

# Principles of PID Controllers

Release date: July 2023

## Introduction

Self-regulating systems using feedback loops, i.e. the routing back of the output of a system to its input, have existed since antiquity and have nowadays become an integral part of modern technology. One of the first attempts to rigorously describe control loops using feedback traces back to more than 150 years ago with James Clark Maxwell's article, *On Governors* [1].

In the context of control strategies, an open-loop control system refers to a controller whose action is determined based on predetermined input values without considering feedback. In contrast, a closed-loop controller incorporates continuous feedback, enabling real-time adjustments to enhance precision, stability, and robustness, making it more suitable for achieving desired control objectives in changing conditions.

Today, the most widespread type of closed-loop control systems is the Proportional–Integral–Derivative (PID) controller. These types of controllers continuously measure and adjust the output of a system to match a desired setpoint, that is, a given target condition for the system or process under consideration. Requiring little prior knowledge or model of the system, PID controllers are extremely versatile, relatively cost effective and straightforward to implement, making their realization possible in a large variety of processes: from hydraulics and pneumatics to analog and digital electronics. For this reason, they have become extensively used in a variety of industries and research applications, including manufacturing, photonics, sensors, material science and nanotechnology.

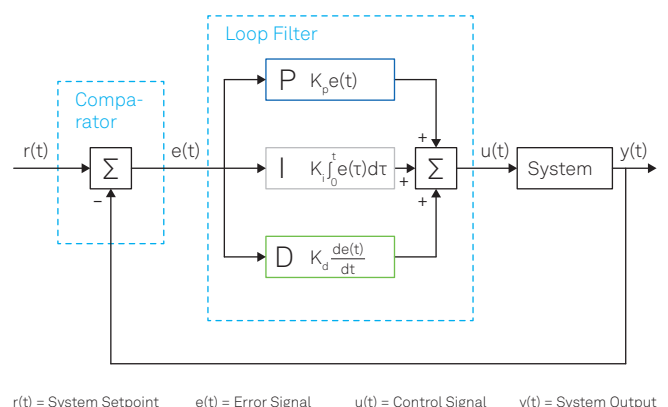
PID control loops are widely employed in various aspects of everyday-life and industrial automation, such as the gyroscopes found in smartphones and self-navigating cars, ovens used for cooking food or samples, flow controllers in pipes, and even in managing the daily vehicle traffic. At the same time, their presence stands out in more advanced research fields as well, for example in the stabilization of laser cavities and interferometers in optics and photonics, in closed-

loop control of MEMS-based (micro-electromechanical systems) gyroscopes, and in the characterization of mechanical resonators in scanning probe microscopy (SPM).

This white paper presents the key functions and principles of PID control loops by analyzing their basic building blocks, by describing their strengths and limitations, and by outlining the tuning and designing strategies and how they can be easily implemented with Zurich Instruments' lock-in amplifiers.

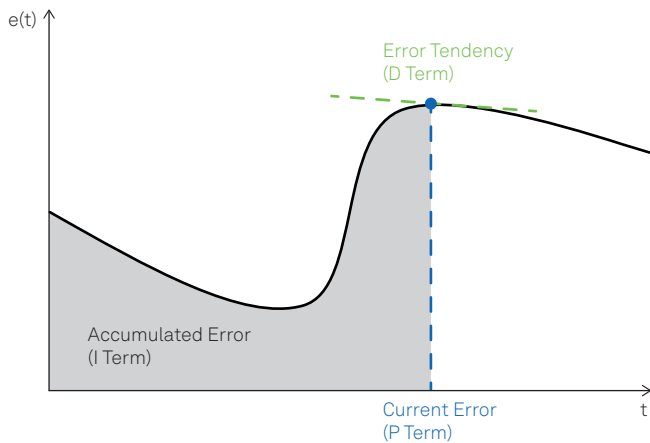
## PID working principle and building blocks

The goal of a PID controller is to produce a control signal that can dynamically minimize the difference between the output and the desired setpoint of a certain system. Let's consider the exemplary scheme depicted in Figure 1. As a first step, the output of the



