# Advanced Mathematical Economics

Paulo B. Brito
PhD in Economics: 2019-2020
ISEG
Universidade de Lisboa
pbrito@iseg.ulisboa.pt

Lecture 1
13.2.2020

# Chapter 1

# Introduction

This is a course on continuous dynamics in economics. By continuous dynamics we mean modelling a function (or jointly determined functions) of one or more independent variables belonging to a **continuous domain** which is endowed with an **order relationship**. In most cases the independent variable represents time, but it can also represent space, age, skills, dimension or the states of nature. The fact that there is an order relationship means that the independent variable is more than an indexing device for modelling heterogeneity. Its value involves a relationship such as present-future, close-far away, high-low skill, small-large, for example.

The applications in economics cover a wide range of subjects: dynamic macroeconomics, growth theory, population economics, spatial economics, finance and dynamic microeconomics.

Most of the economic models in the continuous are related to macroeconomics and growth theory and feature time as the independent variable. In this case we say that we have continuous-time dynamics relating short-run with long-run adjustments. In mathematical finance the major analytical contributions are modelled jointly in continuous-time and in a continuous probability space, allowing to study the random dynamics. Macroeconomic dynamic general equilibrium models with heterogeneous agents and dynamic game theory take both time and another continuous variable as independent variables allowing to study the dynamics of distributions, i.e, the time-varying behavior of distributions of capital and income.

Important theoretical developments in geographical economics or contract theory, for example, consider a continuum of space or of types of agents within a static, i.e., time-independent domain. However, increasingly these fields are joining macroeconomics with heterogeneous agents in featuring the dynamics jointly in a continuous domain for space, types, income, capital and continuous time.

Although the particularities of the main variables may be different, for the different topics, the models, whose solutions we may be interested in characterizing, share some common formal structures. The purpose of this course is to provide an introduction to those common formal structures. The fact that the space of independent variables is endowed with an order allows for modelling the variables of interest by functions whose values at a particular point depend on interactions with neighboring points along a particular gradient of the independent variable. In particular, we will consider models including ordinary differential equations (ODE) , partial differential equations (hyperbolic and parabolic) (PDE) and stochastic differential equations (SDE), and the related optimal control of ODEs, PDEs and SDEs.

Continuous-time dynamic models have been at the core of growth theory and mathematical finance and until the late eighties in macroeconomics[1]. Since the 1990's the dynamic general equilibrium (DGE) and dynamic stochastic general equilibrium models (DSGE) became the dominant paradigm in macroeconomics. Most of DGE and DSGE have been modelled in a discrete time space (see Ljungqvist and Sargent (2012) and Miao (2014)). This allows for a quantitative calibration, simulation and estimation. However, the cost is introducing too much detail associated with the timing of the decisions and renders the qualitative analysis of the models more difficult.

In other areas in economic theory modelling in the continuum was still dominant: in spatial economics (see Fujita and Thisse (2002)), in finance (see Cvitanić and Zapatero (2004) and Stokey (2009)).

Recently, the research in macroeconomics turned to DGE and DSGE models with heterogeneous agents. This is starting to signify the comeback of continuous-time modelling (see Brunnermeier and Sannikov (2016) and Gabaix et al. (2016) and the references therein.)

There are two advantages of using a continuous-time framework [2]

1. obtaining qualitative dynamics results, in particular asymptotic dynamics, is not only easier but can also be done drawing on a large body of results from other disciplines (applied mathematics, physics and mathematical biology);

2. extending the models by including dynamics of distributions and stochastic dynamics and the possibility of obtaining qualitative dynamic results is also made possible for the same reasons.

[1] The state of the art at that time can be seen in v.g. Burmeister and Dobell (1970), Turnovsky (1977), or in Gandolfo (1997)

[2] Discussions continuous versus discrete time modelling Isohätälä et al. (2016) Brunnermeier and Sannikov (2016)

## 1.1 Magnitudes

Functions mapping between number and numbers

Functionals mapping between a function and a number

Operators mapping between functions and functions

Algebrais equation

Functional equation

Operator equation

Next we present the generic structure of the models we will study.

