INTERACTIVE CONTROL ENGINEERING
Principles and Applications

Volume 11 Control in Petroleum Refining

CHAPTER 1

Introduction: Throughput, Quality and Limits in a Refinery

Pages 1–22

Edited by John D. Hedengren

How to cite this chapter. "Introduction: Throughput, Quality and Limits in a Refinery," in *Control in Petroleum Refining*, John D. Hedengren, Ed., vol. 11 of *Interactive Control Engineering: Principles and Applications*. Advanced Process Solutions, LLC, 2026, ch. 1, pp. 1–22.

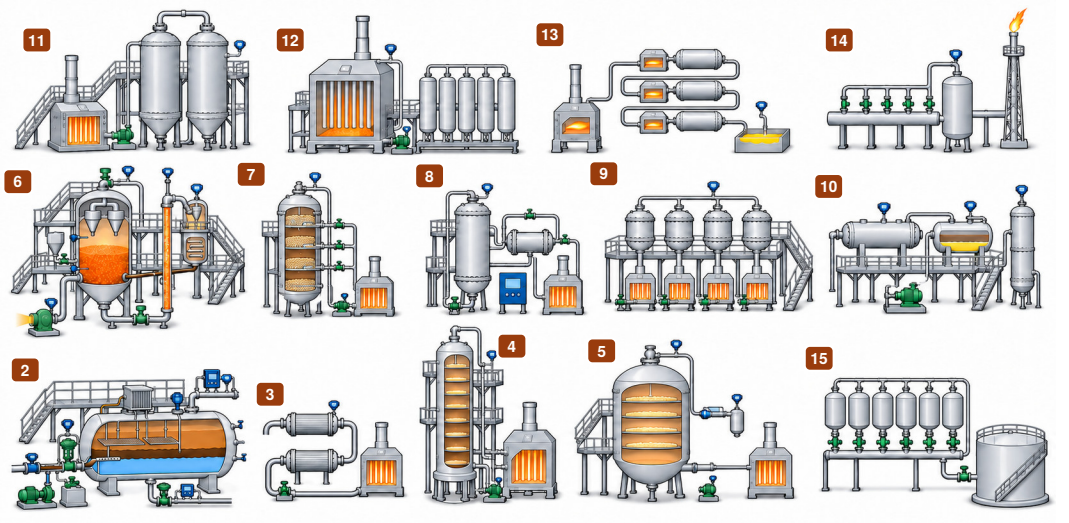
The page numbers in this file are those of the printed volume.

The volume online: contents, other chapters and citation records
https://apmonitor.com/apps/books/v11-refining/

Copyright © 2026 Advanced Process Solutions, LLC. All rights reserved. Produced by Advanced Process Solutions, LLC.

CHAPTER 1

Introduction: Throughput, Quality and Limits in a Refinery



The chapters 2 desalter 3 preheat train 4 atmospheric tower 5 vacuum tower 6 cat cracker 7 hydrocracker 8 diesel hydrotreater 9 catalytic reformer 10 alkylation unit 11 delayed coker 12 hydrogen plant 13 sulfur recovery 14 fuel gas header 15 gasoline blender

Notice that the crude is washed, heated and distilled along the front row, where the gasoline is also blended, that the middle row converts and treats the cuts, and that the rear row holds the coker and the three plants that serve the others. Illustration, not to scale.

IN PLAIN TERMS **Running a refinery against its limits**

The main idea. A refinery heats crude oil, splits it into parts by how easily they boil, changes some parts chemically and blends the results into fuels. Each unit earns more the more oil it takes, so it is pushed until a product or a machine says stop.

The challenge. The limit that stops a unit moves with the oil, the weather and the wear of the equipment. The quality of a product is often known only minutes or hours after the oil has passed.

Why it matters. A unit held short of its limits wastes money, and one pushed past them makes fuel that cannot be sold.

THREE WAYS TO READ THIS CHAPTER

Explore	7 min	This page, Section 1.2 with its timeline and two History boxes, Figure 1.2 and the Did you know? boxes in Sections 1.4.1, 1.4.4 and 1.4.7, and Key points in Section 1.8. No equations.
Study	45 min	The whole chapter in order, working each number of the example R-1, then Exercises 1 to 10.
Practice	10 min	This page, Section 1.4.7 with Table 1.2, Section 1.5 with Figures 1.6 and 1.10, and Section 1.8.

1.1 The Question This Volume Answers

A refinery turns crude oil into fuels by heating it, cutting it into fractions, reacting the fractions over catalysts and blending what comes out. Each unit earns more the more it processes, each product is sold against a specification, and each piece of equipment has a limit that ends the run if it is crossed. This volume answers one question: how does each unit hold its products on specification and its

equipment inside its limits while it processes as much as the market will take, and which control techniques recur from the crude tank to the gasoline tank?

Six control problems define the industry, and each chapter meets at least two of them.

- *Heat is bought with fuel and recovered from products.* Fired heaters supply the heat the units need, exchangers recover what the products carry away, and fouling moves the load from the exchangers to the heaters.
- *Products are cut from a boiling curve.* A tower makes several products from one feed, so it has more specifications than handles, and half of the specifications follow from the other half.
- *Reactors feed back on their own heat.* An exothermic bed amplifies a change in its inlet, a cracking unit burns the coke it made a minute earlier, and an exchanger around a reactor returns part of every degree added.
- *Quality is known late or inferred.* Product analyzers report minutes to half an hour after the oil they describe, laboratories report a few times a day, and much of what the controllers read is computed from temperatures, pressures and densities.
- *Throughput is pushed against limits that move.* Firing, tube metal, compressors, blowers and catalyst all bind at some throughput, and which binds first changes with the crude, the weather and the season.
- *Units are coupled through shared streams.* Off-gases fuel the heaters, hydrogen passes from the units that make it to the units that consume it, acid gas goes to sulfur recovery, and the gasoline sold is the average of everything blended into a tank.

Fourteen chapters answer the question on fourteen plants, each with an app that the reader runs in a browser. Chapters 2 to 5 follow the crude through washing, preheating and atmospheric and vacuum distillation. Chapters 6 to 11 are the conversion and treating units: a catalytic cracker, a hydrocracker, a diesel hydrotreater, a catalytic reformer, an alkylation unit and a delayed coker. Chapters 12 to 15 are the plants that serve the others and finish the products: hydrogen, sulfur recovery, the fuel gas header and gasoline blending.

A reader who has worked through the volume can do the following on a unit that is new to the reader.

- Write a fired heater's heat balance, turn it into a feedforward and say how much throughput a fouled preheat train costs once the firing is at its limit.
- Count a tower's degrees of freedom, place its cut points on the assay and say which draw moves which cut.
- Compute the runaway index of an adiabatic bed and the loop gain of a feed and effluent exchanger, and say how close the unit sits to losing its operating point.
- Correct a temperature for pressure, update an inferred property from a laboratory result and say what an analyzer's delay costs.
- Place each controller in its layer, from regulatory loops to the optimizer, and name the limit that each override or selector protects.

This chapter develops each of these on one small example and names the chapter that uses it. PID forms and tuning are taken as known from Volume 2, cascade and feedforward from Volume 3, selectors and overrides from Volume 4, dead time and estimation from Volume 5, and predictive control and optimization from Volume 9.

1.2 A Short History

The first refineries made one product. A plant of the early 1860s distilled crude oil at atmospheric pressure to make kerosene for lamps, and naphtha and tar were its by-products [1]. The electric light then cut the demand for kerosene, the internal combustion engine created one for gasoline and diesel, and distillation alone gave only a certain amount of gasoline from each barrel [1]. Each later process

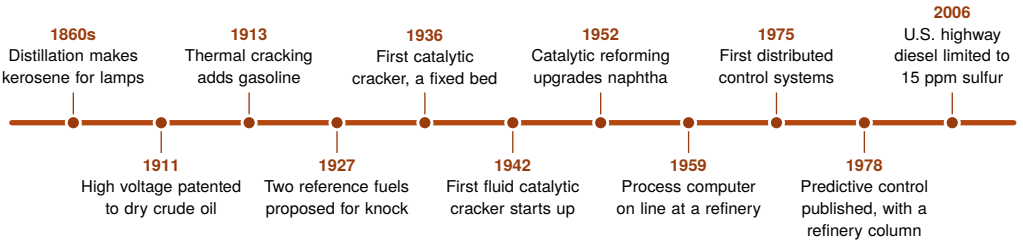


Figure 1.1 Milestones of refining and of its control. The process dates are from [1] (the 1860s, 1913 and 1952), [2] (1911), [3] (1927), [4] (1936 and 1942) and [5] (2006). The control dates are from [6] (1959 and 1975) and [7] (1978). Notice that the three control milestones fall within twenty years, about a century after the first unit. Events are evenly spaced, not to a time scale.

on the timeline of Figure 1.1 answers one of three pressures: more gasoline from a barrel, a better gasoline, or a cleaner product.

Thermal cracking, introduced in 1913, broke heavy fractions into gasoline under heat and pressure [1, 4]. In 1927 two pure hydrocarbons were proposed as reference fuels for knock, and the octane tests still rest on mixtures of those two [3]. Octane is the number that Chapter 9 holds and Chapter 15 blends to. Catalytic cracking arrived in the 1930s, alkylation in 1940 and catalytic reforming in 1952, and each raised the octane or the yield of gasoline [1]. Hydrodesulfurization followed in 1954 and hydrocracking in 1960 [1], the two processes that the plant of Chapter 12 supplies. A rule on highway diesel in the United States then cut the sulfur limit from 500 to 15 parts per million for refiners from June 1, 2006 [5]. Chapter 8 works under that limit with an analyzer half an hour late.

> ⧖ HISTORY **1913 to 1942: Three Ways to Crack a Heavy Oil**
>
> Thermal cracking, developed around 1913, heated distillate fuels and heavy oils under pressure in large drums until they cracked into smaller molecules, and it left large amounts of solid, unwanted coke [1]. Its gasoline was of relatively poor quality [4]. The first catalytic cracker started up in 1936 with its catalyst in a fixed bed and processed 15 000 barrels a day. The first commercial fluid unit followed in mid-1942, with a powdered catalyst that a gas flow lifts until it moves like a liquid between a reactor and a regenerator, and 34 such units ran in the United States at the end of the war [4]. Thermal cracking did not disappear. It evolved into visbreaking, steam cracking and coking [1]. The fluid unit is the plant of Chapter 6, and coking that of Chapter 11.

