Bryan Hart explains the problems involved in using mosfets as analogue switches

Mosfet analogue gating

Operation of a mosfet in a logic gate is relatively straightforward. It normally presents few problems – even to those who have only a sketchy knowledge of mosfet theory. This is because the device can be modelled as a simple, though imperfect, on/off switch.

In the case of a basic cmos inverter stage, a device that is off passes only a small drain-to-source 'leakage' current – typically less than a microamp: for a device that is on, the magnitude of the drain-source voltage does not exceed a few millivolts – typically, 5.

In contrast, the behaviour of a mosfet in a 'transmission gate' is more complex and can cause difficulties. This is because the device can no longer be regarded as a simple switch in this application. The problem is made worse by the superficial treatment given to the topic in the literature.

This review adopts an evolutionary approach and uses a recently proposed mosfet model to clarify the operation of a cmos transmission gate.

Transmission gate characteristics

A 'transmission', 'linear' or 'analogue' gate is one in which a signal at an input terminal appears – ideally – unchanged in shape, at an output terminal during a time interval selected by a gating, or 'strobe', pulse applied to a control terminal. A change in polarity is acceptable and there may be a change in scale through attenuation or amplification.

The transmission gate has been used in a wide miscellany of signal processing applications that includes: pulse radar echo selection; nuclear instrumentation; switch for amplifier gain and polarity control, integrator resetting, digital-analogue converters and active switched-capacitor filters.

Figure 1a) shows a transmission gate connected in series with a load resistor R. The instantaneous values of the input, output and control signals are v_I, v_O, v_C respectively.

Figure **1b)** shows the ideal dc transfer characteristics for the two states of the control signal. Section i) of the illustration corresponds to the off state and ii) to the on state.

Characteristic ii) is shown, arbitrarily, as having a positive slope but it may be negative, indicating signal inversion. Also, the magnitude of the slope can be less than unity, in the case of attenuation, or greater than unity, in the case of amplification.

The general characteristics of a practical circuit are shown, similarly labelled, in Fig **1c)**. Characteristic ii) passes through a point V_I=0, V_O=V_P offset from the origin: furthermore the line is no longer straight.

Figure 2 summarises the operation of the gate circuit, for both ideal and non-ideal characteristics when the input signal, Fig. **2a)**, is a train of pulses of arbitrary shape and a control or gating waveform, **Fig. 2b)**, is applied. Figs **2c)** and **2d)** show the output waveforms for the characteristics of Figs 1b) and 1c) respectively.

In Fig. 2d), the output is distorted and sits on a 'platform' or 'pedestal' of height V_P. Additionally, there are gating transients or 'spikes', S_1, S_2 at the beginning and end of the gating pulse. There are two reasons why these appear. One is capacitive feedthrough from the gate terminal to the output, i.e., via the stray and inter-electrode capacitance C in Fig. 1a). The other is finite rise and fall times and non-coincidence of gating pulse edges in circuit schemes employing complementary gating pulses.

The transmission gate has been implemented using a variety of semiconductor devices. However, the cmos transmission gate is attractive in

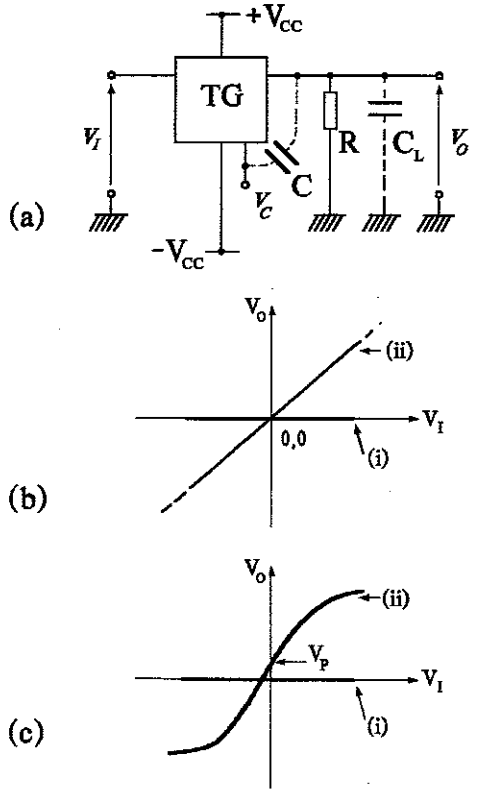

Fig. 1.
a) A series connected transmission gate.
b) Ideal transfer characteristics for a): i) when the transmission gate off, and ii), when the gate is on.
c) Practical counterpart of b).

