon), u^*(t) + \epsilon h(t)) + y(t,\epsilon)\lambda'(t)]\ dt \\ &\qquad -\lambda(T) y(T,\epsilon) + \lambda(0) y(0,\epsilon) \end{aligned}$$

Since u^* is the maximizing control, $J(\epsilon)$ has a local maximum at $\epsilon = 0$. Therefore $J'(0) = 0$.

$$J'(0) = \int_0^T [(f_x + \lambda g_x + \lambda')\ y_\epsilon + (f_u + \lambda g_u)\ h]\ dt = 0\ , \tag{5.4}$$

where $f_x := \frac{\partial}{\partial x} f(t, x^*(t), u^*(t))$, etc.

(5.4) holds for all y, h iff along (x^*, u^*),

$$\lambda'(t) = -\left[f_x(t, x(t), u(t)) + \lambda(t) g_x(t, x(t), u(t))\right]\ , \tag{5.5}$$

$$f_u(t, x(t), u(t)) + \lambda(t)\, g_u(t, x(t), u(t)) = 0\ . \tag{5.6}$$

Define the Hamiltonian

$$H(t, x(t), u(t), \lambda(t)) := f(t, x, u) + \lambda\, g(t, x, u)\ . \tag{5.7}$$

Then

$$\begin{aligned} x' = \frac{\partial H}{\partial \lambda} &\iff (5.3)\ , \\ \lambda' = -\frac{\partial H}{\partial x} &\iff (5.5)\ , \\ \frac{\partial H}{\partial u} = 0 &\iff (5.6)\ . \end{aligned}$$

5.3. Terminal Conditions

Consider the problem (with $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$)

$$\max \int_0^T f(t, \mathbf{x}, \mathbf{u})\, dt + \phi(T, \mathbf{x}(T)) \tag{5.8}$$

subject to

$$x_i'(t) = g_i(t, \mathbf{x}, \mathbf{u})\ ,\ i \in \overline{1, n} \tag{5.9}$$

$$x_i(0) = \text{fixed}\ ,\ i \in \overline{1, n} \tag{5.10}$$

$$x_i(T) = \text{fixed}\ ,\ i \in \overline{1, q} \tag{5.11}$$

$$x_i(T) = \text{free}\ ,\ i \in \overline{q+1, r} \tag{5.12}$$

$$x_i(T) \geq 0\ ,\ i \in \overline{r+1, s} \tag{5.13}$$

$$K(x_{s+1}, \ldots, x_n, t) \geq 0 \text{ at } t = T \tag{5.14}$$

where $1 \leq q \leq r \leq s \leq n$ and K is continuously differentiable. Write

$$J = \int_0^T \left[f + \boldsymbol{\lambda} \cdot (\mathbf{g} - \mathbf{x}') \right] \, dt + \phi(T, \mathbf{x}(T)) \tag{5.15}$$

and integrate by parts to get

$$J = \int_0^T \left[f + \boldsymbol{\lambda} \cdot \mathbf{g} + \boldsymbol{\lambda}' \cdot \mathbf{x} \right] \, dt + \phi(T, \mathbf{x}(T)) + \boldsymbol{\lambda}(\mathbf{0}) \cdot \mathbf{x}(0) - \boldsymbol{\lambda}(\boldsymbol{T}) \cdot \mathbf{x}(T) \tag{5.16}$$

Let $\mathbf{x}^*, \mathbf{u}^*$ be optimal, and let $\mathbf{x}, \mathbf{u}$ be nearby feasible trajectory satisfying (5.9)–(5.14) on $0 \leq t \leq T + \delta T$. Let $J^*, f^*, \mathbf{g}^*, \phi^*$ denote values along an $(t, \mathbf{x}^*, \mathbf{u}^*)$, and $J, f, \mathbf{g}, \phi$ values along $(t, \mathbf{x}, \mathbf{u})$. Then

$$\begin{aligned} J - J^* &= \int_0^T \left[f - f^* + \boldsymbol{\lambda} \cdot (\mathbf{g} - \mathbf{g}^*) + \boldsymbol{\lambda}' \cdot (\mathbf{x} - \mathbf{x}^*) \right] \, dt + \phi - \phi^* \\ &\quad - \boldsymbol{\lambda}(T) \cdot \left[\mathbf{x}(T) - \mathbf{x}^*(T) \right] + \int_T^{T+\delta T} f(t, \mathbf{x}, \mathbf{u}) \, dt \end{aligned} \tag{5.17}$$

The linear part is

$$\begin{aligned} \delta J &= \int_0^T \left[(f_\mathbf{x} + \boldsymbol{\lambda} \mathbf{g}_\mathbf{x} + \boldsymbol{\lambda}') \cdot \mathbf{h} + (f_\mathbf{u} + \boldsymbol{\lambda} \mathbf{g}_\mathbf{u}) \cdot \delta \mathbf{u} \right] \, dt \\ &\quad + \phi_\mathbf{x} \cdot \delta \mathbf{x}_1 + \phi_T \delta T - \boldsymbol{\lambda}(T) \cdot \mathbf{h}(T) + f(T) \delta T \end{aligned} \tag{5.18}$$

where

$$\begin{aligned} \mathbf{h}(t) &= \mathbf{x}(t) - \mathbf{x}^*(t) \\ \delta u(t) &= \mathbf{u}(t) - \mathbf{u}^*(t) \\ \delta \mathbf{x}_1 &= \mathbf{x}(T + \delta T) - \mathbf{x}^*(T) \end{aligned} \tag{5.19}$$

Approximating $\mathbf{h}(T)$ by

$$\mathbf{h}(T) = \mathbf{x}(T) - \mathbf{x}^*(T) \approx \delta \mathbf{x}_1 - \mathbf{x}^{*\prime}(T) \delta T = \delta \mathbf{x}_1 - \mathbf{g}^*(T) \delta T \tag{5.20}$$

and substituting in (5.18) gives

$$\begin{aligned} \delta J &= \int_0^T \left[(f_\mathbf{x} + \boldsymbol{\lambda} \mathbf{g}_\mathbf{x} + \boldsymbol{\lambda}') \cdot \mathbf{h} + (f_\mathbf{u} + \boldsymbol{\lambda} \mathbf{g}_\mathbf{u}) \cdot \delta \mathbf{u} \right] \, dt \\ &\quad + \left[\phi_\mathbf{x} - \boldsymbol{\lambda}(T) \right] \cdot \delta \mathbf{x}_1 + (f + \boldsymbol{\lambda} \cdot \mathbf{g} + \phi_t) \mid_{t=T} \delta T \leq 0 \end{aligned} \tag{5.21}$$