Control followed the processes at a distance. The first process computer went fully on line on a refinery unit in March 1959, and the first distributed control systems were introduced in 1975 [6]. Predictive control was first described at a conference in 1976 and published in 1978 [7], with a distillation column of a catalytic cracker among its three applications [8]. Figure 1.10 sets today's layers on their platforms.

> ⧖ HISTORY **1959: A Computer Goes on Line at a Refinery**
>
> The first computer sold for process use went fully on line in March 1959, on a catalytic polymerization unit of a Texas refinery [6]. It had a memory of 8K, and it raised the conversion of propylene to gasoline from 80 to 91 %. The first process computers were supervisory. They adjusted the setpoints of pneumatic or electric controllers, so the plant kept working when the computer was off. This one closed the loop on product quality every hour and adjusted 19 parameters. The first distributed control system, introduced in 1975, also adjusted independent controllers [6]. This volume still uses that arrangement, a slow layer that moves the setpoints of fast loops (Section 1.4.7).

1.3 Notation and Conventions

Temperatures are in °C, and differences of temperature in kelvin. Mass flows are in t/h, duties in MW, and a stream's heat capacity flow $C = \dot{m}\,c_p$ in MW/K. The coil outlet temperature of a heater is COT, the crude temperature leaving the preheat train T_{pre}, and an exchanger's effectiveness ε. A cut point is the temperature on the crude's true boiling point (TBP) curve at which one product ends and the next begins, and v is the cumulative fraction of the crude distilled. In a reactor, x is the conversion of the reacting lump, ΔT_{ad} the adiabatic rise at complete conversion, r the runaway index and L the gain of a loop of positive feedback. The framework's h is its integration step and its a the controller action, so neither is used here for anything else.

The chapters keep their own symbols, and several collide. Chapter 7 writes the adiabatic rise as K, and Chapter 14 uses K for the integrating gain of its header. Chapter 4 writes the crude's heaviness as h and the internal reflux ratio as R_i, while Chapter 13 uses R for the air ratio and Chapter 2 h for layer heights. Chapter 6, Chapter 5 and Chapter 8 write the controller action as σ. The weighted average bed temperature of Chapter 8 is WABT and the weighted average inlet temperature of Chapter 9 is WAIT. Octane is RON or the anti-knock index AKI, and vapor pressure is Reid vapor pressure (RVP) in psi, the units of the trade in Chapter 15.

1.4 From Crude Tank to Product Tank on One Train

1.4.1 The Refinery Flow Scheme

Figure 1.2 draws a refinery and marks where each chapter's plant sits in it [9]. Crude is washed free of salt, preheated against the hot products and heated in a furnace, and the atmospheric tower cuts it into naphtha, kerosene, diesel, gas oil and a residue. The residue is distilled again under vacuum. Its gas oil feeds the cracking units, and the vacuum residue goes to the coker. Naphtha is reformed into a high-octane component, diesel is hydrotreated to remove sulfur, and the light olefins of the cracker are alkylated. Gasoline is blended from several of these streams. Three plants serve all the others: the hydrogen plant supplies the hydroprocessing units, sulfur recovery takes the hydrogen sulfide that treating removes, and the fuel gas header carries the units' off-gases to every fired heater.

> **DID YOU KNOW? Where the Barrels Go**
>
> On January 1, 2026 the 130 operable refineries of the United States could charge 19.2 million barrels of crude oil per stream day. Their hydrotreaters could take 17.0 million, their vacuum towers 8.4 million, their catalytic crackers 5.3 million, their reformers 3.5 million and their hydrocrackers 2.3 million [10].

The chapters take the plants one at a time, and the techniques they share are developed below on one running example, a small train called R-1. A crude unit charges 720 t/h of crude through a preheat train and a fired heater to an atmospheric tower. The tower's diesel cut feeds a hydroprocessing reactor with a feed and effluent exchanger, a charge heater and one adiabatic bed. The example ends at a gasoline tank. Every number of R-1 is declared here or computed from these declarations by the script kept with this chapter's record. The numbers are chosen for clear arithmetic and are not taken from any one chapter.

1.4.2 Fired Heaters

A fired heater raises a process stream from the temperature at which it arrives to its coil outlet temperature, and the heat it must absorb follows from a balance on the stream, Equation (1.1).

$$Q = C\,(\mathrm{COT} - T_{pre}) + Q_{vap}, \qquad C = \dot{m}\,c_p \tag{1.1}$$

where Q_{vap} is the heat that vaporizes part of the stream in the coil. R-1's crude has $c_p = 2.5\,\mathrm{kJ/(kg\,K)}$, so 720 t/h is $C = 0.5\,\mathrm{MW/K}$. The example leaves out Q_{vap}, which Chapter 4 and Chapter 5 keep.

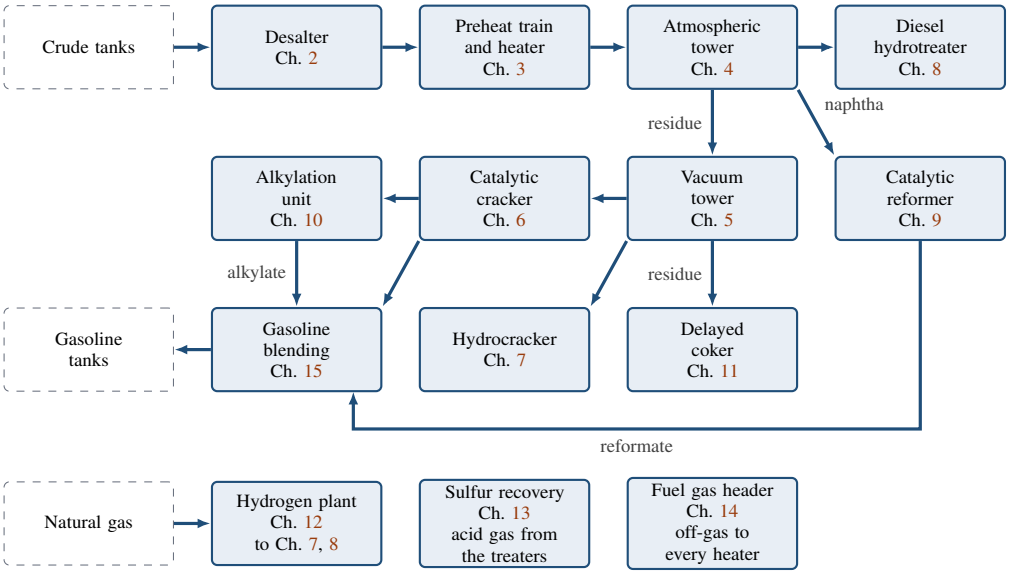


Figure 1.2 A Refinery Flow Scheme. Filled Blocks Are the Plants of Chapters 2 to 15, Dashed Blocks Lie Outside the Volume. Vacuum Gas Oil Feeds the Cracker and the Hydrocracker, and the Cracker's Light Olefins the Alkylation Unit. The Three Plants in the Bottom Row Serve All the Others. Kerosene, Light Ends and Most Product Pools Are Omitted.

The crude arrives at 265 °C and leaves at a COT of 365 °C, so the heater absorbs 50 MW. Its burners can deliver 60 MW, and the heater runs at 83.3 % of full firing.

The balance is also the most useful feedforward in a refinery. A heater whose firing follows Equation (1.2) meets a change in rate or in preheat before the outlet temperature moves.

$$Q_{ff} = C_m \left(\mathrm{COT}_{SP} - T_{pre,m} \right) \tag{1.2}$$

where the subscript m marks a measured value. On R-1 each kelvin of preheat lost needs 0.5 MW more, and a tenth more crude needs 5 MW more at the same preheat. The outlet loop then trims only what the balance does not know. Chapter 3 adds this feedforward to its outlet loop, Chapter 12 subtracts the heat its tail gas brings, and Chapter 11 and Chapter 9 drive their heaters from the duty their balances predict.

The fuel is the balance's least certain term. Refinery fuel gas is a mixture of off-gases whose heating value changes with what the units send. A heater whose fuel valve is held at one flow loses heat in proportion to the heating value. R-1's heater adds 100 K, so each percent of heat lost is a kelvin of COT, and a gas 5 % leaner takes 5 K off the outlet before the loop acts. A burner held at a fixed pressure instead passes heat in proportion to the Wobbe index, the heating value divided by the square root of the specific gravity. Hydrogen lowers both, so 10 % hydrogen in the gas of Chapter 14 lowers the heating value 7.11 % but the Wobbe index only 2.66 %. On R-1 that is 7.1 K against 2.7 K at the coil outlet. Chapter 14 feeds the Wobbe index forward to its burners, and Chapter 3 meets an unmeasured change in heating value through its outlet loop alone.

A heater has two limits. The firing is limited by its burners and its fuel valve. The tubes are limited by their metal temperature, which runs above the fluid by an amount that grows with the heat flux, as shown in Equation (1.3).

$$T_{metal} = T_{fluid} + \theta_m \frac{Q}{Q_{max}} + \Delta T_{coke} \tag{1.3}$$

where θ_m is the metal's rise at full firing and ΔT_{coke} grows as coke lays down inside the tubes. The tube metal is computed, or read from a few thermocouples welded to the tubes, and it binds before the

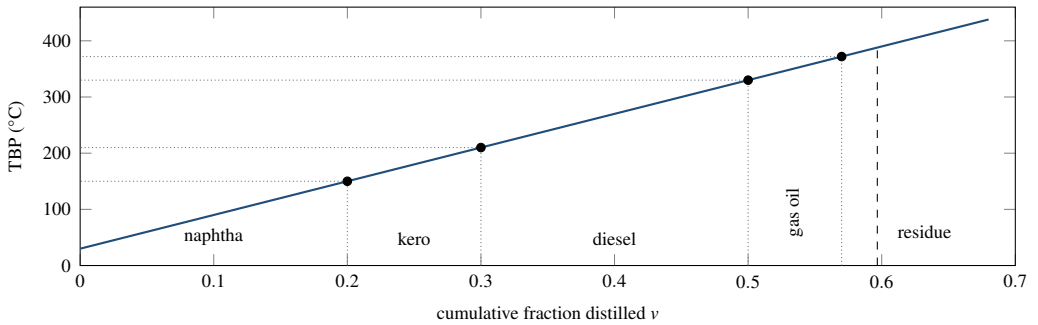


Figure 1.3 R-1's Assay (Thick Line) and Its Four Cuts (Dots), Placed by the Material Balance of Equation (1.4). The Dashed Line Is the Fraction Vaporized at the Flash Zone, and the Gap Between It and the Last Cut Is the Overflash.

firing does once the tubes are coked. Chapter 9 and Chapter 12 write the metal in this form, Chapter 5 and Chapter 11 add a coke layer that grows with temperature, and Chapter 12 protects its tubes with an override on the outlet loop.

