I. BASICS OF HETERODYNE DETECTION

The term *heterodyning* basically describes a process when two signals at (cyclic) frequencies ω_1 and ω_2 (usually $\omega_1 \neq \omega_2$) are converted to a third signal at the sum or difference frequency, $\omega_1 \pm \omega_2$. The main advantage of this procedure is that sometimes it is much better to process or detect a signal in a frequency range which is very different from the frequency of the original signal. For example, in our experiment we want to detect the resonant signal at about 5 MHz from the quartz microbalance by using the lock-in amplifier operating in the range 0-100 kHz. In this case, heterodyning can provide some advantages comparing with AM and FM modulation techniques that we tried earlier. Below, I provide some basic ideas of frequency conversion and heterodyne detection, which we will try in the experiment.

A. Mixer

The frequency conversion is accomplished by the process of *mixing* of the two signals $u_1 \cos(\omega_1 t)$ and $u_2 \cos(\omega_2 t)$ using a non-linear device. For example, in case of *radio-frequency* (RF) signals, that is signals in the frequency range from about 1 kHz to 100 GHz, the mixers are based on semiconductor diodes, that is devices with non-linear current-voltage (I-V) characteristics. Indeed, the AC current $i(t)$ through the diode can be expanded as a series in the applied AC voltage $u(t)$ as

$$i(t) = a_0 + a_1 u(t) + a_2 u^2(t) + \ldots \quad , \tag{1}$$

where a_n, $n = 0, 1, ..$, are some coefficients. If we apply the sum of two signals to the diode, that is $u(t) = u_1 \cos(\omega_1 t) + u_2 \cos(\omega_2 t)$, from the quadratic term in (1) we obtain a signal proportional to

$$u_1^2 \cos^2(\omega_1 t) + u_2^2 \cos^2(\omega_2 t) + 2u_1 u_2 \cos(\omega_1) \cos(\omega_1), \tag{2}$$

which using simple trigonometric relations can be written as

$$\frac{u_1^2}{2}(1 - \cos(2\omega_1 t)) + \frac{u_2^2}{2}(1 - \cos(2\omega_2 t)) + u_1 u_2 (\cos((\omega_1 - \omega_2)t) + \cos((\omega_1 + \omega_2)t)) \,. \tag{3}$$

Thus, the mixer output contains a DC component, which can be always filtered out, signals at original frequencies ω_1 and ω_2, their second harmonics $2\omega_1$ and $2\omega_2$, and signals at frequencies $\omega_1 \pm \omega_2$. In addition, there are various signals coming from the higher order expansion terms in Eq. (1). Usually the mixer is designed such a way as to suppress the latter, so to a good approximation we can ignore them.

B. Frequency down- and up-conversion

There are some standard notations to designate the mixer inputs and outputs. Suppose we mix two signals $u_1 \cos(\omega_1 t)$ and $u_2 \cos(\omega_2 t)$ with similar frequencies $\omega_1 \approx \omega_2$, that is $\Delta\omega << \omega_1, \omega_2$, where $\Delta\omega = |\omega_1 - \omega_2|$. In practice, one of the signal, let's say signal $u_1 \cos(\omega_1 t)$ is a signal that we want to measure, and another signal, $u_2 \cos(\omega_2 t)$, is a signal that we must provide using a generator, which we call the *local oscillator*. We refer to the measured signal as RF signal, and the signal from the local oscillator as LO signal, see Fig. 1a. The power-spectrum of the mixer output is shown in Fig. 1b. The low-frequency component of the mixer output, that is component at the difference frequency $\omega_{RF} - \omega_{LO}$, is referred to as the *intermediate* (IF) signal. Then, the low-frequency component can be easily filtered out by a low-pass filter. Usually, such low pass-filter is already contained in the commercial mixers. This scheme upon which we extract the signal at IF frequency from the original RF signal is called the *frequency down-conversion*. Please note that the amplitude of the IF signal is proportional to the amplitude of the measured RF signal, that is $u_{IF} \propto u_{RF} \cos(\omega_{IF} t)$.

Usually using the same mixer we can do the opposite procedure. Let us apply an LO signal