## 1.2 Generic structure of the models

Consider an (unknown) function $\mathbf{y}(\mathbf{z})$, defined over an independent $\mathbf{z}$, $\mathbf{y}(\mathbf{z})$, mapping

$$\mathbf{y} : \mathcal{Z} \to \mathcal{Y},\ \mathcal{Y} \in \mathbb{R}^n$$

where $\mathcal{Z}$ may have different types of topology (for instance, it can be an Euclidean space or a probability space) but satisfies $\dim(\mathcal{Z}) = m$. Therefore, $\mathbf{z} = (z_1, \dots, z_m)$ and

$$\mathbf{y} = \mathbf{y}(\mathbf{z}) = \begin{pmatrix} y_1(\mathbf{z}) \\ \dots \\ y_n(\mathbf{z}) \end{pmatrix} = \begin{pmatrix} y_1(z_1, \dots, z_m) \\ \dots \\ y_n(z_1, \dots, z_m) \end{pmatrix}.$$

A **differential equation** is an equation of the form

$$\mathbf{F}(D^p_{\mathbf{z}}(\mathbf{y}), \dots, D_{\mathbf{z}}(\mathbf{y}), \mathbf{y}(\mathbf{z}), \mathbf{z}) = 0 \tag{1.1}$$

where the $\mathbf{F}(.)$ is known, where $D_{\mathbf{z}}(\mathbf{y})$ is an appropriately defined gradient

$$D_{\mathbf{z}}(\mathbf{y}) \equiv \left( D_{z_1}(\mathbf{y}), \dots, D_{z_m}(\mathbf{y}) \right)$$

where

$$D_{z_i}(\mathbf{y}) \equiv \begin{pmatrix} \frac{\partial y_1}{\partial z_i} \\ \dots \\ \frac{\partial y_n}{\partial z_i} \end{pmatrix},\ \text{for } i \in \{1, \dots, m\}$$

and $D^p_{\mathbf{z}}(\mathbf{y})$ is the multi-dimensional matrix of higher-order derivatives.

In most economic applications the equations also involve a vector of parameters $\ \in \Phi \subseteq \mathbb{R}^q$ then equation (1.1) becomes

$$\mathbf{F}(D^p_{\mathbf{z}}(\mathbf{y}), \dots, D_{\mathbf{z}}(\mathbf{y}), \mathbf{y}(\mathbf{z},\ ), \mathbf{z},\ ) = 0 \tag{1.2}$$

**Solving** a differential equation means finding at least one function $\ (\mathbf{z})$, mapping $\phi : \mathcal{Z} \to \mathcal{Y} \subseteq \mathbb{R}^n$, that satisfies equation (1.1). That is, if it satisfies

$$\mathbf{F}(D^p_{\mathbf{z}}(\ ), \dots, D_{\mathbf{z}}(\ ),\ (\mathbf{z}), \mathbf{z}) = 0 = 0.$$

we call it **general solution** of equation (1.1). Function $\phi(\mathbf{z})$, if it exists, may be unique or multiple.

**Additional information**: in general we have additional information on the state of the system.

Different types of differential equations are obtained for different sets of independent variables, $\mathcal{Z}$ and $m$, for different $n$, $p$ and for different properties of function $\mathbf{F}(.)$.

Next, we see some of examples that we will be dealt in the course

## 1.3 Ordinary differential equations

**Ordinary differential equation** (ODE) are differential equations in which the independent variable is of dimension one and belongs to the set of real numbers. That is $m = 1$, $\mathbf{z} = z$. and $\mathcal{Z} \equiv (z_0, z_1) \subseteq \mathbb{R}$. Therefore, (1.1) becomes

$$\mathbf{F}(D^p_z(\mathbf{y}), \dots, D_z(\mathbf{y}), \mathbf{y}(z), z) = 0.$$

The value of $p$ gives the order of the equation. If $p = 1$ the equation is called **first order equation**, if $p = 2$ it is called **second order equation**, and so forth. However, all equations with $p \geq 2$ can be transformed into first order equations by defining the derivatives of $y(.)$ as new variables.

This means the an ODE can have the general representation

$$\mathbf{F}(D(\mathbf{y}), \mathbf{y}(z), z) = 0 \tag{1.3}$$

where, again, $\mathbf{F}$ is known, and

$$D(\mathbf{y}) = \left( \frac{dy_1(z)}{dz}, \dots, \frac{dy_n(z)}{dz} \right)$$

is the gradient as regards the independent variable $z$. Equation (1.1) is called **implicit ODE**.

Let the gradient of $\mathbf{F}$ as regards $D(\mathbf{y})$ be

$$\mathbf{A}\ (\mathbf{y}, z) = D_{D(\mathbf{y})}\mathbf{F}.$$

If $\mathbf{F}_{D(\mathbf{y})}(.)$ is monotonic and regular we can transform, at least locally, equation (1.3) into

$$\mathbf{A}\ (\mathbf{y}, z)D(\mathbf{y}) + \mathbf{G}(\mathbf{y}, z) = 0 \tag{1.4}$$

where $\mathbf{A}(.)$ is a $n \times n$ matrix and $\mathbf{G}(.)$ is a $n \times 1$ vector. This equation is called a **quasi-linear equation**, because it is linear in the derivatives. If $\mathbf{A}$ is a matrix of constants, equation (1.4) can be written in the **explicit or normal form**

$$D(\mathbf{y}) = \mathbf{H}(\mathbf{y}, z). \tag{1.5}$$

Intuitively, we can see equation (1.5) as describing a movement within the Euclidean space, while equation (1.3) describes a movement in a generalized surface (which can be regular or not).