1.4.3 Distillation: Cut Points and Degrees of Freedom

A crude is described by its true boiling point curve, the temperature at which each cumulative fraction of it boils [11]. A tower that makes products in series from one feed cuts that curve, and to a first approximation the material balance alone places each cut. The cut below draw i lies at the share of the crude taken by that draw and every draw above it, as shown in Equation (1.4).

$$v_i = \frac{1}{F}\sum_{j\le i} D_j, \qquad T_i = T_{bp}(v_i), \qquad \frac{\partial T_i}{\partial D_j} = \begin{cases} T'_{bp}(v_i)/F, & j \le i \\ 0, & j > i \end{cases} \tag{1.4}$$

where F is the charge, D_j the product drawn at position j counted from the top and T'_{bp} the slope of the curve. R-1's assay is a straight line, $T_{bp} = 30 + 600\,v$, which no crude is, but its slope is in the range of a real curve's middle. Naphtha, kerosene, diesel and gas oil are drawn at 144, 72, 144 and 50.4 t/h, so the cuts fall at 150, 210, 330 and 372 °C, as shown in Figure 1.3. Each tonne an hour of a draw moves every cut below it by 0.83 K and leaves every cut above it alone.

The structure of Equation (1.4) is the lesson. Each product has two ends, and neighbors share a cut, so four cuts fix the eight end specifications of four products. Four of the eight are choices and four are consequences, and which four is an operating decision. The kerosene's front end and the naphtha's back end are the same cut, so a loop on each fights over one variable. The interaction is also one way. A draw moves its own cut and every cut below it, so the gain matrix is lower triangular. To lift the kerosene's end point by 12 K without moving the diesel's, the kerosene draw rises 14.4 t/h and the diesel draw falls by as much.

The tower's heat sets what the material balance does not. Internal reflux sharpens each cut, and a draw that takes liquid off a tray thins the reflux below it. Chapter 4 measures that second effect at about 4 % of the first. The heater sets how much of the crude vaporizes at the flash zone. R-1's flash zone, 2 K below the COT and read 25 K higher on the assay for the stripping steam, vaporizes 59.67 % of the crude, and the draws take 57 %. The difference, 2.67 %, is the overflash, the liquid that wets the wash zone above the flash zone. With a floor of 2 %, the gas oil draw can rise only 4.8 t/h. Each kelvin of COT vaporizes 1.2 t/h more and costs 0.5 MW, so the lowest draw is held by the heater as much as by its own valve.

Chapter 4 counts these degrees of freedom on a four-product tower, measures the triangular interaction and protects the overflash with an override. Chapter 5 cuts the residue by pressure as well as by heat, and holds its wash bed wet with an overflash override. Chapter 11 cuts a fractionator

whose feed swings with the drum cycle, and Chapter 10 sets the purity of its recycle with a column built on the column module shared across the series.

1.4.4 Reactor Heat Balances and Runaway

Most refinery reactors are adiabatic beds of catalyst. The heat of reaction stays in the stream, so the outlet temperature is the inlet plus the reaction's rise, and the rise depends on the temperature. A lumped bed that reacts at the mean of its inlet and outlet obeys Equation (1.5), the form Chapter 7 uses.

$$\tau_b \frac{\mathrm{d}T_{out}}{\mathrm{d}t} = T_{in} + \Delta T_{ad}\, x(\bar{T}) - T_{out}, \qquad \bar{T} = \tfrac{1}{2}(T_{in} + T_{out}) \tag{1.5}$$

where τ_b is the bed's thermal time constant and x the conversion, which rises with temperature. Linearizing about a steady state gives Equation (1.6), in which one group appears twice.

$$\tau_b \frac{\mathrm{d}\,\delta T_{out}}{\mathrm{d}t} = (r-1)\,\delta T_{out} + (1+r)\,\delta T_{in}, \qquad r = \frac{\Delta T_{ad}}{2}\frac{\mathrm{d}x}{\mathrm{d}\bar{T}} \tag{1.6}$$

The runaway index r decides two things. The bed is stable while $r < 1$, and while stable it amplifies its inlet by the steady gain of Equation (1.7) and slows its own response by $1/(1-r)$.

$$G = \frac{\partial T_{out}}{\partial T_{in}} = \frac{1+r}{1-r} \tag{1.7}$$

Studies of parametric sensitivity give the same warning for beds with axial profiles [12].

R-1's bed has $\Delta T_{ad} = 100\,\mathrm{K}$ and a conversion that grows by a factor e every 30 K of mean temperature, which doubles it every 20.8 K, up to complete conversion. At the design inlet of 320 °C it converts $x = 0.2$, rises 20 K to 340 °C and has $\mathrm{d}x/\mathrm{d}\bar{T} = x/30$, so $r = 50 \times 0.2/30 = 1/3$ and $G = 2$ exactly. A kelvin at the inlet is two at the outlet, and the bed's time constant grows from 3 to 4.5 min. Because r grows with the rise, the bed loses its cool operating point where the rise reaches 60 K. That happens at an inlet of 332.96 °C, only 12.96 K above design, as shown in Figure 1.5a. Above it the only steady state is complete conversion, 100 K above the inlet, and the bed stays there until the inlet is cut below 328.28 °C.

A cut in feed moves the bed toward the same edge. Conversion rises as the space velocity falls, and the heat released has less oil to warm. At the design inlet a tenth less feed lifts the rise to 23.6 K and G to 2.30, a fifth less to 29.1 K and 2.88, and a cut of 35.1 % leaves no cool state at all. Chapter 7 reports the same mechanism on its four beds, where the first sits at $r = 0.538$, amplifies its inlet 3.3 times and loses its cool state four degrees above design. Its quench between beds breaks the chain of amplification, and its protection must read a bed rise whose transient sign is wrong for a few minutes after every correction.

An endothermic bed has $r < 0$ and $G < 1$. A hotter bed reacts faster and cools itself more, so it absorbs most of any heat given to it. Chapter 9 measures this on its first bed, which passes 6.3 K of a 15.7 K rise in its inlet to the next bed, which is why a reformer needs a heater before each bed.

The same arithmetic governs any unit that feeds back its own heat. If a fraction L of any change returns to its own input, the steady change is amplified as in Equation (1.8), and a first-order element inside the loop slows by the same factor.

$$A = \frac{1}{1-L}, \qquad \tau_{eff} = \frac{\tau}{1-L} \tag{1.8}$$

The cat cracker of Chapter 6 has no furnace. Its catalyst is reheated by burning the coke the cracking laid on it, and its measured loop gain is 0.523, so a slide valve move arrives 2.09 times larger than it first looks and settles over about 90 min. The loop gain grows with severity, and once it reaches one the unit has more than one steady state [13, 14]. A classic analysis of autothermic processes sets the heat a reaction generates against the heat its stream removes in the same way [15], and Table 1.1 collects the loop gains of the volume.

Table 1.1 Heat Fed Back in the Example and the Chapters

Plant	Path of the feedback	Index or L	Gain
R-1 bed	reaction heat in the bed	$r = 1/3$	2.00
R-1 bed and exchanger	effluent heat returned to feed	$L = 0.60$	2.50
Chapter 7, bed 1	reaction heat in the bed	$r = 0.538$	3.3
Chapter 8	effluent heat returned to feed	$L = 0.728$	3.7
Chapter 6	coke burned to reheat catalyst	$L = 0.523$	2.09
Chapter 9, bed 1	endothermic reaction	$r < 0$	below 1

✦ DID YOU KNOW? Three Units That Multiply Their Own Disturbances

From the simulator (Table 1.1). The first bed of the hydrocracker passes on a change at its inlet 3.3 times larger. The exchanger loop of the diesel hydrotreater multiplies a change by 3.7, and the heat balance of the cat cracker by 2.09.

1.4.5 Heat Integration

A refinery recovers most of its heat. R-1's preheat train heats the crude from 40 °C against products that leave the tower at 340 °C. Because the products are the crude, their heat capacity flow equals the crude's, and the train behaves as one balanced counter-current exchanger with the effectiveness of Equation (1.9).

$$\varepsilon = \frac{N}{1+N}, \qquad N = \frac{UA}{C}, \qquad T_{pre} = T_{cold} + \varepsilon\,(T_{hot} - T_{cold}) \tag{1.9}$$

Clean, the train has $UA = 1.5\,\mathrm{MW/K}$, so $N = 3$, $\varepsilon = 0.75$ and $T_{pre} = 265\,^{\circ}\mathrm{C}$. It recovers 112.5 MW, which is 69.2 % of the heat the crude takes, and the heater supplies the rest.

Fouling lays a deposit on the exchanger walls and lowers UA over weeks [16, 17]. The outlet loop hides it. The COT stays on setpoint while the firing climbs, and only the position of the fuel valve records what happened. On R-1 the firing reaches 95 % when 21 % of the clean UA is lost and the preheat has fallen to 251 °C, as shown in Figure 1.4a. Beyond that point no tuning helps, and the only handle left is the charge. A valve position override takes the charge from the operator's demand when the firing passes its target, as shown in Equation (1.10).

$$F_{cmd} = \min\left(F_{dem},\ \mathrm{PI}(x_{v,SP} - x_v)\right) \tag{1.10}$$

where x_v is the firing and $x_{v,SP}$ its target, 95 % here. The override's upper output limit is the demand, so it rests there until the firing crosses its target and cannot wind up while it waits.

