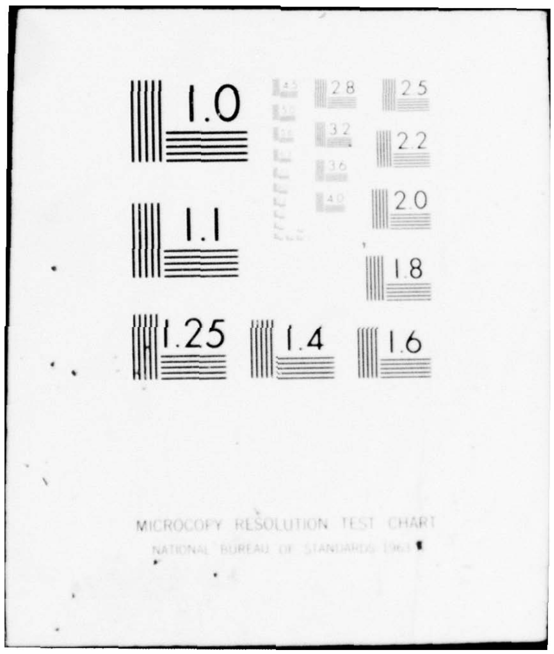
AD-A036 362 ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/G 9/1
MOS TRANSISTOR DRIVERS FOR LARGE CAPACITIVE LOADS.(U)
OCT 76 P F SMITH DAAB07-72-C-0259
UNCLASSIFIED R-748 NL

1 OF 1
AD
A036362

END
DATE
FILMED
8-77


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

REPORT R-748 OCTOBER, 1976 UILU-ENG 76-2236

COORDINATED SCIENCE LABORATORY

MOS TRANSISTOR DRIVERS FOR LARGE CAPACITIVE LOADS

PAUL FIELDING SMITH

UNIVERSITY OF ILLINOIS – URBANA, ILLINOIS

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO. (9)

3. RECIPIENT'S CATALOG NUMBER
Master's Thesis,

4. TITLE (and Subtitle)
(6) MOS TRANSISTOR DRIVERS FOR LARGE CAPACITIVE LOADS.

5. TYPE OF REPORT & PERIOD COVERED
Paul Fielding Smith

6. PERFORMING ORG. REPORT NUMBER
(14) R-748, UILU-ENG-76-2236

7. AUTHOR(s)
(10) Paul Fielding Smith

8. CONTRACT OR GRANT NUMBER(s)
(15) DAAB-07-72-C-0259

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS (11)

12. REPORT DATE
Oct[redacted] 1976

13. NUMBER OF PAGES
53

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)
(12) 60p.

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

MOS Transistor Drivers
Driving Capacitance in LSI
Computer Simulation of MOS Drivers

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The advent of MOS large-scale integration (LSI) has focused considerable research activity into studying the compromises involved in efficient integrated circuit design. These trade-offs may involve circuit size, device parameters, circuit speed, and extra processing steps. Available chip size, manufacturing tolerance, and cost place upper and lower limits on the range each of these can take. There may also be some concern about requiring several voltage sources or complicated clocking sequences which cannot be derived on the chip.

UILU-ENG 76-2236

MOS TRANSISTOR DRIVERS FOR LARGE CAPACITIVE LOADS

by

Paul Fielding Smith

This work was supported in part by the Joint Services Electronics Program (U.S. Army, U.S. Navy and U.S. Air Force) under Contract DAAB-07-72-C-0259.

MOS TRANSISTOR DRIVERS FOR LARGE CAPACITIVE LOADS

BY

PAUL FIELDING SMITH

B.S., University of Kansas, 1975

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1976

Thesis Advisor: Professor Timothy N. Trick

Urbana, Illinois

ACKNOWLEDGMENT

The author is grateful to his advisor Professor Timothy N. Trick for his guidance, suggestions, and efforts during the research and preparation of this thesis.

The author would also like to express his appreciation to his wife for her patience and support.

TABLE OF CONTENTS

LIST OF FIGURES

Figure | Page

1. INTRODUCTION

The advent of MOS large-scale integration (LSI) has focused considerable research activity into studying the compromises involved in efficient integrated circuit design. These trade-offs may involve circuit size, device parameters, circuit speed, and extra processing steps. Available chip size, manufacturing tolerance, and cost place upper and lower limits on the range each of these can take. There may also be some concern about requiring several voltage sources or complicated clocking sequences which cannot be derived on the chip.

At the device level the most expensive parameter in LSI fabrication is the area. Because of this, manufacturing technology has concentrated its efforts into making small devices. Circuit speed is then determined by the rate at which these small devices can drive their capacitive loads (e.g. the gate of another MOS transistor). It is advantageous to study new circuit configurations which allow small MOS devices to drive large capacitive loads in a short time interval. These circuits may then be used as clock drivers, buffer amplifiers, or perhaps just as standard circuit elements.

This thesis examines two circuits suitable for driving large capacitive loads. These circuits are shown in Figure 1.1. The pull-up circuit is currently in use and promises high speed operation when used with clocks of several phases. The source-follower circuit, however, utilizes less space since the large pull-up capacitor C_p does not need to be integrated. In the following chapters a description of the model used to stimulate these circuits is given along with derivations of pertinent

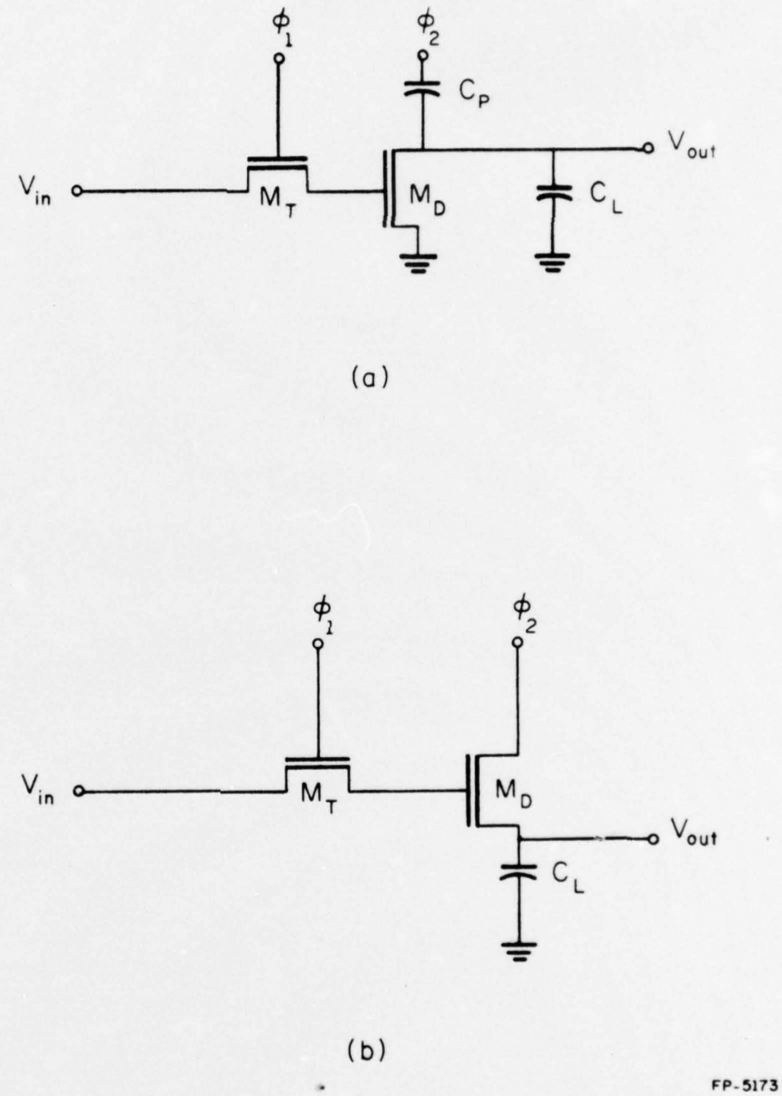


Figure 1.1. a. Pull-up circuit
b. Source-follower circuit

device equations. A discussion of the pull-up and source-follower circuits follow. Included in these discussions are a circuit analysis, formulation of design equations, computer simulation, and a discussion of the results obtained. The results of the two circuits are then compared and summarized.

2. MODELS

Figure 2.1 shows the schematic diagram of a MOS transistor with the more important associated capacitors. All other parasitic capacitors are neglected since they contribute approximately only 10% to the total capacitance [1]. C_{DB} and C_{SB} are the junction capacitances associated with the drain-body and source-body diodes. C_{GD} and C_{GS} consist of gate overlap capacitance and the intrinsic gate MOS capacitance. C_{GB} is the gate-body (or gate-substrate) MOS capacitance.

In order to increase the speed of MOS circuits without increasing the size of the devices, the gate capacitance loads must be made smaller. This is currently accomplished by using self-aligning gate technology. This technique greatly reduces the gate overlap and other constant capacitances leaving the nonlinear intrinsic MOS gate capacitances dominant. The variation of C_{GS}, C_{GD}, and C_{GB} as a function of V_{GS} for a given V_{DS} is shown in Figure 2.2. Modeling these capacitances as constant capacitors may provide results which are not at all in close agreement with experimental results. It has been shown that dynamic and bootstrap circuits generally require a nonlinear capacitor model while static circuit simulation can usually get accurate results with constant capacitors [2].

SPICE 2 was the circuit simulation program used for all simulations. This is a nonlinear analysis program which uses the FET model of Shichman and Hodges [3]. This model, shown in Figure 2.3, includes linear ohmic contacts at the drain and source, diode nonlinearity, junction capacitance nonlinearity (varies as the -½ power of the junction voltage), finite output

