ELECTENG 332 - Control Systems

Lecture Notes

Nicholas Russell

University of Auckland

Department of Electrical, Computer, and Software Engineering

Table of contents

Introduction

This is my "coursebook", or rather, a collection of notes for the course ELECTENG 332: Control Systems. The notes are written in Quarto, a markdown-based document processor that supports LaTeX and code blocks, and requires this introduction file.

Please skip this bit, there's nothing to see here.

Chapter 1

Module 1: Basics of Signals and Systems

Learning Outcomes

▶ Uniqueness of the Exponential Signal

▶ Concept of Engineering Infinity

▶ Concept of Complex Frequency

▶ Classification of Signals: Energy & Power

▶ Classification of Systems

▶ What is a Control System

▶ Classification of a Control System: Open-loop & Closed-loop

1.1   Topic 1: The Importance of the Exponential Function

The Exponential function, written as either e^{ax} or e^{at} depending on whether it is $f(t)$ or $f(x)$, has properties that make it mathematically unique.

1. The derivative (rate of change) of the exponential function is the exponential function itself. More generally, this is a function whose rate of change is proportional to the function itself.

$$\frac{\mathrm{d}e^{ax}}{\mathrm{d}x} = ae^{ax} \tag{1}$$

2. The integral of the exponential function is also the exponential function itself.

$$\int e^{ax}\,\mathrm{d}x = \frac{1}{a}e^{ax} \tag{2}$$

1.2   Topic 2: The Concept of Engineering Infinity

Consider a signal e^{-at}. The time constant for this signal is $T = \frac{1}{a}$. Theoretically, the signal is meant to decay to zero as time approaches infinity, i.e.

$$\lim_{t \to \infty} e^{-at} = 0 \tag{3}$$

But in practice, this is not the case, as its value will be very, very small after five time constants $5T$ (or 5τ). This is the *Concept of Engineering Infinity*. The signal will never reach zero, but it will be so small that it can be considered zero

for all practical purposes. This is a very important concept in control systems, as it allows us to simplify our calculations and analysis.

1.3 Topic 3: The Concept of Complex Frequency

Complex frequency is found commonly in electrical engineering. It is often notated as $j\omega$ or $s = \sigma \pm j\omega$. These frequencies always come in pairs, so the use of $\pm$ is implicit to this, as complex numbers have complex conjugates (normally notated by z^* or $\bar{z}$). i.e. $s = \sigma + j\omega$ has the conjugate $s = \sigma - j\omega$.
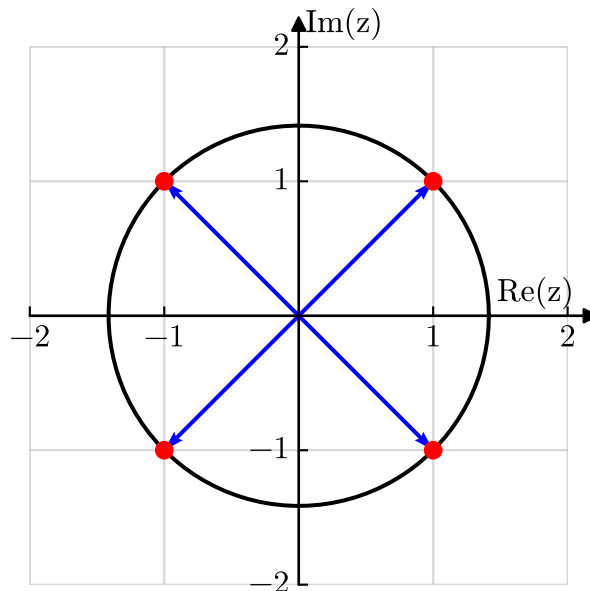

Figure 1.1: Argand diagram of $|z| = \sqrt{2}$

This is also backed up by De Moivre's formula which is defined mathematically as:

$$\forall x \in \mathbb{R}, \quad \forall n \in \mathbb{Z}, \tag{4}$$
$$e^{jnx} = \cos(nx) + j\sin(nx) \tag{5}$$