**Figure 1.** Schematic representation of a general PID control loop in its most general form.

system  $y(t)$  is looped back and measured against the setpoint  $r(t)$  by the comparator, thereby generating the time-dependent error signal  $e(t) = r(t) - y(t)$ . Subse-



**Figure 2.** Example of error function with the highlighted contributions of the P, I and D terms.

quently, this error signal is minimized by the loop filter and then used to generate the control signal  $u(t)$  that is applied to drive the output of the system, initiating closed-loop operation. These steps are continuously executed to minimize the error; hence, apart from considering the current error, it is also relevant to consider its accumulation over time (represented by the integral) and its future tendency (represented by the derivative at time  $t$ ), as shown in Figure 2.

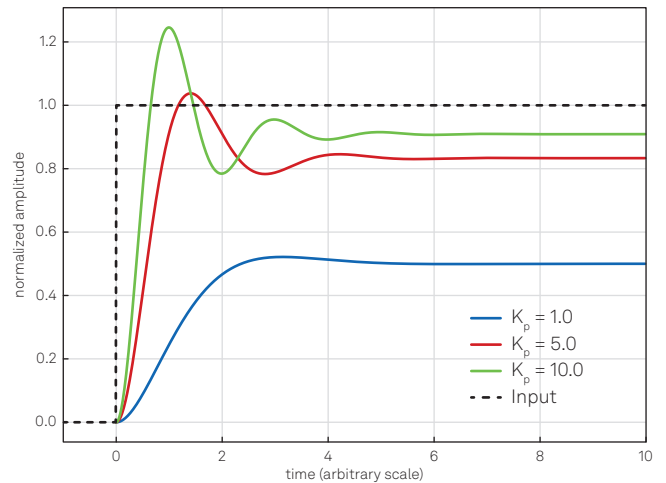
In the most general case, error minimization is accomplished by means of the three primary components of the PID controller loop filter: the proportional, integral, and derivative terms. Mathematically, the complete control function in its most general form can be written as the sum of the three individual contributions:

$$u(t) = u_P(t) + u_I(t) + u_D(t) \\ = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{d}{dt} e(t), \quad (1)$$

where  $K_p$ ,  $K_i$  and  $K_d$  are the gain coefficients related to the proportional, integral, and derivative terms, respectively.

### The proportional term

The proportional term, denoted with P, is based on the current error between the setpoint and the measured output of the system. This term helps bring the output of the system back to the setpoint by applying a correction that is proportional to the amplitude of the error, leading to a reduction of the rise time of the correction signal, see Figure 3. The larger the error, the larger the correction applied by the proportional term – that is, the larger the error with a fixed  $K_p$ , the larger  $u_P(t)$ . Since the P term always requires a non-zero error to generate its output, it cannot nullify the error by itself. In steady-state system conditions, an equilibrium is reached, which includes a steady-state error.



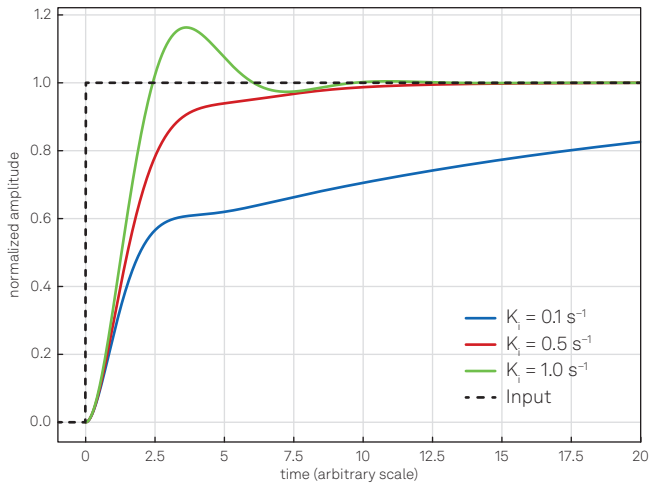
**Figure 3.** Effect of the proportional action. Increasing the  $K_p$  coefficient reduces the rise time, but the error never approaches zero. Additionally, a too high value of the proportional gain might lead to an oscillating output.