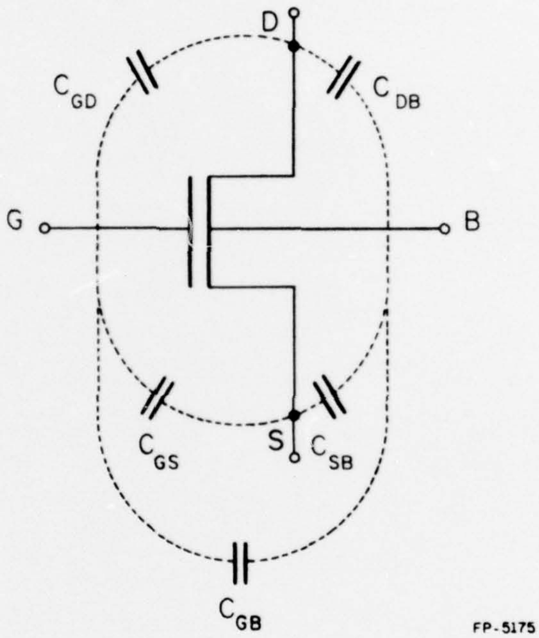


Figure 2.1. Schematic of a MOS transistor with dominant capacitors

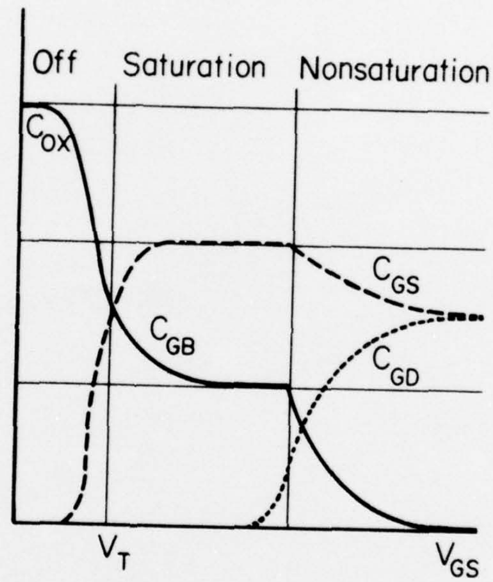


Figure 2.2. Nonlinearity of C_{GB}, C_{GS}, and C_{GD}

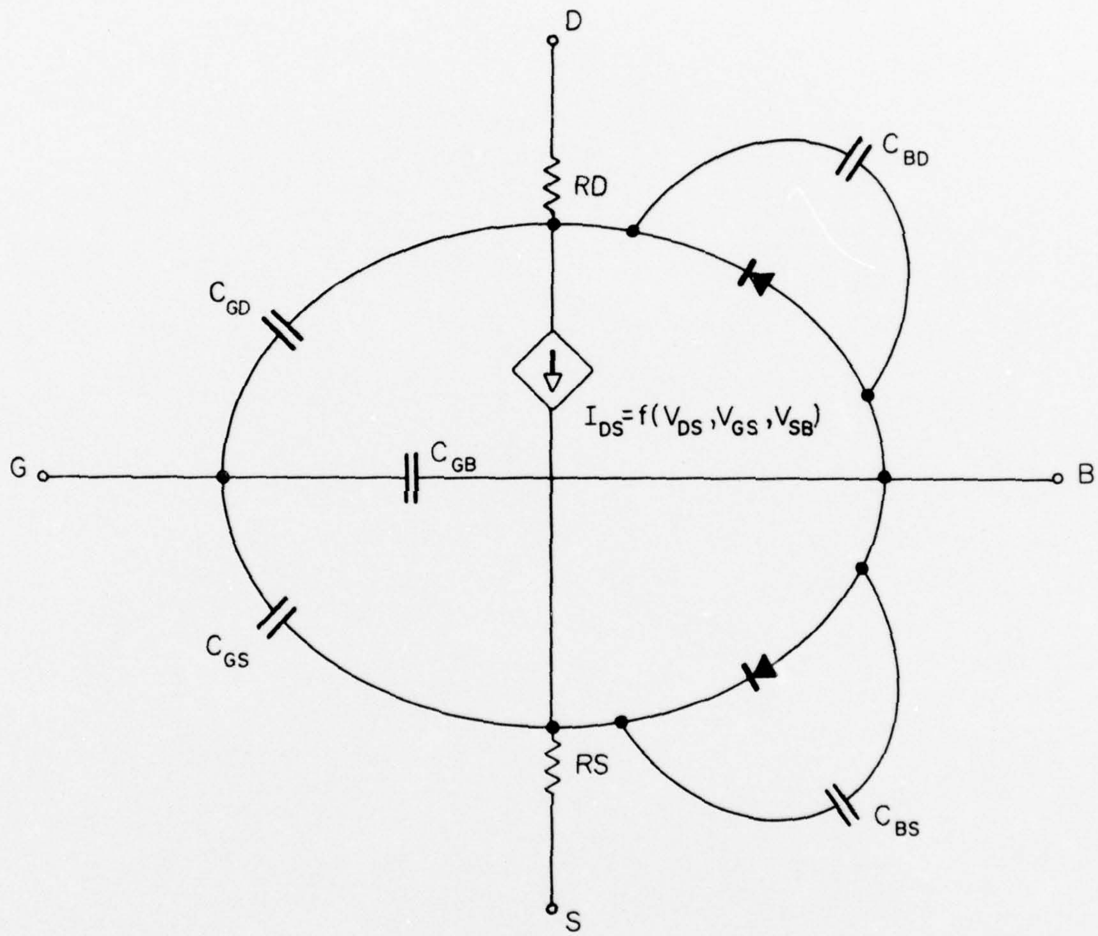


Figure 2.3. Model of a MOS transistor

conductance (channel-length modulation), body effect (variation of threshold voltage with a source-body bias), and constant capacitors C_{GS}, C_{GD}, and C_{GB}. Since SPICE 2 uses a constant gate capacitor model, the results obtained must be interpreted in light of the model inaccuracy described above. The variation of drain-source current with respect to gate-drain voltage (V_{GD}) and drain-source voltage (V_{DS}) is illustrated in Figure 2.4. The device is assumed to be symmetrical, the first and third quadrants being images of one another. Also shown are the different regions in which the device can operate. This becomes important when considering the nonlinear gate capacitances discussed previously.

Simplified equations governing the variation of I_{DS} with V_{GS}, V_{DS}, and several device parameters are given by Equation (2.1) [4].

Nonsaturation region:

$$I_{DS} = \frac{\mu_n \epsilon_{ox}}{2\,t_{ox}}\left(\frac{W}{\ell}\right)[2(V_{GS}-V_T) - V_{DS})](1+\lambda V_{DS})$$

$$= \beta[2(V_{GS}-V_T) - V_{DS})](1+\lambda V_{DS}). \qquad (2.1a)$$

Saturation region:

$$I_{DS} = \frac{\mu_n \epsilon_{ox}}{2\,t_{ox}}\left(\frac{W}{\ell}\right)(V_{GS}-V_T)^2(1+\lambda V_{DS})$$

$$= \beta(V_{GS}-V_T)^2(1+\lambda V_{DS}) \qquad (2.1b)$$

where

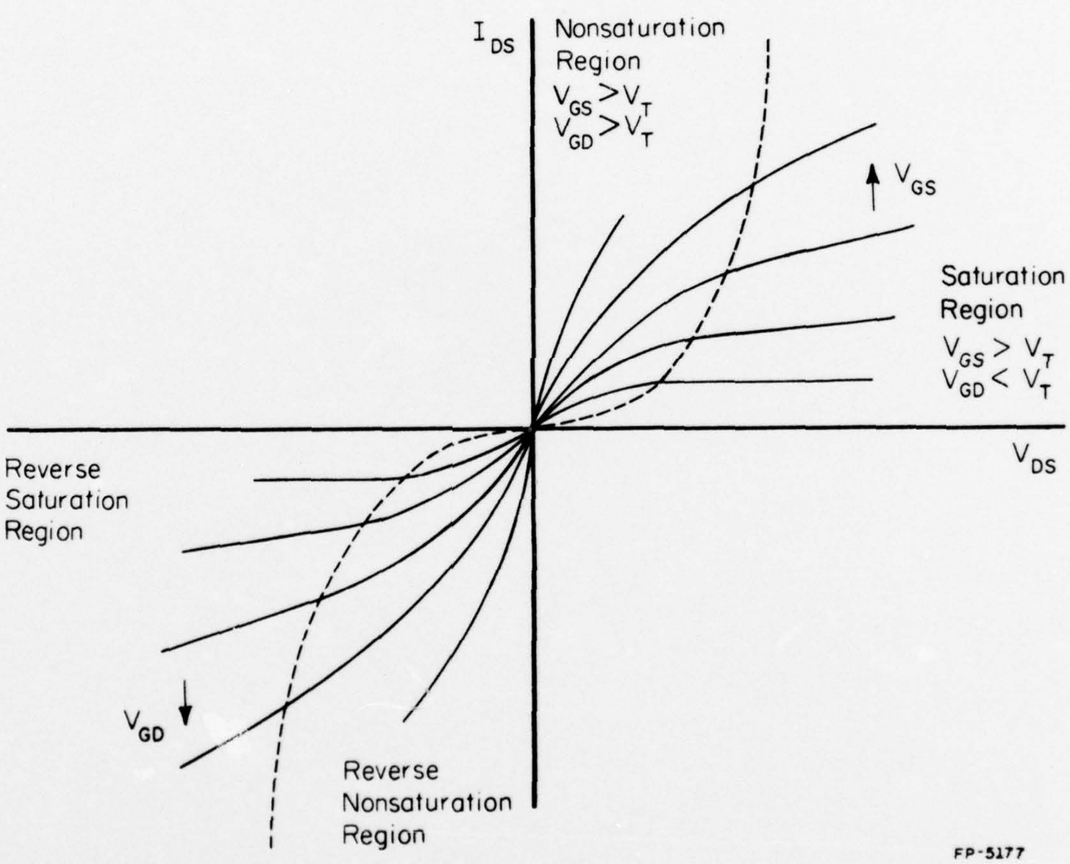


Figure 2.4. Drain-source characteristics of a MOS transistor

μ_n = average minority carrier surface mobility in channel

t_{ox} = thickness of the oxide over channel

ϵ_{ox} = permittivity of the oxide

ℓ = channel length in the direction of current flow

W = width of channel

V_T = threshold voltage

$\beta = \frac{\mu_n \epsilon_{ox}}{2\, t_{ox}} \left(\frac{W}{\ell}\right)$

λ = channel-length modulation parameter (2.2)

Since device geometries can be precisely controlled during fabrication, it is important that current through a device is proportional to the device geometry $\left(\frac{W}{\ell}\right)$, as shown in Equation (2.1). The diode nonlinearity is given by

$$I_D = I_S\left(e^{\frac{qV_D}{kT}} - 1\right)$$

where I_S = bulk junction saturation current

$\frac{kT}{q} \simeq .026$ v at room temperature.

The variation of V_T due to the body effect is given by Equation (2.3).

$$V_T = V_{TO} + \gamma\left(\sqrt{\phi + V_{SB}} - \sqrt{\phi}\right) \qquad (2.3)$$

where V_{TO} = threshold voltage at $V_{SB} = 0$

γ = bulk threshold parameter

ϕ = surface potential.

If the device is operated in its reverse characteristics, then in the above equations V_{GD} replaces V_{GS}, V_{DB} replaces V_{SB}, V_{SD} replaces V_{DS}, and I_{SD} replaces I_{DS}.

The nonlinear junction capacitances are governed by

$$C_{DB} = C_{BDO}\left(1 - \frac{V_{BD}}{\phi_B}\right)^{-\frac{1}{2}}$$

$$C_{SB} = C_{BSO}\left(1 - \frac{V_{BS}}{\phi_B}\right)^{-\frac{1}{2}}$$

where C_{BDO} = zero bias drain-body junction capacitance

C_{BSO} = zero bias source-body junction capacitance

ϕ_B = bulk junction potential.

Device parameters vary with different processing techniques and materials. Typical device parameters are

$$\phi = .7 \text{ v}$$

$$\gamma = 1.0$$

$$\lambda = .1$$

$$\phi_B = .7 \text{ v}$$

$$I_S = 1.0 \times 10^{-14} \text{ A}$$

$$RD = 100 \text{ ohms}$$

$$RS = 100 \text{ ohms.}$$

Typical processing parameters for doping levels between 10^{15} and 10^{16} cm^{-3} are

$$\mu_n = 500 \text{ cm}^2/\text{V-sec}$$

$$t_{ox} = 1000 \text{ Å} = 10^{-5} \text{ cm}$$

$$\epsilon_{ox} = \epsilon_r \epsilon_o = (3.9)(8.85 \times 10^{-14}) = .345 \text{ pF/cm}$$

$$\ell_{min} = W_{min} = .2 \text{ mil} = .508 \times 10^{-3} \text{ cm}$$

Using these values in Equation (2.2),

$$\beta = 8.63 \times 10^{-6} \frac{W}{\ell} \text{ A/v}^2. \tag{2.4}$$

C_{OX} (total gate capacitance per gate area) is given by

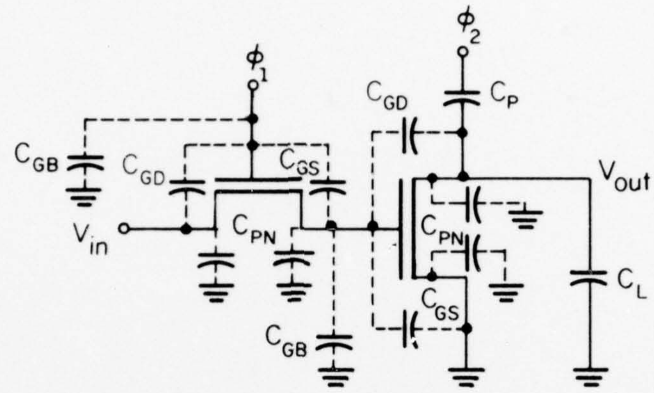
$$C_{OX} = \frac{\epsilon_{OX}}{t_{OX}} = 34.5 \text{ nF/cm}^2 = .222 \text{ pF/mil}^2. \tag{2.5}$$

The junction capacitance per junction area can vary anywhere from 9 nF/cm^2 to 30 nF/cm^2 depending on the doping density (10^{15} to 10^{16} cm^{-3}). Since the junction area (i.e. the drain or source area) is typically larger than the gate, C_{DBO} and C_{SBO} were estimated to be equal to the total gate capacitance.

