Control engineering (LPF and STM feedback)

Naoya Kawakami

October 25, 2022
version 1.0

Contents

Chapter 1 Motivation for writing this text 1

Chapter 2 Evaluation of LPF based on control engineering 2
2.1 LPF . . . 2
2.2 Laplace transform . . . 3
2.3 Transient response of the system . . . 4

Chapter 3 Transit response of LPF for various inputs 8
3.1 Impulse response . . . 8
3.2 Step response . . . 8
3.3 Response for an arbitrary function . . . 9
3.4 Frequency response . . . 9

Chapter 4 Feedback analysis 11
4.1 Basic concept of feedback . . . 11
4.2 Feedback circuit in STM . . . 11
4.3 Responce of feedback circuit . . . 14
4.4 Simulating STM image . . . 14
4.5 Closing remark . . . 16

Chapter 5 Appendix 17
5.1 The basic Laplace transform . . . 17
5.2 Patial fraction decomposition . . . 17

Bibliography 19

Chapter 1

Motivation for writing this text

A low pass filter (LPF) is often used in the electric circuit to cut unnecessary signals. In scanning tunneling microscopy/spectroscopy (STM/S) systems, LPF is used in the feedback circuit and lock-in amplifier. In my document entitled "STM principles," LPF is introduced in section 2.4 as a part of lock-in detection. However, I introduced several formulas without explaining the derivation process because it needs knowledge about control engineering. Therefore, in this document, I will introduce the basics of control engineering to explain the mathematical background of LPF.

Furthermore, control engineering is also useful in analyzing the response in feedback systems. As you know, a feedback circuit keeps the distance between the sample and the tip in the STM system. Unfortunately, the electric circuit in the STM is a kind of "black box," so many users empirically adjust the feedback parameters (such as p-gain and time constant). Understanding the relationship between the parameters and feedback responce would help get better-quality data.[*1] Therefore, a detailed description of the feedback circuit in STM will be discussed. In addition, possible artifacts caused by feedback settings will also be introduced.

[*1] Actually, the characteristics of the STM circuit differ depending on their setup. So even after learning the details of feedback, you still need much experience. In that meaning studying the feedback might not be very helpful. If you are still interested in the feedback in STM, keep reading this text.

Chapter 2

Evaluation of LPF based on control engineering

The primary purpose of control engineering is to understand the relationship between the system's input and output signal. In this chapter, the basics of control engineering will be introduced.

2.1 LPF

Figure 2.1 is the electric circuit of the simplest LPF. An LPF comprises a resistor (with the resistance R) and a capacitor (with the capacitance C). Our interest is the relationship between the input signal v_i and output signal v_o. By applying Ohm's law, we get

$$v_i(t) = v_o(t) + Ri(t). \tag{2.1}$$

As the current (i) is expressed by

$$i(t) = \frac{dQ(t)}{dt} = C\frac{dv_o(t)}{dt}, \tag{2.2}$$

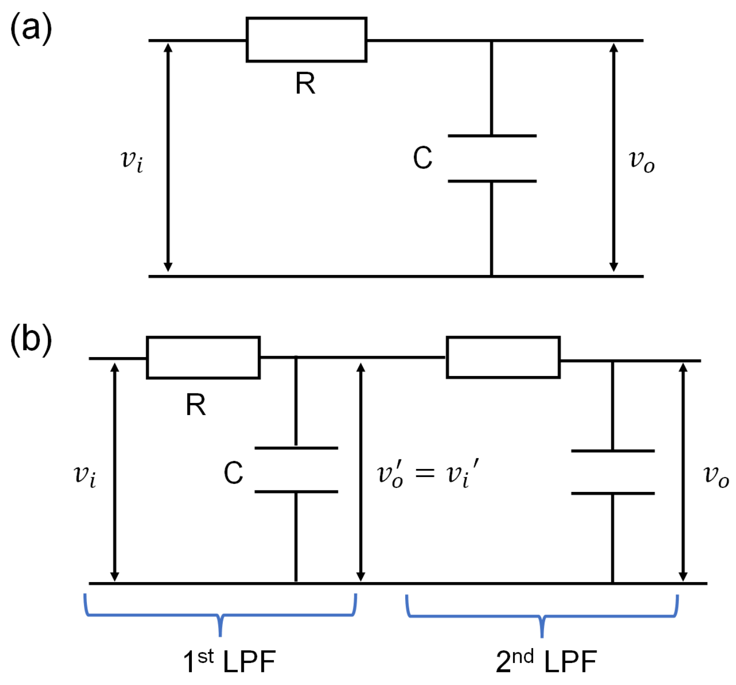


Fig. 2.1: (a) The most simple circuit diagram of LPF. (b) A circuit diagram of a series of two LPFs (called cascade).

the relationship between the input and output is expressed by the following formula:

$$v_i(t) = RC\frac{dv_o(t)}{dt} + v_o(t). \tag{2.3}$$

We got a first-order differential equation. The time evolution of the v_o for v_i is obtained by solving this differential equation. However, solving the differential equation is sometimes tricky. Thus, we will consider the other method, using Laplace transform.

2.2 Laplace transform

We will consider the following transformation:

$$F(s) = \int_0^\infty f(t)e^{-st}dt. \tag{2.4}$$

This transformation from $f(t)$ to $F(s)$ is called Laplace transform (s is a complex number). In the following part, the Laplace transform will be expressed as follows:

$$F(s) = L[f(t)]. \tag{2.5}$$

Laplace transform is the most fundamental technique in control engineering. We will see how it is useful by applying it to the derivative and integral forms. Firstly, consider applying the Laplace transform to the derivative function:

$$\begin{aligned} L\left[\frac{df(t)}{dT}\right] &= \int_0^\infty \frac{df(t)}{dt}e^{-st}dt \\ &= \left[f(t)e^{-st}\right]_0^\infty + \int_0^\infty f(t)se^{-st}dt \\ &= -f(0) + sF(s). \end{aligned} \tag{2.6}$$

Next, consider Laplace transform of a integration ($L[h(t)]$, where $h(t) = \int_0^t f(t)dt$). Firstly, note that

$$\frac{dh(t)}{dt} = f(t), h(0) = 0. \tag{2.7}$$

Using eq. 2.6, we can get the following relations:

$$L\left[\frac{dh(t)}{dt}\right] = -h(0) + sH(s) = sH(s) \tag{2.8}$$

$$L\left[\frac{dh(t)}{dt}\right] = L\left[f(t)\right] = F(s) \tag{2.9}$$

$$\therefore sH(s) = F(s). \tag{2.10}$$

Therefore, the Laplace transform of integration is expressed by

$$\begin{aligned} L\left[\int_0^t f(t)dt\right] &= L\left[h(t)\right] \\ &= H(s) \\ &= \frac{1}{s}F(s). \end{aligned} \tag{2.11}$$

To summarize, if $f(0) = 0$, the Laplace transform of the derivative and integration is expressed as follows:

$$L\left[\frac{df(t)}{dT}\right] = sF(s), \tag{2.12}$$

$$L\left[\int_0^t f(t)dt\right] = \frac{1}{s}F(s). \tag{2.13}$$

These two equations represent the essential features of the Laplace transform: the derivative and integration are expressed by multiplication and division.