Now, we can be more specific about the additional information we referred to in the previous section. Let, again, $z \in [z_0, z_1]$, where $z_0 < z_1$. If we set $\mathbf{y}(z_0) = \mathbf{y}_0$ we call

$$\mathbf{F}(D(\mathbf{y}), \mathbf{y}(z), z) = \mathbf{0},\ \mathbf{y}(z_0) = \mathbf{y}_0$$

**initial-value problem**. The solution of this problem, $\mathbf{y} = \quad (z, \mathbf{y}_0)$, is called **particular solution** and it is a function of the independent variable and the data on the system $\mathbf{y}_0$. Observe that the dimension of $\mathbf{y}_0$ is the same as the number of equations in $\mathbf{F}(.) = \mathbf{0}$. If we keep the same relationship between the number equations, (in the ODE system) but some variables are fixed at the $z_1$ we say we have a **boundary-value** problem.

Next we present some low-dimensional ODE models and related concepts we will deal further in the course.

### 1.3.1 Scalar ODE

In this subsection we set $n = 1$ and consider again the one-dimensional independent variable $z \in \mathcal{Z} \subseteq \mathbb{R}_+$. When convenient we set $\mathcal{Z} = [z_0, z_1]$ (or $[z_0, z_1)$, $(z_0, z_1]$ or $(z_0, z_1)$).

To obtain an intuition, let $y(s+h)$ be the value of $y$ at for $z = s + h$ and let $y(s)$ be the value of $y$ for $z = s$. Let the variation of $y$, be

$$y(s + h) - y(s) = f(y(s), s)(s + h - s) = f(y(s), s)h.$$

Then

$$\frac{y(s+h)-y(s)}{h} = f(y(s))$$

Because the derivative of $y(z)$, taken at the point $z = s$, is

$$\frac{dy(s)}{ds} = \lim_{h\to 0} \frac{y(s+h)-y(s)}{h},$$

if a derivative exists for all $z \in \mathcal{Z}$ we see that a **scalar ODE** in the normal form,

$$y^{'}(z) \equiv \frac{dy(z)}{dz} = f(y(z), z),$$

represents the behavior of a variable, $y$, if its infinitesimal variation is a function of its level and of the value of the independent variable (this is the reason why we need an order structure imposed on $\mathcal{Z}$).

From this point on we will assume that the **independent variable is time**. That is we let $n = 1$ and $z = t \in \mathcal{T} \subseteq \mathbb{R}_+$ where $\mathcal{T}$ is the time interval, usually $\mathcal{T} = [0, T]$ with $T$ finite or $\mathcal{T} = [0, \infty)$. In this case the function $y : \mathcal{T} \to \mathcal{Y} \subseteq \mathcal{R}$ and $D(y) = y^{'}(t)$ is denoted by $\dot{y}$, which is [3]

$$\dot{y} \equiv \frac{dy}{dt} = \lim_{\epsilon\to 0} \frac{y(t+\epsilon)-y(t)}{\epsilon}.$$

The implicit ODE takes the form

$$F(\dot{y}, y, t) = 0.$$

**Solving** an ODE means finding (or proving the existence, multiplicity and characterising) a function $\phi(t)$ such that

$$F(\dot{\phi}(t), \phi(t), t) = 0$$

this means that (non-rigorously) the continuity and differentiability properties of the solution are inherited from the (known) function $f(.)$.

**Properties of** $F(.)$ As we saw the solution of the differential equation (roughly) inherits the properties of $f(.)$, as regards continuity and differentiability. We can consider the following cases:

- $F_{\dot{y}}(\dot{y}, y)$ is non-differentiable as regards $\dot{y}$. Consider the quasi-linear equation

$$a(y)\dot{y} = g(y)$$

[3] We will use the dot notation, $\dot{y}$, for time derivatives and the $y^{'}(z)$ notation for non-time independent variables.

where there are point $y_s$ such that $a(y_s) = 0$ then $f(.)$ is **not regular** in equation

$$\dot{y} = f(y) \equiv a(y)^{-1} g(y)$$

which means that it is not locally Lipschitz (i.e. $\lim_{y \to y_s} \dot{y} = \pm\infty$). In this case we have **constrained** ODEs or **singular** ODEs.

