Dynamic Optimization

Why are we even discussing Dynamic Optimization (DO)?

- We've talked about the price path of NRR.
- We've also see that as $t \to \infty$, $p_t \not\to \infty$ because of the availability of backstop.
- However, until such a backstop becomes available, we need to determine the **optimal time path of the price** of the existing NRR.
- This is done by optimising the resource use along the time $t \to T$.
- For this we need **dynamic optimisation**, for static optimisation does not work here.

Difference between DO and SO.

- Static Optimisation

$$\max\ u = u(X, Y)$$

$$\text{s.t. } P_x.X + P_y.Y = I$$

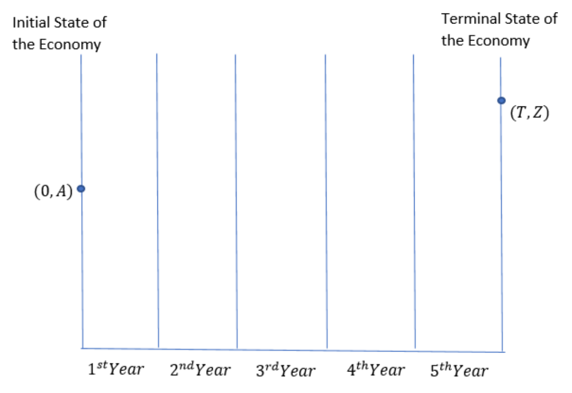
where; X and Y : are goods

I : income

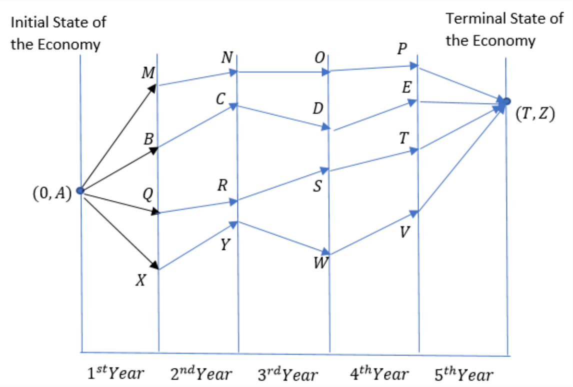
X^* , Y^* : Solution to the optimization problem at a specific time period

- Dynamic Optimisation
    - Objective is to maximise lifetime utility by optimising our consumption of X and Y at each time period.
    - E.g. over our lifetime we can decide on how much to consume, save, invest, etc. to optimise our life.

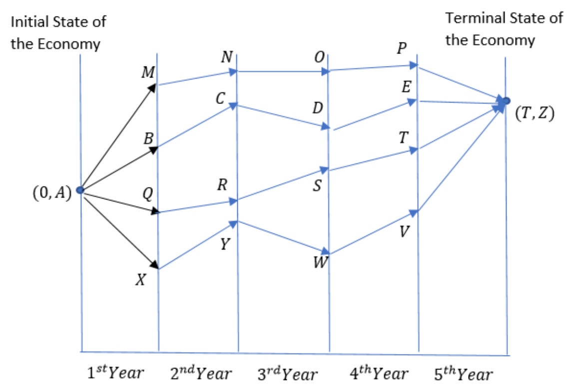
- E.g. the Indian economy is deciding on optimum investment in each period for 5 years so that at the end of the 5^{th} year, the objective of accumulating a certain level of capital stock is realised.
- The cost should be minimum.



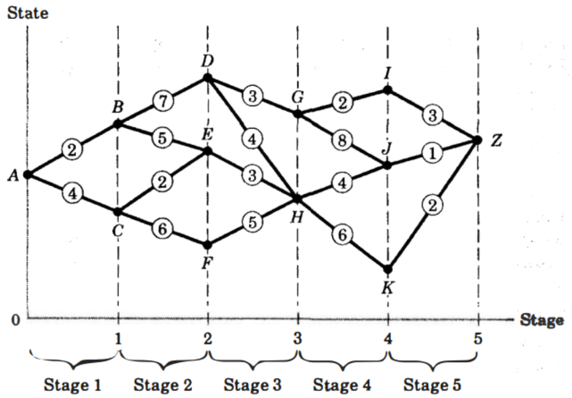
- $(0, A)$: initial state of the economy
- (T, Z) : terminal state of the economy
- 0 : initial time
- T : terminal time
- A : initial capital stock
- Z : terminal capital stock