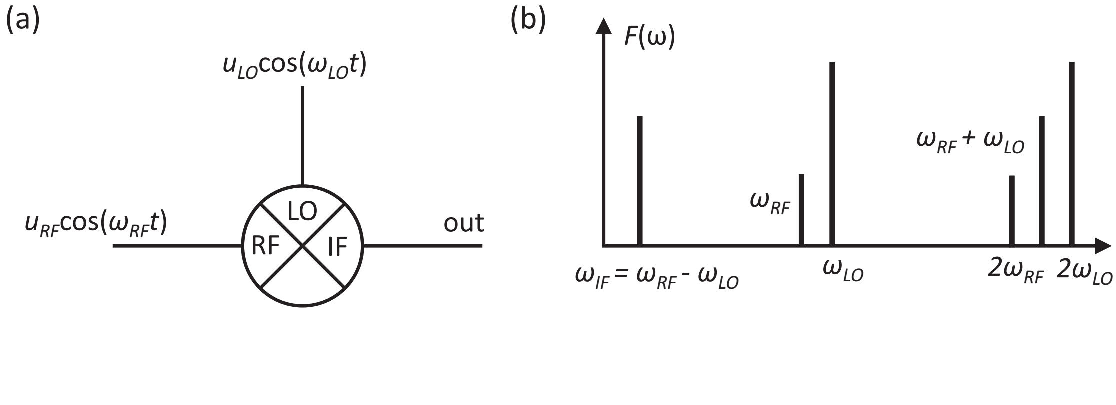

FIG. 1: Down-conversion mixer (a) and the power spectrum of its output (b).

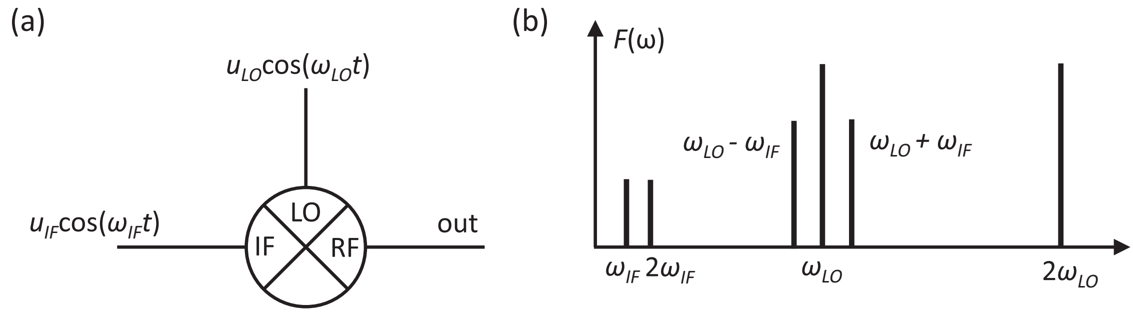

FIG. 2: Up-conversion mixer (a) and the power spectrum of its output (b).

from the local oscillator to LO input and an IF signal at much lower frequency $\omega_{IF} << \omega_{LO}$ to the IF input, see Fig. 2a. Then, at the RF output of the mixer we obtain a mixed signal with the power spectrum shown in Fig. 2b. In particular, we will obtain three signals at close frequencies ω_{LO} and $\omega_{LO} \pm \omega_{IF}$. The spectral components at $\omega_{LO} \pm \omega_{IF}$ are usually called the *side-bands*. Using a pass-band filter (with width comparable to or less than ω_{IF}) we can filter out one of the side-band component to obtain a signal at the frequency either $(\omega_{LO} - \omega_{IF})$ or $(\omega_{LO} + \omega_{IF})$, which is much higher than ω_{IF}. This procedure is called the *frequency up-conversion*. Again, note that the amplitude of the filtered signal is proportional to the amplitude of the applied IF signal, that is $u_{RF} \propto u_{IF}\cos((\omega_{LO} \pm \omega_{IF})t)$.

C. Heterodyne detection

In order to appreciate its numerous advantages, before considering the heterodyne detection it is instructive to consider another detection scheme that is usually referred to as the *homodyne* detection. Consider a scheme shown in Fig. 3. Suppose that we want to measure an amplitude of a low-level harmonic signal $u_0\cos(\omega t)$. We can use, for example, an RF rectifying detector that we used previously. There is a simple way to amplify the detector signal using an additional generator, the local oscillator, generating a large-amplitude signal $u_{LO}\cos(\omega t)$, $u_{LO} >> u_0$, at the same frequency as the measured signal. Let us add the LO signal to the measured one and feed it to the detector, see Fig. 3. Thus, the detector input is $(u_{LO} + u_0)\cos(\omega t)$. Remember that the detector output is proportional to the square of the amplitude (that is the power) of the input RF signal. Thus at the detector output we have a DC signal

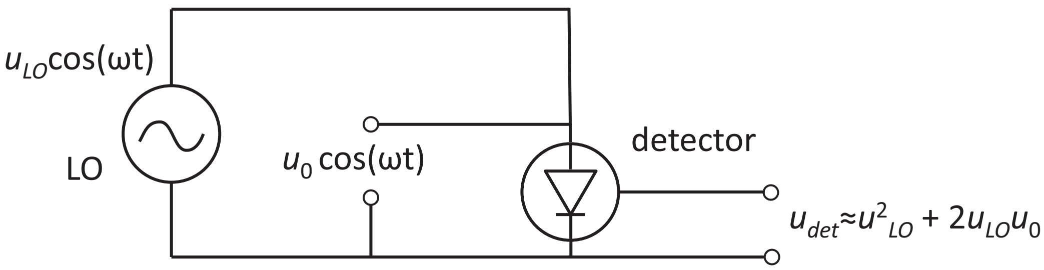

FIG. 3: Schematic of the homodyne detection.