a wide range of applications because it is easy to drive and is free of pedestal. To understand its origin and operation it is necessary to examine first the operation and limitations of a transmission gate using a single mosfet.

Single mosfet transmission gate

Figure 3 shows a single n-channel enhancement-mode mosfet used in the circuit of Fig. 1a). The electrode potentials are shown as dc quantities for the analysis that follows.

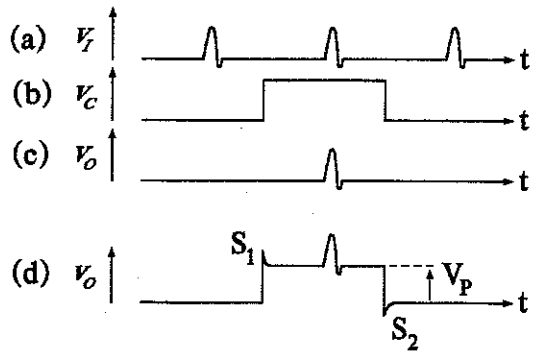

Fig. 2. Transmission-gate operation:
(a) Input pulse train.
(b) Gating pulse.
(c) Ideal output.
(d) Possible practical output. Note that V_P is exaggerated for clarity.

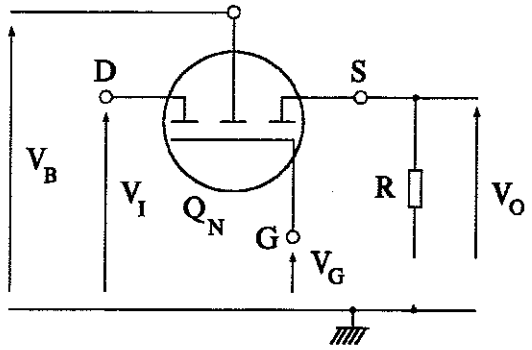

Fig. 3. Transmission gate using an n-channel enhancement-mode mosfet.

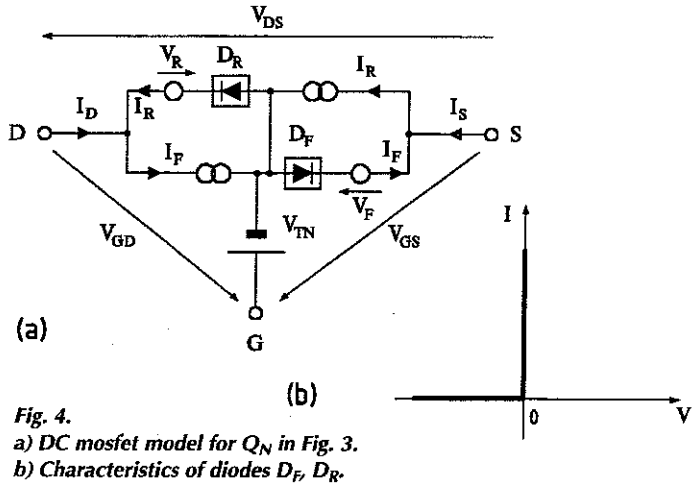

Fig. 4.
a) DC mosfet model for Q_N in Fig. 3.
b) Characteristics of diodes D_F, D_R.

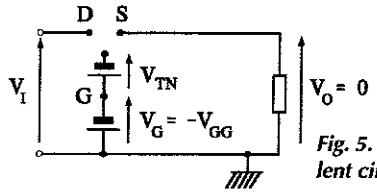

Fig. 5. Transmission-gate equivalent circuit for off state.