Or more generally for our applications (this is also known as Euler's formula):

$$e^{jx} = \cos(x) + j\sin(x) \tag{6}$$
$$\text{Where } x \in \mathbb{R} \ (x \text{ is real}) \tag{7}$$
$$\text{and } j \equiv i = \sqrt{-1} \tag{8}$$

This means that:

> *A complex frequency $j\omega$ represents a pure sinusoidal signal of frequency ω rad/s*

For example, if a signal has a complex frequency $j314$ rad/s, then this responds to a pure sinusoid of frequency 314 rad/s (i.e. 50 Hz).

Furthermore:

> *A complex frequency $s = \sigma + j\omega$ represents an exponentially damped signal of frequency $j\omega$ rad/s, and decays/amplifies at a rate decided by σ.*

For example, the signal $f(t) = e^{-10t}\sin(40\pi t)$ would look like this:
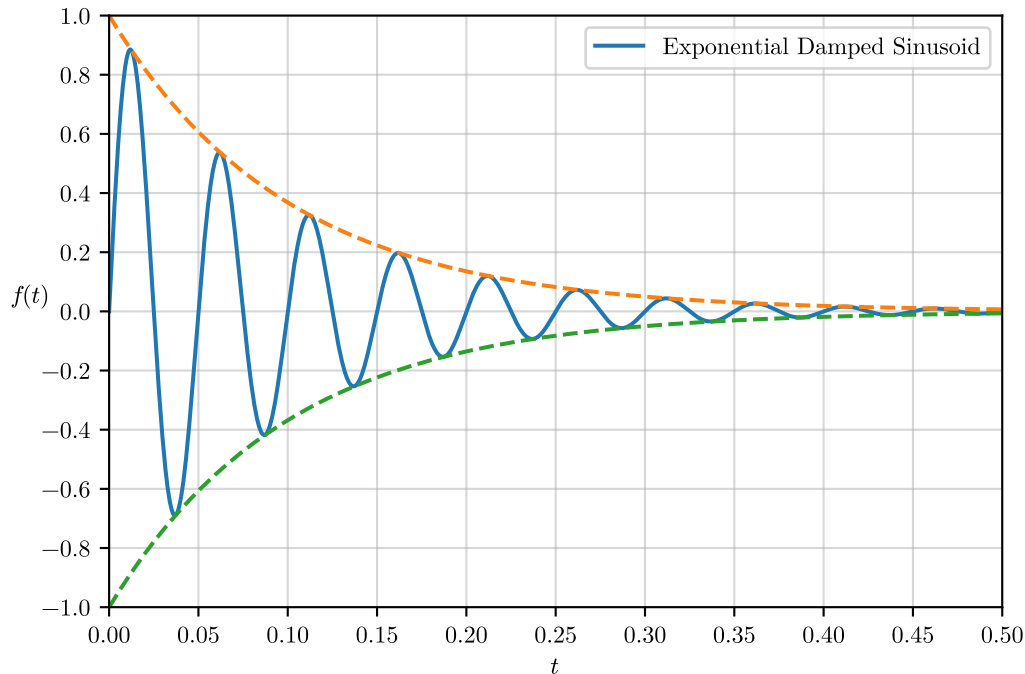

Figure 1.2: Exponentially damped sinusoid signal, $f(t) = e^{-10t}\sin(40\pi t)$

1.4   Topic 4: What are Signals?

1.4.1   Introduction

It is difficult to find a unique definition of a signal. However in the context of this course, we give a workable definition which suits most of out purposes as:

> *A Signal conveys information about a physical phenomenon which evolves in time or space.*

Examples of such signals include: Voltage, Current, Speech, Television, Images from remote space probes, Voltages generated by the heart and brain, Radar and Sonar echoes, Seismic vibrations, Signals from GPS satellites, Signals from human genes, and countless other applications.

1.4.2   Energy and Power Signals

> **Energy Signals**
>
> A signal is said to be an *energy signal* if and only if it has finite energy.

