Chapter 3

CHAPTER 3

Bandgap and LDO

Voltage sources in integrated circuits are usually divided into voltage reference sources and linear regulators. As a stable reference, the bandgap reference and low-dropout(LDO) linear regulator is very important in different modules of analog and mixed-signal integrated circuits. As the voltage and current source the output voltage of bandgap has characteristics independent of temperature. And LDO provides a stable, pure DC voltage reference for on-chip circuits as a basic power supply system.

3.1 Bandgap

As a very important module of analog and mixed-signal IC, the voltage reference source plays a very important role in various electronic systems. With the increasing performance demand for various electronic products, the requirements for voltage reference source are also increasing. The voltage source's output voltage and noise determines the performance of circuits and systems.

The bandgap is fully compatible with standard CMOS process and can work at low power supply voltage. In addition, it has low temperature drift, low noise and high power supply rejection ratio, which can meet the requirements of most electronic systems. With these advantages, the bandgap has been widely studied and applied. In CMOS bandgap, low power supply voltage, low power consumption, high precision and high PSRR are the future development direction.

3.1.1 Basic of bandgap

Many modules in integrated circuits require voltage sources and current sources independent of temperature, which often affects the function of these modules. So how can we get a constant voltage or current reference that has nothing to do with the temperature? We assume that there are two identical physical quantities in the circuit. These two physical quantities have opposite temperature coefficients. When the two physical quantities are added to a certain weight, the voltage reference of zero temperature coefficient can be obtained, as shown in Fig.3.1.

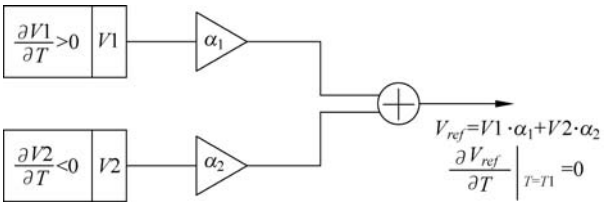

Fig. 3.1 Principle diagram of zero temperature coefficient

In Fig. 3.1, the voltage source $V1$ has a positive temperature coefficient $\left(\frac{\partial V1}{\partial T}>0\right)$, and the voltage source $V2$ has a negative temperature coefficient $\left(\frac{\partial V2}{\partial T}<0\right)$. We choose two weights of α_1 and α_2 to satisfy $\alpha_1 \cdot \frac{\partial V1}{\partial T} + \alpha_2 \cdot \frac{\partial V2}{\partial T} = 0$, then the voltage reference of zero temperature coefficient is obtained: $V_{ref} = V1 \cdot \alpha_1 + V2 \cdot \alpha_2$. The following task is how to get two voltage $V1$ and $V2$ with opposite temperature coefficients. In semiconductor technology, bipolar transistors can provide physical quantities of positive and negative temperature coefficient respectively. They are widely used in the design of bandgap reference. Recently, various literatures have also mentioned that the positive and negative temperature coefficients can be obtained by using MOS transistors working in subthreshold regions, but the accuracy of subthreshold region models needs to be investigated further. And the modern standard CMOS process provides the model of longitudinal PNP bipolar transistor, making the bipolar transistor still the first choice for bandgap reference.

1. Negative temperature coefficient voltage $\left(\frac{\partial \boldsymbol{V2}}{\partial \boldsymbol{T}}<\boldsymbol{0}\right)$

For a bipolar transistor, the relationship between the collector current I_c and the base-emitter voltage V_{BE} is as follows

$$I_c = I_s \cdot \exp(V_{BE}/V_T) \tag{3-1}$$

$$V_{BE} = V_T \cdot \ln(I_C/I_s) \tag{3-2}$$

In equation (3-1) and (3-2), I_s is the saturation current of transistor, V_T is the thermal voltage, $V_T = kT/q$, K is the Boltzmann constant, and Q is electronic charge. The derivative of equation (3-2) for V_{BE} is

$$\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T}\ln\frac{I_c}{I_s} - \frac{V_T}{I_s}\frac{\partial I_s}{\partial T} \tag{3-3}$$

Obtained from the theory of semiconductor physics:

$$I_s = b \cdot T^{4+m}\exp\frac{-E_g}{kT} \tag{3-4}$$

The derivative of (3-4) for temperature:

$$\frac{V_T}{I_s}\frac{\partial I_s}{\partial T} = (4+m)\frac{V_T}{T} + \frac{E_g}{kT^2}V_T \tag{3-5}$$

from equation(3-3) and (3-5):

$$\frac{\partial V_{BE}}{\partial T}=\frac{V_{BE}-(4+m)V_T-E_g/q}{T} \tag{3-6}$$

where $m\approx 1.5$, When the substrate is silicon, E_g = 1.12eV. While V_{BE} = 750mV, T = 300K, $\frac{\partial V_{BE}}{\partial T}$ = −1.5mV/℃。

It is known from equation (3-6) that the temperature coefficient $\frac{\partial V_{BE}}{\partial T}$ of V_{BE} itself is related to the temperature T. If the positive temperature coefficient is a temperature independent value, there will be errors in temperature compensation, resulting in a voltage reference that can only get a zero temperature coefficient at one temperature point.

2. Positive temperature coefficient voltage $\left(\frac{\partial \mathbf{V1}}{\partial \boldsymbol{T}}>\mathbf{0}\right)$

If two identical bipolar transistors are biased at different collector current, the difference between their base-emitter voltage is proportional to the absolute temperature, as shown in Fig.3.2.

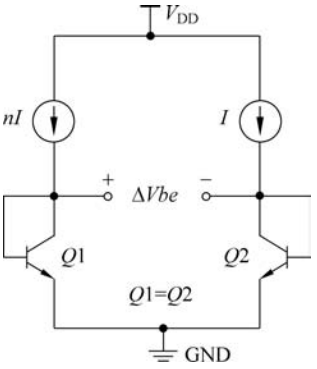

Fig. 3.2 Positive temperature coefficient voltage circuit

As shown in Fig.3.2, two identical bipolar transistors $Q1$ and $Q2$, which are biased at different collector current I_0 and nI_0. Ignoring their base current, there is:

$$\Delta V_{BE}=V_{BE1}-V_{BE2}=V_T\ln\frac{I_{c1}}{I_{s1}}-V_T\ln\frac{I_{c2}}{I_{s2}}=V_T\ln\frac{nI}{I_{s1}}-V_T\ln\frac{I}{I_{s2}} \tag{3-7}$$

There also is $I_{s1}=I_{s2}=I_s$, $I_{c1}=nI_{c2}$, then

$$\Delta V_{BE}=V_T\ln\frac{nI}{I_s}-V_T\ln\frac{I}{I_s}=V_T\ln n=\frac{kT}{q}\ln n \tag{3-8}$$

The derivative of (3-8) for temperature:

$$\frac{\partial \Delta V_{BE}}{\partial T}=\frac{k}{q}\ln n>0 \tag{3-9}$$

We can see that there is a positive temperature coefficient in equation (3-9), which is independent of temperature T.