The override costs throughput, and less than a first estimate suggests. At 40 % of UA lost, the charge held at 720 t/h would need 66.1 MW and would deliver a COT of 352.9 °C with the burners wide open. The override cuts the charge to 653.6 t/h, a cut of 9.2 %, and holds the COT. An estimate at fixed preheat would allow only 621.1 t/h. The difference is the train itself: less crude through the same surface has a larger N, so the preheat rises from 232.9 °C to 239.4 °C and each tonne needs less firing. Chapter 3 measures this trade on a two-exchanger train, where holding temperature through one exchanger's cleaning costs 555 t of crude, and Chapter 4 ties its preheat to the three pumparound duties that recover the tower's heat.

Around a reactor, heat integration closes a loop. R-1's diesel enters its feed and effluent exchanger at 200 °C, and the exchanger with its bypass returns $\varepsilon_u = 0.3$ of the effluent's temperature above the feed. The charge heater adds the rest of the way to the bed, as shown in Equation (1.11).

$$T_{in} = T_f + \varepsilon_u\,(T_{out} - T_f) + \frac{Q_h}{C}, \qquad L = \varepsilon_u\,G \tag{1.11}$$

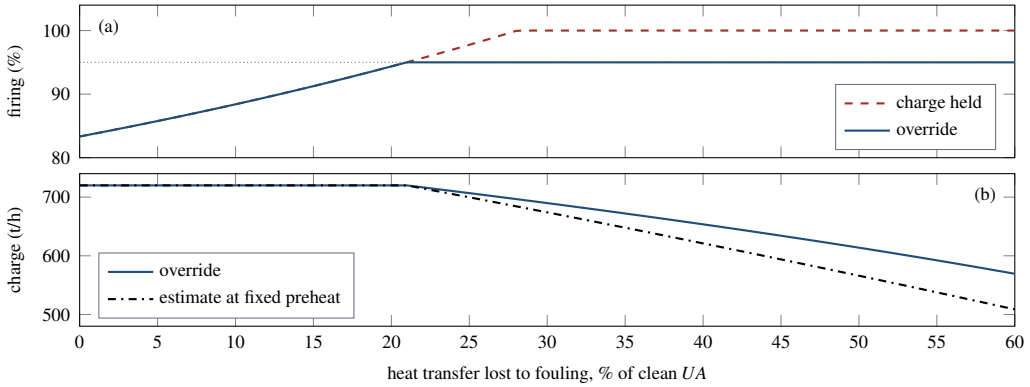


Figure 1.4 R-1's Crude Unit as Its Preheat Train Fouls. (a) Firing With the Charge Held at 720 t/h (Dashed), Which Reaches Its Limit and Then Loses COT, and With the Valve Position Override (Solid), Which Holds the 95 % Target (Dotted). (b) The Charge the Override Allows (Solid) and the Estimate That Ignores the Preheat Gained at Lower Charge (Dash-Dotted).

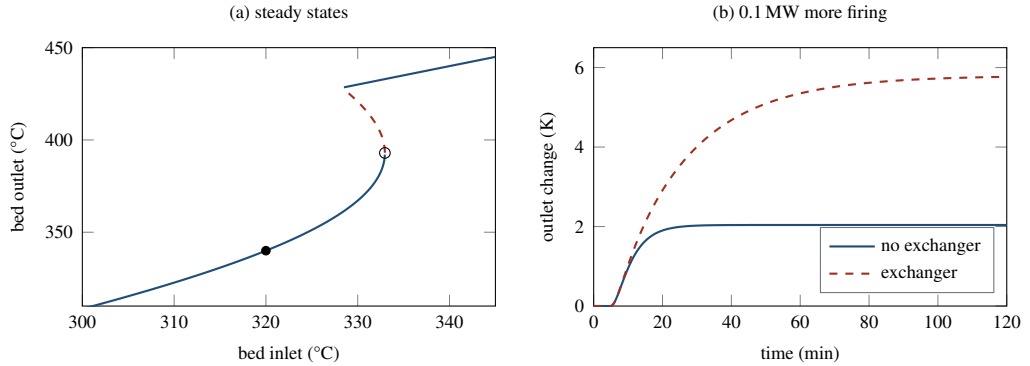


Figure 1.5 R-1's Bed. (a) Steady Outlet Against Inlet: Stable States (Solid) and Unstable States (Dashed). The Dot Is the Design Point and the Circle the Fold Where $r = 1$. (b) Outlet After a 0.1 MW Step in Firing at 5 min, With the Heater Feeding the Bed Directly (Solid) and Through the Exchanger Loop (Dashed).

At design the exchanger delivers 242 °C and the heater adds 78 K, 7.8 MW at $C = 0.1\,\mathrm{MW/K}$. A kelvin more at the inlet is two more at the outlet, and 0.3 of those two come back, so the loop gain is $L = 0.6$ and Equation (1.8) gives an amplification of 2.5. An exchanger that returned half the effluent's heat would bring L to one, the edge of open-loop instability [18]. The loop also brings the bed's own edge closer. With the exchanger in place the unit loses its cool operating point at an inlet of 328.23 °C, 8.23 K above design instead of 12.96 K, and only 0.21 MW more firing takes it there.

Figure 1.5b shows both effects of the loop. With heater and bed lags of 2 and 3 min and an exchanger lag of 3 min, 0.1 MW more firing raises the inlet 1 K and the outlet 2.04 K when the exchanger's return is held, and the outlet reaches 63 % of its change in 7.0 min. Through the loop the outlet settles 5.8 K higher, more than the linear 5.0 because r grows with the rise, and takes 21.5 min to reach 63 %. An operator or a loop tuned on the first ten minutes sees a plant less than half as strong as the one underneath.

Chapter 8 measures its exchanger loop at $L = 0.728$, rising to 0.899 at a cooler inlet, and moves the inlet with a fast bypass while mid-ranging returns the slow heater to the middle of its range. Its bed also shows wrong-way behavior, a cooler inlet that first warms the outlet [19]. Chapter 9 returns

its last bed's effluent to its first heater, and Chapter 12 burns the adsorber's tail gas in the furnace that makes the hydrogen, so what the adsorber recovers changes what the furnace burns.

1.4.6 Analyzers and Inferred Properties

A refinery sells its products on properties that no transmitter reads: a sulfur content, an octane number, a vapor pressure, a 95 percent point. Three sources supply them, and each has a cost. An on-line analyzer takes a sample, carries it down a line and reports after a cycle, so its reading is always old. A sample-and-hold report with transport delay θ and cycle T_s is between θ and $\theta + T_s$ old, and for tuning it acts as the effective delay of Equation (1.12).

$$\theta_{eff} \approx \theta + \tfrac{1}{2} T_s \tag{1.12}$$

R-1's sulfur analyzer, with a 20 min delay and a 10 min cycle, is 25 min late on average. A loop on it can be no faster than that delay allows, as Volume 5 develops. A laboratory is more accurate and far rarer, a few results a day.

An inferred property is computed now from what is measured now. The oldest inference is a tray temperature corrected for pressure, because a cut point is a boiling point and a boiling point moves with pressure. The atmospheric equivalent temperature of Chapter 5 gives the normal boiling point T_b of what boils at a temperature T and pressure p, as shown in Equation (1.13).

$$T_b = T \left(1 + \frac{1}{A} \ln \frac{p_{atm}}{p} \right) \tag{1.13}$$

with temperatures in kelvin and A a constant of the fraction, 16 on the residue of Chapter 5 and declared 12 for R-1's diesel. R-1's diesel draw tray at 2 bar boils at 366.2 °C for a 330 °C cut. A rise of 0.1 bar raises the tray 2.8 K with no change in the cut, so a loop holding the raw temperature would lighten the diesel cut by 2.6 K. Chapter 5 infers its heavy gas oil's 95 percent point this way, and Chapter 4 reads its product properties from inferred cut points.

A model of the property from several measurements is a soft sensor [20, 21]. It has no delay, but it drifts as the plant moves away from where the model was fitted. The laboratory corrects it with the bias update of Equation (1.14).

$$b_k = b_{k-1} + \lambda \left(y_{lab,k} - \hat{y}_{raw}(t_{s,k}) - b_{k-1} \right) \tag{1.14}$$

where $\hat{y}_{raw}(t_{s,k})$ is the model's value at the instant the sample was drawn, not when the result came back, and $0 < \lambda \leq 1$. With $\lambda = 0.5$, half of a step in the true bias remains after one result and an eighth after three. Comparing against the wrong instant adds error of its own. A property drifting by 0.1 an hour, compared at a result that returns 2 h after sampling, puts 0.2 into the bias. Chapter 10 corrects a density soft sensor for acid strength from a titration every four hours, Chapter 9 corrects an analyzer and an inferred octane from the laboratory and gates a reading it cannot trust, and Chapter 15 updates the bias of its blend model.

Three chapters turn on what a measurement does to the problem. Chapter 13 cannot measure recovery, which has a maximum and therefore no sign. Its analyzer reads the excess of hydrogen sulfide over twice the sulfur dioxide, which is zero at the maximum, and an ordinary PI loop holds it there. The instrument of Chapter 2 reads part of the rag layer as water, so a loop that holds the reading lets the true water fall as the rag grows, with no error on the trend. In Chapter 15 the product is certified on the average of the whole tank. With $\bar{y}$ the tank's average so far and f the share of the batch still to blend, the rest must average the value of Equation (1.15) to finish on the target y^*.

$$y_h = y^* + \left(y^* - \bar{y} \right) \frac{1 - f}{f} \tag{1.15}$$

A tank 40 % full at 87.0 against a target of 87.2 needs the header at 87.33. At 90 % full it needs 89.0, which the 1.5-unit clip of Chapter 15 does not allow. The authority of the header over the tank shrinks as the tank fills.

