Technical Report 767

Qualitative Analysis of MOS Circuits

Brian C. Williams

MIT Artificial Intelligence Laboratory

Qualitative Analysis of MOS Circuits

by

Brian Charles Williams

July 1984

This report is a revision of a thesis submitted to the Department of Electrical Engineering and Computer Science in January, 1984 in partial fulfillment of the requirements for the degree of Master of Science.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research Contract numbers N0014-80-C-0505 and N00014-80-C-0622.

Qualitative Analysis of MOS Circuits

by

Brian Charles Williams

Abstract

With the push towards sub-micron technology, transistor models have become increasingly complex. The number of components in integrated circuits has forced designers' efforts and skills towards higher levels of design. This has created a gap between design expertise and the performance demands increasingly imposed by the technology. To alleviate this problem, software tools must be developed that provide the designer with expert advice on circuit performance and design. This requires a theory that links the intuitions of an expert circuit analyst with the corresponding principles of formal theory (i.e., algebra, calculus, feedback analysis, network theory, and electrodynamics), and that makes each underlying assumption explicit.

Temporal Qualitative Analysis is a technique for analyzing the qualitative large signal behavior of MOS circuits that straddle the line between the digital and analog domains. Temporal Qualitative Analysis is based on the following four components: First, a qualitative representation is composed of a set of open regions separated by boundaries. These boundaries are chosen at the appropriate level of detail for the analysis. This concept is used in modeling time, space, circuit state variables, and device operating regions. Second, constraints between circuit state variables are established by circuit theory. At a finer time scale, the designer's intuition of electrodynamics is used to impose a causal relationship among these constraints. Third, large signal behavior is modeled by Transition Analysis, using continuity and theorems of calculus to determine how quantities pass between regions over time. Finally, Feedback Analysis uses knowledge about the structure of equations and the properties of structure classes to resolve ambiguities.

Thesis Advisors:

Howard Shrobe, Principal Research Scientist **Richard Zippel, Asst. Prof. of Comp. Sci. and Eng.**

To my Father and Mother

Acknowledgments

I would like to express my thanks to all of the people who encouraged and aided me in this work:

My advisors, Howie Shrobe and Richard Zippel, for their continued support and valuable recommendations.

Johan de Kleer, whose work on qualitative reasoning layed the foundation for this thesis and whose advice pointed me in the proper direction.

Ramesh Patil, Walter Hamscher and Randy Davis for wading through early drafts of this thesis.

Mark Shirley, Andy Ressler, Ken Forbus, Dan Weld and Dan Carnese for their friendship and many stimulating discussions.

Finally, Leah Ruby who has supported me in every part of my life,

and my family, for many years of love and the belief that anything is possible.

Table of Contents

Chapter One: Introduction 9

1.1 An Explanation for a Complex Bootstrap Driver 10
1.2 Temporal Qualitative Analysis: an Overview 13

Chapter Two: Qualitative Values 16

2.1 The Network Model 16
2.2 Time 18
2.3 State Variables 18
2.4 Qualitative Expressions 19
2.5 Operating Regions 20
2.6 Summary 20

Chapter Three: Causal Propagation 21

3.1 Causality and the Lumped Circuit Approximation 22
3.2 Implementing Causal Propagation 24
3.3 Domain Knowledge 24
3.3.1 Network Laws 25
3.3.2 Device Models 26
3.3.2.1 KCL Applied to Devices 26
3.3.2.2 Network Elements 26
3.4 Example 27
3.5 Ambiguities 28
3.6 Summary 29

Chapter Four: Transition Analysis 30

4.1 Transition Recognition 31
4.1.1 The Intermediate Value Theorem 31
4.1.2 State Variables and Time 32
4.1.3 Mean Value Theorem 35
4.2 Transition Ordering 38
4.2.1 Transition Ordering Rules 40
4.2.2 Applying the Transition Ordering Rules 40
4.3 Example 41
4.3.1 Discontinuous Behavior 42
4.4 Summary 43

Chapter Five: Feedback Analysis 44

5.1 The Effects of Feedback and Simultaneities on TQ Analysis 45
5.2 Qualitative Properties of Feedback 46
5.3 Bidirectional Comparison Points 52

5.4 Localizing the Effects of KVL 55
5.5 Examples 56
5.5.1 RC Ladder Example 56
5.5.2 Wheatstone Bridge Example 58
5.6 Simultaneities Involving Primary Causes 61
5.7 High Order Derivatives 62
5.8 Summary 65

Chapter Six: Operating Regions and the MOS Transistor 66

6.1 Modeling Devices With Multiple Operating Regions 66
6.2 The Mosfet Conduction Model 70
6.3 Mosfet Capacitance Model 72
6.4 Mosfet Example 73
6.5 Summary 78

Chapter Seven: Discussion 79

7.1 Related Work 79
7.1.1 Qual and Envisioning 80
7.1.1.1 Operating Regions and State 80
7.1.1.2 Feedback 82
7.1.2 Qualitative Process Theory 82
7.1.3 Allen's Temporal Intervals 83
7.2 Limitations and Future Directions 85

Appendix A: Appendix: Transition Ordering Rules 87

A.1 Predicates 87
A.2 General 87
A.3 Equality 87
A.4 Sums and Differences 87
A.5 Products 88
A.6 Resistive Feedback 88

Table of Figures

Figure 1-1: Bootstrap Driver 10
Figure 1-2: RC Example 14
Figure 4-1: Relationship between State Variables and Time 33
Figure 4-2: Continuity 34
Figure 4-3: Continuity Example 35
Figure 4-4: Integration Rule 37
Figure 5-1: RR Current Divider 44
Figure 5-2: Causal Relation Graph for RR Current Divider 45
Figure 5-3: Support Classification 47
Figure 5-4: RC Current Divider 50
Figure 5-5: RC High Pass Filter 54
Figure 5-6: RCRC Ladder and Causal Relation Graph 57
Figure 5-7: Wheatstone Bridge and Causal Relation Graph 59
Figure 5-8: Parallel RC Circuit and Causal Relation Graph 62
Figure 6-1: n-Mosfet Large Signal Model 67
Figure 6-2: n-Mosfet Operating Regions 68
Figure 6-3: N-Mosfet Conduction Model 71
Figure 6-4: Enhancement n-Channel Mosfet Conduction Model 72
Figure 6-5: n-Mosfet-Capacitor Example 73
Figure 6-6: Simple Bootstrap Circuit 76

Chapter One

Introduction

Advances in integrated circuit technology in the last decade have presented circuit designers with new problems and challenges not dealt with previously. Interest in the potential of VLSI (Very Large Scale Integration) has grown tremendously in recent years, both in the academic and industrial settings. The ability to place large systems on a single piece of silicon has made more advanced systems realizable. However, this increase in capability brings with it new problems for the designer.

One major problem is how to manage the design and analysis of systems with a large number of components. Potential solutions to this problem include developing methodologies which exploit regularity and hierarchy during the design phase, and computer-aided design tools which take over some of the lower level tasks of design and analysis. Much of the design takes place at an abstract level where, for example, the mosfet is modeled as a charge controlled switch. This allows the designer to ignore unnecessary detail.

An equally important problem arises when designing digital circuits, such as superbuffers, bootstrap clockdrivers, and memory sense amplifiers, which must meet tight performance criteria. These circuits sit on the boundary between the analog and the digital domains; although they are used in digital systems, they must be viewed by the designer as analog circuits in order to meet speed, power and voltage level requirements. These circuits usually consist of a small set of components; however, *the models necessary to analyze the circuit's performance are complex.* The switch level mosfet model used to analyze the digital behavior of circuits is not adequate to describe such characteristics as switching speed, power dissipation, gain, capacitive coupling, or noise immunity, and so the designer must use analog device models to analyze the circuit's behavior.

Few tools are currently available to designers for analyzing the electrical characteristics of high performance digital circuits. Existing tools are primarily electrical simulators which use numerical techniques to produce a set of waveforms showing the circuit's quantitative behavior during successive increments of time in response to a set of inputs.[20] These systems can only provide quantitative answers about what the circuit is doing. The circuit analyst is responsible for using this and other information to provide answers to questions like:

* How does this circuit work?

* Why didn't the circuit behave as I expected?

* Which device parameters should I change to make it work?

* Which parameters should I change to increase the circuit's performance (speed, power, voltage thresholds, etc.)?

The knowledge necessary to answer these and similar questions is usually classified under *designer's intuition.* To produce systems which will assist the designer in answering these questions, we must develop a theory which captures some of these intuitions.

1.1 An Explanation for a Complex Bootstrap Driver

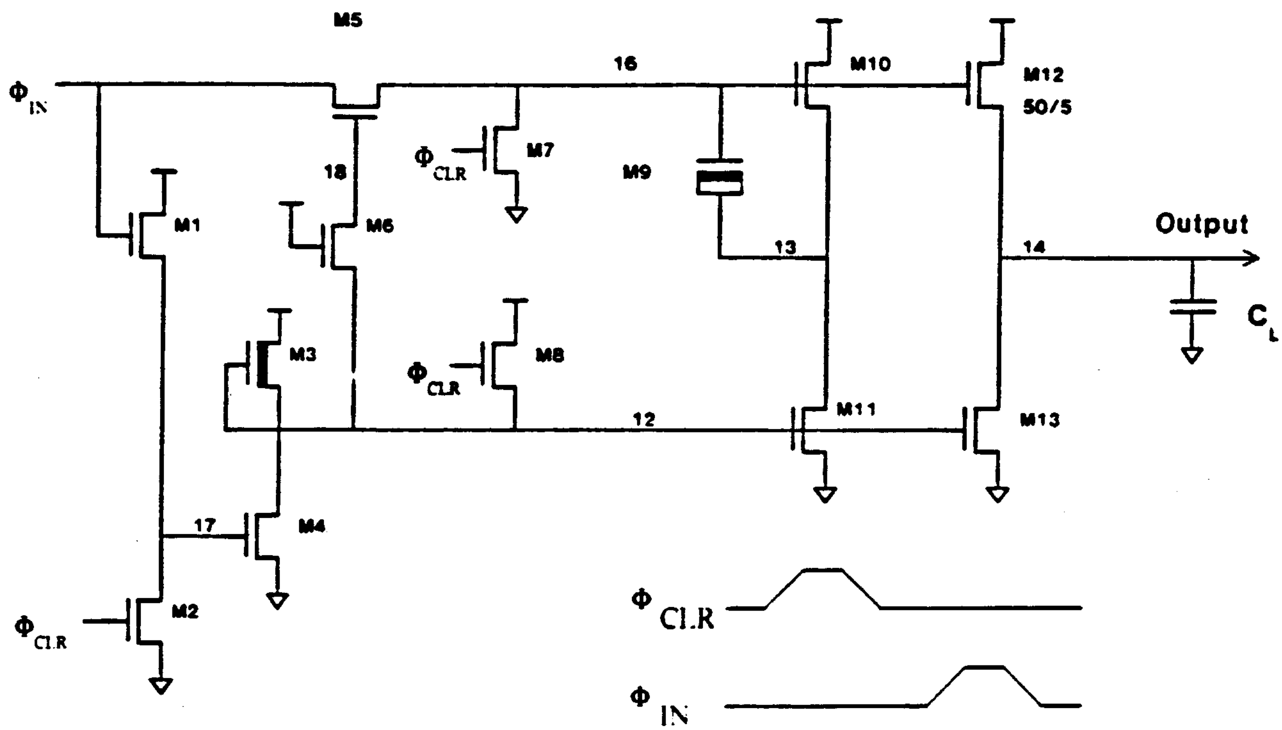


Figure 1-1: Bootstrap Driver

We can examine the type of reasoning involved in describing a circuit's behavior by looking at explanations of circuits which appear in journal articles. The following is a description of a 5-volt bootstrap driver modeled after one used in the INTEL 2118 16K dynamic RAM [4] (Figure 1-1). The reader is not expected to understand the details of this explanation but should just get a feel for the style of reasoning involved. Most of the remaining examples in this paper are very simple and require only a rudimentary knowledge of electronics.

> Initially Φ_{CLR} produces a step which precharges the gates of M11 and M13 high, turning them on and holding the output at ground. When Φ_{IN} starts to rise, it charges capacitor M9 through M5 and starts to turn M10 and M12 on. M6 isolates node 18, which allows that node to bootstrap with M5's gate capacitance, keeping M5 turned on hard. M1 and M4 form a comparator that notices when Φ_{IN} has gone above 2 threshold drops. When this happens, M4 turns on and pulls nodes 12 and 18 down to ground. When node 18 discharges, M5 turns off, isolating node 16. Also, when node 12 discharges, M11 and M13, which had been holding down nodes 13 and 14, now turn off, letting those nodes rise. Capacitor M9 then bootstraps node 16 (which was isolated by M5 when M5's gate fell), turning M10 and M12 on hard. M12 pulls the output node voltage up. Φ_{IN} can now fall without affecting the rest of the circuit because M5 is off. When Φ_{CLR} rises again, M11 and M13 turn on and M10 and M12 turn off, forcing the output low and resetting the circuit. The bootstrapping capacitor M9 is driven from node 13 and not from node 14 to get more gate drive on M12 which significantly improves the output rise time.

Much can be learned by examining this and similar explanations. The terms used in the explanation are primarily qualitative; when a quantity is used, it is usually a symbolic quantity such as "2 threshold drops" ($2V_{Th}$), "high" (V_{DD}), or "ground." The behavior of circuit state variables is described primarily in terms of their rate of change; phrases like "rises," "fall," "starts to rise" and "discharges" are commonly used. In addition, terms like "high" and "low" are used to describe a region in which the circuit state variable lies. High usually identifies the region above V_{DD} minus a noise margin ($x > V_{DD} - NM_H$), while low identifies a region between the threshold voltage minus a noise margin and ground ($V_{Th} - NM_L > x > 0$). The behavior of each device is described in terms of its current region of operation and its movement between these regions (e.g., "on," "off," "turns off," "starts to turn on," and "resetting"). From these phrases we see that the notion of *qualitative analysis* plays a very important role in analyzing circuit behavior.

Many of the qualitative phrases above are connected with words like "affecting," "turning," "causing," "pulls" and "holding." The use of these words gives the explanation a strong sense of cause and effect. Except for the initial inputs, all changes of circuit state variables are described in terms of other circuit state variables which caused the change. We even notice phrases such as ". . . can now fall without affecting . . .," where a lack of effect is made explicit. Thus the notion of *causality* also plays an important role in circuit analysis.

In addition, most of the explanation is given in terms of *local interactions* between devices which are physically connected together. Furthermore, the description of the circuit's overall behavior is inferred from the behavior of each individual component and the way in which they are interconnected. This is a common notion of circuit analysis, which is captured formally by the network laws and device models of circuit theory.

The explanation also uses a number of *cliches*, each of which has a special meaning commonly understood by the designer and the readers. These cliches serve two purposes: to concisely refer to a complex behavior and to assign an intended purpose (teleology) to a set of one or more devices. A cliche like "precharges" conjures up in the mind of the reader a complex set of events whereby turning on a device will charge a particular node up to V_{DD} or some other understood value. A cliche like "bootstrapping capacitor" implies that the function of the capacitor will be to appear as a fixed voltage source during certain periods of circuit operation. This is achieved by isolating one of the capacitor's terminals from the rest of the circuit, causing the voltage at a node connected to that terminal of the capacitor to follow any changes in the voltage of the node connected to the other terminal of the capacitor.

Finally, certain statements reflect particular design decisions and assign purpose to certain devices; for example, "the bootstrapping capacitor M9 is driven from node 13 and not from node 14 to get more gate drive on M12, which significantly improves the output rise time." Notice that node 13 doesn't have nearly as much capacitance as node 14, which is connected to C_L. Thus, node 13 will rise much faster than node 14. By connecting the bootstrap capacitor to node 13, rather than node 14, node 16 rises faster providing M12 with a strong drive quickly.

The explanation also makes a number of implicit assumptions and leaves out parts of the explanation which, hopefully, are obvious to the audience to whom the explanation is addressed. A system which analyzes circuits must be able to make these assumptions explicit. This is important in understanding the limitations of the analysis technique and understanding where exactly these limitations arise. Furthermore, the generality of the components of an analysis technique depends on the specific assumptions made for that component.

A number of properties have been identified above which are important in reasoning about circuits. *Qualitative analysis* is an approach to capturing these properties.

1.2 Temporal Qualitative Analysis: an Overview

The remainder of this paper describes *Temporal Qualitative (TQ) Analysis*, a technique for analyzing MOS circuits whose important behavior straddles the boundary between the analog and the digital domains. TQ Analysis describes the causal qualitative behavior of a circuit in response to an input over "elapsed" time, where time is viewed as a set of intervals in which devices move through different operating regions. A major objective of this work is to show a close link between the intuitions of expert circuit analysts and the formal theories of calculus, circuit theory and feedback analysis.

The analysis of electrical systems involves two steps:

1. Developing models for electrical devices which accurately model their physical behavior.

2. Predicting the behavior of systems which obey these models.

Circuit theory is only concerned with the second step and assumes that the models provided are sound. In this paper a similar assumption is made, although particular properties of the analysis (e.g., continuity) will constrain the models used.

Many of the ideas in TQ Analysis have evolved from work by de Kleer on the causal qualitative analysis of bipolar analog circuits. [7] De Kleer's PhD work concentrated mostly on the Incremental Qualitative (IQ) Analysis of devices within a single operating region. TQ Analysis differs from de Kleer's work in two important ways:

* Because our interest is in analyzing the **electrical performance** of **digital** circuits we must be able to describe the circuit's large signal behavior. This involves providing a mechanism for determining how devices move between operating regions (*transitions*), as well as describing their incremental behavior. Although de Kleer provided a mechanism for recognizing transitions, this was not central to his thesis and is inadequate for the types of circuits which we would like to analyze.

* Unlike the analog bipolar domain, an understanding of **charge flow** and **capacitive storage** is essential in the analysis of digital MOS circuits. A mechanism is provided for representing "capacitive memory," which is based on the continuity of electrical quantities (e.g., charge) over time.

Each of the remaining sections of this paper will describe a major conceptual component of TQ Analysis.

Chapter two provides the basic definitions for qualitative representations of electrical networks, time, state variables, and device operating regions. Each of these representations is based on the notion that a qualitative representation consists of a network of open regions, separated by boundaries. These boundaries are chosen at the appropriate level of detail for the analysis.

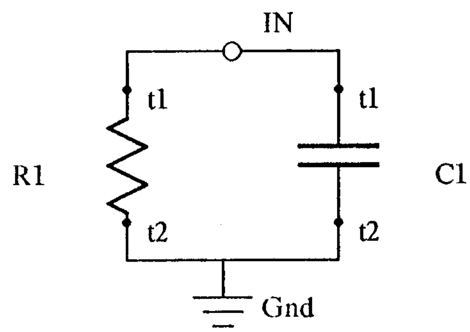


Figure 1-2:RC Example

Chapters three and four discuss the two basic types of reasoning involved in TQ Analysis: *Causal Propagation* and *Transition Analysis* [27]. These two sections will be illustrated by a parallel RC circuit (figure 1-2) which exhibits the following behavior:

> We will assume that at instant t=0 the voltage across the capacitor is positive ([$V_{IN,GND}$] = +).[1] This causes the voltage across the resistor to be positive, producing a positive current through the resistor, which begins to discharge the capacitor, decreasing $V_{IN,GND}$. $V_{IN,GND}$ decreases for an interval of time and eventually reaches zero.[2] At this point the current stops flowing and the circuit has reached a steady state[3] at zero volts.

This description is marked by a series of events such as $V_{IN,GND}$ being initially positive or $V_{IN,GND}$ moving to zero, which break the description into a series of time intervals. Two types of reasoning are required to analyze the circuit during each interval.

One type of reasoning involves determining the instantaneous response of the circuit to a set of primary causes which mark the event; for example, "a positive voltage across the resistor, produces a

---

[1]The notation [$V_{IN,GND}$] indicates the sign of the voltage from the node IN to GND.

[2]Since $V_{IN,GND}$ is a decaying exponential it is positive for $t < \infty$ and reaches zero at ∞.

[3]By steady state we mean that all the voltages and currents in the circuit are constant.

positive current through the resistor . . ." and so on. The mechanism corresponding to this type of reasoning in TQ Analysis is Causal Propagation and is described in chapter three. Causal Propagation occurs at the start of a time interval when a set of qualitative inputs are propagated forward, using a set of causal relations to determine their instantaneous effect on other circuit quantities. This may be viewed as a qualitative small signal analysis. The mechanism in this section is similar to de Kleer's Incremental Qualitative Analysis.

The second type of reasoning determines the long term effects of these qualitative inputs; for example, "$V_{IN,GND}$ decreases for an interval of time and eventually reaches zero." This type of reasoning is modeled by Transition Analysis and is described in chapter four. Transition Analysis determines whether or not a quantity will move between two regions of interest (e.g., moving from positive to zero or saturation to cutoff) at the end of a time interval. This analysis is based on the assumption that real systems are modeled by *continuous functions* and builds on a few simple theorems of calculus. Transition Analysis may be viewed as a qualitative large signal analysis.

An understanding of both positive and negative feedback is essential to understanding digital circuits and restoring logic. Chapter five examines the qualitative properties of feedback and discusses a mechanism for recognizing and dealing with feedback in general.[4]

The examples used in the first five chapters consist only of simple RC networks. In chapter six the above mechanism is extended to describe the behavior of devices with multiple operating regions. A qualitative model is created for an enhancement mode n-channel mosfet and is used to generate an explanation for a simple mosfet circuit.

Chapter seven concludes the paper with a discussion of the material presented, comparing it to other qualitative systems, pointing out its limitations, and suggesting directions for future work.

---

[4]Our feedback mechanism was inspired by a set of feedback heuristics presented in de Kleer's PhD thesis and is a generalization of these concepts.