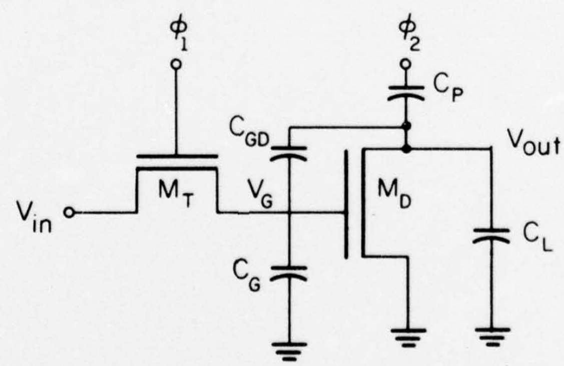
Relationships between ℓ, W, β, and the total gate capacitance ($C_{OX} \times W\ell$) may be obtained by using Equations (2.4) and (2.5). It will be important to note that the gate capacitance is minimum when $\ell = W = .2$ mil. For this case $\beta = 8.63 \times 10^{-6}$ A/v^2 and the gate capacitance is .009 pF.

3. PULL-UP CIRCUIT

A typical pull-up circuit[6] for driving large capacitive loads is shown in Figure 3.1a. The body of each device is connected to ground. In order to obtain a basic understanding of the circuit behavior, first the circuit is analyzed using very simple models. This analysis will be followed by an analysis using the more complete model described in Chapter 2 and computer simulation. The following assumptions are made for the simplified circuit analysis: the transmission gate transistor M_T is small compared to the driver transistor M_D, $C_L > C_{OX}$, $C_P > C_{OX}$, and all voltage sources are ideal with infinite rise times. The variation of the threshold voltage due to the body effect is approximated by $\Delta V_T = \frac{1}{2}\sqrt{V_{SB}}$. Because of the first assumption, all of the capacitors of M_T can be neglected since they are much smaller than those of M_D. The second assumption allows the drain junction capacitor of M_D to be neglected since it is in parallel with C_L. The other driver junction capacitor can also be neglected since it is shorted to ground. C_{GS} and C_{GB} can then be lumped together and called C_G. The approxomate circuit is shown in Figure 3.1b. A typical clocking sequence and corresponding output are shown in Figure 3.2. When clock ϕ_1 is high (time interval A and C), the data is transferred to the driver device M_D and either charges or discharges the 'memory' capacitor C_G. When ϕ_2 is high (time interval B or D), the data output is enabled and can be taken from the drain of M_D. This circuit is an inverting circuit, a logic one input ($V_{IN} > V_T$) produces a logic zero output ($V_{OUT} < V_T$). A characteristic of this circuit is that the logic zero is not produced until sometime after it has been enabled.



(a)



(b)

FP-5178

Figure 3.1. a. Pull-up circuit with all capacitors included
b. Approximate circuit to be analyzed

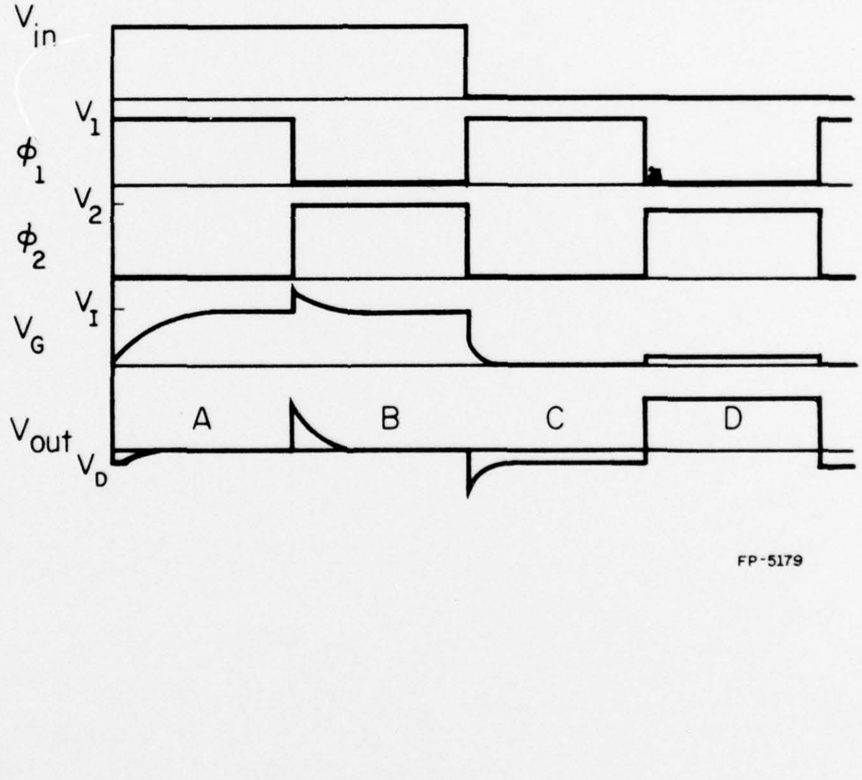


Figure 3.2. Typical clocking sequence and output

A. Circuit Analysis

During time interval A, V_{IN} and ϕ_1 are both high. M_T is turned on (i.e. conducting) if $V_{GS} = V_1 - V_G > V_T$ and is operating either in the saturation or nonsaturation region depending on whether $V_{GD} = V_1 - V_{IN}$ is less than or greater than V_T. M_T charges essentially $C_G + C_{GD}$ since C_P and C_L are very large in comparison. When $V_G = V_T$ M_D turns on discharging C_L (there may be a small negative voltage V_D left on C_L from previous circuit operation). An equivalent model for the circuit during this time interval is given in Figure 3.3. The current source of M_D has the effect of keeping $C_P + C_L$ discharged to ground. During this interval V_G will charge up to V_I where

$$V_I = \begin{cases} V_{IN} & , \text{ when } V_1 - (V_T + \tfrac{1}{2}\sqrt{V_{IN}}) \geq V_{IN} \\ V_1 - (V_T + \tfrac{1}{2}\sqrt{V_I}), & \text{otherwise} \end{cases} \tag{3.1}$$

where the $\frac{1}{2}\sqrt{V}$ term is due to the body effect.

Between intervals A and B clock ϕ_1 goes to zero which turns off M_T as long as V_T, V_{IN}, and V_G are all greater than zero. The clock ϕ_2 goes high in this interval. The jump of ϕ_2 is coupled to the gate and drain of M_D through capacitive voltage dividers. The change in the gate and drain voltages are given by

$$\Delta V_{OUT} \sim V_2 (C_P/(C_P + C_L)) \tag{3.2a}$$

neglecting the small effects of C_G and C_{GD}, and

$$\Delta V_G \sim \Delta V_{OUT}(C_{GD}/(C_{GD}+C_G)) = V_2(C_P/(C_P+C_L))(C_{GD}/(C_{GD}+C_G)). \tag{3.2b}$$

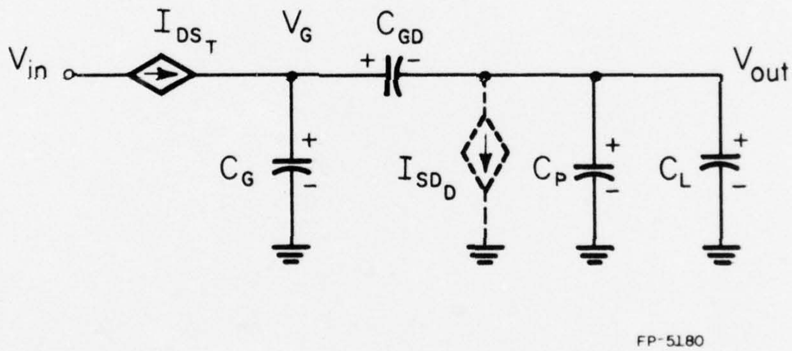


Figure 3.3. Equivalent model during time interval A

During interval B, the equivalent model is shown in Figure 3.4. The current source discharges both C_L and C_G. Just before clock ϕ_2 goes high, $V_G = V_{GD} = V_I$ as given by Equation (3.1). Just after clock ϕ_2 goes high the gate-source and the gate-drain voltages of the driver jump to $V_G = V_I + \Delta V_G$ and $V_{GD} = V_I - \Delta V_G(C_G/C_{GD})$ where ΔV_G is given by Equation (3.2). Since M_D is on in interval B, C_L discharges while $\Delta V_G C_G = -\Delta V_{GD} C_{GD}$ so that at the end of interval B $V_{GD} = V_G$. From these equations it is found that at the end of interval B $V_G = V_{GD} = V_I$. Thus the discharge of V_G just counteracts the initial jump. M_D then stays on during the duration of interval B and reaches the nonsaturation region as V_{OUT} approaches zero.

The ϕ_2 clock discontinuity from interval B to interval C are again transferred to the gate and drain of M_D. The magnitude of the jumps are given by Equation (3.2) but the polarities are reversed, forcing M_D into its reverse characteristics. The jump tries to force V_{OUT} to a large negative value but the magnitude becomes so great that it forward biases the drain-body diode which limits its excursion. The clock ϕ_1 turns on M_T, since $V_{GD} = V_1 - 0 > V_T$, and discharges C_G and the $C_{GD}, C_P + C_L$ combination to ground. As soon as V_{GD} and $V_{GS} < V_T$, M_D turns off. If C_L is still negatively charged, it would discharge through the drain-body pn junction. Above -0.6v, the discharge through the diode is very slow. The circuit model shown in Figure 3.5 is valid for this time interval.

During interval D both devices are held off. Both V_{OUT} and V_G follow the clock waveform reduced in magnitude by the appropriate capacitive voltage dividers given in Equation (3.2). $V_G = \Delta V_G$ and $V_{OUT} = \Delta V_{OUT} + V_D$,

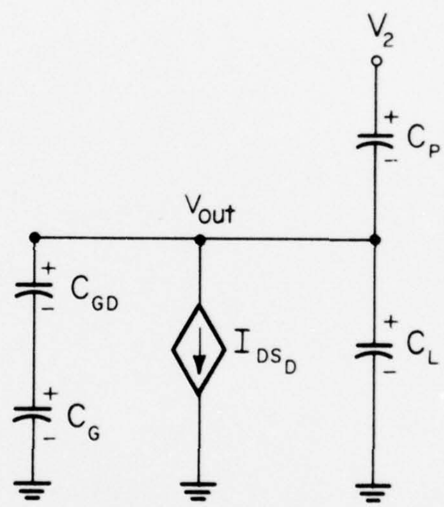


Figure 3.4. Equivalent model during time interval B

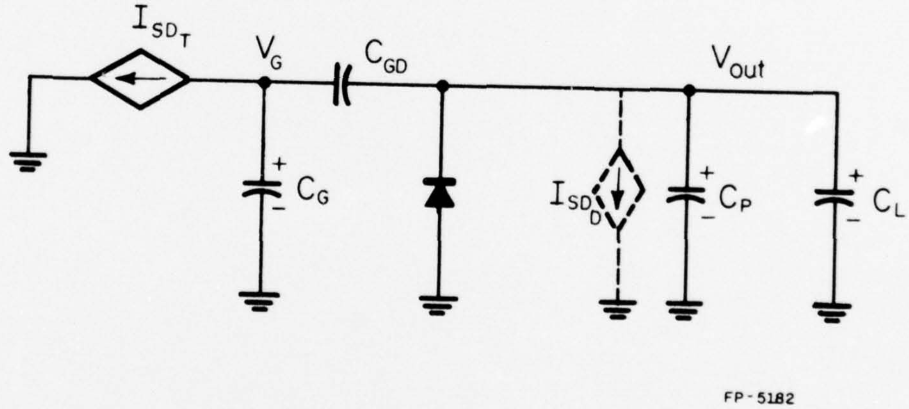


Figure 3.5. Equivalent model during time interval C

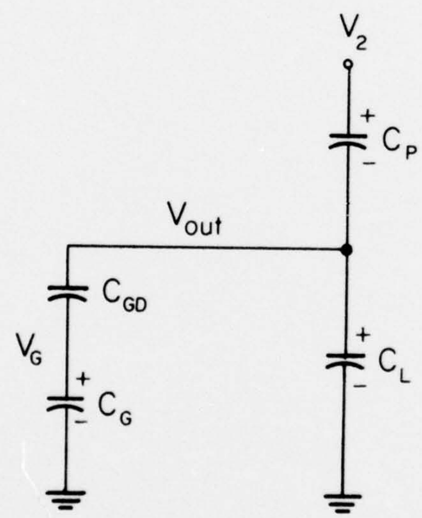


Figure 3.6. Equivalent model during time interval D

where V_D is the almost constant portion of the negative excursion of V_{OUT}. At the end of this interval, the voltages are returned to their original values.