- $F_{\dot{y}}(\dot{y}, y)$ is differentiable as regards $\dot{y}$ but $f(.)$ is locally non-differentiable or non-continuous. We can write the ODE in normal form, but such that

$$\dot{y} = \begin{cases} f_1(y) & \text{if } h(y) \leq 0 \\ f_2(y) & \text{if } h(y) < 0 \end{cases}$$

  where, if $y = \tilde{y}$ such that $h(\tilde{y}) = 0$. We can have two cases

  - **non-differentiable**: $f_1(\tilde{y}) = f_2(\tilde{y})$ and $D(f_1(\tilde{y})) \neq D(f_2(\tilde{y}))$
  - **non-continuous**: $f_1(\tilde{y}) \neq f_2(\tilde{y})$ and $D(f_1(\tilde{y})) \neq D(f_2(\tilde{y}))$

- both $F_{\dot{y}}(\dot{y}, y)$ is differentiable as regards $\dot{y}$ and $f(.)$ are **continuous, differentiable and regular**. In this case the ODE it takes the form

$$\dot{y} = f(y)$$

The quasi-linear equation can be written as

$$a(y,t)\dot{y} + g(y,t) = 0$$

If $a(.)$ is everywhere different from zero, we say have the **non-autonomous** ODE in the normal form [4] and write

$$\dot{y} = \frac{dy}{dt} = f(y,t).$$

If time does not enter explicitly as an argument of $f(.)$, we say we have an **autonomous** ODE (in the normal form)

$$\dot{y} = \frac{dy}{dt} = f(y). \tag{1.6}$$

In most applications we consider a set of parameters $\varphi \in \Phi$

$$\dot{y} = f(y, \varphi).$$

[4] The convention of not writing the time-dependence of the dependent variable $y$ is common in the literature.

**ODEs and problems involving ODEs**

Let $y = y(t)$ for $t \in [0, T]$

**Initial value problem**: defined by an ODE and an initial value for the unknown function

$$\dot{y} = f(y) \text{ and } y(0) = y_0 \text{ known}$$

**Boundary value problem**: defined by an ODE and a terminal value for the unknown function

$$\dot{y} = f(y) \text{ and } y(T) = y_T \text{ known.}$$

**Integral representation** Another intuition can be drawn from the relationship with an **integral equations**, displaying backward or forward deendencies.

Let $0 < t < T$ and let $y(t)$, display **backward** dependency as

$$y(t) = \int_0^t f(y(s))ds$$

that is, is the value at $t$ of a function is an integral (a generalized sum) of a function of its past values. It we take a time derivative, and apply the Leibnitz rule, we have equation (1.6).

In the case in which the value of variable at time $t$, $y(t)$, display **forward** dependency as

$$y(t) = -\int_t^T f(y(s))ds,$$

we get equation (1.6) time-differentiating.

Therefore we can, and this classification is used in the stochastic differential equations (SDE) literature call **forward ODE** when we have a ODE jointly with an initial value and call **backward ODE** when we have a ODE jointly with a terminal value.

**Examples**

of scalar ODE's

**Example 1**: the exponential growth model for population growth

$$\dot{N} = \mu N$$

where $N(t)$ is population at time $t$ and $\mu$ is a parameter representing the instantaneous rate of growth of the population. This ODE is usually interpreted as a forward equation, i.e, we have information on $N(0)$ and want to find the behavior os population $N(t)$ for $t > 0$.

**Example 2**: a generic budget constraint

$$\dot{W} = Y(t) - D(t) + r(t)W$$

where $W$ is the stock of financial assets, $Y$ and $D$ are non-financial income and expenditures, and $r$, is the instantaneous rate of return. This equation can be interpreted as a forward or a backward equation. In the first case we know $W(0)$ and want to determine the future behavior of $W(t)$ and in the second case we fix a, usually bounded, value for $W(T)$ (or for its present value) and want to determine the initial value $W(0)$ which is consistent with it.

**Example 3**: the Solow growth model

$$\dot{k} = s(k) - \delta k$$

where $k$ is the per capital capital stock, $s(k)$ is the savings function, and $\delta$ is the rate of depreciation of capital. This ODE is also usually interpreted as a forward ODE.

### 1.3.2 Planar and higher-dimensional ODE

If $n = 2$, and keeping $z = t$, we have the **planar** ODE, where $\mathbf{y}(t) = (y_1(t), y_2(t)) \in \mathcal{Y} \subseteq \mathbb{R}^2$. The ODE in implicit for is

$$F_1(\dot{y}_1, \dot{y}_2, y_1, y_2) = 0$$
$$F_2(\dot{y}_1, \dot{y}_2, y_1, y_2) = 0$$

where $F_i(.)$, for $i = 1, 2$ can take non continuous, or non-differentiable regular or singular forms as for the planar equation. If $F(.)$ is well-behaved we have the ODE in its normal form

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, y_2) \\ \dot{y}_2 &= f_2(y_1, y_2). \end{aligned}$$

In this case we can also have the initial value (or the forward ODE) and the terminal-value problems (or the backward ODE) cases as in the scalar case. However, we have a new case:

