Automatic control with interactive tools

Supplementary material

Introduction to control system analysis and design in state-space

Second order systems

José Luis Guzmán
(*Professor of Systems Engineering and Automatic Control - UAL*)

Ramón Costa-Castelló
(*Professor of Systems Engineering and Automatic Control - UPC*)

Manuel Berenguel
(*Professor of Systems Engineering and Automatic Control - UAL*)

Sebastián Dormido
(*Professor of Systems Engineering and Automatic Control - UNED*)

Introduction to control system analysis and design in state-space

Introduction

In automatic control theory, two descriptions of dynamical systems are commonly used, the external description and the internal description. The external description establishes an explicit functional relationship between input and output signals, which in the text has been expressed in the form of a transfer function. The internal description is based on the concept of the *state* of a dynamical system:

- The *state* of a dynamical system is understood as the minimum set of variables such that if their value is known at an instant t_0 as well as the inputs for $t \geq t_0$, the behavior of the system is completely determined for any future time $t \geq t_0$ (if the system is deterministic[1]).
- That minimum set of variables defining the state of the dynamical system are known as *state variables*. If at least n variables $x_1(t), x_2(t), ..., x_n(t)$ are required to fully describe the behavior of a dynamical system, such n variables are a set of state variables. In general, the set of dependent variables constituting the state can be chosen in different ways. In physical systems, the state is composed of variables that represent storage of mass, energy and momentum.
- These n variables that configure the state are usually grouped in a vector that is called the *state vector* of the dynamical system. Thus, a state vector is one that uniquely determines the state of a dynamical system $x(t)$ for any time $t \geq t_0$, once the state at $t = t_0$ and the input $u(t)$ for $t \geq t_0$ are known.A dynamical system with one input $u(t)$ and one output $y(t)$ can be represented by an ordinary differential equation (ODE) of the form:

  $$\frac{dx(t)}{dt} = \dot{x}(t) = f(x(t), u(t)), \quad y(t) = g(x(t), u(t))$$

  where $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ and $g : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ are functions without discontinuities, where n is the dimension of the state vector (system *order*). A representation of this type is known as a *state*

[1] A system in which chance is not involved in the future states of the system.

space model. When the functions f and g are not explicitly time dependent, it is a time invariant system. The function f provides the speed of change of the state vector as a function of the state $x(t)$ and the control signal $u(t)$ and the function g provides the measured values of the output as a function of the state $x(t)$ and the input $u(t)$ [1]. These functions determine the time response of the system [4]. ($\bar{u}$,$\bar{x}$) is said to be an *equilibrium state* (*point*) if $f(\bar{u}, \bar{x}) = 0$ (notice that notation (u_e,x_e) will also be used in this framework to describe the equilibrium state). If that condition is met, when all the constituents of the transient response have decayed (under the assumption of a stable system), the system output reaches a *steady-state*. An equilibrium state is a solution of the original differential equation in a steady-state characterized by a constant position or oscillation (in the latter case the system state is called a permanent regime). The equation $\dot{x} = 0$ may have several solutions. Which ones are acceptable depends on their stability, the stable ones are locally considered [2]. If functions f and g are linear in x and u, the system is linear time invariant (LTI) and its representation will be given by:

$$\begin{aligned} \frac{dx(t)}{dt} = \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{0.1}$$

The first equation is the so-called *state equation* and the second *output equation*. In such a representation A, B, C and D are matrices of constant coefficients. The matrix A is is called the *dynamic matrix*, B is the *control matrix*, C is the *sensor matrix*, and D is the *direct term*, which is usually equal to zero because otherwise it would provide non strictly causal systems in which a change in the input would be directly reflected in the output [1].

- As said before, the n-dimensional space whose coordinate axes correspond to the state variables $x_1(t)$, $x_2(t)$, ..., $x_n(t)$ is called *state space*. Any state is represented by a point in the state space, so that the dynamic evolution of a system is a trajectory starting from the initial point (determined by the value of the state vector at $t = t_0$) and arriving at the end point, which is usually an equilibrium state associated with a constant position or oscillation (steady-state or permanent regime). A *trajectory* or *orbit* is the locus of $x(t)$ in the *phase portrait* passing through x_0. Different initial states result in different trajectories. The set of of all trajectories forms the phase portrait of a dynamical system, though in practice, only representative trajectories are considered. *Phase planes* typically arise in the context of two-dimensional autonomous ODEs. The phase plane shows the vector field of the system, which gives the velocity of the state (represented by an arrow) at each point in the state-space. A family of evolution trajectories of the system is often referred to as the *phase plane* (two-dimensions) or *phase portrait* in the general case. The phase plane is a state-space in which the coordinates of a point provide all the information about the state of the system. In the phase plane the arrows shown above are called *isoclines*, since they define the curve in the plane where the trajectories have the same velocity (slope). Indeed, since $\dot{x}(t) = f(x) = Ax(t) + Bu(t)$, $f(x)$ is a tangent vector to the curve $x(t)$ in the phase plane and thus constitutes a vector field whose isoclines are the points where $f(x)$ is constant ($Ax(t) + Bu(t) = c$). For a generic second order *autonomous* system[2]:

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \\ \dot{x}_2 &= f_2(x_1, x_2) \end{aligned}$$

the slope of the trajectory at the point x can be defined as:

$$S(x) = \frac{f_2(x)}{f_1(x)}$$

Equation $S(x) = c$, with constant c, defines the isocline curves for different values of c.

- In this context, the *free response* of a dynamical system (also called *natural response* or *zero-input response*) is the response generated, in the absence of inputs, by the *initial conditions* of the system (non-zero). It is the part of the total response due to the system and the way it acquires or dissipates energy. It is also considered the part of the response determined by the roots of the *characteristic polynomial*. The *forced response* is that due exclusively to variations in the input signal, assuming zero initial conditions. The forced response is the solution when the initial conditions are zero and the system is subjected to an input signal $u(t)$. In LTI systems, the total response of the system is the superposition of the free response and the forced response.

Logically, there is a relationship between the external description and the internal description. Mathematical models of LTI systems are usually represented by linear ODEs of order n. Applying the properties of the Laplace transform with zero initial conditions, a transfer function can be obtained whose order coincides with that of the differential equation from which it comes. The internal description starts from the same differential equation, but transforming it to a system of n first-order differential equations, where each one describes the dynamic evolution of a state. In a way, through the state variables, all the dynamic past of the system at the current instant t is being condensed. In the transformation of the linear differential equation of order n to the system of n first-order equations, there are theoretically infinite ways of defining the state variables and in fact, sometimes they will have a certain physical meaning, but most of the time they will be mathematical variables without physical meaning.