Substrate *B* is biased so that, under all conditions, V_B is less than V_I and V_O. In that case the source-substrate and drain-substrate junctions of the mosfet are reverse-biased and the mosfet can be represented by the recently proposed large-signal model[1] in **Fig. 4a)**. This applies also for alternating signals if inter-electrode capacitances are added.

Battery V_{TN} models the threshold voltage. Diodes D_F and D_R are shown boxed[2] because they have the ideal *I/V* characteristics of Fig. 4b).

The device equations for $V_F>0$ ($V_{GS}>V_{TN}$) and $V_R>0$, ($V_{GD}>V_{TN}$) are,

$$I_F=K_N V_F^2=K_N(V_{GS}-V_{TN})^2 \quad (1)$$

$$I_R=K_N V_R^2=K_N(V_{GD}-V_{TN})^2 \quad (2)$$

Since $I_G = 0$,

$$I_D=-I_S=I_F-I_R=K_N(V_F^2-V_R^2) \quad (3)$$

Device design parameters K_N, V_{TN} can be found experimentally from a plot of $\sqrt{I_D}$ versus V_{GS} with V_{DG} at 0V. The slope of the straight line section gives $\sqrt{K_N}$ and the extrapolated intercept on the V_{GS} axis gives V_{TN}.

The model is attractive in showing clearly the device dc operating mode when arbitrary terminal voltages are applied. Once that mode has been established the model can be reduced to a simpler equivalent form, as will become apparent in what follows.

If, in Fig. 3, V_G is connected to a negative bias source $-V_{GG}$, and if V_I is greater than $-(V_{GG}+V_{TN})$ then D_F and D_R in the model of Fig. 4a) are cut off and $I_F=I_R=I_D=0$.

The transmission gate in Fig. 3 is cut off, its idealised equivalent circuit

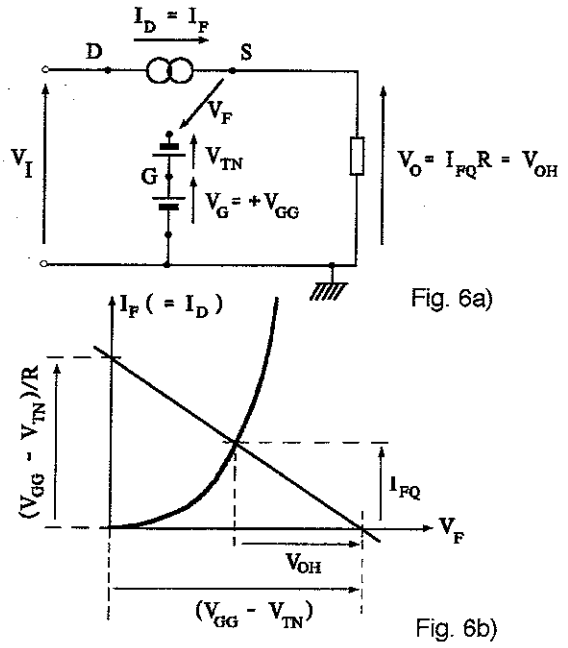

Fig. 6. a) Transmission-gate equivalent circuit for when the gate is on and $V_I>(V_{GG}-V_{TN})$.
b) Load line construction for operating current, I_{FQ}.

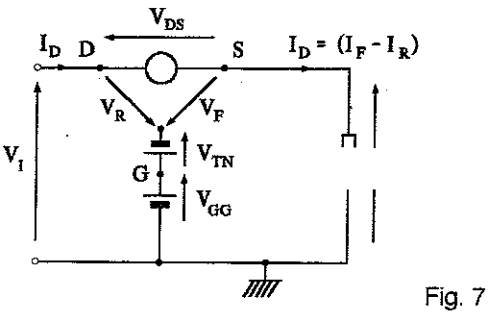

Fig. 7. Transmission-gate equivalent circuit for when the gate is on and $V_I<(V_{GG}-V_{TN})$

eing indicated in **Fig. 5**. In a refinement to the mosfet model under discussion, two diodes arranged in series-opposition can be connected between the terminals D and S to allow for the small cut-off leakage current that occurs with real devices.