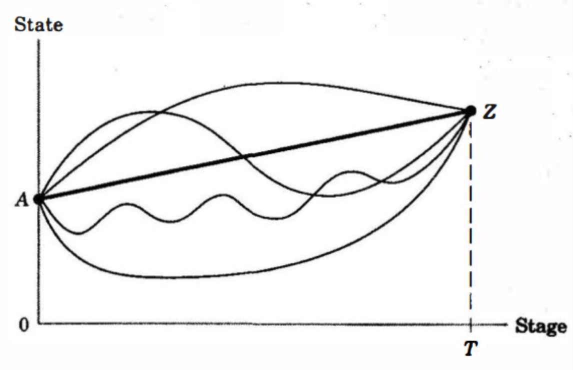
- Suppose the economy starts at $(0, A)$ and wants to reach (T, Z), the decision variable will be **investment in each period**.
- There are different paths to reach the terminal state from the initial state.



- The question is- which path to select?
- The initial thought might be to choose the path $A \to X$.
- However, the total cost along the entire path must be considered while choosing the optimal path.



- Consider a more complex path structure.
- Here costs (in billions of) are represented in circles.
- In this case, the path $ACEHJZ$ gives the optimal solution with 14 billion as the minimum cost.

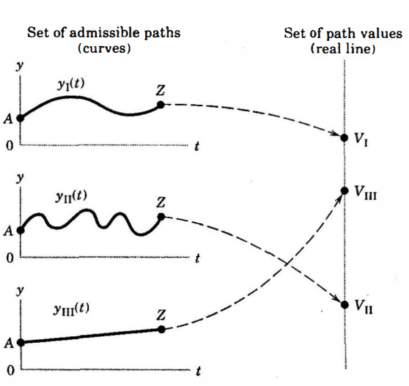


- This is the continuous variable version.
- Each possible path is seen to travel through an infinite number of stages in the interval $[0, T]$.
- E.g. to transport a load of cargo from location A to Z at minimum

Important elements of DO

1. In DO, we have initial state $[0, A]$ and terminal state $[T, Z]$.
2. There are different paths to achieve the terminal state.
3. There should be a **decision variable**. In our example, it's investment.
4. We should have an **objective functional** which we are trying to optimize.

Objective function vs objective functional



- A function maps elements from one set (the domain) to another set (the codomain). For example, $f(x) = x^2$ maps real numbers to real numbers.
- A functional, on the other hand, is a special type of function that takes another function as its input and returns a number (or more generally, a scalar value) as its output. In other words, a functional is a **"function of functions."**

- Examples:-

1. The definite integral is a functional:

- Input: A function $f(x)$
- Output: A single number representing the area under $f(x)$
- Example: $\int_0^1 f(x)dx$ takes any function f and returns its integral from 0 to 1

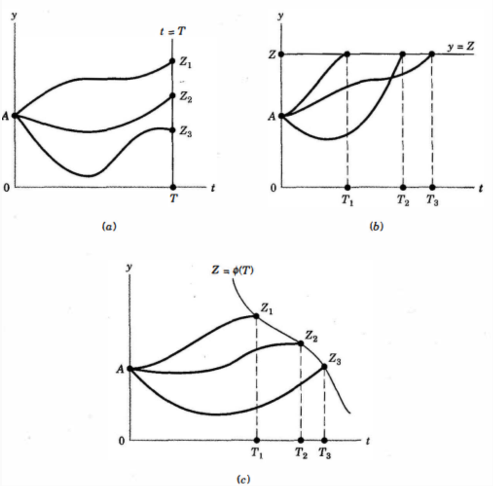
2. The maximum value functional:

- Input: A function $f(x)$ defined on an interval $[a, b]$
- Output: The maximum value of $f(x)$ on that interval
- Example: $max f(x) : x \in [0, 1]$ takes a function and returns its highest value