If the multipliers $\boldsymbol{\lambda}$ satisfy

$$\boldsymbol{\lambda}' = - \left(f_\mathbf{x} + \boldsymbol{\lambda} \mathbf{g}_\mathbf{x} \right) \tag{5.22}$$

along $(\mathbf{x}^*, \mathbf{u}^*)$ then $\delta J \leq 0$ for all feasible $\mathbf{h}, \delta u$ only if (5.6) holds and at $t = T$

$$\left[\phi_\mathbf{x} - \boldsymbol{\lambda}(T) \right] \cdot \delta \mathbf{x}_1 + (f + \boldsymbol{\lambda} \cdot \mathbf{g} + \phi_t) \, \delta T \leq 0 \tag{5.23}$$

for all feasible $\delta\mathbf{x}$, δT. This implies

$$\begin{aligned}
\lambda_i(T) &= \frac{\partial\phi}{\partial x_i}\ ,\ i\in\overline{q+1,r} && (5.24)\\
\lambda_i(T) &\geq \frac{\partial\phi}{\partial x_i}\ ,\ i\in\overline{r+1,s} && (5.25)\\
x_i(T)\left[\lambda_i(T)-\frac{\partial\phi}{\partial x_i}\right] &= 0\ ,\ i\in\overline{r+1,s} && (5.26)\\
\lambda_i(T) &= \frac{\partial\phi}{\partial x_i}+p\frac{\partial K}{\partial x_i}\ ,\ i\in\overline{s+1,n} && (5.27)\\
f+\boldsymbol{\lambda}\cdot\mathbf{g}+\phi_t+p\frac{\partial K}{\partial t} &= 0 \text{ at } T && (5.28)\\
p\geq 0\ ,\ pK &= 0 \text{ at } T\ . && (5.29)
\end{aligned}$$

The last two facts follow from the Farkas Lemma.

If the endpoint $\mathbf{x}(T)$ is free, then the terminal conditions reduce to

$$\boldsymbol{\lambda}(T)=0\ . \tag{5.30}$$

THEOREM 5.1. *If u^* is optimal then $u^*(t)$ maximizes $H(t,x^*(t),u(t),\lambda(t))$, $0\leq t\leq T$, where λ satisfies (5.5) and the appropriate terminal condition.*

EXAMPLE 5.1 (Calculus of Variations). *Consider the problem*

$$\max\left\{\int_0^T f(t,x(t),u(t))\,dt\,:\ x'(t)=u(t)\ ,\ x(0)\ \textit{given}\ ,\ x(T)\ \textit{free}\right\}\ . \tag{5.31}$$

and the Hamiltonian $H(t,x,,,u,\lambda)=f(t,x,u)+\lambda u$.

$$\begin{aligned}
\lambda' &= -\frac{\partial H}{\partial x}=-f_x && (5.32)\\
\lambda(T) &= 0 && (5.33)\\
\frac{\partial H}{\partial u} &= f_u+\lambda=0\ . && (5.34)
\end{aligned}$$

From (5.32) and (5.34) follows the Euler-Lagrange equation

$$f_x=\frac{d}{dt}f_u\ ,$$

and from (5.33) and (5.34) the transversality condition

$$f_u(T)=0\ . \tag{5.35}$$

LECTURE 6

Optimal Control, Constrained Control

6.1. The problem

The simplest such problem is

$$\begin{aligned} \max \quad & \int_0^T f(t, x(t), u(t))\, dt && (6.1) \\ \text{s.t.} \quad & x'(t) = g(t, x(t), u(t))\ ,\ x(0) \text{ given}\ , && (6.2) \\ & a \le u \le b && (6.3) \end{aligned}$$

6.2. Necessary Conditions

Let J be the value of (6.1), J^* the optimal value, and let δJ be the linear part of $J - J^*$.

$$\therefore\ \delta J = \int_o^T \left[(f_x + \lambda g_x + \lambda') + (f_u + \lambda g_u)\, \delta u \right] dt - \lambda(T) \delta x(T)\ . \tag{6.4}$$

If

$$\lambda' = -\left(f_x + \lambda g_x\right)\ ,\ \lambda(T) = 0\ , \tag{6.5}$$

then

$$\delta J = \int_0^T \left(f_u + \lambda g_u\right)\, \delta u\, dt$$

and the optimality condition is

$$\delta J \le 0 \text{ for all feasible } \delta u$$

meaning

$$\begin{aligned} \delta u &\ \ge\ 0 \text{ if } u = a\ , \\ \delta u &\ \le\ 0 \text{ if } u = b\ , \\ \delta u &\ \ \text{unrestricted if } a < u < b\ . \end{aligned}$$

Therefore, for all t,

$$\begin{aligned} u(t) = a &\implies f_u + \lambda g_u \le 0\ , \\ a < u(t) < b &\implies f_u + \lambda g_u = 0\ , \\ u(t) = b &\implies f_u + \lambda g_u \ge 0\ . \end{aligned} \tag{6.6}$$

If (x^*, u^*) is optimal for (6.1)–(6.2) then there is a function λ such that (x^*, u^*, λ) satisfying (6.2),(6.3),(6.5) and (6.6).

$$\begin{aligned} (6.2) &\iff x' = H_\lambda \ , \\ (6.5) &\iff \lambda' = -H_x \ , \end{aligned}$$

for the Hamiltonian

$$H(t, x, u, \lambda) = f(t, x, u) + \lambda\, g(t, x, u) \ . \tag{6.7}$$

Then (6.6) is the solution of the NLP

$$\begin{aligned} \max H &= f + \lambda g \\ \text{s.t.} & \quad a \le u \le b \end{aligned}$$

whose Lagrangian is

$$L := f(t, x, u) + \lambda g(t, x, u) + w_1(b - u) + w_2(u - a)$$

giving the necessary conditions

$$\frac{\partial L}{\partial u} = f_u + \lambda g_u - w_1 + w_2 = 0 \tag{6.8}$$

$$w_1 \ge 0 \ , \ w_1(b - u) = 0 \tag{6.9}$$

$$w_2 \ge 0 \ , \ w_2(u - a) = 0 \tag{6.10}$$

which are equivalent to (6.6). For example, if $u^*(t) = b$ then

$$u^* - a > 0 \ , \ w_2 = 0 \text{ and } f_u + \lambda g_u = w_1 \ge 0 \ .$$

The control u is **singular** if its value does not change H, for example, the middle case in (6.6). A singular control cannot be determined from H.

6.3. Application to Nonrenewable Resources

The model (Hotelling, 1931) is

$$\begin{aligned} \max \quad & \int_0^\infty e^{-rt} u(t) p(t)\, dt \\ \text{s.t.} \quad & \dot{x} = -u \\ & x(0) \text{ given (positive)} \\ & x(t), u(t) \ge 0 \\ \text{and} \quad & 0 \le u(t) \le u_{\max} \text{ (finite)} \ , \ \forall\, t \ . \end{aligned}$$