Consider next what happens in Fig. 3 when V_G is $+V_{GG}$. Using the model of Fig. 4a) it is clear that D_F conducts while V_{GG} is greater than V_{TN}. But there are two cases to consider for D_R. One is that D_R is off when V_I is greater than $V_{GG}-V_{TN}$ and the other is that it is on while V_I is less than $V_{GG}-V_{TN}$.

The equivalent transmission gate circuit for the first case is shown in **Fig. 6a)**. The mosfet behaves as a voltage controlled current generator characterised by equation 1.

The gate-source loop gives,

$$V_O=I_FR=V_{GG}-V_{TN}-V_F \tag{4}$$

Operating current I_{FQ} and V_{OH}, which is equivalent to $I_{FQ}R$, can be calculated using eqns 1 and 4, but the graphical construction of Fig. **6b)** is more instructive.

Equation 4 describes a load line on the plot of I_F, which equals I_D versus V_F. This passes through the points $I_F=0$, $V_F=V_{GG}-V_{TN}$) and $I_F=(V_{GG}-V_{TN})/R$, $V_F=0$. Its intersection with the transfer characteristic of the mosfet gives I_{FQ} and V_{OH}.

Voltage V_{OH} increases with R up to a maximum value os $V_{GG}-V_{TN}$.

When V_{GG} is greater than V_{TN} and V_I is less than $V_{GG}-V_{TN}$, the equivalent circuit is as shown in Fig. 7: diodes D_F and D_R in the model of Fig. 4a) are both on, the device operates in the 'ohmic' (or 'triode') region and equation 3 applies.

Hence,

$$V_O=I_DR=(I_F-I_R)R \tag{5}$$

or,

$$V_O=K_NR(V_F^2-V_R^2)=K_NR(V_F-V_R)(V_F+V_R) \tag{6}$$

But,

$$V_F-V_R=V_{DS} \tag{7}$$

So,

$$V_O=K_NRV_{DS}\,(2V_F-V_{DS}) \tag{8}$$

Substituting $V_{GG}-V_{TN}-V_O$ for V_F in equation 8 gives,

$$V_{DS}^2-2V_{DS}\left(V_{GG}-V_{TN}-V_O\right)+\left(\frac{V_O}{K_NR}\right)=0 \tag{9}$$

For a chosen value of V_O, you can calculate V_{DS} by solving this quadratic and selecting the root with the smaller magnitude. The other root does not satisfy the condition $V_R>0$. So for each value of V_O, you can find a value of V_I, since $V_I=V_O+V_{DS}$. This procedure allows you to make a plot of the transfer characteristic of the circuit.

Even without specific numerical values for V_{GG}, V_{TN}, K_N you can still make two useful statements concerning equation 9. First, if V_O is 0 then V_{DS} is also 0, because the other root $V_{DS}=2(V_{GG}-i_{TN})$ is inadmissible. So $V_I=0$ and all the transfer characteristics pass through the origin – irrespective of the value of R.

Secondly, if R approaches infinity then V_{DS} approaches 0 because $V_O/K_NR\rightarrow 0$. This means the transfer characteristic approximates to a straight line. **Figure 8** shows the characteristic for this condition and for a finite value of R.

Comparing the practical characteristics with the ideal ones in Fig. 1b) it is apparent that the circuit can handle bipolar input signals but there are deficiencies in two respects, range and linearity.

For a given mosfet type, the range can be extended by making V_{GG} larger but there are limits set by gate voltage breakdown and the magnitude of the supply rail voltage.

Linearity can be improved by having a large value of R but there is a limit set by dynamic considerations. When the transmission gate is switched off, the output terminal has associated with it a time constant $(C+C_L)R$, C_L being the load capacitance. If R is very large the circuit takes

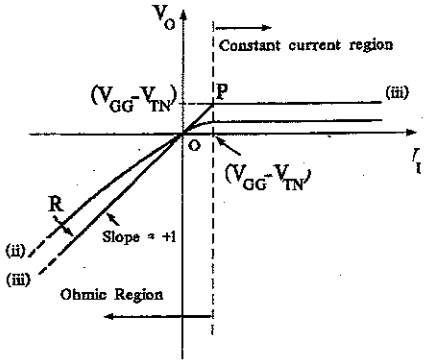

Fig. 8. Transfer characteristics for the circuit of Fig. 3.
i) Gate off
ii) Gate on and R finite
iii) Gate on and R approaching infinity

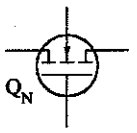

Fig. 9. Alternative symbol for Q_N in Fig. 3.