3. The norm of a function is a functional:

- Input: A function $f(x)$
- Output: A non-negative real number measuring the "size" of the function
- Example: $L_2 norm : ||f|| = \sqrt{}(\int |f(x)|^2 dx)$

- Functionals are particularly important when finding the shortest path between two points on a surface, we're actually minimizing a functional that takes a path (which is a function) as input and returns its length as output.
- A key distinction is that functions operate on points (numbers, vectors, etc.), while functionals operate on entire functions. This makes functionals particularly useful in:
    - Optimization problems where we're looking for optimal functions rather than optimal points



- We might apparently feel that $[0, A]$ and $[T, Z]$ are fixed. But this is not the case.
- Either T or Z or both may be variable in DO.
- There are alternatives regarding the **terminal situation**.

Dynamic Optimization

- Let us assume we have an asset/resource stock from which we want to derive 2 types of benefits:
    1. **Flow benefit**: the value assumed during the use period of the resource.
    2. **Scrap value**: the value derived from a resource after it becomes obsolete. E.g.a car sold after 20 years or more as scrap.
- Our objective is to maximize the **total benefit (flow + scrap value)** from the resource.
- Let us denote

  V : flow benefit

  F : the scrap value

$\therefore$ Our objective is to

$$\text{max}_{[y(t)]} \int_0^T [V(y(t), X(t), t)]dt + F(X(T))$$

$$\text{s.t. } \frac{dX(t)}{dt} = \dot{X}(t) = f(y(t), X(t)) \ \rightarrow \text{equation of motion}$$

or dynamic constraint

$$X(0) = a \ \rightarrow \text{constant}$$

where;

$F(X(T))$: scrap value which is realised at the end of the time i.e. T.

$y(t)$: decision variable or control variable (e.g. rate of extraction)

$X(t)$: state variable i.e. stock of resource at time t.

T : point of time when scrap value is realised.

t : continuous time

To maximise the objective functional we will set a Lagrangian function:

$$
\begin{aligned}
L &= \int_0^T [V(\cdot) + \lambda(t)\{f(\cdot) - \dot{X}(t)\}]dt + F(X(T)) \\
&= \int_0^T [V(\cdot) + \lambda(t) f(\cdot) - \lambda(t)\dot{X}(t)]dt + F(X(T)) \ \ldots (1)
\end{aligned}
$$

Consider;

$$
-\int_0^T \lambda(t)\ \dot{X}(t)dt
$$

The standard integration by parts formula is:

$$\int u \frac{dv}{dt} dt = u \cdot v - \int v \frac{du}{dt} dt \tag{1}$$

Let $u = \lambda(t)$

Let $dv = \dot{X}(t)dt$

$\frac{du}{dt} = \dot{\lambda}(t)$

$v = X(t)$

$$\begin{aligned} -\int_0^T \lambda(t)\ \dot{X}(t)dt &= -[\lambda(t)X(t)]_0^T + \int_0^T X(t)\dot{\lambda}(t)dt \\ &= -\lambda(T)X(T) + \lambda(0)X(0) + \int_0^T X(t)\dot{\lambda}(t)dt \end{aligned}$$

$$\therefore -\int_0^T \lambda(t)\ \dot{X}(t)dt = -[\lambda(t)X(t)]_0^T + \int_0^T X(t)\dot{\lambda}(t)dt$$
$$= -\lambda(T)X(T) + \lambda(0)X(0) +$$
$$\int_0^T X(t)\dot{\lambda}(t)dt \ \ ...(2)$$

Recall;

$$L = \int_0^T [V(\cdot) + \lambda(t)f(\cdot) - \lambda(t)\dot{X}(t)]dt + F(X(T)) \ \ ...(1)$$

Substituting the result from equation (2):

$$L = \int_0^T [V(\cdot) + \lambda(t) f(\cdot)] dt - \lambda(T) X(T) + \lambda(0) X(0) + \int_0^T X(t) \dot{\lambda}(t) dt + F(X(T)) \quad \dots (3)$$