Chapter Two

Qualitative Values

People use a variety of terms which are considered qualitative. Some examples are positive, negative, increasing, decreasing, forward, saturation, yesterday, tomorrow, office and mosfet. What all of these have in common is that they are regions over time, space, or some other set of quantities which the explainer considers "interesting." In addition, these regions, which we call *qualitative regions*, are separated by boundaries. For example, zero is a boundary between plus and minus, midnight is a boundary between today and tomorrow, and the office walls define the space which is called the office.

In Temporal Qualitative Analysis, the space of values, which the quantity of interest can take on is broken into a set of open intervals or regions separated by a set of boundaries. These boundaries are chosen at the appropriate level of detail for the analysis. For a particular domain, the construction of a set of qualitative representations may be viewed as a mapping between continuous functions to functions of discrete intervals. In circuit analysis we are mapping from the continuous equations of device physics to a set of qualitative relations in circuit theory. The following sections describe the qualitative representation used in TQ Analysis for space, time, state variables (e.g., voltage and current), and device behavior. For each section, we make the case that the qualitative representation consists of a set of open regions separated by boundaries. Many of these representations exist in formal circuit theory, while others are implied through common usage.

2.1 The Network Model

An integrated circuit is implemented as a semiconductor wafer with different ions diffused or implanted into its surface. The equations necessary to describe a complex circuit at the device physics level are not easily solvable and, more importantly, would not provide the designer with much insight into the overall circuit behavior.

In circuit theory, the complexity of these equations is reduced by modeling a region of space with a uniform electrical behavior as a single lumped element (e.g., a resistor, capacitor or mosfet). Each element has a set of terminals which allows it to interact with other elements. The electrical behavior of a element is described by a set of *constitutive* relations between state variables associated with the element's terminals.

An electrical circuit is described as a set of elements connected together in a network. Each member of a set of locally interacting elements has a terminal connected to a common node. The interactions between elements are described by a set of network laws known as Kirchoff's Voltage and Current Laws. These network laws only hold as long as the circuit is small enough that electromagnetic waves propagate across it instantaneously, that is:

```
                      d << c*dt

    where:

      d  is the largest distance across the circuit
      dt is the smallest time interval of interest
      c  is the speed of light
```

This is known as the Lumped Circuit Approximation, and is important later in our discussion of causality.

A network is described by a set of devices, nodes, and connections between the two. The field **Devices** specifies the name of of each component in the circuit, along with its corresponding type (e.g., resistor(R1) means that the component R1 is a resistor). **Nodes** is a list of node names used in the circuit. Finally, **Connections** consist of a set of assertions, each of which specifies the device terminals connected to a specific node. For example, if terminal one of C1 (denoted t1(C1)) and t2(R1) are connected to node IN, then this is specified as: connect (IN t1(C1) t2(R1)).

The following is a specification for the parallel RC network in figure 1-2:

```
Network: Parallel RC
  Devices: resistor(R1), capacitor(C1)
  Nodes: IN, GND
  Connections:
    connect(IN t1(R1) t1(C1))
    connect(GND t2(R1) t2(C1))
```

2.2 Time

We represent time as a linear, non-overlapping sequence of alternating instants[5] and open intervals. The duration of each interval is determined by Transition Analysis. During a single time interval, all quantities of interest have a single qualitative value. In other words, each quantity lies within a single qualitative region throughout the duration of an interval. Using this representation, quantities can only interact if they are spatially local and they occur during the same time interval. A linear representation of time has been chosen for simplicity; however, none of the concepts presented in this paper depend strongly on this representation being a linear sequence.

2.3 State Variables

The representation we choose for describing electrical quantities depends on which type of circuit and which properties we are interested in. For example, if we are interested in verifying a circuit's behavior at the digital abstraction level, we might want to segment the range of input and output voltages into the following set of regions:

```
     Valid      Noise       Forbidden      Noise      Valid
      "0"      Margin         Zone        Margin       "1"
      Out       Low                        High        Out
<-----------|------|-----------------|------|----------->

    Qualitative representation of a digital signal
```

To look at the analog performance of digital circuits, the components must be viewed as analog devices. At this level, a state variable is described in terms of its sign and the sign of its derivatives;[6] for example, the voltage is positive or the current is decreasing. Sign separates the real number line into two open intervals, positive and negative, with a boundary at zero. The sign of a quantity, A, will be denoted by [A] and the sign of the quantities nth derivative by $[d^nA/dt^n]$.

---

[5] An instant is a closed interval with zero duration.

[6] Unless otherwise stated, all derivatives discussed in this paper are partials with respect to time.

```
                         A = 0
                           |
-inf <---------------------|----------------------> +inf
                           |
          NEGATIVE        ZERO        POSITIVE
            (-)           (0)           (+)
```

Qualitative representation of a real number A

2.4 Qualitative Expressions

To describe the network laws and the element relations of circuit theory, a qualitative algebra in terms of the signs of quantities is necessary. The arithmetic operations necessary for modeling the MOS domain are negation, addition, and multiplication.[7] Tables describing these operations are shown below:

Addition: [A] + [B]

[A]\[B]	+	0	-
+	+	+	?
0	+	0	-
-	?	-	-

Multiplication: [A] x [B]

[A]\[B]	+	0	-
+	+	0	-
0	0	0	0
-	-	0	+

also ? x 0 = 0

Negation:

[A]	-[A]
+	-
0	0
-	+

where:

+ = positive
0 = zero
- = negative
? = ambiguous

In the table for addition the symbol (?) means that the result of the sum is *ambiguous*. The sign of the sum cannot be determined without additional information. Techniques for resolving this ambiguity are discussed in later sections. Also notice in the multiplication table that the product of two quantities can be deduced even when one quantity is ambiguous, as long as the other quantity is zero. More complex arithmetic operations, such as subtraction, summation, and exponentiation by a

[7] The qualitative arithmetic used here is similar to the one used in de Kleer and Brown's Qual and Envision systems and Forbus' Qualitative Process Theory. [7, 14]

positive integer, can be constructed from these basic arithmetic operations. Finally, it is important to note that the variables participating in a qualitative expression are not limited to quantities and their first derivatives, but may include a mixture of second and higher order derivatives as well.

2.5 Operating Regions

Often non-linear devices, such as diodes, bipolar transistors and mosfets cannot be described by a single set of relations without making these relations overly complex. Instead we consider the behavior of the device in each of several distinct operating regions, each described by a different set of equations. The mechanism for representing operating regions and describing how quantities move between operating regions is presented in chapter six.

2.6 Summary

In TQ Analysis, a qualitative representation consists of a network of open regions, separated by boundaries which are chosen at the appropriate level of detail for the analysis. The following table summarizes the qualitative representations used in this paper for circuit analysis:

Representation	open region	boundary
space	lumped elements	nodes
time	open intervals	instants
state variables	positive,negative	zero
relations	operating regions	edge of op. regions

Chapter Three

Causal Propagation

Causal Propagation is a technique which uses knowledge of circuit theory and qualitative arithmetic to describe the behavior of a circuit during an instant or interval of time; it may be viewed as a qualitative small signal analysis. In circuit analysis, the values of the state variables of a network can be determined at some instant of time from the network laws, each device's behavior, and the initial conditions. Using the qualitative quantities and expressions described in the previous chapter, we can perform a similar analysis at a qualitative level. Returning to the RC example (Figure 1-2), a set of qualitative relations for the network is:

$[V_{IN,GND}]$	$= [I_{t1(R1)}]$	Ohm's Law
$-[I_{t1(R1)}]$	$= [I_{t1(C1)}]$	Kirchoff's Current Law
$[I_{t1(C1)}]$	$= [dV_{IN,GND}/dt]$	Capacitor Law

and the initial condition is:

$$[V_{IN,GND}] = +$$

From these relations and the initial condition we can deduce, for example, that $[I_{t1(R1)}] = +$, $[I_{t1(C1)}] = -$ and $[dV_{IN,GND}/dt] = -$. The qualitative equations act as a set of constraints on the electrical quantities; as long as $[V_{IN,GND}]$ remains positive the other electrical quantities are constrained to be the values shown above. Furthermore, during any time interval, all qualitative quantities must be single valued, i.e., a quantity cannot move between qualitative regions or boundaries during the interval.

The qualitative model described thus far tells us what each qualitative value is, but does not explain how they came about. The qualitative description given by an engineer for the RC circuit (figure 1-2) gives a *causal* account of the circuit behavior. For example, when $V_{IN,GND}$ becomes positive, this causes a positive current through the resistor, discharging the capacitor and causing $V_{IN,GND}$ to decrease. This causality is not provided in a circuit theory model. Where then does this causality come from? The answer lies in the assumptions made in modeling a circuit as a network of lumped elements.

3.1 Causality and the Lumped Circuit Approximation

The Lumped Circuit Approximation (section 2.1) states that the network laws hold only as long as the smallest time interval of interest is sufficiently large that electromagnetic waves appear to propagate across the circuit instantaneously. When generating a causal description, the analyst breaks this assumption by viewing the circuit behavior at a time scale close to the speed of light. On this time scale, for example, there is a finite delay between the time an electromagnetic wave enters one end of a wire and the time the wave exits the other end. One can imagine the wire as a pipe connected to a faucet. When the faucet is turned, on there is a short delay before the water comes out of the pipe. Only after the water has begun to come out of the pipe and the flow has *stabilized*, can we say that the flow rate out of the pipe equals the flow rate out of the faucet. Similarly, once the current into the wire has had time to stabilize, we can say that the current out of the wire equals the current into it.

In general, the analyst uses multiple viewpoints in describing the behavior of a circuit. A microscopic, electrodynamic level model is used initially to describe the effects of a set of changes on the system.[8] Once these effects have stabilized the analyst moves to a macroscopic viewpoint (i.e., the network model), where these effects propagate instantaneously. Using the macroscopic viewpoint, the set of initial changes provides a set of constraints on the rest of the system.

One way of modeling these two different viewpoints is to provide two sets of models, one which describes the electrodynamic behavior of devices in terms of Maxwell's equations and a second set which describes the circuit level behavior in terms of the constraints established by network theory. The former model, however, is both intractable and undesirable. The primary reason for using a circuit model, in the first place, is to avoid the detail and number of interactions which occur in the electrodynamic model. To then reintroduce such a model would be counter productive. Furthermore, supposing we could produce an electrodynamic model, we are still faced with the problem of assigning a causal ordering to the events which occur at the electrodynamics level. Of course, we could produce another model at an even lower level (e.g., quantum physics); however, this simply pushes the problem away one level and doesn't solve it.

The solution which Causal Propagation takes is to build in the intuition, which a designer has about causality from the electrodynamic level into the network level. This is done by initially imposing a

---

[8] de Kleer and Brown refer to this microscopic time scale as "mythical time" since, from the macroscopic view of network theory, this time scale doesn't exist.

sense of causality on the device relations and network laws in response to a set of changes; these will be referred to as *causal relations*. For example, if we know that [V] = [I] by Ohm's Law, and the voltage across the resistor has begun to increase (i.e., [dV/dt] = +), then we say that an increase in V *causes* an increase in I by Ohm's Law.[9] If instead the current through the resistor was changing, then we would say that the change in current causes a change in voltage by Ohm's Law. In the above example, Ohm's Law is a *bidirectional* relation, since the causality can run in either direction. In general, if all but one qualitative value is known in a bidirectional relation, the known values are used to determine their effect on the unknown quantity.

In a few instances the engineer views this causality as occurring in only one direction. (e.g., as we will see in chapter 6, an engineer will say that a voltage across the gate of a mosfet produces a current through the device's channel but not the converse.) In this case we refer to the relation as *unidirectional* and indicate the direction of the causality by replacing = with an arrow (→) pointing from the cause(s) to the effect.[10] Once the circuit has had time to stabilize, the causal relations revert to a set of *constraints* between state variables without imposing a causal ordering. For example, Ohm's Law becomes a constraint between the current and voltage of the resistor and we say that V and I are constrained to be positive by the input.

In TQ Analysis, the beginning of a time interval is marked by the *transition* of one or more quantities from one qualitative value to another. The transitioning quantities are referred to as the *primary causes* for that interval.[11] The microscopic viewpoint is used to determine the effects of these primary causes, i.e., to determine the qualitative value for each state variable in the circuit at the beginning of the interval. Moving to the macroscopic viewpoint, these quantities are then constrained by the

---

[9] Where "A causes B" means that B is functionally dependent on A, i.e., the value that B has is caused by the value that A has.

[10] A bidirectional relation (R) is implemented as a set of unidirectional relations, where each unidirectional relation contains one of R's quantities as an effect and the rest of R's quantities as the cause. For example:

```
[I1] + [I2] + [I3] = 0
```

is equivalent to:

```
[I1] + [I2] → - [I3]
[I1] + [I3] → - [I2]   and
[I2] + [I3] → - [I1]
```

During the Causal Propagation phase of analysis for a particular interval, at most one of a set of unidirectional relations will be used.

[11] Each primary cause is either an externally driven input or the independent variable of a memory element (e.g., the voltage across a capacitor).

qualitative network relations for that time interval. The end of the interval is marked when one or more quantities transition, creating a new set of primary causes for the next interval of time.[12]

3.2 Implementing Causal Propagation

Causal Propagation is implemented as a set of assertions and rules in AMORD [6], a rule-based inference engine with a truth maintenance system. An assertion consists of a fact statement, which is an arbitrary Lisp expression, and a supporting justification. The justification is a reason for the fact to be true, along with a set of facts which support this reason. A rule is composed of a pattern and a rule body which consists of arbitrary Lisp code. The rule body is run whenever a set of assertions is found which matches the rule pattern.

Each unidirectional causal relation is implemented as an AMORD rule . The rule pattern consists of the quantities (i.e., "causes") in the relation's qualitative expression; the expression is evaluated as Lisp code in the rule body, and the effect is an assertion made by the rule body. Each assertion is recorded with a justification describing its cause.

Causal Propagation begins at the start of a time interval by asserting a set of primary causes with their corresponding qualitative values. Rules function in a daemon-like manner. When all of a rule's patterns are matched with a set of assertions, the rule body is run, possibly creating more assertions. Causal Propagation terminates when all of the relevant rules have fired.[13]

3.3 Domain Knowledge

In this section the domain specific knowledge for analyzing electrical networks is discussed. This knowledge is broken into two parts: network laws and device models. The network laws describe how current and voltage quantities of connected devices interact, while the device models describe the behavior of a device via the voltages and currents associated with its terminals. Current is measured going *into* a device's terminals and voltage is measured between network nodes.

---

[12] The duration of the interval is irrelevant to Causal Propagation. The interval may last to infinity or for only an instant. The duration of an interval and the set of transitioning quantities which mark the end of the interval is determined through Transition Analysis.

[13] This is not a complete constraint satisfaction system since some relations can only fire in one direction.

3.3.1 Network Laws

Kirchoff's Voltage and Current Laws describe the network behavior of electrical circuits. Kirchoff's Current Law (KCL) states: the sum of the currents out of a Gaussian surface is zero. If the Gaussian surface is put around a single node, then the sum of the currents out of the node is zero. The *qualitative* KCL rule says that the *signs* of the currents out of a node must sum to zero. Intuitively, this means that a node cannot source or sink current; that is, each node must contain at least one current flowing into the node and one current flowing out (except when all the currents for that node are zero). In addition, the KCL rule says that the signs of the ith derivative of the currents out of a node also sum to zero.

The qualitative KCL rule is shown below. This rule[14] consists of a set of *preconditions* and a set of *relations*. The preconditions are a mix of patterns for assertions which must exist and conditions which must be true in order for the rule to be applicable. KCL has one precondition: an assertion must exist which specifies all of the terminals connected to a particular node. The relations section consists of causal relations as described above.

```
Law: n-Terminal KCL
    Preconditions:
        connect(Node-1 T1 T2 . . . Tn )
    Relations:
        0 = [I_T1] + [I_T2] + . . . + [I_Tn]
        0 = [d^nI_T1/dt^n] + [d^nI_T2/dt^n] + . . . + [d^nI_Tn/dt^n]
```

Kirchoff's Voltage Law (KVL) states that the sum of the branch voltages around any loop in the network graph is zero. The qualitative KVL rule states that the sum of the *signs* of the branch voltages around a loop is zero. For loops containing two nodes, the KVL rule (Voltage Negation) is equivalent to saying that voltage is path independent. For loops containing three nodes, the KVL rule states that the voltage between two nodes is the sum of the voltages between each of the two nodes and an intermediate node.

---

[14] TQ Analysis considers the words Law and Model to be synonymous with Rule and these words are used purely for documentation purposes.

```
Law: Voltage Negation
    Relations:
        [V_{N1,N2}] = - [V_{N2,N1}]
        [d^nV_{N1,N2}/dt^n] = - [d^nV_{N2,N1}/dt^n]

Law: Three Node KVL
    Preconditions:
        When N1 ≠ N2 ≠ N3
    Relations:
        [V_{N1,N2}] + [V_{N2,N3}] = [V_{N1, N3}]
        [d^nV_{N1,N2}/dt^n] + [d^nV_{N2,N3}/dt^n] = [d^nV_{N1, N3}/dt^n]
```

3.3.2 Device Models

In this section we will first discuss a general property of network elements and then present the specific device models for some simple network elements.

3.3.2.1 KCL Applied to Devices

Above KCL was stated as "The sum of the currents out of a Gaussian surface is zero." By placing a Gaussian surface around a device this becomes: The sum of the currents into a device is zero. Qualitatively this means that no device can source or sink current. All of the devices we are interested in have either two or three terminals. KCL for a three terminal device is shown below:

```
Law: Three Terminal Device KCL
    Preconditions:
        three-terminal-device(D)
    Relations:
        0 = [I_{t1(D)}] + [I_{t2(D)}] + [I_{t3(D)}]
        0 = [d^nI_{t1(D)}/dt^n] + [d^nI_{t2(D)}/dt^n] + [d^nI_{t3(D)}/dt^n]
```

3.3.2.2 Network Elements

The basic two terminal elements are resistors, capacitors and inductors. The constitutive relation for each of these elements is $V = IR$, $I = CdV/dt$ and $V = LdI/dt$, respectively. A circuit designer views these relations as being bidirectional; that is, a change in voltage will produce a change in current and vice-versa. The models for the resistor and capacitor are shown below. The fields

Terminals and **Corresponding Nodes** list the device's terminals and the nodes connected to these terminals, respectively. These models have an additional field called **Assertions** consisting of facts which are asserted as a result of running the rule.

```
Model: Resistor(R)
    Terminals: T1 T2
    Corresponding Nodes: N1 N2
    Relations:
```

$$[V_{N1,N2}] = [I_{T1(R)}]$$

$$[d^nV_{N1,N2}/dt^n] = [d^nI_{T1(R)}/dt^n]$$

```
    Assertions:
        two-terminal-device(R)
```

```
Model: Capacitor(C)
    Terminals: T1 T2
    Corresponding Nodes: N1 N2
    Relations:
```

$$[dV_{N1,N2}/dt] = [I_{T1(C)}]$$

$$[d^{n+1}V_{N1,N2}/dt^{n+1}] = [d^nI_{T1(C)}/dt^n]$$

```
    Assertions:
        two-terminal-device(C)
```

3.4 Example

Using the mechanism described thus far, we can determine the behavior of the parallel RC circuit (figure 1-2) for a particular instant of time. The analysis begins by inputting the network description shown in section 2.1. At Instant-0 the initial condition (and the primary cause) is $[V_{IN,GND}@\text{Instant-0}] = +$.[15] This value is asserted and the causal propagator is invoked. We can then ask the system for the qualitative value of any quantity in the network, along with a causal explanation for that quantity:[16]

---

15. The symbol @ means "at time" for example, $[A@t1] = +$ translates to "the sign of A at time t1 is positive."

16. This explanation was generated by the current implementation of TQ Analysis.

```
Explanation for FACT-24 : dV(IN,GND)/dt@Instant-0 is negative:

It was given that V(IN,GND) during Instant-0 is +.
 This causes I(1(R1)) during Instant-0 to be +,
 since from rule RESISTOR : [V12] → [I1].
  This causes I(1(C1)) during Instant-0 to be -,
  since from rule 2-T-KCL : [I2] → - [I1].
   This causes dV(IN,GND)/dt during Instant-0 to be -,
   since from rule CAPACITOR : [I1] → [dV12/dt].
```

3.5 Ambiguities

The analysis technique described thus far is not powerful enough to deduce a set of qualitative values under every condition. We have already seen one example of this ambiguity in the addition table in section 2.4. If $A + B = C$ and A and B have opposite signs, then C is ambiguous; that is, C could be positive, negative or zero.

Each ambiguity which arises in qualitative analysis can be categorized as one of three types:

* Ambiguous effect

* Simultaneity

* Unknown primary cause

First, an ambiguous effect occurs when all of the causes in a qualitative relation are known and the effect cannot be deduced. In the present system such an ambiguity only results from addition. Second, if a quantity (A) is a function of one of its effects (B) then B cannot be deduced without knowing A and A cannot be deduced without B. This cyclic behavior is commonly referred to as a simultaneity. Finally, we need a means of determining how each primary cause changes between time intervals, as they are the inputs to Causal Propagation.

Analysts use a variety of information, both qualitative and quantitative, in resolving these ambiguities. The next two chapters discuss two techniques which use qualitative information to resolve these ambiguities. The first technique, *Transition Analysis*, uses information about continuity to resolve some of these ambiguities. The second technique, *Feedback Analysis*, reasons about the structure of the causal relations to resolve ambiguities which arise from simultaneities along a feedback loop. These two techniques do not resolve all types of ambiguities and other, more quantitative techniques, are needed.

3.6 Summary

Key concepts:

* Causal Propagation models the incremental behavior of a circuit and may be viewed as a qualitative small signal analysis.

* Interactions between circuit state variables are described at two levels:

  - Circuit theory views time at a macroscopic level and describes the interactions between circuit state variables using a set of qualitative network laws and device constitutive relations.

  - Electrodynamics views time at a microscopic level and allows the designer to impose a causality on the network laws and device relations.

