

Internal Model

Abstract. In this chapter we are mainly concerned with modelling. What kind of model to use, where to use it and how is the model implemented? A procedure to decompose the model of an integrative process is presented and the free and forced solution of linear dynamic process is reviewed.

Keywords: modelling, realigned model, independent model, disturbance, decomposition, free solution, forced solution

2.1 Why Is Prediction Necessary?

The primary concern of any financier investing in the construction of an industrial plant is to maximize the return on the financial investment. Consequently, a systems engineer will always strive to design a plant that functions on, or near, the operational boundaries dictated by the safety, technological and budgetary limitations. Ultimately, it is the operational constraints that decide the success or failure of a project. Typically, the constraints involved would arise from:

- **actuators:** *e.g.*, maximum fluid flow, heating power;
- **process limitations:** *e.g.*, maximum temperature gradient, distillation column flooding.

If “tight” process control is to be achieved then the constraints involved must be continuously respected in a dynamic fashion. As we shall see, this will lead to a requirement for some mechanism of predicting the future behaviour of the process.

Returning to the automotive example of Chapter 1, assume some incident occurs that requires the full application of the braking system. From the instant that maximum brakes are applied, the future is completely defined, *i.e.*, the stopping time and distance are fixed by factors such as the car dynamics and road conditions – both of which are out of the driver’s control.

However, common sense would dictate that a more prudent approach would be to anticipate the situation and implement a controlled braking scheme commencing an appropriate distance from the potential hazard. The future prediction of the

process response, *i.e.*, the stopping time and distance of the vehicle in this case, is made with the aid of a mathematical model embedded in the controller from which it derives the name “internal model”.

On the other hand, if infinite deceleration (no constraint) were possible, such anticipation would not be required, as braking immediately prior to the obstacle would suffice! Thus, there is a direct causal link between “constraints” and “prediction”: **Constraints → Prediction**.

In the real world, investors demand optimum profit for their investment that in turn requires optimisation of systems with inevitable constraints thus giving the overall connection between profit and prediction in the following: **Profit → Optimisation → Constraints → Prediction**.

Predictive control possesses many interesting features. However, above all, it is its natural ability to take constraints into account that led to the appearance of predictive control in the 1970s, in the oil industry in particular.

2.2 Model Types

Figure 2.1 shows a first-order process, with a gain K_p and time constant T , subject to an input e . The response of the process s_p is represented, in difference equation form, by:

$$s_p(n) = a_p s_p(n-1) + b_p K_p e(n-1), \text{ where } a_p = e^{-\frac{T_s}{T}} \text{ and } b_p = 1 - a_p,$$

where T_s is the sampling time. This may be modelled by a first-order system with a gain K_m and a time constant T_m as:

$$s_m(n) = a_m s^*(n-1) + b_m K_m e(n-1), \text{ where } a_m = e^{-\frac{T_s}{T_m}} \text{ and } b_m = 1 - a_m.$$

Note that s^* gives rise to two modelling approaches. The first approach uses a model output that is realigned with past measured or estimated process values:

$$s^*(n-1) = s_p(n-1).$$

The second approach adopts an independent process model, where a common input is supplied to both the process and model:

$$s^*(n-1) = s_m(n-1).$$

Both approaches are of use and each possesses its own individual characteristics.

2.2.1 Realigned Model

The above equations give rise to a model of the form:

$$s_m(n) = a_m s_p(n-1) + b_m K_m e(n-1),$$

where $a_m = e^{-\frac{T_s}{T_m}}$, $b_m = 1 - a_m$ and e is the input. Intuitively, we know that using the process output s_p to re-adjust the model would yield the best prediction. This operation is simplified if a difference equation form is used:

$$s_m(n-i) = s_p(n-i),$$

i.e., all past outputs of the process are measured and stored in memory. However, this representation has some limitations. If the order of the system is greater than one, and if the poles are not stable, some numerical imprecision may result and consequently the choice of sampling period becomes critical.

If a state-space representation is used, these difficulties disappear and it becomes necessary to use an “observer”, *i.e.*, a mathematical procedure that reconstructs the unmeasured state variables. But, the application of this technique is not straightforward and, as a consequence, has experienced a reluctant acceptance in the production industry.