**Mixed boundary-initial value problem**: defined by an ODE and a number of initial and terminal value conditions which is equal to the dimension of $y$ ($n$). Example: let $y = (y_1, y_2)$

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, y_2),\ y_1(0) = y_0 \\ \dot{y}_2 &= f_2(y_1, y_2),\ y_2(T) = y_T \end{aligned}$$

The optimality condition for optimal control problems take the form (for $T \to \infty$)

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, y_2),\ y_1(0) = y_0 \\ \dot{y}_2 &= f_2(y_1, y_2),\ \lim_{t\to\infty} h(y_1(t), y_2(t), t) = 0 \end{aligned}$$

We can say that in this case we have a **forward-backward ODE**

**Examples**

**Example 4**: the Ramsey model featuring the couple dynamics of consumption, $c$ and capital, $k$, and is

$$\begin{aligned} \dot{c} &= c(r(k) - \rho) \\ \dot{k} &= y(k) - c \end{aligned}$$

the first equation has been called several names, such as Euler equation, Keynes-Ramsey rule, consumer arbitrage condition, and the second is a budget constraint.

**Example 5**: the Ramsey model with endogenous labor

$$\begin{aligned} \dot{c} &= c(r(k, l) - \rho) \\ \dot{k} &= y(k, l) - c \\ c &= C(k, l) \end{aligned}$$

where $l$ is the labor effort can be transformed into a planar equation.

Those are both cases of forward-backward ODEs.

### 1.3.3 Higher-dimensional ODE's

If $n > 2$ the ODE in normal form is

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, \ldots, y_n) \\ &\ldots \\ \dot{y}_n &= f_n(y_1, \ldots, y_n) \end{aligned}$$

we have a **multidimensional** ODE.

**Example 6**: the Ramsey model with government debt dynamics can have the form

$$\begin{aligned}\dot{c} &= c((1-\tau)r(k,b)-\rho)\\ \dot{k} &= y(k)-c\\ \dot{b} &= g-\tau y(k)+r(k,b)b\end{aligned}$$

where $b$ is the level of government debt, $\tau$ is the tax rate, and $g$ public expenditures. Again this is a forward-backward ODEs, where the dimension of the forward (backward) component depends on the condition on $b$.

### 1.3.4 Optimal control of ODE's

Consider two unknown functions $\mathbf{u} : \mathcal{T} \to \mathbb{R}^u$ and $\mathbf{x} : \mathcal{T} \to \mathbb{R}^n$, and define the **value functional**

$$V(\mathbf{u},\mathbf{x}) = \int_{t_0}^{t_1} h(t,\mathbf{u}(t),\mathbf{x}(t))dt \tag{1.7}$$

the ODE

$$\dot{\mathbf{x}} = f(\mathbf{u},\mathbf{x}). \tag{1.8}$$

and addition conditions on $t=0$, $t=T$ and the associated values $x(t_0)=x_0$ or $x(t_1)=x_{t_1}$, or restrictions upon them. In the simplest problem, we assume we know $(t_0,\mathbf{x}(t_0))=(0,\mathbf{x}_0)$ and $(t_1,\mathbf{x}(t_1))=(T,\mathbf{x}_T)$.

The **optimal control problem** (OCP) is to find the functions $\mathbf{u}^*(.)$ and $\mathbf{x}^*(.)$ that

$$\max_{\mathbf{u}(.)} V(\mathbf{u},\mathbf{x})$$

subject to equation (1.8), given $\mathbf{x}(0)=\mathbf{x}_0$ and other information on $T$ or $\mathbf{x}(T)$.

The most common problem in economics has the value function

$$V(\mathbf{u},\mathbf{x}) = \int_0^{\infty} h(\mathbf{u}(t),\mathbf{x}(t))e^{-\rho t}dt$$

and is called **infinite horizon discounted optimal control problem**.

There are several methods for solving the OCP all leading to a ODE.

From now on let us assume that the control and the state variables are scalar.

**Calculus of variations problem**

If we can write $\dot{x} = f(u, x)$ as $u = g(\dot{x}, x)$ then the problem becomes

$$\max_{x(.)} \int_0^T F(\dot{x}, x, t)dt$$

The optimality condition is the Euler-Lagrange equation

$$\frac{\partial F(\dot{x}, x, t)}{\partial x} + \frac{d}{dt}\left( \frac{\partial F(\dot{x}, x, t)}{\partial \dot{x}} \right) = 0$$

together with initial and terminal conditions (it is a mixed-value problem).