Let

$$H(t, x, u, \lambda) := e^{-rt} p u - \lambda u \tag{6.11}$$

then

$$\lambda' = -H_x = 0 \implies \lambda = \text{constant} \ .$$

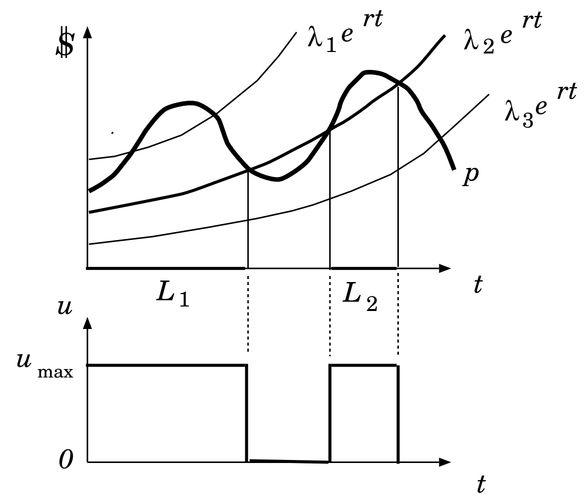

FIGURE 6.1. Maximum extraction in two intervals of lenghts $L_1 + L_2 = L$

Assuming the resource will be exhausted by some time T,

$$\begin{aligned} H(T, x(T), u(T), \lambda(T)) &= \left[e^{-rT} p(T) - \lambda(T)\right] u(T) = 0 \\ \therefore \ \lambda &= e^{-rT} p(T) \\ \therefore \ u(t) &= \begin{cases} 0 & \text{if} \ \ p(t) < \lambda e^{rt} = e^{r(t-T)} p(T) \\ u_{\max} & \text{if} \qquad > \end{cases} \end{aligned}$$

The duration L of the time extraction takes place ($u = u_{\max}$) is

$$L u_{\max} = x(0) \tag{6.12}$$

so λ is adjusted to satisfy (6.12). For example, consider the price function $p(t)$ plotted in Figure 6.1, and three exponential curves $\lambda_i e^{rt}$, $i = 1, 2, 3$. The value λ_1 is too high, allowing for maximal extraction ($p(t) > \lambda_1 e^{rt}$) in a period too short (i.e. shorter than L). Similarly, the value λ_3 is too low. The value λ_2 is just right, where the duration of $p(t) > \lambda_2 e^{rt}$ is $L_1 + L_2 = L$.
Special case: $p(t)$ is exponential.

$$p(t) = p(0) e^{st} \ , \ s = \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} r \implies \text{extract} \left\{ \begin{array}{l} \text{never} \\ \text{anytime} \\ \text{immediately} \end{array} \right.$$

6.4. Corners

The optimality of discontinuous controls raises questions about the continuity of λ, H at points where u is discontinuous. Recall the Weierstrass-Erdmann corner conditions of § 3.9: The functions $F_{x'}$, $F - x' F_{x'}$ are continuous in corners. Returning to optimal control, consider

$$\begin{aligned} \text{opt} \quad & \int_0^T F(t,x,u)\,dt \text{ s.t. } x' = u \\ \text{with} \quad & H = F + \lambda u \\ & H_u = F_u + \lambda = 0 \\ \therefore \quad & \lambda = -F_{x'} \end{aligned}$$

The Weierstrass-Erdmann corner conditions imply

$$\lambda = -F_{x'} \ , \ H = F - x'F_{x'}$$

are continuous in corners of u.

6.5. Control Problems Linear in u

$$\begin{aligned} \max \quad & \int_0^T [F(t,x) + uf(t,x)] \ dt \\ \text{s.t.} \quad & x' = G(t,x) + ug(t,x) \\ & x(0) \text{ given, and } u \text{ satisfies (6.3)} \end{aligned} \tag{6.13}$$

The Hamiltonian is

$$H := (F + \lambda G) + u(f + \lambda g) \tag{6.14}$$

and the necessary conditions include

$$\lambda' = -H_x \tag{6.15}$$

$$u = \begin{Bmatrix} a \\ ? \\ b \end{Bmatrix} \text{ if } f + \lambda g \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 0 \ . \tag{6.16}$$

6.6. A Minimum Time Problem

The position of a moving particle is given by

$$x''(t) = u \ , \ x(0) = x_0 \text{ given} \ , \ x'(0) = 0 \ .$$

The problem is to findn$u$ bounded by

$$-1 \le u(t) \le 1$$

bringing the particle to rest ($x' = 0$) at $x = 0$ in minimal time.
Solution (Pontryagin et al, 1962): Let $x_1 = x, x_2 = x'$. The problem is

$$\begin{aligned} \min \quad & \int_0^T dt \\ \text{s.t.} \quad & x_1' = x_2 \ , \ x_1(0) = x_0 \ , \ x_1(T) = 0 \\ & x_2' = u \ , \ x_2(0) = 0 \ , \ x_2(T) = 0 \\ & u - 1 \le 0 \ , \ -u - 1 \le 0 \end{aligned}$$

with Lagrangian

$$L = 1 + \lambda_1 x_2 + \lambda_2 u + w_1(u-1) - w_2(u+1)$$

and necessary conditions

$$\begin{aligned}
\frac{\partial L}{\partial u} &= \lambda_2 + w_1 - w_2 = 0\ , && (6.17)\\
w_1 &\geq 0\ ,\ w_1(u-1) = 0\ , \\
w_2 &\geq 0\ ,\ w_2(u+1) = 0\ , \\
\lambda_1' &= -\frac{\partial L}{\partial x_1} = 0\ , && (6.18)\\
\lambda_2' &= -\frac{\partial L}{\partial x_2} = -\lambda_1\ , && (6.19)\\
L(T) &= 0\ ,\ \text{since } T \text{ is free}\ . && (6.20)
\end{aligned}$$

Indeed, $L(t) \equiv 0\ ,\ 0 \leq t \leq T$.

$$(6.18) \implies \lambda_1 = \text{constant}\ .\ \therefore\ (6.19) \implies \lambda_2 = -\lambda_1 t + C\ .$$

If $\lambda_2 = 0$ in some interval then $\lambda_1 = 0$ and $L(T) = 1 + 0 + 0$ contradicting (6.20).

$$\therefore\ \lambda_2 \text{ changes sign at most once}\ .$$

- If $\lambda_2 > 0$ then $w_2 > 0$ and $u = -1$.
- If $\lambda_2 < 0$ then $w_1 > 0$ and $u = 1$.

Consider an interval where $u = 1$. Then

$$\begin{aligned}
x_2' &= u = 1\\
\therefore\ x_2 &= t + C_0\\
\therefore\ x_1 &= \frac{t^2}{2} + C_0 t + C - 1\\
&= \frac{(x_2 - C_0)^2}{2} + C_0(x_2 - C_0) + C_1\\
&= \frac{x_2^2}{2} + C_2\ ,\ C_2 = -\frac{C_0^2}{2} + C_1\ .
\end{aligned}$$

Similarly, in an interval where $u = -1$,

$$x_1 = -\frac{x_2^2}{2} + C_3$$

The optimal path must end on one of the parabolas

$$x_1 = \frac{x_2^2}{2} \quad \text{or} \quad x_1 = -\frac{x_2^2}{2}$$

passing through $x_1 = x_2 = 0$. The optimal path begins on a parabola passing through $(x_0, 0)$, and ends up on the switching curve, see Figure 6.3.