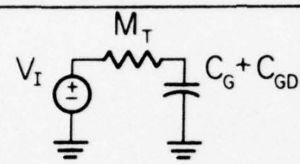
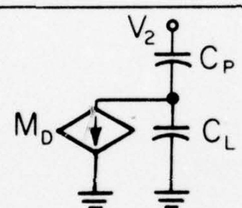
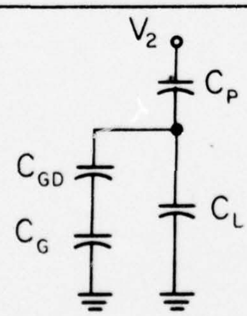
B. Design Equations

A summary of circuit operation is given in Figure 3.7. Included are simplified models for each interval and the design equations which are derived here. Using the simplified model for intervals A and C it is evident that for a fast response V_I must be large, the effective resistance of M_T must be small (i.e. M_T must be large), and the sum of the gate capacitances must be small. V_I may be made large by choosing V_1 so that the first half of Equation (3.1) is valid.

During interval B, the current source of the simplified model sees C_P+C_L as its load. For a fast discharge in this interval M_D must be large and C_P must be as small as possible (C_L cannot usually be arbitrarily chosen since it is a system constraint).

C_P cannot be made arbitrarily small because V_{OUT} must be within a threshold voltage of V_{IN} during interval D (V_{IN} represents both the input variable name and the value of the logic one level). Using the fact that $V_{OUT} = \Delta V_{OUT}+V_D$ must be less than or equal to $V_{IN}-V_T$ and Equation (3.2) it is found that

$$C_P \geq C_L \frac{V_{IN}-V_T-V_D}{V_2-(V_{IN}-V_T-V_D)} \quad . \qquad (3.7)$$

Interval	Simplified Model	Design Equations
A,C		V_I=Large M_T=Large Conductance $C_G + C_{GD}$=Small
B		M_D=Large C_P=Small
D		$C_P \geq C_L \left\{ \dfrac{V_{IN} - V_T - V_D}{V_2 - (V_{IN} - V_T - V_D)} \right\}$ $C_G > C_{GD} \left\{ \dfrac{V_2}{V_T} \left(\dfrac{C_P}{C_P + C_L} \right) - 1 \right\}$

FP-5184

Figure 3.7. Simplified summary of circuit operation

Equality is taken in order to satisfy the condition given for interval B. V_D cannot generally be neglected in the above equation since it is typically of comparable size to V_T. A value of -0.5v will usually be close enough. During this interval, V_G will also make a jump given by Equation (3.2). For proper device operation this jump must not be so large as to turn the device on or a diacharge similar to interval B will result. To insure that this does not happen, $\Delta V_G < V_T$ or, solving for C_G through Equation (3.2),

$$C_G > C_{GD}[(V_2/V_T)(C_P/(C_P+C_L))-1]. \qquad (3.8)$$

The conditions set forth for intervals A and C provide motivation to examine the smallest capacitors possible, the intrinsic MOS gate capacitances. Reference to Figure 2.2 will be useful to the following discussion where $C_G = C_{GS} + C_{GB}$. The design equation presented in interval A is automatically satisfied since $C_G + C_{GD} = C_{OX}$ for all regions of operation. During interval B, M_D operates in the saturation and nonsaturation regions. During the initial jump due to clock ϕ_2 going high, C_{GD} becomes nonzero while C_G gets smaller. This has the effect of increasing the jump of V_{GS} allowing more current to flow through M_D. During interval D the device is off. In this region of operation $C_{GD} \sim 0$ while $C_G \simeq C_{OX}$ fulfilling Equation (3.8). Thus the nonlinearity of these capacitors has a tendency to increase the speed of operation in each interval. These capacitors can be considered as a type of voltage variable feedback which allows smaller devices to be used to achieve a specified operation speed.

C. Computer Simulation

The circuit simulation included the effects of all the capacitors of both transistors. All the gate capacitors of M_T were considered constant and equal to the maximum value of C_{OX}. This can be considered a worst case value since larger capacitors correspond to slower response. Rise times of the voltage sources were 10 nsec. The modeling of the nonlinear M_D gate capacitances as linear capacitors by the simulation program must be examined for validity in each time interval. Intervals A and C will be accurate if $C_{GD} + C_G = C_{OX}$. During interval B, $0 < C_{GD} < \frac{1}{2} C_{OX}$ and $\frac{1}{2} C_{OX} < C_G < C_{OX}$. Average values of $C_{GD} = \frac{1}{4} C_{OX}$ and $C_G = \frac{3}{4} C_{OX}$ were chosen. These averages will not in general satisfy Equation (3.8) and results obtained for interval D will not generally be accurate.

For all simulations of this circuit $V_{IN} = 5v$, $V_1 = 5v$, and $C_L = 10pF$. V_1 was chosen equal to V_{IN} instead of larger, by Equation (3.4), to illustrate the dependence of V_I on the body effect. The variables remaining to be chosen are C_P, β_T, β_D, V_T, and V_2. C_{OX} is not a variable since it is dependent on β. The relation between C_{OX} and β can be found from Equations (2.4) and (2.5) (where C_{OX} now represents the total gate capacitance = $.222\text{pF} \times W\ell$). Multiplying these two equations it is found that

$$C_{OX} = 766 \times 10^{-6}/\beta, \quad \text{for } \beta < 8.63 \times 10^{-6},\ W = .2 \text{ mil},\ C_{OX} \text{ in pF.} \tag{3.9a}$$

Dividing the two equations yields

$$C_{OX} = 1030 \times \beta, \quad \text{for } \beta > 8.63 \times 10^{-6},\ \ell = .2 \text{ mil},\ C_{OX} \text{ in pF.} \tag{3.9b}$$

C_P and V_2 are related by Equation (3.7). It is desirable for V_2 to be a large standard value such as 12v. With this value of V_2, C_P is found to be ($V_D \simeq -.5$v)

$$C_P = \begin{cases} 7\text{pF}, & V_T = .5\text{v} \\ 6\text{pF}, & V_T = 1\text{v} \\ 5\text{pF}, & V_T = 1.5\text{v} \end{cases}$$

The highest frequency of circuit operation is limited by the capacitive discharge of interval B. Because of this, the performance index chosen to evaluate circuit speed was the time involved to discharge V_{OUT} to $\frac{1}{2}$ V_T during this interval. This value was used because a true logic zero must be significantly below V_T in order to discharge the next stage in a reasonable length of time.

If M_D is made larger to increase the rate of discharge, then C_{OX} will become larger as shown in Equation (3.9). M_T may then have to be increased to discharge or charge C_{OX} within the required time. If M_T is made too large, the capacitors associated with M_T become large enough to degrade circuit performance. Figure 3.8 illustrates how the discharge time is affected by the ratio β_D/β_T for several values of β_D. As can be seen, a minimum occurs near $\beta_T = 8.63 \times 10^{-6}$, the value for minimum M_T capacitance, and is relatively insensitive to β_D. For subsequent simulations, β_T was chosen to be this value. For small β_D, the capacitances of M_T become appreciable and degrade performance when M_T deviates from its minimum capacitance size. The nonlinearities of the capacitances of M_T may then need to be taken into account and may alter the shape of the curve.

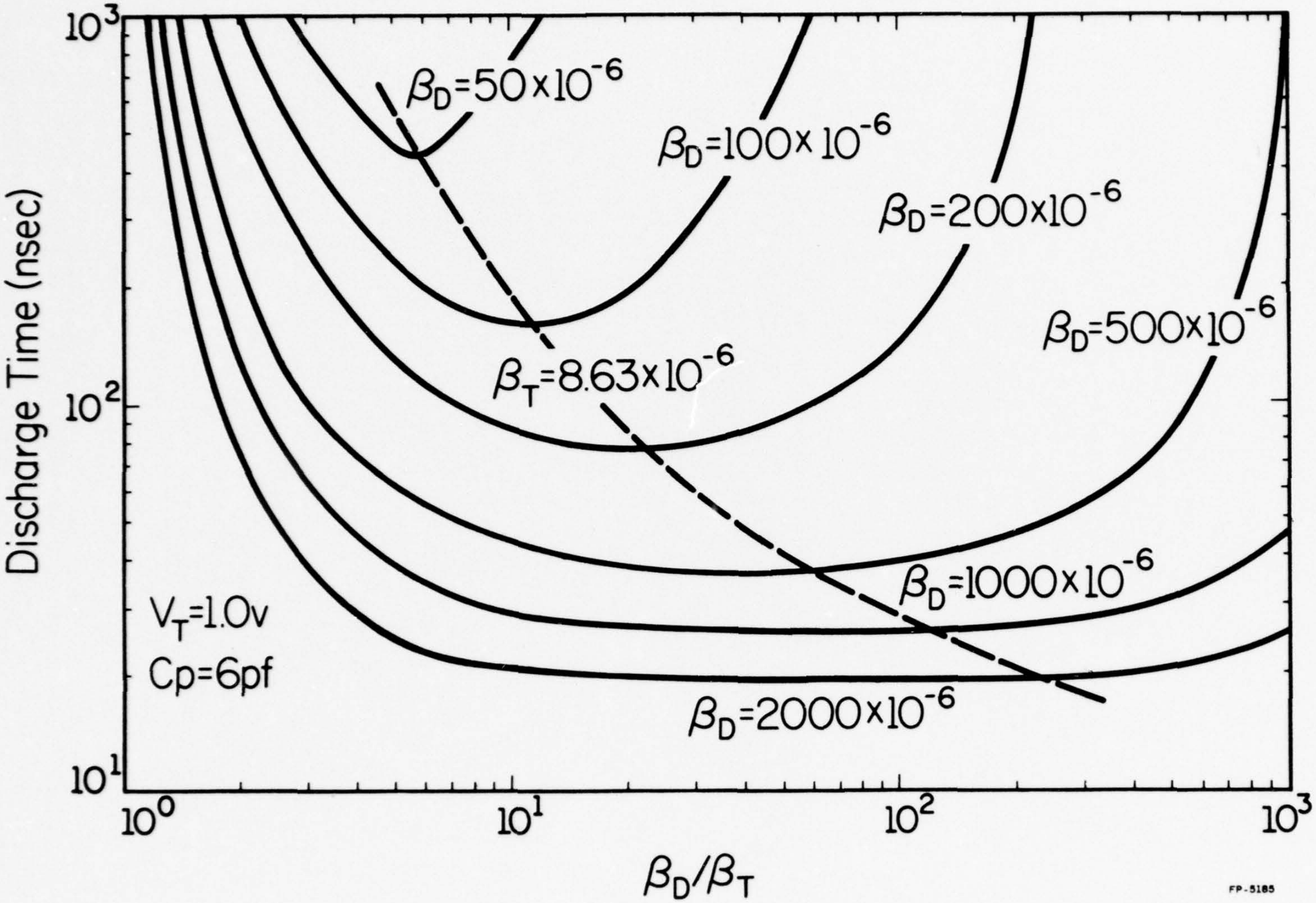


Figure 3.8. Discharge time vs. β_D/β_T

It is expected that if V_T is decreased the discharge rate will be faster. This is because V_I becomes greater and the current for a given V_{GS} becomes greater as shown in Equation (2.1). This equation also shows that a larger device (larger β) will cause a larger current to flow. These relationships are illustrated in Figure 3.9. The circuit speed is limited at the fast end by the clock rise time. The response of $V_T = 1.5v$ increases again at the low end because the high threshold devices cannot drive the large C_{OX} fast enough. The $V_T = .5v$ and $V_T = 1v$ responses cross over at the low end because the discharges are nearly equal but the lower threshold device must discharge V_{OUT} to a lower value. From this figure the proper device size and threshold voltage may be estimated for the circuit to operate at a certain frequency.