Let us apply Laplace transform to LPF. The derivative equation for LPF (eq. 2.3) is translated as follows:

$$V_i(s) = RCsV_o(s) + V_o(s) \tag{2.14}$$

$$\therefore G(s) = \frac{V_o(s)}{V_i(s)} = \frac{1}{1+RCs}. \tag{2.15}$$

$G(s) = V_o(s)/V_i(s)$ is called transfer function. The key point is that the relationship between the input and output is expressed without using the complicated derivative equation. In a series system, the transfer function of the whole system is expressed by multiplying those of each component. For example, let us consider that two LPFs are connected in series (the series connection is called a cascade, as shown in Fig. 2.1(b)). In this system, the output from the first LPF (v_o') is the input of the second LPF (v_i'). Thus, the transfer function of the whole system is expressed as follows:*1

$$\begin{aligned} G(s) = \frac{V_o(s)}{V_i(s)} &= \frac{V_o(s)}{V_i'(s)}\frac{V_o'(s)}{V_i(s)} \\ &= G_1(s)G_2(s) \\ &= \left(\frac{1}{1+RCs}\right)^2, \end{aligned} \tag{2.16}$$

where $G(s)$, $G_1(s)$, and $G_2(s)$ are the transfer function of the whole system, first LPF, and second LPF. By repeating the same process, the transfer function of n-th cascade is expressed by

$$G(s) = \left(\frac{1}{1+RCs}\right)^n. \tag{2.17}$$

Laplace transform is often helpful in clarifying the relationship between the input and output signals in a complicated system. The Laplace transformation for basic functions and related theorem are summarised in Appendix. Refer to it if necessary.

2.3 Transient response of the system

The transient response of the system is often evaluated by using an impulse or step function as an input signal. Therefore, we will consider the response of the system for these basic inputs.

2.3.1 Impulse response

The impulse function ($\delta(t)$) is defined by

$$\delta(t) = \begin{cases} 1 & (t=0) \\ 0 & (else) \end{cases} \tag{2.18}$$

The Laplace transform of the impulse function is calculated as follows:

$$L[\delta(t)] = \int_0^\infty \delta(t)e^{-st}dt = e^0 = 1. \tag{2.19}$$

Therefore, if the impulse function is used as an input ($v_i(t) = \delta(t)$), the output is expressed by

$$V_o(s) = G(s)V_i(s) = G(s). \tag{2.20}$$

*1 This is the basic idea of the transfer function and is often applied to the case of LPFs [1]. However, you may get a different equation if you get the differential equation by applying Ohm's law to the cascade circuit. Some approximation might be applied here, but I still don't know what it is.

2.3.2 Step response

The step function ($u_s(s)$) is defined by

$$u_s(t) = \begin{cases} 1 & (t > 0) \\ 0 & (t < 0) \end{cases} \tag{2.21}$$

The Laplace transform of the step function is calculated as follows:

$$L[u_s(t)] = \int_0^\infty u_s(t)e^{-st}dt = \left[-\frac{1}{s}e^{-st}\right]_0^\infty = \frac{1}{s}. \tag{2.22}$$

Therefore, if the step function is used as an input, the output is expressed by

$$V_o(s) = G(s)V_i(s) = \frac{G(s)}{s}. \tag{2.23}$$

2.3.3 Inverse Laplace transform

We got the output function for impulse and step functions. However, the output is still expressed by s. Therefore, we should convert the function to be expressed by t by inversely applying the Laplace transition (this process is called the inverse Laplace transform). If the function $F(s)$ is expressed by the known form (shown in the table in the Appendix), we can replace it with the corresponding $f(t)$. A helpful method to express $F(s)$ in a known form is to expand the function into the partial fraction. The detail of partial expansion is introduced in Appendix. Here we will consider the impulse and step response of the LPF. The table for inverse Laplace transforms is shown in Appendix.

Firstly consider the impulse response. The impulse response of an LPF is expressed by

$$\begin{aligned} V_o(s) = G(s) &= \frac{1}{1 + RCs} \\ &= \frac{1}{RC}\frac{1}{s + \frac{1}{RC}}. \end{aligned} \tag{2.24}$$

We can apply the relation (v) in the table (see Appendix) for inverse transformation. Thus, the impulse response of LPF is expressed by

$$v_o(t) = \frac{1}{RC}e^{-\frac{1}{RC}t}. \tag{2.25}$$

For the n th cascade, the Laplace function was expressed by

$$\begin{aligned} V_o(s) = G(s) &= \left(\frac{1}{1 + RCs}\right)^n \\ &= \left(\frac{1}{RC}\right)^n \left(\frac{1}{s + \frac{1}{RC}}\right)^n \end{aligned} \tag{2.26}$$

This form is similar to the relation (iv), but the denominator is not a simple s. Therefore, a related theorem $L[e^{at}f(t)] = F(s - a)$ is also applied. By combining this theorem and the relation (iv), the impulse response of n th cascade is expressed as follows:

$$v_o(t) = \left(\frac{1}{RC}\right)^n \frac{1}{(n-1)!}t^{n-1}e^{-\frac{1}{RC}t}. \tag{2.27}$$

Next, consider the step response. There are two methods to calculate the step response. The first is using the inverse Laplace transform as we did for impulse response. The other method is to integrate the impulse response, as is proven as follows:

$$L^{-1}[F(s)u_s(s)] = L^{-1}\left[F(s)\frac{1}{s}\right] \tag{2.28}$$

$$= \int_0^t f(t)dt. \tag{2.29}$$

Here L^{-1} represents the inverse Laplace transform. As $f(t)$ can be calculated by the impulse response, the step response can be calculated by integrating the impulse response.

Let us calculate the step response of an LPF by inverse Laplace transform. The response function is expressed by

$$\begin{aligned} V_o(s) = G(s)\frac{1}{s} &= \frac{1}{1+RCs}\frac{1}{s} \\ &= -\frac{1}{s+\frac{1}{RC}} + \frac{1}{s}. \end{aligned} \tag{2.30}$$

By performing the inverse Laplace transform with referring the table, the step response of an LPF is obtained as follows:

$$v_o(t) = 1 - e^{-\frac{1}{RC}t}. \tag{2.31}$$

For getting the step response of n th cascade, it would be easy to integrate the impulse response (eq. 3.1) as follows:

$$\begin{aligned} v_o(t) &= \int_0^t \left(\frac{1}{RC}\right)^n \frac{1}{(n-1)!} t^{n-1} e^{-\frac{1}{RC}t} dt \\ &= \left(\frac{1}{RC}\right)^n \frac{1}{(n-1)!} \int_0^t t^{n-1} \left[-RTe^{-\frac{1}{RT}t}\right]' dt \\ &= -\left(\frac{1}{RC}\right)^{n-1} \frac{1}{(n-1)!} t^{n-1} e^{-\frac{1}{RT}t} + \left(\frac{1}{RC}\right)^{n-1} \frac{1}{(n-2)!} \int_0^t t^{n-2} e^{-\frac{1}{RT}t} dt \\ &= \cdots . \end{aligned} \tag{2.32}$$