1.4.7 The Layers From Regulatory Control to Optimization

Refinery control is built in layers, each faster than the one above and each setting the targets of the one below [22, 23]. Regulatory loops hold flows, pressures, levels and temperatures in seconds to minutes. Constraint control decides which loop owns a shared actuator when a limit binds: overrides, low and high selectors, valve position control and split ranges, acting in minutes. Quality control trims the regulatory setpoints from analyzers, laboratories and soft sensors over tens of minutes to hours. Multivariable predictive control moves several handles together against several limits [7], and optimization sets the targets from prices and models over hours to days [24].

> ✦ DID YOU KNOW? **Refining Holds the Largest Block of Predictive Controllers**
>
> A vendor survey taken in mid-1999 estimated 4542 applications of linear predictive control by five suppliers. Refining held 1985 of them, the largest single block. The count leaves out the controllers that operating companies built for themselves [7].

The layers share one rule. Most economic optima of a refinery lie on a limit, because throughput and severity both earn money until some piece of equipment says stop. The exceptions sit inside the range, such as the salt minimum of Chapter 2 and the recovery maximum of Chapter 13, and each is found by an optimizer or turned into a setpoint. A layer that pushes throughput must therefore know which limit binds, and the binding limit changes. A low selector over one throughput handle is the classical answer, with external reset so that the loops not selected do not wind up, and its regions can be designed systematically [25]. A coordinating layer does the same across units, raising the plant's rate until its tightest unit binds [26]. R-1's override of Equation (1.10) is the smallest case: one handle, one limit, and a throughput that falls only when the limit binds.

Table 1.2 lists, for each chapter, the handle that sets throughput, the quality the unit holds and how it is known, and the equipment limit with the structure that guards it. Chapter 16 adds the measured results and asks how each unit balances throughput, product quality and equipment limits.

1.5 Instruments and Final Elements in This Volume

Every loop in this volume begins at a sensor and ends at a final element, the device that changes a flow of material or of energy. Figure 1.6 lists what the controllers of each chapter read and what they move, taken from the chapter's variables table and from its instrument model. A flow is read in all fourteen chapters, an on-line analyzer in twelve and a temperature in eleven. A valve or a flow loop is moved in thirteen, and the firing or duty of a heater in ten.

Each device sets a limit that no tuning removes. A thermocouple reaches the stream through a thermowell, drawn in Figure 1.7, and every thermowell adds a first-order lag through its own mass and specific heat [27]. The chapters carry that lag as a time constant, 36 s on the column inlet of Chapter 3. A flow read across an orifice plate loses accuracy below about a third of the meter's range [27].

The analyzers bring the longest delays. A chromatograph separates the components of a sample in time, as shown in Figure 1.8, so cycle times of minutes are not uncommon and some dead time is inevitable [27]. Equation (1.12) turns a cycle into an effective delay. The cycles in this volume include 10 s on the Wobbe meter of Chapter 14 and 25 min on the conversion analyzer of Chapter 7. Where no analyzer serves, a laboratory sample or an inferred property stands in, and eight chapters carry an inference.

The final element is most often a control valve, drawn in Figure 1.9. Static and dynamic friction turn a smooth signal into stair-steps of the stem, an effect called stiction [27]. A valve at the end of its travel has no authority left, which is the case that the override of Equation (1.10) guards.

The platform sets the last limit. In a distributed control system a supervisory controller sends setpoints to distributed field controllers, which move their actuators from sensor feedback, and cycle times run from milliseconds to minutes [28]. Figure 1.10 places the layers of Section 1.4.7 on that

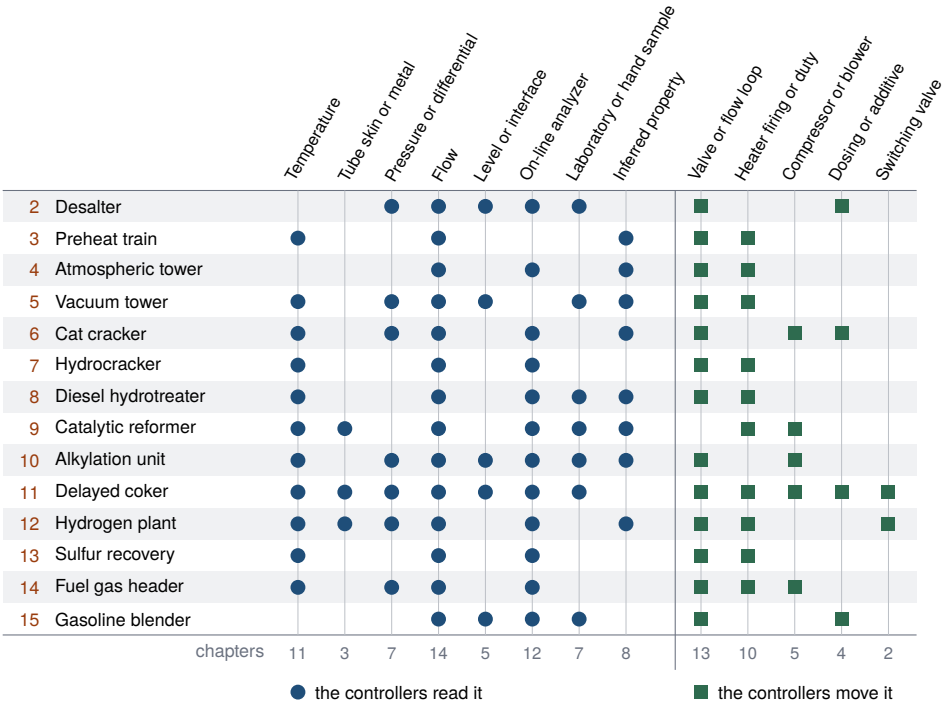


Figure 1.6 Sensors and final elements by chapter. A round mark means that the chapter's controllers or operator read the device, a square mark that they move it. A mark is set only where the chapter's variables table or its instrument section names the device or its reading, so a blank says that the model leaves the device out and not that a real unit lacks it. Notice that every chapter reads a flow and that all but two read an analyzer.

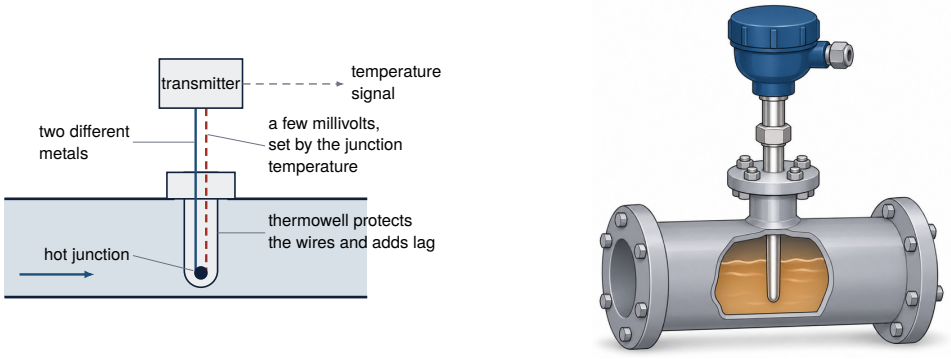


Figure 1.7 Principle of a thermocouple in a thermowell, with a view of the device as installed. Notice that the junction sits inside a metal well, so the reading follows the stream through the well's lag.

Table 1.2 For Each Chapter, the Throughput Handle, the Quality Held and How It Is Known, and the Equipment Limit With Its Guard

Ch.	Throughput	Quality, and how it is known	Limit, and its guard
2	Charge, set upstream	Salt and water in the crude, sampled, and a misread interface	Grid current and trip, an optimizer on the mixing pressure drop
3	Charge	Column inlet temperature, measured	Firing, valve position override on the charge
4	Charge, set upstream	End points and cold properties from inferred cuts	Overflash and heater duty, overflash override
5	Lift by pressure and COT	Heavy gas oil 95 percent point, inferred	Ejector break and tube skin, overrides in a maximizer
6	Feed, at a low selector	Severity, as the riser outlet temperature	Blower, compressor, cyclone and valve limits, four constraint controllers
7	Feed	Conversion on a 25 min analyzer	Bed rise and temperature, override on a compensated rise, trips
8	Feed	Sulfur on a 30 min analyzer	Outlet temperature and heater duty, quench override, mid-ranging
9	Feed	Octane from an analyzer, the laboratory and an inference	Tube metal, regenerator capacity, the operator and the MPC
10	Olefin feed	Acid strength titrated every four hours, a density soft sensor	Refrigeration capacity and the strength floor, a supervisor
11	Charge on a drum cycle	Fractionator cuts	Tube skin and foam, schedule feedforward and MPC
12	Plant rate on header pressure	Steam to carbon, purity from the adsorber	Steam floor and tube metal, lead-lag and an override
13	Acid gas, set upstream	Tail gas ratio on a 3 min analyzer	Sulfur dew point and stack permit, dew point override
14	Off-gas, set by other units	Header pressure and Wobbe index, measured	Import and flare at once, split range with a dead band
15	Blend rate into a tank	Tank-average octane and RVP, analyzer and laboratory	Component stocks and giveaway, an LP and average targeting

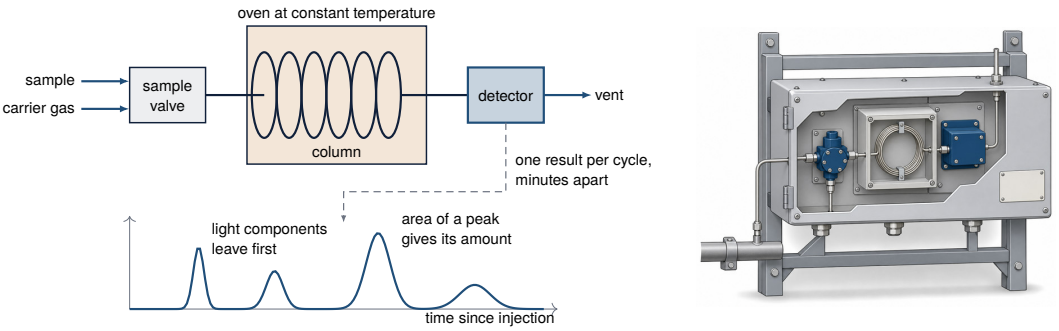


Figure 1.8 Principle of a process gas chromatograph. Notice that one sample gives one result at the end of the cycle, so a loop on this instrument acts on a value that is always old.