[2]A system that is not subject to external signal influences.

State-space

Interactive tool: state_space

Concepts analyzed in the card and learning outcomes

- Concept of state.
- Internal description.
- Differential equation of state.
- State-space.
- Canonical forms.
- Exponential matrix.
- Convolution equation.
- Linear feedback control law of the complete state.

Theory As discussed in the introduction to this document, the two typical representations of time-invariant linear system models are the transfer function (external description) and the state space representation (internal description). The general realization in state-space is given by:

$$\dot{x}(t) = \frac{dx(t)}{dt} = Ax(t) + Bu(t) \tag{0.2}$$
$$y(t) = Cx(t) + Du(t) \tag{0.3}$$

The first equation is the so-called *state equation* and the second *output equation*. In LTI systems, A, B, C and D are matrices of constant coefficients. In this application we will deal with the case of strictly causal systems and therefore we will consider $D = 0$ throughout the development, so that there is no direct transfer of signals from input to output. The vector $x(t)$ is the state vector and its components are the state variables; $u(t)$ and $y(t)$ are respectively the input and output of the system[3].

In this document we will deal exclusively with the complete state vector feedback problem by studying the case of systems with one input and one output where $n = 2$, since it is the simplest case and richest in visual content. The summary, however, will be done for the generic n-dimensional case for systems with one input and one output.

[3]In general, if the system has n-order with m outputs and p inputs the dimensions of the state-space matrices of the model are: $A = n \times n$; $B = n \times p$; $C = m \times n$; $D = m \times p$.

The representations in the state-space are not unique, and there are a number of *canonical forms*, among which the following two stand out. If the transfer function of a strictly causal system[4] is given by:

$$G(s) = \frac{b_{n-1}s^{n-1} + ... + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + ... + a_1 s + a_0} \tag{0.4}$$

the following canonical forms, among others, are obtained:

Controllable Canonical Form (C):

$$A_C = \begin{bmatrix} 0 & \ddots & & \\ 0 & & I_{n-1} & \\ \vdots & & & \ddots \\ -a_0 & -a_1 & \cdots & -a_{n-1} \end{bmatrix}; \quad B_C = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$C_C = [b_0 \; b_1 \; \cdots \; b_{n-1}]; \quad D_C = [0] \tag{0.5}$$

Observable Canonical Form (O):

$$A_O = \begin{bmatrix} 0 & \cdots & 0 & -a_0 \\ \ddots & & & -a_1 \\ & I_{n-1} & & \vdots \\ & & \ddots & -a_{n-1} \end{bmatrix}; \quad B_O = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{bmatrix}$$

$$C_O = [0 \; \cdots \; 0 \; 1]; \quad D_O = [0] \tag{0.6}$$

where I_{n-1} is the identity matrix of dimension $n-1$.

On the other hand, if the Laplace transform is applied to the state-space representation, the relationship linking the matrices of the state-space description with the transfer function is obtained:

$$G(s) = C(sI - A)^{-1}B + D \tag{0.7}$$

The solution of the equations defining the state-space representation of a LTI system is obtained through the *convolution equation*[5]:

$$x(t) = \underbrace{e^{At}x_0}_{\text{free response}} + \underbrace{\int_0^t e^{A(t-\xi)}Bu(\xi)d\xi}_{\text{forced response}}, \quad e^{At} = \mathcal{L}^{-1}\{(sI-A)^{-1}\} = \Phi(t) \tag{0.8}$$

[4]The general case is described in [7], pages 649-651. If the system is not strictly causal, the numerator in (0.4) includes a term $b_n s^n$ added to the polynomial, so that the order of the numerator is equal to that of the denominator. If $b_n \neq 0$, C in the Controllable Canonical Form is given by $C_C = [(b_0 - a_0 b_n) \;\; (b_1 - a_1 b_n) \; ... \; (b_{n-1} - a_{n-1} b_n)]$ and $D_C = b_n$. Notice that in the Observable Canonical Form $B_O = C_C^T$, $\cdot^T$ meaning transpose, $C_O = B_C^T$, and $D_O = D_C$.

[5]Convolution is a complex operation on functions defined by the integral of the two functions multiplied together and shifted in time.

where it has been considered that $t_0 = 0$, $\Phi(t)$ is the *state transition matrix* and $x_0 = x(t_0) = x(0)$. It can be approximated by the series expansion of the exponential matrix[6]:

$$\Phi(t) = e^{At} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + ... = \sum_{k=0}^{\infty} \frac{A^kt^k}{k!} \tag{0.9}$$

Differentiating this expression with respect to t gives:

$$\frac{d}{dt}\Phi(t) = \frac{d}{dt}e^{At} = A + A^2t + \frac{A^3t^2}{2!} + ... = A\sum_{k=0}^{\infty} \frac{A^kt^k}{k!} = Ae^{At} \tag{0.10}$$

When working with a description of LTI system in state-space, the system is said to be *asymptotically stable* if the state $x(t) \to 0$ when $t \to \infty$ with $x(0) = x_0$. This can be shown to be the case when the eigenvalues[7] of matrix A have negative real part. Moreover, if there are no pole-zero cancelations, it can be demonstrated that the eigenvalues of matrix A and the poles of the system are equal. The solution of the autonomous (unforced) system $\dot{x}(t) = Ax(t)$ for a state is given by $x_i(t) = x_i(0)e^{\lambda_i t}$, with λ_i being the eigenvalues of the system or roots of the characteristic equation $J(s) = \det(sI - A) = 0$. Therefore, the stability can be easily evaluated by analyzing the sign of the eigenvalues λ_i, so that the system will be stable if all eigenvalues have negative real part.

An autonomous linear system given by $\dot{x}(t) = Ax(t)$ has a phase plane with non-intersecting trajectories describing the evolution of the system for any initial condition. For a generic system $\dot{x}(t) = f(x(t))$, $\bar{x} = x_e$ is said to be an equilibrium state if $f(x_e) = 0$. For the second-order systems discussed in this card, as long as matrix A is non-singular, six types of responses associated with the equilibrium states can be obtained, depicted in Figure 0.1. The eigenvalues describe how the solution varies over time and this solution is often also called the *mode* of the system, while the eigenvectors provide the *shape* of the solution [1]. This concept of stability is part of a more general one, valid also for non-linear systems. Let $x(t, x_{0_a})$ be a solution to the differential equation with initial condition x_{0_a}. A solution is stable if other solutions whose initial condition is close to x_{0_a} remain close to $x(t, x_{0_a})$ [1]. Formally, the solution $x(t, x_{0_a})$ is said to be stable if for all $\epsilon > 0$, there exists $\delta > 0$ such that:

$$||x_{0_b} - x_{0_a}|| < \delta \to ||x(t, x_{0_b}) - x(t, x_{0_a})|| < \epsilon, \forall t > 0$$

This type of stability is often referred to as *stability in the Lyapunov sense* or *Lyapunov stability*. When $x(t, x_{0_a}) = x_e$ the equilibrium point is said to be stable. In autonomous linear systems, the origin is always an equilibrium point. A solution $x(t, x_{0_a})$ is *asymptotically stable* if it is stable in the Lyapunov sense and it satisfies that $x(t, x_{0_b}) \to x(t, x_{0_a})$ when $t \to \infty$ for x_{0_b} for x_{0_b} is sufficiently close to x_{0_a} [1].