3. Zero temperature coefficient voltage reference $\left(\frac{\partial V_{REF}}{\partial T}=0\right)$

By using the voltage of positive and negative temperature coefficients obtained in the above two sections, a voltage reference V_{REF} independent of temperature can be obtained, as shown in Fig. 3.3, equation (3-10) can de derived:

$$V_{REF}=\alpha_1 \cdot \frac{kT}{q}\ln n+\alpha_2 \cdot V_{BE} \tag{3-10}$$

The following shows how to select α_1 and α_2, then get the zero temperature coefficient voltage V_{REF}. At room temperature(300K), there is a negative temperature coefficient voltage $\frac{\partial V_{BE}}{\partial T}=-1.5\text{mV/K}$, and positive temperature coefficient voltage is $\frac{\partial \Delta V_{BE}}{\partial T}=\frac{k}{q}\ln n=0.087\text{mV/K}\cdot\ln n$

The derivative of (3-10) for temperature

$$\frac{\partial V_{REF}}{\partial T}=\alpha_1 \cdot \frac{k}{q}\ln n+\alpha_2 \cdot \frac{\partial V_{BE}}{\partial T} \tag{3-11}$$

assuming (3-11) equal to zero and $\alpha_2 = 1$, put the positive and negative temperature coefficient voltage in (3-11):

$$\alpha_1 \cdot \ln n = 17.2 \tag{3-12}$$

So the zero temperature coefficient voltage reference is

$$V_{REF} \approx 17.2\frac{kT}{q}+V_{BE} \approx 1.25\text{V} \tag{3-13}$$

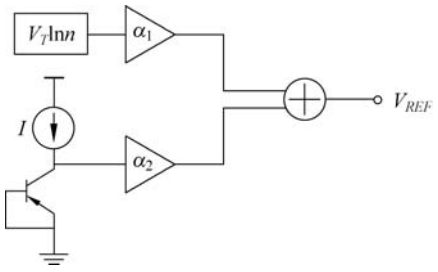

Fig. 3.3 Zero temperature coefficient voltage reference

4. Circuit of Zero temperature coefficient voltage reference

From the analysis of the previous section, the zero temperature coefficient voltage reference is obtained by adding the base-emitter voltage V_{BE} and $17.2 * kT/q$. The zero temperature coefficient voltage reference circuit is shown in Fig. 3.4. Suppose the voltage $V1 = V2$ in Fig. 3.4, then the left and right branches are equation (3-14) and (3-15), respectively.

$$V1 = V_{BE1} \tag{3-14}$$

$$V2 = V_{BE2} + IR \tag{3-15}$$

then

$$V_{BE1} = V_{BE2} + IR \tag{3-16}$$

so

$$IR = V_{BE1} - V_{BE2} = kT/q \cdot \ln n \tag{3-17}$$

put (3-17) into (3-15),

$$V2 = V_{BE2} + kT/q \cdot \ln n \tag{3-18}$$

Comparing (3-18) with (3-13), it is known that this circuit can obtain zero temperature coefficient voltage reference. The problem is how to make the voltage at both ends of the circuit equal in Fig. 3.4, that is, $V1 = V2$?

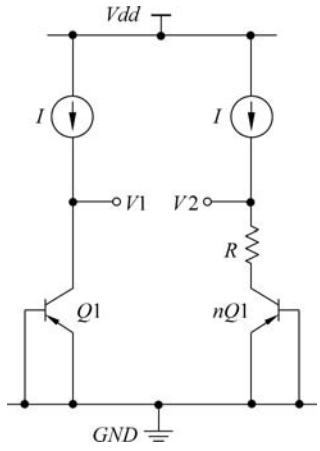

Figure 3.4 zero temperature coefficient voltage reference circuit

We know that when the ideal OPA operates normally, the voltage at the two input is approximately equal, so the following two circuits in Fig. 3.5 and 3.6 can be generated, making $V1 = V2$, respectively. There are two main kinds of circuit structure to complete adding, one is to add the two through an OPA. And its output is the voltage reference. The other is to generate a current proportional to absolute temperature (PTAT), which can be converted into a voltage through a resistor. This voltage naturally has a positive temperature coefficient, which is then added to the base-emitter voltage V_{BE}.

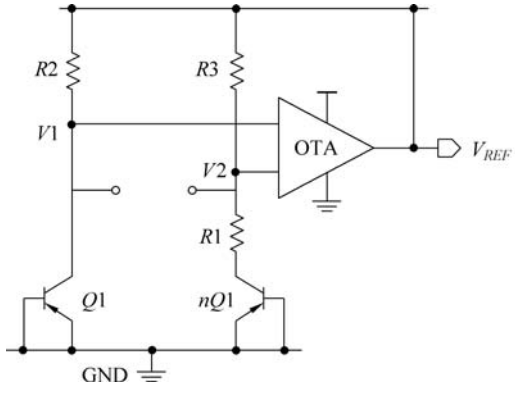

Fig. 3.5 the voltage reference circuit A

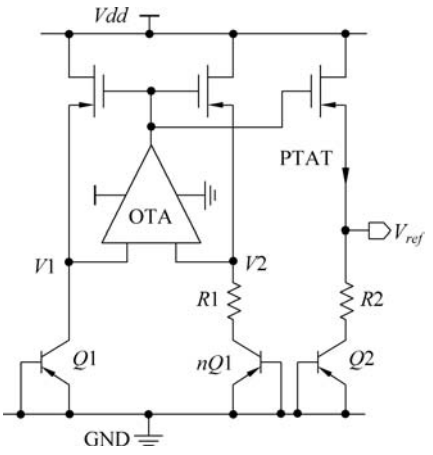

Fig. 3.6 the voltage reference circuit B

In Fig. 3.5, the input voltage of operational transconductance amplifier(OTA) is $V1$ and $V2$, and the output is V_{ref} that drives resistor $R2$ and $R3$. OTA makes the input voltage $V1$ and $V2$ approximately equal. The voltage difference between the base-emitter of the two bipolar transistors is $V_{\mathrm{T}} \ln n$, and the current that flows through $R1$ is

$$I_2 = \frac{V_T \ln n}{R_1} \tag{3-19}$$

get V_{ref} is,

$$V_{ref} = V_{BE,nQ1} + \frac{I_2}{R_1} \cdot (R_1 + R_3) = V_{BE,nQ1} + \left(1 + \frac{R_3}{R_1}\right) V_T \ln n \tag{3-20}$$

Combination of (3-13) and (3-20) shows that at room temperature 300K, zero temperature coefficient voltage can be obtained: $V_{ref} \approx 1.25\mathrm{V}$.

Fig. 3.6 is another circuit for obtaining a voltage reference. The principle of this circuit is to generate a current which is directly proportional to absolute temperature, and then convert it to voltage through resistor. Finally, the voltage is added to V_{BE} of bipolar transistor to get the voltage reference. As in Fig. 3.6, the current produced by the middle branch is still as equation (3-19). The current is PTAT, and the right mirror branch also produces a current of PTAT. This current flows through the resistor to form the PTAT voltage, and finally the base-emitter voltage of bipolar transistor $Q2$ is added to obtain the voltage reference

$$V_{ref} = V_{BE,Q2} + \frac{R_2}{R_1} V_T \ln n \tag{3-21}$$

The combination of (3-13) and (3-21) shows that a voltage reference of zero temperature coefficient can be obtained when $R2(\ln n) / R1 = 17.2$ is used. Then when $R2/R1 = 10$, we can choose $n = 6$.

Figure 3.7 is the third circuit for obtaining a voltage reference. The basic principle of this circuit is similar to that of second kinds. The difference is that two resistors

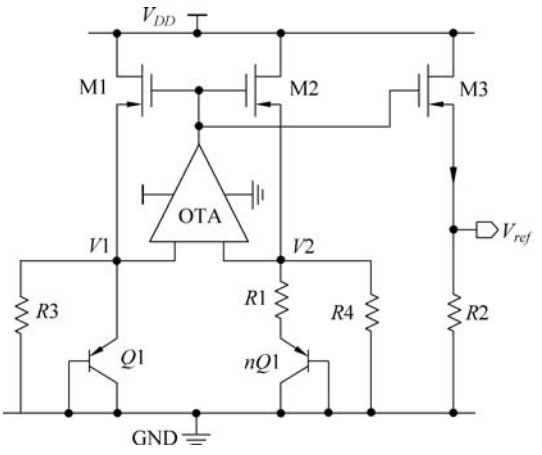

Fig. 3.7 the voltage reference circuit C

($R3 = R4 = R$) are added to the nodes $V1$ and $V2$ respectively, and the current flowing through the resistor is $I_R = V2/R4 = V_{BE1}/R$, So the current I_{M2} flows through the MOS transistor:

$$I_{M2} = I_R + I_{R1} = \frac{V_{BE1}}{R} + \frac{V_T \ln n}{R_1} \tag{3-22}$$

If the MOS transistors size $(W/L)_3 = (W/L)_2$, then there is $I_{M3} = I_{M2}$, and the voltage reference V_{ref} is obtained

$$V_{ref} = I_{M3} R_2 = \left(\frac{V_{BE1}}{R} + \frac{V_T \ln n}{R_1}\right) R_2 \tag{3-23}$$

It is known from equation (3-23) that the voltage reference of this structure can be obtained by adjusting the resistor value of $R2$. The first two structures can only get the reference voltage of 1.25V.