### The integral term

The integral term, denoted with I, applies a correction that is proportional to the time integral of the error, i.e. the history of the error. For example, if the error persists over time, the integral term continues to increase, resulting in a larger correction applied to the output of the system. Unlike the proportional term, the integral term makes it possible for the controller to generate a non-zero control signal even under a zero-error condition at present. This property enables the controller to bring the system exactly to the required setpoint. Its effect is illustrated in Figure 4. Increasing the value of the integral gain coefficient increases the contribution of the accumulated error over time to the control signal. This means that if there is a steady-state error, an integral term with a large gain coefficient will drive the control signal to eliminate the error faster than a smaller integral term. However, increasing the integral term too much can lead to an oscillating output if too much error is accumulated, causing the control signal to overshoot and create oscillations around the setpoint. This phenomenon is sometimes called integral windup [2].

### The derivative term

The derivative term, denoted with D, provides a control over the error tendency, i.e. its future behavior, by applying a correction proportional to the time derivative of the error. This allows to reduce the rate of change of the error and so helps improve the stability and responsiveness of the control loop. The aim is to anticipate the changes in the error signal: if the error shows an upward trend, the derivative action tries to compensate without waiting for the error to become significant (proportional action) or for it to persist for some time (integral action). In real-world implementations of PIDs, the derivative action is sometimes omitted due to its high sensitivity to the quality of the input signal. When



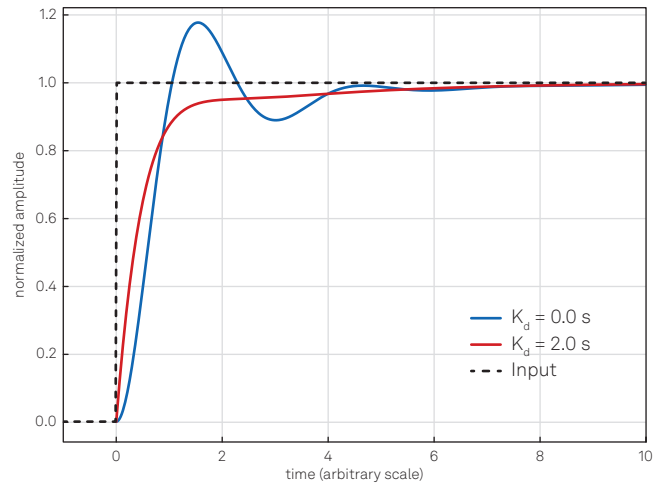
**Figure 4.** Effect of the integral action with constant  $K_p = 1$ . Increasing  $K_i$ , the response will be faster but also lead to larger oscillations and overshoot if the value increases too much (green curve).

the reference value changes rapidly, as in the case of a very noisy control signal, the derivative of the error tends to become very large, causing the PID controller to undergo an abrupt change that can result in instabilities or oscillations in the control loop. To improve the stability, prior low-pass filtering of the error signal is often used as a mitigation strategy; however, low-pass filtering and derivative control neutralize one another, hence only a limited amount of filtering is possible. If it is properly calibrated and if the system is "tolerant" enough, the derivative action can give a decisive contribution to the controller performance. The effect of the derivative term is shown in Figure 5.

The effect of each term on the system's response depends strongly on the system's characteristics. Therefore, the weighting of the  $K_p$ ,  $K_i$ , and  $K_d$  gains can be adjusted to fine-tune the performance of the control loop and achieve the desired responsiveness and accuracy. Some applications or simple systems may only require one or two of the three control terms provided by a PID controller. To operate the controller with only a subset of these terms, the unused terms can be set to zero, thus resulting in a PI, PD, P, or I controller. For instance, the use of a PI controller is common in applications that prioritize steady-state error elimination and stability, rather than fast response times, due to their slow dynamics. A typical example is the control of an oven's temperature, where a PI controller is normally employed to ensure precise temperature regulation and eliminate any steady-state offset, considering the oven's relatively slow response characteristics.