Let us define

$$H = V(\cdot) + \lambda(t) f(\cdot)$$
$$\implies H = H(y(t), X(t), \lambda(t), t)$$

$$\begin{aligned}
\therefore L = & \int_0^T [H(\cdot)] dt - \lambda(T)X(T) + \lambda(0)X(0) + \\
& \int_0^T X(t)\dot{\lambda}(t) dt + F(X(T)) \\
= & \int_0^T [H(\cdot)] dt + \int_0^T X(t)\dot{\lambda}(t) dt + \\
& F(X(T) - \lambda(T)X(T) + \lambda(0)X(0)) \\
= & \int_0^T \left[H(\cdot) + X(t)\dot{\lambda}(t) \right] dt + \\
& F(X(T) - \lambda(T)X(T) + \lambda(0)X(0))
\end{aligned}$$

We want to optimize L by chosing $y(t)$ the control variable.

Let us assume $y(t)$ is changed to $y(t) + \Delta y(t)$

$\{y(t) \to y(t) + \Delta y(t)\}$: **Change in the rate of extraction.**

$\{X(t) \to X(t) + \Delta X(t)\}$: **Change in the stock of resources.**

Step 1: Define L

The given functional is:

$$L = \int_0^T \left[H(\cdot) + X(t)\dot{\lambda}(t) \right] dt + F(X(T) - \lambda(T)X(T) + \lambda(0)X(0))$$

where:

- $H(\cdot)$ is the **Hamiltonian**.
- $X(t)$ is the **state variable**.
- $\lambda(t)$ is the **co-state (Lagrange multiplier)**.
- $F(\cdot)$ is a function depending on the terminal state $X(T)$.

Why is it called the Hamiltonian?

The **Hamiltonian** is named after **Sir William Rowan Hamilton**, who developed **Hamiltonian mechanics** in the 19th century. Originally, it was used in **classical mechanics** to describe the total energy of a system:

In Physics

$$H = T(q, p, t) + V(q, t)$$

$$\implies \text{Total Energy} = \text{Kinetic Energy} + \text{Potential Energy}$$

And in Economics

$$H = V(y(t), X(t), t) + \lambda(t) f(y(t), X(t))$$

$$\implies \text{Total Benefits or Costs} = \text{Flow Benefit or Costs} + \text{Stock Benefit or Costs}$$

Step 2: Compute the Variation ΔL

The total variation of L comes from two parts:

1. **Variation of the Integral Term**:

$$\int_0^T \left[H(\cdot) + X(t)\dot{\lambda}(t) \right] dt$$

2. **Variation of the Terminal Function $F(\cdot)$**:

$$F(X(T) - \lambda(T)X(T) + \lambda(0)X(0))$$

Step 2.1: Variation of the Integral Term

Since L is an integral, its variation follows:

$$\Delta L = \int_0^T \Delta \left[H(\cdot) + X(t)\dot{\lambda}(t) \right] dt + \Delta F(\cdot)$$

Expanding $\Delta H(\cdot)$ using the **first-order Taylor expansion**:

$$\Delta H(\cdot) = \frac{\partial H}{\partial y(t)} \Delta y(t) + \frac{\partial H}{\partial X(t)} \Delta X(t)$$

Also, the variation of $X(t)\dot{\lambda}(t)$ gives:

$$\Delta (X(t)\dot{\lambda}(t)) = \dot{\lambda}(t) \Delta X(t)$$

$$\implies \int_0^T \left[\frac{\partial H}{\partial y(t)}\Delta y(t) + \frac{\partial H}{\partial X(t)}\Delta X(t) + \dot{\lambda}(t)\Delta X(t)\right] dt$$

Step 2.2: Variation of the Terminal Function $F(\cdot)$

The function $F(X(T) - \lambda(T)X(T) + \lambda(0)X(0))$ depends on $X(T)$, $\lambda(T)$, and $X(0)$. Its total variation is:

$$\Delta F = \frac{\partial F}{\partial X(T)}\Delta X(T) - \lambda(T)\Delta X(T) + \lambda(0)\frac{\partial X(0)}{\partial X(T)}\Delta X(T)$$