5. Offset in OTA

In Fig. 3.8, V_{os} is the offset voltage of OTA. The offset voltage makes the voltage reference error, and this error is also related to the temperature. Suppose that the OTA is ideal, so there is

$$V_{BE,Q1} - V_{os} \approx V_{BE,nQ1} + R_1 I_c \tag{3-24}$$

where I_c is the current flowing through resistor $R1$ and bipolar transistor $nQ1$, and the output voltage of OTA is

$$V_{ref} = V_{BE,nQ1} + (R_1 + R_3) I_c = V_{BE,nQ1} + (R_1 + R_3) I_c \tag{3-25}$$

combine(3-24) with (3-25):

$$V_{ref} = V_{BE,nQ1} + \left(1 + \frac{R_3}{R_1}\right)(V_T \ln n - V_{os}) \tag{3-26}$$

$$V_{ref} = V_{BE,nQ1} + \left(1 + \frac{R_3}{R_1}\right) V_T \ln n - \left(1 + \frac{R_3}{R_1}\right) V_{os} \tag{3-27}$$

According to equation (3-27), there is an error in the output voltage reference

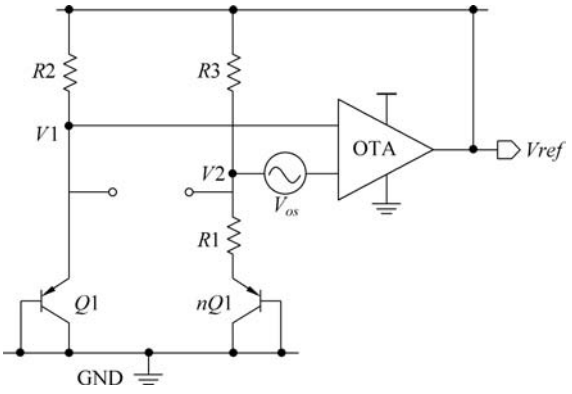

Fig. 3.8 Offset of OTA in Zero temperature coefficient voltage reference circuit

because of the offset voltage in OTA. This error is the result of $(1 + R3/R1)$ times of OTA offset. There are several circuits that can reduce the offset voltage:

(1) Increase the area of input transistors of an OTA, reduce the offset voltage by the design of the common-centroid layout.

(2) In high power supply, two bipolar transistors can be used in series, which makes ΔV_{BE} increase.

(3) Differnent current of two branches makes ΔV_{BE} increase from $\ln n$ to $\ln(MN)$, and reduces the ratio of V_{os} to V_{ref}, of which m and N are positive integers.

6. Start-up of bandgap

As shown in Fig. 3.6, the bandgap actually has two operating points, one is the working point when the circuit is normal, and the other is the zero current point. In power-on process, all the transistors in the circuit do not have current, and this state will be kept forever without no external interference. This is the problem of start-up.

In order to solve the start-up problem, an additional circuit is needed. The basic requirement of the start-up circuit is that after the power supply is stable, when circuit is in the "zero current" working state, the start-up circuit gives a stimulus to the internal circuit node, which forces it to get rid of the "zero current" working state and fall into the normal working mode. And when circuit is in normal operating mode, the start-up circuit stops working. The start-up circuit is in the right part in Fig. 3.9.

In Fig. 3.9, when the supply voltage is normally supplied and there is no current in the reference circuit, that is, PMOS transistor $PM4$ and $PM5$ have no current to pass through. While the node $Net1$ voltage is zero, PMOS transistor $PM3$ turns on, NMOS transistor $NM1$ cuts off. So the voltage of node $net3$ is $Vdd - 2V_{BE}$, which turns on NMOS transistor $NM2$ and makes node $net4$ connect to ground. Finally, the PMOS transistor $PM4$ and $PM5$ are on, and the voltage of node $Net1$ gradually increases to about $2V_{BE}$. The bandgap get to be operating normally.

Then in start-up circuit, the NMOS transistor $NM1$ is on, and $PM3$ is off, which makes the voltage of node $net3$ decrease gradually and $NM2$ is gradually cut off. After

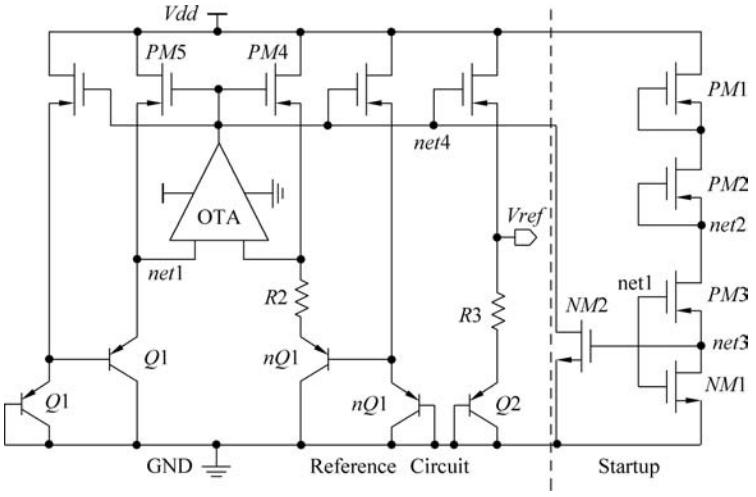

Fig. 3.9 Bandgap with start-up circuit

the normal operation of bandgap circuit, the two branches of start-up circuit stop working(*NM*2/*PM*3 is off).

3.1.2 Bandgap design

The performance of bandgap determines the precision that analog circuit can achieve, so its performance paremeter needs to be set according to the precision of functional module. Because the bandgap circuit mainly provides DC voltage(current), its low-frequency AC (< 10kHz), temperature and voltage characteristics are highly demanded. The main design difficulty is that it has good AC characteristics and stability while obtaining a lower temperature coefficient. The following is a brief introduction to its performance parameter.

1. Parameter

1) Temperature Coefficient(TC)

The temperature coefficient is a performance specification to measure the output voltage as a function of temperature. It is usually represented by ppm(parts per million). The equation is shown in (3-28).

$$\mathrm{TC(ppm/℃)}=\frac{V_{\max}-V_{\min}}{V_{\mathrm{mean}}(T_{\max}-T_{\min})}\times 10^{6} \tag{3-28}$$

where $V_{\max}$ and $V_{\min}$ are the maximum and minimum values of voltage reference obtained at required temperature range. V_{mean} is the average, while $T_{\max}$ and $T_{\min}$ are the maximum and minimum temperature.

2) Power Supply Rejection Ratio(PSRR)

Power supply rejection ratio is a parameter to measure the ability of output voltage to suppress the variation of power supply voltage. Because the power supply voltage is not fixed during the normal operation, there are various kinds of noises. The greater the

PSRR, the stronger the circuit's ability to suppress the power noise. There is

$$\text{PSRR}\mid_{\text{Hz}} = -20\log_{10}(\partial V_{ref}/\partial V_{\text{DD}})(\text{dB}) \tag{3-29}$$

$\partial V_{ref}/\partial V_{DD}$ is the ratio of the voltage reference to the change of power supply voltage at a certain frequency. The frequency we usually care about is 1kHz and 10kHz.

3) Power

Power consumption is a concern for any kind of integrated circuits. The lower power consumption means less power dissipation per unit time, which is especially important for IC packaging. Power consumption is also important for handheld devices, which is related to the lifetime of mobile device. But if the power consumption is too demanding, there may be noise and driving problems, so the power consumption of bandgap circuit should be in a reasonable range.