## Derivation of an initial set of parameters (Tuning)

One of the main advantages of PID controllers is that they can be implemented without knowledge or a detailed model of the system. Thanks to heuristic calibration procedures, it is possible to calculate coefficients



**Figure 5.** The purpose of the derivative action is to increase the damping of the system; however, too large values of  $K_d$  might make the system unstable or oscillatory, as described in the text. The curves are obtained keeping the proportional and integral gain constant ( $K_p = 4$  and  $K_i = 1 \text{ s}^{-1}$ ).

based only on simple experimental tests to be carried out directly on the process. Nevertheless, the initial tuning of the PID parameters can be a delicate task. There exist several well-established methods to derive an initial set of coefficients, many of them involving the measurement of some open-loop parameters of the system.

As pointed out in the introduction, open-loop refers to the behavior of the system without any feedback control, where an input signal is applied to the system and the resulting output is just measured, but not fed back to the input. The input signal can be a step function, a ramp, a sine wave, or any other type of signal that is appropriate for the system being controlled. The output of the system is then recorded as a function of time or frequency and can be analyzed to determine the system's response characteristics, such as its time constant, natural frequency, and damping ratio.

The general strategy to find initial parameters involves typically three steps:

1. Obtain the open-loop response of the system and measure some of the characteristic parameters – e.g., the oscillation period of the system output and the process delay.
2. Calculate coarse values of the gain coefficients  $K_p$ ,  $K_i$  and  $K_d$  based on the measured parameters.
3. Tune the PID gain coefficients to optimize for noise, speed, or robustness.

Some of the most widespread techniques to achieve coarse tuning are the Ziegler-Nichols method [3], the Cohen-Coon method [4], the relay method [5] and the Tyreus-Luyben method [6].

In Figure 6, a practical example of an initial tuning procedure is outlined, constructed from the Ziegler-Nichols method.

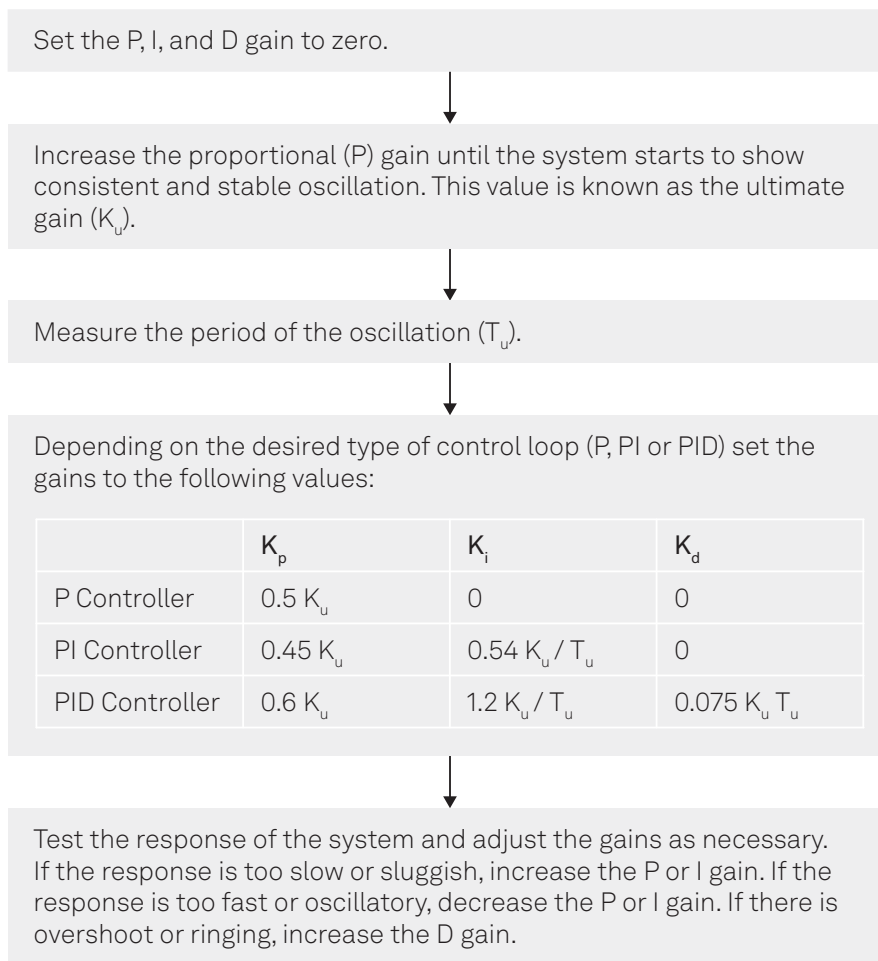


Figure 6. Step-by-step procedure for the initial tuning of a PID controller, based on the Ziegler-Nichols method.