At a 1 megahertz clock rate, interval B is 500 nsec long. If it is desirable for V_{OUT} to be less than V_T within the first half of B, then the figure shows that $\beta_D > 50$. Choosing $V_T = 1v$, for illustration, and a slightly larger β_D of 86.3×10^{-6} ($W/\ell = 10$), the response of the circuit is shown in Figure 3.10. As discussed previously, this solution is fairly accurate except for interval D where V_G should remain near zero. The device parameters used for this example are the same as those given in the section discussing modeling. For M_T, since $\beta_T = 8.63 \times 10^{-6}$ A/v^2, all of the associated capacitors were set to .009pF. For M_D C_{OX} was found from Equation (3.9b) and is equal to .09pF. From this value C_G was chosen to be $\frac{3}{4} C_{OX} = .067pF$ and $C_{GD} = \frac{1}{4} C_{OX} = .023pF$. Typical minimum areas of small transistors such as M_T are 2 mil^2. Larger transistors, such as M_D may be made in 4 mil^2. The size of the pull-up capacitor can be found from

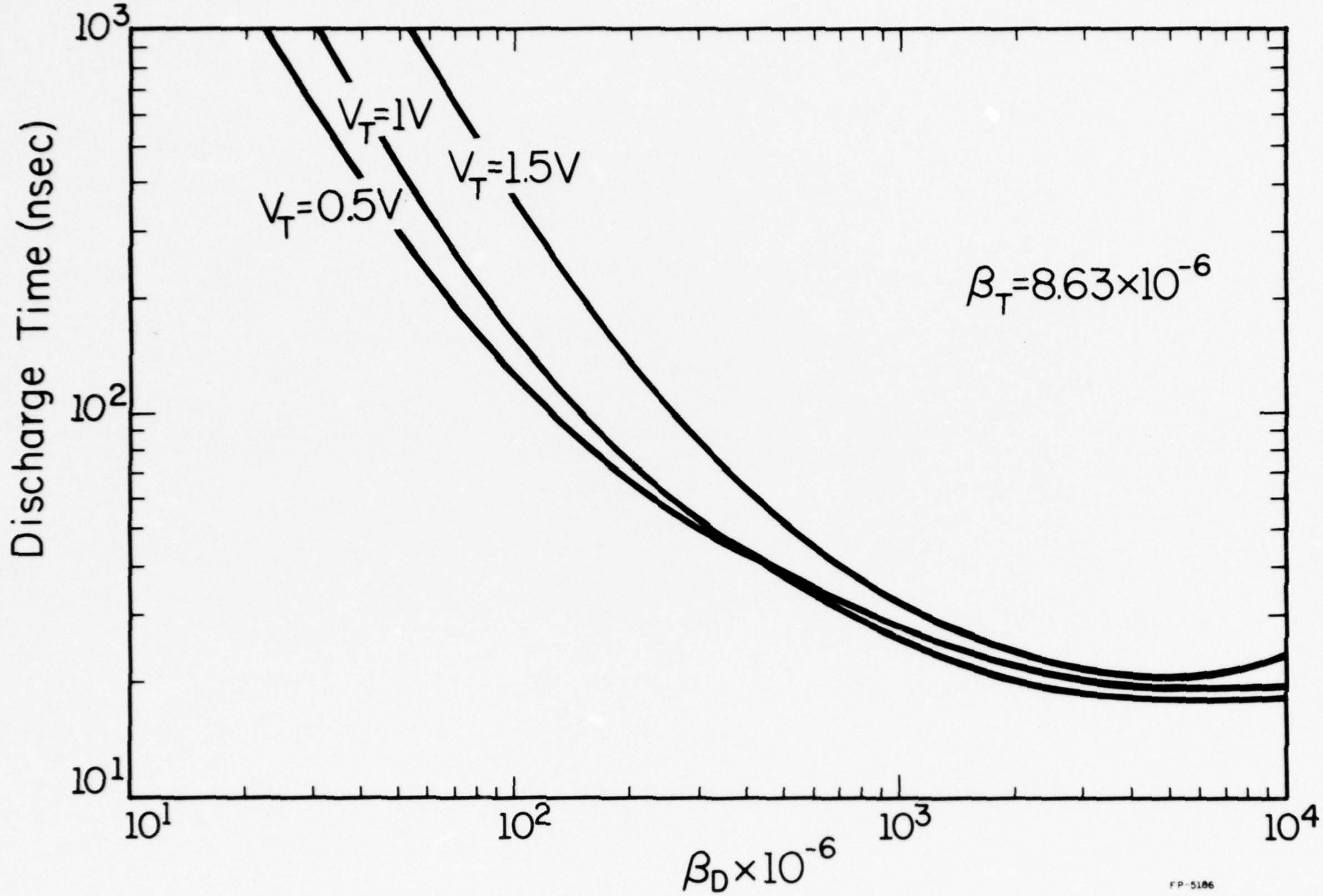


Figure 3.9. Discharge time vs. β_D

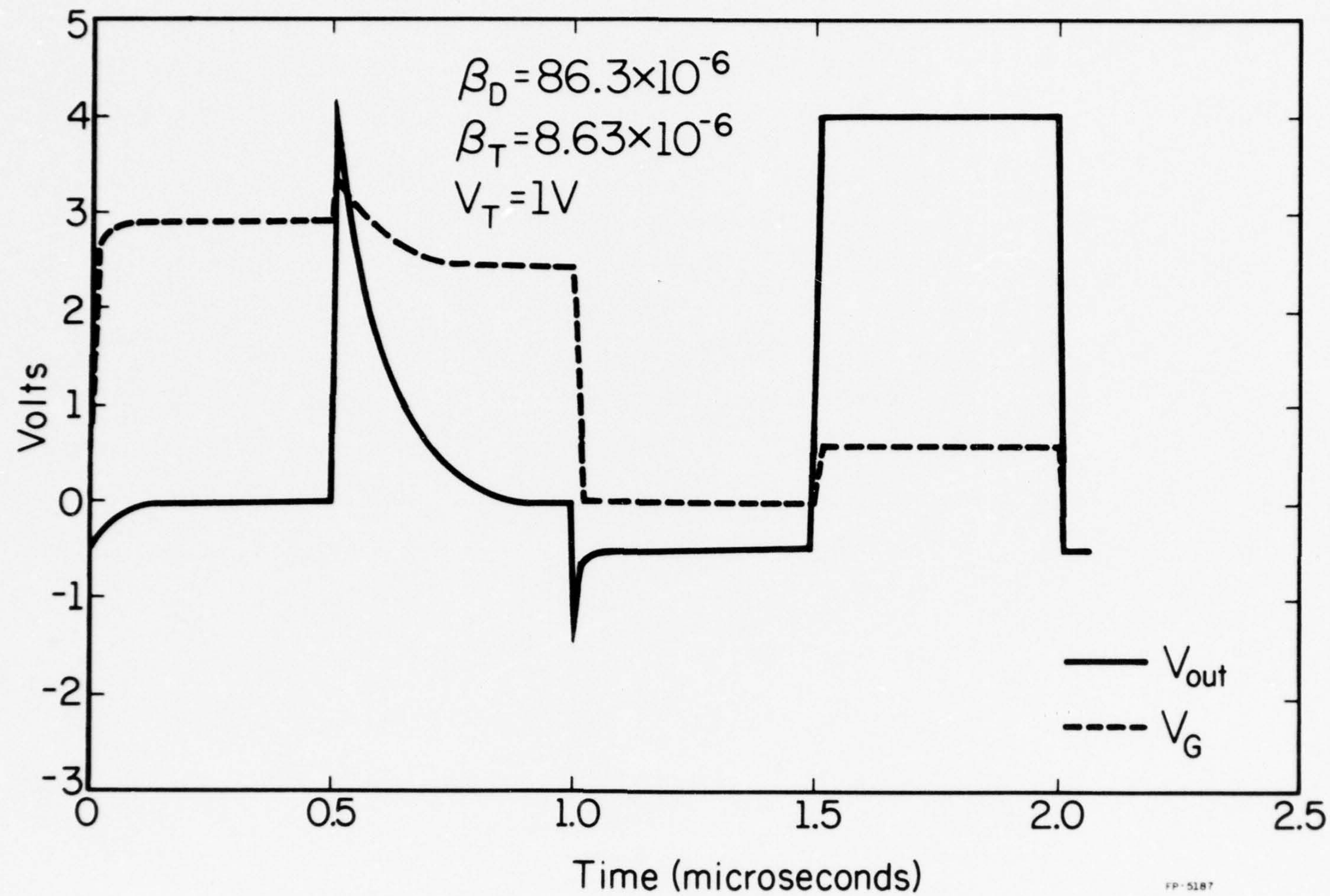


Figure 3.10. Circuit response

Equation (2.5). $W\ell = 6/.222 = 27\ \text{mil}^2$. Thus for this typical example, the circuit size is of the order of 33 mil^2 with over 80% of the area taken up by the pull-up capacitor.

D. Discussion of Results

The major disadvantage of this circuit is due to the forward biasing of the drain-body diode. This action injects minority carriers into the substrate which may cause signal degradation due to the NPN effect [1]. This can be effectively eliminated by placing a negative bias on the substrate. The disadvantage of this, however, is the reduced speed due to the body effect and the requirement of an extra negative voltage source. The body effect may be partially overcome by decreasing the device threshold voltage.

Another drawback of this circuit is the integration of the large pull-up capacitor. From the example presented the size of the 6pF capacitor is $W \times \ell = 27\ \text{mil}^2$ while the size of the largest device M_D is only 4 mil^2. The capacitor is almost 7 times the size of the device. This drawback may be partially overcome if a material with higher permittivity is used, such as silicon nitride (Si_3N_4-MNOS) [5]. The disadvantage of doing this is the resulting slower speed due to the increased C_{OX}.

It is desired to make the zero-level pulse of interval B as small as possible. This is most effectively done by increasing V_1. If it is possible to lower the threshold voltage of M_D while leaving all the other's constant (perhaps through ion implantation), then higher current is available for the interval B discharge while the high level pulse of interval D will remain within the desired range of V_{IN}.

A pulse is needed during the second half of intervals B and D to provide a data enable for the next stage. It provides a signal to the following stages that V_{OUT} is stable. This pulse may be derived on the chip if a clock of twice the frequency of ϕ_2 is provided.

A possible realization of the driver device M_D and the pull-up capacitor C_P is given in reference [6].