* Causal Propagation cannot always deduce the sign of every state variable in the circuit unambiguously. These ambiguities are dealt with by Transition Analysis and Feedback Analysis.

Chapter Four

Transition Analysis

In the previous chapter we discussed the causal qualitative relationship between different state variables over an interval of time. During a time interval it is assumed that each quantity of interest remains within a single qualitative region (e.g., "the voltage is positive" or "the mosfet is in saturation during the interval"). Causal Propagation, however, makes no predictions about if and when a quantity will move to another qualitative region. The goal of *Transition Analysis* is to make these predictions.

Causal Propagation may be viewed as a qualitative small signal analysis; similarly, Transition Analysis may be viewed as a qualitative large signal analysis. In Transition Analysis we are concerned with the way quantities move from one qualitative region to another, such as a mosfet becoming saturated or a current becoming positive and increasing. For each state variable in the circuit, Transition Analysis tries to determine whether or not it will remain in the same qualitative region or transition into another region at the end of a time interval.

As we discussed at the end of the last chapter, Causal Propagation sometimes cannot determine the qualitative value for one or more quantities during a particular time interval. When this occurs, the results of Transition Analysis can often be used to resolve the ambiguous quantity by determining how the quantity has changed (i.e., whether or not it has transitioned) between the previous and current time intervals. In the event that Transition Analysis cannot determine if a quantity has transitioned, other techniques must be used to resolve the ambiguity, such as Feedback Analysis (chapter 5).

Transition Analysis is broken into two steps: *Transition Recognition* and *Transition Ordering*. Transition Recognition attempts to determine whether or not a quantity is moving towards another qualitative region or boundary (e.g., the positive charge on the capacitor is decreasing towards zero, or a mosfet is moving from the boundary between ON and OFF to the region ON). Transition Recognition often determines that more than one quantity is moving towards another region or boundary. Transition Ordering determines which subset of these quantities will *transition* into a new region or boundary first, marking the end of that interval. This chapter only discusses transitions across zero. In chapter 6 the mechanism described here is extended to recognize transitions across boundaries other than zero.

4.1 Transition Recognition

The basic assumption underlying Transition Analysis is:

The behavior of real physical systems is continuous.[17]

More precisely, it is the functions which describe a physical system that are continuous. This is not to say that the models that an engineer uses are always continuous. For example, only the currents, voltages and their first derivatives are continuous in the Shichman-Hodges model [22] of the mosfet. However, an engineer knows that this model is only an approximation and the behavior of a mosfet in the real world is continuous and infinitely differentiable.

There are a number of simple theorems of calculus which describe the behavior of continuous functions over time intervals. In this section we discuss the intuition which these theorems provide in determining how quantities move between and within qualitative regions. These theorems are then used to derive two rules about qualitative quantities: the *Continuity Rule* and the *Integration Rule.* The first rule requires that a quantity is continuous over the interval of interest, while the second assumes that a quantity is both continuous and differentiable.

4.1.1 The Intermediate Value Theorem

When describing the behavior of some quantity over time, we need a set of rules for determining how a quantity changes from one interval or instant to the next. If, for example, a quantity is positive during some interval of time, will it be positive, zero or negative during the next interval of time?

The **Zero-crossing Principle** states that:

> If f is continuous on the closed interval $[a,b]$ and if $f(a) < 0 < f(b)$ then $f(X) = 0$ for some number X in $[a,b]$. [19]

Intuitively, this means that a continuous quantity must cross zero when moving between the positive and negative regions. In the above example, the positive quantity will be positive or zero during the

[17] Continuity: "The function f is continuous if a small change in x produces only a small change in f(x), and if we can keep the change in f(x) as small as we wish by holding the change in x sufficiently small." [19]

next interval of time, however, it cannot be negative.

The Zero-crossing Principle is a specialization of the **Intermediate Value Theorem** which states that:

> If f is continuous on the closed interval $[a,b]$ and if l is any number between $f(a)$ and $f(b)$, then there is at least one point X in $[a,b]$ for which $f(X) = l$. [19]

From this we can infer, in general, that a quantity will *always* cross a boundary when moving from one qualitative open region to another.

4.1.2 State Variables and Time

By assuming that quantities are continuous and by using the results of the Intermediate Value Theorem, a relationship can now be drawn between the representations for state variables and time. Recall that the representation for time consists of a series of *instants* separated by *open intervals*. An instant marks a quantity moving from an open region to a boundary or from a boundary to an open region. Also, recall that the range of a state variable is represented by the open regions *positive* $(0,+\infty)$ and *negative* $(-\infty,0)$ separated by the boundary *zero*, which we denote $+$, $-$ and 0, respectively. If some quantity (Q) is positive at some time instant t1 (Q@t1 $= \varepsilon$ where $\varepsilon > 0$), then there exists some finite open interval $(\varepsilon,0)$ separating Q from zero.[18]

If we assume that Q is described by a continuous function of time, then it will take some finite interval of time {(t1,t2) where t1 $\neq$ t2} to move from ε to 0, traversing the interval $(\varepsilon,0)$. Similarly, it will take a finite interval of time to move from 0 to some positive value ε (Figure 4-1). Furthermore, we can say that a quantity moving from 0 to ε will leave zero at the *beginning* of an open interval of time, arriving at ε at the end of the interval. Conversely, a quantity moving from ε to 0 will leave ε at the beginning of an open interval and arrive at 0 at the *end* of the open interval. Another way of viewing this is that a quantity will move through an open region during an open interval of time, and a quantity will remain on a boundary for some closed interval of time (possibly for only an instant).

The notion of continuity is captured with the following rule (Figure 4-2).

Continuity Rule

* If some quantity Q is positive (negative) at an instant, it will remain positive (negative) for an open interval of time immediately following that instant.

---

[18] Any two distinct points are separated by an open interval.

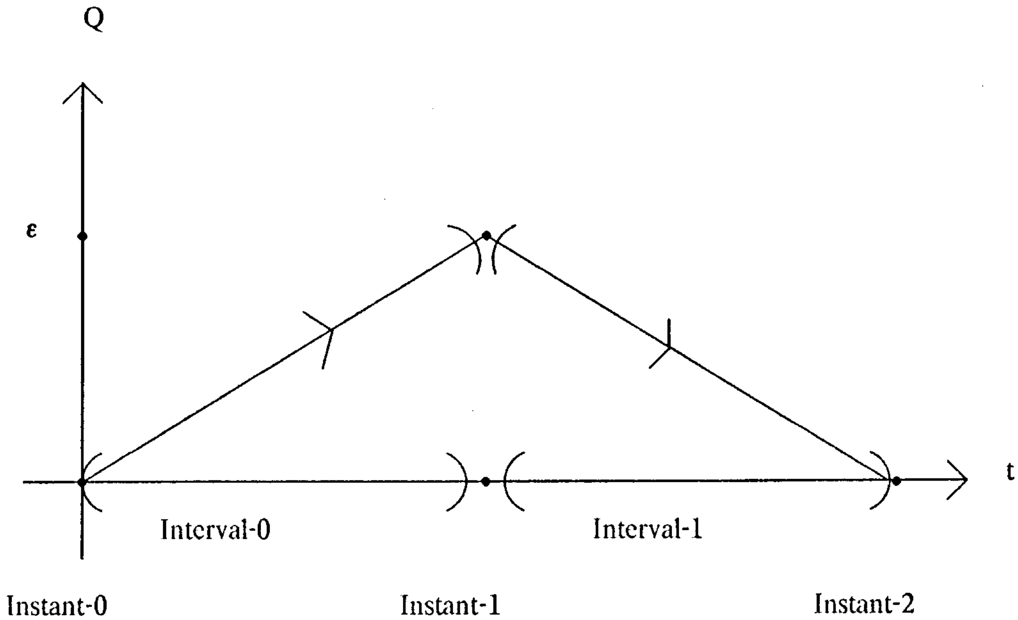


Figure 4-1:Relationship between State Variables and Time

* If some quantity Q is zero during some open interval of time, it will remain zero at the instant following the open interval.

For example, suppose that $A - B = C$, $[A] = +$ and $[B] = 0$ for some instant of time (t1). By Causal Propagation we deduce that $[C@t1] = +$. If B becomes positive for the next open interval (I1), then we cannot deduce C during I1 by the above causal relation, since the sum is ambiguous (i.e., $[C@I1] = (+) - (+) = ?$). Using the first part of the Continuity Rule, however, we predict that C remains positive during I1. This agrees with our intuition since C is the difference between A and B and we know that it will take some interval of time before B "catches" up to A (Figure 4-3).

Using the fact that a state variable will only move off of zero at the beginning of an open interval and will only arrive at zero at the end of an open interval, we can now sketch an outline of the steps involved in TQ Analysis:[19]

[19] In the actual implementation the first four steps are performed concurrently.

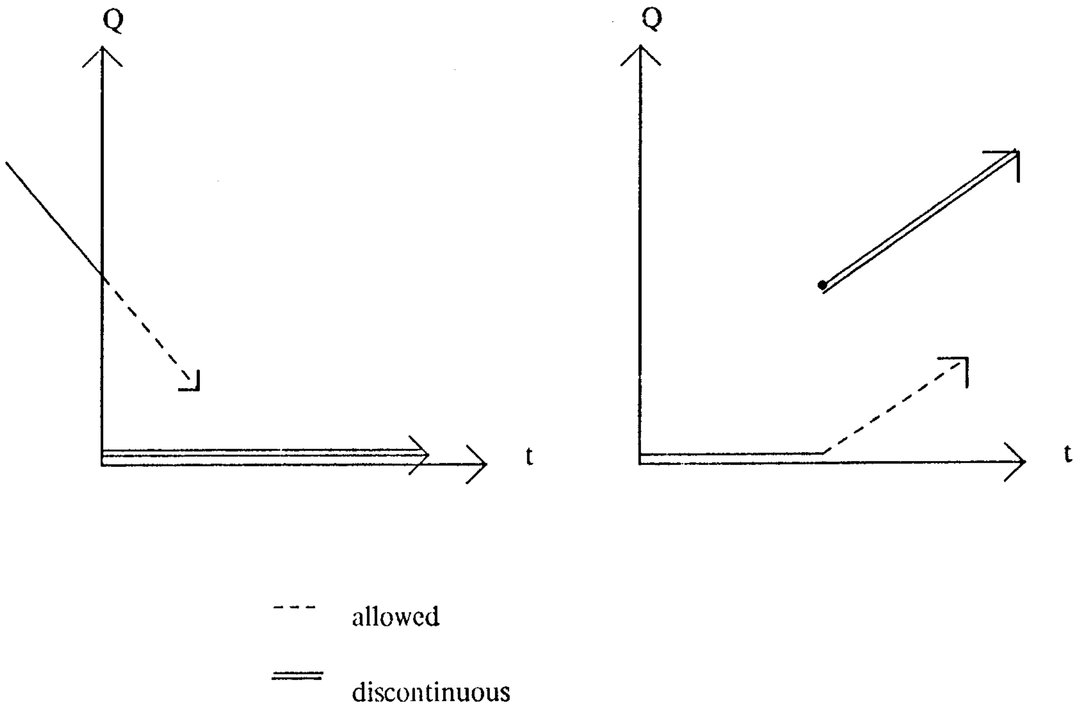


Figure 4-2:Continuity

* Given a set of primary causes for an instant, run Causal Propagation.

* Determine which quantities may transition from zero to a positive or negative region at the beginning of the next open interval.

* Use the results of Transition Analysis to determine the values of the primary causes for the next open interval of time. Run Causal Propagation for that interval.

* Use Transition Recognition to determine which quantities are moving from positive or negative towards zero.

* Use Transition Ordering to determine which quantities will transition to zero first. These transitions define the end of that interval and the beginning of the next instant.

* Repeat this process for the next time instant.

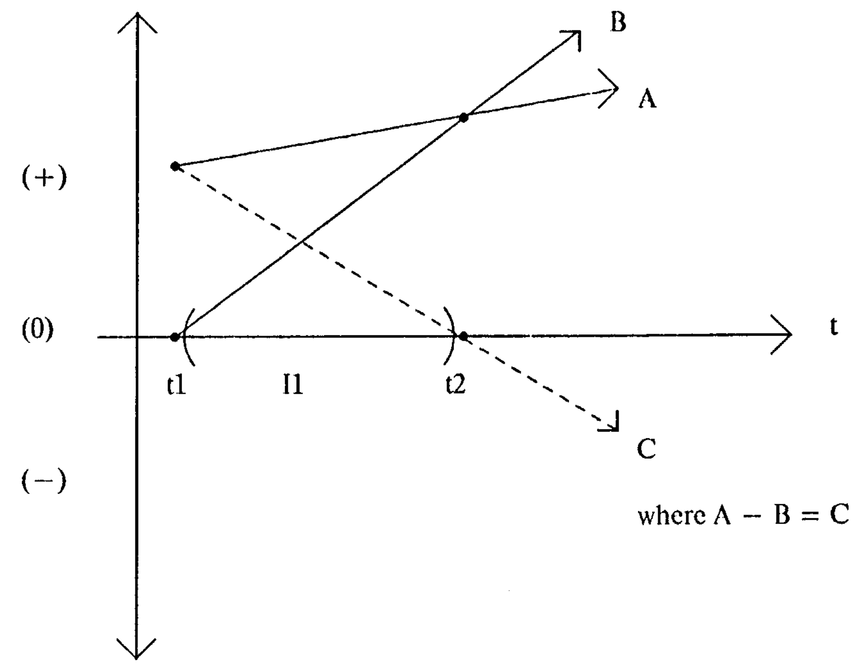


Figure 4-3:Continuity Example

4.1.3 Mean Value Theorem

In addition to looking at the continuity of quantities, information can also be derived by looking at the relationship between quantities and their derivatives. The following two corollaries of the Mean Value Theorem [26] are of particular interest to TQ Analysis:

Corollary 1:

If a function f has a derivative which is equal to zero for all values of x in an interval (a,b), then the function is constant throughout the interval.

Corollary 2:

Let f be continuous on $[a,b]$ and differentiable on (a,b). If $f'(x)$ is positive throughout (a,b), then f is an increasing function on $[a,b]$, and if $f'(x)$ is negative throughout (a,b), then f is decreasing on $[a,b]$.

By combining these two corollaries with the Intermediate Value Theorem, the behavior of a state variable is described over an interval (instant) in terms of its value during the previous instant (interval) and its derivative. At the qualitative level, this is similar to integration and is captured by the following rule (Figure 4-4):

Qualitative Integration Rule

Transitions to Zero

* If a quantity is positive and decreasing[20] (negative and increasing) over an open time interval, then it will move towards zero during that interval and possibly transition to zero at the end of the interval.

* If a quantity is positive but not decreasing[21] (negative and not increasing) over an open time interval, then it cannot transition to zero and will remain positive (negative) during the following instant.

Transitions Off Zero

* If a quantity is increasing (decreasing) during some open time interval and was zero at the previous instant, then it will be positive (negative) during the interval.

* If a quantity is constant during some open time interval and was zero at the previous instant, then it will be zero during that interval.

Note that in the first two parts of the rule the derivative of the quantity affects how it behaves at the following instant, while in the last two parts the derivative of a quantity affects that quantity during the same interval. For example, suppose that a quantity (Q) is resting at zero at some instant (t1) (i.e., [Q]@t1 = 0 and [dQ/dt]@t1 = 0). If dQ/dt becomes positive for the next open interval (I2), then it will cause Q to increase during that interval and become positive. Furthermore, Q moves off zero instantaneously, thus Q is also positive during I2. In the above case, the causal relationship between a quantity and its derivative is similar to that between two different quantities related by a qualitative expression (e.g., in a resistor a change in current instantaneously causes a change in voltage).

If we are interested in analyzing a system which includes a number of higher order derivatives, then the Integration Rule may also be applied between each derivative and the next higher order

---

[20] where "Q is decreasing" means that [dQ/dt] = −.

[21] where "Q is not decreasing" means that [dQ/dt] = + or 0.

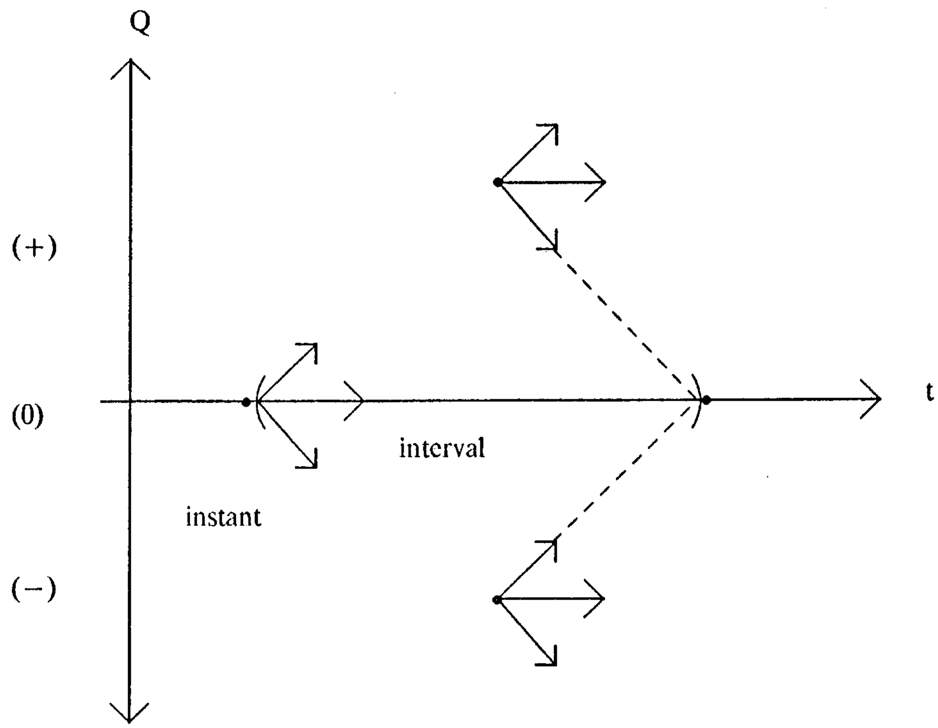


Figure 4-4: Integration Rule

derivative. For example, suppose the system being analyzed involves the position (x), velocity (v) and acceleration (a) of a mass (where dv/dt = a) and that all three quantities are constant at some instant (t1). If a becomes positive for the next open interval (I2), then it will cause an increase in v, making it positive for I2. Similarly, positive v causes an increase in x, making it positive for I2. Thus the Integration Rule uses the relation between each quantity and its derivative to locally propagate the effects of changes along a chain from higher order derivatives down towards the lower order derivatives.

deKleer and Bobrow [12] suggest an alternate formulation of the last two parts of the Integration Rule which, for example, says that: When a quantity (Q) is zero at some instant (t1), if all of its derivatives are zero at t1, then Q will remain zero during the following interval (I1), otherwise [Q] = [dQ^n/dt^n] during I1 (where [dQ^n/dt^n] is the first non-zero derivative). This formulation has a number of problems. First, it is over restrictive since it requires each quantity and all of its higher order derivatives to be continuous. This restriction rarely holds when modeling non-linear systems, such as MOS circuits, where a device model usually contains a discontinuity in at least one level of derivative. Thus the formulation is inadequate for many complex systems. The Integration rule only requires that a quantity is continuous and differentiable, making the rule applicable for a wider class of systems. Second, their formulation is non-local in the sense that it looks at the relationship between a quantity and *all* of it higher order derivatives to determine the behavior of that quantity. The Integration rule only looks at the relationship between a quantity and its *first derivative*, allowing changes in higher order derivatives to propagate up locally towards the lower order derivatives. Finally, deKleer and Bobrow's formulation can produce a description consisting of a sequence of instants which are not separated by open intervals, thus their model of time is not dense.

As we have seen above, the Integration Rule describes the direction a quantity is moving with respect to zero (e.g., towards or away from zero). Recall that if a quantity is zero and increasing or decreasing during the next interval, then the quantity must transition from zero. If, however, a quantity (A) is moving towards zero for some interval of time, it may or may not reach zero by the end of the interval. Suppose some other quantity (B) reaches zero first and B causes dA/dt to become zero, then A will not reach zero. Thus we need a mechanism for determining which quantity or set of quantities will reach zero first during an open interval of time. One mechanism for doing this is called *Transition Ordering* and is described in the following section.

4.2 Transition Ordering

As a result of Transition Recognition we have divided the set of all quantities into 1) those which may transition (they are moving monotonically towards zero) 2) those which can't transition (they are constant or moving monotonically away from zero) and 3) those whose status is unknown (their direction is unknown or is not monotonic).

Next we want to determine which subsets of these quantities can transition by eliminating those transitions which lead to 1) quantities which are inconsistent with the set of qualitative relations (e.g., $[A] = +$ and $[B] = 0$ when $[A] = [B]$) and 2) quantities which violate the Intermediate Value Theorem and thus are discontinuous (e.g., Q is caused to jump from + to − without crossing 0).

The simplest solution to this is to enumerate all sets of possible transitions and test each for the above two criteria. However, the number of sets of possible transitions grows exponentially with the number of quantities which can transition; thus this solution becomes intractable for large systems (deKleer and Bobrow [12] use a similar approach, but only need to consider the transitions of the independent state variables).

Instead, Transition Ordering uses 1) the direction each quantity is moving with respect to zero, and 2) the qualitative relations between these quantities to formulate a set of constraints. These constraints determine which quantities can transition first while still satisfying the criteria of consistency and continuity. If in the worst case, every qualitative relation is used during Transition Ordering, then this solution grows linearly with the number of relations in the system.

