$$I = I_0 = BU_T \ln K/R_3, \tag{8.30}$$

which serves as the start-up current of the oscillator. As the oscillation voltage grows, it is superimposed on the DC component of gate voltages V_{Gdc} for M_1, *but not* for M_2, since it is blocked by the low-pass filter R_2-C_2. Because of the exponential function $I_D(V_G)$ of M_1, its average drain current should increase, which is not compatible with the 1-to-1 ratio imposed by the mirror M_4-M_3. Instead, V_{Gdc} decreases, resulting in a decrease of the output current I.

Assuming that all transistors remain saturated, that M_1 remains in weak inversion even during the peaks of its drain current, and that the residual oscillation amplitude at the gate of M_2 is much smaller than U_T (so that it has no nonlinear effect), the transfer characteristics are given by [31]

$$I = I_0 \cdot \left(1 - \frac{\ln \mathrm{I_{B0}}(\frac{V}{nU_T})}{\ln K}\right), \tag{8.31}$$

where $\mathrm{I_{B0}}$ is the 0-order modified Bessel function. They are plotted in Figure 8.9(b) for several values of K.

The amplitude of oscillation will stabilize when the regulator delivers exactly the bias current I required to produce the amplitude V.

If the peak drain current of M_1 leaves weak inversion, the transfer characteristics will be modified and may eventually loose their monotonicity, which must absolutely be avoided to maintain stable oscillation. A semi-empirical condition to ensure monotonicity is

$$\beta_1 > \frac{2\beta_3/\beta_4}{nU_T R_3}. \tag{8.32}$$

The role of capacitor C_3 is to keep the drain voltage of M_1 sufficiently constant to avoid de-saturation during the positive peaks of current.

At low frequencies, high (non-critical) values may be needed for resistors R_1 and R_2. Very high values have been obtained by using lateral diodes in the polysilicon layer [167, 37]. Lower values can be obtained by means of transistors adequately biased [46].

This amplitude regulator has been applied extensively to quartz oscillators in watches [37, 46, 168, 169, 170], but it can be used in different type of sinusoidal oscillators as well [38].

8.4.3 Translinear Circuits

Discovered for bipolar transistors, the translinear principle [171] is an outstanding application of the exponential characteristics of MOS transistors in weak inversion. Consider the loops of saturated transistors illustrated in Figure 8.10. They include an even number of transistors, half of which have their gate to source "junction" in the clockwise (cw) direction, the other half in the counter-clockwise (ccw) direction. Hence, for the whole loop:

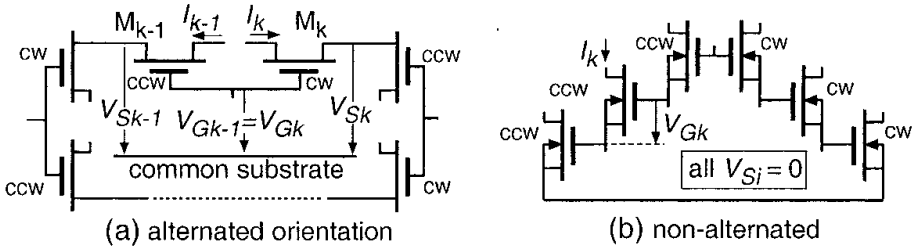

Fig. 8.10. Translinear loops: (a) alternated orientation of transistors in a common substrate; (b) non-alternated orientation: separate local substrates are necessary.

$$\sum_{cw}(V_{Gi} - V_{Si}) = \sum_{ccw}(V_{Gi} - V_{Si}). \tag{8.33}$$

If all transistors are in weak inversion, with negligible reverse current (saturated), then according to (5.41):

$$I_i = I_{D0i} \exp \frac{V_{Gi}/n_i - V_{si}}{U_T} \quad \text{or} \quad \frac{V_{Gi}}{n_i} - V_{Si} = U_T \ln \frac{I_i}{I_{D0i}}. \tag{8.34}$$

Now if cw and ccw transistors are *alternated* [172] as in Figure 8.10(a), then each gate voltage V_{Gi} is common to a pair cw-ccw of transistors. It appears therefore in both sides of equation (8.33), which can thus be rewritten as

$$\sum_{cw}(\frac{V_{Gi}}{n_i} - V_{Si}) = \sum_{ccw}(\frac{V_{Gi}}{n_i} - V_{Si}). \tag{8.35}$$