[6] An excellent summary of definitions and matrix calculus results useful for the study of many questions related to dynamic and control systems can be found in Appendix of the reference [3].

[7] In linear algebra, the eigenvectors of a linear operator are the nonzero v vectors such that, when transformed by the operator, they give rise to a scalar multiple of themselves, thus not changing their direction. This scalar λ is called the eigenvalue. Therefore, v is said to be an eigenvector of A with eigenvalue λ if $Av = \lambda v$. Often, a transformation is completely determined by its eigenvectors and eigenvalues. An eigenspace is the set of eigenvectors with a common eigenvalue.

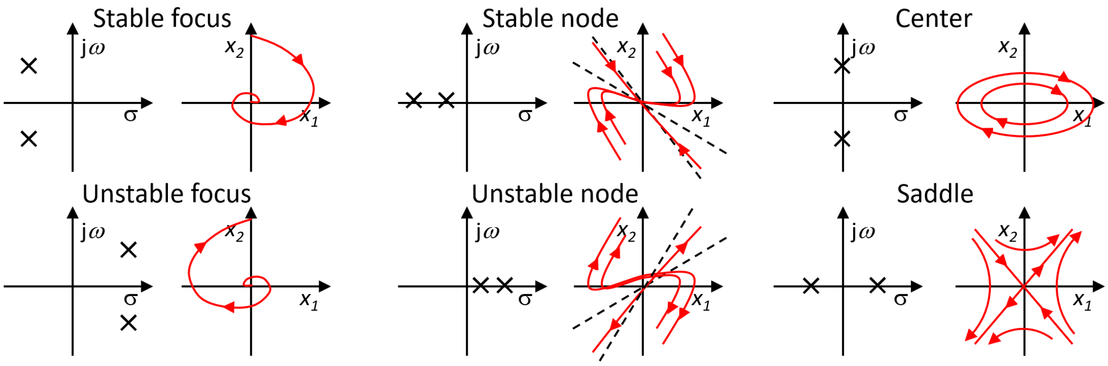


Figure 0.1 Classification of equilibrium points for second order systems

The introduction of control concepts when working with a state-space representation is usually done by means of so-called regulators. Regulators tend to keep the system output constant in the presence of disturbances. For this reason, when designing regulators, transient specifications are taken into account, which translate into achieving a desired position of the poles of the system using feedback. In state-space, the most classical way of approaching the control problem is to assume that the system starts from an initial state $x(t_0) \neq 0$ and that through control it has to be brought to the origin of the state-space (equilibrium state) with dynamic characteristics determined by the assignment or placement of the poles of the closed-loop system. The control applied is generally based on a linear feedback of the state vector, where the measurements or estimates of the system states are used to generate the control signal $u(t) = -Kx(t)$ after multiplication by gains computed to provide the desired behaviour.

A system is said to be *controllable* if, using an appropriate set of inputs (which in state-space control consists of a linear feedback of the state vector $u(t) = -Kx(t)$, as said before), states can be moved in an arbitrary direction in the state space. In other words, a system is fully controllable if there is an unrestricted control $u(t)$ that can take any initial state $x(t_0)$ to any other desired state $x(t)$ in a finite time $t_0 \leq t \leq t_f$ [4]. This concept is equivalent to being able to place the poles of the controlled system at any location in the complex plane. The pair (A, B) is *controllable* if and only if the rank of the *controllability matrix* Ω is n (the determinant of Ω is non-zero), where n is the order of the system, that is, the dimension of A. The controllability matrix is given by: $\Omega = \left[B\ AB\ A^2B\ ...\ A^{n-1}B\right]$.

If the system is linear and completely controllable, there are several formulas to calculate the values of the feedback gains (elements of the vector K). The characteristic polynomial of the system is given by the eigenvalues of the dynamic matrix A: $J(s) = \det(sI - A) = s^n + a_{n-1}s^{n-1} + ... + a_1 s + a_0 = 0$. With $u(t) = -Kx(t)$, if we consider the case of regulation to the origin (without external reference), we have that $\dot{x}(t) = (A - BK)x(t)$, so the characteristic polynomial of the closed-loop is given by the eigenvalues of the matrix $A_{cl} = (A - BK)$. The transient regime specifications can be translated into the location of the dominant poles of the closed-loop system, i.e. the characteristic equation of the closed-loop system: $J_{cl}(s) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_0$. Therefore, a possible synthesis strategy (called *eigenvalue assignment* or *pole placement*) can be to impose that $\det(sI - A_{cl}) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_0$.

The Controllable Canonical Form has the property that the parameters of the system are the coefficients of the characteristic polynomial. Therefore, it is natural to work with this representation when solving the eigenvalue assignment problem, as it produces a vector of feedback gains equal to $K = [(\alpha_{n-1} - a_{n-1}) \ ... \ (\alpha_0 - a_0)]$.

The *Ackermann's formula* [7] allows calculation of the feedback gain matrix of the state vector to place the closed-loop poles in a desired position:

$$K = [0\,0\ \cdots\ 1]\,\Omega^{-1} J_{cl}(A) \tag{0.11}$$

where $J_{cl}(A)$ is the characteristic polynomial of the closed-loop particularized for matrix A ($J_{cl}(A) = A^n + \alpha_{n-1}A^{n-1} + \cdots + \alpha_1 A + \alpha_0 I$).