The EL equation is a second order ODE, evaluated at an optimum,

$$F_x(\dot{x}, x, t) + F_{\dot{x}t}(\dot{x}, x, t) + F_{\dot{x}x}(\dot{x}, x, t)\dot{x} + F_{\dot{x}\dot{x}}(\dot{x}, x, t)\ddot{x} = 0.$$

**Pontriyagin's maximum principle**

Introducing the co-state variable $q(t)$ and the Hamiltonian function

$$H(x, u, t) = h(x, u, t) + q(t)f(x, u, t)$$

the necessary conditions for an optimum involve the modified Hamiltonian dynamic system (MHDS ), which is a system of two first order equations,

$$\begin{aligned} \dot{q} &= -H_x(x, u(q, x), t) \\ \dot{x} &= f(x, u(q, u), t) \end{aligned}$$

if the functions $f(.)$ and $h(.)$ are sufficiently differentiable, together with initial and terminal conditions. Examples 4 and 5 are MHDS

**Dynamic programming**

Optimality conditions are given by an implicit ODE's the **Hamilton-Jacobi-Bellman** (HJB) equation: value function $V(x, t)$ satisfies at the optimum the HJB equation

$$-V_t(x, t) \ = \max_{u(.)} \{ \ h(x, u, t) + V_x(x, t) \ f(x, u, t)\}$$

This is a partial differential equation (PDE) in implicit form.

For infinite-horizon autonomous problems the HJB equation becomes an ODE (in the implicit form)

$$\rho V(x) = \max_{u(.)} \left\{ h(x,u) + V^{'}(x)\ f(x,u) \right\}$$

where $V(x)$ is unknown. The policy function takes the form $u = u(x, V^{'}(x))$ then the HJB function becomes the implicit ODE

$$\rho V(x) = \ h\left(x, u(x, V^{'}(x))\right) + V^{'}(x)\ f\left(x, u(x, V^{'}(x))\right).$$

**Extensions**

The former definition of optimal control problem can be extended in several different ways, for instance:

- by increasing the number of control variables (see example 5)
- by increasing the dimension of the state vector $x = (x_1, \dots, x_n)$ (which doubles the dimension of the ODE representing the first order conditions)
- by introducing a instantaneous value terminal state or control $V(x(t), u(T), T)$
- by introducing constraints on the terminal state or control $H(x(T), u(T), T) \leq 0$
- by introducing constraints on the trajectories of the state or control variables $H(x(t), u(t), t) \leq 0$ .

### 1.3.5 Solving ODEs and problems involving ODEs

Ideally we would like to find function $\phi(t, .)$ explicitly. However this is only possible for a small number of equations.

The three main issues regarding solving differential equations that cannot be solved explicitly are associated with:

- existence of solutions
- uniqueness of solutions

- characterization of solutions, i.e., their behavior relative to the independent variable, $t$, and the other data of the problem (parameters, initial or terminal values). The main tool for this is the **qualitative theory of ODE** or **bifurcation theory**.

## 1.4 Partial differential equations

A **partial differential equation** (PDE) is one equation involving one or more than one function of at least two independent variables together with its derivatives. In economics applications one of the independent variable is time. In equation (1.1) we have $m > 1$, $\mathcal{Z} \subseteq \mathbb{R}^m$, $p = 2$ and $\mathbf{F}(.)$ be monotonic and regular in all the derivative arguments.

### 1.4.1 First-order PDE

Let $m = 2$ and consider the function $\mathbf{y} = \mathbf{y}(t, x)$ where $(t, x) \in \mathbb{R}^2$.

A **first-order partial differential equation** takes the general form

$$\mathbf{F}(D_t(\mathbf{y}), D_x(\mathbf{y}), \mathbf{y}(t, x), t, x) = 0$$

If $n = 1$ a **quasi-linear** hyperbolic PDE is

$$a(t, x)y_t + b(t, x)y_x = f(y),$$

where $y : \mathbb{R}^2 \to \mathcal{Y} \subseteq \mathbb{R}$, $y_t = \frac{\partial y(t,x)}{\partial t}$ and $y_x = \frac{\partial y(t,x)}{\partial x}$. Some times we represent the partial derivatives by $\partial_t y(t, x)$ and $\partial_x y(t, x)$.

This type of equations models for instance transport, conservative, age-dependent distributions along time. It represents a distribution moving along time.

**Example 6** Dynamics of an age-dependent population. The density of population $n = n(a, t)$ is equal to the number of individuals of age $a$ at time $t$ is governed by the McKendrick PDE

$$n_t + n_a = \mu(a)n$$

where $\mu(.)$ is the mortality rate. In general the equation is complemented with a condition for the newborns $n(0, t) = \int_0^{a_{\max}} \beta(a)n(a, t)da$ when fertility is age-dependent.

### 1.4.2 Parabolic partial differential equations

A parabolic PDE has the general form

$$\mathbf{F}(D_x^2(\mathbf{y}), D_t(\mathbf{y}), D_x(\mathbf{y}), \mathbf{y}(t, x), t, x) = 0$$

that is, it involves a second derivative as regards the ”spatial” variable.