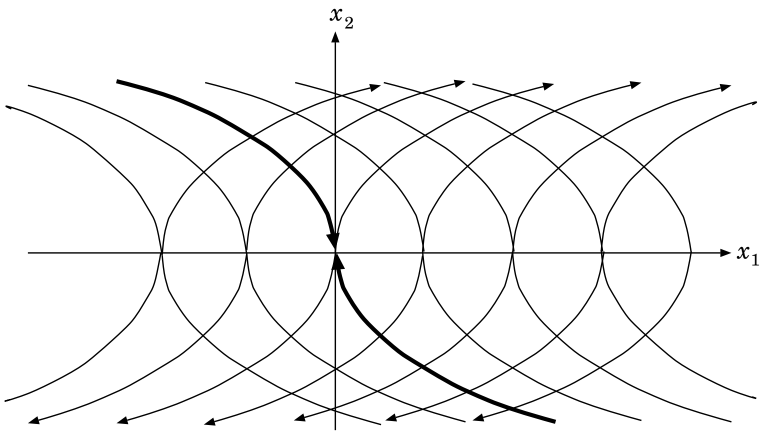

FIGURE 6.2. The parabolas $x_1 = \pm\dfrac{x_2^2}{2} + C$ and the switching curve.

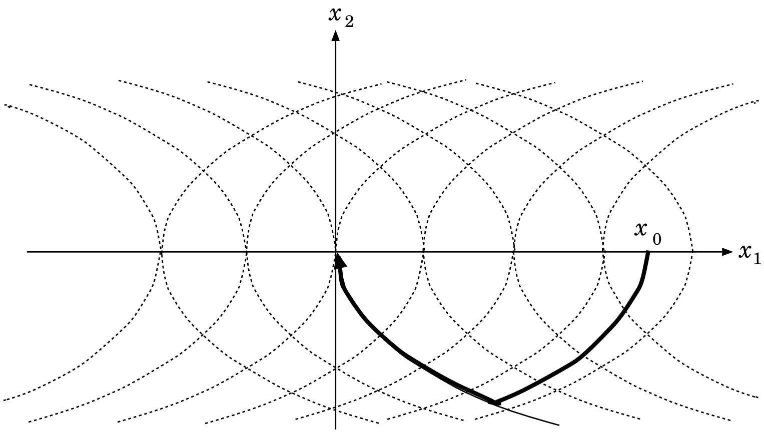

FIGURE 6.3. The optimal trajectory starting at $(x_0, 0)$.

6.7. An Optimal Investment Problem

The model is

$$\begin{aligned} \max_I \quad & \int_0^\infty e^{-rt}\left[p(t)f(K(t)) - c(t)I(t)\right]\,dt \\ \text{s.t.} \quad & K' = I - bK \\ & K(0) = K_0 \text{ given} \\ & I \geq 0 \end{aligned}$$

where: K = capital assets, I = investment, $f(K)$ = output, $p(t)$ = unit price of output, and $c(t)$ = unit price of investment.

The current value Hamiltonian is

$$\mathcal{H} := pf(K) - cI + m(I - bK)$$

Necessary conditions include:

$$\begin{aligned} m(t) &\leq c(t), I(t)\,[c(t) - m(t)] = 0\ , && (6.21) \\ m' &= (r+b)m - pf'(K) && (6.22) \end{aligned}$$

At any interval with $I > 0$, $m = c$ and $m' = c'$. Substituting in (6.22)

$$pf'(K) = (r+b)c - c'\ ,\ \text{while } I > 0\ , \qquad (6.23)$$

a static equation giving the optimal K, with

$$\begin{aligned} \text{LHS}(6.23) &= \text{marginal benefit of investment}\ , \\ \text{RHS}(6.23) &= \text{marginal cost of investment}\ , \end{aligned}$$

Differentiating (6.23)

$$p'f'(K) + pf''(K)K' = (r+b)c' - c''$$

and substituting $K' = I - bK$ we get the singular solution $I > 0$.

Between investment periods we have $I = 0$. To see what this means collect m–terms in (6.22) and integrate

$$e^{-(r+b)t}\,m(t) = \int_t^\infty e^{-(r+b)s} p(s) f'(K(s))\,ds$$

Also

$$\begin{aligned} e^{-(r+b)t}c(t) &= -\int_t^\infty \frac{d}{ds}\left[e^{-(r+b)s}c(s)\right]\,ds \\ &= \int_t^\infty e^{-(r+b)s}\,[c(s)(r+b) - c'(s)]\,ds \end{aligned}$$

The condition $m(t) \leq c(t)$ gives

$$\int_t^\infty e^{-(r+b)s}\,[p(s)f'(K(s)) - (r+b)c(s) + c'(s)]\,ds \leq 0$$

with equality if $I > 0$. Therefore at any interval $[t_1, t_2]$ with $I = 0$ ($I > 0$ for $(t_1)_-$ and $(t_2)_+$)

$$\int_t^{t_2} e^{-(r+b)s}\,[pf' - (r+b)c + c']\,ds \leq 0$$

with equality at $t = t_1$.

6.8. Fisheries Management

The model is

$$\begin{aligned} \max \quad & \int_0^\infty e^{-rt}\,(p(t) - c(x))u(t)\,dt \\ \text{s.t.} \quad & \dot{x} = F(x) - u && (6.24) \\ & x(0) = x_0 \text{ given} \\ & x(t) \geq 0 \\ & u(t) \geq 0 \end{aligned}$$

where: $x(t)$ = stock at time t, $F(x)$ = the growth function, $u(t)$ = the harvest rate, p = unit price, and $c(x)$ = unit cost.

Substituting (6.24) in the objective

$$\max \int_0^\infty e^{-rt}\,(p-c(x))(F(x)-\dot{x})\,dt$$

or $\max \int_0^\infty \phi(t,x,\dot{x})\,dt$ with E-L equation

$$\frac{\partial \phi}{\partial x} = \frac{d}{dt}\frac{\partial \phi}{\partial \dot{x}}$$

where

$$\begin{aligned}
\frac{\partial \phi}{\partial x} &= e^{-rt}\left\{-c'(x)\left[F(x)-\dot{x}\right] + \left[p = c(x)\right]F'(x)\right\} \\
\frac{d}{dt}\frac{\partial \phi}{\partial \dot{x}} &= \frac{d}{dt}\left\{-e^{-rt}\left[p-c(x)\right]\right\} \\
&= e^{-rt}\left\{r\left[p-c(x)\right] + c'(x)\dot{x}\right\} \\
\therefore\ F'(x) - \frac{c'(x)F(x)}{p-c(x)} &= r \qquad (6.25)
\end{aligned}$$

Economic interpretation: Write (6.25) as

$$\begin{aligned}
F'(x)\left[p-c(x)\right] - c'(x)F(x) &= r\left[p-c(x)\right] \\
\text{or } \frac{d}{dx}\left\{\left[p-c(x)\right]F(x)\right\} &= r\left[p-c(x)\right] \qquad (6.26) \\
\text{RHS} &= \text{value, one instant later, of catching fish \# } x+1 \\
\text{LHS} &= \text{value, one instant later, of not catching fish \# } x+1
\end{aligned}$$

LECTURE 7

Stochastic Dynamic Programming

This lecture is based on [**28**].

7.1. Finite Stage Models

A system has finitely many, say N, **states**, denoted $i \in \overline{1,N}$. There are finitely many, say n, **stages**. The state i of the system is observed at the beginning of each stage, and a **decision** $a \in \mathcal{A}$ is made, resulting in a **reward** $R(i,a)$. The state of the system then changes, from i to $j \in \overline{1,N}$, according to the **transition probabilities** $P_{ij}(a)$. Taken together, a **finite–stage sequential decision process** is the quintuple $\{n, N, \mathcal{A}, R, P\}$.