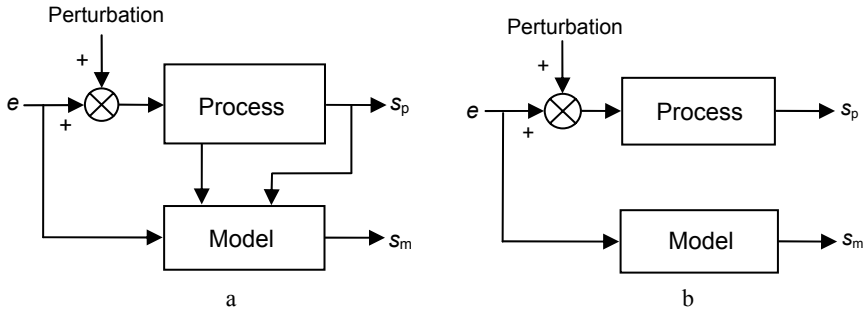


Figure 2.1. **a** Realigned and **b** independent models

2.2.2 Independent Model

The independent model is defined by:

$$s_m(n) = a_m s_m(n-1) + b_m k_m e(n-1).$$

In this configuration the process and independent models are supplied by the same input variable. This value is always available since it is calculated by the controller. As the process may be subjected to unknown disturbances, the resulting outputs $s_p(n)$ and $s_m(n)$ may be very different. If the disturbances are constant,

the process and model outputs will evolve in parallel. Thus, the model is used to calculate a prediction of the *increment* of the process output and not to calculate the absolute response of the process subjected to a particular input.

The advantage of such an approach is that any type of model may be used, *e.g.*, mathematical, logical, or look-up table. The only requirement is that the model is capable of answering the question “What will be the resulting output increment if the process is subjected to a known input?”

Unfortunately, this approach is not valid for unstable models. In this case, using an independent model configuration as the internal model within the regulator produces an output that is open loop and unstable. However, the process may be stabilised using a stable regulator in a feedback loop. We will see later how to solve this problem.

These two approaches (realigned or independent models) result in the following identification strategies, *i.e.*, estimation of the model parameters:

- The realigned model method is equivalent to system identification using the least squares technique where the explicit variables are physical measurements derived from the process. If these variables are noisy, the identified parameters will be biased in general.
- The independent model method [2] uses only the process input and the error between the measured process and model predicted outputs, *i.e.*, the process-model error (PME). In this case the parameters of the model are generally non-biased (see Figure 2.2). Note that the distance criteria for the least-squares and model methods are usually different.

If a convolution-based model is used, the two approaches are identical as past process outputs are not used to calculate future outputs:

$$s_m(n) = a_1 e(n-1) + a_2(n-2) + a_3 e(n-3) + \dots + a_N e(n-N) .$$

This form of representation is not appropriate for unstable systems since N is finite but it is used frequently in the case of multivariable systems.

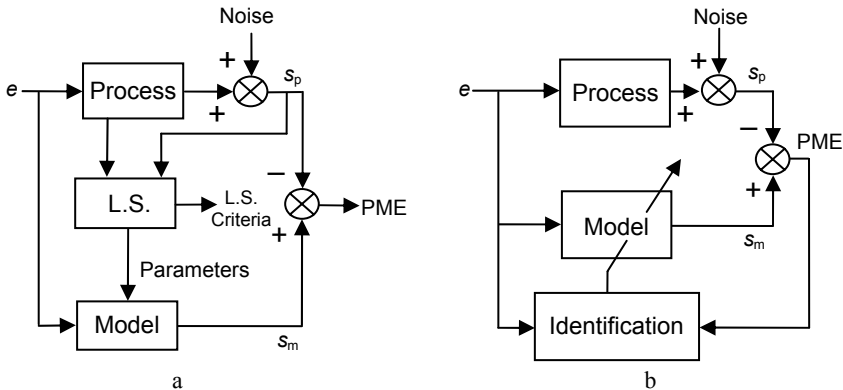


Figure 2.2. Estimation using the **a** realigned least-squares (LS) and **b** independent model methods

2.3 Decomposition of Unstable or Non-asymptotically Stable Systems

The previous comparison of the independent model concept with that of system identification gives an indication of the potential benefits that may be derived from this independent model approach, *i.e.*, absence of bias, universal application and simplicity. The case of unstable or non-minimum phase systems, *e.g.*, integrating and unstable, must be approached with caution. But first, consider the case of taking a measurable, external disturbance into account.

2.3.1 Measurable Disturbance Compensation

The real time, mathematical model implemented within the controller permits the calculation of future responses to potential manipulated variable inputs. This same technique may be used to compensate for the effects of an external, measured disturbance that is not controllable (see Figure 2.3).

Remark 2.1. In Figure 2.3, the regulated (or controlled) variable is referred to as CV and the manipulated variable is denoted by MV. These two notations will be used frequently throughout this book.