4) Start-up

The start-up of bandgap is not a quantitative parameter, but it determines whether its function is normal. Start-up problems are not observed in the usual simulation of transient, DC, and parameter scanning. Since there are two DC operating points in bandgap, if we want bandgap to leave the zero current state and enter normal working mode, we need to add the necessary "stimulus" to start the circuit. The normal operation of bandgap is usually confirmed by adding 0 to V_{DD} slope voltage signals on power supply.

2. Circuit design

A complete circuit of a bandgap is shown in Fig. 3.10.

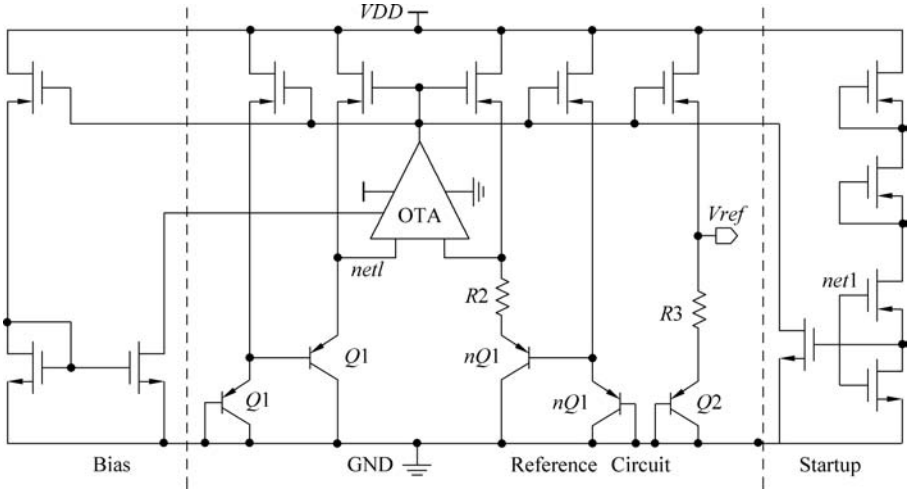

Fig. 3.10 Bandgap circuit

The bandgap is mainly divided into three parts, from left to right, followed by voltage bias circuit (Bias), voltage reference generator (Reference Circuit) and start-up circuit (Start-up). The voltage bias circuit is used to generate bias voltage when OTA works normally. From the Fig. 3.10, we can see that voltage bias comes from voltage

reference generator. The voltage reference generator is used to generate the required voltage reference(1.25V). We usually use the bipolar transistors in series to obtain the positive temperature coefficient voltage, which is helpful to reduce the influence of OTA offset to reference voltage.

As the positive temperature coefficient voltage is obtained by using the bipolar transistors in series, the current of resistor $R2$ is I_2 in Fig.3.10.

$$I_2 = \frac{2V_T \ln n}{R_2} \tag{3-30}$$

The resulting voltage reference V_{ref}

$$V_{ref} = V_{BE,Q2} + \frac{2R_3}{R_4} V_T \ln n \tag{3-31}$$

$$\frac{\partial V_{REF}}{\partial T} = \frac{\partial V_{BE,Q2}}{\partial T} + \frac{2R_3}{R_4} \cdot \frac{k}{q} \ln n \tag{3-32}$$

put $\frac{\partial V_{BE}}{\partial T} = -1.5\text{mV/K}$, $\frac{\partial \Delta V_{BE}}{\partial T} = \frac{k}{q}\ln n = 0.087\text{mV/K} \cdot \ln n$ into (3-32)

$$\frac{R_3}{R_4} \cdot \ln n = 8.6 \tag{3-33}$$

When we choose $n = 7$, $R_3 \approx 5R_4$; And $R_3 = 5\text{k}\Omega$, $R_4 = 25\text{k}\Omega$.

For OTA in a bandgap, the more important performance parameters are DC gain, gain bandwidth product, phase margin, and power supply rejection ratio. The OTA circuit is shown in Fig.3.11.

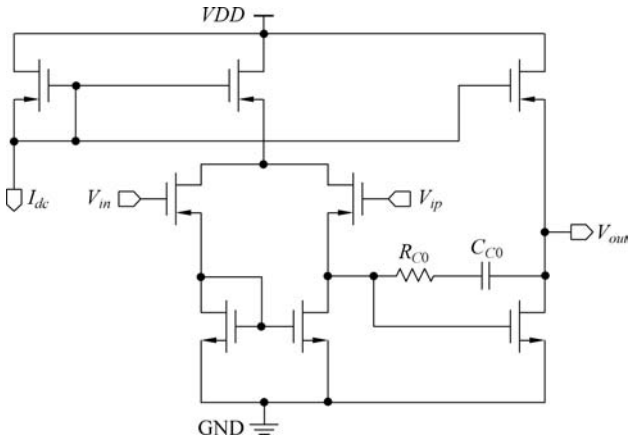

Fig. 3.11 OTA circuit

Because the output voltage accuracy and PSRR of bandgap is related to the DC gain of OTA, the two-stage structure is selected to increase its DC gain. The first stage uses a simple five-transistor structure to gain medium gain, and the second stage uses common source to provide a certain gain while providing a larger output swing. The capacitor C_{C0} between the first and second stage is the Miller-compensation capacitor. It is used to separate the two adjacent poles of OTA, pushing the dominant pole to the origin and

making the second pole leave the unit gain bandwidth. In addition, the second pole is further offset by resistor R_{C0}, which makes OTA have the characteristic of monopole and has a better phase characteristic.

In order to optimize the noise and offset voltage, we must ensure that the input transistors of the first stage have a large W/L and area. The large transconductance can effectively lower its flick noise.

3.2 Low-dropout linear regulator

Whether it is portable consumer electronics or large household electrical appliances, in operation the constant changes in load and the other various reasons make the power supply fluctuate in a large range, which is very harmful to the circuit. Especially for high-precision measurement, conversion and detection equipment, the power supply voltage is often required to be stable and has low noise. In order to meet the requirements, almost all electronic devices are powered by a Low-dropout linear regulator (LDO). The LDO has the advantages of simple structure, low cost, low noise and so on. It has been widely used in portable electronic equipment.

3.2.1 Basis of LDO

As a basic power supply module, LDO plays a very important role in analog integrated circuits. The changes in output load and the fluctuation of power supply voltage itself have a great influence on the performance of integrated circuit system. As a result, LDO is used as a linear regulator and is often used in systems with high performance requirements.

LDO adjusts the output voltage by principle of negative feedback, and obtains the stable DC output voltage on the basis of providing a certain output current capability. In normal working state, the output voltage is independent of load, input voltage change and temperature. The minimum input voltage of LDO is determined by the minimum voltage drop of the adjusting transistor, usually 150-300mV.

The basic structure of LDO is shown in Fig. 3.12. LDO is mainly composed of the following parts: bandgap, error amplifier, feedback / phase compensation network, and adjusting transistor. The error amplifier, feedback resistor network, adjusting transistor and phase compensation network constitute the stable output voltage V_{out} of feedback loop.

In LDO, a bandgap provides a voltage reference independent of temperature and power. The error amplifier amplifies the difference between voltage reference and feedback voltage, so that the feedback voltage is basically equal to the voltage reference. Phase compensation network is used to compensate the phase of whole feedback network to ensure the feedback network is stable. The adjusting transistor outputs a stable voltage

under the control of error amplifier output. The adjusting transistor is a PMOS transistor, and can also be a NMOS ora NPN transistor. In CMOS process, the PMOS is usually selected.