Let $V_k(i)$ denote the **maximum expected return**, with state i and k stages to go. The optimality principle then states

$$V_k(i) = \max_{a\in\mathcal{A}} \left\{ R(i,a) + \sum_{j=1}^{N} P_{ij}(a)\, V_{k-1}(j) \right\} , \quad k = n, n-1, \cdots, 2 , \tag{7.1a}$$

$$V_1(i) = \max_{a\in\mathcal{A}} R(i,a) , \tag{7.1b}$$

a recursive computation of V_k in terms of V_{k-1}, with boundary condition (7.1b). An equivalent boundary condition is $V_0(i) \equiv 0$ for all i, with (7.1a) for $k = n, n-1, \cdots, 1$.

7.1.1. A Gambling Model.

This section is based on [**22**]. A player is allowed to gamble n times, in each he can bet any part of his current fortune. He either wins or loses the amount of the bet with probabilities p or $q = 1 - p$, respectively.

Let $V_k(x)$ be the maximal expected return with present fortune x and k gambles to go. Assuming the gambler bets a fraction α of his fortune,

$$V_k(x) = \max_{0\le\alpha\le 1} \{p\, V_{k-1}(x+\alpha x) + q\, V_{k-1}(x - \alpha x)\} , \tag{7.2a}$$

$$V_0(x) = \log x , \tag{7.2b}$$

logarithmic utility used in (7.2b). If $p \le 1/2$ then it is optimal to bet zero ($\alpha = 0$) and $V_n(x) = \log x$. Suppose $p > 1/2$. Then,

$$\begin{aligned}
V_1(x) &= \max_{\alpha} \{p \log(x + \alpha x) + q \log(x - \alpha x)\} \\
&= \max_{\alpha} \{p \log(1 + \alpha) + q \log(1 - \alpha)\} + \log x \\
\therefore\ \alpha &= q - p && (7.3) \\
\therefore\ V_1(x) &= C + \log x \\
\text{where } C &= \log 2 + p \log p + q \log q && (7.4) \\
\text{Similarly, } V_n(x) &= n\,C + \log x && (7.5)
\end{aligned}$$

and the optimal policy is to bet the fraction (7.3) of the current fortune. See also Exercise 7.1.

7.1.2. A Stock–Option Model. The model considered here is a special case of [**30**]. Let S_k denote the price of a given stock on day k, and suppose

$$S_{k+1} = S_k + X_{k+1} = S_0 + \sum_{i=1}^{k+1} X_i$$

where $X_1, X_2, \cdots$ are i.i.d. RV's with distribution F and finite mean μ_F. You have an option to buy one share at a fixed price c, and N days in which to exercise this option. If the option is exercised (on a day) when the stock price is s, the profit is $s - c$.

Let $V_n(s)$ denote the maximal expected profit if the current stock price is s and n days remain in the life of the option. Then

$$\begin{aligned}
V_n(s) &= \max \left\{ s - c\ ,\ \int V_{n-1}(s + x)\, dF(x) \right\}\ ,\ n \ge 1 && (7.6a) \\
V_0(x) &= \max\{s - c, 0\} && (7.6b)
\end{aligned}$$

Clearly $V_n(s)$ is increasing in n for all s (having more time cannot hurt) and increasing in s for all n (prove by induction).

Lemma 7.1. $V_n(s) - s$ is decreasing in s.

Proof. Use induction on n. The claim is true for $n = 0$: $V_0(s) - s$ is decreasing in s. Assume that $V_{n-1}(s) - s$ is decreasing in s. Then

$$V_n(s) - s = \max \left\{ -c, \int \left[V_{n-1}(s + x) - (s + x)\right]\, dF(x) + \mu_F \right\}$$

is decreasing in s since $V_{n-1}(s + x) - (s + x)$ is decreasing in s for all x. □

Theorem 7.1. There are numbers $s_1 \le s_2 \le \cdots \le s_n \le \cdots$ such that if there are n days to go and the present price is s, the option should be exercised iff $s \ge s_n$

Proof. If the price is s and n days remain, it is optimal to exercise the option if

$$V_n(s) \le s - c\ ,\ \text{by (7.6a)}$$

Let

$$s_n := \min\{s : V_n(s) - s = -c\}\ ,$$

with $s_n := \infty$ if the above set is empty. By Lemma 7.1,

$$V_n(s) - s \;\le\; V_n(s_n) - s_n \;=\; -c \;,\; \forall\, s \ge s_n$$

Therefore it is optimal to exercise the option if $s \ge s_n$. Since $V_n(s)$ is increasing in n, it follows that s_n is increasing. □

Exercise 7.2 shows that it is never optimal to exercise the option if $\mu_F \ge 0$.

7.1.3. Modular Functions and Monotone Policies. This section is based on [**31**]. Let a function $g(x, y)$ be maximized for fixed x, and define $y(x)$ as the maximal value of y where the maximum occurs,

$$\max_y g(x, y) \;=\; g(x, y(x))$$

When is $y(x)$ increasing in x? It can be shown that a sufficient condition is

$$g(x_1, y_1) + g(x_2, y_2) \ge g(x_1, y_2) + g(x_2, y_1) \;,\; \forall\; x_1 > x_2 \;,\; y_1 > y_2 \;. \tag{7.7}$$

A function $g(x, y)$ satisfying (7.7) is called **supermodular**, and is called **submodular** if the inequality is reversed.

LEMMA 7.2. If $\partial^2 g(x, y)/\partial x \partial y$ exists, then g is supermodular iff

$$\frac{\partial^2 g(x, y)}{\partial x \partial y} \;\ge\; 0 \;, \tag{7.8}$$

and submodular iff $\partial^2 g(x, y)/\partial x \partial y \;\le\; 0$.

PROOF. If $\partial^2 g(x, y)/\partial x \partial y \;\ge\; 0$ then for $x_1 > x_2$ and $y_1 > y_2$,

$$\int_{y_2}^{y_1} \int_{x_2}^{x_1} \frac{\partial^2 g(x, y)}{\partial x \partial y}\, dx\, dy \;\;\ge\;\; 0 \;,$$

$$\therefore \;\; \int_{y_2}^{y_1} [g(x_1, y) - g(x_2, y)]\; dy \;\;\ge\;\; 0 \;,$$

and (7.7) follows. Conversely, suppose (7.7) holds. Then for all $x_1 > x_2$ and $y_1 > y$,

$$\frac{g(x_1, y_1) - g(x_1, y)}{y_1 - y} \ge \frac{g(x_2, y_1) - g(x_2, y)}{y_1 - y}$$

and as $y_1 \downarrow y$,

$$\frac{\partial}{\partial y}\, g(x_1, y) \ge \frac{\partial}{\partial y}\, g(x_2, y)$$

implying (7.8). □

EXAMPLE 7.1. An optimal allocation problem with penalty costs. There are I identical jobs to be done consecutively in N stages, at most one job per stage with uncertainty about completing it. If in any stage a budget y is allocated, the job will be completed with probability $P(y)$, an increasing function with $P(0) = 0$. If a job is not completed, the amount allocated is lost. At the end of the Nth stage, if there are i unfinished jobs, a penalty $C(i)$ must be paid.