By repeating the partial integration, we finally get

$$v_o(t) = 1 - e^{-\frac{1}{RC}t} \sum_{k=0}^{n-1} \frac{1}{k!} \left(\frac{t}{RC}\right)^k. \tag{2.33}$$

2.3.4 Response for arbitrary input

In the previous section, we got the response for impulse and step functions. The response for the arbitral function can be obtained by expressing the function by impulse functions as follows:

$$v_i(t) = v_i(0)\delta(t)\Delta t + v_i(\Delta t)\delta(t-\Delta t)\Delta t + v_i(2\Delta t)\delta(t-2\Delta t)\Delta t + \ldots \tag{2.34}$$

For considering the Laplace transform of eq. 2.34, we should know the how the function with time delay ($f(t-a)$) is transformed. By following the definition, the Laplace transform of the function with time delay is expressed as follows:

$$\begin{aligned} L\left[f(t-a)\right] &= \int_0^\infty f(t-a)e^{-st} dt \\ &= \int_0^\infty f(\tau)e^{-s(\tau+a)} d\tau \\ &= e^{-as}F(s). \end{aligned} \tag{2.35}$$

By using this relationship, the Laplace transform of eq. 2.34 is calculated as follows:

$$V_i(s) = v_i(0)\Delta t + v_i(\Delta t)e^{\Delta ts}\Delta t + v_i(2\Delta t)e^{2\Delta ts}\Delta t + \ldots. \tag{2.36}$$

Therefore, the output function is expressed by

$$\begin{aligned} V_o(s) &= G(s)V_i(s) \\ &= v_i(0)\Delta t G(s) + v_i(\Delta t)\Delta t e^{\Delta ts} G(s) + v_i(2\Delta t)\Delta t e^{2\Delta ts} G(s) + \ldots. \end{aligned} \tag{2.37}$$

By inversely using eq. 2.35, the output function in t space is expressed as follows:

$$v_o(t) = v_i(0)g(t)\Delta t + v_i(\Delta t)g(t-\Delta t)\Delta t + v_i(2\Delta t)g(t-2\Delta t)\Delta t + \ldots, \tag{2.38}$$

where $g(t) = L^{-1}\left[G(s)\right]$. Now we got the output signal for arbitrary inputs.

2.3.5 Frequency response

The basic function of LPF is the attenuation of the high-frequency signal. Here we will consider the input $v_i(t) = e^{j\omega t} = \cos(\omega t) + j\sin(\omega t)$ (j is the complex number) for considering the frequency dependence. Because the Laplace transform of the input is expressed by

$$V_i(s) = L\left[e^{j\omega t}\right] = \frac{1}{s - j\omega}, \tag{2.39}$$

the output in t space is expressed as follows:

$$\begin{aligned} v_o(t) &= L^{-1}\left[G(s)V_i(s)\right] \\ &= L^{-1}\left[G(s)\frac{1}{s - j\omega}\right] \\ &= L^{-1}\left[\frac{K_0}{s - j\omega} + \sum_{i=1}^{n}\frac{K_i}{s - p_i}\right] \\ &= K_0 e^{j\omega t} + \sum_{i=1}^{n} K_i e^{p_i t}. \end{aligned} \tag{2.40}$$

If $p_i < 0$, the second term disappears at $t \to \infty$. Thus, the frequency response is determined by the first term. The coefficient K_0 can be obtained by

$$\begin{aligned} K_0 &= \lim_{s \to j\omega}(s - j\omega)G(s)V_i(s) \\ &= G(j\omega) \\ &= |G(j\omega)|e^{j\phi}. \end{aligned} \tag{2.41}$$

In the final part, $G(j\omega)$ is expressed by a polar form. By combining eq. 2.40 and eq. 2.41, at the steady state ($t \to \infty$), the output is expressed as follows:

$$\begin{aligned} v_o(t) &= G(j\omega)e^{j\omega t} \\ &= |G(j\omega)|e^{j(\omega t + \phi)}. \end{aligned} \tag{2.42}$$

Thus, the change of the amplitude and phase after passing the LPF is expressed by $|G(j\omega)|$ and ϕ, respectively.

Let us consider the frequency response of the LPF. For a single LPF, the transfer function was expressed by

$$G(s) = \frac{1}{1 + RCs} \tag{2.43}$$

Therefore, the frequency response is calculated as follows:

$$\begin{aligned} G(j\omega) &= \frac{1}{1 + jRC\omega} \\ &= \frac{1}{1 + (RC\omega)^2} - j\frac{RC\omega}{1 + (RC\omega)^2} \\ \therefore |G(j\omega)| &= \frac{1}{\sqrt{1 + (RC\omega)^2}}, \phi = -\tan^{-1}(RC\omega). \end{aligned} \tag{2.44}$$

For the n th cascade, the frequency response is expressed as follows:

$$|G(j\omega)| = \left\{\frac{1}{\sqrt{1 + (RC\omega)^2}}\right\}^n, \phi = -\tan^{-1}(nRC\omega). \tag{2.45}$$

This equation indicates that the output amplitude is suppressed for higher-frequency inputs. We will see it in detail in the next chapter.

Chapter 3

Transit response of LPF for various inputs

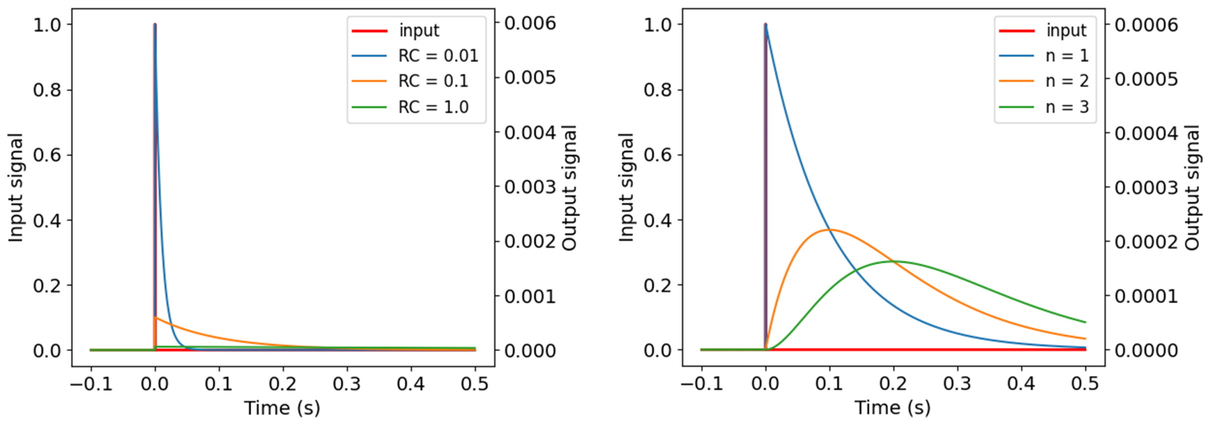
In the previous chapter, we got the formulas of the output from LPF. In this chapter, we will see the simulated response of LPF for various inputs.