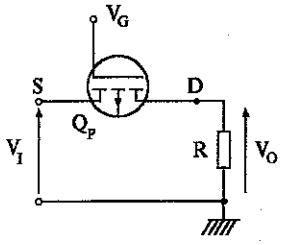

Fig. 10. Transmission gate using a p-channel enhancement-mode mosfet.

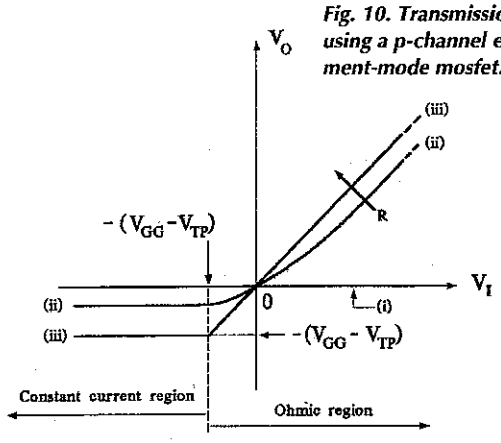

Fig. 11. Transfer characteristics for the circuit of Fig. 10. Compare with Fig. 8, to which the labelling of i), ii), iii) applies.

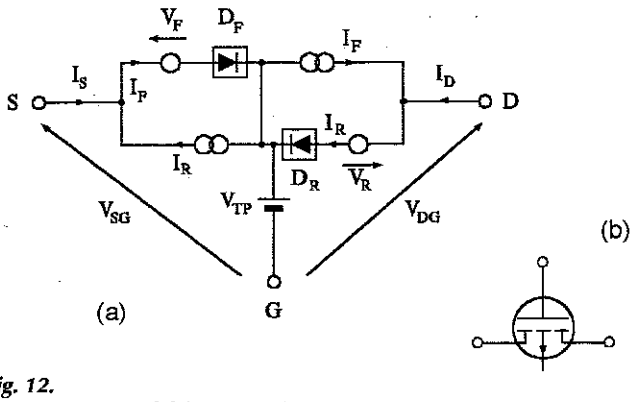

Fig. 12.
a) DC mosfet model for Q_P in Fig. 10.
b) Alternative symbol for Q_P.

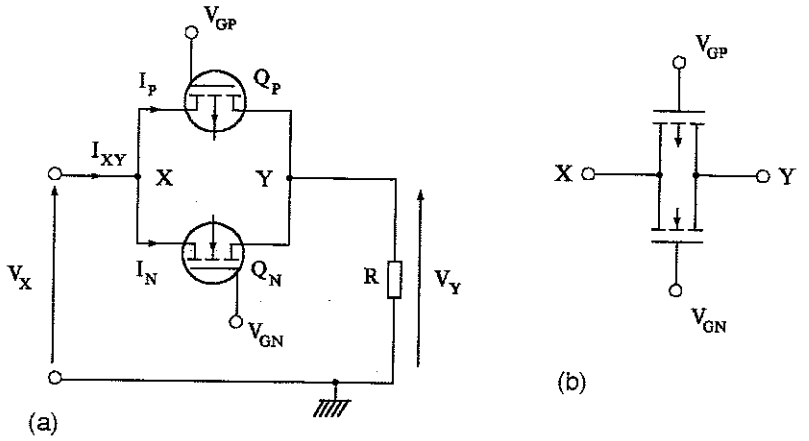

Fig. 13.
a) A complementary mosfet transmission gate.
b) Alternative representation of complementary connection.

a long time to reach the state where V_O is 0.