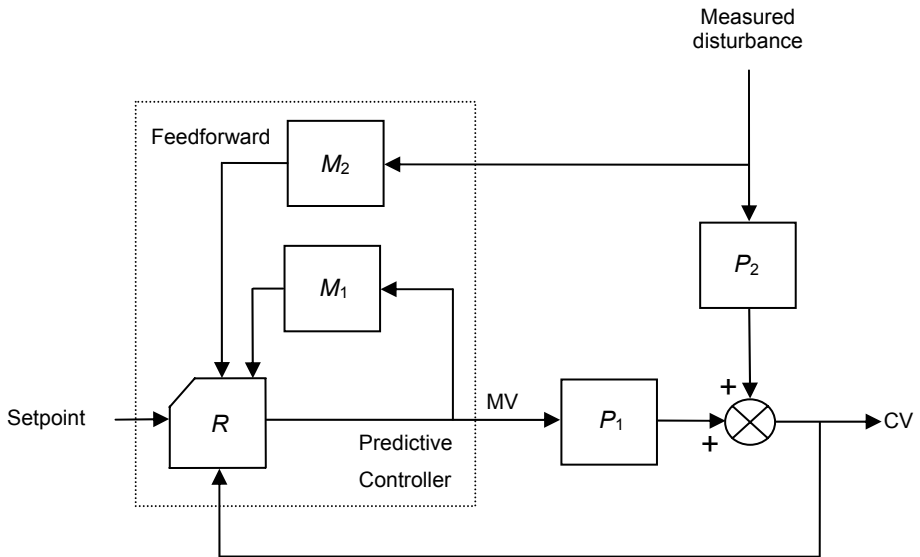


Figure 2.3. Measured disturbance compensation

The procedure consists of making future predictions of the manipulated variable and the measured disturbance using a procedure discussed in Section 2.4. An output increment is then calculated and subsequently taken into account in the control calculation (see Chapter 4).

A practical example of such compensation can be easily found. For example, assume we wish to regulate the interior temperature of a house heated by an electric radiator. The exterior temperature varies randomly but is constantly monitored. A transfer function relating the exterior and interior temperatures may be determined using system-identification techniques. This information may then be used to predict the action required to counteract the influence of the exterior temperature variations at some fixed point in the future. However, any compensatory actions must react faster than any exterior temperature variations.

This open-loop procedure should be used systematically as it does not present any risk to stability and it puts all the necessary information to best use.

2.3.2 Decomposition

For reasons of implementation and initialisation, it is necessary that the closed loop and the regulator be stable. This implies that the internal models of the regulator must also be stable. If the process is an integrating or unstable system the output, for a given steady-state, non-zero input, would escalate over time. This issue is dealt with by decomposing the unstable model into two stable processes. The first model M_1 is supplied by the manipulated variable MV. The second model M_2 effectively takes the form of a compensated input and is supplied by the output of the physical process. The process models may be represented by their continuous or discrete transfer functions.

The key point to note here is that M_0 is an unstable model. A stable equivalent, in the form of the M_1 and M_2 combination, is sought in order to circumvent the inability of any subsequently developed regulator to control the unstable plant.

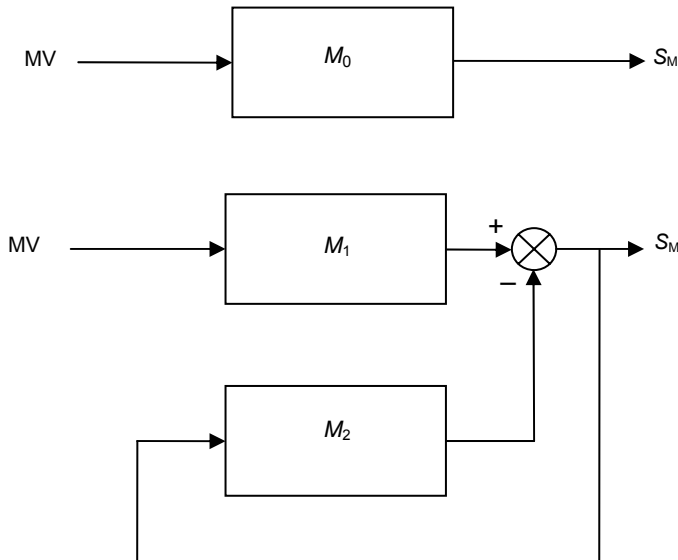


Figure 2.4. Decomposition principle

Referring to Figure 2.4: $S_M = M_1 MV - M_2 S_M$ and re-organising gives:

$S_M = \frac{M_1}{1 + M_2} MV$. Assume we have two transfer functions, M_1 and M_2 , such that

$M_0 = \frac{M_1}{1 + M_2}$, where both M_1 and M_2 are asymptotically stable. In the nominal

case, if the process output is taken into account in the form of a compensator, the blocks may be readily identified (see Figure 2.5).