The problem is to determine the optimal allocation at each stage so as to minimize the total expected cost.

Take the state of the system to be the number i of unfinished jobs, and let $V_n(i)$ be the minimal expected cost in state i with n stages to go. Then

$$\begin{aligned} V_n(i) &= \min_{y \geq 0} \{y + P(y)\, V_{n-1}(i-1) + [1 - P(y)]\; V_{n-1}(i)\} \ , && (7.9a) \\ V_0(i) &= C(i) \ . && (7.9b) \end{aligned}$$

Clearly $V_n(i)$ increases in i and decreases in n. Let $y_n(i)$ denote the minimizing y in (7.9a). One may expect that

$$y_n(i) \text{ increases in } i \text{ and decreases in } n \ . \tag{7.10}$$

It follows that $y_n(i)$ increases in i if

$$\begin{aligned} \frac{\partial^2}{\partial i \partial y} \{y + P(y)\, V_{n-1}(i-1) + [1 - P(y)]\; V_{n-1}(i)\} &\leq 0 \ , \\ \text{or } P'(y)\, \frac{\partial}{\partial i} [V_{n-1}(i-1) - V_{n-1}(i)] &\leq 0 \ , \end{aligned}$$

where i is formally interpreted as continuous, so $\dfrac{\partial}{\partial i}$ makes sense. Since $P'(y) \geq 0$, it follows that

$$y_n(i) \text{ increases in } i \text{ if } V_{n-1}(i-1) - V_{n-1}(i) \text{ decreases in } i \ . \tag{7.11}$$

Similarly,

$$y_n(i) \text{ decreases in } n \text{ if } V_{n-1}(i-1) - V_{n-1}(i) \text{ increases in } n \ . \tag{7.12}$$

It can be shown that if $C(i)$ is convex in i, then (7.11)–(7.12) hold.

7.1.4. Accepting the best offer. This section is based on [**19**]. Given n offers in sequence, which one to accept? Assume:
(a) if any offer is accepted, the process stops,
(b) if an offer is rejected, it is lost forever,
(c) the relative rank of an offer, relative to previous offers, is known, and
(d) information about future offers is unavailable.

The objective is to maximize the probability of accepting the best offer, assuming that all $n!$ arrangements of offers are equally likely.

Let $P(i)$ denote the probability that the ith offer is the best,

$$P(i) \;=\; \text{Prob}\{\text{offer is best of } n |\, \text{offer is best of } i\} \;=\; \frac{1/n}{1/i} \;=\; \frac{i}{n} \ ,$$

and let $H(i)$ denote the best outcome if offer i is rejected, i.e. the maximal probability of accepting the best offer if the first i offers were rejected. Clearly $H(i)$ is decreasing in i.

The maximal probability of accepting the best offer is

$$V(i) \;=\; \max\{P(i), H(i)\} \;=\; \max\left\{\frac{i}{n}, H(i)\right\} \ , \ i = 1, \cdots, n \ .$$

It follows (since i/n is increasing and $H(i)$ is decreasing) that for some j,

$$\begin{aligned} \frac{i}{n} &\leq H(i) \ , \ i \leq j \ , \\ \frac{i}{n} &> H(i) \ , \ i > j \ , \end{aligned}$$

and the optimal policy is: reject the first j offers, then accept the best offer.

Exercises.

EXERCISE 7.1. What happens in the gambling model in § 7.1.1 if the boundary condition (7.2b) is replaced by $V_0(x) = x$?

EXERCISE 7.2. For the stock option problem of § 7.1.2 show: if $\mu_F \geq 0$ then $s_n = \infty$ for $n \geq 1$.

7.2. Discounted Dynamic Programming

Consider the model in the beginning of § 7.1 with two changes:
(a) the number of states is countable, and
(b) the number of stages is countable,
and assume that rewards are bounded, i.e. there exists $B > 0$ such that

$$|R(i,a)| < B \ , \ \forall \ i, a \ . \tag{7.13}$$

A **policy** is a rule $a = f(i,t)$, assigning an action $a :=\in \mathcal{A}$ to a state i at time t. It is: **randomized** if it chooses a with some probability P_a , $a \in \mathcal{A}$, and **stationary** if not randomized, and if the rule is $a = f(i)$, i.e. depends only on the state.

If a stationary policy is used, then the sequence of states $\{X_n : n = 0,1,2,\cdots\}$ is a Markov chain with transition parobabilities $P_{ij} = P_{ij}(f(i))$.

Let $0 < \alpha < 1$ be the **discount factor**. The expected discounted return $V_\pi(i)$ from a policy π and initial state i is the conditional expectation

$$V_\pi(i) \ = \ \mathbf{E}_\pi \left\{ \sum_{n=0}^{\infty} \alpha^n \, R(X_n, a_n) | \, X_0 = i \right\} \tag{7.14}$$

and as consequence of (7.13),

$$V_\pi(i) < \frac{B}{1-\alpha} \ , \ \text{for all policies } \pi \ . \tag{7.15}$$

7.2.1. Optimal policies. The **optimal value function** is defined as

$$V(i) \ := \ \sup_\pi V_\pi(i) \ , \ \forall \ i \ .$$

A policy π^* is **optimal** if

$$V_{\pi^*}(i) = V(i) \ , \ \forall \ i \ .$$

THEOREM 7.2. The optimal value function V satisfies

$$V(i) = \max_a \left\{ R(i,a) + \alpha \sum_j P_{ij}(a) V(j) \right\} \ , \ \forall \ i \tag{7.16}$$

PROOF. Let π be any policy, choosing action a at time 0 with probability P_a, $a \in \mathcal{A}$. Then

$$V_\pi(i) = \sum_{a\in\mathcal{A}} P_a \left[R(i,a) + \sum_j P_{ij}(a)\, W_\pi(j) \right]$$

where $W_\pi(j)$ is the expected discounted return from time 1, under policy π. Since

$$W_\pi(j) \le \alpha V(j)$$

it follows that

$$\begin{aligned}
V_\pi(i) &\le \sum_{a\in\mathcal{A}} P_a \left[R(i,a) + \alpha \sum_j P_{ij}(a)\, V(j) \right\} \\
&\le \sum_{a\in\mathcal{A}} P_a \max_a \left[R(i,a) + \sum_j P_{ij}(a)\, W_\pi(j) \right] \\
&= \max_a \left\{ R(i,a) + \alpha \sum_j P_{ij}(a) V(j) \right\} && (7.17) \\
\therefore\ V(i) &\le \max_a \left\{ R(i,a) + \alpha \sum_j P_{ij}(a) V(j) \right\} && (7.18)
\end{aligned}$$

since π is arbitrary. To prove the reverse inequality, let a policy π begin with a decision a_0 satisfying

$$R(i,a_0) + \alpha \sum_j P_{ij}(a_0) V(j) \ = \ \max_a \left\{ R(i,a) + \alpha \sum_j P_{ij}(a) V(j) \right\} \tag{7.19}$$

and continue in such a way that $V_{\pi_j}(j) \ge V(j) - \epsilon$.

$$\begin{aligned}
\therefore\ V_\pi(i) &= R(i,a_0) + \alpha \sum_j P_{ij}(a_0) V_{\pi_j}(j) \\
&\ge R(i,a_0) + \alpha \sum_j P_{ij}(a_0) V(j) - \alpha\epsilon \\
\therefore\ V(i) &\ge R(i,a_0) + \alpha \sum_j P_{ij}(a_0) V(j) - \alpha\epsilon \\
\therefore\ V(i) &\ge \max_{a\in\mathcal{A}} \left\{ R(i,a) + \alpha \sum_j P_{ij}(a) V(j) \right\} - \alpha\epsilon
\end{aligned}$$

and the proof is completed since ϵ is arbitrary. □

We prove now that a stationary policy satisfying the optimality equation (7.16) is optimal.