4. SOURCE-FOLLOWER CIRCUIT

The approximations used to aid in the analysis of the pull-up circuit are also assumed for the analysis of the source-follower circuit. The resulting approximate model to be analyzed is shown in Figure 4.1. A typical clocking sequence and corresponding output is shown in Figure 4.2. When clock ϕ_1 is high (the second half of intervals A and C), data is transferred to the gate of M_D by charging the gate capacitances of M_D. When clock ϕ_2 goes high (intervals B and D), M_D is turned on allowing the data to be taken from its source. This circuit is noninverting and requires no components other than the two transistors.

A. Circuit Analysis

During the second half of interval A, V_{IN} and ϕ_1 are both high turning on M_T. Assume that in the high state all clock and input voltages have the same amplitude V_1. If $V_{GS} = V_1 > V_T$ and $V_{GD} = 0 < V_T$ them M_T is operating in the saturation region. As the series combination of C_{GS} and C_L is charged, M_D will conduct when $V_{GS} = V_G - V_{OUT} \geq V_T$. This will in turn drive V_{OUT} to ground from its previous nonzero value. During this interval, the equivalent model for the circuit is shown in Figure 4.3. The effect of the current source is to keep C_L discharged. V_G will charge up to V_I, where $V_I = V_1 - (V_T + \frac{1}{2}\sqrt{V_I})$, when M_T turns off. V_I may be solved from this equation to obtain

$$V_I = V_1 - V_T + \frac{1}{8} - \frac{1}{2}\sqrt{V_1 - V_T + \frac{1}{16}} \; . \qquad (4.1)$$

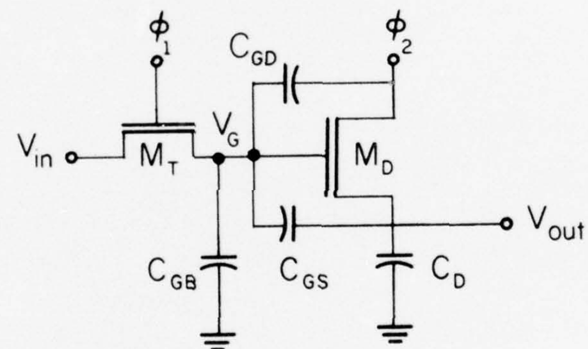


Figure 4.1. Approximate circuit to be analyzed

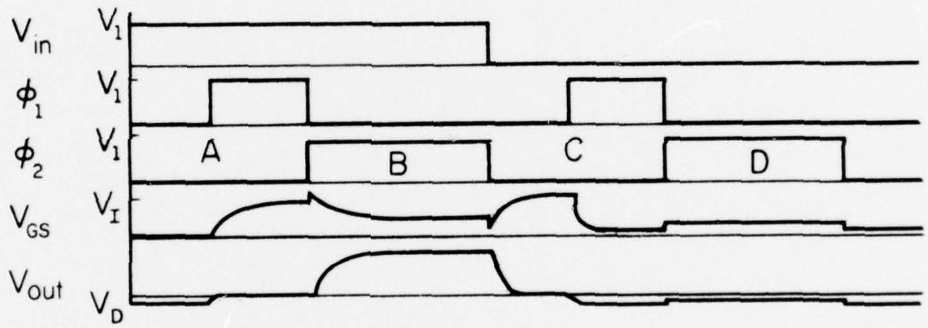


Figure 4.2. Typical clocking sequence and output

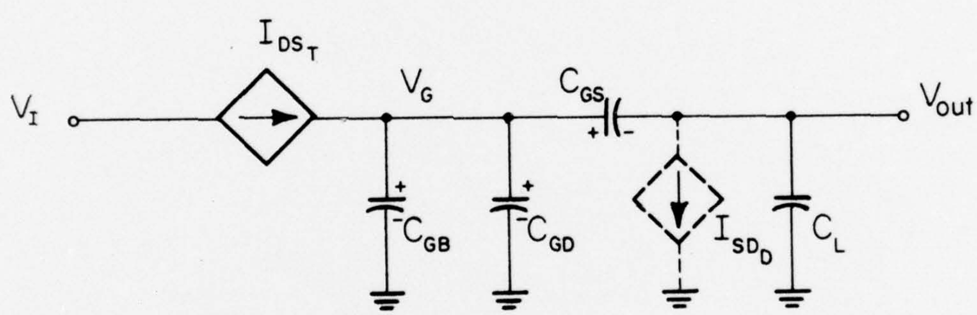


Figure 4.3. Equivalent model during time interval A

Between intervals A and B, Φ_1 goes low turning M_T off ($V_{GS} = -V_I < 0 < V_T$). The discontinuity of clock Φ_2 going high is coupled to the gate and source of M_D by the capacitive voltage dividers

$$\Delta V_{GS} \sim \Delta V_G \sim V_1 (C_{GD}/(C_{GB} + C_{GS} + C_{GD})) \tag{4.2a}$$

$$\Delta V_{OUT} \sim \Delta V_{GS}(C_{GS}/(C_L + C_{GS})) \simeq V_1 (C_{GD}/C_{GB} + C_{GS} + C_{GD}))(C_{GS}/(C_L + C_{GS})) \tag{4.2b}$$

During time interval B the equivalent model is shown in Figure 4.4. The current source charges C_L, C_{GB}, and C_{GD} while discharging C_{GS}. At the beginning of interval B, $V_{GS} = V_I + \Delta V_{GS}$ and $V_{GD} = V_I - \Delta V_{GS}((C_{GS} + C_{GB})/C_{GD})$ where ΔV_{GS} is given by Equation (4.2). During this interval charge is conserved at the node of V_G requiring $\Delta V_{GD}(C_{GB} + C_{GD}) = -\Delta V_{GS} C_{GS}$. Utilizing these equations it can be found that at any given V_{OUT}

$$V_{GS} = V_I + \Delta V_{GS} - V_{OUT}((C_{GB} + C_{GB})/(C_{GB} + C_{GS} + C_{GD})) \tag{4.3a}$$

$$V_{GD} = V_I - \Delta V_{GS}((C_{GS} + C_{GB})/C_{GD}) + V_{OUT}(C_{GS}/(C_{GB} + C_{GS} + C_{GD})) \tag{4.3b}$$

where ΔV_{GS} is given by Equation (4.2). These equations show that as V_{OUT} approaches V_1, V_{GS} and V_{GD} approach $V_I - V_1(C_{GB}/(C_{GB} + C_{GS} + C_{GD}))$ and M_D reaches the nonsaturation region. Comparing this interval with the corresponding interval of the pull-up circuit is it evident that charge is lost through C_{GB} making this circuit inherently slower.

The Φ_2 clock discontinuity from interval B to C are again transferred to the gate and source of M_D. The magnitudes of the jumps are given by Equation (4.2) but the polarities are reversed. The model which is valid during this interval is shown in Figure 4.5. During the first half of this

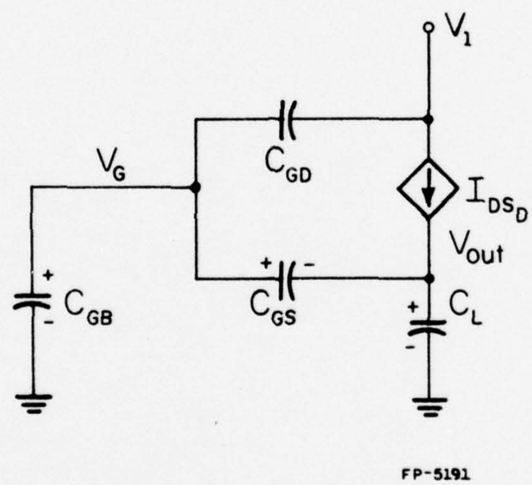


Figure 4.4. Equivalent model during time interval B

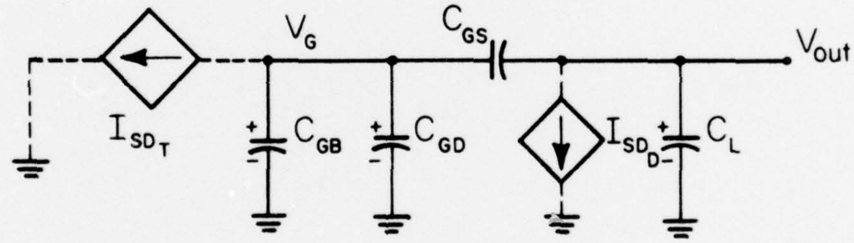


Figure 4.5. Equivalent model during time interval C

interval M_T is off while the current source of M_D discharges C_L. During this discharge, charge is again conserved at the gate of M_D and as V_{OUT} approaches zero, V_{GS} and V_{DS} approach V_I. During the time clock ϕ_2 is high, M_D operates in the nonsaturation region. After C_L is discharged, clock ϕ_1 is allowed to go high. This turns M_T on which discharges C_{GD}, C_{GS}, and C_{GB}. As soon as $V_{GD} = V_{GS} = V_T$, M_D is turned off allowing a small negative voltage to be developed across C_L as M_T continues to discharge C_{GB} until $V_G = 0$. This voltage is given by V_D where

$$V_D = -V_T(C_{GS}/(C_{GS} + C_L)).$$

During interval D both devices are off provided that the jump in ϕ_2 does not cause ΔV_{GS} to exceed V_T. V_{OUT} and V_{GS} follow the clock waveform reduced in magnitude by the capacitive voltage dividers given in Equation (4.2) where $V_{OUT} = \Delta V_{OUT} + V_D$ and $V_{GS} = \Delta V_{GS} - V_D$. After ϕ_2 goes low, the voltages are returned to their original values.

B. Design Equations

A summary of circuit operation during each interval is shown in Figure 4.7. Included are simplified models for each interval and the design equations which are derived here. Using the simplified model for intervals A and C, it is evident that for a fast response the effective resistance of M_T must be small (i.e. M_T must be large), and the sum of the gate capacitance must be small. However, the time constants are so small in these intervals

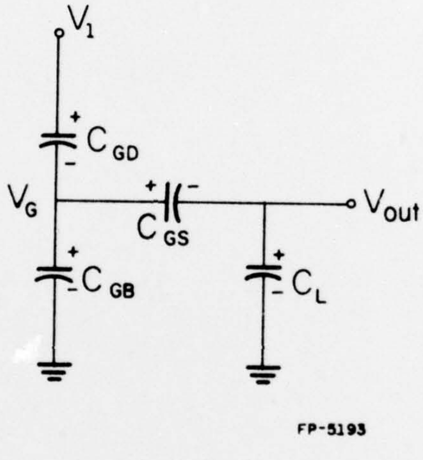
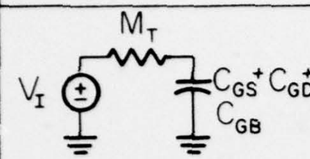
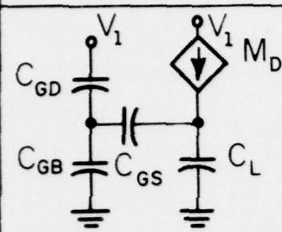
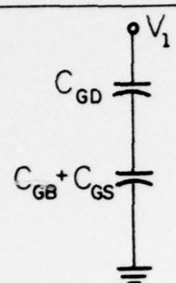


Figure 4.6. Equivalent model during time interval D

Interval	Simplified Model	Design Equations
A,C		M_T=Large Conductance $C_{GS}+C_{GD}+C_{GB}$=Small
B		M_D=Large C_B=Small $V_1\left(\frac{C_{GB}+C_{GD}}{C_{GS}+C_{GD}+C_{GB}}\right) \geq V_T\left(\frac{C_{GS}}{C_{GS}+C_{GD}+C_{GB}}+1\right)+\sqrt{V_1-V_T}$
D		$C_{GD} < \left(\frac{C_{GB}+C_{GS}}{V_1/V_T-1}\right)$

FP-5194

Figure 4.7. Simplified summary of circuit operation

compared to interval B that these parameters are not critical in determining the speed of the circuit.