> **Power Signals**
>
> A signal is said to be a *power signal* if and only if the average power of the signal is finite and non-zero.

Instantaneous Power

The instantaneous power $p(t)$ of a signal $x(t)$ is expressed as:

$$p(t) = x^2(t) \tag{9}$$

Continuous-Time Signal Energy

The total energy of a continuous-time signal $x(t)$ is given by:

$$E = \lim_{T \to \infty} \int_{-T/2}^{T/2} x^2(t)\, \mathrm{d}t = \int_{-\infty}^{\infty} x^2(t)\, \mathrm{d}t \tag{10}$$

Complex Valued Signal Energy

For a complex valued signal:

$$E = \int_{-\infty}^{\infty} |x(t)|^2\, \mathrm{d}t \tag{11}$$

Average Power

Since power equals to the time average of the energy, the average power is given by:

$$P = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t)\, \mathrm{d}t = \frac{E}{T} \tag{12}$$

Note that during calculation of energy, we average the power over an indefinitely large interval.

A signal with finite energy has zero power and a signal with finite power has infinite energy.

Furthermore, some additional concepts of note:

a. A signal can not both be an energy and a power signal. This classification of signals based on power and energy are mutually exclusive.
b. However, a signal can belong to neither of the above two categories.
c. The signals which are both deterministic and non-periodic have finite energy and therefore are energy signals. Most of the signals, in practice, belong to this category.
d. Periodic signals and random signals are essentially power signals.
e. Periodic signals for which the area under $|x(t)|^2$ over one period is finite are power signals.

1.4.3 Examples

1.4.3.1 Example 1: Unit Step Function

Consider a unit step function defined as:

$$u(t) = \begin{cases} 1 & t \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

Determine whether this is an energy signal or a power signal or neither.

Solution: Let us compute the energy of this signal as:

$$E = \int_{-\infty}^{\infty} [u(t)]^2 \, \mathrm{d}t = \int_0^{\infty} [0]^2 \, \mathrm{d}t = \int_0^{\infty} [1]^2 \, \mathrm{d}t = \infty \tag{14}$$

Since the energy of this signal is infinite, it cannot be an energy signal. Let us compute the power of this signal as:

$$P = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} [u(t)]^2 \, \mathrm{d}t = \lim_{T \to \infty} \frac{1}{T} \int_0^{T/2} [u(t)]^2 \, \mathrm{d}t = \frac{1}{2} \tag{15}$$

The power of this signal is finite. Hence, this is a power signal.

1.4.3.2 Example 2: Exponential Function

Consider an exponential function defined as:

$$x(t) = e^{-at}u(t), \text{ where } u(t) \text{ is the unit step signal, } a > 0 \tag{16}$$

Classify this signal as an energy, power, or neither.

Solution: Let us compute the energy of this signal as:

$$E = \int_{-\infty}^{\infty} [x(t)]^2 \, \mathrm{d}t = \int_0^{\infty} [e^{-at}]^2 \, \mathrm{d}t = \int_0^{\infty} e^{-2at} \, \mathrm{d}t = \frac{1}{2a} < \infty \tag{17}$$

Thus, $x(t) = e^{-at}u(t)$ is an energy signal.

1.4.3.3 Example 3: Ramp Function:

Consider a ramp function defined as:

$$r(t) = \begin{cases} At & t \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

Classify this signal as an energy, power, or neither.

Solution: Let us compute the energy of this signal as:

$$E = \int_{-\infty}^{\infty} r(t)^2 \, \mathrm{d}t = \int_{-\infty}^{0} [0]^2 \, \mathrm{d}t = \int_0^{\infty} A^2 t^2 \, \mathrm{d}t = A^2 \frac{T^3}{3} \Big|_0^{\infty} = \infty \tag{19}$$

Since the energy of this signal is infinite, it cannot be an energy signal. Let us compute the power of this signal as:

$$P = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} [r(t)]^2 \, \mathrm{d}t = \lim_{T \to \infty} \frac{1}{T} \int_0^{T/2} A^2 t^2 \, \mathrm{d}t = A^2 \lim_{T \to \infty} \frac{1}{T} \frac{T^3}{3} \Big|_0^{\infty} = \infty \tag{20}$$

The power of this signal is infinite. Hence, this is neither a power nor an energy signal..