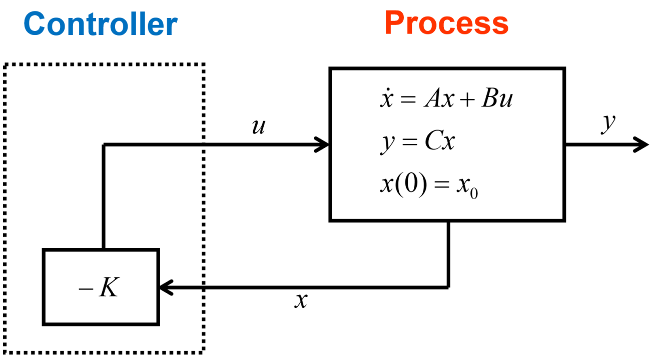


Figure 0.2 State feedback stabilization: Free response

With this idea as a basis, a series of control schemes can be developed depending on the characteristics of the disturbances exogenous to the system (existence of reference and input disturbances). The basic control scheme is the one shown in Figure 0.2, which represents a stabilization (regulation) at the origin with free response of the system. The dynamics of the closed-loop system is described by:

$$u(t) = -Kx(t) \rightarrow \dot{x}(t) = (A - BK)x(t) = A_{cl}x(t) \rightarrow x(t) = e^{A_{cl}t}x_0,\ y(t) = Ce^{A_{cl}t}x_0,\ x_0 = x(0) \tag{0.12}$$

If the control problem is the tracking of a reference r, there are several control structures. The simplest is shown in Figure 0.3, where the system is stabilized by following the input step using a control structure with only one-degree-of-freedom (the gain K). The equations governing the dynamics of the closed-loop system are:

$$u(t) = -Kx(t) + r(t) \rightarrow \dot{x}(t) = (A - BK)x(t) + Br(t) \tag{0.13}$$

If the reference is assumed to be constant at steady-state (as is the usual case and the case discussed in this document), the equilibrium condition $\bar{x} = x_e = x(t \rightarrow \infty), \dot{x}(t) = 0$, is given by:

$$x_e = -(A - BK)^{-1}Br \rightarrow y_e = Cx_e = -C(A - BK)^{-1}Br$$

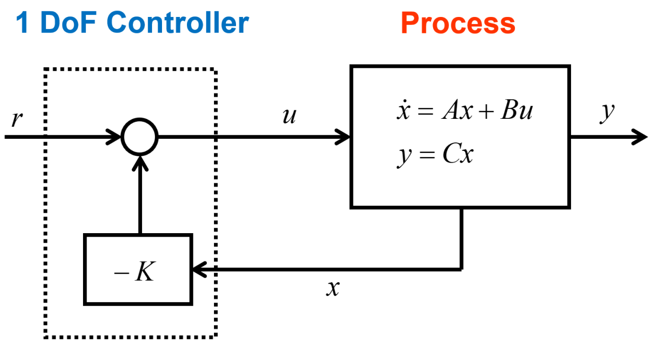


Figure 0.3 State feedback stabilization: Step response - system with one-degree-of-freedom (1DoF)

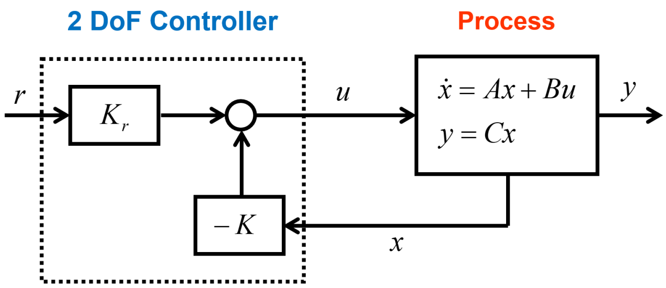


Figure 0.4 State feedback stabilization: Step response - system with two-degree-of-freedom (2DoF)

where in general $y_e \neq r$ and the response to a set point step produces a steady-state error.

To solve this problem in the presence of a non-zero reference signal, a scheme with two-degree-of-freedom is introduced as shown in Figure 0.4. The reference is multiplied by a design constant called *reference gain* K_r which aims to eliminate steady-state error:

$$u(t) = -Kx(t) + K_r r(t) \rightarrow \dot{x}(t) = (A - BK)x(t) + BK_r r(t) \tag{0.14}$$

In steady-state:

$$x_e = -(A - BK)^{-1} BK_r r \rightarrow y_e = Cx_e = -C(A - BK)^{-1} BK_r r$$

$$K_r = \frac{-1}{(C(A - BK)^{-1}B)} \rightarrow y_e = r$$

As can be seen, an appropriate choice of gain K_r allows the steady-state error to be eliminated, but does not affect the stability of the system (which is determined by the eigenvalues of $(A - BK)$).

This scheme works well when only changes in the reference occur, but causes steady-state error when a step load disturbance $d(t)$ is introduced at the input of the system (Figure 0.5). The equations governing the dynamics and steady-state of this system are:

$$u(t) = -Kx(t) + K_r r(t) + d(t) \rightarrow \dot{x}(t) = (A - BK)x(t) + B(K_r r(t) + d(t)) \tag{0.15}$$

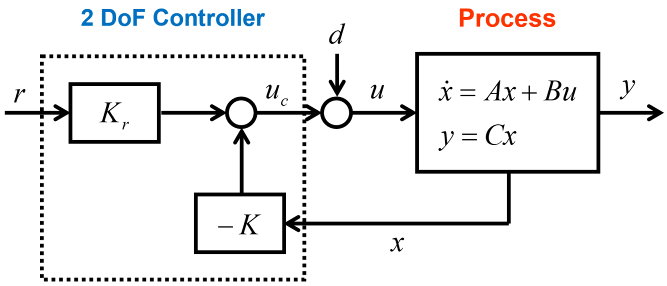


Figure 0.5 State feedback stabilization: Step response - system with two-degree-of-freedom (2DoF) and introduction of a load disturbance $d(t)$ at the system input

$$\left.\begin{aligned} x_e &= -(A-BK)^{-1}(BK_r r + Bd) \\ K_r &= \frac{-1}{(C(A-BK)^{-1}B)} \end{aligned}\right\} \rightarrow y_e = r - C(A-BK)^{-1}d \rightarrow y_e \neq r$$

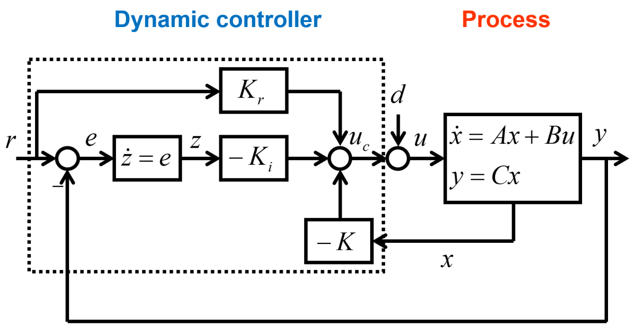


Figure 0.6 State feedback stabilization: Step response - system with two-degree-of-freedom (2DoF). Introduction of a load disturbance $d(t)$ at the system input and compensation by integral action (increasing the order of the system)

Therefore, to compensate for the influence of the load disturbance, it is necessary to introduce integral effect in the control loop, increasing the dimension of the system by one through the inclusion of an additional state variable:

$$\left.\begin{aligned} \begin{bmatrix} \dot{x}(t) \\ \dot{z}(t) \end{bmatrix} &= \begin{bmatrix} Ax(t) + Bu(t) \\ r(t) - y(t) \end{bmatrix} \\ u(t) = -Kx(t) - K_i z(t) &+ K_r r(t) + d(t) \end{aligned}\right\} \rightarrow \begin{aligned} & x_e = -(A-BK)^{-1}B(K_r r - K_i z_e + d) \\ & r - y = 0 \rightarrow r = y \end{aligned} \tag{0.16}$$

The response in the phase plane does not follow the isocline map exactly. This is because the system is of third order and the isocline map only depends on two variables.