If $n = 1$ a **quasi-linear** parabolic PDE is

$$a(t,x)y_t + b(t,x)y_x + c(t,x)y_{xx} = f(y),$$

where $y : \mathbb{R}^2 \to \mathcal{Y} \subseteq \mathbb{R}$ and $y_{xx} = \frac{\partial^2 y}{\partial x^2}$.

In Economics and Finance there are two types of PDE's we should distinguish:

- **forward** parabolic PDE
$$y_t(x,t) - y_{xx}(x,t)) = f(y(x,t))$$
models forward diffusion phenomena starting from one known initial distribution $y(0,x) = \phi(x)$ and diffusing out through time;

- **backward** parabolic PDE
$$y_t(x,t) + y_{xx}(x,t)) = f(y(x,t))$$
is very common on mathematical finance, in which a terminal distribution $y(T,x) = h(x)$ is known and an initial distribution $y(0,x) = \phi(x)$ is to be determined.

**Example 7** The heath equation is the simplest case of a forward equation

$$u_t(x,t) = u_{xx}(x,t)$$

where $u(x,t)$ is the temperature of an one-dimensional rod, at location $x$ at time $t$.

**Example 8** The well known Black-Scholes is an example of a backward equation

$$v_t(S,t) = -\frac{\sigma S^2}{2} v_{SS}(S,t) + (v(S,t) - S v_S(S,t))$$

where $v(S,t)$ is the value of an option over an underlying asset with price $S$ at time $t$, $\sigma$ is the instantaneous volatility of the underlying asset and $r$ is the risk-free interest rate.

### 1.4.3 Optimal control of PDE's

Although much less well known, and usually very hard to solve, there are optimal control problems for systems governed by PDE's. The first order conditions are generally forward-backward PDE's.

### Optimal control of first-order PDE's

Consider two unknown functions $u : \mathcal{D} \rightarrow \mathbb{R}^u$ and $y : \mathcal{D} \rightarrow \mathbb{R}^n$ where $\mathcal{D} = (\underline{x}, \bar{x}) \times (0, \infty)$ An **optimal control problem of hyperbolic PDE** has the form

$$\max_{u(.)} \int_{\underline{x}}^{\bar{x}} \int_0^{\infty} h(u(x,t), y(x,t), x, t) dt dx$$

subject to the first-order PDE

$$y_t(x,t) \ + y_x(x,t) = f(u(x,t), y(x,t))$$

plus initial $y(0,x) = y_0(x)$ and possibly boundary conditions. This models the optimal choice of a distribution along time.

The first order conditions, from the **Pontryagin's maximum principle**

Consider the co-state variable $q(t,x)$ and the Hamiltonian function

$$H(h(u(t,x), y(t,x), x, t) = h(u(t,x), y(t,x), x, t) + q(t,x) f(u(t,x), y(t,x))$$

necessary f.o.c. include involve a system of two first-order PDE (one moving forward and the other backward)

$$\begin{aligned} q_t &= - q_x + H_y(u(q,y), y, .) \\ y_t &= y_x \ + f(u(q,y), y) \end{aligned}$$

the solution, if it exists, features an optimal distribution evolving along time, possibly converging to a bounded asymptotic distribution.

### Parabolic PDE's

Consider two unknown functions $u : \mathcal{Z} \rightarrow \mathbb{R}^u$ and $x : \mathcal{Z} \rightarrow \mathbb{R}^n$ where $\mathcal{Z} = (\underline{x}, \bar{x}) \times (0, \infty)$ An **optimal control problem of parabolic PDE** has the form

$$\max_{u(.)} \int_{\underline{x}}^{\bar{x}} \int_0^{\infty} h(u(x,t), y(x,t), x, t) dt dx$$

subject to the *forward* parabolic PDE

$$y_t(x,t) = y_{xx}(x,t) \ + f(u(x,t), y(x,t))$$

plus initial, $y(0,x) = y_0(x)$, and possibly boundary conditions

The first order conditions, from the **Pontryagin's maximum principle** Consider the co-state variable $q(t, x)$ and the Hamiltonian function

$$H(h(u(t,x), y(t,x), x, t) = h(u(t,x), y(t,x), x, t) + q(t,x) f(u(t,x), y(t,x))$$

necessary f.o.c. include involve a system of two parabolic PDE (one forward and one backward)

$$\begin{aligned} q_t &= -q_{xx} + H_y(u(q,y), y, .) \\ y_t &= y_{xx} + f(u(q,y), y) \end{aligned}$$

the solution, if it exists, features an optimal distribution evolving along time, possibly converging to a bounded asymptotic distribution. Reference: Li and Yong (1995).