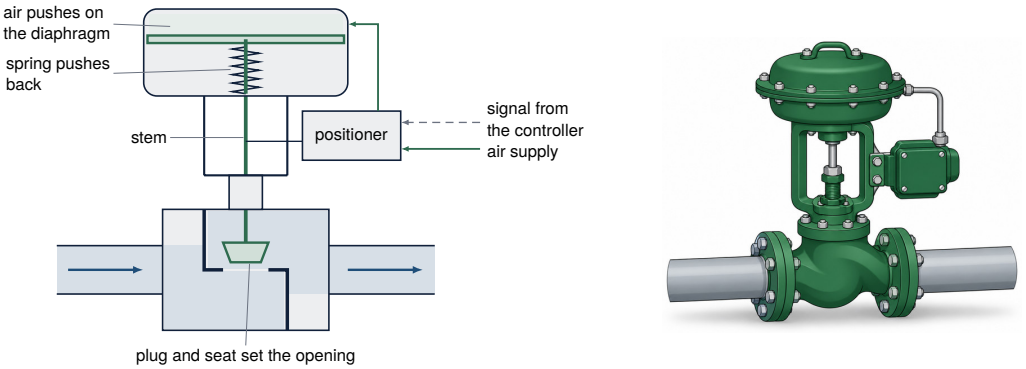


Figure 1.9 Principle of a globe control valve with a spring-and-diaphragm actuator and a positioner, with a view of the device as installed. Notice that the positioner compares the controller's signal with the position of the stem and adjusts the air until the two agree [27].

Layer	What it does	Where it runs	How often
Optimization	sets targets from prices and models	**Supervisory computers** send setpoints to the controllers below	every hour for a unit, every day for the plant
Predictive control	moves several handles together against several limits		every minute
Quality control	trims setpoints from analyzers, the laboratory and soft sensors	**Distributed control system** field controllers with PID, lead-lag and select blocks	at each analyzer or laboratory report
Constraint control	overrides, selectors, split ranges, valve position control		every second to every minute
Regulatory control	holds flows, pressures, levels and temperatures		every second
Field devices	transmitters, analyzers, valves and drives at the unit	point-to-point wiring or a fieldbus	continuously

Figure 1.10 The control layers of Section 1.4.7, the platform that hosts each and how often each acts. The periods of the regulatory, predictive and optimization layers are those of [7, 23], and the platforms those of [28]. Notice that each layer sends setpoints to the layer below and acts less often than it.

hardware. Regulatory loops execute about every second, constraint and predictive control every minute, a unit optimizer every hour and a plant-wide optimizer every day [7, 23]. The apps execute their loops every 6 s in Chapter 11 and every 10 s in Chapter 8.

1.6 The Simulator Framework

Every chapter of this book has an app that runs in a web browser, collected with the others at https://apmonitor.com/apps/. About 240 apps exist across the series. All but a few are built on one shared framework, so the engine, the PID algorithm, the trends and the data export behave the same way in every chapter. The exceptions are standalone studios, such as the cascade studio and the two Temperature Control Lab apps, which follow the same ideas in their own code. This section describes the framework once and states where the standalone apps differ. Table 1.3 lists what a reader meets.

Table 1.3 What Every App Provides. Standalone Apps Are Described in Their Own Chapters.

Element	What it does
Run, Pause, Reset	Start, hold and restart the run. Reset keeps the reader's settings.
Speed	Simulated time per second of real time. It changes the pace and never the result.
Manual and Auto	Hand each actuator to the reader or to the controller, without a bump.
Scenario list	Numbered cases, each with a briefing, a completion rule and a reload button.
Control panel	Setpoints, manual outputs, tuning, options and disturbances, at a Beginner or an Advanced level.
Trends	SP and PV above the MV for each loop, with limits, event marks and a cursor.
Scorecard	IAE, overshoot, settling time, MV travel and time at a limit for each loop.
Learn card	Predict, Try It, two hints and Explain, opened by the reader in that order.
Show Model	Equations, live variable values, editable parameters and assumptions.
Export CSV	Every logged sample of the run as a text file with a header row.
Instructions	A separate page with background, an operating guide and the model's limitations.

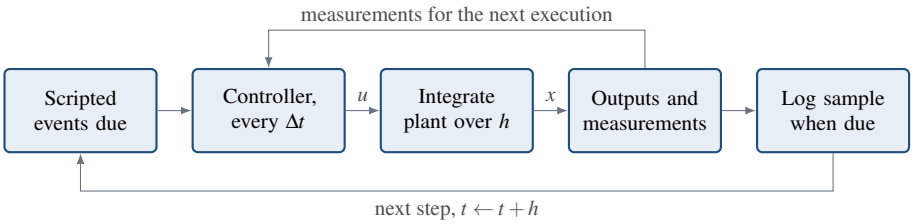


Figure 1.11 Work of the Engine in One Integration Step

1.6.1 The Simulation Engine

Each app carries a plant model written as ordinary differential equations $\frac{dx}{dt} = f(x,u,d,p)$ in the states x, the inputs u, the disturbances d and the parameters p. The engine integrates the model with the classical fourth-order Runge–Kutta method on a fixed step h set for the fastest dynamics of the app. Some apps supply their own step, such as explicit Euler for a discontinuous right-hand side or an exact update for a model driven by events.

Three clocks run at separate rates. The plant advances every step h. The controller executes every control period Δt, equal to h unless the app declares a slower controller, and holds its outputs between executions. The log takes one sample at a fixed interval of simulated time for the trends, the scorecard and the exported file. The order of work in one step is shown in Figure 1.11. The controller also executes once when a scenario is loaded, so every run starts from an initialized controller.

Drawing the screen is kept apart from these clocks. A speed of 5 means five units of the app's time base per second of real time. Speed, frame rate and pauses change how fast a run is shown and never what it computes, so a phone and a workstation produce the same trajectory and the same file.

Noise comes from a deterministic pseudorandom generator. Its seed is declared by the app, or by a scenario that needs its own noise record, and it is restored whenever a scenario is loaded or Reset is pressed. The framework gives the reader no seed setting, so *default seed* in this book means the seed the app ships with. A run repeats bit for bit when the same scenario is loaded and the same settings are made before Run. A change made by hand during a run lands on a different step each time, which is one reason numeric answers carry a tolerance. The standalone simulators draw noise from the browser's own generator, so a noisy run there repeats in its statistics and not sample for sample.

Run and Pause start and hold the clock. Reset restarts the run at $t = 0$, paused, and keeps the reader's settings: every setpoint, tuning value, option, model parameter, disturbance slider and mode the reader changed is put back before the plant and the controller are initialized. Scripted events replay, and a one-shot disturbance button pressed during the run is cleared. The reload button beside

the scenario list restores the scenario's own settings instead.

1.6.2 The Velocity-Form PID Algorithm

One PID implementation serves the PID loops of every framework app. It is the velocity (incremental) form: each execution adds the change in the proportional and derivative terms to a running sum, and the integral is a separate term, as shown in Equations (1.16) to (1.19) for execution k and control period Δt.

$$e_k = a\,(\mathrm{SP}_k - \mathrm{PV}_k), \qquad P_k = K_c\,e_k, \qquad D_k = -a\,K_c\,\tau_D\,\frac{\mathrm{PV}^f_k - \mathrm{PV}^f_{k-1}}{\Delta t} \tag{1.16}$$

$$S_k = S_{k-1} + (P_k - P_{k-1}) + (D_k - D_{k-1}) \tag{1.17}$$

$$v_k = u_b + S_k + I_{k-1}, \qquad u_k = \mathrm{sat}(v_k) \tag{1.18}$$

$$I_k = I_{k-1} + \left[\frac{K_c}{\tau_I}\,e_k + \frac{u_k - v_k}{T_t}\right]\Delta t \tag{1.19}$$

where $a = \pm 1$ is the controller action, PV^f is the measurement after an optional first-order filter, u_b is the output bias, S is the running sum of proportional and derivative increments, v is the output before limiting, u is the output sent to the actuator, I is the integral term and T_t is the tracking time constant, equal to τ_I unless a loop sets another value. The function sat clamps v_k to the actuator range and then applies a rate limit if the loop declares one.

The limit acts on the emitted output u_k only. The sum S_k is never clamped, so it equals $P_k + D_k$ and the output before limiting is the position form $u_b + K_c e_k + D_k + I_{k-1}$. Anti-windup touches the integral alone. While the output is limited, the last term of Equation (1.19) drives I toward the value that reproduces the output applied, which is back-calculation [29]. With no integral action, Equation (1.19) is skipped. Clamping the running sum would instead discard each increment that crosses a limit and keep each one that returns, so symmetric noise would push the output off the limit even in a proportional-only loop.

The derivative acts on the measurement, so a setpoint step gives a proportional kick and no derivative kick. A change of K_c acts on the whole error at the next execution, as in the position form. The action $a = +1$ suits a process whose PV rises with the MV, a reverse-acting controller in the ISA convention, and $a = -1$ gives the direct-acting one.

In Manual the controller keeps executing and sets $I_k = u_{\mathrm{man}} - u_b - S_k$, so the output it would send equals the manual value u_{man}. The first automatic output differs from it by one period's change in P and D. In Auto the engine copies each controller output to the manual station, so a switch to Manual is also bumpless. The standalone studios with PID loops use velocity-form code of their own, with the limit on the output and anti-windup on the integral alone, in which a gain change applies only to later increments.

1.6.3 The Predictive Control and Estimator Options

More than 100 apps offer model predictive control (MPC) as an option beside the structure they open with, which stays the default. Volume 9 develops the theory. In most of them the controller plans with a small linear model in discrete time as shown in Equation (1.20).

$$x_{k+1} = A\,x_k + B\,u_k + E\,w_k + c, \qquad y_k = C\,x_k \tag{1.20}$$

where w is a measured or scheduled disturbance and c is a constant left by linearization. The model comes from a step test at the design point, from the app's balances linearized there or again at every decision, or from the model the reader fits. It never reads the true states or parameters of the run, so a parameter the reader changes in Show Model is model mismatch, as on a real plant.