Since $X(0)$ is **constant**, its derivative with respect to $X(T)$ is:

$$\frac{\partial X(0)}{\partial X(T)} = 0$$

which simplifies the terminal term to:

$$\frac{\partial F}{\partial X(T)}\Delta X(T) - \lambda(T)\Delta X(T)$$

Step 3: Final Expression for ΔL

Now, combining everything:

$$\Delta L = \int_0^T \left[\frac{\partial H}{\partial y(t)} \Delta y(t) + \frac{\partial H}{\partial X(t)} \Delta X(t) + \dot{\lambda}(t) \Delta X(t) \right] dt \\ + \frac{\partial F}{\partial X(T)} \Delta X(T) - \lambda(T) \Delta X(T)$$

For optimization $\Delta L = 0$, we analyze:

$$\Delta L = \int_0^T \left[\frac{\partial H}{\partial y(t)} \Delta y(t) + \frac{\partial H}{\partial X(t)} \Delta X(t) + \dot{\lambda}(t) \Delta X(t) \right] dt \\ + \frac{\partial F}{\partial X(T)} \Delta X(T) - \lambda(T) \Delta X(T)$$

Rearranging;

$$\Delta L = \int_0^T \left[\frac{\partial H}{\partial y(t)} \Delta y(t) + \left(\frac{\partial H}{\partial X(t)} + \dot{\lambda}(t) \right) \Delta X(t) \right] dt \\ + \left(\frac{\partial F}{\partial X(T)} - \lambda(T) \right) \Delta X(T)$$

Since $\Delta L = 0$ must hold for any small variations $\Delta X(t)$ and $\Delta y(t)$, each coefficient must be zero.

Principle 1

Since $\Delta y(t)$ is arbitrary, we must have:

$$\frac{\partial H}{\partial y(t)} = 0$$

This is the **control optimality condition**, ensuring that the Hamiltonian is optimized with respect to the control $y(t)$.

Principle 2

From the integral term:

$$\left[\frac{\partial H}{\partial X(t)} + \dot{\lambda}(t)\right] \Delta X(t)$$

For arbitrary $\Delta X(t)$, we get:

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial X(t)}$$

This is the **co-state equation**, governing the evolution of the costate $\lambda(t)$.

Principle 3

From the terminal variation term:

$$\left[\frac{\partial F}{\partial X(T)} - \lambda(T)\right] \Delta X(T)$$

For arbitrary $\Delta X(T)$, we get:

$$\lambda(T) = \frac{\partial F}{\partial X(T)}$$

This is setting the final value of the costate.

Summary of Maximum Principle Conditions

To satisfy $\Delta L = 0$:

1. **Control Optimality Condition**

$$\frac{\partial H}{\partial y(t)} = 0$$

2. **Co-state Equation**

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial X(t)}$$

3. **Terminal Condition**

$$\lambda(T) = \frac{\partial F}{\partial X(T)}$$

These are the necessary conditions from **Pontryagin's Maximum Principle**.

Pontryagin's Maximum Principle: Interpretation

- Pontryagin's Maximum Principle is named after the Russian mathematician Lev Pontryagin, who formulated this principle in 1956 along with his students.
- The principle was initially developed to solve optimization problems in control theory, specifically for maximizing the terminal speed of a rocket.

1. Control Optimality Condition

$$\frac{\partial H}{\partial y(t)} = 0$$

Interpretation:

- Ensures that the **Hamiltonian is optimized** with respect to the control variable $y(t)$.
- Determines the **optimal control strategy**.
- The best action $y^*(t)$ must satisfy this equation.

Example (Renewable Resource Extraction)

- Managing a **fishery**:
  - $y(t) =$ harvesting rate
  - $X(t) =$ fish population
  - $\frac{\partial H}{\partial y(t)} = 0$ ensures **profit maximization while maintaining sustainability**.

2. Co-State (Shadow Price) Equation

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial X(t)}$$

Interpretation:

- Describes how the **shadow price $\lambda(t)$ evolves over time**.
- $\lambda(t)$ represents **the value of an extra unit of $X(t)$**.
- Shows how resource depletion **affects future value**.