THEOREM 7.3. Let f be a stationary policy, such that $a = f(i)$ is an action maximizing RHS(7.16)

$$R(i, f(i)) + \alpha \sum_j P_{ij}(f(i))V(j) = \max_{a \in \mathcal{A}} \left\{ R(i,a) + \alpha \sum_j P_{ij}(a)V(j) \right\} \ , \ \forall \ i \ .$$

Then

$$V_f(i) = V(i) \ , \ \forall \ i \ .$$

PROOF. Using

$$\begin{aligned} V(i) &= \max_{a \in \mathcal{A}} \left\{ R(i,a) + \alpha \sum_j P_{ij}(a)V(j) \right\} \\ &= R(i, f(i)) + \alpha \sum_j P_{ij}(f(i))V(j) \end{aligned} \tag{7.20}$$

we interpret V as the expected return from a 2–stage process with policy f in stage 1, and terminal reward V in stage 2. Repeating the argument

$$\begin{aligned} V(i) &= \mathbf{E}\{n\text{–stage return under } f |\, X_0 = i\} + \alpha^n \mathbf{E}(V(X_n | X_0 = i) \\ &\to V(i) \ , \ \text{by (7.15)} \ . \end{aligned}$$

□

7.2.2. A fixed point argument. We recall the following

THEOREM 7.4. (The Banach Fixed Point Theorem) Let $T : X \to X$ be defined on a complete metric space $\{X, d\}$, and let $0 < \alpha < 1$ satisfy

$$d(T(x), T(y)) \leq \alpha d(x,y) \ , \ \forall \ x, y \in X \ . \tag{7.21}$$

Then T has a unique fixed point x^*, and the successive approximations $x^{n+1} = T(x^n)$ converge uniformly to x^* on any sphere $d(x^0, x^*) \leq \delta$.

PROOF. For (7.21) it follows that

$$\begin{aligned} d(x^{n+1}, x^n) &= d(T(x^n), T(x^{n-1})) \leq \alpha d(x^n, x^{n-1}) \\ \therefore \ d(x^{n+1}, x^n) &\leq \alpha^n d(x^1, x^0) \\ \therefore \ d(x^m, x^n) &\leq \sum_{i=m}^{n-1} d(x^i, x^{i+1}) \ , \ \text{for } m < n \ , \\ &\leq \alpha^m (1 + \alpha + \alpha^2 + \cdots + \alpha^{n-1-m}) d(x^1, x^0) \ , \\ &\leq \frac{\alpha^m}{1-\alpha} d(x^1, x^0) \ , \qquad (7.22) \\ \therefore \ d(x^m, x^n) &\to 0 \text{ as } m, n \to \infty \ . \end{aligned}$$

Since $\{X, d\}$ is complete, the sequence $\{x^n\}$ converges, say to $x^* \in X$. Then $T(x^*) = x^*$ since

$$d(x^{n+1}, T(x^*)) = d(T(x^n), T(x^*)) \leq \alpha d(x^n, x^*) \to 0 \text{ as } n \to \infty$$

If x^{**} is another fixed point of T then

$$d(x^*, x^{**}) = d(T(x^*), T(x^{**}) \leq \alpha d(x^*, x^{**})$$

and $x^* = x^{**}$ since $\alpha < 1$. Since (7.22) is independent of n it follows that

$$d(x^m, x^*) \leq \frac{\alpha^m}{1-\alpha} d(x^1, x^0) \ , \ m = 1, 2, \cdots \tag{7.23}$$

□

A mapping T satisfying (7.21) is called **contraction**, and Theorem 7.4 is also called the **contraction mapping principle**.

Theorem 7.4 applies here as follows: Let $\mathcal{U}$ be the set of bounded functions on the state space, endowed with the **sup norm**

$$\|u\| \ = \ \sup_i \ |u(i)| \tag{7.24}$$

For any stationary policy f let T_f be an operator mapping $\mathcal{U}$ into itself,

$$(T_f u)(i) := R(i, f(i)) + \alpha \sum_j P_{ij}(f(i)) u(j)$$

Then T_f satisfies for all i,

$$|(T_f u)(i) - (T_f v)(i)| < \alpha \sup_j |u(j) - v(j)| \ , \ \forall \ u, v \in \mathcal{U}$$

and therefore,

$$\sup_i \ |(T_f u)(i) - (T_f v)(i)| < \alpha \sup_j |u(j) - v(j)| \ , \ \forall \ u, v \in \mathcal{U} \tag{7.25}$$

i.e. T_f is a contraction in the sup norm (since $0 < \alpha < 1$). It follows that

$$T_f^n u \to V_f \text{ as } n \to \infty$$

It follows that the optimal value function V is the unique bounded solution of (7.16)

7.2.3. Successive approximations of the optimal value. For $V_0(i)$ be any bounded function, $i \in \overline{1, N}$, and define

$$V_n(i) \ := \ \max_{a \in \mathcal{A}} \{R(i, a) + \alpha \sum_j P_{ij}(a) V_{n-1}(j)\} \ , \ n = 1, 2, \cdots \tag{7.26}$$

Then V_n converges to the optimal value V.

PROPOSITION 7.1.
(a) If $V_0 \equiv 0$ then

$$|V(i) - V_n(i)| \ \leq \ \alpha^{n+1} \frac{B}{1-\alpha} \ .$$

(b) For any bounded V_o, $V_n(i) \to V(i)$ uniformly in i.

Note: (a) is the analog of (7.23).

7.2.4. Policy improvement. Let h be a policy that improves on a given stationary policy g in the sense of (7.27) below. The following proposition shows the seinse in which h is closer to an optimal policy than g.

PROPOSITION 7.2. *Let g be a stationary policy with expected value V_g and define a policy h by*

$$R(i,h(i)) + \alpha \sum_j P_{ij}(h(i))V_g(i) \;=\; \max_a \{R(i,a) + \alpha \sum_j P_{ij}(a)V_g(j)\}\ . \tag{7.27}$$

Then

$$\begin{aligned} V_h(i) &\geq V_g(i)\ ,\ \forall\ i\ , \\ \textit{and if } V_h(i) &= V_g(i)\ ,\ \forall(i)\ ,\ \textit{then } V_g = V_h = V\ . \end{aligned}$$

The **policy improvement method** starts with any policy g, and does (7.27) until no improvement is possible. Finiteness follows if the state and action spaces are finite.