Although the tool does not include a state observer analysis (only the description of the system in Observable Canonical Form), a brief summary of its fundamentals is given below.

One of the main problems with state-space design is that it is usually impractical, either because not all states are measurable or because it is too expensive to place many sensors. Hence the use of so-called *observers* makes sense, by means of which, from a set of inputs and outputs of the system, it is possible to estimate the states of the system. If the estimated state is denoted as $\hat{x}(t)$, it follows that:

$$\frac{dx}{dt} = Ax(t) + Bu(t), \qquad x(0) = x_0 \tag{0.17}$$

$$\frac{d\hat{x}(t)}{dt} = A\hat{x}(t) + Bu(t), \qquad \hat{x}(0) = \hat{x}_0 \tag{0.18}$$

If we define $\breve{x}(t) = x(t) - \hat{x}(t)$, then:

$$\frac{d\breve{x}(t)}{dt} = A\breve{x}(t), \;\; \breve{x}(0) = \breve{x}_0 \rightarrow \breve{x}(t) = e^{At}\breve{x}_0 \tag{0.19}$$

Therefore, the observer error $\breve{x}(t)$ tends to zero if the original system is stable. Since the initial conditions of the system are not known, they have to be estimated. As formulated, this is an open-loop observer. If the known output of the system is used and compared to that predicted by the observer, a closed-loop observer formulation can be obtained:

$$\frac{\hat{x}(t)}{dt} = A\hat{x}(t) + Bu(t) + L(y(t) - C\hat{x}(t)); \quad \hat{x}(0) = \hat{x}_0$$

The dynamics with which the error between observed and real states tends to zero is determined by $(A - LC)$.Therefore, L is chosen so that the dynamics of the evolution of the error given by the eigenvalues of $(A - LC)$ is faster than the evolution of the system. Actually, in order to choose it, it is necessary to previously analyze the possible noise and disturbances to which the output may be subjected, in order to avoid the introduction and amplification of signals not modeled in the scheme. If these eigenvalues can be located at any point in the complex plane, the system is said to be *observable*. Mathematically, a system is observable if the rank of the observability matrix is n, where n is the dimension of the system. The observability matrix is given by: $\Theta = \left[C^T \; (CA)^T \; (CA^2)^T \; ... \; (CA^{n-1})^T\right]^T$.

The calculation of the observer gain vector L is done using a formula similar to the calculation of the feedback gain vector K. If the desired characteristic polynomial for the state observer is denoted by $J_{O_{cl}}$, the vector L can be obtained from:

$$L = J_{O_{cl}}(A) \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-2} \\ CA^{n-1} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

So-called minimum-order observers can also be designed, where it is assumed that only one set of plant states needs to be estimated since the rest (usually the output) can be directly measured by sensors in the system. The design of this type of observers is discussed in [7] and [8].

The idea behind the development of observers is that the estimated states can be used in the control law by linear feedback of the state vector as if they were the true states. The question that arises at this point is whether the stability of the closed-loop system is guaranteed in this way. The answer is that there is a separation between the control and observation problems, i.e., one can find the controller gains assuming that the states are accessible and then design an observer to estimate the states and use them in place of the true states. The poles of the closed-loop system are the union of the controller and observer poles. This leads to the well-known property or separation principle, which is a crucially important result in modern control theory. This separation property can be analyzed by combining the dynamics of the real states and of the estimation errors, obtaining:

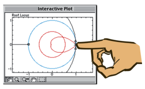
$$\begin{bmatrix} \frac{dx(t)}{dt} \\ \frac{d\breve{x}(t)}{dt} \end{bmatrix} = \begin{bmatrix} A-BK & BK \\ 0 & A-LC \end{bmatrix} \begin{bmatrix} x(t) \\ \breve{x}(t) \end{bmatrix}$$

Note that from matrix theory, the eigenvalues of a piecewise triangular matrix are the union of the eigenvalues of the diagonal blocks, from which the above separation theorem is obtained.