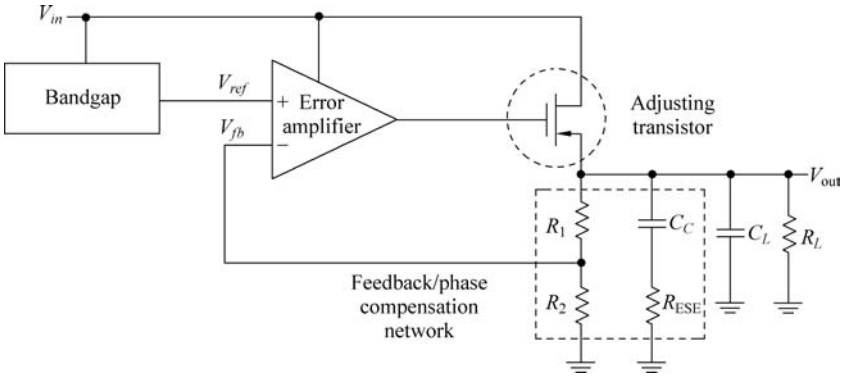

Fig. 3.12 LDO structure

3.2.2 Operational principle

Fig. 3.13 is a LDO curve of the relationship between output voltage and input voltage. The abscissa are the input voltage 0-3.3V, and the ordinate is output voltage. When the input voltage is less than a certain value(1.2-1.8), the output voltage is zero. When the input voltage Vin is greater than it (1.2-1.8V), the output voltage V_{out} increases with V_{in}. When Vin is greater than 2.1V, LDO is in normal operating state and the output voltage is stable at 1.8V.

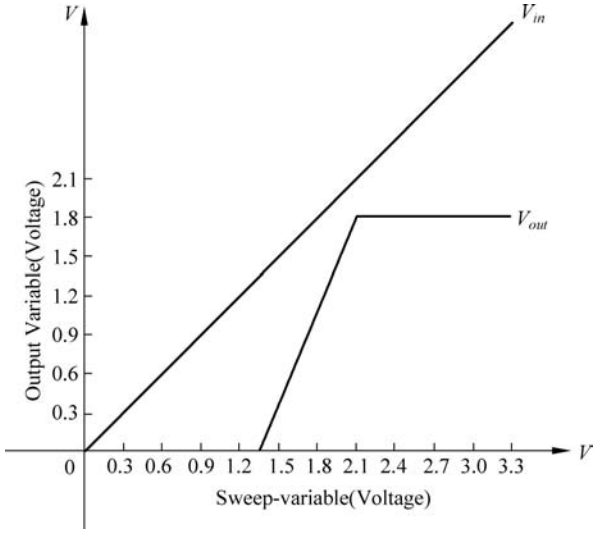

Fig. 3.13 the relationship between output voltage and input voltage

The operational principle of LDO shown in Fig. 3.12 is: when LDO is on power, the start-up circuit in bandgap begins to work to ensure that the whole system begins to

operate normally. The bandgap outputs a stable voltage reference V_{ref} which is independent of supply voltage and temperature, while the $R1$ and $R2$ of the feedback / phase compensation network generate feedback voltage V_{fb}. The two voltages are input to error amplifier for comparison. The error amplifier amplifies the result and adjusts the gate voltage of transistors to control their current, finally stabilizes the output voltage. The whole adjustment loop is a stable negative feedback system. When the input voltage V_{out} increases, the input V_{fb} of feedback resistor network will also increase. V_{fb} and V_{ref} are compared and amplified, so that the gate voltage of the transistors is increased, and the output current is reduced. At last the output voltage V_{out} decreases and stays at a stable voltage value.

The closed loop expression of negative feedback loop is

$$V_{out} = \frac{A_{ol}}{1 + A_{ol}\beta} V_{ref} \tag{3-34}$$

where A_{ol} is the open-loop gain of negative feedback loop, and β is the feedback coefficient, and its expression is

$$\beta = \frac{R_2}{R_1 + R_2} \tag{3-35}$$

when $A_{ol}\beta \gg 1$, equation (3-34) can be expressed as

$$V_{out} \approx \frac{V_{ref}}{\beta} = \frac{R_1 + R_2}{R_2} V_{ref} \tag{3-36}$$

It can be seen from (3-35) and (3-36), the output voltage V_{out} of LDO is only related to the voltage reference V_{ref} and the resistor ratio of feedback resistor network. It has nothing to do with the input voltage V_{in}, load current and temperature. Therefore, the required output voltage can be obtained by adjusting the resistor ratio.

3.2.3 Parameter

The main parameter of LDO are divided into static and dynamic parameter. The static parameter includes dropout voltage、quiescent current and efficiency. And dynamic parameter consists of transient response、line regulation、load regulation and power supply rejection ratio.

1. Dropout Voltage

When input voltage V_{in} is less than a certain value of V_{cutoff}, the LDO output zero voltage, and it is in the cutoff region; While V_{in} is between V_{cutoff} and critical voltage V_{Break}, the output voltage is not fixed, and LDO has no adjustment ability; And when V_{in} is greater than V_{Break}, LDO enters normal operating state and the output voltage remains unchanged. We define that the difference between V_{Break} and output voltage is the dropout voltage of LDO, and the dropout voltage is usually 150mV-300mV. Its expression is

$$V_{drop} = V_{in} - V_{Break} \tag{3-37}$$

In general, the smaller dropout voltage is beneficial to improve the conversion efficiency of LDO. However, a too small dropout voltage may cause a poor phase margin and power supply rejection ratio to the whole feedback loop, so a compromise should be considered in LDO design.

2. Quiescent Current

Quiescent current(I_Q) usually refers to the current consumed by the whole LDO when it is not connected to any load, or it is defined as the difference between input current and output current with output load. The expression is shown

$$I_Q = I_{in} - I_{out} \tag{3-38}$$

The quiescent current is mainly composed of the bias current of active devices such as bandgap, error amplifier, active compensation network (if exist), and the current consumed by the feedback resistor network. The smaller quiescent current is beneficial to improve the conversion efficiency of LDO and the life of the battery. In CMOS integrated circuit, the quiescent current of the LDO is in the dozens of μA or smaller.

3. Efficiency

The efficiency of LDO is defined as the percentage of ratio of output power P_{out} to input power P_{in}

$$\beta = \frac{P_{out}}{P_{in}} \cdot 100\% = \frac{I_{out} \cdot V_{out}}{I_{in} \cdot V_{dd}} \cdot 100\% = \frac{(I_{in} - I_Q) \cdot (V_{dd} - V_{drop})}{I_{in} \cdot V_{dd}} \cdot 100\% \tag{3-39}$$

where I_{in} is input current, V_{dd} is supply voltage, I_Q is quiescent current and V_{drop} is dropout voltage.

We can know that from (3-39), the efficiency of LDO related to quiescent current and dropout voltage. The way to improve efficiency is to reduce I_Q and V_{drop}.

4. Transient Response

The transient response of LDO includes two aspects: ① the linear transient response due to supply voltage change; ② the load transient response when load and output voltage suddenly changes. The linear transient response is more important in frequent power up and power down application, while the load transient response is more important when the load current changes frequently. In specific application, the latter occurs in real time, so the load transient response is paid more attention in design.

Fig. 3.14 is the transient response of LDO. The upper half part is the step change of output load current, and the lower part is the output voltage change with load current. The changes in load transient response can be expressed

$$\Delta V_{\mathrm{TR,max}} = \frac{I_0 \cdot \Delta t}{C_0} + \Delta V_{\mathrm{ESR}} \tag{3-40}$$

I_0 is load current. C_0 is output capacitor. ΔV_{ESR} is the output voltage variation caused by the equivalent series resistor(ESR) of output capacitor. And Δt is step response time. In practical application, the smaller transient response time, the better of LDO performance. By equation (3-40), with the fixed load current increasing the output

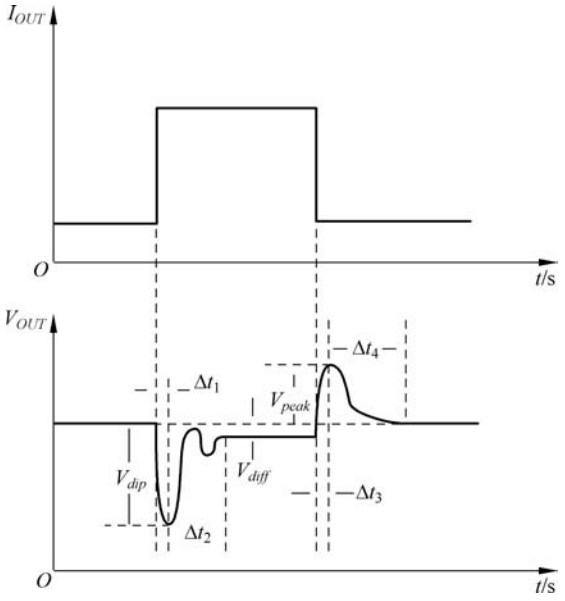

Fig. 3.14 LDO load transient response

capacitor, the closed-loop bandwidth and ESR will reduce the amplitude of load transient response, thereby reducing the load response time. In LDO design, the frequency response method should be adopted to determine the stability of the whole feedback loop first. Because LDO is a closed-loop system with multiple poles and negative feedback, if the design is improper, there may be a stability problem, which will naturally affect its time-domain transient characteristics.