3.1 Impulse response

As we have already discussed in the previous chapter, the impulse response of LPF is expressed as follows:

$$v_o(t) = \left(\frac{1}{RC}\right)^n \frac{1}{(n-1)!} t^{n-1} e^{-\frac{1}{RC}t}. \tag{3.1}$$

Firstly let us consider the case of $n = 1$ and see the dependence on RC. As shown in Fig. 3.1(a), the lower RC leads to stronger responses and fast decay. The n dependence is shown in Fig. 3.1(b). The increase of n leads to a weaker output with a time delay.



(a) (b)

Fig. 3.1: Implpuse response for (a) various RC ($n = 1$) and (b) various n ($RC = 0.1$).

3.2 Step response

The step response of LPF was expressed as follows:

$$v_o(t) = 1 - e^{-\frac{1}{RC}t} \sum_{k=0}^{n-1} \frac{1}{k!} \left(\frac{t}{RC}\right)^k. \tag{3.2}$$

Figure 3.2(a) and (b) shows the dependence of step response on RC and n, respectively. All outputs finally saturate at the input value. However, the response of LPFs with higher RC is slow, resulting in a long time to reach the steady value. Even with the same RC, the larger cascade also takes longer to be saturated.

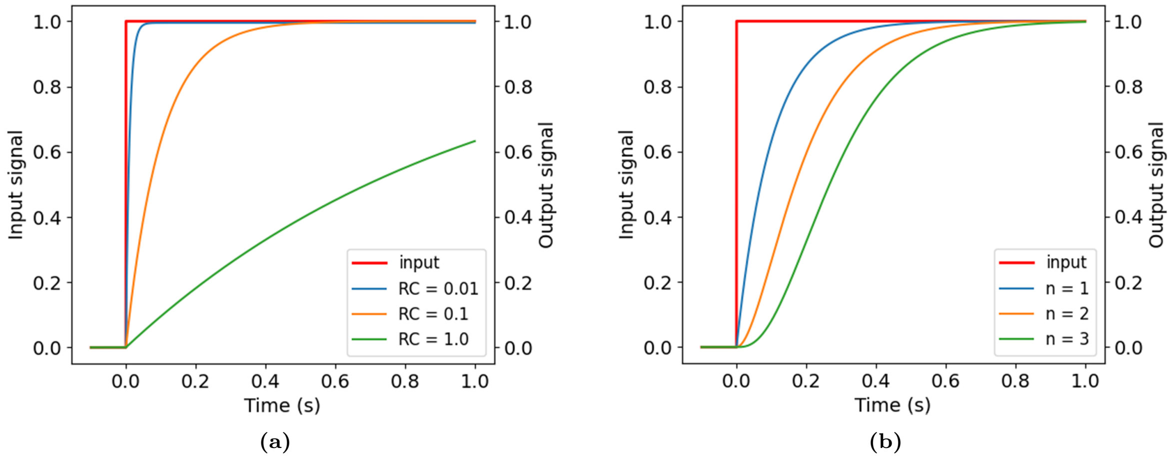


Fig. 3.2: Step response for (a) various RC ($n = 1$) and (b) various n ($RC = 1$).

3.3 Response for an arbitrary function

We also got the formula for the response for arbitrary inputs as follows:

$$v_o(t) = v_i(0)g(t)\Delta t + v_i(\Delta t)g(t - \Delta t)\Delta t + v_i(2\Delta t)g(t - 2\Delta t)\Delta t + \dots. \tag{3.3}$$

We can simulate the response for any inputs by using this formula. For example, Fig. 3.3 shows the response for input with a Gaussian shape (we will call it Gaussian response). While the LPF with a small RC well follows the input signal, the LPF with a large RC shows a weaker signal. In addition, it becomes apparent that the LPF distorts the shape of the original signal, as can be seen in the asymmetric shape of the $RC = 1.0$ case. Thus, keep in mind that the LPF with high RC not only attenuates the high-frequency signal but also distorts the signal.

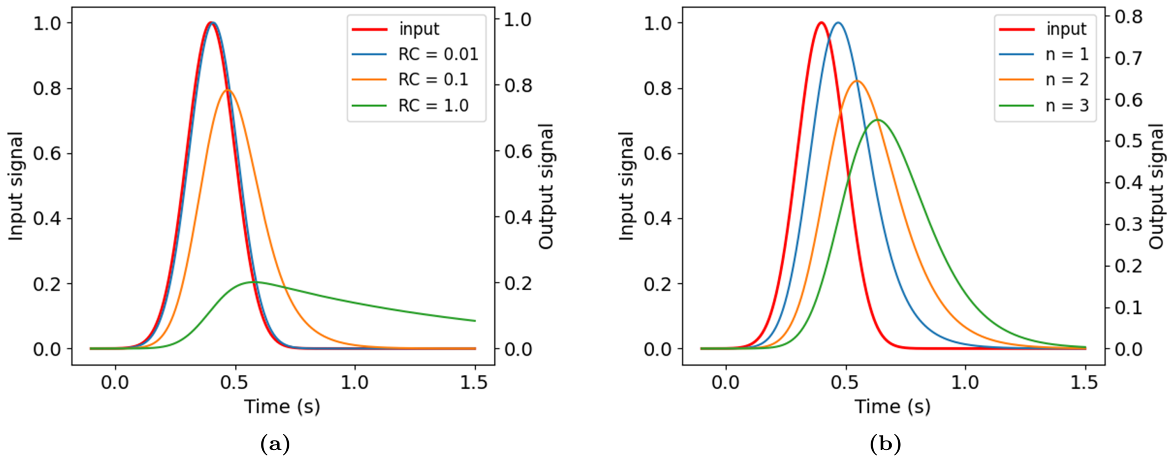
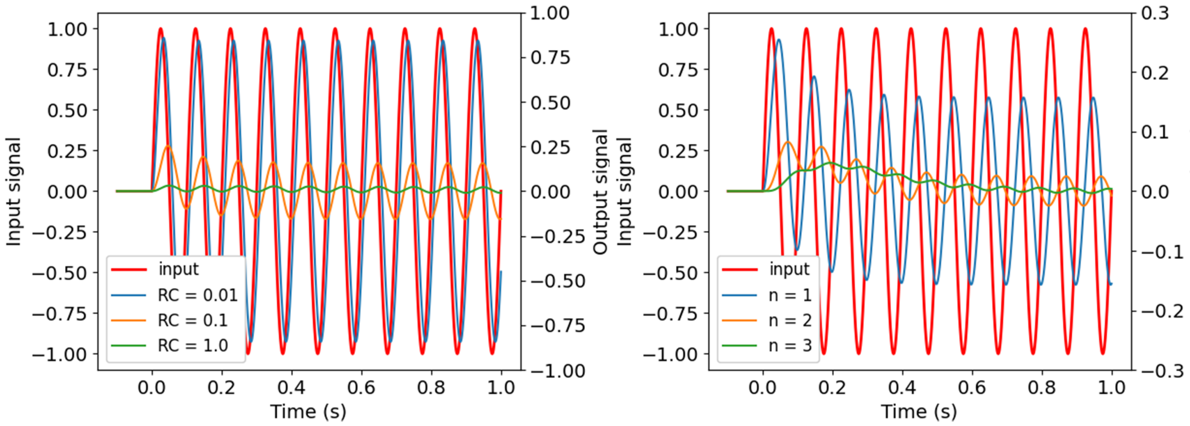


Fig. 3.3: Gaussian response for (a) various RC ($n = 1$) and (b) various n ($RC = 0.1$).