If the derivative of a non-zero quantity (Q) is known then its direction will be monotonic over the interval (Q's derivative has a single qualitative value during that interval) and can always be determined by Transition Recognition. However, even if the derivative of Q is unknown, it is still sometimes possible to determine Q's direction using one of the qualitative relations associated with Q, along with the directions of the other quantities involved in that relation. This is similar to determining the derivative of Q in that, given the value of Q and its direction we can compute dQ/dt for that interval; however, it differs in a number of important respects. When determining the direction of Q we are taking advantage of those times when 1) the value of dQ/dt remains the same during the entire interval of interest (i.e., Q is changing monotonically) and 2) the value of dQ/dt can be computed unambiguously. If the direction of Q cannot be easily determined it is left unknown. On the other hand, when determining the behavior of dQ/dt over time each ambiguity must be resolved. Furthermore, if dQ/dt changes value several times over the interval of interest, then this interval must be broken into a series of sub-intervals using Transition Analysis (which then attempts to determine dQ/dt's direction).

The qualitative relations used in modeling devices are built from equality, negation, addition and multiplication. Thus for each of these operations Transition Ordering contains a set of rules which place constraints on the direction (e.g., toward zero) and transition status (e.g., can't transition) of each quantity involved in the operation. The next section provides a few examples of these rules for each type of operation. In each example we assume that the relation holds over the interval of interest and the succeeding instant. A complete list of Transition Ordering rules is presented in the appendix.

4.2.1 Transition Ordering Rules

If the signs of two continuous quantities are equivalent (i.e., [A] = [B]) over the open interval of interest and the following instant, then we know that if one of the quantities transitions to zero, then the other quantity must transition at the same time. If we know further that A is a monotone increasing function of B, then A and B are moving in the same direction. This may be viewed simply as a consistency check on equality. The above rule also holds for negation (e.g., A = − B), since negating a quantity does not change its direction with respect to zero.

The case where a quantity is the sum or difference of two other continuous quantities is more interesting. For example, assume that quantities A and C are moving towards zero and B is constant, where [C] = [A] + [B]. If A, B and C are positive, then A will transition to zero before C and C can be eliminated from the list of potential transitions.[22] On the other hand, if B is negative, then C will transition before A, and finally, if B is zero, then A and C will transition at the same time (since [C] = [A]). Also, consider the case where A and C are positive and B is negative but the direction of C is unknown (with the further restriction that C = A + B). If B is known to be constant and A is moving towards zero, then C must also be moving towards zero and will reach zero before A.[23] Finally, for multiplication (e.g., [A] x [B] = [C]) we know that, if A and/or B transitions to zero, then C will transition to zero at the same time; otherwise, neither A nor B is transitioning and C won't transition.

Thus, Transition Ordering 1) factors the quantities into sets which transition at the same time and 2) creates an ordering between these sets according to which transitions precede other transitions.

4.2.2 Applying the Transition Ordering Rules

Transition Ordering rules are applied using a constraint propagation mechanism similar to the one used in propagating qualitative values. If as the result of applying these inference rules it is determined that 1) all the remaining potential transitions will occur at the same time, and 2) the

---

[22] If instead we had said that C transitioned to zero first then A would have to jump from plus to minus without crossing zero (i.e., [A] = [C] − [B] = (0) − (−) = −). This violates the Intermediate Value Theorem and, therefore, cannot occur.

[23] In Transition Ordering the constraint requiring monotonicity of qualitative variables can be significantly weakened. Under this weaker constraint we say that a quantity Q(t) *goes to zero* in the interval $I = (t_s, t_f)$ if Q(t) is continuous and if Q(t) > 0 for all t in I and Lim inf Q(t) = 0 or if Q(t) < 0 for all t in I and Lim sup Q(t) = 0. It can be shown that the Transition Ordering rules presented above are still applicable under this weaker definition of *goes to zero*. An indepth discussion of this definition and its ramifications is presented in [29].

direction of these quantities is known to be toward zero, then the transitions occur at the end of the current interval. Otherwise, an ordering may be externally provided for the remaining potential transitions, or the system can try each of the remaining sets of possible transitions. More quantitative techniques which help resolve the remaining sets of possible transitions are currently being explored.

The next section provides an example of how Transition Recognition and Transition Ordering work together to describe how a simple circuit behaves over time.

4.3 Example

Using Transition Analysis, we can now describe the behavior of the RC example (Figure 1-2) after Instant-0.[24] In section 3.4, Causal Propagation was used to determine the values of the circuit's state variables at Instant-0. The results of this propagation were:

	$[V_{IN,GND}]$	= +	
$\rightarrow$	$[I_{t1(R1)}]$	= +	Resistor Model
$\rightarrow$	$[I_{t1(C1)}]$	= -	Kirchoff's Current Law
$\rightarrow$	$[dV_{IN,GND}/dt]$	= -	Capacitor Model
$\rightarrow$	$[dI_{t1(R1)}/dt]$	= -	Resistor Model
$\rightarrow$	$[dI_{t1(C1)}/dt]$	= +	Kirchoff's Current Law

Since each quantity is non-zero at Instant-0, we know by the Continuity Rule that all the values will remain the same for an open interval (Interval-0) following Instant-0.

Next it must be determined whether or not any quantities will transition to zero at the end of Interval-0. By applying the Integration Rule to $[V_{IN,GND}] = +$ and $[dV_{IN,GND}/dt] = -$, we know that $V_{IN,GND}$ is moving towards zero. Using a similar argument, we determine that $[I_{t1(R1)}]$ and $[I_{t1(C1)}]$ are also moving towards zero.

The direction of $[dV_{IN,GND}/dt]$, $[dI_{t1(R1)}/dt]$ and $[dI_{t1(C1)}/dt]$, however, cannot be determined using the Integration Rule, since their derivatives (the second derivatives of V and I) are unknown. The direction of each of these quantities can be determined using the inference rules for equivalences described above. For example, we know that $[dV_{IN,GND}/dt]$ is moving towards zero, since $[I_{t1(C1)}]$ is moving towards zero and $[I_{t1(C1)}] = [dV_{IN,GND}/dt]$ from the capacitor model. In addition, it is

[24] More complex examples of Transition Ordering are found in the example sections of chapters 5 and 6.

deduced from KCL and the resistor model, which are both equivalences, that $[dI_{t1(R1)}/dt]$ and $[dI_{t1(C1)}/dt]$ are also moving towards zero.

Finally, since all of the quantities are qualitatively equivalent, they will all transition to zero at the same time. Since no other potential transitions exist, each of these quantities will transition to zero at the end of Interval-0.

Using the results of Transition Analysis, we know that the primary cause ($V_{IN,GND}$) at Instant-1 is zero, where Instant-1 immediately follows Interval-0. Causal Propagation is then used to generate a causal account of why, for example, $[dV_{IN,GND}/dt]$ is zero at Instant-1.

The discussion, thus far, has assumed that all quantities behave continuously. The next section discusses how TQ Analysis might be extended to deal with discontinuous behavior.

4.3.1 Discontinuous Behavior

Although an engineer believes that circuits in the physical world exhibit continuous behavior, he often wants to model portions of their behavior discontinuously. For example, a voltage may rise sufficiently fast that the engineer wants to idealize the behavior as a step, simplifying his analysis.

Even when a circuit's behavior is modeled by a discontinuous function, the discontinuities are isolated to a few places and the rest of the function behaves continuously (e.g., a step is only discontinuous at one point). If the point at which a quantity is discontinuous can be identified, TQ Analysis can deal with it simply by not applying Transition Analysis to the particular quantity at that point in time.

The remaining task, then is to identify when a quantity may behave discontinuously. A discontinuity in one of the circuit's state variables may result from either a discontinuity in 1) an input, or 2) one of the device models. Discontinuities in state variables can be identified by propagating each discontinuity (or continuity) forward from the input (or device model) to the affected quantities. This propagation is performed using rules like:

> If $A + B = C$ where A is discontinuous at some point and B is not, then C is discontinuous at that point.[25]

[25] Usually we can say that the output of a qualitative expression is guaranteed to be continuous at some time as long as all of its inputs are continuous at that time.

Creating a set of rules which correspond to integration is more difficult since the integral of a discontinuous function may or may not be discontinuous, depending on the order of the singularity. For example, the integral of an impulse (a step) is discontinuous, while the integral of a step (a ramp) is continuous. To deal with integration the propagation mechanism for singularities must keep track of the order of the singularity as well. TQ Analysis is currently being extended with a set of rules similar to the ones above which deal with discontinuities.

4.4 Summary

Key Concepts:

* The behavior of real physical systems is continuous.

* Transition Analysis may be viewed as a qualitative large signal analysis.

* Transition Analysis is built on a few simple theorems of calculus about intervals.

* Transition Recognition determines the direction of a quantity with respect to zero using the Continuity Rule and the Integration Rule.

* Transition Ordering uses the directions deduced during Transition Recognition, along with the qualitative relations between quantities to:

  - eliminate potential transitions which would violate the Zero-crossing Principle

  - determine, when possible, the direction of quantities not deduced by Transition Recognition.

* Transition Analysis can be easily extended to deal with discontinuities.

Chapter Five

Feedback Analysis

Feedback is an important property of most physical systems. Roughly speaking, feedback occurs whenever one of the inputs to a sum is a function of the sum's output. A feedback path then exists between the sum's output and an input. Negative feedback is often used to add stability to amplifier gain and positive feedback is used in digital systems to provide sharp transitions and bistability. This chapter discusses how feedback and equations with simultaneities affect TQ Analysis.

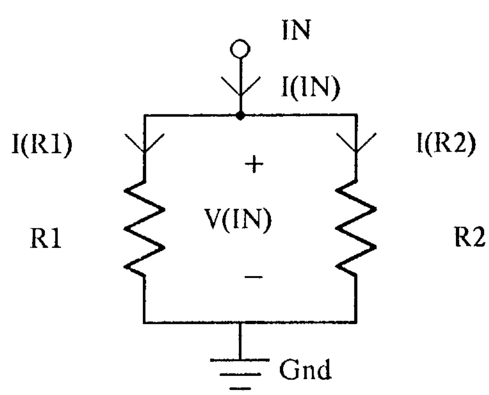


Figure 5-1:RR Current Divider

Instances of feedback can be found in remarkably simple circuits, such as the resistive current divider circuit (RR) shown in Figure 5-1. Assuming that I_{IN} is initially zero, the following is one possible explanation for the response of the circuit to a rise in I_{IN}:

> An increase in I_{IN} produces an increase in I_{R1}, causing V_{IN} to rise. The rise in voltage is applied across R2, increasing I_{R2} and, hence, reducing the effect of the initial current increase on I_{R1}.

This is a simple example of negative feedback, where I_{R2} is the feedback quantity. Circuit analysts usually ignore feedback at this primitive level. Nevertheless, in qualitative analysis it is important to understand feedback at any level for two reasons. First, feedback is a special case which cannot be handled by the TQ mechanism discussed. Second, by understanding the properties which are particular to feedback, the power of TQ Analysis is increased. In the following two sections we discuss the effects of feedback on TQ Analysis, and how TQ Analysis can be augmented to deal with it.

5.1 The Effects of Feedback and Simultaneities on TQ Analysis

If TQ Analysis is run on the RR circuit we immediately run into a problem. Initially all the circuit's state variables and their derivatives are zero. At the beginning of the first open interval, dI_{IN}/dt becomes positive. At this point the only applicable qualitative relation is Kirchoff's Current Law:

$$[dI_{IN}/dt] - [dI_{R2}/dt] = [dI_{R1}/dt]$$

Unfortunately, either $[dI_{R1}/dt]$ or $[dI_{R2}/dt]$ must be known to solve this equation and there is no means of calculating them using purely local information. We can see why this is not possible by looking at the overall structure of the causal relations between the circuit's state variables; this is shown in Figure 5-2 and is called a *causal relation graph.* To deduce $[dI_{R1}/dt]$ from KCL we need to know $[dI_{R2}/dt]$; however, we can see from Figure 5-2 that $[dI_{R2}/dt]$ is a function of $[dI_{R1}/dt]$. This results in a set of simultaneous relations; $[dI_{R2}/dt]$ cannot be calculated without knowing $[dI_{R1}/dt]$ and vice-versa.

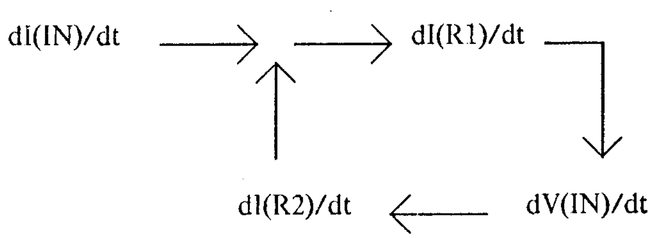


Figure 5-2:Causal Relation Graph for RR Current Divider

The structure of the relations around a binary sum (A + B = C) can be classified as one of two types: *direct sum* or *simultaneity*. A *direct* sum occurs when both inputs are independent of the output C. A *simultancity* occurs whenever one of the inputs, A or B, is a function of its output C (and possibly some other inputs). The simultancity is distinguished when one of the inputs is *only* a function of the output by calling it *feedback*.[26] An example of each type of sum is shown in figure 5-3. The point at which the feedback or simultancity is summed is called a *comparison point*.[27] For the following discussion we will always use C as the *effect*, A as the independent *cause* and B as the *feedback* term.

The mechanism discussed thus far has assumed that all sums are direct. Earlier we found that if the result of a direct sum is ambiguous, the results of Transition Analysis can often be used to resolve this ambiguity. Similarly, if a sum is the comparison point of a simultaneity, the results of Transition Analysis can often be used to determine the value of a quantity which is the effect of a comparison point and continue the propagation based on that value. If the results of Transition Analysis cannot determine this value, we must look for the answer elsewhere. The next section examines the properties of feedback and shows how the resulting information is used to deduce the value for the effect of a comparison point.

5.2 Qualitative Properties of Feedback

Thus far TQ Analysis has only used local information to determine the behavior of a circuit. In this section we examine the overall structure of the relations around a comparison point to determine the value of its effect.

A Feedback loop is described in general by the following two equations:

$$A + B = C$$
$$B = \mathcal{F}(C)$$

If the sign of B is the same as A, then we have an instance of *positive feedback*. In this case the sign of C can be determined unambiguously and is the same as the sign of A. The result of positive feedback is to amplify the effect of any changes in A.

If the sign of B is the opposite of A, then we have an instance of *negative feedback*. The value of C is the difference between the magnitudes of A and B; this results in the sign of C being ambiguous. In a typical use of negative feedback, B dampens the effects of the input A on C, thus stabilizing the

[26] This is a more restrictive definition than the one used in most texts; however, by doing so we are able to describe a number of interesting properties later on.

[27] A comparison point is a point in the relation graph, and does not necessarily correspond to any particular point in the circuit topology.

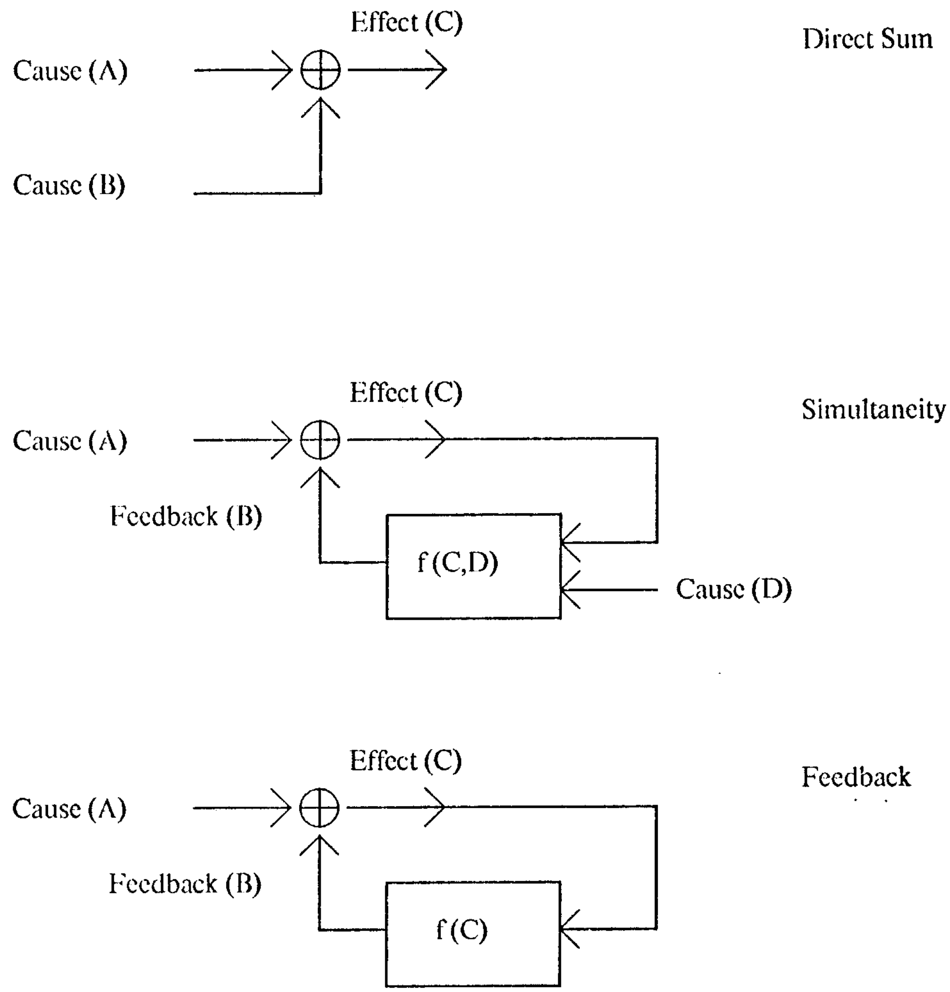


Figure 5-3:Support Classification

output. If $\mathcal{F}$ is a polynomial function, that is $\mathcal{F}$ only involves addition, negation and multiplication, then we say that it is *resistive.*[28] If the feedback is purely resistive then the magnitude of B will be less than A, except when A, B and C are zero; thus, the sign of A and C are the same. Intuitively, this means that, if there are no independent sources in the negative feedback loop, the feedback term will always be weaker than the input to the feedback loop. In order for the magnitude of the feedback

[28] Intuitively, this means that the function is memoryless.

term to be larger than the input, the gain of the function $\mathcal{F}(C)$ would have to be greater than one. An engineers intuition says that such a gain can only be produced using an independent source of power.

Another way of viewing this concept is to consider there to be a finite delay along the feedback path and simulate the results of a change on the input. Assume that A, B and C are initially zero and $\mathcal{F}(C)$ is a negative constant gain such that $[B] = -[C]$. Now, if A becomes positive, then C becomes positive, since B is initially zero and $C = A + B$. This then causes B to become negative, reducing the effect of A on C. Now suppose that the magnitude of B becomes as large as A. This causes C to become zero, and since $[B] = -[C]$, B must also become zero. However, we are now back to the case where A is positive and B is zero, so C must instantaneously jump back to positive. By continuing this argument C appears to oscillate back and forth between positive and zero instantaneously. Such an oscillation violates continuity and cannot occur in real systems, thus C remains positive until A moves to zero.[29]

As we have seen above, for both types of feedback the sign of the effect is the same as the input. This is described by the following rule:

Resistive Feedback Rule

If $A + B = C$ and B is a resistive function of C

then the sign of C will always be the same as A ($[A] = [C]$).

The equivalence between A and B also supplies an additional constraint to Transition Ordering which is used to break simultaneities in the Direction and Transition constraints:

Resistive Feedback Transition Ordering Constraint

If $A + B = C$ and B is a resistive function of C

then A, B and C will transition to (from) zero at the same time.

Returning to the RR example (Figure 5-1), we know that dI_{IN}/dt transitioned from 0 to + at the beginning of Interval-0. $[dI_{R1}/dt]$ cannot be deduced by KCL because $[dI_{R2}/dt]$ is unknown; furthermore, it is assumed that the value of $[dI_{R1}/dt]$ cannot be determined by Transition Analysis. We therefore assume that KCL produces a relation which is a feedback comparison point in the

---

[29] Forbus uses this type of argument, which he refers to as "stutter", to model a person's naive intuition of simple feedback.

current divider's relation graph (Figure 5-2). In addition we assume that dI_{R1}/dt is the effect of the comparison point and dI_{R2}/dt is the feedback term.[30] Under this assumption we assert that $[dI_{R1}/dt] = +$, since $[dI_{IN}/dt] = [dI_{R1}/dt]$ by the Resistive Feedback Rule, and eventually deduce by Causal Propagation that $[dI_{R2}/dt] = +$. This is a valid instance of feedback since the feedback term ($[dI_{R2}/dt]$), is only a function of the effect ($[dI_{R1}/dt]$). Furthermore, it is negative feedback since dI_{IN}/dt and dI_{R2}/dt are both positive, where $[dI_{IN}/dt] - [dI_{R2}/dt] = [dI_{R1}/dt]$.

After a period of time, I_{IN} stops rising and dI_{IN}/dt transitions to zero, marking the end of Interval-0. Using the Transition Ordering constraint for resistive feedback we deduce that dI_{R1}/dt will transition to zero at the same time as dI_{IN}/dt, and dI_{R2}/dt will also transition at the end of Interval-0 since both dI_{IN}/dt and dI_{R1}/dt are transitioning and $[dI_{IN}/dt] - [dI_{R1}/dt] = [dI_{R2}/dt]$. Finally, dV_{IN}/dt will also transition to zero since $[dV_{IN}/dt) = [dI_{R1}/dt]$. Therefore, at the instant following Interval-0 all three currents are positive and constant and remain so until the input changes.

If a circuit includes an inductor or a capacitor along the feedback path then the function $\mathcal{F}$ will involve integration.[31] For a capacitor the relation $[dV/dt] = [I]$ makes it necessary to integrate dV/dt to get V, while the relation for the inductor, $[V] = [dI/dt]$ requires dI/dt to be integrated. A feedback path which requires voltage integration is called *capacitive feedback*, a path requiring current integration is called *inductive feedback*, and a path requiring no integration is called *resistive feedback*. The properties of resistive feedback have already been discussed above. The remainder of this section discusses the properties of feedback paths which involve integration.