During interval B, a large M_D will charge C_L quickly. It was found previously that during this interval V_{GS} droops by $V_1(C_{GB})/(C_{GB}+C_{GS}+C_{GD})$ as V_{OUT} approaches V_1. To minimize this droop, and keep the current through M_T large, C_{GB} must be small in comparison to C_{GS} and C_{GD}. Also during this interval, the effective threshold voltage increases because of the rising source voltage and the body effect. This together with the V_{GS} droop is the main factor involved in limiting circuit speed. Utilizing Equations (4.1), (4.2), and (4.3), it can be found that for $V_{GS} \geq V_T + \frac{1}{2}\sqrt{V_{OUT}}$ when $V_{OUT} = V_1 - V_T$ (i.e. M_D has not turned off when the desired output is reached), then Equation (4.4) must be satisfied.

$$V_1[1 - C_{GB}/(C_{GS}+C_{GD}+C_{GB})] + \frac{1}{8} \geq V_T[1 + C_{GS}/(C_{GS}+C_{GD}+C_{GB})] + \sqrt{V_1 - V_T} \quad (4.4)$$

To gain some insight into the restrictions placed by the above equation, an example is graphed in Figure 4.8 as the V_{Tmax} curve. For this example $C_{GB}/(C_{GS}+C_{GD}+C_{GB}) = \frac{1}{6}$ and $C_{GS}/(C_{GS}+C_{GD}+C_{GB}) = \frac{7}{12}$; the reasons for using these ratios will become clear in the next section. The region of proper operations is in the lower right-hand portion.

During interval D it is required that the jump of V_{GS} not be large enough to turn M_D on. Using Equation (4.2), where $V_{GS} = \Delta V_{GS} < V_T$, and neglecting V_D, M_D will stay off if

$$C_{GD} < (C_{GB}+C_{GS})/(V_1/V_T - 1). \quad (4.5)$$

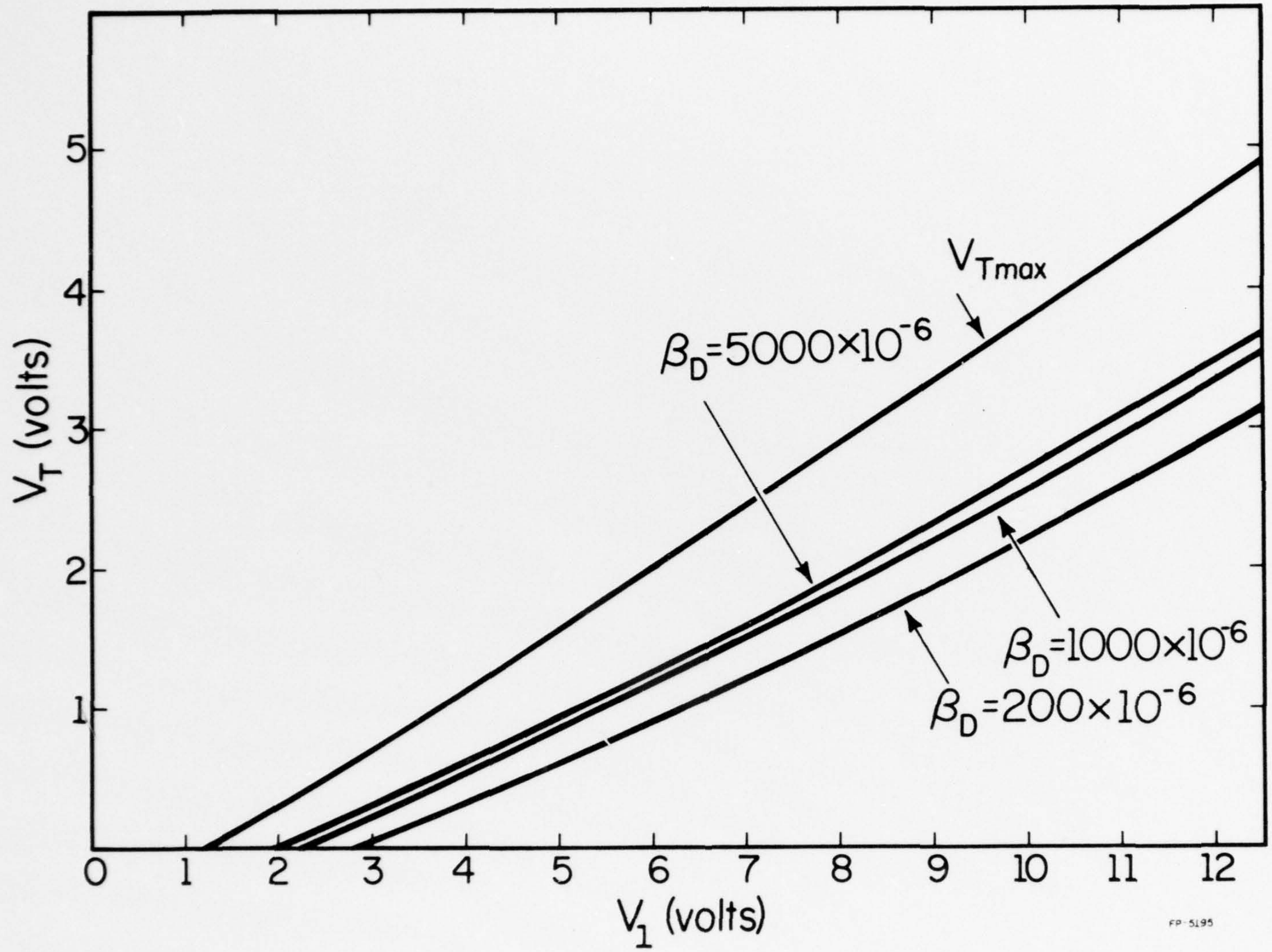


Figure 4.8. V_T vs. V_1 at $V_{OUT} = V_1 - V_T$

The analysis of the pull-up circuit provides motivation to study the effects of the nonlinear MOS gate capacitances in this circuit. The requirements of interval A are satisfied since $C_{GS}+C_{GB}+C_{GD} = C_{OX}$ for all regions of operation. During interval B, M_D operates in the saturation and nonsaturation regions. In these regions of operation C_{GB} is very small while C_{GD} becomes larger. This not only has the tendency to reduce the V_{GS} droop but also increases the initial positive-going V_{GS} jump. These both have the effect of increasing the logic-one rise time of interval B. In interval D the device is off. In this region of operation $C_{GB} \simeq C_{OX}$ while $C_{GS} \simeq C_{GD} \simeq 0$ satisfying Equation (4.5). Thus, just as in the pull-up circuit, the effect of the nonlinear gate capacitances is to increase the overall speed of operation.

C. Computer Simulation

For proper circuit comparisons, the initial approximations used in the simulation of the pull-up circuit are also assumed here. The capacitances of M_T were assumed constant and equal to C_{OX}. The voltage source rise times were 10 nsec. The effects of modeling the nonlinear gate capacitances as linear capacitors must again be examined interval by interval. During intervals A and C the model will provide accurate results if $C_{GB}+C_{GD}+C_{GS} = C_{OX}$. During interval B, $0 < C_{GB} < \frac{1}{3} C_{OX}$, $0 < C_{GD} < \frac{1}{2} C_{OX}$, and $\frac{1}{2} C_{OX} < C_{GS} < \frac{2}{3} C_{OX}$. Average values of $C_{GB} = \frac{1}{6} C_{OX}$, $C_{GD} = \frac{1}{4} C_{OX}$, and $C_{GS} = \frac{7}{12} C_{OX}$ were chosen. These values will not generally satisfy Equation (4.5) and will give false results for interval D. During this interval V_{OUT} and V_{GS} should remain close to the respective previous values since $C_{GD} \simeq 0$ in Equation (4.2). The remaining variables that need to be chosen are V_1, V_T, β_D, and β_T.

An important parameter in the operation of this circuit is the logic-one level pulse height of interval B. It may not reach V_1 during this interval since the effective threshold voltage is increasing and V_{GS} is decreasing. Although the device may not turn off, the current can become so small that V_{OUT} changes very little over the remaining portion of time Figure 4.8 shows limiting values of V_T and V_1 for V_{OUT} to reach within a threshold voltage of the final value (V_1-V_T) in less than 500 nsec. This length of time was chosen because for the device sizes of interest, V_{OUT} changed very little after this. The safe portion of the graph to operate in is below the curve of proper β_D. As can be seen, V_T should be chosen considerably less than the maximum found earlier in order to achieve a reasonable rise time.

The interval limiting the total circuit speed is interval B. The performance index chosen to evaluate total circuit speed was the time V_{OUT} needed to rise to V_1-V_T in interval B. The variation of the charging time with respect to the ratio β_D/β_T is shown in Figure 4.9. For these simulations $V_1 = 5v$ and $V_T = .5v$. When the ratio is small, the capacitors associated with M_T become comparative in size to C_{GB} making the V_{GS} droop of interval B more pronounced. For large ratios, response is limited by two factors. For large β_D, M_T becomes too small to charge the gate capacitances of M_D in the required time. For small β_D, the capacitors of M_T are no longer negligible and degrade circuit performance. For this case the nonlinear effects of the gate capacitance may have to be taken into account for more accurate results.

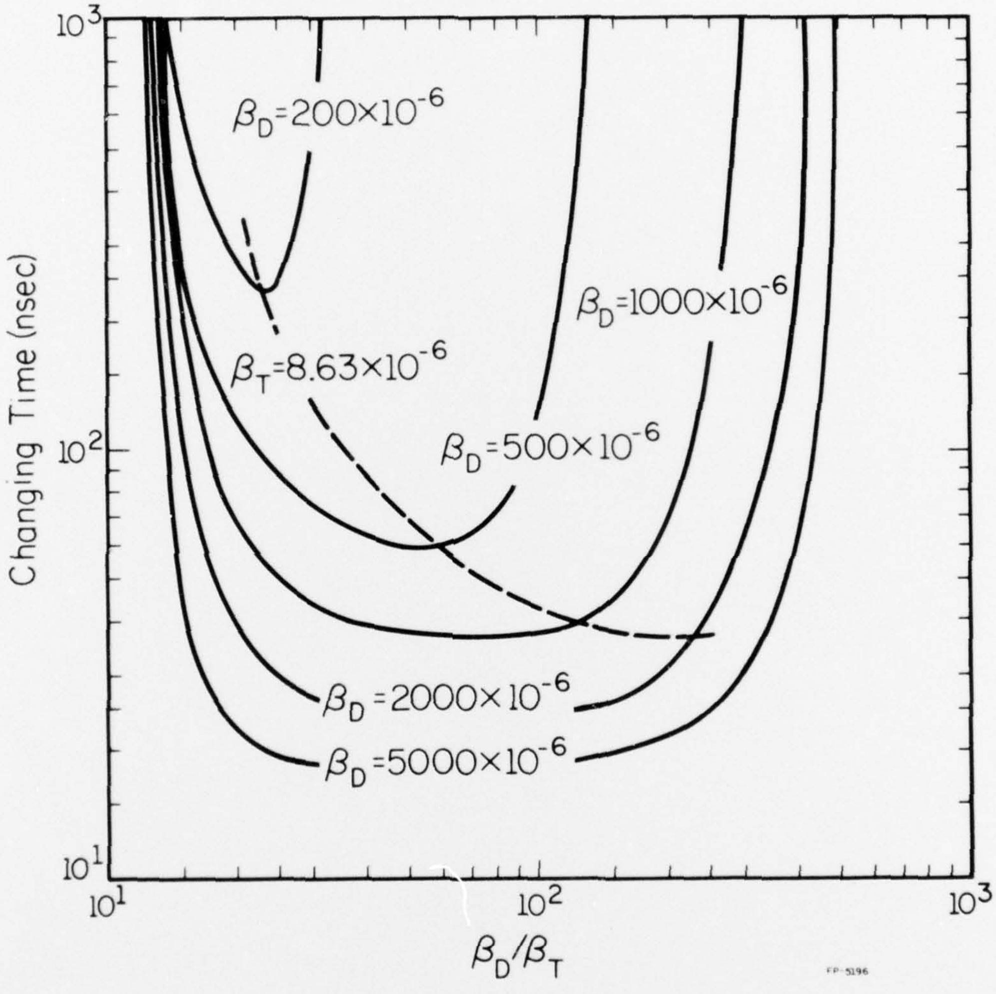


Figure 4.9. Charging time vs. β_D/β_T

From this graph the optimum value of β_T is chosen to be $\max\{8.63 \times 10^{-6}, \beta_D/50\}$. This value of β_T was used for all other simulations.