I will now digress briefly to clarify a point purposely glossed over at the start of this section. It concerns the naming of parts and symbolic representation of the mosfet. The connection of the drain terminal to the input and the source to the output of the circuit in Fig. 3 was an *arbitrary* choice.

Reversing the connections produces exactly the same characteristics as those in Fig. 8. The reason for this is the symmetry in the electrical characteristics of a mosfet. The channel region is uniformly doped in the direction of current flow along it. Also, the device structure is geometrically symmetrical about the plane drawn perpendicular to the semiconductor surface at a point mid-way along the channel. The symmetry is emphasised in the form of the mosfet model.

It is helpful to distinguish between the name given to a terminal and the function performed at that terminal, because a terminal designated 'source' can also function as a 'drain' and *vice-versa*. There is a loose analogy here, in mechanics. What we might call the 'business' end of a double-ended symmetrical spanner is the one being used to tighten or loosen a nut.

When discussing standard circuit configurations, for example a common-source stage and a source-follower, it is preferable to use the symbol shown in Fig. 3. Here, the gate connection is located near what is regarded as the source end of the channel because that aids in the identification of the intended function of the circuit.

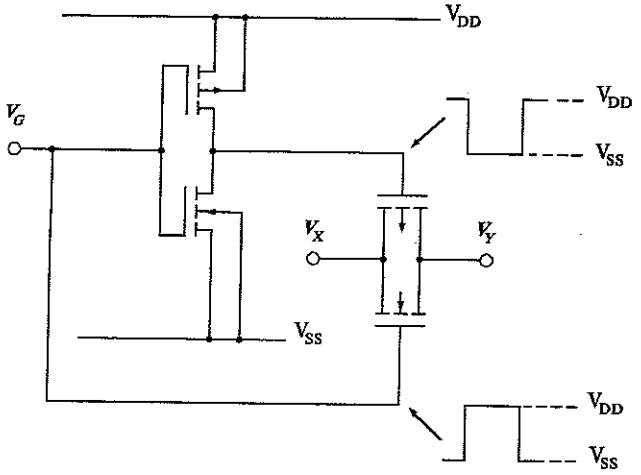

Fig. 15. Data sheet representation of a cmos transmission gate.

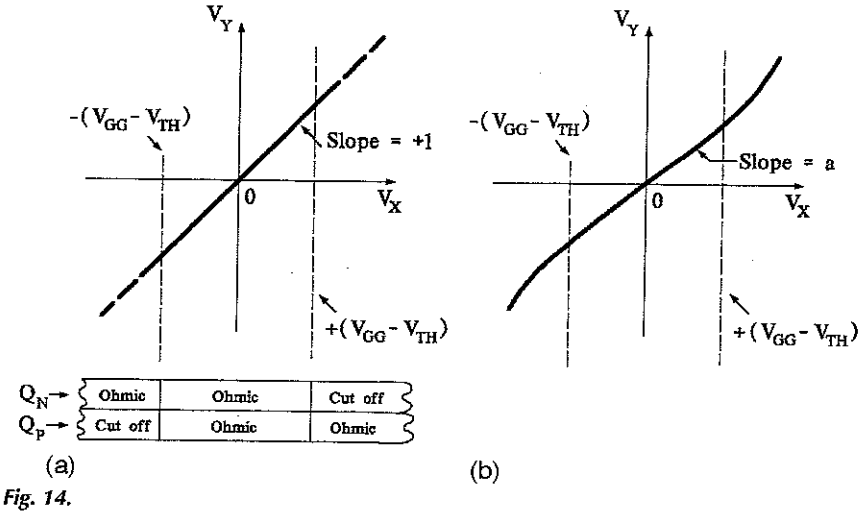

Fig. 14.
a) On characteristic for Fig. 13(a) for R infinite , showing conduction states.
b) On characteristic for finite R. Non-linearity is exaggerated for clarity.

However, when neither of the mosfet terminals in the path of signal flow is at a fixed potential, and when there is no preferred direction of current flow through the device, the symmetrical nature of mosfet operation is indicated better by using the standard alternative symbol in Fig. 9. Here, the gate terminal is located centrally.