## Analog and digital PID controllers

In recent times, thanks to the advent of field-programmable gate array (FPGA) technology, micro-processors and digital signal processing techniques, the digital implementation of PID controllers – together with many other tools – has outperformed its analog counterpart in a number of aspects. In particular, the digital form of PID control has the great advantage that it can be easily implemented as an algorithm and executed by a microcontroller device. This vastly increases its flexibility, performance, and the number of addressable applications.

Digital and analog PID controllers differ in the way they process signals and perform control actions. Analog controllers use continuous signals and analog components such as operational amplifiers and resistors to perform computations and generate control signals. Conversely, digital PID controllers sample and digitize signals, perform calculations and output discrete control signals. The use of digital signal processing techniques allows digital controllers to execute more complex control algorithms without adding extra analog circuitry whose physical properties might drift and thus

degrade the quality of the control action over time. Furthermore, the storing/recalling of settings and the integration with other digital systems can be easily achieved with digital controllers, thereby increasing their adaptability to many more control problems and applications. As with all implementations of digital signals, the most common issues of digital PID controllers are related to quantization and sampling. Operating at discrete time intervals, selecting the sampling rate is critical to avoid artifacts such as aliasing and ensure accurate control. Furthermore, the implementation of the algorithms on digital platforms (e.g. on FPGAs or microprocessors) might be more costly and require additional technical expertise and careful evaluation of the corresponding numerical methods used for signal calculation and conditioning. Ultimately, the choice between analog and digital controllers often depends on the application requirements, the available resources, and the desired performance. Table 1 provides an overview of the advantages and disadvantages of the two approaches.

Zurich Instruments offers digital PID controllers as upgrade options for its lock-in amplifiers, impedance analyzers and boxcar averagers. The integrated architec-

Feature	Analog PID Controller	Digital PID Controller
Cost	+ (less expensive)	- (more expensive)
Precision	- (limited precision)	+ (high precision)
Flexibility	- (difficult to modify)	+ (easy to modify)
Ease of implementation	+ (simple to implement)	- (may require technical expertise)
Noise sensitivity	- (more sensitive to noise)	+ (less sensitive to noise)
Dynamic range	- (limited)	+ (wide)
Stability	- (prone to drifting)	+ (stable)
Response time	+ (fast)	-/+ (can be made faster with faster processors, but generally not as fast as analog ones)

**Table 1.** Comparison between analog and digital PID control loops.

ture of the instruments allows the user to choose from many different inputs and outputs signals for the PIDs. Additionally, by integrating demodulated signals as input choices for PID controllers, the users can benefit from several advantages:

- Enhanced signal-to-noise ratio and stability: demodulation isolates the desired signal from noise and interference, resulting in a cleaner input signal for the controller.
- Increased sensitivity and improved system response: lock-in detection amplifies weak signals, enabling finer adjustments and enhancing the PID controller's ability to respond to subtle changes.
- Elimination of unwanted frequency components: demodulation filters out unwanted frequency components present in the original signal, leading to more accurate and efficient control.

As depicted in Figure 7, the set of signal choices includes amplitude, phase (thereby effectively setting up what is known as a phase-locked loop, or PLL for short [7]), quadrature and in-phase components of demodulated signals, boxcar outputs, as well as auxiliary inputs and outputs. The integrated digital signal processing guarantees maximum signal-to-noise ratio, minimal feedback loop latency and closed-loop operation with high stability. The feedback signal is then available as an analog output and can also be directly applied to an internal signal generator output to control its amplitude, frequency, offset and phase. The digital implementation makes it possible to embed multiple PID controllers on the same instrument, whose inputs and outputs can be either cascaded in a sequential fashion or used in a parallel configuration, that is, independently of one another. Thanks to the type of architecture depicted in Figure 7, the PID option embedded in lock-in amplifiers enhances the versatility of the instruments, making them benefit a large variety of applications – even though the specific application needs might significantly differ from one another.