5. Line Regulation

The line regulation of LDO is defined as the variation ratio of output to input voltage when input voltage changes under constant load. It can reflect the ability of LDO to restrain the change of input voltage. Its expression is

$$S_{LR} = \frac{\Delta V_{out}}{\Delta V_{in}} \Big|_{I_{out}=constant} \tag{3-41}$$

In Figure 3.15, it is assumed that the load current is kept at 50mA, and when the input voltage is 2.1V, the LDO enters adjustment state and outputs a 1.8V voltage. The

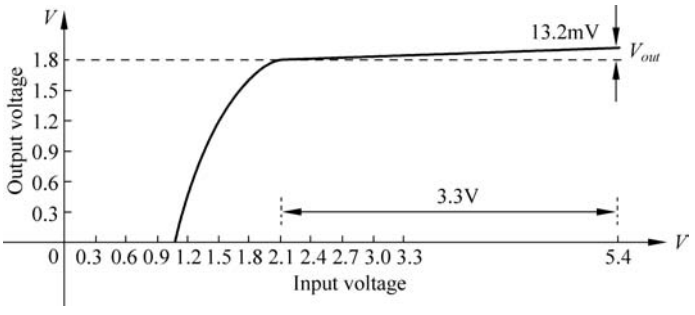

Fig. 3.15 Line regulation of LDO

input voltage is gradually increased to 5.4V, and the output voltage varies from 1.8V to 1.8132V, so the line regulation of LDO is

$$S_{LR}=\frac{0.0132}{(5.4-2.1)}\bigg|_{I_{out}=50\text{mA}}=4\text{mV/V}$$

6. Load Regulation

The load regulation of LDO is defined as the ratio of the output voltage variation ΔV_{out} to the load current variation ΔI_{load} when input voltage V_{in} remains unchanged. The load regulation reflects the influence of load current on output voltage, and the smaller, the better it is. Its expression is:

$$S_{LOR}=\frac{\Delta V_{out}}{\Delta I_{load}}\bigg|_{V_{in}=constant} \tag{3-42}$$

Figure 3.16 gives a schematic of load regulation, where I_{load} and ΔI_{load} are load current and variation value respectively, and V_{out} and ΔV_{out} are output voltage and output voltage variation value respectively.

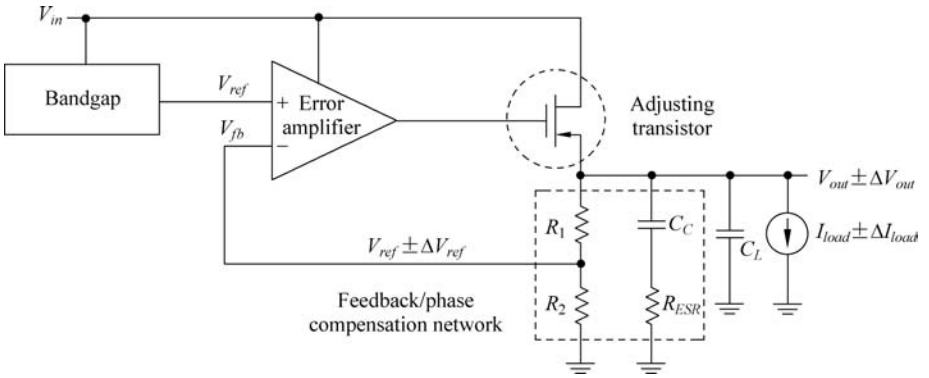

Fig. 3.16 Circuit diagram for calculating load regulation

When the output load current changes to $I_{load}+\Delta I_{load}$, the output voltage changes to $V_{out}+\Delta V_{out}$, which makes the output voltage variation of feedback resistor network is delta ΔV_{fb}

$$\Delta V_{fb}=\Delta V_{out}\cdot\frac{R2}{R1+R2} \tag{3-43}$$

and ΔI_{load} is

$$\Delta I_{load}=g_{mEA}\cdot g_{mp}\cdot\Delta V_{fb} \tag{3-44}$$

g_{mEA} and g_{mp} are the transconductance of error amplifier and adjusting transistor. Putting (3-43) and (3-44) into (3-42), the expression of load regulation is

$$S_{LOR}=\frac{1}{g_{mEA}g_{mp}}\cdot\frac{R_1+R_2}{R_2}\,|_{V_{in}=constant} \tag{3-45}$$

From equation (3-45), we can see that we can improve load regulation by increasing the transconductance of error amplifier and adjusting transistor, and the feedback coefficient of feedback resistor network.

In Fig. 3.17, assuming the input voltage V_{in} = 3.3V, the load current is changed from 0 to 50mA, and the output voltage varies about 1.32mV near 1.8V, then load regulation is

$$S_{LOR} = \frac{1.32\text{mV}}{50\text{mA}} \Big|_{V_{in}=3.3\text{V}} = 0.0264\text{V/A} \Big|_{V_{in}=3.3\text{V}} \tag{3-46}$$

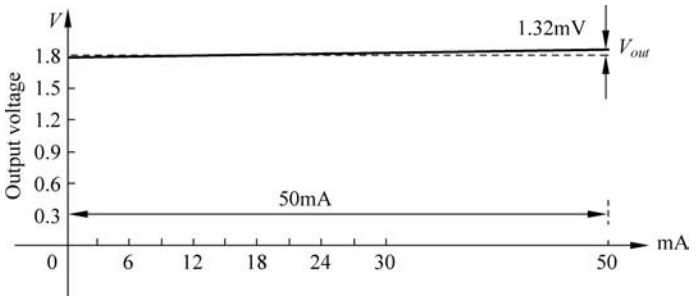

Fig. 3.17 Load regulation

7. Power Supply Rejection Ratio(PSRR)

The PSRR is defined as the ratio of output voltage variation to power supply voltage variation, that is, the small-signal gain of output voltage to supply voltage in a certain frequency range. The PSRR reflects the ability of output voltage to suppress the noise of power supply. Its expression is

$$\text{PSRR} \Big|_{\text{Hz}} = 20 \cdot \log_{10}\left(\frac{\Delta V_{out}}{\Delta V_{dd}}\right) \text{dB} \tag{3-47}$$

The definition of PSRR is similar to line regulation, but there is an essential difference. The line regulation represents a large signal, DC characteristic, and a small-signal and AC characteristic is expressed by PSRR. The PSRR is related to LDO structure, PSRR of bandgap and error amplifier. Fig. 3.18 is a typical PSRR curve. Because LDO mainly provides DC voltage, the PSRR within 10KHz is more important. Usually high-precision circuits are required for PSRR within 100kHz.