7.2.5. Solution by LP.

PROPOSITION 7.3. *If u is a bounded function on the state space such that*

$$u(i) \geq \max_a \{R(i,a) + \alpha \sum_j P_{ij}(a)u(j)\}\ ,\ \forall\ i\ , \tag{7.28}$$

then $u(i) \geq V(i)$ for all i.

PROOF. Let the mapping T on be defined by

$$(Tu)(i) \;:=\; \max_a \{R(i,a) + \alpha \sum_j P_{ij}(a)u(j)\}$$

for bounded functions u. Then (7.28) means $u \geq Tu$, and therefore $u \geq T^n u$ for all u. $\square$

It follows that the optimal value function V is the unique solution of the problem

$$\begin{array}{ll} \min & \sum_i u(i) \\ \text{s.t.} & u(i) \geq \max_a \{R(i,a) + \alpha \sum_j P_{ij}(a)u(j)\}\ ,\ \forall\ i\ , \end{array}$$

or equivalently

$$\begin{array}{ll} \min & \sum_i u(i) \\ \text{s.t.} & u(i) \geq R(i,a) + \alpha \sum_j P_{ij}(a)u(j)\ ,\ \forall\ i\ ,\ \forall\ a \in \mathcal{A}\ . \end{array} \tag{7.29}$$

7.3. Negative DP

Assume countable state space and finite action space. "Negative" means here that the rewards are negative, i.e. the objective is to minimize costs. Let a cost $C(i,a) \geq 0$ result from action a in state i, and for any policy π,

$$\begin{aligned} V_\pi(i) &= E_\pi\left[\sum_{n=0}^{\infty} C(X_n, a_n) :\ X_0 = i\right] \\ V(i) &= \inf_\pi V_\pi(i) \end{aligned}$$

and call a policy π^* optimal if

$$V_{\pi^*}(i) = V(i)\ ,\ \forall\ i\ .$$

The optimality principle here is

$$V(i) = \min_a \{C(i,a) + \sum_j P_{ij}(a)V(j)\}$$

THEOREM 7.5. *Let f be a stationary policy defined by*

$$C(i, f(i)) + \sum_j P_{ij}(f(i))V(j) = \min_a \{C(i,a) + \sum_j P_{ij}(a)V(j)\}\ ,\ \forall\ i\ .$$

Then f is optimal.

PROOF. From the optimality principle it follows that

$$C(i, f(i)) + \sum_j P_{ij}(f(i))V(j) = V(i)\ ,\ \forall\ i\ . \tag{7.30}$$

If $V(j)$ is cost of stopping in state j, (7.30) expresses indifference between stopping and continuing one more stage. Repeating the argument,

$$E_f\left[\sum_{t=0}^{n-1} C(X_t, a_t) :\ X_0 = i\right] + E_f\left[V(X_n) :\ X_0 = i\right] = V(i)$$

and, since costs are nonnegative,

$$E_f\left[\sum_{t=0}^{n-1} C(X_t, a_t) :\ X_0 = i\right] \leq V(i)\ ,$$

and in the limit, $V_f(i) \leq V(i)$, proving optimality. □

Bibliography

[1] G.C. Alway, Crossing the desert, *Math. Gazette* **41**(1957), 209

[2] V.I. Arnold, *Mathematical Methods of Classical Mechanics* (Second Edition, English Translation), Springer, New York, 1989

[3] R.E. Bellman, On the theory of dynamic programming *Proc. Nat. Acad. Sci* **38**(1952), 716–719.

[4] R.E. Bellman, *Dynamic Programming*, Princeton University Press, 1957

[5] R.E. Bellman and S.E. Dreyfus, *Applied Dynamic Programming*, Princeton University Press, 1962

[6] A. Ben-Israel and S.D. Flåm, Input optimization for infinite horizon discounted programs, *J. Optimiz. Th. Appl.* **61**(1989), 347–357

[7] A. Ben-Israel and W. Koepf, Making change with DERIVE: Different approaches, *The International DERIVE Journal* **2**(1995), 72–78

[8] D.P. Bertsekas, *Dynamic Programming and Stochastic Control*, Academic Press, 1976

[9] U. Brauer and W. Brauer, A new approach to the jeep problem, *Bull. Euro. Assoc. for Theoret. Comp. Sci.* **38**(1989), 145–154

[10] J.W. Craggs, *Calculus of Variations*, Problem Solvers # 9, George Allen & Unwin, London, 1973

[11] A.K. Dewdney, Computer recreations, *Scientific American*, June 1987, 128–131

[12] S.E. Dreyfus and A.M. Law, *The Art and Theory of Dynamic Programming*, Academic Press, 1977

[13] L. Elsgolts, *Differential Equations and the Calculus of Variations* (English Translation), MIR, Moscow, 1970

[14] G.M. Ewing, *Calculus of Variations with Applications*, Norton, 1969 (reprinted by Dover, 1985)

[15] N.J. Fine, The jeep problem, *Amer. Math. Monthly* **54**(1947), 24–31

[16] C. Fox, *An Introduction to the Calculus of Variations*, Oxford University Press, 1950 (reprinted by Dover, 1987(

[17] J.N. Franklin, The range of a fleet of aircraft, *J. Soc. Indust. Appl. Math.* **8**(1960), 541–548

[18] D. Gale, The jeep once more or jeeper by dozen, *Amer. Math. Monthly* **77**(1970), 493–501

[19] J. Gilbert and F. Mosteller, *Recognizing the maximum of a sequence*, J. Amer. Statist. Assoc. **61**(1966), 35-73

[20] H. Goldstein, *Classical Mechanics*, Addison-Wesley, Reading, 1950

[21] A. Hausrath, B. Jackson, J. Mitchem, E. Schmeichel, Gale's round-trip jeep problem, *Amer. Math. Monthly* **102**(1995), 299–309

[22] J. Kelly, *A new interpretation of information rate*, Bell System Tech. J. **35**(1956), 917–926

[23] C. Lanczos, *The Variational Principles of Mechanics* (4th Edition), University of Toronto Press, 1970. Reprinted by Dover, 1987.

[24] L.D. Landau and E.M. Lifshitz, *Course of Theoretical Physics, Vol 1: Mechanics* (English Translation), Pergamon Press, Oxford 1960

[25] C.G. Phipps, The jeep problem, *Amer. Math. Monthly* **54**(1947), 458–462

[26] D.A. Pierre, *Optimization Theory with Applications*, Wiley, New York, 1969 (reprinted by Dover, 1986)

[27] W. Ritz, *Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik*, J. Reine Angew. Math. **135**(1908)

[28] S.M. Ross, *Introduction to Stochastic Dynamic Programming*, Academic Press, New York, 1983

[29] M. Sniedovich, *Dynamic Programming*, M. Dekker, 1992

[30] H. Taylor, *Evaluating a call option and optimal timing strategy in the stock market*, Management Sci. **12**(1967), 111-120

[31] D.M. Topkins, *Minimizing a submodular function on a lattice*, Operations Res. **26**(1978), 305-321

[32] R. Weinstock, *Calculus of Variations with Applications to Physics & Engineering*, McGraw Hill, New York, 1952 (reprinted by Dover, 1974)

[33] H.S. Wilf, *Mathematics for the Physical Sciences*, Wiley, New York, 1962 (reprinted by Dover, 1978)