An RC circuit exhibiting capacitive feedback is shown in Figure 5-4, along with its relation graph. KCL, again produces a feedback comparison point in the relation graph with I_{R2} as the feedback term and I_{C1} as the effect. If V_{IN} is initially zero, then the Integration Rule tells us that V_{IN} will have the same sign as dV_{IN}/dt during the following time interval. This is depicted in the relation graph as $dV_{IN}/dt \rightarrow V_{IN}$.

When a negative feedback path involves integration, it is not necessarily true that $[C] = [A]$ for all time. For example, if A is positive and begins to quickly drop, then A may become less than B and C becomes negative. The reason for this is that the integration along the feedback loop makes B

---

[30] In the RR example we also could have called dI_{R2}/dt the effect and dI_{R1}/dt the feedback term. A set of heuristics for selecting the effect and feedback term of a comparison point is discussed in section 5.3.

[31] Intuitively a feedback path involving integration has memory.

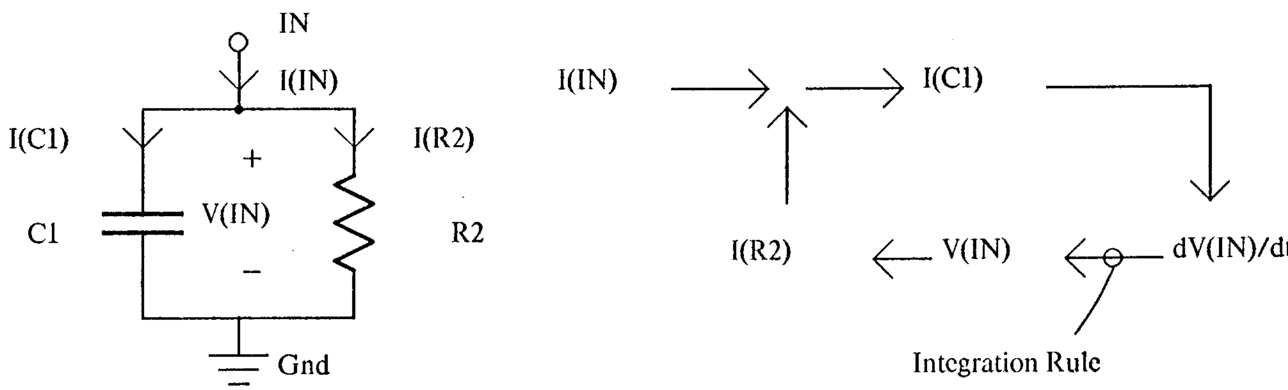


Figure 5-4:RC Current Divider

sluggish and it can't respond quickly enough.

If, however, A, B and C are initially zero at some time instant (t1) and A transitions from zero immediately after t1, then the magnitude of B will be less than A for *an interval of time* (I1) immediately following t1.[32] Therefore, *the sign of C will be the same as A during I1.* After this interval, the magnitude of B may become as large as A, in which case C will cross to zero and [C] ≠ [A]. This is described by the following rule:

Integrating Feedback Rule

IF

1. A + B = C,
2. B is a function of C involving integration,
3. [A] = [B] = [C] = 0 at some time instant and
4. A transitions from 0 immediately after that instant

THEN

[32] The integral of a continuous quantity Q is always less than Q for some interval starting at the beginning of the integration; therefore, the initial effect of the integration is to reduce the magnitude of the feedback term.

the sign of C will be the same as A i.e., ([A] = [C]) for an interval of time immediately following that instant.

This rule can now be used to describe the behavior of the RC circuit (Figure 5-4). Initially, it is assumed that the input current is zero and the capacitor is discharged, so that all the circuit's currents and voltages are constant at zero. At the beginning of some open interval the current I_{IN} becomes positive.[33] Using the Integrating Feedback Rule it is assumed that $[I_{C1}] = +$, allowing the following chain of deductions to be made:

$$([I_{C1}] = +) \rightarrow ([dV(IN)/dt] = +) \rightarrow ([V(IN)] = +) \rightarrow ([I(R2)] = +)$$

Next we want to determine what transitions occur at the end of the interval. By the Integration Rule, V_{IN} is moving away from zero and cannot transition. I_{R2} also cannot transition since $[I_{R2}] = [V_{IN}]$. If I_{IN} is moving towards zero then, I_{C1} must reach zero before I_{IN} to satisfy the Intermediate Value Theorem in the relation $[I_{IN}] = [I_{C1}] + [I_{R2}]$; thus I_{IN} is eliminated. This leaves I_{C1} and dV_{IN}/dt, which must transition at the same time, since $[I_{C1}] = [dV_{IN}/dt]$. Since neither quantity's direction is known, they may or may not transition. The former corresponds to the case where the input current levels off and the capacitor eventually stops charging. This matches our intuition since a capacitor acts like an open circuit at DC. The latter corresponds to the case where the input current continues to rise forever, and the capacitor never stops charging. The Integrating Feedback Rule also holds for inductive circuits; however, inductance is rarely considered in digital MOS circuits. The remainder of this section summarizes the steps involved in resolving ambiguities due to simultaneities.

During Causal Propagation, if one of the inputs to a sum (CP1) cannot be determined and no further deductions can be made, it is assumed that the sum is a simultaneity and the results of Transition Analysis are used, if possible, to determine a value for C. If the value for C cannot be determined by Transition Analysis, it is assumed that the sum is part of a feedback loop and one of the above feedback rules is used to determine C.

Once the value for the feedback term (B) is deduced, this assumption is verified by looking at the causal chain supporting B. If C is encountered along B's causal chain then B is a function of C and the

[33] It is interesting to note that the behavior of this circuit is deduced without knowing the derivative of the input. We will return to this example again in section 5.7, using higher order derivatives to produce a more detailed explanation.

sum is a simultaneity. If all of the causal paths supporting B start at C[34] then B is a function of C alone and the sum is further classified as feedback.[35] Similarly, the feedback is further classified as positive or negative, and resistive, capacitive, inductive or both. Finally, if C is not encountered among B's support, then B is not a function of C and we have a direct sum. If an assumption proves to be false it must be retracted.

Verifying a simultaneity or feedback assumption is more complex for circuits with multiple feedback loops or cross coupled feedback loops since the feedback loops may need to be verified simultaneously. An approach, similar the one discussed above traces B's causal chain with one modification. If the effect of another comparison point (CP2) is encountered along B's causal chain, then it will not be possible to follow the causal chain of CP2's feedback term if it hasn't been determined yet. We know, however, that CP2's feedback term is only a function of CP2's effect and need not be traced. Thus only the input to CP2 is followed (and the feedback term is ignored) under the assumption that CP2 is the comparison point of a valid feedback loop. If at a later point CP2's feedback term is determined and it is found that CP2 is not the comparison point of a feedback loop then the feedback term must be traced to verify CP1.

5.3 Bidirectional Comparison Points

The problem addressed in this section is the determination of the effect and feedback terms of a comparison point. In the previous section we assumed that it was known which quantity was the effect and which quantity was the feedback term. However, if the comparison point of a feedback loop involves a bidirectional sum, the selection of these two quantities is not obvious. Although all of the feedback comparison points seen thus far have been a result of KCL, in general they can be produced by any relation involving a sum, either from a network law or a device model. Both of the network laws (KCL and KVL) contain bidirectional sums; however, none of the device models which

[34] A less restrictive version states that the sum is a feedback comparison point if all of the causal paths supporting B start at a quantity which is a member of C's qualitative equivalence class. C's qualitative equivalence class then consists of all continuous quantities whose qualitative values are equal (or the negation) of C's qualitative value during the interval (thus each quantity in C's qualitative equivalence class can be described as a function of C alone).

[35] B can be a function of another quantity (X) as well as C as long as $[X] = 0$; in other words, X is not driving the feedback term.

we've encountered in the MOS domain contain bidirectional sums.[36]

Returning to the RR current divider example (Figure 5-1) we notice that the circuit is symmetrical; R1 and R2 could be switched without changing the behavior of the circuit. This can also be seen in the relation graph (Figure 5-2). An equally valid description of the circuit behavior would have been to use I_{R1} as the feedback term and I_{R2} as the effect. In this example the selection of a feedback term is arbitrary since the feedback path is bidirectional. This, however, is not the case in the RC current divider example (Figure 5-4). We see from the relation graph that the causality along the feedback path can only run in one direction. I_{R2} must be the feedback term since the Integration Rule is unidirectional, allowing the causality to run only from dV_{C1}/dt to V_{C1}, but not the reverse.

One way of dealing with bidirectional comparison points is simply to try both possible directions. This, however, becomes costly, since most complex circuits have a large number of simultaneities. Furthermore, engineers appear to use a set of heuristics which allow them to significantly reduce the amount of backtracking which is performed while reasoning about feedback circuits.

Returning to the RC current divider, a circuit analyst might describe its behavior as follows:

> When the input current becomes positive, the capacitor initially acts like an incremental short and all the current goes into the capacitor. As the capacitor charges, this produces a positive voltage across the resistor, causing I_{R1} to be positive.

The important part of this dialogue is the viewpoint that C1 acts like an incremental short. We can understand this viewpoint by looking at the impedance of a capacitor. Initial changes in state variables are usually fairly sharp, involving a large high frequency component. At high frequencies (ω), the impedance of the capacitor ($1/j\omega C$) becomes very small, causing the capacitor to act like an incremental short or "battery".

In this example the designer reasons that the current through the capacitor initially dominates over the resistor current, and selects the former as the effect of the comparison point. Looking at the circuit's relation graph, we see that the capacitor "integrates" I_{C1}. The causality can only move from I_{C1} towards V_{In} and not vice-versa; therefore, I_{C1} must be the effect of the comparison point. The opposite case occurs in the RC high pass filter shown in Figure 5-5. This circuit behaves as follows:

> When the input voltage begins to rise the capacitor initially acts like a "battery", transmitting the change in the input voltage directly to the resistor's voltage. This

---

[36] The mosfet model, presented later on, contains a relation involving a sum which can be used as a comparison point; however, the relation is unidirectional.

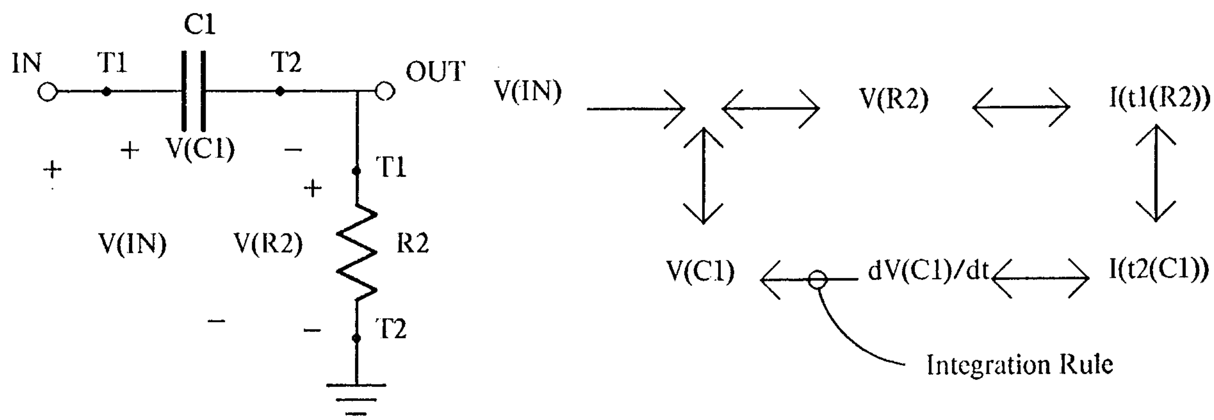


Figure 5-5:RC High Pass Filter

produces a current through the resistor which charges the capacitor and causes V_{C1} to increase.

In this example the designer reasons that the capacitor is initially insensitive to any changes in voltage, and therefore selects it as the feedback term of the comparison point. Looking at this circuit's relation graph, we see that causality can only move from I_{C1} towards V_{C1} and not vice-versa; therefore, I_{C1} must be the feedback term in the comparison point.

If the capacitor is replaced with an inductor in the above two circuits, the behavior is exactly the opposite. At high frequencies (ω) the impedance of the inductor ($j\omega L$) becomes very large, causing the inductor to act incrementally like an open circuit or current source. Using the above analysis, we can construct the following heuristic:

Feedback Direction Heuristic

For each of the two unknown quantities in the comparison point:

* If the quantity is current (or one of its derivatives) then

  - If the relation attached to the quantity is capacitive, the quantity is the effect of the comparison point.

  - If the relation attached to the quantity is inductive, the quantity is the feedback term of the comparison point.

* If the quantity is voltage (or one of its derivatives) then

  - If the relation attached to the quantity is inductive, the quantity is the effect of the comparison point.

  - If the relation attached to the quantity is capacitive, the quantity is the feedback term of the comparison point.

* If no relation, other than the comparison point, is attached to either quantity then Feedback Analysis is not appropriate.

Section 5.5.1 provides an example of how these heuristics are used to describe more complex circuits. The next section discusses a means of restricting the number of sums which are treated as comparison points.

5.4 Localizing the Effects of KVL

The KVL rule from chapter 3 states that the sum of the voltages between any three nodes is zero, no matter how far or close the nodes are spaced. This differs from KCL and the device models in that its effects are non-local. This presents some serious problems when applying the feedback rules, since a change in voltage at one end of the network will produce a plethora of feedback assumptions across the network, few of which are of any use. Something is clearly wrong with this approach. An engineer doesn't suddenly jump back and forth from one end of a circuit to the other when describing its behavior. Instead he prefers to reason about circuit behavior in terms of local interactions. When a voltage is given with respect to a reference, it is often viewed as a node voltage (or potential) and the reference becomes implicit (e.g., $V_{IN,GND}$ becomes V_{IN}). A node voltage at node (N) may then be reasoned about as if it was a quantity local to N. To determine the effects of the node voltage, the analysts will look at the *branch* voltage (i.e. a voltage between two nodes) across devices which are directly connected to node N. In his PhD thesis, de Kleer identified the importance of reasoning about voltage locally when dealing with feedback, calling it the KVL Connection Heuristic. The following is a paraphrased version of the KVL Connection Heuristic:

> If the voltage at a node, which is connected to one terminal of a device, is increasing or decreasing, and nothing else is known to be acting on the device, then the device responds as if the unknown actions are negligible

A statement of the KVL Locality Heuristic used in TQ Analysis is shown below. The effect of the node voltage on the surrounding circuit is not stated in the KVL Locality Heuristic, but is determined by the feedback rules and heuristics outlined above. The resulting behavior of the KVL Locality Heuristic differs from the KVL Connection Heuristic in that the change in the node voltage **may** or **may not** be transferred across a locally connected device, depending on whether it is resistive, capacitive, or inductive. The latter case occurs in the RC example (Figure 5-5) where the capacitor acts initially like a voltage supply, and the change in V_{IN} is produced across R2 rather than C1.

KVL Locality Heuristic

Only apply Feedback Analysis to a comparison point produced by KVL if the input to the comparison point is a node voltage (potential).

5.5 Examples

Using the rules and heuristics described in this chapter, it is now possible to describe the behavior of more complex circuits. In this section, TQ Analysis is used to describe the behavior of two circuits: an RC ladder network and a Wheatstone bridge circuit. The first example combines the Integrating Feedback Rule with the Feedback Direction Heuristic to describe the behavior of a capacitive circuit over an interval of time. The second circuit provides a complex example of resistive feedback, and shows how Feedback Analysis interacts with Transition Ordering.

5.5.1 RC Ladder Example

An RC ladder network is shown in Figure 5-6 along with its causal relation graph. This circuit has three comparison points, two from KVL and one from KCL, producing the three potential feedback loops shown in the network's causal relation graph. We assume that the voltage across the input and each capacitor is zero at Instant-0. At the beginning of the next interval (Interval-0), V_{IN} begins to rise and becomes positive. Using TQ Analysis, we can predict the behavior of the circuit during Interval-0, as described below. Note that each phrase of the explanation is followed by a set of rules which were used to deduce that portion of the behavior:

1. As the voltage at node IN rises,
   (Input, KVL Locality Heuristic)

2. C2 initially acts like a battery
   (Feedback Direction Heuristic for voltages)

3. and the voltage across the resistor (R1) connected to IN begins to increase.

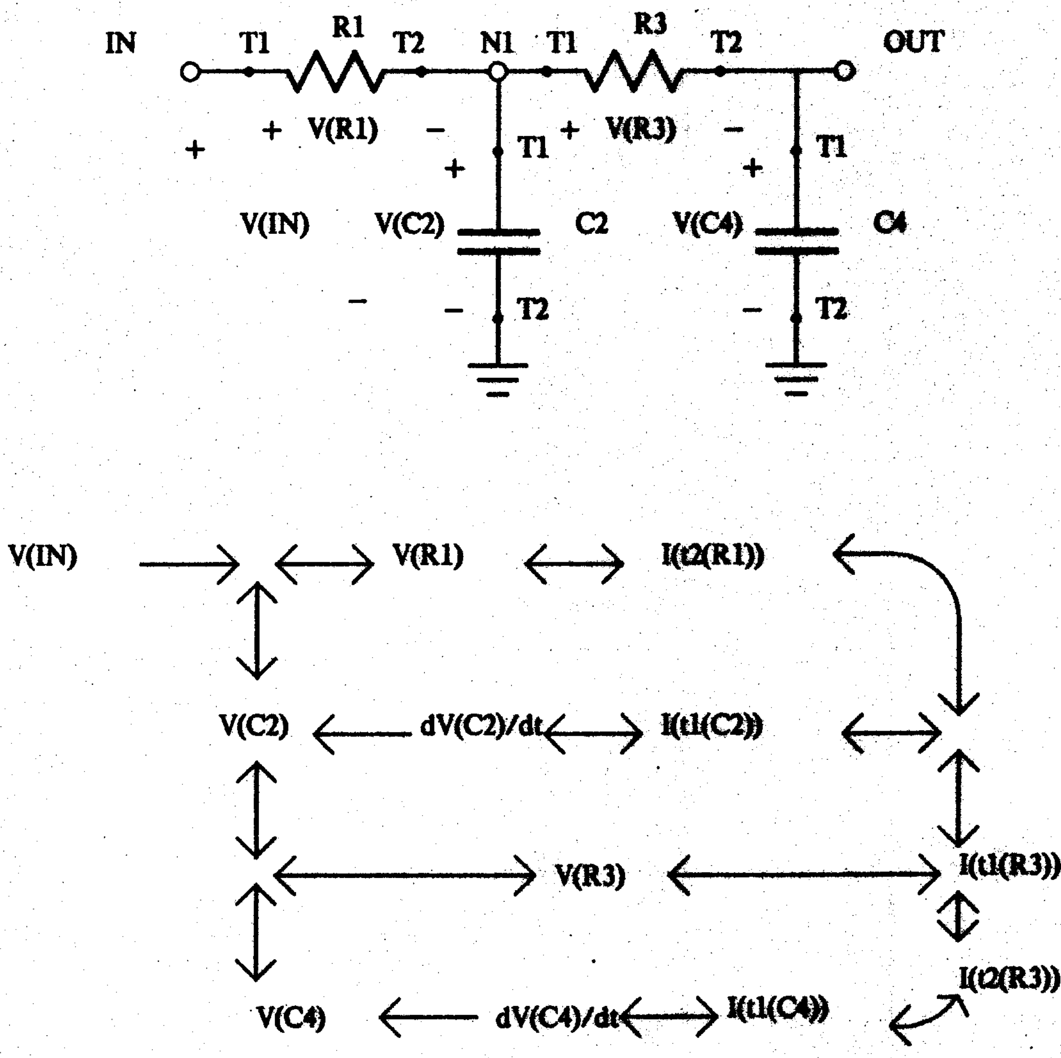


Figure 5-6:RCRC Ladder and Causal Relation Graph

(Integrating Feedback Rule)

4. The positive voltage across R1 produces a current
   (Integration Rule, resistor model)

5. which flows into C2,
   (KCL, Integrating Feedback Rule)

6. since C2 initially acts like an incremental short.
   (Feedback Direction Heuristic for currents)

7. This causes C2 to charge, producing an increase in voltage at node N1.
   (capacitor model)

8. This change in voltage is transferred across R3,
   (same as 1, 2 and 3)

9. producing a current which flows into C4,
   (resistor model, Integration Rule, KCL)

10. causing the capacitor to charge, and raising the voltage at node OUT.
    (capacitor model)

5.5.2 Wheatstone Bridge Example

Figure 5-7 shows an example of a Wheatstone bridge, one of the more complex, purely resistive circuits used in engineering practice. Again, the voltage across the input, and therefore across each resistor, is assumed to be zero at Instant-0. At the beginning of the next interval (Interval-0), V_{IN} begins to rise and becomes positive. Causal Propagation and Feedback Analysis are used to predict the behavior of the circuit during Interval-0. The arrows in the causal relation graph (figure 5-7) for the bridge circuit indicate the direction of causal flow resulting from the propagation.[37] The following is an explanation of I_{R5}'s behavior in response to the input:

1. As the voltage at node IN rises,
   Input

2. the voltage across R2 increases,
   KVL, Locality Heuristic, Resistive Feedback Rule

3. causing an increase in I_{R2},
   resistor model

---

[37] All the relations for the Wheatstone bridge circuit are bi-directional.

