Chapter 2

BACKGROUND ON NONLINEAR CONTROL SYSTEM & OPTIMIZATION

2.1 INTRODUCTION

In general, all physical systems have some kind of nonlinearities. Sometimes it may even be desirable to introduce a nonlinearity deliberately in order to improve the performance of the system to make its operation secure. This may also result in making the system more economical than is possible with linear components alone.

Nonlinear systems differ from the usual linear system in several ways. Perhaps the most significant of these is the fact that the principle of superposition is not applicable to nonlinear systems i.e. altering the size of the input does not change the shape of the response of the linear system, whereas for the nonlinear system there is a considerable change in both the percentage overshoot and the frequency of oscillation. Similar observations may be made about the stability. In linear systems, stability is a characteristic of the system, independent of the magnitude of the input or the initial conditions. In the case of nonlinear systems, stability may depend on the magnitude of the input as well as the initial condition. Furthermore, application of a sinusoidal input to a stable linear system causes the steady-state output to be a sinusoid of the same frequency, which will, in general, differ from the input in phase and magnitude. In nonlinear systems, on the other hand, the steady-state output may contain harmonics of the input, and in some cases even sub-harmonics may arise.

Other unusual feature of nonlinear system includes such limit cycles and jump phenomena. The former means that independent of the magnitude of the input or initial conditions, the system may produce oscillations of a certain period and amplitude, which may not be sinusoidal. The latter imply jumps in magnitude and phase as the frequency is changed near resonance [50].

2.2 TYPES OF NONLINEARITIES

In most control systems we cannot avoid the presence of certain types of nonlinearities. These can be classified as static or dynamic. A device for which there is a nonlinear relationship between the input, $x(t)$, and the output, $y(t)$, that does not involve a differential equation is called a static nonlinearity. On the other hand, the input and the output may be related through a nonlinear differential equation. Such a device is called a dynamic nonlinearity [51-53]. The basic features of some common nonlinearity are:

- Saturation: it is one of the most common static nonlinearities. A simple example is an amplifier for which the output is proportional to the input only for a limited range of values of the input. As the magnitude of the input exceeds the range, the output approaches a constant.
- Dead zone: In many physical devices the output is zero until the magnitude of the input exceeds a certain value. For example, while developing the mathematical model for a dc servo-motor, any voltage applied to the armature windings will cause the armature to rotate, if the field current is maintained constant. In practice, rotation will result only if the torque produced by the motor is sufficient to overcome the static friction.
- Relays: a relay is often used in control systems as it provides large power amplification relatively economical.

- Friction: A frictional force opposes motion whenever there is sliding contact between mechanical surfaces. The predominant part of the frictional forces is called viscous friction, which is proportional to the relative velocity between the moving surfaces. In addition to viscous friction, there are two components of the total friction forces that are nonlinear. One of them is coulomb friction, which produces a constant force opposing motion. The other is called stiction, which is the force required to initiate the motion, and always greater than the force of coulomb friction.

2.3 STABILITY OF NONLINEAR SYSTEMS

The stability of a nonlinear system depends not only on the physical properties of the system but also on the magnitude and nature of the input as well as the initial conditions. Hence, the study of stability of nonlinear systems is more complicated than for linear systems. Several definitions of stability have been used in the literature of nonlinear systems. Here only the case of an autonomous or unforced system is discussed [54]. Consider an autonomous nonlinear continuous-time system, represented by the state equations

$$\dot{x} = f(x, u) \tag{2.1}$$

$$y = g(x, u) \tag{2.2}$$

where $x(t)$ is the n-dimensional state vector, $u(t)$ is the m-dimensional input vector, $y(t)$ is the p-dimensional output vector, and the vectors f and g are nonlinear functions of x and u. The system is said to be autonomous if the input $u(t)$ is identically zero. For this case, Eq. (2.1) is reduced to

$$\dot{x} = f(x) \tag{2.3}$$

A point of equilibrium is obtained for any value of the vector x that $\dot{x} = 0$. In general a nonlinear system may have many points of equilibrium. Some of these may be points of stable equilibrium, while others may be points of unstable equilibrium. A good example is the bi-stable multi-vibrator, an electronics circuit with three states of equilibrium, two of which are stable and one stable. Consequently, it is necessary to examine stability at each point of equilibrium. It is common practice to transform coordinates in the state space so that the origin becomes the point of equilibrium. This is convenient for examining local stability and can be done for each point of equilibrium [55].