At each decision the controller minimizes the objective of Equation (1.21) over a horizon of N steps.

$$J = \sum_{k=1}^{N} (y_k - r_k)^{\top} Q (y_k - r_k) + \sum_{j} \Delta u_j^{\top} R \, \Delta u_j + \sum_{i} \rho_i s_i^2 \tag{1.21}$$

where r is the setpoint, Q weights the tracking error, R is the move suppression on each input change Δu_j and ρ_i is the price of a slack $s_i \geq 0$ that lets an output band be broken when it cannot be held. Actuator limits are hard constraints. The quadratic program is solved in the browser by a dual active-set method [30]. The first move is applied, and the plan is made again at the next decision.

The plan starts from an estimate of the model state. About half of these apps use a steady-state Kalman filter on the model with an added disturbance state, which absorbs what the model did not predict and removes steady offset without an integral term [31]. Others correct the model output with a filtered bias to the same end. About a dozen apps also offer moving horizon estimation (MHE), which solves the estimation problem over a window of recent samples and can hold bounds that a Kalman filter cannot, such as a concentration that is never negative [32].

With the MPC selected, the tuning fields of the loops it replaces leave the Control panel. In most of these apps the Advanced level shows up to three settings, each a multiple of the shipped value: the prediction horizon, the move suppression and the band price. The trend draws the accepted plan ahead of the present time, and a list of constraint prices shows which limits bind. A plan that is not optimal, not finite or outside its limits is rejected, and a fallback, in most apps the shipped controller, holds the actuator until a plan succeeds. The Model Fitting and Tuning Studio builds the same kind of controller from the reader's own data in its tab, MPC from the fit.

1.6.4 Scenarios, Progress and Data

An app teaches through scenarios, the numbered cases at the top of the Control panel, nine to thirteen in most apps. Loading one sets the model parameters, the controller settings, the disturbances, the mode, the initial state and any scripted events. The information button beside the list opens a briefing in four parts: what the case is, why it matters, what to do and one detail for the practicing engineer. The Learn card holds the lesson back, with Predict and Try It before the run, hints on request and Explain afterward.

Each scenario has a completion rule that the app judges from the run on the simulated clock: a run length, the scripted events passed, a setting made by the reader, a state held for a stated time or a target met. Choosing a case earns nothing. A completed case is marked in the list, and a badge row shows the count and a level, Explorer after one case, Practitioner at half and Completed at all. Progress is kept in the reader's browser with no account. A certificate of progress or of completion can be issued from it, as a record of self-paced work and not an assessed credential.

One case, usually the last, is the graded challenge. It states a target, such as an error band to hold through a sequence of upsets. Changing the model parameters or the disturbances voids the attempt in most apps, and a record made with Reveal suggested tuning says so.

Show Model opens the equations, live values of the variables, the editable parameters and the assumptions. The trends plot SP and PV above the MV for each loop, with actuator limits and marks at setpoint moves, disturbances, mode changes and scripted events. The scorecard reports for each loop the integral of the absolute error (IAE), overshoot, settling time, MV travel and share of time at a limit, with the app's own measures, and it does not combine them into one score. IAE and travel accumulate from Reset, so two runs are compared at the same clock reading.

Export CSV saves the log as a text file with a header row of column names that carry their units and one row per logged sample. In most apps the time column is in seconds, minutes or hours according to the run length. The log holds the most recent 60 000 samples unless the app sets another size. Apps with the MPC option begin the file with comment lines that start with #, which a script must skip.

An app exercise in this book names the scenario by its number and name as the app shows them, every setting that differs from the scenario's defaults, the run length in simulated time, the seed

(the default seed unless stated), what to read and where (a trend, a scorecard tile or a column of the exported file) and a tolerance on a numeric answer. Settings are made before Run unless the exercise gives a time.

1.7 Working With AI Agents While Studying

A reader of this book is likely to have an AI assistant within reach, and the apps make its work easy to test. This section gives the editor's view of how to use one while working through a chapter. In short, the assistant may do the algebra and the scripting, and the reader keeps the checking.

What to delegate. Tasks with a clear input and a result that can be tested are good candidates. An assistant can derive the linearization of a chapter's balance at an operating point and the transfer function that follows from it. It can write a short script that reads the exported CSV file and fits a first-order-plus-dead-time model to a step test. It can propose a first tuning from that fit with a named rule, convert units, explain a term in a model or outline a report. The assistant knows nothing about the app until it is told. Give it the equations and parameter values from Show Model, the column names of the exported file and the scenario that produced the data, as described in Section 1.6.4. Ask it to state its assumptions and the units of every result.

What to check, and how. An assistant returns a fluent answer whether or not the answer is right, so each result needs a test that it could fail. Four tests cover most of the errors a chapter will produce.

- *Units.* Every term of an equation carries the same units, and a process gain is in PV units per MV unit. A gain computed with the valve as a fraction is 100 times the gain of a loop whose valve is in percent.
- *Signs.* Decide from the physics which way the PV moves when the MV rises. Compare that with the sign of the fitted gain and with the controller action of Section 1.6.2.
- *Steady state.* Set the derivatives to zero. The model must return the operating point that Show Model reports, and a linearization must give zero deviation there.
- *A run on the app.* Load the named scenario at the default seed, enter the proposed numbers and run. The trend and the scorecard decide. A tuning that reads well on paper and holds the valve at a limit for half the run has failed.

A fitted model gets one more test: compare it with data it was not fitted to, such as a second step in the opposite direction. When a check fails, report the failure to the assistant with the evidence and ask for the cause, then check the correction the same way.

How the exercises train the habit. One exercise in each chapter is marked *Review the agent*. It presents a tuning, a model or a claim as an AI assistant might propose it, with a flaw, and asks the reader to find the flaw in about five minutes. The flaws are ordinary ones: a unit, a sign, a steady state that does not close, a result measured on the wrong scenario or at a different run length, or a limit that was ignored. Each is found with one of the tests above. The purpose is not distrust. An engineer who can say why an answer is right can hand an assistant much larger tasks and still stand behind the result. That larger picture, in which agents do most of the routine work and the engineer supervises and checks, is the subject of Chapter 16.

1.8 Chapter Roadmap

Table 1.4 lists the chapters with the lesson each adds to this one. The level says how much of Section 1.4 a chapter assumes. The time is an estimate for reading the chapter once and running each case of its app once, at about eight minutes a case.

Chapters 3, 13, 14 and 15 form a first course: a heater with one limit, a trim on a late analyzer, an integrating header and a blend. They rest on Sections 1.4.2, 1.4.5 and 1.4.6. Chapters 2, 4, 8 and 9 add

Table 1.4 Roadmap of the Volume

Ch.	Plant	The lesson	Level	Time
2	Crude desalter	A salt minimum inside the mixing range and an interface reading that drifts with the rag.	intermediate	2.5 h
3	Crude preheat train	Fouling hides in the fuel valve until the charge is the only handle left.	introductory	3 h
4	Atmospheric crude tower	Count the degrees of freedom before pairing loops.	intermediate	2.5 h
5	Vacuum tower	The cheapest handle on the cut rests on an ejector curve that breaks.	advanced	2 h
6	Fluid catalytic cracker	A unit that is its own furnace, and a selector over five limits.	advanced	3.5 h
7	Hydrocracker reactor	Beds that amplify their inlet, protected in layers.	advanced	2.5 h
8	Diesel hydrotreater	Two inlet handles on two clocks and a half-hour analyzer.	intermediate	2 h
9	Catalytic reformer	Four heaters on one octane that is measured late.	intermediate	2 h
10	Sulfuric acid alkylation	Three stocks on three clocks behind one octane.	advanced	2.5 h
11	Delayed coker	A drum switch is a disturbance known hours ahead.	intermediate	2 h
12	Hydrogen plant	The steam must lead the feed, and the waste is the fuel.	intermediate	2 h
13	Claus sulfur recovery	An optimum turned into a zero, behind three minutes of delay.	introductory	2.5 h
14	Fuel gas header	An integrator on a split range whose ends cost money both ways.	introductory	2 h
15	Gasoline blender	Authority over a tank average shrinks as the tank fills.	introductory	2 h

degrees of freedom, heat integration and inferred quality, and Chapters 11 and 12 add a scheduled disturbance and a ratio that must lead. Chapters 5, 6, 7 and 10 combine a moving limit with internal feedback or slow measurement and are best left until last. A reader who follows the crude can read the volume in order, from the desalter to the gasoline tank.

KEY POINTS

- A refinery unit earns more the more oil it processes, until a product specification or an equipment limit stops it.
- A fired heater's heat balance is also the most useful feedforward in a refinery.
- Four cuts fix the eight end specifications of four products, so half are choices and half are consequences.
- A unit that returns part of its own heat amplifies every change and slows its own response by the same factor.
- Most economic optima lie on a limit, so the layer that pushes throughput must know which limit binds.

Where else this applies. The same layers and limits return in every industry volume of the series. Selectors and overrides are developed in Volume 4 and predictive control in Volume 9.

References

[1] Occupational Safety and Health Administration, "OSHA technical manual, section IV, chapter 2: Petroleum refining process," U.S. Department of Labor, Washington, DC, osha.gov/otm/section-4-safety-hazards/chapter-2, read November 2026.

[2] F. G. Cottrell and J. B. Speed, "Separating and collecting particles of one liquid suspended in another liquid," U.S. Patent 987 115, Mar. 21, 1911.

[3] D. Splitter, A. Pawlowski, and R. Wagner, "A historical analysis of the co-evolution of gasoline octane number and spark-ignition engines," *Frontiers in Mechanical Engineering*, vol. 1, p. 16, 2016.

[4] E. T. C. Vogt and B. M. Weckhuysen, "Fluid catalytic cracking: recent developments on the grand old lady of zeolite catalysis," *Chemical Society Reviews*, vol. 44, no. 20, pp. 7342–7370, 2015.

[5] U.S. Environmental Protection Agency, "Control of air pollution from new motor vehicles: Heavy-duty engine and vehicle standards and highway diesel fuel sulfur control requirements," *Federal Register*, vol. 66, pp. 5002–5193, Jan. 2001, document 01-2, published January 18, 2001.