1.5 Topic 5: What are Systems?

1.5.1 Introduction

The term *system* is derived from the Greek word *systema*, which means an organised relationship among functioning units or components. It is often used to describe any orderly arrangement of ideas or constructs.

According to the Webster's Dictionary

"A system is an aggregation or assemblage of objects united by some form of regular interaction or interdependence; a group of diverse units so combined by nature or art as to form an integral; whole and to function, operate, or move in unison and often in obedience to some form of control..."

According to the International Council on Systems Engineering (INCOSE)

A system is an arrangement of parts or elements that together exhibit behaviour or meaning that the individual constituents do not.
The elements or parts, can include people, hardware, software, facilities, policies, and documents; that is, all things required to produce system-level results.

- It is difficult to give a single and precise definition of the term *system*, which will suit to different perspectives of different people.

- In practice, what is meant by "the system" depends on the objectives of a particular study.

- From the control engineering perspective, the system is any interconnection of components to achieve desired objectives. It is characterised by its **inputs**, **outputs**, and the rules of operations or laws. For example:
    a. The laws of operation in electrical systems are Ohm's law, which gives the voltage-current relationships for resistors, capacitors and inductors, and Kirchhoff's laws, which govern the laws of interconnection of various electrical components.
    b. Similarly, in mechanical systems, the laws of operation are Newton's laws. These laws can be used to derive mathematical models of the system.

1.5.2 The System as an Operator

The System Operator

A system is defined mathematically as a transformation which maps an input signal $x(t)$ to an output signal $y(t)$. For a continuous time system, the input-output mapping is expressed as:

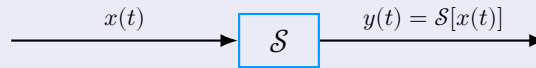$$y(t) = \mathcal{S}[x(t)], \quad \text{where } \mathcal{S} \text{ is an operator.} \tag{21}$$

Figure 1.3: System Operator Diagram

A control system may be defined as an interconnection of components which are configured to provide a desired response.

1.5.3 Classification of Systems

The basis of classifying systems are many. They can be classified according to the following:

Types of System Classifications

     a. **The Time Frame:** (*discrete, continuous or hybrid*);
     b. **System Complexity:** (*physical, conceptual and esoteric*);
     c. **Uncertainties:** (*deterministic and stochastic*);
     d. **Nature and type of components:** (*static or dynamic, linear or nonlinear, time-invariant or time variant, lumped or distributed etc*);
-        1. Linear and nonlinear systems
       2. Time-invariant and time-variant systems
       3. Static (memory less) and dynamic (with memory) systems
       4. Causal and Non-causal systems
       5. Lumped and distributed parameter systems
       6. Deterministic and stochastic systems
       7. Continuous and discrete systems

1.5.4 Linear and Nonlinear Systems

Linear and Nonlinear Systems

A system is said to be linear provided it satisfies the *principle of superposition* which is the combination of the *additive* and *homogeneity* properties. Otherwise, it is *non-linear*

1.5.4.1   Principle of Additivity

Principle of Additivity

Assume the system initially at rest.

▶ Suppose an input $x_1(t)$ to this system produces an output $y_1(t)$ and an input $x_2(t)$ produces an output $y_2(t)$.