Let us now consider a hyper-sphere of finite radius surrounding the origin of the state space (the point of equilibrium), that is the set of points described by the equation

$$x_1^2 + x_2^2 + \cdots + x_n^2 = R^2 \tag{2.4}$$

in the n-dimensional state space. Let this region be denoted by $S(R)$.

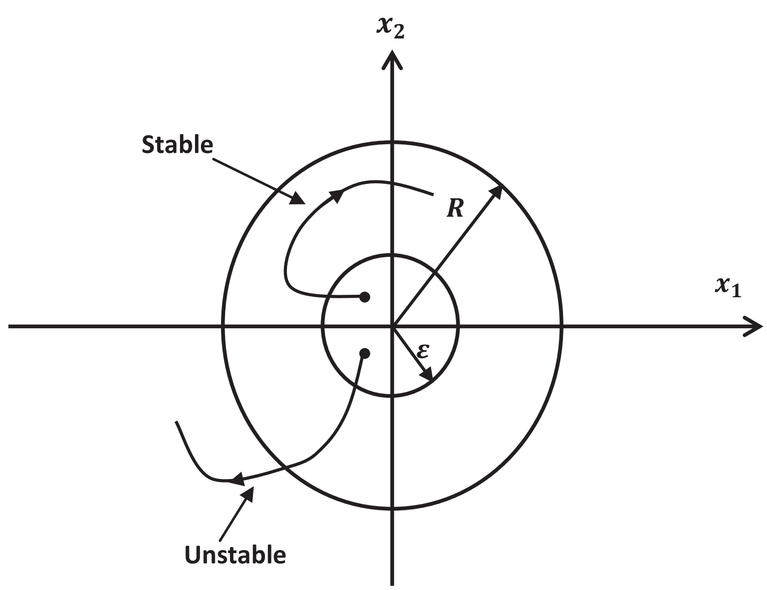

Figure 2.1 Stability in the sense of Lyapunov

The system is said to be *stable* in the sense of Lyapunov if there exist a region $S(\varepsilon)$ such that a trajectory starting from any point $x(0)$ in this region does not go outside the region $S(R)$ as shown in Figure 2.1 for the two-dimensional case.

The system is said to be *asymptotically stable* if there exist a $\delta > 0$ such that the trajectory starting from any point $x(0)$ within $S(\delta)$ does not leave $S(R)$ at any time and finally returns to the origin as shown in Figure 2.2.

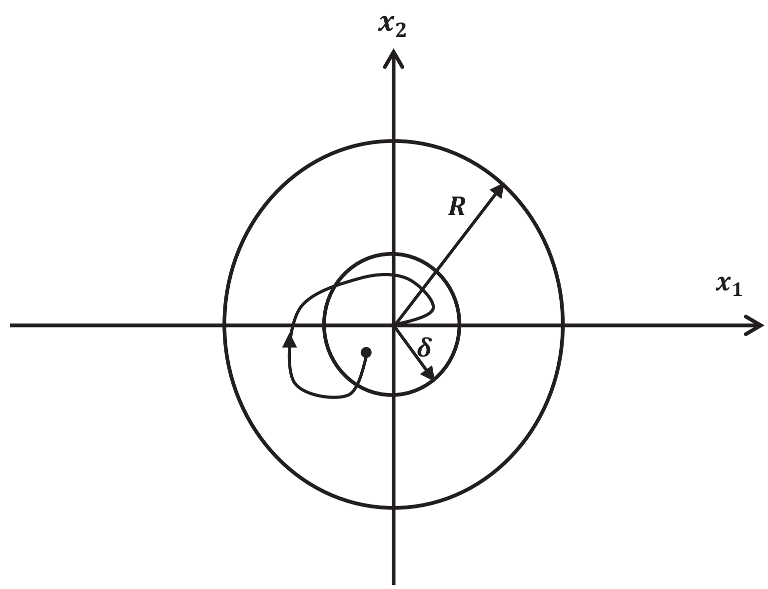

Figure 2.2 Asymptotic stability in the sense of Lyapunov

The system is said to be *monotonically stable* if it is asymptotically stable and the distance of the state from the origin decreases monotonically with time as shown in Figure 2.3.