As previously mentioned, determining appropriate starting conditions for a PID controller and its optimization can be a challenging task. LabOne® control software, the software of Zurich Instruments

interfacing with the hardware, offers a kit of several tools to streamline the process of setting up and optimizing a PID controller, making the task more efficient and straightforward. LabOne tools like the Sweeper, the PID Advisor and Autotune can help the user with the implementation of PID control loops, providing intuitive adjustability of the bandwidth, speed and initial tuning parameters, according to the desired performance. For example, the Sweeper can be used to obtain the open-loop response of the device under test (DUT) through signal sweeping, then the PID Advisor selects initial feedback gain parameters based on predefined transfer functions and the DUT open-loop parameters obtained with the Sweeper measurement. After closing the loop, to minimize residual error signal and enhance control performance, the Auto Tune feature can be applied to the feedback signal to dynamically adjust feedback gain parameters.

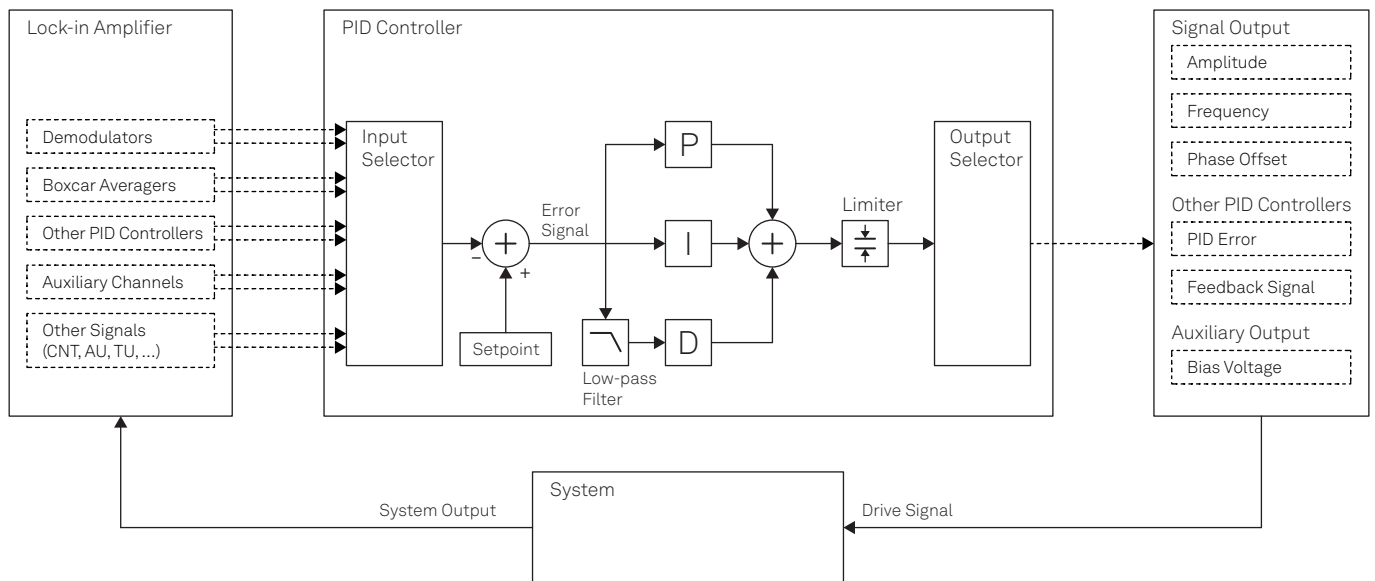
Additionally, thanks to LabOne's support for the most popular programming languages (Python, C/C++, MATLAB, LabVIEW and .NET), maximum versatility is ensured when setting up a control loop and its successive optimization and refinement.

## Challenges and trade-offs

While their versatility and simplicity allow PID controllers to be applied to numerous control problems, they are subject to some limitations and cannot deliver so-called optimal control. Optimal control refers to the set of control signals that minimize or maximize a certain criterion (sometimes known as cost function) used to measure the performance of the controlled system, e.g. minimizing energy consumption.