[6] M. Pitt, "Taking a look back at control: Part 2," *The Chemical Engineer*, no. 989, Nov. 2023, thechemicalengineer.com/features/taking-a-look-back-at-control-part-2.

[7] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[8] M. Schwenzer, M. Ay, T. Bergs, and D. Abel, "Review on model predictive control: an engineering perspective," *The International Journal of Advanced Manufacturing Technology*, vol. 117, no. 5–6, pp. 1327–1349, 2021.

[9] M. J. Kaiser, A. de Klerk, J. H. Gary, and G. E. Handwerk, *Petroleum Refining: Technology, Economics, and Markets*, 6th ed. CRC Press, 2019.

[10] U.S. Energy Information Administration, "Refinery capacity report 2026, table 1: Number and capacity of operable petroleum refineries by PAD district and state as of January 1, 2026," U.S. Department of Energy, Washington, DC, Tech. Rep., 2026.

[11] M. R. Riazi, *Characterization and Properties of Petroleum Fractions*. ASTM International, 2005.

[12] M. Morbidelli and A. Varma, "Parametric sensitivity and runaway in fixed-bed catalytic reactors," *Chemical Engineering Science*, vol. 41, no. 4, pp. 1063–1071, 1986.

[13] W. M. Edwards and H. N. Kim, "Multiple steady states in FCC unit operations," *Chemical Engineering Science*, vol. 43, no. 8, pp. 1825–1830, 1988.

[14] A. Arbel, I. H. Rinard, R. Shinnar, and A. V. Sapre, "Dynamics and control of fluidized catalytic crackers. 2. multiple steady states and instabilities," *Industrial & Engineering Chemistry Research*, vol. 34, no. 9, pp. 3014–3026, 1995.

[15] C. van Heerden, "Autothermic processes," *Industrial & Engineering Chemistry*, vol. 45, no. 6, pp. 1242–1247, 1953.

[16] H. Müller-Steinhagen, "Heat transfer fouling: 50 years after the Kern and Seaton model," *Heat Transfer Engineering*, vol. 32, no. 1, pp. 1–13, 2011.

[17] F. Coletti and S. Macchietto, "A dynamic, distributed model of shell-and-tube heat exchangers undergoing crude oil fouling," *Industrial & Engineering Chemistry Research*, vol. 50, no. 8, pp. 4515–4533, 2011.

[18] F. Reyes and W. L. Luyben, "Steady-state and dynamic effects of design alternatives in heat-exchanger/furnace/reactor processes," *Industrial & Engineering Chemistry Research*, vol. 39, no. 9, pp. 3335–3346, 2000.

[19] V. Pinjala, Y. C. Chen, and D. Luss, "Wrong-way behavior of packed-bed reactors: II. impact of thermal dispersion," *AIChE Journal*, vol. 34, no. 10, pp. 1663–1672, 1988.

[20] B. Joseph and C. B. Brosilow, "Inferential control of processes: Part I. steady state analysis and design," *AIChE Journal*, vol. 24, no. 3, pp. 485–492, 1978.

[21] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[22] M. Morari, Y. Arkun, and G. Stephanopoulos, "Studies in the synthesis of control structures for chemical processes: Part I: Formulation of the problem. process decomposition and the classification of the control tasks. analysis of the optimizing control structures," *AIChE Journal*, vol. 26, no. 2, pp. 220–232, 1980.

[23] S. Skogestad, "Plantwide control: The search for the self-optimizing control structure," *Journal of Process Control*, vol. 10, no. 5, pp. 487–507, 2000.

[24] M. L. Darby, M. Nikolaou, J. Jones, and D. Nicholson, "RTO: An overview and assessment of current practice," *Journal of Process Control*, vol. 21, no. 6, pp. 874–884, 2011.

[25] A. Reyes-Lúa and S. Skogestad, "Systematic design of active constraint switching using classical advanced control structures," *Industrial & Engineering Chemistry Research*, vol. 59, no. 6, pp. 2229–2241, 2020.

[26] E. M. B. Aske, S. Strand, and S. Skogestad, "Coordinator MPC for maximizing plant throughput," *Computers & Chemical Engineering*, vol. 32, no. 1–2, pp. 195–204, 2008.

[27] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[28] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[29] K. J. Åström and L. Rundqwist, "Integrator windup and how to avoid it," in *Proc. 1989 American Control Conference*, 1989, pp. 1693–1698.

[30] D. Goldfarb and A. Idnani, "A numerically stable dual method for solving strictly convex quadratic programs," *Mathematical Programming*, vol. 27, no. 1, pp. 1–33, 1983.

[31] G. Pannocchia and J. B. Rawlings, "Disturbance models for offset-free model-predictive control," *AIChE Journal*, vol. 49, no. 2, pp. 426–437, 2003.

[32] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 246–258, 2003.

1.9 Exercises

1.1 Theory and calculation · 5 min

R-1's crude unit is clean, with $UA = 1.5\,\mathrm{MW/K}$, crude entering the train at 40 °C and products at 340 °C, as in Section 1.4.5. The charge falls to 650 t/h and the COT setpoint rises to 370 °C, with $c_p = 2.5\,\mathrm{kJ/(kg\,K)}$.

(a) Find the preheat temperature. *Checkpoint:* 270.6 °C.

(b) Find the heater duty and the firing as a share of 60 MW. *Checkpoint:* 44.9 MW, 74.8 %.

1.2 Concept check · 2 min

The fuel gas of a refinery takes up hydrogen, and its heating value falls 7 % while its Wobbe index falls 3 %. Two heaters are on the header. Heater A holds its fuel flow on a flow loop, and heater B holds its burner pressure. Before their outlet loops act, which statement is right?

(a) both lose 7 % of their heat, because the heat is the heating value times the gas burned

(b) heater A loses about 3 % and heater B about 7 %, because a lighter gas flows faster through a burner

(c) heater A loses about 7 % and heater B about 3 %, because a burner at fixed pressure passes more of a lighter gas

(d) neither loses heat, because hydrogen burns hotter than the gas it replaces

1.3 Theory and calculation · 4 min

R-1's tower draws naphtha, kerosene, diesel and gas oil at 144, 72, 144 and 50.4 t/h from 720 t/h of crude whose assay is $T_{bp} = 30 + 600\,v$.

(a) The kerosene draw rises 14.4 t/h with the others held. Find the new temperatures of the three cuts below the naphtha. *Checkpoint:* 222, 342 and 384 °C.

(b) Which draw should move, and by how much, so that only the kerosene and diesel cut moves? *Checkpoint:* diesel down 14.4 t/h.

1.4 Concept check · 2 min

An engineer pairs one loop to hold the naphtha end point on the top reflux and a second loop to hold the kerosene flash point on the kerosene draw of an atmospheric tower. What happens?

(a) the loops fight, because both specifications are set mostly by one and the same cut

(b) both loops settle, because each product has its own draw and therefore its own handle

(c) the kerosene loop moves every cut above the kerosene draw, so the naphtha loop saturates

(d) the loops settle only if the kerosene loop is tuned five times slower than the reflux loop

1.5 Theory and calculation · 4 min

R-1's diesel draw tray boils at 366.23 °C at 2.0 bar for a 330 °C cut, with $A = 12$ and $p_{atm} = 1.013\,25$ bar in Equation (1.13). The tower pressure falls to 1.9 bar with the cut unchanged.

(a) Find the new tray temperature. *Checkpoint:* 363.35 °C.

(b) A loop holds the raw tray temperature at 366.23 °C. In which direction does it move the cut, and by about how much? *Checkpoint:* heavier, by about 2.7 K.

1.6 Theory and calculation · 4 min

A bed of the form of Equation (1.5) has $\Delta T_{ad} = 100\,\mathrm{K}$, converts $x = 0.3$ at design and its conversion grows by a factor e every 20 K of mean temperature.

(a) Find the runaway index r and the steady gain G. *Checkpoint:* $r = 0.75$, $G = 7$.

(b) By what factor is the bed's response slower than its time constant τ_b? *Checkpoint:* 4.

1.7 Theory and calculation · 3 min

A feed and effluent exchanger returns $\varepsilon_u = 0.4$ of the effluent's temperature above the feed to the bed inlet, as in Equation (1.11).

(a) Find the loop gain and the amplification of a heater move for a bed with $G = 1.5$. *Checkpoint:* $L = 0.6$, $A = 2.5$.

(b) Find the same for R-1's bed, $G = 2$, and the largest ε_u that keeps that bed open-loop stable. *Checkpoint:* $L = 0.8$, $A = 5$, $\varepsilon_u < 0.5$.

1.8 Concept check · 2 min

A slide valve on a catalytic cracker opens by three percent in Manual. The riser outlet rises 2.1 K in two minutes and about 7.2 K after ninety. Which statement explains the difference?

(a) the riser thermocouple has a ninety-minute lag, so the first reading is simply late

(b) the regenerator's air blower responds to the valve and adds heat after a delay

(c) the cracking reaction is exothermic, so a hotter riser releases more heat of its own

(d) the extra catalyst makes more coke, and burning it reheats the returning catalyst

1.9 Theory and calculation · 5 min

R-1's preheat train has fouled to 60 % of its clean UA, so $UA = 0.9\,\mathrm{MW/K}$. The valve position override holds the firing at 95 % of 60 MW and the COT at 365 °C.

(a) Estimate the charge the override allows if the preheat stayed where it is at 720 t/h. *Checkpoint:* 621 t/h.

(b) Check that 653.6 t/h meets the firing target once the preheat is recomputed at that charge. *Checkpoint:* preheat 239.4 °C, duty 57.0 MW.

1.10 Theory and calculation · 5 min

A gasoline tank's target is 87.2. Use Equation (1.15).

(a) The tank is 75 % full at an average of 87.0. At what value must the header run for the rest of the batch? *Checkpoint:* 87.8.

(b) The tank is 92 % full at 87.0, and the header may run at most 1.5 above target. With the header at that limit until 95 % full and at target after, where does the tank finish? *Checkpoint:* 87.06.