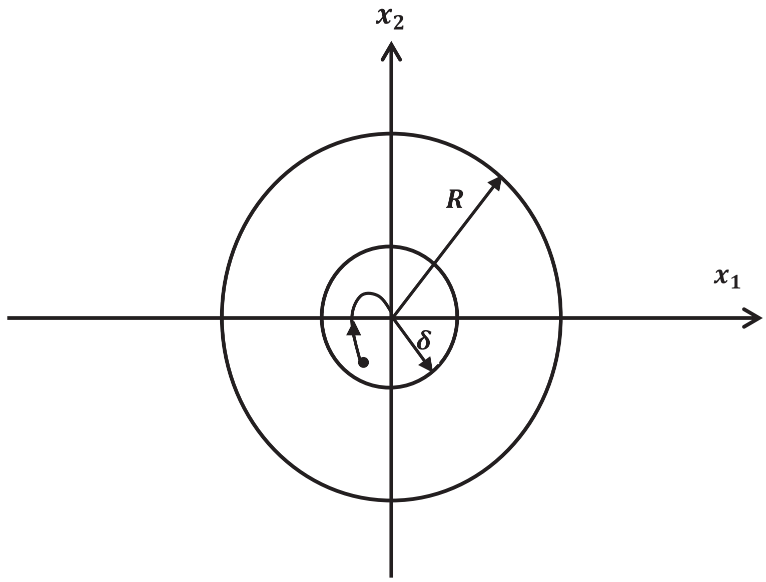

Figure 2.3 Monotonic stability in the sense of Lyapunov

A system is said to be globally stable if the regions $S(\delta)$ and $S(R)$ extend to infinity. A system is said to be locally stable if the region $S(\delta)$ is small and when subjected to small perturbations the state remains within the small specified region $S(R)$.

2.4 LINEARIZATION

Linearization is based on the Taylor series expansion of a nonlinear function about an operating point. For example, consider a nonlinear function, $f(x)$. It can be written as

$$f(x) = f(x_0) + \frac{df}{dx}\bigg|_{x=x_0}(x - x_0) + \frac{d^2f}{dx^2}\bigg|_{x=x_0}\frac{(x-x_0)^2}{2!} + \cdots H.O.T \qquad (2.5)$$

We get a linear approximation of Eq. (2.5) if we ignore all terms except the first two. Clearly, this will be a good approximation if either $(x - x_0)$ is very small, or the higher order derivatives of f are very small. This is the main idea behind the incremental linear models used for the analysis of electronic circuits [56].

We shall now generalize this to the case of the state equations for nonlinear systems. Assuming that the dimension of x is n, therefore Eq. (2.1) can be re-written as

$$\dot{x} = f(x,u) = \begin{bmatrix} f_1(x,u) \\ f_2(x,u) \\ \vdots \\ f_n(x,u) \end{bmatrix} \tag{2.6}$$

Ignoring the higher-order terms in Taylor series expansion of this vector differential equation leads to the linearized model (assuming $x_0 = 0$, that is, the coordinates have been transformed to make the origin the point of equilibrium, and $x_0 = 0$)

$$\dot{x} = Ax + Bu \tag{2.7}$$

where

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \tag{2.8}$$

and

$$B = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} & \cdots & \frac{\partial f_1}{\partial u_m} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} & \cdots & \frac{\partial f_2}{\partial u_m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial u_1} & \frac{\partial f_n}{\partial u_2} & \cdots & \frac{\partial f_n}{\partial u_m} \end{bmatrix} \tag{2.9}$$

A and B are said to be Jacobian matrices. Again, this linear model will be valid only for small deviations around the equilibrium point. Nevertheless, it can be used for investigating local stability around the equilibrium point by simply applying the Routh criterion to the characteristic polynomial of A.