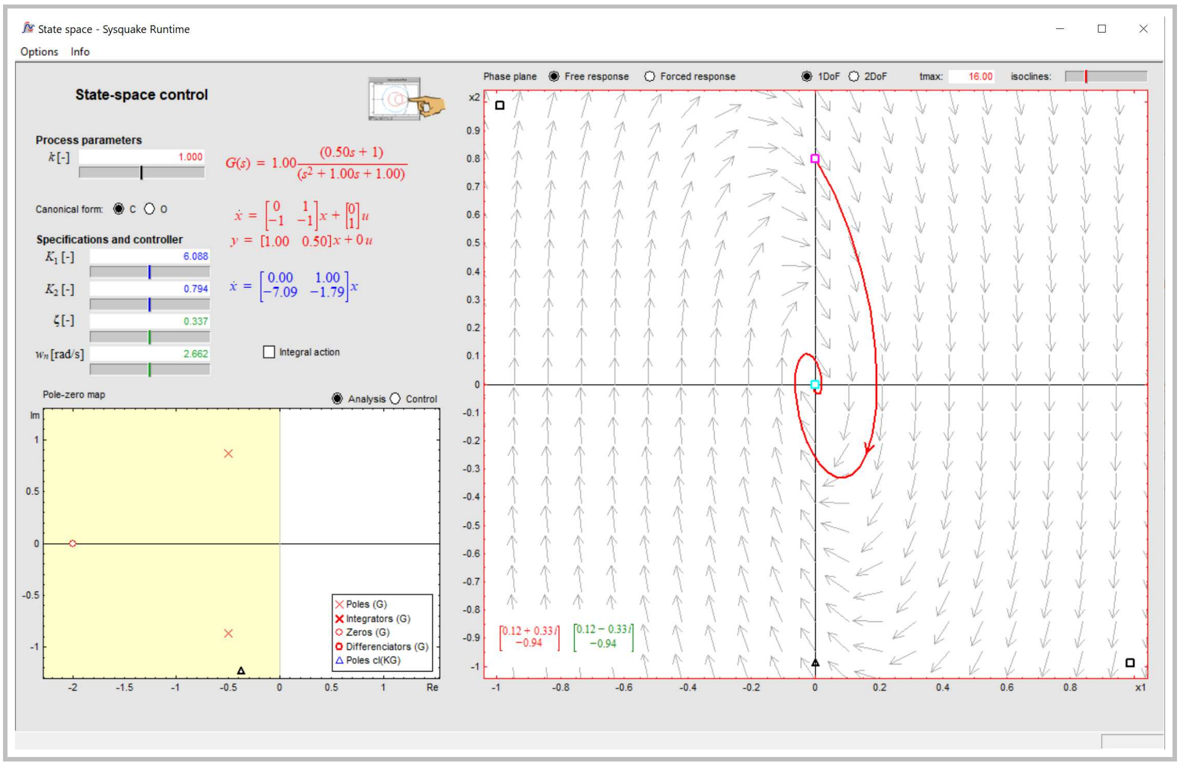
References

- [1] Åström, K. J. and R. M. Murray. Feedback systems: An introduction for scientists and engineers. Second edition. Princeton University Press, ISBN:9780691193984, 2014.
- [3] Bolzern, P., R. Scattolini and N. Schiavoni. Fundamentos de control automático (Fundamentals of automatic control). McGraw-Hill, ISBN: 978-84-481-6640-3, 2009.
- [4] Dorf, R. C. and R. H. Bishop. Modern control systems. Twelfth edition. Prentice Hall, ISBN: 978-0-13-602458-3, 2011.
- [7] Ogata, K. Modern control engineering. Fifth edition. Prentice Hall, ISBN: 978-0-13-615673-4, 2010.
- [8] Ollero, A. Control por computador: Descripción interna y diseño óptimo (Computer control: Internal description and optimal design). Marcombo, ISBN:84-267-0813-7, 1991.
- [9] Shahian, B. and M. Hassul. Control system design using Matlab. Prentice Hall, ISBN: 0-13-174061-X, 1993.

Application



Interactive tool: state_space



State-space control

As indicated in the theory overview, this tool treats multiple concepts in a single application. However, the way it is structured facilitates the user to learn incrementally the ideas derived from the use of a state-space representation of the system. The systems treated are second-order systems with the possibility of incorporating a zero (the degree of the polynomial of the numerator will always be $m \leq 1$). No time delays are considered in this tool so as not to overcomplicate the understanding of the basic concepts.

The Options menu is first analyzed, which contains two groupings. Through the first of these, an (exclusive) selection is made of the type of graphic representation to be used in the right-hand area of the tool:

- Time response: This is the default option when starting the application. The graphs represent the time evolution of the system states (in the case of Free response) or the output of the system (if the Forced response is used).

Application..

- Phase plane: If selected, this disables the Time response option and changes the graph in the right-hand area of the tool to a two-dimensional representation of the phase plane where the abscissa axis contains the coordinates of the first state $x_1(t)$ and the ordinate axis contains the coordinates of the second state $x_2(t)$.

The other group of items in the Options menu allows the introduction of different plant structures. In all cases it is restricted to second-order systems, which are the ones supported by the tool (as visually they are the ones that support a two-dimensional representation in the phase plane). The first elements of this group are the transfer functions $P_1(s)$ $(n = 2)$, $P_3(s)$ $(n = 2)$, $P_5(s)$, $P_6(s)$, $P_7(s)$, $P_{7b}(s) = P_7(s)$ with $beta = 0$, $P_{12}(s)$ and $P_{12b}(s) = P_{12}(s)$ with $beta = 0$, included in the book [6]. Then follows the option to include transfer functions in (NUM,DEN) or ZPK formats, restricted to strictly causal second-order structures. This tool has also incorporated the possibility of representing the system in internal description format (A,B,C), assuming D=0. In the latter case, the matrices and vectors have to be entered in the following format: A=[a11,a12;a21,a22], B=[b1;b2], C=[c1,c2].

There are other important options that determine the configuration of the tool depending on the type of cases to be analyzed. To facilitate interactivity, instead of placing these alternatives in the Options menu, they have been placed over the graphs representing the Time response or the Phase plane and also on the Pole-zero map. These are mutually exclusive pairs of options (exclusionary selection ◉○):

- Free response - Forced response: These two options allow the tool to be configured so that the time response or phase plane shown in the graphs (and the associated parameters), consider only the free response of the system, i.e. the regulation to the origin (null reference) or the forced response (tracking of a reference starting from null initial conditions).
- 1DoF - 2DoF: These options make sense when the Forced response representation has been selected, as they enable the structure of one-degree-of-freedom (1DoF) in Figure 0.3 or two-degree-of-freedom (2DoF) in Figure 0.4, where it is taken into account that the reference may be non-zero. When in the Options menu the Phase plane representation is selected, a scroll bar is added on the graph to facilitate the adjustment of the size of the isoclines.

In the lower left area of the tool the Pole-zero map is placed, where the poles (eigenvalues of the matrix A) and zeros of the system under study are located, represented in red colour. Next to the title, there are two excluding selection buttons called Analysis and Control:

- By selecting the Analysis option, it is possible to study the system response and the location of the closed-loop poles (represented by the non-interactive symbol △) when the feedback gains K_1 and K_2 are modified. The closed-loop zeros coincide with the open-loop zeros.

Application...

- If the Control option is chosen, it allows the closed-loop analysis of the system to be carried out when the location of the poles of the system has been fixed through the specifications and, therefore, the calculation of the feedback gains is carried out automatically through Ackermann's formula. In this case, the desired closed-loop poles are interactive, to allow the user to position them anywhere in the complex plane so that they define certain performance specifications. These poles defining the desired closed-loop behaviour are represented by the symbol □.

The legends indicating the colour code associated with each dynamic element are visible at all times in the bottom right-hand corner of the diagram. The coloured **black** triangle located in the lower area of the graph can be used to change the background scale by clicking with the mouse to the right or left of it.

The top left area of the tool defines the static gain of the system and the parameters of the controller used, based on a linear feedback of the state vector. The **Process parameters** section includes a text box and a slider to modify the value of the static gain of the system whose structure has been selected via the Options menu ($P_i(s)$). The position of the poles and zero (if any) of that transfer function can be modified through the Pole-zero map, where they can be placed in an arbitrary position that completely defines its dynamics.

A symbolic representation of the system is displayed in the right-hand side of the **Process parameters** section in the form of a transfer function $G(s)$ and also in the form of a state space $\dot{x} = Ax + Bu$, $y = Cx + Du$, which can be selected in the Controllable Canonical Form (C) or in the Observable Canonical Form (O) buttons, via exclusionary radio buttons.