## 1.5 Stochastic differential equations

### 1.5.1 The diffusion equation

Let: $m = 2$, $\mathcal{Z} \subseteq \mathbb{R} \times \Omega$ ($\Omega$ is a probability space) , $p = 1$ and $F(.)$ monotonic and regular in all the derivative arguments

Then $y = y(t, \omega)$ where $\omega \in (\Omega, \mathcal{F}, \mathbb{P})$ is a probability space

A **Itô's stochastic differential equation** (SDE) takes the form

$$dy = f(y)dt + \sigma(y)dW$$

where $dW$ is a standard Wiener process, i.e., a stochastic process following a normal distribution with zero mean and variance $dt$: $dW \sim N(0, dt)$. An important fact about $y(t, \omega)$ is that it is not differentiable (in the classic sense) as regards $t$. Therefore, in order to solve a SDE we need to apply the **Itô's or stochastic calculus**. Essentially this

There are relationships between SDE's and PDE's. Ex: the probability distribution as of time $t = 0$ that $y(t, \omega) = x$, $p(x, t)$, satisfies the Kolmogorov forward or Fokker-Planck equation

$$p_t\ (x,t) = \frac{1}{2}\ \partial_{xx}[\sigma(x)p(x,t)]\ - \partial_t[f(y(x,t)p(x,t)],$$

which is a parabolic PDE.

### 1.5.2 Optimal control of SDE's

Consider two unknown functions $u : \mathcal{Z} \to \mathbb{R}^u$ and $x : \mathcal{Z} \to \mathbb{R}^n$ where $\mathcal{Z} = \Omega \times (0, \infty)$ where $\Omega$ is again a probability space

An **optimal control problem of SDE** can take the form

$$\max_{u(.)} \mathbb{E}\left[ \int_0^\infty h(u(t), y(t), t)dt \right]$$

subject to the SDE

$$dy = f(u, y)dt + \sigma(u, y)dB$$

plus initial and boundary conditions.

Note that

$$\mathbb{E}\left[ \int_0^\infty h(u(t), y(t), t)dt \right] = \int_\Omega \int_0^\infty h(u(t, \omega), y(t, \omega), t)\pi(\omega)dtd\omega$$

where $\pi(.)$ is a density function

To find the optimum it is convenient to use the **stochastic dynamic programming principle**

The value function $V(y, t)$ satisfies at the optimum the HJB equation

$$-V_t(y, t) \; = \max_{u(.)} \left\{ \; h(y, u, t) + V_y(y, t) \; f(y, u) + \frac{1}{2}\sigma(u, y)^2 V_{yy}(y, t) \; \right\}$$

which is an implicit parabolic PDE.

Much less known is a version SDE version to the maximum principle. Reference: Peng (1990) and Yong and Zhou (1999).

# Bibliography


Brunnermeier, M. K. and Sannikov, Y. (2016). Macro, Money and Finance: A Continuous Time Approach. NBER Working Papers 22343, National Bureau of Economic Research, Inc.

Burmeister, E. and Dobell, A. R. (1970). *Mathematical Theories of Economic Growth*. Macmillan.

Cvitanić, J. and Zapatero, F. (2004). *Introduction to the Economics and Mathematics of Financial Markets*. MIT Press.

Fujita, M. and Thisse, J.-F. (2002). *Economics of Agglomeration*. Cambridge.

Gabaix, X., Lasry, J., Lions, P., and Moll, B. (2016). The Dynamics of Inequality. *Econometrica*, 84:2071–2111.

Gandolfo, G. (1997). *Economic Dynamics*. Springer-Verlag.

Isohätälä, J., Klimenko, N., and Milne, A. (2016). Post-crisis macrofinancial modeling: Continuous time approaches. In Haven, E., Molyneux, P., Wilson, J. O. S., Fedotov, S., and Duygun, M., editors, *The Handbook of Post Crisis Financial Modeling*, pages 235–282. Palgrave Macmillan UK.

Li, X. and Yong, J. (1995). *Optimal Control Theory for Infinite Dimensional Systems*. Systems & Control: Foundations & Applications. Birkhäuser Basel, 1 edition.

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive Macroeconomic Theory*. MIT Press, Cambridge and London, 3rd edition.

Miao, J. (2014). *Economic Dynamics in Discrete Time*. MIT Press.

Peng, S. (1990). A general stochastic maximum principle for optimal control problems. *SIAM Journal on Control and Optimization*, 28.

Stokey, N. L. (2009). *The Economics of Inaction*. Princeton.

Turnovsky, S. (1977). *Macroeconomic Analysis and Stabilization Policy*. Cambridge University Press.

Yong, J. and Zhou, X. Y. (1999). *Stochastic Controls. Hamiltonian Systems and HJB Equations*. Number 43 in Applcations of Mathematics. Stochastic Modelling and Applied Probability. Springer.