2.5 STABILITY ANALYSIS USING LYAPUNOV'S DIRECT METHOD

The simple stability criteria developed for linear systems are not applicable to linear systems since the concept of roots of a characteristic polynomial are no longer valid. As stated in section 2.3, many different classes of stability have been defined for nonlinear systems. Now, the stability is to be discussed in the sense of Lyapunov.

Consider a region ε in the state space enclosing an equilibrium point x_0. Then this is point of stable equilibrium provided that there is a region $\delta(\varepsilon)$ contained within ε such that any trajectory starting in the region δ does not leave the region ε.

With this definition it is not necessary that the trajectory approach to the equilibrium point. It is only required that the trajectory be within the region ε. This permits the existence of oscillations of limited amplitude, like limit cycles.

Lyapunov's direct method provides a means for determining the stability of a system without actually solving for the trajectories in the state space. It is based on the simple concept that the energy stored in a stable system cannot increase with time. Given a set of nonlinear state equations, one first defines a scalar function $V(x)$ that has properties similar to energy and then examines its derivatives with respect to time.

THEOREM 2.1

A system described by $\dot{x} = fx$ is asymptotically stable in the vicinity of the equilibrium point at the origin of the state space if there exist a scalar function V such that

1. $V(x)$ is continuous and has continuous first partial derivatives at the origin.
2. $V(x) > 0$ for $x \neq 0$ and $V(0) = 0$.

3. $\dot{V}(x) < 0$ for all $x \neq 0$.

These conditions are *sufficient but not necessary* for stability. $V(x)$ is often called a Lyapunov function.

THEOREM 2.2

A system described by $\dot{x} = fx$ is unstable in a region Ω about the equilibrium at the origin of the state space if there exist a scalar function V such that

1. $V(x)$ is continuous and has continuous first partial derivatives at the origin.
2. $V(x) \geq 0$ for $x \neq 0$ and $V(0) = 0$.
3. $\dot{V}(x) > 0$ for all $x \neq 0$.

Again these conditions are *sufficient but not necessary.*

2.6 CONTEMPORARY METHODS OF OPTIMIZATION

In recent years, some optimization methods that are conceptually different from the traditional mathematical programming techniques have been developed. These methods are labeled as modern or nontraditional methods of optimization. Most of these methods are based on certain characteristics and behavior of biological, molecular, swarm of insects, and neurobiological systems. The following methods are:

- Genetic algorithms
- Simulated Annealing
- Particle swarm optimization
- Ant colony optimization
- Grey Wolf Optimizer

- Teaching Learning Based Optimization
- Fuzzy Logic Control

Most of these methods have been developed only in recent years and are emerging as popular methods for the solution of complex engineering problems. Most require only the function values (and not the derivatives). The genetic algorithms are based on the principles of natural genetics and natural selection [57-58]. Simulated annealing is based on the simulation of thermal annealing of critically heated solids. Both genetic algorithms and simulated annealing are stochastic methods that can find the global minimum with a high probability and are naturally applicable for the solution of discrete optimization problems. The particle swarm optimization is based on the behavior of a colony of living things, such as a swarm of insects, a flock of birds, or a school of fish. Ant colony optimization is based on the cooperative behavior of real ant colonies, which are able to find the shortest path from their nest to a food source. The grey wolf optimizer is based on the hunting behavior of wolves and whereas, teaching learning based optimization is based on the outcome of effective learning amongst the learners. The latter two techniques are developed in the recent years and gained popularity amongst the former techniques in providing effective and efficient solution in coping up all the design requirements.

In many practical systems, the objective function, constraints, and the design data are known only in vague and linguistic terms. Fuzzy optimization methods have been developed for solving such problems [59-60].