3.4 Frequency response

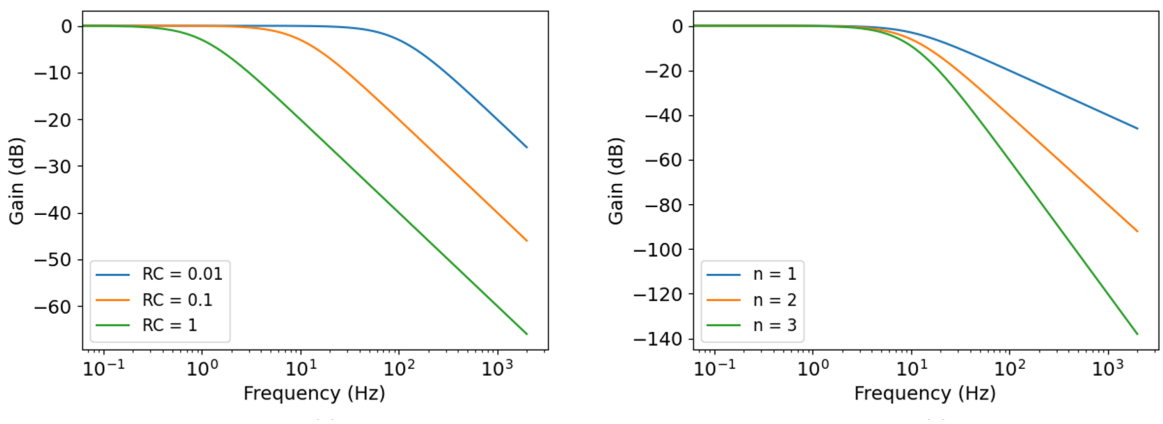
Figure 3.4 shows the response of LPF for a wave function with a frequency of 10 Hz. At $RC = 0.01$, as the cutoff frequency ($\omega_c = 1/RC = 100$ Hz) is much higher than the wave frequency, the LPF does not affect the signal much. Thus, the output signal is almost identical to the input wave. However, the input wave is well attenuated with the phase delay at the higher RC (indicating smaller ω_c). With using the cascade, the attenuation ability gets higher.



(a) **(b)**

Fig. 3.4: Responce for a wave functin for (a) various RC ($n = 1$) and (b) various n ($RC = 0.1$).

The attenuation ability of LPF can be evaluated by plotting the gain as a function of the input frequency, as shown in Fig. 3.5. Note that the gain is expressed in dB unit, defined by $20 \log_{10}(v_o/v_i)$. As shown in Fig. 3.5(a), the gain is almost constant at the low frequency. The curves show a corner at the cutoff frequency ($\omega_c = 1/RC$), and the signal is well attenuated above the corner. This is the primary function of LPF. Figure 3.5(b) shows the gain for a cascade of $RC = 0.1$ LPFs. The position of the corner is almost kept at the same frequency while the slope above the cutoff becomes steep.



(a) **(b)**

Fig. 3.5: Frequency dependent attenuation ability for (a) various RC ($n = 1$) and (b) various n ($RC = 0.1$).

Chapter 4

Feedback analysis

4.1 Basic concept of feedback

Figure 4.1(a) shows a simplified block diagram of a feedback system. We start from $O(s)$, which is the output from the system. The output signal is sent to component B with the transfer function of $G_2(s)$. Therefore, the output from the component B is $M_1(s) = G_2(s)O(s)$. The difference between $M_1(s)$ and the input ($I(s)$) becomes the input to another component A with $G_1(s)$. In other words, $M_2(s) = I(s) - A(s)$. The signal after passing A is the output $O(s)$. Thus, $O(s) = G_1(s)M_2(s)$. To summarize,

$$M_1(s) = O(s)G_2(s), \tag{4.1}$$
$$M_2(s) = I(s) - M_1(s), \tag{4.2}$$
$$O(s) = G_1(s)M_2(s), \tag{4.3}$$
$$\therefore O(s) = G_1(s)\left[I(s) - G_2(s)O(s)\right] \tag{4.4}$$
$$\therefore G(s) = \frac{O(s)}{I(s)} = \frac{G_1(s)}{1 + G_1(s)G_2(s)}. \tag{4.5}$$

The transfer function of the whole system ($G(s)$) is expressed as eq. 4.5. It means that the block diagram can be simplified as shown in Fig. 4.1(b). By using the transfer functions, the response of the feedback system is analyzed. We will apply this analysis to a more complicated STM system.

4.2 Feedback circuit in STM

Figure 4.2(a) shows the schematics of the distances in STM, which will be used to explain the feedback of STM. STM tip is mounted on a piezoelectric material, which extends by applying a high bias. The

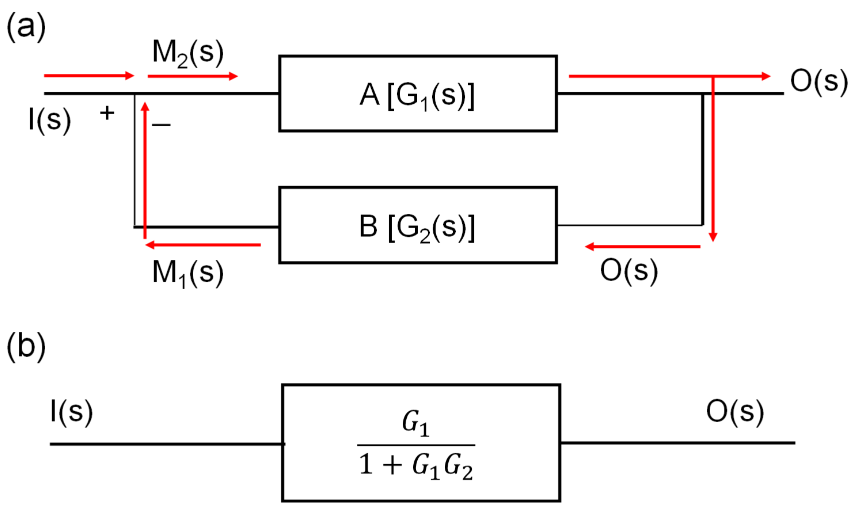


Fig. 4.1: (a) A block diagram of the feedback circuit. (b) A block diagram equivalent to (a).