If a p-channel enhancement-mode mosfet, Q_P, is used instead of Q_N, as shown in **Fig. 10**, then the transmission gate is on for V_G=$-V_{GG}$ and off for V_G=$+V_{GG}$. The resulting transfer characteristics are given in **Fig 11**. These are obtained, by reasoning similar to that already given for a general mosfet, Q_N using the p-channel model of **Fig. 12a)**. Since no new principle is involved, that analysis is not given here.

As with Q_N, the source and drain terminals of Q_P can be interchanged and to reflect this the alternative standard symbol is that in Fig **12b)**.

A complementary mosfet TG

The shapes of the plots in Figs 8 and 11 suggest that the transfer characteristic of a mosfet transmission gate, when on, might be improved in range and linearity by using two complementary devices in parallel, as shown in **Fig. 13a)** or the alternative representation in Fig. **13b)**. The terminals in the signal path, arbitrarily labelled X and Y, refer to the circuit nodes rather than the device terminal designations.

The transmission gate is off for the conditions V_{GN}= $-V_{GG}$ with V_{GP}=$+V_{GG}$ and on for V_{GN}=$+V_{GG}$ with V_{GP}=$-V_{GG}$. The following discussion refers to the on state. Of special interest are the theoretical conditions K_P=K_N=K and V_{TP}=V_{TN}=V_{TH}.

For a specified R, in order to plot a transfer characteristic for Fig. 13a), V_X can be calculated for assumed values of V_Y following the procedure

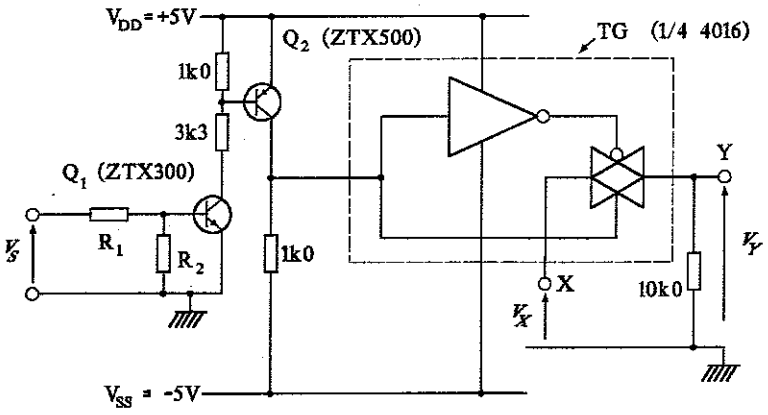

Fig. 16. Driving the cmos transmission gate from ttl-compatible circuitry.

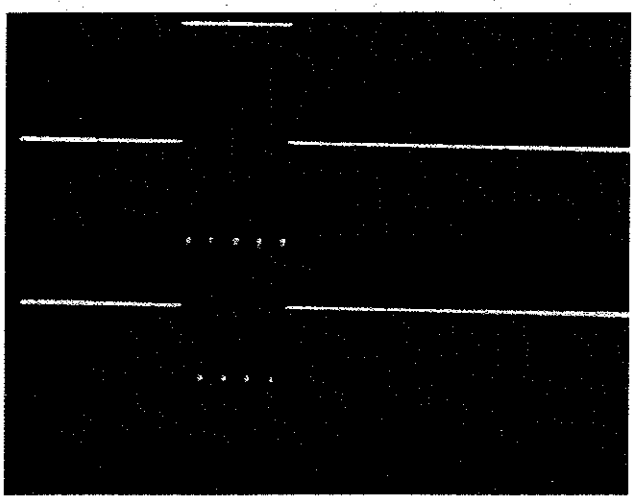

Fig. 17. Oscilloscope display for circuit of Fig. 16. The traces have common time and voltage scales.
Upper trace: v_S, with amplitude 5V and duration 100μs. Lower trace: v_Y, for a sinusoidal input.

described earlier for a single n-channel mosfet. The device equations give I_N and I_P as functions of V_X and V_Y. You can now use the relationship $I_P+I_N=I_{XY}=V_Y/R$.