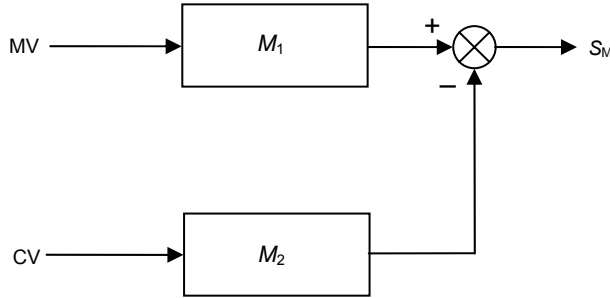


Figure 2.5. Decomposition formulation

Consider the example of a non-asymptotically stable system with an integrator represented by the continuous transfer function $H(s) = K/s$ (where s is the Laplace operator) as illustrated in Figure 2.6.

The plant is decomposed using a first-order transfer function whose dynamics resemble those of the desired closed-loop behaviour:

$$S_M = \frac{KT}{1+sT} MV - \frac{-1}{1+sT} S_M \rightarrow S_M = \frac{KT}{sT} MV = \frac{K}{s} MV.$$

The case where the process is unstable with an “inverse response”² should be dealt with cautiously (see Chapter 9). As we shall see, a satisfactory response is achieved without any particular difficulty when a coincidence point is placed beyond the characteristic ‘dip’ generally associated with the initial response of such systems. A typical example would be the temperature response of a chemical reactor due to the introduction of a cold reagent.

In Chapter 10 we shall see that this decomposition procedure³ can be used beneficially in many practical situations. It should be noted that the decomposition principle is, in fact, a compromise solution between the realigned and independent model approaches.

² A non-minimum phase system with an “unstable zero”.

³ The benefit of this approach is that it only requires external measurements of the system model and process (input and output) and makes no assumptions about the mathematical representation, state observation or system characteristics.

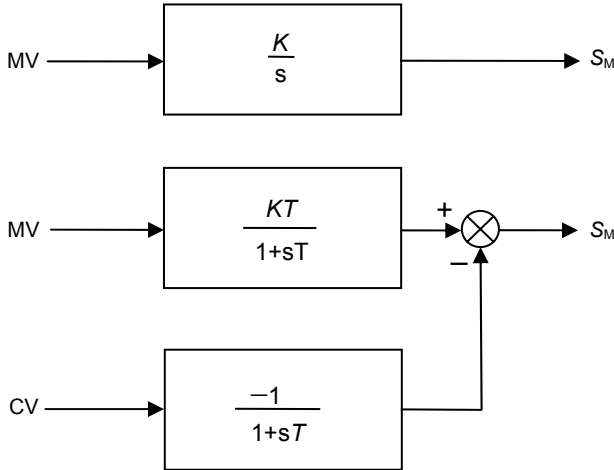


Figure 2.6. Decomposition of a process with a single integrator

2.4 Prediction

The key to any model-based controller lies in its ability to predict the process response using a model that may be physically realised, *i.e.*, a causal model. We will return to the issues of model and system identification in Chapter 7. In fact, determining the response of a process over time is equivalent to the classical problem of solving differential or difference equations. For convenience, we recall the fundamental concepts used in solving such equations.

The solution of a linear differential or finite-difference equation, from the instant $t = 0$ to the present time, consists of two terms; the free solution and the forced solution.

2.4.1 The Free Solution $S_L(t)$

The free solution (also referred to as the complementary, natural or homogeneous solution) is defined as the output when the input $e(t)$ is zero for $t > 0$, but was non-zero in the past, *i.e.*, the initial value of the process at time $t = 0$ is non-zero. This solution represents the output when no further external stimulus is applied. If the system is asymptotically stable the output will eventually decay to a zero state, as illustrated in Figure 2.7.

Example 2.1. Consider the behaviour of the interior temperature of an oven when the gas flow is cut-off. The temperature will decrease, at a rate broadly dictated by the level of insulation in the oven, towards the ambient temperature.