$$u_{det} \propto u_{LO}^2 + 2u_{LO}u_O + u_0^2 \approx u_{LO}^2 + 2u_{LO}u_O, \tag{4}$$

where we can neglect the term small u_0^2. Note that the output of the detector is significantly larger than what we would have if we applied only signal $u_0 \cos(\omega t)$. In particular, there is a component proportional to the product $u_{LO}u_0$, which is much larger than just u_0^2. This simple technique is called the *homodyning*.

Now, we can use a modulation technique to extract the signal proportional to the measured signal amplitude u_0 (just like we did before!). For example, let us use the amplitude modulation (AM) of the measured signal, that is $u_0[1+m\cos(\omega_m t)]\cos(\omega t)$, where ω_m is the modulation frequency which is typically below 100 kHz. Then, the output of the detector reads

$$u_{det} \approx u_{LO}^2 + 2u_{LO}u_0[1 + m\cos\omega_m t], \tag{5}$$

and the component proportional to $u_{LO}u_0\cos(\omega_m t)$ can be detected using the lock-in amplifier with the reference signal at the modulation frequency ω_m. Again, note how the detected signal is amplified by a large factor $u_{LO}/u_0 >> 1$.

Homodyne detection is a very good technique. A slight disadvantage of this method is that by increasing the LO signal, that is the amplification factor u_{LO}/u_0, we also increase the huge back-ground DC signal proportional to u_{LO}^2. Since this signal will always have some fluctuations, this bring more noise to the detected signal. Let us now use the *heterodyne* method. We will use a mixer instead of the detector, and we will use slightly different frequencies for LO and detected signal (whose frequency we will call ω_{RF} in accordance with the notations adopted earlier), see Fig. 4. In this case, the output of the down-converting

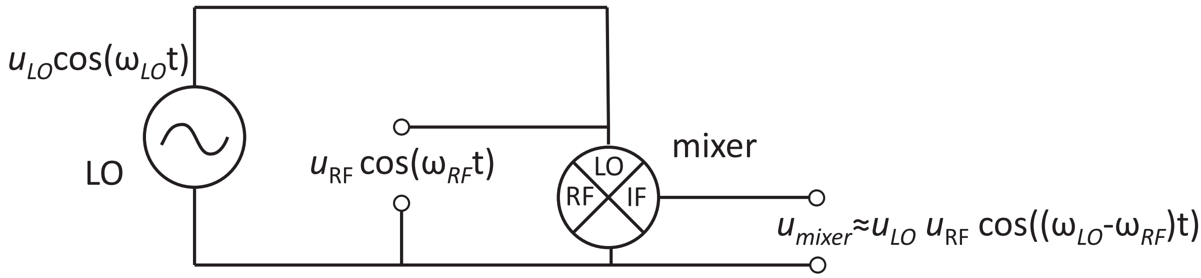

FIG. 4: Schematic of the heterodyne detection.

mixer provides the signal $u_{LO}u_{RF}\cos(\omega_{IF}t)$ at the low IF frequency $\omega_{IF} = \omega_{RF} - \omega_{LO}$ and with amplitude proportional to u_{RF}. The large background signal proportional to u_{LO} comes at much higher frequency ω_{LO} and can be efficiently removed by the low-pass filtering. Note that similar to the homodyne detection, we have achieved amplification of the detected RF signal by the factor u_{LO}/u_{RF}. In fact you might notice that the homodyne detection is a kind of heterodyne detection with $\omega_{LO} = \omega_{RF}$.

We can arrange detection of the signal $u_{LO}u_{RF}\cos(\omega_{IF}t)$ with a lock-in amplifier using its reference signal at ω_{IF}. Even more, we can use the amplitude modulation of the RF signal in the form $u_{RF}[1 + m\cos(\omega_m t)]\cos(\omega_{RF}t)$. Then, the mixer output will read

$$u_{RF}u_{LO}[1 + m\cos(\omega_m t)]\cos(\omega_{IF}t). \tag{6}$$

We can efficiently amplify this signal using an RF amplifier working at frequency ω_{IF}, remove the harmonic dependence $\cos\omega_{IF}t$ by a rectifying detector, and demodulate the signal using lock-in at the reference frequency ω_m. The down-conversion of the received high-frequency signal to an intermediate frequency followed by it amplification and further demodulation/detection is usually referred to as the *superheterodyning* and is a very powerful technique. It is enough to mention that almost all modern radio-frequency receivers are *suprheterodyne receivers.*