For *R infinite*, the on characteristic is the straight line in **Fig. 14a**). The conduction states of Q_N and Q_P are determined by using the models of Figs 4 and 12 in the composite connection.

By inspection, Q_N and Q_P both operate in the ohmic region for,

$$(V_{GG}-V_{TH})\geq V_X\geq-(V_{GG}-V_{TH})$$

Mosfet Q_N is cut off and Q_P is ohmic while V_X is greater than $V_{GG}-V_{TH}$). Similarly, Q_P is cut off and Q_N is ohmic when V_X is less than $-(V_{GG}-V_{TH})$.

The theoretical characteristic for a finite value of R has the general shape shown in Fig. **14b)**. Referring to Fig. 13, it can be shown that for, $+(V_{GG}-V_{TH})\geq V_X\geq-(V_{GG}-V_{TH})$,

$$I_{XY}=4KV_{XY}(V_{GG}-V_{TH}) \quad (10)$$

or,

$$\frac{V_{XY}}{I_{XY}}=R_{ON}=\frac{1}{4}K(V_{GG}-V_{TH}) \quad (11)$$

Resistance R_{ON} is both the dc and the incremental resistance over the V_X range considered. The transfer characteristic of Fig 14b) is a straight line of slope 'a', where 'a' is a linear attenuation factor given by,

$$a=\frac{R}{R+R_{ON}} \quad (12)$$

Outside the V_X range considered the transfer characteristic ceases to be linear because R_{ON} is a function of V_X. This is to be expected since one of the devices cuts off.

Whether the composite arrangement of Fig. 13 is implemented in practice using discrete devices or, as is usually the case, using a monolithic cmos structure, there are at least two reasons why the transfer characteristic is non-linear over the whole range of V_X.

One reason is the impossibility of meeting, precisely, the conditions $K_P=K_N$ and $V_{TP}=V_{TH}$. The other is the variation of V_{TP}, V_{TN} with V_X via their dependencies on substrate bias – an effect so far ignored. In a cmos scheme, the second effect is usually minimised by using additional mosfets. These are not always shown on schematic diagrams.

Non-linearity in the transfer characteristic can be assessed from manufacturers' data, which specifies the variation in R_{ON} over a given range of V_X for a given R.

If R is much greater than R_{ON} and R_{ON} changes by $x\%$, where x is less than 10, then it follows from equation 12 that 'a' varies by $x\%$ also.

To conclude this review I will present a practical gate.

The IC cmos gate in action

Figure 15 is a schematic circuit of an IC cmos transmission gate, as shown on a data sheet. The mosfet symbols in the inverter section, which supplies a complementary gating waveform, are different from those used in the signal-path section.

Figure 16 shows part of a circuit for generating 'flashing-ring' symbols for use in oscilloscope tests of human visual perception. The dotted box contains the circuitry of Fig 15: the parallel mosfets are now represented by a symbol comprising two interlocked oppositely directed arrowheads. This highlights the bilateral nature of this transmission gate. The input can be at X, as here, and the output at Y, or *vice-versa*.

Resistance R_{ON} is typically 250Ω for a 4016 type cmos transmission gate operating between V_{DD} = + 5 V, V_{SS} = -5 V. As a result, the choice 10kΩ for R meets the condition that R should be much greater than R_{ON}.

In this case, the input v_X is a sinusoidal signal. The ttl-compatible drive circuit, comprising Q_1, Q_2 and associated resistors, supplies a gating waveform with the required dc levels. This can be modified for faster operation by connecting a Schottky diode between the collector and base of each transistor, to minimise carrier storage time, and by adding an active pull-down circuit to the collector of Q_2.

Figure 17 shows an oscilloscope display of the time-selection pulse, v_S, applied to the base of Q_1 and the output waveform, v_Y, which is a selected portion of the input v_X. ■

References

1. Hart BL, 'First-order d.c. model of the mosfet', *IJEEE* Vol. 34, pp. 326-330, Oct 1997.
2. Hart BL, 'Introduction to Analogue Electronics', p. 7, pub Arnold 1997