Example (Groundwater Extraction)

- $X(t) =$ amount of water in aquifer
- $y(t) =$ extraction rate
- $\lambda(t) =$ future value of preserving water
- If overuse today **reduces future availability**, then $\lambda(t)$ changes accordingly.

3. Transversality (Terminal) Condition

$$\lambda(T) = \frac{\partial F}{\partial X(T)}$$

Interpretation:

- Determines the **final value of shadow price** $\lambda(T)$.
- If there's a **terminal reward or penalty**, it sets the final condition.
- If no terminal condition exists, often $\lambda(T) = 0$.

Example (Deforestation & Land Use)

- $X(T) =$ remaining forest at time T
- $\frac{\partial F}{\partial X(T)} =$ future value of forest
- Ensures that **future benefits of conservation** are included in today's decisions.

Recall,

$$H = V(\cdot) + \lambda(t) f(\cdot)$$

where;

$V(\cdot)$: flow benefit

$\lambda(t) f(\cdot)$: future benefit or value

$\lambda(t)$: shadow price

$\therefore H =$ Present Benefit $+$ Future Benefit

Future benefit is realised at one point in time.

However, present benefits are a stream of benefits.

So if we have to convert $V(\cdot)$ at present time then we have to multiply it with e^{-rt}, i.e. discounting.

$\because H = V(\cdot) + \lambda(t)f(\cdot)$

$H_p = V(\cdot)e^{-rt} + \lambda(t)f(\cdot)$

where;

H_p: Present Value Hamiltonian because the **flow benefit** is converted into stock benefit.

Now we want to convert the sotck (one period) benefit into a stream of benefits:

$H_c = V(\cdot) + \lambda(t)e^{rt}f(\cdot)$

where;

H_c: Current Value Hamiltonian because the **stock (one period) benefit** is converted to flow benefits.

Let's assume

$$
\begin{aligned}
\lambda(t)e^{rt} &= \rho(t) \\
\implies \lambda(t) &= \rho(t)e^{-rt} \\
\implies \dot{\lambda}(t) &= \dot{\rho}(t)e^{-rt} - r\rho(t)e^{-rt} \\
\implies \dot{\lambda}(t)e^{rt} &= \dot{\rho}(t) - r\rho(t) \\
\implies \dot{\rho}(t) &= \dot{\lambda}(t)e^{rt} + r\rho(t)
\end{aligned}
$$

where,

$\rho(t)$: current value of the co-state variable $\lambda(t)$

Intuition:

The idea behind H_p and H_c is that we are representing two types of benefits that we can actually get from a NRR.

These benefits could be flow or stock; and in the Hamiltonian we have converted either of the benefits (flow or stock) to arrive at a single benefit (present value) or stream of benefits (current value).

Maximizing total (flow + stock) benefits requires

$$
\begin{aligned}
\frac{\partial H_p}{\partial y(t)} &= 0 \\
\implies \frac{\partial [V(\cdot)e^{-rt} + \lambda(t) f(\cdot)]}{\partial y(t)} &= 0 \\
\implies \frac{\partial V(\cdot)}{\partial y(t)} e^{rt} + \lambda(t) \frac{\partial f(\cdot)}{\partial y(t)} &= 0
\end{aligned}
$$

We also know that one of the conditions of the maximum principle is;

$$
\begin{aligned}
\dot{\lambda}(t) &= -\frac{\partial H_p(\cdot)}{\partial X(t)} \\
\implies \dot{\lambda}(t)e^{rt} &= -\frac{\partial H_p(\cdot)}{\partial X(t)}e^{rt} \\
\implies \dot{\rho}(t) - r\rho(t) &= -\frac{\partial H_c(\cdot)}{\partial X(t)} [\because \dot{\rho}(t) = \dot{\lambda}(t)e^{rt} + rp(t)] \\
\implies \dot{\rho}(t) &= r\rho(t) - \frac{\partial H_c(\cdot)}{\partial X(t)}
\end{aligned}
$$