Specifically, the PID controllers' reactive nature, i.e., the fact that no explicit knowledge of the system is used to generate the control coefficients, makes them sensitive to changes in the system dynamics and may require careful tuning to achieve the desired performance. Some typical shortcomings may include:

- Performance degradation if the system has a complex or non-linear transfer function. In these cases, more advanced and computationally heavy model-based control methods, such as state-



**Figure 7.** Architecture of the digital implementation of the PID controllers in Zurich Instruments Lock-in Amplifiers. Depending on the specific application, different input and output signals can be used.

space control or H-infinity control [8], may be needed.

- Considerable sensitivity to noise and disturbance signals of the D term, affecting the stability and accuracy of the control loop. Mitigation strategies can be applied to partly address these issues – e.g. omission or low-pass filtering of the D component.
- Sensitivity to deadtime or phase delay, i.e., the time it takes for the control loop to respond to a change in the system or the setpoint. If the phase delay is too large, the control loop may become unstable and start to oscillate or exhibit other undesirable behaviors.
- Process saturation: this phenomenon affects physical systems by introducing nonlinearities in the process due to the saturation of the actuators (e.g., a motor cannot supply more than its maximum torque). Techniques such as anti-windup or setpoint weighting can help mitigate these issues with relatively minor modifications to the control loop [9].

### Loop bandwidth

Another important factor to consider when designing a PID controller is the bandwidth of the loop or the closed-loop bandwidth. This bandwidth specifies the range of frequencies over which the loop can effectively control the system. A wide-bandwidth loop responds rapidly to changes in the setpoint or the system but may also be more prone to oscillation and overshoot. A narrow-bandwidth loop, in contrast, is less sensitive to changes in the setpoint and can deliver a cleaner and more accurate control signal at the expense of a slower response time. Given this tradeoff between speed and precision, it is important to understand the requirements of the considered application to select an ap-

propriate loop bandwidth. A wide-bandwidth loop may be preferred in applications where a rapid response is required, such as in the control of motors or other mechanical systems. Conversely, a narrow-bandwidth loop may be best in applications where stability is more important, such as in the control of chemical processes or for other systems with long time delays.

## Conclusion

Thanks to their ability to adjust the system's output accurately and quickly without detailed knowledge about its dynamics, PID control loops stand as a powerful and widely-used tool for maintaining a stable and predictable output in a variety of applications. In this paper, we have reviewed the fundamental principles and characteristics of these control systems, providing insight into their functioning, tuning strategies, advantages and trade-offs. As a result of their integrated architecture, Zurich Instruments' lock-in amplifiers allow users to make the most of all the advantages of digital PID control loops, so that their operation can be adapted to match the needs of different use cases.

## References

- [1] J. K. Maxwell. On governors. Proc. R. Soc. Lond., 16:270–283, 1868.
- [2] Y. Tian and D.P. Atherton. Analysis and Design of Reset Control Systems. Institution of Electrical Engineers, 2004.
- [3] J. G. Ziegler and N. B. Nichols. Optimum settings for automatic controllers. Transactions of the A.S.M.E., 64(8):759–768, 1942.

- [4] H. Cohen and G. A. Coon. A linear operator approach to the analysis of closed-loop servomechanisms. *Journal of the Franklin Institute*, 256(6):759–768, 1953.
- [5] J. K. Åström and T. Hägglund. *PID controllers: Theory, design, and tuning* (2nd ed.). Instrument Society of Automation, 1995.
- [6] W. L. Luyben. Plantwide control: Some practical considerations. *Chemical Engineering Progress*, 86(4):49–53, 1990.
- [7] Zurich Instruments. [Phase-Locked Loops for Analog Signals](#), 2022.
- [8] F. Lin, H. Yang, and Z. Lin. *Robust Control Design: An Optimal Control Approach*. CRC Press, 2013.
- [9] T. Hägglund and K.J Astrom. *PID Controllers: Theory, Design, and Tuning*. Instrument Society of America, 1995.