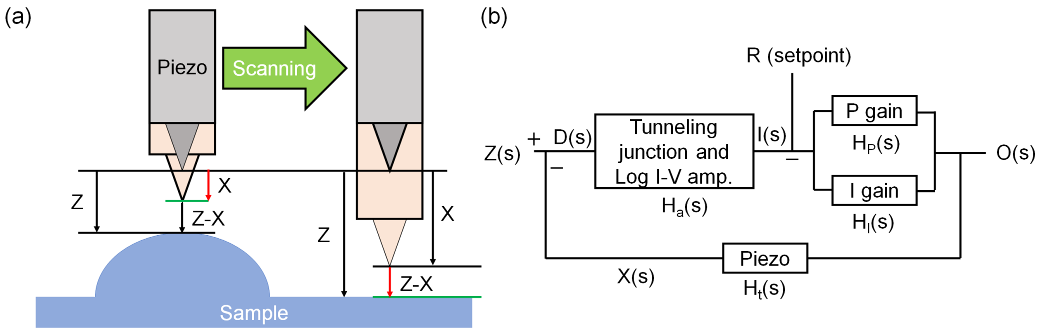


Fig. 4.2: (a) Relationships of distances in STM. (b) A block diagram of the STM circuit.

distance Z indicates the distance between the tip and sample without any extension. The tip-sample distance is adjusted by extending the piezo. X is the extended length of the piezo from the original position. Therefore, the tip-sample distance during the feedback operation is expressed by $Z - X$. The lateral movement of the tip (scanning) changes Z. For keeping $Z - X$ constant, X is modified by the feedback. Figure 4.2(b) shows the block diagram of the feedback system in STM. Let us start with the extension of the piezo ($X(s)$). As explained, the distance between the tip and sample is expressed by $D(s) = Z(s) - X(s)$. In STM, a tunneling current flows between the tip and the sample. The tunneling current is converted to a voltage signal by a log amplifier. We express the transfer function of this part as $H_a(s)$. Thus, $I(s) = H_a(s)D(s)$. Then, $I(s)$ is compared with the setpoint value R. Therefore, the input to the feedback gains is $R - I(s)$. Because R just act as an offset in the system analysis, we will assume $R = 0$, indicating the tip follows the sample surface without a gap. The output from the feedback gains, $O(s) = -(H_p(s) + H_I(s)) \times I(s)$, is recorded as the topography in STM. The response of piezo is expressed by $H_t(s)$, thus $X(s) = H_t(s)O(s)$. To summarize, following relationships were obtained:

$$D(s) = Z(s) - X(s) \tag{4.6}$$

$$I(s) = H_a(s)D(s) \tag{4.7}$$

$$O(s) = -(H_p(s) + H_I(s))I(s) \tag{4.8}$$

$$X(s) = H_t(s)O(s) \tag{4.9}$$

$$\therefore G(s) = \frac{O(s)}{Z(s)} = -\frac{H_a(s)\left[H_P(s) + H_I(s)\right]}{1 - H_t(s)H_a(s)\left[H_P(s) + H_I(s)\right]}. \tag{4.10}$$

Let us think about the transfer function of each component. The output from the proportional gain ($O_P(t)$) is proportional to the error input ($E(t)$). Thus, in the time domain,

$$O_P(t) = K_p E(t). \tag{4.11}$$

K_p is called p-gain. On the other hand, the output from the integral gain ($O_S(t)$) is the integration of the error input from the beginning, expressed as follows:

$$O_I(t) = K_I \int_0^t E(t')dt'. \tag{4.12}$$

K_I is called i-gain. The outputs from two gains are summed as the feedback gain as follows:

$$O'(t) = K_P E(t) + K_I \int_0^t E(t')dt' \tag{4.13}$$

$$\left[= K_P\left(E(t) + \frac{1}{T_I}\int_0^t E(t')dt'\right), where\, K_I = \frac{K_P}{T_I}\right]. \tag{4.14}$$

The i-gain is sometimes expressed by a time constant T_I. Remember that the time constant is inversely proportional to K_I. Following the recent paper in ref. [2], we will regard the output $O'(t)$ as the bias change applied to the piezo. In the mathematical expression, it means

$$O'(t) = \Delta O(t) = K_P E(t) + K_I \int_0^t E(t')dt'$$
$$\therefore O(t) = \int_0^t O'(t) = K_P \int_0^t E(t')dt' + K_I \int_0^t \int_0^{t'} E(t'')dt''dt'. \tag{4.15}$$

The transfer function is obtained by applying the Laplace transform as follows:

$$H_P(s) = \frac{K_P}{s} \tag{4.16}$$
$$H_I(s) = \frac{K_I}{s^2} \tag{4.17}$$

The transfer function of the piezo is treated as the spring-mass-damper system. The equation of motion for the spring-mass-damper system is expressed as follows:

$$F(t) = m\frac{d^2x(t)}{dt^2} + c\frac{dx}{dt} + kx(t), \tag{4.18}$$

where m, c, k represent the mass, coefficient for the force related to the velocity, and spring constant. If we define the following relations

$$\omega_0 = \sqrt{\frac{k}{m}}, \eta = \frac{c}{2\sqrt{km}}, Q = \frac{1}{2\eta}, \tag{4.19}$$

the equation of motion is expressed as follows:

$$F(t) = \frac{1}{\omega_0^2}\frac{d^2x(t)}{dt^2} + \frac{1}{Q\omega_0}\frac{dx}{dt} + x(t), \tag{4.20}$$

where ω_0 and η are called resonance frequency and damping constant, respectively. Q is called the quality factor, representing how sharp the resonance is. By applying the Laplace transform, we get the following equation:

$$H_p(s) = \frac{X(s)}{F(s)} = \frac{K_{piezo}}{\frac{s^2}{\omega_0^2} + \frac{s}{Q\omega_0} + 1}. \tag{4.21}$$

This transfer function is used for the feedback analysis. For simplicity, we will assume $K_{piezo} = 1$.

Then let us consider about $H_a(s)$. According to the basic quantum mechanics, the tunneling current exponentially increases as the tip-sample distance reduces ($I(t) \propto \exp(-D(t))$. The exponential signal is converted to a linear signal by a log amplifier. Thus, $\log[\exp(-D(t))] = -D(t)$. The amplifier is often equipped with an LPF. As explained in previous chapters, the transfer function of LPF is expressed by

$$H_{LPF}(s) = \frac{1}{\frac{s}{\omega_c} + 1}, \tag{4.22}$$

where ω_c is the cutoff frequency of the LPF. Finally, the output from this part is expressed as follows:

$$I(s) = -H_{LPF}D(s)$$
$$\therefore H_a(s) = \frac{I(s)}{D(s)} = -\frac{1}{\frac{s}{\omega_c} + 1} \tag{4.23}$$

By combining eq. 4.10, eq. 4.16, eq. 4.17, eq. 4.21, and eq. 4.23, the transfer function in the STM circuit is obtained. Using the transfer function, we can analyze the relationships between the parameters in feedback and the response of STM.

4.3 Responce of feedback circuit

We got the transfer function of the STM system. Therefore, a simple simulation can be performed to see the effect of the feedback parameters. The following simulations were performed with $\omega_0 = 10$, $Q = 10$, and $\omega_c = 0$.

Firstly, let us see the effect of p-gain with fixing i-gain zero. The step and Gaussian responses are shown in Fig. 4.3. At the low p-gain, the step response of STM is slow, taking a long time to reach the input value. By increasing the p-gain, the responses get faster. However, if the p-gain value is too high, the output signal oscillates and becomes unstable. The same tendency can be seen in the Gaussian response. Although the output traces the input much better when the higher p-gain is set, oscillation occurs at the high p-gain. Thus, it is concluded that the p-gain should be set high for better tracing but should not be too high for avoiding oscillation.