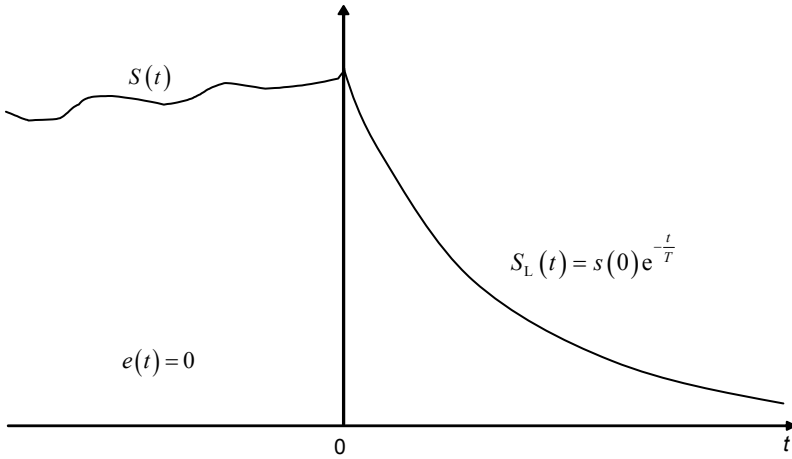


Figure 2.7. Free output response of a first-order system

2.4.2 The Forced Solution $S_F(t)$

The forced solution is also referred to as the particular, forced or inhomogeneous solution. The forced solution makes the opposite assumption to that of the free solution. This implies that all past signals, both input and output, are zero. The future, non-zero input signal is known or determined, and the future output is calculated by the model using only this non-zero input signal.

If the system is linear, the superposition theorem applies, and the future output of the process $s(t)$ is the sum of the free S_L and forced S_F responses:

$$s(t) = S_L(t) + S_F(t) .$$

Thus, the future depends both on the past process responses and its future input. The past is fixed, whereas the future input depends on the operator. Consider a first-order, linear process with a gain K and a time constant T .

Assume, for simplicity, that the future input is constant $e(t) = e_0$ (see Figure 2.8), thus,

$$S_F(t) = Ke_0 \left(1 - e^{-\frac{t}{T}} \right) .$$

The future behaviour is then the sum of the two outputs:

$$s(t) = s(0)e^{-\frac{t}{T}} + Ke_0 \left(1 - e^{-\frac{t}{T}} \right) .$$

Already we can begin to see how a simple control signal such as $MV(n) = e(n) = e_0$ allows the system to be regulated. The first term on the right-hand side takes the effects of the past into account. This, combined with the second term, which includes the effects of the regulator, determines the future output of the system.

As the first term is pre-defined by the past actions we can conclude that the only way to influence the future output is by manipulating the second term. Thus, from an intuitive standpoint, we can see that if we can somehow predict the required control signal necessary to generate the desired change in the forced term, and hence the objective $s(t)$, we have the regulator we require. We will see in the next chapter that it is beneficial to dictate the manner in which the final objective is reached. The name given to this manipulated objective path is the *reference trajectory*.

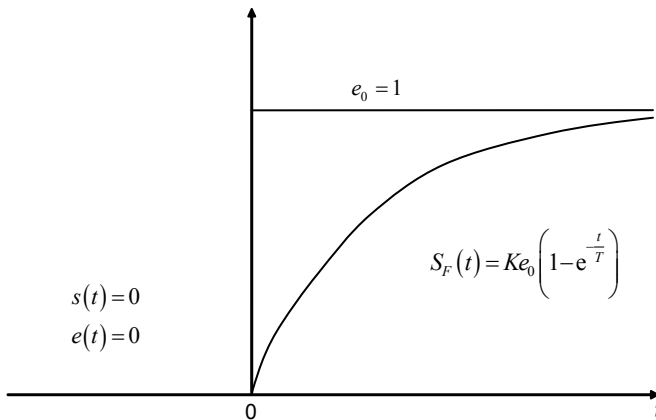


Figure 2.8. Forced output response to a unit step input

2.5 Summary

- Two generic model types are used in the formulation of PFC:
 1. **Realigned:** The input of the model is the process input and the state of the model is realigned with the measured or estimated state of the process.
 2. **Independent:** The input of the model is the input of the process and the state of the model is not realigned with the process state. In this case the only input is the actual MV and so the model output may be different from the process output.
- Integrating processes may be represented using an independent model supplied by the applied process MV. The model is only supplied with the measured output of the process.

- Realigned and independent models are implemented in the control computer.
- The solution of a linear differential equation is composed of two terms:
 1. **Free mode:** Response of the model to a zero input;
 2. **Forced mode:** Response of the model starting from zero state and subjected to the required future input.



<http://www.springer.com/978-1-84882-492-8>

Predictive Functional Control
Principles and Industrial Applications
Richalet, J.; O'Donovan, D.
2009, XXII, 224 p., Hardcover
ISBN: 978-1-84882-492-8