The variation of the charging time in interval B in relation to β_D is illustrated in Figure 4.10 for several V_T ($V_1 = 5v$). The response is limited at the fast end by the clock rise time and increases as β_D decreases. This circuit is much more sensitive to threshold voltage changes than the pull-up circuit. This is due to the dependence of the rise time in interval B on V_T and the body effect. From this curve the proper device size can be found for the circuit to operate at a specific frequency.

As an example, let the desired clock frequency be 1 megahertz. Then if it is desired that V_{OUT} be within a threshold voltage of its final value within 10% of the clock period (100 nsec), then $\beta_D = 325 \times 10^{-6}$ (W/ℓ = 37.5). The corresponding output is shown in Figure 4.11. As discussed previously, this output is valid for all intervals except D. During this interval V_{OUT} and V_{GS} are expected to remain near ground.

D. Discussion of Results

As is typical of most MOS circuits, this circuit will operate faster with a higher power supply voltage V_1. This will also allow devices with higher threshold voltages to be used.

The small transmission gate transistor M_T may be as small as 2 mil^2. Using $\sqrt{2}$ mil as the smallest device length and finding the width of M_D from the required W/ℓ ratio (W = 7.5 mil from the example cited), a typical driver

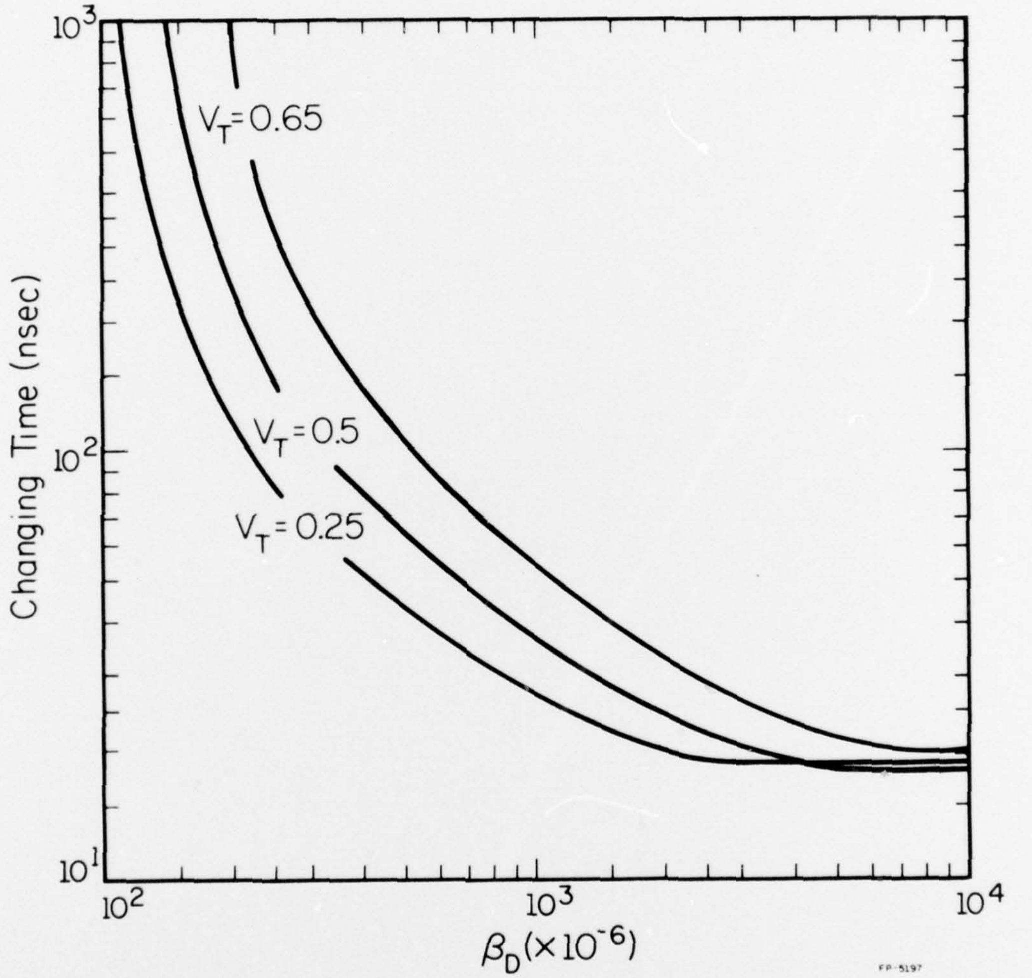


Figure 4.10. Charging time vs. β_D

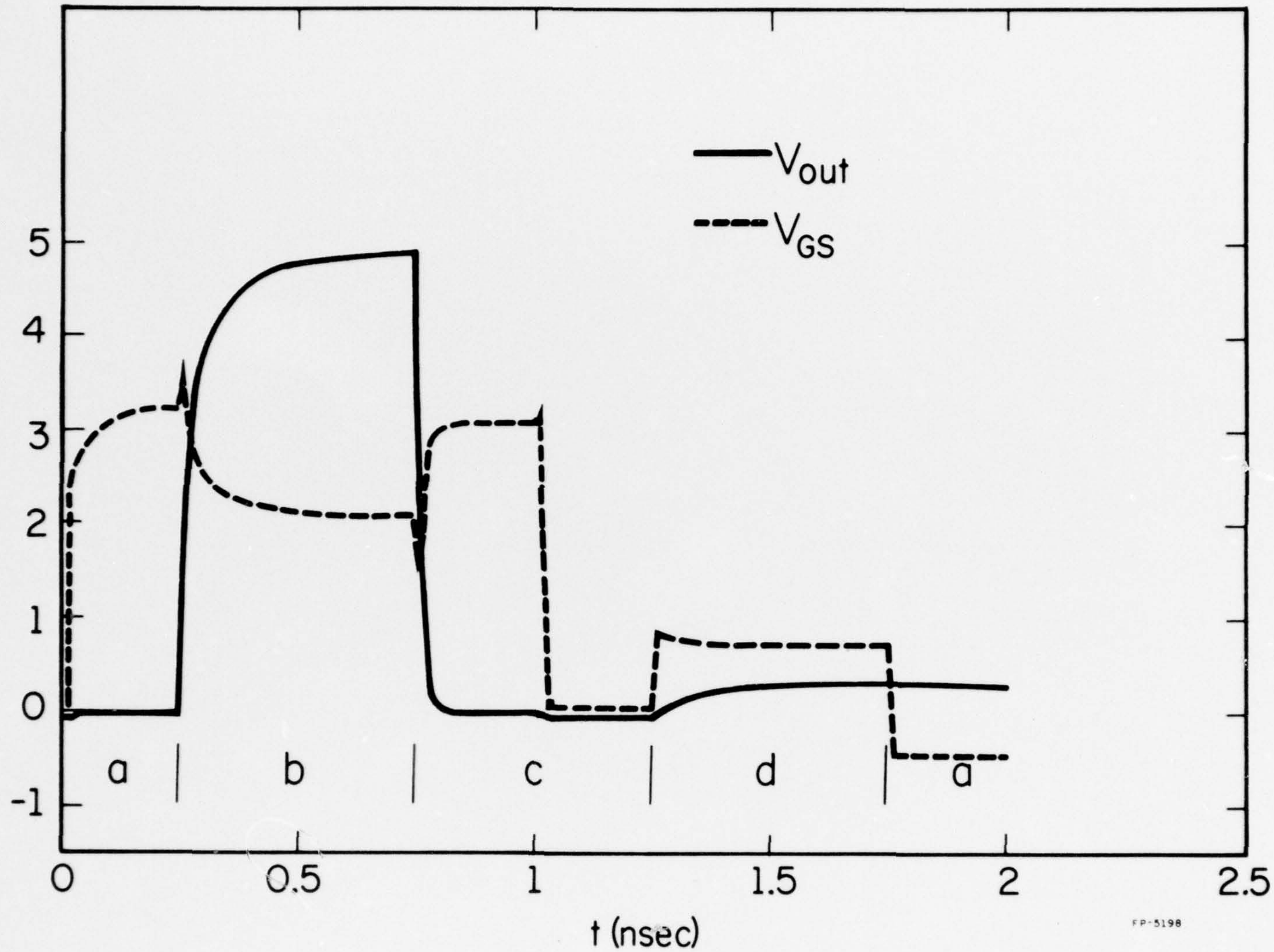


Figure 4.11. Circuit response

device size will be on the order of 10 mil^2. The total circuit area, neglecting interconnections, will be close to 12 mil^2.

A disadvantage of this circuit is the gap required between the logic one levels of the two clocks. From interval C of Figure 4.11 it can be seen that this gap can be made quite small ($\sim$50 nsec) in order to fully discharge C_L. This gap may be eliminated if devices and capacitors of correct sizes are used so that, during interval C, as M_T discharges the gate capacitances of M_D, turning it off, M_D has discharged C_L. Initial inspection reveals that C_{GS} and C_{GD} may be increased to aid in this, but care must be taken so that the device does not turn on during interval D because of these changes.

5. SUMMARY

In this thesis two circuits are presented capable of driving large capacitive loads. These circuits are suitable for dynamic operation and utilize the nonlinear gate capacitances to allow small devices to drive loads larger than normal. Both circuits are capable of operating at frequencies above several megahertz for a load capacitance less than or equal to 10pF.

The pull-up circuit uses devices of standard threshold voltage levels but requires a large voltage clock for a reasonably sized capacitor to be integrated. Another voltage source may be required as a substrate bias to eliminate the NPN effect. A clock of twice the normal operating frequency is required to derive a data enable signal for the following stages. In the typical example operating at 1 megahertz cited, the total circuit size was approximately 33 mil^2.

The source-follower is inherently a slower circuit due to the body effect. It requires devices of low threshold voltages if low supply voltages are to be used. This circuit also requires clocks with gaps between the respective logic-one levels to allow for a fast two stage discharge (interval C). In the example cited, the devices are larger than the pull-up circuit devices but the total circuit size is approximately only 12 mil^2. This is approximately $\frac{1}{3}$ the size of the pull-up circuit.

LIST OF REFERENCES

1. American Microsystems, Inc., MOS Integrated Circuits, Van Nostrand, Reinhold, New York, 1972.

2. F. S. Jenkins, E. R. Lane, W. W. Lattin, and W. S. Richardson, "MOS-Device Modeling for Computer Implementation," IEEE Trans. Circuit Theory, Vol. CT-20, pp. 649-658, Nov. 1973.

3. H. Shichman and D. A. Hodges, "Modeling and Simulation of Insulated-Gate Field-Effect Transistor Switching Circuits," IEEE J. Solid-State Circuits (Special Issue 1968 International Solid-State Circuits Conf.), Vol. SC-3, pp. 258-289, Sept. 1968.

4. L. W. Nagel, "SPICE 2: A Computer Program to Simulate Semiconductor Circuits," Memorandum No. ERL-M520, Electronics Research Lab., University of California at Berkeley, 9 May 1975.

5. D. J. Hamilton and W. G. Howard, Basic Integrated Circuit Engineering, McGraw-Hill, New York, 1975.

6. J. P. Ellul, M. A. Copeland, and C. H. Chan, "MOS Capacitor Pull-up Circuits," IEEE J. Solid-State Circuits, Vol. SC-10, pp. 298-307, Oct. 1975.