Combination with i-gain improves the response speed. The benefit of i-gain becomes clear by seeing the response to the ramp input (the input linearly increases), as shown in Fig. 4.4(a). The response only using p-gain (red curve) follows the input with a constant delay. However, by applying finite i-gain, the response follows the input with almost no delay. Figure 4.4 shows the step response with setting i-gain. Applying the i-gain improves the initial response, while the output shows an overshoot.

4.4 Simulating STM image

We further performed a simple simulation corresponding to the STM scan to visually clarify the effect of feedback parameters on imaging. The simulations were performed for a model surface in which a half sphere is placed on a flat surface (Fig. 4.5(a)). The simulated images consisted of the response to each

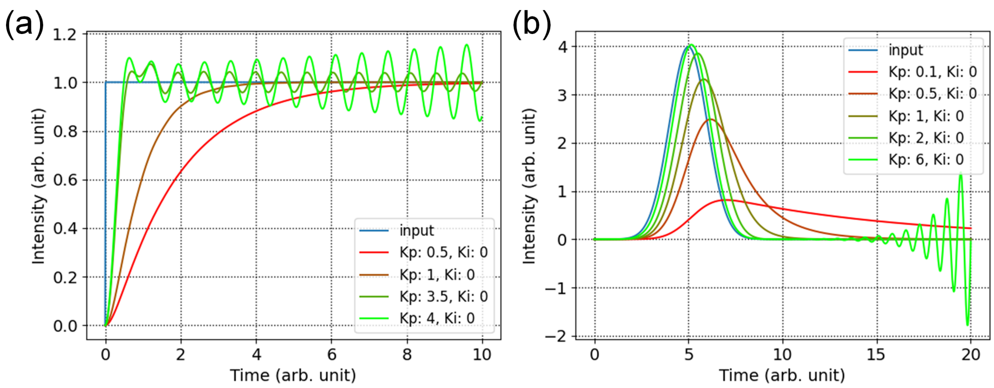


Fig. 4.3: Response of the STM circuit to (a) step input and (b) Gaussian inputs, respectively.

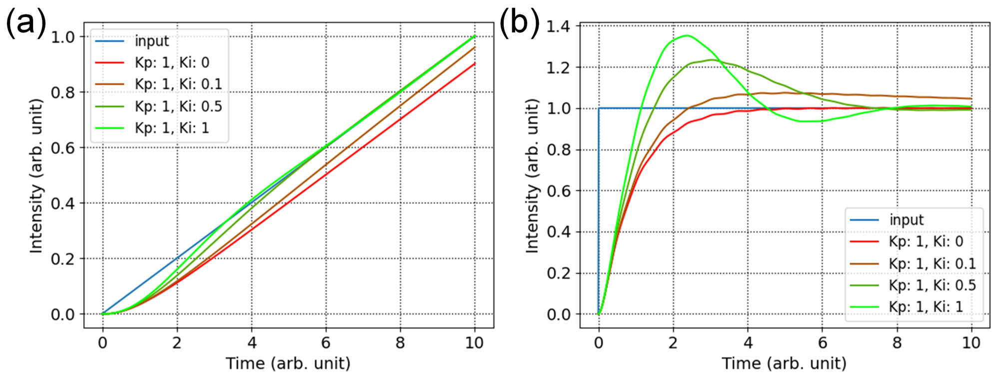


Fig. 4.4: Responce of the STM circuit to (a) ramp and (b) step inputs, respectively.

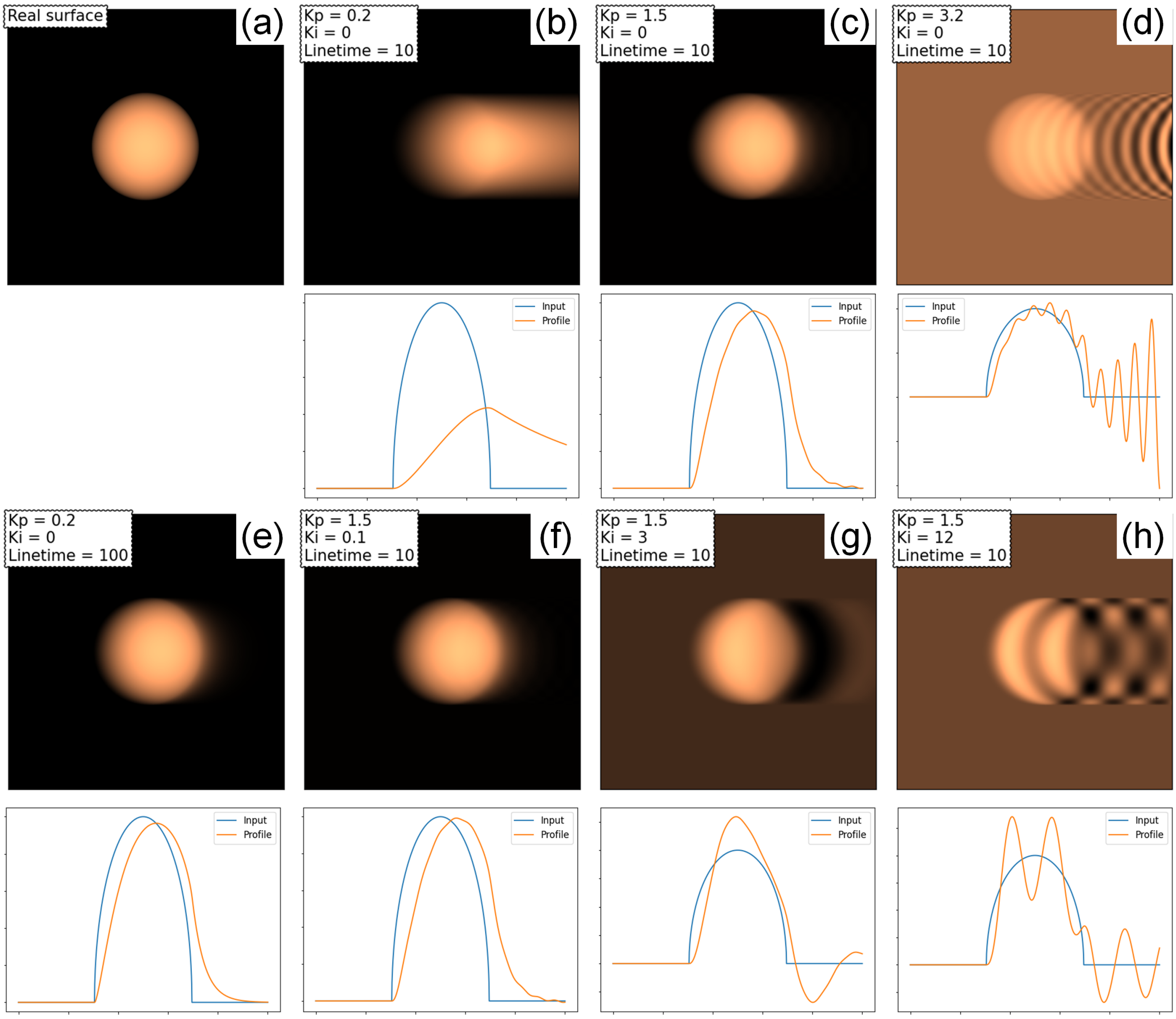


Fig. 4.5: (a) The geometry of the model surface, in which a half sphere is placed at the center. (b)-(h) Simulated STM images using various feedback parameters. The line profiles crossing the center of the images are shown at the lower side of each image. The blue and orange curves in the profiles represent the profile of the original and simulated images, respectively.

horizontal line, imitating the scanning from left to right. The parameters except for the p- and i-gains and line time are fixed at $\omega_0 = 10$, $Q = 10$, and $\omega_c = 0$.