We can now introduce (8.34), divide by U_T (that is common to all transistors) and exponentiate both sides of the equation, which yields

$$\prod_{cw} \frac{I_i}{I_{D0i}} = \prod_{ccw} \frac{I_i}{I_{D0i}} \quad \text{or} \quad \frac{\prod_{cw} I_i}{\prod_{ccw} I_i} = \frac{\prod_{cw} I_{D0i}}{\prod_{ccw} I_{D0i}} = \lambda. \tag{8.36}$$

This result is independent of the temperature. If I_{D0} is the same for all transistors, then $\lambda = 1$. A circuit may include several loops sharing some transistors, each loop characterized by its value of λ. The mismatch of I_{D0i} is dominated by that of V_{T0i} and results in an error in the value of λ, with a standard deviation

$$\frac{\sigma(\Delta\lambda)}{\lambda} = \frac{1}{nU_T} \left[\sum \frac{1}{2}\sigma^2(\Delta V_{T0i})\right]^{1/2}, \tag{8.37}$$

where the factor $1/2$ comes from the fact that $\sigma(\Delta V_{T0})$ is defined for a pair of transistors.

The current-mode multiplier/divider [173] shown in Figure 8.11(a) is a simple example of a single translinear loop with identical transistors ($\lambda = 1$) of alternated orientations.

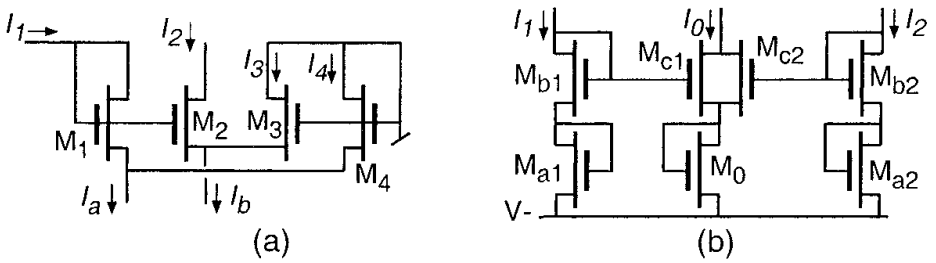

Fig. 8.11. Example of translinear circuits: (a) multiplier/divider; (b) vector length calculation.

The input currents are I_1, I_a and I_b whereas I_2 is the output. Using (8.36), we can write

$$I_1 I_3 = I_2 I_4 \quad \text{hence} \quad I_1(I_b - I_2) = I_2(I_a - I_1), \tag{8.38}$$

which gives after simplification

$$I_2 = I_1 I_b / I_a. \tag{8.39}$$

This result is valid as long as $I_1 < I_a$.

If cw and ccw transistors are *not* alternated, as in the loop example of Figure 8.10(b), then (8.33) cannot be rewritten as (8.35). The only way to make (8.34) compatible with (8.33) is to impose $V_{si} = 0$ by putting each transistor in a separate well connected to its source. Although each n_i is slightly dependent on the particular gate voltage, it can be approximated by a constant n multiplying U_T in (8.34).

In addition to requiring separate wells, non-alternated loops need a higher supply voltage, since they include stacks of at least two gate-to-source voltages.

An interesting example of a loop where transistors cannot be alternated is the circuit of Figure 8.11(b) that calculates the length of a vector in a N-dimensional space [174].

For this example with $N = 2$, the circuit contains two loops that share the transistor M_0. The corresponding current equations are

$$I_1^2 = I_{c1} I_0 \text{ and } I_2^2 = I_{c2} I_0, \text{ thus } I_0 = I_{c1} + I_{c2} = (I_1^2 + I_2^2)/I_0. \tag{8.40}$$

Hence, finally:

$$I_0 = \sqrt{I_1^2 + I_2^2}. \tag{8.41}$$

N loops are needed for N dimensions, each loop made of transistors M_{ai}, M_{bi} and M_{ci}, and all loops sharing M_0.

Although the basic translinear principle assumes that all the loop transistors are saturated, it is possible to include non-saturated transistors [175]. Consider M_k and M_{k-1} in the basic loop of Figure 8.10(a). They have the same gate voltage. Therefore, if they are connected in parallel, they represent the forward and reverse currents of a non-saturated transistor. This may help reducing the supply voltage needed by translinear circuits [176].