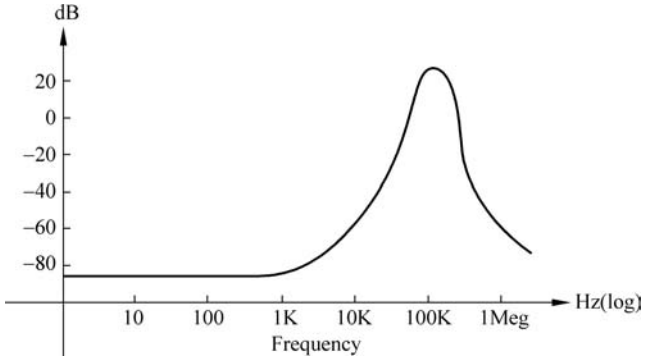

Fig. 3.18 PSRR curve

3.2.4 Stability analysis

LDO is a closed-loop system with negative feedback. The system outputs feedback voltage to input terminal and compares with the voltage reference for adjusting, finally making the output stable on a expected value. Due to the complexity of feedback system and existence of multiple zero-poles, there is a stability problem in itself. If the system is not designed properly, it will cause the whole negative feedback system to oscillate. Therefore we must analyze the open-loop amplitude-frequency and phase-frequency characteristics, so as to ensure the stability.

Usually, the condition of ensuring stability is that the phase margin of open-loop characteristic is greater than 45 degrees. Considering the transient characteristics, overshoot and other properties, the 60 degree is an ideal value.

General LDO compensates the poles to achieve a better open-loop phase margin by using the equivalent series resistor(ESR) of off-chip capacitors, as Fig. 3.19 shows.

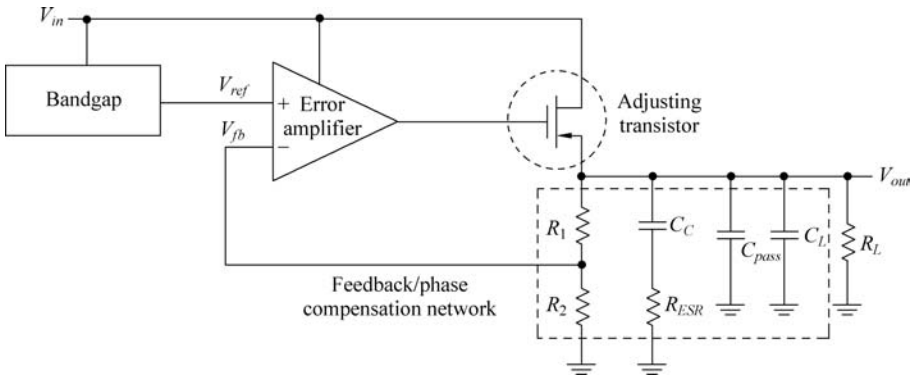

Fig. 3.19 LDO with off-chip capacitors

In Fig. 3.20, $R1$ and $R2$ are feedback network resistor, C_C is off-chip compensation capacitor, and R_{ESR} is the equivalent series resistor of compensation capacitor. The range of R_{ESR} is 10mΩ-1Ω. C_C generally takes a value of 0.1μF-10μF, and a tantalum capacitor with good performance is generally used. Because the value of C_C is large, the output node Vout is the dominant pole of LDO. C_{pass} is a bypass capacitor. Usually, when the load current changes suddenly and the output voltage generates large ripple, C_{pass} will reduce the ripple amplitude and make the output voltage stable quickly. The load capacitor CL is usually smaller than C_C and C_{pass}. For simplicity, it is not considered in the following analysis.

When analyzing the small-signal characteristic of off-chip compensation LDO, we need to break a certain point in the feedback loop. Start from this point, analyzing the components of system sequentially, and finally returning to the starting point, as Fig. 3.20 shows.

In Fig. 3.21, we disconnect the feedback network and define error amplifier input

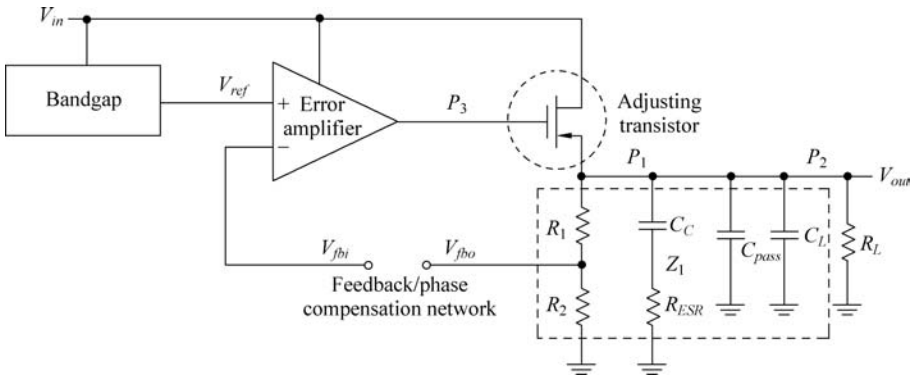

Fig. 3. 20 Analysis of small-signal characteristic for off-chip compensation LDO

V_{fbi} as the starting point of analysis, and the feedback resistor output is V_{fbo}. So the equivalent output resistor of error amplifier is R_{EA}, and the gate equivalent input capacitor of adjusting transistor is C_{GP}, and the source-drain equivalent resistor is R_{DSP}.

From Fig. 3.21, the output resistor of V_{out} is $R_o = (R_1 + R_2) /\!/ R_{DSP} /\!/ R_L \approx R_{DSP}$, its equivalent resistor is Z_{out},

$$Z_{out} = R_o \ /\!/ \ (1/sC_b) \ /\!/ \ (R_{ESR} + 1/sC_c) \tag{3-48}$$

put (3-48) into (3-49)

$$Z_{out} = \frac{R_o \cdot (1 + sR_{ESR}C_c)}{s^2 R_o R_{ESR} C_c C_{pass} + s\,[C_c(R_o + R_{ESR}) + R_o C_{pass}] + 1} \tag{3-49}$$

The open-loop transfer function $H(s)$ is

$$H(s) = \frac{V_{fbo}}{V_{fbi}} = \frac{R_2}{R_1 + R_2} \cdot \frac{g_{EA} R_{EA} g_{MP}}{1 + sR_{EA}C_{GP}} \cdot Z_{out}$$

$$H(s) = \frac{R_2}{R_1 + R_2} \cdot \frac{g_{EA} R_{EA} g_{MP} (1 + sR_{ESR}C_c)}{[1 + s(R_0 + R_{ESR})C_c] \cdot [1 + s(R_0 /\!/ R_{ESR})C_{pass}] \cdot (1 + sR_{EA}C_{GP})} \tag{3-50}$$

Considering $R_0 \approx R_{DSP} \ll R1 + R2$, $R_0 \gg R_{ESR}$, the poles and zeroes of entire LDO open-loop system are

$$P_1 = -\frac{1}{(R_0 + R_{ESR}) \cdot C_c} = -\frac{1}{R_{DSP} \cdot C_c} \tag{3-51}$$

$$P_2 = -\frac{1}{(R_0 /\!/ R_{ESR}) \cdot C_{pass}} = -\frac{1}{R_{ESR} \cdot C_{pass}} \tag{3-52}$$

$$P_3 = -\frac{1}{R_{EA} \cdot C_{GP}} \tag{3-53}$$

$$Z_1 = -\frac{1}{R_{ESR} \cdot C_c} \tag{3-54}$$

In general, $P1 < P2 < P3$, and $P1 < Z1 < P2$. The frequency response is shown in Fig. 3.47.

The system's open-loop phase margin formula shows that if the phase margin is at

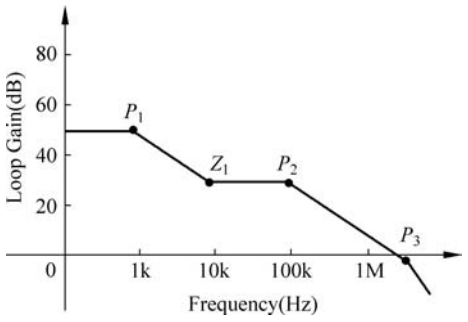

Fig. 3.21 Frequency response

least 60 degrees, there is only one pole in the unit gain bandwidth (UGB). Therefore it shows the characteristics of single pole system, and the non dominant pole is outside 2.2 times of UGB. In Fig. 3.21, if the phase margin is up to 60 degrees, Z1 must be adjacent to P2 and P3 > 2.2UGB, which means that zero Z1 must be used to compensate for the non dominant pole P2. Furthermore the choice of the equivalent series resistance R_{ESR} must also be in a reasonable range. If R_{ESR} is too large, the zero Z1 near the low frequency will cause UGB to increase. It may cause pole P3 to go into UGB, and the phase margin becomes smaller, which results in the instability of system; But If the R_{ESR} is too small, the zero Z1 moves to high frequency, which may move out of UGB and can not achieve compensation. So R_{ESR} has an important impact on system's stability. In fact, due to the large size of adjusting transistor, the gate-source capacitor also constitutes a zero.