Figure 4.5(b) shows the result when the p-gain is low. The tip cannot follow the change in the topography of the image. As a result, the half sphere is imaged with the lower height and leaves traces on the right side. If the features in the images seem to be extended in the scanning direction, the p-gain is too low, or scanning is too fast. By raising the p-gain or slow scanning (shown in 4.5(c) and (e), respectively), the STM image represents the actual surface much better.

However, if the p-gain is set too high, the feedback circuit oscillates, resulting in the wavy pattern in the STM image, as shown in Fig. 4.5(d). The oscillation can be found not only in the image but also in an oscilloscope. The oscillation means the tip continuously crashes into the sample surface. If you find the sign of oscillation, reduce the p-gain or retract the tip as soon as possible.

Figure 4.5(f) shows the result with adding low i-gain to (c). The image better represents the real topography than the image in (c), indicating the proper combination of p- and i-gains results in the best imaging. However, high i-gain also causes artifacts in the images, as shown in Fig. 4.5(g) and (h). In Fig. 4.5(g), although scanning is stable and no oscillation happens, the sphere becomes anisotropic and accompanies depression on the right side. As a result, the sphere looks to be shed lighted from the left side. If i-gain is too high, the circuit also oscillates, as shown in Fig. 4.5(h).

4.5 Closing remark

The motivation for introducing the feedback circuit in STM is to give a clear thought about how the feedback affects the STM images. As clarified, The improper setting in the parameters causes artifacts in STM images, which sometimes are misunderstood as meaningful patterns [3]. In addition, the unstable scanning can damage the tip and sample. Therefore, the STM users should pay attention to whether the feedback is set correctly during the scanning.

However, I think that the experience of the STM users would be much more critical. Firstly, the proper setting of the feedback differs depending on the setup of each STM. Secondly, the STM can be operated without understanding the feedback correctly. It is supported by the fact that although the proper setting of feedback in STM has long been discussed since the development of STM in the 1980s [4, 5], it was pointed out in 2014 that the conventional understanding of STM circuits was wrong! It means almost all the STM users successfully operated STM without the correct understanding of the feedback in STM. Just keep in mind that the feedback in STM can cause artifacts in the images, and check how the artifact appears in your STM.

Chapter 5

Appendix

5.1 The basic Laplace transform

$$
\begin{array}{c|ll|c|ll}
 & f(t) & F(s) & & f(t) & F(s) \\
\hline
(1) & \delta(t) & 1 & (6) & te^{-at} & \dfrac{1}{(s+a)^2} \\
(2) & u_s(t)(=1) & \dfrac{1}{s} & (7) & \sin\omega t & \dfrac{\omega}{s^2+\omega^2} \\
(3) & t & \dfrac{1}{s^2} & (8) & \cos\omega t & \dfrac{s}{s^2+\omega^2} \\
(4) & \dfrac{t^n}{n!} & \dfrac{1}{s^{n+1}} & (9) & e^{-at}\sin\omega t & \dfrac{\omega}{(s+a)^2+\omega^2} \\
(5) & e^{-at} & \dfrac{1}{s+a} & (10) & e^{-at}\cos\omega t & \dfrac{s+a}{(s+a)^2+\omega^2}
\end{array}
\tag{5.1}
$$

5.2 Patial fraction decomposition

In performing the inverse Laplace transform, the equation should be transformed to a known form shown in the table above. Partial fraction decomposition is often used to achieve it.

Here assume that the equation is expressed in the following form:

$$F(s) = \frac{b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}. \tag{5.2}$$

If we assume that the solution of the denominator is expressed by $p_1, p_2, ..., p_n$ (here we also assume that all the solutions are different values), the equation is expressed by

$$F(s) = \frac{b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0}{(s-p_1)(s-p_2)\dots(s-p_n)}. \tag{5.3}$$

p_i is called a pole. This form of the equation can be further converted to the following form:

$$F(s) = \frac{k_1}{s-p_1} + \frac{k_2}{s-p_2} + \cdots + \frac{k_n}{s-p_n} \tag{5.4}$$

This conversion is called partial fraction decomposition. The coefficients k_i can be obtained by

$$
\begin{aligned}
\lim_{s\to p_i}(s-p_i)F(s) &= \lim_{s\to p_i}\left[(s-p_i)\frac{k_1}{s-p_1} + \cdots + k_i + \cdots + (s-p_i)\frac{k_n}{s-p_n}\right] \\
&= k_i
\end{aligned}
\tag{5.5}
$$

Using the relation (5) in the table, the inverse Laplace transform of the decomposed function in eq. 5.4 is expressed as follows:

$$f(t) = L^{-1}[F(s)] = k_1 e^{p_1 t} + k_2 e^{p_2 t} + \cdots + k_n e^{p_n t}. \tag{5.6}$$

Here it is important that the poles determine the stability of the $f(t)$. Consider that a pole is expressed by $p = a + jb$, where a and b are the real numbers, and j is the complex number. In this case,

$$\begin{aligned} e^{pt} = e^{(a+jb)t} &= e^{at} e^{jbt} \\ &= e^{at} \left[\cos(bt) + j\sin(bt)\right]. \end{aligned} \tag{5.7}$$

The real part of the pole determines how the signal decays (or diverges), and the imaginary part is related to the oscillation behavior. Therefore, if $a > 0$, the signal diverges, and the system should be unstable. If b is not zero, the signal should show oscillation. The necessary condition for the system to be stable is that all the poles are located on the negative side of the imaginary axis.

Bibliography

[1] "Principles of lock-in detection and the state of the art" by Zurich Instruments, https://www.zhinst.cn/sites/default/files/documents/2020-06/zi_whitepaper_principles_of_lock-in_detection.pdf/zi_whitepaper_principles_of_lock-in_detection.pdf

[2] J. Stirling, "Control theory for scanning probe microscopy revisited", Beilstein J. Nanotechnol. 5, 337 (2014).

[3] J. Stirling, I. Lekkas, A. Sweetman, P. Djuranovic, Q. Guo, B. Pauw, J. Granwehr, R. Levy, and P. Moriarty, "Critical Assessment of the Evidence for Striped Nanoparticles", PLoS ONE 9, e108482 (2014).

[4] E. Anguiano, A. I. Oliva, and M. Aguilar, "Optimal conditions for imaging in scanning tunneling microscopy: Theory", Rev. Sci. Instrum. 69, 3876 (1998).

[5] S. Park, and R. C. Barrett, Chapter 2 in "Methods in Experimental Physics", by Academic Press in pp. 31-76 (1993).