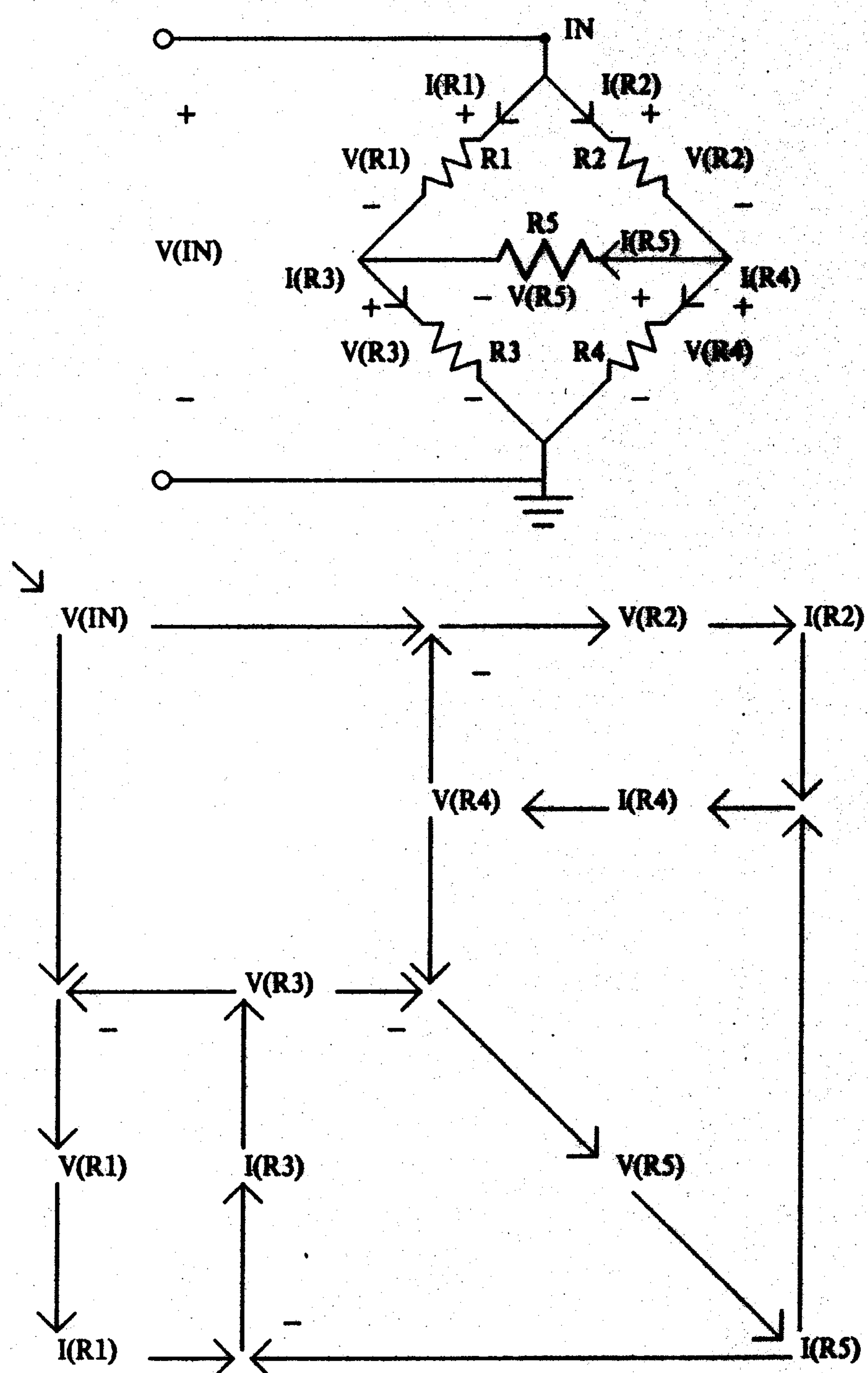


Figure 5-7: Wheatstone Bridge and Causal Relation Graph

4. which flows into R4, producing an increase in I_{R4},
   KCL, Resistive Feedback Rule

5. and causing V_{R4} to rise.
   resistor model

6. Similarly, increasing V_{IN} causes an increase in V_{R1},
   KVL Locality Heuristic, Resistive Feedback Rule

7. increasing I_{R1},
   resistor model

8. which then flows into R3, producing an increase in I_{R3},
   KCL, Resistive Feedback Rule

9. and causing V_{R3} to rise.
   resistor model

At this point an ambiguity arises; the voltage across R5 may become positive, negative, or remain zero. All three possibilities could occur, depending on the relative magnitudes of V_{R4} and V_{R3}. These, in turn, depend on the specific values of the resistors in the network. For now we assume that $|V_{R3}| > |V_{R4}|$. Using this assumption the explanation can be completed:

1. Assuming the increase in V_{R3} dominates over V_{R4}, then
   Assumption

2. this causes an increase in V_{R5},
   KVL

3. producing an increase in I_{R5}.
   resistor model

Next, TQ Analysis tries to determine whether or not any voltage or current can transition back to zero. At first glance, it seems likely that V_{R5} will go to zero, or oscillate back and forth between positive and negative, since its qualitative value was ambiguous in the above analysis. However, intuitively, an engineer knows that none of the circuit's voltages or currents should oscillate; because the circuit is purely resistive, there are no energy storage units to support an oscillation. That is, none of the voltages or currents should transition to zero until the input voltage goes to zero. Combining Transition Ordering with the results of Feedback Analysis, shown below, TQ Analysis is able to make a similar prediction.

Applying the Resistive Feedback Transition Ordering Constraint (section 5.2) to the four feedback comparison points produces the following relations, where t(X) denotes the time that X will transition to zero:

$$t(V_{IN}) = t(V_{R1}) = t(V_{R3}) \tag{5-1}$$
$$t(V_{IN}) = t(V_{R2}) = t(V_{R4}) \tag{5-2}$$
$$t(I_{R2}) = t(I_{R4}) = t(I_{R5}) \tag{5-3}$$
$$t(I_{R1}) = t(I_{R3}) = t(I_{R5}) \tag{5-4}$$

The resistor model provides the following additional constraint:

$$t(V_{Rn}) = t(I_{Rn}) \text{ for n from 1 to 5} \tag{5-5}$$

Using relations 5-1 and 5-2 above, we know that V_{R1} through V_{R4} will transition to zero exactly when V_{IN} transitions. In addition, from relations 5-3 and 5-4, I_{R1} through I_{R5} will also transition together. Finally, applying relation 5-5 to R5 and one of the other resistors (R1 – R5), Transition Ordering determines that all of the currents and voltages in the circuit will transition to zero, exactly when V_{IN} does. This is precisely what we predicted above based on our intuition; thus V_{R5} can only oscillate if the input voltage oscillates.

Thus far we have discussed simultaneities which result in ambiguities at sums; in the next section we discuss how simultaneities can produce ambiguities which originate right at the primary cause.

5.6 Simultaneities Involving Primary Causes

Returning to the parallel RC circuit described in the introduction, it has been determined thus far that $V_{IN,GND}$ is initially positive during Instant-0, and decreases during Interval-0, reaching zero at Instant-1. At Instant-1 the values of the voltage, currents and their derivatives are all zero. Intuitively we know that the voltage will remain zero during Interval-1, the interval following Instant-1 (since any perturbation of $V_{IN,GND}$ off of zero will immediately decay back to zero). If $[dV_{IN,GND}/dt]$ is known during Interval-1, then $V_{IN,GND}$ can be determined using the Integration Rule. Unfortunately $[dV_{IN,GND}/dt]$ can only be deduced from $V_{IN,GND}$; i.e., the relations contain a simultaneity involving $V_{IN,GND}$ (Figure 5-8). This situation is quite analogous to the feedback examples presented earlier. In fact the parallel RC circuit is identical to the RC high pass filter (Figure 5-5) when V_{IN} for the filter is zero. Because there is no independent source driving the feedback loop all quantities along the feedback loop will remain constant at zero. Feedback Analysis deals with this type of simultaneity with the following rule:

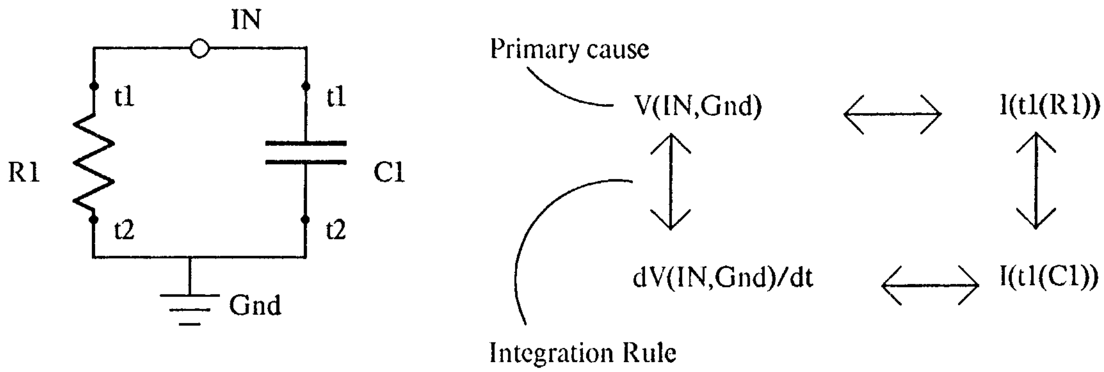


Figure 5-8: Parallel RC Circuit and Causal Relation Graph

Simultaneous Primary Cause Rule

> If the value of a primary cause (Q), and its derivative (dQ/dt) are unknown during an interval and were both zero at the previous instant, then assert that Q is zero during that interval under the assumption that Q is part of a feedback loop with no source.

This rule is applied to any primary cause which is not independently driven as an input (i.e., the independent variable of a memory element, such as the voltage across a capacitor or the current through an inductor). It is then substantiated when dQ/dt is deduced by making sure that 1) [dQ/dt] = 0 and 2) [dQ/dt] is a function of Q alone.

Applying the above rule to the parallel RC example, we assume that $[V_{IN,GND}] = 0$ during Interval-1. From this it is deduced that I_{R1}, I_{C1}, and finally $dV_{IN,GND}/dt$ are zero, thus substantiating the assumption. At this point all of the state variables and their derivatives are zero. Transition Analysis determines that there are no more transitions and the system has reached steady state.

5.7 High Order Derivatives

Thus far, we have described the complete mechanism, provided by TQ Analysis, for analyzing networks of devices which are modeled by a single operating region. The examples presented have only involved voltage, current, and their first derivatives. TQ Analysis, however, is not restricted to these quantities and can use higher order derivatives when available. In this section we return to the RC current divider (figure 5-4) to show how higher order derivatives may be used to provide more detailed predictions.

Recall from section 5.2 that, assuming C1 of the divider is initially discharged, a positive input current, I_{IN}, starting at the beginning of an open interval (I1), causes C1 to charge, making V_{IN} positive, which produces a current through R2. Knowing only the sign of the input current, TQ Analysis is unable to determine whether or not the capacitor stops charging after a period of time, (i.e., does I_{C1} transition to zero?). The resolution of this ambiguity depends on more detailed characteristics of the input waveform which were not provided. Now we will provide some additional constraints on the input waveform and see how they affect the resulting prediction.

Instead of specifying just that the input is positive, it is assumed that the input current is monotonically increasing. This provides the additional constraint that the input current's derivative is positive during I1. The relation graph corresponding to the additional input, dI_{IN}/dt, is the same as the one in figure 5-4, except that each quantity is replaced by its derivative. Again, applying the Integrating Feedback Rule to KCL, followed by constraint propagation, the following *additional* deductions are made:

$$([dI_{IN}/dt] = +) \rightarrow ([dI_{C1}/dt] = +) \rightarrow ([d^2V(IN)/dt^2] = +)$$
$$\rightarrow ([d^2I(R2)/dt^2] = +)$$

Using these deductions, it is now possible to determine what I_{C1} will do. Notice that, since both dI_{C1}/dt and I_{C1} are positive, I_{C1} is moving *away* from zero, thus, resolving the ambiguity about whether or not I_{C1} transitions. As long as dI_{C1}/dt stays positive, current will continue to flow and the capacitor voltage will rise monotonically. This resolves the ambiguity mentioned above; however, another ambiguity arises: dI_{C1}/dt may transition to zero. This transition might occur if the derivative of the input current oscillates sharply, while remaining in the positive region.

To resolve this new ambiguity, one final constraint is placed on the input current: the input waveform is a ramp. This corresponds to the second derivative of the input current being zero during I1. Using Causal Propagation, the following deductions are made:

$$([d^2I_{IN}/dt^2] = 0) - ([d^2I_{R2}/dt^2] = +) \rightarrow ([d^2I_{C1}/dt^2] = -)$$
$$\rightarrow ([d^3V(IN)/dt^3] = -)$$

This resolves the ambiguity about dI_{C1}/dt transitioning. Since its derivative (d^2I_{C1}/dt^2) is negative, dI_{C1}/dt is moving toward zero. Furthermore, Transition Ordering determines that, since dI_{C1} will reach zero first, it must transition. Once, dI_{C1}/dt reaches zero, I_{IN} and dI_{R2}/dt will both be rising ramps, where I_{R2} lags I_{IN} by the constant current I_{C1}. No further transitions will occur as long as the input current remains a ramp.

Above we saw that, by using higher order derivatives, it was possible to resolve all ambiguity in response to the input. Furthermore, higher order derivatives enabled us to describe the behavior of the circuit's state variables more precisely. For example, knowing only the sign of the input it was only possible to determine that I_{C1} was positive. By knowing the first and second derivatives of the input, we determined that I_{C1} was initially rising, but eventually leveled off to a constant value.

Using higher order derivatives does not always reduce the number of ambiguities. If instead $[d^2I_{IN}/dt^2]$ was positive, then it would have been ambiguous whether or not $[d^2I_{C1}/dt^2]$ transitioned to zero. Furthermore, if the input was a rising exponential (e.g., $I_{IN} = e^t$), then all of the nth derivatives of I_{IN}, up to $n = \infty$ would be positive. For the exponential input, the addition of higher order derivatives would result in replacing one ambiguity with another at a more detailed level. For this input one must use some other reasoning technique, such as induction, to resolve the ambiguity.

In this section we have seen that higher order derivatives may be used to add detail to the prediction of the circuit's behavior. Furthermore, this additional information may sometimes be used to resolve ambiguities. However, the use of higher order derivatives does not guarantee that all existing ambiguities will be resolved and may even add more ambiguities. In addition, these derivatives may add a level of detail into the explanation which the user would rather ignore or might find confusing.[38] A good theory of when it is profitable to pursue higher order derivatives is important in qualitative analysis and is a topic of future research.

---

[38]The desired level of detail for an explanation depends on many factors, such as the domain and type of user. In the MOS domain, first and second derivatives have been found adequate for most circuits examined.

5.8 Summary

When TQ Analysis encounters a situation where the input to a sum is a function of the sum's output, Causal Propagation cannot continue. The sum is assumed to be a simultaneity, and the results of Transition Analysis are used to deduce the output. If this is not possible, the sum is assumed to be a case of feedback and the qualitative properties of feedback can be used to determine the output. These properties are summarized by the Resistive Feedback Rule and the Integrating Feedback Rule. Finally, the assumption of simultaneity or feedback is verified.

At a bidirectional comparison point an additional complication arises: The cause and feedback terms must be identified. An engineer uses his intuition about capacitive and inductive relations to resolve this complication. This intuition is summarized in the Feedback Direction Heuristic. At present, bidirectional sums appear only in the KCL and KVL rules.

The number of potential feedback comparison points can be quite large due to the non-local behavior of KVL. A designer, restricts this number by reasoning in terms of local interactions. This notion is captured in the KVL Locality Heuristic.

Finally, by viewing a circular set of relations as a feedback loop with a "mythical cause", Feedback Analysis can be used to resolve the simultaneity.

Chapter Six

Operating Regions and the MOS Transistor

The basic building block of MOS circuits is the Metal Oxide Silicon Field Effect Transistor or mosfet. This chapter develops a qualitative model for an enhancement n-channel mosfet and discusses how TQ Analysis is extended to deal with this and other models which have more than one operating region.

The mosfet can be modeled at a number of levels from device physics to the switch level abstraction. In this chapter we use a very simple analog model [22] which is adequate for modeling most digital circuits at the qualitative level. The mosfet model has three terminals which act as the gate, source and drain. In addition, this model is broken into two parts: *conduction* and *capacitance*, where conduction describes the relationship between $V_{G,S}$, $V_{D,S}$ and I_D, and capacitance describes the relationship between Q_G and the device's terminal voltages. Figure 6-1 shows a high-level model of the mosfet where capacitance is modeled as an ideal capacitor from gate to source and conduction is modeled as a nonlinear dependent current source. The conduction model is broken into several operating regions (e.g., off, on, unsaturated, saturated, forward and reverse), and each region is described using a different set of constitutive relations. In this chapter we first discuss a mechanism for modeling devices with multiple operating regions, then develop the mosfet's conduction and capacitance models, and conclude with the analysis of a simple mosfet circuit.

6.1 Modeling Devices With Multiple Operating Regions

Thus far only the analysis of devices with a single operating region have been discussed. To analyze devices with multiple operating regions (e.g., diodes, bipolar transistors, JFETs, Mosfets etc.), TQ Analysis must be extended in two ways. First, it must be able to determine which region a device is currently operating in. This is necessary to determine which set of device relations is applicable. Second, TQ Analysis must be able to determine when a device transitions from one operating region to another, identifying when one set of device relations must be exchanged for another.

An operating region is described by the set of boundaries which surround it. To determine the

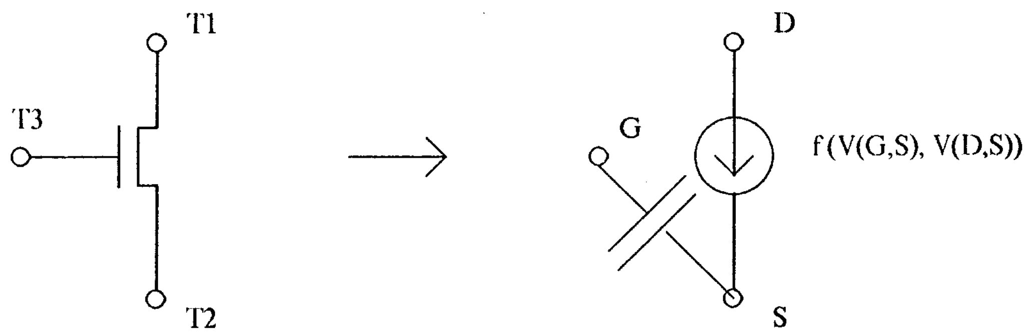


Figure 6-1:n-Mosfet Large Signal Model

operating region which a device is currently in, it is necessary to determine the device's position with respect to each of these boundaries. To determine whether a device may transition between operating regions, one must determine the direction the device is moving with respect to each of these boundaries. When the device crosses one of these boundaries, it moves into a different operating region. In Chapter 4, we discussed a mechanism (Transition Analysis), for determining how a quantity transitions across the boundary *zero*, between the intervals, positive and negative. If a quantity is associated with each operating region boundary, which describes the device's distance from that boundary, then Transition Analysis can be used to determine the device's position and movement with respect to that boundary.

The remainder of this section describes the steps involved in defining the operating regions of a device and specifying their associated device relations. During this explanations each step will be demonstrated using the operating regions of the mosfet as an example.

The operating regions of the mosfet model can be broken into two sets: {Saturated, Unsaturated, Off} and {Forward, Symmetric, Reverse}. These regions are described by the following inequalities:

Forward: $V_{T1,T2} > 0$
Symmetric: $V_{T1,T2} = 0$
Reverse: $V_{T1,T2} < 0$

Off: $(V_{G,S} - V_{Th}) \leq 0$
On: $(V_{G,S} - V_{Th}) > 0$
Unsaturated: $V_{D,S} < (V_{G,S} - V_{Th})$
Saturated: $V_{D,S} \geq (V_{G,S} - V_{Th})$

A graphical representation of these operating regions is shown in figure 6-2.

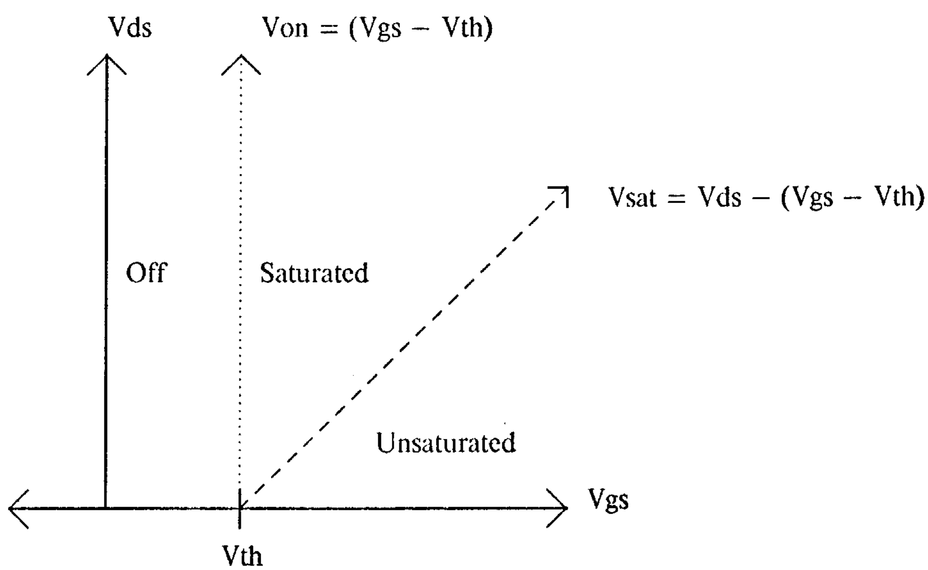


Figure 6-2:n-Mosfet Operating Regions

The first step in creating an operating region involves defining the boundaries that separate it from other regions. A boundary and the half planes above and below it can be described by the sign of a quantity; that is, the boundary's associated state variable.[39] This quantity is defined in terms of other state variables using the arithmetic operations: addition, negation, and multiplication. Graphically,

[39]This state variable represents the distance from the device's current position in state space to the closest point on the boundary.

these operations correspond to rotating and translating, reflecting, and scaling a linear boundary. In the mosfet model, the quantity describing the boundary between Forward and Reverse is just $V_{T1,T2}$ and already exists. For the boundary between On and Off we define the state variable V_{On}, where $V_{On} = V_{G,S} - V_T$. Finally, the state variable V_{Sat} is used to describe the boundary separating Saturated and Unsaturated, where $V_{Sat} = V_{D,S} - V_{On}$.