Below is the section **Specifications and controller**. On the left hand side, a series of text boxes and sliders can be found, corresponding to the feedback gains K_1 and K_2 and the relative damping factor ζ and undamped natural frequency ω_n which define the position of the closed-loop poles that set the transient behaviour specifications (or alternatively the time constants τ_1 and τ_2 when a closed-loop overdamped system is required). When the tool is set to Analysis mode, only the feedback gains K_1 and K_2 can be modified, since in this mode these gains can be modified by the user to analyze their effect on the location of the closed-loop poles and hence on the time response. As the gains K_1 and K_2 are modified, the associated values of ζ and ω_n also vary, since the position of the closed-loop poles changes. If the application is configured in Control mode, the values of ζ and ω_n can be freely modified to impose a dynamic determined by the location of the closed-loop poles and the values of K_1 and K_2 resulting from the application of Ackermann's formula will be automatically updated.

Application...

To the right of these parameters, a symbolic representation of the internal description of the closed-loop system is displayed in the form $\dot{x} = A_{cl}x$, if the tool is in Free response mode, or in the form $\dot{x} = A_{cl}x + B_{cl}r$ if it is in Forced response mode, corresponding to the schemes represented in Figures 0.3, 0.4 and 0.5. In the Forced response mode, it is possible to include disturbances at the input and activate the Integral Action (Figure 0.6), via the selection box below the symbolic representation of the internal dynamics of the closed-loop system. When activated, a slider and a text box become visible which allows to modify the gain value of the integral term K_i to analyze its effect on the behaviour of the closed-loop system.

As indicated at the beginning of the section, the right area of the tool is dedicated to the graphical representations of the evolution of the system. When through the Options menu the Time response is activated, the upper graph represents the evolution in time of the states $x_1(t)$ and $x_2(t)$ if the tool is in Free response mode and the output $y(t)$ if it is in Forced response mode. When the states are plotted, a legend at the bottom right of the graph includes the colour code used for each state.

When using the Forced response, the graph shows a dashed horizontal line in green colour that represents a constant reference. Both the circle above it and the dashed line are interactive elements that facilitate the change of the step amplitude. Similarly, in the middle area of the graph, two other circles (which by default are usually overlapping and located on the time axis) have been included that can be used to impose the amplitude of a step-shaped load disturbance $d(t)$ at the input of the system and to change the instant at which the disturbance is introduced.

Both in the case of free response and forced response, the lower graph represents the evolution of the control signal $u_c(t) = -Kx(t) + K_r r(t)$ (in free response $r(t) = 0$).

When the Phase plane option is chosen in the Options menu, the whole graphic area on the right is occupied by this representation, including the isoclines whose density can be modulated by means of the slider that becomes visible on the figure. Next to the slider there is also a text box (tmax) which allows the simulation time to be changed (it has the same effect as a change in the background scale in the Time response graphs using the △ symbol).

When working in Free response mode, there is always a regulation of the system to the origin of the phase plane. In the lower left corner of the graph, the eigenvectors of the system are shown (the eigenvalues are given by the position of the poles in the Pole-zero map). The initial conditions can be modified interactively by clicking and dragging the magenta square (□) to the desired position. When the closed-loop poles are real, the associated eigenvectors are drawn in the phase plane. The point representing the steady-state condition is represented by a cyan square (□), which in the case of the Free Response will always be located at the origin of the phase plane.

Application...

If the Forced response mode is selected, the box that is fixed is the magenta one (□), as the initial conditions will always be at the origin and the one that can be moved will be a green box (□), which is linked to the amplitude of the reference that has been determined in the Time response mode. The displacement of that box is equivalent to the change of the reference. As the tool is configured for second order systems using the Controllable (C) or Observable (O) canonical forms, that box will always be linked to one of the axes, which facilitates the analysis. When there is no steady-state error, the green square (□) will coincide with the cyan square (□) that determines the steady-state. In this mode the eigenvectors of the system are not represented.

The black squares (□) located in the lower right and upper left corner of the graph, allow the user to change the state variable that is represented on that axis when clicking on them. The black triangle centred at the bottom of the graph makes it easy to change the background scale when the mouse is clicked to the right or left of it.

Exercises

1. A second-order linear system can be written in general state space as:

$$\begin{aligned}\dot{x}(t) = \frac{dx(t)}{dt} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} x(t) + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(t) \\ y(t) &= [c_1 \quad c_2]\, x(t) \\ x(t) &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\end{aligned}$$

   Determine a linear feedback controller of the state vector of the form:

$$u(t) = -Kx(t) = -K_1 x_1(t) - K_2 x_2(t)$$

   Such that the closed-loop characteristic equation is:

$$J(s) = s^2 + 2\zeta\omega_n s + \omega_n^2 = 0$$

2. Particularize the above exercise to the case of a double integrator:

$$G(s) = \frac{1}{s^2}$$

   For the Free response and Analysis configuration, justify the effect on the system response of varying the K_1 parameter of the controller. Also that of parameter K_2.

   Using the Free response and Control modes, design a controller that produces a closed-loop dynamic with:

   (a) $\zeta = 0.5$ y $\omega_n = 2$ rad/s.

   (b) $\zeta = 1.25$ y $\omega_n = 2$ rad/s.

   Calculate the eigenvalues and eigenvectors in both cases. Also obtain the trajectory in the phase plane for generic initial conditions $x_1(0)$ and $x_2(0)$.

3. A plant is described by the following transfer function:

$$G(s) = \frac{1}{s^2 + \mu^2}$$

   Using the Free response and Control configuration, design a controller with specification $\zeta = 1$ and $\omega_n = n\mu, n = 2, 3, ...$ What effect does increasing the value of n have on the response of the system? Calculate the eigenvalues and eigenvectors and the trajectory in the phase plane for generic initial conditions $x_1(0)$ and $x_2(0)$.

4. A plant is described by the following transfer function:

$$G(s) = \frac{s+1}{s^2+5s+6}$$

With the Free response and Control modes:

(a) Find the vector of feedback gains K that places the poles of this system at the roots of the following equation $s^2 + 2\zeta\omega_n s + \omega_n^2$.

(b) Particularize the result for $\zeta = 0.5$ and $\omega_n = 2$ rad/s.

(c) Analyze the effect of increasing ω_n on the feedback gain vector K.