▶ If the system is linear, then the application of the input $x_1(t) + x_2(t)$ will produce an output $y_1(t) + y_2(t)$. Thus if

$$x_1(t) \rightarrow y_1(t) \quad \text{and} \quad x_2(t) \rightarrow y_2(t), \quad \text{then} \quad x_1(t) + x_2(t) \rightarrow y_1(t) + y_2(t) \tag{22}$$
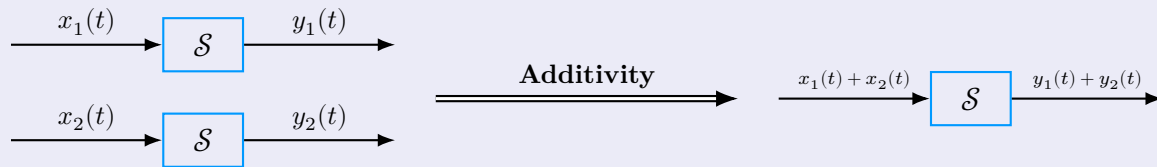

Figure 1.4: Principle of Additivity

1.5.4.2   Principle of Homogeneity or Scaling

Principle of Homogeneity

▶ Let an input *(cause)* $x(t)$ produce an output *(effect)* $y(t)$. If the system is linear then,

▶ Scaling the input *(cause)* $x(t)$ by a factor "a" will scale the output *(effect)* $y(t)$ by the same factor.

▶ Thus, if the input $x(t)$ results in $y(t)$, then the scaled input $ax(t)$ gives the output $ay(t)$ where "a" can either be a real or imaginary number. Thus, for a linear system:

$$x(t) \rightarrow y(t) \quad \Rightarrow ax(t) \rightarrow ay(t) \tag{23}$$

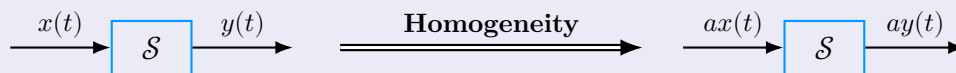Figure 1.5: Principle of Homogeneity

1.5.4.3   Principle of Superposition

> **Principle of Superposition**
>
> This is the combination of both the additive and homogeneity properties for a linear system.
> If the zero state response of a linear system due to a finite N-number of inputs $x_1(t), x_2(t), \dots, x_N(t)$ equals to $y_1(t), y_2(t), \dots, y_N(t)$ respectively, then the response of the system to the linear combination of these inputs
>
> $$a_1 x_1(t) + a_2 x_2(t) + a_3 x_3(t) + \dots \dots + a_N x_N(t) \tag{24}$$
>
> is given by the linear combination of the individual outputs i.e.
>
> $$a_1 y_1(t) + a_2 y_2(t) + a_3 y_3(t) + \dots \dots + a_N y_N(t) \tag{25}$$
>
> where $a_1, a_2, a_3, \dots, a_N$ are arbitrary constants (either real or imaginary). Thus, if
>
> $$x_1(t) \to y_1(t), \quad x_2(t) \to y_2(t), \dots, x_N(t) \to y_N(t) \tag{26}$$
>
> then, for all $a_1, a_2, \dots, a_N$
>
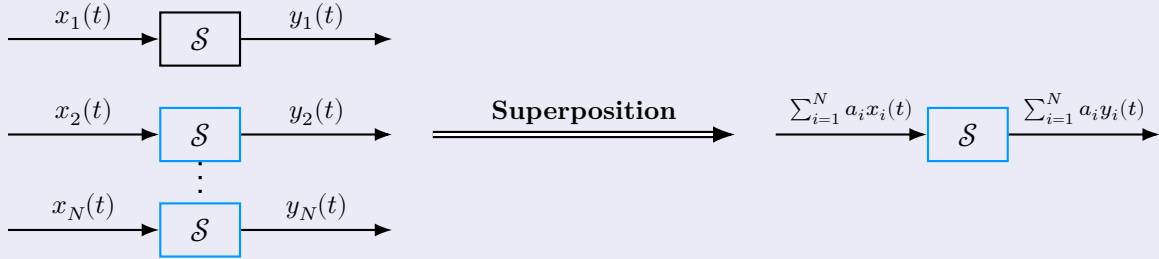> $$a_1 x_1(t) + a_2 x_2(t) + \dots a_N x_N(t) \to a_1 y_1(t) + a_2 y_2(t) + \dots a_N y_N(t) \tag{27}$$
>
>
>
> Figure 1.6: Superposition Principle Diagram