Homeostasis, Linearization, and Integral Control

Ben

February 3, 2026

1 Homeostasis

Let's assume we have a system working to maintain a setpoint. The system is designed to keep some signal as close to constant as possible. I'll call this hopefully constant signal the reguland, y. The system experiences exogenous disturbances, d that might change its course and disrupt the constancy of the reguland. The system has a state, x, a collection of variables that at each fixed time predicts the system's future. The state and disturbance at a particular time determine the next state according to some rule:

$$x_{t+1} = F(x_t, d_t)\,.$$

The reguland can be measured, and is a manifestation of the current system state. That is, we can write the value of reguland as a function of the current state.

$$y_t = G(x_t)\,.$$

For any constant value of the disturbance, we'd like conditions that guarantee the system settles to a state where the reguland equals a specified setpoint level, K. No matter what the disturbance, the system has to converge to the same value of the reguland, but this might require a different state value for every value of the disturbance. Denote the steady-state value of the state by $x(d)$. The steady-state conditions are then

$$\begin{aligned} x(d) &= F(x(d), d) \\ K &= G(x(d)) \end{aligned}$$

The goal of control analysis is to find conditions on the maps F and G that guarantee such a steady state is possible and robust to further disturbances. One of the most basic analyses uses calculus. If we assume that F and G are differentiable, then the implicit function theorem guarantees there is a value of the state that maintains the setpoint. This state value is one determined by the value of the disturbance and can be computed from the derivatives of F and G:

$$(I - \nabla_x F(x(d), d))\nabla_d x(d) - \nabla_u F(x(d), u) = 0 \tag{1a}$$

$$\nabla_x G(x(d))\nabla_d x(d) = 0 \tag{1b}$$

From these equations, we can solve for $\nabla_d x(d)$, and then solve a differential equation to find $x(d)$.

These derivatives also tell us something about the system dynamics near the setpoint. If we start at a fixed point associated with a "normal" environmental disturbance, and nature slightly changes, we can approximate the convergence to the new fixed point using *linearization*. Linearization assumes the dynamics are well approximated by the linear model defined by the Taylor series approximations of F and G at the fixed point. If we assume that nature perturbs the disturbance from d to $d+\Delta d$, then the system will evolve by approximately linear dynamics. To see what these dynamics look like, let's define some auxillary matrices:

$$\begin{aligned} A_d &= \nabla_x F(x(d), d) \\ B_d &= \nabla_u F(x(d), u) \\ C_d &= \nabla_x G(x(d)) \end{aligned}$$

Then we can approximate the dynamics by

$$\begin{aligned} x_{t+1} &= F(x_t, d + \Delta d) \\ &\approx F(x(d), d) + A_d(x_t - x(d)) + B_d \Delta d \\ &= x(d) + A_d(x_t - x(d)) + B_d \Delta d \end{aligned}$$

Defining $\Delta x = x - x(d)$, we thus have the dynamics

$$\Delta x_{t+1} = A_d \Delta x_t + B_d \Delta d \,.$$

Defining $\Delta y = y - K$, a similar calculation reveals

$$\Delta y_t = C_d \Delta x_t$$

These local dynamics are *linear*. In this linear system, the setpoint condition seeks $\Delta y = 0$. This means that x_t must also converge to zero, which implies that A_d must be a stable matrix. That is, the spectral radius of A_d has to be less than one. The linearization analysis tells us another fact about the derivatives of F near the fixed point.

The idea of using static local information to inform temporal behavior is called Lyapunov's direct method or Lyapunov's first method. We transform the problem of general nonlinear control into one of local linear algebra. The linear algebra tells us interesting and surprising things that are generally actionable in engineering design. We just have to be careful to know the limits of these analyses.

We can derive another curious fact from Equations (1a) and (1b). Namely, that there has to exist a vector q such that

$$\begin{bmatrix} q \\ 1 \end{bmatrix}^T \begin{bmatrix} \nabla_x F(x(d), d) & \nabla_u F(x(d), u) \\ \nabla_x G(x(d)) & 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} .$$

If we set $s_t = q^T \Delta x_t$, we find that

$$s_{t+1} = s_t + \Delta y_t \,.$$

That is, the linearized system has a state component that is *integrating* the error in the reguland.

$$s_T = s_0 + \sum_{t=0}^{T-1} \Delta y_t \,.$$

The only way for the variable s to converge to a steady state is for Δy_t to converge to zero. That is, locally, there is always a component of the state that is integrating the error. Control theorists call this integral control, and we'll talk more about it next week. Integral control is an essential tool to maintain setpoints in control design. It turns out that it is in fact a necessary part of any setpoint regulation.

2 Integral Control is Necessary and Sufficient for Setpoint Regulation in Linear Systems

Suppose we have the following linear system:

$$x_{t+1} = Ax_t + Bu_t \tag{2}$$

$$y_t = Cx_t \tag{3}$$

And we know that for *any* constant input $u_t = D$, y_t converges to zero and x_t converges to some $x(D)$. Then there is some linear function of the state

$$s_t = \alpha^T x_t \tag{4}$$

such that

$$s_{t+1} = s_t - y_t \tag{5}$$

That is, s_t is the integral of y_t.

This is a fancy way of saying that the closed loop map has a zero at 1. But it also shows that integral control is necessary for perfect adaptation in linear systems. And we have the fun proof of integral action here: that signal s must converge to a constant, but that means that y must converge to zero.

There is probably a simpler proof, but here's my argument. First, A must be stable, because otherwise the system would not converge to a fixed point. Now assume x_t converges, then

$$x(D) = Ax(D) + BD$$

Which means

$$x(D) = (I - A)^{-1}BD$$

y_t converging to zero implies $C(I - A)^{-1}B = 0$. In other words, the system has a zero at $z = 1$. Set $s_t = C(I - A)^{-1}x_t$. Then

$$s_{t+1} - s_t = C(I - A)^{-1}(Ax_t + Bu_t) - C(I - A)^{-1}x_t = -Cx_t = -y_t \,. \tag{6}$$

Note that the impulse response is

$$g_t = CA^tB$$

The above derivation shows

$$\sum_{t=0}^{\infty} g_t = \sum_{t=0}^{\infty} CA^tB = C(I - A)^{-1}B = 0$$