5. For the plant in exercise 2, choose the Forced response setting and activate the 1DoF (it is active by default). A steady-state error is observed in the response. Calculate this error. Obtain the gain K_r of a two-degree-of-freedom controller that eliminates this steady-state error. Then activate the 2DoF option and verify that you get a response to a setpoint change that has zero error. Apply a constant disturbance $d(t)$ of amplitude 1 to the system input. Analyze the system response and calculate the steady-state error. To eliminate this error due to the load disturbance, activate the Integral action button and analyze the effect of the integral action gain K_i on the system response.

6. Carry out exercise 5 using the plant from exercise 3.

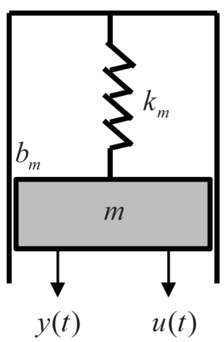
7. Carry out exercise 5 using the plant from exercise 4.

8. For a second order system in normalized form:

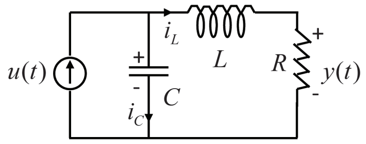
$$\frac{d^2y(t)}{dt^2} + 2\zeta\omega_n \frac{dy(t)}{dt} + \omega_n^2 y(t) = k\omega_n^2 u(t)$$

Calculate a state space description and analyze the phase plane for different values of ζ, ω_n and k.

9. For the spring-mass-damper system shown in the figure, obtain the transfer function and a state-variable model of it (reference [4], pages 163-164), where k_m is the spring constant, b_m is the friction with the wall and m is the supported mass. Consider $y(t)$ to be the position of the mass and choose as state variables the position $y(t)$ and the velocity $\dot{y}(t)$. Particularize the internal description for $m = 1$ kg, $b_m = 20$ N/m/s and $k_m = 100$ N/m. Calculate the roots of the characteristic equation. Perform a system analysis using the interactive tool and a position control using linear feedback of the state vector.



10. For the RLC circuit shown in the figure, obtain the transfer function and a state-variable model of the circuit (reference [4], pages 165-166), where R is the resistance, C is the capacitance and L is the inductance. Consider $y(t)$ to be the voltage across the resistor. Select as state variables the capacitor voltage $v_c(t)$ and the inductor current $i_L(t)$. Particularize the internal description for $R = 3\ \Omega$, $L = 1$ H and $C = 0.5$ F. Calculate the roots of the characteristic equation. Perform a system analysis using the tool and a voltage control using linear feedback of the state vector.



11. Consider the internal description of a pendulum (reference [5], page 466), whose characteristic frequency is ω_0:

$$A = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [1 \quad 0], \quad D = 0$$

Find the control law that places the poles of the closed-loop system at $s = -2\omega_0$. In other words, the aim is to double the natural frequency and increase the relative damping factor of the system from 0 to 1. Demonstrate that the vector of linear gains is $K = [3\omega_0^2 \quad 4\omega_0]$. For a value of $\omega_0 = 1$ rad/s and starting from initial conditions $x_1(0) = 1$ and $x_2(0) = 0$, simulate the behaviour using the interactive tool.

12. Let the transfer function modelling the angular position of a DC motor be ([8], page 39):

$$G(s) = \frac{k}{s(\tau s + 1)}$$

The input is the armature voltage $u(t)$ and the output is the angular position. Choosing as state variables the angular position ($x_1(t)$) and the angular velocity ($x_2(t)$), find the state-space description, as well as the Controllable and Observable Canonical Forms. Integrate the state-space equations and obtain its solution. Particularize for $\tau = 0.25$ s and $k = 1$ rad/V and simulate the behaviour in the interactive tool.

13. Consider the linearized and normalized steering dynamics of a vehicle given by:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \gamma \\ 1 \end{bmatrix}, \quad C = [1 \quad 0], \quad D = 0$$

The full description of the states and equations can be found in [1] (page 6-11). Obtain the controllability matrix. Design a controller that stabilizes the lateral dynamics (y) and follows a reference of value r (the control signal is the steering angle). Calculate the characteristic polynomial of the closed-loop system and impose that the dynamics is second order with a damping factor ζ and a natural frequency ω_n. Particularize for $\gamma = 1$ (although it is actually variable) and different values of ζ and ω_n and simulate the behaviour in the interactive tool.

14. In [1] (page 6-30) the cruise control problem is dealt with in state-space. The linearized dynamics about an equilibrium state (v_e, u_e), where $v(t)$ is the vehicle velocity and $u(t)$ is the throttle position, is given by:

$$\dot{x}(t) = ax(t) - b_g\gamma(t) + b\varpi(t), \quad y(t) = v(t) = x(t) + v_e$$

where $x(t) = v(t) - v_e$, $\varpi(t) = u(t) - u_e$, a, b and b_g are constants that depend on vehicle and road characteristics and $\gamma(t)$ is the curvature of the road (disturbance). We want to control the vehicle speed without error in steady-state, so we need to augment the system dynamics with an integrator. Calculate the gains K, K_i and K_r required to stabilize the system and provide the correct input to achieve the speed reference (assuming $\gamma = 0$). Impose a closed-loop behaviour described by the parameters ζ and ω_n. Particularize in the tool the linearized dynamics for values of $\gamma_e = 0$ rad, $a = -0.0101$, $b = 1.32$, $b_g = g\cos(\gamma_e)$, $\zeta = 0.7$, $\omega_n = 0.5$ rad/s. The velocity reference will be $v_r = 2$ m/s. Enter different values of the perturbation $\gamma(t)$ and analyze the deviation from the equilibrium state.

References

[1] K. J. Åström and R. M. Murray. *Feedback systems: An introduction for scientists and engineers, Second edition*. Princeton University Press, 2014.

[2] Y. Bavafa-Toosi. *Introduction to linear control systems*. Academic Press-Elsevier, 2017.

[3] P. Bolzern, R. Scattolini, and N. Schiavoni. *Fundamentos de control automático (Fundamentals of automatic control)*. McGraw-Hill, 2009.

[4] R. C. Dorf and R. H. Bishop. *Modern control systems. Twelfth edition*. Prentice Hall, 2011.

[5] G. F. Franklin, J. D. Powell, and A. Emani-Naeni. *Feedback control of dynamic systems. Seventh edition*. Pearson, 2015.

[6] J. L. Guzmán, R. Costa-Castelló, M. Berenguel, and S. Dormido. *Automatic control with interactive tools*. Springer, 2021.

[7] K. Ogata. *Modern control engineering. Fifth edition*. Prentice Hall, 2010.

[8] A. Ollero. *Control por computador : Descripción interna y diseño óptimo (Computer controlled systems: internal description and optimal design)*. Marcombo, 1991.

[9] B. Shahian and M. Hassul. *Control system design using MATLAB*. Prentice Hall, 1993.