Next we need to be able to determine whether or not a device is above, below, or at a particular boundary. This corresponds to asking whether or not the boundary's state variable is positive (> 0), negative (< 0) or zero (= 0). In addition, the region name is a boolean which is true if the device is in that region, false if the device is out of that region, and ? if undetermined. The Forward, Symmetric and Reverse regions are now defined as:

$$\begin{aligned}
\texttt{Forward} &= (V_{T1,T2} > 0) \\
\texttt{Symmetric} &= (V_{T1,T2} = 0) \\
\texttt{Reverse} &= (V_{T1,T2} < 0)
\end{aligned}$$

If the behavior in a set of regions and boundaries can be described by the same set of relations, then those regions and boundaries can be combined. These regions are combined using union, intersection and complement, which correspond to ANDing, ORing and Negating the truth values associated with the particular regions being combined. For example, using $\vee$ or $\neg$, two new predicates can be defined which combine a region and a boundary into a semi-closed region:

$$\begin{aligned}
(A \geq 0) &:= (A = 0) \vee (A > 0) \\
(A \leq 0) &:= (A = 0) \vee (A < 0)
\end{aligned}$$

With these boolean operations and predicates we can now define the remaining operating regions Off, On, Unsaturated and Saturated:

$$\begin{aligned}
\texttt{Off} &= (V_{On} \leq 0) \\
\texttt{On} &= (V_{On} > 0) \\
\quad \texttt{Unsaturated} &= (V_{Sat} < 0) \wedge \texttt{On} \\
\quad \texttt{Saturated} &= (V_{Sat} \geq 0) \wedge \texttt{On}
\end{aligned}$$

Finally, each operating region can be associated with its corresponding relations using a set of conditional statements. For example, when a mosfet is off the current is zero and constant. This is expressed in the following statement:

$$\texttt{If Off then } ([I_D] = 0) \wedge ([dI_D/dt] = 0)$$

Next a mechanism must be provided for determining how a device transitions from one operating region to another. Recall that each boundary is described by a corresponding state variable. The sign of this state variable determines where the device is operating with respect to that boundary. By computing the sign of the state variable's derivative, it is possible to determine the direction the device is moving with respect to its associated boundary. Transition Analysis is then applied to determine when a device will transition between operating regions. For example if $[V_{On}] = +$ and $[dV_{On}/dt] = -$ then the mosfet is On and moving towards the On/Off boundary.[40] Transition Analysis assumes that all quantities are modeled by continuous functions. Because of this, the relations in each operating region must be continuous across the region boundaries, or the discontinuity must be made explicit to Transition Analysis (section 4.3.1). In the mosfet model this means that, since the current is zero in the Off region and at the boundary of On/Off, when a device moves from Off to On, the current at the edge of the On region can be 0^+, but could not jump discontinuously to 5 amps. The Shichman-Hodges model, which is used to derive the mosfet model in the next section, is continuous in current, voltage, and their first derivatives. In the next section, the concept of multiple operating regions is used to model conduction in the n-channel mosfet.

6.2 The Mosfet Conduction Model

The n-mosfet is a charge controlled switch in which the three terminals, T1, T2 and T3, act as drain, source and gate respectively. When positive charge is placed on the gate of the n-mosfet a layer of negative charge forms between the gate and the substrate, which creates a channel between the source and drain and allows negative charge to flow from the source to the drain. The voltage at the time the channel is formed is called the threshold voltage (V_{Th}) and is measured between the gate and source. If $V_{G,S} \leq V_{Th}$ then the device is Off and no current flows between source and drain. If $V_{G,S} > V_{Th}$ then the mosfet is On and current may flow along the channel. The On operating region is subdivided into Saturated and Unsaturated regions which are described by the following quantitative relations:

Saturated: $I_D = K(V_{G,S} - V_{Th})^2$

Unsaturated: $I_D = K((V_{G,S} - V_{Th})V_{D,S} - V_{D,S}^2/2)$

If $V_{T1,T2}$ is zero then the source and drain are indistinguishable and no current flows through the channel.

The mental image most electrical engineers have for a mosfet is that voltage between the gate and source produces charge on the gate and in the channel, causing current to flow. Changes in drain and

---

[40] This also requires some modifications to the mechanism for verifying feedback loops which have not been discussed.

```
Model: n-channel-Mosfet(M)
    Terminals: T1 T2 T3
    Corresponding Nodes: N1 N2 N3
    Roles: G S D
    Relations:
        [I_G] = 0
        [dI_G/dt] = 0
        [V_On] ← [V_G,S] - [V_Th]
        [dV_On/dt] ← [dV_G,S/dt] - [dV_Th/dt]
        [V_Sat] ← [V_D,S] - [V_On]
        [dV_Sat/dt] ← [dV_D,S/dt] - [dV_On/dt]
    Region: OFF : [V_On] ≤ 0
        Relations:
            [I_D] = 0
            [dI_D/dt] = 0
    Region: ON : [V_On] > 0
        Region: SATURATED : [V_Sat] ≥ 0
            Relations:
                [I_D] ← [V_On]^2
                [dI_D/dt] ← [V_On] x [dV_On/dt]
        Region: UNSATURATED : [V_Sat] < 0
            Relations:
                [I_D] ← ([V_On] - [V_Sat]) x [V_D,S]
                [dI_D/dt] ← [V_D,S] x [dV_On/dt] - [V_Sat] x [dV_D,S/dt]
    Assertions:
        three-terminal-device(M)
        Assume-initially (is-acting-in-the-role-of (T1(M) D(M)))
        Assume-initially (is-acting-in-the-role-of (T2(M) S(M)))
        Assert-always (is-acting-in-the-role-of (T3(M) G(M)))
    Region: FORWARD : [V_T1,T2] >0
        Assertions:
            is-acting-in-the-role-of (T1(M) D(M))
            is-acting-in-the-role-of (T2(M) S(M))
    Region: REVERSE : [V_T1,T2] < 0
        Assertions:
            is-acting-in-the-role-of (T1(M) S(M))
            is-acting-in-the-role-of (T2(M) D(M))
```

Figure 6-3:N-Mosfet Conduction Model

source current are viewed as the effect of the device's voltages and not the other way around. This means that the two current equations above should be modeled as unidirectional causal relations with the voltages, $V_{G,S}$ and $V_{D,S}$, as causes, and I_D as the effect. Furthermore, positive charge flows from drain to source so these equations have been written in terms of I_D rather than I_S. I_S is determined as an effect of I_D by three terminal device KCL (section 3.3.2.1).

The mosfet is a symmetric device in the sense that T1 and T2 switch between acting as drain and source, depending on whether the mosfet is in the Forward or the Reverse operating region. We, therefore, say that a particular terminal exhibits some behavior when *acting in the role* of source or drain, rather than a terminal *is* the source or drain. T3 always acts in the role of gate; however, the role of T1 and T2 may vary over time. The qualitative model for the n-mosfet is shown in figure 6-3. Note that the mechanism for associating between terminals, nodes and roles is left implicit. For the current I_D, D refers to the terminal acting in the role of drain and in $V_{D,S}$, D and S refer to the nodes connected to the terminals acting in the role of drain and source.

Three types of transistors commonly used in nMOS digital design are enhancement mode, depletion mode and zero-threshold mosfets. The threshold voltage for an enhancement mode device is positive, for a depletion mode device negative and for a zero-threshold device approximately zero. The model for the enhancement mode n-mosfet is shown in figure 6-4. In the next section we discuss the capacitance which results from charge on the gate and along the channel.

```
Model: enhancement-n-channel-mosfet(M)
    Relations:
        [V_Th(M)] = +
        [dV_Th(M)/dt] = 0
    Assertions:
        n-channel-mosfet(M)
```

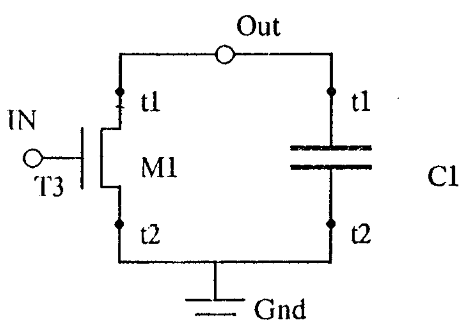
Figure 6-4: Enhancement n-Channel Mosfet Conduction Model

6.3 Mosfet Capacitance Model

Modeling mosfet capacitance is very difficult due to its non-linearities. If MOS capacitance is modeled completely, then a capacitor should be connected between every terminal of the device, including the bias node. A digital designer, however, considers at most two of these capacitances, and usually only one. These capacitances are $C_{G,S}$ and $C_{G,D}$.[41] Usually $C_{G,D}$ is less than $C_{G,S}$ and is ignored. If the designer is interested in the detrimental effects of parasitic capacitance then he considers both of these capacitors, as they both have the ability to bootstrap. If the designer is explaining a bootstrap circuit, then he assumes that $C_{G,S}$ is dominant and ignores $C_{G,D}$. (An explicit capacitor may be added to the circuit to make sure that $C_{G,S}$ is dominant). For more conventional designs, a designer usually explains the gate capacitance as a capacitor from gate to ground. This simplest model is adopted. Whenever a capacitor ($C_{G,S}$) is being used for bootstrapping, it will be made explicit.

As stated earlier, the gate capacitance on a mosfet is nonlinear. Below inversion (Off), there is no channel for charge to move into, and gate charge terminates on the substrate. This produces a very low capacitance. When the channel inverts ($V_{G,S} > V_{Th}$) charge moves into the channel and the capacitance jumps dramatically. To model this effect it would be necessary to add a qualitative mos-gate-capacitor model. In this paper the ideal capacitor model given earlier is used. Next an explanation is shown which TQ Analysis generates for a simple mosfet circuit using the model described above.

6.4 Mosfet Example



[V(In,Gnd)@instant-0] = 0

[V(Out,Gnd)@instant-0] = +

Figure 6-5:n-Mosfet-Capacitor Example

[41] $C_{G,D}$ is often referred to as the miller capacitance.

By replacing the resistor in the parallel RC circuit (figure 1-2) with a mosfet we get the circuit shown in figure 6-5. Again we will assume that the voltage on the capacitor is some positive value; in addition, the input voltage is 0 and the mosfet is off. At some time instant (Instant-0) the input begins to rise. Eventually the mosfet turns on and the capacitor begins to discharge, decreasing $V_{OUT,GND}$. The following is an explanation using TQ Analysis to determine why the $V_{OUT,GND}$ is decreasing:[42]

Explanation for FACT-210: $dV_{IN,GND}$/dt@Interval-1 is decreasing:

(1) It was given that $V_{OUT,GND}$ during Instant-0 is +.
(2) This causes M1 to be Forward,
since from rule NMOS-OP-REGION: Forward if V12 > 0.
(3) This causes T1(M1) to act as Drain and T2(M1) to act as Source, from rule NMOS-FORWARD-BEHAVIOR.
(4) Also, it was given that $V_{IN,GND}$ during Instant-0 is 0.
(5) This causes V_{On} to be -,
since from rule NMOS-BEHAVIOR: [Vgs] - [Vt] → [Von].
(6) This and (1) cause V_{Sat} to be +,
since from rule NMOS-BEHAVIOR: [Vds] - [Von] → [Vsat].
(7) also (5) causes M1 to be Off,
since from rule NMOS-OP-REGION: Off if Von ≤ 0.

Interval-0:

(8) It was given that $dV_{IN,GND}$/dt during Interval-0 becomes +.
(9) This causes $V_{IN,GND}$ to be +,
since from rule INTEGRATION: [dA/dt] → [A].
(10) Fact (8) also causes dV_{ON}/dt to be +.
since from rule NMOS-BEHAVIOR:[dVgs/dt]-[dVt/dt]→[dVon/dt].
(11) This causes V_{ON} to move towards 0
from rule INTEGRATION.

[42] This explanation is similar, in style to the output generated by the current implementation of TQ Analysis

Instant-1:

Eventually, V_{ON} becomes 0 at Instant-1
by Transition Analysis on (11).

Interval-1:

(13) V_{ON} becomes + at Interval-1 by the INTEGRATION rule.
(14) This causes M1 to be On,
since from rule NMOS-OP-REGION: On if Von $\geq$ 0.
(15) This and (6) cause M1 to be Saturated,
since from rule NMOS-OP-REGION: Saturated if On $\wedge$ Vsat>0.
(16) This and (13) cause $I_{t1(M1)}$ to be +,
since from rule NMOS-SAT-BEHAVIOR: [Von] $\rightarrow$ [Id].
(17) This causes $I_{t1(C1)}$ to be -,
since from rule 2-T-SKCL: ([I2] $\rightarrow$ - [I1]).
(18) This causes $dV_{OUT,GND}/dt$ to be -,
since from rule S-CAPACITOR: [I1] $\rightarrow$ [dV12/dt].
(19) This causes $V_{OUT,GND}$ to move towards zero,
from rule INTEGRATION.

The remainder of the explanation describes how the capacitor discharges to 0 volts and will not be given here.

The circuit shown in figure 6-6, is a simplified version of the bootstrap driver described in the introduction. The problem with many digital circuits is that their output rises slowly if the output load capacitance is large. One purpose of the bootstrap driver circuit is to provide a strong current through the pullup (M2) to the output load capacitance, allowing the output to rise quickly. This is achieved by using a bootstrap capacitor to fix a voltage across the gate and source of M2, keeping it turned on hard. To keep the capacitor's voltage constant, the terminal of C_{boot} connected to node N1 is isolated, preventing any current from flowing into the capacitor. This isolation is performed by M3 which turns off when node N1 rises above the supply voltage. The digital behavior of the circuit's

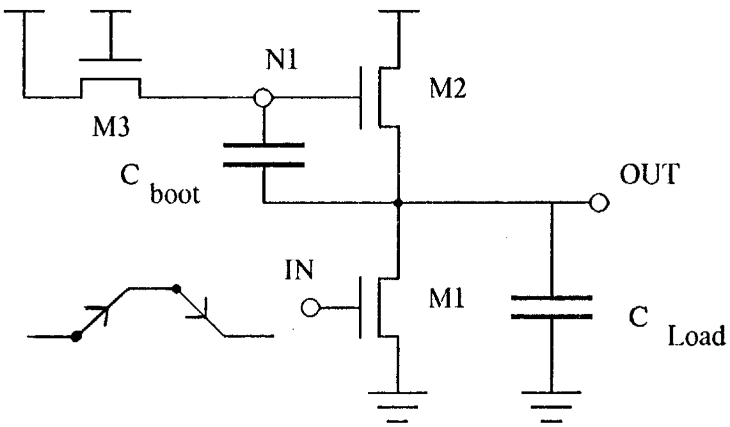


Figure 6-6:Simple Bootstrap Circuit

output is simply to invert the input.

The techniques described above make it possible to analyze the behavior of this circuit. We will examine the behavior of the circuit at two points, when the input rises and when it falls. Initially the input is low and the circuit has stabilized with the following conditions:

M1: Off $V_{G,S} = 0$
M2: Off, $V_{G,S} = V_T$
M3: Off, $V_{G,S} = V_T$
$[V_{Cload}] = +$
$[V_{Cboot}] = +$

Immediately following some instant the input voltage begins to fall ($[dV_{IN}/dt] = +$). TQ Analysis is then used to determine the response of the circuit to this input. Using the results of this analysis it is possible to answer some interesting questions about the circuit's behavior:[43]

[43] The text of the explanation shown here is what we would eventually like to produce and was generated by hand from the rules discussed in this paper.

What happens to V_{OUT} when V_{IN} rises?

V_{OUT} decreases

Why?

An increase in V_{IN} increases $V_{G,S(M1)}$ and eventually turns on M1, producing an increase in $I_{D(M1)}$, which begins to discharge C_{Load} and decreases V_{OUT}. As V_{OUT} drops, C_{Boot} acts like a battery, pulling down N1 and increasing $V_{G,S(M3)}$. This causes M3 to turn on, producing a current out of its Source and into C_{Boot}, which causes $V_{G,S(M2)}$ to increase, turning on M2 and producing a current out of its Source. Both C_{Boot} and M2 supply current to the Drain of M1, decreasing the amount of current flowing out of C_{Load} and decreasing the rate at which V_{OUT} drops by negative feedback.

In the above explanation we see how the rising input causes the output to drop and C_{boot} to charge up, preparing it to bootstrap M2 on the falling edge of the input. The other interesting part of the circuit's behavior occurs when the input is high and begins to fall. Just before the input falls the circuit has stabilized with the following conditions:

```
M1: Unsaturated, [I_D] = +
M2: Saturated, [I_D] = +
M3: Saturated, [I_D] = 0
[V_Cload] = + (but close to zero)
[V_Cboot] = +
```

The results of TQ Analysis is then used to describe the effect of C_{Boot} on M2's drain current when Vin falls:

What happens to $I_{D(M2)}$ when V_{IN} decreases?

$I_{D(M2)}$ remains constant

Why?

A decrease in V_{IN} produces a decrease in $I_{D(M1)}$, causing I_{Cload} to increase and become positive. This causes C_{Load} to charge, thus increasing V_{OUT}. An increase in V_{OUT} causes an increase in V_{N1}, turning off M3. This causes $I_{t1(Cboot)}$ to be zero, holding $V_{G,S}$ across M2 constant, and in turn causing $I_{D(M2)}$ to remain constant.

In this explanation we see how M3 turning off isolates N1 which prevents current from flowing into C_{boot}. This holds V_{Cboot} constant which allows a change in V_{OUT} to be directly transferred to V_{N1}. Thus, C_{boot} is "bootstrapping" V_{N1} from V_{OUT}.

6.5 Summary

This chapter first described a technique for modeling devices with multiple operating regions, using the qualitative arithmetic defined previously, along with a small set of predicates and boolean operations. This technique takes advantage of the mechanism described in Transition Analysis to determine how a device moves between operating regions. However, to use Transition Analysis it is necessary that the relations describing each device is continuous across region boundaries.

We then discussed the qualitative model for an n-mosfet which is broken into two parts: conduction and capacitance. Conduction is modeled as a a non-linear dependent current source described by the enhancement n-channel mosfet conduction model (section 6.2) and the capacitance is modeled as an ideal capacitor (section 3.3.2.2) from gate to source or gate to ground . In the mosfet model, the notion was introduced that a relation can be written in terms of the *role* which a terminal is playing, as well as the terminal itself. Next these models were used to generate a detailed explanation for a simple MOS circuit. Finally, TQ Analysis was used to describe some interesting characteristics of a simplified version of the bootstrap circuit presented in the introduction.

Chapter Seven

Discussion

This paper has presented a technique, Temporal Qualitative Analysis, for analyzing the behavior of MOS circuits whose behavior straddles the analog and digital domains. Throughout this work, we have emphasized the close relationship between an expert's intuition and formal theory. To summarize:

* By assuming that quantities may be represented qualitatively as open regions separated by boundaries, we have been able to unify the representations for space, time, state variables, and device operating regions.

* Using the Lumped Circuit Approximation, circuit behavior has been modeled with a set of network laws and device relations. By looking at the electrodynamics which underlie circuit theory and understanding the limitations of the Lumped Circuit Approximation, we are able to impose a causal viewpoint on these relations (*Causal Propagation*).

* By assuming that physical quantities are modeled by continuous functions, we have been able to use a few simple theorems to determine how state variables move between qualitative regions (*Transition Analysis*). These theorems capture one's intuitive notion of continuity and integration.

* Overall structural patterns of the relations, which describe a circuit's behavior, such as simultaneities and feedback, have been used to derive additional constraints. Furthermore, the intuitions designers use for feedback have been employed to determine the direction of a feedback path (*Feedback Analysis*).

* Because the boundaries of device operating regions are constructed from relations between state variables, the mechanism provided by Transition Analysis for determining state variable transitions is also applicable to determining how devices move between operating regions. In addition open regions and boundaries can be combined to describe more complex regions such as semi-open time intervals and closed operating regions.

7.1 Related Work

7.1.1 Qual and Envisioning

Some of the most notable work on the qualitative analysis of electrical systems has been done by Johan de Kleer and John Seely Brown. In his PhD dissertation, de Kleer discusses Incremental Qualitative (IQ) Analysis, a causal qualitative analysis technique that was the basis of a program, QUAL, for describing and recognizing the functionality of bipolar circuits. [7] de Kleer and Brown's recent work on Envisioning [10, 11, 13] extends this earlier work to other domains, using the methodology provided by system dynamics for describing a variety of physical systems (e.g. electrical, mechanical, fluid and thermal systems) in terms of networks of lumped elements. As discussed below, this work also extends QUAL's theory in the areas of state, time, continuity and transitions.

7.1.1.1 Operating Regions and State

QUAL provides the qualitative arithmetic and some of the basic framework for the propagation mechanism used in TQ Analysis. However, the analog bipolar domain used by de Kleer and the digital MOS domain differ in a number of important respects.

First, when modeling an analog bipolar amplifier, the analyst is primarily concerned with the incremental response of the circuit to a small variation on the input. During the analysis it is assumed that each device in the circuit will remain within a single operating region and the initial perturbation will propagate across the circuit instantaneously; this corresponds to the circuit's small signal behavior. When modeling digital MOS circuits the analyst is still interested in the instantaneous incremental response of the circuit; however, he is equally concerned with the circuit's large signal behavior. This includes the long term effects of a changing input such as a device moving between operating regions or a positive quantity becoming zero.

Second, the modeling of capacitive memory was not important in de Kleer's work since, during the small signal analysis of bipolar circuits, large capacitors become incremental shorts and small, parasitic capacitances become open circuits. In the MOS domain circuit behavior is strongly dependent on charge flow and capacitive memory which results from charge storage.

Because de Kleer was primarily interested in incremental behavior, qualitative relationships were written only in terms of voltage and current derivatives. In TQ Analysis this vocabulary is expanded to include any type of quantity such as charge, current, voltage, their first derivatives, second derivatives and any higher order derivatives.