Because of error amplifier with large output impedance (R_{EA}) and gate capacitor of adjusting transistor (C_{GP}), its output has a large time constant. Since its large and small-signal response time are subject to these restrictions, LDO as shown in Fig. 3.19 is not commonly used. In order to improve the accuracy and transient characteristics of LDO, a buffer is usually inserted between error amplifier and adjusting transistor.

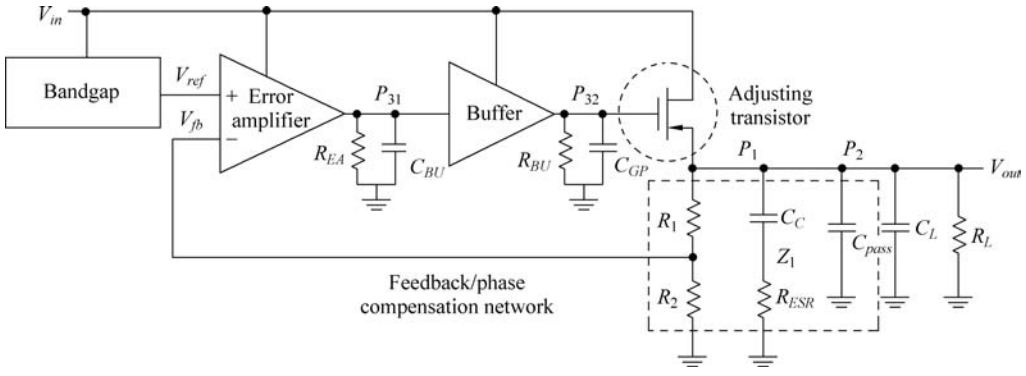

Fig. 3.22 LDO with buffer

As shown in Fig. 3.22, the voltage buffer isolates error amplifier from adjusting transistor, and divides the original pole P3 into two poles: P31 and P32.

$$P_{31}=\frac{1}{R_{EA}C_{BU}} \tag{3-55}$$

$$P_{32}=\frac{1}{R_{BU}C_{GP}} \tag{3-56}$$

Figure 3.23 is the amplitude-frequency characteristic of LDO with buffer. When the buffer is added, the poles $P31$ and $P32$ are greater than the original pole $P3$, so the system is more stable and improves its transient response time greatly.

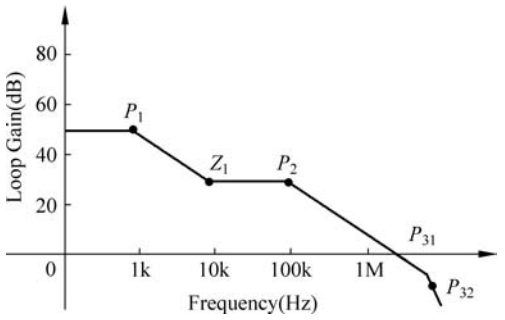

Fig. 3.23 amplitude-frequency characteristic of LDO with buffer

3.3 Technical words and phrases

3.3.1 Terminology

bandgap	带隙基准源
low-dropout linear regulator(LDO)	低压差线性稳压器
positive (negetive) temperature coefficient	正(负)温度系数
subthreshold region	亚阈值区
Boltzmann constant	玻尔兹曼常数
proportional to absolute temperature(PTAT)	与绝对温度成正比
start-up	启动电路
error amplifier	误差放大器
static parameter	静态参数
dynamic parameter	动态参数
dropout voltage	电压差
quiescent current	静态电流
efficiency	转换效率
line regulation	线性调整率
load regulation	负载调整率
equivalent series resistor(ESR)	等效串联电阻

3.3.2 Note to the text

(1) The bandgap is fully compatible with standard CMOS process and can work at low power supply voltage. In addition, it has low temperature drift, low noise and high power supply rejection ratio, which can meet the requirements of most electronic systems.

带隙基准源与标准 CMOS 工艺完全兼容，可以工作于低电源电压下等优点，另外还具有低温度漂移、低噪声和较高的电源抑制比等性能，能够满足大部分电子系统的要求。

(2) In semiconductor technology, bipolar transistors can provide physical quantities of positive and negative temperature coefficient respectively.

在半导体工艺中，双极晶体管能够分别提供正、负温度系数的物理量。

(3) If two identical bipolar transistors are biased at different collector current, the difference between their base-emitter voltage is proportional to the absolute temperature.

如果两个相同的双极晶体管在不同的集电极电流偏置情况下，那么它们的基极-发射极电压的差值与绝对温度成正比。

(4) There are two main kinds of circuit structure to complete adding, one is to add the two through an OPA. And its output is the voltage reference. The other is to generate a current proportional to absolute temperature (PTAT), which can be converted into a voltage through a resistor. This voltage naturally has a positive temperature coefficient, which is then added to the base-emitter voltage V_{BE}.

完成这种相加的电路结构目前主要有两种，一种是通过运算放大器将两者进行相加，输出即为基准电压；另一种是先产生与温度成正比(PTAT)的电流，通过电阻转换成电压，这个电压自然具有正温度系数，然后与二极管的基极-发射极电压 V_{BE} 相加获得。

(5) In order to solve the start-up problem, an additional circuit is needed. The basic requirement of the start-up circuit is that after the power supply is stable, when circuit is in the "zero current" working state, the start-up circuit gives an stimulus to the internal circuit node, which forces it to get rid of the "zero current" working state and fall into the normal working mode.

为了解决电路的启动问题，需要加入额外电路，使得存在启动问题的电路摆脱"零电流"工作状态进入正常工作模式，对启动电路的基本要求是电源电压稳定后，待启动电路处于"零电流"工作状态时，启动电路给内部电路某一节点激励信号，迫使待启动电路摆脱"零电流"工作状态，而进入正常工作模式。

(6) In general, the smaller dropout voltage is beneficial to improve the conversion

efficiency of LDO. However, a too small dropout voltage may cause a poor phase margin and power supply rejection ratio to the whole feedback loop, so a compromise should be considered in LDO design.

通常情况下，要求 LDO 的电压差越小越好，以提高整体电路的转换效率；但是过小电压差可能会造成整个反馈环路的相位裕度、电源抑制比很差，所以在系统设计时要折中考虑。

(7) The line regulation of LDO is defined as the variation ratio of output to input voltage when input voltage changes under constant load.

LDO 的线性调整率定义为在负载保持恒定的情况下，输入电压发生变化时，输出电压变化量与输入电压变化量的比值。

(8) The definition of PSRR is similar to line regulation, but there is an essential difference. The line regulation represents a large signal, DC characteristic, and a small-signal and AC characteristic is expressed by PSRR.

电源抑制比的定义与线性调整率虽然相似，但是有本质的区别，线性调整率表示的是大信号、直流特性，而电源抑制比表示的小信号、交流特性。

(9) Usually, the condition of ensuring stability is that the phase margin of open-loop characteristic is greater than 45 degrees. Considering the transient characteristics, overshoot and other properties, the 60 degree is an ideal value.

通常情况下，确保负反馈系统稳定的条件是其开环特性的相位裕度大于 45°，考虑到瞬态建立特性、过冲以及其他性能，相位裕度达到 60°是一个比较理想的值。

(10) In order to improve the accuracy and transient characteristics of LDO, a buffer is usually inserted between error amplifier and adjusting transistor.

为了提高 LDO 的精度和瞬态特性，一般在误差放大器和调整晶体管之间插入一个缓冲器。