2.6.1 Process of Optimization

The evolutionary optimization process is mainly depends on identifying the parameters that needs to be optimized. Reducing the number of parameters usually reduces the complexity of the optimization task, thereby achieving a faster convergence of the optimization algorithm. However, teaching learning based optimization can typically handle a large number of parameters efficiently. Also, by constructing too many of the parameters may result in eliminating the optimal solution. Therefore, a careful trade off exists between the complexity of the optimization task and convergence of the optimization technique. The optimization process mainly covers following steps as shown in Figure 2.4.

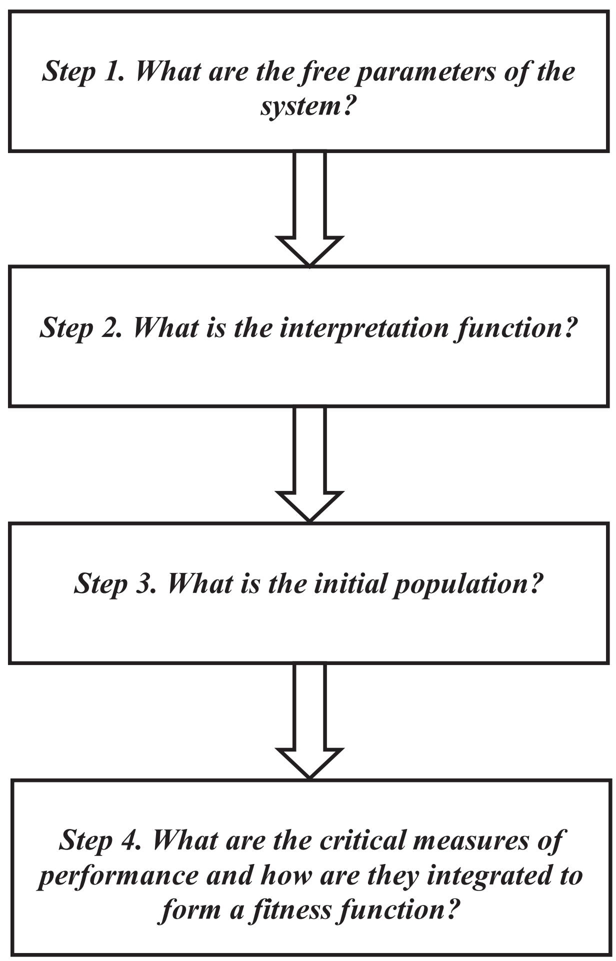

Figure 2.4 Four Stages of Optimization Process

2.7 FUZZY PARAMETERS

Fuzzy expert knowledge can be divided into two basic components: Domain knowledge and Meta Knowledge. The Domain knowledge is generally the conscious operating knowledge about a particular system such as the membership functions and the fuzzy rule set. The Meta knowledge is the unconscious knowledge that is also needed to completely define a fuzzy system such as the mechanism of executing the fuzzy rules, methods of implication, rule aggregation, and defuzzification. Most of the existing methods in evolutionary fuzzy systems attempt to optimize parameters of the domain knowledge only (namely membership functions and rule set) while ignoring the effect of meta-knowledge. Consequently, there are 4 basic methods of optimization as follows

- Automatic optimization of membership functions while there is a fixed and known rule set;
- Automatic selection of the rule set with fixed membership functions;
- Optimization of both the membership functions and rule set in two steps. First selecting the optimal rule set with fixed known membership functions and then tuning the membership functions with the resulting rule set; and
- Simultaneous optimization of fuzzy rule set and membership functions.

Note that the number of membership functions or rules can also be optimized in the algorithm. There may be various reasons for a method to be selected. Some of those advantages and disadvantage are mentioned below:

- Since the rule set and membership functions are codependent, they should be defined simultaneously. This can lead to more optimal solutions. [61, 62];

- Since the performance of a fuzzy system is more dependent on fuzzy rules rather than membership functions, fine tuning of the fuzzy system is better possible by tuning of membership functions. So it seems that it is better first to select the optimal rule set (coarse tuning) and then tune the membership functions (third method);
- Even though various methods exist to encode both the rule base and membership functions, such encoding can have several potential difficulties. In addition to the level of complexity and large number of optimization parameters, the problem of competing conventions may arise and the landscape may unnecessarily become multi-modal.