In QUAL the relationships among a device's different operating regions and the movement between them is represented by explicit statements like "If the Diode is ON and the voltage across the diode is decreasing then the diode will turn off." In addition, the persistence (inertia) exhibited by a device with memory is modeled by breaking the device into several explicit states that include statements describing the device's movement from one state to the next (e.g. two of the states for a capacitor are positively charging and constant). In QUAL, operating regions are not distinguished from state; the term "state" is used to refer to both.

In TQ Analysis a distinction is made between operating regions and state. State is viewed as a property of quantities (i.e., state variables) rather than the device itself; the state of a device is then described by the values of its independent state variables. The persistence exhibited by a device with memory results from the continuity of its state variables. State is then a property of continuous quantities, rather than devices, and is the qualitative region in which the state variable currently lies (e.g., +, 0 or -). For example, the notion that a capacitor is discharging is equivalent to saying that the capacitor's charge is positive and decreasing. Furthermore, if the change in a device's state variables is zero the device exhibits memory (e.g., a capacitor stores charge or a mosfet remains in a single operating region). Operating regions, on the other hand, are properties of the devices themselves and are described as a set of qualitative regions on the device's state variables.[44]

Transition Analysis provides a single mechanism for determining how quantities move from one qualitative region to another. These regions may be as simple as positive and negative, or as complex as the operating regions: Saturated, Unsaturated and Cutoff. The mechanism depends on the properties of continuous quantities and the relationship between them, rather than properties specific to the device. Transition Analysis is, therefore, independent of the model, the domain and even Network Theory.

In recent work by de Kleer and Brown state and operating regions are still indistinguishable, however, boundaries between device "states" are modeled in terms of inequalities between "qualitative variables". The notions of qualitative calculus and continuity are then used to provide a mechanism for recognizing transitions between states (i.e., the inter-state behavior). The strong similarities between this mechanism and Transition Analysis suggests the possibility of unifying the two viewpoints.

---

[44] Recall that a device is broken into several operating regions if the behavior of the device in each region is described by a different set of equations.

7.1.1.2 Feedback

QUAL provided two heuristics for dealing with ambiguities introduced by simultaneities in network laws: the KVL Connection Heuristic (discussed in section 5.4) and the KCL Heuristic. Recent work on ENVISION refers to these heuristics as the Component and Conduit Heuristics, respectively, and provides an additional heuristic for resolving ambiguities introduced by simultaneities in device models (i.e., the Confluence Heuristic). These two heuristics suggest a style of reasoning where ambiguities introduced by simultaneities are resolved heuristically using information local to the ambiguity. This style of reasoning is the motivation behind the Feedback Analysis portion of TQ Analysis. Feedback Analysis separates the general properties of simultaneous relations (e.g., Resistive Feedback Rule) from those properties which are specific to Network Theory (e.g., KVL Locality Heuristic). In addition, the feedback properties of systems with memory are exploited (Integrating Feedback Rule) and an engineers intuition about devices which introduce memory (e.g., capacitors and inductors) is used to determine the direction the feedback flows (Feedback Direction Heuristic).

7.1.2 Qualitative Process Theory

Forbus' Qualitative Process Theory [14] provides a viewpoint where physical interactions are described through properties of processes, rather than properties of devices. A process is the basic vehicle for change in a physical system. Examples of processes are heating, evaporating, stretching, and flowing. A process-centered viewpoint is quite natural for many physical domains where a device centered model would be awkward; on the other hand, many domains, especially circuit analysis, naturally fit into a device centered model and would be difficult to understand in terms of a process centered model. QP Theory and TQ Analysis also differ in intent; Qualitative Process Theory models common-sense reasoning about everyday physics. Temporal Qualitative Analysis models expert reasoning about electronics. TQ Analysis tries to create a close link between the intuitions an expert has and the formal theory that his expertise is built upon (e.g. calculus, algebra, circuit theory, and feedback analysis). For example, QP Theory describes the behavior of negative feedback in terms of an instantaneous oscillation called "stutter" that results from the feedback term oscillating instantaneously between being equal to and less than the input to the feedback loop. An engineer doesn't consider these oscillations in complex circuits since they violate continuity (i.e., a quantity cannot jump instantaneously); however, he may use an argument similar to stutter to describe feedback to someone unfamiliar with the phenomena.[45]

In spite of these differences QP Theory and TQ Analysis have a number of similarities. In QP Theory there is a clear separation between the properties of quantities, as defined by the "Quantity Space", and the properties of processes. The clear division of quantities and processes motivated the separation of TQ Analysis into its basic components. In TQ Analysis the only knowledge that depends on a device centered view is the device model and network laws which have been made explicit through causal rules. Furthermore, the mechanisms for determining transitions between operating regions or processes in each theory are domain independent and have a number of features in common. The similarities between philosophy and underlying mechanism suggest that these two viewpoints can be unified into a single theory.

7.1.3 Allen's Temporal Intervals

In this paper we have discussed the temporal representation that an engineer uses in reasoning about circuits. Allen [2] describes an interval based temporal representation and reasoning mechanism that, among other applications, could be used as the temporal component of Naive Physics. Time is represented as a set of intervals and the relationships between them. Allen argues that zero-width time points are counter-intuitive; therefore, the temporal intervals in his representation are neither open nor closed but are described as "meeting".

Allen supports this argument with the following example [2]:

> " . . . consider the situation where a light is turned on. To describe the world changing we need to have an interval of time during which the light was off, followed by an interval during which it was on. The question arises as to whether these intervals are open or closed. If they are open, then there exists a time (point) between the two where the light is neither on nor off. Such a situation would provide serious semantic difficulties in a temporal logic. On the other hand, if intervals are closed, then there is a time point at which the light is both on and off. This presents even more semantic difficulties than the former case."

The problem in this example is not with intervals being open or closed, but that a continuous process is being modeled discontinuously. Consider what really happens when the light switch is flipped. When the light switch is closed, current begins to flow through the switch into the light bulb and the filament begins to glow at a very small intensity. Initially the light bulb resists this current, due to

---

[45] In fact, the description of the intuition behind the Resistive Feedback Rule in section 5.2 is very similar to the stutter argument in QP Theory.

inductance in the filament; however, eventually the light reaches a steady intensity and is considered "on". This process consists of a closed time interval where the light is off, followed by an open interval during which the light's intensity is increasing and ending with a closed interval when the light is "on". Of course this occurs too fast for the human eye to see. The process, therefore is collapsed into an instant and the light intensity is perceived as stepping from "off" to "on". The reason for collapsing these series of events into an instant is that we are not interested in their details. The price that must be payed for this abstraction is that the process is no longer continuous, and if we look at the process too closely our intuitions about continuity will be violated (after all, how can a light be both on and off during the same instant!).

Allen also argues that ". . . given an event, we can always 'turn up the magnification' and look at its structure." This is certainly true in some cases. In the light switch example " by turning up the magnification" we see that the light intensity doesn't really step from off to on but changes continuously over a finite interval of time. The instant the light switches from off to on is a useful idealization, something which can't be looked inside of without changing magnification (i.e., abstraction level).

The notion that zero-width time points exist is not counter-intuitive to someone with a math background. Intuitions are developed from observations about the surrounding world. For example, it seems intuitively obvious that a person who wants to enter a building cannot walk through its brick walls, but must go through a door or some other opening. Early on in our math background many of us are told of such concepts as infinitely thin, one dimensional lines and points with no dimensions. These are things which don't change appearance, no matter how much the magnification is turned up. Using these concepts as givens, our instructors teach us how they can be manipulated to understand and idealize many things that happen in the real world. Eventually notions such as zero-widths points become part of the intuitions of someone like an engineer or mathematician. For example, when a ball is thrown in the air, it is obvious to a physicist that the ball is at the top of its arc for only an instant independent of the magnification; to hang there longer would defy the laws of gravity.[46]

In spite of these differences the temporal representation used here has a number of similarities to that of Allen's. First, although instants play an important role in TQ Analysis, they are viewed as a subset

[46] Of course, this is an ideal ball which can't be held up by such things as rising gusts of wind!

of closed intervals. Second, except for the parts of Transition Analysis which depend on the continuity of quantities, closed and open intervals are not differentiated (e.g., in Causal Propagation no distinction was made between the types of intervals). Finally, open and closed intervals may be viewed as meeting; however, continuity places the restriction that the closed end of an interval can only be met by the open end of another interval (i.e., time is dense). If continuity were ignored, the distinction between open and closed intervals would not be necessary.

Allen has provided a rich vocabulary for describing the relationships between intervals and a mechanism for reasoning about these relations. As discussed in the next section, a "temporal reasoner" based on open and closed intervals is currently being developed and will be incorporated into TQ Analysis.

7.2 Limitations and Future Directions

As of this writing, all of the parts of TQ Analysis have been implemented and tested except for Feedback Analysis. The system has been tested on simple R,L,C and mosfet circuits. The explanations for the parallel RC and mosfet-capacitor circuits given earlier in this paper were generated by the system. I am currently working towards generating a qualitative description of circuits similar to the bootstrap clock driver given in the introduction.

The representation for time in TQ Analysis is in the process of being modified. The current representation of time as a linear sequence places a total ordering on all events. Such a global viewpoint is often not necessary or desirable, since it requires that an ordering be placed on unrelated events. A more realistic representation breaks time into a set of open and closed intervals in a temporal network. This creates a partial ordering on time intervals, rather than a total ordering. Two quantities, then interact only if they are locally connected in space and their time intervals (or instants) coincide.[47]

There are two major components of circuit analysis which have not been addressed in this work so far: the use of quantitative information and the use of cliches [5, 3].

In TQ Analysis many of the ambiguities which arise during Causal Propagation can be resolved using non-numerical information, such as continuity theorems or properties of feedback. There are,

---

47. This is similar to Hayes' notion of Histories [16].

however, a number of ambiguities which cannot be resolved using these techniques. Many of these remaining ambiguities can only be resolved using some form of quantitative information. A circuit analyst, for example, often considers certain circuit parameters dominant, and might use phrases like the following when analyzing a circuit:

* The current drawn by $C_{G,S}$ is insignificant compared to $I_{D,S}$.

* The rise time of the output capacitance is much slower than the input.

* The pullup is much longer than the pulldown, allowing the inverter to meet a valid logic low level.

Quantitative comparisons such as these, and other types of quantitative knowledge must at some point be integrated with TQ Analysis.

A number of cliches are used by a designer in analyzing a circuit, such as the phrases "isolation", "precharging", and "bootstrapping" used in the bootstrap clockdriver example in the introduction. A cliche can either refer to a set of devices, such as "precharge circuit" and "bootstrap capacitor", or it can refer to a complex behavior, such as "precharging the input node" or "isolating the gate". A cliche which refers to a device can be used to help determine which of a number of possible behaviors the designer intend the circuit to have. Furthermore, the ability to combine a series of events into a cliche is important in generating a qualitative summarization of a circuit's behavior.

If the behavior of a device within a particular operating region during a time interval is viewed as an episode, then a cliche may be described in terms of a sequence of episodes. This is similar to what Forbus refers to as an encapsulated history. [14] Cliches may then be used in analyzing the circuit's behavior to answer questions not yet addressed by TQ Analysis such as:

* Why didn't the circuit behave as I expected?

* Which device parameters should I change to make it work? and

* Which parameters should I change to increase the circuits performance (speed, power, voltage thresholds etc.)?

By answering these and similar questions we hope to create a versatile tool which provides the circuit analyst with expert advice on a wide class of circuits.

Appendix A

Appendix: Transition Ordering Rules

A.1 Predicates

to-zero(Q)	Q is moving monotonically towards zero during the current interval.
not-to-zero(Q)	Q is constant or moving away from zero during the current interval.
transition(Q)	Q will transition at the end of the current interval.
not-transition(Q)	Q cannot transition at the end of the current interval.

A.2 General

* not-to-zero(A) $\rightarrow$ not-transition(A)

A.3 Equality

Transition inferences: $[A] = (+/-)\,[B]$

* transition(A) iff transition(B)

* not-transition(A) iff not-transition(B)

Direction inferences: A is a monotone increasing (decreasing) function of B

* to-zero(A) iff to-zero(B)

* not-to-zero(A) iff not-to-zero(B)

A.4 Sums and Differences

Transition inferences: $[A] + [B] + [C] = 0$ where $[A] = [B] = -\,[C]$

* transition(A) $\wedge$ transition(B) $\rightarrow$ transition(C)

* not-transition(A) $\vee$ not-transition(B) $\rightarrow$ not-transition(C)

* transition(C) → transition(A) ∧ transition(B)

* not-transition(C) → not-transition(A) ∨ not-transition(B)

Direction inferences: A + B = C where [A] = [B] = − [C]

* to-zero(A) ∧ to-zero(B) → to-zero(C)

* not-to-zero(A) ∧ not-to-zero(B) → not-to-zero(C)

* to-zero(C) → to-zero(A) ∨ to-zero(B)

* not-to-zero(C) → not-to-zero(A) ∨ not-to-zero(B)

A.5 Products

Transition inferences: [A] x [B] = [C]

* transition(A) ∨ transition(B) → transition(C)

* not-transition(A) ∧ not-transition(B) → not-transition(C)

* transition(C) → transition(A) ∨ transition(B)

* not-transition(C) → not-transition(A) ∧ not-transition(B)

Direction inferences: A x B = C

* to-zero(A) ∧ to-zero(B) → to-zero(C)

* not-to-zero(A) ∧ not-to-zero(B) → not-to-zero(C)

* to-zero(C) → to-zero(A) ∨ to-zero(B)

* not-to-zero(C) → not-to-zero(A) ∨ not-to-zero(B)

A.6 Resistive Feedback

Relation: [A] + [B] = [C] which is the comparison point of a feedback loop.

Idea: If A is the cause and C is the effect of a feedback loop then [A] = [B] and all the inferences for equivalences apply.

References

1. Allen, J. A General Model of Action and Time. Tech. Rep. 97, University of Rochester Computer Science Department, November, 1981.

2. Allen, J. Maintaining Knowledge about Temporal Intervals. *Communications of the ACM 26*, 11 (November 1983), 832-843.

3. Brotsky, D. Program Understanding through Cliche Recognition. AI working paper 224, Massachusetts Institute of Technology, December, 1981.

4. Cherry, J. and Roylance, G. A One Transistor RAM for MPC Projects. VLSI memo 81-43, Massachusetts Institute of Technology, February, 1981.

5. Cyphers, S. Programming Cliches and Cliche Extraction. AI working paper 223, Massachusetts Institute of Technology, February, 1982.

6. de Kleer, J., Doyle, J. et al. AMORD: A Deductive Procedure System. AI working paper 151, Massachusetts Institute of Technology, August, 1977.

7. de Kleer, J. Causal and Teleological Reasoning in Circuit Recognition. Ph.D. Th., Massachusetts Institute of Technology, September, 1979.

8. de Kleer, J. and Sussman, G. Propagation of Constraints Applied to Circuit Synthesis. *Circuit Theory and Applications 8* (1980), 127-140.

9. de Kleer, J. and Brown, J. Mental Models of Physical Mechanisms and their Acquisition. In *Cognitive Skills and their Acquisition*, J. R. Anderson, Ed., Erlbaum, 1981.

10. de Kleer, J. and Brown, J. Assumptions and Ambiguities in Mechanistic Mental Models. In *Mental Models*, D. Gentner and A. Stevens, Eds., LEA, Inc., 1983.

11. de Kleer, J. and Brown, J. The Origin, Form and Logic of Qualitative Physical Laws. Eighth International Joint Conference on Artificial Intelligence, IJCAI, August, 1983.

12. deKleer, J. and Bobrow, D. Qualitative Reasoning with Higher Order Derivatives. Proceedings of the AAAI, August, 1984.

13. de Kleer, J. and Brown, J. A Qualitative Physics Based on Confluences. In same volume.

14. Forbus, K. Qualitative Process Theory. AI memo 664A, Massachusetts Institute of Technology, May, 1983.

15. Forbus, K. Measurement Interpretation in Qualitative Process Theory. Eighth International Joint Conference on Artificial Intelligence, IJCAI, August, 1983.

16. Hayes, P. The Naive Physics Manifesto. In *Expert Systems in the Micro-Electronic Age*, D. Michie, Ed., Edinburgh University Press, 1979.

17. Kuipers, B. Commonsense Reasoning About Causality: Deriving Behavior from Structure. Tech. Rep. 18, Tufts University Working Papers in Cognitive Science, May, 1982.

18. Kuipers, B. Getting the Envisionment Right. Proceedings of the AAAI, August, 1982.

19. Loomis, L. *Calculus.* Addison-Wesley, Reading, Mass., 1977.

20. Newton, A., Pederson, D., et al. *Spice Version 2G User's Guide.* Dept. of EECS, U of C, Berkeley, 1981.

21. Shearer, J. L., Murphy, A. T. and Richardson, H. H. *Introduction to System Dynamics.* Addison-Wesley, Reading, Mass., 1971.

22. Shichman, H. and Hodges, D. A. Modeling and Simulation of Insulated-Gate Field-Effect Transistor Switching Circuits. *IEEE Journal of Solid State Circuits SC-3*, 3 (1968), 285-289.

23. Simmons, R. The Use of Qualitative and Quantitative Simulations. Proceedings of the National Conference on Artificial Intelligence, AAAI, August, 1983.

24. Stallman, R. and Sussman, G. Forward Reasoning and Dependency-Directed Backtracking in a System for Computer-Aided Circuit Analysis. *Artificial Intelligence* 9 (1977), 135-196.

25. Steele, G. The Definition and Implementation of a Computer Programming Language Based on Constraints. Ph.D. Th., Massachusetts Institute of Technology, August, 1980.

26. Thomas, G. *Calculus and Analytic Geometry.* Addison-Wesley, Reading, Mass., 1968.

27. Williams, B. The Use of Continuity in a Qualitative Physics. Proceedings of the AAAI, August, 1984.

28. Williams, B. Qualitative Analysis of MOS Circuits. *Artificial Intelligence* (To Appear 1984).

29. Williams, B. and Zippel, R. The Mathematical Foundations of Temporal Qualitative Analysis. In progress

CS-TR Scanning Project
Document Control Form

Date : 2 / 6 / 96

Report # AI-TR-767

Each of the following should be identified by a checkmark:
Originating Department:

- [x] Artificial Intellegence Laboratory (AI)
- [ ] Laboratory for Computer Science (LCS)

Document Type:

- [x] Technical Report (TR)
- [ ] Technical Memo (TM)
- [ ] Other: ____________

Document Information

Number of pages: 90 (98 - images)
Not to include DOD forms, printer intstructions, etc... original pages only.

Originals are:

- [x] Single-sided or
- [ ] Double-sided

Intended to be printed as :

- [ ] Single-sided or
- [x] Double-sided

Print type:

- [ ] Typewriter
- [ ] Offset Press
- [ ] Laser Print
- [ ] InkJet Printer
- [ ] Unknown
- [x] Other: ____________

Check each if included with document:

- [x] DOD Form (2)
- [ ] Funding Agent Form
- [x] Cover Page
- [x] Spine
- [ ] Printers Notes
- [ ] Photo negatives
- [ ] Other: ____________

Page Data:

Blank Pages (by page number): ____________

Photographs/Tonal Material (by page number): ____________

Other (note description/page number):

Description : Page Number:

(A) IMAGE MAP: (1-90) UNH'ED TITLE PAGE, 2-90
(91-98) SCANCONTROL, COVER, SPINE, DOD (2), TRGT'S (3)
(B) CUT & PASTE FIG ON PAGE 10

Scanning Agent Signoff:

Date Received: 2 / 6 / 96 Date Scanned: 2 / 8 / 96 Date Returned: 2 / 15 / 96

Scanning Agent Signature: Michael W. Cook

Rev 9/94 DS/LCS Document Control Form cstrform.vsd

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AI-TR-767	2. GOVT ACCESSION NO.
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Qualitative Analysis of MOS Circuits	5. TYPE OF REPORT & PERIOD COVERED: technical report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Brian Charles Williams	8. CONTRACT OR GRANT NUMBER(s): N00014-80-C-0505, N00014-80-C-0622
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, Massachusetts 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Blvd, Arlington, Virginia 22209	12. REPORT DATE: July 1984
	13. NUMBER OF PAGES: 90
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): Office of Naval Research, Information Systems, Arlington, Virginia 22217	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

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

Distribution is unlimited

18. SUPPLEMENTARY NOTES

None

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

causal reasoning
qualitative physics
qualitative cireuit simulation
circuit theory
VLSI
design automation
problem solving
representation of knowledge
expert systems
artificial intelligence

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

With the push towards sub-micron technology, transistor models have become increasingly complex. The number of components in integrated circuits has forced designer's efforts and skills towards higher levels of design. This has created a gap between design expertise and the performance demands increasingly imposed by the technology. To alleviate this problem, software tools must be developed that provide the designer with expert advice on circuit performance and design. This requires a theory that links the (OVER

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 cont.

intuitions of an expert circuit analyst with the corresponding principle of formal theory (i.e., algebra, calculus, feedback analysis, network theory, and electrodynamics), and that makes each underlying assumption explicit.

Temporal Qualitative Analysis is a technique for analyzing the qualitative large signal behavior of MOS circuits that straddle the line between the digital and analog domains. Temporal Qualitative Analysis is based on the following four components: First, a qualitative representation is composed of a set of open regions separated by boundaries. These boundaries are chosen at the appropriate level of detail for the analysis. This concept is used in modeling time, space, circuit state variables, and device operating regions. Second, constraints between circuit state variables are established by circuit theory. At a finer time scale, the designer's intuition of electrodynamics is used to impose a causal relationship among these constraints. Third, large signal behavior is modeled by Transition Analysis, using continuity and theorems of calculus to determine how quantities pass between regions over time. Finally, Feedback Analysis uses knowledge about the structure of equations and the properties of structure classes to resolve ambiguities.